



**HAL**  
open science

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

Bastien Berret, Frédéric Jean

► **To cite this version:**

Bastien Berret, Frédéric Jean. Efficient computation of optimal open-loop controls for stochastic systems. 2019. hal-02158875v1

**HAL Id: hal-02158875**

**<https://ensta-paris.hal.science/hal-02158875v1>**

Preprint submitted on 18 Jun 2019 (v1), last revised 19 Jul 2019 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Bastien Berret<sup>1,2,3,\*</sup> and Frédéric Jean<sup>4</sup>

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](mailto: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 movement neuroscience and robotics, 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 open-loop deterministic control actions. The rationale stems from observations in neural control of movement which highlighted that 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 an approach allowing the efficient computation of optimal open-loop controls for broad class of stochastic systems. We compare the predictions made by this framework to classical deterministic and stochastic frameworks, and illustrate that original departure from classical optimal solutions can be observed. We further exemplify its relevance to the planning of biological movement by showing that a well-known phenomenon, referred to as muscle co-contraction, can be accounted for naturally. Our framework may be relevant to all fields where the optimal choice of open-loop actions is significant.

## Introduction

In fields such as robotics or movement neuroscience, a preliminary motion planning stage is generally considered before movement execution [1, 2]. Optimality is a significant principle to rationalize this stage [3, 4]. 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 [5]. In deterministic optimal control (DOC), a powerful set of theoretical results and numerical tools exist to solve a broad variety of problems [6,7]. 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) [8]. However, this framework poses several issues. Besides the linear-quadratic-Gaussian (LQG) scenario [9, 10], 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 become a real challenge [11, 12]. 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 and sophisticated state estimation procedures. Yet, considering stochastic control variables may not be required or even desired to formulate certain motion planning problems.

Regarding the planning of biological movement for instance, several influential studies have emphasized that the brain can purposely modulate the mechanical impedance of the neuromusculoskeletal systems via feedforward co-contraction mechanisms [13,14]. Humans and animals are able to generate relatively stable motor behaviors even in the absence of on-line feedback circuitry [15,16]. These findings suggest that on-line feedback is not necessarily critical for stability and that open-loop control may be relevant piece of the neural control of movement. Yet, the sensorimotor system is affected by noise [17] which makes it stochastic in essence. These observations led us to consider a more restricted subclass of SOC problems where admissible controls are deterministic open-loop while 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 at this stage of motion planning. Feedback can always be exploited subsequently, for instance during movement execution to track a planned trajectory [10, 18] or in a model predictive control scheme [19]. 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 and methods are lacking overall to solve general SOOC problems. 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, [20, 21]), a critical feature of our approach consists of focusing  
 40 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. [22]). The present philosophy may be relevant to other fields such as robotics, especially  
 45 for robots with variable impedance actuators [23, 24]. More generally, the present SOOC framework may  
 be significant for systems with intermittent, long-latency or even absent on-line feedback.

## Methods

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

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

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  
 50 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)  
 55 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,

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

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  
 60 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 [20] for instance). One obtains in this way a  
 formulation as a DOC problem on a partial differential equation (PDE). However, solving such a problem  
 65 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 appli-

cations), 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  $\mathbf{m}_x(t) = \mathbb{E}[\mathbf{x}_t]$  and  $P_x(t) = \mathbb{E}[(\mathbf{x}_t - \mathbf{m}_x(t))(\mathbf{x}_t - \mathbf{m}_x(t))^\top]$  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  $\varphi(\mathbf{m}_x(T), P_x(T)) \in \mathcal{S}$ ;
- (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,

$$C(\mathbf{u}) = \int_0^T (L(\mathbf{m}_x(t), \mathbf{u}(t), t) + \text{tr}(Q(t)P_x(t))) dt + \psi(\mathbf{m}_x(T)) + \text{tr}(Q_f P_x(T)). \quad (3)$$

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.  $L = L(\mathbf{m}_x, P_x, \mathbf{x}, \mathbf{u}, t)$ , 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.

## Reduction to deterministic optimal control (DOC) problems

### Case of a linear SDE

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

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

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  $\mathbf{m}(t) = \mathbf{m}_y(t)$  and covariance  $P(t) = P_y(t)$ .

The propagation of  $\mathbf{m}$  and  $P$  is given by the following ordinary differential equations (e.g. [25]):

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

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  $(\mathbf{m}^u(t), P^u(t))$  of (5) with  $(\mathbf{m}^u(0), P^u(0)) = (\mathbf{m}^0, P^0)$  satisfies  $\varphi(\mathbf{m}^u(T), P^u(T)) \in \mathcal{S}$ .

