Sample Average Approximations in Optimal Control of Uncertain Systems

Chris Phelps, Johannes O. Royset and Qi Gong

Abstract—This paper focuses on an optimal control problem in which the objective is to minimize the expectation of a cost functional with stochastic parameters. The inclusion of the stochastic parameters in the objective raises new theoretical and computational challenges not present in a standard nonlinear optimal control problem. In this paper, we provide a numerical framework for the solution of this uncertain optimal control problem by taking a sample average approximation approach. An independent random sample is taken from the parameter space, and the expectation is approximated by the sample average. The result is a family of standard nonlinear optimal control problems which can be solved using existing techniques. We provide an optimality function for both the uncertain optimal control problem and its approximation, and show that the approximation based on the sample average approach is consistent in the sense of Polak. We illustrate the uncertain control problem with a numerical example arising in optimal search for a moving target.

I. INTRODUCTION

In this paper we consider an optimal control problem which incorporates uncertainty in the objective function. The goal is to find a control for a constrained nonlinear system which minimizes the expectation of a cost functional that is dependent on a set of stochastic parameters. This framework can be used to model optimization problems in which there is uncertainty about the cost, such as aircraft routing [1], combat modelling [2], and search for a moving target [3], [4]. Given the difficulty in solving standard nonlinear optimal control problems, the inclusion of the stochastic parameters in the objective makes this uncertain optimal control problem particularly challenging. Existing computational and theoretical techniques from the field of optimal control must be extended to this class of problems.

Studies into this uncertain optimal control problem arise in the field of optimal search, where the objective is to find the optimal trajectory for a searcher attempting to detect a non-evading moving target. Previous works focus on special cases which incorporate assumptions such as simplified searcher dynamics [5]–[7] or target dynamics [8]. We consider a broader class of problems with constrained nonlinear dynamics and a general form of the objective functional. Recent work in this area has focused on developing a computational framework to numerically solve the uncertain optimal control problem. Such a framework is provided in [9]–[11], wherein a quadrature-based numerical integration scheme is applied to approximate the expectation in the objective functional. However, this approach is limited, as quadrature-based numerical integration methods are known to be computationally expensive when applied to high-dimensional spaces.

In this paper, we develop a computational framework for the solution of the uncertain optimal control problem based on a sample average approximation method. First, a random, independently distributed sample is taken from the space of stochastic parameters. Then the expectation in the objective functional of the control problem is approximated by the sample average. The resulting family of standard optimal control problems can be solved using existing optimal control algorithms, such as Euler (see [12], Chapter 4), Runge-Kutta [13], [14], and Pseudospectral [15], [16] methods. The use of Monte Carlo methods for integration allows us to avoid computational issues that arise when applying quadrature-based numerical integration schemes to high-dimensional spaces. We refer to [17], [18] for early work on the sample average approximation approach to stochastic optimization, which also provides our foundation. For a treatment of cases in finite dimensions; see [19].

To ensure meaningful results in the proposed computational framework, it is necessary to demonstrate that the sample average scheme produces a consistent approximation of the original uncertain optimal control problem. Even for standard optimal control problems, there are counterexamples showing that an inappropriately designed discretization scheme may not be convergent [20]. We derive optimality functions which provide necessary conditions for optimality for both the original and approximate problems. In addition, we employ an extension of the strong law of large numbers to lower semi-continuous functions (see [17], [18]) to show that the approximation by sample averages is consistent in the sense of Polak [12], Section 3.3. Such a consistency property guarantees that accumulation points of a sequence of optimal solutions to the approximate problem are optimal solutions to the original problem. In addition, accumulation points of a sequence of stationary points of the optimality function of the approximate problem are stationary points for the original problem. Our framework also allows for convex pointwise constraints on the control.

The paper is organized as follows: Section II introduces the uncertain optimal control problem which is the focus of this work. Section III introduces a family of approximate problems based on a sample average scheme, and discusses the convergence properties of the approximation. Section IV introduces optimality functions for both the original and approximate problems. Section V shows that the approximation
using sample averages is consistent in the sense of Polak [12], Section 3.3, which is the main result of the paper. In Section VI, the results are applied to the problem of detecting an intruder in a channel.

II. FORMULATION OF THE UNCERTAIN CONTROL PROBLEM

Before we define the control problem which is the focus of this work, we introduce the spaces on which we will conduct our analysis. To develop optimality conditions, we make use of an inner product on the space of decision variables. Therefore we work in the $L_2$ topology. Let $L^2_2[0,1]$ be the space of all functions $v : [0,1] \mapsto \mathbb{R}^m$ such that $\int_0^1 \|v(t)\|^2 dt < \infty$. We carry out our analysis in a subspace of the Hilbert space

$$H_2 = \mathbb{R}^n \times L^2_2[0,1],$$

where the inner product and norm on $H_2$ are defined for any $\eta = (\xi^n,u^n), \eta' = (\xi'^n,u'^n) \in H_2$ by

$$\langle \eta, \eta' \rangle_{H_2} = \langle \xi^n, \xi'^n \rangle + \langle u^n, u'^n \rangle_2.$$

Therefore the norm in $H_2$ is given by

$$\|\eta\|^2_{H_2} = \|\xi^n\|^2 + \|u^n\|^2_2.$$

Given compact, convex sets $\Xi \subset \mathbb{R}^n$ and $U \subset \mathbb{R}^m$, we define the set of admissible controls

$$U = \{ u \in L^m_2[0,1] : u(t) \in U \text{ for almost every } t \in [0,1] \}.$$ 

The set of all admissible initial state-control-pairs is then given by $H = \{(\xi,u) : \xi \in \Xi, u \in U \}$. This set is a subset of the pre-Hilbert space $H_{\infty,2} = \{(\xi,u) \in H_2 \cap \{ \|u\|_\infty < \infty \}$. Here we have defined the admissible set differently than in Polak [12], Chapter 4, which requires the pointwise control constraint to be satisfied for all $t \in [0,1]$. However, for each $u \in L^m_2[0,1]$, there is a member $\tilde{u}$ of its equivalence class such that $u(t) \in U$ for every $t \in [0,1]$. Therefore we can apply the standard results from the theory of differential equations to controls from our admissible set.

