Optimal Control of Diffusion Processes: Infinite-Order Variational Analysis and Numerical Solution

Roman Chertovskih, Nikolay Pogodaev, Maxim Staritsyn, and A. Pedro Aguiar, *Senior Member, IEEE*

***Abstract*—We tackle a nonlinear optimal control problem for a stochastic differential equation in Euclidean space and its state-linear counterpart for the Fokker-Planck-Kolmogorov equation in the space of probabilities. Our approach is founded on a novel concept of local optimality, stronger than conventional Pontryagin's minimum and originally crafted for deterministic optimal ensemble control problems. A key practical outcome is a rapidly converging numerical algorithm, which proves its feasibility for problems involving Markovian and open-loop strategies.**

***Index Terms*—Stochastic differential equations, Fokker-Planck equation, optimal control, optimality conditions, numerical algorithms.**

I. Introduction

WE EXPLORE an optimal control problem

$$(SP) \quad \min\{\mathcal{I}[\boldsymbol{u}] : \boldsymbol{u} \in \boldsymbol{U}\},$$

where $\mathcal{I}[\boldsymbol{u}]$ is defined based on the solution $X = X[\boldsymbol{u}]$ of a nonlinear Îto stochastic differential equation (SDE)

$$X_t = X_0 + \int_0^t f_s[\boldsymbol{u}](X_s)\, ds + \int_0^t \sigma_s[\boldsymbol{u}](X_s)\, dW_s \qquad (1)$$

with controlled deterministic coefficients $f : I \times \mathbb{R}^n \times \boldsymbol{U} \to \mathbb{R}^n$ and $\sigma : I \times \mathbb{R}^n \times \boldsymbol{U} \to \mathbb{R}^{(k\times n)}$. Here, $W = (W^j)_{j=1}^k : I \times \Omega \to \mathbb{R}^k$, is a standard Wiener process, defined on a complete probability space (p.s.) $(\Omega, \mathcal{F}, \mathbb{P})$ with a natural filtration $t \mapsto \mathcal{F}_t^W$ of the sigma-algebra $\mathcal{F}$, and $X_0 : \Omega \to \mathbb{R}^n$, is a random variable, independent of $\mathcal{F}_T^W$. The state trajectories X are random processes $I \times \Omega \to \mathbb{R}^n$ modulated by control functions $\boldsymbol{u}$. We assume that X are progressively measurable with respect to (w.r.t.) the sigma-algebras $\mathcal{F}_t^{W,X_0}$ generated by random variables W_t and X_0.

Our objective functional reads

$$\mathcal{I}[\boldsymbol{u}] = \mathbb{E}\left[\ell(X_T[\boldsymbol{u}]) + \int_I R_s[\boldsymbol{u}](X_s[\boldsymbol{u}])\, ds\right],$$

where $\ell : \mathbb{R}^n \to \mathbb{R}$ and $R : I \times \mathbb{R}^n \times \boldsymbol{U} \to \mathbb{R}$ are given terminal and running costs, and $\mathbb{E}$ is the $\mathbb{P}$-expectation.

A class $\boldsymbol{U}$ of control inputs, which is pertinent to the problem (SP), consists of random processes $\boldsymbol{u} : I \times \Omega \to U$, defined on the same p.s., adapted to $\mathcal{F}^{W,X_0}$, and valued in a certain set $U \subseteq \mathbb{R}^m$. However, the practice of control engineering usually leans towards more "realizable" options, whose choice is shaped by the nature and quality of online information available to the decision-maker. We investigate two "extreme" scenarios:

1) The state X_t is fully observable throughout the entire interval I, and the guide has technical options to adjust the control strategy accordingly. In this case, $\boldsymbol{u}$ takes the form of so-called Markovian strategy $\boldsymbol{u}_t(\omega) = w_t(X_t(\omega))$ given by a map $w : I \times \mathbb{R}^n \to U$.

2) No observation/intervention is possible during the actual execution of the control process. The guide is expected to predefine an open-loop control strategy $\boldsymbol{u}_t(\omega) \equiv u(t)$, $u : I \to U$, relying on the knowledge of the model and insights gained from preliminary experiments.

The first option is canonical for the stochastic context [1]. The second choice, although somewhat unconventional, finds motivation in different applications where one searches for a "robust" or "broadcast" control signal that compensates for noise or uncertainty. Notably, the related control and optimization problems, termed (optimal) ensemble control, prove to be significantly more intricate compared to those involving Markovian strategies.

For problem (SP) with the specified types of inputs, we introduce a novel approach termed the "∞-order variational analysis". After establishing necessary notations, assumptions, and reviewing essential stochastic analysis in Section I-B, we rigorously present our approach in Section II. Our analysis

Manuscript received 4 March 2024; revised 13 May 2024; accepted 29 May 2024. Date of publication 6 June 2024; date of current version 27 June 2024. This work was supported by the Foundation for Science and Technology (FCT) Portugal, in the Framework of ARISE (DOI 10.54499/LA/P/0112/2020) and R&D Unit SYSTEC (Base UIDB/00147/2020 and Programmatic UIDP/00147/2020 Funds), and Project RELIABLE (DOI 10.54499/PTDC/EEI-AUT/3522/2020) all supported by National Funds through FCT/MCTES (PIDDAC). Recommended by Senior Editor T. Oomen. *(Corresponding author: Maxim Staritsyn.)*

Roman Chertovskih, Maxim Staritsyn, and A. Pedro Aguiar are with the Research Center for Systems and Technologies (SYSTEC), ARISE, Department of Electrical and Computer Engineering, Faculdade de Engenharia, Universidade do Porto, 4200-465 Porto, Portugal (e-mail: roman@fe.up.pt; staritsyn@fe.up.pt; pedro.aguiar@fe.up.pt).

Nikolay Pogodaev is with the Dipartimento di Matematica "Tullio Levi-Civita," University of Padova, 35121 Padua, Italy (e-mail: nickpogo@gmail.com).

Digital Object Identifier 10.1109/LCSYS.2024.3410632

yields "feedback" necessary optimality conditions, offering alternatives to the stochastic Pontryagin's principle, and the corresponding indirect methods for successive approximations derived independently for Markovian (Section II-A) and robust (Section II-B) control inputs. In Section II-C, we recover a new concept of local minimum corresponding to the computed ∞-order variation of the cost functional. We illustrate our approach with a concise mathematical model from neuroscience in Section III. Finally, Section IV spans comparative insights into existing methods regarding numerical performance, followed by concluding remarks.

