$\mathcal{L}_1$-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 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. We further show how the 'size' of the ambiguity tube can be controlled using certain tuning-knobs that $\mathcal{L}_1$-DRAC provides. We demonstrate with a few illustrative examples how $\mathcal{L}_1$-DRAC can operate systems while guaranteeing the existence and tunability of robustness certificates in terms of ambiguity sets/tubes.

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.

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 [18]. 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 [19] 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 [20, 21]. Indeed, distribution shift offers a significant hurdle in using learned models and policies in safety-critical systems, see [22, 23] 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 [24, 25, 26]. 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 [27].

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 methodology for uncertain stochastic systems to produce certificates of robustness that are distributional in nature and thus can facilitate seamless and non-conservative integration with data-driven learned models that can be verified to

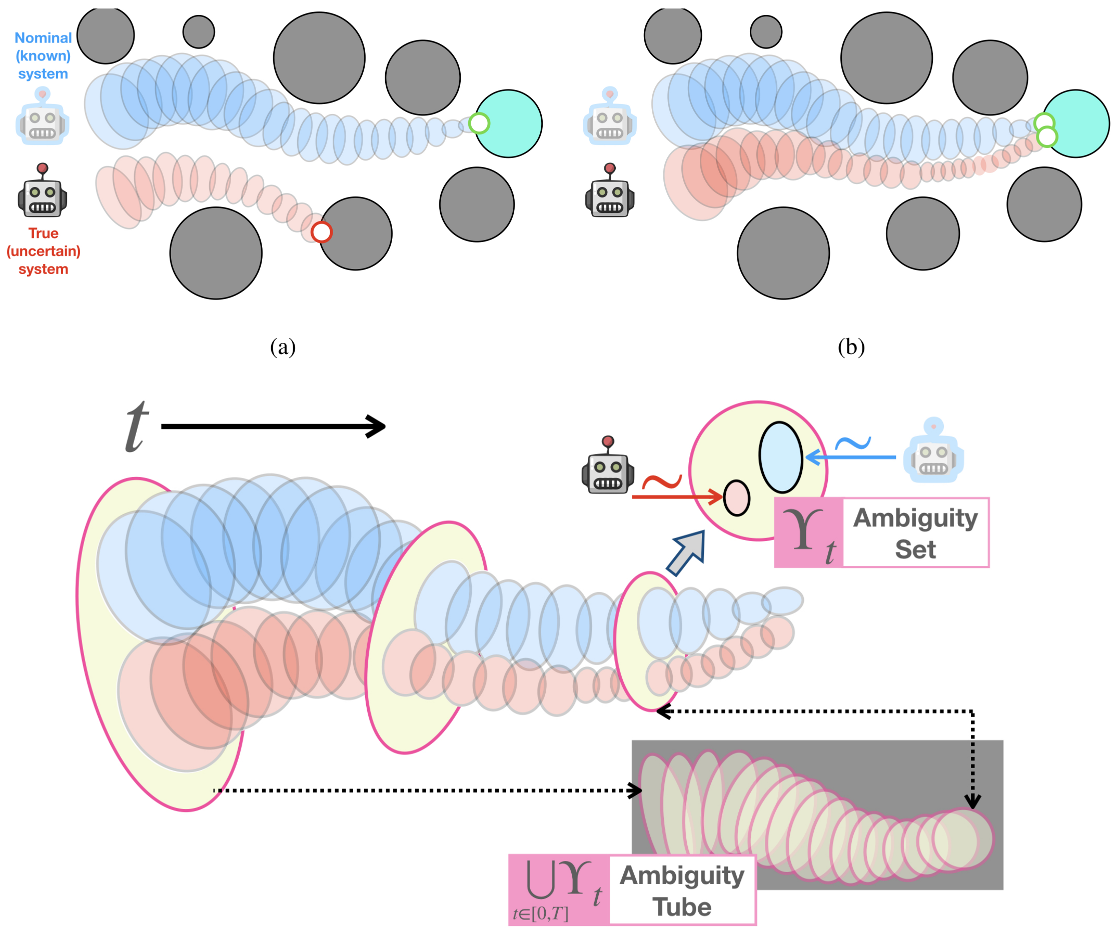

Figure 1: Consider the problem of safely navigating an uncertain stochastic system to a goal set (green circle), avoiding unsafe subsets of the state-space (grey circles). One constructs a control policy for the uncertainty-free nominal (known) system (the robot with faded colors and a blue outline) 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* (Υ_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* ($\cup_{t\in[0,T]}\Upsilon_t$, for any $T \in (0,\infty)$), which enables safe predictive planning.

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 [28, 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 [29], 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% subscale generic transport aircraft model [30], Calspan's Learjet [31], and uncrewed aerial vehicles [32, 33]. 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 [34] and [35] 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 $t \to 0$. Results on finite-time (uniform) guarantees for stochastic systems in discrete-time can be found in [36], and for continuous-time in [37] and [38]. Using a different analysis, the authors in [38] were able to provide bounds on higher-order moments leading to tighter tracking error bounds. Furthermore, the results in [35] and [39] require linearity of the systems under consideration.

(ii) Control-theoretic approaches: Control Lyapunov function (CLF)-based approaches [40, 41] 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 [42, 43, 44, 45] wherein notions like that of globally asymptotically stable in probability are used [46, Chp. 2], [47, 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 [47, Sec. 5.1]. The absence of such constraints prevents the stabilization of the trivial solution and thus requires robust approaches, e.g., see [42, 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 [46, Lem. 2.1], [47, 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 [38] 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 [48]. In fact, contraction theory offers necessary *and* sufficient characterization for the stability of trajectories of nonlinear systems [49]. A few examples of contraction theory-based solutions for control and estimation of nonlinear deterministic systems can be found in, e.g. [50, 51] and in [52, 37] for stochastic systems. The tutorial paper [53] 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 [54, 53] prove the robustness of the stochastic counterpart systems. However, as the authors in [38] 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 [54, 53], which apply only to stochastic systems whose noise (diffusion) vector fields are uniformly spatially and temporally bounded. The authors in [55] developed a learning-based synthesis of contraction theory-based control for nonlinear stochastic systems; however, similar to [54], 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. [56] 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 [57], wherein the authors consider output-feedback tracking for a system with linearly parameterized uncertainties and bounded noise (diffusion) vector field. Using deep-learning methods, the authors in [58] (see also [53, 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 continuous-time random processes like

[1] We shall provide rigorous definitions for ambiguity sets and tubes later in the manuscript; we avoid the rigor here to avoid overburdening the reader in the introduction section. For the moments, ambiguity sets can be considered simply as a set of distributions (Borel measures on vector spaces over the field of reals).

Brownian motion [59], and 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 [60], 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. The result in [61] also uses $\mathcal{L}_1$-adaptive control, albeit for linear systems with additive Brownian motion.

(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 [62] 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 [62] 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. [63] wherein the authors use the Hamilton-Jacobi reachability analysis for systems under unknown but bounded disturbances. Similarly, the authors in [64] 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 [64], 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 [65]. Reachability analysis of systems perturbed by noise with non-compactly supported distributions is a challenging prospect as evidenced by recent results in [66, 67, 68] wherein the authors consider linear systems with additive random disturbances that are possibly correlated, and [69] wherein the authors consider both linear and nonlinear systems with the vector fields are assumed to be known. The work in [70] considers reachability analysis for nonlinear stochastic systems in terms of probability measures corresponding to the state distributions.

(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. [71, 72, 73] 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 [74, 75] 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 [76] provided a distributionally robust integration of perception, planning, and control. While the authors in [77] consider DRC for nonlinear systems, the noise is assumed to perturb the system additively. The authors in [78, 79] 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 [80] 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 [81]. Instead of considering systems with distributionally uncertain disturbances, the DRO formulation can also be used for the case of uncertain environments, see e.g. [82, 83, 84, 85].

1.2 Contributions

The following describes our contributions regarding the features of the $\mathcal{L}_1$-DRAC control compared to existing results 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, suppose the initial distributions have bounded higher moments. In that case, $\mathcal{L}_1$-DRAC can ensure the existence of *a priori* **higher-order Wasserstein ambiguity sets (tubes)**, thus leading to, e.g., tighter tracking errors. However, 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. Finally, in addition to the ambiguity tube guarantees defined as the union of time-dependent ambiguity sets, we also show how $\mathcal{L}_1$-DRAC provides **robust guarantees** in the form of ***a priori* computable ambiguity sets in the space of path measures** (distributions defined directly on the system trajectories over $t \in [0, T]$, for any $T < \infty$). In other words, instead of taking union of ambiguity sets over probability measures on the space of real-valued vectors, we can construct ambiguity tubes as ambiguity sets over probability measures on the set of continuous functions. We posit that such robust path measures can greatly benefit sampling-based planning and control as it will allow one to directly sample trajectories instead of sampling from multiple temporal distributions over the control horizon.

(ii) The $\mathcal{L}_1$-DRAC control relies only on the stability of the deterministic counterpart of the known (nominal) subsystem as in [38, 55, 54]. However, when compared to the existing control-theoretic approaches, $\mathcal{L}_1$-DRAC is applicable to **nonlinear systems with (globally) unbounded 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. Furthermore, $\mathcal{L}_1$-DRAC can also handle **noisy control channels** wherein an instance of the Brownian motion acts multiplicatively on the control input (systems subject to **control multiplicative noise**). Systems with control multiplicative noise are important for modelling processes that are relevant to biomechanics, neuroscience, autonomous systems, and finance [86, 87, 88, 89, 90, 91, 92]. In conclusion, we design $\mathcal{L}_1$-DRAC for **systems with *both* state and control multiplicative noise**.

(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 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 [93, 59], the presence of state-dependent uncertainties in *both* drift and diffusion terms imply the **random perturbations affecting the systems we consider 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 knowledge of the dynamics and the distribution of disturbance/noise, both lead to uncertain state distributions.

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**. 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 both temporal and path measures**, and 3) a certain amount of control on the radius of the ambiguity sets **allowing us to exercise a degree of control on limiting the distribution shift**.

1.3 Notation

1.3.1 Sets, spaces, and norms

Unless specified otherwise, $\|\cdot\|$ denotes the 2-norm on the space $\mathbb{R}^n$, where $n \in \mathbb{N}$ will be clear from context. Similarly, $\|\cdot\|$ denotes the induced operator (Euclidean) norm on the space of linear maps $\mathbb{R}^{n\times m}$, $n, m \in \mathbb{N}$. We denote by $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geq 0}$ the set of positive and non-negative reals, respectively. Furthermore, $\|\cdot\|_F$ denotes the Frobenius norm on $\mathbb{R}^{n\times m}$. For any $A \in \mathbb{R}^{n\times n}$, $\mathrm{Tr}\,[A]$ denotes the trace operator. Note that for any $A \in \mathbb{R}^{n\times m}$, $\|A\|_F = \sqrt{\mathrm{Tr}\,[AA^\top]}$. The space of symmetric matrices $\in \mathbb{R}^{n\times n}$ is denoted by $\mathbb{S}^n$, with $\mathbb{S}^n_{\succ 0}$ and $\mathbb{S}^n_{\succeq 0}$ denoting the set of positive and non-negative symmetric matrices, respectively.

Given any set F, we denote by $\mathcal{B}(F)$, the Borel σ-algebra generated by F (assuming there is a topology or a metric to define open sets in F). Note that, if F is a normed metric space, the open sets included in the Borel σ-algebra are defined using the natural norm on F, unless otherwise specified. We denote by $\mathcal{B}(\mathbb{R}^n)$ as the Borel σ-algebra on $\mathbb{R}^n$. We say that a map is (Ω, Γ)-measurable if the map between the measurable spaces $(\Omega, \mathcal{O})$ and $(\Gamma, \mathfrak{G})$ satisfies the standard definition of measurability (pre-images of sets $\in \mathfrak{G}$ under the map, belong to $\mathcal{O}$). For any finite dimensional vector space V over the field of reals, e.g. $\mathbb{R}^n$, the σ-algebra $\mathcal{V}$ of subsets of V will be the Borel σ-algebra, i.e., $\mathcal{V} = \mathcal{B}(V)$.

Given sets F and G, we denote by $\mathcal{C}^n(F; G)$, $n \in \mathbb{N} \cup \{\infty\}$, the set of maps $F \to G$ that are n-times continuously differentiable, and $\mathcal{C}(F; G) \doteq \mathcal{C}^0(F; G)$. Similarly, we define $\mathcal{C}^n_0(F; G)$ to be the space of compactly supported functions, under the assumption that one can define compacts sets $\subseteq F$. We equip the set $\mathcal{C}(F; G)$, with the norm $\|f\|_{\mathcal{C}} = \sup_{t\in F} \|f(t)\|_G$, if G is a normed space with the norm $\|\cdot\|_G$.

Given any measure space $(\Lambda, \mathfrak{G}, \mathbb{T})$ and the measurable space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, for any $n \in \mathbb{N}$, we define the following space of equivalence classes of functions $\Lambda \to \mathbb{R}^n$ as

$$L_p(\Lambda; \mathbb{R}^n) = \left\{ f : \Lambda \to \mathbb{R}^n,\ (\Lambda, \mathbb{R}^n)\text{-measurable} :\ \|f\|_{L_p} \doteq \left(\int_\Lambda \|f(\lambda)\|^p d\mathbb{T}(\lambda) \right)^{\frac{1}{p}} < \infty \right\},\ p \in [1, \infty), \tag{1a}$$

$$L_\infty(\Lambda; \mathbb{R}^n) = \left\{ f : \Lambda \to \mathbb{R}^n,\ (\Lambda, \mathbb{R}^n)\text{-measurable} :\ \|f\|_{L_\infty} \doteq \operatorname*{ess\,sup}_\Lambda \|f\| < \infty \right\}, \tag{1b}$$

where $\mathrm{ess\,sup}$ denotes the essential supremum, and we choose the Euclidean norm inside the integrand, although, any norm can be used in place due to the equivalence of finite dimensional norms. For the case of functions $\mathbb{R}^m \to \mathbb{R}^n$, $m, n \in \mathbb{N}$, we set $(\Lambda, \mathfrak{G}, \mathbb{T}) \doteq (\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m), \mu_L)$, where μ_L is the Lebesgue measure on $\mathcal{B}(\mathbb{R}^m)$ [94]. Therefore, the norm for functions in $L_p(\mathbb{R}^n; \mathbb{R}^n)$, as in (1) is defined using the Lebesgue integral, and similarly for the $\mathrm{ess\,sup}$ for the case when $p = \infty$. We denote the set of *locally integrable* functions, for $p \in [1, \infty]$, by

$$L_p^{loc}(\Lambda; \mathbb{R}^n) = \{ f : \Lambda \to \mathbb{R}^n \mid f \in L_p(V; \mathbb{R}^n) \text{ for each open } V \subset\subset \Lambda \},$$

where $V \subset\subset \Lambda$ denotes that V is *compactly contained* in Λ, i.e., $V \subset \bar{V} \subset \Lambda$ where the closure $\bar{V}$ is compact [95, Sec. A.2]. Given any open $U \subset\subset \mathbb{R}^n$ whose boundary ∂U is at least continuously differentiable [95, Sec. C.1], and some $\mathbb{N} \ni k > 0$, consider a multiindex $\beta = (\beta_1, \cdots, \beta_n)$ of order $|\beta| = \beta_1 + \cdots + \beta_n \leq k$. Then, we define the Sobolev space [95, Ch. 5] $S^{k,p}(U; \mathbb{R}^m)$ as

$$S^{k,p}(U; \mathbb{R}^m) = \left\{ f : U \to \mathbb{R}^m\ :\ \|f\|_{S^{k,p}} \doteq \left(\sum_{|\beta|\leq k} \left[\left\| D^\beta f \right\|_{L_p} \right]^p \right)^{\frac{1}{p}} < \infty \right\},\ (1 \leq p \leq \infty),$$

where, the notation D^β is defined as

$$\left[D^\beta f(x) \right]_j \doteq \frac{\partial^{\beta_1}}{\partial x^{\beta_1}} \cdots \frac{\partial^{\beta_n}}{\partial x^{\beta_n}} [f(x)]_j, \quad |\beta| \leq k, \quad j \in \{1, \cdots, m\},$$

and denotes the *weak/generalized* derivative [95, Sec. 5.2.1]. Analogous to L_p^{loc}, we define the space of functions with locally integrable weak derivatives over any open $U \subseteq \mathbb{R}^n$ as

$$S^{k,p}_{loc}(U; \mathbb{R}^m) = \left\{ f : \Omega \to \mathbb{R}^n \mid f \in S^{k,p}(V; \mathbb{R}^n) \text{ for each } V \subset\subset \Omega \right\}.$$

We denote by $\mathcal{L}_\infty$ the space of piecewise continuous (in time) and bounded functions $f(t) \in \mathbb{R}^n$ with $\|f\|_{\mathcal{L}_\infty} \doteq \max_{1\leq i\leq n} \{\sup_{t\geq 0} |f_i(t)|\} < \infty$. Similarly, for any $0 < T < \infty$, we define the norm $\|f\|^{[0,T]}_{\mathcal{L}_\infty} \doteq \max_{1\leq i\leq n} \left\{ \sup_{t\in[0,T]} |f_i(t)| \right\} < \infty$, for all piecewise continuous functions $f(t) \in \mathbb{R}^n$ without a finite escape time. In the same spirit we define the space $\mathcal{L}_1$ and $\mathcal{L}_1^{[0,T]}$ of piecewise continuous integrable functions $f(t) \in \mathbb{R}^n$ with $\|f\|_{\mathcal{L}_1} = \int_0^\infty \|f(s)\| \, ds < \infty$ and $\|f\|_{\mathcal{L}_1}^{[0,T]} = \int_0^T \|f(s)\| \, ds < \infty$, respectively, where any finite-dimensional norm is admissible in the integrand.

Given two normed linear spaces F and G, we denote the space of all linear and bounded operators $F \to G$ by $\mathfrak{L}(F,G)$ equipped with the norm $\|\cdot\|_{\mathfrak{L}(F,G)}$ induced by the norms on F and G [96, Defn. A.3.9]. We also write $\mathfrak{L}(F) \doteq \mathfrak{L}(F,F)$.

1.3.2 Probability Theory

Throughout the manuscript, the triple $(\Omega, \mathfrak{F}, \mathbb{P})$ denotes by the underlying complete probability space, where Ω is the sample space, $\mathfrak{F}$ is a σ-algebra of subsets of Ω, and $\mathbb{P}$ is a probability measure on $\mathfrak{F}$. Given any random variable X (collection of sets $\mathfrak{S} \subset \mathfrak{F}$), we denote by $\sigma(X)$ ($\sigma(\mathfrak{S})$) the smallest σ-algebra generated by X ($\mathfrak{S}$). Given any two σ-algebras $\mathfrak{S}$ and $\mathfrak{R}$, we denote by $\mathfrak{S} \vee \mathfrak{R} \doteq \sigma(\mathfrak{S} \cup \mathfrak{R})$ the *join* of the σ-algebras $\mathfrak{S}$ and $\mathfrak{R}$ (the σ-algebra generated by $\mathfrak{S}$ and $\mathfrak{R}$). The product σ-algebra formed from σ-algebras $\mathfrak{S}$ and $\mathfrak{R}$ is denoted by $\mathfrak{S} \otimes \mathfrak{R} \doteq \sigma\left(S \times R \;:\; S \in \mathfrak{S}, \; R \in \mathfrak{R}\right)$.

Given the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ and a measurable space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, we define the spaces $L_p(\Omega; \mathbb{R}^n)$, $p \in [1, \infty]$ using (1) with $(\Lambda, \mathfrak{G}, \mathbb{T}) = (\Omega, \mathfrak{F}, \mathbb{P})$. Given any $f \in L_p(\Omega; \mathbb{R}^n)$, $p \in [1, \infty)$, we define the p^{th}-moment of f as $\mathbb{E}\left[\|f\|^p\right] \doteq \|f\|_{L_p}^p$, where $\mathbb{E}\left[\cdot\right]$ is the *expectation operator*. Note that unless otherwise specified, the underlying probability measure is always chosen to be $\mathbb{P}$ (the measure for the underlying probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ throughout the manuscript).

Given any Polish metric space $(\mathcal{X}, d)$ (complete and separable metric space [97, Defn. 18.1]), and any two probability measures π_1 and π_2, we define the p-Wasserstein metric, $\mathsf{p} \in \mathbb{N}$, as

$$\mathsf{W}_{\mathsf{p}}^{\mathcal{X}}\left(\pi_1, \pi_2\right) \doteq \left(\inf_{\gamma \in \Pi(\pi_1, \pi_2)} \int_{\mathcal{X}} d(x,y)^{\mathsf{p}} d\gamma(x,y) \right)^{\frac{1}{\mathsf{p}}} = \inf_{\gamma \in \Pi(\pi_1, \pi_2)} \left\{ \mathbb{E}_\gamma\left[d(x,y)^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \right\},$$

where $\Pi(\pi_1, \pi_2)$ denotes the set of all possible couplings of π_1 and π_2, see [28, Chp. 6] for further details. Using the Wasserstein metric, we can define the *ambiguity set* $\mathfrak{A}^{\mathsf{p}}(\mu, \epsilon, (\mathcal{X}, d))$ for any probability measure μ on $(\mathcal{X}, d)$ and any scalar $\epsilon > 0$ as

$$\mathsf{A}_{\mathsf{p}}(\mu, \epsilon, (\mathcal{X}, d)) \doteq \left\{ \text{Probability measure } \nu \text{ on } \mathcal{X} \;:\; \mathsf{W}_{\mathsf{p}}^{\mathcal{X}}\left(\nu, \mu\right) \leq \epsilon . \right\}$$

For any two σ-finite measure spaces $(\Lambda_i, \mathfrak{G}_i, \mathbb{T}_i)$, $i \in \{1, 2\}$, we define the product space as $(\Lambda_1 \times \Lambda_2, \mathfrak{G}_1 \times \mathfrak{G}_2, \mathbb{T}_1 \times \mathbb{T}_2)$ and where the product measure is defined as $\mathbb{T} \doteq \mathbb{T}_1 \times \mathbb{T}_2$ such that $\mathbb{T}(G_1 \times G_2) = \mathbb{T}_1(G_1)\mathbb{T}_2(G_2)$. If σ-finite measure spaces are probability spaces, then the product space is a probability space as well with the product measure as the underlying probability measure [97, Sec. 9.2, Thm. 7].

1.3.3 Stochastic Processes

We assume that the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ is equipped with a filtration $(\mathfrak{F}_t)_{t \geq 0}$, and we thus denote the underlying probability space equipped with the filtration by $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \geq 0}, \mathbb{P})$. For statements that hold almost surely with respect to the probability measure $\mathbb{P}$, we write that the statement holds $\mathbb{P}$-a.s.. For example, we say that an $\mathfrak{F}_t$-adapted stochastic process $X_t : \Omega \to \mathbb{R}^n$ is continuous $\mathbb{P}$-a.s. if it is continuous almost surely under the measure $\mathbb{P}$. See [97, Chp. 11] for definitions of filtrations and adapted processes.

Given any $T \in (0, \infty)$ and $n \in \mathbb{R}^n$, let $\mathcal{B}(\mathcal{C}([0,T]; \mathbb{R}^n))$ denote the Borel σ-algebra of open sets in $\mathcal{C}([0,T]; \mathbb{R}^n)$ generated by *cylinder sets* of the form $C = \{f \in \mathcal{C}([0,T]; \mathbb{R}^n) \;:\; f(t_i) \in A_i, \; i = 1, \ldots, k\}$, for all $k \in \mathbb{N}$, all choice of times $0 < t_1 < \cdots < t_k \leq T$, and all $A_i \in \mathcal{B}(\mathbb{R}^n)$ [59, Chp. 2]. Then, an $\mathfrak{F}_t$-adapted and continuous $\mathbb{P}$-a.s. process $X_t : \Omega \to \mathbb{R}^n$, we denote by $\mathbb{X}_{[0,T]}$ the *law of the process* X_t, $t \in [0,T]$, and is the probability measure induced by $X_{[0,T]} \doteq X_{t \in [0,T]}$, when interpreted as a random variable $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \geq 0}, \mathbb{P}) \to (\mathcal{C}([0,T]; \mathbb{R}^n), \mathcal{B}(\mathcal{C}([0,T]; \mathbb{R}^n)))$. That is, for any $A \in \mathcal{B}(\mathcal{C}([0,T]; \mathbb{R}^n))$, $\mathbb{X}_{[0,T]}(A) = \mathbb{P}X_{[0,T]}^{-1}(A)$, the push-forward measure. Similar to $\mathbb{X}_{[0,T]}$, we denote by $\mathbb{X}_{[0,T]}^x$, $x \in \mathbb{R}^n$, the law of the process $X_{t \in [0,T]}$ with $X_0 = x$ $\mathbb{P}$-a.s.. Since the laws are defined on cylinder sets, the law of the process provides us with information like $\mathbb{X}_{[0,T]}(X_{t_1} \in A_1, \cdots, X_{t_k} \in A_k)$, for any $k \in \mathbb{N}$, any choice of times $0 < t_1 < \cdots < t_k \leq T$, and any $A_i \in \mathcal{B}(\mathbb{R}^n)$.

Note that at each time $t \in [0,T]$, X_t induces a probability measure on $\mathbb{R}^n$ (push-forward measure) which we define as $\mathbb{X}_t(A) \doteq \mathbb{P}X_t^{-1}(A)$, for any $A \in \mathcal{B}(\mathbb{R}^n)$. Similarly $\mathbb{X}_t^x$, for when $X_0 = x$ $\mathbb{P}$-a.s.. Thus, we refer to $\mathbb{X}_t$ as the *temporal measure*, while we refer to $\mathbb{X}_{[0,T]}$ as the *path measure*. This allows us to distinguish our interpretation of X_t, $t \in [0,T]$, as a random variable between $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \geq 0}, \mathbb{P}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, or between $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \geq 0}, \mathbb{P}) \to$

$(\mathcal{C}([0,T];\mathbb{R}^n), \mathcal{B}(\mathcal{C}([0,T];\mathbb{R}^n)))$. Thus, we can study properties of the stochastic process either at each point in time, or over the entire temporal horizon. Furthermore, note that for any $t \in [0,T]$, we have the relation $\mathbb{X}_{[0,T]}(X_t \in A) = \mathbb{X}_t(X_t \in A), \forall A \in \mathcal{B}(\mathbb{R}^n)$ since $A \in \mathcal{B}(\mathbb{R}^n)$ is itself a cylinder set.

Throughout the manuscript, we reserve the notation W_t, and its slight variants, to denote a d-dimensional Brownian motion, for any $d \in \mathbb{N}$, that is $\mathfrak{F}_t$-adapted. Since Brownian motion is continuous $\mathbb{P}$-a.s. [59], we use the definitions above with $X_t = W_t$ and thus define $\mathbb{W}_t$, $t \in [0,T]$, as the temporal measure and $\mathbb{W}_{[0,T]}$ as the path measure (law). Note that since by definition $W_0 \equiv 0 \in \mathbb{R}^m$ $\mathbb{P}$-a.s. [95], we have $\mathbb{W}_t \doteq \mathbb{W}_t^0$ and $\mathbb{W}_{[0,T]} \doteq \mathbb{W}_{[0,T]}^0$.

For the $\mathfrak{F}_t$-adapted and d-dimensional Brownian motion W_t, $t \in [0,T]$, we denote the filtration generated by W_t by $\mathfrak{W}_t \doteq \sigma(\{W_s : s \leq t\})$ and we further define $\mathfrak{W}_{[0,T]} \doteq \sigma\left(\cup_{t\in[0,T]}\mathfrak{W}_t\right)$. Additionally, consider an $\mathfrak{F}_t$-adapted and n-dimensional continuous $\mathbb{P}$-a.s. process X_t, $t \in [0,T]$. We define $\mathfrak{X}_t \doteq \sigma(\{X_s : s \leq t\})$ and $\mathfrak{X}_{[0,T]} \doteq \sigma\left(\cup_{t\in[0,T]}\mathfrak{X}_t\right)$. If X_t is $\mathfrak{W}_t$-adapted, then $\mathfrak{X}_t \subseteq \mathfrak{W}_t$, and $\mathbb{X}^x_{[0,T]}(X_{t_1} \in A_1, \cdots, X_{t_k} \in A_k) = \mathbb{W}_{[0,T]}(X^x_{t_1} \in A_1, \cdots, X^x_{t_k} \in A_k)$, for all $k \in \mathbb{N}$, all choice of times $0 < t_1 < \cdots < t_k \leq T$, and all $A_i \in \mathcal{B}(\mathbb{R}^n)$, for any $k \in \mathbb{N}$, any choice of times $0 < t_1 < \cdots < t_k \leq T$, and any $A_i \in \mathfrak{X}_{[0,T]}$ [15, Chp. 7]. Here, X_t^x denotes the fact that $X_0 = x$ $\mathbb{P}$-a.s.. With a slight abuse of notation, we denote by $\mathbb{E}[\cdot]$ to be the expectation w.r.t. to the law $\mathbb{W}_{[0,T]}$ of the Brownian motion $W_t = W_t^0$, $t \in [0,T]$. Similarly, we denote by $\mathbb{E}_x[\cdot]$ to be the expectation w.r.t. to the law $\mathbb{X}^x_{[0,T]}$ of the process X_t^x, $t \in [0,T]$. Thus, if X_t is $\mathfrak{W}_t$-adapted, then $\mathbb{E}_x[f(X_t)] = \mathbb{E}[f(X_t^x)]$, where f is any sufficiently regular function.

1.3.4 General Notation

We define the diagonal set Δ_{2n} of the vector space $\mathbb{R}^{2n}$, $n \in \mathbb{N}$, and the point-to-set distance $|\cdot|_{\Delta_{2n}}$ as

$$\Delta_{2n} \doteq \left\{c \in \mathbb{R}^{2n} \ : \ \exists d \in \mathbb{R}^n \ : \ c = \begin{bmatrix} d^\top & d^\top \end{bmatrix}^\top \right\}, \quad |a|_{\Delta_{2n}} \doteq \inf_{z \in \Delta_{2n}} \|a - z\|.$$

For any $n, m \in \mathbb{N}$, we define 0_n and $0_{n,m}$ to be the vector and matrix of zeros in $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$, respectively. We define 1_n and $1_{n,m}$ analogously. For any $a \in \mathbb{R}$, we define the indicator function $\mathbb{1} : \mathbb{R}_{\geq 0} \to \{0,1\}$ satisfying $\mathbb{1}(a) \doteq 0$, if $a = 0$, and $\mathbb{1}(a) \doteq 1$, otherwise. We define $\mathbb{0}_n : \mathbb{R} \times \{1, \cdots, n\} \to \mathbb{R}^n$ to be a vector-valued map with all zero entries except for the j^{th} element which takes the value of its first argument.

For any $A \in \mathbb{R}^{n\times m}$, we denote by $[A]_{i,j}$ the element at the i^{th}-row and j^{th} column. Similarly, $[A]_{i,\cdot}$ and $[A]_{\cdot,j}$ denote the i^{th}-row and the j^{th} column of A, respectively. We denote the i^{th}-element of a vector $x \in \mathbb{R}^n$ as $[x]_i$.

For any continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, we denote the gradient of f by $\mathbb{R}^n \ni \partial_a f(a) \doteq [\partial_x f(x)]_{x=a} = \begin{bmatrix} \frac{\partial f(x)}{x_1} & \cdots & \frac{\partial f(x)}{x_n} \end{bmatrix}^\top_{x=a}$. Note that for $f : \mathbb{R}^n \to \mathbb{R}^n$, $\partial_a f(a) \in \mathbb{R}^{n\times n}$. We similarly define the Hessian $\partial_a^2 g(a) \in \mathbb{S}^n$ for any twice continuously differentiable $g : \mathbb{R}^n \to \mathbb{R}$.

2 Systems, Assumptions, and Problem Statement

We first begin with the definitions of the processes and stochastic systems we consider.

2.1 The Systems

We begin by defining the known and unknown functions that define the known and unknown drift and diffusion vector fields, respectively.

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

$$F_\mu(t,a,b) \doteq f(t,a) + g(t)b + \Lambda_\mu(t,a) \in \mathbb{R}^n, \quad F_\sigma(t,a) \doteq p(t,a) + \Lambda_\sigma(t,a) \in \mathbb{R}^{n\times d}, \tag{2}$$

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

$$\bar{F}_\mu(t,a) \doteq f(t,a) \in \mathbb{R}^n, \quad \bar{F}_\sigma(t,a) \doteq p(t,a) \in \mathbb{R}^{n\times d}, \tag{3}$$

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

$$F_\mu(t,a,b) = \bar{F}_\mu(t,a) + g(t)b + \Lambda_\mu(t,a) \in \mathbb{R}^n, \quad F_\sigma(t,a) = \bar{F}_\sigma(t,a) + \Lambda_\sigma(t,a) \in \mathbb{R}^{n\times d}. \tag{4}$$

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.

Definition 2 (Processes) *The notations in the following mostly follow the convention in [15]. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space that we consider as the underlying space throughout the manuscript. We denote by W_t and $W_t^\star$ any two independent $\mathbb{P}$-Brownian motions. The filtrations generated by W_t and $W_t^\star$ are denoted by $\mathfrak{W}_t$ and $\mathfrak{W}_t^\star$, respectively. We also define $\mathfrak{W}_\infty = \sigma\left(\cup_{t\geq}\mathfrak{W}_t\right)$ and $\mathfrak{W}_\infty^\star = \sigma\left(\cup_{t\geq}\mathfrak{W}_t^\star\right)$. Let $x_0 \sim \xi_0$ and $x_0^\star \sim \xi_0^\star$ be two $\mathbb{R}^n$-valued random variables that are independent of the σ-algebras $\mathfrak{W}_\infty$ and $\mathfrak{W}_\infty^\star$, respectively, where ξ_0 and $\xi_0^\star$ are the respective distributions (probability measures) on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$. Then, we define $\mathfrak{W}_{0,t} = \sigma(\xi_0) \vee \mathfrak{W}_t$ and $\mathfrak{W}_{0,t}^\star = \sigma(\xi_0^\star) \vee \mathfrak{W}_t^\star$.*

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

$$dX_t = F_\mu(t, X_t, U_t)\,dt + F_\sigma(t, X_t)\,dW_t, \quad X_0 = x_0 \sim \xi_0 \;(\mathbb{P}\text{-a.s.}), \tag{5a}$$

$$dX_t^\star = \bar{F}_\mu(t, X_t^\star)\,dt + \bar{F}_\sigma(t, X_t^\star)\,dW_t^\star, \quad X_0^\star = x_0^\star \sim \xi_0^\star \;(\mathbb{P}\text{-a.s.}), \tag{5b}$$

*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**. We will drop the quantifier $\mathbb{P}$-a.s. from here on unless it is not clear from context.*

*We denote $\mathbb{X}_{[0,T]}$ and $\mathbb{X}_{[0,T]}^\star$ the **true and nominal laws (path/trajectory probability measures)** induced respectively by the process paths $X_{[0,T]}$ and $X_{[0,T]}^\star$ on the σ-algebra $\mathcal{B}(\mathcal{C}([0,T];\mathbb{R}^n))$, i.e., $X_{[0,T]} \sim \mathbb{X}_{[0,T]}$ and $X_{[0,T]}^\star \sim \mathbb{X}_{[0,T]}^\star$.*

Remark 2.1 We do not assume the uniqueness and existence of the strong solutions X_t and $X_t^\star$ of (5a) and (5b), respectively. The well-posedness of (5b) is straightforward to establish under general conditions, see e.g. [59, Definition 5.2.1], [14, Sec. 4.5], and [15, Sec. 5.2]. However, the well-posedness of (5b) is more challenging prospect. The primary reason is that the true (uncertain) process X_t depends on the feedback process U_t, and thus we will have to establish well-posedness once we define U_t.

The separability of the Banach space $\mathcal{C}([0,T];\mathbb{R}^n)$ implies that $\mathcal{B}(\mathcal{C}([0,T];\mathbb{R}^n)) = \sigma(\mathfrak{C}([0,T],\mathbb{R}^n))$ where

$$\mathfrak{C}([0,T],\mathbb{R}^n) \ni C(k) \doteq \{h \in \mathcal{C}([0,T];\mathbb{R}^n) \mid h(t_1) \in B_1, \ldots, h(t_k) \in B_k\}, \quad \forall k \in \mathbb{N}, \tag{6}$$

for all choice of temporal instances $0 < t_1 < \ldots < t_k \leq T$, and all $B_1, \ldots, B_k \in \mathcal{B}(\mathbb{R}^n)$ [98, Sec.2.1, p. 12]. Along with the separability of strong solutions [99, Thm. 5.2.1], the Kolmogorov extension theorem [59, Thm. 2.2] implies that the finite-dimensional cylinder sets form a determining class for the probability measures on $\mathcal{B}(\mathcal{C}([0,T];\mathbb{R}^n))$ [98, Thm 2.0]. Consequently, it suffices to define the laws $\mathbb{X}_{[0,T]}$ and $\mathbb{X}_{[0,T]}^\star$ on the cylinder sets as done by the authors in [15] and [14]. Using the laws as probability measures on the finite-dimensional cylinder sets, we define the distance between the laws $\mathbb{X}_{[0,T]}$ and $\mathbb{X}_{[0,T]}^\star$ using the Wasserstein metric in the following definition.

Definition 3 *For any $C(k) \in \mathfrak{C}([0,T],\mathbb{R}^n)$, $k \in \mathbb{N}$, as in (6), the action of the laws $\mathbb{X}_{[0,T]}$ and $\mathbb{X}_{[0,T]}^\star$ on the finite-dimensional cylinder set $C(k)$ is defined by the respective finite-dimensional distributions $\mathbb{X}_{t_1\cdots t_k}$ and $\mathbb{X}_{t_1\cdots t_k}^\star$ as follows:*

$$\mathbb{X}_{[0,T]}(C(k)) \doteq \mathbb{X}_{t_1\cdots t_k}(B_1 \times \cdots \times B_k) = \mathbb{P}\left[X_{t_1} \in B_1, \cdots, X_{t_k} \in B_k\right]. \tag{7}$$

The map $\mathbb{X}_{t_1\cdots t_k}^\star : C(k) \to [0,1]$ is defined analogously.

For each $k \in \mathbb{N}$, the finite-dimensional laws $\mathbb{X}_{t_1\cdots t_k}$ and $\mathbb{X}_{t_1\cdots t_k}^\star$ are probability measures on the Polish space $\mathbb{R}^{nk}$ equipped with the metric induced by the Euclidean norm $\|\cdot\|$ on $\mathbb{R}^{kn}$. Therefore, we define the distance $\mathfrak{D}_\mathsf{p}$ between the laws $\mathbb{X}_{[0,T]}$ and $\mathbb{X}_{[0,T]}^\star$ using the Wasserstein metric $\mathsf{W}_\mathsf{p}^{nk}\left(\mathbb{X}_{t_1\cdots t_k}, \mathbb{X}_{t_1\cdots t_k}^\star\right)$ as follows:

$$\mathfrak{D}_\mathsf{p}\left(\mathbb{X}_{[0,T]}, \mathbb{X}_{[0,T]}^\star\right) \doteq \sup_{k\in\mathbb{N}} \mathsf{W}_\mathsf{p}^{nk}\left(\mathbb{X}_{t_1\cdots t_k}, \mathbb{X}_{t_1\cdots t_k}^\star\right), \quad 0 < t_1 < \ldots < t_k \leq T, \quad \mathsf{p} \in \mathbb{N}. \tag{8}$$

Remark 2.2 The distance $\mathfrak{D}_\mathsf{p}\left(\mathbb{X}_{[0,T]}, \mathbb{X}_{[0,T]}^\star\right)$ in (8) is the supremum of the set of Wasserstein metrics between all finite-dimensional distributions $\mathbb{X}_{t_1\cdots t_k}$ and $\mathbb{X}_{t_1\cdots t_k}^\star$ corresponding to the laws $\mathbb{X}_{[0,T]}$ and $\mathbb{X}_{[0,T]}^\star$, respectively. It is not, however, a metric between the laws themselves since $\mathbb{X}_{[0,T]}$ and $\mathbb{X}_{[0,T]}^\star$ are probability measures on the infinite-dimensional space $\mathcal{C}([0,T];\mathbb{R}^n)$. As we have previously discussed, the consistent finite-dimensional distributions can be extended to measures on $\mathcal{C}([0,T];\mathbb{R}^n)$ using the Kolmogorov extension theorem [98, Thm 2.0]. For example, such a procedure is used for a construction of the Brownian motion [15, Sec. 2.2]. However, it is not clear to the authors if a similar extension can be performed for the Wasserstein metric from $\mathbb{R}^{nk}$ to $\mathcal{C}([0,T];\mathbb{R}^n)$ equipped with the metric induced by the uniform norm. This a topic of future research.

Remark 2.3 From the perspective of control and planning, the distance $\mathfrak{D}_{\mathsf{p}}\left(\mathbb{X}_{[0,T]}, \mathbb{X}^\star_{[0,T]}\right)$ in (8) provides a bound on the discrepancy between the true and nominal laws $\mathbb{X}_{[0,T]}$ and $\mathbb{X}^\star_{[0,T]}$ over all finite-dimensional cylinder set which has appealing operational interpretations. As an example, the distance in (8) can be used to *a priori* compute the success probability of an uncertain system tracking a nominal trajectory while navigating an obstacle course. The *a priori* quantification of such probabilities can then be used to inform the upstream task of nominal trajectory design that satisfies the desired (probabilistic) performance guarantees.

2.2 Assumptions

We now state the assumptions we place on the systems we defined in Sec. 2.1 and discuss their verifiability and consequences. We begin with the assumptions for the nominal (known) system in (5b).

Assumption 1 (Nominal (Known) System) *The known functions $f(t,a)$ and $p(t,a)$ in Definition 1 are Lipschitz continuous, locally in $a \in \mathbb{R}^n$ and uniformly in $t \in \mathbb{R}_{\geq 0}$, and there exist known $\Delta_f, \Delta_p \in \mathbb{R}_{>0}$ such that*

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

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

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

where $\Delta_g, \Delta_{\dot g} \in \mathbb{R}_{>0}$ are assumed to be known.

Additionally, since $g(t)$ is full rank, we can construct a $g^\perp : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n \times n-m}$ such that $\operatorname{Im} g(t)^\perp = \ker g(t)^\top$ and $\operatorname{rank}\left(\begin{bmatrix} g(t) & g(t)^\perp \end{bmatrix}\right) = n$, $\forall t \in \mathbb{R}_{\geq 0}$. 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 $\left\|g(t)^\perp\right\|_F \leq \Delta_g^\perp$, $\forall t \in \mathbb{R}_{\geq 0}$, where $\Delta_g^\perp \in \mathbb{R}_{>0}$ is assumed to be known.

Remark 2.4 The condition on local, instead of global Lipschitz continuity expands the class of systems we can consider. However, the local Lipschitz continuity along with the linear growth condition requires an analysis like the Khasminskii-type theorem [100, Thm. 3.2] to establish the existence and uniqueness of strong solutions. If instead, one assumes global Lipschitz continuity, then standard results on well-posedness can be used, see e.g., [15, Thm. 5.2.1] and [100, Thm. 2.31]. We assume the uniform boundedness of the known diffusion $p(t,x)$, instead of linear growth, since it is a sufficient condition to transfer the stability of the deterministic counterpart of (5b) to the complete stochastic system via robustness arguments, see e.g., [37, 53, 54]. 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 (5a). Therefore, the stability of (5a) is not guaranteed by the stability of its deterministic counterpart. The requirement that the initial condition $\xi_0 \in L_{2\mathsf{p}}$, for some $\mathsf{p} \in \mathbb{N}$ implies that, at the minimum, $\xi_0 \in L_2$ which is a standard requirement for SDEs, see e.g. [Thm. 5.2.1][15].

We now place assumptions on the existence of certificates of stability that render the nominal (known) system in (5b), and its deterministic counterpart ($W_t^\star \equiv 0$, $\forall t$) stable, if well-posed. We will subsequently discuss the notions of stability that the following assumption endows upon (5b).

Assumption 2 (Nominal (Known) System Stability) *Assume there exist known $\alpha_1, \alpha_2, \lambda \in \mathbb{R}_{>0}$ and a function $V \in \mathcal{C}^3(\mathbb{R}^n \times \mathbb{R}^n; \mathbb{R})$ such that*

$$\alpha_1 \|a-b\|^2 \leq V(a,b) \leq \alpha_2 \|a-b\|^2, \quad V_a(a,b)^\top \bar F_\mu(t,a) + V_b(a,b)^\top \bar F_\mu(t,b) \leq -2\lambda V(a,b), \tag{9}$$

for all $a, b \in \mathbb{R}^n$ and $t \in \mathbb{R}_{\geq 0}$, where $\bar F_\mu$ is defined in (3) and where $V_a(a,b) \doteq \partial_a V(a,b)$ and $V_b(a,b) \doteq \partial_b V(a,b)$.

Furthermore, assume there exist known $\Delta_{\partial V}, \Delta_{\partial^2 V} \in \mathbb{R}_{>0}$ such that

$$\sum_{i=1}^n \left(|V_{a_i}(a,b) - V_{a_i}(a',b')|^2 + |V_{b_i}(a,b) - V_{b_i}(a',b')|^2\right) \leq \Delta_{\partial V}^2 \left|\begin{bmatrix} a \\ b\end{bmatrix} - \begin{bmatrix} a' \\ b'\end{bmatrix}\right|_{\Delta_{2n}}^2, \tag{10a}$$

$$\sum_{i=1}^n\sum_{j=1}^n \left(\left|V_{a_i,a_j}(a,b) - V_{a_i,a_j}(a',b')\right|^2 + \left|V_{b_i,b_j}(a,b) - V_{b_i,b_j}(a',b')\right|^2 + 2\left|V_{a_i,b_j}(a,b) - V_{a_i,b_j}(a',b')\right|^2\right)$$
$$\leq \Delta_{\partial^2 V}^2 \left|\begin{bmatrix} a \\ b\end{bmatrix} - \begin{bmatrix} a' \\ b'\end{bmatrix}\right|_{\Delta_{2n}}^2, \tag{10b}$$

*for all $a, b, a', b' \in \mathbb{R}^n$, and where the point-to-set distance $|\cdot|_{\Delta_{2n}}$ is defined in Sec. 1.3. We refer to $V(\cdot,\cdot)$ as an **incremental Lyapunov function (ILF)**.*

We refer to $V(\cdot,\cdot)$ as an incremental Lyapunov function (ILF) because it is a certificate for the incremental exponential stability (IES) of $\dot{x} = \bar{F}_\mu(t,x)$ (the deterministic counterpart of the nominal (known) system (5b)), see e.g., [53, Defn. 2.2], [11, Defin. 3.3], [101, Defn. 1]. While the sufficiency of (9) for the IES of $\dot{x} = \bar{F}_\mu(t,x)$ is straightforward, its necessity holds only over compact subsets of the state space [49]. However, the authors in [101] showed that the existence of the ILF V satisfying (9), which is sufficient for IES of $\dot{x} = \bar{F}_\mu(t,x)$, is equivalent to the existence of a control contraction metric for the IES system. Thus, the class of ILFs we consider is general enough to represent certificates for IES of deterministic systems. The use of control contraction metrics (CCMs) [51] directly is a subject for future investigation.

Remark 2.5 In Lyapunov function-based analysis for stochastic systems, it is common to assume the Lipschitz continuity of the gradient of the Lyapunov function, see e.g. [38] or [55] for Lipschitz growth on the gradient of the contraction metric. In a similar vein, the condition in (10a) establishes the Lipschitz continuity of the gradient of ILF V with respect to the point-to-set distance $|\cdot|_{\Delta_{2n}}$. We choose the point-to-set distance $|\cdot|_{\Delta_{2n}}$ since the IES stability of a deterministic system is equivalent to the exponential stability of its auxiliary version with respect to the diagonal set $\Delta_{2n} \subset \mathbb{R}^{2n}$ [49, 101]. Note that due to [49, Lem. 2.3], (10a) indeed implies Lipschitz continuity with respect to the standard Euclidean norm. The Lipschitz continuity of the Hessian in (10b) is an additional condition we require. We highlight the fact that the results in the manuscript hold with $|\cdot|_{\Delta_{2n}}$ replaced by the Euclidean norm $\|\cdot\|$, albeit with an amount of additional conservatism.

The next assumption is on the statistical nature of the nominal (known) Itô SDE.

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 (5b) satisfies*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X_t^\star\|^{2\mathsf{p}^\star}\right]^{\frac{1}{2\mathsf{p}^\star}} \leq \Delta_\star,$$

for some $\mathsf{p}^\star \in \mathbb{N}_{\geq 1}$

Remark 2.6 Since we aim to establish the stability of the true (uncertain) process (5a) with respect to the nominal (known) process (5b), 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 conclude the section by stating the assumptions that we place on the true (uncertain) system in (5a).

Assumption 4 (True (Uncertain) System) *Consider the unknown functions Λ_μ and Λ_σ in Definition 1, and define $\Lambda_\mu^{\parallel} : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^m$, $\Lambda_\mu^{\perp} : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^{n-m}$, $\Lambda_\sigma^{\parallel} : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^{m\times d}$, and $\Lambda_\sigma^{\perp} : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^{n-m\times d}$ as*

$$\begin{bmatrix}\Lambda_\mu^{\parallel}(t,a)\\ \Lambda_\mu^{\perp}(t,a)\end{bmatrix} = \begin{bmatrix} g(t) & g(t)^{\perp}\end{bmatrix}^{-1}\Lambda_\mu(t,a), \quad \begin{bmatrix}\Lambda_\sigma^{\parallel}(t,a)\\ \Lambda_\sigma^{\perp}(t,a)\end{bmatrix} = \begin{bmatrix} g(t) & g(t)^{\perp}\end{bmatrix}^{-1}\Lambda_\sigma(t,a), \ \forall (t,a) \in \mathbb{R}_{\geq 0}\times\mathbb{R}^n, \quad (11)$$

where the input operator g is presented in Definition 1, and $g^\perp$ is defined in Assumption 1.

We assume that $\Lambda_\mu^{\parallel}(t,a)$, $\Lambda_\mu^{\perp}(t,a)$, $\Lambda_\sigma^{\parallel}(t,a)$, and $\Lambda_\sigma^{\perp}(t,a)$ are Lipschitz continuous, locally in $a \in \mathbb{R}^n$ and uniformly in $t \in \mathbb{R}_{\geq 0}$, and there exist known $\Delta_\mu^{\parallel}$, $\Delta_\mu^{\perp}$, $\Delta_\sigma^{\parallel}$, $\Delta_\sigma^{\perp} \in \mathbb{R}_{>0}$ such that

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

To avoid burdensome notation, and as an implication of the above, we assume that there exist known $\Delta_\mu, \Delta_\sigma \in \mathbb{R}_{>0}$ such that

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

Finally, there exists a known $\Delta_0 \in \mathbb{R}_{>}0$ such that $\|x_0\|_{L_{2\mathsf{p}}} \leq \Delta_0$, where x_0 is presented in Definition 2, and $\mathsf{p} \in \mathbb{N}$ is introduced in Assumption 1.

Remark 2.7 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 (5b) cannot be extended to the uncertain (true) system in (5a) via robustness arguments as done in [54, 53]. The $\mathcal{L}_1$-DRAC control can, however, accommodate growing uncertainties under certain conditions as we shall later see.

Similar to the decomposition of Λ^μ and Λ^σ in Assumption 4, we can decompose the known diffusion term p which we present below.

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

$$\begin{bmatrix} p^{\|}(t,a) \\ p^{\perp}(t,a) \end{bmatrix} = \begin{bmatrix} g(t) & g(t)^{\perp} \end{bmatrix}^{-1} p(t,a), \; \forall (t,a) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^n, \tag{12}$$

where, we assume that $p^{\|}$ 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 of p in Assumption 1, we assume that there exist known $\Delta_p^{\|}$, $\Delta_p^{\perp} \in \mathbb{R}_{>0}$ such that

$$\left\| p^{\|}(t,a) \right\|_F \leq \Delta_p^{\|}, \quad \left\| p^{\perp}(t,a) \right\|_F \leq \Delta_p^{\perp}, \; \forall (t,a) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^n.$$

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

Definition 4 *We define functions $F_\sigma^{\|} : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^{m \times d}$ and $F_\sigma^{\perp} : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^{n-m \times d}$ as follows:*

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

The final assumption we place is related to the stability of the deterministic counterpart of the true (uncertain) system in (5a).

Assumption 6 *The constants λ and α_1 in (9), and the bound $\Delta_{\partial V}$ in (10a), introduced in Assumption 2 satisfy*

$$\lambda > \frac{1}{\alpha_1} \left(\frac{1}{2} \right)^{\frac{1}{4}} \Delta_g^{\perp} \Delta_\mu^{\perp} \Delta_{\partial V},$$

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

Remark 2.8 The condition in Assumption 6 ensures that the true (uncertain) system in (5a) in the absence of diffusion terms ($F_\sigma \equiv 0$) and the matched drift uncertainty $\Lambda_\mu^{\|} \equiv 0$, i.e.,deterministic counterpart of the true (uncertain) system with only the unmatched uncertainty $\Lambda_\mu^{\perp}$, maintains its incremental exponential stability in the presence of $\Lambda_\mu^{\perp}$. 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.

2.3 Problem Statement

We now state the problem we aim to solve in this manuscript.

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

$$dX_t = F_\mu(t, X_t, U_{\mathcal{L}_1,t})\, dt + F_\sigma(t, X_t)\, dW_t, \quad X_0 = x_0 \sim \xi_0, \forall t \in [0,T], \tag{13}$$

*that is induced by closing the feedback loop on the true (uncertain) Itô SDE in (5a) with the **feedback process** $U_t = U_{\mathcal{L}_1,t}$ defined as*

$$U_{\mathcal{L}_1,t} \doteq \mathcal{F}_{\mathcal{L}_1}(X)(t), \quad \mathcal{F}_{\mathcal{L}_1} : \mathcal{C}([0,T];\mathbb{R}^n) \to \mathcal{C}([0,T];\mathbb{R}^m), \quad \forall t \in [0,T], \tag{14}$$

*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 exists and is a unique strong solution of the closed-loop true (uncertain) Itô SDE in (13), for any $T \in (0, \infty)$.*

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

$$\mathbb{X}_{[0,T]} \in \Upsilon\left(\mathbb{X}^\star_{[0,T]}, \rho\right) \doteq \left\{ \textit{Probability measures } \nu \textit{ on } \mathcal{B}(\mathcal{C}([0,T];\mathbb{R}^n)) \mid \mathfrak{D}_{\mathsf{p}}\left(\mathbb{X}^\star_{[0,T]}, \nu\right) \leq \rho \right\}, \quad \forall T \in (0,\infty), \tag{15}$$

*where $\mathbb{X}_T$ and $\mathbb{X}^\star_T$ are the true and nominal laws of the processes (13) and (5b), respectively, as presented in Definition 2. We refer to $\Upsilon(\mathbb{X}^\star_T, \rho)$ as the **uniform ambiguity set of laws on** $\mathcal{C}([0,T];\mathbb{R}^n)$.*

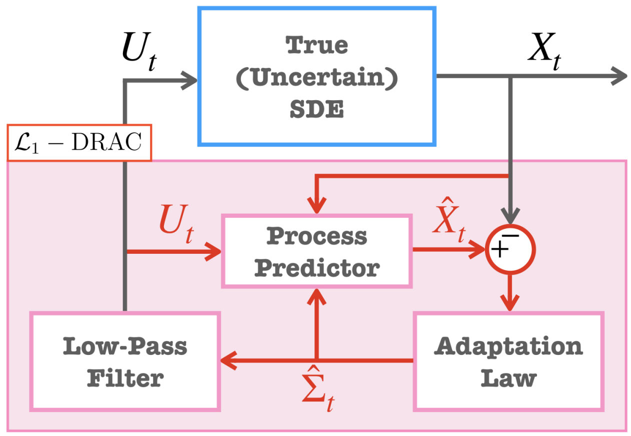

Figure 2: The architecture of the $\mathcal{L}_1$-DRAC controller. The controller has three components: a *process predictor* with output $\hat{X}_t$, an *adaptation law* driven by the prediction error $\hat{X}_t - X_t$, and a *low pass filter* that accepts the adaptive estimate $\hat{\Sigma}_t$ as its input to produce the feedback process U_t.

Condition 3: *Furthermore, there exists an a priori known $\mathbb{R}_{>0} \ni \bar{\rho}(\bar{T}) < \rho$, for any $\bar{T} \in (0, T)$, such that the true law $\mathbb{X}_{[\bar{T},T]}$ satisfies*

$$\mathbb{X}_{[\bar{T},T]} \in \bar{\Upsilon}\left(\mathbb{X}^{\star}_{[\bar{T},T]}, \bar{\rho}\right) \doteq \left\{\text{Probability measures } \nu \text{ on } \mathcal{B}\left(\mathcal{C}\left(\left[\bar{T},T\right];\mathbb{R}^n\right)\right) \mid \mathsf{W}_2\left(\mathbb{X}^{\star}_{[\bar{T},T]}, \nu\right) \leq \bar{\rho}\right\}. \tag{16}$$

*We refer to $\bar{\Upsilon}\left(\mathbb{X}^{\star}_{[\bar{T},T]}, \bar{\rho}\right)$ as the **uniformly ultimate ambiguity set of laws on** $\mathcal{C}\left(\left[\bar{T},T\right];\mathbb{R}^n\right)$.*

Recall that the σ-algebra $\mathcal{B}(\mathcal{C}([0,T];\mathbb{R}^n))$ is generated by the cylinder sets of the form in (6). Consequently, evaluating the conditions in (15) and (16) for the cylinder $\{t, B\}$, for any $(t, B) \in [0,T] \times \mathcal{B}(\mathbb{R}^n)$, implies the existence of analogous pointwise in time ambiguity sets on $\mathcal{B}(\mathbb{R}^n)$. The conclusion is hardly surprising since the uniform conditions in (15) and (16) imply the pointwise conditions.

3 $\mathcal{L}_1$-DRAC Control

In this section we define and analyze the closed loop true (uncertain) system in (13) with the $\mathcal{L}_1$-DRAC feedback control. We begin with the the definition of $\mathcal{L}_1$-DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1}$ that we introduced in (14).

3.1 Architecture and Definition

The design of $\mathcal{F}_{\mathcal{L}_1}$ is based on the $\mathcal{L}_1$-adaptive control methodology [29]. 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 (13) satisfies the conditions we set forth in Sec. 2.3. Following the architecture of $\mathcal{L}_1$-adaptive control [29], 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. 2.

