Learning Optimal Impedance Control During Complex 3D Arm Movements

Abdeldjallil Naceri[1], Tobias Schumacher[2], Qiang Li[2], Sylvain Calinon[3] and Helge Ritter[2]

***Abstract*—In daily life, humans use their limbs to perform various movements to interact with an external environment. Thanks to limb's variable and adaptive stiffness, humans can adapt their movements to unstable dynamics of the external environments. The underlying adaptive mechanism has been investigated, employing a simple planar device perturbed by external 2D force patterns. In the present work, we will employ a more advanced, compliant robot arm to extend previous work to a more realistic 3D-setting. We study the adaptive mechanism and use machine learning to capture the human adaptation behavior. In order to model human's stiffness adaptive skill, we give human subjects the task to reach for a target by moving a handle assembled on the end-effector of a compliant robotic arm. The arm is force controlled and the human is required to navigate the handle inside a non-visible, virtual maze and explore it only through robot force feedback when contacting maze virtual walls. By sampling the hand's position and force data, a computational model based on a combination of model predictive control and nonlinear regression is used to predict participants' successful trials. Our study shows that participants selectively increased the stiffness within the axis direction of uncertainty in order to compensate for instability caused by a divergent external force field. The learned controller was able to successfully mimic this behavior. When it is deployed on the robot for the navigation task, the robot arm successfully adapt to the unstable dynamics in the virtual maze, in a similar manner as observed in the participants' adaptation skill.**

***Index Terms*—Human stiffness learning, Human-Centered Robotics, Modeling and Simulating Humans, Model Learning for Control, Machine Learning for Robot Control.**

I. INTRODUCTION

STABILIZING and controlling movements to adapt to unstable dynamics (i.e., variation in the system dynamics across time) of the external environment is a skill that humans develop and learn since their birth. A skill can be acquired, especially if the task is repetitive by nature or by design, by improving the task outcome of the current trial based on previous trials. However, most of our daily life interactions are not repetitive in the exact same manner across time and are mostly unstable due to external perturbations. For instance, when driving a screw, the task requirements (movements and forces) and dynamics can be totally different if we replace the screw with a different one. Despite these complexities, humans still succeed to learn to perform the task whereas, robots still lack the skill of adapting to unstable dynamics at iteration frame basis when only a small number of trials are possible (within trial or between trials).

Controlling movements in unknown environments is required in various applications, such as robotics, health care, or entertainment [1], [2]. These movements are often unstable because the environment changes frequently [3], while interacting with it. In such cases, the external forces arising from these kinds of interaction must be compensated [4]. This can be achieved by learning a dynamic internal model which represents the relationship between motor commands and motion [5]. Burdet et al. [4] investigated human arm movements within unstable dynamics. They showed that the participants adapt to the dynamics of the movement in a force field and increase their arm stiffness in the respective direction of instability. This latter result and principle were tested only for planar movements (2D). The possibility of extending the latter work to 3D was discussed and investigated by introducing methods to analyze arm stiffness in 3D space [6]–[8]. Specifically, the authors of [8] developed an experimental setup that can measure human arm stiffness in 3D using a robotic arm (Kuka light-weight robot; Kuka-LWR). A similar experimental approach to the latter referenced works was used by [6] to investigate the effect of muscle co-contraction on the arm impedance during arm movements in 3D space. Results in [6] showed individual differences in terms of arm stiffness and their orientation. However, they did not explicitly investigate whether the results of [4] (investigated in 2D movements) could be extended to 3D movements. A recent work has investigated human arm movement along a straight line within unstable dynamics in 3D using Kuka-LWR [7]. In their work they observed stiffness is selectively adjusted independently of the direction of uncertainty.

In the present work we investigate human impedance learning during complex 3D arm movements and propose a computational model based on a combination of model predictive control and nonlinear regression to predict participants' successful trials. By doing so, we aim to transfer to the robot the human's skill of quick adaptation to the unstable dynamics. The target skills are characterized by the predictive controller based on the successful trials.

Our contribution is twofold: (i) In the part of skill learning, we tackle the question of how humans can adapt their arm motion to unstable environments. For this purpose, we analyze arm movement and stiffness control of the participants.

[1]Dr. Abdeldjallil Naceri is with the Advanced Robotics Department, Istituto Italiano di Tecnologia, Genova, 16163, Italy anaceri@iit.it

[2]Tobias Schumacher and Dr. Qiang Li are with the Neuroinformatics Group, Center for Cognitive Interaction Technology (CITEC), Bielefeld University, 33619 Bielefeld, Germany tschumacher,qli@techfak.uni-bielefeld.de

[3]Dr. Sylvain Calinon is with the Idiap Research Institute, CH-1920 Martigny, Switzerland sylvain.calinon@idiap.ch

[2] Prof. Helge Ritter is the head of Neuroinformatics Group, Center for Cognitive Interaction Technology (CITEC), Bielefeld University, 33619 Bielefeld, Germany helge@techfak.uni-bielefeld.de

