arXiv:submit/3363819 [math.OC] 11 Sep 2020

SEQUENTIAL CONVEX PROGRAMMING FOR NON-LINEAR STOCHASTIC OPTIMAL CONTROL

R. BONALLI, T. LEW, M. PAVONE*

Abstract. We introduce a sequential convex programming framework to solve general non-linear stochastic optimal control problems in finite dimension, where uncertainties are modeled by a multidimensional Wiener process. We provide sufficient conditions for the convergence of the method. Moreover, we prove that, when convergence is achieved, sequential convex programming finds a candidate locally optimal solution for the original problem in the sense of the stochastic Pontryagin Maximum Principle. We leverage those properties to design a practical numerical method to solve non-linear stochastic optimal control problems that is based on a deterministic transcription of stochastic sequential convex programming.

1. Introduction. Through the last decades, the applied control community has focused more and more attention to the optimal control of stochastic systems. The general formulation consists of steering a dynamical system, that is affected by uncertainties which are modeled via Wiener processes, from an initial condition to a final configuration by optimizing some prescribed performance criterion, while satisfying specific constraints. Uncertainties may come from unmodeled and/or unpredicted behaviors, as well as from measurement errors. The literature on the subject is rich, yet still at the height of its theoretical and numerical developments, and we can roughly classify the existing works into two main categories.

The first group is composed by contributions that focus on Linear Convex Problems (LCPs), i.e., whose dynamics are linear and costs are convex, both in control and state variables. An important class of LCPs is given by Linear Quadratic Problems (LQPs), whereby costs are quadratic in both control and state variables, and for which the analysis of optimal solutions may be reduced to the study of some, often handier, algebraic relation known as Stochastic Riccati Equation (SRE) [31, 8, 29, 35]. Quite efficient algorithmic frameworks have been conceived to numerically solve LCPs, that range from local search [17] and duality [33, 18], to deterministic-equivalent reformulation [6, 4], to name a few. In the special case of LQPs, those techniques may be further improved by thoroughly combining SRE theory with semidefinite programming [32, 38], finite-dimensional approximation [5, 14] or chaos expansion [23].

The second group of works deals with problems that do not enjoy any specific regularity, in particular, non-linear dynamics and non-convex (therefore non-quadratic) costs are allowed. Throughout this paper, we call those Non-Linear Problems (NLPs). It is unquestionable that NLPs have so far received less attention than LCPs from the community, especially because the analysis of the former is usually more involved. Similarly to what happens in the deterministic case, there are essentially two main theoretical tools that have been developed to analyze NLPs: stochastic Dynamic Programming (DP) [3, 25] and the stochastic Pontryagin Maximum Principle (PMP) [30, 21, 28] (an extensive survey of generalizations of DP and PMP may be found in [39]). Whenever LQPs are considered, one can show that DP and PMP are equivalent to SRE [39]. DP provides optimal policies through the resolution of a partial differential equation, whereas the necessary conditions for optimality offered by the PMP allow to set up a two-point boundary value problem that, if solved, returns candidate locally optimal solutions. Both methods lead to analytical solutions only

*The authors are with the Department of Aeronautics and Astronautics, Stanford University, Stanford, CA 94305. {rbonalli, thomas.lew, pavone}@stanford.edu.

in very few cases, and complex hindrances may quickly appear in the numerical context (the stochastic setting is even more vicious than the deterministic one, the latter being better understood for a quite wide range of problems, see, e.g., [37, 11]). This has fostered the investigation of more tractable approaches to solve NLPs, such as Monte Carlo simulation [34, 16], Markov chain discretization [19, 20] and deterministic (though non-equivalent) reformulation [2, 4], among others. Often based on some sort of approximation of the original formulation, such techniques clearly offer powerful alternatives, especially because they may benefit from practical implementation and may be proved to converge to policies satisfying DP or PMP (see, e.g., [20]).

This last remark motivates the present work. Specifically, our objective consists of introducing and analyzing a general framework to compute candidate optimal solutions for a specific class of NLPs, which is based on Sequential Convex Programming (SCP) [27, 9], probably among the most known (and earliest) approximating techniques in deterministic non-linear optimal control. The simplest SCP scheme (which is the one considered in this work) consists of successively linearizing any non-linear term in the dynamics and any non-convex function in the cost, seeking a solution to the original formulation by rather solving a sequence of LCPs. This entails two main advantages: first, one may rely on the substantial amount of efficient techniques that have been so far designed to deal with LCPs, or even LQRs, whenever allowed by the shape of the original NLP, and second, one may show that, whenever this iterative process converges, it finds a policy that satisfies the PMP related to the original NLP, i.e., a candidate optimum for the original formulation, which makes this approach justified and principled; importantly, though guarantees based on the PMP are weaker than guarantees based on DP, the former allow for the design of cheaper schemes.

We identify three key contributions:

1. we introduce and analyze a new framework to compute candidate optimal solutions for finite-horizon, finite-dimensional non-linear stochastic optimal control problems, with control-affine dynamics and uncontrolled diffusion; this hinges on the basic principle of SCP, i.e., iteratively solving a sequence of LCPs that stem from successive linear approximations of the original problem;
2. through a meticulous study of the continuity of the stochastic Pontryagin cones of variations with respect to linearization, we prove that, whenever the sequence of iterates converges (under specific topologies), a strategy satisfying the PMP related to the original formulation is found; in addition, we prove that, up to some subsequence, the sequence above always has an accumulation point, which in turn provides a "weak" guarantee of success for the method;
3. through an explicit example, we show how to leverage the properties offered by this framework to better understand what approximations may be adopted for the design of cheap and efficient numerical schemes for NLPs.

The paper is organized as follows. Section 2 introduces notation and preliminary results, and defines the stochastic optimal control problem of interest. In Section 3, we introduce the framework of stochastic SCP and the stochastic PMP, and state our main result of convergence. Section 4 retraces the proof of the stochastic PMP to prove our main result of convergence. In Section 5, we show how to leverage our analysis to design a practical numerical method to solve non-linear stochastic optimal control problems, along with numerical experiments. Finally, Section 6 summarizes conclusive remarks and perspectives on future directions.

2. Stochastic Optimal Control Setting. Let $(\Omega, \mathcal{G}, P)$ be a second–countable probability space and $B_t = (B^1_t, \ldots, B^d_t)$ be a d–dimensional Brownian motion with continuous sample paths and starting at zero, whose filtration $\mathcal{F} \triangleq (\mathcal{F}_t)_{t \geq 0} = \big(\sigma(B_s : 0 \leq s \leq t)\big)_{t \geq 0}$ is complete. We consider processes that are defined within bounded time intervals. Hence, for every $n \in \mathbb{N}$, $\ell \geq 2$ and maximal time $T \in \mathbb{R}_+$, we introduce the space $L^\ell_{\mathcal{F}}([0,T] \times \Omega; \mathbb{R}^n)$ of progressive processes $x : [0,T] \times \Omega \to \mathbb{R}^n$ (with respect to the filtration $\mathcal{F}$) such that $\mathbb{E}\left[\int_0^T \|x(s,\omega)\|^\ell \, \mathrm{d}s\right] < \infty$, where $\| \cdot \|$ is the Euclidean norm. In this setting, for every $x \in L^\ell_{\mathcal{F}}([0,T] \times \Omega; \mathbb{R}^n)$ and $i = 1, \dots, d$, the Itô integral of x with respect to B^i is the continuous, bounded in L^2 and n–dimensional martingale in $[0,T]$ (with respect to the filtration $\mathcal{F}$) that starts at zero, denoted $\int_0^\cdot x(s) \, \mathrm{d}B^i_s : [0,T] \times \Omega \to \mathbb{R}^n$. For $\ell \geq 2$, we denote $L^\ell_{\mathcal{F}}(\Omega; C([0,T]; \mathbb{R}^n))$ the space of $\mathcal{F}$–adapted with continuous sample paths processes $x : [0,T] \times \Omega \to \mathbb{R}^n$ that satisfy $\mathbb{E}\left[\sup_{s \in [0,T]} \|x(s,\omega)\|^\ell\right] < \infty$.

2.1. Stochastic Differential Equations. From now on, we fix two integers $n, m \in \mathbb{N}$, a maximal time $T \in \mathbb{R}_+$ and a compact, convex subset $U \subseteq \mathbb{R}^m$. We consider differential equations steered by either deterministic or stochastic controls. Specifically, the set of admissible controls $\mathcal{U}$ may stand for either $L^2([0,T]; U)$ or $L^2_{\mathcal{F}}([0,T] \times \Omega; U)$. Remark that, since U is compact, admissible controls are a.e. (and additionally a.s.) bounded. We are given continuous mappings $b_i : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, $i = 0, \dots, m$ and $\sigma_j : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, $j = 1, \dots, d$ which are (at least) C^2 with respect to the variable x. For a given $u \in \mathcal{U}$, we consider dynamical systems modeled through the following forward stochastic differential equation with uncontrolled diffusion

$$
\begin{aligned}
\mathrm{d}x(t) &= b(t, u(t), x(t)) \, \mathrm{d}t + \sigma(t, x(t)) \, \mathrm{d}B_t, \quad x(0) = x^0 \\
&\triangleq \left(b_0(t, x(t)) + \sum_{i=1}^m u^i(t) b_i(t, x(t)) \right) \mathrm{d}t + \sum_{j=1}^d \sigma_j(t, x(t)) \, \mathrm{d}B^j_t
\end{aligned}
\tag{2.1}
$$

where we assume that the fixed initial condition satisfies $x^0 \in L^\ell_{\mathcal{F}_0}(\Omega; \mathbb{R}^n)$, for every $\ell \geq 2$ (for instance, this holds true whenever x^0 is a deterministic vector of $\mathbb{R}^n$).

The procedure developed in this work is based on the following linearization of (2.1). For $\ell \geq 2$, let $v \in \mathcal{U}$ and $y \in L^\ell_{\mathcal{F}}(\Omega; C([0,T]; \mathbb{R}^n))$. For a given $u \in \mathcal{U}$, we define the linearization of (2.1) around (v, y) to be the following, well-defined forward stochastic differential equation with uncontrolled diffusion

$$
\begin{aligned}
\mathrm{d}x(t) &= b_{v,y}(t, u(t), x(t)) \, \mathrm{d}t + \sigma_y(t, x(t)) \, \mathrm{d}B_t, \quad x(0) = x^0 \\
&\triangleq \left(b_0(t, y(t)) + \sum_{i=1}^m u^i(t) b_i(t, y(t)) \right) \mathrm{d}t + \left(\frac{\partial b_0}{\partial x}(t, y(t)) + \sum_{i=1}^m v^i(t) \frac{\partial b_i}{\partial x}(t, y(t)) \right) \\
&\quad (x(t) - y(t)) \, \mathrm{d}t + \sum_{j=1}^d \sigma_j(t, y(t)) \, \mathrm{d}B^j_t + \sum_{j=1}^d \frac{\partial \sigma_j}{\partial x}(t, y(t))(x(t) - y(t)) \, \mathrm{d}B^j_t.
\end{aligned}
\tag{2.2}
$$

We require the solutions to (2.1), (2.2) to be bounded in expectation, uniformly with respect to u, v and y. For, we consider the following (standard) assumption:

(A_1) Functions b_i, $i = 0, \dots, m$, σ_j, $j = 1, \dots, d$, have compact supports in $[0,T] \times \mathbb{R}^n$.

Under (A_1), for every $\ell \geq 2$ and every $u \in \mathcal{U}$, the stochastic equation (2.1) has a unique (up to stochastic indistinguishability) solution in $L^\ell_{\mathcal{F}}(\Omega; C([0,T]; \mathbb{R}^n))$, whereas for every $u, v \in \mathcal{U}$ and $y \in L^\ell_{\mathcal{F}}(\Omega; C([0,T]; \mathbb{R}^n))$, the stochastic equation (2.1) has a unique (up to stochastic indistinguishability) solution in $L^\ell_{\mathcal{F}}(\Omega; C([0,T]; \mathbb{R}^n))$ (see, e.g., [22, 13]), and the following result holds (whose proof is given in the appendix).

Lemma 2.1. *Fix $\ell \geq 2$, and for $u \in \mathcal{U}$ let x_u denote the solution to (2.1), whereas, for $u, v \in \mathcal{U}$ and $y \in L^\ell_{\mathcal{F}}(\Omega; C([0,T]; \mathbb{R}^n))$, let $x_{u,v,y}$ denote the solution to (2.2). Under (A_1), there exist a constant $C \geq 0$, which does not depend on u, v or y, such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|x_u(t)\|^\ell\right] + \mathbb{E}\left[\sup_{t\in[0,T]} \|x_{u,v,y}(t)\|^\ell\right] \ \leq \ C, \ \ \mathbb{E}\left[\sup_{t\in[0,T]} \|x_{u_1}(t) - x_{u_2}(t)\|^\ell\right] +$$

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|x_{u_1,v,y}(t) - x_{u_2,v,y}(t)\|^\ell\right] \leq C\mathbb{E}\left[\left(\int_0^T \|u_1(s) - u_2(s)\| \ \mathrm{d}s\right)^\ell\right].$$

2.2. Stochastic Optimal Control Problem. Given $q \in \mathbb{N}$, we consider continuous mappings $g : \mathbb{R}^n \to \mathbb{R}^q$, $G : \mathbb{R}^m \to \mathbb{R}$, $H : \mathbb{R}^n \to \mathbb{R}$ and $L : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$, with $L(t,x,u) = L_0(t,x,u) + \sum_{i=1}^m u^i L_i(t,x,u)$. We require g, H and L_i, $i = 0, \dots, m$, to be (at least) C^2 with respect to the variable x, and G, H to be convex. In particular, G is Lipschitz when restricted to the compact and convex set U. We focus on finite-horizon, finite-dimensional non-linear stochastic Optimal Control Problems (OCP), with control-affine dynamics and uncontrolled diffusion, that takes the form

$$\begin{cases} \min\limits_{u,t_f} \mathbb{E}\left[\int_0^{t_f} f^0(s,u(s),x(s)) \ \mathrm{d}s\right] \triangleq \mathbb{E}\left[\int_0^{t_f} \Big(G(u(s)) + H(x(s)) + L(s,u(s),x(s))\Big) \ \mathrm{d}s\right] \\ \mathrm{d}x(t) = b(t,u(t),x(t)) \ \mathrm{d}t + \sigma(t,x(t)) \ \mathrm{d}B_t, \quad x(0) = x^0, \quad \mathbb{E}\left[g(x(t_f))\right] = 0 \end{cases}$$

where we optimize over (either deterministic or stochastic) controls $u \in \mathcal{U}$ and (whenever free) final times $0 \leq t_f \leq T$. We adopt the following (fairly mild) assumption:

(A_2) Mappings g, H and L_i, $i = 0, \dots, m$, either are affine or have compact supports in $\mathbb{R}^n$ and in $[0,T] \times \mathbb{R}^n$, respectively. In the case of free final time, g is affine.

