



Dynamic Modulation of a Learned Motor Skill for Its Recruitment

Kyuengbo Min^{1*}, Jongho Lee² and Shinji Kakei¹

¹ Department of Psychiatry & Behavioral Sciences, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan,

² Department of Clinical Engineering, Faculty of Health Sciences, Komatsu University, Komatsu, Japan

Humans learn motor skills (MSs) through practice and experience and may then retain them for recruitment, which is effective as a rapid response for novel contexts. For an MS to be recruited for novel contexts, its recruitment range must be extended. In addressing this issue, we hypothesized that an MS is dynamically modulated according to the feedback context to expand its recruitment range into novel contexts, which do not involve the learning of an MS. The following two sub-issues are considered. We previously demonstrated that the learned MS could be recruited in novel contexts through its modulation, which is driven by dynamically regulating the synergistic redundancy between muscles according to the feedback context. However, this modulation is trained in the dynamics under the MS learning context. Learning an MS in a specific condition naturally causes movement deviation from the desired state when the MS is executed in a novel context. We hypothesized that this deviation can be reduced with the additional modulation of an MS, which tunes the MS-produced muscle activities by using the feedback gain signals driven by the deviation from the desired state. Based on this hypothesis, we propose a feedback gain signal-driven tuning model of a learned MS for its robust recruitment. This model is based on the neurophysiological architecture in the cortico-basal ganglia circuit, in which an MS is plausibly retained as it was learned and is then recruited by tuning its muscle control signals according to the feedback context. In this study, through computational simulation, we show that the proposed model may be used to neurophysiologically describe the recruitment of a learned MS in novel contexts.

Keywords: motor skill recruitment, muscle synergy, corticospinal tract, reinforcement learning, cortico-basal ganglia circuit, muscle loading, feedback gain control

OPEN ACCESS

Edited by:

Daya Shankar Gupta,
Camden County College,
United States

Reviewed by:

Pietro Mazzoni,
Washington University in St. Louis,
United States
Luca Padua,
Catholic University of the Sacred
Heart, Italy

Giulia Paparella,
IRCCS Neuromed, Italy

*Correspondence:

Kyuengbo Min
min-kb@igakuken.or.jp

Received: 01 March 2019

Accepted: 13 October 2020

Published: 23 December 2020

Citation:

Min K, Lee J and Kakei S (2020)
Dynamic Modulation of a Learned
Motor Skill for Its Recruitment.
Front. Comput. Neurosci. 14:457682.
doi: 10.3389/fncom.2020.457682

INTRODUCTION

Innate and learned motor skills (MSs) are recruited in the central nervous system (CNS) for effective and fast motor control in response to novel external circumstances such as disturbances. To recruit an MS in response to novel contexts, its contextual information must be afferently transmitted to the CNS through feedback control processes. Therefore, the recruitment of an MS should be considered in the feedback control process. However, this mechanism has not been addressed in previous studies related to feedback control, such as proportional integral derivative control (Petkos and Vijayakumar, 2007) and optimal feedback control (Todorov and Jordan, 2002; Liu and Todorov, 2007), because these studies focused only on correcting motor control errors through feedback gain control. In addressing this issue, we hypothesized that an MS is dynamically modulated according to the feedback context to expand its recruitment range into novel contexts,

which do not involve the learning of an MS. The following two sub-issues are taken into account in this article.