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

$$\mathcal{F}_{\mathcal{L}_1}(y) \doteq \mathcal{F}_{\boldsymbol{\omega}} \circ \mathcal{F}_{\boldsymbol{T_s}} \circ \mathcal{F}_{\lambda_s}(y), \quad y \in \mathcal{C}([0,T] : \mathbb{R}^n), \tag{17}$$

where

$$\mathcal{F}_{\boldsymbol{\omega}}\left(\hat{\Lambda}^{\parallel}\right)(t) \doteq -\boldsymbol{\omega} \int_0^t e^{-\boldsymbol{\omega}(t-\nu)} \hat{\Lambda}^{\parallel}(\nu)\, d\nu, \quad \textit{(Low-pass filter)} \tag{18a}$$

$$\begin{aligned}
\hat{\Lambda}^{\parallel}(t) &= \mathcal{F}^{\parallel}_{\boldsymbol{T_s}}\left(\hat{\Lambda}\right)(t) = \sum_{i=0}^{\lfloor \frac{t}{\boldsymbol{T_s}} \rfloor} \Theta_{ad}(i\boldsymbol{T_s}) \hat{\Lambda}(t) \mathbb{1}_{\{[i\boldsymbol{T_s}, (i+1)\boldsymbol{T_s})\}}(t), \\
\hat{\Lambda}(t) &= \mathcal{F}_{\boldsymbol{T_s}}(\hat{y}, y)(t) = 0_n \mathbb{1}_{\{[0,\boldsymbol{T_s})\}}(t) \qquad \textit{(Adaptation Law)} \\
&\quad + \lambda_s \left(1 - e^{\lambda_s \boldsymbol{T_s}}\right)^{-1} \sum_{i=1}^{\lfloor \frac{t}{\boldsymbol{T_s}} \rfloor} \left(\hat{X}_{i\boldsymbol{T_s}} - X_{i\boldsymbol{T_s}}\right) \mathbb{1}_{\{[i\boldsymbol{T_s}, (i+1)\boldsymbol{T_s})\}}(t),
\end{aligned} \tag{18b}$$

$$\hat{y}(t) = \mathcal{F}_{\lambda_s}(y)(t) \Rightarrow \text{ solution to the integral equation:}$$
$$\hat{y}(t) = \int_0^t \left(-\lambda_s \mathbb{I}_n (\hat{y}(\nu) - y(\nu)) + f(\nu, y(\nu)) + g(\nu)\mathcal{F}_{\mathcal{L}_1}(y)(\nu) + \hat{\Lambda}(\nu)\right) d\nu, \quad \textit{(Process Predictor)} \tag{18c}$$

for $t \in [0, T]$, where $\omega, T_s \in \mathbb{R}_{>0}$ are the control parameters and are referred to as the ***filter bandwidth*** and the ***sampling period***, respectively. Additionally, $\Theta_{ad}(t) = \begin{bmatrix}\mathbb{I}_m & 0_{m,n-m}\end{bmatrix} \bar{g}(t)^{-1} \in \mathbb{R}^{m \times n}$, where $\bar{g}(t) = \begin{bmatrix}g(t) & g(t)^\perp\end{bmatrix} \in \mathbb{R}^{n \times n}$, and here $g^\perp$ is defined in Assumption 4.

We perform the analysis to obtain the performance bounds for the $\mathcal{L}_1$-DRAC true (uncertain) Itô SDE in (13) in two steps. First, in Sec. 3.2 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. 3.3 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.

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 (18a) and the sampling period for the adaptation law in (18b), respectively. Suppose that the assumptions in Sec. 2.2 hold. Then, for arbitrarily chosen $\kappa_{r_i} \in \mathbb{R}_{>0}$, $i \in \{1, \ldots, 5\}$, $\epsilon_r \in \mathbb{R}_{>0}$, and also $\kappa_{r_\perp}$, $\alpha_r \in \mathbb{R}_{>0}$ that are chosen to satisfy

$$\kappa_{r_\perp} \in \left(0, 2\alpha_1 - \frac{2}{\lambda}\Delta_\perp^r\right), \quad \alpha_r \in \left(0, \alpha_1 - \frac{1}{\lambda}\Delta_\perp^r - \frac{1}{2}\kappa_{r_\perp}\right), \tag{19}$$

define

$$\rho_r \doteq \frac{1}{2\alpha_r}\Delta_\bullet^r(\kappa_r) + \frac{1}{2\alpha_r}\left(\Delta_\bullet^r(\kappa_r)^2 + 4\alpha_r\left(\alpha_2 \mathsf{W}_2(\xi_0, \xi_0^\star)^2 + \Delta_\circ^r(\kappa_r) + \epsilon_r\right)\right)^{\frac{1}{2}}, \tag{20a}$$
$$\rho \doteq, \tag{20b}$$

where the constants $\Delta_\circ^r$ and $\Delta_\bullet^r$ are defined in (**??**) and constitute of $\kappa_{r_i} \in \mathbb{R}_{>0}$, $i \in \{1, \ldots, 5\}$, and the bounds in Assumptions 1-5. Furthermore, $\alpha_2 \in \mathbb{R}_{>0}$ is defined in Assumption 2, and $\mathsf{W}_2(\xi_0, \xi_0^\star)$ is the 2-Wasserstein distance between the initial distributions ξ_0 and $\xi_0^\star$ of the true (uncertain) process and the known (nominal) process in (5a) and (5b), respectively.

Remark 3.1 The existence of a $\kappa_{r_\perp}$, and thus an α_r, that are feasible as per their respective open intervals in (19) is guaranteed as a consequence of Assumption 6, and the definition of $\Delta_\perp^r \in \mathbb{R}_{>0}$ in (**??**).

The filter bandwidth $\omega \in \mathbb{R}_{>0}$ is chosen such that the following conditions are satisfied:

$$\alpha_1 - \frac{1}{\lambda}\Delta_\perp^r - \frac{\kappa_{r_\perp}}{2} - \alpha_r \geq \Theta_{\mu_3}^r(\omega, \kappa_{r_2}) + \Theta_{\sigma_3}^r(\omega, \kappa_{r_5}), \tag{21a}$$
$$\alpha_r\rho_r^2 - \Delta_\bullet^r(\kappa_r)\rho_r - \alpha_2 \mathsf{W}_2(\xi_0, \xi_0^\star)^2 - \Delta_\circ^r(\kappa_r) > \Theta_{\mu_1}^r(\omega) + \Theta_{\sigma_1}^r(\omega) + \left(\Theta_{\mu_2}^r(\omega, \kappa_{r_1}) + \Theta_{\sigma_2}^r(\omega, \kappa_{r_4})\right)\rho_r, \tag{21b}$$

where, the functions $\Theta_{\{\mu_i, \sigma_i\}}^r$, $i \in \{1, 2, 3\}$, are defined in (**??**)- (**??**).

Remark 3.2 It is evident by their respective definitions in (**??**) and (**??**) that $\left\{\Theta_{\mu_3}^r(\omega, \kappa_{r_2}), \Theta_{\sigma_3}^r(\omega, \kappa_{r_5})\right\} \in \mathcal{O}\left(\frac{1}{\omega}\right)$, for any fixed values of $\kappa_{r_2}, \kappa_{r_5} \in \mathbb{R}_{>0}$. Furthermore, the choice of $\kappa_{r_\perp}$ and α_r in (19) implies that

$$\alpha_1 - \frac{1}{\lambda}\Delta_\perp^r - \frac{\kappa_{r_\perp}}{2} - \alpha_r > 0.$$

Hence, one can always choose an $\omega \in \mathbb{R}_{>0}$ such that the condition in (21a) is satisfied.

Additionally, the definition of ρ_r in (20a) implies that it is the positive root of the quadratic equation $\alpha_r\rho_r^2 - \Delta_\bullet^r(\kappa_r)\rho_r - \alpha_2 \mathsf{W}_2(\xi_0, \xi_0^\star)^2 - \Delta_\circ^r(\kappa_r) - \epsilon_r = 0$. It is trivial to verify the positivity of the discriminant of this quadratic equation, and that it admits a negative and a positive root, the latter of which is ρ_r. Since $\epsilon_r > 0$, the fact that $\alpha_r\rho_r^2 - \Delta_\bullet^r(\kappa_r)\rho_r - \alpha_2 \mathsf{W}_2(\xi_0, \xi_0^\star)^2 - \Delta_\circ^r(\kappa_r) - \epsilon_r = 0$ leads us to the conclusion that

$$\alpha_r\rho_r^2 - \Delta_\bullet^r(\kappa_r)\rho_r - \alpha_2 \mathsf{W}_2(\xi_0, \xi_0^\star)^2 - \Delta_\circ^r(\kappa_r) > 0.$$

Therefore, we can always choose an $\omega \in \mathbb{R}_{>0}$ that, in addition to (21a), also satisfies (21b) since $\left\{\Theta_{\mu_1}^r(\omega), \Theta_{\sigma_1}^r(\omega), \Theta_{\mu_2}^r(\omega, \kappa_{r_1}), \Theta_{\sigma_2}^r(\omega, \kappa_{r_4})\right\} \in \mathcal{O}\left(\frac{1}{\sqrt{\omega}}\right)$, for any fixed values of $\kappa_{r_1}, \kappa_{r_4} \in \mathbb{R}_{>0}$, as evidenced by their respective definitions in (**??**) and (**??**).

While the choice of $\kappa_{r_i} \in \mathbb{R}_{>0}$, $i \in \{1,\ldots,5\}$, and $\epsilon_r \in \mathbb{R}_{>0}$ is permitted to be arbitrary, any particular choice has consequences on the tradeoff between the performance of the $\mathcal{L}_1$-DRAC true (uncertain) process and the robustness in terms of aggressiveness of the input $U_{\mathcal{L}_1,\cdot}$. We discuss such consequences in Sec. 4, along with the fundamental limitations of the closed loop processes, further considerations regarding the effects of the uncertain vector fields on the performance, and the interpretation of the performance bounds of $\mathcal{L}_1$-DRAC as a generalization of the $\mathcal{L}_1$ adaptive control's performance for deterministic systems, e.g. in [102, 103].

3.2 Performance Analysis: Reference Process

We begin with the definition of the reference process.

Definition 5 (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**:*

$$dX_t^r = F_\mu\left(t, X_t^r, U_t^r\right)dt + F_\sigma\left(t, X_t^r\right)dW_t, \quad X_0^r = x_0 \sim \xi_0,\ \forall t \in [0,T], \tag{22}$$

*where the initial condition $\xi_0 \sim \mathbb{X}_0$ and the driving Brownian motion W_t are identical to those for the true (uncertain) process in (13) (and (5a)). The **reference feedback process** U_t^r is defined via the **reference feedback operator** $\mathcal{F}_r$ as follows:*

$$U_t^r = \mathcal{F}_r\left(X^r\right)(t). \quad \mathcal{F}_r\left(X^r\right) \doteq \mathcal{F}_{\boldsymbol{\omega}}\left(\Lambda_\mu^{\parallel}\left(\cdot, X^r\right)\right) + \mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}\left(F_\sigma^{\parallel}\left(\cdot, X^r\right), W\right), \quad t \in [0,T], \tag{23}$$

where the operator $\mathcal{F}_{\boldsymbol{\omega}}$ is the low-pass filter defined in (18a), the vector field $F_\sigma^{\parallel}$ is introduced in Definition 4, and

$$\mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}\left(F_\sigma^{\parallel}\left(\cdot, X^r\right), W\right)(t) = \mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}\left(p^{\parallel}\left(\cdot, X^r\right) + \Lambda_\sigma^{\parallel}\left(\cdot, X^r\right), W\right)(t) \doteq -\boldsymbol{\omega}\int_0^t e^{-\boldsymbol{\omega}(t-\nu)} F_\sigma^{\parallel}\left(\nu, X_\nu^r\right)dW_\nu. \tag{24}$$

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

*Finally, similar to $\mathbb{X}_{[0,T]}$ and $\mathbb{X}^\star_{[0,T]}$ in Definition 2, we denote by $\mathbb{X}^r_{[0,T]}$ the **reference law (path/trajectory probability measures)** induced by the process path $X^r_{[0,T]}$ on $\mathcal{B}\left(\mathcal{C}\left([0,T];\mathbb{R}^n\right)\right)$, i.e., $X^r_{[0,T]} \sim \mathbb{X}^r_{[0,T]}$.*

Remark 3.3 The reference process is obtained by closing the loop of the true (uncertain) process in (5a) with the feedback process U_t^r that is composed of the filtered matched drift uncertainty $\mathcal{F}_{\boldsymbol{\omega}}\Lambda_\mu^{\parallel}\left(\cdot, X^r\right)t$ and the filtered *totality* of the matched diffusion vector field $\mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}p^{\parallel}\left(\cdot, X^r\right) + \Lambda_\sigma^{\parallel}\left(\cdot, X^r\right), Wt$ 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 $\boldsymbol{\omega}$, how the system operates under the non-realizable assumption of perfect knowledge of the uncertainties.

Remark 3.4 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 $p\left(\cdot, X^r\right)$ in its definition. Therefore, the reference feedback process U_t^r further attempts to remove the effects of the known diffusion term $p\left(t, X_t^r\right)$ 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 $p\left(t, X_t^r\right)dW_t$, one cannot in general disambiguate the effects of the known term $p\left(t, X_t^r\right)$ from the total uncertainty $\int_0^t \Lambda_\mu\left(\nu, X_\nu^r\right)d\nu + \int_0^t \left(p\left(\nu, X_\nu^r\right) + \Lambda_\sigma\left(\nu, X_\nu^r\right)\right)dW_\nu$. Importantly, the design of the reference feedback process U_t^r with the included known diffusion term $p\left(t, X_t^r\right)$ leads to the subsequent $\mathcal{L}_1$-DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1}$ with important desirable implications that we will discuss later.

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

Definition 6 (Joint Known (Nominal)-Reference Process) *We say that Y_t, $t \in [0,T]$, 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 $\left(\Omega, \mathcal{F}, \mathfrak{W}_{0,t} \times \mathfrak{W}^\star_{0,t}, \mathbb{P}\right)$ (see Definition 2 for the filtrations):*

$$dY_t = G_\mu\left(t, Y_t\right)dt + G_\sigma\left(t, Y_t\right)d\widehat{W}_t, \quad t \in [0,T], \quad Y_0 = y_0 \sim \zeta_0^r, \quad Y_{[0,T]} \sim \mathbb{Y}_{[0,T]}, \tag{25}$$

where

$$y_0 \doteq \begin{bmatrix} x_0 \\ x_0^\star \end{bmatrix} \in \mathbb{R}^{2n}, \quad Y_t \doteq \begin{bmatrix} X_t^r \\ X_t^\star \end{bmatrix} \in \mathbb{R}^{2n}, \quad \widehat{W}_t \doteq \begin{bmatrix} W_t \\ W_t^\star \end{bmatrix} \in \mathbb{R}^{2d}, \quad \zeta_0^r \doteq \pi_0, \quad \mathbb{Y}_{[0,T]} \doteq \pi_{[0,T]},$$

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

and where π_0 and $\pi_{[0,T]}$ denote arbitrary couplings [28, Chp. 1] of the initial condition distributions ξ_0 and $\xi_0^\star$ on $\mathcal{B}(\mathbb{R}^{2n})$ and the laws $\mathbb{X}_{[0,T]}^r$ and $\mathbb{X}_{[0,T]}^\star$ on $\mathcal{B}(\mathcal{C}([0,T]; \mathbb{R}^{2n}))$, respectively.

We do not *a priori* assume the existence and uniqueness of strong solutions $X_t^\star$ and X_t^r of (5b) and (22), respectively. Instead, we will establish the well-posedness and the other desired results using the Khasminskii-type theorem [100, Thm. 3.2]. For this purpose, we will require a truncated version of (25) so that we can build local solutions and extend them to the global solution using a limiting procedure.

Definition 7 (Truncated Joint Known (Nominal)-Reference Process) *We first define*

$$U_N \doteq \left\{a \in \mathbb{R}^{2n} \; : \; \|a\| < N\right\} \subset\subset \mathbb{R}^{2n}, \quad \forall N \in \mathbb{R}_{>0}, \tag{26}$$

*where $\subset\subset$ $\mathbb{R}^{2n}$ denotes compact containment in $\mathbb{R}^{2n}$ [104, Sec. A.2]. Next, we define the **truncated joint known(nominal)-reference Itô SDE** as*

$$dY_{N,t} = G_{N,\mu}(t, Y_{N,t})\, dt + G_{N,\sigma}(t, Y_{N,t})\, d\widehat{W}_t, \quad Y_{N,0} = Y_0, \tag{27}$$

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

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

*for any $G_{N,\mu}(t,a)$ and $G_{N,\sigma}(t,a)$ that are **uniformly** 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 (27).*

An example of explicit construction of functions of the form $G_{N,\{\mu,\sigma\}}$ can be found in [93, 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 3.1 (Well-Posedness of (27)) *If Assumptions 1 and 4 hold true, then for any $N \in \mathbb{R}_{>0}$, $Y_{N,t}$ is a unique strong solution of (27), $\forall t \in [0,T]$, for any $T \in (0,\infty)$ and is a strong Markov process $\forall t \in \mathbb{R}_{\geq 0}$.*

Furthermore, define

$$\tau_N \doteq T \wedge \inf\left\{t \in [0,T] \; : \; Y_{N,t} \notin U_N\right\}, \tag{28}$$

where the open and bounded set U_N is defined in (26), for an arbitrary $N \in \mathbb{R}_{>0}$. Then, $Y_{N,t}$ uniquely solves (25), in the strong sense, for all $t \in [0, \tau_N]$.

Proof. See Appendix C. ☐

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

Lemma 3.1 *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 (28) for the truncated joint process in Definition 7. For any constant $t^\star \in \mathbb{R}_{>0}$ define*

$$\tau^\star = t^\star \wedge \tau_N, \quad \tau(t) = t \wedge \tau^\star, \tag{29}$$

and let $\pi_\star^0 \doteq \pi_{\tau^\star}^{y_0}$ be the finite-dimensional distribution of the coupling $\pi_{[0,T]}$ at the time instant $\tau^\star$ under the condition $Y_{N,0} = y_0$ $\mathbb{P}$-a.s..

Then, for the truncated joint process $Y_{N,t} = \left(X_{N,t}^r, X_{N,t}^\star\right)$ in Definition 7, the following holds:

$$\left\| \sup_{t \in [0,\tau^\star]} e^{(2\lambda + \boldsymbol{\omega})t} V(Y_{N,t}) \right\|_{\mathsf{p}}^{\pi_\star^0} \leq \left\| e^{\boldsymbol{\omega}\tau^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} V(y_0) + \left\| e^{\boldsymbol{\omega}\tau^\star} \left(\Xi^r(\cdot, Y_N)\right)_{\tau^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} + \left\| \left(\Xi_{\mathcal{U}}^r(\cdot, Y_N; \boldsymbol{\omega})\right)_{\tau^\star} \right\|_{\mathsf{p}}^{\pi_\star^0}, \tag{30}$$

where $\|\cdot\|_{\mathsf{p}}^{\pi_\star^0}$ denotes the norm on $L_{\mathsf{p}}(\pi_{\tau^\star}^{y_0})$ under the probability measure $\pi_{\tau^\star}^{y_0}$ on $\mathcal{B}(\mathbb{R}^{2n})$, and $V(Y_{N,t}) = V\left(X_{N,t}^r, X_{N,t}^\star\right)$ is the incremental lyapunov function (ILF) defined in Assumption 2. Furthermore,

$$\begin{aligned} \Xi^r(\tau(t), Y_N) &= \int_0^{\tau(t)} e^{2\lambda\nu}\left(\phi_\mu^r(\nu, Y_{N,\nu})\, d\nu + \phi_{\sigma_\star}^r(\nu, Y_{N,\nu})\, dW_\nu^\star + \phi_\sigma^r(\nu, Y_{N,\nu})\, dW_\nu\right), \\ \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \boldsymbol{\omega}) &= \int_0^{\tau(t)} \left(\mathcal{U}_\mu^r(\tau(t), \nu, Y_N; \boldsymbol{\omega})\, d\nu + \mathcal{U}_\sigma^r(\tau(t), \nu, Y_N; \boldsymbol{\omega})\, dW_\nu\right), \end{aligned} \tag{31}$$

where

$$\phi^r_\mu(\nu, Y_{N,\nu}) = V_r(Y_{N,\nu})^\top g(\nu)^\perp \Lambda^\perp_\mu(\nu, X^r_{N,t}) + \frac{1}{2} Tr\left[H_\sigma(\nu, Y_{N,\nu}) \nabla^2 V(Y_{N,\nu})\right],$$
$$\phi^r_{\sigma_\star}(\nu, Y_{N,\nu}) = V_\star(Y_\nu)^\top \bar{F}_\sigma(\nu, X^\star_\nu), \quad \phi^r_\sigma(\nu, Y_{N,\nu}) = V_r(Y_\nu)^\top g(\nu)^\perp F^\perp_\sigma(\nu, X^r_{N,\nu}), \tag{32}$$
$$\mathcal{U}^r_\mu(\tau(t), \nu, Y_N; \omega) = \psi^r(\tau(t), \nu, Y_N)\Lambda^\parallel_\mu(\nu, X^r_{N,\nu}), \quad \mathcal{U}^r_\sigma(\tau(t), \nu, Y_N; \omega) = \psi^r(\tau(t), \nu, Y_N) F^\parallel_\sigma(\nu, X^r_{N,\nu}),$$

and

$$\psi^r(\tau(t), \nu, Y_N) = \frac{\omega}{2\lambda - \omega}\left(e^{\omega(\tau(t)+\nu)}\mathcal{P}^r(\tau(t), \nu) - e^{(2\lambda\tau(t)+\omega\nu)} V_r\left(Y_{N,\tau(t)}\right)^\top g(\tau(t))\right)$$
$$+ \frac{2\lambda}{2\lambda - \omega} e^{(\omega\tau(t)+2\lambda\nu)} V_r(Y_{N,\nu})^\top g(\nu) \in \mathbb{R}^{1\times m}. \tag{33}$$

In the expressions above, we have defined $V_r = \nabla_{X^r_{N,t}} V \in \mathbb{R}^n$, $V_\star = \nabla_{X^\star_{N,t}} V \in \mathbb{R}^n$, $H_\sigma = G_\sigma G_\sigma^\top \in \mathbb{S}^{2n}$, and

$$\mathcal{P}^r(\tau(t), \nu) = \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} d_\beta\left[V_r(Y_{N,\beta})^\top g(\beta)\right] \in \mathbb{R}^{1\times m}, \quad 0 \le \nu \le \tau(t), \tag{34}$$

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

Proof. We prove the hypotheses of the lemma only for $y_0 \in U_N$, since for when $y_0 \notin U_N$, $\tau_N = 0$, and thus $\tau^\star = \bar{\tau} = 0$. Therefore, the bound in (30) is trivially satisfied when $y_0 \notin U_N$, $\tau_N = 0$. Hence, the proof of (30) for $y_0 \in U_N$ implies the result for all $y_0 \in \mathbb{R}^{2n}$.

Since $t^\star$ is a constant, the fact that τ_N is a stopping time implies that $\tau^\star$ is a stopping time as well [95, Sec. 6.1]. Additionally, from Proposition 3.1, we know that $Y_{N,t}$ is a unique strong solution of (27), 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 (27), 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 [100, Thm. 2.9.3]. Hence, the process $Y_{N,\tau(t)}$ obtained by stopping the process $Y_{N,t}$ at the first (random) instant it leaves the set U_N or when $t = t^\star$ is a Markov process and is well-defined [47, Lem. 3.2]. We may then apply the Itô lemma [100, Thm. 1.6.4] to $e^{2\lambda\tau(t)}V\left(Y_{N,\tau(t)}\right)$ using the dynamics in (27) to obtain

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

for all $t \in \mathbb{R}_{\ge 0}$, where $H_{N,\sigma}(\nu, Y_{N,\nu}) = G_{N,\sigma}(\nu, Y_{N,\nu}) G_{N,\sigma}(\nu, Y_{N,\nu})^\top$. Next, from Proposition 3.1, $Y_{N,t}$ is also unique strong solution of (25), for all $t \in [0, \tau^\star]$ because $[0, \tau^\star] \subseteq [0, \tau_N]$. Therefore, 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

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

for all $t \in \mathbb{R}_{\ge 0}$, 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.9a) and (C.9b), Proposition C.1, for the last two terms on the right hand side of the above expression and re-arranging terms leads to

$$e^{2\lambda\tau(t)}V\left(Y_{N,\tau(t)}\right)$$
$$\le V(y_0) + \int_0^{\tau(t)} e^{2\lambda\nu}\left(\phi^r_\mu(\nu, Y_{N,\nu})\,d\nu + \phi^r_{\sigma_\star}(\nu, Y_{N,\nu})\,dW^\star_\nu + \phi^r_\sigma(\nu, Y_{N,\nu})\,dW_\nu\right)$$
$$+ \int_0^{\tau(t)} e^{2\lambda\nu}\left(\left[\phi^r_U(\nu, Y_{N,\nu}) + \phi^r_{\mu^\parallel}(\nu, Y_{N,\nu})\right]d\nu + \phi^r_{\sigma^\parallel}(\nu, Y_{N,\nu})\,dW_\nu\right), \quad \forall t \in \mathbb{R}_{\ge 0}, \tag{35}$$

where

$$\phi^r_{\mu^\parallel}(\nu, Y_{N,\nu}) = V_r(Y_{N,\nu})^\top g(\nu)\Lambda^\parallel_\mu(\nu, X^r_{N,t}), \quad \phi^r_{\sigma^\parallel}(\nu, Y_{N,\nu}) = V_r(Y_\nu)^\top g(\nu) F^\parallel_\sigma(\nu, X^r_{N,\nu}),$$

$$\phi_U^r\left(\nu, Y_{N,\nu}\right) = V_r\left(Y_{N,\nu}\right)^\top g(\nu) U_\nu^r.$$

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

$$\begin{aligned}
&\int_0^{\tau(t)} e^{2\lambda\nu}\left(\left[\phi_U^r\left(\nu, Y_{N,\nu}\right) + \phi_{\mu^{\parallel}}^r\left(\nu, Y_{N,\nu}\right)\right] d\nu + \phi_{\sigma^{\parallel}}^r\left(\nu, Y_{N,\nu}\right) dW_\nu\right) \\
&= \int_0^{\tau(t)} e^{2\lambda\nu}\left(\phi_{\mu^{\parallel}}^r\left(\nu, Y_{N,\nu}\right) + \phi_{U_\mu}^r\left(\nu, Y_{N,\nu}; \omega\right)\right) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu}\left(\phi_{\sigma^{\parallel}}^r\left(\nu, Y_{N,\nu}\right) + \phi_{U_\sigma}^r\left(\nu, Y_{N,\nu}; \omega\right)\right) dW_\nu \\
&\quad + \int_0^{\tau(t)}\left(\hat{\mathcal{U}}_\mu^r\left(\tau(t), \nu, Y_N; \omega\right) d\nu + \hat{\mathcal{U}}_\sigma^r\left(\tau(t), \nu, Y_N; \omega\right) dW_\nu\right), \quad \forall t \in \mathbb{R}_{\geq 0}, \quad (36)
\end{aligned}$$

where

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

Using the definitions of $\phi_{\mu^{\parallel}}^r$ and $\phi_{\sigma^{\parallel}}^r$ in (35), we obtain

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

Substituting into (36) thus produces

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

where $\mathcal{U}_\mu^r$ and $\mathcal{U}_\sigma^r$ are defined in (32). Substituting the above expression for the last integral on the right hand side of (35) and using the definitions of Ξ^r and $\Xi_{\mathcal{U}}^r$ in (31) yields the following bound:

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

which in turn produces

$$e^{(2\lambda+\omega)\tau(t)} V\left(Y_{N,\tau(t)}\right) \leq e^{\omega\tau(t)} V\left(y_0\right) + e^{\omega\tau(t)}\Xi^r\left(\tau(t), Y_N\right) + \Xi_{\mathcal{U}}^r\left(\tau(t), Y_N; \omega\right), \quad \forall t \in \mathbb{R}_{\geq 0}, \quad (37)$$

Taking supremum on both sides over the interval $[0, \tau^\star]$ then produces

$$\sup_{t\in[0,\tau^\star]}\left[e^{(2\lambda+\omega)t} V\left(Y_{N,t}\right)\right] \leq e^{\omega\tau^\star} V\left(y_0\right) + e^{\omega\tau^\star}\sup_{t\in[0,\tau^\star]}\left|\Xi^r\left(t, Y_N\right)\right| + \sup_{t\in[0,\tau^\star]}\left|\Xi_{\mathcal{U}}^r\left(t, Y_N; \omega\right)\right|. \quad (38)$$

Then, the bound in (30) is established by invoking the Minkowski's inequality and thus concluding the proof.

□

3.3 Performance Analysis: True (Uncertain) Process

We now consider the true (uncertain) system in (5a) operating under the $\mathcal{L}_1$-DRAC feedback law $\mathcal{F}_{\mathcal{L}_1}$ defined in (17), Sec. 3.1.

Definition 8 (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 (5a) under the $\mathcal{L}_1$-DRAC feedback law $\mathcal{F}_{\mathcal{L}_1}$ in (17):*

$$dX_t = F_\mu\left(t, X_t, U_{\mathcal{L}_1,t}\right)dt + F_\sigma\left(t, X_t\right)dW_t, \quad X_0 = x_0 \sim \xi_0 \; (\mathbb{P}\text{-a.s.}), \tag{39}$$

where

$$U_{\mathcal{L}_1,t} \doteq \mathcal{F}_{\mathcal{L}_1}\left(X\right)(t). \tag{40}$$

The $\mathcal{L}_1$-DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1} : \mathcal{C}([0,T] : \mathbb{R}^n) \to \mathcal{C}([0,T] : \mathbb{R}^m)$, is defined in (17), which we restate next. Using the definition in (17) that

$$\mathcal{F}_{\mathcal{L}_1}\left(X\right) = \mathcal{F}_{\boldsymbol{\omega}} \circ \mathcal{F}_{\boldsymbol{T_s}} \circ \mathcal{F}_{\lambda_s}\left(X\right),$$

which we can re-write as

$$U_{\mathcal{L}_1} = \mathcal{F}_{\boldsymbol{\omega}}\left(\hat{\Lambda}^{\parallel}\right), \quad \hat{\Lambda}^{\parallel} = \mathcal{F}^{\parallel}_{\boldsymbol{T_s}}\left(\hat{\Lambda}\right), \quad \Lambda^{\parallel} = \mathcal{F}_{\boldsymbol{T_s}}\left(\hat{X}, X\right), \quad \hat{X} = \mathcal{F}_{\lambda_s}\left(X\right). \tag{41}$$

*Using (18a), we see that the input $U_{\mathcal{L}_1,t}$ is defined as the output of the **low-pass filter**:*

$$U_{\mathcal{L}_1,t} = \mathcal{F}_{\boldsymbol{\omega}}\left(\hat{\Lambda}^{\parallel}\right)(t) = -\boldsymbol{\omega}\int_0^t e^{-\boldsymbol{\omega}(t-\nu)}\hat{\Lambda}^{\parallel}\left(\nu\right)d\nu, \tag{42}$$

*where $\boldsymbol{\omega} \in \mathbb{R}_{>0}$ is the **filter bandwidth**. The **adaptive estimates** $\hat{\Lambda}^{\parallel}$ and $\hat{\Lambda}$ are obtained via the **adaptation law** operator $\mathcal{F}_{\boldsymbol{T_s}}$ $\left(\mathcal{F}^{\parallel}_{\boldsymbol{T_s}}\right)$ in (18b) as follows:*

$$\begin{aligned}
\hat{\Lambda}^{\parallel}\left(t\right) =& \mathcal{F}^{\parallel}_{\boldsymbol{T_s}}\left(\hat{\Lambda}\right)(t) = \sum_{i=0}^{\lfloor \frac{t}{\boldsymbol{T_s}} \rfloor} \Theta_{ad}(i\boldsymbol{T_s})\hat{\Lambda}\left(t\right)\mathbb{1}_{\{[i\boldsymbol{T_s},(i+1)\boldsymbol{T_s})\}}\left(t\right), \\
\hat{\Lambda}\left(t\right) =& \mathcal{F}_{\boldsymbol{T_s}}\left(\hat{X}_t, X\right)(t) \\
=& 0_n\mathbb{1}_{\{[0,\boldsymbol{T_s})\}}\left(t\right) + \lambda_s\left(1 - e^{\lambda_s \boldsymbol{T_s}}\right)^{-1}\sum_{i=1}^{\lfloor \frac{t}{\boldsymbol{T_s}} \rfloor}\tilde{X}_{i\boldsymbol{T_s}}\mathbb{1}_{\{[i\boldsymbol{T_s},(i+1)\boldsymbol{T_s})\}}\left(t\right), \quad \tilde{X}_{i\boldsymbol{T_s}} \doteq \hat{X}_{i\boldsymbol{T_s}} - X_{i\boldsymbol{T_s}},
\end{aligned} \tag{43}$$

*where $\boldsymbol{T_s} \in \mathbb{R}_{>0}$ is the **sampling period** and $\Theta_{ad}(t) = \begin{bmatrix}\mathbb{I}_m & 0_{m,n-m}\end{bmatrix}\bar{g}(t)^{-1} \in \mathbb{R}^{m\times n}$, with $\bar{g}(t) = \begin{bmatrix}g(t) & g(t)^{\perp}\end{bmatrix} \in \mathbb{R}^{n\times n}$ 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 $\mathcal{F}_{\lambda_s}$ in (18c) which induces the **process predictor** as follows:*

$$\hat{X}_t = \mathcal{F}_{\lambda_s}\left(X\right)(t) \Rightarrow \hat{X}_t = \int_0^t\left(-\lambda_s\mathbb{I}_n\tilde{X}_\nu + f(\nu, X_\nu) + g(\nu)\mathcal{F}_{\mathcal{L}_1}\left(X\right)(\nu) + \hat{\Lambda}\left(\nu\right)\right)d\nu, \quad \tilde{X}_t \doteq \hat{X}_t - X_t. \tag{44}$$

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

*Finally, in Definition 2, $\mathbb{X}_{[0,T]}$ denotes the **true law (path/trajectory probability measures)** induced by the process path $X_{[0,T]}$ on $\mathcal{B}\left(\mathcal{C}\left([0,T];\mathbb{R}^n\right)\right)$, i.e., $X_{[0,T]} \sim \mathbb{X}_{[0,T]}$.*

Our goal in this section is to establish the performance of the $\mathcal{L}_1$-DRAC closed-loop true (uncertain) process in (39) relative to the reference process in Definition 5. Thus, similar to the joint known(nominal)-reference process in Definition 6, we define the following:

Definition 9 (Joint True (Uncertain)-Reference Process) *Consider the reference process $X^r_{[0,T]} \sim \mathbb{X}^r_{[0,T]}$ in Definition 5. We say that Z_t, $t \in [0,T]$, for any $T \in (0,\infty)$, is the **joint tue (uncertain)-reference process**, if it is a unique strong solution of the following **joint true (uncertain)-reference Itô SDE** on $(\Omega, \mathcal{F}, \mathfrak{W}_{0,t}, \mathbb{P})$ (see Definition 2 for the filtrations):*

$$dZ_t = J_\mu\left(t, Z_t\right)dt + J_\sigma\left(t, Z_t\right)dW_t, \quad t \in [0,T], \quad Z_0 = z_0 \sim \zeta_0, \quad Z_{[0,T]} \sim \mathbb{Z}_{[0,T]}, \tag{45}$$

where

$$z_0 \doteq 1_2 \otimes x_0 \in \mathbb{R}^{2n}, \quad Z_t \doteq \begin{bmatrix}X_t \\ X^r_t\end{bmatrix} \in \mathbb{R}^{2n}, \quad \zeta_0 \doteq \bar{\pi}_0, \quad \mathbb{Z}_{[0,T]} \doteq \bar{\pi}_{[0,T]},$$

$$J_\mu\left(t, Z_t\right) \doteq \begin{bmatrix}F_\mu\left(t, X_t, U_{\mathcal{L}_1,t}\right) \\ F_\mu\left(t, X^r_t, U^r_t\right)\end{bmatrix} \in \mathbb{R}^{2n}, \quad J_\sigma\left(t, Z_t\right) \doteq \begin{bmatrix}F_\sigma\left(t, X_t\right) \\ F_\sigma\left(t, X^r_t\right)\end{bmatrix} \in \mathbb{R}^{2n\times d},$$

where $\otimes$ denotes the Kronecker product and $x_0 \sim \xi_0$ is the initial condition for the true (uncertain) process in Definition 2, and restated above in Definition 8. Therefore, $\bar{\pi}_0$ denotes the trivial self-coupling of the initial condition measure ξ_0 [105]. Moreover, $\bar{\pi}_{[0,T]}$ denotes the arbitrary coupling of the the laws $\mathbb{X}_{[0,T]}$ and $\mathbb{X}^r_{[0,T]}$ (see Definition 5) on $\mathcal{B}\left(\mathcal{C}\left([0,T];\mathbb{R}^{2n}\right)\right)$, respectively.

Remark 3.5 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 6. 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 6, where the known (nominal) process and the reference process are driven by *independent Brownian motions* $W_t^\star$ and W_t and have *different initial conditions* $x_0 \sim \xi_0$ and $x_0^\star \sim \xi_0^\star$, respectively. Recall our discussion in Remark 3.3 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.

As in the case of the reference process, we establish the well-posedness of the true (uncertain)-reference process in Definition 9 using the Khasminskii-type theorem [100, Thm. 3.2]. Therefore, similar to the truncated joint known (nominal)-reference process in Definition 7, we require the following:

Definition 10 (Truncated Joint True (Uncertain)-Reference Process) *Recall the definition of the set $U_N \subset\subset \mathbb{R}^{2n}$ in (26) which we restate below:*

$$U_N \doteq \left\{a \in \mathbb{R}^{2n} \ : \ \|a\| < N\right\} \subset\subset \mathbb{R}^{2n}, \quad \forall N \in \mathbb{R}_{>0}. \tag{46}$$

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

$$dZ_{N,t} = J_{N,\mu}\left(t, Z_{N,t}\right) dt + J_{N,\sigma}\left(t, Z_{N,t}\right) dW_t, \quad Z_{N,0} = Z_0, \tag{47}$$

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

$$Z_{N,t} \doteq \begin{bmatrix} X_{N,t} \\ X_{N,t}^r \end{bmatrix}, \quad J_{N,\mu}\left(t,a\right)\left(J_{N,\sigma}\left(t,a\right)\right) = \begin{cases} J_\mu\left(t,a\right)\left(J_\sigma\left(t,a\right)\right), & \|a\| \le N \\ 0_{2n}\left(0_{2n,d}\right), & \|a\| \ge 2N \end{cases}, \quad \forall (a,t) \in \mathbb{R}^{2n} \times [0,T],$$

for any $J_{N,\mu}\left(t,a\right)$ and $J_{N,\sigma}\left(t,a\right)$ that are ***uniformly*** *Lipschitz continuous for all $a \in \mathbb{R}^{2n}$ and $t \in \mathbb{R}_{\ge 0}$. Similar to Z_t, we refer to $Z_{N,t} \in \mathbb{R}^{2n}$ as the* ***truncated joint known(nominal)-reference process*** *if it is a unique strong solution of (47).*

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

Proposition 3.2 (Well-Posedness of (47)) *If Assumptions 1 and 4 hold true, then for any $N \in \mathbb{R}_{>0}$, $Z_{N,t}$ is a unique strong solution of (47), $\forall t \in [0,T]$, for any $T \in (0,\infty)$ and is a strong Markov process $\forall t \in \mathbb{R}_{\ge 0}$.*

Furthermore, define

$$\tau_N \doteq T \wedge \inf\left\{t \in [0,T] \ : \ Z_{N,t} \notin U_N\right\}, \tag{48}$$

where the open and bounded set U_N is defined in (46), for an arbitrary $N \in \mathbb{R}_{>0}$. Then, $Z_{N,t}$ uniquely solves (45), in the strong sense, for all $t \in [0,\tau_N]$.

Proof. See Appendix D. □

Remark 3.6 Recall that τ_N in (28), Proposition 3.1 denotes the first exit time of the process $Y_{N,t}$ from the set U_N. Similarly, with an abuse of notation and from this point onward, τ_N in (48) denotes the first exit time of the process $Z_{N,t}$ from the set U_N.

Next, we study the *uniform in time* bounds on the moments of the truncated joint true (uncertain)-reference process.

Lemma 3.2 *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 (28) for the truncated joint process in Definition 7. For any constant $t^\star \in \mathbb{R}_{>0}$ define*

$$\tau^\star = t^\star \wedge \tau_N, \quad \tau(t) = t \wedge \tau^\star, \tag{49}$$

and let $\bar{\pi}_\star^0 \doteq \bar{\pi}_{\tau^\star}^{z_0}$ be the finite-dimensional distribution of the coupling $\bar{\pi}_{[0,T]}$ at the time instant $\tau^\star$ under the condition $Z_{N,0} = z_0$ $\mathbb{P}$-a.s..

Then, for the truncated joint process $Z_{N,t} = \left(X_{N,t}, X_{N,t}^r\right)$ in Definition 10, the following holds:

$$\left\| \sup_{t\in[0,\tau^\star]} \left[e^{(2\lambda+\omega)\tau(t)} V\left(Z_{N,\tau(t)}\right)\right] \right\|_{\mathsf{p}}^{\bar{\pi}_\star^0} \le \left\| e^{\omega\tau^\star} \right\|_{\mathsf{p}}^{\bar{\pi}_\star^0} V\left(z_0\right) + \left\| e^{\omega\tau^\star} \left(\Xi\left(\cdot, Z_N\right)\right)_{\tau^\star} \right\|_{\mathsf{p}}^{\bar{\pi}_\star^0}$$
$$+ \left\| e^{\omega\tau^\star} \left(\widetilde{\Xi}_{\mathcal{U}}\left(\cdot, Z_N\right)\right)_{\tau^\star} \right\|_{\mathsf{p}}^{\bar{\pi}_\star^0} + \left\| \left(\Xi_{\mathcal{U}}\left(\cdot, Z_N\right)\right)_{\tau^\star} \right\|_{\mathsf{p}}^{\bar{\pi}_\star^0}. \tag{50}$$

where $\|\cdot\|_{\mathsf{p}}^{\bar{\pi}_\star^0}$ *denotes the norm on* $L_{\mathsf{p}}\left(\bar{\pi}_{\tau^\star}^{z_0}\right)$ *under the probability measure* $\bar{\pi}_{\tau^\star}^{z_0}$ *on* $\mathcal{B}\left(\mathbb{R}^{2n}\right)$*, and* $V\left(Z_{N,t}\right) = V\left(X_{N,t}, X_{N,t}^r\right)$ *is the incremental lyapunov function (ILF) defined in Assumption 2.*

Furthermore,

$$\Xi\left(\tau(t), Z_N\right) = \int_0^{\tau(t)} e^{2\lambda\nu}\left(\phi_\mu\left(\nu, Z_{N,\nu}\right) + \phi_\sigma\left(\nu, Z_{N,\nu}\right) dW_\nu\right),$$
$$\widetilde{\Xi}_{\mathcal{U}}\left(\tau(t), Z_N\right) = \int_0^{\tau(t)} e^{2\lambda\nu}\widetilde{\mathcal{U}}\left(\nu, Z_{N,\nu}\right) d\nu, \tag{51}$$
$$\Xi_{\mathcal{U}}\left(\tau(t), Z_N; \boldsymbol{\omega}\right) = \int_0^{\tau(t)} \left(\mathcal{U}_\mu\left(\tau(t), \nu, Z_N; \boldsymbol{\omega}\right) d\nu + \mathcal{U}_\sigma\left(\tau(t), \nu, Z_N; \boldsymbol{\omega}\right) dW_\nu\right),$$

where

$$\phi_\mu\left(\nu, Z_{N,\nu}\right) = \nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2 \otimes g(\nu)^\perp\right)\left(\Lambda_\mu^\perp \odot Z_N\right)(\nu) + \frac{1}{2}Tr\left[K_\sigma\left(\nu, Z_{N,\nu}\right)\nabla^2 V\left(Z_{N,\nu}\right)\right],$$
$$\phi_\sigma\left(\nu, Z_{N,\nu}\right) = \nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2 \otimes g(\nu)^\perp\right)\left(F_\sigma^\perp \odot Z_N\right)(\nu), \tag{52a}$$

$$\mathcal{U}_\mu\left(\tau(t), \nu, Z_N; \boldsymbol{\omega}\right) = \psi(\tau(t), \nu, Z_N)\left(\Lambda_\mu^\parallel \odot Z_N\right)(\nu),$$
$$\mathcal{U}_\sigma\left(\tau(t), \nu, Z_N; \boldsymbol{\omega}\right) = \psi(\tau(t), \nu, Z_N)\left(F_\sigma^\parallel \odot Z_N\right)(\nu), \tag{52b}$$
$$\widetilde{\mathcal{U}}\left(\nu, Z_{N,\nu}\right) = V_{\cdot}\left(Z_{N,\nu}\right)^\top g(\nu)\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)(X_N)(\nu),$$

and

$$\psi(\tau(t), \nu, Y_N) = \frac{\boldsymbol{\omega}}{2\lambda - \boldsymbol{\omega}}\left(e^{\boldsymbol{\omega}(\tau(t)+\nu)}\mathcal{P}\left(\tau(t), \nu\right) - e^{(2\lambda\tau(t)+\boldsymbol{\omega}\nu)}\nabla V\left(Z_{N,\tau(t)}\right)^\top\left(\mathbb{I}_2 \otimes g(\tau(t))\right)\right) + \frac{2\lambda}{2\lambda - \boldsymbol{\omega}}e^{(\boldsymbol{\omega}\tau(t)+2\lambda\nu)}\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2 \otimes g(\nu)\right), \in \mathbb{R}^{1\times 2m}. \tag{53}$$

In the expressions above, we have defined $V_{\cdot} = \nabla_{X_{N,t}} V \in \mathbb{R}^n$*,* $V_r = \nabla_{X_{N,t}^r} V \in \mathbb{R}^n$*,* $K_\sigma = J_\sigma J_\sigma^\top \in \mathbb{S}^{2n}$*, and*

$$\mathcal{P}\left(\tau(t), \nu\right) = \int_\nu^{\tau(t)} e^{(2\lambda - \boldsymbol{\omega})\beta} d_\beta\left[\nabla V\left(Z_{N,\beta}\right)^\top\left(\mathbb{I}_2 \otimes g(\beta)\right)\right] d\beta \in \mathbb{R}^{1\times 2m}, \quad 0 \le \nu \le \tau(t), \tag{54}$$

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

Proof. As in the proof of Lemma 3.1, we consider $z_0 \in U_N$ w.l.o.g., and apply the Itô lemma [100, Thm. 1.6.4] to $e^{2\lambda\tau(t)}V\left(Z_{N,\tau(t)}\right)$ using the dynamics in (47) to obtain

$$e^{2\lambda\tau(t)}V\left(Z_{N,\tau(t)}\right) = V\left(z_0\right) + 2\lambda\int_0^{\tau(t)} e^{2\lambda\nu}V\left(Z_{N,\nu}\right) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{N,\nu}\right)^\top J_{N,\sigma}\left(\nu, Z_{N,\nu}\right) dW_\nu + \int_0^{\tau(t)} e^{2\lambda\nu}\left(\nabla V\left(Z_{N,\nu}\right)^\top J_{N,\mu}\left(\nu, Z_{N,\nu}\right) + \frac{1}{2}\mathrm{Tr}\left[K_{N,\sigma}\left(\nu, Z_{N,\nu}\right)\nabla^2 V\left(Z_{N,\nu}\right)\right]\right) d\nu,$$

for all $t \in \mathbb{R}_{\ge 0}$, where $K_{N,\sigma}\left(\nu, Y_{N,\nu}\right) = J_{N,\sigma}\left(\nu, Y_{N,\nu}\right) J_{N,\sigma}\left(\nu, Y_{N,\nu}\right)^\top$. Next, from Proposition 3.2, $Z_{N,t}$ is also unique strong solution of (45), for all $t \in [0, \tau^\star]$ because $[0, \tau^\star] \subseteq [0, \tau_N]$. Therefore, we may replace $J_{N,\mu}$, $J_{N,\sigma}$, and $K_{N,\sigma}$ with J_μ, J_σ, and J_σ, respectively, in the above inequality and obtain

$$e^{2\lambda\tau(t)}V\left(Z_{N,\tau(t)}\right) = V\left(z_0\right) + 2\lambda\int_0^{\tau(t)} e^{2\lambda\nu}V\left(Z_{N,\nu}\right) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{N,\nu}\right)^\top J_\sigma\left(\nu, Z_{N,\nu}\right) dW_\nu + \int_0^{\tau(t)} e^{2\lambda\nu}\left(\nabla V\left(Z_{N,\nu}\right)^\top J_\mu\left(\nu, Z_{N,\nu}\right) + \frac{1}{2}\mathrm{Tr}\left[K_\sigma\left(\nu, Z_{N,\nu}\right)\nabla^2 V\left(Z_{N,\nu}\right)\right]\right) d\nu,$$

for all $t \in \mathbb{R}_{\ge 0}$, where $K_\sigma\left(\nu, Z_{N,\nu}\right) = J_\sigma\left(\nu, Z_{N,\nu}\right) J_\sigma\left(\nu, Z_{N,\nu}\right)^\top$. Substituting the expressions in (D.17a) and (D.17b), Proposition D.1, for the last two terms on the right hand side of the above expression and re-arranging terms leads to

$$
\begin{aligned}
&e^{2\lambda\tau(t)}V\left(Z_{N,\tau(t)}\right) \\
&\leq V\left(z_0\right) + \int_0^{\tau(t)} e^{2\lambda\nu}\left(\left[\phi_\mu\left(\nu, Z_{N,\nu}\right) + \widetilde{\mathcal{U}}\left(\nu, Z_{N,\nu}\right)\right]d\nu + \phi_\sigma\left(\nu, Z_{N,\nu}\right)dW_\nu\right) \\
&\qquad + \int_0^{\tau(t)} e^{2\lambda\nu}\left(\left[\phi_U\left(\nu, Z_{N,\nu}\right) + \phi_{\mu^{\parallel}}\left(\nu, Z_{N,\nu}\right)\right]d\nu + \phi_{\sigma^{\parallel}}\left(\nu, Z_{N,\nu}\right)dW_\nu\right), \quad \forall t \in \mathbb{R}_{\geq 0}, \quad (55)
\end{aligned}
$$

where

$$
\begin{aligned}
\phi_{\mu^{\parallel}}\left(\nu, Z_{N,\nu}\right) &= \nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2 \otimes g(\nu)\right)\left(\Lambda_\mu^{\parallel} \odot Z_N\right)(\nu), \\
\phi_{\sigma^{\parallel}}\left(\nu, Z_{N,\nu}\right) &= \nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2 \otimes g(\nu)\right)\left(F_\sigma^{\parallel} \odot Z_N\right)(\nu), \\
\phi_U\left(\nu, Z_{N,\nu}\right) &= \nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2 \otimes g(\nu)\right)\left(\mathcal{F}_r \odot Z_N\right)(\nu).
\end{aligned}
$$

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

$$
\begin{aligned}
&\int_0^{\tau(t)} e^{2\lambda\nu}\left(\left[\phi_U\left(\nu, Z_{N,\nu}\right) + \phi_{\mu^{\parallel}}\left(\nu, Z_{N,\nu}\right)\right]d\nu + \phi_{\sigma^{\parallel}}\left(\nu, Z_{N,\nu}\right)dW_\nu\right) \\
&= \int_0^{\tau(t)} e^{2\lambda\nu}\left(\phi_{\mu^{\parallel}}\left(\nu, Z_{N,\nu}\right) + \phi_{U_\mu}\left(\nu, Z_{N,\nu}; \omega\right)\right)d\nu + \int_0^{\tau(t)} e^{2\lambda\nu}\left(\phi_{\sigma^{\parallel}}\left(\nu, Z_{N,\nu}\right) + \phi_{U_\sigma}\left(\nu, Z_{N,\nu}; \omega\right)\right)dW_\nu \\
&\qquad + \int_0^{\tau(t)}\left(\hat{\mathcal{U}}_\mu\left(\tau(t), \nu, Z_N; \omega\right)d\nu + \hat{\mathcal{U}}_\sigma\left(\tau(t), \nu, Z_N; \omega\right)dW_\nu\right), \quad \forall t \in \mathbb{R}_{\geq 0}, \quad (56)
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{\mathcal{U}}_\mu\left(\tau(t), \nu, Z_N; \omega\right) &= e^{-\omega\tau(t)}\frac{\omega}{2\lambda - \omega}\left(e^{\omega\tau(t)}\mathcal{P}\left(\tau(t), \nu\right) - e^{2\lambda\tau(t)}\nabla V\left(Z_{N,\tau(t)}\right)^\top\left(\mathbb{I}_2 \otimes g(\tau(t))\right)\right) \\
&\qquad \times e^{\omega\nu}\left(\Lambda_\mu^{\parallel} \odot Z_N\right)(\nu), \\
\hat{\mathcal{U}}_\sigma\left(\tau(t), \nu, Z_N; \omega\right) &= e^{-\omega\tau(t)}\frac{\omega}{2\lambda - \omega}\left(e^{\omega\tau(t)}\mathcal{P}\left(\tau(t), \nu\right) - e^{2\lambda\tau(t)}\nabla V\left(Z_{N,\tau(t)}\right)^\top\left(\mathbb{I}_2 \otimes g(\tau(t))\right)\right) \\
&\qquad \times e^{\omega\nu}\left(F_\sigma^{\parallel} \odot Z_N\right)(\nu), \\
\phi_{U_\mu}\left(\nu, Z_{N,\nu}; \omega\right) &= \frac{\omega}{2\lambda - \omega}\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2 \otimes g(\nu)\right)\left(\Lambda_\mu^{\parallel} \odot Z_N\right)(\nu), \\
\phi_{U_\sigma}\left(\nu, Z_{N,\nu}; \omega\right) &= \frac{\omega}{2\lambda - \omega}\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2 \otimes g(\nu)\right)\left(F_\sigma^{\parallel} \odot Z_N\right)(\nu).
\end{aligned}
$$

Using the definitions of $\phi_{\mu^{\parallel}}$ and $\phi_{\sigma^{\parallel}}$ in (55), we obtain

$$
\begin{aligned}
\phi_{\mu^{\parallel}}\left(\nu, Z_{N,\nu}\right) + \phi_{U_\mu}\left(\nu, Z_{N,\nu}; \omega\right) &= \frac{2\lambda}{2\lambda - \omega}\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2 \otimes g(\nu)\right)\left(\Lambda_\mu^{\parallel} \odot Z_N\right)(\nu), \\
\phi_{\sigma^{\parallel}}\left(\nu, Z_{N,\nu}\right) + \phi_{U_\sigma}\left(\nu, Z_{N,\nu}; \omega\right) &= \frac{2\lambda}{2\lambda - \omega}\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2 \otimes g(\nu)\right)\left(F_\sigma^{\parallel} \odot Z_N\right)(\nu).
\end{aligned}
$$

Substituting into (56) thus produces

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

where $\mathcal{U}_\mu$ and $\mathcal{U}_\sigma$ are defined in (52). Substituting the above expression for the last integral on the right hand side of (55) and using the definitions of Ξ, $\widetilde{\Xi}_{\mathcal{U}}$, and $\Xi_{\mathcal{U}}$ in (51) yields the following bound:

$$
e^{2\lambda\tau(t)}V\left(Z_{N,\tau(t)}\right) \leq V\left(z_0\right) + \Xi\left(\tau(t), Z_N\right) + \widetilde{\Xi}_{\mathcal{U}}\left(\tau(t), Z_N\right) + e^{-\omega\tau(t)}\Xi_{\mathcal{U}}\left(\tau(t), Z_N; \omega\right), \quad \forall t \in \mathbb{R}_{\geq 0},
$$

which in turn produces

$$
e^{(2\lambda+\omega)\tau(t)}V\left(Z_{N,\tau(t)}\right) \leq e^{\omega\tau(t)}V\left(z_0\right) + e^{\omega\tau(t)}\Xi\left(\tau(t), Z_N\right) + e^{\omega\tau(t)}\widetilde{\Xi}_{\mathcal{U}}\left(\tau(t), Z_N\right) + \Xi_{\mathcal{U}}\left(\tau(t), Z_N; \omega\right),
$$

or all $t \in \mathbb{R}_{\geq 0}$. Taking the supremum on both sides over the interval $[0, \tau^\star]$ then produces

$$\sup_{t\in[0,\tau^\star]} \left[e^{(2\lambda+\omega)\tau(t)} V\left(Z_{N,\tau(t)}\right)\right] \leq e^{\omega\tau^\star} V\left(z_0\right) + e^{\omega\tau^\star}\left(\Xi\left(\cdot, Z_N\right)\right)_{\tau^\star} + e^{\omega\tau^\star}\left(\widetilde{\Xi}_{\mathcal{U}}\left(\cdot, Z_N\right)\right)_{\tau^\star} + \left(\Xi_{\mathcal{U}}\left(\cdot, Z_N\right)\right)_{\tau^\star}.$$

The proof is then concluded by invoking the Minkowski's inequality to obtain (50).

□

4 Discussion

Appendices

A Constants

In this section we collect cumbersome constants we use throughout the manuscript. The subsequent definitions use the the constants $\Delta_{(\cdot)}$ defined in Assumptions 1-5 in Sec. 2.2.

We begin by providing the definitions of constants and maps bespoke to the analysis of the reference process in Sec. 3.2.

