Variable Impedance Learning Control with Faster Re-learning and Reduced Initial Errors in Re-perturbation for Robots Operating in Divergent Force Fields

Shail Jadav
SysIDEA Robotics Lab, IIT Gandhinagar,
Gandhinagar, Gujarat, India

Shubhankar Riswadkar
SysIDEA Robotics Lab, IIT Gandhinagar,
Gandhinagar, Gujarat, India

Sujay D. Kadam
DA-IICT,
Gandhinagar, Gujarat, India

Harish J. Palanthandalam-Madapusi
SysIDEA Robotics Lab, IIT Gandhinagar,
Gandhinagar, Gujarat, India

ABSTRACT

In recent years, robots with higher levels of sophistication in controlling the mechanical impedance of interaction have received increasing attention. These robots can precisely move and manipulate objects in predetermined situations by exploiting impedance control. However, this requires that the desired impedance of interaction be factored into the control design process beforehand, and interaction impedance is typically not dynamically modulated traditionally. In contrast, humans are skilled at modulating mechanical impedance, mainly mechanical stiffness, in their interactions with their surroundings. This ability is vital for successful interaction with the environment. This human behaviour has inspired various variable impedance learning controllers for tool manipulation. However, humans not only demonstrate the ability to stabilise interaction with a divergent force-field but also demonstrate a few higher-level learning features. Savings is one such feature, and it refers to people being able to learn faster when the same divergent-field is applied again. Retention is another essential feature, allowing humans to initially demonstrate lower error when the same exposure to a divergent field (perturbation) is applied again. These features have practical utility in the robotic application and have not been explored, this paper presents a novel approach to variable impedance learning control for robots inspired by human motor learning mechanisms. We combine iterative learning control, feedforward control, and ideas from neuroscience to synthesise the controller. The proposed method aims to vary impedance to stabilise the interaction with divergent-field while incorporating savings and retention features. Simulation results with a two-link serial chain manipulator demonstrate the proposed method's efficacy, savings, and retention.

CCS CONCEPTS

• **Computer systems organization** → **Robotic control**.

KEYWORDS

Variable impedance learning controller, Human motor learning mechanisms, Divergent force-field

ACM Reference Format:
Shail Jadav, Shubhankar Riswadkar, Sujay D. Kadam, and Harish J. Palanthandalam-Madapusi. 2023. Variable Impedance Learning Control with Faster Re-learning and Reduced Initial Errors in Re-perturbation for Robots Operating in Divergent Force Fields. In *Advances In Robotics - 6th International Conference of The Robotics Society (AIR 2023), July 05–08, 2023, Ropar, India.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3610419.3610423

AIR 2023, July 05–08, 2023, Ropar, India

ACM ISBN 978-1-4503-9980-7/23/07…$15.00
https://doi.org/10.1145/3610419.3610423

1 INTRODUCTION

Robots with higher levels of sophistication in controlling the mechanical impedance of interaction have received much more attention in recent years, and this field of robotics is expanding quickly. These robots can precisely move and manipulate objects in predetermined situations by exploiting impedance control. Although methods like impedance control have gained popularity in recent years, these methods require that a desired impedance of interaction be factored into the control design process beforehand. In other words, interaction impedance is typically not dynamically modulated as the robot requires. On the other hand, humans are incredibly skilled at modulating mechanical impedance, mainly mechanical stiffness, in their interactions with their surroundings. In repetitive tasks, humans can modulate the stiffness of interaction and learn the optimal stiffness 'behaviour' most appropriate for the task (without using any explicit strategies for stiffness)[5]. It is worth noting that modulating impedance and modulating forces are independent tasks. This ability to modulate interaction impedance is vital while interacting with the environment.

Several recent works have modulated the impedance of interactions in robots [1, 2, 4, 15, 17, 18] and learning such behaviour from repeated trials. Studies using reinforcement learning to control stiffness are [4, 15], and a recent review on variable impedance control through learning and control approaches is given in [1]. A general overview of impedance control approaches has been given recently in [2]. The authors of [4] propose a reinforcement learning-based gain scheduling scheme that allows the robot controller to track motion trajectories and the associated impedance. A recent study [13] has demonstrated reinforcement learning for robot tasks using force/torque control in the operational space. In [7], the authors

argue that iterative learning control (ILC) based feedforward plus a low-gain feedback action can derive a 'human-like' behaviour for soft actuators. Iterative learning of impedance parameters over time for a human–robot cooperative reaching task by a human operator moving the end-effector between initial and final positions is discussed in [17]. Of particular note is the work in [18] that develops human-like adaptation algorithms for robots, highlights particular challenges due to an unstable environment, and develops an impedance tuning algorithm that can stabilize the interaction dynamics. Nonetheless, this area remains an area of active investigation and is perhaps in its relative infancy. Furthermore, many of these works focus on human-like impedance adaptation [18] without paying as much attention to higher-level learning features of human motor learning. We define first-time exposure to a divergent field as a perturbation, and subsequent exposure to the same divergent field as re-perturbation, in accordance with human motor learning literature [12].