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. [6, 9]).

*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

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

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  $\text{tr}(QP)$  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. [13] 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.

### 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,

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

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

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

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

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

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 [26] and to the monographs [27, 28] 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

150 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  $\tilde{A} = \tilde{A}(\mathbf{m}(t), P(t), \mathbf{u}(t), t)$ ,  $\tilde{\mathbf{b}} = \tilde{\mathbf{b}}(\mathbf{m}(t), P(t), \mathbf{u}(t), t)$  and  $\tilde{H} = \tilde{H}(\mathbf{m}(t), P(t), \mathbf{u}(t), t)$ . Therefore  $(\mathbf{m}(t), P(t))$  is determined as a solution of the following nonlinear control system,

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

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

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

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

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

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

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.

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

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

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 [29] 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 [30] or the discussion in [31]).

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

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

170 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):

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

175 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

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

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.

180 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 [32]). 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. 185 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 190 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, 195 other choices are available (see for instance [28, 33] on the choice of non-Gaussian distributions for the first point, and [29, 34] for computations of expectations).

# Results

## Comparison with DOC

200 A classical thought experiment to illustrate the fundamental difference between DOC and SOC frameworks is often used: the drunken spider example (e.g. [35]). In this example, a spider has to reach its home by either crossing a tiny bridge over a lake (shortest path) or by moving around the lake (safer yet 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  
205 otherwise the spider may fall off the bridge into the lake due to the effects of noise). This example illustrates that fundamentally different predictions may be done by the DOC and SOC frameworks. Here we illustrate that the SOOC framework may exhibit similar drastic changes of optimal trajectory depending on the level of noise.

We constructed a model in the spirit of the above example. Consider a dot moving in a plane, governed  
210 by the stochastic dynamics:

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

As in the rest of the study, the controls  $u$  and  $v$  are assumed to be deterministic. 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  $\varphi(y) = \frac{\sigma}{\sqrt{2\pi}} \exp(\frac{y^2}{2\sigma})$  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  
215 process and the parameter  $g$  specifies noise magnitude (i.e. the spider more or less drunk).

Set  $\mathbf{x}_t = (x_t, y_t)^\top$ . We fix the distribution of the initial state,  $\mathbf{x}_0 \sim \mathcal{N}(\mathbf{m}_0, P_0)$  and we consider two possible terminal states on the mean,  $\mathbf{m}_1$  and  $\mathbf{m}_2$ . The goal is to reach either target in mean at time  $T$  (the final state covariance being left free here) and minimize the cost:

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

This problem can be solved explicitly by Gaussian statistical linearization and by computing the expected  
220 values of (11) analytically. Then, numerical methods to solve a 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 optimal cost. Note that two targets are used for simplicity in this toy example, but only a single target could be used if we modeled the problem on a cylinder instead of a plane (the  $x$  coordinate would become an angle in this case). The results of simulations are shown  
225 in Figure 1 with  $\mathbf{m}_0 = \mathbf{0}$ ,  $P_0 = \text{diag}(10^{-3}, 10^{-3})$ ,  $\mathbf{m}_1 = (1, 0)^\top$  and  $\mathbf{m}_2 = (-2, 0)^\top$ . It is worth noting that a qualitative change of optimal strategy appears depending on the amount of noise  $g$  (optimal costs are emphasized in bold). 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.

230

## Comparison with SOC

To stress the difference between the proposed framework and the usual one in SOC, let us introduce the following simple example. We consider a noisy 1 degree-of-freedom system whose speed is directly controlled by feedback, so its dynamics writes as follows:

$$dx_t = v_t dt + g dw_t, \quad (20)$$

235 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

$$\frac{1}{2} \mathbb{E} \left[ \int_0^T (v_t^2 + qx_t^2) dt + q_f x_T^2 \right], \quad (21)$$

where  $T$  is a fixed time. It is known from LQG theory that the optimal solution is the feedback control  $v_t = k(t)x_t$  where the gain  $k$  is the solution of

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

240 Assume now that we only allow open-loop control, but we add a mechanical device so that the action on the velocity is proportional to the position (in other terms we control the impedance). The dynamics then writes

