

Online sensorimotor learning and adaptation for inverse dynamics control

Xiaofeng Xiong ^{a,*}, Poramate Manoonpong ^{a,b}

^a SDU Biorobotics, the Mærsk Mc-Kinney Møller Institute, the University of Southern Denmark (SDU), Campusvej 55, 5230 Odense M, Denmark

^b Bio-Inspired Robotics and Neural Engineering Lab, the School of Information Science and Technology, Vidyasirimedhi Institute of Science and Technology, Wangchan Valley 555 Moo 1 Payupnai, Wangchan, 21210 Rayong, Thailand

ARTICLE INFO

Article history:

Received 18 December 2020

Revised and accepted 29 June 2021

Available online xxxx

Keywords:

Neural network

Robot control

Variable compliant control

Gaussian model

ABSTRACT

We propose a micro-data (< 10 trials) sensorimotor learning and adaptation (SEED) model for human-like arm inverse dynamics control. The SEED model consists of a feedforward Gaussian motor primitive (GATE) neural network and an adaptive feedback impedance (AIM) mechanism. Sensorimotor weights over trials are learned in the GATE network, while the AIM mechanism is used to online tune impedance gains in a trial. The model was validated by periodic and non-periodic tracking tasks on a two-joint robot arm. As a result, the proposed model enables the arm to stably learn the tasks within 10 trials, compared to thousands of trials required by state-of-art deep learning. This model facilitates the exploration of unknown arm dynamics, in which the elbow joint requires much less active control compared to the shoulder. This control goes below 3% of the overall effort. This finding complies with a proximal-distal control gradient in human arm control. Taken together, the proposed SEED model paves a way for implementing data-efficient sensorimotor learning and adaptation of human-like arm movement.

© 2021 Published by Elsevier Ltd.

1. Introduction

Human sensorimotor systems can learn to online adapt to unknown arm dynamics in object manipulation, e.g., lifting a bottle of water (Hitzler, Meier, Schaal, & Asfour, 2019). To emulate this capability, a robot arm needs to learn its changing dynamics to adapt new tasks (Zhang, Wang, & Masayoshi, 2020). This is the well-known inverse dynamics problem in robot arm control. A learning method is used to predict (output) joint torques given by desired joint trajectories and changing dynamics (Ren & Ben-Tzvi, 2020). Learning to track trajectories is a challenging task in a data-efficient way (<10 trials), which has not yet been realized by state-of-art methods such as inertial parameter estimation, nonlinear regression based on feedforward neural networks, and deep learning (Bogdanovic, Khadiv, & Righetti, 2020; Hitzler et al., 2019). Many experiments have indicated that this learning is required in feedforward and feedback control of stable movement (Seidler, Noll, & Thiers, 2004; Yousif & Diedrichsen, 2012). Inspired by these principles, in this paper we propose a computational model for micro-data (<10 trials) sensorimotor learning and adaptation (see Fig. 1). The proposed model integrates feed-forward and feedback learning to produce stable arm movement through the exploration of natural (i.e., passive) arm dynamics.

The results and findings of the model have not yet been shown and realized by other robot arm learning and computational models.

Prominent robot learning control models in the literature consist of online regression, reinforcement learning, optimization, iteration, and learning from demonstration. Extensive reviews are provided in Chatzilygeroudis, Vassiliades, Stulp, Calinon, and Mouret (2020), Nguyen-Tuong and Peters (2011), Ravichandar, Polydoros, Chernova, and Billard (2020) and Schaal and Atkeson (2010). The proposed SEED outperforms pure data-driven methods such as deep reinforcement learning and generative adversarial networks (GANs) in fewer trials (Ren & Ben-Tzvi, 2020). For instance, Bogdanovic et al. (2020) applied Deep Deterministic Policy Gradient (DDPG) to train online impedance control of a two-jointed hopping robot (Bogdanovic et al., 2020). This train requires more 7000 trials performed in simulation. To reduce training trials, Zhang et al. (2021) developed a neural network based iterative learning control (ILC) for one-joint control of a SCARA robot (Zhang et al., 2020). Although this physics-based learning only requires 14 trials, it does not exhibit faster learning (<10 trials), inverse dynamics exploitation and adaptation. Data-driven learning and physics-based adaptation are integrated to achieve these in the proposed SEED. It shows that online physics-based (impedance) adaptation can boost data-driven learning in inverse arm dynamics control. It exhibits fast learning and gradient exploitation of joint dynamics (see section experiments),

* Corresponding author.

E-mail address: xizi@mmti.sdu.dk (X. Xiong).

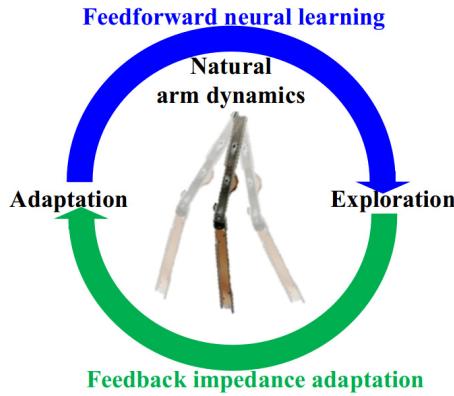


Fig. 1. Sensorimotor learning and adaptation methodology for human-like arm movement control.

which have not yet been realized by the state-of-art fast learning methods such as feedforward neural networks and iterative learning control.

The proposed SEED learning contributes to state-of-art by providing data-efficient (<10 trials) sensorimotor learning and adaptation of inverse dynamics control. This efficiency is facilitated by integrating data-driven (supervised) learning and physics-based (impedance) adaptation in control. Here ‘adaptation’ means an online impedance tuning law is shared between different arm dynamics and tasks (Hitzler et al., 2019). This law allows for online learning, i.e., generalizing its dynamics (impedance) adaptation in different tasks and dynamics. This can increase computation efficiency, compared to incremental learning used in state-of-art methods. Our experiments show that integrating this physics-based adaptation to a data-driven algorithm boosts faster inverse dynamics learning, compared to pure data-driven algorithms. Existing feedforward neural networks (FFNN) and iterative learning control (ILC) algorithms were implemented and compared to our proposed SEED learning (Hitzler et al., 2019; Zhang et al., 2020), because they are simple and data-efficient (<100 trials) in real robot control (see Table 1).

The remainder of this paper is structured as follows. A two-joint robot arm and the proposed model are presented in Section 2. The model consists of online feedforward sensorimotor learning and feedback impedance adaptation mechanisms. Section 3 contains the experimental results of more stable and accurate joint tracking motions driven by the proposed model, compared to classical learning control models. The experimental tasks consist of periodic and non-periodic joint trajectory tracking. A discussion and conclusion, including a review of related works, are provided in Section 4.

2. Hardware and software designs

An online sensorimotor learning and adaptation (SEED) model is developed to control the movement of a two-joint robot arm (see Fig. 2). The online SEED model consists of a Gaussian motor primitive (GATE) neural network for feedforward control (Takiyama & Sakai, 2016; Thoroughman & Shadmehr, 2000), as well as an adaptive impedance (AIM) mechanism for feedback control (Franklin, 2015; Franklin & Wolpert, 2011) (see Fig. 2). The feedback control is assisted by online adapting joint impedance gains (K_p, K_d) $\in \mathbb{R}^{2 \times 1}$ based on joint position and velocity errors, while the sensorimotor weights $\Omega \in \mathbb{R}^{14 \times 20}$ are learned based on regression error minimization in the feedforward control. The learning and adaptation mechanisms were tested on the movement control of a two-joint robot arm (see Fig. 3). The SEED

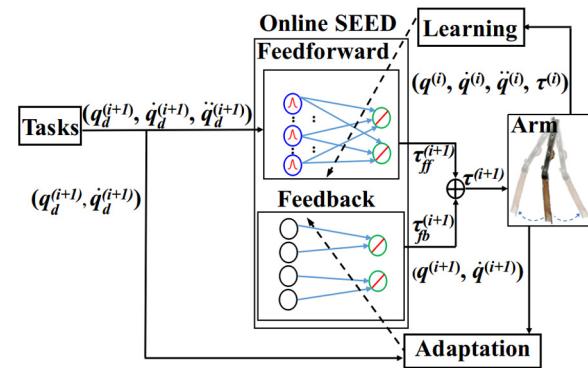


Fig. 2. Online sensorimotor learning and adaptation (SEED) model for human-like arm movement control. The weights and impedance parameter are updated in the feedforward neural learning and feedback impedance adaptation (see the dashed arrow lines).

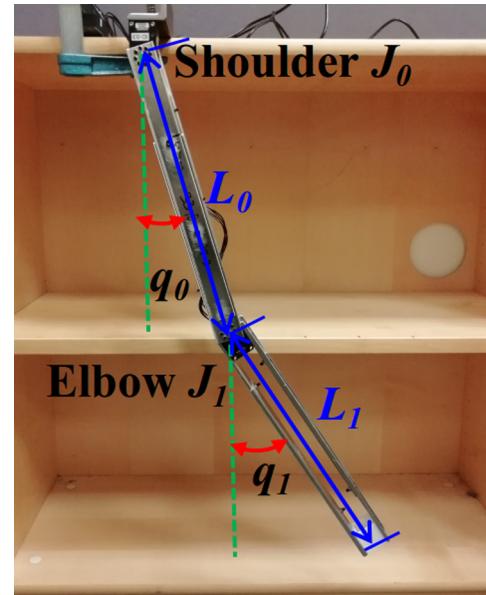


Fig. 3. Two-joint robot arm consisting of the joints $J_{(0,1)}$ for 2D movement control. Its weight is 0.524 kg. Its link lengths $L_{(0,1)}$ are 0.31 m and 0.34 m, as in the case of a human arm model (Burdet et al., 2006).

starts at the feedback AIM control. The feedback impedance gains (K_p, K_d) are subsequently decreased by the learned sensorimotor weights by the GATE feedforward learning.

2.1. Arm dynamics

A two-joint robot arm consists of proximal (i.e., J_0) and distal (i.e., J_1) joints for human-like arm movement control (see Fig. 3). Each joint is actuated by a Dynamixel motor.¹ Dynamixel actuators have been popularly applied to several robots due to their cost-effective characteristics (Karakasiliotis et al., 2016; Pogue, Bianes, Hong, & Iwasaki, 2017; Tanaka & Tanaka, 2015).