Human motor learning, also known as human learning mechanisms, can stabilize and exhibit a range of higher-level learning features. Savings is one such feature, allowing people to learn faster in the re-perturbation phase [12]. Retention is another essential feature, allowing humans to initially demonstrate lower error in the re-perturbation phase [12]. While applicable in robotic systems, these features are yet to be incorporated into human-inspired variable-impedance learning controllers [18]. Human motor learning mechanisms, such as savings and retention, have practical applications in settings where impedance modulation is necessary for optimal performance. An example of this is a pill-picking robot that must adjust its impedance for different batches of pills. By incorporating these features, the robot can more efficiently learn and improve its performance when faced with the same batch of pills again, ultimately leading to an overall increase in performance.

This paper presents a novel approach to learning control for robots inspired by human motor learning. We aim to develop an algorithm that allows robots to autonomously learn impedance and feed-forward force for a given task through repeated interactions with an unstable environment without relying on explicit force-based or impedance-based strategies. In line with human motor learning, we seek a computationally simple algorithm to implement, which requires minimal memory and prior knowledge of disturbances and can demonstrate adaptation, savings, and retention during interaction with the environment.

2 PROBLEM FORMULATION

Consider a robot with n degrees of freedom whose dynamics is expressed in standard manipulator form as

$$M(\theta(t))\ddot{\theta}(t) + C(\theta(t), \dot{\theta}(t))\dot{\theta}(t) + G(\theta) = \tau(t) + \tau_{\mathrm{e}}(t), \quad (1)$$

where $M \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^n$, are the inertia, coriolis and centripetal, and gravitational terms respectively. $\theta \in \mathbb{R}^n$ is the vector of joint variables (most commonly joint angles), $\tau \in \mathbb{R}^n$ is the vector of inputs (most commonly joint torques) to the robot system and $\tau_e(t) \in \mathbb{R}^n$ represents torques/forces of interactions from the environment. For simplicity of discussions, henceforth, we refer to θ as joint angles and τ are the joint torques while recognizing that they may also represent linear displacements/forces (in case of any prismatic joints in the robot). We pay particular attention to unstable environmental interactions, that is, situations in which $\tau_e(t)$ actively acts to destabilize the dynamics, and consider the problem of developing a learning controller that can learn τ (consisting of feedback, feedforward, and impedance terms) to both stabilize the unstable dynamics (due to interaction) and simultaneously help the robot track the desired trajectory. Thus the control inputs τ help the robot perform a simple tracking task in the presence of unstable environmental interactions. Furthermore, we examine the learned impedance to see if it can demonstrate adaptation with higher-level motor learning features like savings and retention.

3 CONTROL DESIGN

For tracking control, since it is a common practice to consider a combination of feedforward and feedback control, and further incorporating the desired impedance terms, we consider a controller structure of the form

$$\tau_k(t) = \tau_{\mathrm{ILC},k}(t) + \tau_{\mathrm{FF},k}(t) + K_k(\theta_d(t) - \theta(t)) + D_k(\dot{\theta}_d(t) - \dot{\theta}(t)). \quad (2)$$

Here, t is the time index and k is the trial (since the robot learns over several repeated trials). All four terms are adapted (learnt) over trials (and hence the subscripts k) with $\tau_{\mathrm{ILC},k}(t)$ being the torque from iterative learning control (ILC), $\tau_{\mathrm{FF},k}$ being the torques from feedforward control (FF), and $K_k$and D_k being the stiffness and damping matrices, respectively. $\theta_d(t)$ and $\dot{\theta}_d(t)$ represent the desired robot trajectory, and $\theta(t)$ and $\dot{\theta}(t)$ represent the measured robot trajectory. While there are conceptual similarities with the structure proposed in [8, 18], there are also some important differences. All four terms in the controller (2) are adapted, and the adaptation laws as relatively simple.

We start with iterative learning control (ILC) and adopt the Arimoto algorithm [3, 14] that is both simple and widely used in the ILC literature to arrive at the ILC control input

$$\tau_{\mathrm{ILC,k}}(t) = \lambda_{\mathrm{ILC}}\tau_{\mathrm{ILC,k-1}}(t) + \gamma_{\mathrm{ILC}}(\theta_{\mathrm{d,k-1}}(t) - \theta_{k-1}(t)), \quad (3)$$

where λ_{ILC} is the scalar retention rate that is used to determine what proportion of the inputs used in the previous trial $k-1$ are to be retained, and γ_{ILC} is the scalar learning rate that determines adaptation based on task error in the previous trial. Inputs and errors from the previous trial $k-1$ are used as the basis for adaptation.

Next, we note that for the feedforward control input for any trajectory tracking controller[9], it is a natural choice to employ a model inversion (or an inverse model-based) controller [10]. We also note that a feedforward, inverse-model-based contribution to the control action is also well reported in the human motor learning community [6, 11]. We thus synthesize our second controller equation to complement (3) as

$$\tau_{\mathrm{FF},k}(t) = \lambda_{\mathrm{FF}}\tau_{\mathrm{FF},k-1}(t) + \gamma_{\mathrm{FF}}\Gamma\left(\mathcal{F}(\tau_{k-1}, \theta_{\mathrm{k-1,i}}) - \hat{\theta}_{\mathrm{model},k-1}(t)\right),$$

where λ_{FF} as the retention rate and γ_{FF} is the learning rate. Note that $\mathcal{F}(\tau_{k-1}, \theta_{\mathrm{k-1,i}})$ in the equation represents the forward dynamics of the system initialised with initial conditions of the previous trial and the entire torque vector of the previous trial as the input, and Γ represents the inverse dynamics of the system, and the entire trajectory vector (including positions, velocities, and accelerations

