

# Stochastic optimal control with variable impedance manipulators in presence of uncertainties and delayed feedback

B. Berret, S. Ivaldi, F. Nori, G. Sandini

**Abstract**—Muscle co-contraction can be modeled as an active modulation of the passive musculo-skeletal compliance. Within this context, recent findings in human motor control have shown that active compliance modulation is fundamental when planning movements in presence of unpredictability and uncertainties. Along this line of research, this paper investigates the link between active impedance control and unpredictability, with special focus on robotic applications. Different types of actuators are considered and confronted to extreme situations such as moving in an unstable force field and controlling a system with significant delays in the feedback loop. We use tools from stochastic optimal control to illustrate the possibility of optimally planning the intrinsic system stiffness when performing movements in such situations. In the extreme case of total feedback absence, different actuators model are considered and their performance in dealing with unpredictability compared. Finally, an application of the proposed theories on planning reaching movements with the iCub humanoid platform is proposed.

## I. INTRODUCTION

Recently, robotic research has shown a growing interest in studying human motor control to understand how humans are capable of performing motor skills well beyond current robot motor capabilities. In particular, recent findings in human motor control have suggested that co-contraction, or the human ability to change the intrinsic musculo-skeletal compliance, might play a crucial role when dealing with uncertainties and unpredictability. It is therefore reasonable to understand *if and consequently why* we need to replicate this capability in the robots of the next generation since the ability of changing the system mechanical impedance is nowadays not available in most of the robots.

The present paper focuses on systems capable of varying their intrinsic compliance. As previously discussed, the human musculo-skeletal system possesses this property via muscular co-contraction. Recently, a number of robotic actuators capable of actively regulating the overall system compliance have been proposed. Different motivations have been behind the design of these actuators: safety [1], energy storing [2] and force regulation [3] to cite a few. We propose here a rather different point of view suggested by recent results from human motor control which has proposed that variable compliance can be interpreted as a tunable, high bandwidth reaction strategy. In this paper we suggest that a relevant advantage of variable compliance becomes

apparent when considering the intrinsic latencies typical of a distributed control system. Within this context, it is worth observing that both humans and robots share qualitatively similar (but quantitatively different) control issues. Generally speaking, signal transmissions are never instantaneous, neither in robots nor in humans. Latencies typically augment with distances and therefore, as a rule of thumb: (a) local (distributed) controls are typically affected by relatively small latencies; (b) global (centralized) controllers are necessarily affected by significant delays in signal propagation.

The idea proposed in this paper is that co-contraction in humans can be interpreted as a distributed and local control strategy not affected by delays. Our goal will be to show that the ability of actively varying the system compliance can prevent the disadvantages of delayed feedback if coupled with a global and centralized feedforward motor plan which exploits muscle co-contraction to achieve (feedback free) disturbance rejection. Similar characteristics will be simulated in robots equipped with passive variable impedance actuators (VIA) where the ability of actively regulating the stiffness can be seen as an analogous of muscle co-contraction.

### A. Previous works

Previous works within the field of motor control have demonstrated that humans have at least two different control strategies to compensate for external disturbances: disturbance compensation via motor plan adaptation [4] and disturbance rejection via muscle co-contraction [5].

When the *disturbance is deterministic and predictable*, it can be compensated by a proper adjustment of the feedforward motor plan. In particular, F. Mussa-Ivaldi and collaborators [4] have shown that humans are able to compensate for the disturbance due to a velocity dependent force field. Evidence suggests that disturbs are compensated by a motor plan adaptation (i.e. feedforward control adjustment) since a sudden removal of the perturbation produces an incorrect movement execution, the so called *after-effect*. If the disturbance compensation were the result of an increased musculo-skeletal compliance (i.e. an increased muscle co-contraction) then the after-effect should not be noticeable; its presence therefore has been classically interpreted as an evidence for the adaptation of the motor plan to the applied disturbance.

When the *disturbance is stochastic and unpredictable*, it can be compensated by augmenting the musculo-skeletal stiffness via muscle co-contraction. E. Burdet and collaborators [5] have shown that when executing a reaching movement in a divergent force field, the initial task failures grad-

All the authors are with the Robotics, Brain and Cognitive Sciences Department, Italian Institute of Technology, Genoa, Italy (corresponding author phone: +39-010-71781420; e-mail: bastien.berret@iit.it).

The authors were supported by the European Projects: VIATORS (FP7-ICT-2007-3, contract 231554) and ITALK (ICT-214668).

ually disappear as a consequence of an increased stiffness which compensates for the instability caused by the force field. Differently from the disturbance proposed by Mussa-Ivaldi, it is likely to expect that a divergent force field is unpredictable as a consequence of the fact that little (i.e. non-measurable) positioning errors either produce a perturbing force in one direction or in the opposite direction. Changing the motor plan would be inefficient in this context, because it is practically impossible to decide a priori on which side of the divergent field the movement will fall. Even motor plan adaptation strategies relying on a (delayed) feedback will not be useful, because the divergent field intensity was tuned so that typical reaction delays in the central nervous system were not sufficiently fast to react on time before the arm was driven away from the target trajectory. The observed behavior, an increase of the muscle co-contraction to modify the intrinsic musculo-skeletal compliance, was therefore reasonable to expect given the task specifications.

Remarkably, classical neurophysiological experiments have already shown that disturbance rejection does not necessarily pass through proprioceptive feedback. Polit and Bizzi [6] have demonstrated that both intact and deafferented monkeys correctly perform goal directed reaching movements even in presence of unexpected displacement of the arm (prior to movement initialization). These results have been classically interpreted as an evidence for the so called “equilibrium point hypothesis” suggesting that the central nervous system controls both the body equilibrium and the stiffness via proper agonist/antagonist muscles co-contraction. It is worth stressing here that even if deafferentation (i.e. total removal of feedback) is just an extreme situation, we just discussed that there might be tasks [5] where relying on feedback would prevent the task achievement in presence of delays and unpredictability.

