# Efficient computation of optimal open-loop controls for stochastic systems

Bastien Berret, Frédéric Jean

# Efficient computation of optimal open-loop controls for stochastic systems

Bastien Berret[1,2,3,⋆] and Frédéric Jean[4]

1. CIAMS, Univ. Paris-Sud, Université Paris-Saclay, F-91405 Orsay, France
2. CIAMS, Université d'Orléans, F-45067 Orléans, France.
3. Institut Universitaire de France, Paris
4. Unité de Mathématiques Appliquées, ENSTA Paris, Institut Polytechnique de Paris, F-91120 Palaiseau, France

⋆ Corresponding author: bastien.berret@u-psud.fr

**Keywords**: motion planning, stochastic optimal control, open-loop deterministic control, muscle co-contraction, neural control of movement

**Abstract**

Optimal control is a prominent approach in robotics and movement neuroscience, among other fields of science. Methods for deriving optimal choices of action have been classically devised either in deterministic or stochastic settings. Here, we consider a setting in-between that retains the stochastic aspect of the controlled system but assumes deterministic open-loop control actions. The rationale stems from observations about the neural control of movement which highlighted that relatively stable behaviors can be achieved without feedback circuitry, via open-loop motor commands adequately tuning the mechanical impedance of the neuromusculoskeletal system. Yet, effective methods for deriving optimal open-loop controls for stochastic systems are lacking overall. This work presents a continuous-time approach for the efficient computation of optimal open-loop controls for a broad class of stochastic optimal control problems. We first consider simple synthetic examples showing that non-trivial departure from the optimal solutions of classical deterministic and stochastic approaches arises, and to stress the originality of the framework. We further exemplify its potential relevance to the planning of biological movement by showing that a well-known phenomenon in motor control, referred to as muscle co-contraction, occurs spontaneously. More generally, this stochastic optimal control framework may be suited to all fields where the design of optimal open-loop actions is relevant.

# 1 Introduction

In fields such as robotics or movement neuroscience, a preliminary motion planning stage is generally considered before movement execution (Todorov, 2004; LaValle, 2006). Optimality is a significant principle to rationalize this stage (Schoemaker, 1991; Berret *et al.*, 2019). Motion planning can thus be stated as an optimal control (OC) problem to figure out how to best drive a system to desired states, according to some optimality criterion or cost function. Originally, either deterministic or stochastic settings have been assumed to mathematically formulate OC problems (Fleming and Rishel, 1975). In deterministic optimal control (DOC), a powerful set of theoretical results and numerical tools exist to solve a broad variety of problems (Kirk, 1970; Trélat, 2008). Regarding motion planning, deterministic modeling is therefore a compelling approach but robustness/stability concerns may significantly hamper its relevance in presence of noise. Stochastic modeling is more appropriate to account for uncertainty in the control system, which naturally leads to stochastic optimal control (SOC) (Yong and Zhou, 1999). However, this framework poses several issues. Besides the linear-quadratic-Gaussian (LQG) scenario (Bryson and Ho, 1969; Athans, 1971), available methods are overall less efficient than in DOC and more sensitive to the curse of dimensionality. Indeed, going beyond dimension three or four may often become a real challenge (Fahim *et al.*, 2011; Falcone and Ferretti, 2014). Part of these difficulties arise because in the usual SOC settings the control is itself a stochastic variable through its dependence on the random observed data. Computation and application of the optimal controls typically require on-line feedback measurements and sophisticated state estimation procedures. Yet, considering stochastic control variables may not be required or even desired to formulate certain motion planning problems.

Indeed, regarding the planning of biological movement for instance, several influential studies have emphasized that our brain can purposely modulate the mechanical impedance of the neuromusculoskeletal system via feedforward co-contraction mechanisms (Hogan, 1984; Burdet *et al.*, 2001; Franklin *et al.*, 2003). Humans and animals are thus able to generate relatively stable motor behaviors even in the absence of on-line feedback circuitry (Polit and Bizzi, 1978, 1979; Ghez *et al.*, 1995). These findings suggest that on-line sensory feedback is not necessarily critical for stability, thereby suggesting that open-loop motion planning may be relevant piece of the neural control of movement. Yet, the sensorimotor system is affected by noise (Faisal *et al.*, 2008) which makes it stochastic by nature. These observations led us to consider a more restricted subclass of SOC problems where admissible controls are deterministic open-loop and the controlled system is stochastic. We refer to it as "stochastic optimal open-loop control" (SOOC). As such, we neglect on purpose the use or availability of on-line feedback from sensory data at this stage of motion planning. Feedback can always be exploited subsequently, for instance during movement execution to track a planned trajectory (Athans, 1971; Todorov and Li, 2005) or in a model predictive control scheme (Mayne *et al.*, 2000). To our knowledge, this hybrid setting where the control is deterministic and the dynamics is stochastic has not been studied in depth in optimal control theory. A difficulty is that the principle of optimality does not simply apply in this case and, therefore, methods for solving general SOOC problems are lacking overall. Here we show that under practical modeling choices it is possible to exploit the powerful and well-established deterministic machinery to efficiently compute solutions for a broad class of SOOC problems. More precisely, we consider stochastic processes described by Itô stochastic differential equations (SDE) and model costs/constraints as expectations. This is a classical approach to reformulate the problem from stochastic to deterministic. However, instead of considering the evolution of the probability density function (which evolves according to a partial differential equation and leads to infinite-dimensional problems, e.g. Palmer and Milutinovic 2011; Annunziato and Borzì 2013), a critical feature of our approach consists of focusing on propagation of the mean and covariance of the stochastic process, and on designing a DOC problem depending only on these variables via appropriate statistical linearization techniques. The rationale is that in many applications these two first moments are quantities of major interest. For example, mean and variance of endpoints are typically investigated to assess the quality of motor behaviors in movement neuroscience (e.g. van Beers *et al.*, 2004). The present considerations may be relevant to other fields such as robotics, especially for robots with variable impedance actuators that mimic the nonlinear spring-like characteristics of biological muscles (Migliore *et al.*, 2005; Vanderborght *et al.*, 2012). More generally, the SOOC framework may be significant to the control of stochastic systems with intermittent, long-latency or even absent on-line sensory feedback.

The study is organized as follows. In the Section 2, we introduce the mathematical settings of the SOOC framework under investigation. We show that a SOOC problem exactly reduces to a DOC problem with augmented mean-covariance state when the motion can be described by linear SDE. For nonlinear SDE, a Gaussian statistical linearization approach is proposed to get an approximately equivalent DOC problem that can be solved efficiently via state-of-the-art numerical techniques for deterministic optimal control. In Section 3, we compare the SOOC framework with DOC and SOC frameworks on synthetic examples that make derivations analytically tractable. In particular, we highlight that SOOC solutions may exhibit drastic qualitative changes of behavior depending on noise magnitude. These simple examples are also used to stress and discuss the differences between the SOOC and SOC approaches. In the Section 4, we consider an application of the SOOC framework to the field of the neural control of movement. This suggests that the SOOC framework may be particularly well-suited to describe motion planning for systems with tunable mechanical impedance like the neuromusculoskeletal system. In Section 5, we provide some conclusions and orientations for future works.

# 2 Stochastic optimal open-loop control framework

## 2.1 Mathematical formulation

The considered motions are modeled by stochastic dynamical systems of the form

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}(t), t)\,dt + G(\mathbf{x}_t, \mathbf{u}(t), t)\,d\mathbf{W}_t, \tag{1}$$

where $\mathbf{x}_t \in \mathbb{R}^n$ denotes the state variable and $\mathbf{u}(t) \in \mathbb{R}^m$ the control variable. The state of this Itô stochastic differential equation (SDE) is subject to deterministic infinitesimal increments driven by the vector valued drift function $\mathbf{f}$, and to random increments proportional to a multi-dimensional Wiener process $\mathbf{W}_t \in \mathbb{R}^k$, with stochastically independent components. The $(n \times k)$ dispersion matrix $G$ is full rank. The initial state is assumed to be a random variable $\mathbf{x}^0$ with known distribution.

The distinctive feature of our approach is that we consider open-loop control, i.e., $\mathbf{u}$ is a (deterministic) function of the time (for instance $\mathbf{u} \in L^2([0,T], \mathbb{R}^m)$ where $T$ is a fixed time). We will look for controls that minimize a cost expressed as an expectation,

$$C(\mathbf{u}) = \mathbb{E}\left[\int_0^T L(\mathbf{x}_t, \mathbf{u}(t), t)\,dt + \psi(\mathbf{x}_T)\right], \tag{2}$$

with possibly terminal constraints also expressed as expectations such as $\mathbb{E}[\phi(\mathbf{x}_T)] \in \mathcal{S}$, where $\mathcal{S}$ is a given target set. To summarize, we consider the following stochastic optimal open-loop control problem:

**(SOOC)** *minimize the cost* $C(\mathbf{u})$ *among all controls laws* $\mathbf{u}(t)$, $t \in [0,T]$, *such that the corresponding solution* $\mathbf{x}_t^{\mathbf{u}}$ *of* (1) *with* $\mathbf{x}_0^{\mathbf{u}} = \mathbf{x}^0$ *satisfies* $\mathbb{E}[\phi(\mathbf{x}_T^{\mathbf{u}})] \in \mathcal{S}$.

Such an optimization problem is in general very difficult to solve, and theoretical tools are lacking. One way to tackle the problem would be to characterize the stochastic process $\mathbf{x}_t$ by its density, whose evolution is modeled by a Fokker–Planck equation (see Annunziato and Borzì, 2013 for instance). One obtains in this way a formulation as a DOC problem on a partial differential equation (PDE). However, solving such a problem requires heavy computational efforts even in small dimension.