A.1 Reference Process Analysis

We begin with the definition of constants $\widehat{\Delta}_1^r$, $\widehat{\Delta}_2^r$, $\widehat{\Delta}_3^r$, $\widehat{\Delta}_4^r$, and $\widehat{\Delta}_{4_\parallel}^r$ that are used in Proposition C.7, Appendix C:

$$\widehat{\Delta}_1^r = \Delta_{\partial V}\left(2\Delta_f + \Delta_\mu\right)\left(1+\Delta_\star\right) + \Delta_{\partial^2 V}\left(\Delta_p^2 + \Delta_\sigma^2\left(1+\Delta_\star\right)\right), \tag{A.1a}$$

$$\widehat{\Delta}_2^r = 2\Delta_p + \Delta_\sigma\left(1+\Delta_\star\right)^{\frac{1}{2}}, \tag{A.1b}$$

$$\widehat{\Delta}_3^r = \sqrt{\frac{n}{2}}\Delta_g\left(\Delta_{\partial V}\left(\Delta_f + \Delta_\mu\right) + \Delta_{\partial^2 V}\Delta_\sigma^2\right) + \frac{\Delta_{\dot{g}}\Delta_{\partial V}}{\sqrt{2}}, \tag{A.1c}$$

$$\widehat{\Delta}_4^r = \Delta_p + \Delta_\sigma\left(1+\Delta_\star\right)^{\frac{1}{2}}, \quad \widehat{\Delta}_{4_\parallel}^r = \Delta_p^\parallel + \Delta_\sigma^\parallel\left(1+\Delta_\star\right)^{\frac{1}{2}}. \tag{A.1d}$$

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

$$\boldsymbol{\Delta}_{\circ_1}^{\boldsymbol{r}} = \frac{\Delta_g\Delta_\mu^\perp}{2}\left(1+\Delta_\star\right)\|\nabla V(0,0)\| + \frac{\sqrt{n}\Delta_{\partial V}}{4\sqrt{2}}\left(2\Delta_p^2 + \Delta_\sigma^2\left(1+\Delta_\star\right)\right), \tag{A.2a}$$

$$\boldsymbol{\Delta}_{\circ_2}^{\boldsymbol{r}} = 2\sqrt{\mathsf{p}}\left[\Delta_g^\perp\left(\Delta_p^\perp + \Delta_\sigma^\perp\left(1+\Delta_\star\right)^{\frac{1}{2}}\right) + \Delta_p\right]\|\nabla V(0,0)\|, \tag{A.2b}$$

$$\begin{aligned}\boldsymbol{\Delta}_{\circ_3}^{\boldsymbol{r}} = {} & 2\Delta_g\left[\left(1+\sqrt{2\lambda\mathsf{p}}\right)\Delta_\mu^\parallel\left(1+\Delta_\star\right) + \sqrt{2\lambda\mathsf{p}}\left(\Delta_p^\parallel + \Delta_\sigma^\parallel\left(1+\Delta_\star\right)^{\frac{1}{2}}\right)\right]\|\nabla V(0,0)\| \\ & + \frac{\Delta_\mu^\parallel}{\sqrt{\lambda}}\left(1+\Delta_\star\right)\left(\frac{\sqrt{n}\Delta_g\widehat{\Delta}_1^r}{2\sqrt{2\lambda}} + \frac{\Delta_{\dot{g}}}{2\sqrt{\lambda}}\|\nabla V(0,0)\| + \sqrt{2n\mathsf{p}}\Delta_g\Delta_{\partial V}\widehat{\Delta}_2^r\right) \\ & + \sqrt{\frac{n}{2}}\frac{\mathsf{p}\Delta_g^2\Delta_{\partial V}}{\lambda}\left(\Delta_\mu^\parallel\right)^2\left(1+2\Delta_\star^2\right),\end{aligned} \tag{A.2c}$$

$$\begin{aligned}\boldsymbol{\Delta}_{\circ_4}^{\boldsymbol{r}} = {} & \Delta_g\left(2m\sqrt{2\mathsf{p}}\|\nabla V(0,0)\| + \left(2n\mathsf{p}^3(4\mathsf{p}-1)\right)^{\frac{1}{2}}\frac{\Delta_{\partial V}\widehat{\Delta}_2^r}{\sqrt{\lambda}}\right)\left(\Delta_p^\parallel + \Delta_\sigma^\parallel\left(1+\Delta_\star\right)^{\frac{1}{2}}\right) \\ & + \left(p^3\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}}\frac{1}{2\lambda}\left(\sqrt{\frac{n}{2}}\Delta_g\widehat{\Delta}_1^r + \Delta_{\dot{g}}\|\nabla V(0,0)\|\right)\left(\Delta_p^\parallel + \Delta_\sigma^\parallel\left(1+\Delta_\star\right)^{\frac{1}{2}}\right),\end{aligned} \tag{A.2d}$$

$$\boldsymbol{\Delta}_{\circ_5}^{\boldsymbol{r}} = \sqrt{\frac{n}{2}}\frac{\Delta_g\Delta_{\partial V}}{2\lambda}\left[\sqrt{m}\widehat{\Delta}_4^r\widehat{\Delta}_{4_\parallel}^r + \mathsf{p}^2(2\mathsf{p}-1)\Delta_g\left(\left(\Delta_p^\parallel\right)^2 + \left(\Delta_\sigma^\parallel\right)^2\left(1+\Delta_\star\right)\right)\right]. \tag{A.2e}$$

Next, we define the constants $\Delta^r_{\circledcirc_i}$, $i \in \{1, \ldots, 4\}$, as follows:

$$\boldsymbol{\Delta^r_{\circledcirc_1}} = 2\sqrt{\mathsf{p}}\Delta_g^\perp \Delta_\sigma^\perp \left\|\nabla V(0,0)\right\|, \tag{A.3a}$$

$$\boldsymbol{\Delta^r_{\circledcirc_2}} = 2\Delta_g \sqrt{2\lambda\mathsf{p}} \left\|\nabla V(0,0)\right\| \Delta_\sigma^\parallel + \sqrt{\frac{2n\mathsf{p}}{\lambda}} \Delta_g \Delta_{\partial V} \Delta_\sigma \Delta_\mu^\parallel \left(1 + \Delta_\star\right), \tag{A.3b}$$

$$\begin{aligned}\boldsymbol{\Delta^r_{\circledcirc_3}} = \left(p^3 \frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{\Delta_\sigma^\parallel}{2\lambda} \left(\sqrt{\frac{n}{2}} \Delta_g \widehat{\Delta}_1^r + \Delta_{\dot{g}} \left\|\nabla V(0,0)\right\|\right) + 2m\sqrt{2\mathsf{p}}\Delta_g \left\|\nabla V(0,0)\right\| \Delta_\sigma^\parallel \\ + \left(2n\mathsf{p}^3(4\mathsf{p}-1)\right)^{\frac{1}{2}} \frac{\Delta_g \Delta_{\partial V}}{\sqrt{\lambda}} \left[\widehat{\Delta}_2^r \Delta_\sigma^\parallel + \Delta_\sigma \left(\Delta_p^\parallel + \Delta_\sigma^\parallel \left(1 + \Delta_\star\right)^{\frac{1}{2}}\right)\right],\end{aligned} \tag{A.3c}$$

$$\boldsymbol{\Delta^r_{\circledcirc_4}} = \sqrt{\frac{mn}{2}} \frac{\Delta_g \Delta_{\partial V}}{2\lambda} \left(\Delta_\sigma \widehat{\Delta}_{4_\parallel}^r + \Delta_\sigma^\parallel \widehat{\Delta}_4^r\right). \tag{A.3d}$$

The constants $\Delta^r_{\odot_i}$, $i \in \{1, \ldots, 5\}$, are defined as:

$$\boldsymbol{\Delta^r_{\odot_1}} = \frac{\Delta_g \Delta_\mu^\perp}{2} \left\|\nabla V(0,0)\right\| + \frac{\Delta_g \Delta_{\partial V} \Delta_\mu^\perp}{2\sqrt{2}} \left(1 + \Delta_\star\right) + \frac{\sqrt{n}\Delta_{\partial V} \Delta_\sigma^2}{4\sqrt{2}}, \tag{A.4a}$$

$$\boldsymbol{\Delta^r_{\odot_2}} = \sqrt{2\mathsf{p}}\Delta_{\partial V} \left(\Delta_g^\perp \left(\Delta_p^\perp + \Delta_\sigma^\perp \left(1 + \Delta_\star\right)^{\frac{1}{2}}\right) + \Delta_p\right), \tag{A.4b}$$

$$\begin{aligned}\boldsymbol{\Delta^r_{\odot_3}} = \sqrt{2}\Delta_g \Delta_\mu^\parallel \left[\sqrt{2} \left\|\nabla V(0,0)\right\| + \Delta_{\partial V} \left(1 + \Delta_\star\right)\right] \left(1 + \sqrt{2\lambda\mathsf{p}}\right) + \frac{\Delta_\mu^\parallel \widehat{\Delta}_3^r}{2\lambda} \left(1 + \Delta_\star\right) \\ + \frac{\Delta_\mu^\parallel}{\sqrt{\lambda}} \left(\frac{\sqrt{n}\Delta_g \widehat{\Delta}_1^r}{2\sqrt{2\lambda}} + \frac{\Delta_{\dot{g}}}{2\sqrt{\lambda}} \left\|\nabla V(0,0)\right\| + \sqrt{2n\mathsf{p}}\Delta_g \Delta_{\partial V} \widehat{\Delta}_2^r\right) \\ + 2\sqrt{\lambda\mathsf{p}}\Delta_g \Delta_{\partial V} \left(\Delta_p^\parallel + \Delta_\sigma^\parallel \left(1 + \Delta_\star\right)^{\frac{1}{2}}\right),\end{aligned} \tag{A.4c}$$

$$\begin{aligned}\boldsymbol{\Delta^r_{\odot_4}} = \left(2m\sqrt{\mathsf{p}}\Delta_g \Delta_{\partial V} + \left(p^3 \frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{\widehat{\Delta}_3^r}{2\lambda}\right) \left(\Delta_p^\parallel + \Delta_\sigma^\parallel \left(1 + \Delta_\star\right)^{\frac{1}{2}}\right) \\ + \left(2n\mathsf{p}^3(4\mathsf{p}-1)\right)^{\frac{1}{2}} \frac{\Delta_g \Delta_{\partial V} \Delta_\sigma \Delta_\sigma^\parallel}{\sqrt{\lambda}},\end{aligned} \tag{A.4d}$$

$$\boldsymbol{\Delta^r_{\odot_5}} = \sqrt{\frac{n}{2}} \frac{\Delta_g \Delta_{\partial V} \Delta_\sigma^\parallel}{2\lambda} \left(\sqrt{m}\Delta_\sigma + \mathsf{p}^2(2\mathsf{p}-1)\Delta_g \Delta_\sigma^\parallel\right), \tag{A.4e}$$

while the constants $\Delta^r_{\otimes_i}$, $i \in \{1, \ldots, 3\}$, are defined as follows:

$$\boldsymbol{\Delta^r_{\otimes_1}} = \sqrt{2\mathsf{p}}\Delta_{\partial V} \Delta_g^\perp \Delta_\sigma^\perp, \tag{A.5a}$$

$$\boldsymbol{\Delta^r_{\otimes_2}} = \sqrt{2\mathsf{p}}\Delta_g \Delta_{\partial V} \left(\sqrt{2\lambda}\Delta_\sigma^\parallel + \sqrt{\frac{n}{\lambda}} \Delta_\sigma \Delta_\mu^\parallel\right), \tag{A.5b}$$

$$\boldsymbol{\Delta^r_{\otimes_3}} = \Delta_\sigma^\parallel \left(2m\sqrt{\mathsf{p}}\Delta_g \Delta_{\partial V} + \left(p^3 \frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{\widehat{\Delta}_3^r}{2\lambda}\right). \tag{A.5c}$$

Finally, we define the constants $\Delta^r_{\circledast_i}$, $i \in \{1, 2\}$, as follows:

$$\boldsymbol{\Delta^r_{\circledast_1}} = \frac{\Delta_g \Delta_{\partial V} \Delta_\mu^\perp}{2\sqrt{2}}, \tag{A.6a}$$

$$\boldsymbol{\Delta^r_{\circledast_2}} = \Delta_\mu^\parallel \left[\sqrt{2}\Delta_g \Delta_{\partial V} \left(1 + \sqrt{2\lambda\mathsf{p}} + \frac{\sqrt{n}\mathsf{p}\Delta_g \Delta_\mu^\parallel}{\lambda}\right) + \frac{\widehat{\Delta}_3^r}{2\lambda}\right]. \tag{A.6b}$$

B Technical Results

We begin with the following result that is a multidimensional version of the stochastic integration by parts formula in [100, Thm. 1.6.5].

Lemma B.1 *Consider any $L \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{l\times l'}|\mathfrak{F}_t\right)$ and $S \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{l'\times n_q}|\mathfrak{F}_t\right)$, 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*

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

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

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

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

As a consequence of the above definition, we have that

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

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

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

where $\hat{L}_{1,\cdots,l} \in \mathbb{R}^{1\times l'}$ and $S_{1,\cdots,l'} \in \mathbb{R}^{1\times n_q}$ 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:

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

Using this definition, we derive the following expression:

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

for $(j, i, t) \in \{1, \ldots, l\} \times \{1, \ldots, l'\} \times [0, T]$.

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

$$\hat{L}^+_{j,i}(t)^+(s) \doteq \begin{cases} \hat{L}_{j,i}(t) = \int_0^t L_{j,i}(\varsigma)d\varsigma, & \text{if } \hat{L}_{j,i}(t) > 0, \\ 0, & \text{otherwise} \end{cases},$$

$$\hat{L}^-_{j,i}(t) \doteq \begin{cases} -\hat{L}_{j,i}(t) = -\int_0^t L_{j,i}(\varsigma)d\varsigma, & \text{if } \hat{L}_{j,i}(t) < 0, \\ 0, & \text{otherwise} \end{cases}$$

that is, $\hat{L}^+_{j,i}(t)$ and $\hat{L}^-_{j,i}(t)$ are the positive and negative parts of $\hat{L}_{j,i}(t)$, respectively. Then, since both $\hat{L}^+_{j,i}(t)$ and $\hat{L}^-_{j,i}(t)$ are increasing processes, we have that $\hat{L}_{j,i}(t)$ is a finite variation process [100, Sec. 1.3]. Therefore we may use the integration by parts formula [100, 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

$$\begin{aligned}\int_0^t \left(\hat{L}_{j,i}(t) - \hat{L}_{j,i}(\nu)\right) S_i(\nu)dQ_\nu &= \hat{L}_{j,i}(t)\hat{S}_i(t) - \int_0^t \hat{L}_{j,i}(\nu)d\hat{S}_i(\nu) \\ &= \int_0^t \hat{S}_i(\nu)d\hat{L}_{j,i}(\nu), \quad (j,i,t) \in \{1,\dots,l\}\times\{1,\dots,l'\}\times[0,T]. \end{aligned} \tag{B.7}$$

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 [106, Thm. 7.7.1] to (B.7) and obtain

$$\begin{aligned}\int_0^t \left(\hat{L}_{j,i}(t) - \hat{L}_{j,i}(\nu)\right) S_i(\nu)dQ_\nu &= \int_0^t \hat{S}_i(\nu)d\hat{L}_{j,i}(\nu) \\ &= \int_0^t \hat{S}_i(\nu)L_{j,i}(\nu)d\nu, \quad (j,i,t) \in \{1,\dots,l\}\times\{1,\dots,l'\}\times[0,T].\end{aligned}$$

The proof of (B.1a) is then concluded by substituting the above expression into (B.4) to obtain

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

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

The following oft used result is a straightforward consequence of the Burkholder-Davis-Gundy inequality [93, Chp. 3.5], [107].

Proposition B.1 *Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\mathfrak{F}_t$, and any $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{m\times n_q}|\mathfrak{F}_t)$, any constant $\theta \in \mathbb{R}$, and an $\mathfrak{F}_t$-adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$. Then, the following bound holds for any constant $t' \in [0,T]$ and $\mathsf{p} \geq 1$,*

$$\left\|\left(M\right)_{t'}\right\|_{\mathsf{p}} \leq 2m\sqrt{\mathsf{p}}\left(\frac{e^{2\theta t'}-1}{\theta}\right)^{\frac{1}{2}} \left\|\left(L\right)_{t'}\right\|_{\mathsf{p}}, \quad M_t \doteq \int_0^t e^{\theta\nu}L(\nu)dQ_\nu \in \mathbb{R}^m, \quad t \in \mathbb{R}_{\geq 0}. \tag{B.8}$$

Proof. The equivalence of the finite-dimensional norms $\|\cdot\| \doteq \|\cdot\|_2 \leq \|\cdot\|_1$ implies that

$$\left(M\right)_{t'} \doteq \sup_{t\in[0,t']}\|M_t\| \leq \sup_{t\in[0,t']}\|M_t\|_1 = \sup_{t\in[0,t']}\sum_{i=1}^m |M_{i,t}| \leq \sum_{i=1}^m \sup_{t\in[0,t']}|M_{i,t}| \doteq \sum_{i=1}^m \left(M_i\right)_{t'},$$

where $M_{i,t} \in \mathbb{R}$ is the i-th, $i \in \{1,\dots,m\}$, component of $M_t \in \mathbb{R}^m$. Applying the Minkowski's inequality leads to the following bound:

$$\left\|\left(M\right)_{t'}\right\|_{\mathsf{p}} \leq \sum_{i=1}^m \left\|\left(M_i\right)_{t'}\right\|_{\mathsf{p}}. \tag{B.9}$$

Next, using the definition of a continuous local martingale [59, Def. 1.5.15], we use the fact that $L \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{m \times n_q} | \mathfrak{F}_t\right)$ and invoke [108, Thm. 5.5.2] to conclude that $M_{i,t}$ is a continuous local martingale with respect to the filtration $\mathfrak{F}_t$ and satisfies $M_{i,0} = 0$, for all $i \in \{1, \ldots, m\}$. Furthermore, since $t' \in \mathbb{R}_{>0}$ is constant it is a stopping time with respect to the filtration $\mathfrak{F}_t$ [95, Sec. 6.1.1]. Thus, we may invoke the Burkholder-Davis-Gundy inequality [107, Thm. 2] to obtain the following bound:

$$\left\| \left(M_i \right)_{t'} \right\|_{\mathsf{p}} \leq \sqrt{8\mathsf{p}} \left\| \left\langle M_i \right\rangle_{t'}^{\frac{1}{2}} \right\|_{\mathsf{p}}, \quad \forall i \in \{1, \ldots, m\}, \tag{B.10}$$

where $\left\langle M_i \right\rangle_{t'}$ is the quadratic variation process of of $M_{i,t}$ evaluated at t' [59, Prop. 3.2.10].

We next bound the quadratic variation process as follows:

$$\left\langle M_i \right\rangle_{t'} = \int_0^{t'} e^{2\theta\nu} \left\| L_i(\nu)^\top \right\|^2 d\nu \leq \int_0^{t'} e^{2\theta\nu} d\nu \left(L_i^\top \right)_{t'}^2 = \frac{e^{2\theta t'} - 1}{2\theta} \left(L_i^\top \right)_{t'}^2, \quad \forall i \in \{1, \ldots, m\},$$

where $L_i(\nu) \in \mathbb{R}^{1 \times n_q}$ denotes the i-th row of $L(\nu) \in \mathbb{R}^{m \times n_q}$. Substituting into (B.10) and using the fact that t' is a constant produces

$$\left\| \left(M_i \right)_{t'} \right\|_{\mathsf{p}} \leq \sqrt{8\mathsf{p}} \left\| \left\langle M_i \right\rangle_{t'}^{\frac{1}{2}} \right\|_{\mathsf{p}} \leq \left(4\mathsf{p} \frac{e^{2\theta t'} - 1}{\theta} \right)^{\frac{1}{2}} \left\| \left(L_i^\top \right)_{t'} \right\|_{\mathsf{p}}, \quad \forall i \in \{1, \ldots, m\},$$

which upon further substitution into (B.9) produces the following bound:

$$\left\| \left(M \right)_{t'} \right\|_{\mathsf{p}} \leq \sum_{i=1}^{m} \left\| \left(M_i \right)_{t'} \right\|_{\mathsf{p}} \leq \left(4\mathsf{p} \frac{e^{2\theta t'} - 1}{\theta} \right)^{\frac{1}{2}} \sum_{i=1}^{m} \left\| \left(L_i^\top \right)_{t'} \right\|_{\mathsf{p}}. \tag{B.11}$$

Now, consider $\vec{L} = \begin{bmatrix} \left\| \left(L_1^\top \right)_{t'} \right\|_{\mathsf{p}} & \cdots & \left\| \left(L_m^\top \right)_{t'} \right\|_{\mathsf{p}} \end{bmatrix}^\top \in \mathbb{R}^m$. Then, using the equivalence of the finite-dimensional norms $\|\cdot\|_1 \leq \sqrt{m} \|\cdot\|_2 \doteq \|\cdot\|$ and the definition of the Frobenius norm, we obtain

We may bound the term $\left\| L_i(t)^\top \right\|$ using the definition of the Frobenius norm as follows:

$$\left\| L_i(t)^\top \right\|^2 \leq \sum_{i=1}^{m} \left\| L_i(t)^\top \right\|^2 = \| L(t) \|_F^2, \quad \forall i \in \{1, \ldots, m\},$$

and thus

$$\left(L_i^\top \right)_{t'} = \sup_{t \in [0,t']} \left\| L_i(t)^\top \right\| \leq \sup_{t \in [0,t']} \| L(t) \|_F = \left(L \right)_{t'}, \quad \forall i \in \{1, \ldots, m\}.$$

The bound in(B.8) is then established by substituting the above bound into the right hand side of (B.11).

□

The next result provides a moment bound for the process M_t defined in Proposition B.1.

Lemma B.2 *Let $\mathsf{p} \in \mathbb{N}_{\geq 1}$, and suppose that $L \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{m \times n_q} | \mathfrak{F}_t\right)$ defined in Proposition B.1 satisfies*

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

Then, we have the following bound on the moments of the process M_t defined in Proposition B.1:

$$\mathbb{E}\left[\left\| \int_0^t e^{\theta\nu} L(\nu) dQ_\nu \right\|^{2\mathsf{p}} \right] \leq \left(\mathsf{p} \frac{2\mathsf{p} - 1}{2} \right)^{\mathsf{p}} \left(\frac{e^{2\theta t} - 1}{\theta} \right)^{\mathsf{p}} \mathbb{E}\left[\left(L \right)_t^{2\mathsf{p}} \right], \quad \forall t \in [0, T], \quad \mathsf{p} \geq 1, \tag{B.12}$$

where $\theta \in \mathbb{R}$ is a constant and $Q_t \in \mathbb{R}^{n_q}$ is an $\mathfrak{F}_t$-adapted Brownian motion as in the statement of Proposition B.1.

Proof. We begin with the case when $\mathsf{p} = 1$, and use [100, Thm. 1.5.21] to obtain

$$\mathbb{E}\left[\left\|\int_0^t e^{\theta\nu} L(\nu) dQ_\nu\right\|^2\right] = \mathbb{E}\left[\int_0^t e^{2\theta\nu} \|L(\nu)\|_F^2 \, d\nu\right].$$

Therefore,

$$\mathbb{E}\left[\left\|\int_0^t e^{\theta\nu} L(\nu) dQ_\nu\right\|^2\right] \leq \mathbb{E}\left[\int_0^t e^{2\theta\nu} d\nu \left(L\right)_t^2\right] = \mathbb{E}\left[\left(e^{2\theta t} - 1\right) \frac{1}{2\theta} \left(L\right)_t^2\right],$$

thus establishing (B.12) for $\mathsf{p} = 1$.

Now, let $\mathsf{p} \in \mathbb{N}_{\geq 2}$, and for $t \in [0, T]$, set

$$N(t) = \int_0^t e^{\theta\nu} L(\nu) dQ_\nu \in \mathbb{R}^m.$$

Then, as in the proof of [100, Thm. 1.7.1], we use Itô's lemma and [100, Thm. 1.5.21] to obtain

$$\begin{aligned} &\mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right] \\ &= \mathsf{p}\mathbb{E}\left[\int_0^t \left(\|N(\nu)\|^{2(\mathsf{p}-1)} e^{2\theta\nu} \|L(\nu)\|_F^2 + 2(\mathsf{p}-1) \|N(\nu)\|^{2(\mathsf{p}-2)} e^{2\theta\nu} \left\|N(\nu)^\top L(\nu)\right\|^2\right) d\nu\right]. \end{aligned} \tag{B.13}$$

Thus, we obtain the following bound

$$\begin{aligned} \mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right] \leq& \mathsf{p}(2\mathsf{p}-1)\mathbb{E}\left[\int_0^t \|N(\nu)\|^{2(\mathsf{p}-1)} e^{2\theta\nu} \|L(\nu)\|_F^2 \, d\nu\right] \\ =& \mathsf{p}(2\mathsf{p}-1)\mathbb{E}\left[\int_0^t e^{2\theta\nu(\mathsf{p}-1)/\mathsf{p}} \|N(\nu)\|^{2(\mathsf{p}-1)} e^{2\theta\nu/\mathsf{p}} \|L(\nu)\|_F^2 \, d\nu\right], \end{aligned}$$

where, in order to obtain the last equality, we have used $e^{2\theta\nu} = e^{2\theta\nu[(\mathsf{p}-1)/\mathsf{p}+1/\mathsf{p}]} = e^{2\theta\nu(\mathsf{p}-1)/\mathsf{p}} e^{2\theta\nu/\mathsf{p}}$. Using Hölder's inequality with conjugates $\mathsf{p}/(\mathsf{p}-1)$ and p, one sees that

$$\begin{aligned} \mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right] &\leq \mathsf{p}(2\mathsf{p}-1)\mathbb{E}\left[\int_0^t e^{2\theta\nu} \|N(\nu)\|^{2\mathsf{p}} d\nu\right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\int_0^t e^{2\theta\nu} \|L(\nu)\|_F^{2\mathsf{p}} d\nu\right]^{\frac{1}{\mathsf{p}}} \\ &= \mathsf{p}(2\mathsf{p}-1)\left(\int_0^t e^{2\theta\nu} \mathbb{E}\left[\|N(\nu)\|^{2\mathsf{p}}\right] d\nu\right)^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\int_0^t e^{2\theta\nu} \|L(\nu)\|_F^{2\mathsf{p}} d\nu\right]^{\frac{1}{\mathsf{p}}}. \end{aligned}$$

Note from (B.13) that $\mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right]$ is non-decreasing in t. It then follows that

$$\begin{aligned} \mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right] \leq& \mathsf{p}(2\mathsf{p}-1)\left(\int_0^t e^{2\theta\nu} d\nu\right)^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\int_0^t e^{2\theta\nu} \|L(\nu)\|_F^{2\mathsf{p}} d\nu\right]^{\frac{1}{\mathsf{p}}} \\ \leq& \mathsf{p}(2\mathsf{p}-1)\left(\int_0^t e^{2\theta\nu} d\nu\right)^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\int_0^t e^{2\theta\nu} d\nu \left(L\right)_t^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \\ =& \mathsf{p}(2\mathsf{p}-1)\left(\int_0^t e^{2\theta\nu} d\nu\right)^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \left(\int_0^t e^{2\theta\nu} d\nu\right)^{\frac{1}{\mathsf{p}}} \mathbb{E}\left[\left(L\right)_t^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \\ =& \mathsf{p}(2\mathsf{p}-1)\mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \left(\int_0^t e^{2\theta\nu} d\nu\right) \mathbb{E}\left[\left(L\right)_t^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}}, \end{aligned}$$

and thus

$$\mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right]^{\mathsf{p}} \leq (\mathsf{p}(2\mathsf{p}-1))^{\mathsf{p}} \, \mathbb{E}\left[\|N(\nu)\|^{2\mathsf{p}}\right]^{\mathsf{p}-1} \left(\int_0^t e^{2\theta\nu} d\nu\right)^{\mathsf{p}} \mathbb{E}\left[\left(L\right)_t^{2\mathsf{p}}\right].$$

The above inequality implies

$$\mathbb{E}\left[\|N(t)\|^{2\mathsf{p}}\right] \leq (\mathsf{p}(2\mathsf{p}-1))^{\mathsf{p}} \left(\int_0^t e^{2\theta\nu} d\nu\right)^{\mathsf{p}} \mathbb{E}\left[\left(L\right)_t^{2\mathsf{p}}\right] = \left(e^{2\theta t} - 1\right)^{\mathsf{p}} \left(\frac{\mathsf{p}(2\mathsf{p}-1)}{2\theta}\right)^{\mathsf{p}} \mathbb{E}\left[\left(L\right)_t^{2\mathsf{p}}\right],$$

thus completing the proof. ☐

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 $S \in \mathcal{M}_2^{loc}(\mathbb{R}^m|\mathfrak{F}_t)$ and $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{m\times n_q}|\mathfrak{F}_t)$ satisfy*

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

Additionally, for any strictly positive constants θ_1, $\theta_2 \in \mathbb{R}_{>0}$ and an $\mathfrak{F}_t$-adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$ define

$$N(t) = \int_0^t e^{(\theta_1-\theta_2)\nu} S(\nu)^\top \left(\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right) d\nu \in \mathbb{R}, \quad t \in [0,T]. \tag{B.14}$$

Then, the following bound holds for any constant $t' \in [0,T]$:

$$\left\|\left(N\right)_{t'}\right\|_{\mathsf{p}} \leq \left(\mathsf{p}^3 \frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{e^{\theta_1 t'}}{\theta_1\sqrt{\theta_2}} \left\|\left(L\right)_{t'}\right\|_{2\mathsf{p}} \left\|\left(S\right)_{t'}\right\|_{2\mathsf{p}}, \quad \forall \mathsf{p} \in \mathbb{N}_{\geq 1}. \tag{B.15}$$

Proof. We begin with the case when $\mathsf{p} = 1$. Note that the definition of $N(t)$ implies that

$$|N(t)| \leq \int_0^t e^{(\theta_1-\theta_2)\nu} \|S(\nu)\| \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\| d\nu, \quad t \in [0,T].$$

It then follows from the non-negativity of the integrand of the Lebesgue integral that for all $t \in [0,t']$

$$|N(t)| \leq \int_0^{t'} e^{(\theta_1-\theta_2)\nu} \|S(\nu)\| \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\| d\nu, \quad \forall t \in [0,t'],$$

which in turn implies that

$$\sup_{t\in[0,t']} |N(t)| \doteq \left(N\right)_{t'} \leq \int_0^{t'} e^{(\theta_1-\theta_2)\nu} \|S(\nu)\| \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\| d\nu.$$

Hence

$$\begin{aligned}\mathbb{E}\left[\left(N\right)_{t'}\right] &\leq \mathbb{E}\left[\int_0^{t'} e^{(\theta_1-\theta_2)\nu} \|S(\nu)\| \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\| d\nu\right] \\ &= \int_0^{t'} e^{(\theta_1-\theta_2)\nu} \mathbb{E}\left[\|S(\nu)\| \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\|\right] d\nu.\end{aligned}$$

Using the Cauchy-Schwarz inequality one then sees that

$$\begin{aligned}\mathbb{E}\left[\left(N\right)_{t'}\right] &\leq \int_0^{t'} e^{(\theta_1-\theta_2)\nu} \mathbb{E}\left[\|S(\nu)\|^2\right]^{\frac{1}{2}} \mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^2\right]^{\frac{1}{2}} d\nu \\ &\leq \left(\int_0^{t'} e^{(\theta_1-\theta_2)\nu} \mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^2\right]^{\frac{1}{2}} d\nu\right) \mathbb{E}\left[\left(S\right)_{t'}^2\right]^{\frac{1}{2}}.\end{aligned}$$

Using Lemma B.2 on the inner expectation produces

$$\begin{aligned}\mathbb{E}\left[\left(N\right)_{t'}\right] &\leq \left(\int_0^{t'} e^{(\theta_1-\theta_2)\nu} \left(e^{2\theta_2\nu}-1\right)^{\frac{1}{2}} \left(\frac{1}{2\theta_2}\right)^{\frac{1}{2}} \mathbb{E}\left[\left(L\right)_\nu^2\right]^{\frac{1}{2}} d\nu\right) \mathbb{E}\left[\left(S\right)_{t'}^2\right]^{\frac{1}{2}} \\ &\leq \left(\int_0^{t'} e^{(\theta_1-\theta_2)\nu} \left(e^{2\theta_2\nu}-1\right)^{\frac{1}{2}} d\nu\right) \left(\frac{1}{2\theta_2}\right)^{\frac{1}{2}} \mathbb{E}\left[\left(L\right)_{t'}^2\right]^{\frac{1}{2}} \mathbb{E}\left[\left(S\right)_{t'}^2\right]^{\frac{1}{2}},\end{aligned}$$

where the last inequality follows from the fact that $[0,\nu] \subseteq [0,t']$ and thus $\sup_{\beta\in[0,\nu]} \|L(\beta)\|_F \leq \sup_{\beta\in[0,t']} \|L(\beta)\|_F$. The bound in (B.15) is then established for $\mathsf{p} = 1$ by solving the integral term as follows:

$$\int_0^{t'} e^{(\theta_1-\theta_2)\nu} \left(e^{2\theta_2\nu}-1\right)^{\frac{1}{2}} d\nu = \int_0^{t'} e^{\theta_1\nu} \left(e^{-2\theta_2\nu}\right)^{\frac{1}{2}} \left(e^{2\theta_2\nu}-1\right)^{\frac{1}{2}} d\nu = \int_0^{t'} e^{\theta_1\nu} \left(1-e^{-2\theta_2\nu}\right)^{\frac{1}{2}} d\nu,$$

which, upon using the strict positivity of θ_1 and θ_2, yields

$$\int_0^{t'} e^{(\theta_1-\theta_2)\nu}\left(e^{2\theta_2\nu}-1\right)^{\frac{1}{2}} d\nu = \int_0^{t'} e^{\theta_1\nu}\left(1-e^{-2\theta_2\nu}\right)^{\frac{1}{2}} d\nu \leq \int_0^{t'} e^{\theta_1\nu} d\nu = \left(e^{\theta_1 t'}-1\right)\frac{1}{\theta_1} \leq e^{\theta_1 t'}\frac{1}{\theta_1}.$$

Now, let $\mathsf{p} \in \mathbb{N}_{\geq 2}$. Since the outer integral in the definition of $N(t)$ is a standard Lebesgue integral with a t-continuous integrand, the fundamental theorem for the Lebesgue integral [106, Thm. 6.4.1] implies that $N(t)$ is absolutely continuous on $[0,T]$ and

$$dN(t) = e^{(\theta_1-\theta_2)t} S(t)^\top \left(\int_0^t e^{\theta_2\beta} L(\beta) dQ_\beta\right) dt, \quad \mu_L\text{-a.e. on } [0,T],$$

where μ_L denotes the Lebesgue measure. Using the chain rule one then sees that

$$\begin{aligned} |N(t)|^{\mathsf{p}} &= \mathsf{p}\int_0^t |N(\nu)|^{\mathsf{p}-2} N(\nu) e^{(\theta_1-\theta_2)\nu} S(\nu)^\top \left(\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right) d\nu \\ &\leq \mathsf{p}\int_0^t |N(\nu)|^{\mathsf{p}-1} e^{(\theta_1-\theta_2)\nu} \|S(\nu)\| \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\| d\nu, \quad t \in [0,T]. \end{aligned}$$

Since the right hand side is a non-decreasing function over $t \in [0,t']$, it follows then

$$|N(t)|^{\mathsf{p}} \leq \mathsf{p}\int_0^{t'} |N(\nu)|^{\mathsf{p}-1} e^{(\theta_1-\theta_2)\nu} \|S(\nu)\| \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\| d\nu, \quad t \in [0,t'],$$

and thus

$$\left(N\right)_{t'}^{\mathsf{p}} \leq \mathsf{p}\int_0^{t'} |N(\nu)|^{\mathsf{p}-1} e^{(\theta_1-\theta_2)\nu} \|S(\nu)\| \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\| d\nu.$$

Using the above bound we obtain the following inequality:

$$\begin{aligned} \mathbb{E}\left[\left(N\right)_{t'}^{\mathsf{p}}\right] &\leq \mathsf{p}\,\mathbb{E}\left[\int_0^{t'} |N(\nu)|^{\mathsf{p}-1} e^{(\theta_1-\theta_2)\nu} \|S(\nu)\| \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\| d\nu\right] \\ &\overset{(i)}{=} \mathsf{p}\,\mathbb{E}\left[\int_0^{t'} \left(|N(\nu)|^{\mathsf{p}-1} e^{\theta_1(\mathsf{p}-1)\nu/\mathsf{p}}\right)\left(e^{(\theta_1/\mathsf{p}-\theta_2)\nu} \|S(\nu)\| \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\|\right) d\nu\right] \\ &\overset{(ii)}{\leq} \mathsf{p}\,\mathbb{E}\left[\int_0^{t'} |N(\nu)|^{\mathsf{p}} e^{\theta_1\nu} d\nu\right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \|S(\nu)\|^{\mathsf{p}} \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{\mathsf{p}} d\nu\right]^{\frac{1}{\mathsf{p}}}, \end{aligned} \tag{B.16}$$

where we obtain (i) by writing

$$e^{(\theta_1-\theta_2)\nu} = e^{\left[\theta_1\left(\frac{\mathsf{p}-1}{\mathsf{p}}+\frac{1}{\mathsf{p}}\right)-\theta_2\right]\nu} = e^{\theta_1(\mathsf{p}-1)\nu/\mathsf{p}} e^{(\theta_1/\mathsf{p}-\theta_2)\nu},$$

and (ii) follows from Hölder's inequality with conjugates $\mathsf{p}/(\mathsf{p}-1)$ and p.

Now, we have the following straightforward inequality:

$$\mathbb{E}\left[\int_0^{t'} |N(\nu)|^{\mathsf{p}} e^{\theta_1\nu} d\nu\right] \leq \mathbb{E}\left[\left(\int_0^{t'} e^{\theta_1\nu} d\nu\right)\left(N\right)_{t'}^{\mathsf{p}}\right] = \left(\int_0^{t'} e^{\theta_1\nu} d\nu\right)\mathbb{E}\left[\left(N\right)_{t'}^{\mathsf{p}}\right], \tag{B.17}$$

where we have used the fact that t' is a constant.

Next, using the Cauchy-Schwarz inequality one sees that

$$\begin{aligned} &\mathbb{E}\left[\int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \|S(\nu)\|^{\mathsf{p}} \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{\mathsf{p}} d\nu\right] \\ &\qquad = \int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \mathbb{E}\left[\|S(\nu)\|^{\mathsf{p}} \left\|\int_0^\nu e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{\mathsf{p}}\right] d\nu \end{aligned}$$

$$\leq \int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \mathbb{E}\left[\|S(\nu)\|^{2\mathsf{p}}\right]^{\frac{1}{2}} \mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{2\mathsf{p}}\right]^{\frac{1}{2}} d\nu.$$

We develop the bound further as follows:

$$\mathbb{E}\left[\int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \|S(\nu)\|^{\mathsf{p}} \left\|\int_0^{\nu} e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{\mathsf{p}} d\nu\right] \leq \left(\int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{2\mathsf{p}}\right]^{\frac{1}{2}} d\nu\right) \mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2}},$$

which, upon using Lemma B.2 on the inner expectation, leads to

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \|S(\nu)\|^{\mathsf{p}} \left\|\int_0^{\nu} e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{\mathsf{p}} d\nu\right] \\
&\leq \left(\mathsf{p}\frac{2\mathsf{p}-1}{2\theta_2}\right)^{\frac{\mathsf{p}}{2}} \left(\int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \left(e^{2\theta_2\nu}-1\right)^{\frac{\mathsf{p}}{2}} \mathbb{E}\left[\left(L\right)_{\nu}^{2\mathsf{p}}\right]^{\frac{1}{2}} d\nu\right) \mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2}} \\
&\leq \left(\mathsf{p}\frac{2\mathsf{p}-1}{2\theta_2}\right)^{\frac{\mathsf{p}}{2}} \left(\int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \left(e^{2\theta_2\nu}-1\right)^{\frac{\mathsf{p}}{2}} d\nu\right) \mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2}} \mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2}}, \quad \text{(B.18)}
\end{aligned}$$

where the last inequality follows from the fact that $[0,\nu] \subseteq [0,t']$ and thus $\sup_{\beta\in[0,\nu]} \|L(\beta)\|_F \leq \sup_{\beta\in[0,t']} \|L(\beta)\|_F$. The integral term can be bounded as

$$\begin{aligned}
\int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \left(e^{2\theta_2\nu}-1\right)^{\frac{\mathsf{p}}{2}} d\nu = \int_0^{t'} e^{\theta_1\nu} \left(e^{-2\theta_2\nu}\right)^{\frac{\mathsf{p}}{2}} \left(e^{2\theta_2\nu}-1\right)^{\frac{\mathsf{p}}{2}} d\nu &= \int_0^{t'} e^{\theta_1\nu} \left(1-e^{-2\theta_2\nu}\right)^{\frac{\mathsf{p}}{2}} d\nu \\
&\leq \int_0^{t'} e^{\theta_1\nu} d\nu,
\end{aligned}$$

where we have used the strict positivity of θ_2 to obtain the last inequality. Substituting this bound into (B.18) produces

$$\mathbb{E}\left[\int_0^{t'} e^{(\theta_1-\theta_2\mathsf{p})\nu} \|S(\nu)\|^{\mathsf{p}} \left\|\int_0^{\nu} e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{\mathsf{p}} d\nu\right] \leq \left(\mathsf{p}\frac{2\mathsf{p}-1}{2\theta_2}\right)^{\frac{\mathsf{p}}{2}} \left(\int_0^{t'} e^{\theta_1\nu} d\nu\right) \mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2}} \mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2}}, \quad \text{(B.19)}$$

Substituting (B.17) and (B.19) into (B.16) produces

$$\mathbb{E}\left[\left(N\right)_{t'}^{\mathsf{p}}\right] \leq \mathbb{E}\left[\left(N\right)_{t'}^{\mathsf{p}}\right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \left[\left(\int_0^{t'} e^{\theta_1\nu} d\nu\right) \left(\mathsf{p}^3\frac{2\mathsf{p}-1}{2\theta_2}\right)^{\frac{1}{2}} \mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}\right].$$

The proof is then completed by concluding from the above inequality that

$$\begin{aligned}
\mathbb{E}\left[\left(N\right)_{t'}^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} &\leq \left(\int_0^{t'} e^{\theta_1\nu} d\nu\right) \left(\mathsf{p}^3\frac{2\mathsf{p}-1}{2\theta_2}\right)^{\frac{1}{2}} \mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \\
&= \left(e^{\theta_1 t'}-1\right) \frac{1}{\theta_1\sqrt{\theta_2}} \left(\mathsf{p}^3\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \\
&\leq e^{\theta_1 t'} \frac{1}{\theta_1\sqrt{\theta_2}} \left(\mathsf{p}^3\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}},
\end{aligned}$$

where we have used the strict positivity of θ_1 to obtain the last inequality.

□

The following corollary to the last lemma considers a particular type of integral that we encounter due to the structure of the control law in Sec. 3.1.

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

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

Additionally, for any strictly positive constants θ_1, $\theta_2 \in \mathbb{R}_{>0}$ and an $\mathfrak{F}_t$-adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$ define

$$\breve{N}(t) = \int_0^t e^{2(\theta_1-\theta_2)\nu}\, \breve{R}(\nu)^\top R(\nu)\, \breve{R}(\nu) d\nu \in \mathbb{R}, \quad t \in [0,T], \tag{B.20}$$

where

$$\breve{R}(t) = \int_0^t e^{\theta_2\beta}\left[S(\beta)d\beta + L(\beta)dQ_\beta\right] \in \mathbb{R}^m.$$

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 any constant $t' \in [0,T]$ and for all $\mathsf{p} \in \mathbb{N}_{\geq 1}$:

$$\left\|\left(\breve{N}\right)_{t'}\right\|_{\mathsf{p}} \leq \frac{\mathsf{p}\Delta_R e^{2\theta_1 t'}}{2\theta_1}\left(\frac{1}{\theta_2}\left\|\left(S\right)_{t'}\right\|_{2\mathsf{p}} + \left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}}\frac{1}{\sqrt{\theta_2}}\left\|\left(L\right)_{t'}\right\|_{2\mathsf{p}}\right)^2. \tag{B.21}$$

Proof. The proof follows the identical line of reasoning as that of Lemma B.3. We begin with the case when $\mathsf{p} = 1$, for which we have

$$\left|\breve{N}(t)\right| \leq \int_0^t e^{2(\theta_1-\theta_2)\nu}\,\|R(\nu)\|_F \left\|\breve{R}(\nu)\right\|^2 d\nu \leq \Delta_R \int_0^t e^{2(\theta_1-\theta_2)\nu}\left\|\breve{R}(\nu)\right\|^2 d\nu, \quad t \in [0,T].$$

Since the right hand side is a non-decreasing function over $t \in [0,t']$,

$$\left|\breve{N}(t)\right| \leq \Delta_R \int_0^{t'} e^{2(\theta_1-\theta_2)\nu}\left\|\breve{R}(\nu)\right\|^2 d\nu, \quad t \in [0,T],$$

and thus

$$\left(\breve{N}\right)_{t'} \leq \Delta_R \int_0^{t'} e^{2(\theta_1-\theta_2)\nu}\left\|\breve{R}(\nu)\right\|^2 d\nu.$$

Consequently,

$$\mathbb{E}\left[\left(\breve{N}\right)_{t'}\right] \leq \Delta_R \int_0^{t'} e^{2(\theta_1-\theta_2)\nu}\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^2\right] d\nu. \tag{B.22}$$

Next, it follows from the definition of $\breve{R}(t)$ that

$$\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^2\right] = \mathbb{E}\left[\left(\breve{R}(\nu)^\top\breve{R}(\nu)\right)\right],$$

which leads to the following bound:

$$\begin{aligned}\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^2\right] &\leq \mathbb{E}\left[\left(\left\|\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right\| + \left\|\int_0^\nu e^{\theta_2\beta}L(\beta)dQ_\beta\right\|\right)^2\right] \\ &= \mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right\|^2\right] + \mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta}L(\beta)dQ_\beta\right\|^2\right]\end{aligned}$$

$$+ 2\mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right\|\left\|\int_0^\nu e^{\theta_2\beta}L(\beta)dQ_\beta\right\|\right],$$

for $0 \le \nu \le t \le T$. It then follows from the Cauchy-Schwarz inequality that

$$\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^2\right] \le \mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right\|^2\right] + \mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta}L(\beta)dQ_\beta\right\|^2\right] + 2\mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right\|^2\right]^{\frac{1}{2}}\mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta}L(\beta)dQ_\beta\right\|^2\right]^{\frac{1}{2}}, \quad \text{(B.23)}$$

for $0 \le \nu \le t \le T$. Now, we have the following inequality:

$$\mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right\|^2\right] \le \mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta}d\beta\right\|^2\left(S\right)_\nu^2\right] \le \left(\frac{e^{\theta_2\nu}}{\theta_2}\right)^2\mathbb{E}\left[\left(S\right)_\nu^2\right], \quad \forall 0 \le \nu \le t \le T. \quad \text{(B.24)}$$

Similarly, using Lemma B.2, we have

$$\mathbb{E}\left[\left\|\int_0^\nu e^{\theta_2\beta}L(\beta)dQ_\beta\right\|^2\right] \le \frac{e^{2\theta_2\nu}}{2\theta_2}\mathbb{E}\left[\left(L\right)_\nu^2\right], \quad \forall 0 \le \nu \le t \le T. \quad \text{(B.25)}$$

Substituting the bounds (B.24) and (B.25) into (B.23) produces

$$\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^2\right] \le \frac{e^{2\theta_2\nu}}{\theta_2^2}\mathbb{E}\left[\left(S\right)_\nu^2\right] + \frac{e^{2\theta_2\nu}}{2\theta_2}\mathbb{E}\left[\left(L\right)_\nu^2\right] + \sqrt{2}\frac{e^{2\theta_2\nu}}{\theta_2\sqrt{\theta_2}}\mathbb{E}\left[\left(S\right)_\nu^2\right]^{\frac{1}{2}}\mathbb{E}\left[\left(L\right)_\nu^2\right]^{\frac{1}{2}},$$

for all $0 \le \nu \le t \le T$. Substituting the above bound into (B.22) yields

$$\begin{aligned}\mathbb{E}\left[\left(\breve{N}\right)_{t'}\right] &\le \Delta_R\int_0^{t'} e^{2\theta_1\nu}\left(\frac{1}{\theta_2^2}\mathbb{E}\left[\left(S\right)_\nu^2\right] + \frac{1}{2\theta_2}\mathbb{E}\left[\left(L\right)_\nu^2\right] + \sqrt{2}\frac{1}{\theta_2\sqrt{\theta_2}}\mathbb{E}\left[\left(S\right)_\nu^2\right]^{\frac{1}{2}}\mathbb{E}\left[\left(L\right)_\nu^2\right]^{\frac{1}{2}}\right)d\nu \\ &\le \Delta_R\left(\int_0^{t'} e^{2\theta_1\nu}d\nu\right)\left(\frac{1}{\theta_2^2}\mathbb{E}\left[\left(S\right)_{t'}^2\right] + \frac{1}{2\theta_2}\mathbb{E}\left[\left(L\right)_{t'}^2\right] + \sqrt{2}\frac{1}{\theta_2\sqrt{\theta_2}}\mathbb{E}\left[\left(S\right)_{t'}^2\right]^{\frac{1}{2}}\mathbb{E}\left[\left(L\right)_{t'}^2\right]^{\frac{1}{2}}\right).\end{aligned}$$

Solving the integral and using the strict positivity of $\theta_1 \in \mathbb{R}_{>0}$ produces

$$\begin{aligned}\mathbb{E}\left[\left(\breve{N}\right)_{t'}\right] &\le \frac{\Delta_R e^{2\theta_1 t'}}{2\theta_1}\left(\frac{1}{\theta_2^2}\mathbb{E}\left[\left(S\right)_{t'}^2\right] + \frac{1}{2\theta_2}\mathbb{E}\left[\left(L\right)_{t'}^2\right] + \sqrt{2}\frac{1}{\theta_2\sqrt{\theta_2}}\mathbb{E}\left[\left(S\right)_{t'}^2\right]^{\frac{1}{2}}\mathbb{E}\left[\left(L\right)_{t'}^2\right]^{\frac{1}{2}}\right) \\ &= \frac{\Delta_R e^{2\theta_1 t'}}{2\theta_1}\left(\frac{1}{\theta_2}\mathbb{E}\left[\left(S\right)_{t'}^2\right]^{\frac{1}{2}} + \frac{1}{\sqrt{2\theta_2}}\mathbb{E}\left[\left(L\right)_{t'}^2\right]^{\frac{1}{2}}\right)^2,\end{aligned}$$

thus proving (B.21) for $\mathsf{p} = 1$.

Next, consider $\mathsf{p} \in \mathbb{N}_{\ge}2$. Since $\breve{N}(t)$ is a standard Lebesgue integral with a t-continuous integrand, the fundamental theorem for the Lebesgue integral [106, Thm. 6.4.1] implies that $\breve{N}(t)$ is absolutely continuous on $[0, T]$ and

$$d\,\breve{N}(t) = e^{2(\theta_1-\theta_2)t}\,\breve{R}(t)^\top R(t)\,\breve{R}(t)dt, \quad \mu_L\text{-a.e. on } [0, T],$$

where μ_L denotes the Lebesgue measure. Using the chain rule one then sees that

$$\begin{aligned}\left|\breve{N}(t)\right|^{\mathsf{p}} &= \mathsf{p}\int_0^t\left|\breve{N}(\nu)\right|^{\mathsf{p}-2}\breve{N}(\nu)e^{2(\theta_1-\theta_2)\nu}\,\breve{R}(\nu)^\top R(\nu)\,\breve{R}(\nu)d\nu \\ &\le \mathsf{p}\int_0^t\left|\breve{N}(\nu)\right|^{\mathsf{p}-1}e^{2(\theta_1-\theta_2)\nu}\left\|R(\nu)\right\|_F\left\|\breve{R}(\nu)\right\|^2 d\nu, \\ &\le \mathsf{p}\Delta_R\int_0^t\left|\breve{N}(\nu)\right|^{\mathsf{p}-1}e^{2(\theta_1-\theta_2)\nu}\left\|\breve{R}(\nu)\right\|^2 d\nu, \quad t \in [0, T].\end{aligned}$$

Since the right hand side is a non decreasing function over $t \in [0, t']$, it then follows that

$$\left|\breve{N}(t)\right|^{\mathsf{p}} \leq \mathsf{p}\Delta_R \int_0^{t'} \left|\breve{N}(\nu)\right|^{\mathsf{p}-1} e^{2(\theta_1-\theta_2)\nu} \left\|\breve{R}(\nu)\right\|^2 d\nu, \quad \forall t \in [0, t'],$$

and thus

$$\left(\breve{N}\right)_{t'}^{\mathsf{p}} \leq \mathsf{p}\Delta_R \int_0^{t'} \left|\breve{N}(\nu)\right|^{\mathsf{p}-1} e^{2(\theta_1-\theta_2)\nu} \left\|\breve{R}(\nu)\right\|^2 d\nu.$$

It then follows that

$$\begin{aligned}\mathbb{E}\left[\left(\breve{N}\right)_{t'}^{\mathsf{p}}\right] \leq & \mathsf{p}\Delta_R \mathbb{E}\left[\int_0^{t'} \left|\breve{N}(\nu)\right|^{\mathsf{p}-1} e^{2(\theta_1-\theta_2)\nu} \left\|\breve{R}(\nu)\right\|^2 d\nu\right] \\ \overset{(i)}{=} & \mathsf{p}\Delta_R \mathbb{E}\left[\int_0^{t'} \left(\left|\breve{N}(\nu)\right|^{\mathsf{p}-1} e^{2\theta_1(\mathsf{p}-1)\nu/\mathsf{p}}\right)\left(e^{2(\theta_1/\mathsf{p}-\theta_2)\nu}\left\|\breve{R}(\nu)\right\|^2\right) d\nu\right] \\ \overset{(ii)}{\leq} & \mathsf{p}\Delta_R \mathbb{E}\left[\int_0^{t'} \left|\breve{N}(\nu)\right|^{\mathsf{p}} e^{2\theta_1\nu} d\nu\right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\int_0^{t'} e^{2(\theta_1-\theta_2\mathsf{p})\nu}\left\|\breve{R}(\nu)\right\|^{2\mathsf{p}} d\nu\right]^{\frac{1}{\mathsf{p}}},\end{aligned} \tag{B.26}$$

where we obtain (i) by writing

$$e^{2(\theta_1-\theta_2)\nu} = e^{2\left[\theta_1\left(\frac{\mathsf{p}-1}{\mathsf{p}}+\frac{1}{\mathsf{p}}\right)-\theta_2\right]\nu} = e^{2\theta_1(\mathsf{p}-1)\nu/\mathsf{p}} e^{2(\theta_1/\mathsf{p}-\theta_2)\nu},$$

and (ii) follows from Hölder's inequality with conjugates $\mathsf{p}/(\mathsf{p}-1)$ and p.
Now, we have the following straightforward inequality:

$$\mathbb{E}\left[\int_0^{t'} \left|\breve{N}(\nu)\right|^{\mathsf{p}} e^{2\theta_1\nu} d\nu\right] \leq \mathbb{E}\left[\left(\int_0^{t'} e^{2\theta_1\nu} d\nu\right)\left(\breve{N}\right)_{t'}^{\mathsf{p}}\right] = \left(\int_0^{t'} e^{2\theta_1\nu} d\nu\right) \mathbb{E}\left[\left(\breve{N}\right)_{t'}^{\mathsf{p}}\right], \tag{B.27}$$

where we have used the fact that t' is a constant.
Next, it follows from the definition of $\breve{R}(t)$ that

$$\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} = \mathbb{E}\left[\left(\breve{R}(\nu)^\top \breve{R}(\nu)\right)^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}},$$

which leads to the following bound:

$$\begin{aligned}&\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \\ &\quad \leq \mathbb{E}\left[\left(\left\|\int_0^{\nu} e^{\theta_2\beta} S(\beta) d\beta\right\|^2 + \left\|\int_0^{\nu} e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^2 + 2\left\|\int_0^{\nu} e^{\theta_2\beta} S(\beta) d\beta\right\|\left\|\int_0^{\nu} e^{\theta_2\beta} L(\beta) dQ_\beta\right\|\right)^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}},\end{aligned}$$

for $0 \leq \nu \leq t \leq T$. It then follows from the Minkowski inequality that

$$\begin{aligned}\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \leq & \mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2\beta} S(\beta) d\beta\right\|^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \\ & + 2\mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2\beta} S(\beta) d\beta\right\|^{\mathsf{p}}\left\|\int_0^{\nu} e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}},\end{aligned}$$

which, upon an application of the Cauchy-Schwarz inequality produces

$$\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \leq \mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2\beta} S(\beta) d\beta\right\|^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2\beta} L(\beta) dQ_\beta\right\|^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}}$$

$$+ 2\mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2 \beta} S(\beta)d\beta\right\|^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2 \beta} L(\beta)dQ_\beta\right\|^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}. \quad \text{(B.28)}$$

Now, we have the following inequality:

$$\mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2 \beta} S(\beta)d\beta\right\|^{2\mathsf{p}}\right] \leq \mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2 \beta} d\beta\right\|^{2\mathsf{p}} \left(S\right)_\nu^{2\mathsf{p}}\right] \leq \left(\frac{e^{\theta_2 \nu}}{\theta_2}\right)^{2\mathsf{p}} \mathbb{E}\left[\left(S\right)_\nu^{2\mathsf{p}}\right], \quad \forall 0 \leq \nu \leq t \leq T. \quad \text{(B.29)}$$

Similarly, using Lemma B.2, we have

$$\mathbb{E}\left[\left\|\int_0^{\nu} e^{\theta_2 \beta} L(\beta)dQ_\beta\right\|^{2\mathsf{p}}\right] \leq \left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\mathsf{p}} \left(\frac{e^{2\theta_2 \nu}}{\theta_2}\right)^{\mathsf{p}} \mathbb{E}\left[\left(L\right)_\nu^{2\mathsf{p}}\right], \quad \forall 0 \leq \nu \leq t \leq T. \quad \text{(B.30)}$$

Substituting (B.29) - (B.30) into (B.28) leads to

$$\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \leq e^{2\theta_2 \nu} \left(\frac{1}{\theta_2^2}\mathbb{E}\left[\left(S\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)\frac{1}{\theta_2}\mathbb{E}\left[\left(L\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + 2\left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{1}{\theta_2\sqrt{\theta_2}}\mathbb{E}\left[\left(S\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left(L\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}\right),$$

and thus

$$\mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^{2\mathsf{p}}\right] \leq e^{2\theta_2 \mathsf{p} \nu} \left(\frac{1}{\theta_2^2}\mathbb{E}\left[\left(S\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)\frac{1}{\theta_2}\mathbb{E}\left[\left(L\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + 2\left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{1}{\theta_2\sqrt{\theta_2}}\mathbb{E}\left[\left(S\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left(L\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}\right)^{\mathsf{p}}. \quad \text{(B.31)}$$

Furthermore, since

$$\mathbb{E}\left[\int_0^{t'} e^{2(\theta_1 - \theta_2 \mathsf{p})\nu} \left\|\breve{R}(\nu)\right\|^{2\mathsf{p}} d\nu\right] = \int_0^{t'} e^{2(\theta_1 - \theta_2 \mathsf{p})\nu} \mathbb{E}\left[\left\|\breve{R}(\nu)\right\|^{2\mathsf{p}}\right] d\nu,$$

we may use (B.31) and conclude that

$$\mathbb{E}\left[\int_0^{t'} e^{2(\theta_1 - \theta_2 \mathsf{p})\nu} \left\|\breve{R}(\nu)\right\|^{2\mathsf{p}} d\nu\right] \leq \int_0^{t'} e^{2\theta_1 \nu} \left(\frac{1}{\theta_2^2}\mathbb{E}\left[\left(S\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)\frac{1}{\theta_2}\mathbb{E}\left[\left(L\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + 2\left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{1}{\theta_2\sqrt{\theta_2}}\mathbb{E}\left[\left(S\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left(L\right)_\nu^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}\right)^{\mathsf{p}} d\nu.$$

We develop the bound further as follows:

$$\mathbb{E}\left[\int_0^{t'} e^{2(\theta_1 - \theta_2 \mathsf{p})\nu} \left\|\breve{R}(\nu)\right\|^{2\mathsf{p}} d\nu\right] \leq \left(\int_0^{t'} e^{2\theta_1 \nu} d\nu\right) \left(\frac{1}{\theta_2^2}\mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)\frac{1}{\theta_2}\mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + 2\left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{1}{\theta_2\sqrt{\theta_2}}\mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}\right)^{\mathsf{p}}. \quad \text{(B.32)}$$

Substituting (B.27) and (B.32) into (B.26) produces

$$\mathbb{E}\left[\left(\breve{N}\right)_{t'}^{\mathsf{p}}\right] \leq \mathsf{p}\Delta_R \mathbb{E}\left[\left(\breve{N}\right)_{t'}^{\mathsf{p}}\right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \left(\int_0^{t'} e^{2\theta_1 \nu} d\nu\right) \left(\frac{1}{\theta_2^2}\mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)\frac{1}{\theta_2}\mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}}\right.$$

$$+ 2\left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{1}{\theta_2\sqrt{\theta_2}}\mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}\right),$$

and thus

$$\mathbb{E}\left[\left(\breve{N}\right)_{t'}^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \leq \mathsf{p}\Delta_R\left(\int_0^{t'} e^{2\theta_1\nu}d\nu\right)\left(\frac{1}{\theta_2^2}\mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)\frac{1}{\theta_2}\mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}}\right.$$
$$\left. + 2\left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{1}{\theta_2\sqrt{\theta_2}}\mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} \mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}\right).$$

Solving the integral and using the strict positivity of $\theta_1 \in \mathbb{R}_{>0}$ then completes the proof.