In many applications of interest, state constraints (in expectation) are often involved. In this case, to make sure the procedure developed in this work still applies, every such constraint must be penalized within the cost of OCP (for example, by including those contributions in H or in L through some penalization function).

3. Stochastic Sequential Convex Programming. We propose the following framework to solve OCP, which is based on the classical SCP principle. Starting from some guess control $u_0 \in \mathcal{U}$ and some guess trajectory $x_0 \in L^\ell_{\mathcal{F}}(\Omega; C([0,T]; \mathbb{R}^n))$, $\ell \geq 2$, we inductively define a sequence of stochastic linear-convex problems whose dynamics and costs stem from successive linearizations of the mappings b, σ and L, and those problems are successively solved via the update of user-defined parameters. Specifically, the user tries to achieve convergence (with respect to some topologies that will be defined shortly) by a good choice of the initial guess (u_0, x_0) and of updates, through iterations, for *trust region constraints*. These constraints are adopted to make the successive linearizations of OCP well-posed. Below, we detail this procedure.

3.1. The Method. At iteration $k+1 \in \mathbb{N}$, by denoting

$$f^0_{v,y}(s,u,x) \triangleq G(u) + H(x) + L(s,u,y) + \frac{\partial L}{\partial x}(s,v,y)(x-y), \tag{3.1}$$

the following stochastic Linearized Optimal Control Problem ($\mathrm{LOCP}^{\Delta}_{k+1}$) is defined

$$\begin{cases} \displaystyle\min_{u,t_f} \mathbb{E}\left[\int_0^{t_f} f^0_{k+1}(s,u(s),x(s))\ \mathrm{d}s\right] \triangleq \mathbb{E}\left[\int_0^{t_f} f^0_{u_k,x_k}(s,u(s),x(s))\ \mathrm{d}s\right] \\ \mathrm{d}x(t) = b_{k+1}(t,u(t),x(s))\ \mathrm{d}t + \sigma_{k+1}(t,x(s))\ \mathrm{d}B_t, \quad x(0) = x^0 \\ \qquad\quad \triangleq b_{u_k,x_k}(t,u(t),x(s))\ \mathrm{d}t + \sigma_{x_k}(t,x(s))\ \mathrm{d}B_t \\ \mathbb{E}\left[g_{k+1}(x(t_f))\right] \triangleq \mathbb{E}\left[g(x_k(t^k_f)) + \dfrac{\partial g}{\partial x}(x_k(t^k_f))(x(t_f) - x_k(t^k_f))\right] = 0 \\ \displaystyle\int_0^T \mathbb{E}\left[\|x(s) - x_k(s)\|^2\right]\ \mathrm{d}s \le \Delta_{k+1}, \quad |t_f - t^k_f| \le \Delta_{k+1} \end{cases}$$

where we optimize over (either deterministic or stochastic) controls $u \in \mathcal{U}$ and (whenever free) final times $0 \le t_f \le T$. The quantity $(t^k_f, u_k, x_k) \in [0,T] \times \mathcal{U} \times L^\ell_{\mathcal{F}}(\Omega; C([0,T];\mathbb{R}^n))$ is defined inductively and denotes a solution to $(\mathrm{LOCP})^{\Delta}_k$.

Each problem $\mathrm{LOCP}^{\Delta}_{k+1}$ consists of linearizing OCP around the *solution at the previous iteration* (t^k_f, u_k, x_k), starting from (t^0_f, u_0, x_0). In order for that to make sense, we must restrict the research of optimal solutions for $\mathrm{LOCP}^{\Delta}_{k+1}$ to neighborhoods of (t^k_f, u_k, x_k), being aware that bounds on the final time are needed only when t_f is free. This is achieved by the last constraints listed in $\mathrm{LOCP}^{\Delta}_{k+1}$, which are called trust-region constraints, whereby the constant $\Delta_{k+1} \ge 0$ is the trust region radius. No such constraint is enforced on controls, those appearing linearly in b, σ and L. The construction of the sequence $(\mathrm{LOCP}^{\Delta}_k)_{k\in\mathbb{N}}$ is well-posed only if, for every $k \in \mathbb{N}$, $\mathrm{LOCP}^{\Delta}_{k+1}$ has a solution. For, we consider the following assumption:

(A_3) For every $k \in \mathbb{N}$, problem $\mathrm{LOCP}^{\Delta}_{k+1}$ is feasible.

By following the argument developed in [39, Theorem 5.2, Chapter 2], one readily checks that, for every $k \in \mathbb{N}$, assumption (A_3) implies the existence of an (at least locally) optimal solution for $\mathrm{LOCP}^{\Delta}_{k+1}$. Under assumptions (A_1)–(A_3), the method consists of iteratively solving those linearized problems through the update of the sequence of trust region radii, producing a sequence of tuples $(t^k_f, u_k, x_k)_{k\in\mathbb{N}}$ such that, for each $k \in \mathbb{N}$, $(t^{k+1}_f, u_{k+1}, x_{k+1})$ solves $\mathrm{LOCP}^{\Delta}_{k+1}$.

The user may in general steer this procedure to convergence (with respect to appropriate topologies) by adequately selecting an initial guess (u_0, x_0) and an update rule for $(\Delta_k)_{k\in\mathbb{N}}$[1], and appropriate choices will be described in Section 5. Assuming that an accumulation point for $(t^k_f, u_k, x_k)_{k\in\mathbb{N}}$ can be found (whose existence will be discussed shortly), our objective consists of proving that this is a candidate locally optimal solution to the original formulation OCP. Specifically, we show that any accumulation point for $(t^k_f, u_k, x_k)_{k\in\mathbb{N}}$ satisfies the stochastic PMP related to OCP. To develop such analysis, we require the absence of state constraints, and in particular of trust region constraints. For, later in our analysis we will be asking that, starting from some (even large) integer $\bar{k} \in \mathbb{N}$, the solution to $\mathrm{LOCP}^{\Delta}_{k+1}$ for $k \ge \bar{k}$ strictly satisfies the trust-region constraints. Under this requirement, those solutions are also optimal for the family of linearized problems without the presence of trust-region constraints. These problems are denoted LOCP_{k+1}.

[1] Whenever state constraint penalization is adopted, SCP procedures may also consider update rules for *penalization weights*, and those must be provided together with update rules for trust region radii. Further details may be found in [27, 9].

3.2. Stochastic Pontryagin Maximum Principle.

THEOREM 3.1 (Stochastic Pontryagin Maximum Principle [39]). *Let (t_f, u, x) be a locally optimal solution to OCP. There exists a tuple $(\mathfrak{p}, p^0, q)$, where $\mathfrak{p} \in \mathbb{R}^q$, $p^0 \le 0$ are constant, and $q = (q_1, \dots, q_d) \in L^2_{\mathcal{F}}([0, t_f] \times \Omega; \mathbb{R}^{n \times d})$, such that the following relations are satisfied (we adopt the same notation as in (2.1)):*

1. *Non-Triviality Condition: It holds $(\mathfrak{p}, p^0) \neq 0$.*
2. *Adjoint Equation:*

$$\mathrm{d}p(t) = -\bigg(p(t)^\top \frac{\partial b}{\partial x}(t, u(t), x(t)) + p^0 \frac{\partial f^0}{\partial x}(t, u(t), x(t)) + \\ q(t)^\top \frac{\partial \sigma}{\partial x}(t, x(t))\bigg)\, \mathrm{d}t + q(t)\, \mathrm{d}B_t, \quad p(t_f) = \mathfrak{p}\mathbb{E}\left[\frac{\partial g}{\partial x}(x(t_f))\right] \in \mathbb{R}^n.$$

3. *Maximality Condition:*

$$u(t) = \underset{v \in U}{\arg\max}\; \mathbb{E}\Big[p(t)^\top b(t, v, x(t)) + p^0 f^0(t, v, x(t))\Big], \text{ a.e. (det. controls)}$$

$$u(t) = \underset{v \in U}{\arg\max}\; \Big(p(t)^\top b(t, v, x(t)) + p^0 f^0(t, v, x(t))\Big), \text{ a.e., a.s. (sto. controls)}$$

4. *Transversality Condition: if the final time is free*

$$\max_{v \in U} \mathbb{E}\Big[p(t_f)^\top b(t_f, v, x(t_f)) + p^0 f^0(t_f, v, x(t_f))\Big] \ge 0 \text{ (det. controls)}$$

$$\mathbb{E}\left[\max_{v \in U} \Big(p(t_f)^\top b(t_f, v, x(t_f)) + p^0 f^0(t_f, v, x(t_f))\Big)\right] \ge 0 \text{ (sto. controls)}$$

where equalities hold in the case $t_f < T$.

The quantity $(t_f, u, x, \mathfrak{p}, p^0, q)$ is called (Pontryagin) extremal.

Theorem 3.1 provides necessary conditions for (local) optimality related to OCP. It is worth noting that, albeit mere final conditions are specified, it turns out that processes satisfying backward stochastic differential equations are adapted with respect to the filtration $\mathcal{F}$ (see, e.g., [13, 39]). Although the proof of this result for fixed final time problems is well-established, we could not find any published proof of Theorem 3.1 when the final time is free. Therefore, we will retrace its proof. Theorem 3.1 may be extended to the linearized problems without trust region constraints LOCP_k. For, we need the following additional requirement to hold, though for free final time problems only (see the proof in Section 4.1):

(A_4) In the case of free final time only, for every $k \in \mathbb{N}$, any optimal control u_k for LOCP_k has continuous sample paths at the optimal final time t^{k+1}_f for LOCP_{k+1}.

Importantly, the validity of (A_4) can be (statistically) checked through iterations.

3.3. Main Result of Convergence. Assumptions (A_1)–(A_4), together with some minor requirement, are sufficient to establish that any accumulation point for the sequence of iterates $(t^k_f, u_k, x_k)_{k \in \mathbb{N}}$ satisfies the stochastic PMP related to OCP. In the case of deterministic admissible controls, we can additionally infer the existence of accumulation points (up to some subsequence) if the following is assumed:

(A_5) For deterministic admissible controls, there exists a finite subset $\mathcal{D} \subseteq [0, T]$ such that, for $k \in \mathbb{N}$, any discontinuity of an optimal control u_{k+1} for $\mathrm{LOCP}^{\Delta}_{k+1}$ lies in $\mathcal{D}$.

Similarly to (A_4), assumption (A_5) can be iteratively checked. Moreover, in the case that optimal controls to LOCP^{Δ}_k are also optimal for LOCP_k, as a direct consequence of the PMP regular enough costs, control constraints and final conditions provide that these controls are time-continuous. Our main result of convergence is as follows.

THEOREM 3.2. *Assume that (A_1)–(A_4) hold, and that we are given a sequence $(\Delta_k, t^k_f, u_k, x_k)_{k\in\mathbb{N}}$ such that $(\Delta_k)_{k\in\mathbb{N}}$ converges to zero and, for every $k \in \mathbb{N}$, the tuple $(t^{k+1}_f, u_{k+1}, x_{k+1})$ locally solves $LOCP^{\Delta}_{k+1}$ with strict trust regions, i.e.,*

$$\int_0^T \mathbb{E}\left[\|x_{k+1}(s) - x_k(s)\|^2\right] \, \mathrm{d}s < \Delta_{k+1}, \quad |t^{k+1}_f - t^k_f| < \Delta_{k+1}. \tag{3.2}$$

1. *By adopting either deterministic or stochastic admissible controls, assume that the sequence $(u_k)_{k\in\mathbb{N}} \subseteq \mathcal{U}$ converges to some control u for the strong topology of L^2, whereas the sequence of times $(t^k_f)_{k\in\mathbb{N}} \subseteq [0,T]$ converges to some time $t_f \in [0,T]$ (in the case of free final time problems). Then, $u \in \mathcal{U}$ and, if we denote x the $\mathcal{F}$–adapted with continuous sample paths process solution to (2.1) related to the admissible control u, the following hold:*
   (a) *There exists a tuple $(\mathfrak{p}, p^0, q)$ such that $(t_f, u, x, \mathfrak{p}, p^0, q)$ is a Pontryagin extremal for the original problem OCP;*
   (b) *For every $\ell \geq 2$, $\mathbb{E}\left[\sup_{s\in[0,t_f]} \|x_k(s) - x(s)\|^{\ell}\right] \to 0$ whenever $k \to 0$;*
   (c) *For every $k \geq 1$, there exists $(\mathfrak{p}_k, p^0_k, q_k)$ such that $(t^k_f, u_k, x_k, \mathfrak{p}_k, p^0_k, q_k)$ is an extremal for $LOCP_k$, and the sequence $(t^k_f, u_k, x_k, \mathfrak{p}_k, p^0_k, q_k)_{k\in\mathbb{N}}$ satisfies, up to some subsequence, $p^0_k \to p^0$, $\mathbb{E}\left[\sup_{s\in[0,t_f]} \|p_k(s) - p(s)\|^2\right] \to 0$ and $\mathbb{E}\left[\int_0^T \|q_k(s) - q(s)\|^2 \, \mathrm{d}s\right] \to 0$, whenever $k \to 0$.*
2. *By adopting deterministic admissible controls, assume that in addition (A_5) holds. If the sequence $(u_k)_{k\in\mathbb{N}} \subseteq \mathcal{U}$ converges to some control u for the weak topology of L^2, whereas the sequence of times $(t^k_f)_{k\in\mathbb{N}} \subseteq [0,T]$ converges to some time $t_f \in [0,T]$ (in the case of free final time problems). Then, $u \in \mathcal{U}$ and the conclusions stated at points 1.a)–1.c) still hold. Furthermore, there always exist a subsequence $(u_{k_j})_{j\in\mathbb{N}} \subseteq (u_k)_{k\in\mathbb{N}} \subseteq \mathcal{U}$ that converges to a control $u \in \mathcal{U}$ for the weak topology of L^2 and a subsequence $(t^{k_j}_f)_{j\in\mathbb{N}} \subseteq (t^k_f)_{k\in\mathbb{N}}$ that converges to a time $t_f \in [0,T]$ (in the case of free final time problems), such that the conclusions stated at points 1.a)–1.c) hold, up to some subsequence.*

The guarantees offer by Theorem 3.2 read as follows. Under (A_1)–(A_3) and by selecting a *shrinking-to-zero* sequence of trust region radii, if iteratively solving problems LOCP^{Δ}_k returns a sequence of strategies that satisfy (3.2) (note, (3.2) needs to hold starting from some large enough iteration only) and whose controls converge with respect to the strong topology of L^2, then there exists a Pontryagin extremal for the original problem, i.e., a candidate (local) solution to OCP (point 1. in Theorem 3.2). Moreover, whenever deterministic controls are considered, under the additional assumption that the generated sequence of controls has a finite amount of time-discontinuities, such a converging sequence of controls always exists (point 2. in Theorem 3.2); this can be clearly interpreted as a “weak” guarantee of success.