We present here an alternative approach based on a slight restriction of the framework motivated by applications in neuroscience and robotics. Indeed in these fields (and in many other engineering applications), it appears that the data of the problem have in general the following distinctive features: first, the cost is formulated as a quadratic function of both state and control (often for the sake of simplicity); second, the terminal constraints are expressed in terms of mean value and covariance in order to specify both the intended target and the requested precision/accuracy. Thus we make the following Assumptions in **(SOOC)**, where $\mathbf{m}_\mathbf{x}(t) = \mathbb{E}[\mathbf{x}_t]$ and $P_\mathbf{x}(t) = \mathbb{E}[(\mathbf{x}_t - \mathbf{m}_\mathbf{x}(t))(\mathbf{x}_t - \mathbf{m}_\mathbf{x}(t))^\top]$ denote respectively the mean and covariance of $\mathbf{x}_t$:

(i) The infinitesimal cost $L$ and the terminal cost $\psi$ are quadratic functions, that is $L(\mathbf{x},\mathbf{u},t) = \mathbf{u}^\top R(t)\mathbf{u} + \mathbf{x}^\top Q(t)\mathbf{x} + \mathbf{u}^\top S(t)\mathbf{x} + \mathbf{c_1(t)}^\top \mathbf{x} + \mathbf{c_2(t)}^\top \mathbf{u}$ and $\psi(\mathbf{x}) = \mathbf{x}^\top Q_f \mathbf{x} + \mathbf{c}_f^\top \mathbf{x}$, where $R(t)$, $Q(t)$, $S(t)$, $Q_f$ and $\mathbf{c_1}(t)$, $\mathbf{c_2}(t)$, $\mathbf{c}_f$ are respectively matrices and vectors of appropriate dimensions;

(ii) the terminal constraint writes as $\varphi(\mathbf{m}_\mathbf{x}(T), P_\mathbf{x}(T)) \in \mathcal{S}$;

(iii) the initial state $\mathbf{x}^0$ has a multi-normal distribution $\mathcal{N}(\mathbf{m}^0, P^0)$.

The crucial observation is that under Assumption (i), a simple computation shows that the cost in (2) takes the following form,

$$C(\mathbf{u}) = \int_0^T (L(\mathbf{m}_\mathbf{x}(t), \mathbf{u}(t), t) + \mathrm{tr}(Q(t)P_\mathbf{x}(t))\ dt + \psi(\mathbf{m}_\mathbf{x}(T)) + \mathrm{tr}(Q_f P_\mathbf{x}(T)). \tag{3}$$

Thus all the data of the problem (both the cost and the initial/terminal constraints) only depends on the mean and covariance of the process $\mathbf{x}_t$. We will see below that within this framework it is possible to reduce **(SOOC)** to a deterministic optimal control problem, up to an approximation of the dynamics.

*Remark* 1. The hypothesis that the infinitesimal cost is quadratic in the control plays actually no role here. In fact, we can consider infinitesimal costs in a much more general form, with explicit dependence on mean value and covariance, i.e. $L = L(\mathbf{m}_\mathbf{x}, P_\mathbf{x}, \mathbf{x}, \mathbf{u}, t)$, and require in Assumption (i) that $L$ is quadratic with respect to $\mathbf{x}$ only. The corresponding cost $C(\mathbf{u})$ is again of the form (3). Moreover, it is noteworthy that further assumptions about dynamics and cost would be necessary to ensure the existence of solutions, but we do not consider this problem here.

## 2.2 Reduction to deterministic optimal control (DOC) problems

### 2.2.1 Case of a linear SDE

Consider the case where the dynamic is a linear SDE, i.e.,

$$d\mathbf{y}_t = \big(A(\mathbf{u}(t),t)\mathbf{y}_t + \mathbf{b}(\mathbf{u}(t),t)\big)\,dt + H(\mathbf{u}(t),t)\,d\mathbf{W}_t, \tag{4}$$

where $A(\mathbf{u}(t), t)$ is an $(n \times n)$ matrix, $H(\mathbf{u}(t), t)$ is an $(n \times k)$ matrix, and $\mathbf{b}(\mathbf{u}(t), t) \in \mathbb{R}^n$. Since the control is deterministic, the process $\mathbf{y}_t$ is a Gaussian process, which is fully determined by its mean $\mathbf{m}(t) = \mathbf{m}_{\mathbf{y}}(t)$ and covariance $P(t) = P_{\mathbf{y}}(t)$.

The propagation of $\mathbf{m}$ and $P$ is given by the following ordinary differential equations (e.g. Stengel, 1986):

$$\begin{cases} \dot{\mathbf{m}}(t) = A(\mathbf{u}(t), t)\mathbf{m}(t) + \mathbf{b}(\mathbf{u}(t), t)), \\ \dot{P}(t) = A(\mathbf{u}(t), t)P(t) + P(t)A(\mathbf{u}(t), t)^\top + H(\mathbf{u}(t), t)H(\mathbf{u}(t), t)^\top, \end{cases} \tag{5}$$

and the initial values $\mathbf{m}(0), P(0)$ are the mean and covariance $\mathbf{m}^0, P^0$ of the initial state $\mathbf{y}^0$. Since moreover the cost $C$ in the form (3) depend only on the state variable $(\mathbf{m}, P)$, the problem **(SOOC)** in the case of a linear SDE is exactly equivalent to the following deterministic one:

**(DOC)** *minimize the cost* (3) *among all controls* $\mathbf{u}(t)$*,* $t \in [0, T]$*, such that the corresponding solution* $(\mathbf{m}^{\mathbf{u}}(t), P^{\mathbf{u}}(t))$ *of* (5) *with* $(\mathbf{m}^{\mathbf{u}}(0), P^{\mathbf{u}}(0)) = (\mathbf{m}^0, P^0)$ *satisfies* $\varphi(\mathbf{m}^{\mathbf{u}}(T), P^{\mathbf{u}}(T)) \in \mathcal{S}$*.*

Importantly, this DOC problem has a nonlinear dynamics but it is in finite dimension and can thus be solved and analyzed with classical tools (e.g. Bryson and Ho, 1969; Kirk, 1970).

*Remark* 2. From a control theory point of view, what is usually called a “linear system” is the case where the drift is linear with respect to both state and control, and the dispersion matrix $H$ is constant. The dynamics then writes as

$$d\mathbf{y}_t = \big(A\mathbf{y}_t + B\mathbf{u}(t)\big)\, dt + H\, dW_t. \tag{6}$$

In this case, the control does not enter at all in the dynamics of the covariance in (4). As a consequence the terms of the form $\mathrm{tr}(QP)$ do not depend on the control and can be removed from the cost. Hence **(DOC)** is reduced to a linear-quadratic optimal control problem on the mean value $\mathbf{m}(t)$, the covariance $P(t)$ remaining uncontrolled. In other terms, for linear control systems our open-loop approach only consists of replacing the SDE by its deterministic part $\dot{\mathbf{y}} = A\mathbf{y} + B\mathbf{u}$, the effect of the noise being completely ignored in the determination of the control. Therefore, our approach is mostly relevant for non-linear dynamics such as bilinear systems where product terms between the components of the control and the state are present in the dynamics (e.g. Hogan, 1984, for a movement neuroscience example). The necessity of working with non-linear dynamics to get non-trivial outcomes may partly explain why relatively little attention has been dedicated to SOOC.

#### 2.2.2 Gaussian statistical linearization

Consider now the general case of (1) and let us focus on the first two moments as motivated above. A classical computation using Itô’s formula shows that mean and covariance satisfy the following differential equations,

$$\begin{cases} \dot{\mathbf{m}}_{\mathbf{x}} = \mathbb{E}\left[\mathbf{f}(\mathbf{x}, \mathbf{u})\right], \\ \dot{P}_{\mathbf{x}} = \mathbb{E}\left[\mathbf{f}(\mathbf{x}, \mathbf{u})(\mathbf{x} - \mathbf{m}_{\mathbf{x}})^\top\right] + \mathbb{E}\left[(\mathbf{x} - \mathbf{m}_{\mathbf{x}})\mathbf{f}(\mathbf{x}, \mathbf{u})^\top\right] + \mathbb{E}\left[G(\mathbf{x}, \mathbf{u}, t)G(\mathbf{x}, \mathbf{u}, t)^\top\right]. \end{cases} \tag{7}$$

Let us introduce the following functions (we omit some of the dependence w.r.t. $t$ to simplify notations),

$$\begin{gathered} A_{\mathbf{u}}(t) = \mathbb{E}\left[\mathbf{f}(\mathbf{x}, \mathbf{u})(\mathbf{x} - \mathbf{m}_{\mathbf{x}}(t))^\top\right] P_{\mathbf{x}}(t)^{-1}, \quad \mathbf{b}_{\mathbf{u}}(t) = \mathbb{E}\left[\mathbf{f}(\mathbf{x}, \mathbf{u})\right], \\ H_{\mathbf{u}}(t),\ (n \times k) \text{ matrix s.t. } H_{\mathbf{u}}(t)H_{\mathbf{u}}(t)^\top = \mathbb{E}\left[G(\mathbf{x}, \mathbf{u}, t)G(\mathbf{x}, \mathbf{u}, t)^\top\right]. \end{gathered} \tag{8}$$

Note that the expectations $\mathbb{E} = \mathbb{E}^{\mathbf{x}_t}$ are taken with respect to the distribution of the solution $\mathbf{x}_t$ of (1). As a consequence, $A_{\mathbf{u}}(t)$, $\mathbf{b}_{\mathbf{u}}(t)$ and $H_{\mathbf{u}}(t)$ depends on all past values $\mathbf{u}(s)$, $s \in [0,t]$, of the control, and not only on $\mathbf{u}(t)$.

With these notations, mean and covariance are the solutions of the control system

$$\begin{cases} \dot{\mathbf{m}}_{\mathbf{x}}(t) = \mathbf{b}_{\mathbf{u}}(t), \\ \dot{P}_{\mathbf{x}}(t) = A_{\mathbf{u}}(t)P_{\mathbf{x}}(t) + P_{\mathbf{x}}(t)A_{\mathbf{u}}(t)^\top + H_{\mathbf{u}}(t)H_{\mathbf{u}}(t)^\top, \end{cases} \tag{9}$$

with $\mathbf{m}_{\mathbf{x}}(0) = \mathbf{m}^0$ and $P_{\mathbf{x}}(0) = P^0$. Since the cost $C$ and the initial/final conditions in **(SOOC)** depend only on the first and second moment of $\mathbf{x}_t$, we obtain that **(SOOC)** is formally equivalent to a DOC problem associated with the above control system in the variable $(\mathbf{m}_{\mathbf{x}}, P_{\mathbf{x}})$ and the cost (3).

Thus we are left with the problem of computing the parameters $A_{\mathbf{u}}(t)$, $\mathbf{b}_{\mathbf{u}}(t)$ and $H_{\mathbf{u}}(t)$, or at least an approximation of them. Indeed, with these coefficients, solving **(SOOC)** amounts to solving a finite-dimensional deterministic optimal control problem. This may be appealing because, for this type of problem, advanced numerical tools and theoretical results are available.