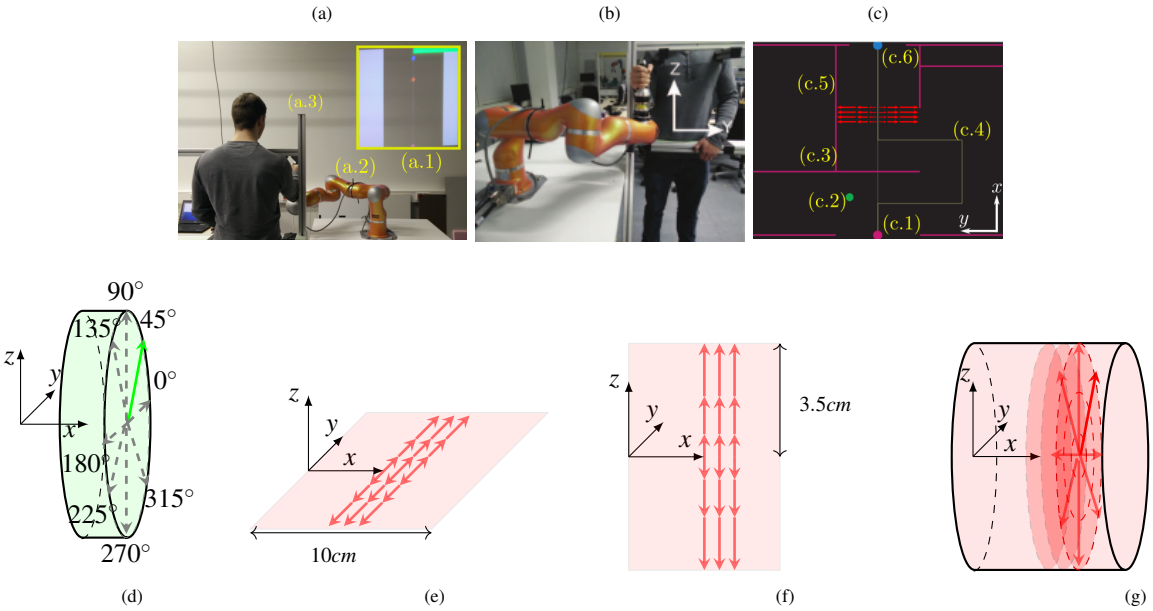


Fig. 1: Experimental materials. (a) Rear view of participants in the experimental site; (a.1): visual feedback of task starting, ending points and current hand location. (a.2) KUKA-LWR. (a.3) Fixation frame for participants' body. (b) front view of participant in experimental site holding a handle placed on Kuka-LWR. (c) Maze layout that was visible only in session 4; (c.1): starting point, (c.2): actual hand pose, (c.3): maze virtual walls, (c.4): maze solution path, (c.5): area location of force field application (detailed in e, f, g panels), (c.6): end point. (d) Constant force-field "CF". (e, f, g panels) Different divergent force-field "DF": (e) transverse condition, (f) sagittal condition, (g) combined condition.

Specifically, we use the same experimental paradigm described in [4]. We answer whether human participants would adopt an adaptation strategy during complex 3D arm movements similar to the one observed by previous studies conducted in simple planar movements. (ii) For the controller, we characterize human skills via experimental trials to encode their reactions and the variations of these reactions within a hidden semi-Markov model, which is combined with linear quadratic tracking to predict participants' trajectories and forces produced during successful trials [9], [10]. After the controller is deployed on a robot arm for a maze navigation task, its adaptive behaviour is similar to a human participant.

II. Materials and Methods

A. Participants

18 participants (3 female and 15 male, age Mean $\pm$ SD: 24 $\pm$ 4 years) took part in the experiment. Between subject design was used in this experiment by assigning each subject to one experimental condition. These participants had no neurological or motor deficits. The testing procedures were approved by the ethics committee of Bielefeld University. A written declaration of consent was obtained from the participants.

B. Hardware

The participants stood behind a frame and held with their right hand a cylindrical handle (length=15cm, diameter=5cm) which was mounted on a Kuka-LWR 4+ Robot (Fig. 1a, Fig. 1b). They were required to navigate the handle in an invisible virtual maze. The robot has 7 degrees of freedom and is working in the joint compliance mode [11]. A six-axis force/-torque sensor from ATI (*Gamma*, IP60) was mounted on the end-effector to measure participants' force data. Participants' hand position (same with position of the arm's end-effector) trajectories were estimated from the arm's forward kinematics. The data were recorded at 250 Hz. The participants received a visual feedback projected on the wall (Fig. 1a). The picture shows the current position of the end-effector, the start and finish positions. The size of the maze, illustrated in Fig. 1c, is $32 \times 30 \times 30 cm^3$ (length $\times$ width $\times$ height). Nearby the maze wall and within the force field, the participants' movements are perturbed by active force generated by the KUKA arm's controller. In the remaining area the robot's end-effector can be freely moved.

The reason behind using an invisible virtual maze is to investigate whether participants distinguish between an event of hitting a maze wall then avoiding it, versus facing a divergent external force field then compensate for it. In other words, in this task participants will face two types of external perturbations: (i) facing a haptic wall acting as an obstacle (maximum force of 150N applied against their direction of movements without bouncing them off the wall by setting the Cartesian stiffness to maximum value), versus (ii) an external divergent force field deviating them from the correct path that leads to the target. This will allow us to check whether participants proceed in straight line regardless to the perturbation type and direction or they would differentiate between the external forces direction and events and react accordingly.

C. Procedure

Participants were instructed to move and explore the task by moving the end-effector within the maze in order to find the solution path towards the target. They were asked to solve the invisible 3D maze by suitably deviating horizontally and/or vertically to avoid walls. They were also asked to execute the movements in a natural fashion as much as they can and

preferably to keep the end-effector height at the same level as their hip. Moreover, participants were informed about their movement speed using the visual feedback on current hand location (Fig. 1a). It changed color from red to green if they produced the required speed: faster than 0.3 m/s and slower than 1.3 m/s. The trial was ended as soon as they reached the target, thus, data recordings were stopped and the participants pulled the end-effector back to the starting position for the next trial. Three experimental conditions were conducted: transverse (Fig. 1e), sagittal (Fig. 1f) and combined (Fig. 1g) perturbation conditions. The participants were divided into 3 different groups. Each group conducted one single condition to avoid learning effects across conditions. Each participant conducted a total of 290 trials divided on four different experimental sessions. Participants had a short break at any time they wanted during the experiment in addition to a compulsory break between sessions.

Session 0: Training Session

Session 0 was provided for participants to get acquainted with the experimental setup and task. To this end, they carried out ten trials of reaching for the target, using the Kuka-LWR without imposing on them any force field perturbations.

