Open-loop stochastic optimal control of a passive noise-rejection variable stiffness actuator: application to unstable tasks

Bastien Berret, I Yung and Francesco Nori

***Abstract*—In this paper we propose a methodology to control a novel class of actuators that we called passive noise rejection variable stiffness actuators (pnrVSA). Differently from nowadays classical VSA designs, this novel class of actuators mimics the human musculoskeletal ability to increase noise rejection without relying on feedback. To fully highlight the potentialities behind these actuators we consider movement planning under two constraints: (1) absence of feedback, i.e. purely open-loop planning[1]; (2) uncertain dynamic model. Under these constraints, movement planning can be formalized as an open-loop stochastic optimal control. Due to the lack of classical methods forcing the open-loop nature of the computed solution, we used here a slight modification of available methodologies based on importance sampling of trajectories using forward diffusion processes. Simulations show that the proposed algorithm can be effectively used to plan open-loop movements with pnrVSA. In particular, two different scenarios are considered: the control of a single joint pnrVSA and the control of a two degrees of freedom planar arm equipped with antagonist pnrVSAs at each joint. In both cases, movement has to be planned in presence of uncertain dynamics for unstable tasks. It is shown that open-loop stochastic optimal control can modulate the intrinsic stiffness of the system to cope with both instability and noise.**

***Index Terms*—open-loop stochastic optimal control, noise rejection, variable stiffness actuator, unstable task**

I. Introduction

Designing and controlling robots with rigid actuators is an idealization that becomes limiting for some tasks or applications. In the field of robotics, current interest in variable stiffness actuators (i.e. actuators with adjustable rigidity) has been mainly concentrated on safety, interaction and mechanical robustness. Nowadays there is enough evidence supporting the idea that humans exploit muscle co-activation (which is related to rigidity regulation) to cope with sensorimotor delays and noise in presence of instabilities [1]. This aspect has been weakly explored in robotics and the current article tries to explore its potential by showing (in a stochastic scenario) that certain types of variable stiffness actuators can cope with instabilities in an open-loop manner (i.e. without relying on feedback). The finding that monkeys actually specify the intrinsic musculoskeletal impedance even when deafferented [2], further reinforced the idea that stiffness regulation is indeed a crucial movement feature that (in biological systems) is not realized with explicit feedback loops. Heavily relying on feedback in artificial agents (such as humanoid robots) might not be a viable strategy especially considering the growing amount of sensors (e.g. below the traditional ones: whole-body distributed tactile sensors [3], whole-body distributed force/torque sensors [4], whole-body distributed gyros and accelerometers [5]) which are currently available and have to be centrally acquired/processed to perform complex actions. Along this direction we recently proposed [6] a coupling between variable stiffness actuation (VSA, e.g. [7]) and open-loop (e.g. [8]) stochastic optimal control (SOC; e.g. [9,10]) as an attractive framework to circumvent problems related to temporal latencies in transmissions or noisy state estimations related to feedback control.

In this paper we further proceed along this line of research. The original contribution of this paper is the derivation of a complete framework to efficiently control pnrVSA in a stochastic optimal control context. First, we present a simple algorithm to compute open-loop controllers from the formulation of generic stochastic optimal control (SOC) problems. Quite a number of solutions exist for closed-loop stochastic planning [11–13]; here we use a recent approach called path-integral SOC to construct controllers that cope with an issue which is often neglected for practical[2] and technical[3] reasons: the absence of feedback (i.e. pure open-loop planning). The effectiveness of the framework is shown in the context of unstable tasks, for which coping with uncertainties without explicitly relying on feedback is very challenging.

The paper is organized as follows. Section II-B describes the class of variable stiffness actuators considered in this paper and gives its dynamical model. Section II-C presents our modification of classical closed-loop approaches to deal with open-loop planning. Section II-D discusses the considered tasks, a single joint and a two degrees of freedom planar arm both controlled under stochastic and unstable conditions. Section III presents the simulation results and

All the authors are with the Robotics, Brain and Cognitive Sciences Department, Italian Institute of Technology, Genoa, Italy (corresponding author: francesco.nori@iit.it).