□

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 *Let $S \in \mathcal{M}_2^{loc}(\mathbb{R}^m|\mathfrak{F}_t)$, $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{m\times n_q}|\mathfrak{F}_t)$, $\mathfrak{F}_t$-adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$, and strictly positive constants θ_1, $\theta_2 \in \mathbb{R}_{>0}$ be as in Lemma B.3. Define*

$$\widehat{N}(t) = \int_0^t e^{(\theta_1-\theta_2)\nu}\left(\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right)^\top L(\nu)dQ_\nu \in \mathbb{R}, \quad t\in[0,T]. \tag{B.33}$$

Then, the following bound holds for any constant $t' \in [0,T]$:

$$\left\|\left(\widehat{N}\right)_{t'}\right\|_{\mathsf{p}} \leq 2\sqrt{\mathsf{p}}\frac{e^{\theta_1 t'}}{\sqrt{\theta_1}\theta_2}\left\|\left(S\right)_{t'}\right\|_{2\mathsf{p}}\left\|\left(L\right)_{t'}\right\|_{2\mathsf{p}}, \quad \forall \mathsf{p}\in\mathbb{N}_{\geq 1}. \tag{B.34}$$

Proof. We begin by noting that the inner Lebesgue integral is absolute continuous on $[0,T]$ by fundamental theorem for the Lebesgue integral [106, Thm. 6.4.1], and $S \in \mathcal{M}_2^{loc}(\mathbb{R}^m|\mathfrak{F}_t)$ implies that

$$\int_0^t e^{\theta_2\beta}S(\beta)d\beta \in \mathcal{M}_2^{loc}(\mathbb{R}^m|\mathfrak{F}_t).$$

It follows then from [108, Thm. 5.5.2] that $\widehat{N}(t)$ is a continuous local martingale with respect to the filtration $\mathfrak{F}_t$ since $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{m\times n_q}|\mathfrak{F}_t)$. Furthermore, $\widehat{N}(0)=0$, and hence we invoke the Burkholder-Davis-Gundy inequality [108, Thm. 5.5.1] for the stopping time t' to obtain

$$\mathbb{E}\left[\left(\widehat{N}\right)_{t'}^{\mathsf{p}}\right] \leq (8\mathsf{p})^{\frac{\mathsf{p}}{2}}\,\mathbb{E}\left[\left\langle\widehat{N}\right\rangle_{t'}^{\frac{\mathsf{p}}{2}}\right] = (8\mathsf{p})^{\frac{\mathsf{p}}{2}}\,\mathbb{E}\left[\left(\int_0^{t'} e^{2(\theta_1-\theta_2)\nu}\left\|\left(\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right)^\top L(\nu)\right\|^2 d\nu\right)^{\frac{\mathsf{p}}{2}}\right]. \tag{B.35}$$

We now develop the bound on the integral term as follows:

$$\int_0^{t'} e^{2(\theta_1-\theta_2)\nu}\left\|\left(\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right)^\top L(\nu)\right\|^2 d\nu \leq \int_0^{t'} e^{2(\theta_1-\theta_2)\nu}\left\|\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right\|^2 \|L(\nu)\|_F^2\,d\nu$$
$$\leq \left[\int_0^{t'} e^{2(\theta_1-\theta_2)\nu}\left(\int_0^\nu e^{\theta_2\beta}d\beta\right)^2\left(S\right)_\nu^2 d\nu\right]\left(L\right)_{t'}^2.$$

It then follows that

$$\int_0^{t'} e^{2(\theta_1-\theta_2)\nu}\left\|\left(\int_0^\nu e^{\theta_2\beta}S(\beta)d\beta\right)^\top L(\nu)\right\|^2 d\nu \leq \left[\int_0^{t'} e^{2(\theta_1-\theta_2)\nu}\left(\int_0^\nu e^{\theta_2\beta}d\beta\right)^2 d\nu\right]\left(S\right)_{t'}^2\left(L\right)_{t'}^2. \tag{B.36}$$

Solving the integrals one sees that

$$\int_0^{t'} e^{2(\theta_1-\theta_2)\nu}\left(\int_0^\nu e^{\theta_2\beta}d\beta\right)^2 d\nu$$

$$= \left(\frac{1}{\theta_2}\right)^2 \int_0^{t'} e^{2(\theta_1-\theta_2)\nu} \left(e^{\theta_2\nu}-1\right)^2 d\nu \overset{(\star)}{\leq} \left(\frac{1}{\theta_2}\right)^2 \int_0^{t'} e^{2(\theta_1-\theta_2)\nu} e^{2\theta_2\nu} d\nu = \left(\frac{1}{\theta_2}\right)^2 \int_0^{t'} e^{2\theta_1\nu} d\nu,$$

where we have utilized the strict positivity of $\theta_2 \in \mathbb{R}_{>0}$ to obtain $(\star)$. It then follows from the strict positivity of $\theta_1 \in \mathbb{R}_{>0}$ that

$$\int_0^{t'} e^{2(\theta_1-\theta_2)\nu} \left(\int_0^{\nu} e^{\theta_2\beta} d\beta\right)^2 d\nu \leq \left(e^{2\theta_1 t'}-1\right) \frac{1}{2\theta_1} \left(\frac{1}{\theta_2}\right)^2 \leq e^{2\theta_1 t'} \frac{1}{2\theta_1} \left(\frac{1}{\theta_2}\right)^2.$$

We then obtain the following upon substituting the above inequality into (B.36):

$$\int_0^{t'} e^{2(\theta_1-\theta_2)\nu} \left\| \left(\int_0^{\nu} e^{\theta_2\beta} S(\beta) d\beta\right)^\top L(\nu) \right\|^2 d\nu \leq e^{2\theta_1 t'} \frac{1}{2\theta_1} \left(\frac{1}{\theta_2}\right)^2 \left(S\right)_{t'}^2 \left(L\right)_{t'}^2.$$

The proof is then concluded by substituting the above bound into (B.35) and performing the following manipulations:

$$\begin{aligned}
\mathbb{E}\left[\left(\widehat{N}\right)_{t'}^{\mathsf{p}}\right] &\leq (4\mathsf{p})^{\frac{\mathsf{p}}{2}} \mathbb{E}\left[\left(\frac{e^{\theta_1 t'}}{\sqrt{\theta_1}\theta_2}\right)^{\mathsf{p}} \left(S\right)_{t'}^{\mathsf{p}} \left(L\right)_{t'}^{\mathsf{p}}\right] \\
&\overset{(i)}{=} (4\mathsf{p})^{\frac{\mathsf{p}}{2}} \left(\frac{e^{\theta_1 t'}}{\sqrt{\theta_1}\theta_2}\right)^{\mathsf{p}} \mathbb{E}\left[\left(S\right)_{t'}^{\mathsf{p}} \left(L\right)_{t'}^{\mathsf{p}}\right] \\
&\overset{(ii)}{\leq} (4\mathsf{p})^{\frac{\mathsf{p}}{2}} \left(\frac{e^{\theta_1 t'}}{\sqrt{\theta_1}\theta_2}\right)^{\mathsf{p}} \mathbb{E}\left[\left(S\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2}} \mathbb{E}\left[\left(L\right)_{t'}^{2\mathsf{p}}\right]^{\frac{1}{2}},
\end{aligned}$$

where (i) and (ii) are due to t' being a constant and the Cauchy-Schwarz inequality, respectively.

□

We need the following result for Lemma B.6.

Lemma B.5 *Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\mathfrak{F}_t$, and let $R_1, R_2 \in \mathcal{M}_2^{loc}(\mathbb{R}^{m\times n_q}|\mathfrak{F}_t)$ satisfy*

$$\mathbb{E}\left[\int_0^T \|R_1(\nu)+R_2(\nu)\|_F^{4\mathsf{p}} d\nu\right] < \infty, \quad \forall \mathsf{p} \geq 1.$$

Additionally, for any strictly positive constants $\kappa_1, \kappa_2 \in \mathbb{R}_{>0}$ and an $\mathfrak{F}_t$-adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$, define for all $t \in [0,T]$

$$\widehat{R}(t) = \int_0^t e^{2(\kappa_1-\kappa_2)\nu} \left\|\widehat{R}_1(\nu)\right\|^2 d\nu \in \mathbb{R}_{\geq 0}, \quad \widehat{R}_1(t) = R_1(t)^\top \widehat{R}_2(t) \in \mathbb{R}^{n_q}, \quad \widehat{R}_2(t) = \int_0^t e^{\kappa_2\beta} R_2(\beta) dQ_\beta \in \mathbb{R}^m.$$

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

$$\mathbb{E}\left[\widehat{R}(t)^{\frac{\mathsf{p}}{2}}\right] \leq \left(\mathsf{p}^2 \frac{4\mathsf{p}-1}{2}\right)^{\frac{\mathsf{p}}{2}} \left(\frac{e^{\kappa_1 t}}{\sqrt{\kappa_1\kappa_2}}\right)^{\mathsf{p}} \mathbb{E}\left[\left(R_2\right)_t^{4\mathsf{p}}\right]^{\frac{1}{4}} \mathbb{E}\left[\left(R_1\right)_t^{4\mathsf{p}}\right]^{\frac{1}{4}}, \quad \forall \mathsf{p} \in \mathbb{N}_{\geq 1}. \tag{B.37}$$

Proof. We begin with the case when $\mathsf{p} = 1$. Since $\widehat{R}(t) \geq 0$, using the Cauchy-Schwarz inequality (or alternatively, Jensen's inequality) one sees that

$$\mathbb{E}\left[\widehat{R}(t)^{\frac{1}{2}}\right]^2 \leq \mathbb{E}\left[\widehat{R}(t)\right] = \mathbb{E}\left[\int_0^t e^{2(\kappa_1-\kappa_2)\nu} \left\|\widehat{R}_1(\nu)\right\|^2 d\nu\right] = \int_0^t e^{2(\kappa_1-\kappa_2)\nu} \mathbb{E}\left[\left\|\widehat{R}_1(\nu)\right\|^2\right] d\nu, \quad t \in [0,T].$$

It follows then from the definition $\widehat{R}_1(t) = R_1(t)^\top \widehat{R}_2(t)$ that

$$\begin{aligned}
\mathbb{E}\left[\widehat{R}(t)^{\frac{1}{2}}\right]^2 &\leq \int_0^t e^{2(\kappa_1-\kappa_2)\nu} \mathbb{E}\left[\|R_1(\nu)\|_F^2 \left\|\widehat{R}_2(\nu)\right\|^2\right] d\nu \\
&\overset{(\star)}{\leq} \int_0^t e^{2(\kappa_1-\kappa_2)\nu} \mathbb{E}\left[\|R_1(\nu)\|_F^4\right]^{\frac{1}{2}} \mathbb{E}\left[\left\|\widehat{R}_2(\nu)\right\|^4\right]^{\frac{1}{2}} d\nu
\end{aligned}$$

$$\leq \left(\int_0^t e^{2(\kappa_1 - \kappa_2)\nu} \mathbb{E}\left[\left\| \widehat{R}_2(\nu) \right\|^4 \right]^{\frac{1}{2}} d\nu \right) \mathbb{E}\left[\left(R_1 \right)_t^4 \right]^{\frac{1}{2}}, \quad t \in [0,T], \quad \text{(B.38)}$$

where $(\star)$ is a consequence of the Cauchy-Schwarz inequality. We now use Lemma B.2 to see that

$$\mathbb{E}\left[\left\| \widehat{R}_2(\nu) \right\|^4 \right]^{\frac{1}{2}} = \mathbb{E}\left[\left\| \int_0^\nu e^{\kappa_2 \beta} R_2(\beta) dQ_\beta \right\|^4 \right]^{\frac{1}{2}} \leq 3 \left(\frac{e^{2\kappa_2 \nu} - 1}{\kappa_2} \right) \mathbb{E}\left[\left(R_2 \right)_\nu^4 \right]^{\frac{1}{2}}, \quad \forall \nu \in [0,t].$$

Substituting the above for the inner expectation in (B.38) then produces

$$\begin{aligned} \mathbb{E}\left[\widehat{R}(t)^{\frac{1}{2}} \right]^2 \leq & 3 \left(\int_0^t e^{2(\kappa_1 - \kappa_2)\nu} \left(\frac{e^{2\kappa_2 \nu} - 1}{\kappa_2} \right) \mathbb{E}\left[\left(R_2 \right)_\nu^4 \right]^{\frac{1}{2}} d\nu \right) \mathbb{E}\left[\left(R_1 \right)_t^4 \right]^{\frac{1}{2}} \\ \leq & 3 \left(\int_0^t e^{2(\kappa_1 - \kappa_2)\nu} \left(\frac{e^{2\kappa_2 \nu} - 1}{\kappa_2} \right) d\nu \right) \mathbb{E}\left[\left(R_2 \right)_t^4 \right]^{\frac{1}{2}} \mathbb{E}\left[\left(R_1 \right)_t^4 \right]^{\frac{1}{2}}, \quad t \in [0,T]. \end{aligned}$$

It then follows that

$$\mathbb{E}\left[\widehat{R}(t)^{\frac{1}{2}} \right] \leq \sqrt{3} \left(\int_0^t e^{2(\kappa_1 - \kappa_2)\nu} \left(\frac{e^{2\kappa_2 \nu} - 1}{\kappa_2} \right) d\nu \right)^{\frac{1}{2}} \mathbb{E}\left[\left(R_2 \right)_t^4 \right]^{\frac{1}{4}} \mathbb{E}\left[\left(R_1 \right)_t^4 \right]^{\frac{1}{4}}, \quad t \in [0,T].$$

Then, one establishes (B.37) for $\mathsf{p} = 1$ by using the strict positivity of $\kappa_1, \kappa_2 \in \mathbb{R}_{>0}$ to obtain

$$\int_0^t e^{2(\kappa_1 - \kappa_2)\nu} \left(\frac{e^{2\kappa_2 \nu} - 1}{\kappa_2} \right) d\nu = \int_0^t e^{2\kappa_1 \nu} \left(\frac{1 - e^{-2\kappa_2 \nu}}{\kappa_2} \right) d\nu \leq \left(\int_0^t e^{2\kappa_1 \nu} d\nu \right) \frac{1}{\kappa_2},$$

and thus

$$\int_0^t e^{2(\kappa_1 - \kappa_2)\nu} \left(\frac{e^{2\kappa_2 \nu} - 1}{\kappa_2} \right) d\nu \leq \frac{e^{2\kappa_1 t} - 1}{2\kappa_1 \kappa_2} \leq \frac{e^{2\kappa_1 t}}{2\kappa_1 \kappa_2}.$$

Next, consider the case when $\mathsf{p} \in \mathbb{N}_{\geq 2}$. The fundamental theorem for the Lebesgue integral [106, Thm. 6.4.1] implies that $\widehat{R}(t)$ is absolutely continuous on $[0,T]$ and

$$d\widehat{R}(t) = e^{2(\kappa_1 - \kappa_2)t} \left\| \widehat{R}_1(t) \right\|^2 dt, \quad \mu_L\text{-a.e. on } [0,T],$$

where μ_L denotes the Lebesgue measure. Therefore, since $\widehat{R}(t) \geq 0$, the chain rule implies that

$$\widehat{R}(t)^{\mathsf{p}} = \mathsf{p} \int_0^t \widehat{R}(\nu)^{\mathsf{p}-1} e^{2(\kappa_1 - \kappa_2)\nu} \left\| \widehat{R}_1(\nu) \right\|^2 d\nu, \quad t \in [0,T].$$

It then follows that

$$\begin{aligned} \mathbb{E}\left[\widehat{R}(t)^{\mathsf{p}} \right] = & \mathsf{p} \mathbb{E}\left[\int_0^t \widehat{R}(\nu)^{\mathsf{p}-1} e^{2(\kappa_1 - \kappa_2)\nu} \left\| \widehat{R}_1(\nu) \right\|^2 d\nu \right] \\ = & \mathsf{p} \mathbb{E}\left[\int_0^t \left(e^{2\kappa_1 (\mathsf{p}-1)\nu/\mathsf{p}} \widehat{R}(\nu)^{\mathsf{p}-1} \right) \left(e^{2(\kappa_1/\mathsf{p} - \kappa_2)\nu} \left\| \widehat{R}_1(\nu) \right\|^2 \right) d\nu \right], \end{aligned} \quad \text{(B.39)}$$

where the last expression is due to

$$e^{2(\kappa_1 - \kappa_2)\nu} = e^{2\left[\kappa_1 \left(\frac{\mathsf{p}-1}{\mathsf{p}} + \frac{1}{\mathsf{p}} \right) - \kappa_2 \right]\nu} = e^{2\kappa_1 (\mathsf{p}-1)\nu/\mathsf{p}} e^{2(\kappa_1/\mathsf{p} - \kappa_2)\nu}.$$

Using Hölder's inequality with conjugates $\mathsf{p}/(\mathsf{p}-1)$ and p, one sees that

$$\begin{aligned} \mathbb{E}\left[\widehat{R}(t)^{\mathsf{p}} \right] \leq & \mathsf{p} \mathbb{E}\left[\int_0^t e^{2\kappa_1 \nu} \widehat{R}(\nu)^{\mathsf{p}} d\nu \right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \left\| \widehat{R}_1(\nu) \right\|^{2\mathsf{p}} d\nu \right]^{\frac{1}{\mathsf{p}}} \\ = & \mathsf{p} \left(\int_0^t e^{2\kappa_1 \nu} \mathbb{E}\left[\widehat{R}(\nu)^{\mathsf{p}} \right] d\nu \right)^{\frac{\mathsf{p}-1}{\mathsf{p}}} \left(\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \widehat{R}_1(\nu) \right\|^{2\mathsf{p}} \right] d\nu \right)^{\frac{1}{\mathsf{p}}} \end{aligned}$$

$$\leq \mathsf{p} \left(\int_0^t e^{2\kappa_1 \nu} d\nu \right)^{\frac{\mathsf{p}-1}{\mathsf{p}}} \mathbb{E}\left[\widehat{R}(t)^{\mathsf{p}} \right]^{\frac{\mathsf{p}-1}{\mathsf{p}}} \left(\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \widehat{R}_1(\nu) \right\|^{2\mathsf{p}} \right] d\nu \right)^{\frac{1}{\mathsf{p}}},$$

where we have used the fact that $\mathbb{E}\left[\widehat{R}(t)^{\mathsf{p}}\right]$ is non-decreasing in t as implied by (B.39). It then follows that

$$\mathbb{E}\left[\widehat{R}(t)^{\mathsf{p}} \right]^{\frac{1}{\mathsf{p}}} \leq \mathsf{p} \left(\int_0^t e^{2\kappa_1 \nu} d\nu \right)^{\frac{\mathsf{p}-1}{\mathsf{p}}} \left(\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \widehat{R}_1(\nu) \right\|^{2\mathsf{p}} \right] d\nu \right)^{\frac{1}{\mathsf{p}}},$$

and thus

$$\mathbb{E}\left[\widehat{R}(t)^{\mathsf{p}} \right] \leq \mathsf{p}^{\mathsf{p}} \left(\int_0^t e^{2\kappa_1 \nu} d\nu \right)^{\mathsf{p}-1} \int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \widehat{R}_1(\nu) \right\|^{2\mathsf{p}} \right] d\nu, \quad \forall t \in [0,T]. \tag{B.40}$$

Now, using the definition $\widehat{R}_1(t) = R_1(t)^\top \widehat{R}_2(t)$, we obtain

$$\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \widehat{R}_1(\nu) \right\|^{2\mathsf{p}} \right] d\nu \leq \int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\| R_1(\nu) \|_F^{2\mathsf{p}} \left\| \widehat{R}_2(\nu) \right\|^{2\mathsf{p}} \right] d\nu.$$

Using the Cauchy-Schwarz inequality, one sees that

$$\begin{aligned}\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \widehat{R}_1(\nu) \right\|^{2\mathsf{p}} \right] d\nu &\leq \int_0^t e^{2\kappa_1 \nu} \mathbb{E}\left[\| R_1(\nu) \|_F^{4\mathsf{p}} \right]^{\frac{1}{2}} \mathbb{E}\left[\left\| \widehat{R}_2(\nu) \right\|^{4\mathsf{p}} \right]^{\frac{1}{2}} d\nu \\ &\leq \left(\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \widehat{R}_2(\nu) \right\|^{4\mathsf{p}} \right]^{\frac{1}{2}} d\nu \right) \mathbb{E}\left[\left(R_1 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}}, \quad \forall t \in [0,T].\end{aligned}$$

It then follows from the definition of $\widehat{R}_2(t)$ that

$$\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \widehat{R}_1(\nu) \right\|^{2\mathsf{p}} \right] d\nu \leq \left(\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \int_0^\nu e^{\kappa_2 \beta} R_2(\beta) dQ_\beta \right\|^{4\mathsf{p}} \right]^{\frac{1}{2}} d\nu \right) \mathbb{E}\left[\left(R_1 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}},$$

for all $t \in [0,T]$. Consequently, using Lemma B.2 to bound the inner expectation leads to

$$\begin{aligned}&\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \widehat{R}_1(\nu) \right\|^{2\mathsf{p}} \right] d\nu \\ &\quad \leq p^{\mathsf{p}} (4p-1)^{\mathsf{p}} \left(\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \left(\frac{e^{2\kappa_2 \nu} - 1}{\kappa_2} \right)^{\mathsf{p}} \mathbb{E}\left[\left(R_2 \right)_\nu^{4\mathsf{p}} \right]^{\frac{1}{2}} d\nu \right) \mathbb{E}\left[\left(R_1 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}} \\ &\quad \leq p^{\mathsf{p}} (4p-1)^{\mathsf{p}} \left(\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \left(\frac{e^{2\kappa_2 \nu} - 1}{\kappa_2} \right)^{\mathsf{p}} d\nu \right) \mathbb{E}\left[\left(R_2 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}} \mathbb{E}\left[\left(R_1 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}}. \end{aligned} \tag{B.41}$$

Using the strict positivity of $\kappa_2 \in \mathbb{R}_{>0}$, one sees that

$$\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \left(\frac{e^{2\kappa_2 \nu} - 1}{\kappa_2} \right)^{\mathsf{p}} d\nu \leq \int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \left(\frac{e^{2\kappa_2 \nu}}{\kappa_2} \right)^{\mathsf{p}} d\nu = \int_0^t e^{2\kappa_1 \nu} d\nu \left(\frac{1}{\kappa_2} \right)^{\mathsf{p}}.$$

and thus, (B.41) can be written as

$$\int_0^t e^{2(\kappa_1 - \kappa_2 \mathsf{p})\nu} \mathbb{E}\left[\left\| \widehat{R}_1(\nu) \right\|^{2\mathsf{p}} \right] d\nu \leq p^{\mathsf{p}} (4p-1)^{\mathsf{p}} \left(\int_0^t e^{2\kappa_1 \nu} d\nu \right) \left(\frac{1}{\kappa_2} \right)^{\mathsf{p}} \mathbb{E}\left[\left(R_2 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}} \mathbb{E}\left[\left(R_1 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}}.$$

Substituting the above bound into (B.40) produces

$$\mathbb{E}\left[\widehat{R}(t)^{\mathsf{p}} \right] \leq \mathsf{p}^{2\mathsf{p}} (4\mathsf{p}-1)^{\mathsf{p}} \left(\int_0^t e^{2\kappa_1 \nu} d\nu \right)^{\mathsf{p}} \left(\frac{1}{\kappa_2} \right)^{\mathsf{p}} \mathbb{E}\left[\left(R_2 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}} \mathbb{E}\left[\left(R_1 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}}, \quad \forall t \in [0,T].$$

Solving the integral and using the strict positivity of $\kappa_1 \in \mathbb{R}_{>0}$, we obtain

$$\mathbb{E}\left[\widehat{R}(t)^{\mathsf{p}} \right] \leq \left(\mathsf{p}^2 \frac{4\mathsf{p}-1}{2} \right)^{\mathsf{p}} \left(\frac{e^{2\kappa_1 t}}{\kappa_1 \kappa_2} \right)^{\mathsf{p}} \mathbb{E}\left[\left(R_2 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}} \mathbb{E}\left[\left(R_1 \right)_t^{4\mathsf{p}} \right]^{\frac{1}{2}}, \quad \forall t \in [0,T].$$

The proof is then concluded by observing that $\widehat{R}(t) \geq 0$, $\forall t$, and Jensen's inequality imply that $\mathbb{E}\left[\widehat{R}(t)^{\frac{\mathsf{p}}{2}}\right]^2 \leq \mathbb{E}\left[\widehat{R}(t)^{\mathsf{p}}\right]$.

□

The following result provides a bound on a class of nested Itô integrals.

Lemma B.6 *Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\mathfrak{F}_t$, and let $L_1, L_2 \in \mathcal{M}_2^{loc}(\mathbb{R}^{m \times n_q}|\mathfrak{F}_t)$ satisfy*

$$\mathbb{E}\left[\int_0^T \|L_1(\nu) + L_2(\nu)\|_F^{4\mathsf{p}}\, d\nu\right] < \infty, \quad \forall \mathsf{p} \geq 1. \tag{B.42}$$

Additionally, for any strictly positive constants θ_1, $\theta_2 \in \mathbb{R}_{>0}$ and an $\mathfrak{F}_t$-adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$ define

$$\widetilde{N}(t) = \int_0^t e^{\theta_2 \nu}\left(\int_\nu^t e^{(\theta_1 - \theta_2)\beta} L_1(\beta) dQ_\beta\right)^\top L_2(\nu) dQ_\nu \in \mathbb{R}, \quad t \in [0, T].$$

Then, the following bound holds for any constant $t' \in [0, T]$:

$$\left\|\left(\widetilde{N}\right)_{t'}\right\|_{\mathsf{p}} \leq 2\mathsf{p}^{\frac{3}{2}}\left(4\mathsf{p} - 1\right)^{\frac{1}{2}} \frac{e^{\theta_1 t'}}{\sqrt{\theta_1 \theta_2}} \left\|\left(L_2\right)_{t'}\right\|_{4\mathsf{p}} \left\|\left(L_1\right)_{t'}\right\|_{4\mathsf{p}} + \frac{\sqrt{m} e^{\theta_1 t'}}{\theta_1} \left\|\left(L_1 L_2^\top\right)_{t'}\right\|_{\mathsf{p}}, \quad \mathsf{p} \in \mathbb{N}_{\geq 1}. \tag{B.43}$$

Proof. We begin by setting

$$N_1(t) = \int_0^t e^{(\theta_1 - \theta_2)\nu} L_1(\nu) dQ_\nu \in \mathbb{R}^m, \quad N_2(t) = \int_0^t e^{\theta_2 \nu} L_2(\nu) dQ_\nu \in \mathbb{R}^m, \quad t \in [0, T].$$

It then follows from the definition of $\widetilde{N}(t)$ that

$$\widetilde{N}(t) = \int_0^t e^{\theta_2 \nu}\left(\int_\nu^t e^{(\theta_1 - \theta_2)\beta} L_1(\beta) dQ_\beta\right)^\top L_2(\nu) dQ_\nu = \int_0^t \left(\int_\nu^t dN_1(\beta)\right)^\top dN_2(\nu).$$

Hence,

$$\begin{aligned} \widetilde{N}(t) = \int_0^t \left(\int_0^t dN_1(\beta)\right)^\top dN_2(\nu) - \int_0^t \left(\int_0^\nu dN_1(\beta)\right)^\top dN_2(\nu) \\ = N_1(t)^\top N_2(t) - \int_0^t N_1(\nu)^\top dN_2(\nu), \quad t \in [0, T]. \end{aligned} \tag{B.44}$$

Next, applying Itô's lemma to $N_1(t)^\top N_2(t)$ (or alternatively using the Itô product rule [95, Sec. 4.4.1] applied element wise) produces

$$N_1(t)^\top N_2(t) = \int_0^t N_1(\nu)^\top dN_2(\nu) + \int_0^t N_2(\nu)^\top dN_1(\nu) + \sum_{i=1}^m \left\langle N_{1,i}, N_{2,i}\right\rangle_t, \quad t \in [0, T],$$

where $N_{1,i}$, $N_{2,i} \in \mathbb{R}$ denote the i^{th} scalar-valued process, $i \in \{1, \ldots, m\}$, 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 $N_{2,i}(t)$ [93, Defn. 2.3.9]. Substituting the above expression for $N_1(t)^\top N_2(t)$ into (B.44) produces

$$\widetilde{N}(t) = \int_0^t N_2(\nu)^\top dN_1(\nu) + \sum_{i=1}^m \left\langle N_{1,i}, N_{2,i}\right\rangle_t, \quad t \in [0, T],$$

which implies that

$$\left(\widetilde{N}\right)_{t'} \leq \sup_{t \in [0,t']} \left|\int_0^t N_2(\nu)^\top dN_1(\nu)\right| + \sup_{t \in [0,t']} \left|\sum_{i=1}^m \left\langle N_{1,i}, N_{2,i}\right\rangle_t\right|.$$

Using Minkowski's inequality one sees that

$$\mathbb{E}\left[\left(\widetilde{N}\right)_{t'}^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \leq \mathbb{E}\left[\sup_{t\in[0,t']}\left|\int_0^t N_2(\nu)^\top dN_1(\nu)\right|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \mathbb{E}\left[\sup_{t\in[0,t']}\left|\sum_{i=1}^m \left\langle N_{1,i}, N_{2,i}\right\rangle_t\right|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}}. \tag{B.45}$$

Now, let

$$\breve{N}(t) \doteq \int_0^t N_2(\nu)^\top dN_1(\nu) = \int_0^t e^{(\theta_1-\theta_2)\nu}\left(\int_0^\nu e^{\theta_2\beta}L_2(\beta)dQ_\beta\right)^\top L_1(\nu)dQ_\nu \in \mathbb{R}.$$

The t-continuity of $N_2(t)$ [15, Thm. 3.2.5], along with $L_1,\ L_2 \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{m\times n_q}|\mathfrak{F}_t\right)$ imply that $\breve{N}(t)$ is a continuous local martingale [108, Thm. 5.5.2]. Furthermore, since $\breve{N}(0) = 0$, we use the Burkholder-Davis-Gundy inequality [107, Thm. 2] to see that

$$\mathbb{E}\left[\left(\breve{N}\right)_{t'}^{\mathsf{p}}\right] \leq (8\mathsf{p})^{\frac{\mathsf{p}}{2}}\,\mathbb{E}\left[\left\langle \breve{N}\right\rangle_{t'}^{\frac{\mathsf{p}}{2}}\right] = (8\mathsf{p})^{\frac{\mathsf{p}}{2}}\,\mathbb{E}\left[\left(\left(\int_0^t e^{2(\theta_1-\theta_2)\nu}\left\|L_1(\nu)^\top\int_0^\nu e^{\theta_2\beta}L_2(\beta)dQ_\beta\right\|^2 d\nu\right)^{\frac{\mathsf{p}}{2}}\right]\right].$$

Note that the integral inside the expectation can be cast as the process $\widehat{R}(t)$ in Lemma B.5 with $\kappa_i = \theta_i$ and $R_i(t) = L_i(t),\ i \in \{1,2\}$. It then follows from Lemma B.5 that

$$\mathbb{E}\left[\left(\breve{N}\right)_{t'}^{\mathsf{p}}\right] \leq 2^{\mathsf{p}}\mathsf{p}^{\frac{3\mathsf{p}}{2}}\left(4\mathsf{p}-1\right)^{\frac{\mathsf{p}}{2}}\left(\frac{e^{\theta_1 t'}}{\sqrt{\theta_1\theta_2}}\right)^{\mathsf{p}}\mathbb{E}\left[\left(L_2\right)_{t'}^{4\mathsf{p}}\right]^{\frac{1}{4}}\mathbb{E}\left[\left(L_1\right)_{t'}^{4\mathsf{p}}\right]^{\frac{1}{4}}.$$

Then, since

$$\mathbb{E}\left[\sup_{t\in[0,t']}\left|\int_0^t N_2(\nu)^\top dN_1(\nu)\right|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} = \mathbb{E}\left[\sup_{t\in[0,t']}\left|\breve{N}(t)\right|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} = \mathbb{E}\left[\left(\breve{N}\right)_{t'}^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}},$$

we get that

$$\mathbb{E}\left[\sup_{t\in[0,t']}\left|\int_0^t N_2(\nu)^\top dN_1(\nu)\right|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \leq 2\mathsf{p}^{\frac{3}{2}}\left(4\mathsf{p}-1\right)^{\frac{1}{2}}\frac{e^{\theta_1 t'}}{\sqrt{\theta_1\theta_2}}\left\|\left(L_2\right)_{t'}\right\|_{4\mathsf{p}}\left\|\left(L_1\right)_{t'}\right\|_{4\mathsf{p}}. \tag{B.46}$$

Next, it follows from the definition of the cross-variation process [93, Defn. 2.3.9] that

$$\sum_{i=1}^m\left\langle N_{1,i}, N_{2,i}\right\rangle_t = \frac{1}{4}\sum_{i=1}^m\left(\left\langle N_{1,i}+N_{2,i}\right\rangle_t - \left\langle N_{1,i}-N_{2,i}\right\rangle_t\right) = \sum_{i=1}^m\int_0^t e^{\theta_1\nu}L_{1,i}(\nu)L_{2,i}(\nu)^\top d\nu,$$

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,

$$\begin{aligned}\left|\sum_{i=1}^m\left\langle N_{1,i}, N_{2,i}\right\rangle_t\right| \leq \int_0^t e^{(\theta_1+\theta_2)\nu}\left|\sum_{i=1}^m L_{1,i}(\nu)L_{2,i}(\nu)^\top\right| d\nu &\leq \sqrt{m}\int_0^t e^{\theta_1\nu}\left(\sum_{i=1}^m\left|L_{1,i}(\nu)L_{2,i}(\nu)^\top\right|^2\right)^{\frac{1}{2}} d\nu \\ &= \sqrt{m}\int_0^t e^{\theta_1\nu}\left\|L_1(\nu)L_2(\nu)^\top\right\|_F d\nu.\end{aligned}$$

We further develop the bound as

$$\left|\sum_{i=1}^m\left\langle N_{1,i}, N_{2,i}\right\rangle_t\right| \leq \sqrt{m}\left(\int_0^t e^{\theta_1\nu}d\nu\right)\left(L_1L_2^\top\right)_{t'}, \quad \forall t \in [0,T],$$

which further implies that

$$\sup_{t\in[0,t']}\left|\sum_{i=1}^m\left\langle N_{1,i}, N_{2,i}\right\rangle_t\right| \leq \sqrt{m}\left(\int_0^{t'} e^{\theta_1\nu}d\nu\right)\left(L_1L_2^\top\right)_{t'} = \sqrt{m}\frac{e^{\theta_1 t'}-1}{\theta_1}\left(L_1L_2^\top\right)_{t'} \leq \frac{\sqrt{m}e^{\theta_1 t'}}{\theta_1}\left(L_1L_2^\top\right)_{t'},$$

where we have used the strict positivity of $\theta_1 \in \mathbb{R}_{>0}$. Taking expectation on both sides and using the fact that t' is a constant produces

$$\mathbb{E}\left[\sup_{t\in[0,t']}\left|\sum_{i=1}^m\left\langle N_{1,i}, N_{2,i}\right\rangle_t\right|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \leq \frac{\sqrt{m}e^{\theta_1 t'}}{\theta_1}\left\|\left(L_1L_2^\top\right)_{t'}\right\|_{\mathsf{p}}. \tag{B.47}$$

Substituting the bounds (B.46) and (B.47) into (B.45) produces the desired result.

□

C Reference Process

We provide the proof of Proposition 3.1 below.

Proof of Proposition 3.1. We consider the case $\mathbb{P}\{Y_{N,0} = Y_0 \in U_N\} = 1$ w.l.o.g. since otherwise $\tau_N = 0$ and the result is trivial.

Note that establishing the well-posedness of $Y_{N,t} = \begin{bmatrix}(X^r_{N.t})^\top & (X^\star_{N.t})^\top\end{bmatrix}$ is equivalent to establishing the individual well-posedness of the following systems:

$$dX^\star_{N,t} = \bar{F}_{N,\mu}\left(t, X^\star_{N,t}\right)dt + \bar{F}_{N,\sigma}\left(t, X^\star_{N,t}\right)dW^\star_t, \quad X^\star_{N,0} = x^\star_0 \sim \xi^\star_0, \tag{C.1a}$$

$$dX^r_{N,t} = F_{N,\mu}\left(t, X^r_{N,t}, U^r_t\right)dt + F_{N,\sigma}\left(t, X^r_{N,t}\right)dW_t, \quad X^r_{N,0} = x_0 \sim \xi_0, \tag{C.1b}$$

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 (27).

The well-posedness of (C.1a) is straightforward to establish using Definition 1, Assumption 1 [47, 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 $z \in \mathcal{M}_2\left([0,T], \mathbb{R}^n \mid \mathfrak{W}_{0,t}\right)$, for any $t \in [0,T]$, and define

$$M(z(t)) = \int_0^t F_{N,\mu}\left(\nu, z(\nu), U^r_\nu\right)d\nu + \int_0^t F_{N,\sigma}\left(\nu, z(\nu)\right)dW_\nu, \quad t \in [0,T]. \tag{C.2}$$

Let us denote by 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 (25). Then, we have that

$$M(z(t)) = \int_0^t \left(f_N\left(\nu, z(\nu)\right) + g(\nu)U^r_\nu + \Lambda_{N,\mu}\left(\nu, z(\nu)\right)\right)d\nu + \int_0^t \left(p_N(\nu, z(\nu)) + \Lambda_{N,\sigma}\left(\nu, z(\nu)\right)\right)dW_\nu,$$

for $t \in [0,T]$, where, from (23), we have that

$$U^r_\nu = \mathcal{F}_{\boldsymbol{\omega}}\Lambda^{\parallel}_{N,\mu}\left(\cdot, z\right)\nu + \mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}p^{\parallel}_N(\cdot, z) + \Lambda^{\parallel}_{N,\sigma}\left(\cdot, z\right), W\nu.$$

Therefore, the previous expression can be expressed as

$$\begin{aligned} M(z(t)) = &\int_0^t \left(f_N\left(\nu, z(\nu)\right) + \Lambda_{N,\mu}\left(\nu, z(\nu)\right)\right)d\nu + \int_0^t g(\nu)\mathcal{F}_{\boldsymbol{\omega}}\Lambda^{\parallel}_{N,\mu}\left(\cdot, z\right)\nu d\nu \\ &+ \int_0^t g(\nu)\mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}p^{\parallel}_N(\cdot, z) + \Lambda^{\parallel}_{N,\sigma}\left(\cdot, z\right), W\nu d\nu + \int_0^t \left(p_N(\nu, z(\nu)) + \Lambda_{N,\sigma}\left(\nu, z(\nu)\right)\right)dW_\nu, \end{aligned} \tag{C.3}$$

for $t \in [0,T]$. Using the definition of $\mathcal{F}_{\boldsymbol{\omega}}$ in (18a), we have that

$$\int_0^t g(\nu)\mathcal{F}_{\boldsymbol{\omega}}\Lambda^{\parallel}_{N,\mu}\left(\cdot, z\right)\nu d\nu = -\boldsymbol{\omega}\int_0^t\int_0^\nu g(\nu)e^{-\boldsymbol{\omega}(\nu-\beta)}\Lambda^{\parallel}_{N,\mu}\left(\beta, z(\beta)\right)d\beta d\nu, \quad t \in [0,T].$$

Changing the order of integration in the double Lebesgue integral produces

$$\int_0^t g(\nu)\mathcal{F}_{\boldsymbol{\omega}}\Lambda^{\parallel}_{N,\mu}\left(\cdot, z\right)\nu d\nu = -\boldsymbol{\omega}\int_0^t\left(\int_\nu^t g(\beta)e^{-\boldsymbol{\omega}\beta}d\beta\right)e^{\boldsymbol{\omega}\nu}\Lambda^{\parallel}_{N,\mu}\left(\nu, z(\nu)\right)d\nu, \quad t \in [0,T]. \tag{C.4}$$

Next, using the definition of $\mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}$ in (23), we have that

$$\begin{aligned} &\int_0^t g(\nu)\mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}p^{\parallel}_N(\cdot, z) + \Lambda^{\parallel}_{N,\sigma}\left(\cdot, z\right), W\nu d\nu \\ &\qquad = -\boldsymbol{\omega}\int_0^t\int_0^\nu g(\nu)e^{-\boldsymbol{\omega}(\nu-\beta)}\left(p^{\parallel}_N\left(\beta, z(\beta)\right) + \Lambda^{\parallel}_{N,\sigma}\left(\beta, z(\beta)\right)\right)dW_\beta d\nu, \quad t \in [0,T]. \end{aligned}$$

Applying Lemma B.1 to the above expression for

$$P(\nu) = g(\nu)e^{-\boldsymbol{\omega}\nu} \in \mathcal{C}\left([0,T]; \mathbb{R}^{n\times m}\right),$$

$$S(\beta) = e^{\boldsymbol{\omega}\beta}\left(p^{\parallel}_N\left(\beta, z(\beta)\right) + \Lambda^{\parallel}_{N,\sigma}\left(\beta, z(\beta)\right)\right) \in \mathcal{M}\left([0,T]; \mathbb{R}^{m\times d} \mid \mathfrak{W}_{0,t}\right), Q_\beta = W_\beta,$$

produces

$$\int_0^t g(\nu)\mathcal{F}_{N,\omega}p_N^{\parallel}(\cdot,z) + \Lambda_{N,\sigma}^{\parallel}(\cdot,z), W\nu d\nu$$
$$= -\omega \int_0^t \left(\int_\nu^t g(\beta)e^{-\omega\beta}d\beta\right) e^{\omega\nu}\left(p_N^{\parallel}(\nu,z(\nu)) + \Lambda_{N,\sigma}^{\parallel}(\nu,z(\nu))\right) dW_\nu, \quad t\in[0,T]. \tag{C.5}$$

Substituting (C.4) and (C.5) into (C.3) yields

$$M(z(t)) = \int_0^t F_{N,\mu}(\nu,z(\nu),U_\nu^r)\,d\nu + \int_0^t F_{N,\sigma}(\nu,z(\nu))\,dW_\nu$$
$$= \int_0^t M_\mu(t,\nu,z(\nu))\,d\nu + \int_0^t M_\sigma(t,\nu,z(\nu))\,dW_\nu, \quad t\in[0,T], \tag{C.6}$$

where

$$M_\mu(t,\nu,z(\nu)) = f_N(\nu,z(\nu)) + \Lambda_{N,\mu}(\nu,z(\nu)) - \omega\left(\int_\nu^t g(\beta)e^{-\omega\beta}d\beta\right)e^{\omega\nu}\Lambda_{N,\mu}^{\parallel}(\nu,z(\nu)),$$
$$M_\sigma(t,\nu,z(\nu)) = p_N(\nu,z(\nu)) + \Lambda_{N,\sigma}(\nu,z(\nu)) - \omega\left(\int_\nu^t g(\beta)e^{-\omega\beta}d\beta\right)e^{\omega\nu}\left(p_N^{\parallel}(\nu,z(\nu)) + \Lambda_{N,\sigma}^{\parallel}(\nu,z(\nu))\right),$$

for $\nu\in[0,t]$, and $t\in[0,T]$. Using Assumption 1, we have that

$$\left\|\left(\int_\nu^t g(\beta)e^{-\omega\beta}d\beta\right)e^{\omega\nu}\right\|_F$$
$$\leq \int_\nu^t \|g(\beta)\|_F\, e^{-\omega\beta}d\beta e^{\omega\nu} \leq \Delta_g\frac{1-e^{-\omega(t-\nu)}}{\omega} \leq \Delta_g\frac{1}{\omega}, \quad \forall\nu\in[0,t],\ t\in[0,T], \tag{C.7}$$

Thus, we conclude that

$$\|M_\mu(t,\nu,z(\nu))\| \leq \|f_N(\nu,z(\nu))\| + \|\Lambda_{N,\mu}(\nu,z(\nu))\| + \Delta_g\left\|\Lambda_{N,\mu}^{\parallel}(\nu,z(\nu))\right\|, \tag{C.8a}$$
$$\|M_\sigma(t,\nu,z(\nu))\|_F = \|p_N(\nu,z(\nu))\|_F + \|\Lambda_{N,\sigma}(\nu,z(\nu))\|_F + \Delta_g\left\|p_N^{\parallel}(\nu,z(\nu)) + \Lambda_{N,\sigma}^{\parallel}(\nu,z(\nu))\right\|_F, \tag{C.8b}$$

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 (27), along with the bounds in (C.8), imply linear growth of the integrands in the Picard iterates above. Therefore, as in the proof of [100, Thm. 2.3.1], we claim the existence of solutions to (C.1b) on $[0,T]$.

Similarly, using the assumptions of local Lipschitz continuity on the functions f, 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 [100, Thm. 2.3.1].

Furthermore, as above, using the definition of the truncation in (27), 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 [100, Thm. 9.3].

Finally, since $G^r_{N,\{\mu,\sigma\}}(t,\cdot) = G^r_{\{\mu,\sigma\}}(t,\cdot)$ for all $t\in[0,\tau_N]$, we may invoke [47, Thm. 3.5], [59, Thm. 5.2.9] to conclude that $Y_{N,t}$ is a unique solution to (25) on $[0,\tau_N]$. □

The next two results help us with the computation of $dV(Y_{N,t})$ in the proof of Lemma 3.1.

Proposition C.1 *Let $Y_{N,t}$ be the strong solution of (27), and let $\tau(t)$ be the stopping time defined in (29), Lemma 3.1. Then,*

$$\int_0^{\tau(t)} e^{2\lambda\nu} \left(\nabla V\left(Y_{N,\nu}\right)^\top G_\mu\left(\nu, Y_{N,\nu}\right) + \frac{1}{2} Tr\left[H_\sigma\left(\nu, Y_{N,\nu}\right) \nabla^2 V\left(Y_{N,\nu}\right)\right]\right) d\nu$$
$$\leq -2\lambda \int_0^{\tau(t)} e^{2\lambda\nu} V\left(Y_{N,\nu}\right) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu} \phi^r_U\left(\nu, Y_{N,\nu}\right) d\nu$$
$$+ \int_0^{\tau(t)} e^{2\lambda\nu} \left(\phi^r_\mu\left(\nu, Y_{N,\nu}\right) + \phi^r_{\mu^\parallel}\left(\nu, Y_{N,\nu}\right)\right) d\nu, \quad \text{(C.9a)}$$

$$\int_0^{\tau(t)} e^{2\lambda\nu} \nabla V\left(Y_{N,\nu}\right)^\top G_\sigma\left(\nu, Y_{N,\nu}\right) d\widehat{W}_\nu$$
$$= \int_0^{\tau(t)} e^{2\lambda\nu} \left(\phi^r_{\sigma_\star}\left(\nu, Y_{N,\nu}\right) dW^\star_\nu + \left[\phi^r_\sigma\left(\nu, Y_{N,\nu}\right) + \phi^r_{\sigma^\parallel}\left(\nu, Y_{N,\nu}\right)\right] dW_\nu\right), \quad \text{(C.9b)}$$

for all $t \in \mathbb{R}_{\geq 0}$, where $H_\sigma\left(\nu, Y_{N,\nu}\right) = G_\sigma\left(\nu, Y_{N,\nu}\right) G_\sigma\left(\nu, Y_{N,\nu}\right)^\top$, $G_\mu\left(\nu, Y_{N,\nu}\right)$ and $G_\sigma\left(\nu, Y_{N,\nu}\right)$ are defined in (25), and the functions ϕ^r_μ, $\phi^r_{\sigma_\star}$, and ϕ^r_σ are defined in (32) in the statement of Lemma 3.1. Additionally, we have defined

$$\phi^r_{\mu^\parallel}\left(\nu, Y_{N,\nu}\right) = V_r\left(Y_{N,\nu}\right)^\top g(\nu) \Lambda^\parallel_\mu\left(\nu, X^r_{N,t}\right), \quad \phi^r_{\sigma^\parallel}\left(\nu, Y_{N,\nu}\right) = V_r\left(Y_\nu\right)^\top g(\nu) F^\parallel_\sigma\left(\nu, X^r_{N,\nu}\right),$$
$$\phi^r_U\left(\nu, Y_{N,\nu}\right) = V_r\left(Y_{N,\nu}\right)^\top g(\nu) U^r_\nu.$$

Proof. Using the definitions of G^r_μ in (25), we have that

$$\nabla V\left(Y_{N,\nu}\right)^\top G_\mu\left(\nu, Y_{N,\nu}\right) = V_\star\left(Y_{N,\nu}\right)^\top \bar{F}_\mu\left(\nu, X^\star_{N,\nu}\right) + V_r\left(Y_{N,\nu}\right)^\top \bar{F}_\mu\left(\nu, X^r_{N,\nu}\right)$$
$$+ V_r\left(Y_{N,\nu}\right)^\top \left(g(\nu) U^r_\nu + \Lambda_\mu\left(\nu, X^r_{N,\nu}\right)\right), \quad \nu \in [0, \tau(t)],$$

which, upon using (9), Assumption 2 can be re-written as

$$\nabla V\left(Y_{N,\nu}\right)^\top G_\mu\left(\nu, Y_{N,\nu}\right) \leq -2\lambda V\left(Y_{N,\nu}\right) + V_r\left(Y_{N,\nu}\right)^\top \left(g(\nu) U^r_\nu + \Lambda_\mu\left(\nu, X^r_{N,\nu}\right)\right), \quad \nu \in [0, \tau(t)]. \quad \text{(C.10)}$$

We develop the expression further by using (11) in Assumption 4 to conclude that

$$\Lambda_\mu\left(\nu, X^r_{N,\nu}\right) = \begin{bmatrix} g(\nu) & g(\nu)^\perp \end{bmatrix} \begin{bmatrix} \Lambda^\parallel_\mu\left(\nu, X^r_{N,\nu}\right) \\ \Lambda^\perp_\mu\left(\nu, X^r_{N,\nu}\right) \end{bmatrix} = g(\nu) \Lambda^\parallel_\mu\left(\nu, X^r_{N,\nu}\right) + g(\nu)^\perp \Lambda^\perp_\mu\left(\nu, X^r_{N,\nu}\right).$$

Substituting into (C.10) yields

$$\nabla V\left(Y_{N,\nu}\right)^\top G_\mu\left(\nu, Y_{N,\nu}\right) \leq -2\lambda V\left(Y_{N,\nu}\right) + V_r\left(Y_{N,\nu}\right)^\top g(\nu)^\perp \Lambda^\perp_\mu\left(\nu, X^r_{N,t}\right)$$
$$+ V_r\left(Y_{N,\nu}\right)^\top g(\nu) \left(U^r_\nu + \Lambda^\parallel_\mu\left(\nu, X^r_{N,t}\right)\right), \quad \nu \in [0, \tau(t)],$$

which then leads to (C.9a).

Next, using the definition of G_σ in (25), we have that

$$\int_0^{\tau(t)} e^{2\lambda\nu} \nabla V\left(Y_{N,\nu}\right)^\top G_\sigma\left(\nu, Y_{N,\nu}\right) d\widehat{W}_\nu$$
$$= \int_0^{\tau(t)} e^{2\lambda\nu} \left(V_\star\left(Y_\nu\right)^\top \bar{F}_\sigma\left(\nu, X^\star_\nu\right) dW^\star_\nu + V_r\left(Y_\nu\right)^\top F_\sigma\left(\nu, X^r_\nu\right) dW_\nu\right)$$
$$= \int_0^{\tau(t)} e^{2\lambda\nu} V_\star\left(Y_\nu\right)^\top \bar{F}_\sigma\left(\nu, X^\star_\nu\right) dW^\star_\nu$$
$$+ \int_0^{\tau(t)} e^{2\lambda\nu} V_r\left(Y_\nu\right)^\top \left(p\left(\nu, X^r_{N,\nu}\right) + \Lambda_\sigma\left(\nu, X^r_{N,\nu}\right)\right) dW_\nu, \quad t \in \mathbb{R}_{\geq 0},$$

where we have used the definition of F_σ in (2). Since (11) and (12) in Assumptions 4 and 5, respectively, along with Definition 4 imply that

$$
\begin{aligned}
&p\left(\nu, X^r_{N,t}\right) + \Lambda_\sigma\left(\nu, X^r_{N,t}\right) \\
&\quad = g(\nu)^\perp p^\perp\left(\nu, X^r_{N,t}\right) + g(\nu)^\perp \Lambda^\perp_\sigma\left(\nu, X^r_{N,t}\right) + g(\nu) p^{\parallel}\left(\nu, X^r_{N,t}\right) + g(\nu)\Lambda^{\parallel}_\sigma\left(\nu, X^r_{N,t}\right) \\
&\quad = g(\nu)^\perp F^\perp_\sigma\left(\nu, X^r_{N,t}\right) + g(\nu) F^{\parallel}_\sigma\left(\nu, X^r_{N,t}\right), \quad \forall \nu \in [0, \tau(t)],
\end{aligned}
$$

the previous integral equality can be re-written as

$$
\begin{aligned}
&\int_0^{\tau(t)} e^{2\lambda\nu} \nabla V\left(Y_{N,\nu}\right)^\top G_\sigma\left(\nu, Y_{N,\nu}\right) d\widehat{W}_\nu \\
&= \int_0^{\tau(t)} e^{2\lambda\nu} V_\star\left(Y_\nu\right)^\top \bar{F}_\sigma\left(\nu, X^\star_\nu\right) dW^\star_\nu + \int_0^{\tau(t)} e^{2\lambda\nu} V_r\left(Y_\nu\right)^\top g(\nu)^\perp F^\perp_\sigma\left(\nu, X^r_{N,\nu}\right) dW_\nu \\
&\qquad + \int_0^{\tau(t)} e^{2\lambda\nu} V_r\left(Y_\nu\right)^\top g(\nu) F^{\parallel}_\sigma\left(\nu, X^r_{N,\nu}\right) dW_\nu, \quad t \in \mathbb{R}_{\geq 0},
\end{aligned}
$$

thus establishing the expression in (C.9b). □

In the subsequent proposition, we derive the expression for how the reference feedback process U^r (23) enters the truncated joint process $Y_{N,t}$.

Proposition C.2 *Let $Y_{N,t}$ be the strong solution of (27), and let $\tau(t)$ be the stopping time defined in (29), Lemma 3.1. Then, for the term ϕ^r_U defined in the statement of Proposition C.1, we have that*

$$
\begin{aligned}
\int_0^{\tau(t)} e^{2\lambda\nu} \phi^r_U\left(\nu, Y_{N,\nu}\right) d\nu &= \int_0^{\tau(t)} \left(\hat{\mathcal{U}}^r_\mu\left(\tau(t), \nu, Y_N; \boldsymbol{\omega}\right) d\nu + \hat{\mathcal{U}}^r_\sigma\left(\tau(t), \nu, Y_N; \boldsymbol{\omega}\right) dW_\nu\right) \\
&\quad + \int_0^{\tau(t)} e^{2\lambda\nu} \left(\phi^r_{U_\mu}\left(\nu, Y_{N,\nu}; \boldsymbol{\omega}\right) d\nu + \phi^r_{U_\sigma}\left(\nu, Y_{N,\nu}; \boldsymbol{\omega}\right) dW_\nu\right), \quad t \in \mathbb{R}_{\geq 0}, \qquad \text{(C.11)}
\end{aligned}
$$

where

$$
\begin{aligned}
&\hat{\mathcal{U}}^r_\mu\left(\tau(t), \nu, Y_N; \boldsymbol{\omega}\right) \\
&\quad = e^{-\boldsymbol{\omega}\tau(t)} \frac{\boldsymbol{\omega}}{2\lambda - \boldsymbol{\omega}} \left(e^{\boldsymbol{\omega}\tau(t)} \mathcal{P}^r\left(\tau(t), \nu\right) - e^{2\lambda\tau(t)} V_r\left(Y_{N,\tau(t)}\right)^\top g(\tau(t))\right) e^{\boldsymbol{\omega}\nu} \Lambda^{\parallel}_\mu\left(\nu, X^r_{N,\nu}\right), \\
&\hat{\mathcal{U}}^r_\sigma\left(\tau(t), \nu, Y_N; \boldsymbol{\omega}\right) \\
&\quad = e^{-\boldsymbol{\omega}\tau(t)} \frac{\boldsymbol{\omega}}{2\lambda - \boldsymbol{\omega}} \left(e^{\boldsymbol{\omega}\tau(t)} \mathcal{P}^r\left(\tau(t), \nu\right) - e^{2\lambda\tau(t)} V_r\left(Y_{N,\tau(t)}\right)^\top g(\tau(t))\right) e^{\boldsymbol{\omega}\nu} F^{\parallel}_\sigma\left(\nu, X^r_{N,\nu}\right), \\
&\phi^r_{U_\mu}\left(\nu, Y_{N,\nu}; \boldsymbol{\omega}\right) = \frac{\boldsymbol{\omega}}{2\lambda - \boldsymbol{\omega}} V_r\left(Y_{N,\nu}\right)^\top g(\nu) \Lambda^{\parallel}_\mu\left(\nu, X^r_{N,\nu}\right), \\
&\phi^r_{U_\sigma}\left(\nu, Y_{N,\nu}; \boldsymbol{\omega}\right) = \frac{\boldsymbol{\omega}}{2\lambda - \boldsymbol{\omega}} V_r\left(Y_{N,\nu}\right)^\top g(\nu) F^{\parallel}_\sigma\left(\nu, X^r_{N,\nu}\right),
\end{aligned}
$$

and where $\mathcal{P}^r\left(\tau(t), \nu\right)$ is defined in (34), Lemma 3.1.

