

Modeling Previous Trial Effect in Human Manipulation through Iterative Learning Control

Lorenzo Cenceschi, Cosimo Della Santina,* Giuseppe Averta, Manolo Garabini, Qiushi Fu, Marco Santello, Matteo Bianchi, and Antonio Bicchi

In the execution of repetitive tasks, humans can capitalize on experience to improve their motor performance. Prominent examples of this ability can be recognized in our capacity of grasping and manipulating in uncertain conditions. With the aim of providing a mathematical description for such behavior, experiments are considered where participants are required to lift an object with an unexpected mass distribution. By repeating multiple times the same lifting action, participants can learn the correct motor command for task accomplishment. Three models are proposed that combine reactive terms and a learned anticipatory action to explain experimental data. The models feature intratrial and intertrial memory, and the effect of slowly and fast adaptive sensory receptors. The architectures' effectiveness in explaining experimental data is compared with a general-purpose state of the art model. The proposed algorithms conspicuously outperform the state of the art in all the considered validation routines. Global and within-trial human behavior is predicted with 88% of accuracy in nominal conditions. When the object's center of mass is moved, the accuracy is maintained up to 83%. Finally, convergence properties of proposed algorithms are analytically discussed, and their stability and robustness against measurement noise are evaluated in simulation.

sensory-motor memory from previous repetitions of the task. This learning behavior was first observed in human upper limb actions, and described in seminal papers, such as refs. [1,2]. In these manuscripts, authors report on experiments involving subjects whose reach-to-grasp movements are perturbed by external force fields. Subjects are capable to learn how to efficiently counteract different perturbations to perform the same kinematic task. This capability is ascribed to the nervous control system as a sum of three components: an internal model for skeletal system movements without perturbation, an internal model for perturbations, and a feedback control to stabilize trajectory, if any error occurs. Internal models were postulated by Kawato and Wolpert in a series of seminal papers.^[3–6] These models are internal representations of the body and environment dynamics, to be used by the central nervous system (CNS) for evaluating the motor command needed to perform an action. These representations are full-fledged models in the original

definition. However, they can also be thought as specific patterns of motor actions necessary to produce a certain behavior (e.g., dynamic primitives^[7,8]).


The matter of which kind of signal is used by the CNS to be fed into the internal model learning mechanism has been largely

1. Introduction

A key characteristic that makes humans capable of complex manipulation actions is the ability of learning compensatory actions in response to external uncertainties, relying on

L. Cenceschi, G. Averta, Prof. M. Garabini, Prof. M. Bianchi, Prof. A. Bicchi
 Research Center "Enrico Piaggio"
 University of Pisa
 Largo Lucio Lazzarino 1, Pisa 56126, Italy

Dr. C. Della Santina
 Computer Science and Artificial Intelligence Laboratory (CSAIL)
 Massachusetts Institute of Technology
 32 Vassar Street, Cambridge, MA 02139, USA
 E-mail: cosimodellasantina@gmail.com

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aisy.201900074>.

© 2020 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.201900074

G. Averta, Prof. M. Garabini, Prof. M. Bianchi, Prof. A. Bicchi
 Department of Information Engineering
 University of Pisa
 Via G. Caruso 16, Pisa 56122, Italy

Prof. Q. Fu
 Department of Mechanical and Aerospace Engineering
 University of Central Florida
 12760 Pegasus Dr., Orlando, FL 32816, USA

Prof. M. Santello, Prof. A. Bicchi
 School of Biological and Health Systems Engineering
 Ira A. Fulton Schools of Engineering
 Arizona State University
 Tempe, AZ, USA

Prof. A. Bicchi
 Istituto Italiano di Tecnologia
 Via Morego 30, Genoa 16163, Italy

investigated in the literature.^[6] One popular candidate is the control signal generated by some sensor-based reactive actions. As also discussed in refs. [3–5], the reactive motor actions (which can be assimilated to what in control language is defined as *feedback*) are identified as a source of information for module updates and adaptation mechanisms. We refer to this hypothesis as *control-based learning*. Note that this hypothesis is sometimes referred in the literature as feedback-error learning. We do not use this terminology to avoid confusions with other hypotheses. Other works^[9,10] introduce the hypothesis that adaptation is driven by the difference between expected and actual outcome of a motor command. This is the error signal in control language. We therefore refer to this hypothesis as *error-based learning*. The authors of ref. [11] propose a learning mechanism emerging from the model of the Purkinje cells in the cerebellum. Learning is guided by outcome error only, without taking into account reactive motor actions from previous trials.

In the study of human manipulation, the learning of a proper motor action through repetition of the same task is called *previous trial effect*. Some of the authors of the current article experimentally investigated this behavior in refs. [12–15]. In these experiments, subjects were asked to lift an object with unexpected mass distribution, maintaining it vertical (see **Figure 1**). In the first execution of the task, the compensatory action appears to be mostly reactive due to the lack of previous experience on the object. In successive repetitions, all subjects implicitly learn the interaction force patterns to correctly lift the object without tilting it. Learning capabilities through repetition in manipulation is an open research topic, made challenging by the high number of environmental and human variables. The effect of anticipatory and reactive actions in decision-making for manipulation is discussed in ref. [16]. Digit placement plays a clear role in exerting effective feedback control actions. This is studied when using thumb and index by ref. [15], and when using multiple digits

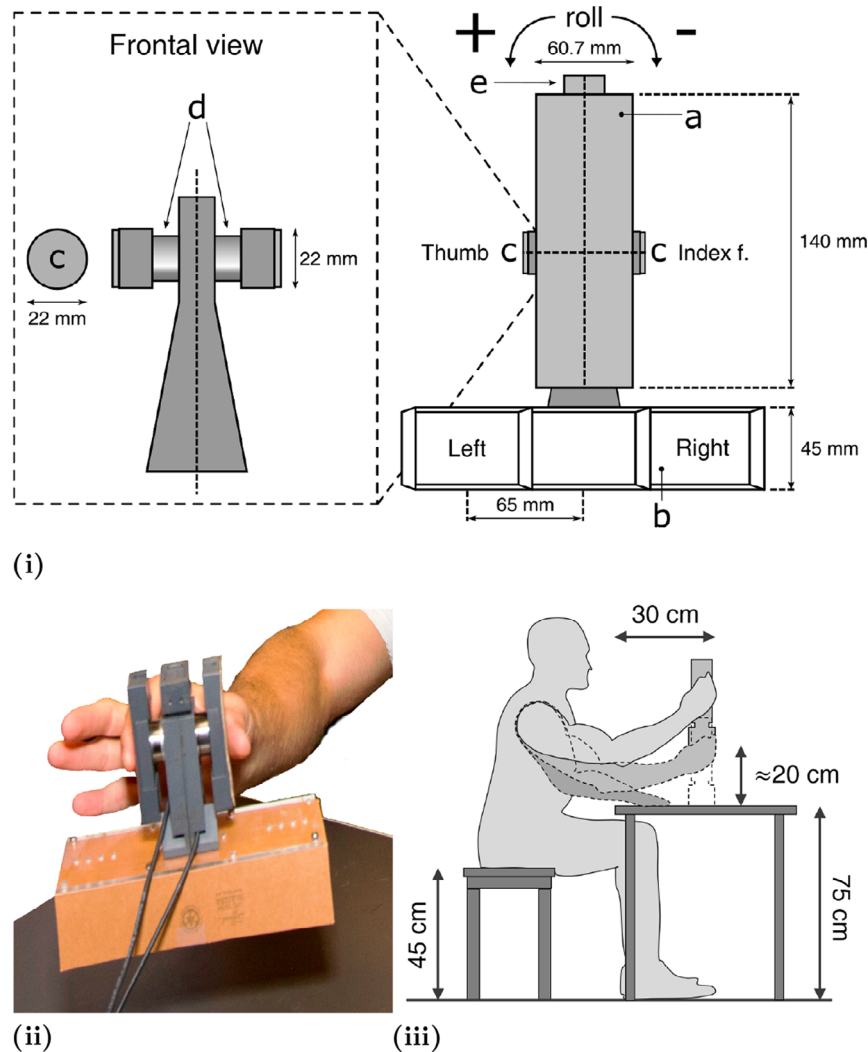


Figure 1. Experimental setup used for human data collection. In panel (i), we present a frontal schematic view of the inverted T device, composed of two parts: a) a handle and b) a slotted box below. c) The subject grasps the object with thumb and index finger at fixed positions, represented by circular plates. d) Two 6-axis force/torque sensors and e) a position sensor to measure the vertical pose of the device were used to collect the experimental data. Panel (ii) depicts the actual device during an experiment. Panel (iii) shows the workspace of the experimental set-up.

in ref. [17]. Anticipatory control affects initial digit position and task performance as well. In case of objects with unknown mass distribution, subjects prefer to put digit in standard position.^[18,19] Suggests there is no significant difference between dominant and nondominant hands when learning a manipulation task with unfamiliar objects. The role of dynamic primitives in manipulation of objects with complex dynamics is discussed in ref. [20].

Despite this lively research activity, no mathematical description of previous trial effect is, as of today, available in the literature. The aim of this work is to provide such characterization. More specifically, we formulate and test the hypothesis that *previous trial effect* can be described through a proper combination of linear iterative learning control^[21,22] and delayed linear feedback.

1.1. Proposed Approach

In an attempt to model both intratrial memory and the physiological role of reactive mechanoreceptors,^[23] we included in the feedback law also an integral and a derivative action. The first term accounts for human ability of reducing error also when no knowledge of the system is available from previous executions of the task. This means having no feed-forward compensatory action available, according to our model. For what concerns the latter action, deriving the signal produces a high pass effect, which produces a signal resembling the kind of high-frequency characteristics that it is provided by reactive mechanoreceptors.^[24] Furthermore, derivative action accounts for the ability observed in our subjects to damp oscillations. Iterative learning control (ILC hereinafter) theory is used to model intertrial memory. The error evolution during the whole lifting action is used here to update the anticipatory motor command used in executing the following repetition of the task. We introduce three control models implementing these principles, each corresponding to a different neuroscientific hypothesis on which learning signal is used to drive the anticipatory action: a control-based learning algorithm, an error-based learning algorithm, and a model combining the two. Note that no explicit internal model is trained here. We instead focus on trial-by-trial approach. Instead of generating a desired control using reverse dynamics, previous trial control signal is used and improved.