The importance of the agonist/antagonist muscle arrangement in dealing with the minimization of uncertainties, has been recently studied by D. Mitrovic et al. [7]. In particular, it was shown that the tool of stochastic optimal control can efficiently simulate the impedance regulation principles observed in humans performing stationary and adaptive tasks. Similarly to the work presented in [7], we will make extensive use of a state of the art optimization tool (ILQG, [8]) to solve the problem of planning movements in presence of uncertainties but extending results to two degrees of freedom models of the human arm. Instead of focusing only on realistic models of human muscles, we will consider also other actuators whose dynamical model covers a number passive variable stiffness actuators recently designed for robotic applications [9], [10]. Moreover, differently from the approach proposed by Mitrovic et al. (see also [11]) we focus on purely feedforward control thus neglecting the possibility of using feedback to correct online the motor plan (see the discussion in Section I-B and Section II).

### B. Summary of the paper

In this paper we are interested in understanding how to deal with delays in presence of unpredictability by exploiting

the ability to change the intrinsic stiffness of a kinematic chain. Unpredictability<sup>1</sup> will be modeled by stochastic noise affecting the system dynamics. In order to emphasize the role of delays, we will often put forward a major working hypothesis by considering motor plans which do not rely on feedback (a sort of infinite delay case). Similarly, we will neglect the use of state estimators typically considered very useful when dealing with uncertainties and delays<sup>2</sup>. As a justification of this major simplification, we will concentrate our analysis on simulating movements where relying on feedback is not an efficient strategy. The general idea is that if delays are sufficiently long and the task sufficiently unstable, feedback information might be available only when the system state has been driven too far from the nominal trajectory and the task execution has become impossible.

The paper is organized as follows. We first provide the necessary materials (Section II) to understand the numerical simulations and experiments presented in Section III. Finally, Section IV summarizes the main implications of the current results.

## II. METHODS

### A. General settings

Throughout this paper, we will consider rigid body dynamics described by the following differential equation [13]:

$$M(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) = \boldsymbol{\tau}_j + J^\top(\mathbf{q})\mathbf{D}(\mathbf{q}), \quad (1)$$

where  $M$ ,  $\mathbf{C}$ ,  $\mathbf{g}$  and  $J$  denote the inertia matrix, the vector of Coriolis/centripetal torques, the gravitational torque vector and the Jacobian matrix, respectively. The vector  $\mathbf{D}$  refers to an external (disturbing) force field possibly applied to the end-effector. The vectors  $\mathbf{q}$  and  $\boldsymbol{\tau}_j$  denote the system generalized coordinates and forces, respectively. Often Eq. 1 will be written in the canonical state space form:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}), \quad (2)$$

where  $\mathbf{x} \in \mathbb{R}^n$  is the state of the system and  $\mathbf{u} \in \mathcal{U} \subset \mathbb{R}^m$  is the control ( $\mathcal{U}$  is convex, typically defined by linear constraints given on  $\mathbf{u}$ ).

### B. Variable impedance control

In the following, we shall consider three types of robotic actuator to control Eq. 2. In robots with rigid joints, the control vector  $\mathbf{u}$  corresponds to  $\boldsymbol{\tau}_j$ , here denoted  $\boldsymbol{\tau}_{ff}$  to stress on its feedforward nature. In this case, the control vector is simply:

$$\text{1st type of actuator : } \mathbf{u} = \boldsymbol{\tau}_{ff}. \quad (3)$$

<sup>1</sup>Within the proposed context, unpredictability is crucial because otherwise delays can be easily compensated by recurring to suitable predictors.

<sup>2</sup>In the field of control system theory, typical solutions for handling delays rely on predictors (e.g. Smith predictors and Kalman filters) which exploit a model of the controlled system to anticipate the effect of delays. Of course predictors can be very beneficial but in the context of this paper the reader should be aware of the fact that a certain level of unpredictability is always present. Mathematically, unpredictability takes the form of the innovation [12], i.e. the stochastic process representing the difference between predictions and measurements. More practically, unpredictability always follows from differences between the system and its model.

The latter actuator does not allow controlling stiffness per se. However, it can be combined with a PD control law to regulate the joint impedance. The second type of actuator we will consider is therefore expressed as a combination of a feedforward torque plus a PD control law. For this kind of actuator, the feedforward torque, the equilibrium configuration and the joint stiffness can be independently controlled. This is similar to the properties of the musculoskeletal system in humans. Hence the joint torque is expressed as:

$$\tau_j = \tau_{ff} - K_P(\mathbf{q} - \mathbf{q}_d) - K_D(\dot{\mathbf{q}} - \dot{\mathbf{q}}_d), \quad (4)$$

where  $K_P$  and  $K_D$  are symmetric positive-definite matrices (respectively the *stiffness* and *damping* gains) and  $\mathbf{q}_d$ ,  $\dot{\mathbf{q}}_d$  define the *desired* equilibrium trajectory. For the sake of simplicity and to make simulations computationally tractable, we will restrict ourselves to the case  $K_D = \alpha K_P$  with constant  $\alpha$ . Moreover, we will assume  $K_P$  diagonal and positive definite (e.g.  $K_P = \text{diag}(k_1, k_2)$  with  $k_1 > 0$  and  $k_2 > 0$ , for a two-link arm). Therefore, the control vector turns out to be:

$$\text{2nd type of actuator : } \mathbf{u}^\top = \left( \tau_{ff}^\top \quad \text{vec}(K_P)^\top \right). \quad (5)$$

Note that this type of actuator captures the main features of an agonist/antagonist configuration such as the one proposed by D. Mitrovic and collaborators for modeling human muscles or the one realized in [14] with electro active polymers.