An inverse dynamics model of the two-joint robot arm is given by Spong, Hutchinson, and Vidyasagar (2006),

$$M(q)\ddot{q} + H(q, \dot{q})\dot{q} + G(q) = \tau_{in} + \tau, \quad (1)$$

where $q, \dot{q}, \ddot{q} \in \mathbb{R}^2$ denote the position, velocity, and acceleration of the two joints of the arm. $M(q) \in \mathbb{R}^{2 \times 2}$, $H(q, \dot{q}) \in$

¹ XM430-W350, <http://emanual.robotis.com/docs/en/dxl/x/xm430-w350/>.

Table 1
Comparison between our proposed (SEED) and state-of-art methods.

Algorithm	SEED	FFNN (Hitzler et al., 2019)	ILC (Zhang et al., 2020)	GANs (Ren & Ben-Tzvi, 2020)	DRL (Bogdanovic et al., 2020)
Year	2021	2019	2021	2020	2020
Predictive learning	Yes	Yes	Yes	Yes	Yes
Online impedance adaptation	Yes	No	No	No	Yes
Physics-based	Yes	No	Yes	No	No
Data-driven	Yes	Yes	No	Yes	Yes
Online learning and adaptation	Yes	No	No	No	No
Dynamics exploitation	Yes	No	No	No	No
Real robot	Yes	Yes	Yes	No	Yes
Incremental learning	No	Yes	No	Yes	Yes
Trials	< 10	100	14	100	>7000

$\mathbb{R}^{2 \times 2}$, and $G(q) \in \mathbb{R}^2$ are unknown inertial, Coriolis/centrifugal, and gravitational matrices. The torques $\tau_{in} \in \mathbb{R}^2$ result from interaction between joints, while the input (i.e., control) torque τ is generated by the online SEED model. Note that all arm dynamic parameters (e.g., $M(q)$) in Eq. (1) are unknown for all implementations presented in this paper.

2.2. Sensorimotor learning and adaptation (SEED)

The SEED model consists of feedforward and feedback control loops (see Fig. 2). Therefore, the control torque τ is given by:

$$\tau = \tau_{ff} + \tau_{fb}, \quad (2)$$

where $\tau_{ff} \in \mathbb{R}^2$ and $\tau_{fb} \in \mathbb{R}^2$ denote the feedforward and feedback torques produced by a feedforward Gaussian motor primitive (GATE) neural network and feedback adaptive impedance (AIM) mechanism.

2.2.1. Feedforward GATE neural network

The network has input, middle, and output layers of neurons organized in a feedforward and fully connected way (see Fig. 4). The seven input neurons are given by:

$$\theta_t^{(i+1)} = [1, q_0, \dot{q}_0, \ddot{q}_0, q_1, \dot{q}_1, \ddot{q}_1]^T \in \mathbb{R}^{7 \times 1}, \quad (3)$$

where $q_{(0,1)}, \dot{q}_{(0,1)}, \ddot{q}_{(0,1)}$ are the joint positions, velocities, and accelerations of the two-joint robot arm (see Fig. 3) at the time t of the $(i+1)$ th trial.

The input neurons $\theta_t^{(i+1)}$ are weightedly transmitted to 20 middle neurons,

$$\Omega^{(i+1)} = \begin{bmatrix} \omega_{0,0} & \omega_{0,1} & \dots & \omega_{0,18} & \omega_{0,19} \\ \omega_{1,0} & \omega_{1,1} & \dots & \omega_{1,18} & \omega_{1,19} \end{bmatrix} \in \mathbb{R}^{14 \times 20}, \quad (4)$$

where $\omega_{s,j} \in \mathbb{R}^{7 \times 1}$ denotes the weights of the input neurons $\theta_t^{(i+1)}$ with respect to the s th joint and the j th Gaussian motor primitive (GATE). Each GATE neuron is given by a multivariate Gaussian function:

$$\sigma_j(\theta_t^{(i+1)}, \mu_j^{(i+1)}) \sim N(\mu_j^{(i+1)}, \Sigma_j^{(i+1)}), j = 0, 1, \dots, 18, 19, \quad (5)$$

where σ_j denotes the j th GATE neuron output with respect to the input neurons $\theta_t^{(i+1)}$. The number of the GATE neurons is set to 20 for stable and fast learning track control.

The two output neurons are the feedforward torques $\tau_{ff,s}^{(i+1)}$ ($s = 0, 1$) at the time t of the $(i+1)$ th trial,

$$\tau_{ff,s}^{(i+1)}(\theta_t^{(i+1)}) = \frac{\sum_{j=0}^{j=19} (\sigma_j(\theta_t^{(i+1)}, \mu_j^{(i+1)}) (\theta_t^{(i+1)})^T \omega_{s,j}^{(i+1)})}{\sum_{j=0}^{j=19} (\sigma_j(\theta_t^{(i+1)}, \mu_j^{(i+1)}))}, \quad (6)$$

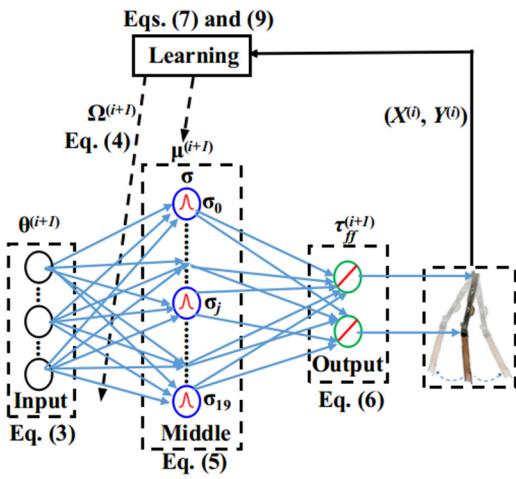


Fig. 4. Feedforward learning control based on a Gaussian motor primitive (GATE) neural network. It learns to generate the feedforward joint torques $\tau_{ff,s}^{(i+1)}$ ($s = 0, 1$) during trials.

where the feedforward torques are the normalized (i.e., $\sigma_j(\theta_t^{(i+1)}, \mu_j^{(i)})$) and weighted (i.e., $\omega_{s,j}^{(i+1)}$) sum of the GATE and input neurons. Due to its online learning, the proposed GATE feedforward learning is more computationally efficient, compared to classical regression models (e.g., RBF or LWPR) characterized by the incremental learning (Sigaud, Salaün, & Padois, 2011; Stulp & Sigaud, 2015).

The mean vector $\mu_j^{(i+1)}$ of the j th GATE neuron in the $(i+1)$ th trial is given by:

$$\mu_j^{(i+1)} = (\bar{X}_{(j\Delta t \rightarrow (j+1)\Delta t)}^{(i)})^T \in \mathbb{R}^{7 \times 1}, \quad (7)$$

where Δt (see its determinant in Appendix A.2) denotes the mean row length of the trained input matrix $X^{(i)}$ during t time steps in the i th trial,

$$X^{(i)} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{(t-2)} \\ x_{(t-1)} \end{bmatrix} \in \mathbb{R}^{t \times 7}, x_t = (\theta_t^{(i)})^T, \quad (8)$$

Where X is the trained (sensor) input matrix, in which each row x_t consists of joint positions, velocities, and accelerations (see Eq. (3)) at the time step t . The weights $\omega_{s,j}^{(i+1)}$ (see Eq. (6)) in the $(i+1)$ th trial is updated by:

$$\omega_{s,j}^{(i+1)} = ((X^{(i)})^T \delta_j^{(i+1)} X^{(i)})^{-1} ((X^{(i)})^T \delta_j^{(i+1)} Y_s^{(i)}), s = 0, 1, \quad (9)$$

where $\delta_j^{(i+1)} \in \mathbb{R}^{t \times t}$ denotes the weighted diagonal matrix based on the j th GATE neuron (see Eq. (5)),

$$\delta_j^{(i+1)} = \text{diag}(\sigma_j(x_0, \mu_j^{(i+1)}), \sigma_j(x_1, \mu_j^{(i+1)}), \dots, \sigma_j(x_{(t-2)}, \mu_j^{(i+1)}), \sigma_j(x_{(t-1)}, \mu_j^{(i+1)})), \quad (10)$$

and $Y_s^{(i)}$ the trained torque matrix with respect to the s th joint given by,

$$Y_s^{(i)} = [\tau_0, \tau_1, \dots, \tau_{(t-2)}, \tau_{(t-1)}]^T \in \mathbb{R}^{t \times 1}. \quad (11)$$

Updating the neural weights $\omega_{s,j}^{(i+1)}$ in the $(i+1)$ th trial minimizes the error $J(\omega_{s,j}^{(i+1)})$ of the weighted regression (Ting, Meier, Vijayakumar, & Schaal, 2016) based on the trained input $X^{(i)}$ and output $Y_s^{(i)}$ matrices,

$$J(\omega_{s,j}^{(i+1)}) = \frac{1}{2}(X^{(i)}\omega_{s,j}^{(i+1)} - Y_s^{(i)})^T \delta_j^{(i)}(X^{(i)}\omega_{s,j}^{(i+1)} - Y_s^{(i)}). \quad (12)$$

The minimization and derivation of Eqs. (12) and (9) can be seen in Appendix A.3. All parameters and their determinations of the GATE neural network are given in Table 2.

2.2.2. Feedback Adaptive Impedance Mechanism (AIM)

The torques $\tau_{fb} \in \mathbb{R}^2$ produced by the PD control, are given by Xiong and Manoonpong (2018):

$$\begin{aligned} \tau_{fb} &= -K_p(t)e(t) - K_d(t)\dot{e}(t), \quad K_p = [k_0, k_1]^T, \quad K_d = [d_0, d_1]^T, \\ e(t) &= q(t) - q_d(t), \quad \dot{e}(t) = \dot{q}(t) - \dot{q}_d(t), \\ \varepsilon(t) &= e(t) + \beta\dot{e}(t), \quad \beta = 0.05, \end{aligned} \quad (13)$$

where $K_p(t)$ and $K_d(t) \in \mathbb{R}^{2 \times 1}$ denote joint impedance parameters, $e(t)$ and $\dot{e}(t) \in \mathbb{R}^{2 \times 1}$ joint position and velocity errors, and $\varepsilon(t)$ joint tracking error. Note that the integral (I) control part is replaced and learned by the GATE neural network. This replacement facilitates feedforward learning for unknown arm dynamics. $\varepsilon(t)$ is the tracking error usually used in robot control (Li et al., 2018; Yang et al., 2011), in which geometry and timing errors are included in motion minimization (see Eq. (14)). Note that 'geometry' means a path (position) that joints need to follow, while 'timing' is a velocity profile along the path.