We identify the free parameters in our models using the experimental data collected in ref. [14]. See Section 2.1 for more details on the dataset. Convergence characteristics of the proposed controllers are analytically discussed, leveraging on ILC literature. Then, we test the capability of the three models to reproduce human learning behavior, and the temporal evolution of the torques generated within each trial. We are not aware of any previous work in the literature tackling the mathematical description of this behavior in human grasping and manipulation tasks. We consider as benchmark one of the few analytical descriptions of human learning by repetition, introduced in ref. [9]. There, subjects were asked to walk on a treadmill, with an unknown force field applied to their ankle by a robot. The autoregressive model they propose is able to correctly describe the adaptation of motor action across steps. We opportunely modify this algorithm to tailor it to the problem under exam, and we use it as a benchmark.

Three validation tests are considered. The first evaluates the ability of the proposed models to predict motor actions, when the mass distribution is the same as in the identification dataset. The second validation targets the case of different mass distribution between the identification and the testing phases. Finally, the stability and robustness of the proposed strategies against sensory noise are also evaluated in simulation. Performance indexes include normalized root mean square error (RMSE) (between experiment and model data) computed across all subjects, and Bayesian Information Criterion (BIC), to evaluate model complexity. Our findings suggest that a model combining control- and error-based learning is the best choice to describe human control actions, according to both the performance indexes.

The work is organized as follows. In Section 2, experimental data collection is reported and qualitatively discussed, a mathematical formalization is introduced, and the problem is stated. In Section 3, we present our three architectures. Section 4 analytically describes the convergence properties of the proposed models. In Section 5, we introduce the identification and validation procedures, and we present their outcomes. Section 7.2 reports a comparative study of closed loop accuracy, performed in a simulative environment.

2. Problem Definition

2.1. Experimental Task

The previous trial effect behavior we consider in this work was observed in the experiment described in ref. [14]. The experiments involving human subjects have been performed with the full, informed consent of the volunteers. Data from nine right-handed volunteers (aged 20–26 years) were collected. Participants had no history of musculoskeletal or neurological disorders, and they were all naive to the experimental purpose of the study. Without loss of generality, in this work we focus on the learning process for force generation. For this reason, hereinafter we will consider the outcomes of the experiments where finger positions are constrained, as explained later and shown in Figure 1.

The device used for the manipulation task is shown in Figure 1i. It is an “inverted T”-shaped object consisting of a sensorized graspable handle and a bottom box with “left” and “right” slots. A 0.400 kg load is hidden in one of the slots (either right or left), thus resulting in a change in the device’s center of mass (CM) and mass symmetry. The total device mass is 0.396 kg. With the addition of the hidden load, it reaches 0.796 kg.

Two force/torque sensors are placed below the plates where the subjects were instructed to place thumb and index (d in Figure 1i), records the applied digit forces while an electromagnetic position/orientation sensor (Polhemus Fastrack, orientation resolution of 0.05° , e in Figure 1i) was used to record the orientation (roll) of the device and the elevation along the vertical axis, as well as the horizontal displacement w.r.t. the initial position.

2.2. Protocol

Subjects were asked to perform two distinct series of ten consecutive trials, one series per hidden load position. The sequence of load placement was randomized.

In each trial, the subject comfortably seated in front of the table as in Figure 1iii, with the device placed at 30 cm from the hand rest position and aligned with the subject's right shoulder. For each trial, the subject was asked to 1) reach, grasp, and lift the object at a natural speed, with the only thumb and index fingertips placed on the plates; 2) lift the object vertically to a comfortable height of 15 – 20 cm maintaining its vertical alignment; 3) hold it for about 1 s and replace it on the table.

Subjects had no previous experience on the device mass distribution before any series of trials. Changes to the object's CM were performed out of view, to prevent subjects from anticipating object's CM location. Therefore, for the first trial of a series, they relied only on vision, touch, and proprioception with no other nonimmanent feedback information about the device. Subsequent trials were used by subjects to learn how to cope with magnitude and direction of the external torque caused by the added mass. Experimental outcomes show that participants were able to reduce the device peak roll error by trial 3, following the pattern shown in Figure 2. Note that during the first trial the error (in terms of roll, i.e., deviation from the vertical axis) is high because subjects' action is entirely reactive. The exponential decay of the curve suggests an underlying first-order linear learning mechanism.

2.3. Mathematical Formalization

Given the task requirement of object roll minimization, we study here the 2D problem on the vertical plane. Figure 3 shows the corresponding free-body diagram. We consider as control input the forces exerted at the fingertips, assuming a potential deviation of the contact centroid w.r.t. the circular plates of Figure 1. (Note that our approach does not assume which variables are directly controlled or represented by the Central Nervous System. For a discussion about this topic, the interested reader can refer to refs. [25,26].)

Each digit exerts an independent force distribution on the object, which we express here as the sum of grip forces (*LFO* and *RFO*) and tangential forces (*LFV* and *RFV*) (Figure 3). Note that the contact forces are assumed to be applied in contact points that may be different w.r.t. the geometrical centers of the graspable plates. Indeed, the actual points of application of forces

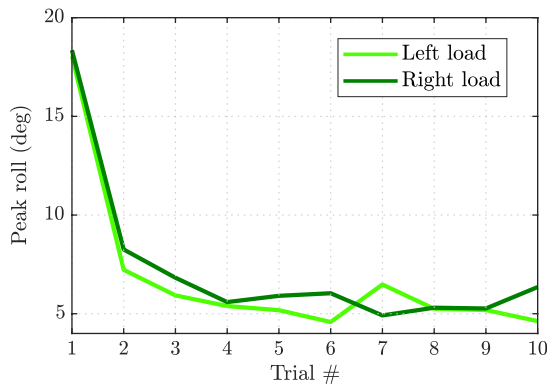


Figure 2. The experimental performance averaged across all subjects and all trials. Trials 1:10, all subjects, left/right hidden load. Average peak roll behavior.

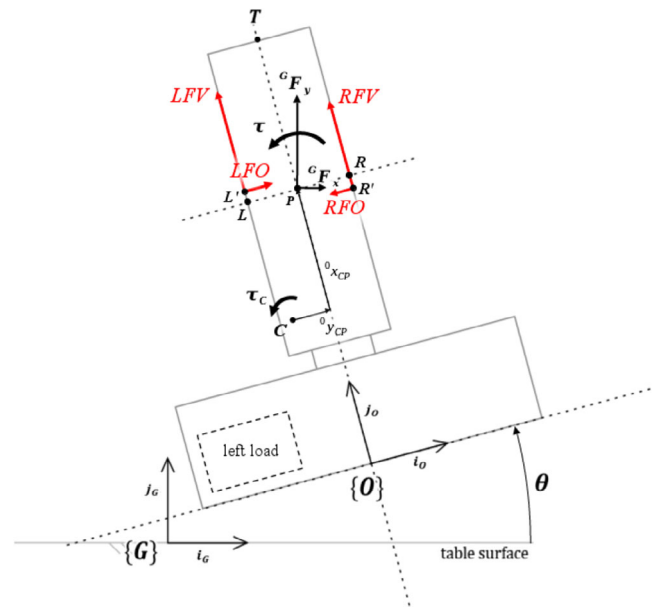


Figure 3. Schematic representation of the exerted forces with the reference frames. $\{O\}$ and $\{G\}$ are, respectively, local and global frames. C is the CM, P the external wrench reference point, L and R are the centers of the graspable plates (which also represents the origin of the reference frames of the two force/torque sensors), while L' and R' are the real digit pressure points, T is the origin of the position sensor. *LFO*, *RFO* are the grip digit forces and *LFV*, *RFV* are the load digit forces..

are free to vary within the circular plate. Therefore, we schematized this assumption in the free-body model through the definition of two contact points, namely, L' and R' , which can in general differ from L and R (i.e., the centers of the circular plates), although this variation is bounded by the radius of the circular plates (11 mm). The relationship between the total wrench evaluated on a fixed point P (Figure 3), and digit forces is

$$\begin{bmatrix} {}^G F_x \\ {}^G F_y \\ \tau \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & -\cos \theta & -\sin \theta \\ \sin \theta & \cos \theta & -\sin \theta & \cos \theta \\ -{}^O \gamma_{PL'} & {}^O x_{PL'} & {}^O \gamma_{PR'} & {}^O x_{PR'} \end{bmatrix} \begin{bmatrix} LFO \\ LFV \\ RFO \\ RFV \end{bmatrix} \quad (1)$$

where $[{}^G F_x \quad {}^G F_y \quad \tau]^T \in \mathbb{R}^3$ is the wrench, in global frame $\{G\}$ components, applied in an arbitrary point $P \in \mathbb{R}^2$; $LFO, LFV \in \mathbb{R}^+$ are, respectively, the grip and load contact forces of the thumb in the pressure centroid $L' \in \mathbb{R}^2$; $RFO, RFV \in \mathbb{R}^+$ are the same forces of the index finger in $R' \in \mathbb{R}^2$; θ is the object roll; $PL = [{}^O x_{PL'} \quad {}^O \gamma_{PL'}] \in \mathbb{R}^2$ and $PR = [{}^O x_{PR'} \quad {}^O \gamma_{PR'}] \in \mathbb{R}^2$ are the distances between the arbitrary point P and the digit contact points L', R' , while the torque applied w.r.t. the CM is

$$\tau_C = [-{}^O \gamma_{CP} \quad {}^O x_{CP} \quad 1] \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^G F_x \\ {}^G F_y \\ \tau \end{bmatrix} \quad (2)$$

From data recorded through the position/orientation sensor, we observed that the translation of the device in the horizontal

line was 1.47 mm (median value, interquartile range 1.01 mm). This supports the consideration that the horizontal displacement of the device during the experiments is negligible, thus enabling the introduction of the following simplifying assumption ${}^G F_x \approx 0$. Equation (2) then becomes

$$\tau_C = \tau + \begin{bmatrix} -{}^O \gamma_{CP} & {}^O x_{CP} \\ \cos \theta \end{bmatrix} {}^G F_y \quad (3)$$

where $\tau_C, \tau \in \mathbb{R}$ are the torques applied, respectively, at the CM $C \in \mathbb{R}^2$ and at the arbitrary external wrench application point P . $CP = \begin{bmatrix} {}^O x_{CP} & {}^O \gamma_{CP} \end{bmatrix} \in \mathbb{R}^2$ is the distance of P from C in local frame components. ${}^G F_y \in \mathbb{R}$ is the global lift force.