In robots with flexible joints [1], the joint stiffness can be modulated as well. A standard model for such a kind of actuators is given by:

$$\tau_j = -K(\mathbf{q} - \boldsymbol{\theta}) \quad \text{with} \quad \tau_m = B\ddot{\boldsymbol{\theta}} + K(\boldsymbol{\theta} - \mathbf{q}). \quad (6)$$

In these equations,  $\boldsymbol{\theta}$  and  $\dot{\boldsymbol{\theta}}$  are the internal state of the motors,  $B$  is the matrix of the effective rotor inertias of the motors. The controlled variables are the torques supplied by the motors  $\tau_m$  and the stiffness matrix  $K$  typically diagonal and positive definite (e.g.  $K = \text{diag}(k_1, k_2)$ ). For this type of actuators, the control vector is thus augmented by the stiffness components:

$$\text{3rd type of actuator : } \mathbf{u}^\top = \left( \tau_m^\top, \quad \text{vec}(K)^\top \right). \quad (7)$$

### C. Stochastic optimal control

As previously pointed out we are interested in modeling unpredictability in dynamical systems by means of stochastic differential equations [15] of the following form:

$$d\mathbf{x} = \mathbf{f}(\mathbf{x}, \mathbf{u})dt + F(\mathbf{x}, \mathbf{u})d\mathbf{w}, \quad (8)$$

where  $\mathbf{w} \in \mathbb{R}^p$  is a standard Brownian motion (i.e. the noise). All the quantities are assumed to have compatible dimensions. Functions  $\mathbf{f}$  and  $F$  respectively characterize the nonlinear deterministic dynamics (see Eq. 2) and the state/control-dependent disturbances.

We address the planning problem in the context of stochastic optimal control (SOC) [15], [16]. This mathematical framework is used to derive approximate optimal control

laws driving the system from an initial state  $\mathbf{x}(t = 0)$  to a final state  $\mathbf{x}(t = T)$ . Optimality is defined with respect to expected value  $C$  of a function written as follows:

$$C = \mathbb{E}[g(\mathbf{x}(T)) + \int_0^T l(\mathbf{x}, \mathbf{u}, t)dt],$$

with  $g$  the final cost (typically accuracy constraint in task space) and  $l$  the running cost (typically control energy and/or trajectory tracking constraints).

SOC problems we here consider are non-linear and constrained and, therefore, they do not fall in the well-known Linear Quadratic Gaussian (LQG) framework [17]. We therefore tackled the problem using an iterative-LQG algorithm (ILQG) [8], recently developed to solve such non-linear SOC problems with linear constraints on the control. The solution  $\mathbf{u}^*$  computed by this iterative algorithm has the following form:

$$\mathbf{u}^*(t) = \bar{\mathbf{u}}(t) + L(t)(\mathbf{x}(t) - \bar{\mathbf{x}}(t)). \quad (9)$$

where  $\bar{\mathbf{u}}(t)$  represents the optimal open-loop control policy and where  $L(t)$  is the optimal feedback strategy to correct local deviation of the current state  $\mathbf{x}(t)$  from the optimal state trajectory  $\bar{\mathbf{x}}(t)$ . Clearly the optimal feedback matrix  $L(t)$  has some analogies with the impedance matrix but there are some important differences partly discussed in the Appendix and partly discussed in the next section.

### D. System with time-delays

As motivated in the introduction, we will sometimes consider dynamical systems with temporal latencies in the feedback loop. Introducing time delays in a discrete dynamic system<sup>3</sup>, is quite straightforward following the well known procedure of state augmentation.

Let a discrete approximation of Eq. 8 be described as follows:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta f(\mathbf{x}_k, \mathbf{u}_k) + \sqrt{\Delta} F(\mathbf{x}_k, \mathbf{u}_k) \epsilon_k, \quad (10)$$

where  $\epsilon_k$  is a Gaussian random vector with zero mean and covariance equal to identity. As it is classical in such discrete time formulation, time latencies can be represented with a totally equivalent equation:

$$\tilde{\mathbf{x}}_{k+1} = \tilde{F}\tilde{\mathbf{x}}_k + \tilde{f}(\tilde{\mathbf{x}}_k, \tilde{\mathbf{u}}_k) + \tilde{F}(\tilde{\mathbf{x}}_k, \tilde{\mathbf{u}}_k)\epsilon_k, \quad (11)$$

by simply augmenting the state space with the process history, e.g. with  $h$ -steps delay we have  $\tilde{\mathbf{x}}_k^\top = (\mathbf{x}_k^\top, \dots, \mathbf{x}_{k-h+1}^\top)$  (see the Appendix for details). The ILQG algorithm applied to this augmented system provides an approximately optimal control law of the form:

$$\tilde{\mathbf{u}}_k = \bar{\tilde{\mathbf{u}}}_k + \tilde{L}_k(\tilde{\mathbf{x}}_k - \bar{\tilde{\mathbf{x}}}_k). \quad (12)$$

In practice, delays are present in the feedback as well as in the feedforward loops. Therefore, it is not possible to apply the feedback to the whole process history from the current state. If the feedback is also delayed of  $h$  steps, we may only control the system using:

<sup>3</sup>The ILQG algorithm relies on a discrete time formulation. Typically, time-discretization is achieved by Euler integration, setting  $\Delta = T/N$  and defining  $\mathbf{x}_k = \mathbf{x}(k\Delta)$ , such that  $\mathbf{x}_0 = \mathbf{x}(t = 0)$  and  $\mathbf{x}_N = \mathbf{x}(t = T)$ .

$$\tilde{\mathbf{u}}_k = \bar{\mathbf{u}}_k + \tilde{L}_k^{[h]}(\mathbf{x}_{k-h+1} - \bar{\mathbf{x}}_{k-h+1}). \quad (13)$$

This means it is only possible to use a part of the optimal feedback gain  $\tilde{L}$ , the one corresponding to first available previous state.

### III. RESULTS

In this section, we want to show that unpredictability and significant delays can be efficiently overcome through impedance control. The subsequent numerical simulations will show that exploiting stiffness is a means to accurately control the system where feedback control schemes are likely to fail (because of delays and/or uncertainty about the state/task).

#### A. Two-link arm in a divergent force field

*Numerical settings:* In this example, we consider the three types of actuators defined above for controlling a two-link arm moving in a horizontal plane. The joint angle vector is denoted by  $\mathbf{q} = (q_1, q_2)^\top$ , where the subscript 1 corresponds to the shoulder. A divergent force field is possibly applied to the end-effector (i.e. the vector  $\mathbf{D}$  in Eq. 1). Therefore, two situations are considered:

$$\mathbf{D}(\mathbf{q}) = 0 \text{ and } \mathbf{D}(\mathbf{q}) = (\beta x \ 0)^\top,$$

where  $\beta = 50 \text{ N/m}$  is a constant that determines the strength of the divergent force field and  $x = l_1 \cos q_1 + l_2 \cos(q_1 + q_2)$  is the  $x$ -coordinates of the end-effector in the horizontal plane  $(x, y)$ . The origin of the frame of reference is centred at the shoulder. The numerical parameters we used for the two-link arm can be found in [18]. The  $\beta$  parameter means that a force of 0.5 N is applied to the arm when the end-effector is 1 cm away from the straight line defined by  $x = 0$ . The magnitude of the force field thus scales with the distance from the straight line.