The authors were supported by the European Projects: VIACTORS (FP7-ICT-2007-3, contract 231554) and ITALK (ICT-214668). I Yung was supported by Erasmus Mundus Scholarship (European Master on Advanced Robotics).

[1] A more realistic scenario would consider a mixture of open-loop and closed-loop control but in this paper we stress on the first to fully understand the features behind pnrVSA.

[2] Robots nowadays rely on fast feedback loops with delays and latencies of approximately $1ms$; it makes therefore no sense to force solutions which do not take into account this efficient feedback. Our motivation is slightly different and therefore the request of open-loop solutions becomes a necessity.

[3] To our knowledge, focus on open-loop control is difficult in this kind of literature mainly because the Hamilton-Jacobi-Bellman equation naturally leads to a feedback solution.

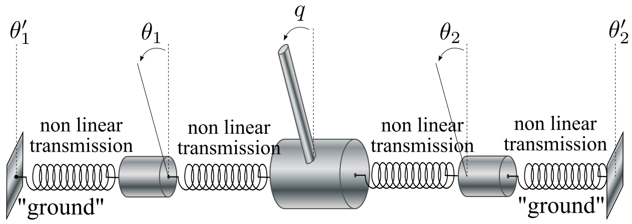

Figure 1. *Variable stiffness actuator with passive noise rejection.* The variable θ_1 and θ_2 are the motor angles, the variable q is the joint angle. The constants θ_1' and θ_2' are fixed angles of reference.

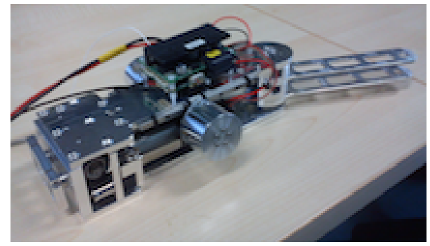

Figure 2. *Variable stiffness actuator with passive noise rejection.* A picture of the first prototype.

proves that the proposed VSA actuator model can deal in open-loop with the extreme tasks considered. Finally Section IV draws the conclusion.

II. Methods

A. Framework and rationale

This work relies on several conceptual considerations that are important to describe. First, throughout this paper, we shall consider a variable stiffness actuator allowing to emulate the human-like property of muscle co-contraction in order to reject disturbances. This VSA device has been already described thoroughly in [14–16] and thus only a brief description of it will be given in Section B. The relevance of such property for VSA can be exemplified in the context of unknown dynamics/environment, for which uncertainty is often model mathematically as noise. All physical and biological systems are actually affected by such a noise to some extent. We argue here that one fundamental advantage of VSA is to cope with all unpredictable fluctuations affecting such a stochastic control system. Additionally, we will voluntarily prevent us from using feedback control in order to emphasize that VSA can play the role of an immediate feedback via adequate co-contraction. We choose this extreme scenario to prove the possible usefulness of noise-rejection VSA in the context of stochastic dynamics. Therefore, we will focus on open-loop stochastic control. Since co-contraction usually costs energy, an additional goal is to minimize energy consumption and to automatically find the minimal stiffness required to perform the task. Thus, the controller must optimally trade-off stiffness and compliance depending on the task and the noise magnitude. For all these reasons, our framework will be the one of *open-loop stochastic optimal control*. The chosen approach will be reviewed in Section C. Finally, all the above-mentioned ideas will be emphasized in the case of unstable tasks, i.e. tasks in which noise can make the system diverge very rapidly if feedback is delayed. The reference case that motivated most of this work is presented in [1]. Section D will present the unstable tasks we consider here.

B. Variable stiffness actuator

Here we consider the variable stiffness actuator that we previously developed in [14–16] and that we named passive noise rejection variable stiffness actuator: pnrVSA. A thorough discussion of various classes of VSA is out of the scope of this article, but one crucial difference in the chosen VSA design is its ability to "passively reject noise" via co-contraction. See for instance [17] for a similar class of VSA, yet lacking this property. Here, for one joint, the actuation system is constituted of 4 non-linear flexible transmission elements, two of which being connected to a fixed reference. Figure 1 sketches the pnrVSA considered in this paper.

