

Exploiting Sensorimotor Stochasticity for Learning Control of Variable Impedance Actuators

Djordje Mitrovic, Stefan Klanke, Matthew Howard, Sethu Vijayakumar

Abstract—Novel anthropomorphic robotic systems increasingly employ variable impedance actuation in order to achieve robustness to uncertainty, superior agility and efficiency that are hallmarks of biological systems. Controlling and modulating impedance profiles such that it is optimally tuned to the controlled plant is crucial to realise these benefits. In this work, we propose a methodology to generate optimal control commands for variable impedance actuators under a prescribed trade-off of task accuracy and energy cost. In contrast to classical optimal control methods that typically require an accurate analytical plant dynamics model, we employ a supervised learning paradigm to acquire *both* the process dynamics as well as the stochastic properties. This enables us to prescribe an optimal impedance and command profile (i) tuned to the hard-to-model stochastic characteristics of a plant and (ii) adapt to the systematic changes such as a change in load.

I. INTRODUCTION

Humans have remarkable abilities in controlling their limbs in a fashion that outperforms most artificial systems in terms of versatility, compliance and energy efficiency. The fact that biological motor systems suffer from significant noise, sensory delays and other sources of stochasticity [3] makes its performance even more impressive. Therefore, it comes as no surprise that biological motor control is often used as a benchmark for robotic systems. Biological motor control characteristics, on the one hand, are a result of the inherent biophysical properties of human limbs and on the other hand, are achieved through a framework of learning and adaptation processes [18], [7].

In this paper, we focus on issues related to adaptive motor control of antagonistically actuated robots. Antagonistic actuator designs are based on the biological principle of opposing muscle pairs. Therefore, the joint torque motors, for example, of a robotic arm are replaced by opposing actuators, typically using mechanical springs [10]. Such *series elastic actuators (SEA)* have found increasing attention over the last decades [16] as they provide certain beneficial properties over classic joint torque actuated systems.

Through the use of antagonistic actuation, the system is able to vary co-contraction levels which in turn change the system's mechanical properties – this is commonly referred to as *impedance control* [5]. The impedance in a mechanical system is defined as a measure of force response to a movement exerted on the system and is made of constituent components such as *inertia*, *damping*, and *stiffness*. This additional degree of freedom in the limb dynamics, i.e., the same joint torque can be achieved by different muscle activations, can be used beneficially in many motion tasks, especially those involving interaction with tools or manipulation. It has been shown through many studies (e.g. [2])

that humans are capable of modulating this impedance in an optimal way with respect to the task demands, trading off selectively against energy consumption. For example, when you use a drilling machine to drill holes into a wall, you will *learn* to co-contract your muscles such that the random perturbations of the drilling has minimal impact on your task. In general, impedance modulation is an efficient way to control systems that suffer from noise, disturbances or sensorimotor delays.

On the down-side, the introduction of antagonistic actuation imposes higher demands on the redundancy resolution capabilities of a motor controller. Optimality principles have successfully been used in biological [12] and in artificial systems [9] as a principled strategy to resolve redundancies in a way that is beneficial for the task at hand. More specifically, *stochastic optimal control (SOC)* [1], [11], [13] appears to be an especially appealing theory as it studies optimality principles under the premise of noisy and uncertain dynamics. Another important aspect when studying stochastic systems is how the information, for example, about noise or uncertainty is obtained without prior knowledge. Supervised learning methods can provide a viable solution to this problem as they can be used to extract information from the plant's sensorimotor data directly.

Here, we propose a control strategy for antagonistic systems which is based on stochastic optimal control theory under the premise of a minimal energy cost. We propose to extend SOC by learning the dynamics model of the plant, which enables us (i) to adapt to systematic changes of the plant and (ii) extract its stochastic properties. By incorporating this stochastic information into the optimisation process, we show for the first time, how impedance modulation and co-contraction behaviour emerges as an optimal control strategy from first principles.

II. A NOVEL ANTAGONISTIC ACTUATOR DESIGN FOR IMPEDANCE CONTROL

To study impedance control, we developed an antagonistic joint with a simple mechanical setup. Our design is based on the SEA approach in which the driven joint is connected via spring(s) to a stiff actuator (e.g., a servo-motor). A variety of SEA designs have been proposed in the past (for a recent review see [16]). *Antagonistic SEA* have one motor per opposing spring and the stiffness is controlled through a combination of both motor commands which means that the relationship between motor commands and stiffness must be resolved by the controller. This additional computational cost is the trade-off for a biologically plausible architecture.