The adaptive impedance (AIM) mechanism co-minimizes the impedance efforts and motion errors of the s th ($s = 0, 1$) joint of the human-like arm during the time period T (see Fig. 5) (simplified from Xiong & Manoonpong, 2018),

$$\begin{aligned} J_o(t) &= J_c(t) + J_p(t), \\ J_c(t) &= \frac{1}{2} \int_{t-T}^t (k_s(t))^2 + (d_s(t))^2, \quad J_p(t) = \frac{1}{2} \int_{t-T}^t V_s(t), \\ V_s(t) &= M_s(q_s)(\varepsilon_s(t))^2, \end{aligned} \quad (14)$$

where $M_s(q_s)$ denotes the inertial scalar of the s th joint.

The co-minimization of AIM leads to human-like impedance adaptation of the s th ($s = 0, 1$) joint given by,

$$\begin{aligned} K_{p,s}(t) &= k_s = f_s(t)e_s(t), \quad K_{d,s}(t) = d_s = f_s(t)\dot{e}_s(t), \\ f_s(t) &= \frac{\varepsilon_s(t)}{\gamma_s(t)}, \quad \gamma_s(t) = \frac{a}{1 + b\varepsilon_s(t)^2}, \quad s = 0, 1, \end{aligned} \quad (15)$$

where $\gamma_s(t)$ is an adaptation scalar with the positive scalars $a = 0.2$ and $b = 5$. All scalars, as well as the derivation of Eqs. (14) and (15) refer to our developed human-like impedance controller (Xiong & Manoonpong, 2018). The stability proof of the online impedance adaptation law (i.e., Eq. (15)) can be seen in Xiong and Manoonpong (2018). Note that the law has been validated in finger and elbow exoskeletons for exercise (Xiong & Manoonpong, 2020a, 2020b), respectively. Note that the presented impedance adaptation is implemented based on the joint position and velocity errors in a trial, rather than between trials.

Table 2 Parameters and their determinations of GATE neural network.

Parameter	Denotation	Determination
θ	Input vector consisting of joint positions, velocities, and accelerations, Eq. (3)	Joint feedback (from motor encoders)
Ω	Learned weight matrix, Eq. (4)	Minimization based on non-linear regression, Eq. (9)
μ	Mean vector of Gaussian motor primitive (GATE), Eq. (5)	Trained input X , Eq. (8)
σ	GATE output vector, Eq. (5)	Input vector θ and learned weight matrix Ω
τ_{ff}	Feedforward torque vector, Eq. (6)	Input vector θ , GATE output vector σ , and learned weight matrix Ω
X	Trained input matrix, Eq. (8)	Joint feedback (from motor encoders)
Y	Trained torque matrix, Eq. (11)	Control torque vector

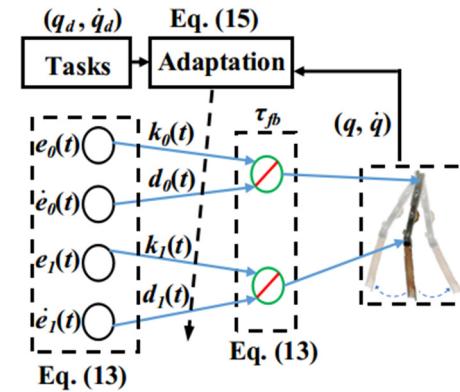


Fig. 5. Feedback control based on an adaptive impedance (AIM) mechanism. It learns to generate the feedback joint torques $\tau_{fb,s}$ ($s = 0, 1$) in a trial.

2.2.3. Online SEED model

The GATE is used to learn inverse dynamics control, while its online adaptation is achieved by the AIM mechanism. This integration facilitates micro-data (< 10 trials) supervised learning for learning and adaptation of inverse dynamics control (Hitzler et al., 2019). Its detailed computation is depicted in Algorithm 1.

Note that the online SEED is a trial-based learning model starting at the proposed AIM feedback controller, in which the impedance parameters K_p and K_d are online tuned to track the desired joint positions q_d and velocities \dot{q}_d (see Figs. B.1 and B.2).

3. Experiments

The online SEED model was tested on the two-joint robot arm (see Fig. 3). The joint feedforward torques τ_{ff} are learned by the GATE neural network of the SEED during trials (see Fig. 4), while their feedback torques τ_{fb} are online adapted by the AIM mechanism (see Fig. 5). Specifically, the joint impedance gains K_p and K_d are online tuned by the AIM mechanism in a trial (see Eq. (15)). During trials, the sensorimotor weights Ω of the GATE neural network are learned to minimize the weighted regression (see Eqs. (9) and (12)). Additionally, the proposed SEED model was compared to classical feedforward and trial error-based learning models, based on periodic and non-periodic tasks (Hitzler et al., 2019; Zhang et al., 2020). Tracking control is a typical experiment used to validate the sensorimotor learning control models (Helwa, Heins, & Schoellig, 2019; Nguyen-Tuong & Peters, 2011; Schaal & Atkeson, 2010). Each model was run for 20 trials in each task.

Algorithm 1: Online SEED

Inputs: $X^{(i)}$ and $Y_s^{(i)}$ (Eqs. (8) and (11)); $q^{(i+1)}$ and $q_d^{(i+1)}$ (Eq. (13)) ($i \geq 1$) ;
Outputs: $\tau^{(i+1)}$, $\tau_{ff}^{(i+1)}$, and $\tau_{fb}^{(i+1)}$ (Eq. (2)) ;
Learned: $\tau_{ff}^{(i+1)}$, $K_p^{(i+1)}$, and $K_d^{(i+1)}$ (Eq. (15)) ;
while $trial\ i < i_{max}$ **do**
 Obtain the trained data $X^{(i)}$ and $Y_s^{(i)}$ in the i th trial;
 Update the weights $\omega_{s,j}^{(i+1)}$ (Eq. (9)) ;
 Learn the feedforward torque $\tau_{ff}^{(i+1)}$ in the $(i + 1)$ th trial
 (Eq. (6)) ;
 while **time** $t < t_{max}$ **do**
 Adapt the impedance parameters $K_p^{(i+1)}(t)$ and
 $K_d^{(i+1)}(t)$ (Eq. (15)) ;
 Compute the feedback torques $\tau_{fb}^{(i+1)}(t)$ (Eq. (13)) ;
 Compute the control torques
 $\tau^{(i+1)}(t) = \tau_{ff}^{(i+1)}(t) + \tau_{fb}^{(i+1)}(t)$;
 end
 Save the trained data $X^{(i+1)}$ and $Y_s^{(i+1)}$ in the $(i + 1)$ th
 trial;
end

1 3.1. Periodic task

2 Compared to the adaptive feedback torques τ_{fb} , the learned
3 feedforward torques τ_{ff} contribute mainly to the joint torques τ
4 (see Fig. 6(A)). This is because τ_{fb} is determined by the reduced
5 impedance gains K_p and K_d . The reductions in K_p and K_d result
6 from the decreased tracking errors e (see the 'After learning' area
7 in Fig. 6(A)) owing to the increasing and learned feedforward
8 torques τ_{ff} . The increasing feedforward and decreasing feedback
9 torques quickly and stably reduce the position error e after trial
10 learning. After the 5th trial, the feedforward learning converges
11 where the sensorimotor weights Ω of the GATE neural network
12 are learned (see Fig. 6(B) and (C)). The fast (i.e., <10 trials)
13 sensorimotor learning control results from integrating feedforward
14 neural learning and feedback impedance adaptation control (see
15 Figs. 4 and 5). The integration contributes to conventional learning
16 control models (Franklin, Burdet, Peng Tee, Osu, Chew, Milner,
17 & Kawato, 2008; Li et al., 2018; Proietti, Morel, Roby-Brami, &
18 Jarrassé, 2017; Sigaud et al., 2011) by using the AIM (see Fig. 5)
19 to speed up sensorimotor learning. The joint impedance gains
20 K_p and K_d are initially high to accelerate sensorimotor learning,
21 subsequently decreasing due to increasing feedforward neural
22 learning control (see K_p , K_d , and τ_{ff} in Fig. 6). This is comparable
23 to human joint impedance modulation in sensorimotor learning
24 control (Osu et al., 2002; Santina et al., 2017). It is believed
25 that the strategy guarantees fast and stable human sensorimotor
26 learning control (Franklin & Wolpert, 2011; Wolpert & Flanagan,
27 2016). The strategy increases the suitability of our proposed feed-
28 forward model for more efficient online learning, compared to
29 the incremental learning used by classic sensorimotor regression
30 models (Sigaud et al., 2011; Stulp & Sigaud, 2015).