Assuming cubic springs for simplicity, the dynamical system can be derived from Lagrangian mechanics, and are as follows:

$$\begin{aligned} I\ddot{q} &= k_1(\theta_1 - q)^3 + k_2(\theta_2 - q)^3 - b\dot{\theta} + \tau, \\ I_1\ddot{\theta}_1 &= k_1(q - \theta_1)^3 + k_1'(\theta_1' - \theta_1) - b_1\dot{\theta}_1 + \tau_1, \\ I_2\ddot{\theta}_2 &= k_2(q - \theta_2)^3 + k_2'(\theta_2' - \theta_2) - b_2\dot{\theta}_2 + \tau_2. \end{aligned} \tag{1}$$

Remarkably, in this formulation there is no direct control on q which is anyway the variable to be controlled. The quantity τ is only used to represent the external interaction with the environment. The internal driving torques, which are antagonist, are denoted by τ_1 and τ_2. Their combined action is used to indirectly control q via the net torque $\tau_q = k_1(\theta_1 - q)^3 + k_2(\theta_2 - q)^3$ which in a sense represents the internal torque acting on q. Remarkably the quantity $K_q = -\partial\tau_q/\partial q = 3\big(k_1(\theta_1 - q)^2 + k_2(\theta_2 - q)^2\big)$ can be used to represent the stiffness of the variable q due the internal torques. The fact that this stiffness can be tuned with the variables θ_1 and θ_2 accounts for the possibility to reject disturbances by "co-contracting" the two antagonist actuators.

When dealing with a multi-joint system, each joint is equipped with such pnrVSA system. In the present work we neglect bi-articular actuators whose role will be investigated in future studies. In our simulations, we will consider a 2-joint planar arm equipped with two pnrVSA, as depicted in Figure 3.

Still considering cubic springs, the complete system dynamics writes as:

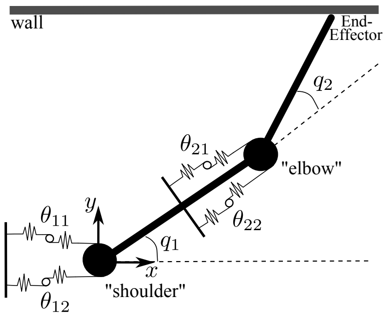

Figure 3. 2-dof arm equipped with antagonist pnrVSAs and illustration of the unstable task considered in Subsection D.

$$\begin{array}{rcl} M(\mathbf{q})\ddot{\mathbf{q}} & = & \big(\tau_1 - \tau_2 + \tau_{ext} - \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\big), \\ \tau_i & = & k_{i1}(\theta_{i1} - q_i)^3 + k_{i2}(\theta_{i2} - q_i)^3, \\ I_{ij}\ddot{\theta}_{ij} & = & k_{ij}(q_i - \theta_{ij})^3 + k'_{11}(\theta'_{ij} - \theta_{ij})^3 - b_{ij}\dot{\theta}_{ij} + F_{ij}, \\ i & = & 1, 2 \end{array}$$

The first row represents a standard rigid body dynamics (with the mass matrix M and the Coriolis/centripetal/friction term $\mathbf{C}$), the second row represents the torques caused by the spring elongations and the last row represents the internal actuation dynamics of the spring contraction. Importantly, the stiffness of variable q_i will be expressed as $K_{q_i} = 3k_{i1}(\theta_{i1} - q_i)^2 + 3k_{i2}(\theta_{i2} - q_i)^2$ for $i = 1, 2$.

It is straightforward to write these dynamical systems in state space and to see that they fall in the class of system considered in Subsection C.