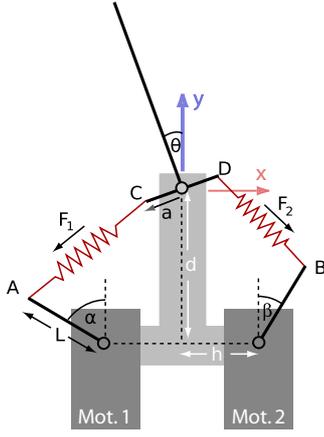


Fig. 1. Schematic of the variable stiffness actuator. The robot dimension are: $a = 15\text{mm}$, $L = 26\text{mm}$, $d = 81\text{mm}$, $h = 27\text{mm}$.

In this paper, when we refer to *impedance control*, we will solely address a *change in stiffness* and ignore variable damping or variable inertia.

For antagonistic SEA, nonlinearity of the springs is essential to obtain a variable impedance [15]. Because forces produced through springs with linear tension to force characteristics tend to cancel out in an antagonistic setup, an increase in the tension of both springs (i.e., co-contraction) does not change the stiffness of the system. Commercially available springs usually have linear tension to force characteristics and consequently most antagonistic SEA require relatively *complex mechanical structures* to achieve a nonlinear tension to force curve [8], [14]. These mechanisms typically increase construction and maintenance effort but also can complicate the system identification and controllability, for example, due to added drag and friction properties. We directly addressed this aspect in our design of the SEA, which primarily aims to achieve variable stiffness characteristics using a simple mechanical setup.

A. Variable stiffness with linear springs

Here we propose a SEA design which does not rely on complex mechanisms to achieve variable stiffness but achieves the desired properties through a specific geometric arrangement of the springs. While the emphasis of this paper is not on the mechanical design of actuators, we will explain the salient dynamic properties of our testbed. Fig. 1 shows a sketch of the robot, which is mounted horizontally and consists of a single joint and two antagonistic servomotors that are connected to the joint via *linear springs*. The springs are mounted offset with a moment arm a at the joints and an offset of L at the motors. Therefore, the spring-endpoints move along a circular paths at the joints and at the motors. Under the assumption that the servo motors are infinitely stiff, we can calculate the torque τ acting on the arm as follows. Let \mathbf{s}_1 denote the vector from point C to A, and \mathbf{s}_2 the vector from D to B, and s_1 and s_2 their respective length. Putting the origin of the coordinate system at the arm joint,

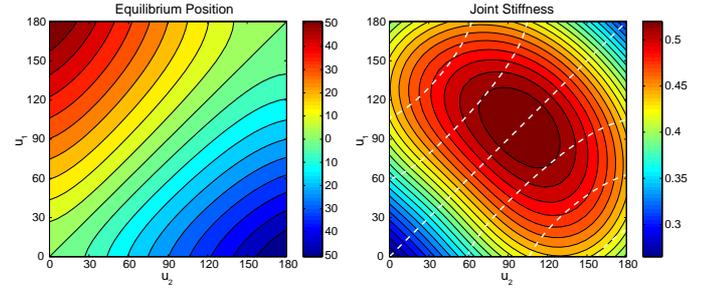


Fig. 2. Left: Equilibrium position as a function of the motor positions (in degrees), with contour lines spaced at 5 degree intervals. Right: Stiffness profile of the arm, as calculated from (4).

we have

$$\begin{aligned} \mathbf{s}_1 &= \begin{pmatrix} -h - L \sin \alpha \\ -d + L \cos \alpha \\ 0 \end{pmatrix} - \underbrace{\begin{pmatrix} -a \cos \theta \\ -a \sin \theta \\ 0 \end{pmatrix}}_{=\mathbf{a}_1}, \quad (1) \\ \mathbf{s}_2 &= \begin{pmatrix} h + L \sin \beta \\ -d + L \cos \beta \\ 0 \end{pmatrix} - \underbrace{\begin{pmatrix} a \cos \theta \\ a \sin \theta \\ 0 \end{pmatrix}}_{=\mathbf{a}_2}. \end{aligned}$$

Denoting the spring constant by κ and the rest length by s_0 , this yields forces

$$\mathbf{F}_1 = \kappa(s_1 - s_0) \frac{\mathbf{s}_1}{s_1} \quad \text{and} \quad \mathbf{F}_2 = \kappa(s_2 - s_0) \frac{\mathbf{s}_2}{s_2}. \quad (2)$$