In Theorem 3.2, there are also insightful statements concerning the convergence of Pontryagin extremals. Let us outline how those statements may be leveraged to speed up the convergence, though it is not our goal to provide a precise numerical strategy. For, adopt the notation of Theorem 3.1 and assume that we are in the situation for which applying the maximality condition of the PMP to problem OCP leads to smooth expressions for candidate optimal controls, as functions of the variables x, p^0 and p (be aware that this might not be straightforward to obtain). We are then in position to define two-points boundary value problems to solve OCP, also known as *shooting methods*, for which the decision variables become t_f, p^0, $p(t_f)$ and q. In particular, the core of the method consists of iteratively choosing $(t_f, p^0, p(t_f), q)$ making the adjoint equation evolves, until some given final condition is met (see, e.g., [12, 7] for a more detailed explanation of shooting methods). In the context of deterministic optimal control, when convergence is achieved, shooting methods terminate quite fast (at least quadratically). However, here the bottleneck are: 1) to deal with the presence of the infinite dimensional variable q, and 2) to find a good guess for the initial value of $(t_f, p^0, p(t_f), q)$ to make the whole procedure converge. In the setting of Theorem 3.2, a valid option to design well-posed shooting methods might be as follows. With the notation and the assumptions of Theorem 3.2, up to some subsequence it holds $(t^k_f, p^0_k, p_k(t^k_f), q_k) \to (t_f, p^0, p(t_f), q)$ (with respect to appropriate topologies), whenever $k \to \infty$. Therefore, assuming we have access to Pontryagin extremals along iterations and given some large enough iteration k, we may fix $q = q_k$ and initialize with $(t^k_f, p^0_k, p_k(t^k_f))$ a shooting method for OCP that operates on the finite dimensional variable $(t_f, p^0, p(t_f)) \in \mathbb{R}^{q+2}$. If successful, this strategy would speed up the convergence of the entire scheme.

4. Proof of the Main Result. We split the proof of Theorem 3.2 in three main steps. First, we retrace the proof of the stochastic PMP to introduce necessary notation and expressions. In addition, we leverage this step to provide a novel insight on how to prove the stochastic PMP by following the lines of the original work of Pontryagin and his group (see, e.g., [30, 15, 1]), proof that we could not find in the stochastic literature. Second, the convergence of trajectories and controls is showed, together with the convergence of variational inequalities (see Section 4.1.3 for a definition). The latter represents the cornerstone of the proof and paves the way for the final step, which consists of proving the convergence of the Pontryagin extremals. For sake of clarity and concision, and without loss of generality, we carry out the proof in the case of scalar Brownian motion, i.e., we assume $d = 1$. For the same reason, we consider free final times only, that are in addition strictly lower than the maximal time T. Finally, for any $x \in \mathbb{R}^\ell$ with $n \in \mathbb{N}$, we adopt the notation $\tilde{x} \triangleq (x, x^{\ell+1}) \in \mathbb{R}^{\ell+1}$.

4.1. Main Steps of the Proof of the Stochastic Maximum Principle. Before getting started, we need to introduce the notion of *Lebesgue point* for a stochastic control $u \in L^2_{\mathcal{F}}([0,T] \times \Omega; \mathbb{R}^m)$. For, we adopt the theory of Bochner integrals, showing that $L^2_{\mathcal{F}}([0,T] \times \Omega; \mathbb{R}^m) \subseteq L^2([0,T]; L^2_{\mathcal{G}}(\Omega; \mathbb{R}^m))$, where the latter is the space of Bochner integrable mappings $u : [0,T] \to L^2_{\mathcal{G}}(\Omega; \mathbb{R}^m)$. Let $u \in L^2_{\mathcal{F}}([0,T] \times \Omega; \mathbb{R}^m)$. First of all, by definition we see that, for almost every $t \in [0,T]$, it holds $\mathbb{E}\left[\|u(t)\|^2\right] < \infty$, and thus this control is well-defined as a mapping $u : [0,T] \to L^2_{\mathcal{G}}(\Omega; \mathbb{R}^m)$. Since Ω is second-countable, $L^2_{\mathcal{G}}(\Omega; \mathbb{R}^m)$ is separable, and therefore the claim follows from the Pettis measurability theorem once we prove that $u : [0,T] \to L^2_{\mathcal{G}}(\Omega; \mathbb{R}^m)$ is strongly measurable with respect to the Lebesgue measure of $[0,T]$. For, it is sufficient to show that, for every $A \in \mathcal{B}([0,T]) \otimes \mathcal{G}$ and $\alpha \in L^2_{\mathcal{G}}(\Omega; \mathbb{R}^m)$, the mapping $t \mapsto \mathbb{E}\left[\mathbb{1}_A(t,\omega)\alpha(\omega)\right]$ is Lebesgue measurable. By fixing $\alpha \in L^2_{\mathcal{G}}(\Omega; \mathbb{R}^m)$, this can be achieved by proving that the family $\{A \in \mathcal{B}([0,T]) \otimes \mathcal{G} :\ t \mapsto \mathbb{E}\left[\mathbb{1}_A(t,\omega)\alpha(\omega)\right]$ is Lebesgue measurable$\}$ is a monotone class, and then using standard monotone class arguments (the details are left to the reader). At this step, the Lebesgue differentiation theorem provides that, for almost every $t \in [0,T]$, the following relations hold:

$$\lim_{\eta \to 0} \frac{1}{\eta} \int_t^{t+\eta} \mathbb{E}\Big[\|u(s)-u(t)\|\Big] \, \mathrm{d}s = 0, \quad \lim_{\eta \to 0} \frac{1}{\eta} \int_t^{t+\eta} \mathbb{E}\Big[\|u(s)-u(t)\|^2\Big] \, \mathrm{d}s = 0.$$

Such a time $t \in [0,T]$ is called Lebesgue point for the control $u \in L^2_{\mathcal{F}}([0,T] \times \Omega; \mathbb{R}^m)$.

4.1.1. Linear Stochastic Differential Equations. Define the stochastic matrices $A(t) \triangleq \frac{\partial(b,f^0)}{\partial x}(t,u(t),x(t))$ and $D(t) \triangleq \frac{\partial(\sigma,0)}{\partial x}(t,u(t),x(t))$. For any time $r \in [0,t_f]$ and any bounded initial condition $\tilde{\xi}_r \in L^2_{\mathcal{F}_r}(\Omega;\mathbb{R}^{n+1})$, the following problem

$$\begin{cases} \mathrm{d}z(t) = A(t)z(t) \, \mathrm{d}t + D(t)z(t) \, \mathrm{d}B_t \\ z(t) = 0, \; t \in [0,r), \quad z(r) = \tilde{\xi}_r \end{cases} \tag{4.1}$$

is well-posed [39]. Its unique solution is the $\mathcal{F}$–adapted with right-continuous sample paths process $z : [0,T] \times \Omega \to \mathbb{R}^{n+1} : (t,\omega) \mapsto \mathbb{1}_{[r,T]}(t)\phi(t,\omega)\psi(r,\omega)\xi_r(\omega)$, where the matrix-valued $\mathcal{F}$–adapted with continuous sample paths processes ϕ and ψ satisfy

$$\begin{cases} \mathrm{d}\phi(t) = A(t)\phi(t)\mathrm{d}t + D(t)\phi(t)\mathrm{d}B_t \\ \phi(0) = I, \end{cases} \qquad \begin{cases} \mathrm{d}\psi(t) = -\psi(r)\left(A(t) - D(t)^2\right)\mathrm{d}t - \psi(t)D(t)\mathrm{d}B_t \\ \psi(0) = I, \end{cases} \tag{4.2}$$

respectively. In particular, a straightforward application of Itô formula show that $\phi(t)\psi(t) = \psi(t)\phi(t) = I$, and therefore $\psi(t) = \phi(t)^{-1}$, for every $t \in [0,T]$.

4.1.2. Needle-like Variations and End-point Mapping. One way to prove the PMP comes from the analysis of specific variations, called *needle-like variations*, on a certain mapping, called *end-point mapping*. Those concepts are introduced below in the context of stochastic controls, the deterministic case being easier.

Given an integer $j \in \mathbb{N}$, fix j times $0 < t_1 < \cdots < t_j < t_f$ which are Lebesgue points for u and j random variables $u_1,\dots,u_j$, such that $u_i \in L^2_{\mathcal{F}_{t_i}}(\Omega;U)$. For fixed scalars $0 \le \eta_i < t_{i+1} - t_i$, $i = 1,\dots,j-1$ and $0 \le \eta_j < t_f - t_j$, the needle-like variation $\pi = \{t_i,\eta_i,u_i\}_{i=1,\dots,j}$ of the control u is defined to be the admissible control $u_\pi(t) = u_i$, if $t \in [t_i, t_i + \eta_i]$, and $u_\pi(t) = u(t)$, otherwise. Denote $\tilde{x}_v$ the solution, related to an admissible control v, of the augmented system

$$\begin{cases} \mathrm{d}x(t) = b(t,v(t),x(t)) \, \mathrm{d}t + \sigma(t,x(t)) \, \mathrm{d}B_t, & x(0) = x^0 \\ \mathrm{d}x^{n+1}(t) = f^0(t,v(t),x(t)) \, \mathrm{d}t, & x^{n+1}(0) = 0 \end{cases} \tag{4.3}$$

and define the mapping $\tilde{g} : \mathbb{R}^{n+1} \to \mathbb{R}^{q+1} : \tilde{x} \mapsto (g(x), x^{n+1})$. For every fixed time $t \in (t_j, t_f]$, by denoting $\Delta_t \triangleq \min\{t_{i+1} - t_i, t - t_j, T - t : i = 1,\dots,j-1\} > 0$ the end-point mapping at time t is defined to be the function

$$\begin{aligned} F^j_t : \mathcal{C}^j_t \triangleq B^j_{\Delta_t}(0) \cap (\mathbb{R} \times \mathbb{R}^j_+) &\longrightarrow \mathbb{R}^{q+1} \\ (\delta,\eta_1,\dots,\eta_j) &\mapsto \mathbb{E}\left[\tilde{g}(\tilde{x}_{u_\pi}(t+\delta))\right] - \mathbb{E}\left[\tilde{g}(\tilde{x}_u(t))\right] \end{aligned} \tag{4.4}$$

where B^j_ρ is the open ball in $\mathbb{R}^j$ of radius $\rho > 0$. Variation on the variable δ are necessary only if free final time problems are considered, in which case (A_2), and in particular the fact that g is an affine function, plays a crucial role for computations.

Thanks to Lemma 2.1 (and to (A_2) in the case of free final time problems), it is not difficult to see that F^j_t is Lipschitz (see also the argument developed to prove Lemma 4.1 below). In addition, this mapping may be Gateaux differentiated at zero along admissible directions of the cone $\mathcal{C}^j_t$. For, denote $\tilde{b} = (b^\top, f^0)^\top$, $\tilde{\sigma} = (\sigma^\top, 0)^\top$ and let z_{t_i,u_i} be the unique solution to (4.1) with $\xi_{t_i} = \tilde{b}(t_i, u_i, x(t_i)) - \tilde{b}(t_i, u(t_i), x(t_i))$.

LEMMA 4.1. *Let $(\delta, \eta_1, \dots, \eta_j) \in \mathcal{C}^j_t$ and assume (A_2) holds (in particular, $\tilde{g}$ is an affine function whenever $\delta \neq 0$). If $t > t_j$ is a Lebesgue point for u, then it holds*

$$\Bigg\| \mathbb{E}\Bigg[\tilde{g}(\tilde{x}_{u_\pi}(t+\delta)) - \tilde{g}(\tilde{x}_u(t)) - \delta \frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t)) \tilde{b}(t, u(t), x_u(t)) - \\ - \sum_{i=1}^{j} \eta_i \frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t)) z_{t_i,u_i}(t) \Bigg] \Bigg\| = o\left(\delta + \sum_{i=1}^{j} \eta_i \right).$$

The proof of this result is technical (it requires an intense use of stochastic inequalities) but not difficult. We provide an extensive proof of Lemma 4.1 in the appendix.

4.1.3. Variational Inequalities. The main step in the proof of the PMP goes by contradiction, leveraging Lemma 4.1. For, from now on we will be assuming that, whenever free, the final time is a Lebesgue point for the optimal control. Otherwise, one may proceed by mimicking the argument developed in [15, Section 7.3].

Assume that (A_1) and (A_2) hold. For every $j \in \mathbb{N}$, define the linear mapping $dF^j_{t_f}(\delta, \eta) = \delta \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f)) \tilde{b}(t_f, u(t_f), x_u(t_f)) \right] + \sum_{i=1}^{j} \eta_i \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f)) z_{t_i,u_i}(t_f) \right]$ which, thanks to Lemma 4.1, satisfies $\lim_{\alpha > 0, \alpha \to 0} \frac{F^j_{t_f}(\alpha(\delta,\eta))}{\alpha} = dF^j_{t_f}(\delta, \eta)$, for every $(\delta, \eta) \in \mathbb{R} \times \mathbb{R}^j_+$. Finally, consider the closed, convex cone of $\mathbb{R}^{q+1}$ given by

$$K \triangleq \mathrm{Cl}\Bigg(\mathrm{Cone} \left\{ \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f)) z_{t_i,u_i}(t_f) \right] : \ u_i \in U, \ t_i \in (0, t_f) \text{ is Lebesgue for } u, \\ \pm\, \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f)) \tilde{b}(t_f, u(t_f), x_u(t_f)) \right] \right\} \Bigg).$$

If $K = \mathbb{R}^{q+1}$, it would hold $dF^j_{t_f}(\mathbb{R} \times \mathbb{R}^j_+) = K = \mathbb{R}^{q+1}$, and by applying [1, Lemma 12.1] one would obtain that the origin is an interior point of $F^j_{t_f}(\mathcal{C}^j_{t_f})$. In turn, this would imply that (t_f, u, x) cannot be optimal for OCP, giving a contradiction.