31 The online SEED model resembles human sensorimotor learning
32 and adaptation (Maeda, Cluff, Gribble, & Pruszyński, 2018;
33 Maeda, Gribble, & Pruszyński, 2020; Yeo, Franklin, & Wolpert,
34 2016). The process starts at feedback impedance control (see
35 the sensorimotor weights $\Omega = 0$ in trial 1 of Fig. 6(C)). It
36 drives feedforward sensorimotor learning, which in turn tunes
37 (i.e., decreases) feedback impedance responses during trials of the
38 task. The learned feedforward torques and feedback impedance

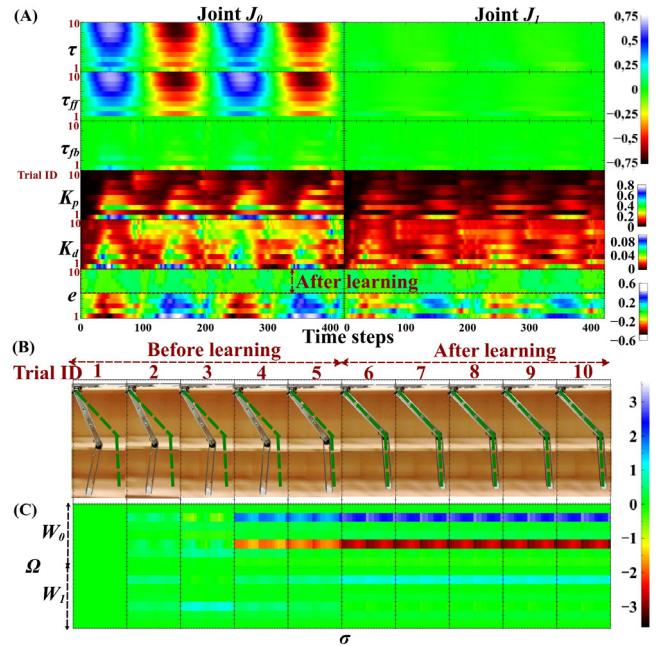


Fig. 6. Periodic task (see Eq. (B.1) and Fig. B.1) undertaken by the online SEED model during 10 trials. The learning and adaptation applies to the two-joint arm joints J_0 and J_1 (see Fig. 3). (A) Overall torques τ consist of the feedforward τ_{ff} and feedback τ_{fb} torques. τ_{ff} is learned by the GATE neural network (see Fig. 4 and Eq. (9)), while τ_{fb} are determined by impedance gains K_p and K_d tuned by the AIM mechanism based on the position error e (see Fig. 5 and Eq. (15)). (B) Snapshots during the learning trials. (C) Sensorimotor weights Ω learned during 10 trials ($W_i = \omega_{i,0-19} \in \mathbb{R}^{7 \times 20}, i = 0, 1$, see Eq. (4)). Its learning is based on the 20 GATE neurons σ (see Fig. 4).

gains are displayed in Fig. C.2. The experimental video can be seen at footnote.² 39
40

3.2. Non-periodic task 41

It can be seen that the online SEED model also enables the two-joint arm to reduce the position error e (see the 'After learning' area in Fig. 7) by increasing feedforward and decreasing feedback torques during trials. However, the model learns and adapts faster (i.e., after the 4th trial) in the non-periodic task, compared to the periodic task. No matter whether the task is periodic or non-periodic, the feedback impedance contribution is reduced through increased feedforward learning. This is because the increased learned feedforward control reduces the tracking errors e that lead to a decrease in the impedance gains K_p and K_d . This is comparable to those found in human motor control and learning. Humans usually use higher arm joint stiffness in the learning phase of a new motor task. Afterwards, the joint stiffness is gradually reduced while the feedforward control is learned (Osu et al., 2002; Santina et al., 2017). It is believed that this adaptation guarantees human sensorimotor learning control stability and versatility. The experimental video can be seen at footnote.³ The learned feedforward torques and feedback impedance gains are displayed in Fig. C.3. 42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Another set of experiments was performed, in which a bottle of (incompletely full) water was attached at the robot arm tip, acting as external disturbances (see Fig. 8). The robotic joints were controlled to learn to track the same minimal jerk trajectory when subjected to these disturbances. We can see that

² <https://www.youtube.com/watch?v=e6UD3aYljaw>

³ <https://www.youtube.com/watch?v=0nbN-3YA184>

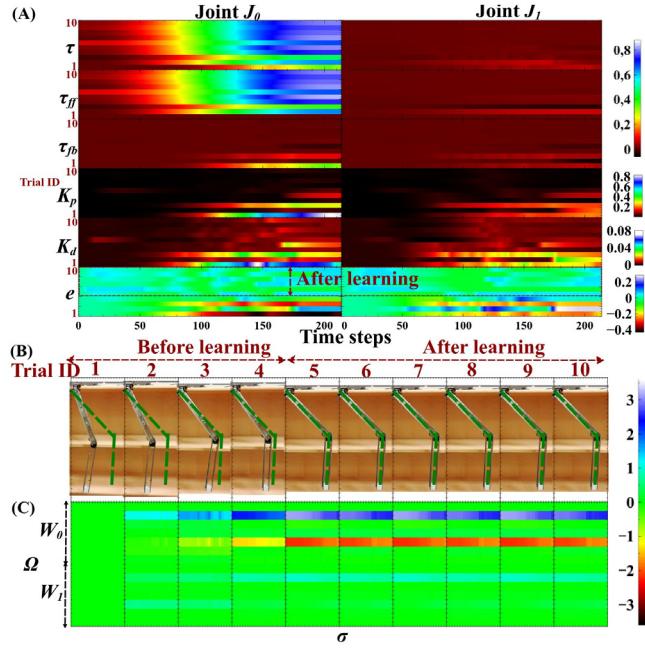


Fig. 7. Non-periodic task (see Eq. (B.2) and Fig. B.2) performed by the online SEED model over 10 trials. Learning and adaptation apply to the two-joint arm joints J_0 and J_1 (see Fig. 3). (A) Overall torques τ consist of the feedforward τ_{ff} and feedback τ_{fb} torques. τ_{ff} is learned by the GATE neural network (see Fig. 4 and Eq. (9)), while τ_{fb} are determined by impedance gains K_p and K_d tuned by the AIM mechanism based on the position error e (see Fig. 5 and Eq. (15)). (B) Snapshots during the learning trials. (C) Sensorimotor weights Ω learned during 10 trials ($W_i = \omega_{i,0-19} \in \mathbb{R}^{7 \times 20}$, $i = 0, 1$, see Eq. (4)). Its learning is based on the 20 GATE neurons σ (see Fig. 4).

this learning remains stable in the new task, and the learned joint torques are larger than those without a bottle of water. In feedback control, the impedance parameters are online tuned to fight against uncertain disturbances caused by water fluid dynamics. Here the learning performance is measured by the average absolute position error $|\bar{e}|(i)$ in trial i ,

$$\bar{e}(i) = \sum_{s=0}^{t=m} \frac{|e|_s}{m}, s = 0, 1, \quad (16)$$

where t denotes the time steps, s is the joint $J_{0,1}$ IDs (see Fig. 3) (see Fig. 9).

In addition to human-like sensorimotor learning and adaptation, the online SEED model enables the two-joint robot arm (see Fig. 3) to explore its natural dynamics.

3.3. Exploring arm dynamics

We can see that the learned J_0 torque is much greater than the J_1 torque (see τ in Fig. 6(A)). This is because the proximal joint J_0 takes on more static and dynamic loads in the tracking tasks (see Fig. 3). For instance, the proximal joint J_0 is subjected to heavier loads (the two links $L_{0,1}$ and joint J_1), compared to the distal joint J_1 (the link L_1). The tracking control of the distal joint J_1 is driven mainly by the interaction torque $\tau_{in,1}$ between the joints $J_{0,1}$,

$$\tau_{in,1} = \tau_{a,1} - \tau_1, \quad (17)$$

where $\tau_{a,1}$ and τ_1 denote the sensed and control joint torques (see Fig. 10). Interestingly, this finding is comparable to the leading joint hypothesis (LJH) of human motor control (Amabile & Schmideler, 2013; Dounskaya, 2010). Therefore, the proposed SEED model enables the two-joint robot arm to explore



Fig. 8. Experimental setup with a bottle of (incompletely full) water. A 0.11 kg weight bottle was attached at the robot arm tip, acting as external disturbances.

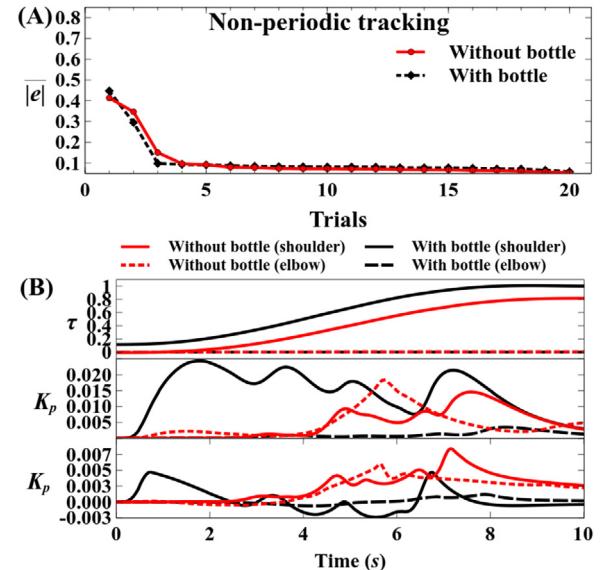


Fig. 9. Learning comparison in the non-periodic task (see Eq. (B.2) and Fig. B.2). The comparison was performed when the robotic tip was attach with and without a bottle of water (0.11 kg), respectively. (A) Average absolute position errors $|\bar{e}|$ during the trials. (B) Learned torques τ , impedance parameters K_p and K_d at the 20th trial.

natural arm dynamics (see Fig. 11). The exploration reveals a proximal-distal control gradient over joints (see Section 3.3).

In reducing the position errors e , the online SEED model makes the two-joint arm (see Fig. 3) explore the joint control dynamics in the periodic and non-periodic tracking tasks (see Figs. B.1 and B.2). Specifically, the dynamics exploration results in the joint control effort gradient of the robot arm. The joint control effort is

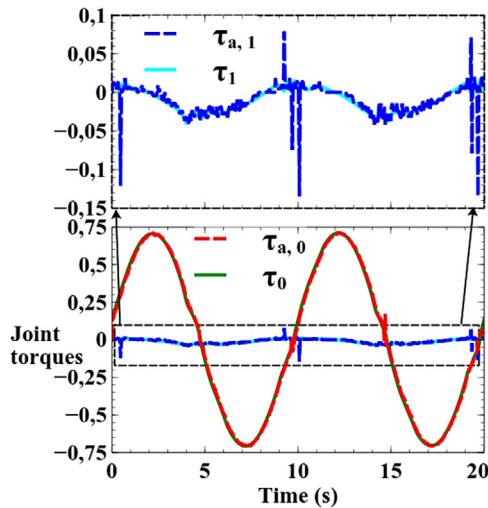


Fig. 10. Joint torques of the proximal J_0 and distal J_1 joints (see Eq. (17)).

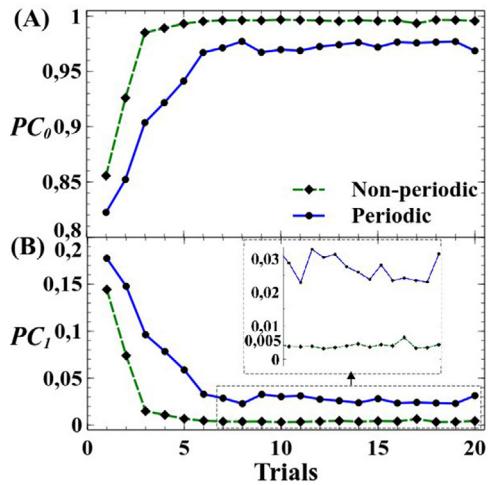


Fig. 11. Proximal-distal control gradient. The gradient is quantified by the joint control power consumption ratios PC_s (see Eq. (18)).