Proof. Using the definition of ϕ^r_U in the statement of Proposition C.1, we have that

$$
\begin{aligned}
&\int_0^{\tau(t)} e^{2\lambda\nu} \phi^r_U\left(\nu, Y_{N,\nu}\right) d\nu \\
&= \int_0^{\tau(t)} e^{2\lambda\nu} V_r\left(Y_{N,\nu}\right)^\top g(\nu) U^r_\nu d\nu \\
&\quad = \int_0^{\tau(t)} e^{2\lambda\nu} V_r\left(Y_{N,\nu}\right)^\top g(\nu) \left(\mathcal{F}_{\boldsymbol{\omega}}\left(\Lambda^{\parallel}_\mu\left(\cdot, X^r\right)\right)(\nu) + \mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}\left(p^{\parallel}\left(\cdot, X^r\right) + \Lambda^{\parallel}_\sigma\left(\cdot, X^r\right), W\right)(\nu)\right) d\nu,
\end{aligned}
$$

for all $t \geq 0$, where we have incorporated the definition of U^r in (23). Next, using the definitions of $\mathcal{F}_{\boldsymbol{\omega}}$ and $\mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}\left(\cdot, W\right)$ in (18a) and (23), respectively, we can re-write the previous expression as

$$
\begin{aligned}
&\int_0^{\tau(t)} e^{2\lambda\nu} \phi^r_U\left(\nu, Y_{N,\nu}\right) d\nu \\
&= \int_0^{\tau(t)} \int_0^{\nu} \left(-\boldsymbol{\omega} e^{(2\lambda - \boldsymbol{\omega})\nu}\right) V_r\left(Y_{N,\nu}\right)^\top g(\nu) \left(e^{\boldsymbol{\omega}\beta} \Lambda^{\parallel}_\mu\left(\beta, X^r_{N,\beta}\right) d\beta\right) d\nu
\end{aligned}
$$

$$+ \int_0^{\tau(t)} \int_0^{\nu} \left(-\omega e^{(2\lambda-\omega)\nu}\right) V_r\left(Y_{N,\nu}\right)^\top g(\nu) \left(e^{\omega\beta} F_\sigma^{\parallel}\left(\beta, X^r_{N,\beta}\right) dW_\beta\right) d\nu,$$

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:

$$\begin{aligned}
&\int_0^{\tau(t)} e^{2\lambda\nu} \phi_U^r\left(\nu, Y_{N,\nu}\right) d\nu \\
&= \int_0^{\tau(t)} \left(-\int_\nu^{\tau(t)} \omega e^{(2\lambda-\omega)\beta} V_r\left(Y_{N,\beta}\right)^\top g(\beta) d\beta\right) e^{\omega\nu} \Lambda_\mu^{\parallel}\left(\nu, X^r_{N,\nu}\right) d\nu \\
&\qquad + \int_0^{\tau(t)} \left(-\int_\nu^{\tau(t)} \omega e^{(2\lambda-\omega)\beta} V_r\left(Y_{N,\beta}\right)^\top g(\beta) d\beta\right) e^{\omega\nu} F_\sigma^{\parallel}\left(\nu, X^r_{N,\nu}\right) dW_\nu,
\end{aligned} \tag{C.12}$$

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, observe that

$$d_\beta\left[e^{(2\lambda-\omega)\beta} V_r\left(Y_{N,\beta}\right)^\top g(\beta)\right] = \left(\frac{d}{d\beta} e^{(2\lambda-\omega)\beta}\right) V_r\left(Y_{N,\beta}\right)^\top g(\beta) + e^{(2\lambda-\omega)\beta} d_\beta\left[V_r\left(Y_{N,\beta}\right)^\top g(\beta)\right],$$

where $d_\beta[\cdot]$ denotes the stochastic differential with respect to the variable β. Multiplying both sides by $-\frac{\omega}{2\lambda-\omega}$ yields

$$\begin{aligned}
&-\frac{\omega}{2\lambda-\omega} d_\beta\left[e^{(2\lambda-\omega)\beta} V_r\left(Y_{N,\beta}\right)^\top g(\beta)\right] \\
&\quad = -\frac{\omega}{2\lambda-\omega}\left(\frac{d}{d\beta} e^{(2\lambda-\omega)\beta}\right) V_r\left(Y_{N,\beta}\right)^\top g(\beta) - \frac{\omega}{2\lambda-\omega} e^{(2\lambda-\omega)\beta} d_\beta\left[V_r\left(Y_{N,\beta}\right)^\top g(\beta)\right] \\
&\quad = -\omega e^{(2\lambda-\omega)\beta} V_r\left(Y_{N,\beta}\right)^\top g(\beta) - \frac{\omega}{2\lambda-\omega} e^{(2\lambda-\omega)\beta} d_\beta\left[V_r\left(Y_{N,\beta}\right)^\top g(\beta)\right].
\end{aligned}$$

We can alternatively write the expression above as

$$\begin{aligned}
&-\omega e^{(2\lambda-\omega)\beta} V_r\left(Y_{N,\beta}\right)^\top g(\beta) \\
&\quad = -\frac{\omega}{2\lambda-\omega} d_\beta\left[e^{(2\lambda-\omega)\beta} V_r\left(Y_{N,\beta}\right)^\top g(\beta)\right] + \frac{\omega}{2\lambda-\omega} e^{(2\lambda-\omega)\beta} d_\beta\left[V_r\left(Y_{N,\beta}\right)^\top g(\beta)\right],
\end{aligned}$$

which upon integration over the interval $[\nu, \tau(t)]$ produces

$$\begin{aligned}
&-\int_\nu^{\tau(t)} \omega e^{(2\lambda-\omega)\beta} V_r\left(Y_{N,\beta}\right)^\top g(\beta) d\beta \\
&\quad = -\frac{\omega}{2\lambda-\omega} e^{(2\lambda-\omega)\tau(t)} V_r\left(Y_{N,\tau(t)}\right)^\top g(\tau(t)) + \frac{\omega}{2\lambda-\omega} e^{(2\lambda-\omega)\nu} V_r\left(Y_{N,\nu}\right)^\top g(\nu) \\
&\qquad + \frac{\omega}{2\lambda-\omega} \mathcal{P}^r\left(\tau(t), \nu\right),
\end{aligned}$$

where we have used the definition of $\mathcal{P}^r\left(\tau(t), \nu\right)$ from (34). 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. □

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 $\mathcal{P}^r\left(\tau(t), \nu\right)$ in (34) in the statement of Lemma 3.1 which we restate below:*

$$\mathcal{P}^r\left(\tau(t), \nu\right) = \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} d_\beta\left[V_r\left(Y_{N,\beta}\right)^\top g(\beta)\right] \in \mathbb{R}^{1\times m}, \quad 0 \leq \nu \leq \tau(t), \tag{C.13}$$

where the $\tau(t)$ is defined in (29), and $V_r\left(Y_{N,t}\right) \doteq \nabla_{X^r_{N,t}} V\left(X^\star_{N,t}, X^r_{N,t}\right) \in \mathbb{R}^n$. Then, $\mathcal{P}^r\left(\tau(t), \nu\right)$ admits the following representation:

$$\mathcal{P}^r\left(\tau(t), \nu\right) = \mathcal{P}^r_\circ\left(\tau(t), \nu\right) + \mathcal{P}^r_{ad}\left(\tau(t), \nu\right) \in \mathbb{R}^{1\times m}, \quad 0 \leq \nu \leq \tau(t),\ t \in \mathbb{R}_{\geq 0}, \tag{C.14}$$

where

$$\mathcal{P}^r_{\circ}(\tau(t),\nu) = \sum_{i=1}^{3}\int_{\nu}^{\tau(t)} e^{(2\lambda-\omega)\beta}\mathcal{P}^r_{\mu_i}(\beta)^\top d\beta + \int_{\nu}^{\tau(t)} e^{(2\lambda-\omega)\beta}\left[\mathcal{P}^r_{\sigma}(\beta)dW_\beta + \mathcal{P}^r_{\sigma_\star}(\beta)dW^\star_\beta\right]^\top \in \mathbb{R}^{1\times m}, \tag{C.15a}$$

$$\mathcal{P}^r_{ad}(\tau(t),\nu) = \int_{\nu}^{\tau(t)} e^{(2\lambda-\omega)\beta}\mathcal{P}^r_{\mathcal{U}}(\beta)^\top d\beta \in \mathbb{R}^{1\times m}, \tag{C.15b}$$

and where $\mathcal{P}^r_{\mathcal{U}}(\beta)$*,* $\mathcal{P}^r_{\mu_1}(\beta) \in \mathbb{R}^m$*,* $i \in \{1,2,3\}$*, and* $\mathcal{P}^r_{\sigma}(\beta)$*,* $\mathcal{P}^r_{\sigma_\star}(\beta) \in \mathbb{R}^{m\times d}$ *are defined as*

$$\mathcal{P}^r_{\mathcal{U}}(\beta) = g(\beta)^\top V_{r,r}(Y_{N,\beta})\, g(\beta) U^r_\beta,$$
$$\mathcal{P}^r_{\mu_1}(\beta) = \dot{g}(\beta)^\top V_r(Y_{N,\beta}),$$
$$\mathcal{P}^r_{\mu_2}(\beta) = g(\beta)^\top \left[V_{r,r}(Y_{N,\beta})\left(\bar{F}_\mu\left(\beta, X^r_{N,\beta}\right) + \Lambda_\mu\left(\beta, X^r_{N,\beta}\right)\right) + V_{\star,r}(Y_{N,\beta})^\top \bar{F}_\mu\left(\beta, X^\star_{N,\beta}\right)\right],$$
$$\mathcal{P}^r_{\mu_3}(\beta) = \frac{1}{2} g(\beta)^\top \vec{Tr}\left[H_\sigma(\beta, Y_{N,\beta})\nabla^2 V_{r_i}(Y_{N,\beta})\right]_{i=1}^n$$

and

$$\mathcal{P}^r_{\sigma}(\beta) = g(\beta)^\top V_{r,r}(Y_{N,\beta}) F_\sigma\left(\beta, X^r_{N,\beta}\right), \quad \mathcal{P}^r_{\sigma_\star}(\beta) = g(\beta)^\top V_{\star,r}(Y_{N,\beta})^\top \bar{F}_\sigma\left(\beta, X^\star_{N,\beta}\right).$$

Additionally, we have defined $H_\sigma(\beta, Y_{N,\beta}) \doteq G_\sigma(\beta, Y_{N,\beta}) G_\sigma(\beta, Y_{N,\beta})^\top \in \mathbb{S}^{2n}$ *and*

$$\vec{Tr}\left[H_\sigma(\beta, Y_{N,\beta})\nabla^2 V_{r_i}(Y_{N,\beta})\right]_{i=1}^n \doteq \left[\mathfrak{T}_1(\beta, Y_{N,\beta}) \quad \cdots \quad \mathfrak{T}_n(\beta, Y_{N,\beta})\right]^\top \in \mathbb{R}^n,$$
$$\mathfrak{T}_i(\beta, Y_{N,\beta}) = Tr\left[H_\sigma(\beta, Y_{N,\beta})\nabla^2 V_{r_i}(Y_{N,\beta})\right] \in \mathbb{R}.$$

Proof. We begin by writing $V_r(Y_{N,\beta})^\top g(\beta)$ as

$$V_r(Y_{N,\beta})^\top g(\beta) = \left[\nabla_r V(Y_{N,\beta})^\top g_{\cdot,1}(\beta) \quad \cdots \quad \nabla_r V(Y_{N,\beta})^\top g_{\cdot,m}(\beta)\right] = \left[\sum_{i=1}^n V_{r_i}(Y_{N,\beta}) g_{i,1}(\beta) \quad \cdots \quad \sum_{i=1}^n V_{r_i}(Y_{N,\beta}) g_{i,m}(\beta)\right] \in \mathbb{R}^{1\times m}, \tag{C.16}$$

where $g_{\cdot j}(\beta) \in \mathbb{R}^n$ is the j-th column of $g(\beta)$. Applying Itô's lemma to $V_{r_i}(Y_{N,\beta}) g_{i,j}(\beta) \in \mathbb{R}$, $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$, and using the truncated dynamics in (27) we get

$$d_\beta\left[V_{r_i}(Y_{N,\beta}) g_{i,j}(\beta)\right] = \left[V_{r_i}(Y_{N,\beta})\dot{g}_{i,j}(\beta) + \left(\nabla V_{r_i}(Y_{N,\beta})^\top G_\mu(\beta, Y_{N,\beta}) + \frac{1}{2}\text{Tr}\left[H_\sigma(\beta, Y_{N,\beta})\nabla^2 V_{r_i}(Y_{N,\beta})\right]\right) g_{i,j}(\beta)\right] d\beta + \nabla V_{r_i}(Y_{N,\beta})^\top G_\sigma(\beta, Y_{N,\beta}) g_{i,j}(\beta) d\widehat{W}_\beta, \tag{C.17}$$

where we have replaced $G_{N,\mu}$ and $H_{N,\sigma} = G_{N,\sigma}G_{N,\sigma}^\top \in \mathbb{S}^{2n}$ with G_μ and H_σ because from Proposition 3.1, $Y_{N,\beta}$ is also a strong solution of the joint process (25) for all $\beta \in [\nu, \tau(t)] \subseteq [0, \tau^\star] \subseteq [0, \tau_N]$. See (29) for the definition of the stopping times $\tau^\star$ and τ_N. Since (C.16) implies that

$$d_\beta\left[\nabla_r V(Y_{N,\beta})^\top g_{\cdot j}(\beta)\right] = \sum_{i=1}^n d_\beta\left[V_{r_i}(Y_{N,\beta}) g_{i,j}(\beta)\right] \in \mathbb{R}, \quad j \in \{1,\ldots,m\},$$

we may substitute the expression in (C.17) to obtain

$$d_\beta\left[\nabla_r V(Y_{N,\beta})^\top g_{\cdot j}(\beta)\right] = V_r(Y_{N,\beta})^\top \dot{g}_{\cdot j}(\beta) d\beta + \left(\nabla V_r(Y_{N,\beta})^\top G_\mu(\beta, Y_{N,\beta}) + \frac{1}{2}\vec{\text{Tr}}\left[H_\sigma(\beta, Y_{N,\beta})\nabla^2 V_{r_i}(Y_{N,\beta})\right]_{i=1}^n\right)^\top g_{\cdot j}(\beta) d\beta + g_{\cdot j}(\beta)^\top \nabla V_r(Y_{N,\beta})^\top G_\sigma(\beta, Y_{N,\beta}) d\widehat{W}_\beta \in \mathbb{R},$$

for $j \in \{1, \ldots, m\}$. Once again. from (C.16) we have that

$$V_r\left(Y_{N,\beta}\right)^\top g(\beta) = \begin{bmatrix} \nabla_r V\left(Y_{N,\beta}\right)^\top g_{\cdot,1}(\beta) & \cdots & \nabla_r V\left(Y_{N,\beta}\right)^\top g_{\cdot,m}(\beta) \end{bmatrix} \in \mathbb{R}^{1\times m},$$

we therefore use the previous expression to write

$$\begin{aligned} d_\beta\left[V_r\left(Y_{N,\beta}\right)^\top g(\beta)\right] =& V_r\left(Y_{N,\beta}\right)^\top \dot{g}(\beta) d\beta \\ &+ \left(\nabla V_r\left(Y_{N,\beta}\right)^\top G_\mu\left(\beta, Y_{N,\beta}\right) + \frac{1}{2}\vec{\operatorname{Tr}}\left[H_\sigma\left(\beta, Y_{N,\beta}\right)\nabla^2 V_{r_i}\left(Y_{N,\beta}\right)\right]_{i=1}^n\right)^\top g(\beta) d\beta \\ &+ \left(g(\beta)^\top \nabla V_r\left(Y_{N,\beta}\right)^\top G_\sigma\left(\beta, Y_{N,\beta}\right) d\widehat{W}_\beta\right)^\top \in \mathbb{R}^{1\times m}. \end{aligned} \tag{C.18}$$

Now, using the definition of the Hessian of vector valued functions in Sec. 1.3, we observe that

$$\mathbb{R}^{2n\times n} \ni \nabla V_r\left(Y_{N,\beta}\right) = \begin{bmatrix} \nabla V_{r_1} & \cdots & \nabla V_{r_n} \end{bmatrix} = \nabla\cdot\left(\nabla_r V\right)^\top = \begin{bmatrix} \nabla_r\cdot\left(\nabla_r V\right)^\top \\ \nabla_\star\cdot\left(\nabla_r V\right)^\top \end{bmatrix} = \begin{bmatrix} \nabla_r^2 V \\ \nabla_{\star,r}^2 V \end{bmatrix} = \begin{bmatrix} V_{r,r}\left(Y_{N,\beta}\right) \\ V_{\star,r}\left(Y_{N,\beta}\right) \end{bmatrix},$$

where $V_{\star,r}\left(Y_{N,\beta}\right) \in \mathbb{R}^{n\times n}$ and $V_{r,r}\left(Y_{N,\beta}\right) \in \mathbb{S}^n$. Therefore, using the definition of G_μ and G_σ in (25), we get

$$\begin{aligned} \nabla V_r\left(Y_{N,\beta}\right)^\top G_\mu\left(\beta, Y_{N,\beta}\right) &= \begin{bmatrix} V_{r,r}\left(Y_{N,\beta}\right) & V_{\star,r}\left(Y_{N,\beta}\right)^\top \end{bmatrix} \begin{bmatrix} F_\mu\left(\beta, X^r_{N,\beta}, U^r_\beta\right) \\ \bar{F}_\mu\left(\beta, X^\star_{N,\beta}\right) \end{bmatrix} \\ &= V_{r,r}\left(Y_{N,\beta}\right) F_\mu\left(\beta, X^r_{N,\beta}, U^r_\beta\right) + V_{\star,r}\left(Y_{N,\beta}\right)^\top \bar{F}_\mu\left(\beta, X^\star_{N,\beta}\right) \in \mathbb{R}^n, \\ \nabla V_r\left(Y_{N,\beta}\right)^\top G_\sigma\left(\beta, Y_{N,\beta}\right) d\widehat{W}_\beta &= \begin{bmatrix} V_{r,r}\left(Y_{N,\beta}\right) & V_{\star,r}\left(Y_{N,\beta}\right)^\top \end{bmatrix} \begin{bmatrix} F_\sigma\left(\beta, X^r_{N,\beta}\right) & 0_{n,d} \\ 0_{n,d} & \bar{F}_\sigma\left(\beta, X^\star_{N,\beta}\right) \end{bmatrix} \begin{bmatrix} dW_\beta \\ dW^\star_\beta \end{bmatrix} \\ &= V_{r,r}\left(Y_{N,\beta}\right) F_\sigma\left(\beta, X^r_{N,\beta}\right) dW_\beta + V_{\star,r}\left(Y_{N,\beta}\right)^\top \bar{F}_\sigma\left(\beta, X^\star_{N,\beta}\right) dW^\star_\beta \in \mathbb{R}^n. \end{aligned}$$

Next, note that the decomposition (4) in Definition 1 states that

$$F_\mu\left(\beta, X^r_{N,\beta}, U^r_\beta\right) = \bar{F}_\mu\left(\beta, X^r_{N,\beta}\right) + g(\beta) U^r_\beta + \Lambda_\mu\left(\beta, X^r_{N,\beta}\right).$$

Therefore, the previous expressions can be re-written as

$$\begin{aligned} &\nabla V_r\left(Y_{N,\beta}\right)^\top G_\mu\left(\beta, Y_{N,\beta}\right) \\ &\quad = V_{r,r}\left(Y_{N,\beta}\right)\left(\bar{F}_\mu\left(\beta, X^r_{N,\beta}\right) + \Lambda_\mu\left(\beta, X^r_{N,\beta}\right)\right) + V_{\star,r}\left(Y_{N,\beta}\right)^\top \bar{F}_\mu\left(\beta, X^\star_{N,\beta}\right) \\ &\qquad + V_{r,r}\left(Y_{N,\beta}\right) g(\beta) U^r_\beta \in \mathbb{R}^n, \\ &\nabla V_r\left(Y_{N,\beta}\right)^\top G_\sigma\left(\beta, Y_{N,\beta}\right) d\widehat{W}_\beta = \begin{bmatrix} V_{r,r}\left(Y_{N,\beta}\right) & V_{\star,r}\left(Y_{N,\beta}\right)^\top \end{bmatrix} \begin{bmatrix} F_\sigma\left(\beta, X^r_{N,\beta}\right) & 0_{n,d} \\ 0_{n,d} & \bar{F}_\sigma\left(\beta, X^\star_{N,\beta}\right) \end{bmatrix} \begin{bmatrix} dW_\beta \\ dW^\star_\beta \end{bmatrix} \\ &\quad = V_{r,r}\left(Y_{N,\beta}\right) F_\sigma\left(\beta, X^r_{N,\beta}\right) dW_\beta + V_{\star,r}\left(Y_{N,\beta}\right)^\top \bar{F}_\sigma\left(\beta, X^\star_{N,\beta}\right) dW^\star_\beta \in \mathbb{R}^n. \end{aligned}$$

Substituting the above identities into (C.18) produces

$$d_\beta\left[V_r\left(Y_{N,\beta}\right)^\top g(\beta)\right] = \left[\sum_{i=1}^3 \mathcal{P}^r_{\mu_i}(\beta) + \mathcal{P}^r_{\mathcal{U}}(\beta)\right]^\top d\beta + \left[\mathcal{P}^r_\sigma(\beta) dW_\beta + \mathcal{P}^r_{\sigma_\star}(\beta) dW^\star_\beta\right]^\top \in \mathbb{R}^{1\times m}.$$

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 *Consider the functions $\mathcal{P}^r_{\mu_i}(t) \in \mathbb{R}^m$, $i \in \{1,2,3\}$, and $\mathcal{P}^r_\sigma(t)$, $\mathcal{P}^r_{\sigma_\star}(t) \in \mathbb{R}^{m\times d}$ defined in the statement of Proposition C.3. If the stopping time $\tau^\star$, defined in (29), Lemma 3.1, satisfies $\tau^\star = t^\star$, then*

$$\sum_{i=1}^3 \left\|\left(\mathcal{P}^r_{\mu_i}\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi^0_\star} \leq \sqrt{\frac{n}{2}}\Delta_g\left(\Delta_{\partial V}\Delta^r_{\mathcal{P}_\mu} + \frac{1}{2}\Delta_{\partial^2 V}\Delta^r_{\mathcal{P}_\sigma}(4\mathsf{p}, 2\mathsf{p})\right) + \Delta_{\dot{g}}\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi^0_\star}, \tag{C.19a}$$

$$\sum_{i\in\{\sigma,\sigma_\star\}} \left\| \left(\mathcal{P}_i^r\right)_{t^\star} \right\|_{\mathsf{q}}^{\pi_\star^0} \leq \sqrt{\frac{n}{2}} \Delta_g \Delta_{\partial V} \Delta_{\mathcal{P}_\sigma}^r(\mathsf{q},\mathsf{q}), \quad \mathsf{q} \in \{2\mathsf{p}, 4\mathsf{p}\}, \tag{C.19b}$$

where

$$\Delta_{\mathcal{P}_\mu}^r = \left\| \left(\bar{F}_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\bar{F}_\mu\left(\cdot, X_N^\star\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\Lambda_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0},$$

$$\Delta_{\mathcal{P}_\sigma}^r(\mathsf{r},\mathsf{s}) = \mathbb{E}_{\pi_\star^0}\left[\left(\bar{F}_\sigma\left(\cdot, X_N^r\right)\right)_{t^\star}^{\mathsf{r}}\right]^{\frac{1}{\mathsf{s}}} + \mathbb{E}_{\pi_\star^0}\left[\left(\bar{F}_\sigma\left(\cdot, X_N^\star\right)\right)_{t^\star}^{\mathsf{r}}\right]^{\frac{1}{\mathsf{s}}} + \mathbb{E}_{\pi_\star^0}\left[\left(\Lambda_\sigma\left(\cdot, X_N^r\right)\right)_{t^\star}^{\mathsf{r}}\right]^{\frac{1}{\mathsf{s}}},$$

for $(\mathsf{r},\mathsf{s}) \in \{2\mathsf{p}, 4\mathsf{p}\} \times \{2\mathsf{p}, 4\mathsf{p}\}$.

Proof. We begin by using the definition of $\mathcal{P}_{\mu_1}^r$ in (C.15a) to obtain

$$\left\|\mathcal{P}_{\mu_1}^r(t)\right\| \leq \|\dot{g}(t)\|_F \left\|V_r\left(Y_{N,t}\right)\right\| \leq \Delta_{\dot{g}} \left\|V_r\left(Y_{N,t}\right)\right\|, \quad \forall t \in [0,T],$$

where we have used the bound on $\dot{g}(t)$ in Assumption 1. It then follows that

$$\left(\mathcal{P}_{\mu_1}^r\right)_{t^\star} \leq \Delta_{\dot{g}}\left(V_r\left(Y_N\right)\right)_{t^\star},$$

and thus

$$\left\| \left(\mathcal{P}_{\mu_1}^r\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \leq \Delta_{\dot{g}} \left\| \left(V_r\left(Y_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0}. \tag{C.20}$$

Similarly, using the bound on $g(t)$ in Assumption 1, we obtain

$$\left\|\mathcal{P}_{\mu_2}^r(t)\right\| \leq \Delta_g \left(\left\|V_{r,r}\left(Y_{N,t}\right)\right\|_F \left(\left\|\bar{F}_\mu\left(t, X_{N,t}^r\right)\right\| + \left\|\Lambda_\mu\left(t, X_{N,t}^r\right)\right\|\right) + \left\|V_{\star,r}\left(Y_{N,t}\right)\right\|_F \left\|\bar{F}_\mu\left(t, X_{N,t}^\star\right)\right\|\right), \quad \forall t \in [0,T].$$

Using the bound in (E.1b), Proposition E.1, produces the following bound:

$$\left\|\mathcal{P}_{\mu_2}^r(t)\right\| \leq \sqrt{\frac{n}{2}} \Delta_g \Delta_{\partial V} \left(\left\|\bar{F}_\mu\left(t, X_{N,t}^r\right)\right\| + \left\|\bar{F}_\mu\left(t, X_{N,t}^\star\right)\right\| + \left\|\Lambda_\mu\left(t, X_{N,t}^r\right)\right\|\right), \quad \forall t \in [0,T].$$

Therefore, we conclude that

$$\left(\mathcal{P}_{\mu_2}^r\right)_{t^\star} \leq \sqrt{\frac{n}{2}} \Delta_g \Delta_{\partial V} \left(\left(\bar{F}_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} + \left(\bar{F}_\mu\left(\cdot, X_N^\star\right)\right)_{t^\star} + \left(\Lambda_\mu\left(\cdot, X_N^r\right)\right)_{t^\star}\right).$$

It then follows due to the Minkowski's inequality that

$$\left\| \left(\mathcal{P}_{\mu_2}^r\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \leq \sqrt{\frac{n}{2}} \Delta_g \Delta_{\partial V} \left(\left\| \left(\bar{F}_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\bar{F}_\mu\left(\cdot, X_N^\star\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\Lambda_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \right). \tag{C.21}$$

Next, we consider the term $\mathcal{P}_{\mu_3}^r$ defined in (C.15a), using which we obtain the following bound:

$$\left\|\mathcal{P}_{\mu_3}^r(t)\right\| \leq \frac{1}{2}\Delta_g \left\| \vec{\mathrm{Tr}}\left[H_\sigma\left(t, Y_{N,t}\right) \nabla^2 V_{r_i}\left(Y_{N,t}\right)\right]_{i=1}^n \right\| = \frac{1}{2}\Delta_g \left(\sum_{i=1}^n \left|\mathfrak{T}_i\left(t, Y_{N,t}\right)\right|^2\right)^{\frac{1}{2}}, \tag{C.22}$$

where

$$\mathfrak{T}_i\left(t, Y_{N,t}\right) = \mathrm{Tr}\left[H_\sigma\left(t, Y_{N,t}\right) \nabla^2 V_{r_i}\left(Y_{N,t}\right)\right] \in \mathbb{R}, \quad H_\sigma\left(t, Y_{N,t}\right) \doteq G_\sigma\left(t, Y_{N,t}\right) G_\sigma\left(t, Y_{N,t}\right)^\top \in \mathbb{S}^{2n}.$$

Now, for $\nabla^2 V_{r_i}\left(Y_{N,t}\right) \in \mathbb{S}^{2n}$, it is straightforward to establish that $\left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F \cdot \mathbb{I}_{2n} - \nabla^2 V_{r_i}\left(Y_{N,t}\right) \in \mathbb{S}^{2n}_{\succeq 0}$. Moreover, by definition $\mathbb{S}^{2n} \ni H_\sigma\left(t, Y_{N,t}\right) = G_\sigma\left(t, Y_{N,t}\right) G_\sigma\left(t, Y_{N,t}\right)^\top \in \mathbb{S}^{2n}_{\succeq 0}$. It then follows from [109, Thm. 7.5] that

$$\mathrm{Tr}\left[H_\sigma\left(t, Y_{N,t}\right)\left(\left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F \cdot \mathbb{I}_{2n} - \nabla^2 V_{r_i}\left(Y_{N,t}\right)\right)\right] \geq 0.$$

Thus, as a consequence of the linearity of the trace operator

$$\begin{aligned} 0 \leq& \mathrm{Tr}\left[H_\sigma\left(t, Y_{N,t}\right)\left(\left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F \cdot \mathbb{I}_{2n} - \nabla^2 V_{r_i}\left(Y_{N,t}\right)\right)\right] \\ =& \left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F \mathrm{Tr}\left[H_\sigma\left(t, Y_{N,t}\right)\right] - \mathrm{Tr}\left[H_\sigma\left(t, Y_{N,t}\right) \nabla^2 V_{r_i}\left(Y_{N,t}\right)\right], \quad \forall (t,i) \in [0,T] \times \{1,\ldots,n\}, \end{aligned}$$

and therefore

$$\text{Tr}\left[H_\sigma\left(t, Y_{N,t}\right) \nabla^2 V_{r_i}\left(Y_{N,t}\right)\right] \leq \left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F \text{Tr}\left[H_\sigma\left(t, Y_{N,t}\right)\right] = \left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F \left\|G_\sigma\left(t, Y_{N,t}\right)\right\|_F^2,$$

for all $(t, i) \in [0, T] \times \{1, \ldots, n\}$. Using the definition of G_σ in (25), we get

$$\begin{aligned}\text{Tr}\left[H_\sigma\left(t, Y_{N,t}\right) \nabla^2 V_{r_i}\left(Y_{N,t}\right)\right] &\leq \left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F \left\|G_\sigma\left(t, Y_{N,t}\right)\right\|_F^2 \\ &= \left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F \left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F^2 + \left\|\bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\|_F^2\right), \end{aligned} \tag{C.23}$$

and thus

$$\left|\mathfrak{T}_i\left(t, Y_{N,t}\right)\right| = \left|\text{Tr}\left[H_\sigma\left(t, Y_{N,t}\right) \nabla^2 V_{r_i}\left(Y_{N,t}\right)\right]\right| \leq \left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F \left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F^2 + \left\|\bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\|_F^2\right),$$

for all $(t, i) \in [0, T] \times \{1, \ldots, n\}$. Substituting into (C.22) then leads to

$$\begin{aligned}\left\|\mathcal{P}_{\mu_3}^r(t)\right\| &\leq \frac{1}{2}\Delta_g \left(\sum_{i=1}^n \left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F^2\right)^{\frac{1}{2}} \left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F^2 + \left\|\bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\|_F^2\right) \\ &\overset{(i)}{\leq} \frac{1}{2}\Delta_g \left(\sum_{i=1}^n \left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F^2\right)^{\frac{1}{2}} \left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F^2 + \left\|\bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\|_F^2\right) \\ &\overset{(ii)}{\leq} \frac{1}{2}\Delta_g \left(\sum_{i=1}^n \left\|\nabla^2 V_{r_i}\left(Y_{N,t}\right)\right\|_F\right) \left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F^2 + \left\|\bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\|_F^2\right),\end{aligned}$$

for all $t \in [0, T]$, where (i) is due to $\|\cdot\| \doteq \|\cdot\|_2 \leq \|\cdot\|_F$ for matrices, and (ii) is due to the equivalence $\|\cdot\| \doteq \|\cdot\|_2 \leq \|\cdot\|_1$ for vectors. Substituting the bound in (E.1c), Proposition E.1 produces

$$\left\|\mathcal{P}_{\mu_3}^r(t)\right\| \leq \frac{1}{2}\sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial^2 V} \left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F^2 + \left\|\bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\|_F^2\right), \quad \forall t \in [0, T].$$

Consequently

$$\left(\mathcal{P}_{\mu_3}^r\right)_{t^\star} \leq \frac{1}{2}\sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial^2 V} \left(\left(F_\sigma\left(\cdot, X_N\right)\right)_{t^\star}^2 + \left(\bar{F}_\sigma\left(\cdot, X_N^\star\right)\right)_{t^\star}^2\right).$$

It then follows due to the Minkowski's inequality that

$$\left\|\left(\mathcal{P}_{\mu_3}^r\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \leq \frac{1}{2}\sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial^2 V} \left(\mathbb{E}_{\pi_\star^0}\left[\left(F_\sigma\left(\cdot, X_N\right)\right)_{t^\star}^{4\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} + \mathbb{E}_{\pi_\star^0}\left[\left(\bar{F}_\sigma\left(\cdot, X_N^\star\right)\right)_{t^\star}^{4\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}\right).$$

Using the decomposition $F_\sigma = \bar{F}_\sigma + \Lambda_\sigma$ in (4) followed by the Minkowski's inequality, we obtain

$$\begin{aligned}&\left\|\left(\mathcal{P}_{\mu_3}^r\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \\ &\leq \frac{1}{2}\sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial^2 V} \left(\mathbb{E}_{\pi_\star^0}\left[\left(\bar{F}_\sigma\left(\cdot, X_N^r\right)\right)_{t^\star}^{4\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} + \mathbb{E}_{\pi_\star^0}\left[\left(\bar{F}_\sigma\left(\cdot, X_N^\star\right)\right)_{t^\star}^{4\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} + \mathbb{E}_{\pi_\star^0}\left[\left(\Lambda_\sigma\left(\cdot, X_N^r\right)\right)_{t^\star}^{4\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}\right).\end{aligned} \tag{C.24}$$

Adding the bounds in (C.20), (C.21), and (C.24), establishes (C.19a).

Next, using the definitions of $\mathcal{P}_\sigma^r$ and $\mathcal{P}_{\sigma_\star}^r$, we obtain

$$\left\|\mathcal{P}_\sigma^r(t)\right\|_F \leq \sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial V} \left\|F_\sigma\left(t, X_{N,t}^r\right)\right\|_F, \quad \left\|\mathcal{P}_{\sigma_\star}^r(t)\right\|_F \leq \sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial V} \left\|\bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\|_F, \quad \forall t \in [0, T],$$

where we have used the bound on $g(t)$ in Assumption 1, and the shared bound on $\left\|V_{r,r}\left(Y_{N,t}\right)\right\|_F$ and $\left\|V_{\star,r}\left(Y_{N,t}\right)\right\|_F$ in (E.1b), Proposition E.1. It then follows from the decomposition $F_\sigma = \bar{F}_\sigma + \Lambda_\sigma$ in (4) that

$$\left\|\mathcal{P}_\sigma^r(t)\right\|_F \leq \sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial V} \left(\left\|\bar{F}_\sigma\left(t, X_{N,t}^r\right)\right\|_F + \left\|\Lambda_\sigma\left(t, X_{N,t}^r\right)\right\|_F\right),$$

$$\left\|\mathcal{P}^r_{\sigma_\star}(t)\right\|_F \leq \sqrt{\frac{n}{2}}\Delta_g\Delta_{\partial V}\left\|\bar{F}_\sigma\left(t, X^\star_{N,t}\right)\right\|_F, \quad \forall t \in [0,T],$$

and hence,

$$\left(\mathcal{P}^r_\sigma\right)_{t^\star} \leq \sqrt{\frac{n}{2}}\Delta_g\Delta_{\partial V}\left(\left(\bar{F}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star} + \left(\Lambda_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right), \quad \left(\mathcal{P}^r_{\sigma_\star}\right)_{t^\star} \leq \sqrt{\frac{n}{2}}\Delta_g\Delta_{\partial V}\left(\bar{F}_\sigma\left(\cdot, X^\star_N\right)\right)_{t^\star}.$$

Applying the Minkowski's inequality for $\mathsf{q} \in \{2\mathsf{p}, 4\mathsf{p}\}$

$$\left\|\left(\mathcal{P}^r_\sigma\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{q}} \leq \sqrt{\frac{n}{2}}\Delta_g\Delta_{\partial V}\left(\left\|\left(\bar{F}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{q}} + \left\|\left(\Lambda_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{q}}\right),$$

$$\left\|\left(\mathcal{P}^r_{\sigma_\star}\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{q}} \leq \sqrt{\frac{n}{2}}\Delta_g\Delta_{\partial V}\left\|\left(\bar{F}_\sigma\left(\cdot, X^\star_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{q}}.$$

Therefore, we conclude that

$$\sum_{i\in\{\sigma,\sigma_\star\}}\left\|\left(\mathcal{P}^r_i\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{q}} \leq \sqrt{\frac{n}{2}}\Delta_g\Delta_{\partial V}\left(\left\|\left(\bar{F}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{q}} + \left\|\left(\Lambda_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{q}} + \left\|\left(\bar{F}_\sigma\left(\cdot, X^\star_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{q}}\right),$$

for $\mathsf{q} \in \{2\mathsf{p}, 4\mathsf{p}\}$, thus establishing (C.19b) and concluding the proof. □

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:*

$$\begin{aligned} N^r_1(t) &= \int_0^t e^{\omega\nu}M_\mu(t,\nu)\Lambda^{\parallel}_\mu\left(\nu, X^r_{N,\nu}\right)d\nu, \quad N^r_2(t) = \int_0^t e^{\omega\nu}M_\mu(t,\nu)F^{\parallel}_\sigma\left(\nu, X^r_{N,\nu}\right)dW_\nu, \\ N^r_3(t) &= \int_0^t e^{\omega\nu}M_\sigma(t,\nu)\Lambda^{\parallel}_\mu\left(\nu, X^r_{N,\nu}\right)d\nu, \quad N^r_4(t) = \int_0^t e^{\omega\nu}M_\sigma(t,\nu)F^{\parallel}_\sigma\left(\nu, X^r_{N,\nu}\right)dW_\nu, \end{aligned} \tag{C.25}$$

where

$$M_\mu(t,\nu) = \sum_{i=1}^{3}\int_\nu^t e^{(2\lambda-\omega)\beta}\mathcal{P}^r_{\mu_i}(\beta)^\top d\beta, \quad M_\sigma(t,\nu) = \int_\nu^t e^{(2\lambda-\omega)\beta}\left[\mathcal{P}^r_\sigma(\beta)dW_\beta + \mathcal{P}^r_{\sigma_\star}(\beta)dW^\star_\beta\right]^\top,$$

for $0 \leq \nu \leq t \leq T$, and where where $\mathcal{P}^r_{\mu_i}(\beta) \in \mathbb{R}^m$, $i \in \{1,2,3\}$, and $\mathcal{P}^r_\sigma(\beta)$, $\mathcal{P}^r_{\sigma_\star}(\beta) \in \mathbb{R}^{m\times d}$ are defined in the statement of Proposition C.3.

If the stopping time $\tau^\star$, defined in (29), Lemma 3.1, satisfies $\tau^\star = t^\star$, then we have the following bound for all $\mathsf{p} \in \mathbb{N}_{\geq 1}$:

$$\sum_{i=1}^{4}\left\|\left(N^r_i\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}} \leq \frac{e^{2\lambda t^\star}}{\sqrt{\lambda}\omega}\Delta^r_{\mathcal{P}_1}(t^\star)\left\|\left(\Lambda^{\parallel}_\mu\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}} + \frac{e^{2\lambda t^\star}}{\sqrt{\lambda}\omega}\Delta^r_{\mathcal{P}_2}(t^\star)\left\|\left(F^{\parallel}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{4\mathsf{p}}$$
$$+ \frac{\sqrt{m}e^{2\lambda t^\star}}{2\lambda}\left\|\left(\mathcal{P}^r_\sigma F^{\parallel}_\sigma\left(\cdot, X^r_N\right)^\top\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}}, \tag{C.26}$$

where

$$\Delta^r_{\mathcal{P}_1}(t^\star) = \frac{1}{2\sqrt{\lambda}}\left(\sum_{i=1}^{3}\left\|\left(\mathcal{P}^r_{\mu_i}\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}}\right) + 2\sqrt{\mathsf{p}}\left(\sum_{i\in\{\sigma,\sigma_\star\}}\left\|\left(\mathcal{P}^r_i\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}}\right),$$

$$\Delta^r_{\mathcal{P}_2}(t^\star) = \left(p^3\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}}\frac{1}{2\sqrt{\lambda}}\left(\sum_{i=1}^{3}\left\|\left(\mathcal{P}^r_{\mu_i}\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}}\right) + 2\mathsf{p}^{\frac{3}{2}}\left(4\mathsf{p}-1\right)^{\frac{1}{2}}\left(\sum_{i\in\{\sigma,\sigma_\star\}}\left\|\left(\mathcal{P}^r_i\right)_{t^\star}\right\|^{\pi^0_\star}_{4\mathsf{p}}\right).$$

Proof. We begin with the process N^r_1 and use the definition of M_μ to obtain

$$\left(N_1^r\right)_{t^\star} \doteq \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} M_\mu(t,\nu) \Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \right|$$
$$\leq \sum_{i=1}^{3} \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} \int_\nu^t e^{(2\lambda-\omega)\beta} \left[\mathcal{P}_{\mu_i}^r(\beta)\right]^\top d\beta \Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \right|,$$

which can be further bounded as

$$\left(N_1^r\right)_{t^\star} \leq \sum_{i=1}^{3} \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} \int_\nu^t e^{(2\lambda-\omega)\beta} \left[\mathcal{P}_{\mu_i}^r(\beta)\right]^\top d\beta \Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \right|$$
$$\leq \sum_{i=1}^{3} \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} \int_\nu^t e^{(2\lambda-\omega)\beta} d\beta d\nu \right| \left(\mathcal{P}_{\mu_i}^r\right)_{t^\star} \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}.$$

Solving the integrals in the above expression yields

$$\left(N_1^r\right)_{t^\star} \leq \sum_{i=1}^{3} \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} \int_\nu^t e^{(2\lambda-\omega)\beta} d\beta d\nu \right| \left(\mathcal{P}_{\mu_i}^r\right)_{t^\star} \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}$$
$$\leq \frac{1}{|2\lambda-\omega|} \sum_{i=1}^{3} \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} \left(e^{(2\lambda-\omega)t} - e^{(2\lambda-\omega)\nu}\right) d\nu \right| \left(\mathcal{P}_{\mu_i}^r\right)_{t^\star} \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}$$
$$= \frac{1}{|2\lambda-\omega|} \sum_{i=1}^{3} \sup_{t\in[0,t^\star]} \left| e^{(2\lambda-\omega)t}\left(e^{\omega t}-1\right)\frac{1}{\omega} - \left(e^{2\lambda t}-1\right)\frac{1}{2\lambda} \right| \left(\mathcal{P}_{\mu_i}^r\right)_{t^\star} \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}$$
$$\leq e^{2\lambda t^\star} \frac{1}{2\lambda\omega} \sum_{i=1}^{3} \left(\mathcal{P}_{\mu_i}^r\right)_{t^\star} \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star},$$

where we have used the following to obtain the last inequality:

$$\sup_{t\in[0,t^\star]} \left| e^{(2\lambda-\omega)t}\left(e^{\omega t}-1\right)\frac{1}{\omega} - \left(e^{2\lambda t}-1\right)\frac{1}{2\lambda} \right| = \sup_{t\in[0,t^\star]} \left(e^{2\lambda t} \left| \frac{2\lambda\left(1-e^{-\omega t}\right) - \omega\left(1-e^{-2\lambda t}\right)}{2\lambda\omega} \right| \right)$$
$$\leq e^{2\lambda t^\star} \frac{|2\lambda-\omega|}{2\lambda\omega}.$$

Using Minkowski's inequality one sees that

$$\left\| \left(N_1^r\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} \overset{(i)}{\leq} \frac{e^{2\lambda t^\star}}{2\lambda\omega} \sum_{i=1}^{3} \left\| \left(\mathcal{P}_{\mu_i}^r\right)_{t^\star} \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0}$$
$$\overset{(ii)}{\leq} \frac{e^{2\lambda t^\star}}{2\lambda\omega} \sum_{i=1}^{3} \left\| \left(\mathcal{P}_{\mu_i}^r\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \left\| \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0}, \quad \text{(C.27)}$$

where (i) is a consequence of the fact that $t^\star$ is a constant, and (ii) is due to the Cauchy-Schwarz inequality.
Next, from the definitions of $M_\mu(t,\nu)$ and N_2^r we obtain

$$N_2^r(t) = \int_0^t e^{\omega\nu} M_\mu(t,\nu) F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu$$
$$= \sum_{i=1}^{3} \int_0^t e^{\omega\nu} \int_\nu^t e^{(2\lambda-\omega)\beta} \left[\mathcal{P}_{\mu_i}^r(\beta)\right]^\top d\beta F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu$$
$$= \sum_{i=1}^{3} \int_0^t \int_\nu^t e^{(2\lambda-\omega)\beta} \left[\mathcal{P}_{\mu_i}^r(\beta)\right]^\top e^{\omega\nu} \left[F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) \quad 0_{m,d}\right] d\beta d\widehat{W}_\nu,$$

where recall from Definition 6 that $\widehat{W}_t = \left[(W_t)^\top \quad (W_t^\star)^\top\right]^\top \in \mathbb{R}^{2d}$. Using the definition of $\mathcal{P}_{\mu_i}^r$ in Proposition C.3, and the regularity assumptions in Sec. 2.2, it is straightforward to show that $\mathcal{P}_{\mu_i}^r \in \mathcal{M}_2^{loc}\left(\mathbb{R}^m | \mathfrak{W}_t \times \mathfrak{W}_t^\star\right)$

and $\begin{bmatrix} F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) & 0_{m,d}\end{bmatrix} \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{m\times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^\star\right)$. Therefore, we may apply Lemma B.1 to the right hand side of the previous expression and obtain

$$N_2^r(t) = \sum_{i=1}^3 \int_0^t e^{(2\lambda - \omega)\nu} \left[\mathcal{P}_{\mu_i}^r(\nu)\right]^\top \left(\int_0^\nu e^{\omega\beta} \begin{bmatrix} F_\sigma^{\parallel}\left(\beta, X_{N,\beta}^r\right) & 0_{m,d}\end{bmatrix} d\widehat{W}_\beta\right) d\nu \in \mathbb{R}.$$

Note that $N_2^r(t)$ can be cast in the form of the process $N(t)$ in Lemma B.3 by setting

$$Q_t = \widehat{W}_t \in \mathbb{R}^{2d}\,(n_q = 2d), \quad \mathfrak{F}_t = \mathfrak{W}_t \times \mathfrak{W}_t^\star, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega,$$
$$S(t) = \sum_{i=1}^3 \mathcal{P}_{\mu_i}^r(t) \in \mathcal{M}_2^{loc}\left(\mathbb{R}^m | \mathfrak{W}_t \times \mathfrak{W}_t^\star\right), \quad L(t) = \begin{bmatrix} F_\sigma^{\parallel}\left(t, X_{N,t}^r\right) & 0_{m,d}\end{bmatrix} \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{m\times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^\star\right).$$

It then follows from Lemma B.3 that

$$\begin{aligned}\left\|\left(N_2^r\right)_{t^\star}\right\|_{\mathsf{p}}^{\pi_\star^0} &\leq \left(\mathsf{p}^3 \frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{e^{2\lambda t^\star}}{2\lambda\sqrt{\omega}} \sum_{i=1}^3 \left\|\left(\mathcal{P}_{\mu_i}^r\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \left\|\left(F_\sigma^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \\ &\leq \left(\mathsf{p}^3 \frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{e^{2\lambda t^\star}}{2\lambda\sqrt{\omega}} \sum_{i=1}^3 \left\|\left(\mathcal{P}_{\mu_i}^r\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \left\|\left(F_\sigma^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{4\mathsf{p}}^{\pi_\star^0}, \end{aligned} \tag{C.28}$$

where the last inequality is due to Jensen's inequality.

Next, we consider the process $N_3^r(t)$ given by

$$\begin{aligned} N_3^r(t) &= \int_0^t e^{\omega\nu} M_\sigma(t,\nu) \Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \\ &= \int_0^t e^{\omega\nu} \left(\int_\nu^t e^{(2\lambda-\omega)\beta} \left[\mathcal{P}_\sigma^r(\beta) dW_\beta + \mathcal{P}_{\sigma_\star}^r(\beta) dW_\beta^\star\right]^\top\right) \Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \\ &= \int_0^t \int_\nu^t e^{\omega\nu} \Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right)^\top e^{(2\lambda-\omega)\beta} \begin{bmatrix}\mathcal{P}_\sigma^r(\beta) & \mathcal{P}_{\sigma_\star}^r(\beta)\end{bmatrix} d\widehat{W}_\beta d\nu. \end{aligned}$$

Using Lemma B.1, one sees that

$$\begin{aligned} N_3^r(t) &= \int_0^t \int_0^\nu e^{\omega\beta} \Lambda_\mu^{\parallel}\left(\beta, X_{N,\beta}^r\right)^\top e^{(2\lambda-\omega)\nu} \begin{bmatrix}\mathcal{P}_\sigma^r(\nu) & \mathcal{P}_{\sigma_\star}^r(\nu)\end{bmatrix} d\beta d\widehat{W}_\nu \\ &= \int_0^t e^{(2\lambda-\omega)\nu} \left(\int_0^\nu e^{\omega\beta} \Lambda_\mu^{\parallel}\left(\beta, X_{N,\beta}^r\right) d\beta\right)^\top \begin{bmatrix}\mathcal{P}_\sigma^r(\nu) & \mathcal{P}_{\sigma_\star}^r(\nu)\end{bmatrix} d\widehat{W}_\nu \end{aligned}$$

Note that $N_3^r(t)$ can be cast in the form of the process $\widehat{N}(t)$ in Lemma B.4 by setting

$$Q_t = \widehat{W}_t \in \mathbb{R}^{2d}\,(n_q = 2d), \quad \mathfrak{F}_t = \mathfrak{W}_t \times \mathfrak{W}_t^\star, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega,$$
$$S(t) = \Lambda_\mu^{\parallel}\left(t, X_{N,t}^r\right) \in \mathcal{M}_2^{loc}\left(\mathbb{R}^m | \mathfrak{W}_t \times \mathfrak{W}_t^\star\right), \quad L(t) = \begin{bmatrix}\mathcal{P}_\sigma^r(t) & \mathcal{P}_{\sigma_\star}^r(t)\end{bmatrix} \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{m\times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^\star\right).$$

Therefore, we use Lemma B.4 and obtain

$$\begin{aligned}\left\|\left(N_3^r\right)_{t^\star}\right\|_{\mathsf{p}}^{\pi_\star^0} &\leq 2\sqrt{\mathsf{p}} \frac{e^{2\lambda t^\star}}{\sqrt{2\lambda}\omega} \left\|\left(\begin{bmatrix}\mathcal{P}_\sigma^r & \mathcal{P}_{\sigma_\star}^r\end{bmatrix}\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \left\|\left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \\ &\leq 2\sqrt{\mathsf{p}} \frac{e^{2\lambda t^\star}}{\sqrt{\lambda}\omega} \left(\left\|\left(\mathcal{P}_\sigma^r\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\|\left(\mathcal{P}_{\sigma_\star}^r\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}\right) \left\|\left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}, \end{aligned} \tag{C.29}$$

where the last inequality follows from the Minkowski's inequality subsequent to the following manipulations using the definition of the Frobenius norm that

$$\left\|\begin{bmatrix}\mathcal{P}_\sigma^r(t) & \mathcal{P}_{\sigma_\star}^r(t)\end{bmatrix}\right\|_F = \left(\left\|\mathcal{P}_\sigma^r(t)\right\|_F^2 + \left\|\mathcal{P}_{\sigma_\star}^r(t)\right\|_F^2\right)^{\frac{1}{2}} \leq \sqrt{2}\left(\left\|\mathcal{P}_\sigma^r(t)\right\|_F + \left\|\mathcal{P}_{\sigma_\star}^r(t)\right\|_F\right),$$

and thus

$$\left(\begin{bmatrix}\mathcal{P}_\sigma^r(t) & \mathcal{P}_{\sigma_\star}^r(t)\end{bmatrix}\right)_{t^\star} \doteq \sup_{t\in[0,t^\star]} \left\|\begin{bmatrix}\mathcal{P}_\sigma^r(t) & \mathcal{P}_{\sigma_\star}^r(t)\end{bmatrix}\right\|_F \leq \sup_{t\in[0,t^\star]} \sqrt{2}\left(\left\|\mathcal{P}_\sigma^r(t)\right\|_F + \left\|\mathcal{P}_{\sigma_\star}^r(t)\right\|_F\right)$$

$$\leq \sqrt{2}\left(\sup_{t\in[0,t^\star]}\|\mathcal{P}^r_\sigma(t)\|_F + \sup_{t\in[0,t^\star]}\left\|\mathcal{P}^r_{\sigma_\star}(t)\right\|_F\right). \tag{C.30}$$

Finally, consider the process $N^r_4(t)$ given as

$$\begin{aligned}
N^r_4(t) &= \int_0^t e^{\omega\nu} M_\sigma(t,\nu) F^{\parallel}_\sigma\left(\nu, X^r_{N,\nu}\right) dW_\nu \\
&= \int_0^t e^{\omega\nu}\int_\nu^t e^{(2\lambda-\omega)\beta}\left[\mathcal{P}^r_\sigma(\beta)dW_\beta + \mathcal{P}^r_{\sigma_\star}(\beta)dW^\star_\beta\right]^\top F^{\parallel}_\sigma\left(\nu, X^r_{N,\nu}\right) dW_\nu \\
&= \int_0^t e^{\omega\nu}\left(\int_\nu^t e^{(2\lambda-\omega)\beta}\left[\mathcal{P}^r_\sigma(\beta) \quad \mathcal{P}^r_{\sigma_\star}(\beta)\right] d\widehat{W}_\beta\right)^\top \left[F^{\parallel}_\sigma\left(\nu, X^r_{N,\nu}\right) \quad 0_{m,d}\right] d\widehat{W}_\nu,
\end{aligned}$$

which can be cast in the form of the process $\widetilde{N}(t)$ in Lemma B.6 by choosing

$$Q_t = \widehat{W}_t \in \mathbb{R}^{2d}\,(n_q = 2d), \quad \mathfrak{F}_t = \mathfrak{W}_t \times \mathfrak{W}^\star_t, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega,$$
$$L_1(t) = \left[\mathcal{P}^r_\sigma(t) \quad \mathcal{P}^r_{\sigma_\star}(t)\right] \in \mathcal{M}^{loc}_2\left(\mathbb{R}^{m\times 2d}|\mathfrak{W}_t \times \mathfrak{W}^\star_t\right),$$
$$L_2(t) = \left[F^{\parallel}_\sigma\left(t, X^r_{N,t}\right) \quad 0_{m,d}\right] \in \mathcal{M}^{loc}_2\left(\mathbb{R}^{m\times 2d}|\mathfrak{W}_t \times \mathfrak{W}^\star_t\right).$$

It then follows from Lemma B.6 that

$$\left\|\left(N^r_4\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}} \leq 2\mathsf{p}^{\frac{3}{2}}\left(4\mathsf{p}-1\right)^{\frac{1}{2}}\frac{e^{2\lambda t^\star}}{\sqrt{2\lambda\omega}}\left\|\left(\left[\mathcal{P}^r_\sigma \quad \mathcal{P}^r_{\sigma_\star}\right]\right)_{t^\star}\right\|^{\pi^0_\star}_{4\mathsf{p}}\left\|\left(F^{\parallel}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{4\mathsf{p}} + \frac{\sqrt{m}e^{2\lambda t^\star}}{2\lambda}\left\|\left(\mathcal{P}^r_\sigma F^{\parallel}_\sigma\left(\cdot, X^r_N\right)^\top\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}}.$$

Using (C.30) on the first term on the right hand side produces

$$\left\|\left(N^r_4\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}} \leq 2\mathsf{p}^{\frac{3}{2}}\left(4\mathsf{p}-1\right)^{\frac{1}{2}}\frac{e^{2\lambda t^\star}}{\sqrt{\lambda\omega}}\left\|\left(\mathcal{P}^r_\sigma\right)_{t^\star} + \left(\mathcal{P}^r_{\sigma_\star}\right)_{t^\star}\right\|^{\pi^0_\star}_{4\mathsf{p}}\left\|\left(F^{\parallel}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{4\mathsf{p}} + \frac{\sqrt{m}e^{2\lambda t^\star}}{2\lambda}\left\|\left(\mathcal{P}^r_\sigma F^{\parallel}_\sigma\left(\cdot, X^r_N\right)^\top\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}}.$$

Using the Minkowski's inequality then leads to

$$\left\|\left(N^r_4\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}} \leq 2\mathsf{p}^{\frac{3}{2}}\left(4\mathsf{p}-1\right)^{\frac{1}{2}}\frac{e^{2\lambda t^\star}}{\sqrt{\lambda\omega}}\left(\left\|\left(\mathcal{P}^r_\sigma\right)_{t^\star}\right\|^{\pi^0_\star}_{4\mathsf{p}} + \left\|\left(\mathcal{P}^r_{\sigma_\star}\right)_{t^\star}\right\|^{\pi^0_\star}_{4\mathsf{p}}\right)\left\|\left(F^{\parallel}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{4\mathsf{p}} + \frac{\sqrt{m}e^{2\lambda t^\star}}{2\lambda}\left\|\left(\mathcal{P}^r_\sigma F^{\parallel}_\sigma\left(\cdot, X^r_N\right)^\top\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}}. \tag{C.31}$$

The bound in (C.26) is then established by adding together (C.27), (C.28), (C.29), and (C.31)

□

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:*

$$N^r_{\mathcal{U}}(t) = \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t,\nu)\left(\Lambda^{\parallel}_\mu\left(\nu, X^r_{N,\nu}\right)d\nu + F^{\parallel}_\sigma\left(\nu, X^r_{N,\nu}\right)dW_\nu\right), \tag{C.32}$$

where

$$M_{\mathcal{U}}(t,\nu) = \int_\nu^t e^{(2\lambda-\omega)\beta}\mathcal{P}^r_{\mathcal{U}}(\beta)^\top d\beta,$$

for $0 \leq \nu \leq t \leq T$, and where where $\mathcal{P}^r_{\mathcal{U}}(\beta) \in \mathbb{R}^m$ is defined in the statement of Proposition C.3.