C. Open-loop stochastic optimal control

In all the reported simulations, we consider stochastic nonlinear control-affine systems:

$$d\mathbf{x} = \mathbf{a}(\mathbf{x}(t), t)dt + B(\mathbf{x}(t), t)\mathbf{u}(\mathbf{x}(t), t)dt + C(\mathbf{x}, t)d\mathbf{w}, \quad (2)$$

where the variable $\mathbf{w}$ denotes a standard m-dimensional Brownian motion, $\mathbf{x}(t) \in \mathbb{R}^n$ is the state of the system and $\mathbf{u}(\mathbf{x}, t) \in \mathbb{R}^m$ is the control (at this point, the control is a function of both the time and the state). The controlled stochastic differential Equation (2) is defined in the sense of Itō's integral. In the following, we will often omit explicit dependencies on $\mathbf{x}$ or t to simplify notations.

Additionally, we consider an expected cost of the form:

$$J(\mathbf{x}_0, t_0) = \mathbb{E}[\phi\big(\mathbf{x}(t_f)\big) + \int_{t_0}^{t_f} q(\mathbf{x}, t) + \frac{1}{2}\mathbf{u}^\top R\mathbf{u}\, dt], \quad (3)$$

in which the source state $\mathbf{x}_0$ and initial time t_0 are assumed to be known and fixed.

Equations (2)-(3) form a quite general class of stochastic optimal control (SOC) problems, whose optimal control $\mathbf{u}$ can be expressed as $\mathbf{u} = -R^{-1}B^\top \nabla_{\mathbf{x}} J(\mathbf{x}, t)$, where the optimal cost-to-go function J satisfies the stochastic version of the Hamilton-Jacobi-Bellman equation:

$$-J_t = \min_{\mathbf{u}}[q + \frac{1}{2}\mathbf{u}^\top R\mathbf{u} + (\mathbf{a} + B\mathbf{u})^\top J_{\mathbf{x}} + \frac{1}{2}tr(CC^\top J_{\mathbf{xx}})]. \quad (4)$$

It is now well-known (see [11–13]) that this partial differential equation can be made linear if we rewrite it in terms of the desirability function $\psi = \exp(-\frac{1}{\lambda}J)$ and if we further assume that $C = B\sqrt{\lambda R^{-1}}$ for some chosen λ. It can be then shown that the optimal control can be expressed at each state/time as a path integral [11, 18], that can be approximated via importance sampling methods. This is in this framework that the $\mathbf{PI}^2$ algorithm was developed [13]. Here we use a pretty similar algorithm, though we will not make use of the dynamic movement primitives. Although the optimal control is theoretically a feedback control law, the algorithm is designed such as to find an open-loop control policy. In order to present the numerical algorithm used to derive open-loop controls, we now consider a discrete-time representation of Eq. (2):

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{a}_i dt + B_i\mathbf{u}_i dt + \sqrt{dt}C_i\epsilon_i, \quad (5)$$

with $\epsilon_i = \mathcal{N}(\mathbf{0}, I_m)$ (centered and normalized Gaussian noise) and with $\epsilon_i = \mathcal{N}(\mathbf{0}, I_m)$ (centered and normalized Gaussian noise) and $dt = \frac{T}{n}$, $T = t_f - t_0$, $\mathbf{x}_i = \mathbf{x}(idt + t_0)$ with $i \in [0, n]$. We denote a path starting at state $\mathbf{x}_i$ by $\boldsymbol{\tau}_i = (\mathbf{x}_i, ..., \mathbf{x}_n)$.

Algorithm 1 Main steps of the path integral SOC algorithm used in simulations.