The data collected include digit grip forces (LFO, RFO), digit load forces (LFV, RFV), device roll and elevation (θ, h), and digit center of pressure distance ($\Delta L', \Delta R'$) from plates centers (L, R).

2.4. Problem Statement

The aim of this work is to provide an analytical model connecting the entire motor action τ generated by subjects in each iteration, to the entire evolution of the roll error. We are not interested in finding the optimal algorithm in terms of the task to be performed (i.e., object tilting θ stabilization). On the contrary, we are interested in understanding and modeling using our theoretical framework, the mechanism used by the human CNS during the execution of the experimental action under investigation.

This goal can be expressed in compact form as finding an algorithm able to match as close as possible the following relationship, evaluated on k subsequent recorded trials

$$(U^0, \dots, U^{k-1}, E^0, \dots, E^k) \rightarrow U^k \quad (4)$$

where

$$U^k \triangleq \begin{bmatrix} \tau^k[1] & \dots & \tau^k[N] \end{bmatrix} \in \mathbb{R}^{1 \times N} \quad (5)$$

collects the torque samples, $\tau^k[i] \in \mathbb{R}$ is the torque as in Equation (1)–(3), evaluated from the measures LFO, LFV, RFO, RFV at the i th time step; $U^0 = 0$

$$E^k(\Delta) \triangleq \begin{bmatrix} 0 & \dots & 0 & \epsilon^k[1] & \dots & \epsilon^k[N - \Delta] \end{bmatrix} \in \mathbb{R}^{3 \times N} \quad (6)$$

is the error signal delayed of Δ time steps; $E^0(\Delta) = 0, \forall \Delta, N \in \mathbb{N}$ is the number of samples per trial. The parameter Δ quantifies the delay (in terms of number of time frames) between the occurrence of an unwanted behavior (object tilt) and the instant in which the subjects' CNS produces a compensatory response. In other terms, Δ is the sum of a perception, transmission, and processing delays. It is one of the problem-free parameters that need to be identified from data. The task error at trial k is quantified as

$$\epsilon^k(t) = - \begin{bmatrix} \theta^k(t), & \int \theta^k(t), & \dot{\theta}^k(t) \end{bmatrix}^T \in \mathbb{R}^{3 \times 1} \quad (7)$$

where we consider the roll θ^k , its integral, and its derivative. Integral terms are included to account for human ability of reducing the error to zero also in the first iteration (i.e., without

any feed-forward compensatory action), whereas the derivative accounts for the ability to damp oscillations. These signals could be roughly related to slowly adaptive and fast adaptive receptors placed in the fingertip (see, e.g.,^[27]).

The models considered in this work will be identified on a subset of the collected data, and will be tested on other, unseen, datasets. Details on the split of recorded signals are reported in the next sections.

2.4.1. Identification and Validation Sets

We split here the collected data into an identification set \mathbf{I} and two validation sets \mathbf{V} . The first will be used to train the proposed algorithms, and the second to evaluate their ability of predicting human behavior.

For each subject we choose three trials belonging to the same block of trials with a given CM distribution (left or right). In this article, we choose the same trials for every subject as

$$\mathbf{T}_I = \{1, 3, 8\} \quad (8)$$

For each chosen trial, except for the first one, we need current and previous data iterations to account for the entirety of the learning process. The data identification set is

$$\mathbf{I} = \left\{ (U^k, U^{k-1}, E^k(\Delta), E^{k-1}(0)) \mid k \in \mathbf{T}_I \right\} \quad (9)$$

Note that the identification set includes the first trial $1 \in \mathbf{T}_I$. This is done to guarantee enough information during the fitting phase about the purely feedback behavior, when no learning has occurred yet.

Two validation sets are considered. The first one tests predictions within the same learning process (i.e., same mass distribution for the prediction and the identification phase), whereas the second one targets the learning under different mass distributions. The first validation set is composed of all the trials with same mass distribution that are excluded from the identification phase

$$\mathbf{T}_V = (\{1, \dots, 10\} \setminus \mathbf{T}_I) = \{2, 4, 5, 6, 7, 9, 10\} \quad (10)$$

from which it results the following validation set

$$\mathbf{V} = \left\{ (U^k, U^{k-1}, E^k(\Delta), E^{k-1}(0)) \mid k \in \mathbf{T}_V \right\} \quad (11)$$

The second validation set is composed of experiments executed when the load is placed in the left slot while maintaining the identification set with the load on the right. The set of trials used in validation is thus

$$\mathbf{T}_V = \{1, \dots, 10\} \quad (12)$$

The validation set is defined as shown in Equation (11). The aim of this latter validation set is to measure if the proposed mechanisms are independent from the asymmetries of human hand anatomy and control or not.

2.4.2. Performance Evaluation

We consider normalized RMSE across all subjects to measure the ability of an architecture to explain data. We define it as

$$\bar{R} = \frac{1}{S} \sum_{s=1}^S R(s), \quad \text{with} \quad R(s) = \frac{\sqrt{\frac{1}{KN} \sum_{k \in T_V} \|\hat{U}_s^k - U_s^k\|^2}}{\max_{k \in T_V} \hat{U}_s^k - \min_{k \in T_V} \hat{U}_s^k} \quad (13)$$

where $S, K, N \in \mathbb{N}$ are, respectively, the number of subjects, iterations, and collected time samples. $\hat{U}_s^k, U_s^k \in [1, N]$ are, respectively, the measured and identified signals vectors for every s th subject and t th trial.

We introduce also the following BIC^[28] for evaluating the quality of the model also from the point of view of complexity

$$\text{BIC} = \ln \left(\frac{\sum_{s=1}^S \sum_{k \in T_V} \|\hat{U}_s^k - U_s^k\|^2}{KNS} \right) + \kappa \frac{\ln(KNS)}{KNS} \quad (14)$$

where κ is the number of parameters of the model. Lower is the BIC, better is the model in terms of the complexity-accuracy trade-off it represents.

3. Proposed Models of Previous Trial Effect

As stated in Section 1, we use *error-based learning* and *control-based learning* to refer to adaptation, respectively, resulting from reactive motor actions^[3–5] and outcome error.^[9,10]

A first mathematical formalization of the error-based learning theory is given in the form of an autoregressive model in ref. [9]

$$\tau^k(t) = f \tau^{k-1}(t) + \alpha e^{k-1}(t) \quad (15)$$

where f, α are two constants, called forgetting and learning factors, respectively: $\tau^k: [t_0, t_f] \rightarrow \mathbb{R}^m$ and $e^k: [t_0, t_f] \rightarrow \mathbb{R}^m$ are the control action and error evolution at the k th execution of the task, called also iteration hereinafter. $t \in [t_0, t_f]$ is a generic time instant. Note that a similar learning rule also resulted from the direct modeling of human cerebellum in ref. [11]. In ref. [29] we observed that such rule can be seen in the more general theory of Iterative Learning Control (ILC).^[22] Indeed, ILC exploits the error evolution in the whole interval $[t_0, t_f]$ of a previous iteration, to update a feedforward command, according to the law

$$\tau^k = Q(\tau^{k-1}) + R(e^{k-1}) \quad (16)$$

where the learning function $R(\cdot)$ identifies the ILC algorithm and generalizes f of Equation (15), whereas the forgetting function $Q(\cdot)$ maps the old control in the new one generalizing the α in Equation (15).

To extend (16) to the case of control-based learning,^[3,5] we can still leverage on the ILC framework, which allows introducing this effect by considering the current-iteration version of the algorithm. In the general case, this corresponds to substituting $R(e^{k-1})$ with $R(e^k)$.^[22] The introduction of the term e^k considers the current feedback action for the feed-forward estimation in future iterations.

It is therefore natural to introduce a more general hypothesis, combining error- and control-based learning. Leveraging on it we introduce the following general form of motor learning

$$\tau^k(t) = f(e^k(t), e^{k-1}(t), \dots, e^{k-m}(t), \tau^{k-1}(t), \dots, \tau^{k-l}(t)) \quad (17)$$

where $m, l \in \mathbb{N}$ are the number of past iterations used to evaluate the new control action, and $f: \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a generic smooth function combining current and past iteration errors and the past control action.

In Section 2, we observed that the decreasing exponential trend in Figure 2 suggests an underlying first-order linear learning mechanism. (Or, in other terms, average error in Figure 2 can be fitted with a single exponential function.) We propose thus a linear form of (17) for modeling the learning behavior underlying previous trial effect

$$\tau^k(t) = K_{fb} e^k(t - \delta) + K_{ff} e^{k-1}(t) + Q \tau^{k-1}(t) \quad (18)$$

where $\tau^k(t) \in \mathbb{R}$ is the global torque applied in P (as in Equation (3)), $k \in \mathbb{N}$ is the iteration index, $K_{fb}, K_{ff} \in \mathbb{R}^{1 \times 3}$ are the coefficients matrices and $0 < Q \leq 1$ is a linear forgetting factor. The term $\delta \in \mathbb{R}^+$ is the continuous-time counterpart of the quantity Δ introduced in Section 2.4, and is used to account for perception, transmission, and processing delays present in the human control loop,^[30] representing the delay between the instant the subject realizes that an unwanted behavior occurs and when she/he exerts the compensatory response. This value will be identified from data in Section 5. The composite error signal e^k is the task error defined in Equation (7), which is reported here to improve readings

$$e^k(t) = - \left[\theta^k(t), \quad f \theta^k(t), \quad \dot{\theta}^k(t) \right]^T \in \mathbb{R}^{3 \times 1} \quad (19)$$

To recap, in this article we propose three novel control architectures to describe the experimental observation. The first two are built leveraging on preexisting neuroscientific hypotheses: control-based learning and error-based learning. The third one is the mathematical translation of the here introduced generalized hypothesis.