Dynamic modulation of an MS in response to the feedback context is a mechanism that allows rapid recruitment of an MS in a novel context. In validating this hypothesis, the Synergy strategy-based muscle Control (SC) proposed in our previous study (Min et al., 2018) is a valuable concept because it contributes to the dynamic modulation of an MS to regulate the functional redundancy of individual muscle units for the feedback context. To achieve this SC-driven MS (SC-MS), all muscle units contributing to an MS need to be classified into multiple group units according to their innervation by peripheral nerves derived from the brachial plexus. Consequently, these group units cause the contracting sets of muscles, termed motor primitives (MPs) (Giszter et al., 1993), to effectively suppress the control redundancy of muscle units in feedback control. In our previous study (Min et al., 2018), this muscle control policy was defined as the group control policy (GCP) that outputs the same control signal to all muscle units constituting a group unit. Although the GCP is an effective control policy for suppressing the control redundancy of muscle units in feedback control, it needs assistance to generate novel patterns of muscle activities that cannot be produced by combining group units. To assist GCP, the individual control policy (ICP) was defined as the control policy for outputting identified individual control signals to individual muscle units. These two control policies synergistically combine to optimally control muscle units according to the feedback context. This synergy may neurophysiologically correspond to the combination of corticospinal neurons (CSTs) in the primary motor cortex (M1) and its second type of CSTs, termed cortico-motoneuronal cells (CMs), through the corticospinal tract, which was addressed in a previous study (Rathelot and Strick, 2009). In this previous study, it was suggested that the MPs activated by the CSTs in M1 through their connection with interneurons in the spinal cord may be adjusted by the signals that are produced from CMs through their monosynaptic connection with motor neurons (MNs) in the spinal cord. This adjusting of MPs may sculpt novel motor output patterns for highly skilled movements that cannot be produced by combining MPs. Consequently, the combination of two ways of controlling muscles in the corticospinal tract is more plausible in neurophysiologically representing an MS than CSTs driven one way, which is the route for controlling MPs termed muscle synergies (Tresch et al., 1999; d'Avella et al., 2003; Torres-Oviedo et al., 2006; Safavynia and Ting, 2012; Ting and Macpherson, 2012; Barroso et al., 2014; Suzuki et al., 2017; Amundsen Huffmaster et al., 2018; De Marchis et al., 2018; Kibushi et al., 2018; Toma and Santello, 2019), whose individual MP units are composed of spatiotemporally fixed muscle activities. These studies demonstrate that SC may be neurophysiologically suitable for characterizing the dynamic modulation of an MS.

Even if an MS is dynamically modulated for its recruitment in novel feedback contexts, this modulation is trained in the dynamics under the MS learning context. Therefore, this learning condition of an MS naturally brings about movement control deviation from the desired state in a novel context. To overcome

this handicap, a learned MS needs to be modified in response to a novel context. In addressing this issue, we propose a Tuned Synergy strategy-based muscle Control (T-SC) model, in which the SC-MS is tuned in response to the feedback context. Through this tuning, the aforementioned deviation is supposedly reduced. In designing this model, we assumed that the tuning signals of the SC-MS are cumulatively modified to tune SC-MS-produced muscle activities according to the deviation from the desired movement, which is recognized through feedback control. This hypothesis is based on experimental evidence demonstrating that the response through sensorimotor control is coupled with ongoing decision processes, which are reflected by the accumulated feedback information (Selen et al., 2012). In a previous related study (d'Avella and Pai, 2010), this issue was also addressed with regard to the limited recruitment range of existing modules such as muscle synergies in novel contexts. However, an alternative solution, apart from learning a new MS, has not been suggested so far. The proposed T-SC may be an alternative motion control strategy for novel contexts because it is more efficient for the rapid adaptation of motion control in novel contexts than learning a new MS.

The neurophysiological architecture and mathematical description of the T-SC model are presented in sections Neurophysiological Architecture and Mathematical Model, respectively. To validate this model, we simulated the recruitment of the SC-MS in novel contexts that were not present when the MS was learned (section Results).

MATERIALS AND METHODS

Neurophysiological Architecture

The neurophysiological architecture of T-SC is based on experimental evidence (Spraker et al., 2007) showing that the cortico-basal ganglia (cortico-BG) circuit is involved in scaling the force generation according to the external environment. Accordingly, this evidence is applicable to validating the recruitment of a learned SC-MS through tuning its muscle force control signals according to the feedback context. In this architecture, the operation of T-SC in the CNS may be achieved as follows.

Based on the experimental evidence (Pruszynski et al., 2011) for involvement of the M1 region in modulating the proprioceptive response related to the knowledge of limb mechanics, we surmised that the sensory feedback signals, s_{fb} , including the contextual information for the dynamic states of the skeletal joint, are transferred to M1 through its somatosensory pathway (London and Miller, 2013). These feedback signals, s_{fb} , are inputted to the basal ganglia (BG) through M1. In the cortico-BG loop (Barto, 1995; Doya, 2000, 2007, 2008; Ito and Doya, 2011), the BG selectively disinhibits the activities of both M1 and the brainstem to select the optimal tactic for motion control (Hikosaka et al., 2000). The extent of this disinhibition is controlled via dopamine release (Shinnamon, 1993) during reinforcement learning (Houk et al., 1995). Therefore, the BG is assumed to dynamically produce a trade-off between inhibition and disinhibition of the activity in M1 during sequential motion control (Nambu et al., 2002).