In developing optimality conditions, we evaluate derivatives with respect to the decision variable $\eta$. In order to guarantee these derivatives exist, we work on the slightly larger space $H^0$. Let $\rho_1, \rho_2 \in \mathbb{R}$ be constants large enough so that $\|\xi^n\| < \rho_1, \|u^n\|_\infty < \rho_2$, for all $\eta \in H$. The existence of these constants is guaranteed by the compactness of $\Xi$ and $U$. We define the larger space $H^0 = \{(\xi,u) \in \mathbb{R}^n \times L^2_2[0,1] \cap \|\xi\| < \rho_1, \|u\|_\infty < \rho_2 \}$. The space $H^0$ is then open in the $L_\infty$ topology and the inclusion $H \subset H^0$ holds. It is important to note that all convergence results on the sets $H$ and $H^0$ are with respect to the $L_2$ topology.

We can now state the uncertain optimal control Problem $B$: find an initial state and control pair $\eta \in H$ to minimize the objective functional

$$J(\eta) = \mathbb{E}^P \left[ G \left( \int_0^1 r(x^n(t), u^n(t), t, \omega) dt \right) \right], \quad (1)$$

where $x^n(t)$ is the solution to the differential equation

$$\frac{d}{dt} x(t) = h(x(t), u^n(t)), \quad t \in [0,1], \quad x(0) = \xi^n, \quad (2)$$

which is assumed to be unique. Here $\omega \in \Omega$, with $\{\Omega, \Sigma, P\}$ a probability space where $\Sigma$ is $P$-complete, and $\mathbb{E}^P$ is the expectation on $\Omega$ with respect to the probability measure $P$. The functions $r : \mathbb{R}^n \times \mathbb{R}^m \times [0,1] \times \Omega \mapsto R, G : R \mapsto R$ and $h : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$.

Problem $B$ differs from the standard Bolza problem in the inclusion of an expectation over the space of stochastic parameters in (1), which raises new theoretical and computational challenges. The focus of this paper is the development of a computational framework for solving Problem $B$ based on sample average approximations. This framework avoids computational difficulties that arise when attempting to apply previous quadrature-based approaches ([9]–[11]) to problems with high-dimensional parameter spaces. In order to conduct theoretical analysis of this computational method, we assume the following regularity conditions.

Assumption 1. There exists a compact set $X_0 \subset \mathbb{R}^m$ such that for each $\eta \in H^0$, $x^n(t) \in X_0$ for all $t \in [0,1]$, where $x^n$ is the solution to (2) for $\eta = (u^n, \xi^n)$.

This assumption essentially requires that there is no admissible control for which the dynamical system has a finite escape time. This assumption will be satisfied for a variety of systems frequently encountered in control problems, such as input-to-state stable systems.

Assumption 2. For the set $X_0$ defined in Assumption 1 and the set $V = \{ v \in \mathbb{R}^m \cap \|v\| < \rho_2 \}$, the function $h$ is continuously differentiable on $X_0 \times V$ and there exists a constant $K \in [1, \infty)$ such that for all $x', x'' \in X_0$, and $v', v'' \in V$, the following inequalities hold:

$$\|h(x', v') - h(x'', v'')\| \leq K \|(x'-x'')\| + \|(v'-v'')\|, \quad (3)$$

$$\|h_x(x', v') - h_x(x'', v'')\| \leq K \|x'-x''\| + \|v'-v''\|, \quad (4)$$

$$\|h_u(x', v') - h_u(x'', v'')\| \leq K \|x'-x''\| + \|v'-v''\|.$$ 

Assumption 3. For the set $X_0$ defined in Assumption 1 and the set $V$ defined in Assumption 2, the function $r(\cdot, v', \cdot, \cdot)$ is continuously differentiable on $X_0 \times V \times [0,1]$ for each $\omega \in \Omega$, and $r(x, v, t, \cdot)$ is measurable and uniformly bounded for each $x \in X_0, v \in V, t \in [0,1]$. There exists a constant $L_r \in [1, \infty)$ such that for all $x', x'' \in X_0, v', v'' \in V$ the following inequalities hold for every $t \in [0,1], \omega \in \Omega$:

$$\|r(x', v', t, \omega) - r(x'', v'', t, \omega)\| \leq L_r \|x'-x''\| + \|v'-v''\|,$$ 

$$\|r_x(x', v', t, \omega) - r_x(x'', v'', t, \omega)\| \leq L_r \|x'-x''\| + \|v'-v''\|,$$ 

$$\|r_u(x', v', t, \omega) - r_u(x'', v'', t, \omega)\| \leq L_r \|x'-x''\| + \|v'-v''\|.$$ 

Furthermore, $G$ is continuously differentiable and Lipschitz continuous with Lipschitz constant $L_G$. 

1959
Assumptions 2 and 3 require that the functions \( r \) and \( h \), and their derivatives, are Lipschitz continuous on the set of all admissible states and controls.

III. APPROXIMATION OF THE UNCERTAIN CONTROL PROBLEM

In this section we introduce a family of optimal control problems, Problem \( B^M \), which approximates Problem \( B \). For a given \( M \), we take an independent \( P \)-distributed sample \( \{\omega_1, \omega_2, \ldots, \omega_M\} \) from the parameter space \( \Omega \), and approximate the objective functional (1) by the sample average

\[
J^M(\eta) = \frac{1}{M} \sum_{i=1}^{M} G \left( \int_0^1 r(x^n(t), u^n(t), t, \omega_i)dt \right). \tag{3}
\]

Problem \( B^M \) can then be stated as follows: find \( \eta \in H \) to minimize the objective functional (3), where \( x^n \) is the solution to the differential equation (2). This approximate problem can be transformed into a standard optimal control problem (see [9]–[11]), which can be solved using existing techniques [12]–[16]. This allows us to leverage previous work on the numerical solution of nonlinear constrained optimal control problems in the solution of Problem \( B \). For such an approach to work, it is necessary to show that accumulation points of a sequence of optimal solutions to the approximate problem are optimal solutions to the original problem. To carry out this analysis, we first introduce results from the fields of variational analysis and stochastic programming.

A. Preliminary Results

The concept of epiconvergence provides a natural framework to analyze the approximation of an optimization problem, as it allows us to discuss the convergence of the inf and arg min operators.

**Definition 1.** [21] Let \( (X, d) \) be a separable complete metric space. Consider the sequence of lower semi-continuous functions \( f^M : X \mapsto \mathbb{R} \). We say that \( f^M \) epiconverges to \( f \), denoted \( f^M \xrightarrow{\text{epi}} f \), if and only if

i) \( \liminf_{M \to \infty} f^M(x_M) \geq f(x) \) whenever \( x_M \to x \),

ii) \( \lim f^M(x_M) = f(x) \) for at least one sequence \( x_M \to x \).