1 quantified by the joint control power consumption ratios PC_s ,

$$2 \quad PC_s = \frac{P_s}{\sum_{s=0}^{t=m} P_s}, \quad P_s = \frac{\sum_{t=1}^{t=m} VI(t)}{m}, \quad s = 0, 1, \quad (18)$$

3 where P_s denotes the average control power consumption of the
4 joint J_s during time steps m .

5 We can see that the online SEED model enables the two-
6 joint arm to evolve into different control effort gradients in the
7 periodic and non-periodic tasks (see Fig. 11). The proximal joint
8 J_0 consumes much more power, compared to the distal joint J_1
9 (see Fig. 3). This is because the proximal joint J_0 takes on more
10 static and dynamic loads in the tracking tasks. For instance, the
11 proximal joint J_0 is subjected to heavier loads (i.e., the two links
12 $L_{0,1}$ and joint J_1), compared to the distal joint J_1 (i.e., the link
13 L_1). Therefore, the control effort needed for the distal joint J_1 is
14 relatively less (i.e., $PC_1 < 0.04$, see Fig. 11(B)) after trial learning.
15 Interestingly, the control power consumption ratio of the distal
16 joint J_1 is comparatively lower (i.e., $PC_1 < 0.005$, see Fig. 11(B))
17 in the non-periodic task (see Eq. (B.2)), compared to the periodic

task (see Eq. (B.1)). This is because the periodic task is more dynamic, therefore requiring greater distal joint control effort. Here we hypothesize a proximal-distal control gradient of human-like arm movement control: more proximal joints are, the greater control effort required. In contrast, distal joints are controlled mainly by interaction torques between joints and limbs resulting from natural arm dynamics (see Fig. 10). The presented proximal-distal gradient is comparable to the leading joint hypothesis (LJH) of human motor control (Ambike & Schmiedeler, 2013; Dounskaya, 2010).

3.4. Comparisons to classic learning models

The compared models are feedforward learning with various impedance gains, and error-based learning using the AIM. Specifically, the feedforward learning models consist of the GATE neural network learning with low (GATE-L), intermediate (GATE-I), and high (GATE-H) impedance gains. This set of feedforward learning models characterizes classical sensorimotor regression learning (Hitzler et al., 2019; Sigaud et al., 2011; Stulp & Sigaud, 2015). Comparison with another compared model highlights trial-and-error based learning (Franklin et al., 2008; Li et al., 2018; Proietti et al., 2017; Zhang et al., 2020). The model uses iterative learning control with the AIM (IT-AIM). More details of compared model implementations can be seen in Appendix C.

We can see that the proposed online SEED model enables the human-like arm to reduce the error $|\epsilon|$ faster during trials, compared to other compared models (see Figs. 12 and 13). This is because the SEED model not only learns joint feedforward torques during trials, but also adapts the joint impedance gains K_p and K_d in a trial. In contrast, the low impedance constant gains result in a low learning speed, while unstable learning results from high impedance constant gains (see GATE-L and GATE-H in Figs. 12 and 13). Even if the different impedance constant gains are applied to the different arm joints (see GATE-I in Figs. 12 and 13), the learning speed is still slower than that of the online SEED model. Taken together, the compared GATE models with the three sets of the joint impedance constants cause a dilemma between fast and stable convergences, like classic regression models (e.g., RBF and LWPR) do (Sigaud et al., 2011; Stulp & Sigaud, 2015). Therefore, low joint impedance gains are utilized to obtain slow and stable convergences in robot sensorimotor learning (Helwa et al., 2019; Nguyen-Tuong & Peters, 2011; Santina et al., 2017). Moreover, conventional regression models such as RBF and LWPR are computationally inefficient owing to their incremental learning, compared to the online learning used in our proposed SEED model (see Fig. 2). This is because they do not have the adaptive impedance (AIM) mechanism (see Fig. 5) presented in this paper.

Although the same feedback adaptation (i.e., the AIM mechanism) is applied, the online SEED model reduces the position error $|\epsilon|$ faster, compared to the IT-AIM. This is because the GATE neural network of the SEED learns the sensorimotor mapping (i.e., weights Ω) through integrating joint accelerations, velocities, and positions (see Fig. 6). However, the IT-AIM individually learns the joint feedforward torques based on the joint position error e (see Eq. (C.1)). This leads to slower learning and joint jerk behaviors (see Fig. C.2(B) and (C)), compared to the online SEED model. This is because the joint control is interacted (see Fig. 10), and such interaction effects (e.g., torques) should be integrally taken into account. Moreover, the learned impedance parameters K_p and K_d are decreased through the increased feedforward contribution from the GATE neural network learning. More predictive feedforward and less feedback contributions guarantee sensorimotor learning control stability and smoothness. Furthermore, the learned proximal joint (i.e., J_0) torque is much greater than that of the distal joint (i.e., J_1) by applying the online SEED

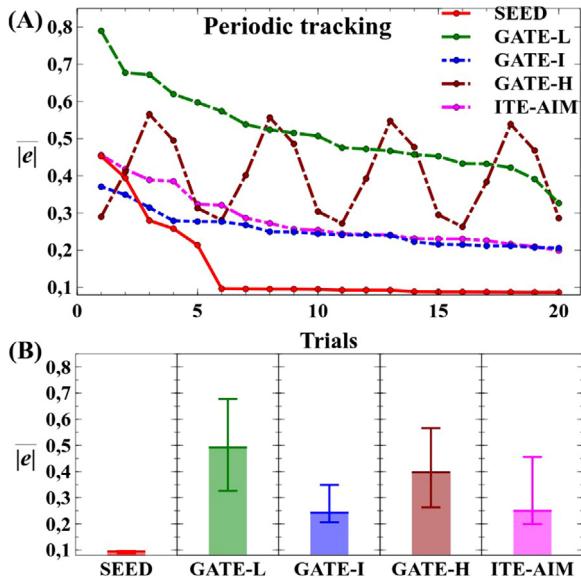


Fig. 12. Learning comparison in the periodic task (see Eq. (B.1) and Fig. B.1). (A) Average absolute position errors $|e|$ during the trials. (B) Statistics of $|e|$. Each model was run 20 trials. The impedance constant gains of the GATE-L, GATE-I, and GATE-H can be seen in Table 3. The IT-AIM model combines iterative learning control with the AIM mechanism (see Eqs. (15) and (C.1)).

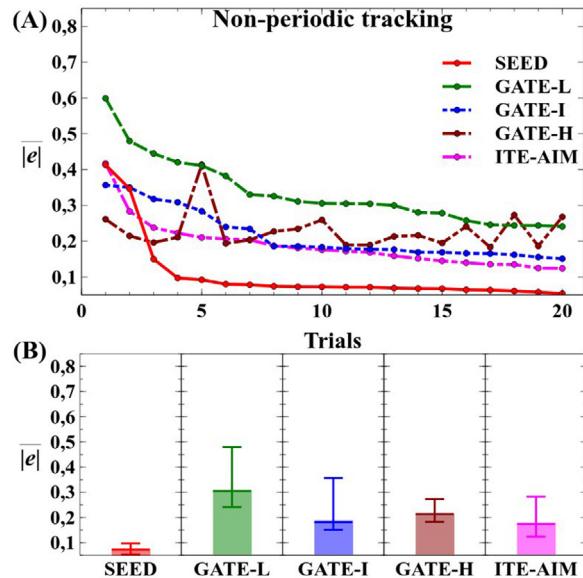


Fig. 13. Learning comparison in the non-periodic task (see Eq. (B.2) and Fig. B.2). (A) Average absolute position errors $|e|$ during the trials. (B) Statistics of $|e|$. Each model was run 20 trials. The impedance constant gains of the GATE-L, GATE-I, and GATE-H can be seen in Table 3. The IT-AIM model combines iterative learning control with the AIM mechanism (see Eqs. (15) and (C.1)).

model. The model allows the human-like arm to explore its natural dynamics by evolving different control effort gradients (see Fig. 11). The above-mentioned comparisons in the periodic task are similar to those in the non-periodic (see Fig. C.3). The experimental videos can be seen at footnotes.^{4,5}

4. Discussion and conclusion

As the results show, the proposed SEED model paves a way forward for achieving fast (e.g., <10 trials) sensorimotor learning control. Its learning task is faster and more stable, compared to the state-of-the-art (SOA) models for sensorimotor learning control in robotics and computational neuroscience. Moreover, the proposed model allows for arm dynamics exploration, indicating a proximal-distal control gradient comparable to the leading joint hypothesis (LJH) of human motor control (Ambike & Schmiedeler, 2013; Dounskoia, 2010). Such an indication has not been presented by conventional control models (Hitzler et al., 2019; Pan & Yu, 2017; Santina et al., 2017; Yu, Shi, Lin, & Yu, 2020).

4.1. Robotics

Integrating feedforward and feedback control in robotic arms increases control smoothness and robustness (Kuo, 2002; Takiyama & Sakai, 2016), compared to pure feedforward or feedback control. However, most SOA models focus mainly on learning feedforward control policies. The major feedforward learning models involve supervised sensorimotor regression (Gao et al., 2018), iterative learning control (ILC) (Proietti et al., 2017), neural learning (Pan & Yu, 2017), and reinforcement learning (Haith & Krakauer, 2013). However, the SOA learning models usually choose low gain feedback control. For example, Santina et al. (2017) proposed an ILC model with low

impedance gains for soft robot arm control (Santina et al., 2017). Unfortunately, it is difficult to select proper learning rates and impedance gains to guarantee learning control stability in different tasks. Alternatively, Helwa et al. (2019) presented a Gaussian Process (GP) regression model with low impedance gains for high accuracy robot arm tracking (Helwa et al., 2019). This is because low gain feedback control guarantees learning stability. However, the GP regression learning control is computationally expensive owing to its incremental data learning. Moreover, the low gain feedback control slows down regression learning, leading to learning control inefficiency (see GATE-L in Figs. 12 and 13). In contrast, an increase in impedance gains can initially speed up sensorimotor learning. However, it may cause learning control instability (see GATE-H in Figs. 12 and 13). To address this dilemma, Mathew et al. (2019) integrated a Gaussian Mixture Model (GMM) and variable impedance control for robotic arm learning continuous contact tasks (Mathew et al., 2019). However, the free and maximum joint stiffness gains need to be carefully predefined to guarantee learning control stability. Moreover, their proposed model does not show the exploration of natural arm dynamics. Exploring such dynamics can uncover a novel control strategy that reduces control redundancy (see Fig. 11). Taken together, most SOA models for robot learning impedance gains fail to guarantee sensorimotor learning control stability (Kronander & Billard, 2016). Extensive reviews on robot arm learning control can be seen in Chatzilygeroudis et al. (2020), Nguyen-Tuong and Peters (2011), Ravichandar et al. (2020), Schaal and Atkeson (2010), Sigaud et al. (2011) and Stulp and Sigaud (2015).