A. Contribution and Novelty

In the realm of optimal control for diffusion processes, three mainstream frameworks basically stand out: Stochastic Pontryagin's Maximum Principle (SPMP) [2], [3], [4], [5], Distributed Pontryagin's Principle [6], [7], [8], [9], [10], [11], [12], [13], [14], [15] assuming a shift to an optimal control of a Fokker-Planck-Kolmogorov equation (FP-PMP), and Dynamic Programming (DP) [1], [5], [16], [17] (the bibliography is vast, we only mention a few).

All these approaches pose significant challenges in numerical contexts, addressed, e.g., by [8], [9], [10], [11], [14], [16], [18]: SPMP deals with a sophisticated system of coupled forward-backward SDEs (FBSDEs) [19], prompting the search for indirect algorithms. The latter two demand a numerical solution to (respectively, linear Fokker-Planck and nonlinear Hamilton-Jacobi) partial differential equations (PDEs), a persistent issue practically viable only for small dimensions.

In general, Pontryagin's principle (PMP) gives rise to standard indirect algorithms reminiscent of the conventional gradient descent, which incorporate internal "step" parameters subjected to line search through specific backtracking, see, e.g., [11], [20]. Typically, such algorithms take numerous iterations to achieve a satisfactory approximation of a local solution. DP yields a global solution but, in practice, it is applicable exclusively to deterministic initial states: when X_0 is uncertain, the corresponding Hamilton-Jacobi equation is formulated on the space of measures, even if $\sigma \equiv 0$ [21].

In this letter, we promote an alternative approach that falls somewhere between FP-PMP and DP. In line with the former, a numerical method derived from our approach involves solving a linear parabolic PDE or extracting statistics through Monte Carlo as pivotal steps in each iteration. Nevertheless, the method is devoid of free intrinsic parameters and shows noticeably fast convergence.

Recently, our approach showcased its efficacy in numerical optimization involving random ordinary differential equations (ODEs) [22], [23], [24]. The ongoing extension to the stochastic framework, further complicated by the "arrow of time" burden, remained a natural and intriguing challenge. To our knowledge, this letter is pioneering on this line.

B. Notations, Standing Assumptions, and Preliminaries

We use the following notations: $\mathbb{R}^n$ is the n-dimensional Euclidean space with a fixed norm $|\cdot|$, and the corresponding matrix norm is denoted by the same symbol; $\mathcal{L}^n$ stands for the Lebesgue measure on $\mathbb{R}^n$. $C(\mathcal{X};\mathcal{Y})$ denotes the space of continuous functions between normed spaces $\mathcal{X}$ and $\mathcal{Y}$, $C(\mathcal{X}) \doteq C(\mathcal{X};\mathbb{R})$, and $\|\cdot\|_\infty$ is the natural sup-norm on C. C^k is the space of k times continuously differentiable functions, and C_c^∞ the class of smooth and compactly supported functions between the corresponding sets. $C^{1,2}$ stands for the class of functions $(t,x) \mapsto \eta_t(x)$, which are C^1 w.r.t. t and C^2 w.r.t. x. $L_{\rm p}$, ${\rm p} \geq 1$, are Lebesgue quotient spaces endowed with the corresponding norms $\|\cdot\|_{L_{\rm p}}$. For $A \subseteq \mathbb{R}^n$ and ${\rm p} \geq 1$, the class $W_{\rm p}^{1,2}(I\times A)$ is introduced as a completion of $C_c^\infty(I\times A)$ w.r.t. the norm $\|\varphi\|_{W_{\rm p}^{1,2}} \doteq \|\varphi\|_{L_{\rm p}} + \|\partial_t\varphi\|_{L_{\rm p}} + \|\nabla_x\varphi\|_{L_{\rm p}} + \|\nabla_{xx}^2\varphi\|_{L_{\rm p}}$. Given a measure space $(\nu,\mathcal{X},\mathcal{B})$, we denote by $\mathbb{E}^\nu\varphi$ the expectation $\langle\nu,\varphi\rangle$ of a ν-integrable function φ w.r.t. ν, and abbreviate $\mathbb{E} \doteq \mathbb{E}^{\mathbb{P}}$. $\mathcal{P}(\mathcal{X})$ stands for the collection of probability measures on $\mathcal{X}$, and $\mathcal{P}_c(\mathcal{X})$ for the set of probability measures with a compact support in $\mathcal{X}$. For a Borel measurable function $F:\mathcal{X}\to\mathcal{Y}$ between metric spaces, the pushforward of a measure $\nu\in\mathcal{P}(\mathcal{X})$ through F, $F_\sharp:\mathcal{P}(\mathcal{X})\to\mathcal{P}(\mathcal{Y})$, is introduced via the action on functions $\varphi\in L^1_{F_\sharp\nu}(\mathcal{Y})$ as $\mathbb{E}^{F_\sharp\nu}\varphi \doteq \mathbb{E}^\nu[\varphi\circ F]$.

We designate a compact and convex set $U\subset\mathbb{R}^m$ to represent the feasible range of control actions. The classes $\boldsymbol{U}_M$ and $\boldsymbol{U}_r$ of admissible Markovian and open-loop (robust) control strategies are composed of Borel measurable functions $I\times\mathbb{R}^n\to U$ and $I\to U$, respectively. We endow these sets with the weak* topologies of the corresponding dual Banach spaces $L_\infty(I\times\mathbb{R}^n;\mathbb{R}^m)$ and $L_\infty(I;\mathbb{R}^m)$, provided by the duality $(L_1)^* = L_\infty$. Note that any $u\in\boldsymbol{U}_r$ can be identified with an element $w\in\boldsymbol{U}_M$ such that $w_t(x)\equiv u(t)$.