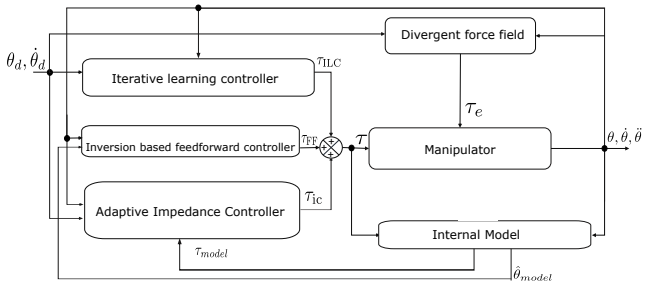

Figure 1: Schematic block diagram representing the control logic. The total control torque $\tau_k(t)$ consists of $\tau_{\text{ILC},k}(t)$ which is the torque from iterative learning control, $\tau_{\text{FF},k}$ which is the torques from feed-forward control and $\tau_{\text{ic},k}$ which is a torque from impedance controller. K_k and D_k are the impedance stiffness and damping matrices that adapt in the iteration domain (over repeated trials) along with an iterative update of an internal model representation marked as 'Model'.

of joint variables) are used as input to inverse dynamics for generating torques that realize the commanded trajectory. This work assumes that the system dynamics are known while perturbations to the system are unknown. If the system dynamics are unknown, online parameter identification algorithms can be utilized [16].

$\hat{\theta}_k$ is the augmented joint variables vector defined by $\hat{\theta} = [\theta_{1,k}, \theta_{2,k}, \dot{\theta}_{1,k}, \dot{\theta}_{2,k}, \ddot{\theta}_{1,k}, \ddot{\theta}_{2,k}]$. Similar to human-motor-learning literature, we propose the following internal model adaptation equations along the lines of the previous two update equations

$$\hat{\theta}_{\text{model},k}(t) = \lambda_{\text{model},\theta}\hat{\theta}_{\text{model},k-1}(t) + \gamma_{\text{model},\theta}(\hat{\theta}_{k-1}(t) - \hat{\theta}_{\text{model},k-1}(t)), \tag{4}$$

and

$$\begin{aligned}\tau_{\text{model},k}(t) = \lambda_{\text{model},\tau}\tau_{\text{model},k-1}(t) + \gamma_{\text{model},\tau} \\ \times \left(\tau_{k-1}(t) - \tau_{\text{model},k-1}(t)\right).\end{aligned} \tag{5}$$

Here the corresponding λ_{model} and γ_{model} are retention and learning rates, respectively. These internal model updates are consistent with observations in human-motor-learning literature that there is evidence that humans continually adapt their internal model representations [12]. These internal models act as the diffuse memory of the tasks and filter out any sudden changes of input torques and measured joint space variables.

The difference between $\hat{\theta}_{k-1}$ and $\hat{\theta}_{\text{model},k-1}$ acts as the input to the inverse dynamics model of the robot denoted by Γ to generate the required adaptation term for the feedforward torques. The forward dynamics of the robot provide the anticipated effect of the control input on the system outputs. Therefore the error between the anticipated outputs and internal model outputs drives the adaptation in the feedforward control inputs.

Before we proceed further, we first note that all four update equations (3) - (5) have a simple and intuitive update law for online adaptation. As highlighted earlier, several works utilize simple updates (particularly ILC) for various robot control problems. However, as highlighted in [6], such approaches fail to stabilize the system in unstable interaction dynamics. It has been well established that explicit impedance adaptation is needed to stabilize such unstable interaction dynamics. We thus arrive at the last two terms in (2). As is the case in humans and many robots, we consider the stiffness and damping matrices to consist of two parts, active and passive stiffness, and damping. The active stiffness and damping adapt to counter the disturbances over trials where passive stiffness and damping remains constant. Thus, the total stiffness K is given by,

$$K = K_a + K_p, \tag{6}$$

where K_a and K_p respectively denote the active and passive stiffness. Also, the total damping is given by,

$$D = D_a + D_p, \tag{7}$$

where the active and passive damping are denoted by D_a and D_p respectively. The update in the active stiffness and damping is again chosen to be a straightforward update with

$$K_{a,k} = \lambda_k K_{a,k-1} + \gamma_k \Delta K_{k-1}, \tag{8}$$

$$D_{a,k} = \lambda_d D_{a,k-1} + \gamma_d \Delta D_{k-1}, \tag{9}$$

where

$$\Delta K_k = (\theta_{d,k-1}(t) - \theta_{k-1}(t))^{\dagger}(\tau_{k-1}(t) - \tau_{\text{model,k}}(t)), \tag{10}$$

$$\Delta D_k = (\dot{\theta}_{d,k-1}(t) - \dot{\theta}_{k-1}(t))^{\dagger}(\tau_{\text{k-1}}(t) - \tau_{\text{model,k}}(t)), \tag{11}$$

and † is Moore–Penrose inverse. The controller equations are represented in a schematic block diagram as shown in Figure 1. It is also worth noting here that [6] considered mean force over the past several trials (also requiring memory of actions from past several trials) to find the update of the stiffness matrix. The current set of proposed controller equations does not suffer from these limitations.