Control-based learning (CL)

In (18), we impose $K_{ff} := 0$, and we keep Q and K_{fb} to be identified. The control rule is

$$\tau^k(t) = K_{fb} e^k(t - \delta) + Q \tau^{k-1}(t) \quad (20)$$

Error-based learning (EL)

The control rule is

$$\begin{cases} \tau^k(t) &= K_{fb} e^k(t - \delta) + \tau_{ff}^k(t) \\ \tau_{ff}^k(t) &= \bar{K}_{ff} e^{k-1}(t) + Q \tau_{ff}^{k-1}(t) \end{cases} \quad (21)$$

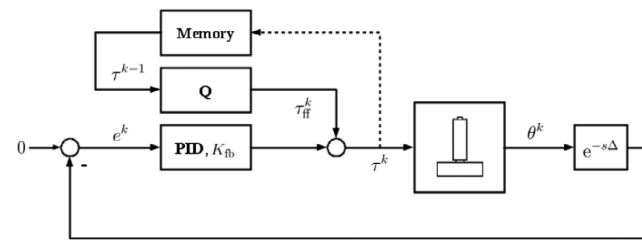
where $K_{fb} e^k(t - \delta)$ is the feedback contribution, $\tau_{ff}^k(t)$ is the feedforward contribution—still applied in P —calculated through the error-based learning rule. Note that if $\delta = 0$, Equation (21) can be rewritten in the generalized form (18) via algebraic manipulation (see Appendix).

Generalized model (GM)

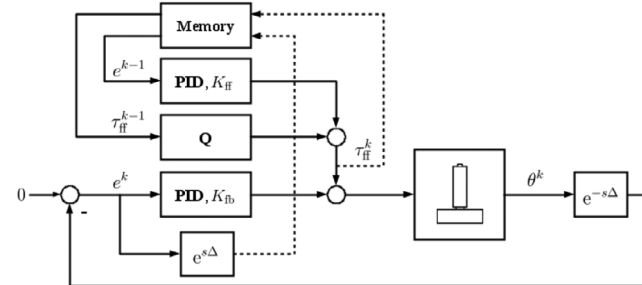
This architecture is based on the full generalized control rule (18), keeping all gains Q , K_{fb} , and K_{ff} to be identified.

Figure 4 graphically shows the three models. It is worth noting that, while two of them are inspired by preexisting neuroscientific literature, the three detailed algorithms are proposed here for the first time.

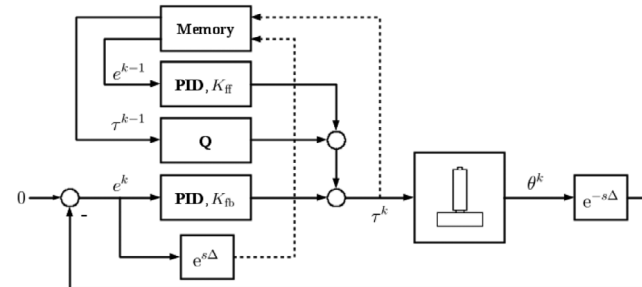
Finally note that due to the larger number of parameters, GM is expected to present equal or better performance w.r.t. the other two proposed architectures in identification phase. We will thus test in the following our architectures using a cross-validation approach, to detect possible occurrences of over-fitting.



(i) Control-based learning



(ii) Error-based learning



(iii) Generalized model

Figure 4. Block diagrams of the proposed learning models. The first two models use either the motor command (i) or the error signal (ii) to drive the learning processes. In (iii) we consider the simultaneous learning of control action from error and control signals. The perception delay Δ is shown with a dedicated block. The anticipation performed before storing the error signal is shown too. Note that—despite the proposed architectures are general in their formulation—the reference tilting angle considered here is zero.

4. Convergence Properties of the Proposed Models

The dynamic model of the inverted T object in Figure 3 is

$$\begin{cases} J\ddot{\theta} &= ({}^Ox_{CP} \cos \theta - {}^Oy_{CP} \sin \theta) {}^G F_y + \tau \\ m {}^G \ddot{y}_C &= {}^G F_y - mg \end{cases} \quad (22)$$

where $m, J \in \mathbb{R}^+$ are the object mass and inertia, respectively, $\tau \in \mathbb{R}$ is the compensatory torque (generated by the proposed algorithm CL/ EL/GM), ${}^G F_y \in \mathbb{R}$ is the time-variant lift force, g is the earth gravity acceleration, $\dot{\theta} \in \mathbb{R}$ is the object angular acceleration, $\theta \in \mathbb{R}$ is the object roll, ${}^G \ddot{y}_C \in \mathbb{R}$ is the vertical linear acceleration of the CM C , and ${}^Ox_{CP}, {}^Oy_{CP} \in \mathbb{R}$ are the horizontal and vertical distances of P from C in local frame coordinates, respectively.

Leveraging on the ILC framework, we can analytically discuss the convergence of (18), (20), and (21) when applied to (22). Let's consider, for example, the case of $\delta = 0$, $K_{ff} = [0, 0, 0]$, and $K_{fb} = [0, 0, D_{fb}]$, with $D_{fb} \in \mathbb{R}$. By applying Theorem 3.1 in ref. [31]—derived for a more general class of nonlinear system—the following convergence condition results

$$\frac{D_{fb}}{J + 1} > 1 \quad (23)$$

which, if fulfilled, assures that the error goes to 0 if $Q = 1$. If instead $Q \in (0, 1)$, the error remains bounded in a neighborhood of 0. A nonunitary forgetting factor is therefore coherent with the experimental results shown in Figure 2.

Remark 1. Note that the only physical parameter affecting the learning convergence is the object inertia J . The convergence is assured independently from the position of the CM (${}^Ox_{CP}, {}^Oy_{CP}$). This confirms that the same mechanism, with the same parameters, can be used to achieve the control goal independently from the mass distribution. This hypothesis will be confirmed in Section 7.

Similar conclusions can be drawn using results provided in ref. [32], when $\delta = 0$, $K_{ff} = [0, 0, D_{ff}]$, and $K_{fb} = [0, 0, 0]$, with $D_{ff} \in \mathbb{R}$. In ref. [33], authors consider the case with $\delta \neq 0$, $K_{ff} = [0, 0, D_{ff}]$, $K_{fb} = [0, 0, 0]$, and $Q = 1$, showing that the delay does not affect the convergence of the algorithm.

Apart from combining the effects considered in refs. [31–33], the proposed learning rule (18) includes also the use of $\theta^i(t)$ and $f\theta^i(t)$. As discussed in ref. [34] despite the convergence of ILC algorithms is typically dictated by the higher order derivative of the error involved in the learning rule, the use of proportional and integral terms can ensure a smooth and monotonic convergence. Unfortunately, we are not aware of any comprehensive convergence result for the full control law (18). Deriving it is beyond the scope of this article and it will be considered in future work.

5. Identification

5.1. Identification Procedure

To validate the effectiveness of the proposed models, we identify here their free parameters. As quality measure for the identification problem, we consider here

$$\sum_{k \in \mathbf{T}_1} \|\hat{U}^k(K_{fb}, K_{ff}, Q, \Delta) - U^k\|^2 \quad (24)$$

where $\hat{U}^k \in \mathbb{R}^{1 \times N}$ collects control action estimation at the k th iteration. We derive it from Equation (18) as

$$\hat{U}^k(K_{fb}, K_{ff}, Q, \Delta) = K_{fb}E^k(\Delta) + K_{ff}E^{k-1}(0) + QU^{k-1} \quad (25)$$

where $K_{fb}, K_{ff} \in \mathbb{R}^{1 \times 3}$ are the control gains, $Q \in \mathbb{R}$ is the forgetting factor, and $\Delta \in \mathbb{N}$ is the delay expressed in time steps. Note that K_{fb}, K_{ff}, Q are directly the unknown parameters of the proposed models, whereas Δ can be converted into δ by multiplication for the sampling rate.

We split the problem of finding K_{fb} and Δ from the one of finding K_{ff} and Q that minimize (24), by assuming the first trial to be driven by pure feedback action, i.e., $U^0 = 0$. This can be justified by considering that at the first iteration subjects do not yet experience of the object, which could be used for generating the anticipatory action. A small effect from the previous block of trials may still occur. We will neglect this component, being the block order randomized (refer to Section 2.2 for more details). We thus focus on trial-by-trial evolution within a block of trials instead of block-by-block interferences.

Leveraging on this assumption we identify feedback gains K_{fb} as

$$\arg \min_{K_{fb}} \|U^1 - K_{fb}E^1(\Delta)\|^2 \quad (26)$$

that can be solved in closed form as

$$K_{fb}(\Delta) := U^1(E^1(\Delta))^\dagger \quad (27)$$

where $(\cdot)^\dagger$ is the Moore-Penrose right pseudoinversion.^[35] From all the delays we select $\bar{\Delta}$, minimizing $\|U^1 - U^1(E^1(\bar{\Delta}))^\dagger E^1(\bar{\Delta})\|^2$. We derive it as direct evaluation of (24), $\forall \Delta \in \{0, \dots, N\}$.