4.2. Computational neuroscience

Franklin (2015) proposed that feedback-based modulated joint impedance provides a novel way to ensure human sensorimotor learning control stability (Franklin, 2015). Recent experimental works have demonstrated that the interaction between feedforward and feedback is bidirectional (Maeda et al., 2018, 2020). However, as yet, there is little computational understanding

⁴ <https://www.youtube.com/watch?v=yHeqmGPABlY>
⁵ <https://www.youtube.com/watch?v=mPR-EqkTolg>

on how feedback and feedforward learning bidirectionally contributes to human-like sensorimotor learning control stability. Here, we hypothesize the learning strategy used in our proposed SEED model (see Fig. 2) (Wolpert, Diedrichsen, & Flanagan, 2011; Wolpert & Flanagan, 2016; Yeo et al., 2016):

- Sensorimotor learning control starts at feedback impedance adaptation control (i.e., the AIM mechanism, see Fig. 5);
- Feedback impedance control drives feedforward learning sensorimotor mapping (i.e., the GATE neural network, see Fig. 4);
- Increased feedforward contribution decreases feedback impedance responses to guarantee learning control stability (see impedance gains K_p and K_d in Figs. 6 and 7).

Our proposed SEED model allows the two-joint robot arm to explore unknown natural arm dynamics, revealing a novel sensorimotor control strategy, i.e., proximal-distal gradient (see Fig. 11) (Ambike & Schmiedeler, 2013; Dounskaya, 2010). Besides, some variations of the proximal-distal control gradient have been found in the three-jointed limb control of an animal's movement (Charles, Cappellari, & Hutchinson, 2018; Lai, Biewener, & Waking, 2019). Experiments show that their proximal joints are driven mainly by active (neural) control, while a distal one resists against external perturbations via passive dynamics (e.g., tendons). This is because central neural control is too intrinsically slow to control distal joints against quick perturbations. Therefore, this control gradient allows for stable limb movement, which has been proved in legged robots (Lee & Biewener, 2011; Xiong, Wörgötter, & Manoonpong, 2016). This gradient paves a way forward for reducing human sensorimotor control redundancy. The above-mentioned contributions by our proposed SEED model are not shown and realized by the prominent computational models such as optimal feedback control, Bayesian integration, and motor primitives. Extensive reviews on human arm learning control models can be seen in Campos and Calado (2009), Franklin and Wolpert (2011), Haith and Krakauer (2013) and Wolpert and Flanagan (2016).

4.3. Conclusion

Micro-data (<10 trials) in learning inverse dynamics is a challenging problem, which pure data-driven and physics-based methods fail to solve (Bogdanovic et al., 2020; Hitzler et al., 2019; Ren & Ben-Tzvi, 2020; Zhang et al., 2020; Zhou, Tao, Paszke, Stojanovic, & Yang, 2020). To address this challenge, both methods are integrated to achieve micro-data supervised learning in our proposed framework. One of its limitation is learning instability caused by online impedance adaptation when being applied to high degree-of-freedom robotic arms. However, this limitation can be tackled by applying the proposed method to an arm end-effector, rather than joints in this paper.

In spite of its impressiveness and novelty, there are still an open question and missing validations remain for further research to answer and provide. First, how does the number of the GATE neurons impact the performance of more complex learning control tasks (e.g., polishing)? Currently the number is set to 20 neurons for the stable and fast learning track control presented in this paper. Second, the modified SEED model will be developed for Cartesian space control, rather than the joint space control presented in this paper. Such modification will enhance versatile motor skill learning (e.g., writing) in high DOF robot arms (e.g., hand writing and dual arm manipulation tasks). Third, the SEED model will be further developed for redundant muscle

space control, such that the computational understanding of muscle functions will be advanced in complex motor skill learning for human-robot interaction and collaboration tasks. Last but not least, a source of errors results from the direct-actuated robot arm, in which cheap servo motors were used to drive joints. The arm will be extended to a quasi drive-drive design for accurate tracking in future.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Cao Danh Do for robotic arm design. This work is supported by the Human Frontier Science Program (Grant No. RGP0002/2017, P.M. = Project Co-PI), the Brødrene Hartmanns Fund (No. A36775, X.X. = Project PI), the Thomas B. Thriges Fund (No. 7648-2106, X.X. = Project PI), and a startup grant on bio-inspired robotics from the Vidyasirimedhi Institute of Science and Technology (VISTEC, P.M. = Project PI).

Appendix A. Feedforward learning of the online SEED

A.1. GATE neural network parameters

All parameters and their determinations of the GATE neural network are given in Table 2.

A.2. Determining Δt in Eq. (7)

The row length Δt is determined by the overall row amount t of the trained input matrix $X^{(i)}$,

$$\Delta t = \begin{cases} Quo(t/20), & j \neq 19 \\ Quo(t/20) + Re(t/20), & j = 19 \end{cases}, \quad (A.1)$$

where the j th part of the trained input matrix $X^{(i)}$ is used to calculate the mean vector $\mu_j^{(i+1)}$ with respect to the j th GATE neuron (see details in Eq. (7), $j = 0, 1, \dots, 18, 19$). Re and Quo denote the remainder and quotient functions, respectively.

A.3. Minimization and derivation of Eqs. (9) and (12)

The feedforward GATE neural learning is designed to minimize the error $J(\omega_{s,j}^{(i+1)})$ of the weighted regression based on the trained input $X^{(i)}$ and output $Y_s^{(i)}$ matrices (see Eqs. (3)–(12)),

$$J(\omega_{s,j}^{(i+1)}) = \frac{1}{2}(X^{(i)}\omega_{s,j}^{(i+1)} - Y_s^{(i)})^T \delta_j^{(i)} (X^{(i)}\omega_{s,j}^{(i+1)} - Y_s^{(i)}). \quad (A.2)$$

Considering the definition of $J(\omega_{s,j}^{(i+1)})$, we obtain

$$J(\omega_{s,j}^{(i+1)}) = \frac{1}{2} \left((\omega_{s,j}^{(i+1)})^T (X^{(i)})^T \delta_j^{(i)} X^{(i)} \omega_{s,j}^{(i+1)} - (\omega_{s,j}^{(i+1)})^T (X^{(i)})^T \delta_j^{(i)} Y_s^{(i)} - (Y_s^{(i)})^T \delta_j^{(i)} X^{(i)} \omega_{s,j}^{(i+1)} + (Y_s^{(i)})^T \delta_j^{(i)} Y_s^{(i)} \right), \quad (A.3)$$

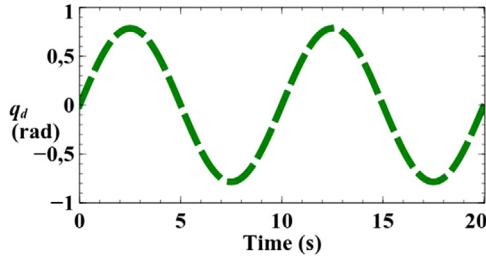


Fig. B.1. Desired joint positions $q_d(t)$ for the periodic task (see Eq. (B.1)).

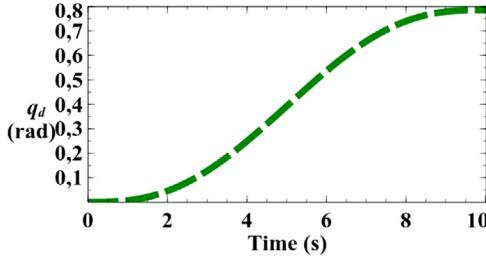


Fig. B.2. Desired joint position $q_d(t)$ of the non-periodic task (see Eq. (B.2)).

1 and its derivative is given by,

$$\begin{aligned} \frac{\partial J(\omega_{s,j}^{(i+1)})}{\partial (\omega_{s,j}^{(i+1)})} &= \frac{1}{2} \left((X^{(i)})^T \delta_j^{(i)} X^{(i)} \omega_{s,j}^{(i+1)} + \right. \\ &\quad \left((X^{(i)})^T \delta_j^{(i)} X^{(i)} \right)^T \omega_{s,j}^{(i+1)} \\ &\quad (X^{(i)})^T \delta_j^{(i)} Y_s^{(i)} - (X^{(i)})^T (\delta_j^{(i)})^T Y_s^{(i)} \right) . \\ \delta_j^{(i)T} = \delta_j^{(i)} &= \frac{1}{2} \left(2(X^{(i)})^T \delta_j^{(i)} X^{(i)} \omega_{s,j}^{(i+1)} - \right. \\ &\quad \left. 2(X^{(i)})^T \delta_j^{(i)} Y_s^{(i)} \right) \end{aligned} \quad (A.4)$$

3 Let $\frac{\partial J(\omega_{s,j}^{(i+1)})}{\partial (\omega_{s,j}^{(i+1)})} = 0$, then we get,

$$\omega_{s,j}^{(i+1)} = ((X^{(i)})^T \delta_j^{(i)} X^{(i)})^{-1} ((X^{(i)})^T \delta_j^{(i)} Y_s^{(i)}), s = 0, 1, \quad (A.5)$$

5 where $\delta_j^{(i)}$ is a diagonal matrix (see Eq. (10)).