The argument above (together with an application of the separation plane theorem) provides the existence of a non-zero vector, denoted $\tilde{\mathfrak{p}} = (\mathfrak{p}^\top, \mathfrak{p}^0) \in \mathbb{R}^{q+1}$, such that the following variational inequalities hold

$$\begin{cases} \tilde{\mathfrak{p}}^\top \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f)) \tilde{b}(t_f, u(r), x_u(t_f)) \right] = 0 \\ \tilde{\mathfrak{p}}^\top \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f)) z_{r,v}(t_f) \right] \leq 0, \ r \in [0, t_f] \text{ is Lebesgue for } u, \ v \in L^2_{\mathcal{F}_r}(\Omega; U). \end{cases} \tag{4.5}$$

In the case of deterministic controls, the random variables $v \in L^2_{\mathcal{F}_r}(\Omega; U)$ in (4.5) are replaced by deterministic vectors $v \in U$. Moreover, whenever $t_f = T$, only negative variations on the final time are allowed, hence in this case the first equality of (4.5) actually becomes a greater-or-equal-than-zero inequality (see also [15, Chapter 7]).

4.1.4. Conclusion of the Proof of the Stochastic Maximum Principle. The conditions of the PMP are derived by working out the variational inequalities (4.5), finding expressions of some appropriate conditional expectation. The main details are developed below for stochastic controls, the deterministic case being easier.

We start by analyzing the second inequality of (4.5). First, by appropriately developing solutions to (4.1), this inequality can be reshaped as

$$\mathbb{E}\left[\left(\tilde{\mathfrak{p}}^\top \frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f))\phi_{t_f}\psi_r\right)^\top \left(\begin{pmatrix} b \\ f^0 \end{pmatrix}(r,v,x(r)) - \begin{pmatrix} b \\ f^0 \end{pmatrix}(r,u(r),x(r))\right)\right] \le 0$$

for every $r \in [0,t_f]$ Lebesgue point for u, and every $v \in L^2_{\mathcal{F}_r}(\Omega;U)$. Second, again from the structure of (4.1), it can be readily checked that, by denoting

$$p(t) \triangleq \left(\tilde{\mathfrak{p}}^\top \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f))\phi_{t_f}\middle|\mathcal{F}_t\right]\psi_t\right)_{1,\dots,n}, \quad p^0 \triangleq \left(\tilde{\mathfrak{p}}^\top \frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f))\phi_{t_f}\psi_t\right)_{n+1}, \tag{4.6}$$

the quantity p^0 is constant in $[0,t_f]$ (in addition, its negativity can be shown through a standard reformulation of problem OCP, as done in [1, Section 12.4]). Notice that the stochastic process $p : [0,t_f] \times \Omega \to \mathbb{R}^n$ is by definition $\mathcal{F}$–adapted.

The quantities so far introduced allow to reformulate the inequality above as

$$\mathbb{E}\Big[p(t)^\top \Big(b(t,u(t),x(t)) - b(t,v,x(t))\Big) + p^0\Big(f^0(t,u(t),x(t)) - f^0(t,v,x(t))\Big)\Big] \ge 0$$

for every $t \in [0,t_f]$ Lebesgue point for u, and every $v \in L^2_{\mathcal{F}_t}(\Omega;U)$, from which we infer the maximality condition of the PMP. In addition, together with the equality in (4.5), this entails the transversality condition of the PMP (t_f is Lebesgue for u).

It remains to show the existence of the process $q \in L_{\mathcal{F}}(\Omega; L^2([0,t_f];\mathbb{R}^n))$, the continuity of the sample paths of the process p and the validity of the adjoint equation. For, remark that, thanks to Jensen inequality and Lemma 2.1, the martingale $\mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f))\phi_{t_f}\middle|\mathcal{F}_t\right]$ is bounded in L^2. Hence, the martingale representation theorem provides the existence of a process $\mu \in L^2_{\mathcal{F}}([0,t_f]\times\Omega;\mathbb{R}^{(q+1)\times(n+1)})$ such that $\mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f))\phi_{t_f}\middle|\mathcal{F}_t\right] = \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_u(t_f))\phi_{t_f}\middle|\mathcal{F}_0\right] + \int_0^t \mu(s)\,\mathrm{d}B_s \triangleq N + \chi(t)$, where $N \in \mathbb{R}^{(q+1)\times(n+1)}$ is a constant matrix. The definition in (4.6) immediately gives that the sample paths of the process p are continuous. Next, an application of Itô formula (component-wise) readily shows that the product $\chi\psi$ satisfies, for $t \in [0,t_f]$,

$$\begin{aligned}(\chi\psi)(t) = \left(\int_0^t \mu(s)\,\mathrm{d}B_s\right)\psi(t) &= \int_0^t \mu(s)\psi(s)\,\mathrm{d}B_s - \int_0^t \mu(s)\psi(s)D(s)\,\mathrm{d}s \\ &- \int_0^t \chi(s)\psi(s)\left(A(s) - D(s)^2\right)\,\mathrm{d}s - \int_0^t \chi(s)\psi(s)D(s)\,\mathrm{d}B_s.\end{aligned}$$

Denoting $q(t) \triangleq \left(\tilde{\mathfrak{p}}^\top\Big(\mu(t)\psi(t) - (N\psi(t) + \chi(t)\psi(t))\,D(t)\Big)\right)_{1,\dots,n}$, the computations above entail the adjoint equation of the PMP. This may also be leveraged to show $p \in L^2_{\mathcal{F}}(\Omega;C([0,t_f];\mathbb{R}^n))$ and $q \in L^2_{\mathcal{F}}([0,t_f]\times\Omega;\mathbb{R}^n)$ (see, e.g., [39, Section 7.2]).

4.2. Proof of the Convergence Result. Here, we enter the core of the proof of Theorem 3.2. The convergences of trajectories and controls are addressed first. We devote the last two sections for the convergence of the variational inequalities and the Pontryagin extremals. For sake of concision, we only consider free final time problems.

4.2.1. Convergence of Controls and Trajectories. Thanks to (A_1)–(A_4), there exists a sequence of tuples $(t_f^k, u_k, x_k)_{k\in\mathbb{N}}$ such that, for every $k \in \mathbb{N}$, the tuple $(t_f^{k+1}, u_{k+1}, x_{k+1})$ solves LOCP_{k+1} in $[0,T]$, with $t_f^{k+1} \le T$. If $u \in \mathcal{U}$ denotes an admissible control for OCP that fulfills the conditions of Theorem 3.2, we denote by $\tilde{x} : [0,T] \times \Omega \to \mathbb{R}^{n+1}$ the $\mathcal{F}$–adapted with continuous sample path process solution to the augmented system (4.3), related to OCP, with control u. The following holds.

Lemma 4.2. *Assume that the sequence $(\Delta_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}_+$ converges to zero. If the sequence $(u_k)_{k\in\mathbb{N}}$ converges to u either for the strong topology of L^2, or, in the case of deterministic controls, for the weak topology of L^2, for every $\ell \ge 2$ it holds* $\mathbb{E}\left[\sup_{t\in[0,T]} \|x_k(t) - x(t)\|^\ell\right] \to 0,$ *whenever $k \to \infty$.*

Proof. We only consider the case of weak convergence of deterministic controls, being the other case easier. For every $t \in [0,T]$, we have (below, $C \ge 0$ represents some, often overloaded, appropriate constant)

$$
\begin{aligned}
\|x_{k+1}(t) - x(t)\|^\ell &\le C \left\| \int_0^t \Big(b_0(s,x_k(s)) - b_0(s,x(s)) \Big) \, \mathrm{d}s \right\|^\ell \\
&+ C \sum_{i=1}^m \left\| \int_0^t \Big(u^i_{k+1}(s) b_i(s,x_k(s)) - u^i(s) b_i(s,x(s)) \Big) \, \mathrm{d}s \right\|^\ell \\
&+ C \left\| \int_0^t \left(\frac{\partial b_0}{\partial x}(s,x_k(s)) + \sum_{i=1}^m u^i_k(s) \frac{\partial b_i}{\partial x}(s,x_k(s)) \right) (x_{k+1}(s) - x_k(s)) \, \mathrm{d}s \right\|^\ell \\
&+ C \left\| \int_0^t \left(\sigma(s,x_k(s)) - \sigma(s,x(s)) + \frac{\partial \sigma}{\partial x}(s,x_k(s))(x_{k+1}(s) - x_k(s)) \right) \, \mathrm{d}B_s \right\|^\ell
\end{aligned}
\tag{4.7}
$$

Now, we take expectations. For the last term, we compute

$$
\begin{aligned}
&\mathbb{E}\left[\left\| \int_0^t \left(\sigma(s,x_k(s)) - \sigma(s,x(s)) + \frac{\partial \sigma}{\partial x}(s,x_k(s))(x_{k+1}(s) - x_k(s)) \right) \, \mathrm{d}B_s \right\|^\ell \right] \le \\
&\le C \left(\int_0^t \mathbb{E}\left[\sup_{r\in[0,s]} \|x_{k+1}(r) - x(r)\|^\ell \right] \mathrm{d}s + \mathbb{E}\left[\int_0^T \|x_{k+1}(s) - x_k(s)\|^{1+(\ell-1)} \, \mathrm{d}s \right] \right) \\
&\le C \int_0^t \mathbb{E}\left[\sup_{r\in[0,s]} \|x_{k+1}(r) - x(r)\|^\ell \right] \mathrm{d}s \\
&\quad + C \left(\int_0^T \mathbb{E}\left[\|x_{k+1}(s) - x_k(s)\|^2 \right] \mathrm{d}s \right)^{\frac{1}{2}} \left(\sup_{k\in\mathbb{N}} \mathbb{E}\left[\sup_{s\in[0,T]} \|x_k(s)\|^{2(\ell-1)} \right] \right)^{\frac{1}{2}} \\
&\le C \left(\int_0^t \mathbb{E}\left[\sup_{r\in[0,s]} \|x_{k+1}(r) - x(r)\|^\ell \right] \mathrm{d}s + \Delta_{k+1} \right)
\end{aligned}
$$

thanks to Hölder and Burkholder–Davis–Gundy inequalities, and the last inequality comes from Lemma 2.1. Similar computations may be carried out for the first and third terms. To handle the second term, we proceed as follows, we see that, since $\mathbb{E}\left[\int_0^T \|b_i(s,x(s))\|^2 \, \mathrm{d}t\right] < \infty$ implies $\int_0^T \|b_i(s,x(s))\|^2 \, \mathrm{d}t < \infty$, $\mathcal{G}$–a.s., for every fixed $i = 1, \dots, m$ and $t \in [0,T]$, the convergence for the weak topology of L^2 of the sequence of controls entails that $\left\|\int_0^t \left(u^i_k(s) - u^i(s)\right) b_i(s,x(s)) \, \mathrm{d}s\right\| \longrightarrow 0$, $\mathcal{G}$–a.s., whenever $k \to \infty$. In addition, assumption (A_1) gives that $\left\|\int_a^b \left(u^i_k(s) - u^i(s)\right) b_i(s,x(s)) \, \mathrm{d}s\right\| \le$

$C|b-a|$, for every $a,b \in [0,T]$, $\mathcal{G}$–almost surely. Hence, [36, Lemma 3.4] and the dominated convergence theorem finally provide that

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left\| \int_0^t \left(u_k^i(s) - u^i(s)\right) b_i(s,x(s)) \, \mathrm{d}s \right\|^\ell \right] \longrightarrow 0, \quad k \longrightarrow \infty.$$

From this, it is easy to conclude by applying a routine Gronwäll inequality argument.□

Let $t_f \in [0,T]$ fulfill the conditions of Theorem 3.2. The sought convergence of trajectories is a consequence of Lemma 4.2, as soon as the conditions of Theorem 3.2 are met. In addition, when we consider deterministic controls, limiting points for the sequences of controls and of final times fulfilling the conditions of Theorem 3.2 always exist (up to some subsequence). Indeed, in this case, the set of admissible controls $\mathcal{U}$ is closed and convex for the strong topology of L^2, hence it is closed for the weak topology of L^2. Therefore, since $(u_k)_{k\in\mathbb{N}}$ is bounded in L^2, there exists $u \in \mathcal{U}$ such that, up to some subsequence, $(u_k)_{k\in\mathbb{N}}$ weakly converges to u for the weak topology of L^2. Also, there is $t_f \in [0,T]$ such that, up to some subsequence, $(t_f^k)_{k\in\mathbb{N}}$ converges to t_f. It remains to show that the process x is feasible for OCP. For, we compute

$$\begin{aligned}
\Big|\mathbb{E}[g(x(t_f))]\Big| &\le \mathbb{E}\Big[\|g(x(t_f)) - g(x_k(t_f^k))\|\Big] + \mathbb{E}\left[\left\|\frac{\partial g}{\partial x}(x_k(t_f^k))(x_{k+1}(t_f^{k+1}) - x_k(t_f^k))\right\|\right] \\
&\le C\left(\mathbb{E}\Big[\|x(t_f) - x(t_f^k)\|\Big] + 2\mathbb{E}\left[\sup_{t\in[0,T]} \|x_k(t) - x(t)\|^2\right]^{\frac{1}{2}}\right) \\
&\quad + C\left(\mathbb{E}\Big[\|x(t_f^{k+1}) - x(t_f^k)\|\Big] + \mathbb{E}\left[\sup_{t\in[0,T]} \|x_{k+1}(t) - x(t)\|^2\right]^{\frac{1}{2}}\right) \longrightarrow 0
\end{aligned}$$

thanks to Lemma 4.2 and the dominated convergence theorem (here, $C \ge 0$ is a constant coming from (A_2), and we use the continuity of the sample paths of x).

4.2.2. Convergence of Variational Inequalities. Let us start with a crucial result on linear stochastic differential equations. Recall the notation of Section 4.1.