Further, section Recruiting a Learned MS via the Cortico-Basal Ganglia Loop discusses the neurophysiological evidence for the involvement of the BG in kinematic control through the recruitment of a learned MS. Based on this experimental evidence, in the BG, we assumed that an SC-MS is dynamically modulated by inhibiting or disinhibiting the GCP and ICP, which regulate the functional redundancy of individual muscle units for \mathbf{s}_{fb} , as discussed in section Introduction. This modulation is mathematically described in section Dynamic Modulation of an SC-MS. Through this modulation, the SC-MS produces a muscle control signal, $\mathbf{u}^{P_t^{SC}}$, that is efferently copied, as \mathbf{u}^{SC} , to the spinal cord through the corticospinal neurons (CSTs) in M1. Through this process, a learned SC-MS retained in the BG is dynamically modulated to produce \mathbf{u}^{SC} according to the feedback context \mathbf{s}_{fb} . A discussion of this modulation is introduced in sections Recruiting a learned MS via the Cortico-Basal Ganglia Loop and Muscle Control Scheme of the Corticospinal Tract in Recruiting an SC-MS.

The signals \mathbf{u}^{SC} are assumed to be tuned in M1 because experimental evidence (Herter et al., 2009) has shown that neural activity in M1 is broadly tuned to novel contexts, such as mechanical perturbations applied to the shoulder and elbow, and reflects knowledge of joint–limb dynamics (Pruszynski et al., 2011). Based on this supposition, the \mathbf{u}^{SC} may be tuned with the following dynamic modulation process for the feedback context. Concurrent with the aforementioned input of \mathbf{s}_{fb} to M1, the goal states \mathbf{s}^o of \mathbf{s}_{fb} are also input to M1 from the association cortex. Both \mathbf{s}_{fb} and \mathbf{s}^o are inputs to the muscle loading tuner and the difference between the two signals is transferred to the tuning gain (TG), \mathbf{G}^{tuning} , of \mathbf{u}^{SC} . The \mathbf{G}^{tuning} consists of agonist and antagonist loading signals, which disinhibit the activities of loaded muscles and inhibit the activities of unloaded muscles by properly scaling them (Nashed et al., 2015). This tuning process generates the optimal muscle control signals \mathbf{u}^* , which descend to MNs in the spinal cord to control the muscles. This tuning is mathematically described in section Tuning of a Learned SC-MS.

Mathematical Model

Dynamic Modulation of an SC-MS

As mentioned in section Introduction, we have defined the group units as muscle control units, which produce the contracting sets of muscles, termed MPs. These group units and their belongings are determined according to the peripheral nerves innervating them (Table 1). Based on this neurophysiological definition, the GCP is defined as the control policy considering individual muscle units as a component of the group unit, in which all components respond to the feedback context with one common signal. In contrast to the GCP, the ICP is defined as the control policy considering individual muscle units as independent units of the group units, thereby controlling individual muscle units with their identified signals. Therefore, by optimizing the synergy between these two control policies for a feedback context, an SC-MS is dynamically modulated for the feedback context. This modulation is mathematically defined by the following model based on our previous study (Min et al., 2018):

$$\mathbf{P}_t^{SC} = (v^G, v^I), v^G = \sigma(\mathbf{s}^t) = \exp(-0.5V(\mathbf{s}^t)), v^I = 1 - v^G \quad (1)$$

where \mathbf{P}_t^{SC} is the synergy coordinate of the GCP weight v^G and ICP weight v^I . This is determined by the critic value (CV) $V(\mathbf{s}^t)$, which evaluates the potential of the feedback contextual vector \mathbf{s}^t at time t for reaching the goal state. As the \mathbf{s}^t is produced through the performance of the SC-MS, the $V(\mathbf{s}^t)$ presents an evaluation of the performance of the SC-MS for the goal state. Therefore, the \mathbf{P}_t^{SC} is dynamically optimized according to the performance of the SC-MS at time t for the goal state. The synergy between v^G and v^I is simulated in Figures 4, 5A, 6A, 7A (see section Results). By applying the \mathbf{P}_t^{SC} to Equation (2b), the SC-MS, using the actor model in Equation (2a), is dynamically modulated. This achievement is rewarded by functionally improving the $V(\mathbf{s}^t)$ (Min et al., 2018). Consequently, this improvement reinforces the SC-MS to achieve its goal state. This CV-driven reinforcement learning is based on the actor–critic model (Barto, 1995; Sutton and Barto, 1998), which is designed to simulate reinforcement learning (Houk et al., 1995) in the BG. The simulation condition of this learning is precisely described in section Learning and Recruitment Condition of an SC-MS.