Session 1: Stiffness in constant force field

In session 1, participants carried out 40 trials. This session was used to calculate participants' stiffness by activating random perturbations in a constant force field "CF" causing a positional displacement of the robot end-effector (this is similar to the condition was referred as "null field" by applying positional perturbation in [4]) within a predefined area in the maze as shown in Fig. 1d. Perturbations occurred in one of eight different directions within the $y-z$-plane (i.e. the "coronal" plane) with a step size of 45 degrees: $\{0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°\}$ with angle measured counterclockwise against y-axis. Each perturbation direction was repeated five times, giving a total of 40 trials that were randomly shuffled and presented to the participants. The total force was $F_{y,z}^{ext} = 40N$ and was calculated using Eq. 1

$$\begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} = \begin{bmatrix} 0 \\ F_y^{ext} \\ F_z^{ext} \end{bmatrix}. \tag{1}$$

Session 2: Stiffness and trajectory learning in divergent force field

In session 2, a position dependent divergent force field "DF" was applied in 120 trials performed by each participant. Here, "divergent" means that the magnitude of the force perturbation increased in proportion to the distances Δy, Δz from the x-axis using Eq. 2,

$$\begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} = \begin{bmatrix} 0 \\ K_y^{ext} \cdot \Delta y \\ K_z^{ext} \cdot \Delta z \end{bmatrix}, \tag{2}$$

where $K_y^{ext} = 700\ Nm^{-1}$ and $K_z^{ext} = 0\ Nm^{-1}$ for transverse condition (Fig. 1e). $K_y^{ext} = 0\ Nm^{-1}$ and $K_z^{ext} = 700\ Nm^{-1}$ for sagittal condition (Fig. 1f) and $K_y^{ext} = K_z^{ext} = 700\ Nm^{-1}$ for the combined condition (Fig. 1g).

Session 3: Stiffness learning evaluation

In session 3, both DF and CF were used and randomly presented to the participant. In total, 100 trials were performed in this session.

Session 4: Record the ideal trajectory

In the last session, participants were able to see the complete maze on the projected image (Fig. 1c). They performed 20 trials of navigating through the maze as best as possible. The aim of this session was to record each participant's "ideal" trajectories. There were no perturbations to the force fields presented in this session. The trajectories were averaged for each participant in order to extract their positional deviation when force fields are present to calculate the arm stiffness for trials recorded in session 3.

III. Data processing and statistics

A. Data processing

Participants' hand positions and forces were recorded through the forward kinematic computation of the robot arm and the assembled force/torque sensor. These signals were then normalized to the movement time (for each participant's trial individually) and were then filtered using a Butterworth filter (3rd order, cut-off frequency 15 Hz). Participants' learning was evaluated through both the hand trajectory and path error at specific periods. Hand path error was computed as the deviation from ideal trajectories recorded in both y- and z-coordinates in all conditions. In addition, we evaluated stiffness learning effect using stiffness measurements in y- and z-axis.

The linear relation between force and displacements might be expressed in vector/matrix notation as

$$F = K\,\Delta X, \tag{3}$$

where K is a 3×3 stiffness matrix, F is a 3×1 force vector and X is a 3×1 displacement vector. The stiffness matrix can K be separated into symmetric and antisymmetric matrix components

$$K = K^s + K^a, \tag{4}$$

where $K^s = \frac{1}{2}(K + K^T)$ and $K^a = \frac{1}{2}(K - K^T)$.

It is important to point out that $F = F(x,y,z)$ is a differentiable nonlinear function of the position which permits to express the stiffness as differential relationship between small variations of force and small displacements as expressed in [12]. The physical meaning of the symmetric property is that the force field $F(x,y,z)$ is conservative and the antisymmetric property represents the curl of the force field that is mainly produced by the participants arm. To graphically depict the stiffness matrix we use the ellipsoid representation as described in [6], [12], [13]. The stiffness ellipsoids indicate how stiffness magnitude varies with angular direction, with the ellipsoid axes identifying the directions of maximal and minimal stiffness. Therefore, the degree of alignment of the major ellipsoid axis with the direction of instability reflects how much stiffening is oriented along the instability direction. The stiffness was measured perpendicularly to the movement

direction (x-axis) along with both possible directions of the external perturbation (i.e., y- and z-axis). By doing so, we aimed to investigate whether participants would adapt by increasing selectively stiffness along the direction of uncertainty after their exposure to DFs.

B. Statistics

The present study investigated whether participants learned to navigate haptically within the maze and also to stabilize for DF. Moreover, we also checked whether this learning was specific to the perturbation direction. For this purpose we used both linear regression and a Linear Mixed Model (LMM) analysis [14] to fit the data of the hand trajectory error "Err" and arm stiffness "K" respectively. LMM is a function of fixed predictors, the stiffness directions y and z, as

$$Y = X\boldsymbol{\beta} + Z\boldsymbol{b} + \varepsilon, \tag{5}$$

where Y represents the dependent variable, X is the model matrix for the vector of fixed effects, Z is the model matrix for the random effects. $\boldsymbol{\beta}$, and $\boldsymbol{b}$ are vectors of unknown regression coefficients for fixed and random effects respectively, ε is a residual random error. We modeled our dependent variables (either error in hand path "Err" or stiffness "K") and we included stiffness direction (combination of y- and z directions) and trials as fixed effects, and participants as random effect (random pick from a population). An LMM was run for each session separately. We evaluated the effect of perturbation direction on the dependent variables (effect of the adaptation to the CF and DF) by testing the significance of the corresponding fixed-effect parameter $\boldsymbol{\beta}$ direction with the Likelihood Ratio test (LR; [15]). The LR test compares the maximized log-likelihood functions of two nested models, M1 and M0, with and without the parameter of interest, respectively. Under the null hypothesis that the simpler model, M0, is better than M1, the LR has a large sampled distribution χ^2_{df-1} [16]. Finally, a pairwise Tukey test was performed to compare dependent variable K_i, with $i = y, z$, across perturbation conditions: transverse, sagittal and combined (denoted by their directions y, z and yz respectively).

C. Hypothesis

In our null hypothesis $\boldsymbol{H}_0$, a significant effect would be found between stiffness in y- and z-axis in transverse and sagittal conditions. This hypothesis is based on the results found in [4] when they investigated 2D movements. However, since the force perturbation prediction might be harder for the participants in 3D compared to 2D, there could be an alternative coping pattern to overcome the increased task difficulty, namely to instead produce stiffness regardless to the direction of uncertainty to successfully achieve the task. This would be our alternative hypothesis denoted $\boldsymbol{H}_1$.