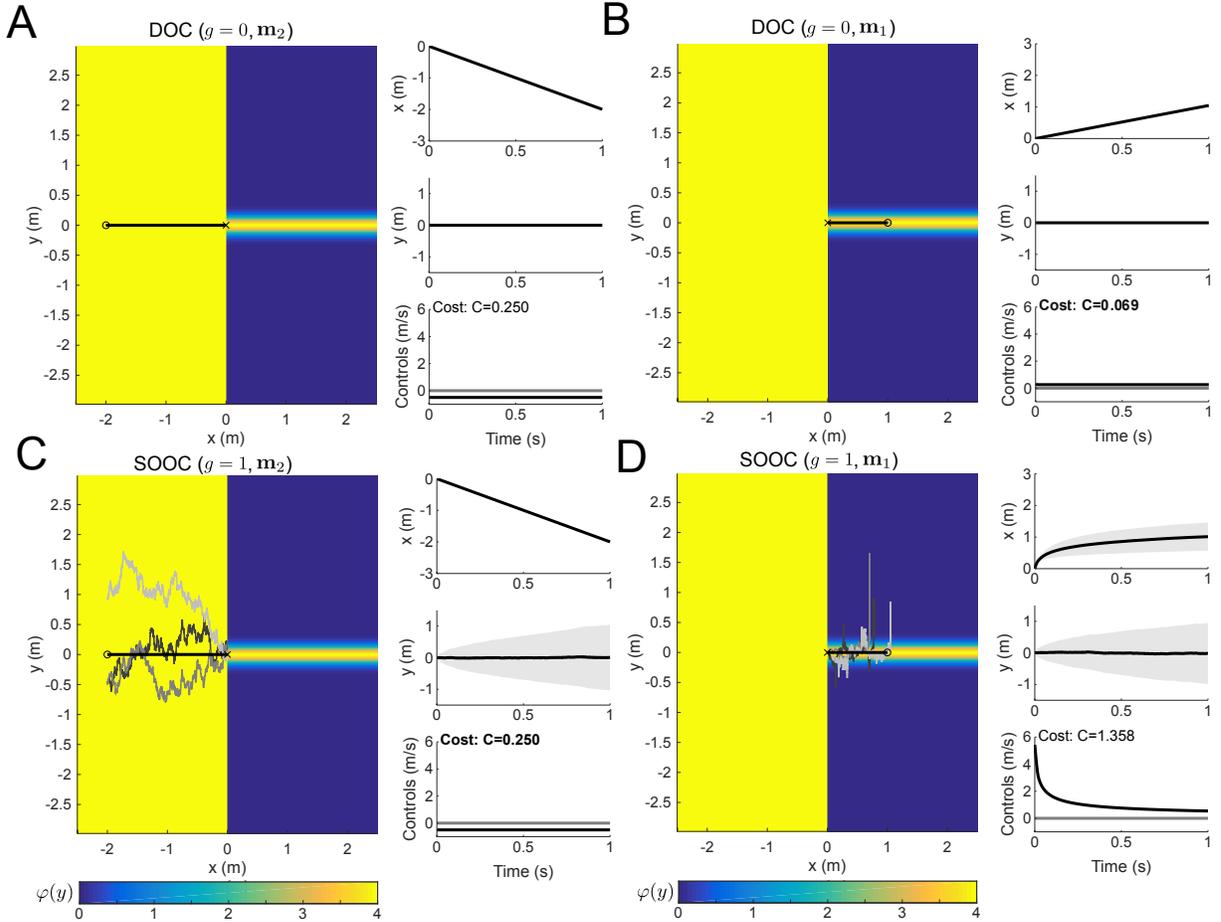
$$dx_t = u(t)x_t dt + g dw_t, \quad (23)$$

where the control  $u(t)$  is a (deterministic) function of the time, and the aim is to minimize the same compromise between position and velocity as before, that is:

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

245 The optimal solution of this problem is obviously  $u(t) = k(t)$  since this problem can be embedded into the first one by setting  $v_t = u(t)x_t$ . This can also be checked directly by solving (using Pontryagin's maximum principle, [36]) the equivalent (DOCP) which is associated with the dynamics

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



**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  $T = 1$ .

and with the cost

$$C(u) = \int_0^T \left( (m^2 + P)(u^2 + q) \right) dt + q_f(m(T)^2 + P(T)). \quad (26)$$

Hence, applying the respective optimal controls in (20) and in (23) will produce exactly the same trajectories.

250 Finally let us compare the SOC and SOOC frameworks when we minimize a cost that penalizes the control quadratically for each system, i.e. we now use the cost

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

in the open-loop framework.