**Proposition 1.** [21, Theorem 2.5] Theorem 2.5. Let \( (X, d) \) be a separable complete metric space. Consider the sequence of lower semi-continuous functions \( f^M : X \mapsto \mathbb{R} \). Suppose that \( f^M \) epiconverges to \( f \). If \( \{x^M\}_{M \in \mathbb{N}} \subseteq X \) is a sequence of global minimizers to \( f^M \), and \( \bar{x} \) is any accumulation point of this sequence (along a subsequence indexed by a set \( K \subseteq \mathbb{N} \)), then \( \bar{x} \) is a global minimizer of \( f \) and

\[
\liminf_{M \to \infty} f^M(x_M) = \inf_{x \in X} f(x).
\]

Because (1) involves an expectation over a space of stochastic, we use results on random lower semi-continuous functions to establish these epiconvergence properties.

**Definition 2.** [18] Let \( (X, d) \) be a separable complete metric space with \( \mathcal{B} \) the Borel field generated by the open subsets of \( X \). Let \( P \) be a probability measure on the measurable space \( (\Omega, \Sigma) \) such that \( \Sigma \) is \( P \)-complete. A function \( f : X \times \Omega \mapsto \mathbb{R} \) is a random lower semi-continuous if and only if:

i) for all \( \omega \in \Omega \), the function \( x \mapsto f(x, \omega) \) is lower semi-continuous,

ii) \( (x, \omega) \mapsto f(x, \omega) \) is \( \mathcal{B} \otimes \Sigma \) measurable.

In probability theory, the strong law of large numbers guarantees the almost sure convergence of the sample average as the number of samples drawn approaches infinity. The following proposition extends this result to random lower semi-continuous functions.

**Proposition 2.** [17, Theorem 2.3] Let \( (\Omega, \Sigma, P) \) be a probability space such that \( \Sigma \) is \( P \)-complete. Let \( (X, d) \) be a separable complete metric space. Suppose that the function \( f : X \times \Omega \mapsto \mathbb{R} \) is a random lower semi-continuous function and there exists an integrable function \( a_0 : \Omega \mapsto \mathbb{R} \) such that \( f(x, \omega) \geq a_0(\omega) \) almost surely. Let \( \{\omega_1, \ldots, \omega_M\} \) be an independent \( P \)-distributed random draw and define

\[
\hat{f}(x, \omega_1, \ldots, \omega_M) = \frac{1}{M} \sum_{i=1}^{M} f(x, \omega_i).
\]

Then, as \( M \to \infty, \hat{f}(x, \omega_1, \ldots, \omega_M) \) epiconverges almost surely to \( E^P f(x, \omega) \).

B. The augmented state \( z(t, \eta, \omega) \)

In this section we introduce an augmented state which allows simplification so that we can express the objective functional \( J \) as an expectation of a random lower semi-continuous function. With this goal in mind, we introduce the function \( z : [0, 1] \times H^0 \times \Omega \mapsto \mathbb{R} \) such that

\[
z(t, \eta, \omega) = \int_0^t r(x^\eta(s), u^\eta(s), s, \omega)ds. \tag{4}
\]

This integral is finite by Assumption 3. To further simplify the notation, let us introduce the functional \( G : H^0 \times \Omega \mapsto \mathbb{R} \) by

\[
G(\eta, \omega) = G(z(1, \eta, \omega)).
\]

The continuity and measurability of \( z \) is established using the following lemma. This demonstrates that \( G \) is a random lower semi-continuous function.

**Lemma 1.** Suppose that Assumption 1 is satisfied, and let \( V \) be the set defined in Assumption 2. Let \( \kappa : \mathbb{R}^l \times X \times \Omega \mapsto \mathbb{R}^l \) be such that \( \kappa(\cdot, \cdot, \cdot) \) is continuously differentiable for each \( \omega \in \Omega \) and \( \kappa(x, u, \cdot) \) is measurable for each \( x \in \mathbb{R}^l, u \in V \). Suppose also that there exists a \( K \subseteq [1, \infty) \) such that for every \( x, x' \in \mathbb{R}^n \), and \( u, u' \in V \), and \( \omega \in \Omega \),

\[
\|\kappa(x, u, \omega) - \kappa(x', u', \omega)\| \leq K \left( \|x - x'\| + \|u - u'\| \right).
\]

For each \( \eta = (\xi^n, u^n) \in H^0, \omega \in \Omega \), let \( \chi^n : [0, 1] \times \Omega \mapsto \mathbb{R}^l \) be the solution to

\[
\dot{\chi}^n(t, \omega) = \kappa(\chi^n(t, \omega), u^n(t), \omega), \quad \chi^n(0) = \xi^n,
\]

which is assumed to be unique. Then \( \chi^n \) is measurable and uniformly bounded.

**Proof.** See Appendix.
Remark 1. Note that a dynamical system of the form
\[ \dot{x}(t) = g(x(t), u(t), t), \quad x(0) = x_0, \] (5)
where \( g \) is continuously differentiable with respect to \( t \), can be transformed into the form \( \dot{x}^T = \bar{g}(\bar{x}(t), u(t)) \) by letting \( \dot{\bar{x}} = \dot{x} + \bar{X} \). Therefore Lemma 1 can be applied to a dynamical system of the form (5).

Proposition 3. Suppose that Assumptions 1-3 are satisfied. Then \( z(t, \cdot, \omega) : H^0 \to \mathbb{R} \) defined in (4) is Lipschitz for each \( t \in [0,1], \omega \in \Omega \), with Lipschitz constant \( L_z \). Furthermore, \( z(t, \cdot, \cdot) : H^0 \times \Omega \to \mathbb{R} \) is \( B(H^0) \otimes \Sigma \) measurable for each \( t \in [0,1] \), where \( B(H^0) \) is the Borel sigma field on \( H^0 \).

Proof. The Lipschitz continuity follows from Assumptions 1-3 and Lemma 5.6.7 of [12]. Because the Lipschitz constant \( L_z \) does not depend on \( \omega \), neither does \( L_z \). By Lemma 1 and Remark 1, \( z(t, \eta, \cdot) : \Omega \to \mathbb{R} \) is measurable for each \( t \in [0,1], \eta \in H^0 \). Then \( z(t, \cdot, \cdot) \) is a Carathéodory function for each \( t \in [0,1] \) and is therefore \( B(H^0) \otimes \Sigma \) measurable by Lemma 4.51 of [22].