Finally, we evaluate the minimum of (24) for the three proposed architectures as

(i) feedback K_{fb} plus control-based learning Q

$$Q = (U^+ - K_{fb}E_{\bar{\Delta}}^+)(U^-)^\dagger \quad (28)$$

(ii) feedback K_{fb} plus error-based learning K_{ff}

$$[K_{ff} \quad Q] = (U^+ - K_{fb}E_{\bar{\Delta}}^+) \left[\begin{array}{c} E^- \\ (U^- - K_{fb}E_{\bar{\Delta}}^-) \end{array} \right]^\dagger \quad (29)$$

(iii) feedback plus both error- and control-based learnings

$$[K_{ff} \quad Q] = (U^+ - K_{fb}E_{\bar{\Delta}}^+) \left[\begin{array}{c} E^- \\ U^- \end{array} \right]^\dagger \quad (30)$$

where

$$\left\{ \begin{array}{ll} U^+ \triangleq [U^{k_1} \quad \dots \quad U^{k_T}] & \in \mathbb{R}^{1 \times NT} \\ U^- \triangleq [U^{k_1-1} \quad \dots \quad U^{k_T-1}] & \in \mathbb{R}^{1 \times NT} \\ E_{\bar{\Delta}}^+ \triangleq [E^{k_1}(\bar{\Delta}) \quad \dots \quad E^{k_T}(\bar{\Delta})] & \in \mathbb{R}^{1 \times NT} \\ E_{\bar{\Delta}}^- \triangleq [E^{k_1-1}(\bar{\Delta}) \quad \dots \quad E^{k_T-1}(\bar{\Delta})] & \in \mathbb{R}^{1 \times NT} \\ E^- \triangleq [E^{k_1-1}(0) \quad \dots \quad E^{k_T-1}(0)] & \in \mathbb{R}^{1 \times NT} \end{array} \right. \quad (31)$$

where $1 < k_1 < \dots < k_i < \dots < k_T$, $k_i \in \mathbf{T}_1 \setminus \{1\}$, and $T = |\mathbf{T}_1| - 1$.

Note that gains so evaluated could produce an unstable feedback, and/or a nonconvergent iterative behavior. This is not a issue per se because the identified values are tested against two validations sets (see Section 2). Thus, in the case an unstable behavior is identified which is not actually coherent with the observed human behavior, this should result in decreased performance in validation.

Nonetheless—for the sake of comparison—we also test a second choice of the parameters, defined so to be inherently stable in time and across the iterations. First, we impose $\Delta = 0$, and we evaluate the linearized system around the desired equilibrium configuration. Applying the Routh–Hurwitz stability criterion to this system yields to the following necessary and sufficient conditions for the closed loop stability in time

$$K_{fb,I} > 0, \quad K_{fb,D} > 0, \quad (\tilde{G} + K_{fb,P})K_{fb,D} > JK_{fb,I} \quad (32)$$

where $K_{fb} = [K_{fb,P}, K_{fb,I}, K_{fb,D}]$ and $\tilde{G} = \frac{\partial G}{\partial \theta}$. This latter value is positive for the aforementioned experimental setup, i.e., gravity produces a stabilizing effect once the correct feedforward action is produced. Therefore, we propose the following regularization of the feedback gains, forcing the closed loop to be stable in time

$$\begin{aligned} \tilde{K}_{fb,P} &:= \max\left(\frac{K_{fb,I}}{K_{fb,D}}J - \tilde{G} + \epsilon_{th}, K_{fb,P}\right) \\ \tilde{K}_{fb,I} &:= \max(\epsilon_{th}, K_{fb,I}) \\ \tilde{K}_{fb,D} &:= \max(\epsilon_{th}, K_{fb,D}) \end{aligned} \quad (33)$$

with ϵ_{th} small positive constant, that we take here equal to 10^{-3} .

For what concerns the convergence of the learning process, we consider the condition in ref. [36], which—in the authors' best knowledge—is the closest to the proposed algorithms among the ones available in the literature (see Section 4). To apply it, we must first impose $Q = 1$. Under this assumption, simple steps yield the following convergence conditions—one for each architecture

$$\begin{aligned} \text{Control learning:} & \quad \tilde{K}_{fb,D} < -2J \text{ or } \tilde{K}_{fb,D} > 0 \\ \text{Error learning:} & \quad 0 < K_{ff,D} < 2J + 2K_{fb,D}, \text{ for } K_{fb,D} > 0 \\ \text{Generalized model:} & \quad -K_{fb,D} < K_{ff,D} < 2J + K_{fb,D}, \text{ for } K_{fb,D} > 0 \end{aligned} \quad (34)$$

where J is the rotational inertia of the object, and $K_{ff} = [K_{ff,P}, K_{ff,I}, K_{ff,D}]$. Note that we used the equivalence in appendix for expressing the error learning case. Based on the derived conditions, we introduce the following projections assuring the convergence of the learning process in the iteration domain

$$\begin{aligned}
 \text{Error learning : } \quad \tilde{K}_{ff,D} &:= \begin{cases} \epsilon_{th} & \text{if } K_{ff,D} < 0 \\ 2J + 2K_{fb,D} - \epsilon_{th} & \text{if } K_{ff,D} > 2J + 2K_{fb,D} \\ K_{ff,D} & \text{otherwise} \end{cases} \\
 \text{Generalized model : } \quad \tilde{K}_{ff,D} &:= \begin{cases} -K_{fb,D} + \epsilon_{th} & \text{if } K_{ff,D} < -K_{fb,D} \\ 2J + K_{fb,D} - \epsilon_{th} & \text{if } K_{ff,D} > 2J + K_{fb,D} \\ K_{ff,D} & \text{otherwise} \end{cases}
 \end{aligned} \tag{35}$$

No projection is needed for the control learning scheme because the fulfillment of the convergence condition is already assured by (32). The remaining terms of K_{fb} and K_{ff} are equal to the corresponding ones in K_{fb} and K_{ff} .

Both K_{fb} , K_{ff} —standard identification hereinafter—and \tilde{K}_{fb} , \tilde{K}_{ff} —projected or inherently stable identification hereinafter—will be tested and compared within all the considered validation sets.

5.2. Identification Results

We report in Table 1 the control gains resulting from the application of (27)–(30) on the experimental data. The forgetting factor Q is greater than 0.8 for each subject and for all the proposed models. Relatively large integral gains (I) result too. This suggests a dominant role of both long-term intertrials and short-term intratrial memory w.r.t. other components. Indeed, derivative terms (D) are instead in the order of 10^{-2} , with values very consistent among subjects. Proportional terms (P) are small too, but they show a larger variability among subjects. Note that—as we already discussed in Section 4—when a $Q \neq 1$ is identified, the three ILC algorithms cannot ensure the perfect tracking performance, and a steady-state error different from zero should be expected. Interestingly, this characteristic can be observed in the average regulation error in the experimental data, as shown in Figure 2.

Table 1 also shows the results of the stability (32) and convergence (34) tests applied when Q is imposed equal to one and δ to zero. All identified parameters passed these checks, except for

the feedback gains of subject 2, the error-based architecture gains of subject 5, and the gains of the generalized architecture for subject 1. The new gains for these three subjects—once projected using (33) and (35)—are shown in Table 2.

The feedforward proportional contribution is often larger than the feedback one. This is a consequence of the nature of the problem as stated in Section 3. Indeed, the feedback action suffers from a physiological delay δ that we took into account with the parameter Δ during the fitting phase. Instead the feedforward action does not suffer from any real-time delay.

6. Closed Loop Accuracy

Previous validation is *open loop* in the sense that the system is not simulated, and only the relationship between measured errors and torques is considered.

However, this validation procedure does not inform us about the ability of the strategies to work properly under perturbations of the state or under uncertainties. These are essential characteristics that an hypothetical neural controller must possess to function properly. It is thus paramount to test these important characteristics with a dedicated validation routine.

All three proposed algorithms CL, EL, and GM will be evaluated.

6.1. Performance Evaluation

To evaluate the ability of the architectures to actually perform the task, we simulate the system (22), controlled through the four

Table 1. Gains resulting from the identification phase—performed using right load experimental data. For each architecture and subject we also indicate if the corresponding test (33) is verified (T) or not (F), and if the feedback action is stable (T) or not (F), according to (31).

Sub	Control-based					Error-based							Generalized							FB stab.				
	K_{fb}			Q	Conv test	K_{fb}			\tilde{K}_{ff}				Q	Conv test	K_{fb}			K_{ff}				Q	Conv test	
	P	I	D			P	I	D	P	I	D	P			I	D	P	I	D		P			I
1	0.68	4.12	0.02	0.81	T	0.68	4.12	0.02	0.45	3.22	0.02	0.82	T	0.68	4.12	0.02	0.19	-0.19	0.03	0.82	F	T		
2	0.01	3.68	0.03	0.95	T	0.01	3.68	0.03	0.07	2.89	0.02	1.02	T	0.01	3.68	0.03	0.11	-0.77	0.01	1.01	T	F		
3	0.12	2.23	0.04	0.83	T	0.12	2.23	0.04	0.33	1.68	0.02	0.85	T	0.12	2.23	0.04	0.22	-0.11	0.01	0.84	T	T		
4	0.17	2.16	0.09	1.06	T	0.17	2.16	0.09	0.34	2.23	0.01	1.06	T	0.17	2.16	0.09	0.12	0.20	0.01	1.04	T	T		
5	-0.14	0.83	0.02	0.92	T	-0.14	0.83	0.02	0.05	0.75	-0.00	0.92	F	-0.14	0.83	0.02	0.14	0.11	-0.00	0.88	T	T		
6	-0.09	1.27	0.05	0.89	T	-0.09	1.27	0.05	0.30	0.98	0.01	0.88	T	-0.09	1.27	0.05	0.36	0.07	0.01	0.82	T	T		
7	0.02	1.25	0.05	0.93	T	0.02	1.25	0.05	0.37	1.23	0.03	0.86	T	0.02	1.25	0.05	0.31	0.27	0.01	0.85	T	T		
8	0.08	2.10	0.04	1.02	T	0.08	2.10	0.04	0.31	2.51	0.03	0.96	T	0.08	2.10	0.04	0.22	0.63	0.01	0.94	T	T		
9	0.19	2.91	0.04	1.04	T	0.19	2.91	0.04	0.28	2.85	0.02	1.06	T	0.19	2.91	0.04	0.12	-0.13	0.01	1.05	T	T		

Table 2. Gains resulting from the identification phase when projections (32) and (34) are applied related to right load experimental data. Only the gains of the three subjects that failed the checks (31) and (33) are shown. The other subjects have unmodified K_{fb} , K_{ff} , and $Q = 1$.