Such a dynamical system can be written in state space form as in Eq. 8, with either  $\mathbf{x} = (q_1, q_2, \dot{q}_1, \dot{q}_2)^\top$  or  $\mathbf{x} = (q_1, q_2, \dot{q}_1, \dot{q}_2, \theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2)^\top$  depending on the type of actuators employed (this latter state vector is for the 3rd type of actuator). A state-dependent noise was then integrated as  $F(\mathbf{x}) = \sigma q_1 I$ . The noise thus reflects on all state components and is proportional to the shoulder velocity. A desired trajectory to be tracked  $\mathbf{q}_d$  was also imposed and derived from a minimum jerk-like trajectory in Cartesian space (i.e. the inverse kinematics of a straight path for the end-effector and a bell-shaped velocity profile).

The expected cost under minimization was precisely:

$$C = \mathbb{E}[g(\mathbf{x}(T)) + \int_0^T \|\mathbf{u}\|_W^2 + w_p \|\mathbf{q} - \mathbf{q}_d\|^2 + w_v \|\dot{\mathbf{q}} - \dot{\mathbf{q}}_d\|^2 dt], \quad (14)$$

with  $g(\mathbf{x}) = w_p \|\mathbf{q} - \mathbf{q}_d(T)\|^2 + w_v \|\dot{\mathbf{q}}\|^2$  and  $\|\mathbf{u}\|_W^2 = \mathbf{u}^\top W \mathbf{u}$ ,  $W$  diagonal positive-definite.

Numerical settings were  $w_{to} = 1$ ,  $w_{st} = 10^{-4}$ ,  $W = \text{diag}(w_{to}, w_{st})$ ,  $w_p = 10^6$ ,  $w_v = 0.1w_p$ ,  $\sigma = 0.05$  and  $\alpha = 0.1$ . A smaller weight was given to the stiffness to allow setting large stiffness for cheap. For the second type

of actuator, we used  $B = \text{diag}(2)[\text{kg.m}^2]$  (a reasonable value considering that this is the *joint-reflected* motor inertia) and the weight given to the torque was decreased to  $w_{to} = 10^{-2}$  to compensate for larger torque commands compared to the other type of actuators. Numerical experiments in presence of noise, with and without the force field were tested and are described in Section III.

*Results:* We replicated in simulation an experiment conducted by E. Burdet and collaborators in humans [5]. It has been observed that a selective use of muscle co-contraction was used by subjects to adapt to a divergent (and thus unpredictable) force field. Indeed, let's consider a forward movement of the end-effector along a straight line. When a force field acts laterally (i.e. leftward or rightward) with a magnitude proportional to the distance of the end-effector from the straight line, the presence of sensorimotor noise and delays makes it impossible to use a feedback control strategy. It was concluded that controlling stiffness is an efficient means to cope with the unpredictability and the delays inherent to the human nervous system. Here we emulate the above experiment using three different actuators (see Section II). Note that we voluntarily restrict our analysis to feedforward control strategies.

Figure 1 depicts the performance of the different actuators in the divergent force field. For the deterministic case (1st column), the presence of a divergent force field is transparent. This is because the desired straight trajectory can be tracked exactly. Therefore, the optimal solution is to use only torque control for the 1st and 2nd types of actuator (no need of gain scheduling). For the 3rd type of actuator, the optimal solution in the deterministic case is a trade-off between torque and stiffness control (not reported here). This is because both control components consume energy, and executing the movement with very small stiffness imply large torques and vice versa.

The role of stiffness is actually highlighted in the stochastic context. The 2nd column of Fig. 1 demonstrates that using only  $\tau_{ff}$  to execute the movement in a divergent force field leads to divergent trajectories. This is because it is basically impossible to track the straight line perfectly in presence of noise. The force field will push to the left or to the right in an unpredictable manner. Since the force increases with the distance from the straight line, the trajectory diverges very quickly. Note that relying on the feedback gain may be possible in simulation, but in practice, the residual uncertainty and the delays would yield divergent trajectories in strong enough force fields.

The third column shows that when using the 2nd type of actuator (with the PD control law), it is possible to counteract the divergent force field and perform the task very accurately. This result is quite obvious but it illustrates that, if the stiffness level is sufficiently high, it is possible to perform the task using the nominal feedforward torque control. This shows that, when using the more human-like kind of actuators, the task can then be performed with high precision and without using feedback at all. If we were using the optimal feedback gain  $L$  without relying on stiffness