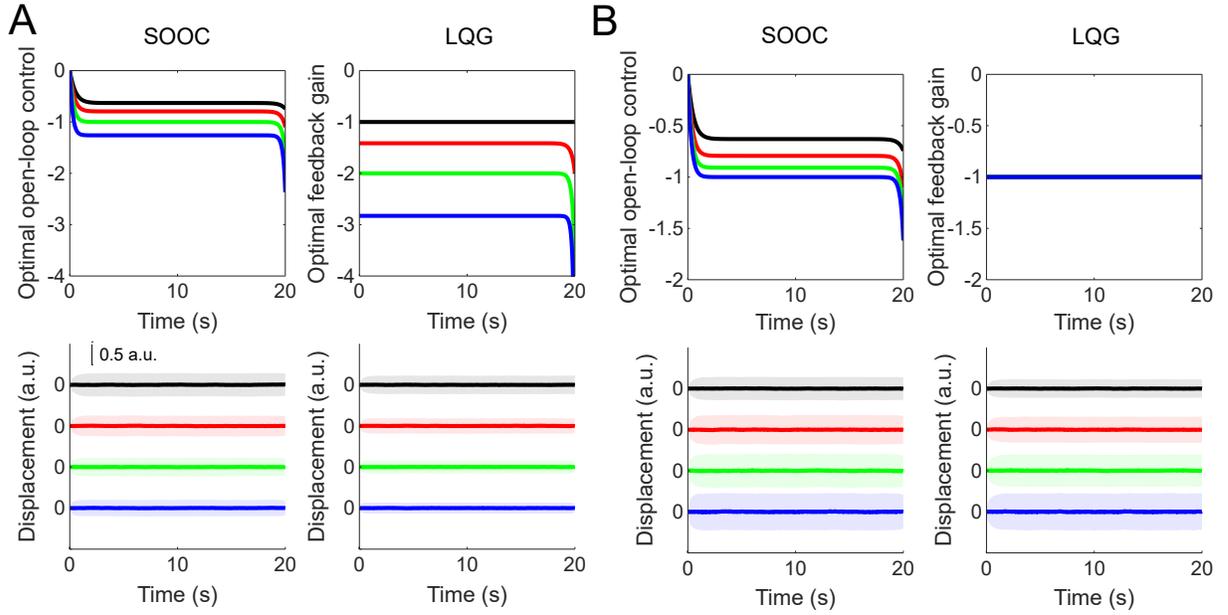
It is not possible anymore to transform the open-loop problem into an LQG problem by the change of control variable  $v = ux$ . We can however compute again the solution using Pontryagin's Maximum  
255 Principle, and we obtain that the optimal control is solution of an equation of the form

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

where the constant  $\mathcal{H}$  depends on  $q$ ,  $q_f$ , and  $g$  (see Appendix ). This must be compared with (22) characterizing the gain of the LQG solution, which does not depend on noise magnitude (via parameter  $g$ ). Figure 2 compares numerically the differences between the SOOC and SOC approaches.

260 In summary, the SOOC approach can give in some cases exactly the same trajectories as the LQG framework. There are however several important differences between our open-loop approach and LQG theory or more generally SOC.

- In general settings, SOC produces a feedback control such that measuring/estimating the state  $x_t$  is required throughout the process, which can be difficult for certain unstable tasks and systems with delays. In the human sensorimotor system for instance, feedback delays are relatively long  
265 (about 20-40 ms for short latency reflexes to more than 50 ms for feedback involving supraspinal centers, e.g. [37]). This may cause issues for the control of very fast movements and the achievement of unstable tasks. In such situations, humans tend to regulate their limb's impedance via muscle co-contraction mechanisms (e.g. [14, 38]). This is similar to having a mechanical system able to  
270 modify its intrinsic mechanical impedance, which may be implemented in artificial systems such as in robotic systems with variable stiffness actuators [24, 39, 40]. Note that the toy example of Figure 2 captures this characteristic via a bilinear control system.
- 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



**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.

275 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  
280 numerical scheme for the resolution of the Hamilton-Jacobi equation [11,12], iterative techniques  
such as iLQG [41] or differential dynamic programming [42] 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 [43–46]. Some  
limitations apply to the latter framework: re-computing the optimal control at each current state  
285 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,  
often, whereas the control is stochastic by nature, open-loop approximations are found numerically  
290 to avoid having to re-sample at every time instant of the real trajectory (e.g. [46,47]).