Using \mathbf{P}_t^{SC} optimized through the aforementioned learning, the SC-MS is dynamically modulated to generate the muscle control signals \mathbf{u}^{SC} as follows:

$$\begin{aligned} U_i^{SC}(\mathbf{s}^t) &\cong u_i(\mathbf{s}^t; \mathbf{W}) \\ &= u_i^{\max} \text{sig}\left(\sum_{k=1}^K W_k^i b_k(\mathbf{s}^t) + \sigma(\mathbf{s}^t) n_i(t) - B\right), \\ u_i^{\max} &= 1.0, \sigma(\mathbf{s}^t) = \sigma_0 \exp(-0.5V(\mathbf{s}^t)), \\ b_k(\mathbf{s}^t) &= \frac{A_k(\mathbf{s}^t)}{\sum_{l=1}^K A_l(\mathbf{s}^t)}, A_k(\mathbf{s}^t) = \exp\left[-\sum_{i=1}^n \left(\frac{s_i^t - c_i^k}{\sigma_i^k}\right)^2\right] \end{aligned} \quad (2a)$$

$$W_k^i = v^I w_k^i + v^G w_k^g \quad (2b)$$

where the $U_i^{SC}(\mathbf{s}^t)$ functions as the actor generating the control signal of the i th muscle of \mathbf{u}^{SC} and $\text{sig}(x)$ is the sigmoid function. The function $n_i(t)$ produces the white noise in determining the activities of individual muscles. The magnitude of $n_i(t)$ is determined according to $\sigma(\mathbf{s}^t)$ by considering $V(\mathbf{s}^t)$. σ_0 is a constant parameter. This noise is designed to enhance the learning dynamic of an MS, thereby being suppressed by setting σ_0 to zero in simulating its recruitment. B is the parameter controlling the baseline of $\text{sig}(x)$, i.e., the value of $\text{sig}(x = 0.0)$. The base function $b_k(\mathbf{s}^t)$ is the k th element of a normalized Gaussian network (NGSN). K is the total number of base functions. The node of $b_k(\mathbf{s}^t)$ is defined as the parameter c_i^k , which is the i th element of the center of $b_k(\mathbf{s}^t)$ and σ_i^k is its range. This c_i^k is determined before the learning takes place. As the state vector \mathbf{s}^t comprises the joint angle and its velocity, the predetermined format of the NGSN is designed based on the grid distribution of the two-dimensional state by setting the total number of state elements n to 2. The symbol s_i^t is the i th element of the contextual vector \mathbf{s}^t . W_k^i is the network weight of $b_k(\mathbf{s}^t)$ in producing the i th muscle activity. As described in

TABLE 1 | Elements of the neuromuscular system controlling the elbow joint.

Flexors	
Group 1 (radial nerve)	Brachioradialis
Group 2 (musculocutaneous nerve)	Biceps brachii (long head), Biceps brachii (short head), and brachialis
Group 3 (median nerve)	Pronator teres
Extensors	
Group 4 (radial nerve)	Triceps brachii (lateral head), Triceps brachii (long head), Triceps brachii (medial head), and anconeus

Equation (2b), the W_k^j is the summation of w_k^j and w_k^g , which are, respectively, weighted by v^j and v^g of \mathbf{P}_t^{SC} . The parameter w_k^g is the k th NGSN weight of the g th group unit, which is governed by the GCP, whereas w_k^i is the k th NGSN weight of the i th muscle affiliated to the g th group unit, which is governed by the ICP. The weights w_k^i and w_k^g are optimized through the aforementioned SC-MS learning. For further information, including the optimizing process of w_k^i and w_k^g regarding Equation (2b), refer to our previous study (Min et al., 2018).

Tuning of a Learned SC-MS

As shown in **Figure 1**, the learned SC-MS-produced \mathbf{u}^{SC} is additionally tuned to \mathbf{u}^* with the TG signal $\mathbf{G}^{\text{tuning}}$ from the muscle loading tuner, which is cumulatively modified as the feedback gain parameters of \mathbf{u}^{SC} according to the deviation from the desired state recognized through feedback control. This tuning is mathematically modeled as follows:

$$\begin{aligned} \mathbf{u}^* &= \mathbf{G}^{\text{tuning}} \mathbf{u}^{SC}, \mathbf{G}^{\text{tuning}} = \text{diag} \left(G_0^{\text{tuning}}, \dots, G_{n-1}^{\text{tuning}} \right), \\ \mathbf{u}^{SC} &= [u_0, \dots, u_{n-1}]^T, \text{ If } u_i \text{ is a flexor, } G_i^{\text{tuning}} = G^F. \\ \text{If } u_i &\text{ is an extensor, } G_i^{\text{tuning}} = G^E, i = 0, \dots, n-1. \end{aligned} \quad (3)$$

where $\mathbf{G}^{\text{tuning}}$ functions as a feedback gain parameter that is the diagonal matrix composed of G_i^{tuning} . The symbol n represents the total number of muscles involved in the motion control. The flexor gain G^F or the extensor gain G^E is determined by G_i^{tuning} according to the function of the individual muscles u^i in controlling the joint.

The $\mathbf{G}^{\text{tuning}}$ is modified by its incremental signal $\Delta \mathbf{G}^{\text{tuning}}$ as follows:

$$\begin{aligned} \Delta \mathbf{G}^{\text{tuning}}(t) &= [\Delta G^F(t), \Delta G^E(t)]^T = \mathbf{k}(t) \cdot \Delta \mathbf{s}(t), \\ \mathbf{k}(t) &= \begin{bmatrix} \mathbf{k}^F(t) \\ \mathbf{k}^E(t) \end{bmatrix} = \begin{bmatrix} k_p(t) & k_d(t) & k_a(t) \\ -k_p(t) & -k_d(t) & k_a(t) \end{bmatrix}, \\ \Delta \mathbf{s}(t) &= \mathbf{s}^G - \mathbf{s}(t) = [\Delta \Theta_t^G, \Delta \dot{\Theta}_t^G, \Delta \ddot{\Theta}_t^G]^T, \\ \Delta \Theta_t^G &= \Theta^G - \Theta_t, \Delta \dot{\Theta}_t^G = \dot{\Theta}^G - \dot{\Theta}_t, \Delta \ddot{\Theta}_t^G = \ddot{\Theta}^G - \ddot{\Theta}_t \end{aligned} \quad (4)$$

where $\Delta \mathbf{G}^{\text{tuning}}(t)$ is composed of the flexor and extensor components, $\Delta G^F(t)$, $\Delta G^E(t)$. These two components are estimated by $\Delta \mathbf{s}(t)$ and its gain matrix $\mathbf{k}(t)$. $\Delta \mathbf{s}(t)$ is the difference between the feedback state $\mathbf{s}(t) = (\Theta_t, \dot{\Theta}_t, \ddot{\Theta}_t)$ and its desired state $\mathbf{s}^G = (\Theta^G, \dot{\Theta}^G, \ddot{\Theta}^G)$, in which both $\dot{\Theta}^G$ and $\ddot{\Theta}^G$ are zero.