Sub	Control-based					Error-based						Generalized						FB stab				
	K_{fb}			Q	Conv test	K_{fb}			\bar{K}_{ff}			Q	Conv test	K_{fb}			K_{ff}			Q	Conv test	
	P	I	D			P	I	D	P	I	D			P	I	D	P		I			D
1	0.68	4.12	0.02	1.00	T	0.68	4.12	0.02	0.45	3.22	0.02	1.00	T	0.68	4.12	0.02	0.19	-0.19	0.02	1.00	T	
2	0.12	3.68	0.03	1.00	T	0.12	3.68	0.03	0.07	2.89	0.02	1.00	T	0.12	3.68	0.03	0.05	-0.73	0.01	1.00	T	T
5	-0.14	0.83	0.02	1.00	T	-0.14	0.83	0.02	0.05	0.75	0.00	1.00	T	-0.14	0.83	0.02	0.14	0.11	-0.00	1.00	T	T

architectures as introduced in Section 3. We are not interested here in modeling the generation of ${}^G F_y$, which we already discussed. We extract it from experimental data, and we directly feed it into (22).

We also account for the well-known high level of noise in human sensory system, testing robustness of the architectures to noises with different amplitudes. Instead of evaluating the controllers directly through the state $(\theta, \dot{\theta})$, we use the following noisy measurements

$$\begin{aligned} \hat{\theta}(t) &= \theta(t) + Z_{\lfloor \frac{t}{T} \rfloor}^1 \\ \hat{\dot{\theta}}(t) &= \dot{\theta}(t) + Z_{\lfloor \frac{t}{T} \rfloor}^2 \\ Z_k^i &= Z_{k-1}^i + W_{k-1}^i, Z_0^i = 0 \end{aligned} \quad (36)$$

where Z_k^i is the k th step of the i th random walk, with $W_k^i \sim N(0, \sigma)$ independent and equally distributed random variables with normal distribution, null mean, and variance σ . We test the algorithms performance for $\sigma \in \{0, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$. We use as index the average RMSE as defined in Section 2.4, calculated between the measured and simulated torques. Each simulation is performed 30 times.

We consider as validation set the one described by (10), i.e., the set of all the trials that were excluded during the identification with the load placed in the left spot.

7. Results

In this section we assess the ability of all the proposed strategies (namely, CL, EL, and GM) of explaining experimental data. Their control gains are identified as discussed in Section 5.

Together with the three algorithms, we also include here the state of the art learning rule (15), serving as a benchmark. Hereinafter, we will call it *modified Emken* model, from the name of the author who originally proposed it. We added the adjective *modified* to stress that this rule was developed for motor learning of a quite different task, so to apply it to the problem under exam we had to combine (15) with the here proposed force mapping (3).

7.1. First Set: Same Mass Distribution

Figure 5i shows the average performances as defined by (13), for the three proposed architectures and for the benchmark,

identified with (27)–(30). Figure 5ii shows the same results when the projections (33) and (35) are applied. As expected, the two trends are similar, and the projected ones perform always worst than the ones obtained with standard identification. Figure 6 shows a comparison of torques produced by the generalized model when projected—assuring inherent stability in time and in iterations—and not projected parameters are considered. The behavior is very similar in the two cases, with the nonprojected model presenting an offset at steady state w.r.t. the projected one, that makes it closer to the measured data.

Emken model presents the worst performance among the proposed ones, whereas EL, CL, and GM learning schemes have similar performance with a slight advantage of the latter. The average validation error is below 15% for all of them. To verify statistical difference between conditions, we used nonparametric tests given the non-Gaussian distribution of samples, as demonstrated by the application of Shapiro–Wilk test. In Table 1 we collect the results of a P -value analysis among RMSE R (see Equation (13)) with Wilcoxon signed rank test with Bonferroni correction. The differences between the proposed architectures and the benchmark are all statistically significant, whereas differences between the three architectures proposed in this work are not. Table 1 shows a comparison of BIC for the four architectures on this validation. Results confirm a slight advantage of GM w.r.t. the other four architectures even when model complexity is penalized. Note that the parameter κ in (14) is two for Emken, four for CL, seven for EL, and seven for GM. Other parameters are defined as earlier.

Differences appear more evident considering temporal evolution. As an example, Figure 7 shows the measured and estimated torques for subject 6. The proposed algorithms (EL, CL, GM) are able to correctly reproduce salient characteristics of the measured signals, e.g., peak torque, steady state behavior. The first iteration shows that the pure feedback assumption fits well, except for the small initial opposite peak. This is probably due to an unmodeled anticipatory action, learned in a previous block of trials with an opposite load position. As soon as the subject realizes that the mass distribution is not the one experienced before, she/he nullifies this action and the pure feedback model starts working well. This happens around 1 s. Of note, as no feedback action is included in the Emken model, no control action is produced at the first iteration. The second iteration also presents a slight lower estimation of the control for CL, and EL schemes. The latter and the GM show some initial oscillations of the control action, which are rapidly mitigated in the subsequent iterations. However, the most evident observation is the strong

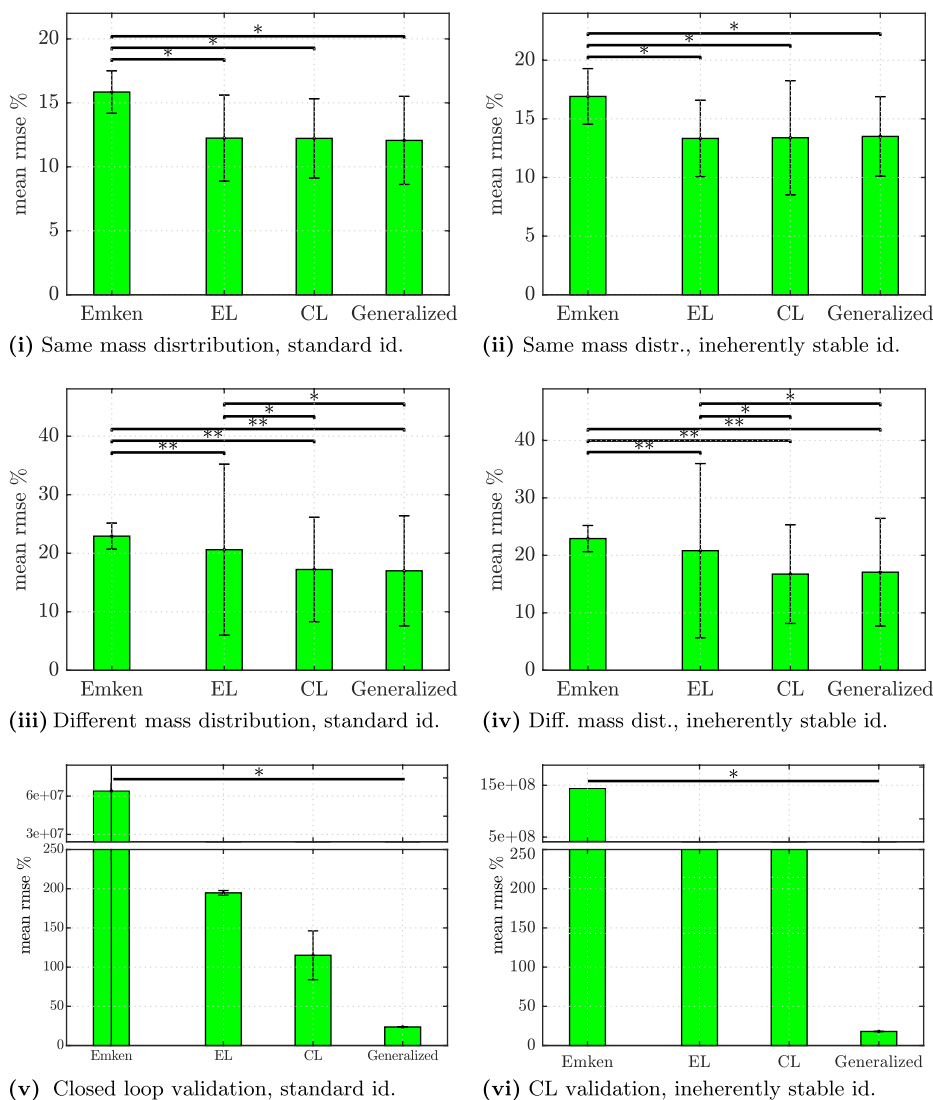


Figure 5. Global performance of the three proposed architectures (CL, EL, GM), compared with the modified Emken model serving as benchmark, for the three validation conditions. Bars indicate the average error (RMSE) across subjects as defined in (13), with their standard deviation.

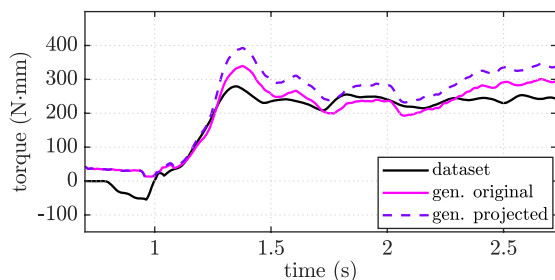


Figure 6. Experimental and predicted torques for subject 1, iteration 5, during validation with same mass distribution as the one used in the identification. We consider here the GM architecture for data identified with and without projection.

mismatching between torques predicted by Emken model and the measured ones.

7.2. Second Set: Different Mass Distribution

Figure 5iii shows average performance of the four algorithms when a variation of the mass distribution is considered, identified with (27)–(30). Figure 5ii shows the same results when the projections (33) and (35) are applied. As observed for the previous validation set, the two trends are similar, and the projected ones perform always worst than the ones obtained with standard identification.

The differences between architectures is here more evident w.r.t. to previous section. Emken model exhibits the worst performances. It is followed by the EL which, in turn, presents this time a very high variance among subjects. Note that the two models come from the same neuroscientific hypothesis. CL and GL schemes have the best performance. These two architectures are able to predict the subjects' behavior with an error comparable with the one attained in the first validation test, i.e., $\approx 17\%$.