*Remark* 3. Actually we have shown that, as far as only the first two moments are concerned, we can replace the nonlinear stochastic dynamics (1) by the linear SDE

$$d\mathbf{y}_t = \big(A_{\mathbf{u}}(t)\mathbf{y}_t + \mathbf{d}_{\mathbf{u}}(t)\big)\,dt + H_{\mathbf{u}}(t)\,d\mathbf{W}_t,\ \mathbf{y}_0 = \mathbf{x}^0, \tag{10}$$

where $\mathbf{d}_{\mathbf{u}}(t) = \mathbf{b}_{\mathbf{u}}(t) - A_{\mathbf{u}}(t)\mathbf{m}_{\mathbf{x}}(t)$. This kind of method for probabilistic analysis of nonlinear stochastic dynamical systems is called *statistical (or stochastic) linearization*. We direct the reader to the recent survey (Elishakoff and Crandall, 2017) and to the monographs (Roberts and Spanos, 2003; Socha, 2008) for details.

It should be stressed that the averaging operations appearing in formula (8) are with respect to the distribution of $\mathbf{x}_t$. Since this distribution is not known exactly, we cannot calculate the parameters $A_{\mathbf{u}}$, $\mathbf{b}_{\mathbf{u}}$, and $H_{\mathbf{u}}$. Furthermore, even if the distribution is known, it is not always possible to calculate analytically these parameters especially when $\mathbf{f}$ is nonlinear. The difficulty in the method based on statistical linearization is to solve these two problems: first find a good approximation of the distribution of $\mathbf{x}_t$, and second compute efficiently the corresponding parameters $A_{\mathbf{u}}$, $\mathbf{b}_{\mathbf{u}}$, and $H_{\mathbf{u}}$.

As for the first problem different solutions exist, we present here the most popular one called *Gaussian statistical linearization*. This technique has been widely used for sixty years in the field of mechanics and has shown to be in most cases a very efficient numerical method. The Gaussian statistical linearization consists in approximating the distribution of $\mathbf{x}_t$ by the one of the solution $\mathbf{y}_t$ of (10). The latter being a Gaussian process, its distribution is the multi-normal distribution $\mathcal{N}(\mathbf{m}(t), P(t))$ characterized by the mean and covariance of $\mathbf{y}_t$. Replacing in (8) the averaging $\mathbb{E} = \mathbb{E}^{\mathbf{x}_t}$ by the averaging $\mathbb{E}^{\mathbf{y}_t}$ with respect to $\mathcal{N}(\mathbf{m}(t), P(t))$, we obtain new parameters $\widetilde{A} = \widetilde{A}(\mathbf{m}(t), P(t), \mathbf{u}(t), t)$, $\widetilde{\mathbf{b}} = \widetilde{\mathbf{b}}(\mathbf{m}(t), P(t), \mathbf{u}(t), t)$ and $\widetilde{H} = \widetilde{H}(\mathbf{m}(t), P(t), \mathbf{u}(t), t)$. Therefore $(\mathbf{m}(t), P(t))$ is determined as a solution of the following nonlinear control system,

$$\begin{cases} \dot{\mathbf{m}} = \widetilde{\mathbf{b}}(\mathbf{m}, P, \mathbf{u}, t), \\ \dot{P} = \widetilde{A}(\mathbf{m}, P, \mathbf{u}, t)P + P\widetilde{A}(\mathbf{m}, P, \mathbf{u}, t)^\top + \widetilde{H}(\mathbf{m}, P, \mathbf{u}, t)\widetilde{H}(\mathbf{m}, P, \mathbf{u}, t)^\top, \end{cases} \tag{11}$$

and the control $\mathbf{u}(t)$ is obtained by solving the **(DOC)** problem associated with the above dynamics.

Note that, using the properties of the multi-normal distribution $\mathcal{N}(\mathbf{m}(t), P(t))$, the formula for $\widetilde{A}$ simplifies as

$$\widetilde{A} = \mathbb{E}^{\mathbf{y}_t}\left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x},\mathbf{u})\right]. \tag{12}$$

*Remark* 4. Thus the Gaussian linearization consists in replacing the drift term $\mathbf{f}$ in the SDE (1) by

$$\mathbb{E}^{\mathbf{y}_t}\left[\mathbf{f}(\mathbf{x},\mathbf{u})\right] + \mathbb{E}^{\mathbf{y}_t}\left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x},\mathbf{u})\right](\mathbf{y}_t - \mathbf{m}(t)). \tag{13}$$

As a consequence, a first-order Taylor series expansion of $\mathbf{f}$ at $\mathbf{m}(t)$ appears as a particular case of Gaussian linearization with zero covariance.

The second problem that occurs now is to compute efficiently $\widetilde{A}$, $\widetilde{\mathbf{b}}$ and $\widetilde{H}$, that is to obtain an analytic approximate expression of the expectations

$$\mathbb{E}^{\mathbf{y}_t}\left[\mathbf{F}(\mathbf{x})\right] \quad \text{for } \mathbf{F}(\mathbf{x}) = \mathbf{f}(\mathbf{x},\mathbf{u}),\ \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x},\mathbf{u}),\ \text{or } G(\mathbf{x},\mathbf{u})G(\mathbf{x},\mathbf{u})^\top. \tag{14}$$

When $\mathbf{F}(\mathbf{x})$ is a polynomial function of $\mathbf{x}$, its expectation $\mathbb{E}^{\mathbf{y}_t}$ may be computed explicitly as a function of $\mathbf{m}$ and $P$ since the distribution of $\mathbf{y}_t$ is Gaussian (moments for Gaussian distributions can be computed by integral by parts). When $\mathbf{F}(\mathbf{x})$ is not polynomial, different techniques exist, such as sigma-point approximations and Taylor expansions (see Särkkä and Solin, 2019 for a survey). Let us describe for instance the second approach which is very common in filtering theory (see the description of the 2nd order compensated extended Kalman filter in Maybeck, 1982, or the discussion in Gustafsson and Hendeby, 2012).

The method consists in approximating $\mathbf{F}$ by its Taylor expansion at a given order $N$ around the mean value, i.e.

$$\mathbf{F}(\mathbf{x}) \sim \sum_{k=0}^{N} \frac{1}{k!}\frac{\partial^k \mathbf{F}}{\partial \mathbf{x}^k}(\mathbf{m})\cdot(\mathbf{x}-\mathbf{m},\ldots,\mathbf{x}-\mathbf{m}). \tag{15}$$

Thus we are reduced to calculating the expectation of a polynomial function. Such an approximation is justified when the covariance is sufficiently small (and is exact when the functions are polynomial with respect to the state).

In our case, taking for instance Taylor expansions at order $N = 1$, we obtain the following approximations (an alternative approximation which higher-order Taylor expansions is provided in Appendix):

$$\widetilde{\mathbf{b}} \sim \mathbf{f}(\mathbf{m},\mathbf{u}),\ \widetilde{A} \sim \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{m},\mathbf{u}),\ \text{and } \widetilde{H}\widetilde{H}^\top \sim G(\mathbf{m},\mathbf{u})G(\mathbf{m},\mathbf{u})^\top. \tag{16}$$

By putting everything together we are finally led to approximate the solutions of **(SOOC)** by the ones of the **(DOC)** problem associated with the cost (3) and the dynamics

$$\begin{cases} \dot{\mathbf{m}}(t) &= \mathbf{f}(\mathbf{m}(t),\mathbf{u}(t)), \\ \dot{P}(t) &= \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{m}(t),\mathbf{u}(t))P(t) + P(t)\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{m}(t),\mathbf{u}(t))^\top + \\ & \qquad G(\mathbf{m}(t),\mathbf{u}(t),t)G(\mathbf{m}(t),\mathbf{u}(t),t)^\top. \end{cases} \tag{17}$$

The method described above is based on two approximations whose validity must be discussed. The first one occurs from the replacement of the distribution of $\mathbf{x}_t$ by the one of $\mathbf{y}_t$. If these distributions are close to each other, then the method will predict accurately the mean and covariance of the system. It is worth noting that this condition does not even seem necessary since it seems that the parameters $A$ and $\mathbf{b}$ are in many cases not very sensitive to the assumed form of the distribution (see Beaman, 1984). The second approximation is to replace the functions $\mathbf{f}$, $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$, and $GG^\top$ by their Taylor expansion at order 1 (or more) for computing the expectations, which is meaningful when the covariance is sufficiently small. This is a reasonable assumption in our context since we minimize costs that indeed penalize the covariance. Moreover, our open-loop approach is designed to plan motions over relatively short times (relative to the time constants of the dynamics), starting from positions that are a priori fairly well known. Under these conditions the covariance will remain small.

The quality of these approximations was tested during simulations (see next sections). Our approach was as follows. Once the optimal control of the above **(DOC)** problem was calculated, we introduced the corresponding open-loop optimal control into the original SDE (1) and estimated by a Monte Carlo procedure the mean value and the covariance of the process $\mathbf{x}_t$. We could then verify that the control propagated the estimated mean and covariance of $\mathbf{x}_t$ in agreement with the approximations. We refer the reader to the human motor control application below for such a comparison.

Finally note that the **(DOC)** problem proposed above is convenient but it is not the only way to approximate the solutions of **(SOOC)** via statistical linearization. Indeed, for both steps of approximation, other choices are available (see for instance Crandall, 2006; Socha, 2008 on the choice of non-Gaussian distributions for the first point, and Ghusinga *et al.*, 2017; Särkkä and Solin, 2019 for computations of expectations).

# 3 Comparison with classical deterministic and stochastic frameworks

## 3.1 Comparison with DOC

To illustrate the fundamental difference between the DOC and SOC frameworks, a enlightening thought experiment is often used: the drunken spider problem (see Kappen, 2005*b*). In this example, a spider has to reach its home by either crossing a tiny bridge over a lake (shortest path) or by safely moving around the lake (longest path). In the deterministic case, the shortest path is the optimal solution. However, when there is noise (induced by alcohol in this thought example), the optimal solution is to take the longest path (as otherwise the spider may fall off the bridge into the lake due to the effects of noise). This example illustrates that significant qualitative changes of behavior between the DOC and SOC frameworks may be observed depending on noise magnitude. Here we highlight that the SOOC framework can exhibit similar qualitative changes of behavior.