C. Epiconvergence of \( J^M \) to \( J \)

Let \( \{\omega_1, \omega_2, \ldots, \omega_M\} \) be an independent \( P \)-distributed draw from \( \Omega \). Using the notation introduced in the previous section, we now restate our optimal control problems.

Problem B: Find the initial state and control pair \( \eta \in H \) to minimize the objective functional
\[ J(\eta) = \mathbb{E}_P^\mu \left[ \bar{G}(\eta, \omega) \right]. \] (6)

Problem \( B^M \): Find the initial state and control pair \( \eta \in H \) to minimize the objective functional
\[ J^M(\eta) = \frac{1}{M} \sum_{i=1}^{\infty} \bar{G}(\eta, \omega_i). \] (7)

We now use the results on convergence of random lower semi-continuous functions to address the convergence of \( J^M \to J \) as \( M \to \infty \). First we introduce the following Lemma.

Lemma 2. The space \( H \) is a complete, separable metric space.

Proof. Proof is omitted due to space constraints.

Theorem 1. Suppose that Assumptions 1-3 hold. Then \( J^M \) epiconverges almost surely to \( J \) on \( H \) as \( M \to \infty \).

Proof. By Lemma 2, \( H \) is a complete separable metric space. By Proposition 3, \( \bar{G} \) is a random lower semi-continuous function and for each \( \omega \in \Omega \), \( z(1, \cdot, \omega) \) is Lipschitz with constant \( L_z \) independent of \( \omega \). Therefore \( \bar{G}(\cdot, \cdot) \) is Lipschitz continuous with constant \( L_G + L_z \) for each \( \omega \in \Omega \). Because \( H \) is bounded, this implies that there exists \( a_0 \in R \) such that \( \bar{G}(\eta, \omega) \geq a_0 \) for each \( \eta \in H, \omega \in \Omega \). The result then follows from Proposition 2.

The epiconvergence of the approximated objective functional is ensured by Theorem 1, which is an essential property to demonstrate before using the proposed computation framework to solve Problem B.

IV. Optimality Conditions

Optimality functions provide necessary conditions which must be satisfied by a solution to an optimal control problem. Such a necessary condition is useful in assessing the optimality of a numerically computed solution. In this section we derive optimality functions for Problems B and \( B^M \) in the sense of Polak [12], Chapter 4.

Definition 3. An upper semi-continuous function \( \theta : X \to \mathbb{R} \) is an optimality function for a problem B if:
\begin{enumerate}
  \item \( \theta(x) \leq 0 \) for all \( x \in X \).
  \item If \( x \) is a local minimizer of \( B \), then \( \theta(x) = 0 \).
\end{enumerate}

Before discussing the optimality conditions, we make an additional assumption about \( r_x \).

Assumption 4. Let \( X_0 \) and \( V \) be defined as in Assumptions 1-2. For all \( x \in X_0, v \in V, t \in [0,1], r_x(x, v, t, \cdot) : \Omega \to \mathbb{R} \) is measurable and bounded.

Note that this assumption is valid for the scenario considered in Section VI, wherein the parameter space \( \Omega \) is a compact subspace of \( \mathbb{R}^n \) and the function \( r_x \) is continuous with respect to \( \omega \). For each \( \omega \in \Omega, t \in [0,1], \eta \in H^0 \), we define a vector \( \tilde{x}^\eta(t, \omega) \in \mathbb{R}^{n+1} \), containing the state vector and the augmented state vector. That is
\[ \tilde{x}^\eta(t, \omega) = \left[ x^\eta(t)^T, z(t, \eta, \omega) \right]^T. \]

Similarly we define the dynamics \( \tilde{h} : \mathbb{R}^{n+1} \times \mathbb{R}^m \times [0,1] \times \Omega \) by
\[ \tilde{h}(\tilde{x}^\eta(t, \omega), u^\eta(t), \omega) = \left[ h(x^\eta(t), u^\eta(t))^T, r(x^\eta(t), u^\eta(t), t, \omega) \right]^T. \]

Then for each \( \omega \in \Omega \), \( \tilde{x}^\eta(t, \omega) \) is the solution to the dynamical system
\[ \dot{\tilde{x}}^\eta(t, \omega) = \tilde{h}(\tilde{x}^\eta(t), u(t), \omega), \quad \tilde{x}(0, \omega) = (\xi, 0)^T. \] (8)

Proposition 4. Suppose that Assumptions 1-4 are satisfied.
\begin{enumerate}
  \item For any \( \omega \in \Omega, \eta \in H^0 \) and \( \delta \in H_{\infty,2}, \bar{G}(\cdot, \omega) \) has a Gateaux derivative \( D\bar{G}(\eta, \cdot, \omega) \) at \( \eta \) given by

  \[ \langle \nabla_{\eta} \bar{G}(\eta, \omega), \delta \rangle_H = \left( \nabla_{\eta} \bar{G}(\eta, \omega), \nabla_{u} \bar{G}(\eta, \omega) \right)^T \in H_{\infty,2} \]

  and \( \nabla_{\eta} \bar{G}(\eta, \omega) \) is the solution to the adjoint equation

  \[ p^\eta(s, \omega) = -\bar{h}_z(\tilde{x}^\eta(s, \omega), u(s), \omega)p^\eta(s, \omega), \quad s \in [0,1] \]

  and \( p^\eta(s, \omega) = [0, \ldots, 0, \nabla G(z(1, \eta, \omega))]^T. \] (11)
\end{enumerate}
ii) The gradient $\nabla_\eta \tilde{G}(\cdot, \omega)$ is Lipschitz continuous on $H$.

iii) For any $\eta \in H^0$ and $\delta \eta \in H_{\infty, 2}$, $\tilde{G}(\cdot, \omega)$ has a Frechet differential $D\tilde{G}(\eta; \delta \eta; \omega)$ at $\eta$.

Proof. The proposition follows directly from Corollary 5.6.9 of [12], Remark 1, and (8).

Remark 2. Because the Lipschitz constants of $r$ and $r_x$ do not depend on $\omega$, it follows from the proof of Corollary 5.6.9 in [12] that the Lipschitz constant of $\nabla_\eta \tilde{G}(\eta, \omega)$ does not depend on $\omega$.

The existence of the Gateaux derivative in Proposition 4 allows us to introduce the Gateaux derivatives of $J$ and $J^M$, which yield the optimality conditions for Problem $B$ and Problem $B^M$.