Algorithm 1 presents the pseudocode for implementing the controller equations in a trial domain. In the above algorithm, N denotes the total number of trials. The first stage of the algorithm involves updating the active stiffness and damping terms for the subsequent trial based on the position and velocity errors of the previous trial using the Arimoto algorithm. The next stage of the algorithm involves updating the internal model used for adapting the feedforward torques. Finally, the iterative learning, feedforward, and impedance controller are updated to generate the torque input required for the subsequent trial.

4 SIMULATION STUDY: PLANAR RR ELBOW MANIPULATOR

We now apply the controller equations to a common robotic system - a planar two-link elbow manipulator. The dynamics of the robot is expressed in standard manipulator form (1) with $M \in \mathbb{R}^{2\times2}$, $C \in \mathbb{R}^{2\times2}$ and $G = 0$ (assuming horizontal plane operations). θ consists of the two joint angles and τ consists of the two joint torques applied to the system. The details of the parameters used in the simulations are provided in appendix. The system is fully actuated and the disturbance acting on the system is chosen to be an unstable force-field disturbance given as

$$\tau_{\text{e}}(t) = -\mathbf{J}(\theta(\mathbf{t}))^{\intercal}K_{\text{u}}(X_d(t) - X(t)), \tag{12}$$

where $\mathbf{J}(\theta(\mathbf{t})) \in \mathbb{R}^2$ is the velocity Jacobian matrix, $\mathbf{K}_{\text{u}} \in \mathbb{R}^2$ is a positive definite matrix such that $-K_{\text{u}}(X_d(t) - X(t))$ results in a repulsive (divergent) force field pointing away from the desired

Algorithm 1: Iterative Impedance Controller

input : $\theta_d(t), \dot{\theta}_d(t), \ddot{\theta}_d(t), K_\mathrm{p}, D_\mathrm{p}, N$
output: τ_k

for $k = 1$ **to** N **do**

———— Stiffness Update ————

$\Delta K_k = (\theta_d(t) - \theta_{k-1}(t))^\dagger (\tau_{k-1}(t) - \hat{\tau}_{\mathrm{model},k-1}(t))$

$\Delta D_k = (\dot{\theta_d}(t) - \dot{\theta}_{k-1}(t))^\dagger (\tau_{k-1}(t) - \hat{\tau}_{\mathrm{model},k-1}(t))$

$K_{\mathrm{a},k} = \lambda_{ka} K_{\mathrm{a},k-1} + \gamma_{ka} \Delta K_{k-1}$

$D_{\mathrm{a},k} = \lambda_{da} D_{\mathrm{a},k-1} + \gamma_{da} \Delta D_{k-1}$

$K_k = K_\mathrm{p} + K_{\mathrm{a},k}$

$D_k = D_\mathrm{p} + D_{\mathrm{a},k}$

———— Model Update ————

$\hat{\theta} = [\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2, \ddot{\theta}_1, \ddot{\theta}_2]$

$\hat{\tau}_{\mathrm{model},k}(t) = \lambda_{\mathrm{IM}} \hat{\tau}_{\mathrm{model},k-1}(t) + \gamma_{\mathrm{IM}} (\tau_{k-1}(t) - \hat{\tau}_{\mathrm{IM},k-1}(t))$

$\hat{\theta}_{\mathrm{model},k}(t) = \lambda_{\mathrm{model}} \hat{\theta}_{\mathrm{model},k-1}(t) + \gamma_{\mathrm{model}} (\hat{\theta}_{k-1}(t) - \hat{\theta}_{\mathrm{model},k-1}(t))$

———— Torque Update ————

$\tau_{\mathrm{ILC},k}(t) = \lambda_{\mathrm{ILC}} \tau_{\mathrm{ILC},k-1}(t) + \gamma_{\mathrm{ILC}} (\theta_d(t) - \theta_{k-1}(t))$

$\tau_{\mathrm{FF},k}(t) = \lambda_{\mathrm{FF}} \tau_{\mathrm{FF},k-1}(t) + \gamma_{\mathrm{FF}} \Gamma(\mathcal{F}(\tau_{k-1}, \theta_{\mathrm{k-1,i}}) - \hat{\theta}_{\mathrm{model},k-1}(t))$

$\tau_{ic}(t) = K_k(\theta_d(t) - \theta(t)) + D_k(\dot{\theta_d}(t) - \dot{\theta}(t))$

$\tau_k = \tau_{\mathrm{ILC},k} + \tau_{\mathrm{FF},k} + \tau_{\mathrm{ic}}$

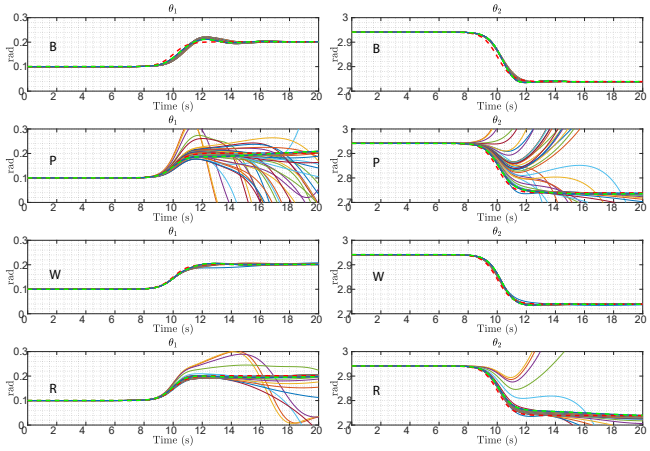

Figure 2: Evolution of the joint-space variable θ_1 and θ_2 over repeated trials in the presence of an active impedance, iterative learning, and feedforward controller. The red dotted line represents the reference trajectory, and the green dotted line represents the output trajectory at the end of the trails of a particular phase. The simulation consists of four phases, including baseline (B), perturbation (P), washout (W) and re-perturbation (R). In the baseline and washout phase, the external unstable force field is off, and in the perturbation and the re-perturbation phase, it is on. The controller is able to learn and achieve trajectory tracking in the absence of the diverging external force field (in presence of noise and disturbances) in B and W trials. The controller is however able to adapt its impedance to stabilize and achieve tracking in P and R trials.