Given a function ${\rm g}: I\times\mathbb{R}^n\times U\to\mathcal{X}$ to some set $\mathcal{X}$, we define an operator $g: I\times\mathbb{R}^n\times\boldsymbol{U}_M\to\mathcal{X}$ via the relation $g_t[w](x)\doteq{\rm g}(t,x,w_t(x))$. This operator is instrumental in formulating the data (R,f,σ) for the problem (SP), using functions $({\rm R},{\rm f},\varsigma)$ that adhere to one of the following sets of assumptions (some of which could be significantly relaxed):

- (A_M) ${\rm f}: I\times\mathbb{R}^n\times U\to\mathbb{R}^n$ and ${\rm R}: I\times\mathbb{R}^n\times U\to\mathbb{R}$ are bounded and Borel measurable, and $\varsigma\equiv\sqrt{2\beta}E_n$, where E_n is the unit matrix $(n\times n)$ and $\beta>0$.
- (A_r) f is as above; $\varsigma: I\times\mathbb{R}^n\times U\to\mathbb{R}^{k\times n}$ and ${\rm R}: I\times\mathbb{R}^n\times U\to\mathbb{R}$ are continuous and Lipschitz in the second variable uniformly w.r.t. the other variables.

In addition, we assume that $\mathbb{E}[|X_0|^2]<\infty$, while ℓ is in $C^2(\mathbb{R}^n)$ and satisfies the quadratic growth condition: $|\ell(x)|\leq {\rm C}(1+|x|^2)$ $\forall(t,x)\in I\times\mathbb{R}^n$ for some ${\rm C}>0$.

Assumptions (A_r) ensure the existence of a strongly unique strong solution for any $\boldsymbol{u}=u\in\boldsymbol{U}_r$ [5, Th. 1.3.15]. This result has been extended to the case of Borel measurable drift by [25, Th. 1]. Notably, conditions (A_M) are sufficient to establish the discussed property for all $\boldsymbol{u}=w\in\boldsymbol{U}_M$.

We complete this section by presenting some necessary preliminary facts from stochastic analysis and stochastic control: Let $t\in I$, $\boldsymbol{u}\in\boldsymbol{U}_M$, and consider the second-order differential operator $L_t[\boldsymbol{u}]:\mathcal{D}(L)\to C(\mathbb{R}^n)$

$$L_t[\boldsymbol{u}]\ \varphi \doteq \nabla_x\varphi\cdot f_t[\boldsymbol{u}]\ + {\rm Tr}\Big(\nabla_{xx}^2\varphi\, D_t[\boldsymbol{u}]\Big). \tag{2}$$

Here, Tr denotes the trace of a matrix, and $D\doteq\frac{1}{2}\sigma^{\rm T}\sigma$. Given a solution $X=X[\boldsymbol{u}]$ to (1), classical Îto's lemma says that, for any $\eta\in C^{1,2}(I\times\mathbb{R}^n)$, the composition $Y=\eta\circ X$ is also an Îto process satisfying

$$Y_t = \eta_0(X_0) + \in t_0^t\{\partial_s + L_s[\boldsymbol{u}]\}\eta(X_s)\,ds + M_t, \tag{3}$$

where $M_t \doteq M_t[\boldsymbol{u}] = \int_0^t \nabla_x \eta_s(X_s)^{\mathrm{T}} \sigma_s[\boldsymbol{u}](X_s)\, dW_s$ is a martingale (in particular, $M_0 = 0$ implies $\mathbb{E}M = 0$ on I). The result remains valid for $\eta \in W^{1,2}_{\mathrm{p}}(I \times A)$, where $A \subset \mathbb{R}^n$ is a bounded domain, and $\mathrm{p} > n + 2$ [26, Th. 3].

Another useful fact is a version [5, Remark 3.5.5] of the Feynman-Kac formula [27, Th. 8.2.1]: given $u \in \boldsymbol{U}_r$ and $q \in C(I \times \mathbb{R}^n \times \boldsymbol{U}_r)$, suppose that $p \in C^{1,2}([0,T) \times \mathbb{R}^n) \cap C(I \times \mathbb{R}^n)$ satisfies the quadratic growth condition, and solves the inhomogeneous backward Cauchy problem

$$\{\partial_s + L_s[\boldsymbol{u}]\}p = q_s[\boldsymbol{u}]; \quad p_T = \ell, \tag{4}$$

where the first relation holds for almost all (a.a.) $s \in I$. Then, p admits the probabilistic representation

$$p_t(x) = \mathbb{E}\left[\ell(X^x_{t,T}) - \int_t^T q_s[\boldsymbol{u}](X^x_{s,T})\right] \tag{5}$$

where $t \mapsto X^x_{s,t}$ is a solution of the SDE (1) on the interval $[s,T]$, $s \in [0,T)$, with a deterministic initial condition $X_{s,s}(x) = x \in \mathbb{R}^n$. We recommend [5, Sec. 1.3.3] for a brief overview of sufficient conditions for the existence of a $C^{1,2}$ solution to (4), which is applicable to the case $\boldsymbol{U} = \boldsymbol{U}_r$. Under the assumptions of [26, Th. 1] and with $q \equiv 0$, the function (5) is a unique solution to (4), which lies in $W^{1,2}_{\mathrm{p}}(I \times A)$ for each bounded $A \subset \mathbb{R}^n$. Furthermore, the mapping $(t,x) \mapsto \nabla_x p_t(x)$ is uniformly continuous (see also [25, Th. 3']). These last two facts are crucial for our analysis in the case $\boldsymbol{U} = \boldsymbol{U}_M$.

The final point to emphasize is the transformation of the nonlinear stochastic problem into a state-linear and deterministic one. This reformulation involves an abstract PDE, specifically, the Fokker-Planck-Kolmogorov (FPK) equation in the space of probability measures: Let $\boldsymbol{U}$ be one of the classes $\boldsymbol{U}_M$ or $\boldsymbol{U}_r$. The corresponding problem (SP) is equivalent [15] to a deterministic optimization

$$(DP) \quad \min\left\{\mathbb{E}^{\mu_T[\boldsymbol{u}]}\ell + \int_I \mathbb{E}^{\mu_s[\boldsymbol{u}]} R_s[\boldsymbol{u}]\ ds : \boldsymbol{u} \in \boldsymbol{U}\right\}$$

over measure-valued functions $\mu = \mu[\boldsymbol{u}] \in C(I; \mathcal{P}(\mathbb{R}^n))$, representing the dynamics $t \mapsto \mu_t$ of the distribution law

$$\mu_t \doteq (X_t)_\sharp \mathbb{P} \tag{6}$$

of X_t. It is a simple conjecture of Îto's lemma that μ satisfies the forward Cauchy system for a linear PDE

$$\partial_t \mu = L^*_t[\boldsymbol{u}]\ \mu, \quad \mu_0 = \vartheta \doteq (X_0)_\sharp \mathbb{P}, \tag{7}$$

where the operator L^* is a formal adjoint of L.