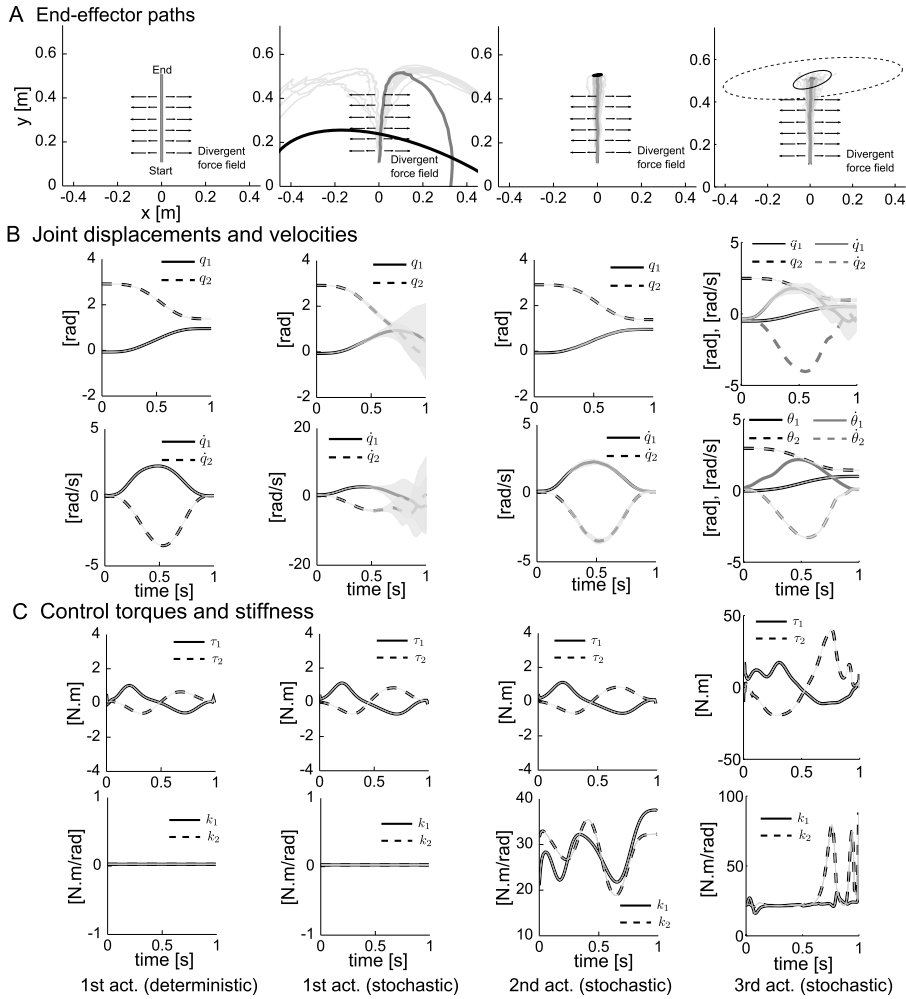


Fig. 1. Two-link arm moving in a divergent force field. The four columns depict the result obtained with the different type of actuators. A. End-effector trajectories. Light gray lines depict trials (20 are depicted). The thicker gray line is the mean path, and a shaded area depicts the standard deviation around the mean when possible. At the end, 95% confidence ellipses are displayed to show the variability of task accuracy. B. Joint displacements and velocities. Mean values and standard deviation (shaded areas) are plotted. C. Feedforward control commands for both the torque and stiffness. Note that, for all trials (i.e. noise instances) the same control is used (no use of feedback). For the 4th column, note the dashed-line ellipse shows the variability we obtain when the inertia matrix  $B$  is reduced by a factor 10.

control, this would have resulted in extensive changes of the feedforward torques to correct disturbances. Assuming that the cost related to the torque or the torque change is higher, it would have resulted in a costly solution, likely more costly than increasing stiffness adequately.

In the fourth column, it is apparent that the 3rd type of actuator yields better results than pure feedforward torque control (2nd column). The end-effector is more robust to perturbations, but the final accuracy is subject to some variability (remind we only use open-loop control). This terminal inaccuracy is related to the properties of the actuator itself. It is clear from Eq. 5 that increasing stiffness implies that a direct transfer of perturbations to the motor torques  $\tau_m$ . Since it is impossible to plan whether the end-effector is going to be pushed leftward or rightward, adjusting the motor torques in open-loop is impossible. Tuning the inertia matrix  $B$  may also improve accuracy as illustrated in the last column in Fig. 1 where it is shown that increasing  $B$  by a

factor of ten results in augmented robustness to perturbations.

### B. Implementation on the iCub robot

In this example we consider a simple example that is a single-joint arm moving in the vertical plane but subject to time delays. Only the 2nd and apparently most efficient actuator is considered.

*Numerical settings:* The system dynamics is a 1-D version of Eq. 1, including the gravitational torque:

$$J\ddot{q} + mgl_c \cos q = \tau_{ff} - K_P(q - q_d) - \alpha K_P(\dot{q} - \dot{q}_d), \quad (15)$$

where,  $J$ ,  $m$ ,  $g$  and  $l_c$  are the moment of inertia, the mass, the gravity acceleration and the length to the center of mass of the arm, respectively. Because of gravity, the dynamics is nonlinear. Time delays were integrated using the discrete time formulation provided above. Parameters compatible with the forearm of the robot iCub are chosen here:  $J = 5.08 \times 10^{-3}$  kg.m<sup>2</sup>,  $m = 0.812$  kg,  $l_c = 0.103$  m and  $l = 0.273$  m (the total forearm length). The noise

magnitude is scaled by  $\sigma = 0.03$  and the damping gain was related to the stiffness gain by  $\alpha = 0.2$ . Here, the 2-dimensional state vector is  $\mathbf{x} = (q, \dot{q})^\top$ . The function  $F$  is again state-dependent and similar to the previous one, that is  $F(\mathbf{x}) = \sigma \dot{q} I$ . The cost function is the same as in Eq. 14 and the same parameters are used. Note that in this case  $\tau_{\text{ff}}$  and  $K_{\text{P}}$  are scalar.

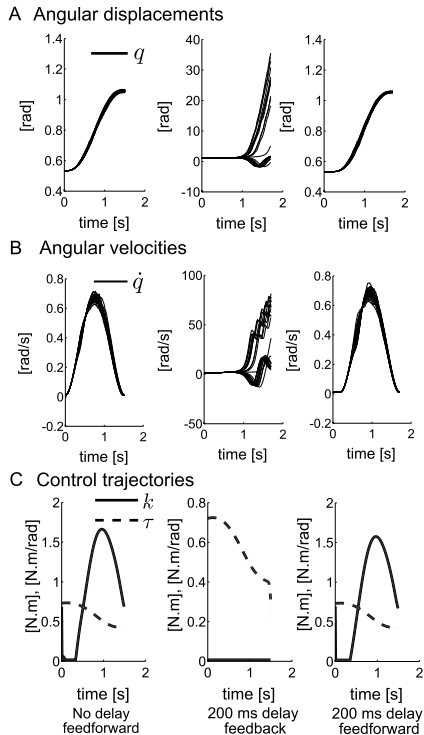


Fig. 2. Single-joint arm subject to time delays. A. Angular displacements for 20 trials. 1st col.: feedforward control strategy without delays. 2nd col.: delayed feedback control strategy without stiffness. 3rd col.: Feedforward control strategy with 200 ms of delay, without using the optimal feedback gain  $L$  but using stiffness control. B. Angular velocities. C. Control trajectories.