- 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 [48]. SOOC, on the other hand, depends explicitly on the noise level even for a simple additive  
295 noise (because it deals with non-linear dynamics, e.g. bilinear system of Fig. 2).
- 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 could yield to meaningless optimal trajectories and controls. This problem does not appear in the  
300 open-loop framework as hard constraints can be set, that is, we can deal with explicit terminal constraints on the mean and the covariance.

Actually, despite their differences, both approaches are very complementary. A natural scheme in many applications would be to use the present open-loop approach at the stage of the planning process, and then to use SOC to track a planned trajectory (for instance a LQG method after linearization around the mean  
305 trajectory like in iterative LQG schemes) at the execution stage. We next illustrate the SOOC approach in the context of human motor control. This relatively complex application is interesting because the central nervous system was shown to have the ability to tune the intrinsic impedance of the neuromusculoskeletal system in a feedforward manner, in particular via muscle co-contraction mechanisms.

## Application to neural control of movement

310 Here we focus on the control of reaching movements in humans with a model of the arm taken from [49]. 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 we aimed to use this model to test our modeling framework on a realistic and relatively complex system.

315 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:

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

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:

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

and

$$\begin{aligned} \mathcal{C}_{11}(\mathbf{q}, \dot{\mathbf{q}}) &= -2M_2 L_1 L_{g2} \sin(q_2) \dot{q}_2 \\ \mathcal{C}_{12}(\mathbf{q}, \dot{\mathbf{q}}) &= -M_2 L_1 L_{g2} \sin(q_2) \dot{q}_2 \\ \mathcal{C}_{21}(\mathbf{q}, \dot{\mathbf{q}}) &= M_2 L_1 L_{g2} \sin(q_2) \dot{q}_1 \\ \mathcal{C}_{22}(\mathbf{q}, \dot{\mathbf{q}}) &= 0 \end{aligned} \quad (31)$$

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:

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

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

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

where  $K(\mathbf{u}) = \text{diag}(k_0 + k\mathbf{u})$ ,  $B(\mathbf{u}) = \text{diag}(b_0 + b\mathbf{u})$  and  $\mathbf{l}_r(\mathbf{u}) = \mathbf{l}_0 + \text{diag}(r_1, \dots, r_6)\mathbf{u}$ . 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 [49].

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:

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

335 with

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

and

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

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 ones could be implemented (for instance, if noise acts at torque levels, we would need to replace the second component of  $G$  by  $\mathcal{M}^{-1}(\mathbf{q}_t)\text{diag}(\sigma_1, \sigma_2)$ , 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), and many non-linear interactions between state and control elements.

345 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:

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

with  $L(\mathbf{m}_{\mathbf{x}}, \mathbf{u}) = \mathbf{u}^\top \mathbf{u} + \alpha(a_x^2 + a_y^2)$  where  $a_x$  and  $a_y$  are the Cartesian accelerations of the endpoint along the  $x$  and  $y$  axes respectively (i.e. functions of  $\mathbf{m}_{\mathbf{x}}, \mathbf{u}$ , 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 [50–52].

350 The 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. The final term of the cost corresponds to a penalization of the state covariance, the magnitude of which is tuned by the weight  $q_{var}$ . 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.