1. Initialize $\mathbf{u} = (\mathbf{u}_0, ..., \mathbf{u}_{n-1})$.
2. Sample K paths $\boldsymbol{\tau}_0^{(k)}$ starting from $\mathbf{x}_0$ using Eq. 5 (forward sampling using the guided diffusion)
3. For all $i = 0..n-1$, compute the control update $\delta\mathbf{u}_i = \sum_{k=1}^{K} P(\boldsymbol{\tau}_i^{(k)})\mathbf{u}_L(\boldsymbol{\tau}_i^{(k)})$ given:
   - $S(\boldsymbol{\tau}_i^{(k)}) = \phi(\mathbf{x}_n^{(k)}) + \sum_{j=i}^{n-1} \mathbf{u}_j^\top B_j^\top H_j^{-1}[C_j\epsilon_j^{(k)}\sqrt{dt} + \frac{1}{2}B_j\mathbf{u}_j dt)] + \sum_{j=i}^{n-1} q_j^{(k)}dt]$ (generalized cost)
   - $P(\boldsymbol{\tau}_i^{(k)}) = \frac{\exp\big(-\frac{1}{\lambda}S(\boldsymbol{\tau}_i^{(k)})\big)}{\sum_{k=1}^{K}\exp\big(-\frac{1}{\lambda}S(\boldsymbol{\tau}_i^{(k)})\big)}$ (probability of a path)
   - $\mathbf{u}_L(\boldsymbol{\tau}_i^{(k)}) = R^{-1}B_i^\top H_i^{-1}\epsilon_i^{(k)}\sqrt{dt}$ and $H_j = B_jR^{-1}B_j^\top$ (local control induced by noise instance)
4. Update $\mathbf{u}_i$ using $\mathbf{u}_i dt \leftarrow M_i\mathbf{u}_i dt + \delta\mathbf{u}_i$, with $M_i = R^{-1}B_i^\top H_i^{-1}B_i$.
5. Reiterate from step 2 and continue until convergence or maximum number of iterations

Justifications for the above algorithm can be found in [11] and [13].

D. Unstable tasks

We shall illustrate the effectiveness of pnrVSA for unstable tasks. The first task is a simple proof-of-concept example involving a single-joint system evolving a divergent force field as inspired by [1]. The second task is a multi-joint arm that is pushing against a wall with its endpoint. This more complex example can be viewed as a simplified modeling of screw driving, which is a typical daily life unstable task. Note that all simulations will be conducted in the framework of open-loop stochastic optimal control.

The first task thus involves stabilizing one joint in a divergent force-field environment and in open-loop. In state space, setting $\mathbf{x} = (q, \theta_1, \theta_2, \dot{q}, \dot{\theta}_1, \dot{\theta}_2)^\top$ we can rewrite the system given by Eq. (1) in the control-affine form given by Eq. 2. We consider a state-dependent cost $\phi(\mathbf{x}) = q(\mathbf{x}) = \mathbf{x}^T Q \mathbf{x}$, with $Q = \mathrm{diag}(10^5, 0, 0, 10^3, 10^3, 10^3)$, which essentially penalizes deviations of the variable q from zero. We set $\lambda = 0.1$. The initial and final configurations were defined as $\mathbf{x}(t_f) = (0.5, 0, 0, 0, 0)^\top$ and $\mathbf{x}(0) = (0, -0.5, 0.5, 0, 0, 0)^\top$ and the task duration was $t_f = 1\,\mathrm{s}$. The nonlinear transmission elements (i.e. cubic springs) were characterized by the following parameters: $I = I_i = 0.1$, $k_i = k'_i = 2$ and $b = b_i = 2$, $i = 1, 2$. The ground position was set to $\theta'_1 = \theta'_2 = 0$. Note that we also added a cost to penalize unrealistic behaviors, such as having a spring crossing the reference into the term $q(\mathbf{x})$.

For the simulations we chose a step size $dt = 0.01\,\mathrm{s}$. We used $K = 1000$ samples at each update for a total of 200 main iterations. We checked that the number of iterations was large enough to get a stabilized solution (a posteriori). The task was made more challenging by introducing a specific force field acting on the end-effector composed of a divergent torque and a constant one: $\tau = K_{\mathrm{div}} q + \tau_c$ with $K_{\mathrm{div}} = 5$ N.m/rad and $\tau_c = -5$ N.m. The goal of the task was to move from the initial state to a final end-effector resting position (the choice of $\theta_1(t_f)$ and $\theta_2(t_f)$ being free and found automatically) using an open-loop control and minimizing the amount of control effort.