*Simulation results:* We performed numerical simulations to test the effectiveness of variable impedance control in presence of time delays, where naive feedback schemes are likely to fail. Here we restrict our analysis to the 2nd type of actuator, given in Eq. 5. Figure 2 (first column) shows that, without delay, a feedforward motor command can be used to perform the task accurately despite the noise perturbing the system, confirming the previous results in the divergent force field. There is no need of using the optimal feedback gain in such a case. The second and third columns reveal that stiffness is actually the crucial feature to perform the task accurately in presence of non-negligible latencies. In fact, it is very inefficient to rely only on a feedback control scheme in presence of large time delays in the feedback loop when there is no stiffness. Controlling accurately the end-effector becomes impossible under our working hypothesis. The third column shows that, however, we do not need the feedback gain to accurately control the plant. A feedforward control of both torques and stiffness is in fact sufficient to overcome the problem of delays and uncertainty. These simulations

demonstrate how variable impedance control can be used to compensate the unpredictability inherent to physical systems, which cannot be handled by means of feedback loops in some cases.

*Experimental results on the iCub:* Experiments on the iCub platform confirmed that the ability of changing the stiffness is beneficial to the execution of movements in presence of modeling errors and delays. iCub is a 53DOF full-body humanoid robot [19], [20], developed by the Italian Institute of Technology within the *RobotCub* consortium, as an open-source platform for research in embodied cognition. Even though the robot is not equipped with passive variable stiffness actuators but only with an active torque control, results from previous sections still apply because the control structure of the robot is composed of two main layers: on one side localized control boards implementing torque control at 1KHz; on the other side a centralized controller running at 50Hz. Local control boards are responsible for feedback controlling the joint stiffness and torque, thus simulating within certain bandwidth limits the ability of regulating the system intrinsic stiffness. The centralized controller on the other side, is responsible for planning the movement by sending joint torques and stiffnesses in a feedforward manner (feedback control is more proficiently performed at the control board level). In a sense, the implementation of theoretical results presented in the previous sections, allows to plan the feedforward stiffness and torque commands to be sent by the centralized controller to the local control boards in order to optimally compensate for disturbance exploiting only the efficient local feedback as opposed to exploiting the delayed centralized feedback.

In order to simplify the planning problem, we followed the problem formulation of the current subsection and considered a single degree of freedom movement, nominally the elbow joint of the right arm moving upward in the vertical plane. Exploiting the joint impedance control interface available in the iCub [21], it was possible to reproduce exactly the control action described in Eq. 15, with arbitrary control on both  $\tau_{\text{ff}}$  and  $K_{\text{P}}$ , realized by the localized control boards.

The desired movement was set from  $q_0^* = -60$  to  $q_T^* = -30$  degrees, which were remapped in the iCub elbow joint range. The movement duration  $T$  was set to 1.5s, while the control rate was set to 10ms. System dynamics, modeled by Eq. 15, have been modified with a state-dependent noise  $e$ :

$$e = (a + b\dot{q})\eta \quad (16)$$

where  $\eta$  is a normally distributed stochastic variable ( $\eta \sim N(0, 1)$ ). In order to set reasonable values of the parameters  $a$  and  $b$ , the modeling error has been identified. A sequence of movements was performed with the iCub elbow, varying all the control variables  $\tau_{\text{ff}}$  and  $K_{\text{P}}$ . The associated errors  $e$  were then computed as follows:

$$e = -J\ddot{q} - mgl_c \cos q + \tau_{\text{ff}} - K_{\text{P}}(q - q_d) - \alpha K_{\text{P}}(\dot{q} - \dot{q}_d),$$

and state-dependent characteristics evaluated (see Fig. 3(a) and associated caption). Admissible values for  $a$  and  $b$  were then obtained by running a normality test on  $\eta = \frac{a+b\dot{q}}{e}$ .

Precisely, the Jarque-Bera test [22] available in Matlab, was used. The best set of parameters found was  $(a^\circ, b^\circ) = (0.15, 0.37)$ . Similarly, the dependence between  $K_P$  and  $K_D$  was experimentally identified on the real robot by defining a suitable linear stability region in the  $K_P$ - $K_D$  space (Fig. 3(b)).

After this identification procedure, the associated stochastic optimal control problem was solved in order to compute the optimal control strategy relying on pure feedforward. Associated nominal optimal control and trajectory are visualized with a red line in Fig. 4(a) and Fig. 4(b), respectively. Tests on the real platform resulted in slightly different trajectories (blue line in Fig. 4); differences are motivated by slight imprecision in the desired torque tracking. However the experimentally evaluated expected cost ( $C = 0.41$ ) results quite close to the one predicted by the simulation ( $C^* = 0.29$ ).

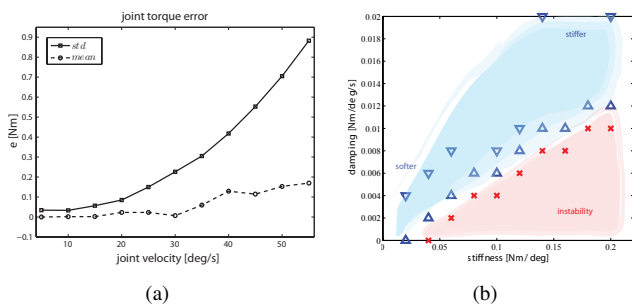


Fig. 3. Left picture: mean  $\bar{e}$  and standard deviation  $\sigma_e$  of the joint torque error from the elbow joint of iCub's right arm. Standard deviation is evidently bilinear with the increase of velocity, while the error mean is linearly proportional to the velocity. Polynomial fitting models are  $\sigma_e = 1.254\dot{q}^2 - 0.351\dot{q} + 0.057$  and  $\bar{e} = 0.213\dot{q} - 0.049$  respectively. Right picture: experimental evaluation of the stiffness-damping relationship. The lower region highlighted in red causes unstable behaviors of the arm when reaching its equilibrium point after an external force is applied. Boundary values were detected by overshoot of a step response. The light blue region shows feasible values, which make the joint behave like a softer or stiffer spring. The boundary region between the two can be ideally described by the following relationship:  $K_D = 0.075K_P - 0.0015$ .