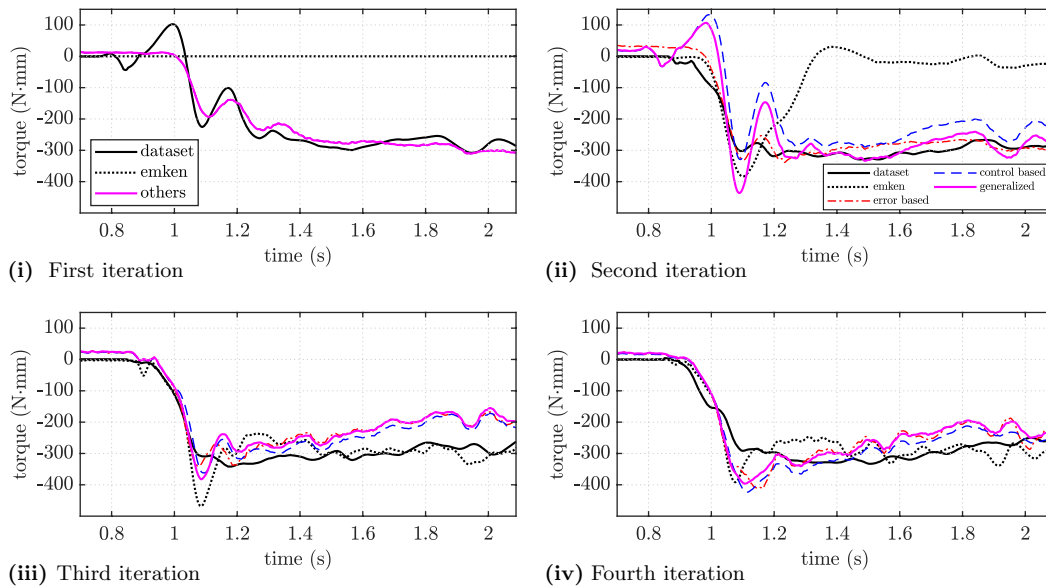


Figure 7. Experimental and predicted torques for subject 6, during validation with same mass distribution as the one used in the identification. CL, EL, and GM can reproduce both general trend and distinctive features of the measured values (solid black lines).

Interestingly, the generalized scheme seems to perform worst in this validation w.r.t. control-learning scheme, when the parameters are projected to be inherently stable. This could be due to the larger amount of projections that the first undergoes

Table 3. *P*-value comparisons for the errors achieved with the validations considered in this work. Each enter of the tables reports a “yes” if the null hypothesis is rejected, i.e., if the difference between the corresponding two architectures is significant, “no” otherwise. Panel (c) refers to significances evaluated for all the considered values of σ . Wilcoxon signed rank test with Bonferroni correction and $\alpha=0.05$ is used to test the hypotheses. The error index is calculated as specified by (13). EL, CL, and GM represent error-based learning, control-based learning, and generalized model, respectively. Emken is the benchmark as defined in (3) and (15).

(a) Same mass distribution			(b) Different mass distr.			(c) Closed loop validation		
EL	CL	GM	EL	CL	GM	EL	CL	GM
Emken	Yes	Yes	Emken	Yes	Yes	Emken	Yes	Yes
EL	–	No	EL	–	Yes	EL	–	Yes
CL	–	–	CL	–	No	CL	–	Yes

Table 4. BIC comparisons for the validations considered in this work. The index is calculated as specified by (14). EL, CL, and GM represent error-based learning, control-based learning, and generalized model, respectively. Emken is the benchmark as defined in (3) and (15).

BIC	BIC		BIC							
	σ	0	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}			
GM	–5.82	GM	–5.1662	GM	–4.6	–4.6	–4.5	–3.4	2.4	6.3
CL	–5.68	CL	–5.1596	EL	1.6	1.6	1.1	2.0	10.8	13.6
EL	–5.77	Emken	–5.0754	CL	1.8	1.8	2.6	2.6	13.2	15.7
Emken	–5.13	EL	–4.8098	Emken	24.0	24.0	24.0	24.3	27.8	29.7

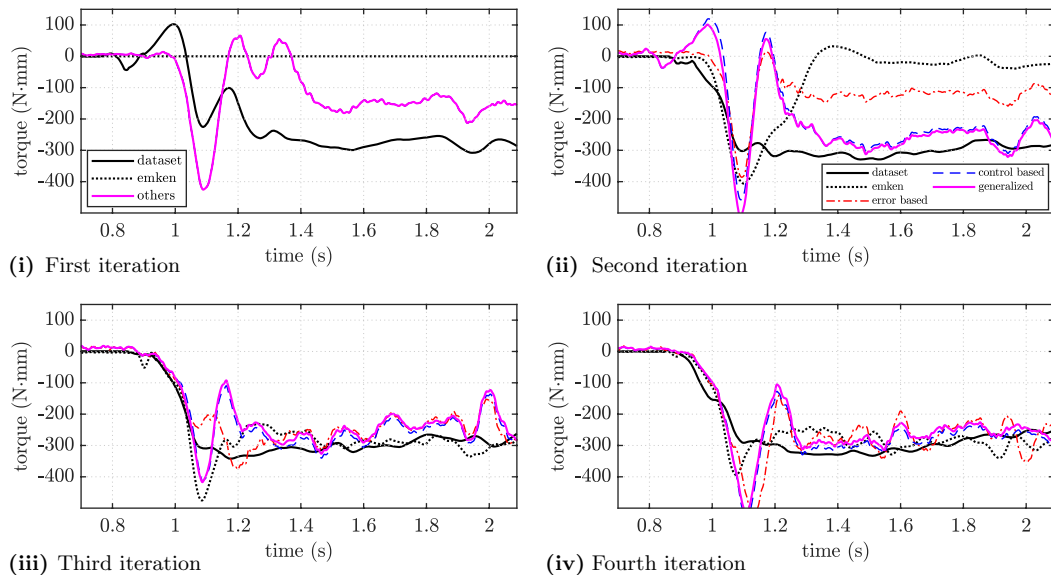


Figure 8. Experimental and predicted torques for subject 6, during validation with opposite mass distribution w.r.t. the one used in the identification. Thus none of the evolutions presented here is part of the identification set.

shows the same results when the projections (33) and (35) are applied. The same trend observed before manifests itself once more; the architectures which are made inherently stable by projection of their gains perform worst than their not projected counterparts. This is particularly relevant considering that this validation should maximally exacerbate instabilities because it is the only one involving an actual simulation.

The differences between models exacerbate, when the distance from the ideal nominal conditions learned in the identification phase increases. The generalized model confirms to be the best in terms of accuracy between the four, with a performance index of $\approx 18\%$. It is followed by control-based learning and error-based learning, with a mean RMSE of more than 200%. Emken model shows here very poor performance, with an error several degrees of magnitude bigger ($\approx 10^7\%$). **Figure 9** shows an example of the evolution in time of the tilting angle θ , for subject 6. Evolution produced by Emken model is characterized by a marked instability arising already at the second iteration. (The evolution produced in the fourth iteration is barely visible, since it soon diverges beyond the plot limits.) This is coherent with the very high RMSE error reported earlier. Error- and control-based learning generate shaky but limited trajectories. Oscillations increase with the continuation of the learning process. The generalized model instead brings the system on trajectories resembling both the general trend and distinctive features of the measurements. This result appears even more relevant when considering that the model is not explicitly identified to match θ , but τ_p instead (see Section 3 and 5).

Figure 10i,ii report averages and variances of the results when noise is introduced. As expected the performance decreases with the increasing of noise variance for all the architectures. Among them Emken consistently shows very poor performance. Interestingly, pure error-based learning works well for small σ , while it is surpassed by control-based learning for higher

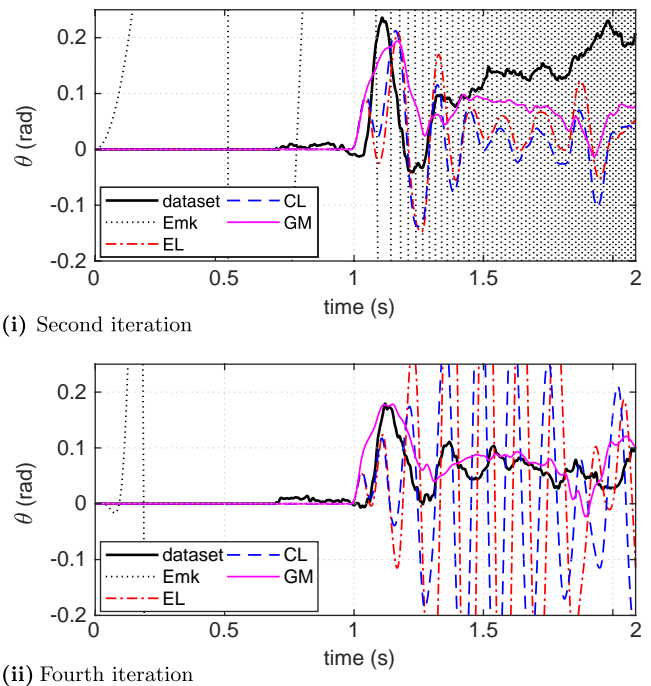


Figure 9. Experimental and predicted angles for subject 6. GM can accurately reproduce the subject behavior (solid black lines). EL and CL learning schemes generate strong oscillations, which increase with the continuation of the learning process. Emken model presents a strongly unstable behavior already at the second iteration.

noise. The generalized model combines the best characteristics of the two learning approaches, and it is the one that copes better with the noise both with small and high values σ . Finally, Table 4c shows a comparison of BIC indexes for the four

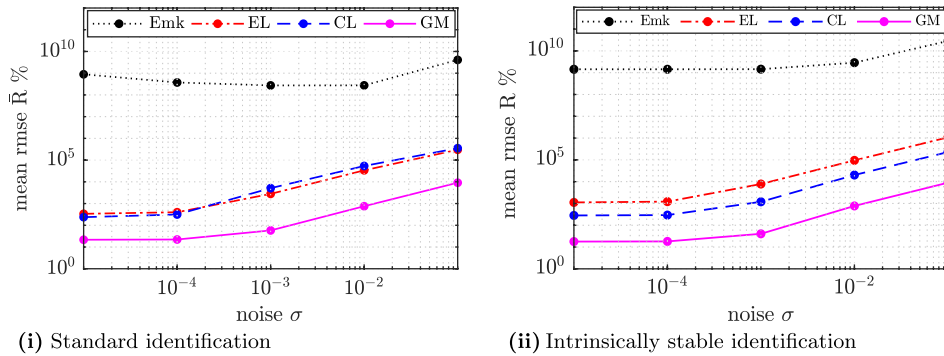


Figure 10. Noisy measurements: average prediction accuracy for both the case with standard identification and intrinsically stable one. System (22) is controlled with the four architectures (the three proposed here plus the modified Emken), and the resulting torques are compared with the measured values. Noise (35) is added to the measurements. Performance are evaluated through the index in (13). The standard identification performs better in all the considered conditions, with a marked difference for the EL case.

architectures on this validation. The same trend discussed earlier is maintained also when complexity is penalized, with the generalized model strongly outperforming the other architectures for all the considered noise levels.