6 Appendix B. Tasks

8 B.1. Periodic

9 The task involves controlling the two arm joints $J_{0,1}$ (see Fig. 3) to follow a periodic trajectory (see Fig. B.1) given by,

$$11 q_d(t) = \frac{\pi}{4.0} \sin(4\pi \frac{t}{20.0}), t \in [0, 20.0]. \quad (B.1)$$

12 B.2. Non-periodic

13 The non-periodic task involves tracking a minimum-jerk trajectory given by Flash and Hogan (1985) (see Fig. B.2),

$$14 q_d(t) = \frac{\pi}{4.0} [10(\frac{t}{10.0})^3 - 15(\frac{t}{10.0})^4 + 6(\frac{t}{10.0})^5], t \in [0, 10.0]. \quad (B.2)$$

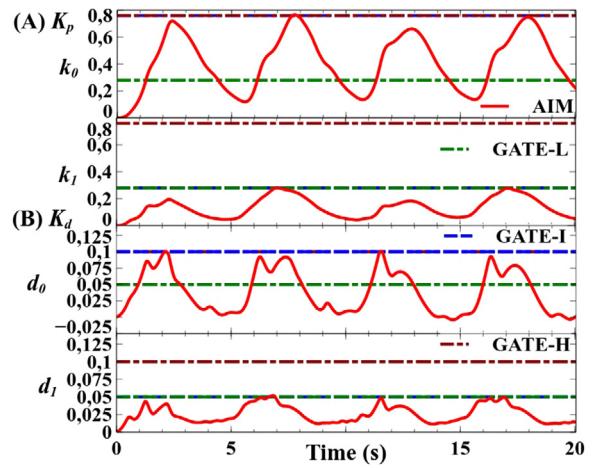


Fig. C.1. Selection of constant impedance gains based on the AIM mechanism (see Fig. 5 and Eq. (15)). (A) Impedance parameter $K_p = [k_0, k_1]$. (B) Impedance parameter $K_d = [d_0, d_1]$.

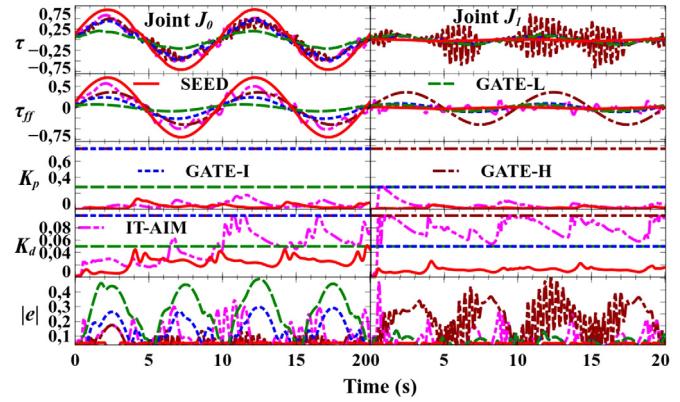


Fig. C.2. Feedforward (i.e., τ_{ff}) learning and impedance (i.e., K_p and K_d) adaptation in the periodic task (see Eq. (B.1) and Fig. B.1). The Joint $J_{0,1}$ arguments are learned in the 20th trial.

Table 3

Constant Impedance Gains (see Fig. C.1).

Models	$K_p = [k_0, k_1]$	$K_d = [d_0, d_1]$
GATE-L	$k_0 = k_1 = 0.28$	$d_0 = d_1 = 0.05$
GATE-I	$k_0 = 0.76, k_1 = 0.28$	$d_0 = 0.1, d_1 = 0.05$
GATE-H	$k_0 = k_1 = 0.76$	$d_0 = d_1 = 0.1$

Appendix C. Compared learning models

16

The first compared learning model consists of iterative learning and AIM mechanisms (IT-AIM). The iterative learning mechanism is used to learn the joint feedforward torques $\tau_{ff}^{(i+1)}$ in trial $(i+1)$,

$$17 \tau_{ff,s}^{(i+1)}(t) = \tau_{ff,s}^{(i)}(t) + (\alpha e_s(t)), \alpha = -0.15, s = 0, 1, \quad (C.1)$$

where α denotes the iterative rate, s is the joint ID. In addition to the learned feedforward torques τ_{ff} , the feedback torques τ_{fb} are tuned by the AIM mechanism in a trial (see Fig. 5 and Eq. (15)). Note that the IT-AIM starts at AIM feedback control (see K_p and K_d of the first learning trial in Fig. C.1).

Other feedforward learning models are achieved by using the GATE neural network with low (GATE-L), intermediate (GATE-I), and high (GATE-H) impedance gains (see Table 3 and Fig. C.1). The constant gains are chosen based on the impedance parameters K_p and K_d tuned by the AIM mechanism.

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

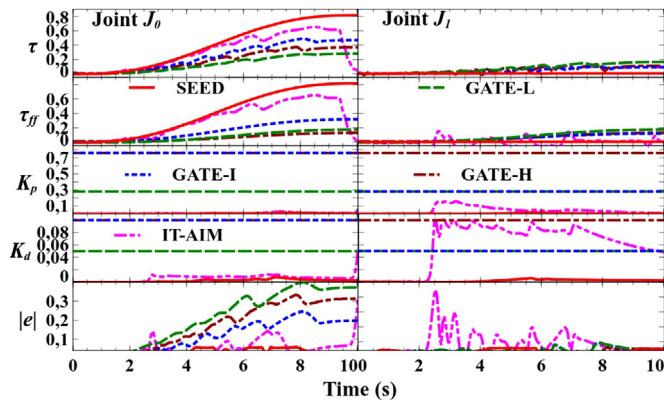


Fig. C.3. Feedforward (i.e., τ_{ff}) learning and impedance adaptation (i.e., K_p and K_d) in the non-periodic task (see Eq. (B.2) and Fig. B.2). The impedance constant gains of the GATE-L, GATE-I, and GATE-H can be seen in Table 3. The Joint $J_{\theta,1}$ arguments are learned in the 20th trial.