To this end, we designed a toy example that captures the essence of the above thought experiment. Consider a dot moving in a plane, governed by the following first-order stochastic dynamics:

$$\begin{aligned} dx_t &= \varphi(y_t)u(t)dt, \\ dy_t &= v(t)dt + gd\omega_t. \end{aligned} \tag{18}$$

where the function $\varphi(y)$ defines a landscape in the plane as follows: $\varphi(y) = c$ where $c$ is a constant if $x \leq 0$ and $\varphi(y) = \dfrac{\sigma}{\sqrt{2\pi}} \exp(\dfrac{y^2}{2\sigma})$ otherwise. The parameter $\sigma$ can be chosen in order to make this (Gaussian) landscape more or less sharp (i.e. the bridge more or less narrow in the spider example). The noise $\omega_t$ is supposed to be a Wiener process and the parameter $g$ specifies noise magnitude (i.e. the spider more or less drunk). As in the rest of the study, the controls $u$ and $v$ are assumed to be deterministic.

We denote by $\mathbf{x}_t = (x_t, y_t)^\top$ the stochastic state of the system and fix the distribution of the initial state, $\mathbf{x_0} \sim \mathcal{N}(\mathbf{m}_0, P_0)$. For our purpose, we consider two possible terminal mean states, $\mathbf{m}_1$ and $\mathbf{m}_2$. The first one implements the close target location (i.e. shortest path solution) whereas the second one implements the far target location (i.e. longest path solution). The goal is thus to reach the state $\mathbf{m}_1$ or $\mathbf{m}_2$ in time $T$ (the final state covariance being left free) that yields the minimal energy cost

$$C(u, v) = \int_0^T \left(u(t)^2 + v(t)^2\right) dt. \tag{19}$$

This problem can be solved explicitly by Gaussian statistical linearization and by computing the expected values of (11) analytically. Then, numerical methods to solve the associated nonlinear DOC problem can be employed to get the optimal solution for each target separately. As there are two possible targets in this example, we just need to find the one that yields the smallest cost. The results of simulations are shown in Figure 1 with $\mathbf{m}_0 = \mathbf{0}$, $P_0 = \mathrm{diag}(10^{-3}, 10^{-3})$, $\mathbf{m}_1 = (1, 0)^\top$ and $\mathbf{m}_2 = (-2, 0)^\top$. It is noteworthy that, while we used two targets here, only a single target could have been used if we had modeled the problem on a cylinder instead of a plane (the $x$ coordinate would then become an angle in this case). However, this does not change the important conclusion which is about the qualitative change of optimal strategy that arises depending on the amount of noise $g$. In particular, when noise level is large enough, it becomes optimal to reach the far (leftward) target whereas the close (rightward) target was optimal for lower noise levels and in deterministic settings (optimal costs are emphasized in bold in Fig. 1).

## 3.2 Comparison with SOC

To stress the difference between the proposed SOOC framework and the classical SOC framework –standard LQG here–, let us introduce another toy example. This example is inspired by the considerations given in Introduction about the role of co-contraction on regulating the mechanical impedance of a limb (e.g. its stiffness). To keep calculations tractable we consider a noisy 1 degree-of-freedom system, the speed of which is directly controlled either by a stochastic –feedback– control (emulating "stiffness" control through an optimal feedback gain) or by a deterministic –open-loop– control (setting an optimal intrinsic "stiffness" directly).

For a stochastic control, the first-order dynamics we consider writes

$$dx_t = v_t dt + g dw_t, \tag{20}$$

where the control $v_t$ is a random process ($w_t$ is a Wiener process and $g$ is a scalar parameter to set the level of noise). The aim is to find the control strategy which minimizes the cost

$$C(v) = \frac{1}{2}\mathbb{E}\left[\int_0^T \left(v_t^2 + q x_t^2\right) dt + q_f x_T^2\right], \tag{21}$$

where $T$ is a fixed time. Therefore, the aim of the problem is to stabilize the system around its origin with minimal control effort and variance. It is known from LQG theory that the optimal solution is the feedback control $v_t = k(t)x_t$ where the gain $k$ (which is an analog of stiffness) is the solution of the following ordinary differential equation:

$$\dot{k} = q - k^2, \qquad k(T) = -q_f. \tag{22}$$

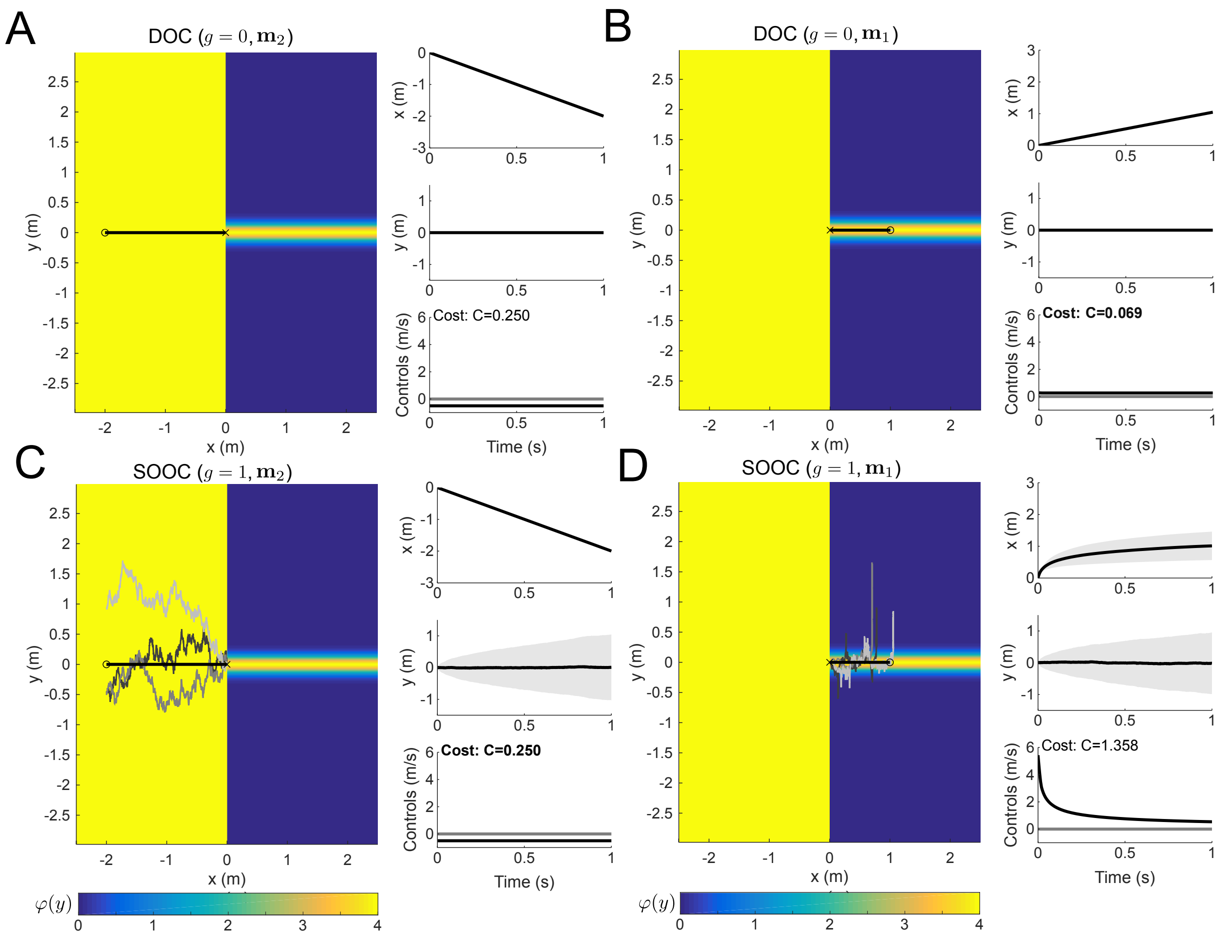


Figure 1: Comparison of DOC and SOOC. A. Deterministic case ($g = 0$) and final state $\mathbf{m}_2$. The colormap illustrates the landscape defined by the function $\varphi(y)$. The thick black path indicates the optimal mean trajectory. The 3 subplots on the right depict the time evolution of mean positions along the $x$ and $y$ axes, and the optimal open-loop controls (black for $u(t)$ and gray for $v(t)$). B. Same information for the deterministic case ($g = 0$) and final state $\mathbf{m}_1$. In the deterministic settings, as the optimal cost to move to $\mathbf{m}_1$ is smaller than the optimal cost to move to $\mathbf{m}_2$, the optimal strategy is thus to use the shortest path (i.e. to move rightward, the optimal cost being emphasized in bold). C. Stochastic case ($g = 1$) and final state $\mathbf{m}_2$. In addition to the colormap and the optimal mean trajectory, here we also display three instances of noisy trajectories to illustrate the stochastic nature of the paths. The 3 subplots on the right depict the same information as before but, here, mean trajectories were estimated from 1 000 samples (depicted in thick traces). Shaded areas depict standard deviations. D. Same information for the stochastic case ($g = 1$) and final state $\mathbf{m}_1$. Here, we see that the optimal strategy is to reach to the left target because the cost is smaller (emphasized in bold). This is because, in the right part of the plane ($x \geq 0$), the controller has to steer quickly the system to the target before it is shifted away from the midline, toward locations where it will be more difficult (i.e. energy consuming) to move. Parameters were as follows: $c = 4$, $\sigma = 0.01$ and $T = 1$.

Assume now that we only allow open-loop control but that we add a mechanical device such that the action on the velocity is proportional to the position (in other terms we control stiffness directly). For the open-loop control, the first-order dynamics of the system then writes

$$dx_t = u(t)x_t dt + g dw_t, \tag{23}$$

where the control $u(t)$ is a (deterministic) function of the time. Assume that the aim is to minimize the same cost as before, that is:

$$C(u) = \mathbb{E}\left[\int_0^T \left((u(t)x_t)^2 + qx_t^2\right) dt + q_f x_T^2\right]. \tag{24}$$

The optimal solution of this problem is obviously $u(t) = k(t)$ since it can be embedded into the first one by setting $v_t = u(t)x_t$. This can also be checked directly by solving the equivalent **(DOCP)** –via Pontryagin's maximum principle (Pontryagin *et al.*, 1964), see Section 6– which is associated with the dynamics

$$\begin{cases} \dot{m} &= um \\ \dot{P} &= 2uP + g^2 \end{cases}, \quad \text{where } m = \mathbb{E}[x],\ P = \mathbb{E}[(x-m)^2], \tag{25}$$

and with the cost

$$C(u) = \int_0^T \Big((m^2 + P)(u^2 + q)\Big) dt + q_f\big(m(T)^2 + P(T)\big). \tag{26}$$

The result is that applying the respective optimal controls in (20) and in (23) will produce exactly the same trajectories (although the underlying philosophy is totally different).