Cartesian trajectory $X_d(t)$ with $X(t)$ being the actual robot trajectory. The Jacobian transformation $(-J(\theta(t))^\intercal)$ transforms this force field to the joint space. In addition to the force field (12), zero-mean normally distributed disturbance torque is added to the simulation. The other controller equations (2) - (11) remain the same.

5 RESULTS AND DISCUSSION

To study the performance of the synthesized controller, we adopt a standard paradigm used in human motor learning experiments with a sequence of a number of repeated trial in four sessions, namely; baseline [B] (external force field is off), perturbation [P] (the field of external force is activated), washout [W] (the external force field is deactivated) and re-perturbation [R] (the external force field is again activated). We choose 50 repeated trials of a simple reaching task in each of these four sessions. Following this paradigm, we observe stabilization and adaptation patterns, along with presence of savings and retention characteristics (akin to human motor learning).

The controller demonstrates learning behavior in the baseline phase without an external diverging force field to reduce trajectory errors that arise due to the unknown internal models, as shown in Figs. 5. In Fig. 2, the joint variables θ_1 and θ_2 converge to the desired trajectories. This behavior is evident in Fig. 3 as the Cartesian error remains almost zero during the trials of the baseline phase, which is consistent with the human motor learning literature. Another key point to note is that there is no significant adaptation of active stiffness and damping parameters of the mechanical impedance of the robot in the baseline phase. Fig. 5 provides further evidence to corroborate this fact, where one can observe that the norm of the stiffness matrix increases slightly, followed by a reduction in the norm of the damping matrix over the trials in the baseline phase. One can observe the final stiffness and damping ellipsoids in Fig. 6, where there has been minor adaptation of the impedance parameters highlighting the aforementioned fact.

In the perturbation phase, the robot may initially appear unstable, as there has been no significant adaptation of active stiffness and damping parameters in trials of the baseline phase. As the trials progressed, the robot learned the required impedance for the reaching task due to the presence of the adaptive impedance controller. The above observation is evident in Fig. 2, where the joint-space trajectories are almost similar to the desired ones over the last few trials of the perturbation phase. One can also observe this in Fig. 3, where the last five Cartesian trajectories are almost identical to the reaching task indicating the adaptation of the stiffness and damping parameters of the robot. Fig. 5 elaborates on the aforementioned fact, further demonstrating the increase in the norm of the damping and stiffness matrices. One interesting point to observe is that the rate of increase in the damping matrix's norm is higher than that of the stiffness matrix. This observation may indicate that damping may play a more significant role in stabilizing the robot in an unstable environment than previously thought. Fig. 6 shows

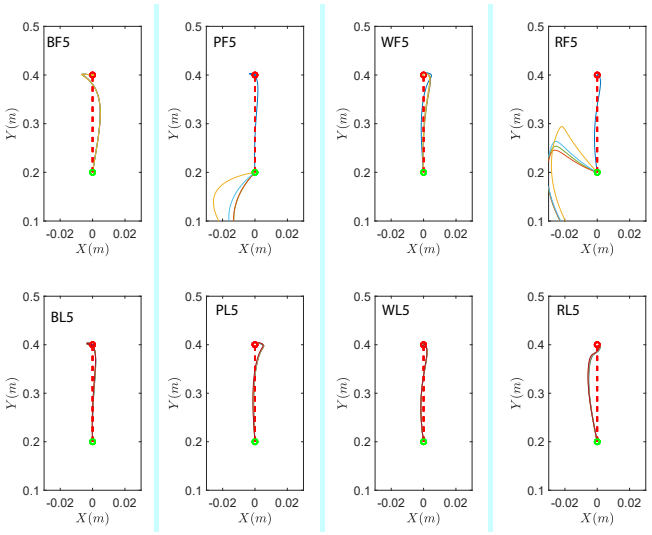

Figure 3: Evolution of the end effector position over repeated trials with the presence of an active impedance controller. Here, *BF*5, *BL*5, *PF*5, *PL*5, *WF*5, *WL*5, *RF*5, *RL*5 stand for the first and last five trials of the baseline (B), perturbation (P), washout(W), reperturbation (R) respectively. The red dotted-line represents the reference trajectory for the reaching task and the green dotted-line represents output trajectory at the end of the trials of particular segment of a particular phase, i.e. the first five or the last five. The system demonstrates stable behavior and reaches the endpoint in the absence and presence of the diverging external force field with variable initial conditions and noise.