Proposition 5. Suppose that Assumptions 1-4 are satisfied. Then for any $\eta \in H^0$ and $\delta \eta \in H_{\infty, 2}$:

i) $J$ has a Gateaux differential $DJ(\eta; \delta \eta)$ at $\eta$ given by

$$DJ(\eta; \delta \eta) = \langle \nabla J^M(\eta), \delta \eta \rangle_{H_2},$$

with the gradient given by

$$\nabla J(\eta) = \mathbb{E}^P \left[ \nabla_\eta \tilde{G}(\eta, \omega) \right],$$

(12)

ii) The gradient $\nabla J$ is Lipschitz continuous on $H$.

iii) $J^M$ has a Gateaux differential $DJ^M(\eta; \delta \eta)$ at $\eta$ given by

$$DJ^M(\eta; \delta \eta) = \langle \nabla J^M(\eta), \delta \eta \rangle_{H_2},$$

with the gradient given by

$$\nabla J^M(\eta) = \frac{1}{M} \sum_{i=1}^{M} \nabla_\eta \tilde{G}(\eta, \omega_i),$$

(13)

iv) The gradient $\nabla J^M$ is Lipschitz continuous on $H$.

Proof. See Appendix.

Remark 3. It follows from Remark 2 and the proof of Proposition 5 that the Lipschitz constant of $\nabla J^M$ does not depend on $M$.

We now define non-positive optimality functions for Problem $B$ and Problem $B^M$.

$$\theta(\eta) = \min_{\eta' \in H} DJ(\eta; \eta' - \eta) + \frac{1}{2} \|\eta' - \eta\|_{H_2},$$

(14)

$$\theta^M(\eta) = \min_{\eta' \in H} DJ^M(\eta; \eta' - \eta) + \frac{1}{2} \|\eta' - \eta\|_{H_2}.$$  

(15)

Proposition 6. Suppose that Assumptions 1-4 hold.

i) If $\hat{\eta} \in H$ is a local minimizer of $B$, then $\theta(\hat{\eta}) = 0$.

ii) $\theta$ is a continuous optimality function for $B$.

iii) If $\hat{\eta}_M \in H$ is a local minimizer of $B^M$, then $\theta^M(\hat{\eta}_M) = 0$.

iv) $\theta^M$ is a continuous optimality function for $B^M$.

Proof. This proof follows directly from Proposition 5 and the arguments used in the proof of Theorem 4.2.3c in [12].

Here $J$ or $J^M$ replaces $J^0$, $H$ replaces $H_c$, and Proposition 5 replaces Corollary 5.6.9.

V. CONSISTENT APPROXIMATION OF PROBLEM $B$

When approximating an optimal control problem, it is desirable to select a scheme which approximates both the objective functional and the optimality condition well. One way to establish such a condition is to show that the approximation scheme is consistent in the sense of Polak [12], Section 3.3.

Definition 4. [12] Let $X$ be a complete separable metric space, let $J^M : X \rightarrow \mathbb{R}, J : X \rightarrow \mathbb{R}$ be lower semi-continuous functions, and let $\theta^M : X \rightarrow \mathbb{R}, \theta : X \rightarrow \mathbb{R}$ be non-positive upper semi-continuous functions. We say that the pair $\{J^M, \theta^M\}_{M\in\mathbb{N}}$ is a consistent approximation to the pair $\{J, \theta\}$ if:

i) $J_M \rightarrow \theta^M J$, 

ii) If $\{x_M\}_{M=1}^{\infty}$ is a sequence converging to $x$, then

$$\lim_{M \rightarrow \infty} \theta^M(x_M) \leq \theta(x).$$

We have already shown the almost sure epiconvergence of the approximate objective functional $J^M$ to the objective functional $J$ in Theorem 1. In order to address the convergence of the optimality functions we introduce the following lemma.

Lemma 3. Suppose that Assumptions 1-4 hold. Let $\eta \in H^0$, then $\nabla_\eta \tilde{G}(\eta, \cdot)(\cdot) : \Omega \times [0, 1] \rightarrow \mathbb{R}$ is measurable. Furthermore, there exists a compact set $F \subset \mathbb{R}^m$ such that $\nabla_u \tilde{G}(\eta, \omega)(t) \in F$ for all $\eta \in H^0, \omega \in \Omega, t \in [0, 1]$.

Proof. First note that

$$\tilde{h}_x = \left( \begin{array}{c} h_x \quad r_x \\ 0 \quad 0 \end{array} \right).$$

(16)

Then write $p^n(t, \omega) = [(p^n_1(t, \omega))^T p^n_2(t, \omega)]^T$ where $p^n_1$ represents the first $n$ components of $p_n$. From (11) and (16), $p^n_2$ is constant with respect to $t$ and equal to $\nabla G(z(1, \eta, \omega))$.

The adjoint equation can then be written as

$$p^n_1(s, \omega) = -h_x(x^n(s), u^n(s))p^n_2(s, \omega) + r_x(x^n(s), u^n(s), s, \omega)\nabla G(z(1, \eta, \omega)),$$

(17)

$$p^n_2(1, \omega) = 0.$$  

(18)

The result then follows from Lemma 1 and Remark 1.

To simplify notation, for a given $\eta^* \in H$, we introduce the following functions:

i) $\kappa^M_{\eta^*} : H \rightarrow \mathbb{R}, \eta \rightarrow \langle \nabla J^M(\eta^*), \eta \rangle_{H_2},$

ii) $\kappa_{\eta^*} : H \rightarrow \mathbb{R}, \eta \rightarrow \langle \nabla J(\eta^*), \eta \rangle_{H_2},$

iii) $\mu^M_{\eta^*} : H \rightarrow \mathbb{R}, \eta \rightarrow \langle \nabla J^M(\eta^*), \eta \rangle_{H_2},$

iv) $\mu_{\eta^*} : H \rightarrow \mathbb{R}, \eta \rightarrow \langle \nabla J(\eta^*), \eta \rangle_{H_2}.$

Lemma 4. Suppose that Assumptions 1-4 are satisfied. For a given $\eta^* \in H$ we have:

i) $\kappa^M_{\eta^*} \rightarrow \kappa_{\eta^*}$ uniformly almost surely.

ii) $\mu^M_{\eta^*} \rightarrow \mu_{\eta^*}$ almost surely.