Equation (7) is understood in the sense of distributions: for any $\varphi \in C^\infty_c(\mathbb{R}^n)$ and $\tau, t \in I$, $\tau \le t$, it holds

$$\mathbb{E}^{(\mu_t - \mu_\tau)}\varphi = \int_\tau^t \mathbb{E}^{\mu_s} L_s[\boldsymbol{u}]\ \varphi\, ds,$$

and the initial condition means that $\lim_{t\to 0} \mu_t = \vartheta$ in the corresponding weak* topology. We refer to [28, Sec. 9] for sufficient conditions of the uniqueness of the measure-valued solution (6) to (7).

Note that, if μ_t, $t \in I$, are absolutely continuous w.r.t. $\mathcal{L}^n$, i.e., $\mu_t = \rho_t \mathcal{L}^n$ with $\rho_t \in L_1(\mathbb{R}^n; \mathbb{R})$, then $\rho : t \to \rho_t$ is a weak solution [29, Def. 9.2] of the parabolic PDE

$$\partial_t \rho = L^*_t[\boldsymbol{u}_t]\rho, \quad \rho_t\big|_{t=0} = \rho_0, \tag{8}$$

and this solution is (weakly) unique under relatively mild assumptions [30]. Some results on the existence of a solution to the optimization problems (SP) and (DP) in the class $\boldsymbol{U}_M$ can be found in [31, Th. 6.3] and [15, Th. 4]. Under the assumptions (A_r), the existence of a minimizer in the class $\boldsymbol{U}_r$ can be established by the classical continuity-compactness argument, using standard moment estimates [5, Th. 1.3.16] and the Banach-Alaoglu theorem.

II. ∞-Order Variational Analysis

To be short, we first focus on the Mayer-type functional and uncontrolled diffusion, assuming that $R \equiv 0$ and σ is independent of $\boldsymbol{u}$. A brief discussion of the general nonlinear Bolza problem is provided in Section II-D.

For the rest of this letter, we adhere to a given reference control $\bar{\boldsymbol{u}} \in \boldsymbol{U} \in \{\boldsymbol{U}_M, \boldsymbol{U}_r\}$. Our preliminary task is to derive a suitable representation for the increment $\Delta\mathcal{I} \doteq \mathcal{I}[\boldsymbol{u}] - \mathcal{I}[\bar{\boldsymbol{u}}]$ of the objective functional concerning another control $\boldsymbol{u} \in \boldsymbol{U}$. In the optimization framework, $\boldsymbol{u}$ takes the role of the unknown target strategy to be devised for orchestrating the $\mathcal{I}$-descent from $\bar{\boldsymbol{u}}$.

To streamline the notation, we use a bar to indicate the dependence on $\bar{\boldsymbol{u}}$ and omit mentioning $\boldsymbol{u}$, e.g., $\bar{X} \doteq X[\bar{\boldsymbol{u}}]$, and $X \doteq X[\boldsymbol{u}]$.

Our reasoning hinges on a simple yet non-standard class of needle-shaped variations: for any $s \in I$, we construct a new control $\boldsymbol{u} \diamond_s \bar{\boldsymbol{u}} \in \boldsymbol{U}$ as

$$t \mapsto (\boldsymbol{u} \diamond_s \bar{\boldsymbol{u}})_t \doteq \begin{cases} \boldsymbol{u}_t, & t \in [0,s) \\ \bar{\boldsymbol{u}}_t, & t \in [s,T], \end{cases} \tag{9}$$

and denote

$$\gamma_s \doteq X_T[\boldsymbol{u} \diamond_s \bar{\boldsymbol{u}}] = \bar{X}^{X_s}_{s,T}. \tag{10}$$

In view of the moment estimates [5, Th. 1.3.16] and the martingale property, it is evident that the map $s \mapsto \mathbb{E}\ell(\gamma_s)$ is Lipschitz continuous on I. Furthermore, by construction,

$$\Delta\mathcal{I} \doteq \mathcal{I}[\boldsymbol{u}] - \mathcal{I}[\bar{\boldsymbol{u}}] = \mathbb{E}\ell(\gamma_s)\Big|_{s=0}^{s=T} = \int_I \frac{d}{ds}\mathbb{E}\ell(\gamma_s)\, ds. \tag{11}$$

Now, let $\bar{p}$ be a solution to (4) with $\boldsymbol{u} = \bar{\boldsymbol{u}}$ and $q \equiv 0$. By the (generalized) Feynman-Kac formula,

$$\bar{p}_t(x) \doteq \mathbb{E}\ell\big(\bar{X}^x_{t,T}\big) \doteq \mathbb{E}\big[\ell(\gamma_s) \,|\, X_{0,t} = x\big].$$

Applying the law of total expectation, we can express

$$\mathbb{E}\bar{p}_s(X_s) \doteq \mathbb{E}\mathbb{E}\big[\ell(\gamma_s) \,|\, X_s\big] = \mathbb{E}\ell(\gamma_s). \tag{12}$$

Then, utilizing the (generalized) Îto's formula (3), the stochastic process $s \mapsto \bar{p}_s(X_s)$ is demonstrated to satisfy

$$\bar{p}_t(X_t) = \bar{p}_0(X_0) + \int_0^t \{\partial_s + L_s[\boldsymbol{u}]\}\bar{p} \circ X_s\, ds + M_t,$$

where M is defined as in Section I-B. By taking $\mathbb{E}$, applying (12), using Fubini's theorem, and differentiating w.r.t. s, we compute, for a.a. $s \in I$:

$$\begin{aligned} \frac{d}{ds}\mathbb{E}\ell(\gamma_s) &= \mathbb{E}\big[\{\partial_s + L_s[\boldsymbol{u}]\}\bar{p} \circ X_s\big] \\ &= \mathbb{E}\big[\{L_s[\boldsymbol{u}] - L_s[\bar{\boldsymbol{u}}]\}\bar{p}_s \circ X_s\big] \\ &= \mathbb{E}\big[\nabla_x \bar{p}_s(X_s) \cdot (f_s[\boldsymbol{u}] - f_s[\bar{\boldsymbol{u}}])(X_s)\big]. \end{aligned}$$