Given the motor positions α and β and the arm position θ , the torque generated by the springs is

$$\tau(\alpha, \beta, \theta) = \hat{\mathbf{z}}^T (\mathbf{F}_1 \times \mathbf{a}_1 + \mathbf{F}_2 \times \mathbf{a}_2). \quad (3)$$

To calculate the equilibrium position θ_{eq} for given motor positions α and β , we need to solve $\tau(\alpha, \beta, \theta_{eq}) = 0$, which in practice we do by numerical optimisation. At this position, we can calculate the joint stiffness as

$$K(\alpha, \beta) = \left. \frac{\partial}{\partial \theta} \tau(\alpha, \beta, \theta) \right|_{\theta=\theta_{eq}}. \quad (4)$$

Please note that K depends linearly on the spring stiffness κ , but that the geometry of the arm induces a nonlinear dependency on α and β . Fig. 2 shows the analytically computed profiles of the equilibrium position and stiffness, respectively.

Further denoting the arm's inertia around the z -axis by I_z and a damping torque given by $\tau(\dot{\theta}) = -D\dot{\theta}$, the dynamics equation can be analytically written as:

$$I_z \ddot{\theta} = \tau(\alpha, \beta, \theta) - D\dot{\theta}. \quad (5)$$

B. Actuator hardware

Fig. 3 depicts our prototype SEA hardware implementation of the discussed design. For actuation, we employ two servo motors (Hitec HSR-5990TG), each of which is connected to the arm via a spring mounted on two low friction ball bearings. To avoid excessive oscillations, the joint is attached to a rotary viscous damper. The servos are controlled

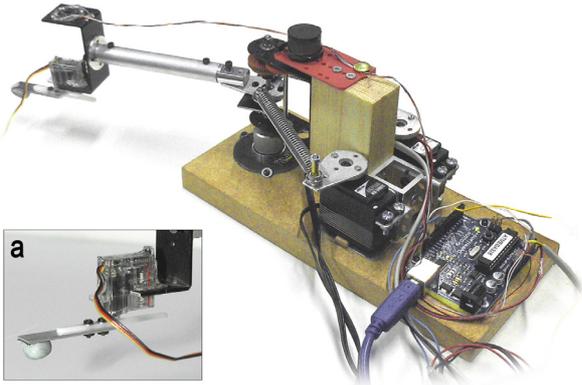


Fig. 3. Photograph of our antagonistic robot. Inset panel (a): Separate servo motor mounted at the end of the arm to create stochastic perturbations (see Section IV-B).

using 50 Hz PWM signals by an Arduino Duemilanove microcontroller board (Atmel ATmega328). That board also measures the arm’s joint angle θ with a contact-free rotary position encoder (Melexis MLX90316GO), as well as its angular acceleration $\ddot{\theta}$ using a LilyPad accelerometer (Analog Devices ADXL330). Finally, we also measure the servo motor positions by feeding a signal from their internal potentiometer to the AD converters of the Arduino. While the operating frequency is limited to 50 Hz due to the PWM control, all measurements are taken at a 4x higher frequency and averaged on the board to reduce the amount of noise, before sending the results to a PC via a serial connection (RS232/USB).

III. STOCHASTIC OPTIMAL CONTROL

In many control scenarios it is desirable to be able to perform in the “best way possible”. For example, one may wish to move the system to a desired posture and consume as little energy as possible during the movement. This type of problem is studied in *stochastic optimal control (SOC)* theory [11], the central ingredient of which is the minimisation of an optimality criterion

$$J(\mathbf{u}_{0..T}) = \int_0^T c(\mathbf{x}(t), \mathbf{u}(t), t)dt + h(\mathbf{x}(T)) \quad \text{or} \quad (6)$$

$$J(\mathbf{u}_{0..T}) = \int_0^\infty c(\mathbf{x}(t), \mathbf{u}(t), t)dt,$$

for a task with a finite or infinite horizon. Apart from the optional final cost $h(\cdot)$, the criterion integrates a cost rate $c(\mathbf{x}, \mathbf{u})$ over the course of the movement. That cost may depend on both the system’s state \mathbf{x} and control commands \mathbf{u} , where the initial state of the system is given as $\mathbf{x}(0)$, and $\mathbf{x}(t)$ evolves depending on the commands $\mathbf{u}(t)$. The movement of the system in the most general form is governed by

$$d\mathbf{x} = \mathbf{f}(\mathbf{x}, \mathbf{u})dt + \mathbf{F}(\mathbf{x}, \mathbf{u})d\xi \quad , \quad \xi \sim \mathcal{N}(0, \mathbf{I}). \quad (7)$$