The second task is a simplification of screw driving and was illustrated in Figure 3. The properties of 2-dof arm are based on human-like measures: $l_1 = 0.3\,\mathrm{m}$, $l_2 = 0.4\,\mathrm{m}$, $m_1 = 1.4\,\mathrm{kg}$, $m_2 = 1.1\,\mathrm{kg}$, center of mass positions $lc_1 = 0.11\,\mathrm{m}$, $lc_2 = 0.16\,\mathrm{m}$, and moment of inertia of each joint is computed by $m_i lc_i^2$. The friction of internal motors are chosen to reduce the oscillation in the system, $b_{11} = b_{12} = b_{21} = b_{22} = 4$Nm/sec and the inertia of the motors are assumed to be $2\,\mathrm{Nm/sec}^2$. The springs stiffness are $k_{11} = k_{12} = k_{21} = k_{22} = 100\,\mathrm{Nm/rad}$ The wall reaction force enters into the dynamics via a term $\tau_{ext} = J_y^\top(\mathbf{q})\lambda_{physical}$. The task is defined such as to push against the wall with a constant force equal to 10 N in the y-axis and this constraint is added to the state-dependent cost, along with the constraint of maintaining the end-effector at the position $x = 0.3$ and $y = 0.4$. The goal is to maintain this arm configuration for two seconds in open-loop and by minimizing the amount of control energy in order to find the minimal level of co-contraction. Instability in this case results from the fact that pushing against the wall with a constant normal force causes unwanted tangential dragging when the link is not normal to the wall itself.

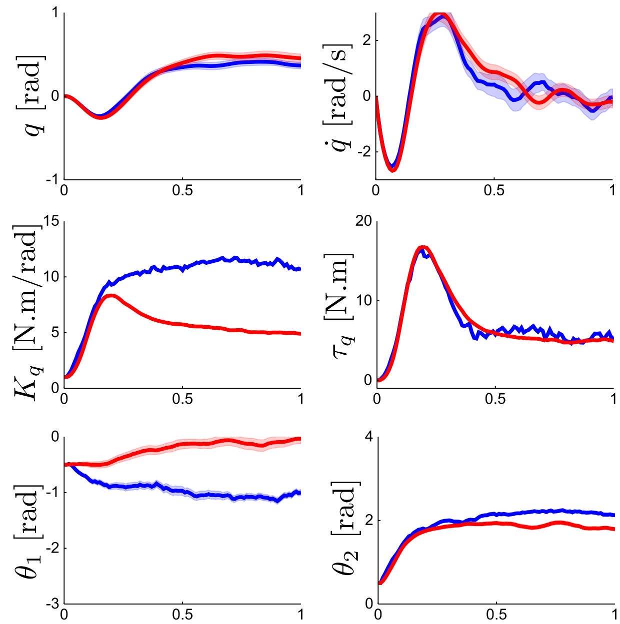

Figure 4. Control of the pnrVSA in an unstable task. 1st row: position and speed of the end-effector. 2nd row: Intrinsic rotational stiffness and feedforward torque applying to the end-effector. 3rd row: Displacements of the internal motors. Red: $R = 0.1 \times \mathrm{Id}_{2\times 2}$, Blue: $R = 0.01 \times \mathrm{Id}_{2\times 2}$ (higher noise entering into the system and cheaper control cost). The mean trajectories across 50 trials are depicted (shaded areas are standard errors). Note that $\tau_q = k_1(\theta_1 - q)^3 + k_2(\theta_2 - q)^3$ and $K_q = 3\big(k_1(\theta_1 - q)^2 + k_2(\theta_2 - q)^2\big)$.

III. Results

A. Open-loop single-joint stabilization in an unstable task

Before describing the simulation results, we should stress that minimizing the control energy makes perfectly sense for pnrVSA because a large stiffness implies a large control energy which turns into passive noise-rejection property: this is not available in other VSA designs. This feature is similar to the large metabolic energy expenditure associated with co-contraction of human muscles, which thus motivates the design of “just stiff enough” control laws. The fact that we stick to open-loop control laws is just to illustrate the properties of pnrVSA for such an extreme case, but this does not prevent the use of feedback laws or model predictive control in real applications. The simulation results presented below can just be improved by using feedback laws.