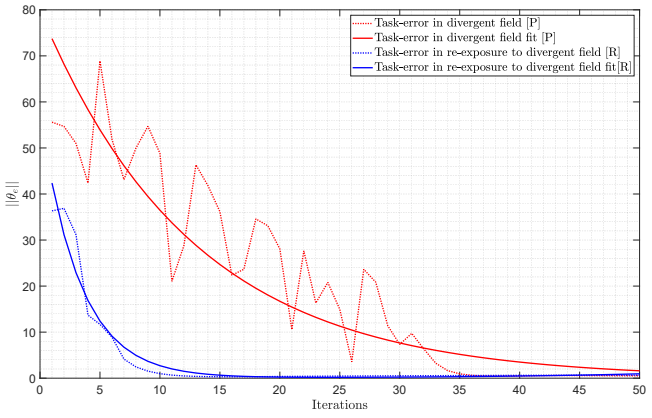

Figure 4: Evolution of error in perturbation and re-perturbation exhibits faster relearning and reduced initial error in the re-perturbation phase compared to the perturbation phase. Further second-order exponential fit also exhibits savings and retention.

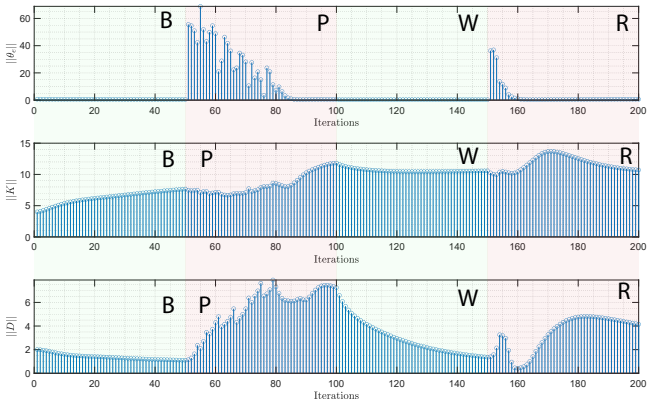

Figure 5: Norm of tracking error and adaptation of stiffness and damping. The impedance controller is able to stabilize the unstable dynamics and reduce errors in all four phases. Note that the stiffness adaptation is dominant in the baseline phase while the damping adaptation is dominant in the perturbation phase.

the stiffness and damping ellipsoids at the beginning and end of the perturbation phase. One can observe that the initial stiffness and damping ellipsoids of the perturbation phase are precisely the same as the final stiffness and damping ellipsoids of the baseline phase. Thus, the adaptation in stiffness and damping parameters is carried forward to the subsequent phase, which aids the robot in rejecting environmental disturbances. The feedforward and iterative learning controllers also help the robot reject environmental disturbances encountered in the presence of a stable force field. Due to the retention rate being less than one, the adapted impedance parameters decay over time in the presence of no perturbation. However, as the position and velocity errors increase, there is also an increase in the adaptation of the active stiffness and damping terms over the trials. This significantly improves trajectory tracking performance towards the last five trials of the perturbation phase.

One can observe that there is no significant change in the robot's joint trajectories in the washout phase trials. Fig. 2 highlights that no crucial change in the robot's joint-space trajectories occurs over the washout phase trials. The reason behind no significant change in joint trajectories is that the robot has learned the required torque inputs for the reaching task through the adaptation of the feed-forward and iterative learning control components. Fig. 3 further corroborates this observation, where the washout phase's first and last five trials demonstrate similar task performance. The first panel of Fig. 3 provides further evidence behind the similarities in the robot's joint trajectories, as the norm of the joint space error is almost zero throughout the washout phase. The second panel of Fig. 3 shows a significant decrease in the norm of the damping matrix along with a slight decrease in the norm of the stiffness matrix. The significant reduction in the damping matrix is because the velocity errors are almost zero, implying a lower change in the damping matrix. As the damping matrix is adapted through the Arimoto algorithm, the retention rate being less than one leads to a decrease in the damping matrix over the trials of the washout phase. Fig. 6 further corroborates this result as one can see the de-adaptation

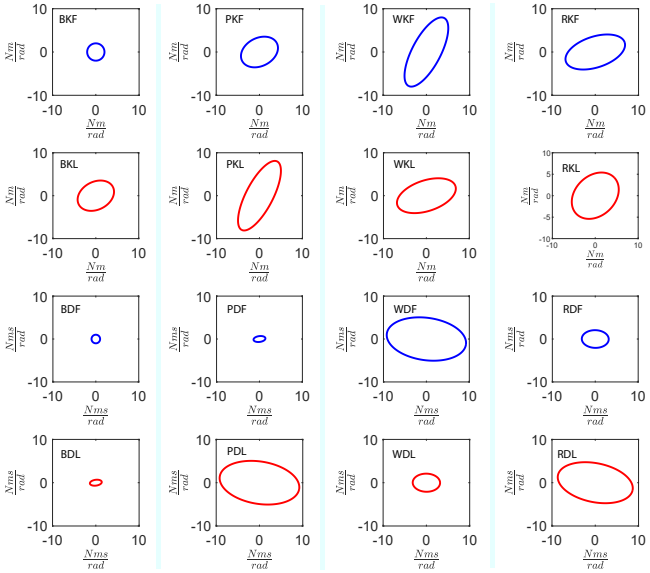

Figure 6: Joint stiffness and damping ellipsoids in all four phases of trials. The blue ellipsoids represent the stiffness and damping ellipsoids at the initial trial of each phase. The red ellipsoids represent the stiffness and damping ellipsoids at the final trial of each phase. One can observe the stiffness and damping ellipsoids at the end of each phase being carried forward to the initial trial of the subsequent phase. The increase in the stiffness and damping ellipsoid at the end of the perturbation and reperturbation phase indicates the increase in the adaptation of the active damping and stiffness terms due to velocity and positional errors respectively. Furthermore, there is a slight decrease in the stiffness ellipsoid of the washout phase followed by a significant decrease in the damping ellipsoid. This is suggestive of the fact that perhaps damping plays a more pivotal role in stabilization in the presence of interactions with unstable environments.

of the stiffness and damping matrices through the initial and final ellipsoids. The damping ellipsoid has decreased significantly in the final trial of the washout phase, along with a slight decrease in the stiffness ellipsoid.