The difference under the sign of expectation shortly writes

$$\bar{H}_s[\boldsymbol{u}](X_s) - \bar{H}_s[\boldsymbol{u}](X_s),$$

where $\bar{H}_s[\boldsymbol{u}](x) \doteq H_s[\boldsymbol{u}](x, \nabla_x \bar{p}_s(x))$ is a contraction to $\psi = \nabla_x \bar{p}_s(x)$ of the Hamilton-Pontryagin functional

$$H_s[\boldsymbol{u}](x, \psi) \doteq \psi \cdot f_s[\boldsymbol{u}](x).$$

Substituting the resulting expression into (11), we arrive at the desired representation:

$$\Delta \mathcal{I} = \int_I \mathbb{E}\big[\bar{H}_s[\boldsymbol{u}](X_s) - \bar{H}_s[\bar{\boldsymbol{u}}](X_s)\big]\, ds, \tag{13}$$

which can be reformulated in terms of the laws μ_t of X_t as

$$\Delta \mathcal{I} = \int_I \mathbb{E}^{\mu_s}\big[\bar{H}_s[\boldsymbol{u}] \; - \bar{H}_s[\bar{\boldsymbol{u}}]\big]\, ds. \tag{14}$$

Two noteworthy features of the expressions (13) and (14) should be highlighted: i) they are applicable to any pair $(\bar{\boldsymbol{u}}, \boldsymbol{u})$ of inputs, without any proximity constraints, and ii) they are explicit and exact, i.e., devoid of any residual terms to be neglected. In parallel with 1st- and **2nd**-order variations arising from Taylor's expansion of $\mathcal{I}$, these formulas can be regarded as ∞-order variations of the cost functional at $\bar{\boldsymbol{u}}$, justifying the terminology "∞-order variational analysis".

Similar to the mentioned finite-order variations, our increment formulas provide a characterization of the optimality of the reference control via a particular pointwise optimization problem, detailed below. It should be stressed, however, that utilizing (13) and (14) for optimization purposes implies operating with a specific feedback mechanism, as X and μ correspond to the target control $\boldsymbol{u}$.

A. Markovian Controls

We start with the class $\boldsymbol{U}_M$ of Markovian strategies $\boldsymbol{u} \doteq w$, and adopt hypotheses (A_M). By $\bar{\mathrm{H}}_s$ we denote a contraction of the usual Hamiltonian $\mathrm{H}_s(x, \psi, \upsilon) = \psi \cdot \mathrm{f}_s(x, \upsilon)$ to $\psi = \nabla_x \bar{p}_s(x)$. In this notation, $\bar{H}_s[w](x) \doteq \bar{\mathrm{H}}_s(x, w_s(x))$. (In the pseudo-codes below, H^k stands for $\bar{\mathrm{H}}$ with $\bar{p} = p^k$.)

Given (13), the following assertion is straightforward.

Proposition 1: Assume that $\bar{w}$ is optimal for (SP) within the class $\boldsymbol{U}_{\mathrm{M}}$, and let $w \in \boldsymbol{U}_{\mathrm{M}}$ be a solution to the mathematical programming problem

$$\mathbb{E}\bar{H}_t[w](X_t[w]) = \mathbb{E}\bar{H}_t[\bar{w}](X_t[w]) \text{ for a.a. } t \in I. \tag{15}$$

Then, the following relation holds:

$$\bar{H}_t[w](x) = \min_{\upsilon \in U} \bar{\mathrm{H}}_t(x, \upsilon) \text{ for a.a. } t \in I, \;\; \forall\, x \in \mathbb{R}^n. \tag{16}$$

Proof: Using any control w from (16) in (13), we have: $\Delta\mathcal{I} \leq 0$. The optimality of $\bar{w}$ thus implies $\Delta\mathcal{I} = 0$, and the assertion follows from non-positivity of the integrand. ■

This result marks a conceptual difference with the PMP: Proposition 1 involves not only the tested control $\bar{w}$ but also an additional comparison control w, derived from $\bar{w}$. This feature complicates the use of Proposition 1 per se. However, construction (16) gives rise to Algorithm 1, which is free of intrinsic parameters, while generating a sequence $\{w^k\}$ with a monotonicity property: $\mathcal{I}[w^{k+1}] \leq \mathcal{I}[w^k] \doteq \mathcal{I}^k$.

Denote by $\mathcal{E}[\bar{w}, w]$ the negative right-hand side of (13). The map $(\bar{w}, w) \mapsto \mathcal{E}[\bar{w}, w]$ defines a functional $\mathcal{E} : \boldsymbol{U}_M \times \boldsymbol{U}_M \to \mathbb{R}$, which measures the violation of the condition (15). Observing that $\{\mathcal{I}^k\}$ is bounded, and therefore, $\mathcal{E}[w^{2k}, w^{2k+1}] \doteq \mathcal{I}^{2k} - \mathcal{I}^{2k+1} \to 0$, we can apply the Tychonoff and Banach-Alaoglu theorems to conclude that $\{(w^{2k}, w^{2k+1})\}$ converges, up to a subsequence, to a pair satisfying (15).

Algorithm 1: Optimal Markovian Control

Data: $\bar{w} \in \boldsymbol{U}_M$ (initial guess), $\varepsilon > 0$ (tolerance)
Result: $\{w^k\}_{k \geq 0} \subset \boldsymbol{U}_M$ such that $\mathcal{I}[w^{k+1}] < \mathcal{I}[w^k]$
$k \leftarrow 0$; $w^0 \leftarrow \bar{w}$;
repeat
  $p^k \leftarrow p[w^k]$; $w \in \arg\min \mathrm{H}^k_s$;
  $w^{k+1} \leftarrow w$; $k \leftarrow k + 1$;
until $\mathcal{I}[w^{k-1}] - \mathcal{I}[w^k] < \varepsilon$;

B. Open-Loop Controls

Now, we turn to the class $\boldsymbol{U}_r$ of robust control strategies. Though this choice implies significantly reduced controllability compared to $\boldsymbol{U}_M$, its practical implementation is much simpler. Moreover, as discussed in [11], [32], open-loop controls can serve as reasonable approximations of Markovian strategies. A numerical verification of this thesis is conducted by [11] using an affine-bilinear approximation.

A narrative akin to the previous paragraph can be extended to the law dynamics $t \mapsto \mu_t$. Assuming (A_r), the formula (14) becomes instrumental in assessing the optimality of $\bar{\boldsymbol{u}} = \bar{u} \in \boldsymbol{U}_{\mathrm{r}}$ proved similar to Proposition 1.