Note that in the method we used the relationship $C = B\sqrt{\lambda R^{-1}}$ is required, and therefore we have a link between the control cost and the noise magnitude. In Figure 4 we tested different levels of noise affecting the system by changing R. A decrease of R results in an increase of the noise magnitude, but also reduces the control cost, thus reinforcing the relative weight of state-dependent costs. In other words, the more the noise the more co-contraction is allowed at equivalent control-cost. Another way to vary the noise but without affecting the cost is to vary λ directly. Similar behaviors can be observed, the main observation being that the intrinsic stiffness of the system is significantly increased when uncertainty and/or instability increase. This low-dimensional example illustrates the possibility to perform challenging tasks in open-loop, while this would not be possible without the pnrVSA. This also show that the controller can adapt the degree of uncertainty/noise by adjusting the stiffness optimally.

The second task is the case of an 2-dof arm equipped with our pnrVSA pushing against a wall. In this case, we set $R = Id_{4\times 4}$ and $\lambda = 1$, thus fixing the noise covariance matrix via the above-mentioned equation for C. The results of depicted in Figure 5.

The simulations showed that the task could be approximately performed in open-loop with a minimal level of stiffness of the system despite the additive noise acting on the system. Of course, a greater accuracy could be achieved if we used feedback control as well. Nevertheless, our framework provides good reasons to optimally set stiffness of such systems in order to optimally reject noise disturbances and ensure a reasonable behavior even in open-loop.

IV. Conclusion

Recently, a number of novel actuator designs with variable passive properties have been proposed. One of the proposed designs (pnrVIA; [15]) allows for simultaneous control [16] of passive stiffness and passive noise rejection. The specific features of this new actuator principle calls for novel models for movement planning: in this paper we proposed open-loop stochastic optimal control. Given the lack of available algorithms for numerically solving this class of problems, the paper relied on a slight modification of available methodologies. The proposed methodology falls into a class which relies on stochastic sampling of diffusion processes to approximate path integrals. To our knowledge, focus on open-loop control is unique in this kind of literature, mainly because the Hamilton-Jacobi-Bellman equation naturally leads to a feedback solution. The proposed algorithm has been used for open-loop movement and stiffness planning with pnrVSA. Experiments have shown that the idea of open-loop controllers nicely fits with the proposed numerical simulation, where unstable and uncertain tasks were performed by exploiting the nice property of a class of VSA actuators, nominally the ability to change the system passive stiffness and noise rejection by actuators co-activation (and therefore without relying on sensory feedback). The development of such actuators providing intrinsic and immediate feedback via co-contraction stresses the need for developing more specific methods of open-loop stochastic optimal control. This could be done in the spirit of the present work, by making use of an infinite-dimensional Pontryagin Maximum Principle [8] or by using the Stochastic Maximum Principle [19].

References

[1] E. Burdet, R. Osu, D. W. Franklin, T. E. Milner, and M. Kawato, "The central nervous system stabilizes unstable dynamics by learning optimal impedance." *Nature*, vol. 414, no. 6862, pp. 446–449, Nov 2001.

[2] A. Polit and E. Bizzi, "Characteristics of motor programs underlying arm movements in monkeys," *Journal of Neurophysiology*, vol. 42, no. 1, pp. 183–194, Jan. 1979.

[3] A. Del Prete, F. Nori, G. Metta, and L. Natale, "Control of Contact Forces: the Role of Tactile Feedback for Contact Localization," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, 2012.

[4] M. Fumagalli, S. Ivaldi, M. Randazzo, L. Natale, G. Metta, G. Sandini, and F. Nori, "Force feedback exploiting tactile and proximal force/torque sensing. Theory and implementation on the humanoid robot iCub," *Autonomous Robots*, vol. 33, no. 4, pp. 381–398, 2012.

[5] E. Wieser, P. Mittendorfer, and G. Cheng, "Accelerometer based robotic joint orientation estimation," in *Humanoids*, 2011, pp. 67–74.

[6] B. Berret, S. Ivaldi, F. Nori, and G. Sandini, "Stochastic optimal control with variable impedance manipulators in presence of uncertainties and delayed feedback," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2011)*, San Francisco, USA, September 2011, pp. 4354–4359.

[7] R. Schiavi, G. Grioli, S. Sen, and A. Bicchi, "Vsa-ii: A novel prototype of variable stiffness actuator for safe and performing robots interacting with humans," in *Proceedings of the IEEE International Conference on Robotics and Automation, 2008*, 2008.