Now, let us compare the SOC and SOOC frameworks if we minimize a cost that penalizes the control quadratically for each system. That is, we now use the cost

$$\widetilde{C}(u) = \mathbb{E}\left[\int_0^T \left(u(t)^2 + qx_t^2\right) dt + q_f x_T^2\right] \tag{27}$$

in the open-loop case (i.e. the aim of the task is now to stabilize the system around its origin in open-loop with minimal effort –measured as stiffness– and variance).

It is not possible anymore to transform the open-loop problem into a LQG problem by the change of control variable $v = ux$. We can however compute again the solution using Pontryagin's Maximum Principle, and we obtain that the optimal control is solution of an ordinary differential equation of the form

$$\dot{u} = \mathcal{H} - u^2 + q(m^2 + P), \tag{28}$$

where the constant $\mathcal{H}$ depends on $q$, $q_f$, and $g$ (see Section 6). This equation can be compared with (22) characterizing the gain of the LQG solution. Whereas $k(t)$ does not depend on noise magnitude (parameter $g$), the optimal control $u(t)$ is sensitive to noise magnitude. Figure 2 illustrates the differences between the SOOC and SOC approaches. For clarity we compared stiffness-like quantities, i.e. $u(t)$ and $k(t)$. The important result is about the dependence of the optimal "stiffness" on the cost weight and the noise magnitude in SOOC but not in LQG. In SOOC, when noise magnitude increases, the magnitude of the optimal open-loop control increases. On the other hand, both solutions do depend on the weight of the variance cost: when $q$ and $q_f$ increase, the magnitude of the optimal control gain and of the optimal open-loop control also increase).

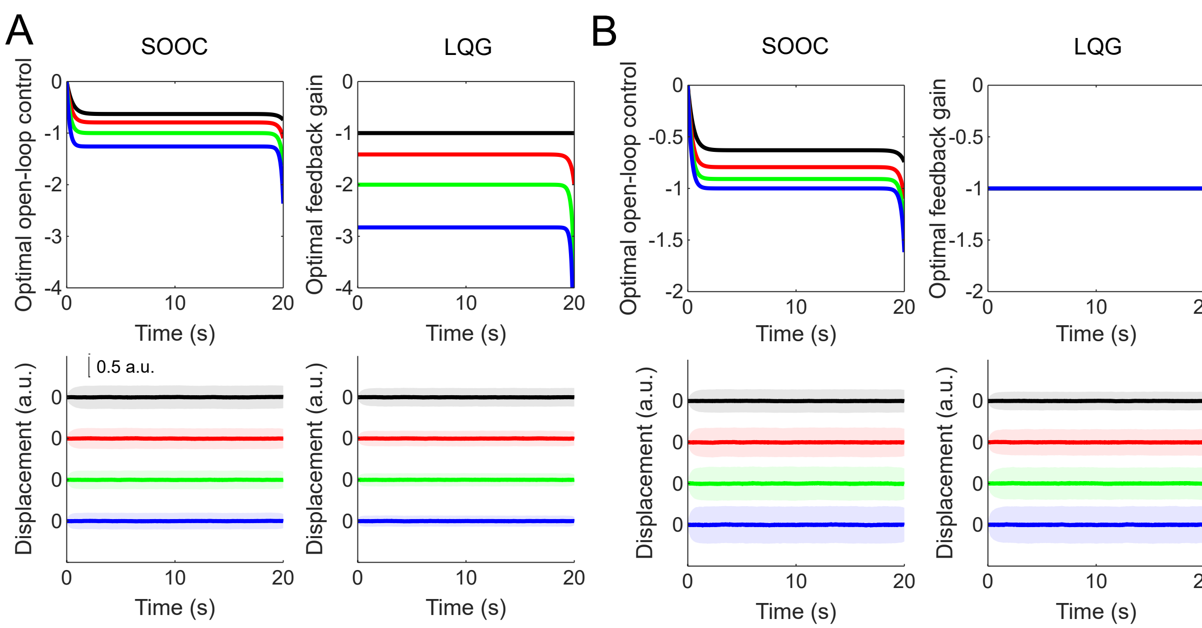


Figure 2: Illustration of the differences between SOOC and SOC (classical LQG) with quadratic costs. The first-order systems $dx_t = u(t)x_t dt + gdw_t$ (open-loop control) and $dx_t = v_t dt + gdw_t$ (feedback control) are considered. A. Influence of increasing the cost weight $q = q_f = 10, 20, 40, 80$ for the open-loop SOC case and $q = q_f = 1, 2, 4, 8$ for the LQG case (in black, red, green and blue respectively). The top graphs represent the controls (i.e. open-loop control or feedback gain for LQG). The bottom graphs represent the displacement around 0 (mean in thick solid line and standard deviation as a shaded area). B Influence of increasing noise covariance ($g^2 = 0.1$, 0.2, 0.3, and, 0.4, in black, red, green and blue respectively). Note that the LQG feedback gain does not depend on this parameter. Parameters were as follows: $P(0) = 0$, $m(0) = 0$, $T = 20$ seconds.

In summary, the SOOC approach can yield optimal solutions that differ in a non-trivial way from the solutions given by the LQG (or SOC) framework. In particular, we emphasized that SOOC solutions may significantly depend on noise magnitude. In the following, we stress important differences (both at conceptual and computational levels) between the SOOC framework and the LQG or SOC frameworks.

- In general settings, SOC produces a feedback control that requires state estimation from sensory measurements throughout motion execution. This loop inevitably induce certain delays that may reveal critical in certain situations, especially for unstable tasks or systems with long transmission latencies. In the human sensorimotor system for instance, feedback delays are relatively long compared to that of robots (muscle responses to a mechanical perturbation occur after ~20 ms for the stretch reflex; however, long-latency responses occur after >50 ms for proprioceptive and visual feedback loops passing through supraspinal centers, which are thought to be crucial for state estimation processes, e.g. Scott, 2012). These delays may cause issues for the control of accuracy in fast movements and for the realization of unstable tasks. In such situations, humans tend to regulate the impedance of their musculoskeletal system via feedforward muscle co-contraction mechanisms (e.g. Gribble *et al.*, 2003; Burdet *et al.*, 2001). The analog in robotics is to design mechanical systems with modifiable mechanical impedance, which may be implemented via variable stiffness actuators (Berret *et al.*, 2011*b*, 2012; Vanderborght *et al.*, 2012; Fiorio *et al.*, 2019). Note that the toy example illustrated in Figure 2 captures this characteristics via the open-loop control of stiffness (hence the bilinear control system in (23)).

- Aside from the LQG case, it can be very difficult to solve a SOC problem numerically in dimensions greater than 3 or 4 (a case frequently encountered in human or robot movement control). The assumptions of linear dynamics and quadratic cost are therefore often crucial for efficient numerical SOC methods. Conversely, the present SOOC approach applies to nonlinear systems and a wide class of cost functions. Efficient numerical and theoretical methods exist once an approximate equivalent DOC problem is defined and are only slightly limited by the dimension of the problem. Some frameworks more general than LQG have also been developed within SOC theory. Besides numerical scheme for the resolution of the Hamilton-Jacobi-Bellman equation (Fahim *et al.*, 2011; Falcone and Ferretti, 2014), iterative techniques such as iLQG (Li and Todorov, 2007) or differential dynamic programming (Theodorou *et al.*, 2010*b*) exist, based on successive first or second order approximations of the dynamics around a current trajectory. More recently, a class of non-linear SOC problems leading to a linear Bellman equation has been described (Kappen, 2005*a*, 2011; Todorov, 2009; Theodorou *et al.*, 2010*a*). Some limitations apply to the latter framework: re-computing the optimal control at each current state is necessary (e.g. using some Monte Carlo sampling techniques to evaluate path integrals), the control-dependent term of the cost function must be quadratic, the noise model is limited such as it must act in control space, the dynamics must be control-affine, and some relation between the noise magnitude and the control penalization in the cost function must hold. It is worth noting that in certain works, whereas the control is stochastic by nature, open-loop approximations were found numerically to avoid having to re-sample at each time of the actual trajectory (e.g. Berret *et al.*, 2013).

- The gain of the optimal control in the LQG solutions does not depend on the amplitude of additive noise, though this amplitude will affect the control indirectly via the feedback. Extensions with multiplicative noise are needed to get optimal feedback gains that depends on the noise magnitude (Todorov, 2005). SOOC, on the other hand, depends explicitly on the noise level even for a simple additive noise (because it deals with non-linear dynamics, e.g. bilinear system, see Fig. 2 and (23)).

- In SOC, and more specifically in the LQG theory, the constraints on the terminal point must be taken into account indirectly by adding a penalty term in the cost. This may be problematic since tuning the compromise between this penalty and the original cost may be tricky. Inadequate choices of cost weights can lead to irrelevant controls and trajectories. This issue does not appear in our open-loop framework as hard constraints can be set, that is, we can deal with explicit terminal constraints on the mean and the covariance of the stochastic process.

Despite their differences, the SOOC and SOC approaches may be complementary. A natural scheme in many applications could be to use the present open-loop approach at the stage of the planning process, and then to use the SOC framework to track a planned trajectory (for instance, one could employ a LQG method after linearization around the mean trajectory –and quadratization of the cost– like in the iterative LQG method of Li and Todorov, 2007) at the execution stage. Moreover, it is worth noting that the SOOC method at least requires knowledge of the initial state (or its distribution), which means that sensory feedback is at least required for that purpose (but it is not required in real-time as the control is open-loop). We next illustrate the SOOC approach in the context of human motor control. This application is interesting because it shows that the method can be applied to relatively complex problems. Moreover, it shows that it can account for the ability of our central nervous system to tune the intrinsic impedance of the neuromusculoskeletal system in a feedforward manner, in particular via co-contraction of antagonist muscles.