Proposition 2: Assume that the pair (u, μ), $u \in \boldsymbol{U}_r$, $\mu = \mu[u]$, satisfies the relation

$$\mathbb{E}^{\mu_t}\, \bar{\mathrm{H}}_t(u_t) = \min_{\upsilon \in U} \mathbb{E}^{\mu_t}\, \bar{\mathrm{H}}_t(\upsilon) \text{ for a.a. } t \in I. \tag{17}$$

If $\bar{u}$ is optimal, then (17) also holds for $u = \bar{u}$.

Note that (17) is, in fact, a type of operator equation on $\boldsymbol{U}_r$. Due to space limitations, we omit a rigorous proof of the existence of its solution, which is grounded in Kakutani-Fan's fixed-point argument. In pursuit of a numerical solution, we instead adopt a constructive approach to approximately solve (17) using the feedback principle: let $\nu \in \mathcal{P}(\mathbb{R}^n)$, and $\bar{v}_t[\nu]$ be a solution to the problem

$$\mathbb{E}^{\nu}\bar{\mathrm{H}}_t(\bar{v}_t[\nu]) = \min_{\upsilon \in U} \mathbb{E}^{\nu}\bar{\mathrm{H}}_t(\upsilon) \text{ for a.a. } t \in I. \tag{18}$$

The mapping $\bar{v} : (t, \nu) \mapsto \bar{v}_t[\nu]$ acts as a feedback control of (7), or a law-feedback control of (1). Substituting $\bar{v}$ into (7) results in a nonlocal FPK equation. Assuming that the latter possesses a unique solution $\hat{\mu} = \hat{\mu}[\bar{v}]$, and setting $u(t) \doteq \bar{v}_t[\hat{\mu}_t]$, we obtain a new control $u \in \boldsymbol{U}_r$ with the property (17) followed by the desired inequality $\Delta\mathcal{I} \leq 0$. Note that, in general, $\bar{v}[\cdot]$ is discontinuous. A solution to the corresponding backfed PDE can be designed by the Krasovskii-Subboting sampling method in the spirit of [22].

An iterative implementation of this idea is outlined in Algorithm 2. The convergence analysis can be elaborated similarly to [23, Appendix B].

Algorithm 2: Optimal Robust Control

Data: $\bar{u} \in \boldsymbol{U}_r$ (initial guess), $\varepsilon > 0$ (tolerance)
Result: $\{u^k\}_{k\geq 0} \subset \boldsymbol{U}_r$ such that $\mathcal{I}[u^{k+1}] < \mathcal{I}[u^k]$
$k \leftarrow 0$; $u^0 \leftarrow \bar{u}$;
repeat
  $p^k \leftarrow p[u^k]$; $v_s^k[\nu] \in \arg\min \mathbb{E}^\nu \mathrm{H}_s^k$;
  $\mu^{k+1} \leftarrow \hat{\mu}[v^k]$; $u^{k+1} \leftarrow v^k[\mu^{k+1}]$; $k \leftarrow k+1$;
until $\mathcal{I}[u^{k-1}] - \mathcal{I}[u^k] < \varepsilon$;

C. Local Optimality. Geometric Intuition. Relation to PMP

For fixed $\omega \in \Omega$, the parametrization $s \mapsto \gamma_s(\omega)$, defined in (10), represents, a.s., a curve in the set

$$\mathcal{R}_T(\omega) \doteq \{X_T[\boldsymbol{u}](\omega) : \boldsymbol{u} \in \boldsymbol{U}\},$$

of points, reachable by all ω-paths of the controlled SDE (1) ("sample-reachable" set). Certainly, since $\boldsymbol{u} \diamond_s \bar{\boldsymbol{u}} \in \boldsymbol{U}$, it is evident that, a.s., $\gamma_s(\omega) \in \mathcal{R}_T(\omega)$ for any $s \in I$, while $\gamma_0(\omega) = \bar{X}_T(\omega)$ and $\gamma_T(\omega) = X_T(\omega)$ by construction.

Building on this observation, we introduce an original concept of local minimum for problem (SP) that applies to both types $\boldsymbol{U}_r$ and $\boldsymbol{U}_M$ of control inputs. This concept does not rely on any norm on the spaces of state trajectories and/or control functions, and consequently, proves to be stronger than the usual "strong" or Pontryagin's minimum associated with the standard class of needle-shaped variations.

Definition 1: A process $\gamma : I \times \Omega \to \mathbb{R}^n$, is said to be a curve of expected monotone decrease from $x \in \mathbb{R}^n$ w.r.t. ℓ if the following holds: i) γ is $\mathcal{F}_T^{W,X_0}$-measurable in ω for all $s \in I$, ii) γ is a.s. continuous in s, iii) $\gamma_0 = x$ a.s., and iv) $s \mapsto \mathbb{E}\ell(\gamma_s)$ is decreasing on I.

We call a control $\bar{u} \in \boldsymbol{U}$ locally optimal for (SP) if there are no curves γ of the expected monotone decrease from $\bar{X}_T$ such that $\gamma_s \in \mathcal{R}_T$ a.s., for all $s \in I$.

Essentially, our approach boils down to suggesting a somewhat simplest class of a.s. Lipschitz curves on the sample-reachable set that perform a guaranteed "expected non-ascent" from $\bar{X}_T$. In the non-stochastic framework, there are toy examples [22] showing that necessary conditions, relying on this idea, can discard non-optimal PMP extrema. To construct such examples for the stochastic problem would be an interesting challenge.

D. Bolza Problem. Nonlinear Functionals

Generalizing the presented results to the Bolza problem is straightforward: by setting $\bar{p}$ in the form (5) with $\boldsymbol{u} = \bar{\boldsymbol{u}}$ and $q = R$, one replicates the logic of Section II with the corresponding re-definition of the maps $\bar{\mathrm{H}}$ and $\bar{H}$. In fact, it is possible to extend (14) to a wide class of nonlinear cost functionals, $\boldsymbol{\ell}, \boldsymbol{R}_s : \mathcal{P}_c(\mathbb{R}^n) \to \mathbb{R}$, possessing a differentiable intrinsic derivative (the derivative along vector fields) [33, Definition 2.2.2]. For the case of random ODEs, such an extension is performed in [24].