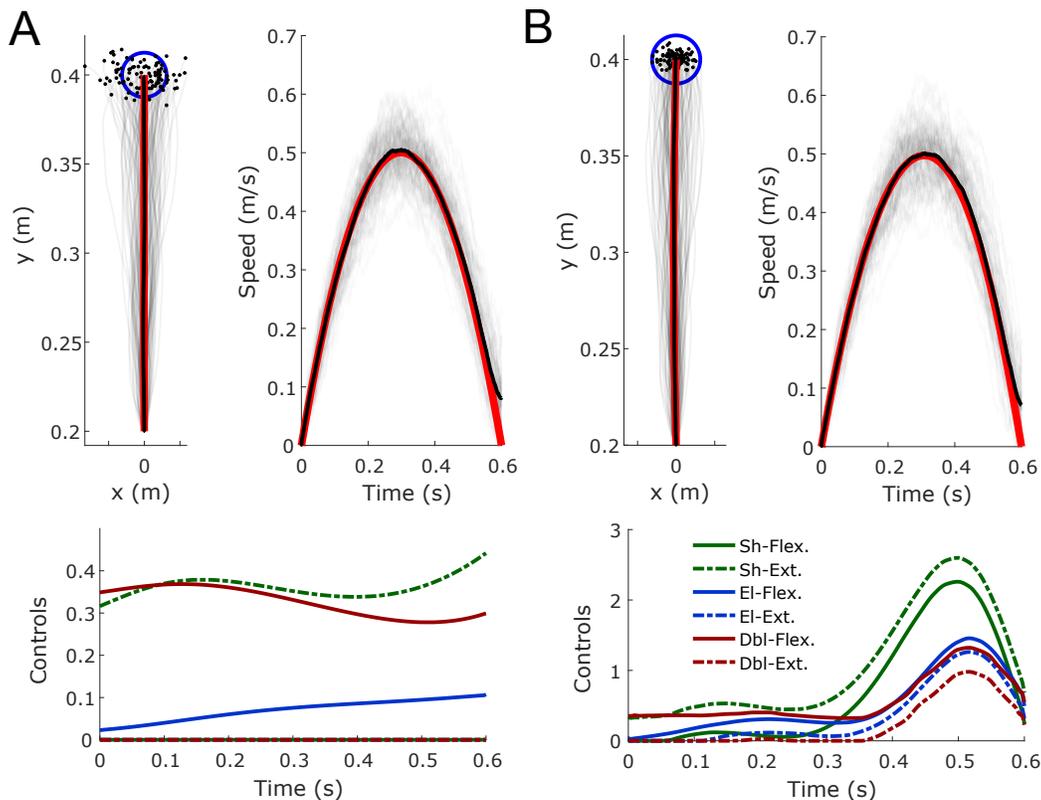
355

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.

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 [38]. These simulations illustrate the basis principles and motivations underlying the SOOC framework. As the role of co-contraction as a constitutive element of the motor plan is still elusive in neuroscience [53], such a framework may reveal itself useful for investigating the nature of descending motor commands in the neural control of movement. The SOOC framework may thus extent influential optimal control models for human movement (e.g. minimum variance model [54] and minimum jerk model [50]). Furthermore, it may provide a breeding-ground to better understand an ubiquitous phenomenon that is usually not easily accounted for by existing deterministic and stochastic models.

## 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 [13, 14, 55]. 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 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 laws after linearization around the planned (mean) trajectory. The latter was found to depend qualitatively on noise magnitude



**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:  $\alpha = 1$ ,  $\sigma_1 = \sigma_2 = 0.5$ .

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 feedback is available, the framework may also be used within a model predictive control approach. Future work will aim at exploiting the present framework to better decipher the role of muscle co-contraction in the neural control of movement.

400 **Appendix**

**Comparison with LQG: computation of the control**

Consider the problem of minimizing the cost

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

among the trajectories of

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

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

$$C(u) = \int_0^T (u^2 + q(m^2 + P)) dt + q_f(m(T)^2 + P(T)), \quad (40)$$

405 among the trajectories of

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

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:

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

with terminal constraints

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

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

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

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

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

The necessary conditions become

$$\begin{cases} \dot{\lambda}_m &= -u\lambda_m + 2m(q + u^2), \\ \dot{\lambda}_P &= -2u\lambda_P + q + u^2, \end{cases} \quad (m^2 + P)u = \frac{\lambda_m m}{2} + \lambda_P P, \quad (46)$$

415 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

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

## Alternative approximation with higher order Taylor expansions

An alternative approximation could be as follows (e.g. [30, Chap. 12]). Taking the expansions at order 420  $N = 2$  for  $\mathbf{f}$  and at order  $N = 1$  for  $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$  and  $GG^\top$ , so that:

$$\begin{aligned} \tilde{\mathbf{b}} &\sim \mathbf{f}(\mathbf{m}, \mathbf{u}) + \frac{1}{2} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^2}(\mathbf{m}, \mathbf{u}) \bullet P, \quad \text{where } \frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^2} \bullet P = \left( \text{tr} \left( \frac{\partial^2 \mathbf{f}_1}{\partial \mathbf{x}^2} P \right), \dots, \text{tr} \left( \frac{\partial^2 \mathbf{f}_n}{\partial \mathbf{x}^2} P \right) \right)^\top \\ \tilde{A} &\sim \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{m}, \mathbf{u}), \quad \text{and} \quad \tilde{H} \tilde{H}^\top \sim G(\mathbf{m}, \mathbf{u}) G(\mathbf{m}, \mathbf{u})^\top, \end{aligned} \quad (48)$$

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

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

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 [28] or [56]).