# 4 Application to the neural control of movement

We now apply the SOOC framework to a realistic movement planning problem to emphasize its potential relevance in this field. Here we focus on the control of reaching movements in humans with a model of the arm taken from Katayama and Kawato (1993). The arm is viewed as a two-link system moving in the horizontal plane and actuated by 6 muscles. Two pairs of muscles involve single-joint muscles acting around the shoulder and elbow joints respectively. The last two muscles are double-joint muscles. This model was shown to predict realistic limb stiffness. Here use this model to test our modeling framework on a realistic and relatively complex system. More precisely, the state of the arm is modeled as $\mathbf{x}^\top = (\mathbf{q}^\top, \dot{\mathbf{q}}^\top) \in \mathbb{R}^4$ where $\mathbf{q} = (q_1, q_2)^\top$ denotes the joint angle vector (1st component for shoulder and 2nd component for elbow) and $\dot{\mathbf{q}} = (\dot{q}_1, \dot{q}_2)^\top$ denote the corresponding joint velocity vector.

The skeletal dynamics of the arm follows a rigid body equation of the form:

$$\ddot{\mathbf{q}} = \mathcal{M}^{-1}(\mathbf{q})\big(\tau(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{u}) - \mathcal{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}}\big) \tag{29}$$

where $\mathcal{M}$ is the inertia matrix, $\mathcal{C}$ is the Coriolis/centripetal matrix, $\tau$ is the net joint torque vector produced by muscles and $\mathbf{u} \in \mathbb{R}^6$ is the muscle activation vector (i.e. the deterministic control variable here).

Precisely, the terms of the inertia and Coriolis/centripetal matrices are:

$$\begin{aligned} \mathcal{M}_{11}(\mathbf{q}) &= I_1 + I_2 + M_2 L_1^2 + 2M_2 L_1 L_{g2}\cos(q_2) \\ \mathcal{M}_{12}(\mathbf{q}) &= I_2 + M_2 L_1 L_{g2}\cos(q_2) \\ \mathcal{M}_{21}(\mathbf{q}) &= \mathcal{M}_{12}(\mathbf{q}) \\ \mathcal{M}_{22}(\mathbf{q}) &= I_2 \end{aligned} \tag{30}$$

and

$$
\begin{array}{rcl}
\mathcal{C}_{11}(\mathbf{q},\dot{\mathbf{q}}) & = & -2M_2L_1L_{g2}\sin(q_2)\dot{q}_2 \\
\mathcal{C}_{12}(\mathbf{q},\dot{\mathbf{q}}) & = & -M_2L_1L_{g2}\sin(q_2)\dot{q}_2 \\
\mathcal{C}_{21}(\mathbf{q},\dot{\mathbf{q}}) & = & M_2L_1L_{g2}\sin(q_2)\dot{q}_1 \\
\mathcal{C}_{22}(\mathbf{q},\dot{\mathbf{q}}) & = & 0
\end{array}
\tag{31}
$$

with $I_i$, $L_i$, $L_{gi}$ and $M_i$ denoting moments of inertia, lengths of segments, lengths to the center of mass and mass of the segments.

Regarding the net joint torque vector, we have $\tau(\mathbf{q},\dot{\mathbf{q}},\mathbf{u}) = -A^\top\mathcal{T}(\mathbf{l},\dot{\mathbf{l}},\mathbf{u})$ where $A$ is the moment arm matrix (constant here), $\mathcal{T}$ is the 6-D muscle tension vector and $\mathbf{l} = \mathbf{l}_m - A\mathbf{q}$ is the muscle length vector ($\mathbf{l}_m$ being the muscle length when the joint angle is zero; hence $\dot{\mathbf{l}} = -A\dot{\mathbf{q}}$). The matrix $A$ is defined as follows to define how the 6-D muscle tensions convert to 2-D joint torques:

$$
A = \begin{pmatrix} a_1 & a_2 & 0 & 0 & a_5 & a_6 \\ 0 & 0 & a_3 & a_4 & a_7 & a_8 \end{pmatrix}. \tag{32}
$$

In this model, the tension vector generated by muscles comes from the following function:

$$
\mathcal{T}(\mathbf{l},\dot{\mathbf{l}},\mathbf{u}) = K(\mathbf{u})\big(\mathbf{l}_r(\mathbf{u}) - \mathbf{l}\big) - B(\mathbf{u})\dot{\mathbf{l}}, \tag{33}
$$

where $K(\mathbf{u}) = \mathrm{diag}(k_0 + k\mathbf{u})$, $B(\mathbf{u}) = \mathrm{diag}(b_0 + b\mathbf{u})$ and $\mathbf{l}_r(\mathbf{u}) = \mathbf{l}_0 + \mathrm{diag}(r_1, ..., r_6)\mathbf{u}$. All the parameters of the model ($k_0$, $k$, $b_0$, $b$, $\mathbf{l}_m - \mathbf{l}_0$, $r_i$, $a_i$, $I_i$, $L_i$, $L_{gi}$, $M_i$) can be found in the Tables 1, 2, and 3 in Katayama and Kawato (1993).

Finally, by introducing noise ($\mathbf{W}_t$ a 2-D Wiener process with unit covariance), we obtain the following SDE modeling the noisy musculoskeletal dynamics of the human arm:

$$
d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}(t))dt + Gd\mathbf{W}_t \tag{34}
$$

with

$$
\mathbf{f}(\mathbf{x}_t, \mathbf{u}(t)) = \begin{pmatrix} \dot{\mathbf{q}}_t \\ \mathcal{M}^{-1}(\mathbf{q}_t)\big(\tau(\mathbf{q}_t, \dot{\mathbf{q}}_t, \mathbf{u}(t)) - \mathcal{C}(\mathbf{q}_t, \dot{\mathbf{q}}_t)\dot{\mathbf{q}}_t\big) \end{pmatrix} \tag{35}
$$

and

$$
G = \begin{pmatrix} \mathrm{diag}(0,0) \\ \mathrm{diag}(\sigma_1, \sigma_2) \end{pmatrix}. \tag{36}
$$

The parameters $\sigma_i$ are used to set the magnitude of additive noise. Here we used a simple additive noise model for simplicity but more complex noise models including state or control dependent noise could be implemented (for instance, if noise is assumed to act at torque level, we would need to replace the second component of $G$ by $\mathcal{M}^{-1}(\mathbf{q}_t)\mathrm{diag}(\sigma_1, \sigma_2)$, which would make the noise state-dependent, $G(\mathbf{x}_t)$). Importantly, it must be noted that the drift term of this SDE is relatively complex. For instance, $\mathbf{f}$ includes quadratic terms in the control variable (i.e. it is not a control-affine system) as well as many non-linear interactions between state and control elements.

In simulations, the goal was to move the arm from an initial position to a given target in fixed time $T$, while minimizing a cost defined as follows:

$$
C(\mathbf{u}) = \mathbb{E}\left[\int_0^T L(\mathbf{m}_\mathbf{x}, \mathbf{u})\,dt + (\mathbf{x}_f - \mathbf{m}_{\mathbf{x}_f})^\top \bar{Q}(\mathbf{x}_f - \mathbf{m}_{\mathbf{x}_f})\right] \tag{37}
$$

with $L(\mathbf{m_x}, \mathbf{u}) = \mathbf{u}^\top \mathbf{u} + \alpha(a_x^2 + a_y^2)$ where $a_x$ and $a_y$ are the Cartesian accelerations of the endpoint along the $x$ and $y$ axes respectively (i.e. functions of $\mathbf{m_x}, \mathbf{u}$, which can be easily computed from the forward kinematic function), and $\bar{Q} = q_{var}\text{diag}(1, 1, 10^{-3}, 10^{-3})$. This Lagrangian implements a compromise between minimizing effort and maximizing smoothness, which is in accordance with the literature (Flash and Hogan, 1985; Berret *et al.*, 2008, 2011*a*). The hand-acceleration cost favors straight hand paths and bell-shaped velocity profiles for the average trajectory as typically observed for such arm reaching movements. The control cost aims at minimizing the total amount of muscle activation used to achieve the task (i.e. effort). The final term of the cost corresponds to a penalization of the state covariance, the magnitude of which is tuned by the weight $q_{var}$. A variance cost has been proposed for the control of movement in Harris and Wolpert (1998). It must be noted that this cost is relatively complex due to the squared acceleration terms which introduce non-trivial dependencies between the mean state and the open-loop controls.

Results of simulations are given in Figure 3. The initial mean state was given [position (0,0.2) in Cartesian coordinates and zero velocity] and the initial covariance was zero. The final mean state was also given [position (0,0.4) in Cartesian coordinates and zero velocity] but the final covariance was left free. The simulated movement was a forward movement of 20 cm performed in 600 ms. Two different cases were tested: no penalization of the covariance term and strong penalization of the covariance term. It is worth noting that drastically different optimal controls and resulting trajectories are obtained depending on the consideration or not of a covariance cost ($q_{var}$ weight).

These simulations highlight that co-contraction may be an open-loop optimal strategy to reduce variability without the need for on-line sensory feedback and advanced state estimation procedures. In particular, the co-contraction observed at the end of the reaching movement agrees with experimental observations and the supposed role of co-contraction to improve movement accuracy (Gribble *et al.*, 2003). These simulations illustrate the basis principles and motivations underlying the SOOC framework. The role of co-contraction as a constitutive element of the motor plan is still elusive in neuroscience (Latash, 2018) and we believe that such a framework may reveal itself useful for investigating the nature of descending motor commands in the neural control of movement (without on-line sensory feedback).