Proof. Proof of i): For a given $t \in [0, 1]$, because the $\nabla_u \tilde{G}(\eta, \omega_i)(t)$, for $i = 1, \ldots, M$ are identically distributed, the strong law of large numbers, (12), and (13) imply that $\nabla J^M(\eta^*)(t) \rightarrow \nabla J(\eta^*)(t)$ almost surely. Therefore
\[ \nabla J^M(\eta^*) \to \nabla J(\eta^*) \] pointwise almost surely as \( M \to \infty \).
Recall that \( \| \eta \|_{H^2} \leq \rho_1 + \rho_2 \) for all \( \eta \in H \). Therefore for each \( \epsilon > 0 \), there exists \( K \in \mathbb{N} \) such that for each \( M > K \), we have \( \| \nabla J^M(\eta^*) - \nabla J(\eta^*) \|_{H^2} < \frac{\epsilon}{\rho_1 + \rho_2} \) by the dominated convergence theorem.
Then
\[
\begin{align*}
| \kappa^M_\eta(\eta) - \kappa^*_\eta(\eta) | &= \| (\nabla J^M(\eta^*) - \nabla J(\eta^*)) \|_{H^2} \\
&\leq \| \nabla J^M(\eta^*) - \nabla J(\eta^*) \|_{H^2} \| \eta \|_{H^2} \\
&< \epsilon \left( \rho_1 + \rho_2 \right) \\
&= \epsilon.
\end{align*}
\]

Proof of ii): First note that by Lemma 3, \( \langle \nabla \tilde{G}(\eta, \omega), \eta^* \rangle_{H^2} \) is continuous in \( \eta \) and measurable in \( \omega \) and therefore is a random lower semi-continuous function by Lemma 4.51 of [22]. Because \( \mu^*_\eta(\cdot) = \mathbb{P} \langle \nabla \tilde{G}(\cdot, \omega), \eta^* \rangle_{H^2} \) by the proof of Proposition 5, \( \mu^*_\eta \) is the expectation of a bounded random lower semi-continuous function. The result then follows from (12), (13) and Proposition 2.

Lemma 4 allows us to state the main result of this work, which is the almost sure consistent approximation of Problem B by Problem \( B^M \).

**Theorem 2.** Suppose that Assumptions 1-4 hold. Then the sequence \( \{ J^M, \theta^M \}_{M \in \mathbb{N}} \) is almost surely a consistent approximation to the pair \( (J, \theta) \).

**Proof.** The epicovergence of \( J^M \) to \( J \) is established in Theorem 1. It remains to show that \( \lim \sup_{M \to \infty} \theta^M(\eta^M) \leq \theta(\eta) \) whenever \( \eta^M \to \eta \).

Suppose that \( \eta^M \in H \) and \( \eta^M \to \eta \). First we write
\[
\begin{align*}
\theta^M(\eta^M) &= \min_{\eta^M \in H} \left\{ \langle \nabla J^M(\eta^M), \eta^M - \eta^M \rangle_{H^2} + \frac{1}{2} \| \eta^M - \eta^M \|_{H^2} \right\} \\
&= \min_{\eta^M \in H} \left\{ \langle \nabla J^M(\eta^M), \eta^M \rangle_{H^2} + \frac{1}{2} \| \eta^M - \eta^M \|_{H^2} \right\} \\
&- \langle \nabla J^M(\eta^M), \eta^M \rangle_{H^2} \\
&= \min_{\eta^M \in H} \left\{ \langle \nabla J^M(\eta^M) - \nabla J^M(\eta^M), \eta^M \rangle_{H^2} + \langle \nabla J^M(\eta^M), \eta^M \rangle_{H^2} \right\} \\
&+ \frac{1}{2} \| \eta^M - \eta^M \|_{H^2} \\
&= \min_{\eta^M \in H} \left\{ \langle \nabla J^M(\eta^M) - \nabla J^M(\eta^M), \eta^M \rangle_{H^2} + \kappa^M_\eta(\eta^M) \right\} \\
&+ \frac{1}{2} \| \eta^M - \eta^M \|_{H^2} \\
&= \min_{\eta^M \in H} \left\{ \langle \nabla J^M(\eta^M) - \nabla J^M(\eta^M), \eta^M \rangle_{H^2} \right\} - \mu^M_\eta(\eta^M).
\end{align*}
\]

We examine the behavior of \( \lim \sup_{M \to \infty} \theta^M(\eta^M) \) by looking at each expression in (19). First, by Remark 3 and the fact that \( \eta^M \to \eta \),
\[
\langle \nabla J^M(\eta^M) - \nabla J^M(\eta^M), \eta^M \rangle_{H^2} \to 0
\]
uniformly in \( \eta^M \), so that
\[
\lim_{M \to \infty} \min_{\eta^M \in H} \langle \nabla J^M(\eta^M) - \nabla J^M(\eta^M), \eta^M \rangle_{H^2} = 0.
\]

By the uniform convergence \( \kappa^M_\eta \to \kappa_\eta \),
\[
\lim_{M \to \infty} \min_{\eta^M \in H} \kappa^M_\eta(\eta^M) = \min_{\eta^M \in H} \kappa_\eta(\eta^M).
\]

Similarly
\[
\lim_{M \to \infty} \min_{\eta^M \in H} \frac{1}{2} \| \eta^M - \eta^M \|_{H^2} = \min_{\eta^M \in H} \frac{1}{2} \| \eta^M - \eta^M \|_{H^2}.
\]

Because \( \nabla J^M(\eta^M) \) is bounded,
\[
\lim_{M \to \infty} \langle \nabla J^M(\eta^M), \eta^M - \eta \rangle_{H^2} = 0.
\]

Finally by the epicovergence \( \mu^M_\eta \to \mu_\eta \), we have
\[
\lim_{M \to \infty} \inf \mu^M_\eta(\eta^M) \leq \mu_\eta(\eta).
\]

Now note that
\[
\theta(\eta) = \min_{\eta \in H} \left[ \kappa_\eta(\eta^M) + \frac{1}{2} \| \eta^M - \eta \|_{H^2} \right] - \mu_\eta(\eta).
\]

Then, by (19)-(25), we have
\[
\lim_{M \to \infty} \sup_{\eta^M \in H} \theta^M(\eta^M) \leq \theta(\eta) \text{ almost surely.}
\]

**VI. APPLICATION TO INTRUDER DETECTION IN A CHANNEL**

In this section we apply the computational framework developed in Sections II-V to an intruder detection problem inspired by [8]. A single searcher is attempting to detect a non-evading target moving down a channel. We assume the searcher has imperfect sensors and a turn-rate constraint. The objective is to find a trajectory for the searcher which maximizes the probability of detecting the target in the time horizon \([0, 75]\) (note that the change in time horizon can be handled by rescaling constants in the problem formulation).