This formulation, besides the deterministic process dynamics $\mathbf{f}(\mathbf{x}, \mathbf{u})$ incorporates stochasticity as a Gaussian noise process

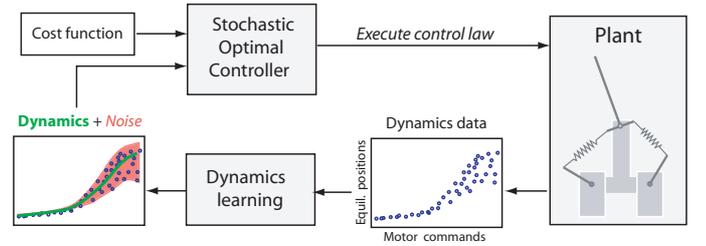


Fig. 4. Schematic diagram of our proposed combination of stochastic optimal control (SOC) and learning. The dynamics model used in SOC is acquired and constantly updated with data from the plant. The learning algorithm extracts the dynamics as well as stochastic information contained (noise model from confidence intervals). SOC takes into account both measures in the optimisation.

$d\xi$, scaled by $\mathbf{F}(\cdot)$, which tells us how strongly the noise affects which parts of the state and control space.

Solving such an SOC problem then corresponds to minimising the *expected cost*¹ for the given process dynamics. If the plant dynamics is linear and the cost function is quadratic the optimisation problem is convex and can be solved analytically². However for high dimensional and non-linear problems, the solutions are much harder to obtain and typically approximative methods must be employed (e.g., [6]).

In order to simplify the presentation as much as possible, in this work we control the functional relation from motor positions $\mathbf{u} = (u_1, u_2)^T = (\alpha, \beta)^T$ to joint angle θ , ignoring velocities and accelerations. During stationary conditions and in the absence of perturbations, this mapping reflects the equilibrium position of the arm (Fig. 2, left). In correspondence to the general dynamics equation (7), the state $\mathbf{x} = \theta_{eq}$ represents the current equilibrium position, \mathbf{u} the applied motor action, and $d\mathbf{x}$ the resulting change in equilibrium position. Therefore the reduced dynamics used here, only depends on the control signals, i.e.,

$$dx = f(\mathbf{u})dt + F(\mathbf{u})d\xi \quad , \quad \xi \sim \mathcal{N}(0, 1). \quad (8)$$

A. Modelling dynamics and noise through learning

Analytical dynamics formulations as described in Section II-A or in eq. (8) have the tremendous advantage of being compact and quick to evaluate numerically, but they also suffer from drawbacks. First, their accuracy is limited to the level of detail put into the physical model. For example, our model is based on the assumption that the robot is completely symmetric, that both motors are perfectly calibrated, and that the two springs are identical, but in reality we cannot avoid small errors in all of these. Second, the analytical model does not provide obvious ways to model changes in the dynamics, such as from wear and tear, or more *systematic* changes due to the weight of an added tool.

While these problems can to some extent be alleviated by a more involved and repeated system identification process,

¹This means we put expectation brackets around the integrals and $h(\cdot)$ in (6).

²So called LQ-problem.

the situation is worse if we look at the noise model $F(\cdot)$, or at *stochastic* changes to the dynamics. For example, an arm might be randomly perturbed by tool interactions such as when drilling into a wall, with stronger effects for certain postures, and milder effects for others. How should we model this “noise landscape” analytically?

We therefore propose to include a supervised learning component and to acquire both the dynamics and the noise model in a data-driven fashion (Fig. 4). Our learning method of choice in this paper is LWPR, or *Locally Weighted Projection Regression* [17], because that algorithm allows us to adapt the models incrementally and online, and it is able to reflect heteroscedastic noise in the training data through localised confidence intervals around its predictions.

Learning the deterministic mapping \tilde{f} from sensorimotor data, we can account for asymmetries right away. More interestingly, when we collect data from the perturbed system, we can acquire a model of the arm’s kinematic variability as a function of the motor positions. Therefore the learned dynamics can be used in (slow) position control tasks as described in the next section.