If the stopping time $\tau^\star$, defined in (29), Lemma 3.1, satisfies $\tau^\star = t^\star$, then we have the following bound for all $\mathsf{p} \in \mathbb{N}_{\geq 1}$:

$$\left\|\left(N^r_{\mathcal{U}}\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}} \leq \mathsf{p}\sqrt{\frac{n}{2}}\Delta^2_g\Delta_{\partial V}\frac{e^{2\lambda t^\star}}{2\lambda}\left(\frac{1}{\sqrt{\omega}}\left\|\left(\Lambda^{\parallel}_\mu\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}} + \left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}}\left\|\left(F^{\parallel}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}}\right)^2. \tag{C.33}$$

Proof. We begin by decomposing the process $N^r_{\mathcal{U}_1}(t)$ as follows:

$$N^r_{\mathcal{U}}(t) = N^r_{\mathcal{U}_1}(t) + N^r_{\mathcal{U}_2}(t), \tag{C.34}$$

where

$$N^r_{\mathcal{U}_1}(t) = \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t,\nu)\Lambda^{\parallel}_{\mu}\left(\nu, X^r_{N,\nu}\right) d\nu, \quad N^r_{\mathcal{U}_2}(t) = \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t,\nu) F^{\parallel}_{\sigma}\left(\nu, X^r_{N,\nu}\right) dW_\nu$$

Now, consider the process $N^r_{\mathcal{U}_1}(t)$, which, using the definition of $M_{\mathcal{U}}$ can be written as

$$\begin{aligned}
N^r_{\mathcal{U}_1}(t) &= \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t,\nu)\Lambda^{\parallel}_{\mu}\left(\nu, X^r_{N,\nu}\right) d\nu \\
&= \int_0^t e^{\omega\nu} \int_\nu^t e^{(2\lambda-\omega)\beta} \mathcal{P}^r_{\mathcal{U}}(\beta)^\top d\beta \Lambda^{\parallel}_{\mu}\left(\nu, X^r_{N,\nu}\right) d\nu \\
&= \int_0^t \int_\nu^t e^{\omega\nu} e^{(2\lambda-\omega)\beta} \mathcal{P}^r_{\mathcal{U}}(\beta)^\top \Lambda^{\parallel}_{\mu}\left(\nu, X^r_{N,\nu}\right) d\beta d\nu.
\end{aligned}$$

Since the integral above is a nested Lebesgue integral with t-continuous integrands, we may change the order of integration as follows:

$$\begin{aligned}
N^r_{\mathcal{U}_1}(t) &= \int_0^t \int_0^\beta e^{\omega\nu} e^{(2\lambda-\omega)\beta} \mathcal{P}^r_{\mathcal{U}}(\beta)^\top \Lambda^{\parallel}_{\mu}\left(\nu, X^r_{N,\nu}\right) d\nu d\beta \\
&= \int_0^t e^{(2\lambda-\omega)\beta} \mathcal{P}^r_{\mathcal{U}}(\beta)^\top \left(\int_0^\beta e^{\omega\nu} \Lambda^{\parallel}_{\mu}\left(\nu, X^r_{N,\nu}\right) d\nu\right) d\beta \\
&= \int_0^t e^{(2\lambda-\omega)\nu} \mathcal{P}^r_{\mathcal{U}}(\nu)^\top \left(\int_0^\nu e^{\omega\beta} \Lambda^{\parallel}_{\mu}\left(\beta, X^r_{N,\beta}\right) d\beta\right) d\nu,
\end{aligned} \tag{C.35}$$

where in the last integral we have switched between the variables β and ν.
Next, consider the process $N^r_{\mathcal{U}_2}(t)$, which, using the definition of $M_{\mathcal{U}}$ can be written as

$$\begin{aligned}
N^r_{\mathcal{U}_2}(t) &= \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t,\nu) F^{\parallel}_{\sigma}\left(\nu, X^r_{N,\nu}\right) dW_\nu \\
&= \int_0^t \int_\nu^t \left(e^{(2\lambda-\omega)\beta} \mathcal{P}^r_{\mathcal{U}}(\beta)^\top\right) \left(e^{\omega\nu} F^{\parallel}_{\sigma}\left(\nu, X^r_{N,\nu}\right)\right) d\beta dW_\nu \\
&= \int_0^t \int_\nu^t \left(e^{(2\lambda-\omega)\beta} \mathcal{P}^r_{\mathcal{U}}(\beta)^\top\right) \left(e^{\omega\nu} \begin{bmatrix} F^{\parallel}_{\sigma}\left(\nu, X^r_{N,\nu}\right) & 0_{m,d}\end{bmatrix}\right) d\beta d\widehat{W}_\nu,
\end{aligned}$$

where once again recall from Definition 6 that $\widehat{W}_t = \begin{bmatrix}(W_t)^\top & (W^\star_t)^\top\end{bmatrix}^\top \in \mathbb{R}^{2d}$. Using the definition of $P^r_{\mathcal{U}}$ in Proposition C.3, and the regularity assumptions in Sec. 2.2, it is straightforward to show that $P^r_{\mathcal{U}} \in \mathcal{M}^{loc}_2\left(\mathbb{R}^m | \mathfrak{W}_t \times \mathfrak{W}^\star_t\right)$ and $\begin{bmatrix} F^{\parallel}_{\sigma}\left(\nu, X^r_{N,\nu}\right) & 0_{m,d}\end{bmatrix} \in \mathcal{M}^{loc}_2\left(\mathbb{R}^{m\times 2d} | \mathfrak{W}_t \times \mathfrak{W}^\star_t\right)$. Therefore, we invoke Lemma B.1 to obtain

$$\begin{aligned}
N^r_{\mathcal{U}_2}(t) &= \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t,\nu) F^{\parallel}_{\sigma}\left(\nu, X^r_{N,\nu}\right) dW_\nu \\
&= \int_0^t \left(e^{(2\lambda-\omega)\nu} \mathcal{P}^r_{\mathcal{U}}(\nu)^\top\right) \int_0^\nu \left(e^{\omega\beta} \begin{bmatrix} F^{\parallel}_{\sigma}\left(\beta, X^r_{N,\beta}\right) & 0_{m,d}\end{bmatrix}\right) d\widehat{W}_\beta d\nu \\
&= \int_0^t e^{(2\lambda-\omega)\nu} \mathcal{P}^r_{\mathcal{U}}(\nu)^\top \left(\int_0^\nu e^{\omega\beta} F^{\parallel}_{\sigma}\left(\beta, X^r_{N,\beta}\right) dW_\beta\right) d\nu.
\end{aligned} \tag{C.36}$$

Substituting (C.35) and (C.36) into (C.34) yields

$$N^r_{\mathcal{U}}(t) = \int_0^t e^{(2\lambda-\omega)\nu} \mathcal{P}^r_{\mathcal{U}}(\nu)^\top \left(\int_0^\nu e^{\omega\beta} \left[\Lambda^{\parallel}_{\mu}\left(\beta, X^r_{N,\beta}\right) d\beta + F^{\parallel}_{\sigma}\left(\beta, X^r_{N,\beta}\right) dW_\beta\right]\right) d\nu. \tag{C.37}$$

We now obtain the following expression by using the definition of $\mathcal{P}^r_{\mathcal{U}}$ in Proposition C.3, followed by the definition of U^r in (23) for the truncated process (27):

$$\begin{aligned}\mathcal{P}^r_{\mathcal{U}}(t) =& g(t)^\top V_{r,r}\left(Y_{N,t}\right) g(t) U^r_t \\ =& g(t)^\top V_{r,r}\left(Y_{N,t}\right) g(t) \left(\mathcal{F}_{\boldsymbol{\omega}} \Lambda^{\parallel}_{\mu}\left(\cdot, X^r_N\right) t + \mathcal{F}_{\mathcal{N},\boldsymbol{\omega}} F^{\parallel}_{\sigma}\left(\cdot, X^r_N\right), Wt\right) \\ =& g(t)^\top V_{r,r}\left(Y_{N,t}\right) g(t) \left(-\boldsymbol{\omega} \int_0^t e^{-\boldsymbol{\omega}(t-\nu)} \Lambda^{\parallel}_{\mu}\left(\nu, X^r_{N,\nu}\right) d\nu - \boldsymbol{\omega} \int_0^t e^{-\boldsymbol{\omega}(t-\nu)} F^{\parallel}_{\sigma}\left(\nu, X^r_{N,\nu}\right) dW_\nu\right) \\ =& -\boldsymbol{\omega} e^{-\boldsymbol{\omega} t} g(t)^\top V_{r,r}\left(Y_{N,t}\right) g(t) \left(\int_0^t e^{\boldsymbol{\omega}\nu} \left[\Lambda^{\parallel}_{\mu}\left(\nu, X^r_{N,\nu}\right) d\nu + F^{\parallel}_{\sigma}\left(\nu, X^r_{N,\nu}\right) dW_\nu\right]\right).\end{aligned}$$

Substituting the above expression into (C.37) produces

$$\begin{aligned}N^r_{\mathcal{U}}(t) = -\boldsymbol{\omega} \int_0^t e^{2(\lambda-\boldsymbol{\omega})\nu} &\left(\int_0^\nu e^{\boldsymbol{\omega}\beta} \left[\Lambda^{\parallel}_{\mu}\left(\beta, X^r_{N,\beta}\right) d\beta + F^{\parallel}_{\sigma}\left(\beta, X^r_{N,\beta}\right) dW_\beta\right]\right)^\top g(\nu)^\top V_{r,r}\left(Y_{N,\nu}\right) g(\nu) \\ &\times \left(\int_0^\nu e^{\boldsymbol{\omega}\beta} \left[\Lambda^{\parallel}_{\mu}\left(\beta, X^r_{N,\beta}\right) d\beta + F^{\parallel}_{\sigma}\left(\beta, X^r_{N,\beta}\right) dW_\beta\right]\right) d\nu. \quad \text{(C.38)}\end{aligned}$$

Observe that $N^r_{\mathcal{U}}(t)$ can be expressed as the process $\breve{N}(t)$ in the statement of Corollary B.1 by setting

$$\begin{aligned}Q_t = \widehat{W}_t \in \mathbb{R}^{2d}\,(n_q = 2d), \quad \mathfrak{F}_t = \mathfrak{W}_t \times \mathfrak{W}^\star_t, \quad \theta_1 = \lambda, \quad \theta_2 = \boldsymbol{\omega}, \\ R(t) = -\boldsymbol{\omega} g(t)^\top V_{r,r}\left(Y_{N,t}\right) g(t) \in \mathcal{M}^{loc}_2\left(\mathbb{S}^m | \mathfrak{W}_t \times \mathfrak{W}^\star_t\right), \quad S(t) = \Lambda^{\parallel}_{\mu}\left(t, X^r_{N,t}\right) \in \mathcal{M}^{loc}_2\left(\mathbb{R}^m | \mathfrak{W}_t \times \mathfrak{W}^\star_t\right), \\ L(t) = \begin{bmatrix} F^{\parallel}_{\sigma}\left(t, X^r_{N,t}\right) & 0_{m,d}\end{bmatrix} \in \mathcal{M}^{loc}_2\left(\mathbb{R}^{m\times 2d} | \mathfrak{W}_t \times \mathfrak{W}^\star_t\right).\end{aligned}$$

Furthermore, as a consequence of Assumption 1 and (E.1b) in Proposition E.1, we may set Δ_R in the hypothesis of Corollary B.1 as

$$\Delta_R = \sqrt{\frac{n}{2}} \Delta_g^2 \Delta_{\partial V} \boldsymbol{\omega}.$$

Hence, the proof is concluded by applying Corollary B.1 to the process $N^r_{\mathcal{U}}(t)$ in (C.38), thereby producing the desired result in (C.33).

□

The next lemma establishes the bound on Ξ^r.

Lemma C.1 *If the stopping time $\tau^\star$, defined in (29), Lemma 3.1, satisfies $\tau^\star = t^\star$, then the term $\Xi^r\left(\tau(t), Y_N\right)$ defined in (31), Lemma 3.1, satisfies the following bound for all $\mathsf{p} \in \mathbb{N}_{\geq 1}$:*

$$\left\|\left(\Xi^r\left(\cdot, Y_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}} \leq e^{2\lambda t^\star}\left(\frac{\Delta^r_{\Xi_1}}{\lambda} + \frac{\Delta^r_{\Xi_2}}{\sqrt{\lambda}}\right), \quad \text{(C.39)}$$

where

$$\begin{aligned}\Delta^r_{\Xi_1} &= \frac{\Delta_g}{2}\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}} \left\|\left(\Lambda^{\perp}_{\mu}\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}} + \frac{1}{4}\left\|\left(Tr\left[H_\sigma\left(\cdot, Y_N\right)\nabla^2 V\left(Y_N\right)\right]\right)_{t^\star}\right\|^{\pi^0_\star}_{\mathsf{p}}, \\ \Delta^r_{\Xi_2} &= 2\Delta^{\perp}_g \sqrt{\mathsf{p}}\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}} \left\|\left(F^{\perp}_{\sigma}\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}} + 2\sqrt{\mathsf{p}}\left\|\left(V_\star\left(Y_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}} \left\|\left(\bar{F}_\sigma\left(\cdot, X^\star_N\right)\right)_{t^\star}\right\|^{\pi^0_\star}_{2\mathsf{p}}.\end{aligned}$$

Proof. We begin by writing

$$\Xi^r\left(\tau(t), Y_N\right) = \int_0^{\tau(t)} e^{2\lambda\nu} \phi^r_{\mu}\left(\nu, Y_{N,\nu}\right) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu}\left[\phi^r_{\sigma}\left(\nu, Y_{N,\nu}\right) \quad \phi^r_{\sigma_\star}\left(\nu, Y_{N,\nu}\right)\right] d\widehat{W}_\nu,$$

and thus

$$\left(\Xi^r\left(\cdot, Y_N\right)\right)_{t^\star} \leq \sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}\phi^r_{\mu}\left(\nu, Y_{N,\nu}\right) d\nu\right| + \sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}\left[\phi^r_{\sigma}\left(\nu, Y_{N,\nu}\right) \quad \phi^r_{\sigma_\star}\left(\nu, Y_{N,\nu}\right)\right] d\widehat{W}_\nu\right|.$$

Using the Minkowski's inequality one sees that

$$\left\|\left(\Xi^r\left(\cdot, Y_N\right)\right)_{t^\star}\right\|_{\mathsf{p}}^{\pi_\star^0} \leq \left\|\sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}\phi_\mu^r\left(\nu, Y_{N,\nu}\right)d\nu\right|\right\|_{\mathsf{p}}^{\pi_\star^0} + \left\|\sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}\left[\phi_\sigma^r\left(\nu, Y_{N,\nu}\right) \quad \phi_{\sigma_\star}^r\left(\nu, Y_{N,\nu}\right)\right]d\widehat{W}_\nu\right|\right\|_{\mathsf{p}}^{\pi_\star^0}. \tag{C.40}$$

Using the definition of ϕ_μ^r in (32), we obtain

$$\int_0^t e^{2\lambda\nu}\phi_\mu^r\left(\nu, Y_{N,\nu}\right)d\nu = \int_0^t e^{2\lambda\nu}\left(V_r\left(Y_{N,\nu}\right)^\top g(\nu)^\perp\Lambda_\mu^\perp\left(\nu, X_{N,t}^r\right) + \frac{1}{2}\mathrm{Tr}\left[H_\sigma\left(\nu, Y_{N,\nu}\right)\nabla^2 V\left(Y_{N,\nu}\right)\right]\right)d\nu,$$

which implies that

$$\begin{aligned}
\int_0^t e^{2\lambda\nu}\phi_\mu^r\left(\nu, Y_{N,\nu}\right)d\nu
&\leq \int_0^t e^{2\lambda\nu}\left(\left\|V_r\left(Y_{N,\nu}\right)\right\|\left\|g(\nu)\right\|_F\left\|\Lambda_\mu^\perp\left(\nu, X_{N,t}^r\right)\right\| + \frac{1}{2}\left|\mathrm{Tr}\left[H_\sigma\left(\nu, Y_{N,\nu}\right)\nabla^2 V\left(Y_{N,\nu}\right)\right]\right|\right)d\nu \\
&\leq \int_0^t e^{2\lambda\nu}\left(\left\|V_r\left(Y_{N,\nu}\right)\right\|\Delta_g\left\|\Lambda_\mu^\perp\left(\nu, X_{N,t}^r\right)\right\| + \frac{1}{2}\left|\mathrm{Tr}\left[H_\sigma\left(\nu, Y_{N,\nu}\right)\nabla^2 V\left(Y_{N,\nu}\right)\right]\right|\right)d\nu,
\end{aligned}$$

where we have used the bound on $g(t)$ from Assumption 1. We develop this bound further as follows

$$\begin{aligned}
\int_0^t e^{2\lambda\nu}\phi_\mu^r\left(\nu, Y_{N,\nu}\right)d\nu
&\leq \left(\int_0^t e^{2\lambda\nu}d\nu\right)\left(\Delta_g\left(V_r\left(Y_N\right)\right)_t\left(\Lambda_\mu^\perp\left(\cdot, X_N^r\right)\right)_t + \frac{1}{2}\left(\mathrm{Tr}\left[H_\sigma\left(\cdot, Y_N\right)\nabla^2 V\left(Y_N\right)\right]\right)_t\right) \\
&\leq \frac{e^{2\lambda t}}{2\lambda}\left(\Delta_g\left(V_r\left(Y_N\right)\right)_t\left(\Lambda_\mu^\perp\left(\cdot, X_N^r\right)\right)_t + \frac{1}{2}\left(\mathrm{Tr}\left[H_\sigma\left(\cdot, Y_N\right)\nabla^2 V\left(Y_N\right)\right]\right)_t\right),
\end{aligned}$$

where in the last inequality we have used the strict positivity of $\lambda \in \mathbb{R}_{>0}$. Hence,

$$\sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}\phi_\mu^r\left(\nu, Y_{N,\nu}\right)d\nu\right| \leq \frac{e^{2\lambda t^\star}}{2\lambda}\left(\Delta_g\left(V_r\left(Y_N\right)\right)_{t^\star}\left(\Lambda_\mu^\perp\left(\cdot, X_N^r\right)\right)_{t^\star} + \frac{1}{2}\left(\mathrm{Tr}\left[H_\sigma\left(\cdot, Y_N\right)\nabla^2 V\left(Y_N\right)\right]\right)_{t^\star}\right).$$

Since $t^\star$ is a constant, it then follows that

$$\begin{aligned}
&\left\|\sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}\phi_\mu^r\left(\nu, Y_{N,\nu}\right)d\nu\right|\right\|_{\mathsf{p}}^{\pi_\star^0} \\
&\overset{(i)}{\leq} \frac{e^{2\lambda t^\star}}{2\lambda}\left(\Delta_g\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\left(\Lambda_\mu^\perp\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{\mathsf{p}}^{\pi_\star^0} + \frac{1}{2}\left\|\left(\mathrm{Tr}\left[H_\sigma\left(\cdot, Y_N\right)\nabla^2 V\left(Y_N\right)\right]\right)_{t^\star}\right\|_{\mathsf{p}}^{\pi_\star^0}\right) \\
&\overset{(ii)}{\leq} \frac{e^{2\lambda t^\star}}{2\lambda}\left(\Delta_g\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}\left\|\left(\Lambda_\mu^\perp\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} + \frac{1}{2}\left\|\left(\mathrm{Tr}\left[H_\sigma\left(\cdot, Y_N\right)\nabla^2 V\left(Y_N\right)\right]\right)_{t^\star}\right\|_{\mathsf{p}}^{\pi_\star^0}\right),
\end{aligned} \tag{C.41}$$

where (i) and (ii) are due to the Minkowski's and the Cauchy-Schwarz inequalities, respectively.

Next, using the definitions of $\phi_{\sigma_\star}^r$ and ϕ_σ^r in (32), we obtain

$$\int_0^t e^{2\lambda\nu}\left[\phi_\sigma^r\left(\nu, Y_{N,\nu}\right) \quad \phi_{\sigma_\star}^r\left(\nu, Y_{N,\nu}\right)\right]d\widehat{W}_\nu$$

$$= \int_0^t e^{2\lambda\nu} \left[V_r\left(Y_\nu\right)^\top g(\nu)^\perp F_\sigma^\perp\left(\nu, X_{N,\nu}^r\right) \quad V_\star\left(Y_\nu\right)^\top \bar{F}_\sigma\left(\nu, X_\nu^\star\right)\right] d\widehat{W}_\nu.$$

The regularity assumptions and the t-continuity of the strong solutions $X_{N,t}^r$ and $X_{N,t}^\star$ imply that $\left[\phi_\sigma^r\left(\nu, Y_{N,\nu}\right) \quad \phi_{\sigma_\star}^r\left(\nu, Y_{N,\nu}\right)\right] \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{1\times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^\star\right)$. Thus, we may use Proposition B.1 and the strict positivity of $\lambda in \mathbb{R}_{>0}$ to obtain

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t e^{2\lambda\nu} \left[\phi_\sigma^r\left(\nu, Y_{N,\nu}\right) \quad \phi_{\sigma_\star}^r\left(\nu, Y_{N,\nu}\right)\right] d\widehat{W}_\nu \right| \right\|_{\mathsf{p}}^{\pi_\star^0}$$
$$\leq \frac{2\sqrt{\mathsf{p}} e^{2\lambda t^\star}}{\sqrt{2\lambda}} \left\| \left(\left[V_r\left(Y_N\right)^\top g^\perp F_\sigma^\perp\left(\cdot, X_N^r\right) \quad V_\star\left(Y_N\right)^\top \bar{F}_\sigma\left(\cdot, X_N^\star\right)\right]\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0}. \quad \text{(C.42)}$$

Next, observe that

$$\left\| \left[V_r\left(Y_{N,t}\right)^\top g(t)^\perp F_\sigma^\perp\left(t, X_{N,t}^r\right) \quad V_\star\left(Y_{N,t}\right)^\top \bar{F}_\sigma\left(t, X_t^\star\right)\right]\right\|$$
$$= \left(\left\| V_r\left(Y_{N,t}\right)^\top g(t)^\perp F_\sigma^\perp\left(t, X_{N,t}^r\right)\right\|^2 + \left\| V_\star\left(Y_{N,t}\right)^\top \bar{F}_\sigma\left(t, X_t^\star\right)\right\|^2 \right)^{\frac{1}{2}}$$
$$\leq \sqrt{2}\left(\left\| V_r\left(Y_{N,t}\right)^\top g(t)^\perp F_\sigma^\perp\left(t, X_{N,t}^r\right)\right\| + \left\| V_\star\left(Y_{N,t}\right)^\top \bar{F}_\sigma\left(t, X_t^\star\right)\right\| \right),$$

and thus

$$\left(\left[V_r\left(Y_N\right)^\top g^\perp F_\sigma^\perp\left(\cdot, X_N^r\right) \quad V_\star\left(Y_N\right)^\top \bar{F}_\sigma\left(\cdot, X_N^\star\right)\right]\right)_{t^\star}$$
$$\doteq \sup_{t\in[0,t^\star]} \left\| \left[V_r\left(Y_{N,t}\right)^\top g(t)^\perp F_\sigma^\perp\left(t, X_{N,t}^r\right) \quad V_\star\left(Y_{N,t}\right)^\top \bar{F}_\sigma\left(t, X_t^\star\right)\right]\right\|$$
$$\leq \sup_{t\in[0,t^\star]} \sqrt{2}\left(\left\| V_r\left(Y_{N,t}\right)^\top g(t)^\perp F_\sigma^\perp\left(t, X_{N,t}^r\right)\right\| + \left\| V_\star\left(Y_{N,t}\right)^\top \bar{F}_\sigma\left(t, X_t^\star\right)\right\| \right).$$

Hence,

$$\left(\left[V_r\left(Y_N\right)^\top g^\perp F_\sigma^\perp\left(\cdot, X_N^r\right) \quad V_\star\left(Y_N\right)^\top \bar{F}_\sigma\left(\cdot, X_N^\star\right)\right]\right)_{t^\star}$$
$$\leq \sqrt{2}\left(\sup_{t\in[0,t^\star]} \left\| V_r\left(Y_{N,t}\right)^\top g(t)^\perp F_\sigma^\perp\left(t, X_{N,t}^r\right)\right\| + \sup_{t\in[0,t^\star]} \left\| V_\star\left(Y_{N,t}\right)^\top \bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\| \right)$$
$$\leq \sqrt{2}\left(\Delta_g^\perp \sup_{t\in[0,t^\star]} \left\| V_r\left(Y_{N,t}\right)\right\| \sup_{t\in[0,t^\star]} \left\| F_\sigma^\perp\left(t, X_{N,t}^r\right)\right\| + \sup_{t\in[0,t^\star]} \left\| V_\star\left(Y_{N,t}\right)\right\| \sup_{t\in[0,t^\star]} \left\| \bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\| \right),$$

where we have used the bound on $g(t)^\perp$ from Assumption 1. It then follows from the definition of $\left(\cdot\right)_{t^\star}$ that

$$\left(\left[V_r\left(Y_N\right)^\top g^\perp F_\sigma^\perp\left(\cdot, X_N^r\right) \quad V_\star\left(Y_N\right)^\top \bar{F}_\sigma\left(\cdot, X_N^\star\right)\right]\right)_{t^\star}$$
$$\leq \sqrt{2}\left(\Delta_g^\perp \left(V_r\left(Y_N\right)\right)_{t^\star} \left(F_\sigma^\perp\left(\cdot, X_N^r\right)\right)_{t^\star} + \left(V_\star\left(Y_N\right)\right)_{t^\star} \left(\bar{F}_\sigma\left(\cdot, X_N^\star\right)\right)_{t^\star} \right).$$

Using the Minkowski's inequality, one sees that

$$\left\| \left(\left[V_r\left(Y_N\right)^\top g^\perp F_\sigma^\perp\left(\cdot, X_N^r\right) \quad V_\star\left(Y_N\right)^\top \bar{F}_\sigma\left(\cdot, X_N^\star\right)\right]\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0}$$
$$\leq \sqrt{2}\left(\Delta_g^\perp \left\| \left(V_r\left(Y_N\right)\right)_{t^\star} \left(F_\sigma^\perp\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} + \left\| \left(V_\star\left(Y_N\right)\right)_{t^\star} \left(\bar{F}_\sigma\left(\cdot, X_N^\star\right)\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} \right)$$
$$\leq \sqrt{2}\left(\Delta_g^\perp \left\| \left(V_r\left(Y_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \left\| \left(F_\sigma^\perp\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(V_\star\left(Y_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \left\| \left(\bar{F}_\sigma\left(\cdot, X_N^\star\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \right),$$

where the last inequality is due to the Cauchy-Schwarz inequality. Substituting the above into (C.42) yields

$$
\begin{aligned}
&\left\| \sup_{t \in [0,t^\star]} \left| \int_0^t e^{2\lambda\nu} \left[\phi^r_\sigma \left(\nu, Y_{N,\nu}\right) \quad \phi^r_{\sigma_\star} \left(\nu, Y_{N,\nu}\right) \right] d\widehat{W}_\nu \right| \right\|_{\mathsf{p}}^{\pi^0_\star} \\
&\leq \frac{2\sqrt{\mathsf{p}} e^{2\lambda t^\star}}{\sqrt{\lambda}} \left(\Delta_g^\perp \left\| \left(V_r\left(Y_N\right) \right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \left\| \left(F_\sigma^\perp \left(\cdot, X_N^r\right) \right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} + \left\| \left(V_\star\left(Y_N\right) \right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \left\| \left(\bar{F}_\sigma \left(\cdot, X_N^\star\right) \right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \right). \qquad \text{(C.43)}
\end{aligned}
$$

Substituting (C.41) and (C.43) into (C.40) and re-arranging terms leads to (C.39), thus concluding the proof.

□

We next derive the bound on $\Xi^r_{\mathcal{U}}$ in the following lemma.

Lemma C.2 *If the stopping time $\tau^\star$, defined in (29), Lemma 3.1, satisfies $\tau^\star = t^\star$, then the term $\Xi^r_{\mathcal{U}}\left(\tau(t), Y_N; \omega\right)$ defined in (31), Lemma 3.1, satisfies the following bound for all $\mathsf{p} \in \mathbb{N}_{\geq 1}$:*

$$
\left\| \left(\Xi^r_{\mathcal{U}} \left(\cdot, Y_N; \omega\right) \right)_{t^\star} \right\|_{\mathsf{p}}^{\pi^0_\star} \leq \frac{e^{(2\lambda + \omega)t^\star}}{|2\lambda - \omega|} \left(\Delta^r_{\mathcal{U}_1} + \sqrt{\omega} \Delta^r_{\mathcal{U}_2} + \omega \Delta^r_{\mathcal{U}_3} \right), \qquad \text{(C.44)}
$$

where

$$
\begin{aligned}
\Delta^r_{\mathcal{U}_1} = &\left(\frac{\Delta^r_{\mathcal{P}_1}(t^\star)}{\sqrt{\lambda}} + 2\Delta_g \left(1 + \sqrt{2\lambda\mathsf{p}}\right) \left\| \left(V_r\left(Y_N\right) \right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \right) \left\| \left(\Lambda^{\parallel}_\mu \left(\cdot, X_N^r\right) \right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \\
&+ 2\Delta_g \sqrt{2\lambda\mathsf{p}} \left\| \left(V_r\left(Y_N\right) \right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \left\| \left(F^{\parallel}_\sigma \left(\cdot, X_N^r\right) \right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \\
&+ \sqrt{\frac{n}{2} \frac{\mathsf{p}\Delta_g^2 \Delta_{\partial V}}{\lambda}} \mathbb{E}_{\pi^0_\star} \left[\left(\Lambda^{\parallel}_\mu \left(\cdot, X_N^r\right) \right)_{t^\star}^{2\mathsf{p}} \right]^{\frac{1}{\mathsf{p}}}, \\
\Delta^r_{\mathcal{U}_2} = &\left(\frac{\Delta^r_{\mathcal{P}_2}(t^\star)}{\sqrt{\lambda}} + 2\Delta_g m \sqrt{2\mathsf{p}} \left\| \left(V_r\left(Y_N\right) \right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \right) \left\| \left(F^{\parallel}_\sigma \left(\cdot, X_N^r\right) \right)_{t^\star} \right\|_{4\mathsf{p}}^{\pi^0_\star}, \\
\Delta^r_{\mathcal{U}_3} = &\frac{\sqrt{m}}{2\lambda} \left\| \left(\mathcal{P}^r_\sigma F^{\parallel}_\sigma \left(\cdot, X_N^r\right)^\top \right)_{t^\star} \right\|_{\mathsf{p}}^{\pi^0_\star} + \sqrt{\frac{n}{2} \frac{\mathsf{p}^2(2\mathsf{p}-1)\Delta_g^2 \Delta_{\partial V}}{2\lambda}} \mathbb{E}_{\pi^0_\star} \left[\left(F^{\parallel}_\sigma \left(\cdot, X_N^r\right) \right)_{t^\star}^{2\mathsf{p}} \right]^{\frac{1}{\mathsf{p}}},
\end{aligned}
$$

and where the function $\mathcal{P}^r_\sigma(t)$, and the constants $\Delta^r_{\mathcal{P}_1}(t^\star)$ and $\Delta^r_{\mathcal{P}_2}(t^\star)$, are defined in the statements of Propositions C.3 and C.5, respectively.

Proof. We begin by recalling the definition of $\psi^r(\tau(t), \nu, Y_N)$ from (33)

$$
\begin{aligned}
\psi^r(\tau(t), \nu, Y_N) = &\frac{\omega}{2\lambda - \omega} \left(e^{\omega(\tau(t)+\nu)} \mathcal{P}^r\left(\tau(t), \nu\right) - e^{(2\lambda\tau(t) + \omega\nu)} V_r\left(Y_{N,\tau(t)}\right)^\top g(\tau(t)) \right) \\
&+ \frac{2\lambda}{2\lambda - \omega} e^{(\omega\tau(t) + 2\lambda\nu)} V_r\left(Y_{N,\nu}\right)^\top g(\nu),
\end{aligned}
$$

which, upon using the decomposition (C.14) in Proposition C.3, can be re-written as

$$
\begin{aligned}
\psi^r(\tau(t), \nu, Y_N) = &\frac{\omega}{2\lambda - \omega} \left(e^{\omega(\tau(t)+\nu)} \mathcal{P}^r_\circ\left(\tau(t), \nu\right) - e^{(2\lambda\tau(t) + \omega\nu)} V_r\left(Y_{N,\tau(t)}\right)^\top g(\tau(t)) \right) \\
&+ \frac{\omega}{2\lambda - \omega} e^{\omega(\tau(t)+\nu)} \mathcal{P}^r_{ad}\left(\tau(t), \nu\right) + \frac{2\lambda}{2\lambda - \omega} e^{(\omega\tau(t) + 2\lambda\nu)} V_r\left(Y_{N,\nu}\right)^\top g(\nu),
\end{aligned}
$$

We thus decompose $\psi^r(\tau(t), \nu, Y_N)$ in (33) as follows:

$$
\psi^r(\tau(t), \nu, Y_N) = \sum_{i=1}^{3} \psi^r_i(\tau(t), \nu, Y_N) + \psi^r_{ad}(\tau(t), \nu, Y_N) \in \mathbb{R}^{1 \times m}, \qquad \text{(C.45)}
$$

where

$$
\begin{aligned}
\psi^r_1(\tau(t), \nu, Y_N) &= \frac{\omega}{2\lambda - \omega} e^{\omega(\tau(t)+\nu)} \mathcal{P}^r_\circ\left(\tau(t), \nu\right), \\
\psi^r_2(\tau(t), \nu, Y_N) &= -\frac{\omega}{2\lambda - \omega} e^{(2\lambda\tau(t) + \omega\nu)} V_r\left(Y_{N,\tau(t)}\right)^\top g(\tau(t)),
\end{aligned}
$$

$$\psi_3^r(\tau(t),\nu,Y_N) = \frac{2\lambda}{2\lambda-\omega}e^{(\omega\tau(t)+2\lambda\nu)}V_r\left(Y_{N,\nu}\right)^\top g(\nu),$$

and

$$\psi_{ad}^r(\tau(t),\nu,Y_N) = \frac{\omega}{2\lambda-\omega}e^{\omega(\tau(t)+\nu)}\mathcal{P}_{ad}^r\left(\tau(t),\nu\right).$$

Next, using the definitions of $\mathcal{U}_\mu^r\left(\tau(t),\nu,Y_N;\omega\right)$ and $\mathcal{U}_\sigma^r\left(\tau(t),\nu,Y_N;\omega\right)$ in (32), the decomposition in (C.45) produces the following expression:

$$\mathcal{U}_\mu^r\left(\tau(t),\nu,Y_N;\omega\right) = \left(\sum_{i=1}^{3}\psi_i^r(\tau(t),\nu,Y_N)+\psi_{ad}^r(\tau(t),\nu,Y_N)\right)\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)\in\mathbb{R},$$

$$\mathcal{U}_\sigma^r\left(\tau(t),\nu,Y_N;\omega\right) = \left(\sum_{i=1}^{3}\psi_i^r(\tau(t),\nu,Y_N)+\psi_{ad}^r(\tau(t),\nu,Y_N)\right)F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)\in\mathbb{R}^{1\times d}.$$

Then, we may re-write $\Xi_{\mathcal{U}}^r\left(\tau(t),Y_N;\omega\right)$ in (31) as

$$\begin{aligned}\Xi_{\mathcal{U}}^r\left(\tau(t),Y_N;\omega\right) &= \int_0^{\tau(t)}\left(\mathcal{U}_\mu^r\left(\tau(t),\nu,Y_N;\omega\right)d\nu+\mathcal{U}_\sigma^r\left(\tau(t),\nu,Y_N;\omega\right)dW_\nu\right)\\ &=\sum_{i=1}^{3}\int_0^{\tau(t)}\psi_i^r(\tau(t),\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu+F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\\ &\quad+\int_0^{\tau(t)}\psi_{ad}^r(\tau(t),\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu+F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right],\end{aligned}$$

and thus

$$\begin{aligned}\left(\Xi_{\mathcal{U}}^r\left(\cdot,Y_N;\omega\right)\right)_{t^\star} &\leq \sum_{i=1}^{3}\sup_{t\in[0,t^\star]}\left|\int_0^t\psi_i^r(t,\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu+F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|\\ &\quad+\sup_{t\in[0,t^\star]}\left|\int_0^t\psi_{ad}^r(t,\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu+F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|.\end{aligned}$$

Applying the Minkowski's inequality produces the following bound:

$$\begin{aligned}&\left\|\left(\Xi_{\mathcal{U}}^r\left(\cdot,Y_N;\omega\right)\right)_{t^\star}\right\|_{\mathsf{p}}^{\pi_\star^0}\\ &\leq\sum_{i=1}^{3}\left\|\sup_{t\in[0,t^\star]}\left|\int_0^t\psi_i^r(t,\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu+F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|\right\|_{\mathsf{p}}^{\pi_\star^0}\\ &\quad+\left\|\sup_{t\in[0,t^\star]}\left|\int_0^t\psi_{ad}^r(t,\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu+F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|\right\|_{\mathsf{p}}^{\pi_\star^0}.\end{aligned}\tag{C.46}$$

Next, using the definition of ψ_1^r in (C.45) produces

$$\begin{aligned}&\sup_{t\in[0,t^\star]}\left|\int_0^t\psi_1^r(t,\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu+F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|\\ &\leq\frac{\omega}{|2\lambda-\omega|}\sup_{t\in[0,t^\star]}\left|e^{\omega t}\int_0^t e^{\omega\nu}\mathcal{P}_\circ^r\left(t,\nu\right)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu+F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|\\ &\leq\frac{\omega e^{\omega t^\star}}{|2\lambda-\omega|}\sup_{t\in[0,t^\star]}\left|\int_0^t e^{\omega\nu}\mathcal{P}_\circ^r\left(t,\nu\right)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu+F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|.\end{aligned}$$

Since $t^\star$ is a constant, we conclude that

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t \psi_1^r(t,\nu,Y_N) \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}$$
$$\leq \frac{\omega e^{\omega t^\star}}{|2\lambda - \omega|} \left\| \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} \mathcal{P}_\circ^r(t,\nu) \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}. \quad \text{(C.47)}$$

Using the definition of $\mathcal{P}_\circ^r$ in (C.15a), one sees that

$$\int_0^t e^{\omega\nu} \mathcal{P}_\circ^r(t,\nu) \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] = \sum_{i=1}^4 N_i^r(t), \quad \forall t \in [0,T],$$

where the processes $N_i^r(t)$, $i \in \{1, \cdots, 4\}$, are defined in (C.25) in the statement of Proposition C.5. Hence,

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} \mathcal{P}_\circ^r(t,\nu) \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0} \leq \sum_{i=1}^4 \left\| \left(N_i^r\right)_{t^\star} \right\|_p^{\pi_\star^0},$$

thus allowing us to write (C.47) as

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t \psi_1^r(t,\nu,Y_N) \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0} \leq \frac{\omega e^{\omega t^\star}}{|2\lambda - \omega|} \sum_{i=1}^4 \left\| \left(N_i^r\right)_{t^\star} \right\|_p^{\pi_\star^0}.$$

It then follows from Proposition C.5 that

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t \psi_1^r(t,\nu,Y_N) \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}$$
$$\leq \frac{e^{(2\lambda+\omega)t^\star}}{\sqrt{\lambda}\,|2\lambda - \omega|} \Delta_{\mathcal{P}_1}^r(t^\star) \left\| \left(\Lambda_\mu^{\shortparallel}\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \frac{\sqrt{\omega} e^{(2\lambda+\omega)t^\star}}{\sqrt{\lambda}\,|2\lambda - \omega|} \Delta_{\mathcal{P}_2}^r(t^\star) \left\| \left(F_\sigma^{\shortparallel}\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{4\mathsf{p}}^{\pi_\star^0}$$
$$+ \frac{\omega\sqrt{m} e^{(2\lambda+\omega)t^\star}}{2\lambda\,|2\lambda - \omega|} \left\| \left(\mathcal{P}_\sigma^r F_\sigma^{\shortparallel}\left(\cdot, X_N^r\right)^\top\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0}. \quad \text{(C.48)}$$

Next, using the definition of ψ_2^r in (C.45) we see that

$$\sup_{t\in[0,t^\star]} \left| \int_0^t \psi_2^r(t,\nu,Y_N) \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right|$$
$$\leq \frac{\omega e^{2\lambda t^\star}}{|2\lambda - \omega|} \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} V_r\left(Y_{N,\tau(t)}\right)^\top g(\tau(t)) \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right|$$
$$\leq \frac{\omega e^{2\lambda t^\star}}{|2\lambda - \omega|} \Delta_g \Big(V_r\left(Y_N\right)\Big)_{t^\star} \sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu} \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right\|,$$

where we have used the bound on $g(t)$ from Assumption 1. Developing the bound further leads to

$$\sup_{t\in[0,t^\star]} \left| \int_0^t \psi_2^r(t,\nu,Y_N) \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right|$$
$$\leq \frac{\omega e^{2\lambda t^\star}}{|2\lambda - \omega|} \Delta_g \Big(V_r\left(Y_N\right)\Big)_{t^\star} \left(\sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu} \Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu \right\| + \sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu} F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu \right\| \right).$$

Using the Cauchy-Schwarz inequality one sees that

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t \psi_2^r(t,\nu,Y_N) \left[\Lambda_\mu^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\shortparallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}$$

$$\leq \frac{\omega e^{2\lambda t^\star}}{|2\lambda-\omega|}\Delta_g \left\| \left(V_r\left(Y_N\right)\right)_{t^\star} \left(\sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu}\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \right\| + \sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu} F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu \right\| \right) \right\|_p^{\pi_\star^0}$$

$$\leq \frac{\omega e^{2\lambda t^\star}}{|2\lambda-\omega|}\Delta_g \left\| \left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2p}^{\pi_\star^0} \left\| \left(\sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu}\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \right\| + \sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu} F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu \right\| \right) \right\|_{2p}^{\pi_\star^0}.$$

It then follows from the Minkowski's inequality that

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t \psi_2^r(t,\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}$$
$$\leq \frac{\omega e^{2\lambda t^\star}}{|2\lambda-\omega|}\Delta_g \left\| \left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2p}^{\pi_\star^0} \left(\left\| \sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu}\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \right\| \right\|_{2p}^{\pi_\star^0} + \left\| \sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu} F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu \right\| \right\|_{2p}^{\pi_\star^0} \right). \quad \text{(C.49)}$$

Now,

$$\left\| \int_0^t e^{\omega\nu}\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \right\| \leq \int_0^t e^{\omega\nu} \left\|\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right)\right\| d\nu \leq \left(\int_0^t e^{\omega\nu} d\nu\right) \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_t \leq \frac{e^{\omega t}}{\omega}\left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_t,$$

where, in the last inequality, we have used the fact the strict positivity of $\omega \in \mathbb{R}_{>0}$. Therefore,

$$\sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu}\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \right\| \leq \frac{e^{\omega t^\star}}{\omega}\left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}.$$

Since $t^\star$ is a constant, it then follows that

$$\left\| \sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu}\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu \right\| \right\|_{2p}^{\pi_\star^0} \leq \frac{e^{\omega t^\star}}{\omega} \left\| \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2p}^{\pi_\star^0}. \quad \text{(C.50)}$$

Next, observe that the regularity assumptions, along with the t-continuity of the strong solution $X_{N,t}^r$, it is straightforward to establish that $F_\sigma^{\parallel}\left(\cdot, X_N^r\right) \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{m\times d}|\mathfrak{W}_t\right)$. Therefore, we may use Proposition B.1 and the strict positivity of $\omega \in \mathbb{R}_{>0}$ to obtain the following bound:

$$\left\| \sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu} F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu \right\| \right\|_{2p}^{\pi_\star^0} \leq 2\sqrt{2}m\sqrt{\mathsf{p}} \left(\frac{e^{2\omega t^\star}-1}{\omega}\right)^{\frac{1}{2}} \left\| \left(F_\sigma^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2p}^{\pi_\star^0}$$
$$\leq \frac{2\sqrt{2}m\sqrt{\mathsf{p}}e^{\omega t^\star}}{\sqrt{\omega}} \left\| \left(F_\sigma^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2p}^{\pi_\star^0}.$$

Using Jensen's inequality one sees that

$$\left\| \sup_{t\in[0,t^\star]} \left\| \int_0^t e^{\omega\nu} F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu \right\| \right\|_{2p}^{\pi_\star^0} \leq 2\sqrt{2}m\sqrt{\mathsf{p}} \left(\frac{e^{2\omega t^\star}-1}{\omega}\right)^{\frac{1}{2}} \left\| \left(F_\sigma^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2p}^{\pi_\star^0}$$
$$\leq \frac{2\sqrt{2}m\sqrt{\mathsf{p}}e^{\omega t^\star}}{\sqrt{\omega}} \left\| \left(F_\sigma^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{4p}^{\pi_\star^0}. \quad \text{(C.51)}$$

Substituting (C.50) and (C.51) into (C.49) produces

$$\left\|\sup_{t\in[0,t^\star]}\left|\int_0^t \psi_2^r(t,\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu + F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|\right\|_p^{\pi_\star^0}$$
$$\leq \frac{e^{(2\lambda+\omega)t^\star}}{|2\lambda-\omega|}\Delta_g\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2p}^{\pi_\star^0}\left\|\left(\Lambda_\mu^{\parallel}\left(\cdot,X_N^r\right)\right)_{t^\star}\right\|_{2p}^{\pi_\star^0}$$
$$+\frac{\sqrt{\omega}e^{(2\lambda+\omega)t^\star}}{|2\lambda-\omega|}2\sqrt{2}m\sqrt{\mathsf{p}}\Delta_g\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2p}^{\pi_\star^0}\left\|\left(F_\sigma^{\parallel}\left(\cdot,X_N^r\right)\right)_{t^\star}\right\|_{4p}^{\pi_\star^0}. \quad \text{(C.52)}$$

Next, using the definition of ψ_3^r in (C.45) we see that

$$\sup_{t\in[0,t^\star]}\left|\int_0^t \psi_3^r(t,\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu + F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|$$
$$\leq \frac{2\lambda e^{\omega t^\star}}{|2\lambda-\omega|}\sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}V_r\left(Y_{N,\nu}\right)^\top g(\nu)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu + F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|.$$

Applying the Minkowski's inequality then leads to

$$\left\|\sup_{t\in[0,t^\star]}\left|\int_0^t \psi_3^r(t,\nu,Y_N)\left[\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu + F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right]\right|\right\|_p^{\pi_\star^0}$$
$$\leq \frac{2\lambda e^{\omega t^\star}}{|2\lambda-\omega|}\left\|\sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}V_r\left(Y_{N,\nu}\right)^\top g(\nu)\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu\right|\right\|_p^{\pi_\star^0}$$
$$+\frac{2\lambda e^{\omega t^\star}}{|2\lambda-\omega|}\left\|\sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}V_r\left(Y_{N,\nu}\right)^\top g(\nu)F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right|\right\|_p^{\pi_\star^0}. \quad \text{(C.53)}$$

Upon performing similar manipulations to (C.50) and applying the Cauchy-Schwarz inequality, we obtain the following bound:

$$\left\|\sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}V_r\left(Y_{N,\nu}\right)^\top g(\nu)\Lambda_\mu^{\parallel}\left(\nu,X_{N,\nu}^r\right)d\nu\right|\right\|_p^{\pi_\star^0}$$
$$\leq \frac{e^{2\lambda t^\star}}{2\lambda}\Delta_g\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\left(\Lambda_\mu^{\parallel}\left(\cdot,X_N^r\right)\right)_{t^\star}\right\|_p^{\pi_\star^0}$$
$$\leq \frac{e^{2\lambda t^\star}}{2\lambda}\Delta_g\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2p}^{\pi_\star^0}\left\|\left(\Lambda_\mu^{\parallel}\left(\cdot,X_N^r\right)\right)_{t^\star}\right\|_{2p}^{\pi_\star^0}. \quad \text{(C.54)}$$

Next, we may write

$$\int_0^t e^{2\lambda\nu}V_r\left(Y_{N,\nu}\right)^\top g(\nu)F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu = \int_0^t e^{2\lambda\nu}V_r\left(Y_{N,\nu}\right)^\top g(\nu)\left[F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)\quad 0_{m,d}\right]d\widehat{W}_\nu.$$

Furthermore, the regularity assumptions and the t-continuity of the strong solutions $X_{N,t}^r$ and $X_{N,t}^\star$ imply that $\left(V_r\left(Y_{N,t}\right)^\top g(t)\left[F_\sigma^{\parallel}\left(t,X_{N,t}^r\right)\quad 0_{m,d}\right]\right) \in \mathcal{M}_2^{loc}\left(\mathbb{R}^{1\times 2d}|\mathfrak{W}_t\times\mathfrak{W}_t^\star\right)$. Thus, we may use Proposition B.1 and the strict positivity of $\lambda\in\mathbb{R}_{>0}$ to obtain the following bound:

$$\left\|\sup_{t\in[0,t^\star]}\left|\int_0^t e^{2\lambda\nu}V_r\left(Y_{N,\nu}\right)^\top g(\nu)F_\sigma^{\parallel}\left(\nu,X_{N,\nu}^r\right)dW_\nu\right|\right\|_{\mathsf{p}}^{\pi_\star^0}$$
$$\leq 2\sqrt{\mathsf{p}}\Delta_g\frac{e^{2\lambda t^\star}}{\sqrt{2\lambda}}\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\left(F_\sigma^{\parallel}\left(\cdot,X_N^r\right)\right)_{t^\star}\right\|_{\mathsf{p}}^{\pi_\star^0}$$
$$\leq 2\sqrt{\mathsf{p}}\Delta_g\frac{e^{2\lambda t^\star}}{\sqrt{2\lambda}}\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}\left\|\left(F_\sigma^{\parallel}\left(\cdot,X_N^r\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}, \quad \text{(C.55)}$$

where the last inequality follows from the Cauchy-Schwarz inequality. Substituting (C.54) and (C.55) into (C.53) produces

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t \psi_3^r(t,\nu,Y_N) \left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}$$
$$\leq \frac{e^{(2\lambda+\omega)t^\star}}{|2\lambda-\omega|}\Delta_g \left\| \left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2p}^{\pi_\star^0} \left\| \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{2p}^{\pi_\star^0}$$
$$+ \frac{\sqrt{2\lambda}e^{(2\lambda+\omega)t^\star}}{|2\lambda-\omega|} 2\sqrt{\mathsf{p}}\Delta_g \left\| \left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \left\| \left(F_\sigma^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}. \quad \text{(C.56)}$$

Finally, using the definition of ψ_{ad}^r in (C.45) we see that

$$\sup_{t\in[0,t^\star]} \left| \int_0^t \psi_3^r(t,\nu,Y_N) \left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right|$$
$$\leq \frac{\omega e^{\omega t^\star}}{|2\lambda-\omega|} \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu}\mathcal{P}_{ad}^r(t,\nu) \left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right|,$$

and thus

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t \psi_3^r(t,\nu,Y_N) \left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}$$
$$\leq \frac{\omega e^{\omega t^\star}}{|2\lambda-\omega|} \left\| \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu}\mathcal{P}_{ad}^r(t,\nu) \left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}$$
$$= \frac{\omega e^{\omega t^\star}}{|2\lambda-\omega|} \left\| \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} \left(\int_\nu^t e^{(2\lambda-\omega)\beta}\mathcal{P}_{\mathcal{U}}^r(\beta)^\top d\beta\right) \left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0},$$

where, in the last expression we have used the definition of $\mathcal{P}_{ad}^r$ in (C.15b), Proposition C.3. Observe that the last inequality can be expressed as

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t \psi_3^r(t,\nu,Y_N) \left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}$$
$$\leq \frac{\omega e^{\omega t^\star}}{|2\lambda-\omega|} \left\| \sup_{t\in[0,t^\star]} \left| \int_0^t e^{\omega\nu} \left(\int_\nu^t e^{(2\lambda-\omega)\beta}\mathcal{P}_{\mathcal{U}}^r(\beta)^\top d\beta\right) \left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}$$
$$= \frac{\omega e^{\omega t^\star}}{|2\lambda-\omega|} \left\| \left(N_{\mathcal{U}}^r\right)_{t^\star}\right\|_p^{\pi_\star^0},$$

where the process $N_{\mathcal{U}}^r$ is defined in (C.32), Proposition C.6. Thus, we invoke Proposition C.6 and obtain

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t \psi_3^r(t,\nu,Y_N) \left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu\right] \right| \right\|_p^{\pi_\star^0}$$
$$\leq \frac{\omega e^{(2\lambda+\omega)t^\star}}{2\lambda\,|2\lambda-\omega|}\mathsf{p}\sqrt{\frac{n}{2}}\Delta_g^2\Delta_{\partial V} \left(\frac{1}{\sqrt{\omega}} \left\| \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} + \left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \left\| \left(F_\sigma^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}\right)^2$$
$$= \frac{e^{(2\lambda+\omega)t^\star}}{2\lambda\,|2\lambda-\omega|}\mathsf{p}\sqrt{\frac{n}{2}}\Delta_g^2\Delta_{\partial V} \left(\left\| \left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} + \sqrt{\omega}\left(\mathsf{p}\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \left\| \left(F_\sigma^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}\right)^2.$$

It then follows from [110, Prop. 3.1.10-(iii)] that

$$\left\| \sup_{t\in[0,t^\star]} \left| \int_0^t \psi_3^r(t,\nu,Y_N) \left[\Lambda_\mu^{\parallel}\left(\nu, X_{N,\nu}^r\right) d\nu + F_\sigma^{\parallel}\left(\nu, X_{N,\nu}^r\right) dW_\nu \right] \right| \right\|_p^{\pi_\star^0}$$
$$\leq \frac{e^{(2\lambda+\omega)t^\star}}{\lambda\left|2\lambda-\omega\right|} \mathsf{p}\sqrt{\frac{n}{2}}\Delta_g^2\Delta_{\partial V}\left(\mathbb{E}_{\pi_\star^0}\left[\left(\Lambda_\mu^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}^{2\mathsf{p}} \right]^{\frac{1}{\mathsf{p}}} + \omega \mathsf{p}\frac{2\mathsf{p}-1}{2}\mathbb{E}_{\pi_\star^0}\left[\left(F_\sigma^{\parallel}\left(\cdot, X_N^r\right)\right)_{t^\star}^{2\mathsf{p}} \right]^{\frac{1}{\mathsf{p}}} \right). \tag{C.57}$$

To conclude the proof we substitute the bounds in (C.48), (C.52), (C.56), and (C.57) into (C.46) and obtain the desired bound in (C.44) upon grouping terms that are $\propto \left\{\omega, \sqrt{\omega}, 1\right\} / \left|2\lambda-\omega\right|$.

∎

Next, we derive the expressions for the terms that constitute the bound in Proposition C.5.

Proposition C.7 *Suppose there exists a strictly positive $\varrho \in \mathbb{R}_{>0}$ such that*

$$\mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]} \left\| X_{N,t}^r - X_{N,t}^\star \right\|^{2\mathsf{p}^\star} \right] \leq \varrho^{2\mathsf{p}^\star}, \tag{C.58}$$

where the constant $t^\star$ is defined in (29) and $\mathsf{p}^\star$ is defined in Assumption 3. Then, the following bound holds $\forall\, \mathbb{N}_{\geq 1} \ni \mathsf{p} \leq \mathsf{p}^\star$:

$$\Delta_{\mathcal{P}_1}^r(t^\star) \leq \frac{1}{2\sqrt{\lambda}}\left(\sqrt{\frac{n}{2}}\Delta_g\widehat{\Delta}_1^r + \Delta_{\dot{g}}\left\|\nabla V(0,0)\right\| \right) + \sqrt{2n\mathsf{p}}\Delta_g\Delta_{\partial V}\widehat{\Delta}_2^r$$
$$+\sqrt{2n\mathsf{p}}\Delta_g\Delta_{\partial V}\Delta_\sigma\sqrt{\varrho} + \frac{\widehat{\Delta}_3^r}{2\sqrt{\lambda}}\varrho, \tag{C.59a}$$

$$\Delta_{\mathcal{P}_2}^r(t^\star) \leq \left(p^3\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{1}{2\sqrt{\lambda}}\left(\sqrt{\frac{n}{2}}\Delta_g\widehat{\Delta}_1^r + \Delta_{\dot{g}}\left\|\nabla V(0,0)\right\| \right) + \left(2n\mathsf{p}^3(4\mathsf{p}-1)\right)^{\frac{1}{2}}\Delta_g\Delta_{\partial V}\widehat{\Delta}_2^r$$
$$+ \left(2n\mathsf{p}^3(4\mathsf{p}-1)\right)^{\frac{1}{2}}\Delta_g\Delta_{\partial V}\Delta_\sigma\sqrt{\varrho} + \left(p^3\frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}}\frac{\widehat{\Delta}_3^r}{2\sqrt{\lambda}}\varrho, \tag{C.59b}$$

$$\left\| \left(\mathcal{P}_\sigma^r F_\sigma^{\parallel}\left(\cdot, X_N^r\right)^\top \right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} \leq \sqrt{\frac{n}{2}}\Delta_g\Delta_{\partial V}\widehat{\Delta}_4^r\widehat{\Delta}_{4_\parallel}^r + \sqrt{\frac{n}{2}}\Delta_g\Delta_{\partial V}\left(\Delta_\sigma\widehat{\Delta}_{4_\parallel}^r + \Delta_\sigma^{\parallel}\widehat{\Delta}_4^r \right)\sqrt{\varrho}$$
$$+\sqrt{\frac{n}{2}}\Delta_g\Delta_{\partial V}\Delta_\sigma\Delta_\sigma^{\parallel}\varrho, \tag{C.59c}$$

where $\Delta_{\mathcal{P}_1}^r(t^\star)$ and $\Delta_{\mathcal{P}_2}^r(t^\star)$ are defined in the statement of Proposition C.5, $\mathcal{P}_\sigma^r$ is defined in Proposition C.3, and where the constants $\widehat{\Delta}_1^r$, $\widehat{\Delta}_2^r$, $\widehat{\Delta}_3^r$, $\widehat{\Delta}_4^r$, and $\widehat{\Delta}_{4_\parallel}^r$, are defined in (A.1), Appendix A.