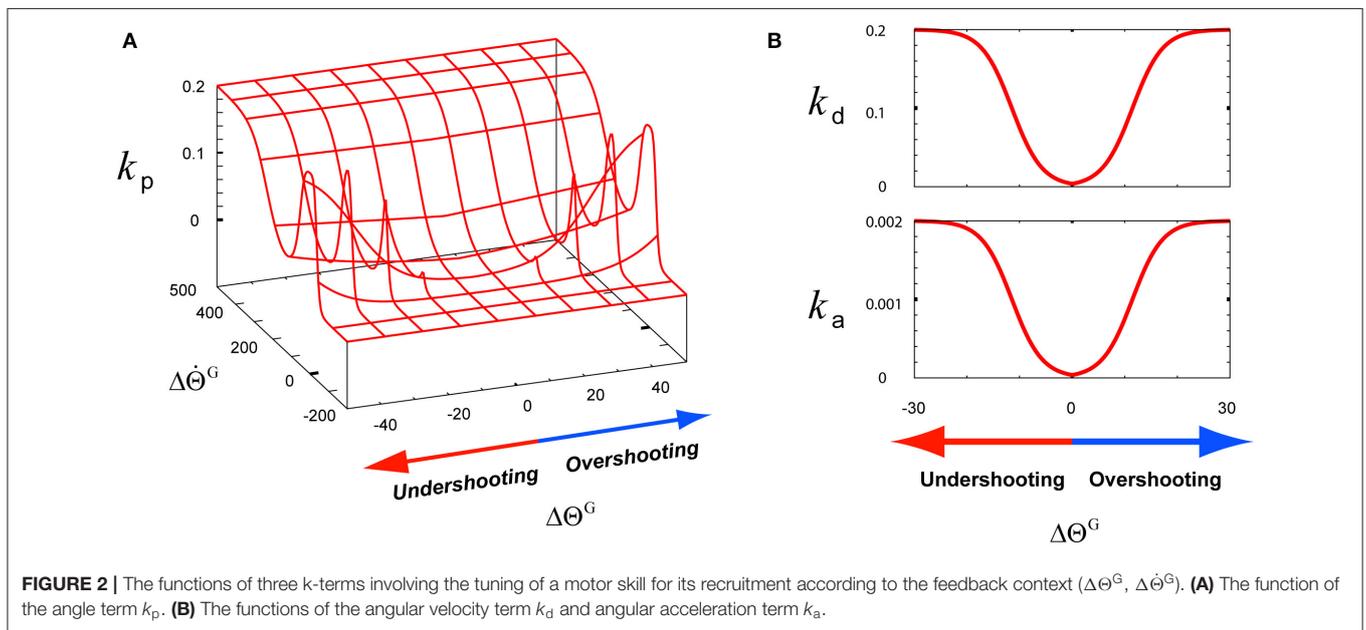
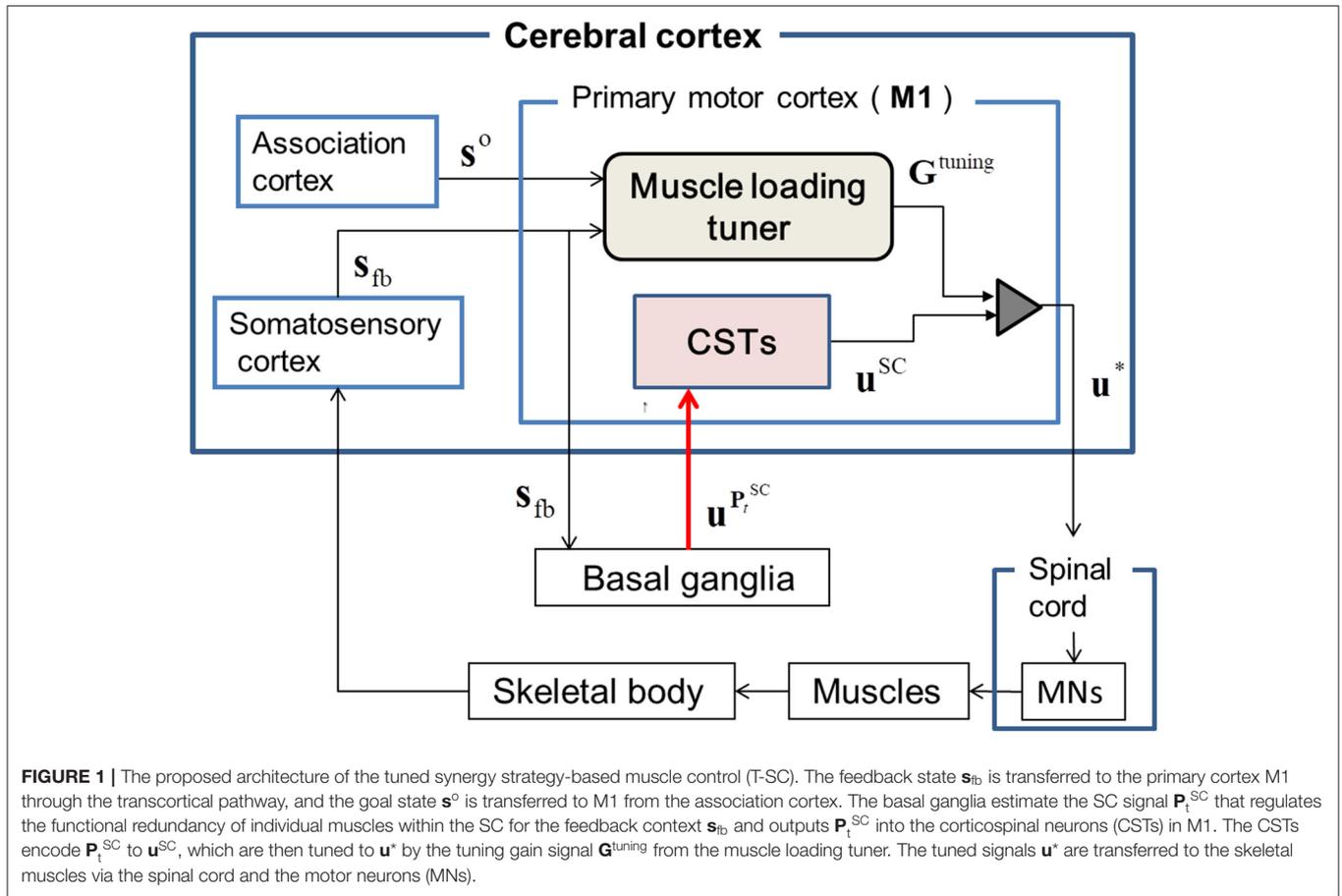
The matrix $\mathbf{k}(t)$ is composed of the following three components: the angle term $k_p(t)$, the angular velocity term $k_d(t)$, and the angular acceleration term $k_a(t)$. These terms contribute to the flexor part $\mathbf{k}^F(t) = (k_p(t), k_d(t), k_a(t))$ and the extensor part $\mathbf{k}^E(t) = (-k_p(t), -k_d(t), k_a(t))$. The components $k_p(t)$ and $k_d(t)$ of \mathbf{k}^E are designed as minus terms of \mathbf{k}^F to simulate the activities of extensors. However, the acceleration term $k_a(t)$ is set to the same value for both the agonist and the antagonist because the direction of the angular acceleration frequently changes, thus it needs to be suppressed to maintain stable motion control during the co-contraction of both the agonist and the antagonist. These three \mathbf{k} components are optimally modulated to make the joint angular state approach the goal state by using Equation (5). To achieve this modulation, the three \mathbf{k} terms of $\Delta \mathbf{G}^{\text{tuning}}(t)$ are modeled to be proportional to $\|\Delta \Theta_t^G\|$ using the function $\text{sig}(x)$ as follows:

$$\begin{aligned} k_p(t) &= C^{k_p} \cdot \text{sig} \left(1.5 \cdot (\Delta \dot{\Theta}_t^G + (A \cdot \exp(-\|\Delta \dot{\Theta}_t^G\|)) \cdot \|\Delta \Theta_t^G\| - B) \right) \\ k_d(t) &= C^{k_d} \cdot \text{sig} \left(D \|\Delta \Theta_t^G\| - B \right), \\ k_a(t) &= C^{k_a} \cdot \text{sig} \left(D \|\Delta \Theta_t^G\| - B \right) \end{aligned} \quad (5)$$

where B ($B = 0.4$) is the parameter controlling the baseline of $\text{sig}(x)$, that is, the value for $\text{sig}(x = 0.0)$, whereas the parameters C^{k_p} ($C^{k_p} = 0.2$), C^{k_d} ($C^{k_d} = 0.2$), and C^{k_a} ($C^{k_a} = 0.002$) are the constant values of $\text{sig}(x)$. The parameter A ($A = 10.0$) is the constant gain of the Gaussian function for modulating $k_p(t)$, and the parameter D ($D = 20.0$) is the constant value for modulating $k_d(t)$ and $k_a(t)$. These k components are dynamically modulated considering $\Delta \mathbf{s}(t)$, which was described in Equation (4). As shown in **Figure 2A**, to model the gain term of $\Delta \dot{\Theta}_t^G$, $k_p(t)$ is modeled to mainly function as $C^{k_p} \cdot \text{sig}(1.5 \cdot (A \|\Delta \Theta_t^G\| - B))$, which is the sigmoid function of $\Delta \dot{\Theta}_t^G$ in the first term of $\text{sig}()$. In low-speed undershooting or overshooting, $k_p(t)$ functions as $C^{k_p} \cdot \text{sig}(1.5 \cdot (A \|\Delta \Theta_t^G\| - B))$, which is the sigmoid function of $\|\Delta \Theta_t^G\|$ in the second term of $\text{sig}()$. Using the hybridization of these two terms, $k_p(t)$ is modeled as shown in **Figure 2A**. Due to this modeling, $k_p(t)$ responds to $\Delta \dot{\Theta}_t^G$ under consideration of $\|\Delta \Theta_t^G\|$. To optimally modulate $k_d(t)$ and $k_a(t)$ as the gain terms of $\Delta \dot{\Theta}_t^G$ and $\Delta \ddot{\Theta}_t^G$, their corresponding $\|\Delta \Theta_t^G\|$ needs to be considered as the feedback context responding to undershooting and overshooting, as shown in **Figure 2B**. Owing to this consideration, the response of these k terms to the feedback context is slower than $k_p(t)$ considering $\Delta \dot{\Theta}_t^G$ in high-speed undershooting.