B. Energy optimal (equilibrium) position control

Let the task be to hold the arm at a certain position $\hat{\theta}$, while consuming as little energy as possible. Let us further assume that we have no feedback from the system³, but that the arm is perturbed randomly. We can state this mathematically as the minimisation of a cost

$$J = \langle w_p (f(\mathbf{u}) - \hat{\theta})^2 + |\mathbf{u}|^2 \rangle, \quad (9)$$

where w_p is a factor that weights the importance of being at the right position against the energy consumption which for simplicity we model by $|\mathbf{u}|^2$. Taking into account that the motor commands \mathbf{u} are deterministic, and decomposing the expected position error into an error of the mean plus the variance, we can write the expected cost J as

$$J = w_p (\langle f(\mathbf{u}) \rangle - \hat{\theta})^2 + w_p \langle (f(\mathbf{u}) - \langle f(\mathbf{u}) \rangle)^2 \rangle + |\mathbf{u}|^2, \quad (10)$$

which based on the LWPR learned model becomes

$$J = w_p (\tilde{f}(\mathbf{u}) - \hat{\theta})^2 + w_p \sigma^2(\mathbf{u}) + |\mathbf{u}|^2. \quad (11)$$

Here $\tilde{f}(\mathbf{u})$ and $\sigma(\mathbf{u})$ denote the prediction and the one-standard-deviation based confidence interval of the LWPR model of $f(\mathbf{u})$. The constant w_p represents the importance of the accuracy demand in our task. We then can easily minimise J with respect to $\mathbf{u} = (u_1, u_2)^T$ numerically, taking into account the box constraints $0^\circ \leq u_i \leq 180^\circ$ ⁴.

³Alternatively, assume the feedback loop is so slow that it is practically useless.

⁴For our SEA this optimisation can be performed in real time, i.e., at least 50 times per second, which corresponds to the maximum control frequency of our system (50Hz).

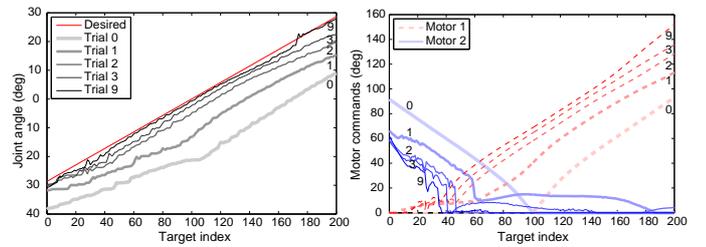


Fig. 5. Visualisation of the adaptation process during 10 trials. Left: Desired (red) and observed arm positions. Right: Motor commands for the corresponding trials. Darker and thinner lines indicate later stages of learning.

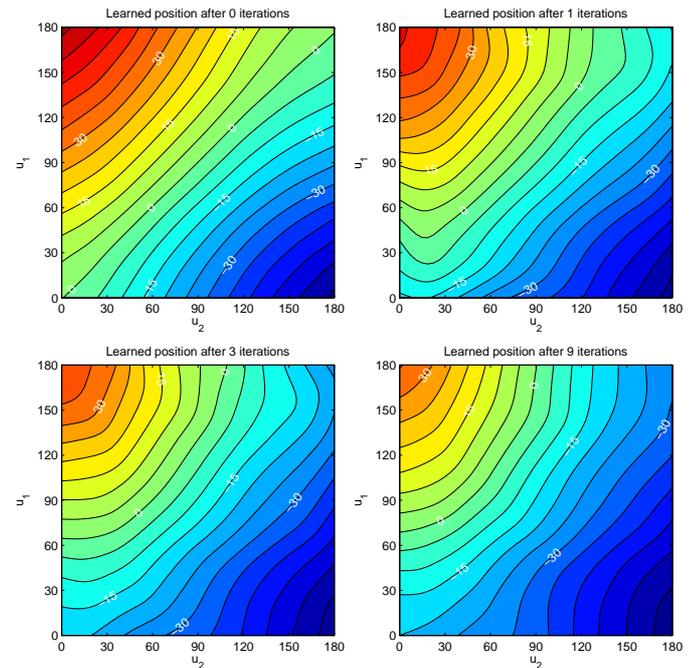


Fig. 6. Learned position models during the adaptation process. The white numbers represent the equilibrium point positions.

IV. RESULTS

In this section we present results from the optimal control model applied to the hardware described earlier in Section II. We first highlight the adaptation capabilities of this framework (Experiment 1) and then show how the learned stochastic information leads to an improved control strategy by varying impedance of the arm through co-contraction (Experiment 2).