IV. Results

Fig. 2 shows all trajectories in all sessions for a representative participant in sagittal condition. This figure shows that participants spent some trials to learn haptically the maze.

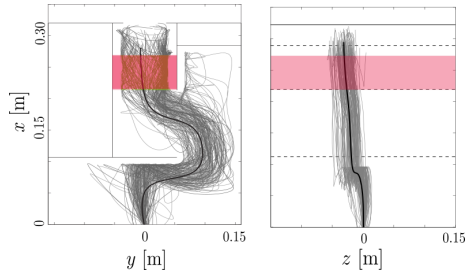


Fig. 2: All trials for representative participants. Green and blue lines represents single trial in all sessions. Black line represents the optimal trajectory recorded in session 4 without force field. Red area shows the interval where force field perturbations were applied.

Fig. 3 shows the measured dependent variables: hand trajectory and forces for a representative subject of last 10 trials within session 2. This figure shows that participant produced a successfully hand path and deliberately deviated their movements from the invisible maze wall faced at the early stage of the trial (as also shown in Fig. 2). This might hint that participant was able to learn how to haptically navigate inside the invisible maze and to compensate for the external perturbation by applying forces in the opposite direction of instability in order to counter balance the robot arm. Next, we analyze participants' hand trajectories and forces in order to validate or negate the effect observed in single trial for a representative subject (Fig. 3).

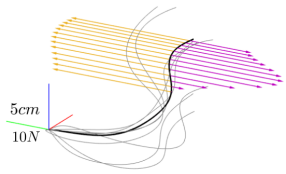


Fig. 3: Sample of individual trials for a representative participant. Solid lines represent trajectories of last 10 trials in session 2 (after learning). Black solid line is the actual trial whereas grey lines represent the rest of trials. Yellow arrows represent the applied DF on y-axis condition and purple lines represent participant's forces to compensate for the external perturbation in actual trial.

A. Trajectories: hand path error

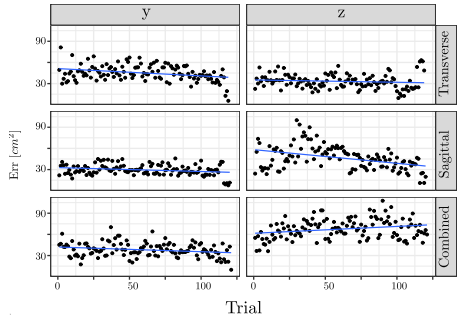
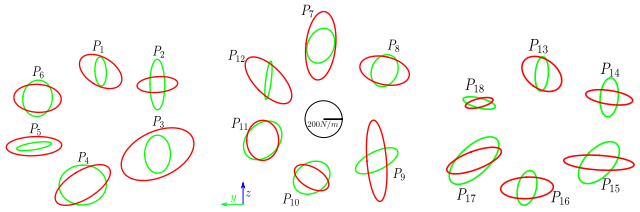


Fig. 4: Error in hand path trajectory with linear fitting in session 2 in all conditions. Each data point represent a median of all participants.

Participants showed a learning trend during the second session in some conditions only (Fig. 4). LMM analysis revealed a learning effect of y-coordinate in combined condition



(a) Transverse (b) Sagittal (c) Combined

Fig. 5: Ellipsoids representation of end point stiffness for each participant in all conditions. Green ellipses represents stiffness data of CF and the red ellipse represents stiffness data of DF. P_i denote participant i, where $i \in \{1,\ldots,6\}$ are participants in transverse condition, $i \in \{7,\ldots,12\}$ are participants in sagittal condition and $i \in \{13,\ldots,18\}$ are participants in the combined condition.

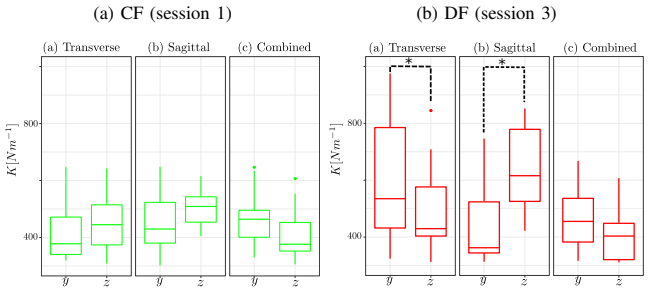


Fig. 6: Boxplot of participants stiffness in CF and DF data for each condition (transverse, sagittal and combined). Asterisks indicated that the statistical test was significant.

and of z-coordinate in both transverse and sagittal conditions ($p = 0.03$, $p = 0.01$, $p =$ 1e$-$5, respectively). In contrast there was no significant effect of y-coordinate in both transverse and sagittal conditions and of z-coordinate in the combined condition ($p = 0.07$, $p = 0.14$, $p = 0.08$, respectively). Finally, in order to shed light on the adaptation of participants we used our LMM that it was modeled on trial basis and revealed that the learning rate in hand error was reduced by -0.5, -0.23 and -0.18 cm^2 per trial for transverse, sagittal and combined condition respectively within all sessions. Despite the hand trajectory error effect was not consistent in all conditions, there was a learning rate per trial indicating that participants adapted their trajectory within the invisible maze.

B. Stiffness

1) Transverse plane: Fig. 5a shows stiffness ellipses for each participant in each condition for both CF and DF trials. For most participants, the alignment of the ellipse changed anisotropically on the y-axis within DF trials (session 2 and 3). For CF trials (sessions 1 and 3), LMM showed that the participants produced a slightly higher stiffness along the z-axis instead of the y-axis ($\Delta K = K_z - K_y$; mean $\pm$ standard deviation: $M \pm SD = 42.79 \pm 22.22 \frac{N}{m}$). The LR test revealed no significance ($\chi^2_1 = 1.27$, $p = 0.52$; see Fig. Fig. 6a). In DF (sessions 2 and 3) data, the LMM showed that the participants had a higher stiffness along the y-axis in comparison to the z-axis ($\Delta K = K_y - K_z$; $M \pm SD = 102.53 \pm 65.01 \frac{N}{m}$) and with the LR test revealed a significant effect ($\chi^2_1 = 6.34$, $p = 0.04$; see Fig. 6b). The pairwise Tukey test was performed on the dependent variable K_y revealed non-significant effect in CF data (y vs. z: $p = 0.761$, y vs. yz: $p = 0.61$, yz vs. z: $p = 0.96$). In contrast, a significant effect was found in DF of y-axis (y vs. z: $p = 0.04$, y vs. yz: $p = 0.11$, yz vs. z: $p = 0.94$).