# 5 Conclusions

A framework for the optimal open-loop control of stochastic systems has been presented. When focusing on propagation of mean and covariance of the stochastic process, resolving SOOC problems can be approached via powerful deterministic optimal control methods. As such, this formulation lies in-between the well-documented deterministic and stochastic optimal control theories, in the sense that noise effects on the dynamics are taken into account while only deterministic open-loop controllers are devised. The resulting framework is versatile and may prove useful in many fields as illustrated here for the neural control of movement. Indeed, the mechanical impedance of the neuromuscular system can be tuned via feedforward co-contraction mechanisms. This was emphasized in particular in unstable tasks where the detrimental effects of neural noise and delayed feedback are more apparent (Hogan, 1984; Burdet *et al.*, 2001; Franklin *et al.*, 2003). Interestingly, the proposed framework can handle non-linear dynamics, general cost functions, and various types of signal-dependent noise which is well suited to investigate complex systems such as the neuromusculoskeletal dynamics. After some modeling choices and statistical linearization, we showed

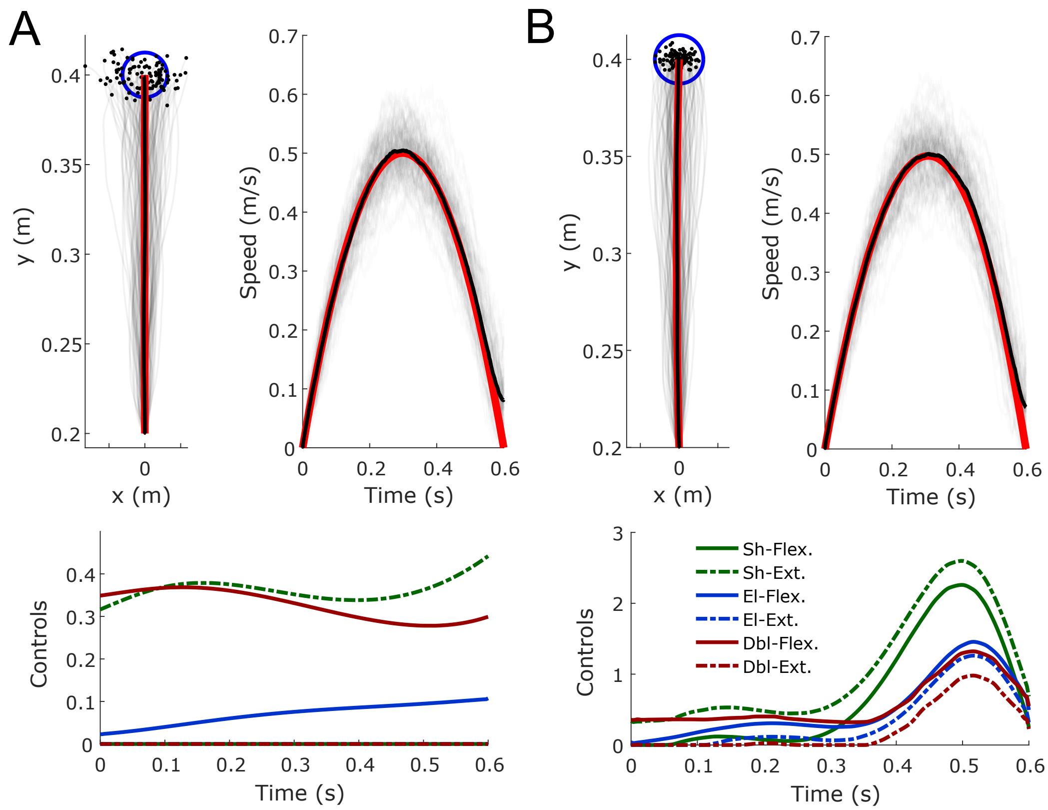


Figure 3: Planar arm reaching experiment. A. Hand trajectories in the horizontal plane when there is no variance cost ($q_{var} = 0$). *Left*: Hand paths. *Right*. Speed profiles. Transparent traces represent 100 sampled trajectories. Thick black traces are the corresponding (Monte Carlo) means. Red trajectories represent the theoretical trajectories predicted by the approximate equivalent deterministic optimal control problem. The blue circle represents the target to reach to. It can be seen that the target is often missed in this case. *Bottom*: Corresponding open-loop optimal controls. B. Hand trajectories in the horizontal plane when there is a variance cost ($q_{var} = 10^4$). Same panel organization as in A. It can be seen that most trajectories reach the target in this case although the control is open-loop. The price to pay is to use greater effort as seen by the larger optimal controls (note that the scale is different for readability). Note that co-contraction of antagonists is largely exploited to tune the mechanical impedance of the arm and resist disturbances intrinsically. Other parameters were as follows: $\alpha = 1$, $\sigma_1 = \sigma_2 = 0.5$.

that efficient numerical tools from DOC can readily be used to find approximate solutions of the original SOOC problem. This open-loop approach may be well complemented by the design of optimal feedback control laws after linearization around the planned (mean) trajectory. Interestingly, the latter was found to depend qualitatively on noise magnitude and cost function design. In conclusion, optimal open-loop control of stochastic systems may be a useful piece of a general motion planning scheme as it goes beyond standard deterministic formulations by taking into account robustness/stability issues and may be more convenient/efficient to use than general SOC for problems involving non-linear systems and non-quadratic costs. In certain cases where sensory delays are long, feedback intermittent or even unavailable on-line, this open-loop restriction may even be the correct way to model the problem at hand. In other cases where high-bandwidth sensory feedback is available, the framework may also be used within a model predictive control approach (Mayne *et al.*, 2000). Future work will aim at exploiting the present framework to investigate more deeply the role of muscle co-contraction in the neural control of movement or in robots with variable impedance actuators.

# 6 Appendix

## 6.1 Comparison with LQG: computation of the control

Consider the problem of minimizing the cost

$$C(u) = \mathbb{E}\left[\int_0^T \left(u(t)^2 + qx_t^2\right) dt + q_f x_T^2\right]. \tag{38}$$

among the trajectories of

$$dx_t = u(t)x_t dt + g dw_t,\ x_0 \sim \mathcal{N}(m^0, P^0) \tag{39}$$

As we have seen before, this is equivalent to minimizing the cost

$$C(u) = \int_0^T \left(u^2 + q(m^2 + P)\right) dt + q_f(m(T)^2 + P(T)), \tag{40}$$

among the trajectories of

$$\begin{cases} \dot{m} &= um \\ \dot{P} &= 2uP + g^2 \end{cases}, \qquad m(0) = m^0, \quad P(0) = P^0. \tag{41}$$

Define the Hamiltonian $H = \lambda_m um + \lambda_P(2uP + g^2) - (u^2 + q(m^2 + P))$. Pontryagin's Maximum Principle yields the following necessary conditions:

$$\begin{cases} \dot{\lambda}_m &= -u\lambda_m + 2qm, \\ \dot{\lambda}_P &= -2u\lambda_P + q, \end{cases} \qquad u = \frac{\lambda_m m}{2} + \lambda_P P, \tag{42}$$

with terminal constraints

$$\begin{cases} m(0) = m^0, \\ P(0) = P^0, \end{cases}, \begin{cases} \lambda_m(T) = -2q_f m(T), \\ \lambda_P(T) = -q_f. \end{cases} \tag{43}$$

We can compute $\dot{u}$ and, using that the Hamiltonian is constant along the time and satisfies $H = u^2 + \lambda_P g^2 - q(m^2 + P)$, we obtain

$$\dot{u} = H - u^2 + q(m^2 + P). \tag{44}$$

This equation allows to compute the optimal $u$ as soon as the constants $H$ and $u(0)$ are known. The latter only depend on the parameters $\lambda_m(0), \lambda_P(0)$ which have to be adjusted in such a way that the terminal conditions are satisfied.

It is interesting as well to see how the Maximum Principle allows to recover the equation of the LQG gain in the case where the cost is

$$\begin{aligned} C(u) &= \mathbb{E}\left[\int_0^T \left((x_t u(t))^2 + qx_t^2\right) dt + q_f x_T^2\right] \\ &= \int_0^T (m^2 + P)(u^2 + q) dt + q_f(m(T)^2 + P(T)). \end{aligned} \tag{45}$$

The necessary conditions become

$$\left\{\begin{array}{lcl}\dot{\lambda}_m & = & -u\lambda_m + 2m(q+u^2),\\ \dot{\lambda}_P & = & -2u\lambda_P + q + u^2,\end{array}\right. \qquad (m^2+P)u = \frac{\lambda_m m}{2} + \lambda_P P, \tag{46}$$

the terminal constraints remaining unchanged. A simple computation then shows that $\lambda_m - 2m\lambda_P \equiv 0$ (it is 0 at $t = T$ and solution of a linear equation). As a consequence, we obtain from the necessary condition above that $u = \lambda_P$ and

$$\dot{u} = \dot{\lambda}_P = -2u\lambda_P + q + u^2 = q - u^2. \tag{47}$$

## 6.2 Alternative approximation with higher order Taylor expansions

An alternative approximation could be as follows (e.g. Maybeck, 1982, Chap. 12). Taking the expansions at order $N = 2$ for $\mathbf{f}$ and at order $N = 1$ for $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ and $GG^\top$, so that:

$$\begin{array}{c}\widetilde{\mathbf{b}} \sim \mathbf{f}(\mathbf{m},\mathbf{u}) + \frac{1}{2}\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^2}(\mathbf{m},\mathbf{u}) \bullet P, \text{ where } \frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^2} \bullet P = \left(\mathrm{tr}(\frac{\partial^2 \mathbf{f}_1}{\partial \mathbf{x}^2}P), \ldots, \mathrm{tr}(\frac{\partial^2 \mathbf{f}_n}{\partial \mathbf{x}^2}P)\right)^\top \\ \widetilde{A} \sim \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{m},\mathbf{u}), \quad \text{ and } \quad \widetilde{H}\widetilde{H}^\top \sim G(\mathbf{m},\mathbf{u})G(\mathbf{m},\mathbf{u})^\top,\end{array} \tag{48}$$

we obtain the following dynamics for **(DOC)**,

$$\begin{cases}\dot{\mathbf{m}}(t) & = \mathbf{f}(\mathbf{m}(t),\mathbf{u}(t)) + \frac{1}{2}\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^2}(\mathbf{m}(t),\mathbf{u}(t)) \bullet P(t),\\ \dot{P}(t) & = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{m}(t),\mathbf{u}(t))P(t) + P(t)\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{m}(t),\mathbf{u}(t))^\top + \\ & \qquad G(\mathbf{m}(t),\mathbf{u}(t),t)G(\mathbf{m}(t),\mathbf{u}(t),t)^\top.\end{cases} \tag{49}$$

Note that an equivalent way to obtain the above approximation is to close the dynamics of $(\mathbf{m}_\mathbf{x}, P_\mathbf{x})$ by using a cumulant-neglect closure method at order 2 (see Socha, 2008 or Wojtkiewicz *et al.*, 1996).

# References


Annunziato, M. and A. Borzì(2013). 'A Fokker-Planck control framework for multidimensional stochastic processes'. *J Comput Appl Math* **237**(1), 487–507.

Athans, M. (1971). 'The role and use of the stochastic linear-quadratic-gaussian problem in control system design'. *IEEE Trans Autom Control* **16**(6), 529–552.