LEMMA 4.3. *Fix $\ell \ge 2$ and consider sequences of times $(r_k)_{k\in\mathbb{N}} \subseteq [0,T]$ and of uniformly bounded variables $(\tilde{\xi}_k)_{k\in\mathbb{N}}$ such that, for every $k \in \mathbb{N}$, $\tilde{\xi}_k \in L^2_{\mathcal{F}_{r_k}}(\Omega;\mathbb{R}^{n+1})$. Assume that $r_k \to r$ with $r_k \le r$ for $k \in \mathbb{N}$ and $\tilde{\xi}_k \overset{L^\ell}{\to} \tilde{\xi} \in L^2_{\mathcal{F}_r}(\Omega;\mathbb{R}^{n+1})$ with $\tilde{\xi}$ bounded. Denote $\tilde{y}_{k+1}$, $\tilde{y}$ the stochastic processes solutions to, respectively,*

$$\begin{cases}
\mathrm{d}y(t) = \left(\dfrac{\partial b_0}{\partial x}(t,x_k(t)) + \displaystyle\sum_{i=1}^m u_k^i(t)\frac{\partial b_i}{\partial x}(t,x_k(t))\right) y(t)\,\mathrm{d}t + \dfrac{\partial \sigma}{\partial x}(t,x_k(t))y(t)\,\mathrm{d}B_t \\
\mathrm{d}y^{n+1}(t) = \left(\dfrac{\partial H}{\partial x}(x_{k+1}(t)) + \dfrac{\partial L}{\partial x}(t,u_k(t),x_k(t))\right) y(t)\,\mathrm{d}t \\
\tilde{y}(t) = 0, \ t \in [0,r_k), \quad \tilde{y}(r_k) = \tilde{\xi}_{k+1},
\end{cases}$$

$$\begin{cases}
\mathrm{d}y(t) = \left(\dfrac{\partial b_0}{\partial x}(t,x(t)) + \displaystyle\sum_{i=1}^m u^i(t)\frac{\partial b_i}{\partial x}(t,x(t))\right) y(t)\,\mathrm{d}t + \dfrac{\partial \sigma}{\partial x}(t,x(t))y(t)\,\mathrm{d}B_t \\
\mathrm{d}y^{n+1}(t) = \left(\dfrac{\partial H}{\partial x}(x(t)) + \dfrac{\partial L}{\partial x}(t,u(t),x(t))\right) y(t)\,\mathrm{d}t \\
\tilde{y}(t) = 0, \ t \in [0,r), \quad \tilde{y}(r) = \tilde{\xi}.
\end{cases}$$

Under the assumptions of Lemma 4.2, $\mathbb{E}\left[\sup_{t\in[r,T]} \|\tilde{y}_k(t)-\tilde{y}(t)\|^\ell\right] \to 0$*, for* $k\to\infty$.

Proof. From $r_k \le r$, $k\in\mathbb{N}$, for $t\in[r,T]$ we have (below, $C\ge 0$ is a constant)

$$
\begin{aligned}
(4.8)\qquad & \|\tilde{y}_{k+1}(t)-\tilde{y}(t)\|^\ell \le C\|\tilde{\xi}_{k+1}-\tilde{\xi}\|^\ell \\
&+ C\left\|\int_{r_k}^t \left(\frac{\partial H}{\partial x}(x_{k+1}(s))y_{k+1}(s) - \frac{\partial H}{\partial x}(x(s))y(s)\right)\,\mathrm{d}s\right\|^\ell \\
&+ C\left\|\int_{r_k}^t \left(\frac{\partial b_0}{\partial x}(s,x_k(s))y_{k+1}(s) - \frac{\partial b_0}{\partial x}(s,x(s))y(s)\right)\,\mathrm{d}s\right\|^\ell
\end{aligned}
$$

$$
\begin{aligned}
&+ C\left\|\int_0^t \mathbb{1}_{[r_k,T]}(t)\left(\frac{\partial \sigma}{\partial x}(s,x_k(s))y_{k+1}(s) - \frac{\partial \sigma}{\partial x}(s,x(s))y(s)\right)\,\mathrm{d}B_s\right\|^\ell \\
&+ C\left\|\int_{r_k}^t \left(\frac{\partial L}{\partial x}(t,u_k(t),x_k(t))y_{k+1}(s) - \frac{\partial L}{\partial x}(t,u(t),x(t))y(s)\right)\,\mathrm{d}s\right\|^\ell \\
&+ C\sum_{i=1}^m \left\|\int_{r_k}^t \left(u_k^i(s)\frac{\partial b_i}{\partial x}(s,x_k(s))y_{k+1}(s) - u^i(s)\frac{\partial b_i}{\partial x}(s,x(s))y(s)\right)\,\mathrm{d}s\right\|^\ell .
\end{aligned}
$$

We consider the expectation of the fourth term on the right-hand side (a similar argument can be developed for the first three terms, which are left to the reader). The Burkholder–Davis–Gundy and Holdër inequalities give

$$
\begin{aligned}
\mathbb{E}\left[\left\|\int_0^t \mathbb{1}_{[r_k,T]}(t)\left(\frac{\partial \sigma}{\partial x}(s,x_k(s))y_{k+1}(s) - \frac{\partial \sigma}{\partial x}(s,x(s))y(s)\right)\,\mathrm{d}B_s\right\|^\ell\right] \le \\
\le C|r-r_k| + C\int_r^t \mathbb{E}\left[\sup_{s'\in[0,s]} \|y_{k+1}(s')-y(s')\|^\ell\right]\mathrm{d}s \\
+ C\mathbb{E}\left[\sup_{s\in[0,T]}\|y(s)\|^{2\ell}\right]\mathbb{E}\left[\sup_{s\in[0,T]}\|x_k(s)-x(s)\|^{2\ell}\right]
\end{aligned}
$$

where the last inequality makes sense thanks to Lemma 4.2, and because Lemma 2.1 may be straightforwardly extended to $\tilde{y}$. Finally, thanks to the fact that L is affine with respect to the control variable, we may handle the fifth and the last terms in (4.8) by combining the argument above with the final steps in the proof of Lemma 4.2. Finally, a routine Gronwäll inequality argument provides the conclusion. ∎

We first consider the case of deterministic admissible controls, for which the sequence $(u_k)_{k\in\mathbb{N}} \subseteq \mathcal{U}$ converges to $u\in\mathcal{U}$ for the weak convergence of L^2. Since we assume (A_5), we denote $\mathcal{L}\subseteq[0,T]$ the full measure subset such that $r\in\mathcal{L}$ if and only if r is a Lebesgue point for u and $r\notin\mathcal{D}$. We prove the existence of a non-zero vector $\tilde{\mathfrak{p}}\in\mathbb{R}^{q+1}$, whose last component is non-positive, such that, for every $r\in\mathcal{L}$, $v\in U$,

$$
(4.9)\qquad \tilde{\mathfrak{p}}^\top \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}(t_f))z_{r,v}(t_f)\right] \le 0,
$$

where, by denoting $A(t) = \begin{pmatrix} \frac{\partial b_0}{\partial x}(t,x(t)) + \sum_{i=1}^m u^i(t)\frac{\partial b_i}{\partial x}(t,x(t)) \\ \frac{\partial H}{\partial x}(x(t)) + \frac{\partial L}{\partial x}(t,u(t),x(t)) \end{pmatrix}$ and $D(t) = \begin{pmatrix} \frac{\partial \sigma}{\partial x}(t,x(t)) \\ 0 \end{pmatrix}$, the $\mathcal{F}$–adapted with continuous sample paths stochastic process

$z_{r,v} : [0,T] \times \Omega \to \mathbb{R}^{n+1}$ solves (4.1) with $\tilde{\xi}_r = \tilde{\xi}_{r,v} \triangleq \tilde{b}(r,v,x(r)) - \tilde{b}(r,u(r),x(r))$ (where we denote $\tilde{b} = (b^\top, f^0)^\top$. We will use this notation from now on).

For, thanks to (3.2), for every $k \in \mathbb{N}$, the optimality of $(t_f^{k+1}, u_{k+1}, x_{k+1})$ for LOCP_{k+1} provides a non-zero vector $\tilde{\mathfrak{p}}_{k+1} \in \mathbb{R}^{q+1}$, whose last component is nonpositive, so that, for $r \in \mathcal{L} \cap [0, t_f^{k+1}]$, $v \in U$, it holds $\tilde{\mathfrak{p}}_{k+1}^\top \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_k(t_f^k)) z_{r,v}^{k+1}(t_f^{k+1})\right] \le 0$, where, with the notation $A_{k+1}(t) = \begin{pmatrix} \frac{\partial b_0}{\partial x}(t,x_k(t)) + \sum_{i=1}^m u_k^i(t) \frac{\partial b_i}{\partial x}(t,x_k(t)) \\ \frac{\partial H}{\partial x}(x_{k+1}(t)) + \frac{\partial L}{\partial x}(t,u_k(t),x_k(t)) \end{pmatrix}$ and $D_{k+1}(t) = \begin{pmatrix} \frac{\partial \sigma}{\partial x}(t,x_k(t)) \\ 0 \end{pmatrix}$, the $\mathcal{F}$–adapted with continuous sample paths stochastic process $z_{r,v}^{k+1} : [0,T] \times \Omega \to \mathbb{R}^{n+1}$ solves (4.1) with $A = A_{k+1}$, $D = D_{k+1}$ and $\tilde{\xi}_r = \tilde{\xi}_{r,v}^{k+1} \triangleq \tilde{b}_{k+1}(r,v,x_{k+1}(r)) - \tilde{b}_{k+1}(r,u_{k+1}(r),x_{k+1}(r))$ (again, we denote $\tilde{b}_{k+1} = (b_{k+1}^\top, f_{k+1}^0)^\top$). Now, fix $r \in \mathcal{L} \cap (0,t_f)$ and $v \in U$. We may assume $r \le t_f^k$ for every $k \in \mathbb{N}$ large enough. The following comes from extending [10, Lemma 3.11].

Lemma 4.4. *Under (A_5), there exists $(r_k)_{k\in\mathbb{N}} \subseteq (0,r)$, such that, for every $k \in \mathbb{N}$, r_k is a Lebesgue point for u_k, and $r_k \longrightarrow r$, $u_k(r_k) \longrightarrow u(r)$, whenever $k \to \infty$.*

If $(r_k)_{k\in\mathbb{N}} \subseteq (0,r)$ denotes the sequence given by Lemma 4.4, we define $\tilde{\xi}_k = \tilde{\xi}_{r_k,v}^k$ and $\tilde{\xi} = \tilde{\xi}_{r,v}$. Straightforward computations give (below, $C \ge 0$ is a constant)

$$\begin{aligned}(4.10)\quad \mathbb{E}\left[\|\tilde{\xi}_{k+1} - \tilde{\xi}\|^2\right] \le & C\Bigg(\|u_{k+1}(r_{k+1}) - u(r)\|^2 + |r_{k+1} - r|^2 \\ & + \mathbb{E}\left[\|x(r_{k+1}) - x(r)\|^2\right] + \mathbb{E}\left[\sup_{s\in[0,T]} \|x_k(s) - x(s)\|^2\right] \Bigg),\end{aligned}$$

from which $\mathbb{E}\left[\|\tilde{\xi}_{k+1} - \tilde{\xi}\|^2\right] \to 0$ for $k \to \infty$, thanks to Lemma 4.2 and Lemma 4.4. Therefore, from Lemma 4.3 we infer that $\mathbb{E}\left[\sup_{s\in[r,T]} \|z_{r_k,v}^k(s) - z_{r,v}(s)\|^2\right] \to 0$, whenever $k \to \infty$. In particular, $\mathbb{E}\left[\|z_{r_{k+1},v}^{k+1}(t_f^{k+1}) - z_{r,v}(t_f)\|^2\right] \longrightarrow 0$, for $k \longrightarrow \infty$. At this step, we point out that the variational inequality in (4.5) still holds if we take multipliers of norm one. Specifically, we may assume that $\|\tilde{\mathfrak{p}}_k\| = 1$, for every $k \in \mathbb{N}$. Therefore, up to some subsequence, there exists a vector $\tilde{\mathfrak{p}} = (\mathfrak{p}^\top, p^0)^\top \in \mathbb{R}^{q+1}$ such that $\tilde{\mathfrak{p}}_k \to \tilde{\mathfrak{p}}$, for $k \to \infty$, and satisfying $\tilde{\mathfrak{p}} \ne 0$, $p^0 \le 0$. We use this remark to conclude as follows. The definition of $\tilde{g}$ and Hölder inequality give ($C \ge 0$ is a constant)

$$\begin{aligned}\tilde{\mathfrak{p}}^\top \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}(t_f)) z_{r,v}(t_f)\right] &\le \left| \tilde{\mathfrak{p}}^\top \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}(t_f)) z_{r,v}(t_f)\right] - \tilde{\mathfrak{p}}_{k+1}^\top \mathbb{E}\left[\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_k(t_f^k)) z_{r_{k+1},v}^{k+1}(t_f^{k+1})\right] \right| \\ &\le C\left(\|\tilde{\mathfrak{p}} - \tilde{\mathfrak{p}}_{k+1}\| + \mathbb{E}\left[\|z_{r_{k+1},v}^{k+1}(t_f^{k+1}) - z_{r,v}(t_f)\|^2\right]^{\frac{1}{2}} \right) \\ &\quad + C\mathbb{E}\left[\|z_{r,v}(t_f)\|^2\right]^{\frac{1}{2}} \left(\mathbb{E}\left[\sup_{s\in[0,T]} \|x_k(s) - x(s)\|^2\right]^{\frac{1}{2}} + \mathbb{E}\left[\|x(t_f^k) - x(s)\|^2\right]^{\frac{1}{2}} \right),\end{aligned}$$

and, in this case, (4.9) follows from Lemma 4.2 and the convergences obtained above.

Finally, we turn to the case for which the sequence $(u_k)_{k\in\mathbb{N}} \subseteq \mathcal{U}$ converges to $u \in \mathcal{U}$ for the strong topology of L^2 (in particular, controls may be either deterministic or stochastic). For, fix $v \in U$ and define the stochastic processes $\tilde{\xi}_k(s) = \tilde{\xi}_{s,v}^k$ and $\tilde{\xi}(s) = \tilde{\xi}_{s,v}$, where $s \in [0,T]$. Similar computations to the ones developed to compute the bound (4.10) provide that $\int_0^T \mathbb{E}\left[\|\tilde{\xi}_k(s) - \tilde{\xi}(s)\|^2\right] \, \mathrm{d}s \to 0$, whenever $k \to \infty$, and therefore, up to some subsequence, the quantity $\mathbb{E}\left[\|\tilde{\xi}_k(s) - \tilde{\xi}(s)\|^2\right]$ converges to zero for $k \to 0$, a.e. in $[0,T]$. By taking countable intersections of sets of Lebesgue points (one for each control u_k, for all $k \in \mathbb{N}$), it follows that the argument above can be iterated exactly in the same manner (via Lemma 4.3), leading to the same conclusion.

4.2.3. Convergence of Multipliers and Conclusion. By applying the construction developed in Section 4.1.4 to the variational inequality (4.9), we retrieve a tuple $(\mathfrak{p}, p^0, q)$, where $\mathfrak{p} \in \mathbb{R}^q$, $p^0 \leq 0$ are constant and $q \in L^2_{\mathcal{F}}([0,T] \times \Omega; \mathbb{R}^n)$, such that $(t_f, u, x, \mathfrak{p}, p^0, q)$ that satisfies conditions 1.,2. and 3. of Theorem 3.1, relatively to OCP. To conclude the proof of Theorem 3.2, it only remains to prove that:

1. up to some subsequence, $\mathbb{E}\left[\sup_{s\in[0,t_f]} \|p_k(s) - p(s)\|^2\right] \to 0$, for $k \to \infty$, where p solves the adjoint equation of Theorem 3.1 relatively to OCP,
2. $(t_f, u, x, \mathfrak{p}, p^0, q)$ satisfies condition 4. of Theorem 3.1, relatively to OCP,
3. up to some subsequence, $\int_0^{t_f} \mathbb{E}\Big[\|q_k(s) - q(s)\|^2\Big] \, \mathrm{d}s \to 0$, for $k \to \infty$,

where each tuple $(t^k_f, u_k, x_k, \mathfrak{p}_k, p^0_k, q_k)$ is the extremal of LOCP$_k$ that we have introduced above, and p_k solves the adjoint equation of Theorem 3.1 relatively to LOCP$_k$.