Recall that our null hypothesis $\boldsymbol{H}_0$ is the existence of a significant effect on participants' arm stiffness in the direction of the applied external perturbation. Alternatively, rejection of $\boldsymbol{H}_0$ will consequently approve $\boldsymbol{H}_1$ where there will be no effect indicating that participants increased their arm stiffness regardless to the external perturbation. There was an effect in this condition, thus, we validate $\boldsymbol{H}_0$ (reject $\boldsymbol{H}_1$) indicating that participants learned to compensate the instability of the divergent field by increasing stiffness in the direction of uncertainty (y-axis).

2) Sagittal plane: Some participants showed a higher stiffness in the sagittal plane in DF trials (see Fig. 5b). The anisotropic alignment of the ellipse has rotated along the z-axis [17]. The LMM for the CF revealed that the participants produced a higher stiffness along the z-axis ($\Delta K = K_z - K_y$; $M \pm SD = 41.98 \pm 25.13 \frac{N}{m}$) but this difference did not reach significance level, measured with the LR test ($\chi^2_1 = 0.88$, $p = 0.64$; see Fig. 6a). In the DF, LMM showed that the participants achieved a higher stiffness in the z-axis than in the y-axis ($\Delta K = K_z - K_y$; $M \pm SD = 205.14 \pm 61.99 \frac{N}{m}$). In addition, this result was significant, which was confirmed by the LR test ($\chi^2_1 = 8.24$, $p = 0.02$; see Fig. 6b). Then, a pairwise Tukey test was performed on the dependent variable K_z and no significant effect was found in CF data (y vs. z: $p = 0.88$, y vs. yz: $p = 0.93$, yz vs. z: $p = 0.87$). In contrast, a significant effect was found in DF data of z-axis (z vs. y: $p = 0.03$, z vs. yz: $p < 0.01$, yz vs. y:: $p = 0.55$). Similarly to the transverse condition, based on this result we validate $\boldsymbol{H}_0$ (reject $\boldsymbol{H}_1$) indicating that participants learned to compensate the instability of the divergent field (DF) by increasing the stiffness in the direction of uncertainty (z-axis).

3) Combined planes: The orientation of the stiffness ellipses turned in both directions (see Fig. 5c). The participants showed a similar stiffness through the LMM in CF data ($\Delta K = K_y - K_z$;$M \pm SM = -62.44 \pm 20.19 \frac{N}{m}$) and DF data ($\Delta K = K_y - K_z$; $M \pm SD = -55.28 \pm 50.37 \frac{N}{m}$). Also there was no significant effect which is measured with the LR test in CF data ($\chi^2_1 = 2.141$, $p = 0.34$; see Fig. 6a) and DF data ($\chi^2_1 = 4.1074$, $p = 0.13$; see Fig. 6b). The pairwise Tukey revealed no significant effect CF data (y vs. z: $p = 0.41$, y vs. yz: $p = 0.99$, yz vs. z: $p = 0.52$) also in DF data (y vs. z: $p = 0.96$, y vs. yz: $p = 0.14$, yz vs. z: $p = 0.28$).

Through this result of the combined condition, we can conclude that the increase of stiffness was in both y- and z-axis responding to the external perturbation that was applied randomly across trials in both directions (Fig. 6b). This control condition credit results found in the other two conditions and validates our null hypothesis $\boldsymbol{H}_0$ indicating that stiffness control is specific to the perturbation direction.

Overall, our results showed that participants corrected their hand trajectories in at least one direction but their arm stiffness was increased in the direction of the uncertainty. Next, we present a controller that learns the adaptation strategy to the DF of the participants, by exploiting the data collected during the successful trials (in both session 2 and 3 after adaptation).

C. Simulation with varying perturbations

First of all, we need to mention that in the human experimental part we used the robot as tool to generate external perturbations and to measure participants' trajectories and forces. The maze walls were simulated using Kuka-LWR robot stiffness control (working in compliance mode). In this subsection, we present our adaptive controller learned from human successful trials. The controller is evaluated and applied on the robot arm within the Gazebo simulation, by implementing the robot end-effector navigation in the same maze space.

Based on the collected data, we built a controller capable of both anticipation and reaction to perturbations with an approach detailed in [9], [10], which combines linear quadratic tracking (LQT) with a hidden semi-Markov model (HSMM). The learned controller relies on the participants' measured dependent variables and associated (co)variability. By doing so, we aim to model and encode both the adaptation strategy to the DF of the participants' observed during the successful trials.

The approach is implemented with an augmented state space composed of 3D position, 3D velocity and 3D force profiles, whose evolution in time is encoded in the form of an HSMM, representing the positions, velocities and forces as a set of states, together with transition and duration information, all represented as probabilistic distributions [18]. This generative model is employed to generate the parameters of the cost function of an extended LQT controller in task space driven by a double integrator (virtual unit mass system). In this way, the HSMM generates a stepwise signal as a reference to be tracked, together with the tracking precision (as full precision matrices), which is used to define the LQT cost function. Such problem can be solved analytically, yielding a trajectory distribution in both control space and state space [10]. The extended LQT problem that we propose to solve is formulated as the minimization of the cost

$$\begin{aligned} c &= \sum_{t=1}^{T} (\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t)^\top \boldsymbol{Q}_t (\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t) + (\hat{\boldsymbol{u}}_t - \boldsymbol{u}_t)^\top \boldsymbol{R}_t (\hat{\boldsymbol{u}}_t - \boldsymbol{u}_t), \\ &\text{s.t.} \quad \boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}\boldsymbol{u}_t, \end{aligned} \tag{6}$$