In the re-perturbation phase, one may anticipate that the robot would likely become unstable after the first few trials due to the re-perturbation. The initial decrease in performance is further clarified in Fig. 3, where one can observe a significant decrease in performance while carrying out the reaching task. However, it is closer to the final reaching point compared to the first perturbation phase. One key observation from the first panel of Fig. 3 is the decrease in the norm of the joint error as compared to that in the perturbation phase. The reduction in the initial joint error is known as retention, which is consistent with the human-motor learning literature. In addition, the joint error converges at a faster rate as compared to that in the perturbation phase, as shown in Fig. 4. Faster convergence of the joint error is known as savings, which is another fundamental property of our controller and is consistent with the human-motor learning literature. The reduction in the initial joint error is that the de-adaptation of the stiffness in the washout phase is not significantly large, which is evident in the second panel of Fig. 3. Fig. 6 shows that the final stiffness ellipsoid of the washout and the initial stiffness ellipsoid of the re-perturbation phase are similar. Thus, the stiffness matrix from the final trial of the washout phase is carried forward to the initial trials of the re-perturbation phase, indicating a decrease in the norm of the joint error. The stiffness and damping parameters are then adapted to counter the environmental disturbances from the divergent force field evident in the second and third panels of Fig. 3. Significant adaptation of the stiffness and damping parameters occurs only until the joint error norm becomes zero, followed by a decrease in the stiffness and damping parameters, respectively. This observation is evident from the second and third panels of Fig. 3, showing an increase in the norm of the stiffness and damping parameters followed by a subsequent decrease.

It is also interesting to note from Fig. 5 that the active impedance controller adapts significantly for the stiffness component in the baseline phase while showing minimal adaptation in the damping component. This appears to be driven by a higher magnitude of positional errors than velocity errors in the stable environment. However, it is seen that adaptation in damping is significantly more dominant in the trials in the presence of unstable interaction dynamics, as seen in the perturbation phases. This indicates that the damping term may play a more dominant role than previously thought in stabilizing unstable interaction dynamics. Perhaps velocity errors have relatively higher magnitudes than positional errors in the unstable environment, leading to greater damping changes in the perturbation and re-perturbation phase. On further consideration, one may also speculate that significant adaptation in damping in the presence of unstable force field is perhaps to be expected. One may argue that velocity errors have the ability to anticipate positional errors and damping has the role of taking preventive action before positional errors increase. In contrast, the stiffness works as a corrective action once errors manifest in position. This argument is also obvious once we note that the position is integral of velocity.

6 CONCLUSION

In this work, we developed a learning control algorithm for robots to autonomously learn impedance and feedforward torque for a given task through repeated interactions with a stable or unstable environment without using an explicit force-based or impedance-based strategy. Our control algorithm is inspired by human motor learning and has several notable features, including simplicity of the control update equations and the incorporation of savings and retention from human motor learning. The proposed method improves performance in tasks involving impedance modulation, as demonstrated by simulation results with a two-link serial chain manipulator. Overall, this work makes a significant contribution to the field of robot learning control by combining ideas from iterative learning control, feedforward control, and neuroscience to synthesize a controller that incorporates higher-level human motor learning features.

REFERENCES

[1] Fares J Abu-Dakka and Matteo Saveriano. 2020. Variable Impedance Control and Learning—A Review. *Frontiers in Robotics and AI* 7 (2020).

[2] Hayder FN Al-Shuka, Steffen Leonhardt, Wen-Hong Zhu, Rui Song, Chao Ding, and Yibin Li. 2018. Active impedance control of bioinspired motion robotic manipulators: An overview. *Applied bionics and biomechanics* 2018 (2018).

[3] Suguru Arimoto, Sadao Kawamura, and Fumio Miyazaki. 1984. Bettering operation of robots by learning. *Journal of Robotic systems* 1, 2 (1984), 123–140.

[4] Jonas Buchli, Evangelos Theodorou, Freek Stulp, and Stefan Schaal. 2011. Variable impedance control a reinforcement learning approach. *Robotics: Science and Systems VI* 153 (2011).

[5] Etienne Burdet, Rieko Osu, David W Franklin, Theodore E Milner, and Mitsuo Kawato. 2001. The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature* 414, 6862 (2001), 446–449.

[6] Etienne Burdet, Keng Peng Tee, I Mareels, Theodore E Milner, Chee-Meng Chew, David W Franklin, Rieko Osu, and Mitsuo Kawato. 2006. Stability and motor adaptation in human arm movements. *Biological cybernetics* 94, 1 (2006), 20–32.