Let us start with the first assertion. For, fix $k \in \mathbb{N}$ and consider the process $\mathbb{E}\left[\sup_{s\in[0,t_f]} \left\|\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_k(t^k_f))\phi_{k+1}(t^{k+1}_f)\psi_{k+1}(s) - \frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}(t_f))\phi(t_f)\psi(s)\right\| \middle| \mathcal{F}_t\right]$ for $t \in [0,t_f]$, where ϕ_{k+1}, ψ_{k+1} solve (4.2) with matrices A_{k+1}, D_{k+1}, whereas ϕ, ψ solve (4.2) with matrices A, D, those matrices being defined above. Thanks to a straightforward extension of Lemma 2.1 to equations (4.2), this process is a martingale, bounded in L^2. Hence, the martingale representation theorem allows us to infer that it is a martingale with continuous sample paths, and Doob and Jensen inequalities give

(4.11)
$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,t_f]} \mathbb{E}\left[\sup_{s\in[0,t_f]} \left\|\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_k(t^k_f))\phi_{k+1}(t^{k+1}_f)\psi_{k+1}(s) - \frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}(t_f))\phi(t_f)\psi(s)\right\| \middle| \mathcal{F}_t\right]^2\right] \leq \\
&\quad \leq 4\mathbb{E}\left[\sup_{s\in[0,t_f]} \left\|\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_k(t^k_f))\phi_{k+1}(t^{k+1}_f)\psi_{k+1}(s) - \frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}(t_f))\phi(t_f)\psi(s)\right\|^2\right],
\end{aligned}$$

this holding for every $k \in \mathbb{N}$. By combining (4.11) with (4.6), we compute

$$\begin{aligned}
\mathbb{E}\left[\sup_{s\in[0,t_f]} \|p_{k+1}(s) - p(s)\|^2\right] &\leq C\mathbb{E}\left[\sup_{t\in[0,t_f]} \left\|\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}(t_f))\phi(t_f)\psi(s)\right\|^2\right] \|\tilde{\mathfrak{p}}_{k+1} - \tilde{\mathfrak{p}}\|^2 \\
&+ C\mathbb{E}\left[\sup_{s\in[0,t_f]} \left\|\frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}_k(t^k_f))\phi_{k+1}(t^{k+1}_f)\psi_{k+1}(s) - \frac{\partial \tilde{g}}{\partial \tilde{x}}(\tilde{x}(t_f))\phi(t_f)\psi(s)\right\|^2\right]
\end{aligned}$$

where $C \geq 0$ is a constant. Up to some subsequence, the first term on the right-hand side converges to zero. Moreover, the definition of $\tilde{g}$ and Hölder inequality give

$$\mathbb{E}\left[\sup_{s\in[0,t_f]}\left\|\frac{\partial\tilde{g}}{\partial\tilde{x}}(\tilde{x}_k(t_f^k))\phi_{k+1}(t_f^{k+1})\psi_{k+1}(s)-\frac{\partial\tilde{g}}{\partial\tilde{x}}(\tilde{x}(t_f))\phi(t_f)\psi(s)\right\|^2\right]\le$$

$$\le C\mathbb{E}\left[\|\phi_{k+1}(t_f^{k+1})\|^4\right]^{\frac{1}{2}}\mathbb{E}\left[\sup_{s\in[0,T]}\|\psi_{k+1}(s)-\psi(s)\|^4\right]^{\frac{1}{2}}$$

$$+C\mathbb{E}\left[\sup_{s\in[0,T]}\|\psi(s)\|^4\right]^{\frac{1}{2}}\left(\mathbb{E}\left[\sup_{s\in[0,T]}\|\phi_{k+1}(s)-\phi(s)\|^4\right]^{\frac{1}{2}}+\mathbb{E}\left[\|\phi(t_f^{k+1})-\phi(t_f)\|^4\right]^{\frac{1}{2}}\right)$$

$$+C\mathbb{E}\left[\sup_{s\in[0,T]}\|\phi(t_f)\psi(s)\|^4\right]^{\frac{1}{2}}\left(\mathbb{E}\Big[\sup_{s\in[0,T]}\|x_k(s)-x(s)\|^4\Big]^{\frac{1}{2}}+\mathbb{E}\Big[\|x(t_f^k)-x(t_f)\|^4\Big]^{\frac{1}{2}}\right)$$

and a straightforward extension of Lemma 4.3 to equations (4.2) entails that all the terms on the right-hand side tend to zero. The convergence of $(p_k)_{k\in\mathbb{N}}$ is proved.

The fact that $(t_f,u,x,\mathfrak{p},p^0,q)$ satisfies condition 4. of Theorem 3.1, relatively to OCP, is now a direct consequence of this last convergence and Lemma 4.2.

It remains to prove that, up to some subsequence, $\int_0^{t_f}\mathbb{E}\Big[\|q_k(s)-q(s)\|^2\Big]\,\mathrm{d}s\to 0$, whenever $k\to 0$. For, we apply Itô formula to $\|p_{k+1}(t)-p(t)\|^2$, which, thanks to (4.6), (4.11) and Lemma 2.1 extended to (4.2), gives (below, $C\ge 0$ is a constant)

$$\mathbb{E}\Big[\|p_{k+1}(0)-p(0)\|^2\Big]+\int_0^{t_f}\mathbb{E}\Big[\|q_{k+1}(s)-q(s)\|^2\Big]\,\mathrm{d}s=\mathbb{E}\Big[\|p_{k+1}(t_f)-p(t_f)\|^2\Big]$$

$$+2\mathbb{E}\Bigg[\int_0^{t_f}\Big(p_{k+1}(s)-p(s)\Big)^\top$$

$$\left(p(s)_{k+1}^\top\frac{\partial b_{k+1}}{\partial x}(s,u_{k+1}(s),x_{k+1}(s))-p(s)^\top\frac{\partial b}{\partial x}(s,u(s),x(s))\right)\mathrm{d}s\Bigg]$$

$$+2\mathbb{E}\left[\int_0^{t_f}\Big(p_{k+1}(s)-p(s)\Big)^\top\left(p_{k+1}^0\frac{\partial f_{k+1}^0}{\partial x}(s,u_{k+1}(s),x_{k+1}(s))-p^0\frac{\partial f^0}{\partial x}(s,u(s),x(s))\right)\mathrm{d}s\right]$$

$$+2\mathbb{E}\left[\int_0^{t_f}\Big(p_{k+1}(s)-p(s)\Big)^\top\left(q(s)_{k+1}^\top\frac{\partial\sigma_{k+1}}{\partial x}(s,x_{k+1}(s))-q(s)^\top\frac{\partial\sigma}{\partial x}(s,x(s))\right)\mathrm{d}s\right]$$

$$\le C\mathbb{E}\left[\sup_{s\in[0,t_f]}\|p_{k+1}(s)-p(s)\|^2\right]^{\frac{1}{2}}\left(\int_0^{t_f}\mathbb{E}\Big[\|q_{k+1}(s)-q(s)\|^2\Big]\,\mathrm{d}s\right)^{\frac{1}{2}}$$

$$+C\mathbb{E}\left[\sup_{s\in[0,t_f]}\|p_{k+1}(s)-p(s)\|^2\right]^{\frac{1}{2}}\left(1+\left(\int_0^{t_f}\mathbb{E}\Big[\|q(s)\|^2\Big]\,\mathrm{d}s\right)^{\frac{1}{2}}\right).$$

The conclusion finally follows from Young inequality and the convergence of $(p_k)_{k\in\mathbb{N}}$.

5. An Example of Numerical Scheme.

Although the procedure detailed previously provides methodological steps to tackle OCP through successive linearizations, numerically solving LCPs that depend on stochastic coefficients still remains a challenge. In this last section, under appropriate assumptions we propose an approximate, though very handy, numerical scheme to effectively solve each subproblem LOCP_k^{Δ}. We stress the fact that our main goal is not the development of an ultimate algorithm, but rather consists of demonstrating how one may leverage the theoretical insights provided by Theorem 3.2 to design efficient strategies to practically solve OCP.

5.1. A Simplified Context.

The proposed approach relies on a specific shape of the cost and the dynamics of OCP. Specifically, we assume we control the system with deterministic admissible controls $u \in \mathcal{U} = L^2([0,T];U)$ only and cost functions f^0 are such that $G = 0$. Moreover, we assume the state variable to be given by two components $(x,z) \in \mathbb{R}^{n_x+n_z}$, for $n_x, n_z \in \mathbb{N}$, satisfying the following system of forward stochastic differential equations (b^x and b^z are accordingly defined as in (2.1))

$$\begin{cases} \mathrm{d}x(t) = b^x(t,u(t),x(t),z(t))\ \mathrm{d}t + \sigma(t,z(t))\ \mathrm{d}B_t, \quad x(0) = x^0 \\ \mathrm{d}z(t) = b^z(t,u(t),z(t))\ \mathrm{d}t, \quad z(0) = z^0 \in \mathbb{R}^{n_z}. \end{cases} \tag{5.1}$$

In particular, any $\mathcal{F}$-adapted with continuous sample paths process (x,z) solution to (5.1) (for a given control $u \in \mathcal{U}$) is such that z is deterministic. For the sake of clarity, to avoid cumbersome notation from now on we assume that b^x does not explicitly depend on z, and $g = 0$. This is clearly done without loss of generality.

5.2. The Proposed Approach. With the assumptions adopted previously, we see that the diffusion in the dynamics of OCP is now forced to be deterministic. This fact is at the root of our method, which mimic the procedure proposed in [4]. Specifically, we transcribe every stochastic subproblem LOCP^{Δ}_k into a deterministic and convex optimal control problem, whereby the variables are the mean and the covariance of the solution x_k to LOCP^{Δ}_k. The main advantage in doing so is that deterministic reformulations of the subproblems LOCP^{Δ}_k can be efficiently solved by off-the-shelf convex solvers. Unlike [4], here we rely on some *upstream information* for the design of this numerical scheme which is entailed by Theorem 3.2 as follows.

By recalling the notation introduced in the previous sections, we denote $\mu(t) \triangleq \mathbb{E}[x(t)]$, $\Sigma(t) \triangleq \mathbb{E}[(x(t)-\mu(t))(x(t)-\mu(t))^\top]$, and, for $k \in \mathbb{N}$, $\mu_k(t) \triangleq \mathbb{E}[x_k(t)]$, $\Sigma_k(t) \triangleq \mathbb{E}[(x_k(t)-\mu_k(t))(x_k(t)-\mu_k(t))^\top]$. Heuristically, assuming that solutions to LOCP^{Δ}_k have *small variance* $\|\mathrm{tr}\ \Sigma_k\|_{L^2} \ll 1$, we might compute the linearization of b^x, b^z, σ and L at μ_k rather than at x_k. In doing so, for the cost of $\mathrm{LOCP}^{\Delta}_{k+1}$ we obtain the following approximation (the notation goes accordingly as in (3.1))

$$\mathbb{E}\left[\int_0^{t_f} f^0_{k+1}(s,u(s),x(s),z(s))\ \mathrm{d}s\right] \approx \int_0^{t_f} f^0_{u_k,\mu_k}(s,u(s),\mu(s),z(s))\ \mathrm{d}s, \tag{5.2}$$

whereas for the dynamics of $\mathrm{LOCP}^{\Delta}_{k+1}$ (the notation goes accordingly as in (2.2))

$$\begin{cases} \mathrm{d}x(t) \approx b^x_{u_k,\mu_k}(t,u(t),x(t))\ \mathrm{d}t + \sigma_{z_k}(t,z(t))\ \mathrm{d}B_t \\ \qquad\quad = \Big(\mathcal{A}_{k+1}(t)x(t) + \mathcal{B}_{k+1}(t,u(t))\Big)\ \mathrm{d}t + \mathcal{C}_{k+1}(t,z(t))\ \mathrm{d}B_t \\ \mathrm{d}z(t) \approx b^z_{u_k,z_k}(t,u(t),z(t))\ \mathrm{d}t = \Big(\mathcal{D}_{k+1}(t)z(t) + \mathcal{E}_{k+1}(t,u(t))\Big)\ \mathrm{d}t \end{cases} \tag{5.3}$$

where $\mathcal{A}_{k+1}(t) \in \mathbb{R}^{n_x \times n_x}$, $\mathcal{B}_{k+1}(t,u) \in \mathbb{R}^{n_x}$, $\mathcal{C}_{k+1}(t,z) \in \mathbb{R}^{n_x}$, $\mathcal{D}_{k+1}(t) \in \mathbb{R}^{n_z \times n_z}$ and $\mathcal{E}_{k+1}(t,u) \in \mathbb{R}^{n_z}$ are deterministic, with $\mathcal{C}_{k+1}(t,z)$ affine in z. Accordingly, by introducing $\mu^{e_k}(t) \triangleq \mu(t) - \mu_k(t)$ and $\Sigma^{e_k}(t) \triangleq \mathbb{E}[(x(t)-x_k(t)-\mu(t)+\mu_k(t))(x(t)-x_k(t)-\mu(t)+\mu_k(t))^\top]$, slightly tighter trust region constraints are given by

$$\int_0^T \mathrm{tr}\ \Sigma^{e_k}(t)\ \mathrm{d}t + \int_0^T \|\mu^{e_k}(t)\|^2\ \mathrm{d}t \le \Delta_{k+1}, \quad |t^{k+1}_f - t^k_f| \le \Delta_{k+1}. \tag{5.4}$$

At this point, the crucial remark is that, since all the coefficients are deterministic, solutions to (5.3) are Gaussian processes whose dynamics the the form (see, e.g., [26])

$$\begin{cases} \dot{\mu}(t) = \mathcal{A}_{k+1}(t)\mu(t) + \mathcal{B}_{k+1}(t,u(t)), \quad \dot{z}(t) = \mathcal{D}_{k+1}(t)z(t) + \mathcal{E}_{k+1}(t,u(t)) \\ \dot{\Sigma}(t) = \mathcal{A}_{k+1}(t)\Sigma(t) + \Sigma(t)\mathcal{A}_{k+1}(t)^\top + \mathcal{C}_{k+1}(t,z(t))\mathcal{C}_{k+1}(t,z(t))^\top. \end{cases} \tag{5.5}$$

The system above is not linear because of $\mathcal{C}_{k+1}(t, z(t))\mathcal{C}_{k+1}(t, z(t))^\top$. Nevertheless, we may call upon the convergences of Theorem 3.2 to replace (5.5) with

$$
(5.6) \quad \begin{cases} \dot{\mu}(t) = \mathcal{A}_{k+1}(t)\mu(t) + \mathcal{B}_{k+1}(t, u(t)), \quad \dot{z}(t) = \mathcal{D}_{k+1}(t)z(t) + \mathcal{E}_{k+1}(t, u(t)) \\ \dot{\Sigma}(t) = \mathcal{A}_{k+1}(t)\Sigma(t) + \Sigma(t)\mathcal{A}_{k+1}(t)^\top + \mathcal{C}_{k+1}(t, z_k(t))\mathcal{C}_{k+1}(t, z(t))^\top. \end{cases}
$$

Finally, with the results of Theorem 3.2, we may heuristically replace every $\mathrm{LOCP}^{\Delta}_{k+1}$ with the deterministic convex optimal control problem whose variables are μ and Σ (and additionally μ^{e_k} and Σ_{e_k}), and whose dynamics are (5.6), whose trust region constraints are (5.4), and whose cost consists of replacing (5.2) with

$$
\int_0^{t_f} \left(f^0_{u_k,\mu_k}(s, u(s), \mu(s), z(s)) + \mathrm{tr}\ \Sigma(t) \right) \mathrm{d}s
$$

to force solutions to have small variances, which in turn justify the whole approach.