A. Experiment 1: Adaptation towards a systematic change in the system

A tremendous advantage of the learned dynamics paradigm is that it allows to account for systematic changes without prior knowledge of the shape or source of the perturbation. To demonstrate such an adaptation scenario we setup a systematic change in the hardware by replacing the left spring, between motor 1 and the joint (i.e., between points A and C in Fig. 1), with one that has a lower, “unknown” spring constant. The aim is to hold a certain

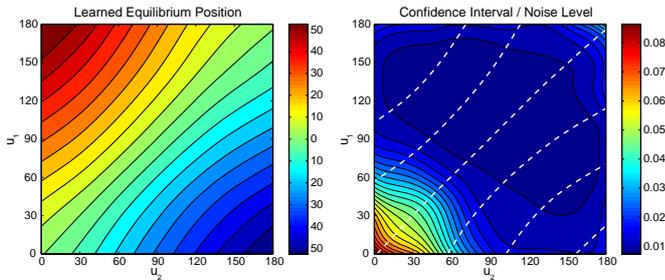


Fig. 7. Left: Learned equilibrium position as a function of the motor positions (in degrees), with contour lines spaced at 5 degree intervals. One can observe small asymmetries, e.g., along $u_1 = u_2$, which could arise from construction or sensor calibration errors. Right: “Noise landscape” given by the heteroscedastic confidence intervals of LWPR.

equilibrium position using the energy optimal position controller described in Section III-B. Expectedly the prediction about the equilibrium points (i.e., $\tilde{f}(\mathbf{u})$) does not match the real changed system properties. Next, we demonstrate how the system can adapt online and increase the performance trial by trial. We specified a target trajectory that is a linear interpolation of 200 steps between start position $\theta_0 = -30^\circ$ and target position $\hat{\theta} = 30^\circ$. We tracked this trajectory by recomputing the equilibrium positions (at a rate of 50Hz) as defined in eq. (11). At the same time we updated $\tilde{f}(\mathbf{u})$ during reaching. Due to the nature of local learning algorithms \tilde{f} is only updated in the neighbourhood of the current trajectory and therefore shows limited generalisation. To account for this, after each trial, we additionally updated the model with 400 training data points, collected from a 20-by-20 grid of the motor’s range $u_1 = u_2 = [0^\circ, 180^\circ]$. Fig. 5 depicts the outcome of this adaptation experiment. One can observe that the controller initially (lighter lines) fails to track the desired trajectory (red). However there is significant improvement between each trial, especially between trials 1 to 5. After about 9 trials the internal model has been updated and manages to track the desired trajectory well (up to the hardware’s level of precision). A look at the equilibrium position predictions in Fig. 6 confirms that the the systematic shift has been successfully learned, which is visible by the asymmetric shape. Analysing the motor commands shows that the optimal controller, for all trials, chooses the motor commands with virtually no co-contraction. This is a sensible choice as it would contradict the minimum energy cost function that we have specified.

B. The role of stochastic information for impedance control

Because co-contraction and energy consumption are opposing properties our controller will hardly make use of the redundant degree of freedom in the actuation. Even though minimum energy optimal control in an antagonistic system seems to be “unable to co-contrast” it remains our favourite choice of performance index as it also implies compliant movement and as it follows the biological motivation. But when should the optimal controller co-contrast? If we take a closer look at the stochastic information that would arise from task involving random perturbations we can see that

the produced stochasticity holds valuable information about stability of the system under the discussed learning paradigm. If the uncertainty can be reduced by co-contraction it will be reflected in the data as a control dependent “noise landscape”. Therefore the answer to the previous question is that the controller should co-contrast if and only if it can reduce the expected noise/stochasticity in the system.

Now imagine our system experiences some form of small random perturbations during control. In the hardware we realise such a scenario by adding a perturbation motor at the end of the arm, which mimics for example a drilling tool (panel “a” in Fig. 3). The perturbation on the arm is created by alternating the servo motor positions quickly every 200ms from $A = 40^\circ$ to $A = -40^\circ$ degrees. The inertia of the additional weight produces deflections on the arm from the current equilibrium position. With these perturbations we collected new training data and updated the existing LWPR model \tilde{f} . A look at the collected data reveals that the arm stabilises in regions with higher co-contraction, where the stiffness is higher. This information is contained in the learned confidence bounds (Fig. 7, right) and therefore the optimal controller computes the trade-off, between accuracy, energy and control dependent noise in the data.