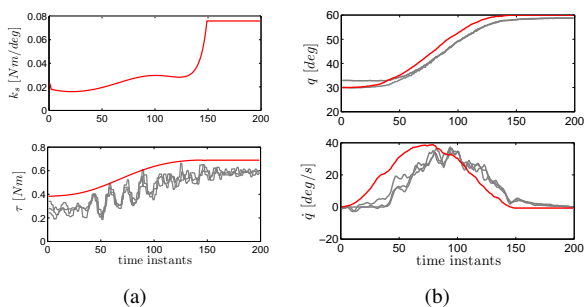


Fig. 4. Plots of optimal control policy (left column) and corresponding system trajectories (right column). In red, the optimal nominal control policy; in blue, the trajectories on the real system. Differences between nominal and real trajectories follow from differences between commanded and executed torques (left bottom panel) due to imperfect torque tracking.

#### IV. CONCLUSIONS AND FUTURE WORKS

In this paper we used numerical methods for optimal movement planning with variable impedance manipulators in presence of uncertainties. In order to evaluate the importance of variable compliance actuation, we considered extreme examples of movement planning in presence of unpredictable disturbances and delayed feedback. The underlying idea was to prove that the ability to regulate the system intrinsic stiffness allows to rely less on feedback control strategies, typically affected by large delays and noise. Simulations were performed in a two degrees of freedom planar reaching task. Preliminary tests on a real robotic platform were performed on the iCub, exploiting its two layer control structure which actively simulates passive variable stiffness within certain bandwidth limits.

In the view of designing novel actuators for robotic application, future works will concentrate on understanding the role of poly-articular muscles in regulating cross joint stiffness (off diagonal terms in the matrix  $K_P$ ), a control variable which was neglected in our current formulation since this possibility is nowadays not implementable in available actuator designs.

#### APPENDIX

##### ON THE ROLE OF OPTIMAL GAIN $L$ AND STIFFNESS $K$

The locally-optimal control law used as an input for the real system is eventually given by:

$$\mathbf{u}^*(t) = \bar{\mathbf{u}}(t) + L(t)(\mathbf{x}(t) - \bar{\mathbf{x}}(t)). \quad (17)$$

It has to be noted that the feedforward control law  $\bar{\mathbf{u}}$  is optimized for the stochastic system, and therefore, may differ from the one obtained using deterministic optimal control. The ILQG method is “blind” to additive noise but control-dependent or state-dependent noise is usually present in physical systems, which allows the algorithm adjusting the optimal control law with respect to the noise statistics (the noise covariance depends on  $F$ ). The optimal feedback gain  $L$  is also necessary to control the system optimally as soon as the actual state differs from the nominal trajectory  $\bar{\mathbf{x}}$ . This is because the cost-to-go function actually depends on the real state  $\mathbf{x}$ . Sub-optimal but accurate enough movements can nevertheless be obtained with a correct gain scheduling (specification of  $K_P$ ) as it can be seen in the Section III. In this paper, we restrict ourselves to systems with fully observable states, but this is not a restriction of the method [23]. Unless stated differently, we have no recourse to the optimal feedback gain  $L$  in the simulations, for the reasons mentioned in the Introduction.

At this point, it should be noted that, for the 2nd type of actuator (Eq. 5), the feedback control term  $L(t)(\mathbf{x}(t) - \bar{\mathbf{x}}(t))$  and the stiffness control term  $K_P(\mathbf{q} - \mathbf{q}_d) - K_D(\dot{\mathbf{q}} - \dot{\mathbf{q}}_d)$  may have relatively similar effects on the mechanical plant. First of all, it can be observed that, for the 1st type of actuator, using the feedback gain is a way to modulate the stiffness through feedback control (using  $L$ ). For the 2nd type of actuator, the total joint torque  $\tau_j$  applied to the joints using the

ILQG controller  $\mathbf{u}^*$  is given by the equation (neglecting the feedback terms correcting  $K_P$ , which involves “quadratic” errors):

$$\begin{aligned} \tau_j = & \tau_{ff} + L^{[q]}(\mathbf{q} - \bar{\mathbf{q}}) + L^{[\dot{q}]}(\dot{\mathbf{q}} - \dot{\bar{\mathbf{q}}}) \\ & - K_P(\mathbf{q} - \mathbf{q}_d) - K_D(\dot{\mathbf{q}} - \dot{\mathbf{q}}_d). \end{aligned} \quad (18)$$

In the latter equation,  $L^{[q]}$  and  $L^{[\dot{q}]}$  refer to the parts of  $L$  scaling with  $(\mathbf{q} - \bar{\mathbf{q}})$  and  $(\dot{\mathbf{q}} - \dot{\bar{\mathbf{q}}})$ , respectively. They correspond to the feedback law correcting the feedforward torque  $\tau_{ff}$ . It is thus evident that the absence of stiffness in the physical plant might be compensated by this “virtual stiffness” specified through the optimal feedback gain.

Nevertheless, some differences exist between  $L$  and  $K_P$ :

- The trajectories  $\mathbf{x}_d$  and  $\bar{\mathbf{x}}$  are generally different, even slightly. The first one is a desired trajectory that is chosen by the experimenter or that could be chosen using an automatic process. The second one is the optimized trajectory depending on the cost function being minimized.