5.3. Uncertain Car Trajectory Planning Problem. We now provide numerical experiments. We consider the problem of planning the trajectory of a car, whose state and control inputs are $x = (r_x, r_y, \theta, v, \omega)$, and $u = (a_v, a_\omega)$, and whose dynamics are

$$
b(t, x, u) = (v\cos(\theta), v\sin(\theta), \omega, a_v, a_\omega), \ \ \sigma(t, x) = \mathrm{diag}([\alpha\omega v, \alpha\omega v, \beta\omega v, 0, 0]),
$$

where $\alpha = 0.1$ and $\beta = 0.01$ quantify the effect of slip. The evolution of (v, ω) is deterministic: given actuator commands, the change in velocity is known exactly, but uncertainty in positional variables subsists. By defining $z = (v, \omega)$, this problem setting matches the derivations presented in the previous section. We consider minimizing control effort $G(u(s)) = \|u(s)\|_2^2$. The initial state x^0 is fixed, and the final state constraint consists of reaching the goal x^f in expectation, such that $g(x(t_f)) = x(t_f) - x^f$.

5.4. Results. To verify the reliability of our method, we run a batch of experiments with randomized parameters, and sample $x^0_i \sim \mathrm{Unif}(-\bar{x}_i, \bar{x}_i)$, with $\bar{x}_1 = \bar{x}_2 = 2$, $\bar{x}_3 = 5$, $\bar{x}_4 = 0.1$, and $\bar{x}_5 = 0.05$. We set the final condition as $x^f_i \sim x^0_i + \mathrm{Unif}(-\bar{x}_i, \bar{x}_i)$. We also sample fixed final times as $t_f \sim \mathrm{Unif}(6, 30)$. We use a forward Euler discretization scheme with $N = 41$ discretization nodes, set $\Delta_0 = 100$, keep Δ_k constant for all SCP iterations, and use IPOPT to solve each convexified problem. We check convergence of SCP by verifying that $\max_s \|u_{k+1}(s) - u_k(s)\|_\infty / \max_s \|u_{k+1}(s)\|_\infty < \epsilon$, and that the left-hand side of (3.2) goes to zero through iterations. Since $\Delta_k = \Delta_0$, for $k \in \mathbb{N}$, this implies the validity of (3.2), and that the assumptions of Theorem 3.2 are satisfied. We run it on 100 randomized experiments, and report sample problems, iterations, and solutions in Figure 1. We observe that the number of SCP iterations depends on the convergence tolerance ϵ. Its variability is mostly dependent on the varying difficulty of the problems. For these problems, our method converges successfully 100% of the time, although it is initialized with unfeasible initial trajectories.

Further, we consider a cylindrical obstacle of radius 0.2 centered at $r_o = [1.5, 1.9]$. We set an obstacle avoidance constraint using the potential function $c_o : \mathbb{R}^n \to \mathbb{R}$, defined as $c_o(x) = (\|r - r_o\|^2 - (0.2 + \varepsilon)^2$ if $\|r - r_o\| < (0.2 + \varepsilon)$, and 0 otherwise, where $r = [r_x, r_y] \subset x$ and $\varepsilon = 0.1$. We penalize it directly within the cost, defining $L(s, x(s), u(s)) = \lambda c_o(x)$, with $\lambda = 500$. For this problem, SCP converges in 8 iterations, with a trajectory avoiding the obstacle in expectation. Indeed, after evaluating 1000 sample paths of the system, only 14 trajectories intersect with the obstacle.

ϵ	# SCP iter.
10%	4.9 ± 1.5
5%	5.6 ± 1.9
2.5%	6.2 ± 2.1
1%	7.0 ± 2.5
0.1%	8.7 ± 3.0

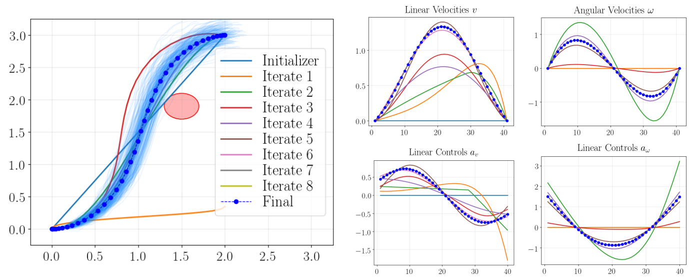

Fig. 1: Left: dependence of the number of SCP iterations on the precision tolerance ϵ, with 1 *standard deviation. Middle: Sample trajectories of a planning problem with obstacle avoidance, and* 100 *sample paths of the resulting trajectories. Right: velocities and control inputs of each SCP iteration of the obstacle avoidance planning problem.*

6. Conclusion and Perspectives. In this paper we introduce and analyze convergences properties for sequential convex programming in a stochastic setting, from which we derive a practical numerical framework to solve non-linear stochastic optimal control problems.

Future work may consider extending this analysis to tackle more general problem formulations, e.g., risk measures as costs, and state (chance) constraints. In this context, some preliminary results using SCP only exist for discrete time problem formulations [24]. However, tackling continuous time formulations will require more sophisticated necessary conditions for optimality. Finally, we plan to further leverage our theoretical insights to design new and more efficient numerical schemes for non-linear stochastic optimal control.

REFERENCES

[1] A. Agrachev and Y. Sachkov, *Control theory from the geometric viewpoint*, vol. 87, Springer Science & Business Media, 2013.

[2] M. Annunziato and A. Borzì, *A fokker–planck control framework for multidimensional stochastic processes*, Journal of Computational and Applied Mathematics, 237 (2013), pp. 487–507.

[3] R. Bellman, *Dynamic Programming*, Princeton Univ. Press, Princeton, New Jersey, 1957.

[4] B. Berret and F. Jean, *Efficient computation of optimal open-loop controls for stochastic systems*, Automatica, 115 (2020).

[5] D. Bertsimas and D. B. Brown, *Constrained stochastic lqc: a tractable approach*, IEEE Transactions on automatic control, 52 (2007), pp. 1826–1841.

[6] C. Bes and S. Sethi, *Solution of a class of stochastic linear-convex control problems using deterministic equivalents*, Journal of optimization theory and applications, 62 (1989), pp. 17–27.

[7] J. Betts, *Survey of numerical methods for trajectory optimization*, Journal of guidance, control, and dynamics, 21 (1998), pp. 193–207.

[8] J.-M. Bismut, *Linear quadratic optimal stochastic control with random coefficients*, SIAM Journal on Control and Optimization, 14 (1976), pp. 419–444.

[9] P. T. Boggs and J. W. Tolle, *Sequential quadratic programming*, Acta numerica, 4 (1995), pp. 1–51.

[10] R. Bonalli, B. Hérissé, and E. Trélat, *Continuity of pontryagin extremals with respect to delays in nonlinear optimal control*, SIAM Journal on Control and Optimization, 57 (2019), pp. 1440–1466.

[11] R. Bonalli, B. Hérissé, and E. Trélat, *Optimal control of endo-atmospheric launch vehicle systems: Geometric and computational issues*, IEEE Transactions on Automatic Control, 65 (2019), pp. 2418–2433.

[12] A. Bryson, *Applied optimal control: optimization, estimation and control*, CRC Press, 1975.
[13] R. Carmona, *Lectures on BSDEs, stochastic control, and stochastic differential games with financial applications*, vol. 1, SIAM, 2016.
[14] T. Damm, H. Mena, and T. Stillfjord, *Numerical solution of the finite horizon stochastic linear quadratic control problem*, Numerical Linear Algebra with Applications, 24 (2017).
[15] R. Gamkrelidze, *Principles of optimal control theory*, vol. 7, Springer Science & Business Media, 2013.
[16] E. Gobet, *Monte-Carlo methods and stochastic processes: from linear to non-linear*, CRC Press, 2016.
[17] P. Kleindorfer and K. Glover, *Linear convex stochastic optimal control with applications in production planning*, IEEE Transactions on Automatic Control, 18 (1973), pp. 56–59.
[18] D. Kuhn, W. Wiesemann, and A. Georghiou, *Primal and dual linear decision rules in stochastic and robust optimization*, Mathematical Programming, 130 (2011), pp. 177–209.
[19] H. J. Kushner, *Numerical methods for stochastic control problems in continuous time*, SIAM Journal on Control and Optimization, 28 (1990), pp. 999–1048.
[20] H. J. Kushner and L. F. Martins, *Numerical methods for stochastic singular control problems*, SIAM journal on control and optimization, 29 (1991), pp. 1443–1475.
[21] H. J. Kushner and F. C. Schweppe, *A maximum principle for stochastic control systems*, Journal of Mathematical Analysis and Applications, 8 (1964), pp. 287–302.
[22] J.-F. Le Gall, *Brownian motion, martingales, and stochastic calculus*, vol. 274, Springer, 2016.
[23] T. Levajković, H. Mena, and L.-M. Pfurtscheller, *Solving stochastic lqr problems by polynomial chaos*, IEEE control systems letters, 2 (2018), pp. 641–646.
[24] T. Lew, R. Bonalli, and M. Pavone, *Chance-constrained sequential convex programming for robust trajectory optimization*, in European Control Conference, 2020.
[25] P.-L. Lions, *Optimal control of diffusion processes and hamilton-jacobi-bellman equations, part i*, Comm. Partial Differential Equations, 8 (1983), pp. 1101–1174.
[26] P. S. Maybeck, *Stochastic models, estimation, and control*, Academic press, 1982.
[27] F. Palacios-Gomez, L. Lasdon, and M. Engquist, *Nonlinear optimization by successive linear programming*, Management science, 28 (1982), pp. 1106–1120.
[28] S. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM Journal on control and optimization, 28 (1990), pp. 966–979.
[29] S. Peng, *Stochastic hamilton–jacobi–bellman equations*, SIAM Journal on Control and Optimization, 30 (1992), pp. 284–304.
[30] L. Pontryagin, *Mathematical theory of optimal processes*, Routledge, 2018.
[31] J. Potter, *A matrix equation arising in statistical filter theory*, Rep. RE-9, Experimental Astronomy Laboratory, Massachusetts Institute of Technology, (1965).
[32] M. A. Rami and X. Y. Zhou, *Linear matrix inequalities, riccati equations, and indefinite stochastic linear quadratic controls*, IEEE Transactions on Automatic Control, 45 (2000), pp. 1131–1143.
[33] R. T. Rockafellar and R. J. B. Wets, *Generalized linear-quadratic problems of deterministic and stochastic optimal control in discrete time*, SIAM Journal on control and optimization, 28 (1990), pp. 810–822.
[34] A. Shapiro and A. Nemirovski, *On complexity of stochastic programming problems*, in Continuous optimization, Springer, 2005, pp. 111–146.
[35] S. Tang, *General linear quadratic optimal stochastic control problems with random coefficients: linear stochastic hamilton systems and backward stochastic riccati equations*, SIAM journal on control and optimization, 42 (2003), pp. 53–75.
[36] E. Trélat, *Some properties of the value function and its level sets for affine control systems with quadratic cost*, Journal of Dynamical and Control Systems, 6 (2000), pp. 511–541.
[37] E. Trélat, *Optimal control and applications to aerospace: some results and challenges*, Journal of Optimization Theory and Applications, 154 (2012), pp. 713–758.
[38] D. D. Yao, S. Zhang, and X. Y. Zhou, *Stochastic linear-quadratic control via semidefinite programming*, SIAM Journal on Control and Optimization, 40 (2001), pp. 801–823.
[39] J. Yong and X. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations*, vol. 43, Springer Science & Business Media, 1999.

Appendix.

6.1. Proof of Lemma 2.1.

Proof of Lemma 2.1. For seek of clarity in the notation, we only consider the case for which $d = 1$, being the case with multivariate Brownian motion similar.

Let us start with the first inequality. For, by denoting $x = x_{u,v,y}$, for every $t \in [0,T]$ we compute (below, $C \geq 0$ is an appropriate constant)

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[0,t]} \|x(s)\|^\ell\right] \leq C\mathbb{E}\left[\|x^0\|^\ell\right] + C\mathbb{E}\left[\sup_{r\in[0,t]} \left\|\int_0^r \sigma(s,y(s))\,\mathrm{d}B_s\right\|^\ell\right] \\
&+ C\mathbb{E}\left[\sup_{r\in[0,t]} \left\|\int_0^r \big(b_0(s,y(s)) + u^i(s)b_i(s,y(s))\big)\,\mathrm{d}s\right\|^p\right] \\
&+ C\mathbb{E}\left[\sup_{r\in[0,t]} \left\|\int_0^r \left(\frac{\partial b_0}{\partial x}(s,y(s)) + v^i(s)\frac{\partial b_i}{\partial x}(s,y(s))\right)(x(s)-y(s))\,\mathrm{d}s\right\|^\ell\right] \\
&+ C\mathbb{E}\left[\sup_{r\in[0,t]} \left\|\int_0^r \frac{\partial \sigma}{\partial x}(s,y(s))(x(s)-y(s))\,\mathrm{d}B_s\right\|^\ell\right].
\end{aligned}$$

For the last term, by denoting $S_\sigma \triangleq \left\{(s,\omega) \in [0,T]\times\Omega : \ (s,y(s,\omega)) \in \mathrm{supp}\ \sigma\right\}$, Burkholder–Davis–Gundy, Hölder and Young inequalities give

$$\begin{aligned}
&\mathbb{E}\left[\sup_{r\in[0,t]} \left\|\int_0^r \frac{\partial \sigma}{\partial x}(s,y(s))(x(s)-y(s))\,\mathrm{d}B_s\right\|^p\right] \\
&\leq C\left(\int_0^t \mathbb{E}\left[\sup_{s\in[0,r]} \|x(r)\|^p\right]\mathrm{d}s + \int_{S_\sigma}\left\|\frac{\partial \sigma}{\partial x}(s,y(s))\right\|^p \|y(s)\|^p\ \mathrm{d}(s\times P)\right) \\
&\leq C\left(1 + \int_0^t \mathbb{E}\left[\sup_{s\in[0,r]} \|x(r)\|^p\right]\mathrm{d}s\right).
\end{aligned}$$

Similar computations apply to the other terms, and when considering solutions to (2.1). Therefore, we conclude from a routine Grönwall inequality argument.