8. Conclusions

Humans are able to efficiently learn the proper control action for lifting an object with unknown mass distribution, leveraging on the sensory-motor memory acquired in few trials. In this work we formulated and tested the hypothesis that this behavior, called previous trial effect, can be explained as proper combination of linear feedback and iterative learning control. The proposed architectures successfully reproduce the human behavior in the generation of compensatory actions on a trial-by-trial basis. Leveraging on ILC theory we were able to discuss analytically the stability and performance of resulting control strategies.

We tested the three proposed models against a benchmark with three validation tests of increasing complexity. In the first one, the model's ability to predict evolutions when the weight is placed in the same position as in the identification set is tested. The three proposed architectures show better performances w.r.t. the benchmark, but similar between each others. In the second validation test the weight distribution changes between the identification and the validation set: the performance of the benchmark degrades, and differences between the proposed architectures start to appear. Finally, the third validation tests the model ability to control the system, both with and without noisy measurements. Here, the proposed architectures show performances which are several degrees of magnitudes better than the benchmark. Among them, the model mixing error and control-based learning (i.e., generalized model) presents again the best performance, comparable with the ones obtained in the first two validations.

The fact that the generalized model had the best performance is consistent with the recent trend in neuroscience research, where motor adaptation can be attributed to a combination of parallel neural mechanisms which differ in time-space and sensorimotor memory usage.^[5,37,38] In addition, our model can capture within-trial temporal evolution of human motor

control during trial-by-trial adaptations, whereas most previous studies consider the motor output and error signals only as scalars for each trial.

Future work will further extend the proposed architectures, to include changes in digit positioning. We will also test different instances of (17), including nonlinear terms and extending the algorithms memory beyond one trial back. We will also devise mechanical apparatuses able to impose on the object more complex disturbance profiles. In this way, we could go further in the exploitation of the mathematical formalization of the problem by optimally exciting disturbances for identification (i.e., persistent excitation).

The results presented in this work open promising perspective in robot control, e.g., to devise control laws in autonomous robotic manipulation of external environment as preliminary depicted in refs. [29,39]. We believe that including the modeling of human behavior could be a key step forward for a new generation of robotic manipulators successfully and autonomously interacting with the external world, in a way similar to the human one and taking advantage from *sensory-motor memory* collected during task execution.

Appendix

An Error-Based Learning to Generalized Model Equivalence

We rewrite error-based learning rule (21) in the generalized model form (18) when $\delta = 0$. Through algebraic manipulation

$$\begin{aligned} \tau^k &= K_{fb} \epsilon^k + \bar{K}_{ff} \epsilon^{k-1} + \tau_{ff}^{k-1} \\ &= K_{fb} \epsilon^{k+1} + \bar{K}_{ff} \epsilon^k + Q(\tau^k - K_{fb} \epsilon^k) \\ &= K_{fb} \epsilon^{k+1} + (\bar{K}_{ff} - QK_{fb}) \epsilon^k + Q\tau^k \\ &= K_{fb} \epsilon^{k+1} + K_{ff} \epsilon^k + Q\tau^k \end{aligned} \quad (37)$$

where $K_{ff} = \bar{K}_{ff} - QK_{fb}$.

Acknowledgements

This work was supported by European Union's Horizon 2020 Research and Innovation Program under grant agreement 871237 (Sophia).

ERC Synergy Grant 810346 (Natural BionicS), and the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence).

Conflict of Interest

The authors declare no conflict of interest.

Keywords

grasping and manipulation, iterative learning control, mathematical models of human motor control, previous trial effect

Received: July 10, 2019

Revised: April 22, 2020

Published online: July 1, 2020

-
- [1] R. Shadmehr, F. A. Mussa-Ivaldi, *J. Neurosci.* **1994**, *14*, 3208.
[2] J. R. Lackner, P. Dizio, *J. Neurophysiol.* **1994**, *72*, 299.
[3] M. Kawato, K. Furukawa, R. Suzuki, *Biol. Cybern.* **1987**, *57*, 169.
[4] H. Miyamoto, M. Kawato, T. Setoyama, R. Suzuki, *Neural Networks* **1988**, *1*, 251.
[5] D. M. Wolpert, R. C. Miall, M. Kawato, *Trends Cognit. Sci.* **1998**, *2*, 338.
[6] T. Kitago, J. W. Krakauer, *Handb. Clin. Neurol.* **2013**, *110*, 93.
[7] N. Hogan, D. Sternad, *Biol. Cybern.* **2012**, *106*, 727.
[8] S. W. Park, H. Marino, S. K. Charles, D. Sternad, N. Hogan, *J. Neurophysiol.* **2017**, *118*, 69.
[9] J. L. Emken, R. Benitez, A. Sideris, J. E. Bobrow, D. J. Reinkensmeyer, *J. Neurophysiol.* **2007**, *97*, 3997.
[10] R. A. Scheidt, J. B. Dingwell, F. A. Mussa-Ivaldi, *J. Neurophysiol.* **2001**, *86*, 971.
[11] I. Herreros, X. Arsiwalla, P. Verschure, *Adv. Neural Inf. Process. Syst.* **2016**, 3828.
[12] J. R. Lukos, C. Ansuini, M. Santello, *J. Neurosci.* **2008**, *28*, 12765.
[13] J. R. Lukos, J. Y. Choi, M. Santello, *J. Neurophysiol.* **2013**, *109*, 2937.
[14] Q. Fu, W. Zhang, M. Santello, *J. Neurosci.* **2010**, *30*, 9117.
[15] P. J. Parikh, J. M. Fine, M. Santello, *Cereb. Cortex* **2019**, *30*, 3087.
[16] J. P. Gollwitzer, C. S. Chapman, D. M. Wolpert, J. R. Flanagan, *Nat. Rev. Neurosci.* **2018**, *19*, 519.
[17] M. Liu, K. Li, N. Zhang, N. Wei, in *2019 12th Int. Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, IEEE, Suzhou, China **2019**, pp. 1–5.
[18] D. Voudouris, M. D. Broda, K. Fiehler, *J. Vision.* **2019**, *19*, 4.
[19] W. Hu, N. Wei, K. Li, S. Yue, in *2018 14th IEEE Int. Conf. on Signal Processing (ICSP)*, IEEE, Beijing, China **2018**, pp. 1010–1014.
[20] H. Guang, S. Bazzi, D. Sternad, N. Hogan, in *2019 Int. Conf. on Robotics and Automation (ICRA)*, IEEE, Montreal, QC, Canada **2019**, pp. 3783–3789.
[21] Casalino G., Bartolini G. in *IATED Symp. on Robotics and Automation*, IATED, Amsterdam **1984**, pp. 108–111.
[22] D. A. Bristow, M. Tharayil, A. G. Alleyne, *IEEE Control Syst. Mag.* **2006**, *26*, 96.
[23] R. Johansson, G. Westling, *Exp. Brain Res.* **1984**, *56*, 550.
[24] K. O. Johnson, *Curr. Opin. Neurobiol.* **2001**, *11*, 455.
[25] M. Gabbicini, A. Bicchi, D. Prattichizzo, M. Malvezzi, *Auton. Robots* **2011**, *31*, 235.
[26] A. G. Feldman, *J. Motor Behav.* **1986**, *18*, 17.
[27] R. S. Johansson, J. R. Flanagan, *Nat. Rev. Neurosci.* **2009**, *10*, 345.
[28] G. Schwarz, *Ann. Stat.* **1978**, *6*, 461.
[29] C. Della Santina, M. Bianchi, G. Grioli, F. Angelini, M. Catalano, M. Garabini, A. Bicchi, *IEEE Robot. Autom. Mag.* **2017**, *24*, 75.
[30] D. A. Winter, *Biomechanics and Motor Control of Human Movement*, John Wiley & Sons, Hoboken, NJ **2009**.
[31] C. J. Chien, J. S. Liu, *Int. J. Control* **1996**, *64*, 319.
[32] S. S. Saab, *IEEE Trans. Autom. Control* **1994**, *39*, 2298.
[33] M. Sun, D. Wang, *Dyn. Control* **2000**, *10*, 341.
[34] Y. Chen, K. L. Moore, in *Proc. of the 2002 American Control Conf.*, IEEE, Anchorage, AK, USA **2002**, Vol. 3, pp. 2138–2143.
[35] A. Albert, *Regression and the Moore-Penrose Pseudoinverse*, Elsevier, Amsterdam/New York **1972**.
[36] J. Shou, D. Pi, W. Wang, in *IEEE Int. Conf. on Systems, Man and Cybernetics*, IEEE, Washington, DC, USA **2003**, Vol. 3 pp. 2557–2562.
[37] J. Diedrichsen, O. White, D. Newman, N. Lally, *J. Neurosci.* **2010**, *30*, 5159.
[38] Q. Fu, M. Santello, *J. Neurophysiol.* **2015**, *113*, 144.
[39] F. Angelini, C. Della Santina, M. Garabini, M. Bianchi, G. M. Gasparri, G. Grioli, M. G. Catalano, A. Bicchi, *IEEE Trans. Robot.* **2018**, *31*, 924.