where $\hat{\boldsymbol{x}}_t \in \mathbb{R}^6$ is a state reference (composed of position and velocity), $\hat{\boldsymbol{u}}_t \in \mathbb{R}^3$ is a force reference, $\boldsymbol{Q}_t \in \mathbb{R}^{6\times 6}$ is a precision matrix for the state profile to track, $\boldsymbol{R}_t \in \mathbb{R}^{3\times 3}$ is a precision matrix for the control profile and T is the total average duration of the task. The states $\hat{\boldsymbol{x}}_t$, forces $\hat{\boldsymbol{u}}_t$, together with their associated covariance matrices $\boldsymbol{Q}_t^{-1}$ and $\boldsymbol{R}_t^{-1}$ are learned and retrieved by the HSMM, see [10] for details. $\boldsymbol{A}$ and $\boldsymbol{B}$ describe the linear evolution of a point mass system (corresponding to a double integrator). The cost function (6) can be minimized by recursion (see [10] for details), and results in a controller of the form

$$\hat{\boldsymbol{u}}_t = \hat{\boldsymbol{u}}_t^{\text{fb}} + \hat{\boldsymbol{u}}_t^{\text{ff}}, \quad \text{with} \quad \hat{\boldsymbol{u}}_t^{\text{fb}} = \hat{\boldsymbol{K}}_t (\hat{\boldsymbol{x}}_t - \boldsymbol{x}_t), \tag{7}$$

characterized by feedback $\hat{\boldsymbol{u}}_t^{\text{fb}}$ and feedforward $\hat{\boldsymbol{u}}_t^{\text{ff}}$ control commands, with adaptive feedback gain matrices $\hat{\boldsymbol{K}}_t \in \mathbb{R}^{3\times 6}$ (concatenation of stiffness and damping matrices).

When a perturbation occurs, (7) will come back smoothly to the reference trajectory, with a reactivity that depends on the tradeoff between tracking accuracy and control effort characterized by $\boldsymbol{Q}_t$ and $\boldsymbol{R}_t$, which are defined based on the regularities observed within the paths and the force perturbations (which can vary along different directions). Indeed, the feedback gain matrices $\hat{\boldsymbol{K}}_t$ in LQT are adaptive and depend on the ratio between $\boldsymbol{Q}_t$ and $\boldsymbol{R}_t$ along different directions (full matrices). If $\boldsymbol{Q}_t$ is fixed and if we vary $\boldsymbol{R}_t$, the controller will automatically adapt the roles of $\hat{\boldsymbol{u}}_t^{\text{fb}}$ and $\hat{\boldsymbol{u}}_t^{\text{ff}}$ in the tracking problem. High $\boldsymbol{R}_t$ along certain directions will result in a controller favoring $\hat{\boldsymbol{u}}_t^{\text{ff}}$ over $\hat{\boldsymbol{u}}_t^{\text{fb}}$ in these directions, and vice-versa. For our problem formulation, it means that if the learning system observes regular perturbation forces over several consecutive trials, the HSMM will result in high $\boldsymbol{R}_t$, which will lower the feedback gains $\hat{\boldsymbol{K}}_t$. Namely, the proposed controller can anticipate the force by generating an open loop command $\hat{\boldsymbol{u}}_t^{\text{ff}}$, while remaining compliant. In contrast, if the learning system observes varied random perturbation forces over several consecutive trials, the controller cannot anymore anticipate the force, and will instead adopt a stiffening behavior to reject perturbations (resulting in higher feedback gains $\hat{\boldsymbol{K}}_t$ in the directions of the random perturbations).[1]

We tested the proposed approach with recordings in two conditions: with regular and irregular perturbation forces, from a set of 5 trials per condition (last trials from different participants). Fig. 7-8 present the results. We can see that in the two conditions, the proposed controller can exploit the observed regularities in the paths and in reaction forces efficiently, by providing controllers that can efficiently reproduce the task. For better illustration, we implemented both controllers on Kuka-LWR navigating in the maze and receiving external perturbations within the simulation environment (Gazebo) as we did for participants (Fig. 9). These results show the efficiency of the proposed controller, which can anticipate the movement while adapting to the external perturbation, by relying on both trajectory and force data of the participants.

V. DISCUSSION

In this work, we investigated the human strategy of quick adaptation to unstable dynamics in order to conceive a controller for a robot with similar behaviour. In the human experiment, participants adapted to unstable dynamics and successfully navigated within the non-visible virtual maze relying only on the haptic modality. Moreover, they adapted the stiffness within the direction of uncertainty in a single plane. In the combined plane condition, participants increased the stiffness in both directions which confirms and credit our results. We then proposed to exploit trajectory and force data to learn controllers with natural anticipation and reaction behaviors. The proposed approach can compensate both regular and irregular force perturbations, and successfully reproduced the demonstrated paths to the target. In contrast, the baseline controller relying only on path information failed to efficiently compensate for the external force perturbations.

[1]Matlab/Octave source code examples for the proposed controller can be found at https://gitlab.idiap.ch/rli/pbdlib-matlab/.

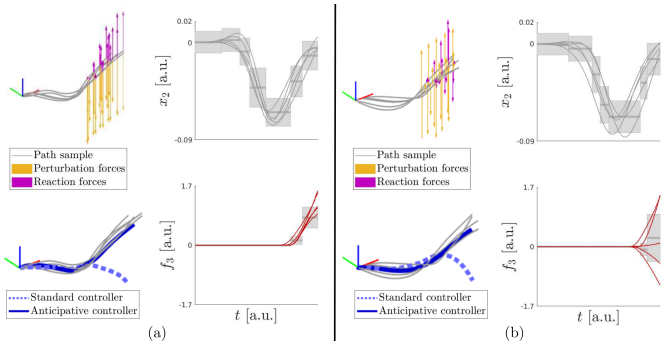


Fig. 7: HSMM-LQT approach with regular *(a)* and irregular *(b)* perturbation forces (yellow arrows). The coordinate system depicts x_1, x_2 and x_3 directions as red, green, blue segments. The forces recorded by the user compensating for these perturbations are displayed in purple arrows and purple lines. An HSMM with 7 components (number selected here experimentally) is used in each condition to learn the state and compensation force profiles used as reference, together with the (co)variations (depicted as blocks in the timeline graphs, representing the mean and the contour of one standard deviation). The path resulting from the proposed anticipative controller is depicted in blue lines. As a baseline, we compare the proposed approach to a standard controller ignoring the learning and anticipation of the compensation forces. The resulting path is represented in blue dotted lines. This baseline is implemented as a standard LQT controller without force anticipation, by solving (7) with a constant isotropic control weight $\boldsymbol{R}_t = \boldsymbol{R}$ and a force reference $\hat{\boldsymbol{u}}_t = \boldsymbol{0}$.