Let us prove the second inequality of the lemma. For $t \in [0,T]$, we compute

$$\begin{aligned}
\mathbb{E}\Bigg[\sup_{s\in[0,t]} \|x_{u_1}(s) - x_{u_2}(s)\|^p\Bigg] &\leq C\mathbb{E}\left[\sup_{s\in[0,t]} \left\|\int_0^r \big(b_0(s,x_{u_1}(s)) - b_0(s,x_{u_2}(s))\big)\,\mathrm{d}s\right\|^p\right] \\
&+ C\mathbb{E}\left[\sup_{s\in[0,t]} \left\|\int_0^r \big(u_1^i(s)b_i(s,x_{u_1}(s)) - u_2^i(s)b_i(s,x_{u_2}(s))\big)\,\mathrm{d}s\right\|^p\right] \\
&+ C\mathbb{E}\left[\sup_{s\in[0,t]} \left\|\int_0^r \big(\sigma(s,x_{u_1}(s)) - \sigma(s,x_{u_2}(s))\big)\,\mathrm{d}B_s\right\|^p\right].
\end{aligned}$$

For the second term on the right-hand side, Hölder inequality gives

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[0,t]} \left\|\int_0^r \big(u_1^i(s)b_i(s,x_{u_1}(s)) - u_2^i(s)b_i(s,x_{u_2}(s))\big)\,\mathrm{d}s\right\|^p\right] \leq \\
&\leq C\left(\int_0^t \mathbb{E}\left[\sup_{r\in[0,s]} \|x_{u_1}(r) - x_{u_2}(r)\|^p\right]\mathrm{d}s + \mathbb{E}\left[\left(\int_0^T \|u_1(s)-u_2(s)\|\,\mathrm{d}s\right)^p\right]\right),
\end{aligned}$$

and similar computations hold for the remaining terms, and when considering. solutions to (2.1). Again, we conclude by a Grönwall inequality argument. $\square$

6.2. Proof of Lemma 4.1. We start with a preliminary result.

LEMMA 6.1. *Let* $(\eta_1,\dots,\eta_j)\in \Pr_{\mathbb{R}^j_+}(\mathcal{C}^j_t)$ *(in particular, no* (A_2) *or any assumption on* $t>t_j$ *are required). For* $\varepsilon\in[0,t-t_j)$*, uniformly for* $\delta\in[-\varepsilon,T-t]$*,*

$$\mathbb{E}\left[\left\|\tilde g(\tilde x_{u_\pi}(t+\delta))-\tilde g(\tilde x_u(t+\delta))-\sum_{i=1}^j\eta_i\frac{\partial\tilde g}{\partial\tilde x}(\tilde x_u(t+\delta))z_{t_i,u_i}(t+\delta)\right\|^2\right]^{\frac12}=o\left(\sum_{i=1}^j\eta_i\right).$$

Proof. We only consider the case $j=1$, being the most general case $j>1$ done by adopting a classical induction argument (see, e.g., [1]). We only need to prove that

$$\mathbb{E}\left[\|\tilde x_{u_\pi}(t+\delta)-\tilde x_u(t+\delta)-\eta_1 z_{t_1,u_1}(t+\delta)\|^2\right]=o(\eta_1^2)\tag{6.1}$$

uniformly for every $\delta\in[-\varepsilon,T-t]$. For, first we remark that $\tilde x_{u_\pi}(t)=\tilde x_u(t)$, for $t\in[0,t_1]$. Since $t\in(t_j,t_f]$ and $\varepsilon\geq 0$ are fixed, and we take the limit $\eta_1\to 0$, we may assume that $\eta_1+t_1<t-\varepsilon$. Therefore, without loss of generality, we replace $t+\delta$ with t, assuming $t\geq t_1+\eta_1$, uniformly. We have (below, $C\geq 0$ denotes a constant)

$$\begin{aligned}
\mathbb{E}\left[\|\tilde x_{u_\pi}(t)-\tilde x_u(t)-\eta_1 z_{t_1,u_1}(t)\|^2\right]&\leq C\mathbb{E}\left[\left\|\int_{t_1}^{t_1+\eta_1}\eta_1A(s)z_{t_1,u_1}(s)\mathrm{d}s\right\|^2\right]\\
&+C\mathbb{E}\left[\left\|\int_{t_1+\eta_1}^{t}\left(\tilde b(s,u_\pi(s),x_{u_\pi}(s))-\tilde b(s,u(s),x_u(s))-\eta_1A(s)z_{t_1,u_1}(s)\right)\mathrm{d}s\right\|^2\right]\\
&+C\mathbb{E}\left[\left\|\int_0^t\mathbb{1}_{(t_1,T]}(s)\left(\tilde\sigma(s,x_{u_\pi}(s))-\tilde\sigma(s,x_u(s))-\eta_1D(s)z_{t_1,u_1}(s)\right)\mathrm{d}B_s\right\|^2\right]\\
&+C\mathbb{E}\left[\left\|\int_{t_1}^{t_1+\eta_1}\left(\tilde b(s,u_\pi(s),x_{u_\pi}(s))-\tilde b(s,u(s),x_u(s))\right)\mathrm{d}s-\eta_1z_{t_1,u_1}(t_1)\right\|^2\right].
\end{aligned}$$

Let us analyze those integrals separately. Starting with the last one, we have

$$\begin{aligned}
&\mathbb{E}\left[\left\|\int_{t_1}^{t_1+\eta_1}\left(\tilde b(s,u_\pi(s),x_{u_\pi}(s))-\tilde b(s,u(s),x_u(s))\right)\mathrm{d}s-\eta_1z_{t_1,u_1}(t_1)\right\|^2\right]\leq\\
&\leq C\mathbb{E}\left[\left\|\int_{t_1}^{t_1+\eta_1}\left(\tilde b(s,u_1,x_{u_\pi}(s))-\tilde b(s,u_1,x_u(s))\right)\mathrm{d}s\right\|^2\right]\\
&\quad+C\mathbb{E}\left[\left\|\int_{t_1}^{t_1+\eta_1}\left(\tilde b(s,u_1,x_u(s))-\tilde b(s,u(s),x_u(s))\right)\mathrm{d}s-\eta_1z_{t_1,u_1}(t_1)\right\|^2\right].
\end{aligned}$$

Lemma 2.1 immediately gives that the first term on the right-hand side is $o(\eta_1^2)$. In addition, from Hölder inequality, it follows that

$$\begin{aligned}
&\frac{1}{\eta_1^2}\mathbb{E}\left[\left\|\int_{t_1}^{t_1+\eta_1}\left(\tilde b(s,u_1,x_u(s))-\tilde b(s,u(s),x_u(s))\right)\mathrm{d}s-\eta_1z_{t_1,u_1}(t_1)\right\|^2\right]\leq\\
&\leq\frac{C}{\eta_1}\int_{t_1}^{t_1+\eta_1}\mathbb{E}\left[\left\|\tilde b(s,u_1,x_u(s))-\tilde b(t_1,u_1,x_u(t_1))\right\|^2\right]\mathrm{d}s\\
&\quad+\frac{C}{\eta_1}\int_{t_1}^{t_1+\eta_1}\mathbb{E}\left[\left\|\tilde b(s,u(s),x_u(s))-\tilde b(t_1,u(t_1),x_u(t_1))\right\|^2\right]\mathrm{d}s.
\end{aligned}$$

Since t_1 is a Lebesgue point for u, the two terms above go to zero whenever $\eta_1 \to 0$. Next, by the BurkholderDavisGundy inequality and a Taylor development, we have

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\int_0^t \mathbb{1}_{(t_1,T]}(s)\left(\tilde{\sigma}(s,x_{u_\pi}(s)) - \tilde{\sigma}(s,x_u(s)) - \eta_1 D(s) z_{t_1,u_1}(s)\right) \mathrm{d}B_s\right\|^2\right] \\
&\leq C\mathbb{E}\left[\int_{t_1}^t \left\|D(s)\left(\tilde{x}_{u_\pi}(s) - \tilde{x}_u(s) - \eta_1 z_{t_1,u_1}(s)\right)\right\|^2 \mathrm{d}s\right] \\
&\quad + C\mathbb{E}\left[\int_{t_1}^t \left\|\int_0^1 \theta \frac{\partial^2 \tilde{\sigma}}{\partial \tilde{x}^2}\left(s, \theta \tilde{x}_u(s) + (1-\theta)(\tilde{x}_{u_\pi} - \tilde{x}_u)(s)\right)\left(\tilde{x}_{u_\pi} - \tilde{x}_u\right)^2(s)\, \mathrm{d}\theta\right\|^2 \mathrm{d}s\right] \\
&= C\int_{t_1}^t \mathbb{E}\left[\|\tilde{x}_{u_\pi}(s) - \tilde{x}_u(s) - \eta_1 z_{t_1,u_1}(s)\|^2\right]\, \mathrm{d}s + o(\eta_1^2),
\end{aligned}
$$

thanks to Lemma 2.1. Similar estimates hold for the remaining terms in the first inequality of the proof. Summarizing, there exists a constant $C \geq 0$ such that

$$
\mathbb{E}\Big[\|\tilde{x}_{u_\pi}(t) - \tilde{x}_u(t) - \eta_1 z_{t_1,u_1}(t)\|^2\Big] \leq o(\eta_1^2) + C\int_{t_1}^t \mathbb{E}\left[\|\tilde{x}_{u_\pi}(s) - \tilde{x}_u(s) - \eta_1 z_{t_1,u_1}(s)\|^2\right]\, \mathrm{d}s
$$

and the conclusion follows from a routine Gronwäll inequality argument. □

Proof of Lemma 4.1. Whenever $\delta = 0$, the claim follows from Lemma 6.1. Therefore, assume $\delta \neq 0$ and (A_2). Hence, $\tilde{g}$ is an affine function, and in the rest of this proof we will be denoting $\tilde{g}(\tilde{x}) = M\tilde{x} + d$, where $M \in \mathbb{R}^{(q+1)\times(n+1)}$ and $d \in \mathbb{R}^{q+1}$.

Developing, we have

$$
\begin{aligned}
&\left\|\mathbb{E}\Big[\tilde{g}(\tilde{x}_{u_\pi}(t+\delta)) - \tilde{g}(\tilde{x}_u(t)) - \delta M\tilde{b}(t,u(t),x_u(t)) - \sum_{i=1}^j \eta_i A z_{t_i,u_i}(t)\Big]\right\| \leq \\
&\quad \leq \left\|\mathbb{E}\Big[\tilde{g}(\tilde{x}_u(t+\delta)) - \tilde{g}(\tilde{x}_u(t)) - \delta M\tilde{b}(t,u(t),x_u(t))\Big]\right\| \\
&\quad\quad + \left\|\mathbb{E}\Big[\tilde{g}(\tilde{x}_{u_\pi}(t+\delta)) - \tilde{g}(\tilde{x}_u(t+\delta)) - \sum_{i=1}^j \eta_i M z_{t_i,u_i}(t+\delta)\Big]\right\| \\
&\quad\quad + \sum_{i=1}^j \eta_i \|M\| \left\|\mathbb{E}\Big[z_{t_i,u_i}(t+\delta) - z_{t_i,u_i}(t)\Big]\right\|.
\end{aligned}
$$

Thanks to Lemma 6.1, the second term on the right-hand side is $o\left(\sum_{i=1}^j \eta_i\right)$. For the last summand, from the property of the stochastic integral, we have

$$
\begin{aligned}
&\left\|\mathbb{E}\Big[z_{t_i,u_i}(t+\delta) - z_{t_i,u_i}(t)\Big]\right\| \leq \left\|\mathbb{E}\left[\int_t^{t+\delta} A(s) z_{t_i,u_i}(s)\, \mathrm{d}s\right]\right\| \\
&\quad + \left\|\mathbb{E}\left[\int_0^{t+\delta} \mathbb{1}_{[t,T]}(s) D(s) z_{t_i,u_i}(s)\, \mathrm{d}B_s\right]\right\| \leq C\mathbb{E}\left[\sup_{s\in[0,T]} \|z_{t_i,u_i}(s)\|\right]\delta,
\end{aligned}
$$

and the last term is $o\left(\delta + \sum_{i=1}^j \eta_i\right)$. It is worth pointing out the importance for g to be affine to provide the last claim. Finally, for the remaining term, we apply Itô formula to each coordinate $h = 1, \ldots, q+1$, obtaining

$$\Big[\tilde{g}(\tilde{x}_u(t+\delta)) - \tilde{g}(\tilde{x}_u(t)) - \delta M\tilde{b}(t, u(t), x_u(t))\Big]_h = \sum_{k=1}^{n+1} M_{hk} \int_0^{t+\delta} \mathbb{1}_{[t,T]}(s)\tilde{\sigma}_k(s, x_u(s)) \, \mathrm{d}B_s$$
$$+ \sum_{k=1}^{n+1} M_{hk} \left(\int_t^{t+\delta} \tilde{b}_k(s, u(s), x_u(s)) \, \mathrm{d}s - \delta\tilde{b}_k(t, u(t), x_u(t)) \right),$$

and therefore

$$\frac{1}{\delta + \sum_{i=1}^{j} \eta_i} \Big\| \mathbb{E}\Big[\tilde{g}(\tilde{x}_u(t+\delta)) - \tilde{g}(\tilde{x}_u(t)) - \delta M\tilde{b}(t, u(t), x_u(t))\Big] \Big\| \leq$$
$$\leq \frac{\|M\|}{\delta} \int_t^{t+\delta} \mathbb{E}\left[\Big\| \tilde{b}_k(s, u(s), x_u(s)) - \tilde{b}_k(t, u(t), x_u(t)) \Big\| \right] \, \mathrm{d}s.$$

But t is Lebesgue for u, thus this last quantity goes to zero for $\delta \to 0$, an we conclude.$\square$