Beaman, J. J. (1984). 'Non-linear quadratic gaussian control'. *Int J Control* **39**(2), 343–361.

Berret, B., C. Darlot, F. Jean, T. Pozzo, C. Papaxanthis and J. P. Gauthier (2008). 'The inactivation principle: mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements.'. *PLoS Comput Biol* **4**(10), e1000194.

Berret, B., Enrico Chiovetto, Francesco Nori and Thierry Pozzo (2011*a*). 'Evidence for composite cost functions in arm movement planning: an inverse optimal control approach.'. *PLoS Comput Biol* **7**(10), e1002183.

Berret, B., G. Sandini and F. Nori (2012). Design principles for muscle-like variable impedance actuators with noise rejection property via co-contraction. In 'IEEE-RAS International Conference on Humanoid Robots (HUMANOIDS2012)'.

Berret, B., I. Yung and F. Nori (2013). Open-loop stochastic optimal control of a passive noise-rejection variable stiffness actuator: Application to unstable tasks. In 'Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems'. pp. 3029–3034.

Berret, B., Ioannis Delis, Jérémie Gaveau and Frédéric Jean (2019). *Optimality and Modularity in Human Movement: From Optimal Control to Muscle Synergies*. Springer International Publishing. Cham. pp. 105–133.

Berret, B., Serena Ivaldi, Francesco Nori and Giulio Sandini (2011*b*). Stochastic optimal control with variable impedance manipulators in presence of uncertainties and delayed feedback. In 'Proc. IEEE/RSJ Int Intelligent Robots and Systems (IROS) Conf'. pp. 4354–4359.

Bryson, A. E. and Y. C. Ho (1969). *Applied Optimal Control*. Blaisdell. New York.

Burdet, E., R. Osu, D. W. Franklin, T. E. Milner and M. Kawato (2001). 'The central nervous system stabilizes unstable dynamics by learning optimal impedance.'. *Nature* **414**(6862), 446–449.

Crandall, S. H. (2006). 'A half-century of stochastic equivalent linearization'. *Structural Control and Health Monitoring* **13**(1), 27–40.

Elishakoff, I. and Stephen H. Crandall (2017). 'Sixty years of stochastic linearization technique'. *Meccanica* **52**(1), 299–305.

Fahim, A., Nizar Touzi and Xavier Warin (2011). 'A probabilistic numerical method for fully nonlinear parabolic PDEs'. *Ann. Appl. Probab.* **21**(4), 1322–1364.

Faisal, A. A., L. P. J. Selen and D. M. Wolpert (2008). 'Noise in the nervous system.'. *Nat Rev Neurosci* **9**(4), 292–303.

Falcone, M. and Roberto Ferretti (2014). *Semi-Lagrangian approximation schemes for linear and Hamilton-Jacobi equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Fiorio, L., Francesco Romano, Alberto Parmiggiani, Bastien Berret, Giorgio Metta and Francesco Nori (2019). *Design and Control of a Passive Noise Rejecting Variable Stiffness Actuator*. Springer International Publishing. Cham. pp. 235–262.

Flash, T. and N. Hogan (1985). 'The coordination of arm movements: an experimentally confirmed mathematical model.'. *J Neurosci* **5**(7), 1688–1703.

Fleming, W. and R.W. Rishel (1975). *Deterministic and stochastic optimal control*. Applications of mathematics. Springer-Verlag.

Franklin, D. W., Rieko Osu, Etienne Burdet, Mitsuo Kawato and Theodore E Milner (2003). 'Adaptation to stable and unstable dynamics achieved by combined impedance control and inverse dynamics model.'. *J Neurophysiol* **90**, 3270–3282.

Ghez, C., J Gordon and M F Ghilardi (1995). 'Impairments of reaching movements in patients without proprioception. ii. effects of visual information on accuracy.'. *J Neurophysiol* **73**, 361–372.

Ghusinga, K. R., M. Soltani, A. Lamperski, S. V. Dhople and A. Singh (2017). Approximate moment dynamics for polynomial and trigonometric stochastic systems. In '2017 IEEE 56th Annual Conference on Decision and Control (CDC)'. pp. 1864–1869.

Gribble, P. L., L. I. Mullin, N. Cothros and A. Mattar (2003). 'Role of cocontraction in arm movement accuracy.'. *J Neurophysiol* **89**(5), 2396–2405.

Gustafsson, F. and Gustaf Hendeby (2012). 'Some relations between extended and unscented kalman filters'. *IEEE Trans Signal Process* **60**(2), 545–555.

Harris, C. M. and D. M. Wolpert (1998). 'Signal-dependent noise determines motor planning.'. *Nature* **394**(6695), 780–784.

Hogan, N. (1984). 'Adaptive control of mechanical impedance by coactivation of antagonist muscles'. *IEEE Trans Autom Control* **29**(8), 681–690.

Kappen, H. J. (2005*a*). 'A linear theory for control of non-linear stochastic systems'. *Phys Rev Lett* **95**, 200–201.

Kappen, H. J. (2005*b*). 'Path integrals and symmetry breaking for optimal control theory'. *J Stat Mech: Theory Exp* **2005**(11), P11011.

Kappen, H. J. (2011). Optimal control theory and the linear bellman equation. In D. Barber, A. T. Cemgil and S. Chiappa (Eds.). 'Bayesian Time Series Models'. Cambridge University Press. pp. 363–387. Cambridge Books Online.

Katayama, M. and M Kawato (1993). 'Virtual trajectory and stiffness ellipse during multijoint arm movement predicted by neural inverse models.'. *Biol Cybern* **69**, 353–362.

Kirk, D. E. (1970). *Optimal control theory: An Introduction.*. Prentice-Hall, New Jersey.

Latash, M. L. (2018). 'Muscle coactivation: definitions, mechanisms, and functions.'. *J Neurophysiol* **120**, 88–104.

LaVelle, S. M. (2006). *Planning algorithms*. Cambridge university press.

Li, W. and E. Todorov (2007). 'Iterative linearization methods for approximately optimal control and estimation of non-linear stochastic system'. *Int J Control* **80**(9), 1439–1453.

Maybeck, P. S. (1982). *Stochastic models, estimation, and control*. Vol. 2. Academic press.

Mayne, D. Q., James B Rawlings, Christopher V Rao and Pierre OM Scokaert (2000). 'Constrained model predictive control: Stability and optimality'. *Automatica* **36**(6), 789–814.

Migliore, S. A., E. A. Brown and S. P. DeWeerth (2005). Biologically inspired joint stiffness control. In 'Proc. IEEE Int. Conf. Robotics and Automation'. pp. 4508–4513.

Palmer, A. and D. Milutinovic (2011). A hamiltonian approach using partial differential equations for open-loop stochastic optimal control. In 'Proc. American Control Conf. (ACC)'. pp. 2056–2061.

Polit, A. and E Bizzi (1978). 'Processes controlling arm movements in monkeys.'. *Science (New York, N.Y.)* **201**, 1235–1237.

Polit, A. and E Bizzi (1979). 'Characteristics of motor programs underlying arm movements in monkeys.'. *J Neurophysiol* **42**, 183–194.

Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko (1964). *The Mathematical Theory of Optimal Processes.*. Pergamon Press.

Roberts, J. B. and P. D. Spanos (2003). *Random vibration and statistical linearization*. Dover Publications, Inc., Mineola, NY. Revised reprint of the 1990 original [Wiley, Chichester; MR1076193].

Särkkä, S. and Arno Solin (2019). *Applied Stochastic Differential Equations*. Institute of Mathematical Statistics Textbooks. Cambridge University Press.

Schoemaker, P. J. (1991). 'The quest for optimality: A positive heuristic of science?'. *Behav Brain Sci* **14**(2), 205–15.

Scott, S. H. (2012). 'The computational and neural basis of voluntary motor control and planning.'. *Trends in cognitive sciences* **16**, 541–549.

Socha, L. (2008). *Linearization methods for stochastic dynamic systems*. Vol. 730 of *Lecture Notes in Physics*. Springer, Berlin.

Stengel, R. (1986). *Optimal Control and Estimation*. Dover books on advanced mathematics. Dover Publications.

Theodorou, E., Jonas Buchli and Stefan Schaal (2010*a*). 'A generalized path integral control approach to reinforcement learning'. *journal of machine learning research* **11**(Nov), 3137–3181.

Theodorou, E., Yuval Tassa and Emo Todorov (2010*b*). Stochastic differential dynamic programming. In 'Proceedings of the 2010 American Control Conference'. IEEE. pp. 1125–1132.

Todorov, E. (2004). 'Optimality principles in sensorimotor control.'. *Nat Neurosci* **7**(9), 907–915.

Todorov, E. (2005). 'Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system.'. *Neural Comput* **17**(5), 1084–1108.

Todorov, E. (2009). 'Efficient computation of optimal actions.'. *Proc Natl Acad Sci U S A* **106**(28), 11478–11483.

Todorov, E. and Weiwei Li (2005). A generalized iterative lqg method for locally-optimal feedback control of constrained nonlinear stochastic systems. In 'American Control Conference, 2005. Proceedings of the 2005'. pp. 300 – 306 vol. 1.

Trélat, E. (2008). *Contrôle optimal : Théorie & applications*.

van Beers, R. J., P. Haggard and D. M. Wolpert (2004). 'The role of execution noise in movement variability.'. *J Neurophysiol* **91**(2), 1050–1063.

Vanderborght, B., A. Albu-Schaeffer, A. Bicchi, E. Burdet, D. Caldwell, R. Carloni, M. Catalano, G. Ganesh, M. Garabini, M. Grebenstein, G. Grioli, S. Haddadin, A. Jafari, M. Laffranchi, D. Lefeber, F. Petit, S. Stramigioli, N. Tsagarakis, M. Van Damme, R. Van Ham, L. C. Visser and S. Wolf (2012). Variable impedance actuators: Moving the robots of tomorrow. In 'Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems'. pp. 5454–5455.

Wojtkiewicz, S. F., B. F. Spencer and L. A. Bergman (1996). New insights on the application of moment closure methods to nonlinear stochastic systems. In A. Naess and S. Krenk (Eds.). 'IUTAM Symposium on Advances in Nonlinear Stochastic Mechanics'. Springer Netherlands. Dordrecht. pp. 479–488.

Yong, J. and Xun Y. Zhou (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations*. 1 edn. Springer.