[7] Cosimo Della Santina, Matteo Bianchi, Giorgio Grioli, Franco Angelini, Manuel Catalano, Manolo Garabini, and Antonio Bicchi. 2017. Controlling soft robots: balancing feedback and feedforward elements. *IEEE Robotics & Automation Magazine* 24, 3 (2017), 75–83.

[8] Sujay D. Kadam, Shail Jadav, Anadi Mehta, and Harish J. Palanthandalam-Madapusi. 2022. A Model-based Feedforward and Iterative Learning Controller Exhibiting Features of Human Motor Learning. *IEEE Transactions on Cognitive and Developmental Systems* 14 (2022). Under review.

[9] Sujay D. Kadam and Harish J. Palanthandalam-Madapusi. 2021. Trackability for Discrete-Time Linear Time-Invariant Systems: A Brief Review and New Insights. *Journal of Dynamic Systems, Measurement, and Control* 144, 3 (12 2021). https://doi.org/10.1115/1.4052918 arXiv:https://asmedigitalcollection.asme.org/dynamicsystems/article-pdf/144/3/031007/6818642/ds_144_03_031007.pdf 031007.

[10] Sujay D Kadam and Harish J Palanthandalam-Madapusi. 2022. A note on invertibility and relative degree of MIMO LTI systems. *IFAC Journal of Systems and Control* 20 (2022), 100193.

[11] Mitsuo Kawato. 1999. Internal models for motor control and trajectory planning. *Current opinion in neurobiology* 9, 6 (1999), 718–727.

[12] John W Krakauer, Alkis M Hadjiosif, Jing Xu, Aaron L Wong, and Adrian M Haith. 2019. Motor learning. *Comprehensive Physiology* 9, 2 (2019), 613–663.

[13] Jianlan Luo, Eugen Solowjow, Chengtao Wen, Juan Aparicio Ojea, Alice M Agogino, Aviv Tamar, and Pieter Abbeel. 2019. Reinforcement learning on variable impedance controller for high-precision robotic assembly. In *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 3080–3087.

[14] Kevin L Moore, YangQuan Chen, and Hyo-Sung Ahn. 2006. Iterative learning control: A tutorial and big picture view. In *Proceedings of the 45th IEEE Conference on Decision and Control*. IEEE, 2352–2357.

[15] Masahide Oikawa, Kyo Kutsuzawa, Sho Sakaino, and Toshiaki Tsuji. 2020. Assembly robots with optimized control stiffness through reinforcement learning. *arXiv preprint arXiv:2002.12207* (2020).

[16] Abdelhamid Tayebi. 2004. Adaptive iterative learning control for robot manipulators. *Automatica* 40, 7 (2004), 1195–1203.

[17] Tasuku Yamawaki, Hiroki Ishikawa, and Masahito Yashima. 2016. Iterative learning of variable impedance control for human-robot cooperation. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 839–844.

[18] Chenguang Yang, Gowrishankar Ganesh, Sami Haddadin, Sven Parusel, Alin Albu-Schaeffer, and Etienne Burdet. 2011. Human-like adaptation of force and impedance in stable and unstable interactions. *IEEE transactions on robotics* 27, 5 (2011), 918–930.

7 APPENDIX

The retention rates and learning rates used for simulation are mentioned in Table 1. The parameters used in two-link elbow manipulator dynamics are as follows: $M = \begin{bmatrix} M_{11} & M_{12}\ M_{21} & M_{22} \end{bmatrix}, C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$, *and* $\tau = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix}$,where $M_{11} = m_1 l_{c1}^2 + m_2(l_1^2 + l_{c2}^2 + 2l_1 l_{c2}^2 + 2l_1 l_{c2}\cos(\theta_2))$,

$$M_{12} = M_{21} = m_2(l_{c2}^2 + l_1 l_{c2} cos(\theta_2)) + I_2, M_{22} = m_2 l_{c2}^2 + I_2$$

$C_{11} = -l_1 l_{c2} m_2 \sin(\theta_2)\dot{\theta_2}$, $C_{12} = -l_1 l_{c2} m_2 \sin(\theta_2)\dot{\theta_2} - l_1 l_{c2} m_2 \sin(\theta_2)\dot{\theta_1}$, $C_{21} = l_1 l_{c2} m_2 \sin(\theta_2)\dot{\theta_1}, C_{22} = 0$. Here m_1, m_2, l_1, l_2, l_{c1}, l_{c2} are the masses, lengths, and distance of centre of mass from the joints for link-1 and link-2 respectively. τ_1, τ_2 are the joint torques and θ_1, θ_2 are joint angles. For the simulation study we consider unit mass of 1 kg for link-1 and link-2 with unit link lengths of 1 m for both links. The passive stiffness and damping matrix along with the environmental stiffness matrix used in the simulations are as follows $K_p = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $D_p = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. $K_u = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$,

Table 1: Learning and Retention rates

$\lambda_{ka} = 0.95$	$\gamma_{ka} = 0.6$
$\lambda_{da} = 0.95$	$\gamma_{da} = 0.6$
$\lambda_{\text{IM}} = 0.9$	$\gamma_{\text{IM}} = 0.1$
$\lambda_{\text{model}} = 0.5$	$\gamma_{\text{model}} = 0.1$
$\lambda_{\text{ILC}} = 0.7$	$\gamma_{\text{ILC}} = 0.1$
$\lambda_{\text{FF}} = 0.7$	$\gamma_{\text{FF}} = 0.1$