The searcher is assumed to be a Dubin’s vehicle with known constant velocity \( v \). The dynamics of the searcher are given by
\[
\begin{align*}
x_1(t) &= v \cos x_3(t), \\
x_2(t) &= v \sin x_3(t), \\
x_3(t) &= u(t).
\end{align*}
\]
where \((x_1, x_2)\) represents the position of the searcher and \( x_3 \) is the heading angle. The control, \( u \), is the turning rate of the vehicle. In the simulation, we set \( v = 1 \), and \( K = 25 \). Let the channel be given by the rectangle \( R = [-20, 20] \times [-10, 10] \). To model the target, for each \( \omega = (\omega_1, \omega_2, \cdots, \omega_{10}) \in \mathbb{R}^{10} \), we define the trajectory \( y(t, \omega) \) in \( \mathbb{R}^2 \) by
\[
\begin{align*}
y_1(t) &= \omega_1 + \omega_2 t + \frac{1}{2} \omega_3 t^2 + \frac{1}{6} \omega_4 t^3 + \frac{1}{24} \omega_5 t^4, \\
y_2(t) &= \omega_6 + \omega_7 t + \frac{1}{2} \omega_8 t^2 + \frac{1}{6} \omega_9 t^3 + \frac{1}{24} \omega_{10} t^4.
\end{align*}
\]
Let \( A \subseteq \mathbb{R}^{10} \) be the rectangle defined by: \( \omega_1 \in [0, 20], \omega_5 \in [-10, 10], \omega_2, \omega_7 \in [-\frac{1}{6}, \frac{1}{6}], \omega_3, \omega_5 \in [-\frac{1}{24}, \frac{1}{24}], \omega_4, \omega_9 \in [-\frac{1}{2000}, \frac{1}{2000}], \omega_6, \omega_{10} \in [-\frac{1}{10}, \frac{1}{10}] \).

Let \( B \subseteq \mathbb{R}^{10} \) be the set of all parameter values for which the corresponding target trajectory is in the channel, and moving
left down the channel, for all times \( t \in [0, 75] \). We then consider parameter values from the set \( \Omega = A \cap B \).

To determine the effectiveness of the search we use a model in which \( \hat{r} \) is the detection rate and is independent of \( u(t) \), and \( G \) is given by \( G(z) = \exp(-z) \). The specific form of the detection rate function is given by the Poisson scan model:

\[
\hat{r}(x(t), y(t, \omega)) = \beta \Phi \left( \frac{F^K - D \| x(t) - y(t, \omega) \|^2 - b}{\sigma} \right),
\]

where \( \Phi \) is the standard normal cumulative distribution function, \( \beta \) is the scan opportunity rate, \( F^K \) is the so-called “figure of merit” (a sonar characteristic), and \( \sigma \) reflects the variability in the “signal excess”. In the simulation we use the values \( \beta = 1, F^K = 20, b = 20, D = 1, \) and \( \sigma = 10 \). For more information about the formulation of this model, see [7]-[11].

Problem C is then to minimize the objective functional

\[
E^P \left[ \exp \left( - \int_0^{75} \hat{r}(x(t), y(t, \omega)) dt \right) \right],
\tag{27}
\]

subject to the dynamics (26), where \( P \) is the uniform distribution on \( \Omega \).

**Remark 4.** Problem C differs from Problem B in that the pointwise control constraint is applied for all \( t \) in Problem C as opposed to almost all \( t \) in Problem B. However this is of no practical significance because an optimal solution to Problem C will be in the equivalence class of an optimal solution to Problem B.

Due to the irregular shape of the parameter space, this problem would be particularly challenging if we were to apply quadrature-based methods. The proposed computational framework of this paper is applied to this search problem by taking a random \( P \)-distributed draw of size \( M \) from the parameter space using an acceptance-rejection method, and approximating (27) by the sample average. The resulting standard optimal control problem is approximated using a direct method based on an LGL-pseudospectral direct discretization scheme with 54 nodes in the time domain. The NLP package SNOPT [23] is used to calculate the solution to NLP problem produced by this sequence of approximations. This yields a numerical approximation to the optimal trajectory for the searcher. A sample computed trajectory for \( M = 5000 \) is shown in Figure 1.

Note that the number of nodes \( M \) determines only the accuracy of the approximation of the objective functional and not the dimension of the resulting discretized NLP problem. Therefore it is possible to use a high sample size for the discretization of the parameter space. Increasing the sample size improves the accuracy of the calculated approximate optimal control but also requires that the corresponding target trajectory for each given node be stored and evaluated at each iteration of the NLP algorithm. Although a large number of nodes are used in the approximation of the given example problem, the computational cost of a similar approximation using a quadrature scheme is prohibitive. For this problem with 10 parameters, a quadrature scheme with 5 nodes in each dimension of the parameter space would require the storage and evaluation of \( 5^{10} \) possible target trajectories.

An extensive stochastic programming literature exists which aims to address the question of what sample size is appropriate to obtain a satisfactory approximation for problems with finite-dimensional decision spaces (see [24], and references therein). Such a discussion for the uncertain optimal control problem is beyond the scope of the current paper.

![Computed trajectory for a searcher attempting to detect an intruder in the channel for \( M = 5000 \). For reference, 10 possible target trajectories are shown. The targets move left down the channel and the searcher starts at \( (0, 0) \) at time \( t = 0 \). The arrows in the figure indicate the orientation of the trajectories.](image)

Fig. 1. Computed trajectory for a searcher attempting to detect an intruder in the channel for \( M = 5000 \). For reference, 10 possible target trajectories are shown. The targets move left down the channel and the searcher starts at \( (0, 0) \) at time \( t = 0 \). The arrows in the figure indicate the orientation of the trajectories.

**VII. Conclusions**

In this paper we provide a computational framework for the numerical solution of an uncertain optimal control problem by using a sample average approach to produce a family of approximating problems. We provided necessary conditions for optimality for both the original and approximating problems. We also showed that this approximation is consistent in the sense of Polak [12], Section 3.3. We applied this approach to the problem of optimizing the search trajectory for a searcher attempting to detect an intruder in a channel.