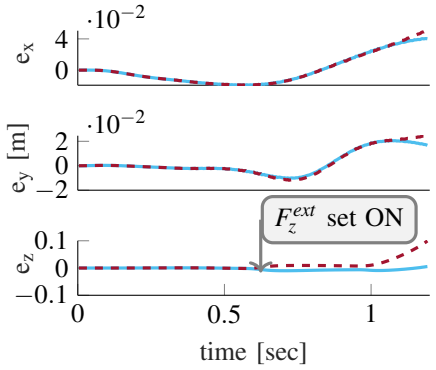


Fig. 8: Error profiles between mean participants' trajectory and model reproduction results. Dashed lines represents the error when using data of standard controller. Solid lines represents the error when using data of anticipative controller.

The latter result credit the coupling between both adaptation of hand trajectory within invisible maze and the produced hand stiffness toward the uncertainty direction.

A. Human Experiment

The experiment showed that all participants successfully reached the goal of the maze after learning. However, high errors were committed per trial, which might be due to different muscle activation in different phases inside the maze [19]. Hand path error effect was not consistent in both directions in y- and z-axis within all conditions. The latter result indicate that participants had to correct hand path error, in given direction that is not necessary along with the direction of uncertainty, to adapt their trajectory within the invisible maze. One possible explanation of this result is that the arm movement produced during the planar force field perturbation was unconstrained in 3D inducing errors in both y- and z-axis.

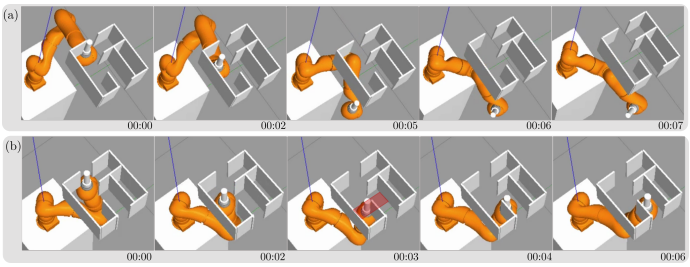


Fig. 9: Snapshots of the reproduced trajectories represented in Fig. 7. Panel (a) represents frames from the reproduced trajectories with standard controller (without force data). Panel (b) represents frames from the reproduced trajectories with anticipative controller (with force data).

The latter result confirm and extend previous findings showing that endpoint stiffness was controlled independently of the arm joint torques [4].

For arm stiffness, we showed that the stiffness ellipsoids adapted to the direction of uncertainty. Our results confirm and extend previous studies conducted in 2D arm movements [4], [20]. Specifically, the direction of the largest stiffness axis was along the axis of the external perturbation. Here we also found that arm stiffness was directed toward the axis of the external perturbation although the movement was not restricted in specific direction. This result corroborated the idea that optimal movement required a proper control of both arm viscosity and stiffness [21]. The stiffness modulation of a given limb is a property of active muscles control. Specifically, it has been shown that muscle co-activation contribute directly to the increase of endpoint stiffness [22]. Endpoint stiffness, viscosity and inertia depend on posture. While viscosity behaves in a similar fashion as stiffness, inertia aligns with the forearm [17], [23], [24]. In our experiments, participants' elbow was unconstrained to some extent whereas their shoulder was fixed by the experimental frame. Releasing the elbow constraint allowed us to tackle the question of stiffness control during complex 3D arm movements. On the other side, this might explain individual differences recorded in participants' stiffness amplitudes. For this purpose, we used LMM in our statistical tests that cope with this kind of problems in participants' data (random effect).

B. Controller model and simulation

A controller was developed based on the participants' successful trials in both transverse and sagittal conditions to encode their reactions and the variations of these reactions using an approach combining HSMM and LQT, which was used to replicate participants' adaptation strategy in trajectories and forces. Results of this part showed that the controller was successful only when including the participants' forces in the learning. The proposed controller showed a capability of adapting to unstable dynamics of the external environment requiring only few trials from different participants and conditions. Moreover, the controller was further tested and successfully compensated for the irregular perturbations which is close to daily life interactions scenarios. Prior work [3] has proposed a controller adapting force and impedance in presence of external perturbation; this controller successfully predicted participants' adaptation behavior published in [4]

and it does not require force sensing data in its input. Our proposed controller uses force data which might be crucial for the interaction and might confirm the coupling between both adaptations in hand trajectory and stiffness direction. Using only trajectory hand path error data for the controller is interesting but adapts only to the invisible maze but not to the external perturbation (Fig. 7). Another example for a controller based on iterative learning was proposed by [25] which was able to predict participants adaptation behaviour using data observed by [4] as well. This latter requires linearization of human body dynamics to demonstrate stability and convergence.

VI. CONCLUSIONS

In this work, we showed that participants adapted to the external unstable dynamics during complex 3D arm movements. With this work, we extend findings of works performed in 2D planar reaching movements. Our results suggest that participants selectively adapt their stiffness to the direction of instability even in complex 3D arm movements. Only few trials from participants' data were enough for our controller to predict a successful trajectory and adapt to both regular and irregular perturbations. As a future direction, we aim to extend our controller even richer adaptation learning skills and transfer them to a real robotic arm [26].

ACKNOWLEDGMENT

This research was in part supported by the Cluster of Excellence "Cognitive Interaction Technology" (CITEC) EXC277 at Bielefeld University funded by the German Research Foundation (DFG), by the "DEXMAN" project (Project number: 410916101) funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), by the Istituto Nazionale per l'Assicurazione contro gli Infortuni sul Lavoro (INAIL), and by European Commission's Horizon 2020 Programme (COLLABORATE project, grant agreement 820767).