## References

1. Todorov E. Optimality principles in sensorimotor control. *Nat Neurosci.* 2004 Sep;7(9):907–915.
2. LaValle SM. *Planning algorithms.* Cambridge university press; 2006.
3. Schoemaker PJ. The quest for optimality: A positive heuristic of science? *Behav Brain Sci.* 1991 Jun;14(2):205–15.
4. Berret B, Delis I, Gaveau J, Jean F. In: Venture G, Laumond JP, Watier B, editors. *Optimality and Modularity in Human Movement: From Optimal Control to Muscle Synergies.* Cham: Springer International Publishing; 2019. p. 105–133.
5. Fleming WH, Rishel RW. *Deterministic and stochastic optimal control. Applications of mathematics.* Springer-Verlag; 1975.
6. Kirk DE. *Optimal control theory: An Introduction.* Prentice-Hall, New Jersey; 1970.
7. Trélat E. *Contrôle optimal : Théorie & applications.* Vuibert, editor; 2008.
8. Yong J, Zhou XY. *Stochastic Controls: Hamiltonian Systems and HJB Equations.* 1st ed. Springer; 1999.
9. Bryson AE, Ho YC. *Applied Optimal Control.* New York: Blaisdell; 1969.
10. Athans M. The Role and Use of the Stochastic Linear-Quadratic-Gaussian Problem in Control System Design. *IEEE Trans Autom Control.* 1971;16(6):529–552.
11. Fahim A, Touzi N, Warin X. A probabilistic numerical method for fully nonlinear parabolic PDEs. *Ann Appl Probab.* 2011;21(4):1322–1364.
12. Falcone M, Ferretti R. *Semi-Lagrangian approximation schemes for linear and Hamilton-Jacobi equations.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; 2014.
13. Hogan N. Adaptive control of mechanical impedance by coactivation of antagonist muscles. *IEEE Trans Autom Control.* 1984;29(8):681–690.
14. Burdet E, Osu R, Franklin DW, Milner TE, Kawato M. The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature.* 2001 Nov;414(6862):446–449.
15. Polit A, Bizzi E. Processes controlling arm movements in monkeys. *Science (New York, NY).* 1978 Sep;201:1235–1237.
16. Polit A, Bizzi E. Characteristics of motor programs underlying arm movements in monkeys. *J Neurophysiol.* 1979 Jan;42:183–194.
17. Faisal AA, Selen LPJ, Wolpert DM. Noise in the nervous system. *Nat Rev Neurosci.* 2008 Apr;9(4):292–303.

18. Todorov E, Li W. A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. In: American Control Conference, 2005. Proceedings of the 2005; 2005. p. 300 – 306 vol. 1.
19. Mayne DQ, Rawlings JB, Rao CV, Scolaert PO. Constrained model predictive control: Stability and optimality. *Automatica*. 2000;36(6):789–814.
20. Annunziato M, Borzi A. A Fokker-Planck control framework for multidimensional stochastic processes. *J Comput Appl Math*. 2013;237(1):487–507.
21. Palmer A, Milutinovic D. A Hamiltonian approach using partial differential equations for open-loop stochastic optimal control. In: Proc. American Control Conf. (ACC); 2011. p. 2056–2061.
22. van Beers RJ, Haggard P, Wolpert DM. The role of execution noise in movement variability. *J Neurophysiol*. 2004 Feb;91(2):1050–1063.
23. Migliore SA, Brown EA, DeWeerth SP. Biologically Inspired Joint Stiffness Control. In: Proc. IEEE Int. Conf. Robotics and Automation; 2005. p. 4508–4513.
24. Vanderborght B, Albu-Schaeffer A, Bicchi A, Burdet E, Caldwell D, Carloni R, et al. Variable impedance actuators: Moving the robots of tomorrow. In: Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems; 2012. p. 5454–5455.
25. Maybeck PS. Stochastic models, estimation, and control. vol. 1. Academic press; 1979.
26. Elishakoff I, Crandall SH. Sixty years of stochastic linearization technique. *Meccanica*. 2017 Jan;52(1):299–305.
27. Roberts JB, Spanos PD. Random vibration and statistical linearization. Dover Publications, Inc., Mineola, NY; 2003. Revised reprint of the 1990 original [Wiley, Chichester; MR1076193].
28. Socha L. Linearization methods for stochastic dynamic systems. vol. 730 of Lecture Notes in Physics. Springer, Berlin; 2008.
29. Särkkä S, Solin A. Applied Stochastic Differential Equations. Institute of Mathematical Statistics Textbooks. Cambridge University Press; 2019.
30. Maybeck PS. Stochastic models, estimation, and control. vol. 2. Academic press; 1982.
31. Gustafsson F, Hendeby G. Some Relations Between Extended and Unscented Kalman Filters. *IEEE Trans Signal Process*. 2012;60(2):545–555.
32. Beaman JJ. Non-linear quadratic gaussian control. *Int J Control*. 1984;39(2):343–361.
33. Crandall SH. A half-century of stochastic equivalent linearization. *Structural Control and Health Monitoring*. 2006 1;13(1):27–40.