References

- Ambike, S., & Schmiedeler, J. P. (2013). The leading joint hypothesis for spatial reaching arm motions. *Experimental Brain Research*, 224(4), 591–603. <http://dx.doi.org/10.1007/s00221-012-3335-x>.
- Bogdanovic, M., Khadiv, M., & Righetti, L. (2020). Learning variable impedance control for contact sensitive tasks. *IEEE Robotics and Automation Letters*, 5(4), 6129–6136. <http://dx.doi.org/10.1109/LRA.2020.3011379>.
- Burdet, E., Tee, K. P., Mareels, I., Milner, T. E., Chew, C. M., Franklin, D. W., et al. (2006). Stability and motor adaptation in human arm movements. *Biological Cybernetics*, 94(1), 20–32. <http://dx.doi.org/10.1007/s00422-005-0025-9>.
- Campos, F., & Calado, J. M. F. (2009). Approaches to human arm movement control-A review. *Annual Reviews in Control*, 33(1), 69–77. <http://dx.doi.org/10.1016/j.arcontrol.2009.03.001>.
- Charles, J. P., Cappellari, O., & Hutchinson, J. R. (2018). A dynamic simulation of musculoskeletal function in the mouse hindlimb during trotting locomotion. *Frontiers in Bioengineering and Biotechnology*, 6(61), <http://dx.doi.org/10.3389/fbioe.2018.00061>.
- Chatzilygeroudis, K., Vassiliades, V., Stulp, F., Calinon, S., & Mouret, J. (2020). A survey on policy search algorithms for learning robot controllers in a handful of trials. *IEEE Transactions on Robotics*, 36(2), 328–347. <http://dx.doi.org/10.1109/TRO.2019.2958211>.
- Dounskaja, N. (2010). Control of human limb movements: the leading joint hypothesis and its practical applications. *Exercise and Sport Sciences Reviews*, 38(4), 201–208. <http://dx.doi.org/10.1097/JES.0b013e3181f45194>.
- Flash, T., & Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *The Journal of Neuroscience*, 5(7), 1688–1703. <http://dx.doi.org/10.1523/JNEUROSCI.05-07-01688.1985>.
- Franklin, D. W. (2015). Impedance control: Learning stability in human sensorimotor control. In 2015 37th annual international conference of the IEEE engineering in medicine and biology society (pp. 1421–1424). <http://dx.doi.org/10.1109/EMBC.2015.7318636>.
- Franklin, D. W., Burdet, E., Peng Tee, K., Osu, R., Chew, C.-M., Milner, T. E., et al. (2008). CNS learns stable, accurate, and efficient movements using a simple algorithm. *The Journal of Neuroscience*, 28(44), 11165–11173. <http://dx.doi.org/10.1523/JNEUROSCI.3099-08.2008>.
- Franklin, D. W., & Wolpert, D. M. (2011). Computational mechanisms of sensorimotor control. *Neuron*, 72(3), 425–442. <http://dx.doi.org/10.1016/j.neuron.2011.10.006>.
- Gao, T., Yin, S., Gao, H., Yang, X., Qiu, J., & Kaynak, O. (2018). A locally weighted project regression approach-aided nonlinear constrained tracking control. *IEEE Transactions on Neural Networks and Learning Systems*, 29(12), 5870–5879. <http://dx.doi.org/10.1109/TNNLS.2018.2808700>.
- Haith, A. M., & Krakauer, J. W. (2013). *Model-based and model-free mechanisms of human motor learning - progress in motor control* (pp. 1–21). New York, NY: Springer New York.
- Helwa, M. K., Heins, A., & Schoellig, A. P. (2019). Provably robust learning-based approach for high-accuracy tracking control of Lagrangian systems. *IEEE Robotics and Automation Letters*, 4(2), 1587–1594. <http://dx.doi.org/10.1109/LRA.2019.2896728>.
- Hitzler, K., Meier, F., Schaal, S., & Asfour, T. (2019). Learning and adaptation of inverse dynamics models: A comparison. In 2019 IEEE-RAS 19th international conference on humanoid robots (pp. 491–498). <http://dx.doi.org/10.1109/Humanoids43949.2019.9035048>.
- Karakasiliotis, K., Thandiackal, R., Melo, K., Horvat, T., Mahabadi, N. K., Tsitkov, S., et al. (2016). From cineradiography to biorobots: an approach for designing robots to emulate and study animal locomotion. *Journal of the Royal Society Interface*, 13(119), Article 20151089. <http://dx.doi.org/10.1098/rsif.2015.1089>.
- Kronander, K., & Billard, A. (2016). Stability considerations for variable impedance control. *IEEE Transactions on Robotics*, 32(5), 1298–1305. <http://dx.doi.org/10.1109/TRO.2016.2593492>.
- Kuo, A. D. (2002). The relative roles of feedforward and feedback in the control of rhythmic movements. *Motor Control*, 6(2), 129–145. <http://dx.doi.org/10.1123/mcj.6.2.129>.
- Lai, A. K. M., Biewener, A. A., & Wakeling, J. M. (2019). Muscle-specific indices to characterise the functional behaviour of human lower-limb muscles during locomotion. *Journal of Biomechanics*, 89, 134–138. <http://dx.doi.org/10.1016/j.jbiomech.2019.04.027>.
- Lee, D. V., & Biewener, A. A. (2011). Bigdog-inspired studies in the locomotion of goats and dogs. *Integrative and Comparative Biology*, 51(1), 190–202. <http://dx.doi.org/10.1093/icb/icr061>.
- Li, Y., Ganesh, G., Jarrassé, N., Haddadin, S., Albu-Schaeffer, A., & Burdet, E. (2018). Force, impedance, and trajectory learning for contact tooling and haptic identification. *IEEE Transactions on Robotics*, 34(5), 1170–1182. <http://dx.doi.org/10.1109/TRO.2018.2830405>.
- Maeda, R. S., Cluff, T., Gribble, P. L., & Pruszyński, J. A. (2018). Feedforward and feedback control share an internal model of the arm's dynamics. *The Journal of Neuroscience*, 38(49), 10505–10514. <http://dx.doi.org/10.1523/JNEUROSCI.1709-18.2018>.
- Maeda, R. S., Gribble, P. L., & Pruszyński, J. A. (2020). Learning new feedforward motor commands based on feedback responses. *Current Biology*, 30(10), 1941–1948. <http://dx.doi.org/10.1016/j.cub.2020.03.005>.
- Mathew, M., Sidhik, S., Sridharan, M., Azad, M., Hayashi, A., & Wyatt, J. (2019). Online learning of feed-forward models for task-space variable impedance control. In 2019 IEEE-RAS 19th international conference on humanoid robots (pp. 222–229). <http://dx.doi.org/10.1109/Humanoids43949.2019.9035063>.
- Nguyen-Tuong, D., & Peters, J. (2011). Model learning for robot control: a survey. *Cognitive Processing*, 12(4), 319–340. <http://dx.doi.org/10.1007/s10339-011-0404-1>.
- Osu, R., Franklin, D. W., Kato, H., Gomi, H., Domen, K., Yoshioka, T., et al. (2002). Short- and long-term changes in joint co-contraction associated with motor learning as revealed from surface EMG. *Journal of Neurophysiology*, 88(2), 991–1004. <http://dx.doi.org/10.1152/jn.2002.88.2.991>.
- Pan, Y., & Yu, H. (2017). Biomimetic hybrid feedbackforward neural-network learning control. *IEEE Transactions on Neural Networks and Learning Systems*, 28(6), 1481–1487. <http://dx.doi.org/10.1109/TNNLS.2016.2527501>.
- Pogue, A., BIANES, A., Hong, D., & Iwasaki, T. (2017). NABI-S: A compliant robot with a CPG for locomotion. In 2017 IEEE/RSJ international conference on intelligent robots and systems (pp. 3366–3371). <http://dx.doi.org/10.1109/IROS.2017.8206175>.
- Proietti, T., Morel, G., Roby-Brami, A., & Jarrassé, N. (2017). Comparison of different error signals driving the adaptation in assist-as-needed controllers for neurorehabilitation with an upper-limb robotic exoskeleton. In 2017 IEEE international conference on robotics and automation (pp. 6645–6650). <http://dx.doi.org/10.1109/ICRA.2017.7989784>.
- Ravichandar, H., Polydoros, A. S., Chernova, S., & Billard, A. (2020). Recent advances in robot learning from demonstration. *Annual Review of Control, Robotics, and Autonomous Systems*, 3(1), 297–330. <http://dx.doi.org/10.1146/annurev-control-100819-063206>.
- Ren, H., & Ben-Tzvi, P. (2020). Learning inverse kinematics and dynamics of a robotic manipulator using generative adversarial networks. *Robotics and Autonomous Systems*, 124, Article 103386. <http://dx.doi.org/10.1016/j.robot.2019.103386>.
- Santini, C. D., Bianchi, M., Grioli, G., Angelini, F., Catalano, M., Garabini, M., et al. (2017). Controlling soft robots: Balancing feedback and feedforward elements. *IEEE Robotics & Automation Magazine*, 24(3), 75–83. <http://dx.doi.org/10.1109/MRA.2016.2636360>.
- Schaal, S., & Atkeson, C. G. (2010). Learning control in robotics. *IEEE Robotics & Automation Magazine*, 17(2), 20–29. <http://dx.doi.org/10.1109/MRA.2010.936957>.
- Seidler, R. D., Noll, D. C., & Thiers, G. (2004). Feedforward and feedback processes in motor control. *NeuroImage*, 22(4), 1775–1783. <http://dx.doi.org/10.1016/j.neuroimage.2004.05.003>.
- Sigaud, O., Salaün, C., & Padois, V. (2011). On-line regression algorithms for learning mechanical models of robots: A survey. *Robotics and Autonomous Systems*, 59(12), 1115–1129. <http://dx.doi.org/10.1016/j.robot.2011.07.006>.

ARTICLE IN PRESS

X. Xiong and P. Manoonpong

Neural Networks xxx (xxxx) xxx

- 1 Spong, M. W., Hutchinson, S., & Vidyasagar, M. (2006). Robot modeling and
2 control. *Industrial Robot*, 30(5), 403. <http://dx.doi.org/10.1108/ir.2006.33.5.403.1>.
- 3 Stulp, F., & Sigaud, O. (2015). Many regression algorithms, one unified model: A
4 review. *Neural Networks*, 69, 60–79. <http://dx.doi.org/10.1016/j.neunet.2015.05.005>.
- 5 Takiyama, K., & Sakai, Y. (2016). Balanced motor primitive can explain
6 generalization of motor learning effects between unimanual and bimanual
7 movements. *Scientific Reports*, 6(1), 23331. <http://dx.doi.org/10.1038/srep23331>.
- 8 Tanaka, M., & Tanaka, K. (2015). Control of a snake robot for ascending and
9 descending steps. *IEEE Transactions on Robotics*, 31(2), 511–520. <http://dx.doi.org/10.1109/TRO.2015.2400655>.
- 10 Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive
11 combination of motor primitives. *Nature*, 407(6805), 742–747. <http://dx.doi.org/10.1038/35037588>.
- 12 Ting, J.-A., Meier, F., Vijayakumar, S., & Schaal, S. (2016). *Locally weighted
13 regression for control - encyclopedia of machine learning and data mining* (pp.
14 1–14). Boston, MA: Springer US. http://dx.doi.org/10.1007/978-1-4899-7502-7_493-1.
- 15 Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor
16 learning. *Nature Reviews Neuroscience*, 12(12), 739–751. <http://dx.doi.org/10.1038/nrn3112>.
- 17 Wolpert, D. M., & Flanagan, J. R. (2016). Computations underlying sensorimotor
18 learning. *Current Opinion in Neurobiology*, 37, 7–11. <http://dx.doi.org/10.1016/j.conb.2015.12.003>.
- 19 Xiong, X., & Manoonpong, P. (2018). Adaptive motor control for human-like
20 spatial-temporal adaptation. In *IEEE international conference on robotics and
21 biomimetics* (pp. 2107–2112). <http://dx.doi.org/10.1109/ROBIO.2018.8665222>.
- 22 Xiong, X., & Manoonpong, P. (2020a). Resistance-as-needed (RAN) control for
23 a wearable and soft hand exoskeleton. *Gait & Posture*, 81, 398–399. <http://dx.doi.org/10.1016/j.gaitpost.2020.08.098>.
- 24 Xiong, X., & Manoonpong, P. (2020b). Online adaptive resistance control of an
25 arm exercise exoskeleton. In V. G. Gradetsky, N. N. Bolotnik, M. O. Tokhi,
26 M. Silva, & G. S. Virk (Eds.), *Robots in human life- proceedings of the 23rd
27 international conference on climbing and walking robots and the support
28 technologies for mobile machines* (pp. 31–38). CLAWAR Association Ltd, <http://dx.doi.org/10.13180/clawar.2020.24-26.08.38>.
- 29 Xiong, X., Wörgötter, F., & Manoonpong, P. (2016). Adaptive and energy ef-
30 ficient walking in a hexapod robot under neuromechanical control and
31 sensorimotor learning. *IEEE Transactions on Cybernetics*, 46(11), 2521–2534.
32 <http://dx.doi.org/10.1109/TCYB.2015.2479237>.
- 33 Yang, C., Ganesh, G., Haddadin, S., Parusel, S., Albu-Schaeffer, A., & Burdet, E.
34 (2011). Human-like adaptation of force and impedance in stable and unstable
35 interactions. *IEEE Transactions on Robotics*, 27(5), 918–930. <http://dx.doi.org/10.1109/TRO.2011.2158251>.
- 36 Yeo, S.-H., Franklin, D. W., & Wolpert, D. M. (2016). When optimal feedback
37 control is not enough: Feedforward strategies are required for optimal
38 control with active sensing. *PLoS Computational Biology*, 12(12), <http://dx.doi.org/10.1371/journal.pcbi.1005190>.
- 39 Yousif, N., & Diedrichsen, J. (2012). Structural learning in feedforward and
40 feedback control. *Journal of Neurophysiology*, 108(9), 2373–2382. <http://dx.doi.org/10.1152/jn.00315.2012>.
- 41 Yu, J., Shi, P., Lin, C., & Yu, H. (2020). Adaptive neural command filtering
42 control for nonlinear MIMO systems with saturation input and unknown
43 control direction. *IEEE Transactions on Cybernetics*, 50(6), 2536–2545. <http://dx.doi.org/10.1109/TCYB.2019.2901250>.
- 44 Zhang, D., Wang, Z., & Masayoshi, T. (2020). Neural-network-based iterative
45 learning control for multiple tasks. *IEEE Transactions on Neural Networks and
46 Learning Systems*, 1–13. <http://dx.doi.org/10.1109/TNNLS.2020.3017158>.
- 47 Zhou, L., Tao, H., Paszke, W., Stojanovic, V., & Yang, H. (2020). PD-type iterative
48 learning control for uncertain spatially interconnected systems. *Mathematics*,
49 8(9), 1528. <http://dx.doi.org/10.3390/math8091528>.