C. Experiment 2: Impedance control for varying accuracy demands

Based on the learned LWPR model \tilde{f} from the previous section we can demonstrate the improved control behaviour of the stochastic optimisation with emerging impedance control. We formulate a task to hold the arm at two fixed positions $\hat{\theta} = 15^\circ$ and $\hat{\theta} = 0^\circ$ respectively. While minimising for the stochastic cost function in eq. (11), we continuously and slowly increased the position penalty within the range $w_p = [10^{-2}, 10^5]$. The perturbation motor is switched on at all times. The left columns in Fig. 8 summarises the results we discuss next: At $w_p = 10^{-2}$ to approximately $w_p = 10^0$ the optimisation neglects position accuracy and minimises mainly for energy, i.e., $u_1 = u_2 = 0$. The actual joint positions, because of the perturbations, oscillate around the mean $\theta = 0^\circ$ as indicated by the shaded area. Between $w_p = 10^0$ and $w_p = 10^2$ the position constraint starts to “catch up” with the energy constraint; a shift in the mean position towards $\hat{\theta}$ can be observed. At about $w_p = 5 * 10^1$ the variance in the positions increases as the periodic perturbation seems to hit the resonance frequency of the system. For $w_p > 10^2$ the stochastic information is weighted sufficiently such that the optimal solution co-contracts increasingly and that the accuracy further improves.

If in contrast we run the same experiment while ignoring the stochastic part in the cost function, i.e., we minimize for the deterministic cost function $J = w_p(\tilde{f}(\mathbf{u}) - \hat{\theta})^2 + |\mathbf{u}|^2$ only, we see (Fig. 9) that the system does expectedly not co-contrast and hardly improves accuracy performance.

V. CONCLUSION AND OUTLOOK

In this paper we have presented a stochastic optimal control model for antagonistically actuated systems. We pro-

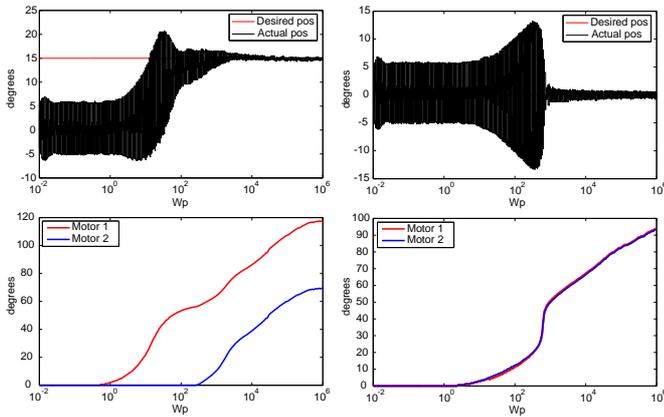


Fig. 8. Experiment with increasing position penalty w_p for two targets. Left plot column: $\hat{\theta} = 15^\circ$; right plot column: $\hat{\theta} = 0^\circ$. The plots show the desired vs. measured position and corresponding motor commands as a function of the accuracy demand (pre-factor w_p). The shaded area results from the perturbation motor.

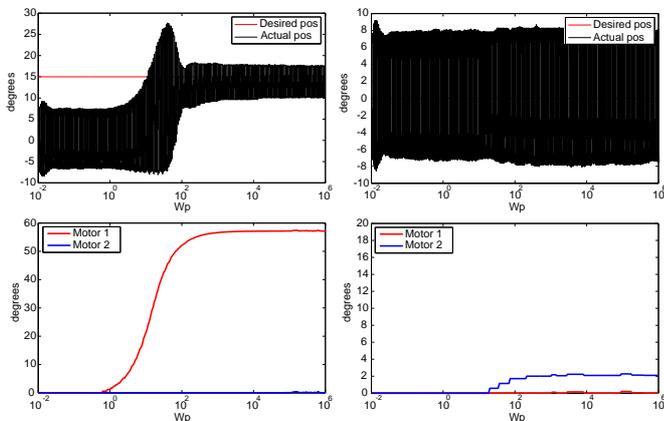


Fig. 9. The same experiment as in Fig. 8 where the stochastic information was not incorporated into the optimisation.

posed to learn, both, the *dynamics* as well as the *stochastic information* of the controlled system from sensorimotor feedback of the plant. This control architecture can account for a systematic change in the system properties (Experiment 1) and furthermore is able, by incorporating the heteroscedastic prediction variances into the optimisation, to compensate for stochastic perturbations that were induced to the plant. Doing so, our control model demonstrated significantly better accuracy performance than the deterministic optimisation (Experiment 2). The improved behaviour was achieved by co-activating antagonistic motors, i.e., by using the redundant degree of freedom in the system based on the first principles of optimality. The presented results demonstrate that this is a viable optimal control strategy for real hardware systems that exhibit hard to model system properties (e.g., asymmetries, systematic changes) as well as stochastic characteristics (e.g., using a power tool) that may be unknown a priori.