34. Ghusinga KR, Soltani M, Lamperski A, Dhople SV, Singh A. Approximate moment dynamics for polynomial and trigonometric stochastic systems. In: 2017 IEEE 56th Annual Conference on Decision and Control (CDC); 2017. p. 1864–1869.
35. Kappen HJ. Path integrals and symmetry breaking for optimal control theory. *J Stat Mech: Theory Exp.* 2005;2005(11):P11011.
36. Pontryagin LS, Boltyanskii VG, Gamkrelidze RV, Mishchenko EF. *The Mathematical Theory of Optimal Processes.* Pergamon Press; 1964.
37. Scott SH. The computational and neural basis of voluntary motor control and planning. *Trends in cognitive sciences.* 2012 Nov;16:541–549.
38. Gribble PL, Mullin LI, Cothros N, Mattar A. Role of cocontraction in arm movement accuracy. *J Neurophysiol.* 2003 May;89(5):2396–2405.
39. Berret B, Ivaldi S, Nori F, Sandini G. 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;* 2011. p. 4354–4359.
40. Fiorio L, Romano F, Parmiggiani A, Berret B, Metta G, Nori F. In: Venture G, Laumond JP, Watier B, editors. *Design and Control of a Passive Noise Rejecting Variable Stiffness Actuator.* Cham: Springer International Publishing; 2019. p. 235–262.
41. Li W, Todorov E. Iterative linearization methods for approximately optimal control and estimation of non-linear stochastic system. *Int J Control.* 2007;80(9):1439–1453.
42. Theodorou E, Tassa Y, Todorov E. Stochastic differential dynamic programming. In: *Proceedings of the 2010 American Control Conference.* IEEE; 2010. p. 1125–1132.
43. Kappen HJ. A linear theory for control of non-linear stochastic systems. *Phys Rev Lett.* 2005;95:200–201.
44. Kappen HJ. Optimal control theory and the linear Bellman equation. In: Barber D, Cemgil AT, Chiappa S, editors. *Bayesian Time Series Models.* Cambridge University Press; 2011. p. 363–387. Cambridge Books Online.
45. Todorov E. Efficient computation of optimal actions. *Proc Natl Acad Sci U S A.* 2009 Jul;106(28):11478–11483.
46. Theodorou E, Buchli J, Schaal S. A generalized path integral control approach to reinforcement learning. *journal of machine learning research.* 2010;11(Nov):3137–3181.
47. Berret B, Yung I, Nori F. 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;* 2013. p. 3029–3034.

48. Todorov E. Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural Comput.* 2005 May;17(5):1084–1108.
49. Katayama M, Kawato M. Virtual trajectory and stiffness ellipse during multijoint arm movement predicted by neural inverse models. *Biol Cybern.* 1993;69:353–362.
50. Flash T, Hogan N. The coordination of arm movements: an experimentally confirmed mathematical model. *J Neurosci.* 1985 Jul;5(7):1688–1703.
51. Berret B, Darlot C, Jean F, Pozzo T, Papaxanthis C, Gauthier JP. The inactivation principle: mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements. *PLoS Comput Biol.* 2008 Oct;4(10):e1000194.
52. Berret B, Chiovetto E, Nori F, Pozzo T. Evidence for composite cost functions in arm movement planning: an inverse optimal control approach. *PLoS Comput Biol.* 2011 Oct;7(10):e1002183.
53. Latash ML. Muscle coactivation: definitions, mechanisms, and functions. *J Neurophysiol.* 2018 Jul;120:88–104.
54. Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. *Nature.* 1998 Aug;394(6695):780–784.
55. Franklin DW, Osu R, Burdet E, Kawato M, Milner TE. Adaptation to stable and unstable dynamics achieved by combined impedance control and inverse dynamics model. *J Neurophysiol.* 2003 Nov;90:3270–3282.
56. Wojtkiewicz SF, Spencer BF, Bergman LA. New Insights on the Application of Moment Closure Methods to Nonlinear Stochastic Systems. In: Naess A, Krenk S, editors. *IUTAM Symposium on Advances in Nonlinear Stochastic Mechanics*. Dordrecht: Springer Netherlands; 1996. p. 479–488.