#### TIME DELAYED DYNAMICS

Here we now consider that there is a time latency equal to  $h$  steps in the discrete dynamics described in Eq. 10, so that we have:

$$\begin{aligned} \mathbf{x}_{k+1} = & \mathbf{x}_k + \Delta f(\mathbf{x}_k, \mathbf{u}_{k-h+1}) \\ & + \sqrt{\Delta} F(\mathbf{x}_k, \mathbf{u}_{k-h+1}) \epsilon_k, \end{aligned} \quad (19)$$

with  $\mathbf{x}_0 = \mathbf{x}_1 = \dots = \mathbf{x}_{h-1}$  and  $k \geq h-1$ . As it is classical in such a setting, the practical strategy is to augment the state vector such that it includes the history of the process over  $h$  steps:

$$\tilde{\mathbf{x}}_k^\top = \left( \mathbf{x}_k^\top \quad \mathbf{x}_{k-1}^\top \quad \dots \quad \mathbf{x}_{k-h+1}^\top \right).$$

Note that we have  $\tilde{\mathbf{x}}_k \in \mathbb{R}^{nh}$ . The augmented dynamics can now be expressed as in Eq. 11 with:

$$\begin{aligned} \tilde{I} = & \begin{pmatrix} I & \dots & 0 & 0 \\ I & 0 & \dots & \\ & \ddots & \ddots & \\ 0 & \dots & I & 0 \end{pmatrix}, \\ \tilde{f}^\top = & \left( \Delta f(\tilde{\mathbf{x}}_k^{[h]}, \mathbf{u}_{k-h+1})^\top \quad \dots \quad 0 \quad 0 \right), \\ \tilde{F}^\top = & \left( \sqrt{\Delta} F(\tilde{\mathbf{x}}_k^{[h]}, \mathbf{u}_{k-h+1})^\top \quad \dots \quad 0 \quad 0 \right). \end{aligned}$$

Above,  $I$  denotes the identity matrix of dimension  $n$ . We may rewrite  $\mathbf{u}_{k-h+1}$  as  $\tilde{\mathbf{u}}_k$ .

#### REFERENCES

[1] R. Alami, A. Albu-Schaeffer, A. Bicchi, R. Bischoff, R. Chatila, A. D. Luca, A. D. Santis, G. Giralt, J. Guiochet, G. Hirzinger, F. Ingrand, V. Lippiello, R. Mattone, D. Powell, S. Sen, B. Siciliano, G. Tonietti, and L. Villani, “Safe and dependable physical human-robot interaction in anthropic domains: State of the art and challenges,” in *Proc. IROS’06 Workshop on pHRI - Physical Human-Robot Interaction in Anthropic Domains*, A. Bicchi and A. D. Luca, Eds. IEEE, 2006.

[2] S. K. Au, J. Weber, and H. Herr, “Powered ankle-foot prosthesis improves walking metabolic economy,” *Transaction on Robotics*, vol. 25, pp. 51–66, February 2009. [Online]. Available: <http://portal.acm.org/citation.cfm?id=1653071.1653076>

[3] G. Pratt and M. Williamson, “Series elastic actuators,” in *1995 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots*, vol. 1. Los Alamitos, CA, USA: IEEE Comput. Soc. Press, 1995, pp. 399–406.

[4] R. Shadmehr and A. Mussa-ivaldi, “Adaptive representation of dynamics during learning of a motor task,” *Journal of Neuroscience*, vol. 14, pp. 3208–3224, 1994.

[5] 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. [Online]. Available: <http://dx.doi.org/10.1038/35106566>

[6] 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.

[7] 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.

[8] E. Todorov and W. Li, “A generalized iterative lqg method for locally-optimal feedback control of constrained nonlinear stochastic systems,” in *American Control Conference, 2005. Proceedings of the 2005*, jun. 2005, pp. 300 – 306 vol. 1.

[9] S. Wolf and G. Hirzinger, “A new variable stiffness design: Matching requirements of the next robot generation,” in *ICRA*, 2008, pp. 1741–1746.

[10] A. Jafari, N.G.Tsagarakis, B. Vanderborght, and D. Caldwell, “A novel actuator with adjustable stiffness (AwAS),” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010)*, 2010, taipei, Taiwan.

[11] D. Mitrovic, S. Nagashima, S. Klanke, T. Matsubara, and S. Vijayakumar, “Optimal feedback control for anthropomorphic manipulators,” in *ICRA*, 2010, pp. 4143–4150.

[12] Y. Bar-Shalom, T. Kirubarajan, and X.-R. Li, *Estimation with Applications to Tracking and Navigation*. New York, NY, USA: John Wiley & Sons, Inc., 2002.

[13] R. M. Murray, S. S. Sastry, and L. Zexiang, *A Mathematical Introduction to Robotic Manipulation*, 1st ed. Boca Raton, FL, USA: CRC Press, Inc., 1994.

[14] M. Randazzo, M. Fumagalli, F. Nori, G. Metta, and G. Sandini, “Force control of a tendon driven joint actuated by dielectric elastomers,” in *Proceedings of the 12th International Conference on New Actuators (Actuator 2010)*, 2010.

[15] B. Oksendal, *Stochastic Differential Equations*, 4th ed. Springer Berlin, 1995.

[16] E. Todorov, *Optimal control theory*. Bayesian Brain: Probabilistic Approaches to Neural Coding, Doya K (ed), 2006, ch. 12, pp. 269–298.

[17] M. Athans, “The role and use of the stochastic linear-quadratic-gaussian problem in control system design,” *IEEE transactions on automatic control*, vol. 16, no. 6, pp. 529–552, 1971.

[18] E. Todorov and W. Li, “Optimal control methods suitable for biomechanical systems,” in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Biology and Medicine Society, IEEE/EMB*, 2003.

[19] G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori, “The iCub humanoid robot: an open platform for research in embodied cognition,” in *PerMIS: Performance Metrics for Intelligent Systems Workshop*, Washington DC, USA, 2008.

[20] G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. von Hofsten, K. Rosander, J. Santos-Victor, A. Bernardino, and L. Montesano, “The icub humanoid robot: An open-systems platform for research in cognitive development,” *Neural Networks, special issue on Social Cognition: From Babies to Robots*, vol. 23, pp. 1125–1134, 2010.

[21] S. Ivaldi, M. Fumagalli, and U. Pattacini, “Doxygen documentation of the iDyn library,” [www, http://eris.liralab.it/iCub/main/dox/html/group\\_\\_iDyn.html](http://eris.liralab.it/iCub/main/dox/html/group__iDyn.html).

[22] C. Jarque and A. Bera, “A test for normality of observations and regression residuals,” *International Statistical Review*, vol. 55, no. 2, pp. 163–172, 1987.

[23] W. Li and E. Todorov, “Iterative linearization methods for approximately optimal control and estimation of non-linear stochastic system,” *International Journal of Control*, vol. 80, no. 9, pp. 1439–1453, 2007.