REFERENCES

[1] E. Colgate, W. Wannasuphoprasit, and M. Peshkin, "Cobots: robots for collaboration with human operators," in *Proceedings of the ASME Dynamic Systems and Control Division*, Y. Kwon, D. Davis, and H. Chung, Eds., vol. 58. ASME, 12 1996, pp. 433–439.

[2] O. Lambercy, L. Dovat, R. Gassert, E. Burdet, C. L. Teo, and T. Milner, "A haptic knob for rehabilitation of hand function," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 3, p. 356–366, 2007.

[3] C. Yang, G. Ganesh, S. Haddadin, S. Parusel, A. Albu-Schaeffer, and E. Burdet, "Human-like adaptation of force and impedance in stable and unstable interactions," *IEEE Transactions on Robotics*, vol. 27, no. 5, pp. 918–930, Oct 2011.

[4] 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, p. 446–449, 2001.

[5] R. Shadmehr and F. Mussa-Ivaldi, "Adaptive representation of dynamics during learning of a motor task," *The Journal of Neuroscience*, vol. 14, no. 5, p. 3208–3224, Jan 1994.

[6] H. Patel, G. O'Neill, and P. Artemiadis, "Regulation of 3d human arm impedance through muscle co-contraction," in *Dynamic Systems and Control Conference*, 2013.

[7] F. Steinbeck, "Learning optimal impedance during arm movements in three dimensional space," Master's thesis, University Bielefeld, 2016.

[8] van der Smagt Patrick, C. Castellini, and H. Urbanek, "Human arm impedance and emg in 3d," in *SKILLS, International Conference on Multimodal Interfaces for Skills Transfer*, December 2009. [Online]. Available: https://elib.dlr.de/62664/

[9] S. Calinon, "Robot learning with task-parameterized generative models," in *Robotics Research*, A. Bicchi and W. Burgard, Eds. Springer International Publishing, 2018, vol. 3, pp. 111–126.

[10] S. Calinon and D. Lee, "Learning control," in *Humanoid Robotics: a Reference*, P. Vadakkepat and A. Goswami, Eds. Springer, 2019, pp. 1261–1312.

[11] R. Bischoff, J. Kurth, G. Schreiber, R. Koeppe, A. Albu-Schäffer, A. Beyer, O. Eiberger, S. Haddadin, A. Stemmer, G. Grunwald *et al.*, "The kuka-dlr lightweight robot arm-a new reference platform for robotics research and manufacturing," in *ISR 2010 (41st international symposium on robotics) and ROBOTIK 2010 (6th German conference on robotics)*. VDE, 2010, pp. 1–8.

[12] F. A. Mussa-Ivaldi, N. Hogan, and E. Bizzi, "Neural, mechanical, and geometric factors subserving arm posture in humans," *Journal of Neuroscience*, vol. 5, no. 10, pp. 2732–2743, 1985.

[13] L. Masia and V. Squeri, "A modular mechatronic device for arm stiffness estimation in human robot interaction," *IEEE/ASME Transactions on Mechatronics*, vol. 20, no. 5, p. 2053–2066, 2015.

[14] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software, Articles*, vol. 67, no. 1, pp. 1–48, 2015. [Online]. Available: https://www.jstatsoft.org/v067/i01

[15] J. C. Pinheiro and D. M. Bates, "Mixed-effects models in sand s-plus," *Statistics and Computing*, 2000.

[16] B. M. Bolker, M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White, "Generalized linear mixed models: a practical guide for ecology and evolution," *Trends in Ecology & Evolution*, vol. 24, no. 3, pp. 127 – 135, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0169534709000196

[17] F. Mussa-Ivaldi, N. Hogan, and E. Bizzi, "Neural, mechanical, and geometric factors subserving arm posture in humans," *Journal of Neuroscience*, vol. 5, no. 10, pp. 2732–2743, 1985. [Online]. Available: https://www.jneurosci.org/content/5/10/2732

[18] M. J. A. Zeestraten, S. Calinon, and D. G. Caldwell, "Variable duration movement encoding with minimal intervention control," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 497–503.

[19] T. Verstynen and P. N. Sabes, "How each movement changes the next: An experimental and theoretical study of fast adaptive priors in reaching," *Journal of Neuroscience*, vol. 31, no. 27, pp. 10 050–10 059, 2011. [Online]. Available: http://www.jneurosci.org/content/31/27/10050

[20] D. W. Franklin, G. Liaw, T. E. Milner, R. Osu, E. Burdet, and M. Kawato, "Endpoint stiffness of the arm is directionally tuned to instability in the environment," *Journal of Neuroscience*, vol. 27, no. 29, pp. 7705–7716, 2007. [Online]. Available: http://www.jneurosci.org/content/27/29/7705

[21] N. Hogan, "Impedance control: An approach to manipulation," in *1984 American Control Conference*, June 1984, pp. 304–313.

[22] D. W. Franklin, R. Osu, E. Burdet, M. Kawato, and T. E. Milner, "Adaptation to stable and unstable dynamics achieved by combined impedance control and inverse dynamics model," *Journal of Neurophysiology*, vol. 90, no. 5, pp. 3270–3282, 2003, pMID: 14615432. [Online]. Available: https://doi.org/10.1152/jn.01112.2002

[23] E. J. Perreault, R. F. Kirsch, and P. E. Crago, "Multijoint dynamics and postural stability of the human arm," *Experimental Brain Research*, vol. 157, no. 4, pp. 507–517, Aug 2004. [Online]. Available: https://doi.org/10.1007/s00221-004-1864-7

[24] S. Togo, T. Yoshioka, and H. Imamizu, "Control strategy of hand movement depends on target redundancy," in *Scientific reports*, 2017.

[25] S. Zhou, D. Oetomo, Y. Tan, E. Burdet, and I. Mareels, "Modeling individual human motor behavior through model reference iterative learning control," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 7, pp. 1892–1901, July 2012.

[26] N. Jaquier, L. Rozo, and S. Calinon, "Analysis and transfer of human movement manipulability in industry-like activities," 2020.