In practical scenarios featuring affine dependence on $\boldsymbol{u}$, geometric bounds, $\boldsymbol{u}_t \in U$ a.e. $t \in I$, are often omitted (U is assumed to be a large-radius ball). Instead, the control is penalized by an "energetic" running cost functional. For $\boldsymbol{u} = w$, the penalty is defined either through an auxiliary function $\mathrm{R} \equiv \frac{\alpha}{2}|\upsilon|^2$ as in Section I-B, or directly as $\frac{\alpha}{2}\int_I \|w_t\|_{L_2}^2\, dt$, $\alpha > 0$. In the latter case, if the initial distribution ϑ is absolutely continuous, a comparison control w is defined via the unique minimizer $w_t[\rho](x)$ of $\upsilon \mapsto \frac{\alpha}{2}|\upsilon|^2 + \bar{\mathrm{H}}_t(x, \upsilon)\rho(x)$, which depends on the density ρ of the probability law.

III. Numeric Experiment

To exemplify our approach, we tackle a stochastic version of an optimal control problem [23, Sec. III.A] for the Theta model, a simple model capturing bursting behavior of excitable neurons. The model features a phase variable $x \in \mathbb{S}^1$ and an excitability parameter $\eta \in \mathbb{R}$, governed by the SDE (1), with $f[\boldsymbol{u}](x) = (1 - \cos x) + (1 + \cos x)(\eta + \boldsymbol{u})$, and $\sigma \equiv \sqrt{2\beta}$. To highlight novel aspects in the stochastic setting, we formulate the problem using Markovian strategies $\boldsymbol{u} = w$, representing external excitations. Our objective is to steer the phase of a neuron from random initial data, characterized by a joint probability distribution in (x, η), to a specified value $\check{x} \in \mathbb{S}^1$ at a given moment T, while optimizing resource utilization: $\ell(x) = 1 - \cos(x - \check{x})$, and the running cost functional is $\frac{\alpha}{2}\int_I \|w_t\|_{L_2}^2\, dt$. The problem is cast into the form (DP) and solved by Algorithm 1.

For our experimental setup, we choose $T = 6$, $\alpha = 1$, $\beta = 0.5$, $\check{x} = \pi$, and $u^0 \equiv 0$. The PDEs (7) and (4) are solved by the standard pseudospectral method, and time integration is performed using the 4^{th}-order Runge-Kutta scheme. Note that the backfed equation (7) turns out to be nonlocal.

The probability density function (PDF) of the initial distribution ϑ is depicted in Fig. 1 (upper panel). The minimization history spans 3 iterations: $\mathcal{I} \approx 11.46 \to 2.88 \to 2.4 \to 2.31$. Notably, rapid convergence towards a meaningful solution with a reasonable control cost is evident. Fig. 1 (bottom panel) illustrates the resulting terminal PDF.

IV. Comparison With Existing Approaches. Conclusion

Recently, the central place in the field of indirect algorithms for controlling differential systems is occupied by the sequential quadratic Hamiltonian method (SQH). Described comprehensively in [34], SQH employs a quadratic penalty for the deviation from the reference control and backtracking to ensure both convergence of the iterative process and robustness against small perturbations of initial data. Notably, SQH demonstrates swift performance in numerical experiments. However, akin to all PMP-based successive approximation algorithms, it suffers from multiple recalculation of the descent control to adjust the penalty parameter and satisfy the Armijo condition. Furthermore, convergence is only guaranteed to Pontryagin's stationary point. Finally, while quadratic penalization of control updates enhances regularity, it may also lead to localization around the reference control, potentially slowing convergence.

Nonlocal Algorithms 1 and 2, on the other hand, hinge on a stronger notion of local optimality (Def. 1), and their iteration cost mirrors that of SQH without the need for recalculating the penalty parameter (assuming analytical solutions to problems (16) and (18)). These aspects suggest theoretical advantages over PMP-inspired algorithms, which still require validation through extensive numerical experiments.

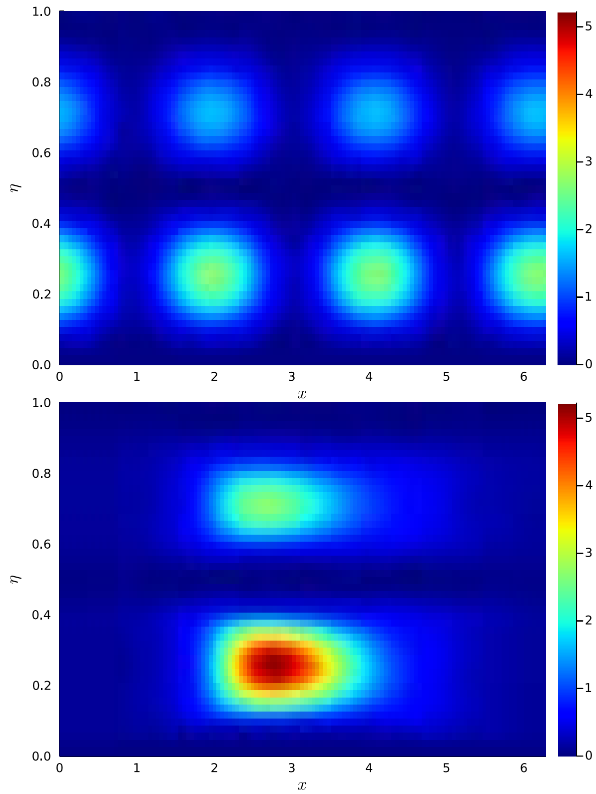

Fig. 1. Theta model: Snapshots of the "optimal" PDF at the initial (top) and final (bottom) time moments.

Another avenue for analysis is to compare the computational performance of our algorithms with (semi-) direct approaches, such as the Markov chain approximation method [35]. The comparison should be conducted at the experimental level, requiring precise and uniform in-code implementation of all utilized methods, along with relevant benchmarks. This challenging technical endeavor stands as a crucial component of our future work.

Acknowledgment

A part of the computations was carried out on the OBLIVION Supercomputer (Évora University) under FCT computational project 2022.15706.CPCA.

References

[1] N. Krylov and A. Aries, *Controlled Diffusion Processes* (Stochastic modelling and applied probability). Berlin, Germany: Springer, 2008.

[2] A. Bensoussan, "Stochastic maximum principle for distributed parameter systems," *J. Frankl. Inst.*, vol. 315, no. 5, pp. 387–406, 1983.

[3] S. Peng, "A general stochastic maximum principle for optimal control problems," *SIAM J. Control Optim.*, vol. 28, no. 4, pp. 966–979, 1990.