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

[9] D. Mitrovic, S. Klanke, R. Osu, M. Kawato, and S. Vijayakumar, "A computational model of limb impedance control based on principles of internal model uncertainty," *PLoS ONE*, vol. 5, no. 10, 2010.

[10] J. Buchli, F. Stulp, E. Theodorou, and S. Schaal, "Learning variable impedance control," *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 820–833, Jun. 2011.

[11] H. J. Kappen, "A linear theory for control of non-linear stochastic systems," *Physical Review Letters*, vol. 95, pp. 200–201, 2005.

[12] E. Todorov, "Efficient computation of optimal actions." *Proc Natl Acad Sci U S A*, vol. 106, no. 28, pp. 11 478–11 483, Jul 2009.

[13] E. Theodorou, J. Buchli, and S. Schaal, "A generalized path-integral control approach to reinforcement learning," 2010.

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

[15] L. Fiorio, A. Parmiggiani, B. Berret, G. Sandini, and F. Nori, "pnrvsa: human-like actuator with non-linear springs in agonist-antagonist configuration," in *IEEE-RAS International Conference on Humanoid Robots (HUMANOIDS2012)*, 2012.

[16] F. Nori, B. Berret, L. Fiorio, A. Parmiggiani, and G. Sandini, "Control of a single degree of freedom noise-rejecting variable impedance," in *Proceedings of the 10th international IFAC symposium on Robot Control (SYROCO)*, 2012.

[17] R. Schiavi, G. Grioli, S. Sen, and A. Bicchi, "Vsa-ii: a novel prototype of variable stiffness actuator for safe and performing robots interacting with humans," in *Proc. IEEE Int. Conf. Robotics and Automation ICRA 2008*, 2008, pp. 2171–2176.

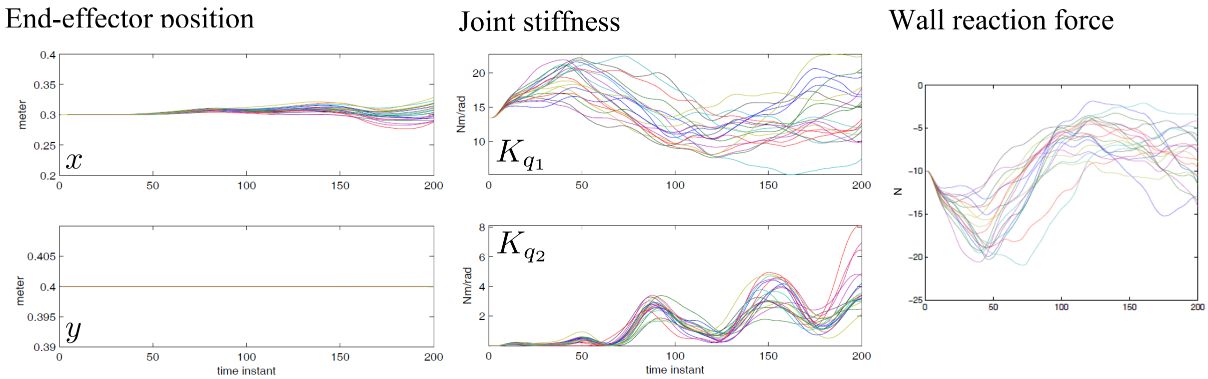

Figure 5. Control of a 2-dof arm equipped with pnrVSA during an unstable task (pushing against a wall). With low stiffness, a diverging behavior was initially observed after 2 seconds. By setting adequately the joint stiffness thanks to co-activation of the antagonist actuators, divergence can be annihilated and the pushing force can be approximately maintained. Note that the goal of keeping a constant pushing force of 10 N is weighted with the goal of preserving the end-effector position at the original location.

[18] R. P. Feynman, *Statistical Mechanics: A Set Of Lectures (Advanced Books Classics)*, 2nd ed. Westview Press, Mar. 1998.
[19] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, 1st ed. Springer, Jun. 1999.