An advantage of this approach is that motor co-activation (or impedance) does not need to be specified explicitly as a control variable but that it emerges from the actual learned

stochasticity within the system (scaled with the specified accuracy demands of the task). Therefore impedance control, since it is energetically expensive, will only emerge if it actually is beneficial for the accuracy of the task. Notably such a learned optimal tradeoff between energy, accuracy and impedance has been repeatedly observed in human impedance control studies [2], [4].

The exploitation of sensorimotor stochasticity through learning is not solely applicable to govern impedance of antagonistic systems but could be applied to any kind of control or state dependent uncertainties. For example, if we would control a robot arm that suffers from oscillation or perturbations in certain joint angles or velocities, this would be visible in the noise landscape (given one has learned state dependent dynamics) and consequently those regions would be “avoided” by the optimal controller.

REFERENCES

- [1] D. P. Bertsekas. *Dynamic programming and optimal control*. Athena Scientific, Belmont, Mass., 1995.
- [2] 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*, 414:446–449, 2001.
- [3] A. A. Faisal, L. P. J. Selen, and D. M. Wolpert. Noise in the nervous system. *Nature Reviews Neuroscience*, 9:292–303, April 2008.
- [4] D. W. Franklin, E. Burdet, K. P. Tee, R. Osu, C. M. Chew, T. E. Milner, and M. Kawato. Cns learns stable, accurate, and efficient movements using a simple algorithm. *Journal of Neuroscience*, 28(44):11165–11173, 2008.
- [5] N. Hogan. Adaptive control of mechanical impedance by coactivation of antagonist muscles. *IEEE Transactions on Automatic Control*, 29(8):681–690, 1984.
- [6] D. H. Jacobson and D. Q. Mayne. *Differential Dynamic Programming*. Elsevier, New York, 1970.
- [7] M. Kawato. Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9(6):718–727, 1999.
- [8] S. A. Migliore, E. A. Brown, and S. P. DeWeerth. Biologically inspired joint stiffness control. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 4508–4513, April 2005.
- [9] Y. Nakamura and H. Hanafusa. Optimal redundancy control of robot manipulators. *The International Journal of Robotics Research*, 6(1):32–42, 1987.
- [10] G. A. Pratt and M. M. Williamson. Series elastic actuators. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 399–406, 1995.
- [11] R. F. Stengel. *Optimal control and estimation*. Dover Publications, New York, 1994.
- [12] E. Todorov. Optimality principles in sensorimotor control. *Nature Neuroscience*, 7(9):907–915, 2004.
- [13] E. Todorov. Optimal Control Theory. In K. Doya, editor, *Bayesian Brain: Probabilistic Approaches to Neural Coding*, pages 269–298. MIT Press, 2006.
- [14] G. Tonietti, R. Schiavi, and A. Bicchi. Design and control of a variable stiffness actuator for safe and fast physical human/robot interaction. *IEEE International Conference Robotics and Automation (ICRA)*, pages 526–531, 2005.
- [15] R. van Ham, T. Sugar, B. Vanderborght, K. Hollander, and D. Lefeber. Compliant actuator designs. *IEEE Robotics & Automation Magazine*, 16(3):81–94, 2009.
- [16] B. Vanderborght, R. Van Ham, D. Lefeber, T. G. Sugar, and Hollander K. W. Comparison of mechanical design and energy consumption of adaptable, passive-compliant actuators. *The International Journal of Robotics Research*, 28(1):90–103, 2009.
- [17] S. Vijayakumar, A. D’Souza, and S. Schaal. Incremental online learning in high dimensions. *Neural Computation*, 17:2602–2634, 2005.
- [18] D. M. Wolpert, Z. Ghahramani, and M. I. Jordan. An internal model for sensorimotor integration. *Science*, 269(5232):1880–1882, 1995.