Proof. We begin with the term $\Delta_{\mathcal{P}_\mu}^r$ defined in Proposition C.4 which is defined as

$$\Delta_{\mathcal{P}_\mu}^r = \left\| \left(\bar{F}_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\bar{F}_\mu\left(\cdot, X_N^\star\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\Lambda_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0}.$$

Using the bounds (E.4b) and (E.4c) in Proposition E.2, we obtain

$$\Delta_{\mathcal{P}_\mu}^r \leq \left(2\Delta_f + \Delta_\mu\right)\left(1+\Delta_\star\right) + \left(\Delta_f + \Delta_\mu\right)\varrho. \tag{C.60}$$

Next, we consider the term $\Delta_{\mathcal{P}_\sigma}^r(\mathsf{r},\mathsf{s})$, $(\mathsf{r},\mathsf{s}) \in \{2\mathsf{p}, 4\mathsf{p}\}\times\{2\mathsf{p}, 4\mathsf{p}\}$, defined in Proposition C.4 and use the bounds in (E.5c) and (E.5e), to obtain

$$\Delta_{\mathcal{P}_\sigma}^r(4\mathsf{p}, 2\mathsf{p}) \leq 2\Delta_p^2 + \left(\Delta_\sigma\left(1+\Delta_\star\right)^{\frac{1}{2}} + \Delta_\sigma\sqrt{\varrho} \right)^2 \overset{(\star)}{\leq} 2\Delta_p^2 + 2\Delta_\sigma^2\left(1+\Delta_\star\right) + 2\Delta_\sigma^2\varrho, \tag{C.61a}$$

$$\left\{ \Delta_{\mathcal{P}_\sigma}^r(2\mathsf{p}, 2\mathsf{p}), \Delta_{\mathcal{P}_\sigma}^r(4\mathsf{p}, 4\mathsf{p}) \right\} \leq 2\Delta_p + \Delta_\sigma\left(1+\Delta_\star\right)^{\frac{1}{2}} + \Delta_\sigma\sqrt{\varrho}, \tag{C.61b}$$

where $(\star)$ is due to [110, Prop. 3.1.10-(iii)]. Next, recall (C.19a) in the statement of Proposition C.4:

$$\sum_{i=1}^{3}\left\| \left(\mathcal{P}_{\mu_i}^r \right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \leq \sqrt{\frac{n}{2}}\Delta_g\left(\Delta_{\partial V}\Delta_{\mathcal{P}_\mu}^r + \frac{1}{2}\Delta_{\partial^2 V}\Delta_{\mathcal{P}_\sigma}^r(4\mathsf{p}, 2\mathsf{p}) \right) + \Delta_{\dot{g}}\left\| \left(V_r\left(Y_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0}.$$

Using (C.60), (C.61a), and (E.3) in Proposition E.2, one sees that

$$\sum_{i=1}^{3} \left\| \left(\mathcal{P}^r_{\mu_i}\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \leq \sqrt{\frac{n}{2}} \Delta_g \widehat{\Delta}^r_1 + \Delta_{\dot{g}} \left\|\nabla V(0,0)\right\| + \widehat{\Delta}^r_3 \varrho. \tag{C.62}$$

Similarly, from (C.19a) and (C.61b), we obtain

$$\sum_{i \in \{\sigma, \sigma_\star\}} \left\| \left(\mathcal{P}^r_{i}\right)_{t^\star} \right\|_{\{2\mathsf{p},4\mathsf{p}\}}^{\pi^0_\star} \leq \sqrt{\frac{n}{2}} \Delta_g \Delta_{\partial V} \widehat{\Delta}^r_2 + \sqrt{\frac{n}{2}} \Delta_g \Delta_{\partial V} \Delta_\sigma \sqrt{\varrho}. \tag{C.63}$$

Now, recall the definition of $\Delta^r_{\mathcal{P}_1}(t^\star)$ from the statement of Proposition C.5:

$$\Delta^r_{\mathcal{P}_1}(t^\star) = \frac{1}{2\sqrt{\lambda}} \left(\sum_{i=1}^{3} \left\| \left(\mathcal{P}^r_{\mu_i}\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \right) + 2\sqrt{\mathsf{p}} \left(\sum_{i \in \{\sigma, \sigma_\star\}} \left\| \left(\mathcal{P}^r_{i}\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \right),$$

which then leads to (C.59a) by using the bounds in (C.62) - (C.63).
Next, recall the definition of $\Delta^r_{\mathcal{P}_2}(t^\star)$:

$$\Delta^r_{\mathcal{P}_2}(t^\star) = \left(p^3 \frac{2\mathsf{p}-1}{2}\right)^{\frac{1}{2}} \frac{1}{2\sqrt{\lambda}} \left(\sum_{i=1}^{3} \left\| \left(\mathcal{P}^r_{\mu_i}\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \right) + 2\mathsf{p}^{\frac{3}{2}} \left(4\mathsf{p}-1\right)^{\frac{1}{2}} \left(\sum_{i \in \{\sigma, \sigma_\star\}} \left\| \left(\mathcal{P}^r_{i}\right)_{t^\star} \right\|_{4\mathsf{p}}^{\pi^0_\star} \right),$$

which yields (C.59b) upon substitution of the bounds in (C.62) - (C.63).
Finally, the definition of $\mathcal{P}^r_\sigma$ in Proposition C.3 and the submultiplicativity of the Frobenius norm imply that

$$\begin{aligned} \left\| \mathcal{P}^r_\sigma(t) F^{\shortparallel}_\sigma \left(t, X^r_{N,t}\right)^\top \right\|_F &\leq \left\|\mathcal{P}^r_\sigma(t)\right\|_F \left\| F^{\shortparallel}_\sigma \left(t, X^r_{N,t}\right)^\top \right\|_F \\ &= \left\| g(t)^\top V_{r,r}\left(Y_{N,t}\right) F_\sigma \left(t, X^r_{N,t}\right) \right\|_F \left\| F^{\shortparallel}_\sigma \left(t, X^r_{N,t}\right)^\top \right\|_F, \end{aligned}$$

and thus, using the bounds on $g(t)$ and $V_{r,r}\left(Y_{N,t}\right)$ in Assumption 1 and (E.1b), respectively, produces

$$\begin{aligned} \left\| \mathcal{P}^r_\sigma(t) F^{\shortparallel}_\sigma \left(t, X^r_{N,t}\right)^\top \right\|_F &\leq \left\|g(t)\right\|_F \left\|V_{r,r}\left(Y_{N,t}\right)\right\|_F \left\|F_\sigma \left(t, X^r_{N,t}\right)\right\|_F \left\| F^{\shortparallel}_\sigma \left(t, X^r_{N,t}\right)^\top \right\|_F \\ &\leq \sqrt{\frac{n}{2}} \Delta_g \Delta_{\partial V} \left\|F_\sigma \left(t, X^r_{N,t}\right)\right\|_F \left\| F^{\shortparallel}_\sigma \left(t, X^r_{N,t}\right)^\top \right\|_F, \quad \forall t \in [0,T]. \end{aligned}$$

Hence,

$$\left(\mathcal{P}^r_\sigma F^{\shortparallel}_\sigma \left(\cdot, X^r_N\right)^\top\right)_{t^\star} \leq \sqrt{\frac{n}{2}} \Delta_g \Delta_{\partial V} \left(F_\sigma \left(\cdot, X^r_N\right)\right)_{t^\star} \left(F^{\shortparallel}_\sigma \left(\cdot, X^r_N\right)^\top\right)_{t^\star}.$$

It then follows from the Cauchy-Schwarz inequality that

$$\begin{aligned} \left\| \left(\mathcal{P}^r_\sigma F^{\shortparallel}_\sigma \left(\cdot, X^r_N\right)^\top\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi^0_\star} &\leq \sqrt{\frac{n}{2}} \Delta_g \Delta_{\partial V} \left\| \left(F_\sigma \left(\cdot, X^r_N\right)\right)_{t^\star} \left(F^{\shortparallel}_\sigma \left(\cdot, X^r_N\right)^\top\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi^0_\star} \\ &\leq \sqrt{\frac{n}{2}} \Delta_g \Delta_{\partial V} \left\| \left(F_\sigma \left(\cdot, X^r_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star} \left\| \left(F^{\shortparallel}_\sigma \left(\cdot, X^r_N\right)^\top\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi^0_\star}. \end{aligned}$$

The proof is then concluded by observing that one obtains (C.59c) by substituting the bound (E.5b) into the inequality above. □

The following result derives the joint effect of the bounds in Lemmas C.1 and C.2.

Lemma C.3 *Consider the the stopping times $\tau^\star$ and $t^\star$ defined in (29), Lemma 3.1.and assume that $\tau^\star = t^\star$. Furthermore, suppose there exists a strictly positive $\varrho \in \mathbb{R}_{>0}$ such that*

$$\mathbb{E}_{\pi^0_\star}\left[\sup_{t \in [0,t^\star]} \left\| X^r_{N,t} - X^\star_{N,t} \right\|^{2\mathsf{p}^\star} \right] \leq \varrho^{2\mathsf{p}^\star}, \tag{C.64}$$

where $\mathsf{p}^\star$ *is defined in Assumption 3. Then, the following bound holds* $\forall\, \mathbb{N}_{\geq 1} \ni \mathsf{p} \leq \mathsf{p}^\star$*:*

$$\left\| e^{\omega \tau^\star} \left(\Xi^r\left(\cdot, Y_N\right)\right)_{\tau^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} + \left\| \left(\Xi^r_{\mathcal{U}}\left(\cdot, Y_N; \omega\right)\right)_{\tau^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} \\ \leq e^{(2\lambda + \omega)t^\star} \left(\Delta^r_{\circ} + \Delta^r_{\circledcirc} \sqrt{\varrho} + \Delta^r_{\odot} \varrho + \Delta^r_{\otimes} \varrho \sqrt{\varrho} + \Delta^r_{\circledast} \varrho^2 \right), \quad \text{(C.65)}$$

where

$$\Delta^r_{\circ} = \frac{\Delta^r_{\circ_1}}{\lambda} + \frac{\Delta^r_{\circ_2}}{\sqrt{\lambda}} + \frac{1}{|2\lambda - \omega|}\left(\Delta^r_{\circ_3} + \sqrt{\omega}\Delta^r_{\circ_4} + \omega \Delta^r_{\circ_5}\right),$$

$$\Delta^r_{\circledcirc} = \frac{\Delta^r_{\circledcirc_1}}{\sqrt{\lambda}} + \frac{1}{|2\lambda - \omega|}\left(\Delta^r_{\circledcirc_2} + \sqrt{\omega}\Delta^r_{\circledcirc_3} + \omega \Delta^r_{\circledcirc_4}\right),$$

$$\Delta^r_{\odot} = \frac{\Delta^r_{\odot_1}}{\lambda} + \frac{\Delta^r_{\odot_2}}{\sqrt{\lambda}} + \frac{1}{|2\lambda - \omega|}\left(\Delta^r_{\odot_3} + \sqrt{\omega}\Delta^r_{\odot_4} + \omega \Delta^r_{\odot_5}\right),$$

$$\Delta^r_{\otimes} = \frac{\Delta^r_{\otimes_1}}{\sqrt{\lambda}} + \frac{1}{|2\lambda - \omega|}\left(\Delta^r_{\otimes_2} + \sqrt{\omega}\Delta^r_{\otimes_3}\right), \quad \Delta^r_{\circledast} = \frac{\Delta^r_{\circledast_1}}{\lambda} + \frac{\Delta^r_{\circledast_2}}{|2\lambda - \omega|},$$

and where the constants Δ^r_i*,* $i \in \{\circ_1, \ldots, \circ_5, \circledcirc_1, \ldots, \circledcirc_4, \odot_1, \ldots, \odot_5, \otimes_1, \ldots, \otimes_3, \circledast_1, \circledast_2\}$*, are defined in (A.2) - (A.6) in Appendix A.*

Proof. Since $\tau^\star = t^\star$, we use Lemmas C.1 and C.2, and the fact that $t^\star$ is a constant, to see that

$$\left\| e^{\omega \tau^\star} \left(\Xi^r\left(\cdot, Y_N\right)\right)_{\tau^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} + \left\| \left(\Xi^r_{\mathcal{U}}\left(\cdot, Y_N; \omega\right)\right)_{\tau^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} \\ = e^{\omega t^\star} \left\| \left(\Xi^r\left(\cdot, Y_N\right)\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} + \left\| \left(\Xi^r_{\mathcal{U}}\left(\cdot, Y_N; \omega\right)\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0} \\ \leq e^{(2\lambda + \omega)t^\star} \left(\frac{\Delta^r_{\Xi_1}}{\lambda} + \frac{\Delta^r_{\Xi_2}}{\sqrt{\lambda}} + \frac{1}{|2\lambda - \omega|}\left(\Delta^r_{\mathcal{U}_1} + \sqrt{\omega}\Delta^r_{\mathcal{U}_2} + \omega \Delta^r_{\mathcal{U}_3}\right)\right). \quad \text{(C.66)}$$

Now, recall the definitions of $\Delta^r_{\Xi_1}$ and $\Delta^r_{\Xi_2}$ in Lemma C.1:

$$\Delta^r_{\Xi_1} = \frac{\Delta_g}{2} \left\| \left(V_r\left(Y_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \left\| \left(\Lambda^\perp_\mu\left(\cdot, X^r_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \frac{1}{4} \left\| \left(\mathrm{Tr}\left[H_\sigma\left(\cdot, Y_N\right) \nabla^2 V\left(Y_N\right)\right]\right)_{t^\star} \right\|_{\mathsf{p}}^{\pi_\star^0},$$

$$\Delta^r_{\Xi_2} = 2\Delta^\perp_g \sqrt{\mathsf{p}} \left\| \left(V_r\left(Y_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \left\| \left(F^\perp_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + 2\sqrt{\mathsf{p}} \left\| \left(V_\star\left(Y_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \left\| \left(\bar{F}_\sigma\left(\cdot, X^\star_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0}.$$

It then follows from the bounds (E.3), (E.4b), (E.5b), (E.5d) and (E.5e), in Proposition E.2, that

$$\Delta^r_{\Xi_1} \leq \Delta^r_{\circ_1} + \Delta^r_{\odot_1} \varrho + \Delta^r_{\circledast_1} \varrho^2, \quad \text{(C.67)}$$

and

$$\Delta^r_{\Xi_2} \leq \Delta^r_{\circ_2} + \Delta^r_{\circledcirc_1} \sqrt{\varrho} + \Delta^r_{\odot_2} \varrho + \Delta^r_{\otimes_1} \varrho \sqrt{\varrho}. \quad \text{(C.68)}$$

Next, recall the term $\Delta^r_{\mathcal{U}_1}$ in Lemma C.2:

$$\Delta^r_{\mathcal{U}_1} = \left(\frac{\Delta^r_{\mathcal{P}_1}(t^\star)}{\sqrt{\lambda}} + 2\Delta_g \left(1 + \sqrt{2\lambda \mathsf{p}}\right) \left\| \left(V_r\left(Y_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \right) \left\| \left(\Lambda^{\parallel}_\mu\left(\cdot, X^r_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \\ + 2\Delta_g \sqrt{2\lambda \mathsf{p}} \left\| \left(V_r\left(Y_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \left\| \left(F^{\parallel}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \\ + \sqrt{\frac{n}{2}} \frac{\mathsf{p}\Delta_g^2 \Delta_{\partial V}}{\lambda} \mathbb{E}_{\pi_\star^0}\left[\left(\Lambda^{\parallel}_\mu\left(\cdot, X^r_N\right)\right)_{t^\star}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}}.$$

Using (C.59a) in Proposition C.7, and the bounds (E.3), (E.4a), (E.4b), and (E.5b) in Proposition E.2, we obtain:

$$\Delta^r_{\mathcal{U}_1} \leq \Delta^r_{\circ_3} + \Delta^r_{\circledcirc_2} \sqrt{\varrho} + \Delta^r_{\odot_3} \varrho + \Delta^r_{\otimes_2} \varrho \sqrt{\varrho} + \Delta^r_{\circledast_2} \varrho^2. \quad \text{(C.69)}$$

Next, recall the term $\Delta^r_{\mathcal{U}_2}$ in Lemma C.2:

$$\Delta^r_{\mathcal{U}_2} = \left(\frac{\Delta^r_{\mathcal{P}_2}(t^\star)}{\sqrt{\lambda}} + 2\Delta_g m \sqrt{2\mathsf{p}} \left\| \left(V_r\left(Y_N\right) \right)_{t^\star} \right\|^{\pi^0_\star}_{2\mathsf{p}} \right) \left\| \left(F^{\parallel}_\sigma\left(\cdot, X^r_N\right) \right)_{t^\star} \right\|^{\pi^0_\star}_{4\mathsf{p}},$$

which, upon using (C.59b) in Proposition C.7, and the bounds (E.3) and (E.5b) in Proposition E.2, produces:

$$\Delta^r_{\mathcal{U}_2} \leq \Delta^r_{\circ_4} + \Delta^r_{\circledcirc_3}\sqrt{\varrho} + \Delta^r_{\odot_4}\varrho + \Delta^r_{\otimes_3}\varrho\sqrt{\varrho}. \tag{C.70}$$

Finally, recall the term $\Delta^r_{\mathcal{U}_3}$ in Lemma C.2:

$$\Delta^r_{\mathcal{U}_3} = \frac{\sqrt{m}}{2\lambda} \left\| \left(\mathcal{P}^r_\sigma F^{\parallel}_\sigma\left(\cdot, X^r_N\right)^\top \right)_{t^\star} \right\|^{\pi^0_\star}_{\mathsf{p}} + \sqrt{\frac{n}{2} \frac{\mathsf{p}^2(2\mathsf{p}-1)\Delta_g^2 \Delta_{\partial V}}{2\lambda}} \mathbb{E}_{\pi^0_\star}\left[\left(F^{\parallel}_\sigma\left(\cdot, X^r_N\right) \right)^{2\mathsf{p}}_{t^\star} \right]^{\frac{1}{\mathsf{p}}},$$

which, upon using (C.59c) in Proposition C.7, and the bound (E.5a) in Proposition E.2, yields:

$$\Delta^r_{\mathcal{U}_3} \leq \Delta^r_{\circ_5} + \Delta^r_{\circledcirc_4}\sqrt{\varrho} + \Delta^r_{\odot_5}\varrho. \tag{C.71}$$

The proof is then concluded by substituting (C.67) - (C.71) into (C.66).

□

D True (Uncertain) Process

We provide the proof of Proposition 3.2 below.

Proof of Proposition 3.2. We consider the case $\mathbb{P}\left\{Z_{N,0} = Z_0 \in U_N\right\} = 1$ w.l.o.g. since otherwise $\tau_N = 0$ and the result is trivial. Furthermore, since Proposition 3.1 establishes the well-posedness of $X^r_{N,t}$, we only need to show that $X_{N,t}$ is a unique strong solution of

$$dX_{N,t} = F_{N,\mu}\left(t, X_{N,t}, U_{\mathcal{L}_1,t}\right)dt + F_{N,\sigma}\left(t, X_{N,t}\right)dW_t, \quad U_{\mathcal{L}_1} = \mathcal{F}_{\mathcal{L}_1}\left(X_N\right), \quad X_{N,0} = x_0 \sim \xi_0, \tag{D.1}$$

for $t \in [0,T]$, where $\bar{F}_{N,\{\mu,\sigma\}}$ and $F_{N,\{\mu,\sigma\}}$ are defined analogously to $J_{N,\{\mu,\sigma\}}$ in (47).

We begin by defining

$$\begin{aligned} N(z(t)) &= \int_0^t F_{N,\mu}\left(\nu, z(\nu), \mathcal{F}_{\mathcal{L}_1}\left(z\right)\left(\nu\right)\right)d\nu + \int_0^t F_{N,\sigma}\left(\nu, z(\nu)\right)dW_\nu \\ &= \int_0^t \left(f_N\left(\nu, z(\nu)\right) + g(\nu)\mathcal{F}_{\mathcal{L}_1}\left(z\right)\left(\nu\right) + \Lambda_{N,\mu}\left(\nu, z(\nu)\right)\right)d\nu \\ &\qquad + \int_0^t \left(p_N(\nu, z(\nu)) + \Lambda_{N,\sigma}\left(\nu, z(\nu)\right)\right)dW_\nu, \quad t \in [0,T], \end{aligned} \tag{D.2}$$

for any $z \in \mathcal{M}_2\left([0, \tau_N(T)), \mathbb{R}^n \mid \mathfrak{W}_{0,t}\right)$, where f_N, $\Lambda_{N,\mu}$, p_N, and $\Lambda_{N,\sigma}$ denote the truncated versions of the functions f, Λ_μ, p, and Λ_σ, respectively, and where the truncation is defined as in (45). We now add and subtract $\mathcal{F}_r\left(z\right)$, where $\mathcal{F}_r$ is defined in (24), to obtain

$$\begin{aligned} N(z(t)) &= \int_0^t \left(f_N\left(\nu, z(\nu)\right) + g(\nu)\left(\mathcal{F}_{\mathcal{L}_1}\left(z\right)\left(\nu\right) - \mathcal{F}_r\left(z\right)\left(\nu\right) + \mathcal{F}_r\left(z\right)\left(\nu\right)\right) + \Lambda_{N,\mu}\left(\nu, z(\nu)\right)\right)d\nu \\ &\qquad + \int_0^t \left(p_N(\nu, z(\nu)) + \Lambda_{N,\sigma}\left(\nu, z(\nu)\right)\right)dW_\nu \\ &= \int_0^t \left(f_N\left(\nu, z(\nu)\right) + g(\nu)\mathcal{F}_r\left(z\right)\left(\nu\right) + \Lambda_{N,\mu}\left(\nu, z(\nu)\right)\right)d\nu \\ &\qquad + \int_0^t \left(p_N(\nu, z(\nu)) + \Lambda_{N,\sigma}\left(\nu, z(\nu)\right)\right)dW_\nu + \int_0^t g(\nu)\left(\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)\left(z\right)\right)\left(\nu\right)d\nu \\ &= \int_0^t F_{N,\mu}\left(\nu, z(\nu), \mathcal{F}_r\left(z\right)\left(\nu\right)\right)d\nu + \int_0^t F_{N,\sigma}\left(\nu, z(\nu)\right)dW_\nu \\ &\qquad + \int_0^t g(\nu)\left(\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)\left(z\right)\right)\left(\nu\right)d\nu, \quad t \in [0,T]. \end{aligned}$$

We now use an identical line of reasoning up to (C.6) in the proof of Proposition 3.1 to obtain

$$N(z(t)) = \int_0^t M_\mu(t,\nu,z(\nu))\,d\nu + \int_0^t M_\sigma(t,\nu,z(\nu))\,dW_\nu \\ + \int_0^t g(\nu)\left(\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)(z)\right)(\nu)d\nu, \quad t \in [0,T], \tag{D.3}$$

where $M_{\{\mu,\sigma\}}$ are defined in (C.6).

Let us now set $x_0(t) \equiv x_0$, and defined the Picard iterates for (D.1) as

$$x_k(t) = x_0 + \int_0^t F_{N,\mu}\left(\nu, x_{k-1}(\nu), \mathcal{F}_{\mathcal{L}_1}(x_{k-1})(\nu)\right)d\nu + \int_0^t F_{N,\sigma}(\nu, x_{k-1}(\nu))\,dW_\nu, \quad k \in \mathbb{N}, \quad t \in [0,T]. \tag{D.4}$$

It then follows from (D.3) that

$$x_k(t) = x_0 + \int_0^t M_\mu(t,\nu,x_{k-1}(\nu))\,d\nu + \int_0^t M_\sigma(t,\nu,x_{k-1}(\nu))\,dW_\nu \\ + \int_0^t g(\nu)\left(\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)(x_{k-1})\right)(\nu)d\nu, \quad k \in \mathbb{N}, \quad t \in [0,T]. \tag{D.5}$$

Next, we formulate the truncations of the feedback operators $\mathcal{F}_{\mathcal{L}_1}$ and $\mathcal{F}_r$ in (17) and (23), respectively, to obtain

$$\begin{aligned}&\left(\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)(x_{k-1})\right)(t) \\ &\quad = -\boldsymbol{\omega}\int_0^t e^{-\boldsymbol{\omega}(t-\nu)}\hat{\Lambda}_N^{\parallel}(\nu)\,d\nu - \mathcal{F}_{\boldsymbol{\omega}}\left(\Lambda_{N,\mu}^{\parallel}(\cdot,x_{k-1})\right)(t) - \mathcal{F}_{N,\boldsymbol{\omega}}\left(F_{N,\sigma}^{\parallel}(\cdot,x_{k-1}),W\right)(t) \\ &\quad = \boldsymbol{\omega}\int_0^t e^{-\boldsymbol{\omega}(t-\nu)}\left(\Lambda_{N,\mu}^{\parallel}(\nu,x_{k-1}(\nu))\,d\nu + F_{N,\sigma}^{\parallel}(\nu,x_{k-1}(\nu))\,dW_\nu - \hat{\Lambda}_N^{\parallel}(\nu)\,d\nu\right),\end{aligned}$$

for $(k,t) \in \mathbb{N} \times [0,T]$. Writing the expression above in its differential form leads to

$$d\widehat{\mathcal{F}}_{k-1}(t) = \left(-\omega\widehat{\mathcal{F}}_{k-1}(t) + \Lambda_{N,\mu}^{\parallel}(t,x_{k-1}(t)) - \hat{\Lambda}_N^{\parallel}(t)\right)dt + F_{N,\sigma}^{\parallel}(t,x_{k-1}(t))\,dW_t, \tag{D.6}$$

where

$$\widehat{\mathcal{F}}_{k-1}(0) = 0_m, \quad \widehat{\mathcal{F}}_{k-1} \doteq \left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)(x_{k-1}).$$

Since the differential equation above is linear in $\widehat{\mathcal{F}}$, its solution for any stopping time $\tau \in [0,T]$ can be computed as (see e.g. [95, Sec. 5.4.2])

$$\widehat{\mathcal{F}}_{k-1}(t) = e^{-\omega(t-\tau)}\widehat{\mathcal{F}}_{k-1}(\tau) + \omega\int_0^t e^{-\omega(t-\nu)}\left(\Sigma_\nu^{\parallel} - \hat{\Lambda}_N^{\parallel}(\nu)\,d\nu\right), \quad (k,t) \in \mathbb{N} \times [\tau,T], \tag{D.7}$$

where, we have (*formally*) defined

$$\Sigma_t^{\parallel} \doteq \Lambda_{N,\mu}^{\parallel}(t,x_{k-1}(t))\,dt + F_{N,\sigma}^{\parallel}(t,x_{k-1}(t))\,dW_t, \quad (k,t) \in \mathbb{N} \times [0,T].$$

Since the temporal instances $i\boldsymbol{T_s}$, $i \in \mathbb{N}$, are constant, and hence stopping times, we may decompose $\widehat{\mathcal{F}}_{k-1}(t)$ as follows:

$$\widehat{\mathcal{F}}_{k-1}(t) = \omega\int_0^t e^{-\omega(t-\nu)}\left(\Sigma_\nu^{\parallel} - \hat{\Lambda}_N^{\parallel}(\nu)\,d\nu\right), \quad t \in [0,\boldsymbol{T_s}), \tag{D.8a}$$

$$\widehat{\mathcal{F}}_{k-1}(t) = e^{-\omega(t-\boldsymbol{T_s})}\widehat{\mathcal{F}}_{k-1}(\boldsymbol{T_s}) + \omega\int_{\boldsymbol{T_s}}^t e^{-\omega(t-\nu)}\left(\Sigma_\nu^{\parallel} - \hat{\Lambda}_N^{\parallel}(\nu)\,d\nu\right), \quad t \in [\boldsymbol{T_s},2\boldsymbol{T_s}), \tag{D.8b}$$

$$\widehat{\mathcal{F}}_{k-1}(t) = e^{-\omega(t-i\boldsymbol{T_s})}\widehat{\mathcal{F}}_{k-1}(i\boldsymbol{T_s}) + \omega\int_{i\boldsymbol{T_s}}^t e^{-\omega(t-\nu)}\left(\Sigma_\nu^{\parallel} - \hat{\Lambda}_N^{\parallel}(\nu)\,d\nu\right), \\ t \in [i\boldsymbol{T_s},(i+1)\boldsymbol{T_s}), \quad i \in \left\{2,\ldots,\lfloor t/\boldsymbol{T_s}\rfloor\right\}, \quad (k,t) \in \mathbb{N} \times [0,T]. \tag{D.8c}$$

Next, we derive the expression for the truncated adaptive estimate $\hat{\Lambda}_N^{\|}(t)$. Let $\hat{x}_{k-1} = \mathcal{F}_{\lambda_s}(x_{k-1})$, where the operator $\mathcal{F}_{\lambda_s}$ is defined in (18c). Since we are considering the truncated vector fields, we define the Picard iterates of $\hat{x}$ using the operator $\mathcal{F}_{\lambda_s}$ by setting $\hat{x}_0(t) = 0_n$ and

$$\hat{x}_k(t) = \int_0^t \left(-\lambda_s \mathbb{I}_n \tilde{x}_{k-1}(\nu) + f_N(\nu, x_{k-1}(\nu)) + g(\nu)\mathcal{F}_{\mathcal{L}_1}(x_{k-1})(\nu) + \hat{\Lambda}_N(\nu)\right) d\nu,$$
$$(k,t) \in \mathbb{N} \times [0,T], \quad \tilde{x} \doteq \hat{x} - x.$$

Since $\tilde{x}_k = \hat{x}_k - x_k$, it follows from (D.4) that

$$\begin{aligned}\tilde{x}_k(t) = \int_0^t &\left(-\lambda_s \mathbb{I}_n \tilde{x}_{k-1}(\nu) + f_N(\nu, x_{k-1}(\nu)) + g(\nu)\mathcal{F}_{\mathcal{L}_1}(x_{k-1})(\nu) + \hat{\Lambda}_N(\nu)\right) d\nu \\ &- x_0 - \int_0^t F_{N,\mu}(\nu, x_{k-1}(\nu), \mathcal{F}_{\mathcal{L}_1}(x_{k-1})(\nu))\, d\nu - \int_0^t F_{N,\sigma}(\nu, x_{k-1}(\nu))\, dW_\nu \\ &= -x_0 + \int_0^t \left(-\lambda_s \mathbb{I}_n \tilde{x}_{k-1}(\nu) + \hat{\Lambda}_N(\nu)\right) d\nu - \int_0^t \Sigma_\nu, \quad (k,t) \in \mathbb{N} \times [0,T], \end{aligned} \tag{D.9}$$

where, we have (*formally*) defined

$$\Sigma_t \doteq \Lambda_{N,\mu}(t, x_{k-1}(t))\, dt + F_{N,\sigma}(t, x_{k-1}(t))\, dW_t, \quad (k,t) \in \mathbb{N} \times [0,T].$$

As before, since the temporal instances $i\boldsymbol{T_s}$, $i \in \mathbb{N}$, are constant, and hence stopping times, we may write

$$\tilde{x}_k(t) = -x_0 + \int_0^{\boldsymbol{T_s}} \left(-\lambda_s \mathbb{I}_n \tilde{x}_{k-1}(\nu) + \hat{\Lambda}_N(\nu)\right) d\nu - \int_0^t \Sigma_\nu, \quad t \in [0, \boldsymbol{T_s}), \tag{D.10a}$$

$$\tilde{x}_k(t) = \tilde{x}_k(i\boldsymbol{T_s}) + \int_{i\boldsymbol{T_s}}^t \left(-\lambda_s \mathbb{I}_n \tilde{x}_{k-1}(\nu) + \hat{\Lambda}_N(\nu)\right) d\nu - \int_{i\boldsymbol{T_s}}^t \Sigma_\nu, \quad t \in [i\boldsymbol{T_s}, (i+1)\boldsymbol{T_s}), \tag{D.10b}$$

for $(k,t) \in \mathbb{N} \times [0,T]$ and $i \in \left\{1, \ldots, \lfloor t/\boldsymbol{T_s} \rfloor\right\}$. Using the definition of the adaptation law in (18b) we have that

$$\begin{aligned}\hat{\Lambda}(t) = \mathcal{F}_{\boldsymbol{T_s}}(\hat{x}_{k-1}, x_{k-1})(t) =& 0_n \mathbb{1}_{\{[0,\boldsymbol{T_s})\}}(t) \\ &+ \lambda_s \left(1 - e^{\lambda_s \boldsymbol{T_s}}\right)^{-1} \sum_{i=1}^{\lfloor \frac{t}{\boldsymbol{T_s}} \rfloor} \tilde{x}_{k-1}(i\boldsymbol{T_s}) \mathbb{1}_{\{[i\boldsymbol{T_s},(i+1)\boldsymbol{T_s})\}}(t), \quad t \in [0,T]. \end{aligned} \tag{D.11}$$

Hence, substituting the above into (D.10) leads to

$$\tilde{x}_k(t) = -x_0 - \int_0^{\boldsymbol{T_s}} \lambda_s \mathbb{I}_n \tilde{x}_{k-1}(\nu) d\nu - \int_0^{\boldsymbol{T_s}} \Sigma_\nu, \quad t \in [0, \boldsymbol{T_s}),$$

$$\tilde{x}_k(t) = \tilde{x}_k(i\boldsymbol{T_s}) + \int_{i\boldsymbol{T_s}}^t \left(-\lambda_s \mathbb{I}_n \tilde{x}_{k-1}(\nu) + \lambda_s \left(1 - e^{\lambda_s \boldsymbol{T_s}}\right)^{-1} \tilde{x}_{k-1}(i\boldsymbol{T_s})\right) d\nu - \int_{i\boldsymbol{T_s}}^t \Sigma_\nu, \quad t \in [i\boldsymbol{T_s}, (i+1)\boldsymbol{T_s}).$$

The two expressions above represent the Picard iterates of the following respective differential equations:

$$\begin{aligned} d\tilde{x}_{k-1}(t) &= \left[-\lambda_s \mathbb{I}_n \tilde{x}_{k-1}(t) - \Lambda_{N,\mu}(t, x_{k-1}(t))\right] dt \\ &\qquad - F_{N,\sigma}(t, x_{k-1}(t))\, dW_t, \quad \tilde{x}_{k-1}(0) = -x_0, \quad t \in [0, \boldsymbol{T_s}), \\ d\tilde{x}_{k-1}(t) &= \left[-\lambda_s \mathbb{I}_n \tilde{x}_{k-1}(t) + \lambda_s \left(1 - e^{\lambda_s \boldsymbol{T_s}}\right)^{-1} \tilde{x}_{k-1}(i\boldsymbol{T_s}) - \Lambda_{N,\mu}(t, x_{k-1}(t))\right] dt \\ &\qquad - F_{N,\sigma}(t, x_{k-1}(t))\, dW_t, \quad t \in [i\boldsymbol{T_s}, (i+1)\boldsymbol{T_s}), \end{aligned}$$

for $(k,t) \in \mathbb{N} \times [0,T]$ and $i \in \left\{1, \ldots, \lfloor t/\boldsymbol{T_s} \rfloor\right\}$, where we have used the definition of Σ_t is defined in (D.9). The equations above are linear in $\tilde{x}_{k-1}$, while the exogenous drift and diffusion inputs $\Lambda_{N,\mu}(t, x_{k-1}(t))$ and $F_{N,\sigma}(t, x_{k-1}(t))$ are continuous and uniformly bounded over $[0,T]$ due to the truncation and the assumed regularity. Thus, we invoke [100, Thm. 2.3.1] to establish the well-posedness of the equations above (in the variable $\tilde{x}_{k-1}$). Furthermore, the linearity in $\tilde{x}_{k-1}$ implies that we may write (see e.g. [95, Sec. 5.4.2])

$$\tilde{x}_{k-1}(t) = -e^{-\lambda_s t} x_0 - \int_0^t e^{-\lambda_s (t-\nu)} \Sigma_\nu, \quad t \in [0, \boldsymbol{T_s}), \tag{D.14a}$$

$$\tilde{x}_{k-1}(t) = e^{-\lambda_s(t-i\boldsymbol{T_s})}\tilde{x}_{k-1}(i\boldsymbol{T_s}) + \lambda_s\left(1-e^{\lambda_s\boldsymbol{T_s}}\right)^{-1}\tilde{x}_{k-1}\left(i\boldsymbol{T_s}\right)\int_{i\boldsymbol{T_s}}^{t} e^{-\lambda_s(t-\nu)}d\nu - \int_{i\boldsymbol{T_s}}^{t} e^{-\lambda_s(t-\nu)}\Sigma_\nu, \quad t\in[i\boldsymbol{T_s},(i+1)\boldsymbol{T_s}), \tag{D.14b}$$

for $(k,t)\in\mathbb{N}\times[0,T]$ and $i\in\left\{1,\ldots,\lfloor t/\boldsymbol{T_s}\rfloor\right\}$, where we have substituted back the definition of Σ_t is defined in (D.9). Simplifying the expression above by solving the second integral in the last equation yields

$$\tilde{x}_{k-1}(t) = -e^{-\lambda_s t}x_0 - \int_0^t e^{-\lambda_s(t-\nu)}\Sigma_\nu, \quad t\in[0,\boldsymbol{T_s}),$$

$$\tilde{x}_{k-1}(t) = e^{-\lambda_s(t-i\boldsymbol{T_s})}\tilde{x}_{k-1}(i\boldsymbol{T_s}) + \left(1-e^{\lambda_s\boldsymbol{T_s}}\right)^{-1}\left(1-e^{-\lambda_s(t-i\boldsymbol{T_s})}\right)\tilde{x}_{k-1}\left(i\boldsymbol{T_s}\right) - \int_{i\boldsymbol{T_s}}^{t} e^{-\lambda_s(t-\nu)}\Sigma_\nu, \quad t\in[i\boldsymbol{T_s},(i+1)\boldsymbol{T_s}),$$

for $(k,t)\in\mathbb{N}\times[0,T]$ and $i\in\left\{1,\ldots,\lfloor t/\boldsymbol{T_s}\rfloor\right\}$. Hence, we conclude that

$$\tilde{x}_{k-1}(\boldsymbol{T_s}) = -e^{-\lambda_s\boldsymbol{T_s}}x_0 - \int_0^{\boldsymbol{T_s}} e^{-\lambda_s(\boldsymbol{T_s}-\nu)}\Sigma_\nu, \tag{D.16a}$$

$$\tilde{x}_{k-1}(i\boldsymbol{T_s}) = -\int_{(i-1)\boldsymbol{T_s}}^{i\boldsymbol{T_s}} e^{-\lambda_s(i\boldsymbol{T_s}-\nu)}\Sigma_\nu, \quad i\in\left\{2,\ldots,\lfloor t/\boldsymbol{T_s}\rfloor\right\}, \quad (k,t)\in\mathbb{N}\times[0,T]. \tag{D.16b}$$

□

The next two results help us with the computation of $dV\left(Z_{N,t}\right)$ in the proof of Lemma 3.2.

Proposition D.1 *Let $Z_{N,t}$ be the strong solution of (47), and let $\tau(t)$ be the stopping time defined in (49), Lemma 3.2. Then,*

$$\int_0^{\tau(t)} e^{2\lambda\nu}\left(\nabla V\left(Z_{N,\nu}\right)^\top J_\mu\left(\nu,Z_{N,\nu}\right) + \frac{1}{2}Tr\left[K_\sigma\left(\nu,Z_{N,\nu}\right)\nabla^2 V\left(Z_{N,\nu}\right)\right]\right)d\nu \leq -2\lambda\int_0^{\tau(t)} e^{2\lambda\nu}V\left(Z_{N,\nu}\right)d\nu + \int_0^{\tau(t)} e^{2\lambda\nu}\widetilde{\mathcal{U}}\left(\nu,Z_{N,\nu}\right)d\nu + \int_0^{\tau(t)} e^{2\lambda\nu}\phi_U\left(\nu,Z_{N,\nu}\right)d\nu + \int_0^{\tau(t)} e^{2\lambda\nu}\left(\phi_\mu\left(\nu,Z_{N,\nu}\right)+\phi_{\mu^{\shortparallel}}\left(\nu,Z_{N,\nu}\right)\right)d\nu, \tag{D.17a}$$

$$\int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{N,\nu}\right)^\top J_\sigma\left(\nu,Z_{N,\nu}\right)dW_\nu = \int_0^{\tau(t)} e^{2\lambda\nu}\left(\phi_\sigma\left(\nu,Z_{N,\nu}\right)+\phi_{\sigma^{\shortparallel}}\left(\nu,Z_{N,\nu}\right)\right)dW_\nu, \tag{D.17b}$$

for all $t\in\mathbb{R}_{\geq 0}$, where $K_\sigma\left(\nu,Z_{N,\nu}\right) = J_\sigma\left(\nu,Z_{N,\nu}\right)J_\sigma\left(\nu,Z_{N,\nu}\right)^\top$, and where $J_\mu\left(\nu,Z_{N,\nu}\right)$ and $J_\sigma\left(\nu,Z_{N,\nu}\right)$ are defined in (45), and the functions $\widetilde{\mathcal{U}}$, ϕ_μ^r, and ϕ_σ^r are defined in (52) in the statement of Lemma 3.2. Additionally, we have defined

$$\phi_{\mu^{\shortparallel}}\left(\nu,Z_{N,\nu}\right) = \nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2\otimes g(\nu)\right)\left(\Lambda_\mu^{\shortparallel}\odot Z_N\right)(\nu),$$
$$\phi_{\sigma^{\shortparallel}}\left(\nu,Z_{N,\nu}\right) = \nabla V\left(Z_\nu\right)^\top\left(\mathbb{I}_2\otimes g(\nu)\right)\left(F_\sigma^{\shortparallel}\odot Z_N\right)(\nu),$$
$$\phi_U\left(\nu,Z_{N,\nu}\right) = \nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2\otimes g(\nu)\right)\left(\mathcal{F}_r\odot Z_N\right)(\nu).$$

Proof. Using the definitions of J_μ in (45), we have that

$$\nabla V\left(Z_{N,\nu}\right)^\top J_\mu\left(\nu,Z_{N,\nu}\right) = V_\cdot\left(Z_{N,\nu}\right)^\top\bar{F}_\mu\left(\nu,X_{N,\nu}\right) + V_r\left(Z_{N,\nu}\right)^\top\bar{F}_\mu\left(\nu,X_{N,\nu}^r\right) + V_\cdot\left(Z_{N,\nu}\right)^\top\left(g(\nu)U_{\mathcal{L}_1,\nu}+\Lambda_\mu\left(\nu,X_{N,\nu}\right)\right) + V_r\left(Z_{N,\nu}\right)^\top\left(g(\nu)U_\nu^r+\Lambda_\mu\left(\nu,X_{N,\nu}^r\right)\right), \quad \nu\in[0,\tau(t)],$$

which, upon using (9), Assumption 2 can be re-written as

$$\nabla V\left(Z_{N,\nu}\right)^{\top} J_{\mu}\left(\nu, Z_{N,\nu}\right) \leq -2\lambda V\left(Z_{N,\nu}\right) + V_{\cdot}\left(Z_{N,\nu}\right)^{\top}\left(g(\nu)U_{\mathcal{L}_1,\nu} + \Lambda_{\mu}\left(\nu, X_{N,\nu}\right)\right) + V_r\left(Z_{N,\nu}\right)^{\top}\left(g(\nu)U^r_{\nu} + \Lambda_{\mu}\left(\nu, X^r_{N,\nu}\right)\right), \quad \nu \in [0,\tau(t)].$$

Adding and subtracting $\mathcal{F}_r\left(X_N\right)$ then leads to

$$\begin{aligned}
&\nabla V\left(Z_{N,\nu}\right)^{\top} J_{\mu}\left(\nu, Z_{N,\nu}\right) \\
&\leq -2\lambda V\left(Z_{N,\nu}\right) + V_{\cdot}\left(Z_{N,\nu}\right)^{\top} g(\nu)\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)\left(X_N\right)(\nu) \\
&\quad + V_{\cdot}\left(Z_{N,\nu}\right)^{\top}\left(g(\nu)\mathcal{F}_r\left(X_N\right)(\nu) + \Lambda_{\mu}\left(\nu, X_{N,\nu}\right)\right) + V_r\left(Z_{N,\nu}\right)^{\top}\left(g(\nu)U^r_{\nu} + \Lambda_{\mu}\left(\nu, X^r_{N,\nu}\right)\right),
\end{aligned} \tag{D.18}$$

for $\nu \in [0,\tau(t)]$, where we have used the definition that $U_{\mathcal{L}_1} \doteq \mathcal{F}_{\mathcal{L}_1}\left(X_N\right)$. We develop the expression further by using (11) in Assumption 4 to conclude that

$$\Lambda_{\mu}(\nu,\cdot) = \begin{bmatrix} g(\nu) & g(\nu)^{\perp} \end{bmatrix} \begin{bmatrix} \Lambda^{\parallel}_{\mu}(\nu,\cdot) \\ \Lambda^{\perp}_{\mu}(\nu,\cdot) \end{bmatrix} = g(\nu)\Lambda^{\parallel}_{\mu}(\nu,\cdot) + g(\nu)^{\perp}\Lambda^{\perp}_{\mu}(\nu,\cdot).$$

Substituting into (D.18) yields

$$\begin{aligned}
&\nabla V\left(Z_{N,\nu}\right)^{\top} J_{\mu}\left(\nu, Z_{N,\nu}\right) \\
&\leq -2\lambda V\left(Z_{N,\nu}\right) + V_{\cdot}\left(Z_{N,\nu}\right)^{\top} g(\nu)\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)\left(X_N\right)(\nu) + \nabla V\left(Z_{N,\nu}\right)^{\top}\left(\mathbb{I}_2 \otimes g(\nu)^{\perp}\right)\left(\Lambda^{\perp}_{\mu} \odot \begin{bmatrix} X_N \\ X^r_N \end{bmatrix}\right)(\nu) \\
&\quad + \nabla V\left(Z_{N,\nu}\right)^{\top}\left(\mathbb{I}_2 \otimes g(\nu)\right)\left(\left(\mathcal{F}_r + \Lambda^{\parallel}_{\mu}\right) \odot \begin{bmatrix} X_N \\ X^r_N \end{bmatrix}\right)(\nu), \quad \nu \in [0,\tau(t)],
\end{aligned}$$

where we have used the definition that $U^r \doteq \mathcal{F}_r\left(X^r_N\right)$. The expression in (D.17a) then follows from the last inequality in the straightforward manner.

Next, using the definition of J_{σ} in (45), we have that

$$\begin{aligned}
\int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{N,\nu}\right)^{\top} J_{\sigma}\left(\nu, Z_{N,\nu}\right) dW_{\nu} &= \int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{\nu}\right)^{\top}\left(F_{\sigma} \odot \begin{bmatrix} X_N \\ X^r_N \end{bmatrix}\right)(\nu) dW_{\nu} \\
&= \int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{\nu}\right)^{\top}\left(\left(p + \Lambda_{\sigma}\right) \odot \begin{bmatrix} X_N \\ X^r_N \end{bmatrix}\right)(\nu) dW_{\nu}, \quad t \in \mathbb{R}_{\geq 0},
\end{aligned}$$

where we have used the definition of F_{σ} in (2). Since (11) and (12) in Assumptions 4 and 5, respectively, along with Definition 4 imply that

$$\begin{aligned}
p(\nu,\cdot) + \Lambda_{\sigma}(\nu,\cdot) &= g(\nu)^{\perp}p^{\perp}(\nu,\cdot) + g(\nu)^{\perp}\Lambda^{\perp}_{\sigma}(\nu,\cdot) + g(\nu)p^{\parallel}(\nu,\cdot) + g(\nu)\Lambda^{\parallel}_{\sigma}(\nu,\cdot) \\
&= g(\nu)^{\perp}F^{\perp}_{\sigma}(\nu,\cdot) + g(\nu)F^{\parallel}_{\sigma}(\nu,\cdot),
\end{aligned}$$

the previous integral equality can be re-written as

$$\begin{aligned}
&\int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{N,\nu}\right)^{\top} J_{\sigma}\left(\nu, Z_{N,\nu}\right) dW_{\nu} \\
&= \int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{\nu}\right)^{\top}\left(\mathbb{I}_2 \otimes g(\nu)^{\perp}\right)\left(F^{\perp}_{\sigma} \odot \begin{bmatrix} X_N \\ X^r_N \end{bmatrix}\right)(\nu) dW_{\nu} \\
&\quad + \int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{\nu}\right)^{\top}\left(\mathbb{I}_2 \otimes g(\nu)\right)\left(F^{\parallel}_{\sigma} \odot \begin{bmatrix} X_N \\ X^r_N \end{bmatrix}\right)(\nu) dW_{\nu},
\end{aligned}$$

for $t \in \mathbb{R}_{\geq 0}$, thus establishing the expression in (D.17b). □

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.2 *Let $Z_{N,t}$ be the strong solution of (47), and let $\tau(t)$ be the stopping time defined in (49), Lemma 3.2. Then, for the term ϕ_U defined in the statement of Proposition D.1, we have that*

$$\int_0^{\tau(t)} e^{2\lambda\nu}\phi_U\left(\nu, Z_{N,\nu}\right) d\nu = \int_0^{\tau(t)}\left(\hat{\mathcal{U}}_{\mu}\left(\tau(t), \nu, Z_N; \boldsymbol{\omega}\right) d\nu + \hat{\mathcal{U}}_{\sigma}\left(\tau(t), \nu, Z_N; \boldsymbol{\omega}\right) dW_{\nu}\right)$$

$$+\int_0^{\tau(t)} e^{2\lambda\nu}\left(\phi_{U_\mu}\left(\nu, Z_{N,\nu};\boldsymbol{\omega}\right)d\nu + \phi_{U_\sigma}\left(\nu, Z_{N,\nu};\boldsymbol{\omega}\right)dW_\nu\right), \quad t\in\mathbb{R}_{\geq 0}, \quad \text{(D.19)}$$

where

$$\hat{\mathcal{U}}_\mu\left(\tau(t),\nu,Z_N;\boldsymbol{\omega}\right) = e^{-\boldsymbol{\omega}\tau(t)}\frac{\boldsymbol{\omega}}{2\lambda-\boldsymbol{\omega}}\left(e^{\boldsymbol{\omega}\tau(t)}\mathcal{P}\left(\tau(t),\nu\right) - e^{2\lambda\tau(t)}\nabla V\left(Z_{N,\tau(t)}\right)^\top\left(\mathbb{I}_2\otimes g(\tau(t))\right)\right) \times e^{\boldsymbol{\omega}\nu}\left(\Lambda_\mu^{\parallel}\odot Z_N\right)(\nu),$$

$$\hat{\mathcal{U}}_\sigma\left(\tau(t),\nu,Z_N;\boldsymbol{\omega}\right) = e^{-\boldsymbol{\omega}\tau(t)}\frac{\boldsymbol{\omega}}{2\lambda-\boldsymbol{\omega}}\left(e^{\boldsymbol{\omega}\tau(t)}\mathcal{P}\left(\tau(t),\nu\right) - e^{2\lambda\tau(t)}\nabla V\left(Z_{N,\tau(t)}\right)^\top\left(\mathbb{I}_2\otimes g(\tau(t))\right)\right) \times e^{\boldsymbol{\omega}\nu}\left(F_\sigma^{\parallel}\odot Z_N\right)(\nu),$$

$$\phi_{U_\mu}\left(\nu,Z_{N,\nu};\boldsymbol{\omega}\right) = \frac{\boldsymbol{\omega}}{2\lambda-\boldsymbol{\omega}}\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2\otimes g(\nu)\right)\left(\Lambda_\mu^{\parallel}\odot Z_N\right)(\nu),$$

$$\phi_{U_\sigma}\left(\nu,Z_{N,\nu};\boldsymbol{\omega}\right) = \frac{\boldsymbol{\omega}}{2\lambda-\boldsymbol{\omega}}\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2\otimes g(\nu)\right)\left(F_\sigma^{\parallel}\odot Z_N\right)(\nu),$$

and where $\mathcal{P}\left(\tau(t),\nu\right)$ *is defined in* (54)*, Lemma* 3.2*.*

Proof. Using the definition of ϕ_U in the statement of Proposition D.1, we have that

$$\int_0^{\tau(t)} e^{2\lambda\nu}\phi_U\left(\nu,Z_{N,\nu}\right)d\nu$$
$$= \int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2\otimes g(\nu)\right)\left(\mathcal{F}_r\odot Z_N\right)(\nu)d\nu$$
$$= \int_0^{\tau(t)} e^{2\lambda\nu}\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2\otimes g(\nu)\right)\left(\mathcal{F}_{\boldsymbol{\omega}}\odot\left(\Lambda_\mu^{\parallel}\odot Z_N\right)(\nu) + \mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}\odot\left(F_\sigma^{\parallel}\odot Z_N, W\right)(\nu)\right)d\nu,$$

for all $t\geq 0$, where we have used the definition of $\mathcal{F}_r$ in (23). Next, using the definitions of $\mathcal{F}_{\boldsymbol{\omega}}$ and $\mathcal{F}_{\mathcal{N},\boldsymbol{\omega}}(\cdot,W)$ in (18a) and (23), respectively, we can re-write the previous expression as

$$\int_0^{\tau(t)} e^{2\lambda\nu}\phi_U\left(\nu,Z_{N,\nu}\right)d\nu = \int_0^{\tau(t)}\int_0^{\nu}\left(-\boldsymbol{\omega}e^{(2\lambda-\boldsymbol{\omega})\nu}\right)\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2\otimes g(\nu)\right)\left(e^{\boldsymbol{\omega}\beta}\left(\Lambda_\mu^{\parallel}\odot Z_N\right)(\beta)d\beta\right)d\nu$$
$$+\int_0^{\tau(t)}\int_0^{\nu}\left(-\boldsymbol{\omega}e^{(2\lambda-\boldsymbol{\omega})\nu}\right)\nabla V\left(Z_{N,\nu}\right)^\top\left(\mathbb{I}_2\otimes g(\nu)\right)\left(e^{\boldsymbol{\omega}\beta}\left(F_\sigma^{\parallel}\odot Z_N\right)(\beta)dW_\beta\right)d\nu,$$

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:

$$\int_0^{\tau(t)} e^{2\lambda\nu}\phi_U\left(\nu,Z_{N,\nu}\right)d\nu$$
$$= \int_0^{\tau(t)}\left(-\int_\nu^{\tau(t)}\boldsymbol{\omega}e^{(2\lambda-\boldsymbol{\omega})\beta}\nabla V\left(Z_{N,\beta}\right)^\top\left(\mathbb{I}_2\otimes g(\beta)\right)d\beta\right)e^{\boldsymbol{\omega}\nu}\left(\Lambda_\mu^{\parallel}\odot Z_N\right)(\nu)d\nu$$
$$+\int_0^{\tau(t)}\left(-\int_\nu^{\tau(t)}\boldsymbol{\omega}e^{(2\lambda-\boldsymbol{\omega})\beta}\nabla V\left(Z_{N,\beta}\right)^\top\left(\mathbb{I}_2\otimes g(\beta)\right)d\beta\right)e^{\boldsymbol{\omega}\nu}\left(F_\sigma^{\parallel}\odot Z_N\right)(\nu)dW_\nu, \quad \text{(D.20)}$$

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.

□

Similar to Proposition C.3, we derive an alternative representation for the term $\mathcal{P}\left(\tau(t),\nu\right)$ next.