[4] Y. Hu, *Stochastic Maximum Principle*. London, U.K.: Springer Int. Publ., 2021, pp. 2186–2190.

[5] H. Pham, *Continuous-time Stochastic Control and Optimization with Financial Applications*. Berlin, Germany: Springer, 2009.

[6] E. Weinan, J. Han, and Q. Li, "A mean-field optimal control formulation of deep learning," *Res. Math. Sci.*, vol. 6, no. 1, p. 10, 2018.

[7] J.-F. Jabir, D. Siska, and L. Szpruch, "Mean-field neural odes via relaxed optimal control," 2019, *arXiv:1912.05475*.

[8] M. Annunziato and A. Borzì, "A Fokker–planck control framework for multidimensional," *J. Comput. Appl. Math.*, vol. 237, no. 1, pp. 487–507, 2013.

[9] S. Roy, M. Annunziato, and A. Borzì, "A Fokker–Planck feedback control-constrained approach for modeling crowd motion A Fokker–Planck feedback control-constrained approach," *J. Comput. Theor. Transp.*, vol. 45, no. 6, pp. 442–458, Jul. 2016.

[10] S. Roy, M. Annunziato, A. Borzì, and C. Klingenberg, "A Fokker–Planck approach to control collective motion," *Comput. Optim. Appl.*, vol. 69, no. 2, pp. 423–459, 2018.

[11] T. Breitenbach and A. Borzì, "The Pontryagin maximum principle for solving Fokker–Planck optimal control problems," *Comput. Optim. Appl.*, vol. 76, no. 2, pp. 499–533, 2020.

[12] A. Fleig and R. Guglielmi, "Bilinear optimal control of the Fokker-Planck equation," *IFAC-PapersOnLine*, vol. 49, no. 8, pp. 254–259, 2016.

[13] A. Fleig and R. Guglielmi, "Optimal control of the Fokker-Planck equation with space-dependent controls," *J. Optim. Theory Appl.*, vol. 174, pp. 408–427, Aug. 2017.

[14] A. Borzì and V. Schulz, *Computational Optimization of Systems Governed by Partial Differential Equations*, Philadelphia, PA, USA: Soc. Ind. Appl. Math., 2011.

[15] S. L. Aniţa, "Optimal control of stochastic differential equations via Fokker–Planck equations," *Appl. Math. Optim.*, vol. 84, pp. 1555–1583, Dec. 2021.

[16] M. Annunziato, A. Borzì, F. Nobile, and R. Tempone, "On the connection between the Hamilton-Jacobi-Bellman and the Fokker-Planck control frameworks," *Appl. Math.*, vol. 5, no. 16, pp. 2476–2484, Sep. 2014.

[17] J. Yong and X. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations* (Stochastic modelling and applied probability). New York, NY, USA: Springer, 1999.

[18] C. Sinigaglia, F. Braghin, and S. Berman, "Optimal control of velocity and nonlocal interactions in the mean-field Kuramoto model," in *Proc. Am. Control Conf. (ACC)*, 2022, pp. 290–295.

[19] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications* (Forward-backward stochastic differential equations and their applications 1702). Berlin, Germany: Springer, 1999.

[20] R. Chertovskih, N. Pogodaev, and M. Staritsyn, "Optimal control of nonlocal continuity equations: Numerical solution," *Appl. Math. Optim.*, vol. 88, no. 3, p. 86, 2023.

[21] A. Marigonda and M. Quincampoix, "Mayer control problem with probabilistic uncertainty on initial positions," *J. Differ. Equat.*, vol. 264, no. 5, pp. 3212–3252, 2018.

[22] M. Staritsyn, N. Pogodaev, R. Chertovskih, and F. L. Pereira, "Feedback maximum principle for ensemble control of local continuity equations: An application to supervised machine learning," *IEEE Control Syst. Lett.*, vol. 6, pp. 1046–1051, Jun. 2021.

[23] M. Staritsyn, N. Pogodaev, and F. L. Pereira, "Linear-quadratic problems of optimal control in the space of probabilities," *IEEE Control Syst. Lett.*, vol. 6, pp. 3271–3276, Jun. 2022.

[24] R. Chertovskih, N. Pogodaev, M. Staritsyn, and A. P. Aguiar, "Optimal control of distributed ensembles with application to Bloch equations," *IEEE Control Syst. Lett.*, vol. 7, pp. 2059–2064, Jun. 2023.

[25] A. J. Veretennikov, "On strong solutions and explicit formulas for solutions of stochastic integral equations," *Math. USSR-Sbornik*, vol. 39, no. 3, p. 387, 1981.

[26] A. K. Zvonkin, "A transformation of the phase space of a diffusion process that removes the drift," *Math. USSR-Sbornik*, vol. 22, no. 1, p. 129, 1974.

[27] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications* (Universitext). Berlin, Germany: Springer, 2003.

[28] V. Bogachev, N. Krylov, M. Röckner, and S. Shaposhnikov, *Fokker-Planck-Kolmogorov Equations* (Mathematical surveys and monographs). Providence, RI, USA: Am. Math. Soc., 2015.

[29] A. Bressan, *Lecture Notes on Functional Analysis: With Applications to Linear Partial Differential Equations* (Graduate studies in mathematics). Providence, RI, USA: Am. Math. Soc., 2013.

[30] M. Röckner and X. Zhang, "Weak uniqueness of Fokker–Planck equations with degenerate and bounded coefficients," *Comptes Rendus Mathematique*, vol. 348, no. 7, pp. 435–438, 2010.

[31] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control, Applications of Mathematices*, vol. 1. New York, NY, USA: Springer, 1975.

[32] R. Brockett, "Optimal control of the Liouville equation," in *Proc. Int. Conf. Complex Geom. Rel. Fields, Stud. Adv. Math., Amer. Math. Soc.*, 2007, pp. 23–35.

[33] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, *The Master Equation and the Convergence Problem in Mean Field Games* (Annals of Mathematics Studies), vol. 201. Princeton, NJ, USA: Princeton Univ. Press, 2019.

[34] A. Borzì, *The Sequential Quadratic Hamiltonian Method: Solving Optimal Control Problems* (Numerical analysis and scientific computing series). Boca Raton, FL, USA: CRC Press, 2023.

[35] H. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time* (Applications of mathematics, stochastic modelling and applied probability). New York, NY, USA: Springer, 1992.