**APPENDIX**

**Proof of Lemma 1**

Let \( \eta = (\xi^0, u^0) \in \mathbb{H}^0 \). Let \( \chi^0_n : [0, 1] \times \Omega \to \mathbb{R}^l \) be such that \( \chi^0_0(0, \omega) = \xi^0 \) for each \( \omega \in \Omega \). \( \chi^0(\cdot, \omega) \) is absolutely continuous, and \( \chi^0(t, \cdot) \) is measurable. Then we define a sequence of functions \( \{\chi_n^0\}_{n=0}^{\infty} \) satisfying

\[
\chi^0_{n+1}(t, \omega) = \xi^0 + \int_0^t \kappa(\chi^0_n(s, \omega), u^0(s), \omega) ds.
\]

We demonstrate the measurability of \( \chi^0_n \), for each \( n \in \mathbb{N} \) by induction. For a given \( n \in \mathbb{N} \), \( t \in [0, 1] \), consider the function

\[
\psi_n : [0, t] \times \Omega \to \mathbb{R}^l,
\]

\[
\psi_n(s, \omega) = \kappa(\chi^0_n(s, \omega), u(s), \omega).
\]
For each $n \in \mathbb{N}$, if $\chi_\eta^n$ is measurable, then $\psi_n$ is a Carathéodory function and therefore measurable by Lemma 4.51 of [22], and thus $\chi_{\eta+1}^n(t, \cdot) = \xi^n + \int_0^t \psi_n(s, \cdot) \, ds$ is measurable. $\chi_\eta$ is a Carathéodory and therefore measurable, which implies that $\chi_\eta^n$ is measurable for each $n \in \mathbb{N}$ by induction. By the proof of Picard’s Lemma (see Polak [12], Lemma 5.6.3), we have $\chi_\eta(\cdot, \omega) \rightarrow \chi_\eta^*(\cdot, \omega)$ pointwise for each $\omega \in \Omega$. $\chi_\eta^n$ is then a pointwise limit of measurable functions, and is therefore measurable. It follows from the proof of Lemma 5.6.7 of [12] that there exists a $L \in [1, \infty)$ such that for all $\eta', \eta'' \in \mathbb{H}_0, \omega \in \Omega$, and $t \in [0, 1]$,
\[
\left| \chi_{\eta'}(t, \omega) - \chi_{\eta''}(t, \omega) \right| \leq L \| \eta' - \eta'' \|_{H_2}.
\]

Because $\mathbb{H}_0$ is bounded, this implies that $\chi_\eta^n$ is uniformly bounded for each $\eta \in \mathbb{H}_0$.

**Proof of Proposition 5** We will prove i) and ii); iii) and iv) will follow by an identical argument with $\Omega$ replaced by $\{\omega_1, \ldots, \omega_M\}$ and $P$ replaced by the counting measure normalized to 1.

Proof of i): Let $\delta \eta \in H^\infty, \eta \in \mathbb{H}_0, \omega \in \Omega$. Because $\mathbb{H}_0$ is open in the $L_\infty$ topology there exists a $\lambda^* > 0$ such that $\eta + \lambda \delta \eta \in \mathbb{H}_0$ for all $\lambda \in [0, \lambda^*]$. From Assumption 3 and Proposition 3, $\tilde{G}(\cdot, \omega)$ is Lipschitz continuous in $\eta$ with Lipschitz constant $L_G L_2$ for each $\omega \in \Omega$. From this fact we have
\[
\left| \tilde{G}(\eta + \lambda \delta \eta, \omega) - \tilde{G}(\eta, \omega) \right| \leq \left( L_G L_2 \| \delta \eta \|_{H_2} \right) \lambda,
\]
therefore for each $\omega \in \Omega, \eta \in \mathbb{H}_0, \lambda \in [0, \lambda^*]$,
\[
\left| \tilde{G}(\eta + \lambda \delta \eta, \omega) - \tilde{G}(\eta, \omega) \right| \leq L_G L_2 \| \delta \eta \|_{H_2}.
\]

Then the Gateaux derivative of $J$ is given by:
\[
DJ(\eta; \delta \eta) = \lim_{\lambda \downarrow 0} \frac{E^P \left[ \tilde{G}(\eta + \lambda \delta \eta, \omega) \right] - E^P \left[ \tilde{G}(\eta, \omega) \right]}{\lambda} = \lim_{\lambda \downarrow 0} \frac{E^P \left[ \tilde{G}(\eta + \lambda \delta \eta, \omega) - \tilde{G}(\eta, \omega) \right]}{\lambda} = E^P \left[ \frac{\tilde{G}(\eta + \lambda \delta \eta, \omega) - \tilde{G}(\eta, \omega)}{\lambda} \right] = E^P \left[ D\tilde{G}(\eta, \delta \eta)|_\omega \right],
\]
where we have used the dominated convergence theorem. Let $\delta \eta = (\xi^{\delta \eta}, u^{\delta \eta})$. Note that
\[
E^P \left[ \int_0^1 \langle \nabla_u \tilde{G}(\eta(t), \omega(t)), u^{\delta \eta}(t) \rangle \, dt \right] \leq E^P \left[ \| \nabla \tilde{G}(\eta, \omega) \|_{H_2} \| \delta \eta \|_{H_2} \right]
\]
is bounded, so that we can write
\[
DJ(\eta; \delta \eta) = E^P \left[ \nabla_x \tilde{G}(\eta, \omega), \xi^{\delta \eta} \right] + E^P \left[ \int_0^1 \langle \nabla_u \tilde{G}(\eta(t), \omega(t)), u^{\delta \eta}(t) \rangle \, dt \right] = E^P \left[ \nabla_x \tilde{G}(\eta, \omega), \xi^{\delta \eta} \right] + \int_0^1 E^P \left[ \nabla_u \tilde{G}(\eta(t), \omega(t)), u^{\delta \eta}(t) \right] \, dt = E^P \left[ \nabla_x \tilde{G}(\eta, \omega), \xi^{\delta \eta} \right] + \int_0^1 E^P \left[ \nabla_u \tilde{G}(\eta(t), \omega(t)), u^{\delta \eta}(t) \right] \, dt = \langle E^P \left[ \nabla \tilde{G}(\eta, \omega), \delta \eta \right] \rangle_{H_2},
\]
where we have used Fubini’s theorem.

The proof of ii) follows directly from the Lipschitz continuity of $\nabla \tilde{G}(\eta, \omega)$.

**References**