Proposition D.3 *Recall the expression for* $\mathcal{P}\left(\tau(t),\nu\right)$ *in* (54) *in the statement of Lemma* 3.2 *which we restate below:*

$$\mathcal{P}\left(\tau(t),\nu\right) = \int_\nu^{\tau(t)} e^{(2\lambda-\boldsymbol{\omega})\beta}d_\beta\left[\nabla V\left(Z_{N,\beta}\right)^\top\left(\mathbb{I}_2\otimes g(\beta)\right)\right]d\beta\in\mathbb{R}^{1\times 2m}, \quad 0\leq\nu\leq\tau(t), \quad \text{(D.21)}$$

where the $\tau(t)$ is defined in (49). Then, $\mathcal{P}(\tau(t),\nu)$ admits the following representation:

$$\mathcal{P}(\tau(t),\nu) = \mathcal{P}_\circ(\tau(t),\nu) + \mathcal{P}_{ad}(\tau(t),\nu) + \widetilde{\mathcal{P}}(\tau(t),\nu) \in \mathbb{R}^{1\times 2m}, \quad 0 \le \nu \le \tau(t),\ t \in \mathbb{R}_{\ge 0}, \tag{D.22}$$

where

$$\mathcal{P}_\circ(\tau(t),\nu) = \sum_{i=1}^{3} \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} \mathcal{P}_{\mu_i}(\beta)^\top d\beta + \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} \left(\mathcal{P}_\sigma(\beta) dW_\beta\right)^\top, \tag{D.23a}$$

$$\mathcal{P}_{ad}(\tau(t),\nu) = \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} \mathcal{P}_{\mathcal{U}}(\beta)^\top d\beta, \tag{D.23b}$$

$$\widetilde{\mathcal{P}}(\tau(t),\nu) = \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} \widetilde{\mathcal{P}}_{\mathcal{U}}(\beta)^\top d\beta, \tag{D.23c}$$

and where $\mathcal{P}_{\mathcal{U}}(\beta)$, $\widetilde{\mathcal{P}}_{\mathcal{U}}(\beta) \in \mathbb{R}^{2m}$ are defined as

$$\mathcal{P}_{\mathcal{U}}(\beta) = (\mathbb{I}_2 \otimes g(\beta))^\top \nabla^2 V(Z_{N,\beta}) (\mathbb{I}_2 \otimes g(\beta)) (\mathcal{F}_r \odot Z_N)(\beta),$$
$$\widetilde{\mathcal{P}}_{\mathcal{U}}(\beta) = (\mathbb{I}_2 \otimes g(\beta))^\top \nabla^2 V(Z_{N,\beta}) \begin{bmatrix} g(\beta)(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(\beta) \\ 0_n \end{bmatrix},$$

the terms $\mathcal{P}_{\mu_1}(\beta) \in \mathbb{R}^{2m}$, $i \in \{1,2,3\}$ are defined as

$$\mathcal{P}_{\mu_1}(\beta) = (\mathbb{I}_2 \otimes \dot{g}(\beta))^\top \nabla V(Z_{N,\beta}),$$
$$\mathcal{P}_{\mu_2}(\beta) = (\mathbb{I}_2 \otimes g(\beta))^\top \nabla^2 V(Z_{N,\beta}) \left(\left[\bar{F}_\mu + \Lambda_\mu\right] \odot Z_N\right)(\beta),$$
$$\mathcal{P}_{\mu_3}(\beta) = \frac{1}{2} (\mathbb{I}_2 \otimes g(\beta))^\top \vec{Tr}\left[K_\sigma(\beta, Z_{N,\beta}) \nabla^2 V_i(Z_{N,\beta})\right]_{i=1}^{2n},$$

and $\mathcal{P}_\sigma(\beta) \in \mathbb{R}^{2m\times d}$ is defined as

$$\mathcal{P}_\sigma(\beta) = (\mathbb{I}_2 \otimes g(\beta))^\top \nabla^2 V(Z_{N,\beta}) J_\sigma(\beta, Z_{N,\beta}).$$

Additionally, we have defined $K_\sigma(\beta, Z_{N,\beta}) \doteq J_\sigma(\beta, Z_{N,\beta}) J_\sigma(\beta, Z_{N,\beta})^\top \in \mathbb{S}^{2n}$ and

$$\vec{Tr}\left[K_\sigma(\beta, Z_{N,\beta}) \nabla^2 V_i(Z_{N,\beta})\right]_{i=1}^{2n} \doteq \begin{bmatrix}\mathfrak{K}_1(\beta, Z_{N,\beta}) & \cdots & \mathfrak{K}_{2n}(\beta, Z_{N,\beta})\end{bmatrix}^\top \in \mathbb{R}^{2n},$$
$$\mathfrak{K}_i(\beta, Z_{N,\beta}) = Tr\left[K_\sigma(\beta, Z_{N,\beta}) \nabla^2 V_i(Z_{N,\beta})\right] \in \mathbb{R}.$$

Proof. We closely follow the proof of Proposition C.3 and begin by defininig

$$\widehat{g}(\cdot) \doteq \mathbb{I}_2 \otimes g(\cdot) = \begin{bmatrix} g(\cdot) & 0_{n,m} \\ 0_{n,m} & g(\cdot) \end{bmatrix} \in \mathbb{R}^{2n\times 2m}. \tag{D.24}$$

Then, we write $\nabla V(Z_{N,\beta})^\top (\mathbb{I}_2 \otimes g(\beta))$ as

$$\begin{aligned}\nabla V(Z_{N,\beta})^\top (\mathbb{I}_2 \otimes g(\beta)) &= \nabla V(Z_{N,\beta})^\top \widehat{g}(\beta) \\ &= \begin{bmatrix}\nabla V(Z_{N,\beta})^\top \widehat{g}_{\cdot,1}(\beta) & \cdots & \nabla V(Z_{N,\beta})^\top \widehat{g}_{\cdot,2m}(\beta)\end{bmatrix} \\ &= \begin{bmatrix}\sum_{i=1}^{2n} V_i(Z_{N,\beta})^\top \widehat{g}_{i,1}(\beta) & \cdots & \sum_{i=1}^{2n} V_i(Z_{N,\beta})^\top \widehat{g}_{i,2m}(\beta)\end{bmatrix} \in \mathbb{R}^{1\times 2m},\end{aligned} \tag{D.25}$$

where $\widehat{g}_{\cdot j}(\beta) \in \mathbb{R}^{2n}$ is the j-th column of $\widehat{g}(\beta)$, $j \in \{1,\ldots,2m\}$, and

$$V_i(Z_{N,\beta}) \doteq \partial V(Z_{N,\beta}) / \partial [Z_{N,\beta}]_i \in \mathbb{R}, \quad i \in \{1,\ldots,2n\}.$$

Applying Itô's lemma to $V_i(Z_{N,\beta}) \widehat{g}_{i,j}(\beta) \in \mathbb{R}$, $(i,j) \in \{1,\ldots,2n\}\times\{1,\ldots,2m\}$, and using the truncated dynamics in (47) we get

$$\begin{aligned} & d_\beta \left[V_i(Z_{N,\beta}) \widehat{g}_{i,j}(\beta)\right] \\ &= \left[V_i(Z_{N,\beta}) \dot{\widehat{g}}_{i,j}(\beta) + \left(\nabla V_i(Z_{N,\beta})^\top J_\mu(\beta, Z_{N,\beta}) + \frac{1}{2}\mathrm{Tr}\left[K_\sigma(\beta, Z_{N,\beta}) \nabla^2 V_i(Z_{N,\beta})\right]\right) \widehat{g}_{i,j}(\beta)\right] d\beta \\ &\quad + \nabla V_i(Z_{N,\beta})^\top J_\sigma(\beta, Z_{N,\beta}) \widehat{g}_{i,j}(\beta) dW_\beta, \end{aligned} \tag{D.26}$$

where we have replaced $J_{N,\mu}$ and $K_{N,\sigma} = J_{N,\sigma}J_{N,\sigma}^\top \in \mathbb{S}^{2n}$ with J_μ and K_σ because from Proposition 3.2, $Z_{N,\beta}$ is also a strong solution of the joint process (45) for all $\beta \in [\nu, \tau(t)] \subseteq [0, \tau^\star] \subseteq [0, \tau_N]$. See (49) for the definition of the stopping times $\tau^\star$ and τ_N. Since (D.25) implies that

$$d_\beta \left[\nabla V\left(Z_{N,\beta}\right)^\top \widehat{g}_{\cdot,j}(\beta)\right] = \sum_{i=1}^{2n} d_\beta \left[V_i\left(Z_{N,\beta}\right)^\top \widehat{g}_{i,j}(\beta)\right] \in \mathbb{R}, \quad j \in \{1, \ldots, 2m\},$$

we may substitute the expression in (D.26) to obtain

$$\begin{aligned} &d_\beta \left[\nabla V\left(Z_{N,\beta}\right)^\top \widehat{g}_{\cdot,j}(\beta)\right] \\ &= \nabla V\left(Z_{N,\beta}\right)^\top \dot{\widehat{g}}_{\cdot,j}(\beta)d\beta + \left(\nabla^2 V\left(Z_{N,\beta}\right) J_\mu\left(\beta, Z_{N,\beta}\right) + \frac{1}{2}\vec{\text{Tr}}\left[K_\sigma\left(\beta, Z_{N,\beta}\right)\nabla^2 V_i\left(Z_{N,\beta}\right)\right]_{i=1}^{2n}\right)^\top \widehat{g}_{\cdot,j}(\beta)d\beta \\ &\qquad + \widehat{g}_{\cdot,j}(\beta)^\top \nabla^2 V\left(Z_{N,\beta}\right) J_\sigma\left(\beta, Z_{N,\beta}\right) dW_\beta \in \mathbb{R}, \end{aligned}$$

for $j \in \{1, \ldots, 2m\}$. Once again from (D.25) we have that

$$\nabla V\left(Z_{N,\beta}\right)^\top \left(\mathbb{I}_2 \otimes g(\beta)\right) = \begin{bmatrix}\nabla V\left(Z_{N,\beta}\right)^\top \widehat{g}_{\cdot,1}(\beta) & \cdots & \nabla V\left(Z_{N,\beta}\right)^\top \widehat{g}_{\cdot,2m}(\beta)\end{bmatrix}.$$

It then follows from the previous expression that

$$\begin{aligned} &d_\beta \left[\nabla V\left(Z_{N,\beta}\right)^\top \left(\mathbb{I}_2 \otimes g(\beta)\right)\right] \\ &= \nabla V\left(Z_{N,\beta}\right)^\top \dot{\widehat{g}}(\beta)d\beta + \left(\nabla^2 V\left(Z_{N,\beta}\right) J_\mu\left(\beta, Z_{N,\beta}\right) + \frac{1}{2}\vec{\text{Tr}}\left[K_\sigma\left(\beta, Z_{N,\beta}\right)\nabla^2 V_i\left(Z_{N,\beta}\right)\right]_{i=1}^{2n}\right)^\top \widehat{g}(\beta)d\beta \\ &\qquad + \left(\widehat{g}(\beta)^\top \nabla^2 V\left(Z_{N,\beta}\right) J_\sigma\left(\beta, Z_{N,\beta}\right) dW_\beta\right)^\top \in \mathbb{R}^{1\times 2m}. \end{aligned} \tag{D.27}$$

Next, it follows from the defintion of J_μ in (45) that

$$J_\mu\left(\beta, Z_{N,\beta}\right) = \begin{bmatrix} F_\mu\left(\beta, X_{N,\beta}, U_{\mathcal{L}_1,\beta}\right) \\ F_\mu\left(\beta, X_{N,\beta}^r, U_\beta^r\right)\end{bmatrix} = \begin{bmatrix} \bar{F}_\mu\left(\beta, X_{N,\beta}\right) + g(\beta)U_{\mathcal{L}_1,\beta} + \Lambda_\mu\left(\beta, X_{N,\beta}\right) \\ F_\mu\left(\beta, X_{N,\beta}^r, U_\beta^r\right)\end{bmatrix} \in \mathbb{R}^{2n},$$

where we have used the decomposition (4) in Definition 1. It then follows from the definitions $U_{\mathcal{L}_1} \doteq \mathcal{F}_{\mathcal{L}_1}\left(X_N\right)$ and $U^r \doteq \mathcal{F}_r\left(X_N^r\right)$ that

$$J_\mu\left(\beta, Z_{N,\beta}\right) = \begin{bmatrix} \bar{F}_\mu\left(\beta, X_{N,\beta}\right) + g(\beta)\mathcal{F}_{\mathcal{L}_1}\left(X_N\right)(\beta) + \Lambda_\mu\left(\beta, X_{N,\beta}\right) \\ F_\mu\left(\beta, X_{N,\beta}^r, \mathcal{F}_r\left(X_N^r\right)(\beta)\right)\end{bmatrix} \in \mathbb{R}^{2n},$$

which, upon adding and subtracting $\mathcal{F}_r\left(X_N\right)$ yields

$$\begin{aligned} J_\mu\left(\beta, Z_{N,\beta}\right) &= \begin{bmatrix} \bar{F}_\mu\left(\beta, X_{N,\beta}\right) + g(\beta)\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)\left(X_N\right)(\beta) + g(\beta)\mathcal{F}_r\left(X_N\right)(\beta) + \Lambda_\mu\left(\beta, X_{N,\beta}\right) \\ F_\mu\left(\beta, X_{N,\beta}^r, \mathcal{F}_r\left(X_N^r\right)(\beta)\right)\end{bmatrix} \\ &= \begin{bmatrix} F_\mu\left(\beta, X_{N,\beta}, \mathcal{F}_r\left(X_N\right)(\beta)\right) \\ F_\mu\left(\beta, X_{N,\beta}^r, \mathcal{F}_r\left(X_N^r\right)(\beta)\right)\end{bmatrix} + \begin{bmatrix} g(\beta)\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)\left(X_N\right)(\beta) \\ 0_n\end{bmatrix} \in \mathbb{R}^{2n}. \end{aligned}$$

Appealing to the decomposition (4) in Definition 1 once again leads to

$$J_\mu\left(\beta, Z_{N,\beta}\right) = \left(\left[\bar{F}_\mu + \Lambda_\mu\right] \odot Z_N\right)(\beta) + \widehat{g}(\beta)\left(\mathcal{F}_r \odot Z_N\right)(\beta) + \begin{bmatrix} g(\beta)\left(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r\right)\left(X_N\right)(\beta) \\ 0_n\end{bmatrix} \in \mathbb{R}^{2n}. \tag{D.28}$$

Similarly, the defintion of J_σ in (45) implies that

$$J_\sigma\left(\beta, Z_{N,\beta}\right) \doteq \begin{bmatrix} F_\sigma\left(\beta, X_{N,\beta}\right) \\ F_\sigma\left(\beta, X_{N,\beta}^r\right)\end{bmatrix} = \left(\left[\bar{F}_\sigma + \Lambda_\sigma\right] \odot Z_N\right)(\beta) \in \mathbb{R}^{2n\times d}, \tag{D.29}$$

where we have once more used the decomposition (4) in Definition 1. Substituting (D.28) and (D.29) into (D.27) produces

$$d_\beta \left[\nabla V\left(Z_{N,\beta}\right)^\top \left(\mathbb{I}_2 \otimes g(\beta)\right)\right] = \left[\sum_{i=1}^{3} \mathcal{P}_{\mu_1}(\beta) + \mathcal{P}_{\mathcal{U}}(\beta) + \widetilde{\mathcal{P}}_{\mathcal{U}}(\beta)\right]^\top d\beta + \left(\mathcal{P}_\sigma(\beta)dW_\beta\right)^\top \in \mathbb{R}^{1\times 2m}.$$

Then, (D.22) is established by substituting the above into (D.21).

□

The next result establishes the bounds for the pertinent entities in the last proposition.

Proposition D.4 *Consider the functions $\mathcal{P}_{\mu_i}(t) \in \mathbb{R}^{2m}$, $i \in \{1,2,3\}$, and $\mathcal{P}_\sigma(t) \in \mathbb{R}^{2m\times d}$ defined in the statement of Proposition D.3. If the stopping time $\tau^\star$, defined in (49), Lemma 3.2, satisfies $\tau^\star = t^\star$, then*

$$\sum_{i=1}^{3} \left\| \left(\mathcal{P}_{\mu_i}\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \leq \sqrt{n}\Delta_g \left(\Delta_{\partial V}\Delta_{\mathcal{P}_\mu} + \frac{1}{\sqrt{2}}\Delta_{\partial^2 V}\Delta_{\mathcal{P}_\sigma}(4\mathsf{p}, 2\mathsf{p}) \right) + \sqrt{2}\Delta_{\dot{g}} \left\| \left(\nabla V\left(Z_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0}, \tag{D.30a}$$

$$\left\| \left(\mathcal{P}_\sigma\right)_{t^\star} \right\|_{\mathsf{q}}^{\pi_\star^0} \leq \sqrt{n}\Delta_g\Delta_{\partial V}\Delta_{\mathcal{P}_\sigma}(\mathsf{q},\mathsf{q}), \quad \mathsf{q} \in \{2\mathsf{p}, 4\mathsf{p}\}, \tag{D.30b}$$

where

$$\Delta_{\mathcal{P}_\mu} = \left\| \left(\bar{F}_\mu\left(\cdot, X_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\bar{F}_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\Lambda_\mu\left(\cdot, X_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\Lambda_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0},$$

$$\Delta_{\mathcal{P}_\sigma}(\mathsf{r},\mathsf{s}) = \mathbb{E}_{\pi_\star^0}\left[\left(\bar{F}_\sigma\left(\cdot, X_N\right)\right)_{t^\star}^{\mathsf{r}}\right]^{\frac{1}{\mathsf{s}}} + \mathbb{E}_{\pi_\star^0}\left[\left(\bar{F}_\sigma\left(\cdot, X_N^r\right)\right)_{t^\star}^{\mathsf{r}}\right]^{\frac{1}{\mathsf{s}}} + \mathbb{E}_{\pi_\star^0}\left[\left(\Lambda_\sigma\left(\cdot, X_N\right)\right)_{t^\star}^{\mathsf{r}}\right]^{\frac{1}{\mathsf{s}}} + \mathbb{E}_{\pi_\star^0}\left[\left(\Lambda_\sigma\left(\cdot, X_N^r\right)\right)_{t^\star}^{\mathsf{r}}\right]^{\frac{1}{\mathsf{s}}},$$

for $(\mathsf{r},\mathsf{s}) \in \{2\mathsf{p}, 4\mathsf{p}\} \times \{2\mathsf{p}, 4\mathsf{p}\}$.

Proof. we closely follow the proof of Proposition C.4. We begin with the term $\mathcal{P}_{\mu_1}$ defined in (D.23a), which we may bound as follows:

$$\left\|\mathcal{P}_{\mu_1}(t)\right\| \leq \left\|\mathbb{I}_2 \otimes \dot{g}(t)\right\|_F \left\|\nabla V\left(Z_{N,t}\right)\right\| \leq \sqrt{2}\Delta_{\dot{g}} \left\|\nabla V\left(Z_{N,t}\right)\right\|, \quad \forall t \in [0,T],$$

where we have used the bound on $\dot{g}(t)$ in Assumption 1. It then follows that

$$\left(\mathcal{P}_{\mu_1}\right)_{t^\star} \leq \sqrt{2}\Delta_{\dot{g}}\left(\nabla V\left(Z_N\right)\right)_{t^\star},$$

and thus

$$\left\| \left(\mathcal{P}_{\mu_1}\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \leq \sqrt{2}\Delta_{\dot{g}} \left\| \left(\nabla V\left(Z_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0}. \tag{D.31}$$

Similarly, using the bound on $g(t)$ in Assumption 1 and using the bound in (E.1b), Proposition E.1, we obtain

$$\begin{aligned}\left\|\mathcal{P}_{\mu_2}(t)\right\| &\leq \left\|\mathbb{I}_2 \otimes g(t)\right\|_F \left\|\nabla^2 V\left(Z_{N,t}\right)\right\| \left\|\left(\left[\bar{F}_\mu + \Lambda_\mu\right] \odot Z_N\right)(t)\right\| \\ &\leq \sqrt{n}\Delta_g\Delta_{\partial V} \left\| \begin{bmatrix} \bar{F}_\mu\left(t, X_{N,t}\right) + \Lambda_\mu\left(t, X_{N,t}\right) \\ \bar{F}_\mu\left(t, X_{N,t}^r\right) + \Lambda_\mu\left(t, X_{N,t}^r\right) \end{bmatrix} \right\|,\end{aligned}$$

and thus

$$\left\|\mathcal{P}_{\mu_2}(t)\right\| \leq \sqrt{n}\Delta_g\Delta_{\partial V}\left(\left\|\bar{F}_\mu\left(t, X_{N,t}\right)\right\| + \left\|\Lambda_\mu\left(t, X_{N,t}\right)\right\| + \left\|\bar{F}_\mu\left(t, X_{N,t}^r\right)\right\| + \left\|\Lambda_\mu\left(t, X_{N,t}^r\right)\right\|\right), \quad \forall t \in [0,T].$$

Therefore, we can conclude that

$$\left(\mathcal{P}_{\mu_2}\right)_{t^\star} \leq \sqrt{n}\Delta_g\Delta_{\partial V}\left(\left(\bar{F}_\mu\left(t, X_{N,t}\right)\right)_{t^\star} + \left(\Lambda_\mu\left(t, X_{N,t}\right)\right)_{t^\star} + \left(\bar{F}_\mu\left(t, X_{N,t}^r\right)\right)_{t^\star} + \left(\Lambda_\mu\left(t, X_{N,t}^r\right)\right)_{t^\star}\right) \tag{D.32}$$

It then follows due to the Minkowski's inequality that

$$\begin{aligned}\left\| \left(\mathcal{P}_{\mu_2}\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} &\leq \sqrt{n}\Delta_g\Delta_{\partial V}\left(\left\| \left(\bar{F}_\mu\left(\cdot, X_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\bar{F}_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \right) \\ &\quad + \sqrt{n}\Delta_g\Delta_{\partial V}\left(\left\| \left(\Lambda_\mu\left(\cdot, X_N\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} + \left\| \left(\Lambda_\mu\left(\cdot, X_N^r\right)\right)_{t^\star} \right\|_{2\mathsf{p}}^{\pi_\star^0} \right).\end{aligned} \tag{D.33}$$

Next, consider the term $\mathcal{P}_{\mu_3}$ defined in (D.23a), using which we obtain the following bound:

$$\left\|\mathcal{P}_{\mu_3}(t)\right\| \leq \frac{1}{\sqrt{2}}\Delta_g \left\| \vec{\mathrm{Tr}}\left[K_\sigma\left(t, Z_{N,t}\right)\nabla^2 V_i\left(Z_{N,t}\right)\right]_{i=1}^{2n} \right\| = \frac{1}{\sqrt{2}}\Delta_g \left(\sum_{i=1}^{2n} \left|\mathfrak{K}_i\left(t, Z_{N,t}\right)\right|^2 \right)^{\frac{1}{2}}, \tag{D.34}$$

where

$$\mathfrak{K}_i\left(t, Z_{N,t}\right) = \text{Tr}\left[K_\sigma\left(t, Z_{N,t}\right)\nabla^2 V_i\left(Z_{N,t}\right)\right] \in \mathbb{R}, \quad K_\sigma\left(t, Z_{N,t}\right) \doteq J_\sigma\left(t, Z_{N,t}\right) J_\sigma\left(t, Z_{N,t}\right)^\top \in \mathbb{S}^{2n}.$$

Following the reasoning that leads to (C.23) and using the definition of J_σ in (45), it can be shown that

$$\left|\mathfrak{K}_i\left(t, Z_{N,t}\right)\right| \leq \left\|\nabla^2 V_i\left(Z_{N,t}\right)\right\|_F \left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F^2 + \left\|F_\sigma\left(t, X^r_{N,t}\right)\right\|_F^2\right),$$

for all $(t, i) \in [0, T] \times \{1, \ldots, 2n\}$. Substituting the above bound into (D.34) then leads to

$$\begin{aligned}\left\|\mathcal{P}_{\mu_3}(t)\right\| &\leq \frac{1}{\sqrt{2}}\Delta_g \left(\sum_{i=1}^{2n}\left\|\nabla^2 V_i\left(Z_{N,t}\right)\right\|_F^2\right)^{\frac{1}{2}} \left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F^2 + \left\|F_\sigma\left(t, X^r_{N,t}\right)\right\|_F^2\right) \\ &\leq \frac{1}{\sqrt{2}}\Delta_g \left(\sum_{i=1}^{2n}\left\|\nabla^2 V_i\left(Z_{N,t}\right)\right\|_F\right) \left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F^2 + \left\|F_\sigma\left(t, X^r_{N,t}\right)\right\|_F^2\right),\end{aligned}$$

for all $t \in [0, T]$. Substituting the bound in (E.1c), Proposition E.1 produces

$$\left\|\mathcal{P}_{\mu_3}(t)\right\| \leq \sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial^2 V}\left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F^2 + \left\|F_\sigma\left(t, X^r_{N,t}\right)\right\|_F^2\right), \quad \forall t \in [0, T].$$

Consequently,

$$\left(\mathcal{P}_{\mu_3}\right)_{t^\star} \leq \sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial^2 V}\left(\left(F_\sigma\left(\cdot, X_N\right)\right)^2_{t^\star} + \left(F_\sigma\left(\cdot, X^r_N\right)\right)^2_{t^\star}\right).$$

By applying the Minkowski's inequality, it follows that

$$\left\|\left(\mathcal{P}_{\mu_3}\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi^0_\star} \leq \sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial^2 V}\left(\mathbb{E}_{\pi^0_\star}\left[\left(F_\sigma\left(\cdot, X_N\right)\right)^{4\mathsf{p}}_{t^\star}\right]^{\frac{1}{2\mathsf{p}}} + \mathbb{E}_{\pi^0_\star}\left[\left(F_\sigma\left(\cdot, X^r_N\right)\right)^{4\mathsf{p}}_{t^\star}\right]^{\frac{1}{2\mathsf{p}}}\right).$$

Then, using the decomposition $F_\sigma = \bar{F}_\sigma + \Lambda_\sigma$ in (4) and the Minkowski's inequality, we obtain

$$\begin{aligned}\left\|\left(\mathcal{P}_{\mu_3}\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi^0_\star} \leq{}& \sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial^2 V}\left(\mathbb{E}_{\pi^0_\star}\left[\left(\bar{F}_\sigma\left(\cdot, X_N\right)\right)^{4\mathsf{p}}_{t^\star}\right]^{\frac{1}{2\mathsf{p}}} + \mathbb{E}_{\pi^0_\star}\left[\left(\bar{F}_\sigma\left(\cdot, X^r_N\right)\right)^{4\mathsf{p}}_{t^\star}\right]^{\frac{1}{2\mathsf{p}}}\right) \\ &+ \sqrt{\frac{n}{2}}\Delta_g \Delta_{\partial^2 V}\left(\mathbb{E}_{\pi^0_\star}\left[\left(\Lambda_\sigma\left(\cdot, X_N\right)\right)^{4\mathsf{p}}_{t^\star}\right]^{\frac{1}{2\mathsf{p}}} + \mathbb{E}_{\pi^0_\star}\left[\left(\Lambda_\sigma\left(\cdot, X^r_N\right)\right)^{4\mathsf{p}}_{t^\star}\right]^{\frac{1}{2\mathsf{p}}}\right). \end{aligned} \tag{D.35}$$

Adding the bounds in (D.31), (D.33), and (D.35), establishes (D.30a).

Next, using the definition of $\mathcal{P}_\sigma$, we obtain

$$\left\|\mathcal{P}_\sigma(t)\right\|_F \leq \sqrt{n}\Delta_g \Delta_{\partial V}\left(\left\|F_\sigma\left(t, X_{N,t}\right)\right\|_F + \left\|F_\sigma\left(t, X^r_{N,t}\right)\right\|_F\right), \forall t \in [0, T].$$

Thus, the decomposition $F_\sigma = \bar{F}_\sigma + \Lambda_\sigma$ in (4) implies that

$$\begin{aligned}\left\|\mathcal{P}_\sigma(t)\right\|_F \leq \sqrt{n}\Delta_g \Delta_{\partial V}\Big(&\left\|\bar{F}_\sigma\left(t, X_{N,t}\right)\right\|_F + \left\|\bar{F}_\sigma\left(t, X^r_{N,t}\right)\right\|_F + \\ &\left\|\Lambda_\sigma\left(t, X_{N,t}\right)\right\|_F + \left\|\Lambda_\sigma\left(t, X^r_{N,t}\right)\right\|_F\Big), \quad \forall t \in [0, T],\end{aligned}$$

and hence,

$$\left(\mathcal{P}_\sigma\right)_{t^\star} \leq \sqrt{n}\Delta_g \Delta_{\partial V}\left(\left(\bar{F}_\sigma\left(\cdot, X_N\right)\right)_{t^\star} + \left(\bar{F}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star} + \left(\Lambda_\sigma\left(\cdot, X_N\right)\right)_{t^\star} + \left(\Lambda_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right). \tag{D.36}$$

Applying the Minkowski's inequality to (D.36) with $\mathsf{q} \in \{2\mathsf{p}, 4\mathsf{p}\}$, we can conclude that

$$\begin{aligned}\left\|\left(\mathcal{P}_\sigma\right)_{t^\star}\right\|_{\mathsf{q}}^{\pi^0_\star} \leq{}& \sqrt{n}\Delta_g \Delta_{\partial V}\left(\left\|\left(\bar{F}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|_{\mathsf{q}}^{\pi^0_\star} + \left\|\left(\bar{F}_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|_{\mathsf{q}}^{\pi^0_\star}\right) \\ &+ \sqrt{n}\Delta_g \Delta_{\partial V}\left(\left\|\left(\Lambda_\sigma\left(\cdot, X_N\right)\right)_{t^\star}\right\|_{\mathsf{q}}^{\pi^0_\star} + \left\|\left(\Lambda_\sigma\left(\cdot, X^r_N\right)\right)_{t^\star}\right\|_{\mathsf{q}}^{\pi^0_\star}\right),\end{aligned}$$

for $\mathsf{q} \in \{2\mathsf{p}, 4\mathsf{p}\}$, thus establishing (D.30b) and concluding the proof.

□

E Supporting Results

The following result establishes a few consequences of Assumption 2 that we utilize throughout the manuscript.

Proposition E.1 *The ILF $V \in \mathcal{C}^3(\mathbb{R}^n \times \mathbb{R}^n; \mathbb{R})$, defined in Assumption 2, satisfies the following bounds:*

$$\{\|\nabla V(a,b)\|, \|\nabla_a V(a,b)\|, \|\nabla_b V(a,b)\|\} \leq \|\nabla V(0,0)\| + \frac{1}{\sqrt{2}}\Delta_{\partial V}\|a-b\| \tag{E.1a}$$
$$\leq \|\nabla V(0,0)\| + \frac{1}{\sqrt{2\alpha_1}}\Delta_{\partial V} V(a,b)^{\frac{1}{2}},$$
$$\left\{\left\|\nabla_a^2 V(a,b)\right\|_F, \left\|\nabla_b^2 V(a,b)\right\|_F, \left\|\nabla_{a,b}^2 V(a,b)\right\|_F\right\} \leq \sqrt{\frac{n}{2}}\Delta_{\partial V}, \tag{E.1b}$$
$$\left\{\sum_{i=1}^n \left\|\nabla_a^2 V_{a_i}(a,b)\right\|_F, \sum_{i=1}^n \left\|\nabla_b^2 V_{b_i}(a,b)\right\|_F, \sum_{i=1}^n \left\|\nabla_a^2 V_{b_i}(a,b)\right\|_F, \sum_{i=1}^n \left\|\nabla_b^2 V_{a_i}(a,b)\right\|_F\right\} \leq \sqrt{\frac{n}{2}}\Delta_{\partial^2 V}, \tag{E.1c}$$

for all $a, b \in \mathbb{R}^n$.

Proof. Setting $a' = b' = 0_n$ in (10a), give us

$$\sum_{i=1}^n \left(|V_{a_i}(a,b) - V_{a_i}(0,0)|^2 + |V_{b_i}(a,b) - V_{b_i}(0,0)|^2\right) \leq \Delta_{\partial V}^2 \left|\begin{bmatrix} a \\ b \end{bmatrix}\right|_{\Delta_{2n}}^2 = \frac{1}{2}\Delta_{\partial V}^2 \|a-b\|^2,$$

where we have used [49, Lem. 2.3]. The above inequality can be re-written as

$$\|\nabla V(a,b) - \nabla V(0,0)\|^2 \leq \frac{1}{2}\Delta_{\partial V}^2\|a-b\|^2 \Rightarrow \|\nabla V(a,b) - \nabla V(0,0)\| \leq \frac{1}{\sqrt{2}}\Delta_{\partial V}\|a-b\|,$$

which, upon using the reverse triangle inequality yields

$$\|\nabla V(a,b)\| - \|\nabla V(0,0)\| \leq \|\nabla V(a,b) - \nabla V(0,0)\| \leq \frac{1}{\sqrt{2}}\Delta_{\partial V}\|a-b\|.$$

Thus, this bound, along with the fact that $\{\|\nabla_a V(a,b)\|, \|\nabla V_b(a,b)\|\} \leq \|\nabla V(a,b)\|$ establishes the penultimate bound in (E.1a). The ultimate bound in (E.1a) is established by invoking (9) in Assumption 2.

Now, for any $\delta \in \mathbb{R}$, (10a) and [49, Lem. 2.3] imply

$$\sum_{i=1}^n (V_{a_i}(a+\delta e_j, b) - V_{a_i}(a,b))^2 \leq \frac{1}{2}\Delta_{\partial V}^2\delta^2 \Rightarrow \sum_{i=1}^n \left(\frac{V_{a_i}(a+\delta e_j,b) - V_{a_i}(a,b)}{\delta}\right)^2 \leq \frac{1}{2}\Delta_{\partial V}^2,$$

where $e_j \in \mathbb{R}^n$ is the j^{th} canonical basis vector for $\mathbb{R}^n$. Thus, we conclude

$$\sum_{i=1}^n \lim_{\delta\to 0}\left(\frac{V_{a_i}(a+\delta e_j,b) - V_{a_i}(a,b)}{\delta}\right)^2 = \sum_{i=1}^n \left(\lim_{\delta\to 0}\frac{V_{a_i}(a+\delta e_j,b) - V_{a_i}(a,b)}{\delta}\right)^2 = \sum_{i=1}^n \left(V_{a_i,a_j}(a,b)\right)^2 \leq \frac{1}{2}\Delta_{\partial V}^2,$$

and thus

$$\left\|\nabla_a^2 V(a,b)\right\|_F^2 = \sum_{j=1}^n \left(\sum_{i=1}^n \left(V_{a_i,a_j}(a,b)\right)^2\right) \leq \frac{n}{2}\Delta_{\partial V}^2.$$

The bounds for $\left\|\nabla_b^2 V(a,b)\right\|_F^2$ and $\left\|\nabla_{a,b}^2 V(a,b)\right\|_F^2$ are established similarly.

Finally, the identities in (E.1c) are established *mutatis mutandis* by using (10b). □

Next, we compute the individual constituent parts for the terms $\Delta_{\Xi_i}^r$ and $\Delta_{\mathcal{U}_j}^r$, $(i,j) \in \{1,2\} \times \{1,2,3\}$, in Lemmas C.1 and C.2, respectively.

Proposition E.2 *Suppose there exists a strictly positive $\varrho \in \mathbb{R}_{>0}$ such that*

$$\mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X^r_{N,t}-X^\star_{N,t}\right\|^{2\mathsf{p}^\star}\right] \leq \varrho^{2\mathsf{p}^\star}, \tag{E.2}$$

where the constant $t^\star$ is defined in (29) and $\mathsf{p}^\star$ is defined in Assumption 3. Then, the following bound holds $\forall\, \mathbb{N}_{\geq 1} \ni \mathsf{p} \leq \mathsf{p}^\star$:

$$\left\{\left\|\left(V_r\left(Y_N\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}, \left\|\left(V_\star\left(Y_N\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}\right\} \leq \|\nabla V(0,0)\| + \frac{1}{\sqrt{2}}\Delta_{\partial V}\varrho. \tag{E.3}$$

Furthermore, the drift vector field satisfies the following bounds $\forall\, \mathbb{N}_{\geq 1} \ni \mathsf{p} \leq \mathsf{p}^\star$:

$$\mathbb{E}_{\pi_\star^0}\left[\left(\Lambda_\mu^{\{\cdot,\parallel,\perp\}}(\cdot,X^r_N)\right)_{t^\star}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \leq \left(\Delta_\mu^{\{\cdot,\parallel,\perp\}}\right)^2\left(1+2\Delta_\star^2\right) + 2\left(\Delta_\mu^{\{\cdot,\parallel,\perp\}}\right)^2\varrho^2, \tag{E.4a}$$

$$\left\|\left(\Lambda_\mu^{\{\cdot,\parallel,\perp\}}(\cdot,X^r_N)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \leq \Delta_\mu^{\{\cdot,\parallel,\perp\}}\left(1+\Delta_\star\right) + \Delta_\mu^{\{\cdot,\parallel,\perp\}}\varrho, \tag{E.4b}$$

$$\left\|\left(\bar{F}_\mu\left(\cdot,X^r_N\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \leq \Delta_f\left(1+\Delta_\star\right) + \Delta_f\varrho, \quad \left\|\left(\bar{F}_\mu\left(\cdot,X^\star_N\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0} \leq \Delta_f\left(1+\Delta_\star\right). \tag{E.4c}$$

Finally, the diffusion vector field satisfies the following bounds $\forall\, \mathbb{N}_{\geq 1} \ni \mathsf{p} \leq \mathsf{p}^\star$:

$$\mathbb{E}_{\pi_\star^0}\left[\left(F_\sigma^{\{\cdot,\parallel,\perp\}}(\cdot,X^r_N)\right)_{t^\star}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \leq \left(\Delta_p^{\{\cdot,\parallel,\perp\}}\right)^2 + \left(\Delta_\sigma^{\{\cdot,\parallel,\perp\}}\right)^2\left(1+\Delta_\star\right) + \left(\Delta_\sigma^{\{\cdot,\parallel,\perp\}}\right)^2\varrho, \tag{E.5a}$$

$$\begin{aligned}\left\{\left\|\left(F_\sigma^{\{\cdot,\parallel,\perp\}}(\cdot,X^r_N)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}, \left\|\left(F_\sigma^{\{\cdot,\parallel,\perp\}}(\cdot,X^r_N)\right)_{t^\star}\right\|_{4\mathsf{p}}^{\pi_\star^0}\right\} &\leq \Delta_p^{\{\cdot,\parallel,\perp\}} + \Delta_\sigma^{\{\cdot,\parallel,\perp\}}\left(1+\Delta_\star\right)^{\frac{1}{2}} \\ &\quad + \Delta_\sigma^{\{\cdot,\parallel,\perp\}}\sqrt{\varrho},\end{aligned} \tag{E.5b}$$

$$\left\{\left\|\left(\Lambda_\sigma^{\{\cdot,\parallel,\perp\}}(\cdot,X^r_N)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}, \left\|\left(\Lambda_\sigma^{\{\cdot,\parallel,\perp\}}(\cdot,X^r_N)\right)_{t^\star}\right\|_{4\mathsf{p}}^{\pi_\star^0}\right\} \leq \Delta_\sigma^{\{\cdot,\parallel,\perp\}}\left(1+\Delta_\star\right)^{\frac{1}{2}} + \Delta_\sigma^{\{\cdot,\parallel,\perp\}}\sqrt{\varrho}, \tag{E.5c}$$

$$\left\|\left(Tr\left[H_\sigma\left(\cdot,Y_N\right)\nabla^2 V\left(Y_N\right)\right]\right)_{t^\star}\right\|_{\mathsf{p}}^{\pi_\star^0} \leq \sqrt{\frac{n}{2}}\Delta_{\partial V}\left(2\Delta_p^2+\Delta_\sigma^2+\Delta_\sigma^2\Delta_\star\right) + \sqrt{\frac{n}{2}}\Delta_{\partial V}\Delta_\sigma^2\varrho, \tag{E.5d}$$

$$\left\{\left\|\left(\bar{F}_\sigma\left(\cdot,X^\star_N\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}, \left\|\left(\bar{F}_\sigma\left(\cdot,X^r_N\right)\right)_{t^\star}\right\|_{2\mathsf{p}}^{\pi_\star^0}, \left\|\left(\bar{F}_\sigma\left(\cdot,X^\star_N\right)\right)_{t^\star}\right\|_{4\mathsf{p}}^{\pi_\star^0}, \left\|\left(\bar{F}_\sigma\left(\cdot,X^r_N\right)\right)_{t^\star}\right\|_{4\mathsf{p}}^{\pi_\star^0}\right\} \leq \Delta_p. \tag{E.5e}$$

Proof. Using (E.1a) in Proposition E.1 one sees that

$$\left\{\left\|V_r\left(Y_{N,t}\right)\right\|, \left\|V_\star\left(Y_{N,t}\right)\right\|\right\} \leq \|\nabla V(0,0)\| + \frac{1}{\sqrt{2}}\Delta_{\partial V}\left\|X^r_{N,t}-X^\star_{N,t}\right\|, \quad \forall t\in[0,T],$$

which implies

$$\left\{\left(V_r\left(Y_N\right)\right)_{t^\star}, \left(V_\star\left(Y_N\right)\right)_{t^\star}\right\} \leq \|\nabla V(0,0)\| + \frac{1}{\sqrt{2}}\Delta_{\partial V}\sup_{t\in[0,t^\star]}\left\|X^r_{N,t}-X^\star_{N,t}\right\|.$$

Then (E.3) follows from (C.64).

Next, Assumption 4 implies that $\forall t \in [0,T]$,

$$\begin{aligned}\left\|\Lambda_\mu^{\{\parallel,\perp\}}(t,X^r_{N,t})\right\|^2 &\leq \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2\left(1+\left\|X^r_{N,t}\right\|^2\right) \\ &= \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2 + \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2\left(\left\|X^\star_{N,t}\right\|+\left\|X^r_{N,t}-X^\star_{N,t}\right\|\right)^2,\end{aligned}$$

and thus

$$\left(\Lambda_\mu^{\{\parallel,\perp\}}(\cdot,X^r_N)\right)_{t^\star}^2 \leq \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2 + \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2\left(\sup_{t\in[0,t^\star]}\left\|X^\star_{N,t}\right\| + \sup_{t\in[0,t^\star]}\left\|X^r_{N,t}-X^\star_{N,t}\right\|\right)^2.$$

Hence, the p^{th}-norm of the above inequality with respect to the measure $\pi_\star^0$ is as as follows:

$$\begin{aligned}
&\mathbb{E}_{\pi_\star^0}\left[\left(\Lambda_\mu^{\{\parallel,\perp\}}(\cdot, X_N^r)\right)_{t^\star}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \\
&\leq \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2 + \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2 \mathbb{E}_{\pi_\star^0}\left[\left(\sup_{t\in[0,t^\star]}\left\|X_{N,t}^\star\right\| + \sup_{t\in[0,t^\star]}\left\|X_{N,t}^r - X_{N,t}^\star\right\|\right)^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \\
&\leq \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2 + \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2 \left(\mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^\star\right\|^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}} + \mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^r - X_{N,t}^\star\right\|^{2\mathsf{p}}\right]^{\frac{1}{2\mathsf{p}}}\right)^2, \qquad \text{(E.6)}
\end{aligned}$$

where the last inequality is the consequence of the Minkowski's inequality. Now, Proposition (3.1) establishes that $X_{N,t}^\star = X_t^\star$, for all $t \in [0, \tau_N]$, where $X_t^\star$ is the unique strong solution of (5b). Then, since $\tau^\star = t^\star \leq \tau_N$, Assumption (3) and Hölder's inequality imply that

$$\mathbb{E}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^\star\right\|^{2\mathsf{p}}\right] \leq \Delta_\star^{2\mathsf{p}}, \quad \forall\, \mathbb{N}_{\geq 1} \ni \mathsf{p} \leq \mathsf{p}^\star. \qquad \text{(E.7)}$$

Substituting the bounds in (E.7) and (C.64) into (E.6) leads to

$$\mathbb{E}_{\pi_\star^0}\left[\left(\Lambda_\mu^{\{\parallel,\perp\}}(\cdot, X_N^r)\right)_{t^\star}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \leq \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2 + \left(\Delta_\mu^{\{\parallel,\perp\}}\right)^2 (\Delta_\star + \varrho)^2. \qquad \text{(E.8)}$$

Then, we obtain (E.4a) by applying [110, Prop. 3.1.10-(iii)] to the right hand side of the above inequality. Furthermore, (E.4b) is obtained from (E.8) by using the subadditivity of the square root operator.

Using the definition $\bar{F}_\mu = f$ in (3) and the bound f in Assumptions 2.4, we use the same approach as above to obtain (E.4c).

Next, Definition 4, followed by Assumptions 4 and 5, lead to the following:

$$\begin{aligned}
\left\|F_\sigma^{\{\parallel,\perp\}}(t, X_{N,t}^r)\right\|_F^2 &= \left\|p^{\{\parallel,\perp\}}(t, X_{N,t}^r)\right\|_F^2 + \left\|\Lambda_\sigma^{\{\parallel,\perp\}}(t, X_{N,t}^r)\right\|_F^2 \\
&\leq \left(\Delta_p^{\{\parallel,\perp\}}\right)^2 + \left(\Delta_\sigma^{\{\parallel,\perp\}}\right)^2 \left(1 + \left\|X_{N,t}^r\right\|^2\right)^{\frac{1}{2}}, \quad \forall t \in [0, T],
\end{aligned}$$

and thus, the subadditivity of the square root operator implies that

$$\begin{aligned}
\left(F_\sigma^{\{\parallel,\perp\}}(\cdot, X_N^r)\right)_{t^\star}^2 &\leq \left(\Delta_p^{\{\parallel,\perp\}}\right)^2 + \left(\Delta_\sigma^{\{\parallel,\perp\}}\right)^2 \left(1 + \sup_{t\in[0,t^\star]}\left\|X_{N,t}^r\right\|\right) \\
&\leq \left(\Delta_p^{\{\parallel,\perp\}}\right)^2 + \left(\Delta_\sigma^{\{\parallel,\perp\}}\right)^2 \left(1 + \sup_{t\in[0,t^\star]}\left\|X_{N,t}^\star\right\| + \sup_{t\in[0,t^\star]}\left\|X_{N,t}^r - X_{N,t}^\star\right\|\right). \qquad \text{(E.9)}
\end{aligned}$$

Hence, taking the p^{th}-norm of the above inequality with respect to the measure $\pi_\star^0$, and using the Minkowski's inequality, we obtain

$$\begin{aligned}
&\mathbb{E}_{\pi_\star^0}\left[\left(F_\sigma^{\{\parallel,\perp\}}(\cdot, X_N^r)\right)_{t^\star}^{2\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} \\
&\leq \left(\Delta_p^{\{\parallel,\perp\}}\right)^2 + \left(\Delta_\sigma^{\{\parallel,\perp\}}\right)^2 \left(1 + \mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^\star\right\|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^r - X_{N,t}^\star\right\|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}}\right). \qquad \text{(E.10)}
\end{aligned}$$

Now, using Jensen's inequality and the bounds in (C.64) and (E.7), we deduce that

$$\mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^\star\right\|^{\mathsf{p}}\right] \leq \mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^\star\right\|^{2\mathsf{p}}\right]^{\frac{1}{2}} \leq \Delta_\star^{\mathsf{p}}, \qquad \text{(E.11a)}$$

$$\mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^r - X_{N,t}^\star\right\|^{\mathsf{p}}\right] \leq \mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^r - X_{N,t}^\star\right\|^{2\mathsf{p}}\right]^{\frac{1}{2}} \leq \varrho^{\mathsf{p}}. \qquad \text{(E.11b)}$$

Substituting these bounds into (E.10) yields (E.5a). Furthermore, the 2p-norm in (E.5b) is obtained by using the subadditivity of the square root operator on (E.5a).

The 4p-norm in (E.5b) is obtained by taking the 2pth-norm of the inequality in (E.9) with respect to the measure $\pi_\star^0$, and following identical steps as above.

Inequality (E.5c) is a direct consequence of (E.5b) and(**??**) and the decomposition $F_\sigma = \bar{F}_\sigma + \Lambda_\sigma$ in (4).

Next, we see that by using an identical line of reasoning that produced (C.23) in the proof of Proposition C.4 and the definition of G_σ in (25), we get that

$$\text{Tr}\left[H_\sigma\left(t, Y_N\right)\nabla^2 V\left(Y_{N,t}\right)\right] \leq \left\|\nabla^2 V\left(Y_{N,t}\right)\right\|_F \left(\left\|F_\sigma\left(t, X_{N,t}^r\right)\right\|_F^2 + \left\|\bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\|_F^2\right), \quad \forall t \in [0,T],$$

which, upon using (E.1b) in Proposition E.1, yields

$$\text{Tr}\left[H_\sigma\left(t, Y_N\right)\nabla^2 V\left(Y_{N,t}\right)\right] \leq \sqrt{\frac{n}{2}}\Delta_{\partial V}\left(\left\|F_\sigma\left(t, X_{N,t}^r\right)\right\|_F^2 + \left\|\bar{F}_\sigma\left(t, X_{N,t}^\star\right)\right\|_F^2\right), \quad \forall t \in [0,T].$$

It then follows from the decomposition $F_\sigma = \bar{F}_\sigma + \Lambda_\sigma$ in (4) that

$$\text{Tr}\left[H_\sigma\left(t, Y_N\right)\nabla^2 V\left(Y_{N,t}\right)\right] \leq \sqrt{\frac{n}{2}}\Delta_{\partial V}\left(\left\|p\left(t, X_{N,t}^r\right)\right\|_F^2 + \left\|p\left(t, X_{N,t}^\star\right)\right\|_F^2 + \left\|\Lambda_\sigma\left(t, X_{N,t}^r\right)\right\|_F^2\right), \quad \forall t \in [0,T],$$

where we have used the definition $\bar{F}_\sigma = p$ in (3). Using the bounds on p and Λ_σ in Assumptions 2.4 and 4, respectively, we obtain

$$\begin{aligned}
\text{Tr}\left[H_\sigma\left(t, Y_N\right)\nabla^2 V\left(Y_{N,t}\right)\right] &\leq \sqrt{\frac{n}{2}}\Delta_{\partial V}\left(2\Delta_p^2 + \Delta_\sigma^2\left(1 + \left\|X_{N,t}^r\right\|^2\right)^{\frac{1}{2}}\right) \\
&\leq \sqrt{\frac{n}{2}}\Delta_{\partial V}\left(2\Delta_p^2 + \Delta_\sigma^2 + \Delta_\sigma^2\left\|X_{N,t}^r\right\|\right) \\
&\leq \sqrt{\frac{n}{2}}\Delta_{\partial V}\left(2\Delta_p^2 + \Delta_\sigma^2 + \Delta_\sigma^2\left\|X_{N,t}^\star\right\| + \Delta_\sigma^2\left\|X_{N,t}^r - X_{N,t}^\star\right\|\right), \quad \forall t \in [0,T].
\end{aligned}$$

Hence, we conclude that

$$\left(\text{Tr}\left[H_\sigma\left(\cdot, Y_N\right)\nabla^2 V\left(Y_N\right)\right]\right)_{t^\star} \leq \sqrt{\frac{n}{2}}\Delta_{\partial V}\left(2\Delta_p^2 + \Delta_\sigma^2 + \Delta_\sigma^2\sup_{t\in[0,t^\star]}\left\|X_{N,t}^\star\right\| + \Delta_\sigma^2\sup_{t\in[0,t^\star]}\left\|X_{N,t}^r - X_{N,t}^\star\right\|\right).$$

Taking the pth-norm of the above inequality with respect to the measure $\pi_\star^0$, and using the Minkowski's inequality, we obtain

$$\begin{aligned}
&\left\|\left(\text{Tr}\left[H_\sigma\left(\cdot, Y_N\right)\nabla^2 V\left(Y_N\right)\right]\right)_{t^\star}\right\|_{\mathsf{p}}^{\pi_\star^0} \\
&\quad \leq \sqrt{\frac{n}{2}}\Delta_{\partial V}\left(2\Delta_p^2 + \Delta_\sigma^2 + \Delta_\sigma^2\mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^\star\right\|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}} + \Delta_\sigma^2\mathbb{E}_{\pi_\star^0}\left[\sup_{t\in[0,t^\star]}\left\|X_{N,t}^r - X_{N,t}^\star\right\|^{\mathsf{p}}\right]^{\frac{1}{\mathsf{p}}}\right).
\end{aligned}$$

Then, substituting the bounds (E.11) into the above inequality leads to (E.5d).

Finally, using the definition $\bar{F}_\sigma = p$ in (3) and the bound on p in Assumptions 2.4 yields (E.5e). □

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 adaptive control," in *Robust and adaptive control: With aerospace applications*. Springer, 2012, pp. 317–353.

[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] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.

[19] N. Boffi, S. Tu, N. Matni, J.-J. Slotine, and V. Sindhwani, "Learning stability certificates from data," in *Conference on Robot Learning*. PMLR, 2021, pp. 1341–1350.

[20] J. Quinonero-Candela, M. Sugiyama, A. Schwaighofer, and N. D. Lawrence, *Dataset shift in machine learning*. Mit Press, 2008.

[21] M. Sugiyama and M. Kawanabe, *Machine learning in non-stationary environments: Introduction to covariate shift adaptation*. MIT press, 2012.

[22] O. Wiles, S. Gowal, 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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. 10 920–10 926.

[27] D. Pfrommer, 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.

[28] C. Villani, *Optimal transport: Old and new*. Springer, 2009, vol. 338.

[29] N. Hovakimyan and C. Cao, *$\mathcal{L}_1$ adaptive control theory: Guaranteed robustness with fast adaptation*. SIAM, 2010.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] E. D. Sontag, "A Lyapunov-like characterization of asymptotic controllability," *SIAM journal on control and optimization*, vol. 21, no. 3, pp. 462–471, 1983.

[41] H. K. Khalil, *Nonlinear systems*. London, UK: Pearson, 2002.

[42] 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.

[43] 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.

[44] H. Deng and M. Krstić, "Stochastic nonlinear stabilization—I: A backstepping design," *Systems & Control Letters*, vol. 32, no. 3, pp. 143–150, 1997.

[45] H. Deng and M. Krstic, "Output-feedback stochastic nonlinear stabilization," *IEEE Transactions on Automatic Control*, vol. 44, no. 2, pp. 328–333, 1999.

[46] H. J. Kushner, *Stochastic stability and control*. Academic Press, 1967.

[47] R. Khasminskii, *Stochastic stability of differential equations*. Springer Science & Business Media, 2011, vol. 66.

[48] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for non-linear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.

[49] D. Angeli, "A Lyapunov approach to incremental stability properties," *IEEE Transactions on Automatic Control*, vol. 47, no. 3, pp. 410–421, 2002.

[50] 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.

[51] 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.

[52] 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.

[53] 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.

[54] 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.

[55] 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.

[56] 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.

[57] 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.

[58] 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.

[59] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*. Springer Science & Business Media, 1991, vol. 113.

[60] S. Richards, N. Azizan, J.-J. Slotine, and M. Pavone, "Adaptive-control-oriented meta-learning for nonlinear systems," in *Robotics science and systems*, 2021.

[61] 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.

[62] 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.

[63] 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.

[64] 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.

[65] D. Applebaum, *Lévy processes and stochastic calculus*. Cambridge university press, 2009.

[66] 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.

[67] ——, "Scenario-based probabilistic reachable sets for recursively feasible stochastic model predictive control," *IEEE Control Systems Letters*, vol. 4, no. 2, pp. 450–455, 2019.

[68] M. Fiacchini and T. Alamo, "Probabilistic reachable and invariant sets for linear systems with correlated disturbance," *Automatica*, vol. 132, p. 109808, 2021.

[69] A. P. Vinod and M. M. Oishi, "Stochastic reachability of a target tube: Theory and computation," *Automatica*, vol. 125, p. 109458, 2021.

[70] K. Sivaramakrishnan, V. Sivaramakrishnan, and M. M. Oishi, "Stochastic reachability of discrete-time stochastic systems via probability measures," *arXiv preprint arXiv:2304.00598*, 2023.

[71] 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.

[72] H. Rahimian and S. Mehrotra, "Distributionally robust optimization: A review," *arXiv preprint arXiv:1908.05659*, 2019.

[73] 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.

[74] 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.

[75] 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.

[76] 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.

[77] 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.

[78] I. Yang, "Wasserstein distributionally robust stochastic control: A data-driven approach," *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3863–3870, 2020.

[79] A. Hakobyan and I. Yang, "Distributionally robust differential dynamic programming with Wasserstein distance," *IEEE Control Systems Letters*, 2023.

[80] ——, "Wasserstein distributionally robust control of partially observable linear stochastic systems," *IEEE Transactions on Automatic Control (to appear), arXiv preprint arXiv:2212.04644*, 2022.

[81] 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.

[82] 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.

[83] 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.

[84] 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.

[85] 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.

[86] K. S. Bakshi, D. D. Fan, and E. A. Theodorou, "Stochastic control of systems with control multiplicative noise using second order FBSDEs," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 424–431.

[87] M. A. Pereira, Z. Wang, T. Chen, E. Reed, and E. A. Theodorou, "Deep 2FBSDEs for systems with control multiplicative noise," *arXiv preprint arXiv:1906.04762*, 2019.

[88] E. Todorov and W. Li, "A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems," in *2005 American Control Conference (ACC)*. IEEE, 2005, pp. 300–306.

[89] D. Mitrovic, S. Klanke, R. Osu, M. Kawato, and S. Vijayakumar, "A computational model of limb impedance control based on principles of internal model uncertainty," *PloS one*, vol. 5, no. 10, p. e13601, 2010.

[90] J. A. Primbs, "Portfolio optimization applications of stochastic receding horizon control," in *2007 American Control Conference (ACC)*. IEEE, 2007, pp. 1811–1816.

[91] P. McLane, "Optimal stochastic control of linear systems with state-and control-dependent disturbances," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 793–798, 1971.

[92] Y. Phillis, "Controller design of systems with multiplicative noise," *IEEE Transactions on Automatic Control*, vol. 30, no. 10, pp. 1017–1019, 1985.

[93] R. Durrett, *Stochastic calculus: A practical introduction*. CRC press, 1996, vol. 6.

[94] S. Axler, *Measure, integration & real analysis*. Springer Nature, 2020.

[95] L. C. Evans, *An introduction to stochastic differential equations*. American Mathematical Soc., 2012, vol. 82.

[96] R. F. Curtain and H. Zwart, *An introduction to infinite-dimensional linear systems theory*. Springer Science & Business Media, 2012, vol. 21.

[97] B. E. Fristedt and L. F. Gray, *A modern approach to probability theory*. Springer Science & Business Media, 2013.

[98] A. V. Balakrishnan, *Stochastic Differential Systems I: Filtering and Control-A Function Space Approach*. Springer Science & Business Media, 2012, vol. 84.

[99] 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.

[100] X. Mao, *Stochastic differential equations and applications*. Elsevier, 2007.

[101] B. Yi, R. Wang, and I. R. Manchester, "On necessary conditions of tracking control for nonlinear systems via contraction analysis," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 2000–2005.

[102] 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.

[103] 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.

[104] L. C. Evans, *Partial differential equations*. American Mathematical Society, 2022, vol. 19.

[105] H. Thorisson, "Coupling methods in probability theory," *Scandinavian journal of statistics*, pp. 159–182, 1995.

[106] F. E. Burk, *A garden of integrals*. American Mathematical Soc., 2007, vol. 31.

[107] Y.-F. Ren, "On the Burkholder–Davis–Gundy inequalities for continuous martingales," *Statistics & probability letters*, vol. 78, no. 17, pp. 3034–3039, 2008.

[108] H. Kuo, *Introduction to Stochastic Integration*. Springer New York, 2006.

[109] F. Zhang, *Matrix theory: Basic results and techniques*. Springer Science & Business Media, 2011.

[110] K. B. Athreya and S. N. Lahiri, *Measure theory and probability theory*. Springer, 2006, vol. 19.