The TG increments generated according to the aforementioned calculations in Equations (4) and (5) are accumulated to modify the corresponding TG signals as follows:

$$\begin{aligned} \mathbf{G}^{\text{tuning}}(t) &= \mathbf{G}^{\text{tuning}}(t - \Delta t) + \Delta \mathbf{G}^{\text{tuning}}(t), \\ \mathbf{G}^{\text{tuning}}(t) &= (G^F(t), G^E(t)), \mathbf{G}^{\text{tuning}}(0) = (1.0, 1.0). \\ \text{While } G^F(t) &> 0.0, \Delta G^F(t) \text{ is available.} \\ \text{While } G^E(t) &> 0.0, \Delta G^E(t) \text{ is available.} \end{aligned} \quad (6)$$



where the initial TG, $\mathbf{G}^{tuning}(0)$, is set to 1.0 to simulate non-interference by the TG. To achieve this modification, $G^F(t)$ and $G^E(t)$, termed the flexor and extensor components of $\mathbf{G}^{tuning}(t)$,

respectively, must be above zero. Therefore, if $G^F(t)$ or $G^E(t)$ is modified to be below zero, the corresponding signal is set to zero by suppressing its increment.

Simulation Architecture

The simulation architecture has been described in detail in our previous study (Min et al., 2018). This architecture is composed of the musculoskeletal finite-element (FE) model (Figure 3), its motion control agent model, and the interface model, which integrates both of the aforementioned models. Precise descriptions are as follows.

Musculoskeletal Finite-Element Model

The musculoskeletal model proposed in our previous study (Min et al., 2018) was used for simulating the motion control of a musculoskeletal system. This model was designed using LS-DYNA (Livermore Software Technology Corporation, Livermore, CA, USA), which is an explicit FE code developed for dynamic analyses through simulation. To consider the trade-offs between analytical precision and calculation costs in simulating the motion control of the musculoskeletal FE model, the muscles are designed with FE modeling using multiple bar elements of the muscles that formulate muscle paths between the origin and the insertion points in LS-DYNA. The characteristic features of the muscle forces, which change according to the length of the muscle and its contraction velocity, were modeled using a Hill-type model (Hill and Sec, 1938; Zajac et al., 1985; Thelen, 2003). Anatomical references (Neumann, 2002) were used to align the origin and insertion points and the via points, and to represent the appropriate muscle moment arms using the wrapping contacts (Hada et al., 2007). The predicted muscle moment arms were well-validated against data from several experimental studies (Amis et al., 1979; Murray et al., 1995). As shown in Figure 3A, the proposed FE model consists of two rigid body parts: one representing the upper arm and shoulder, and the other representing the lower arm and hand. The two body parts are linked using a joint constraint that represents the ulnar-humeral joint. The mass of the lower arm was 1.7 kg. The principal moments of inertia of the lower arm body were $I_{11} = 7.66 \times 10^{-3} \text{ kg m}^2$, $I_{22} = 7.36 \times 10^{-3} \text{ kg m}^2$, and $I_{33} = 0.34 \times 10^{-3} \text{ kg m}^2$.

Integration of the Musculoskeletal Finite-Element Model and Its Motion Control Agent Model

The entire architecture was implemented through software programming, in which the agent model of the SC-MS was programmed with C++ code to perform the learning and recruitment of the LS-DYNA-coded musculoskeletal FE model. This performance was achieved through a C++ code interface model, which was programmed to allow the coding difference between the aforementioned two models.

Learning and Recruitment Condition of an SC-MS

To validate the recruitment of the SC-MS under novel conditions involving transient and sustained disturbances, the learning condition of the SC-MS was not affected by any external interference as follows.

In the simulation architecture, the agent model reinforced an MS to be dynamically modulated by the SC, described in Equation (1), for controlling the forearm to reach a goal without

any disturbances. Through this reinforcement learning, the agent model learned an SC-MS. During this learning process, the control range of the elbow joint was limited to 30–140°. The aim of this task was to move the hand to its goal position, where the elbow joint angle was at 70°, and to maintain this position. The degree of freedom of the joint was 1. The nine muscles listed in Table 1 were activated to control the elbow in the simulation, as shown in Figure 3A. The time step t was 0.01 s. If the total learning time in a trial exceeded 2.0 s or if the angle of the elbow joint was out of the defined control range, a new trial was started after randomly changing the initial position. This process was repeated 780 times.

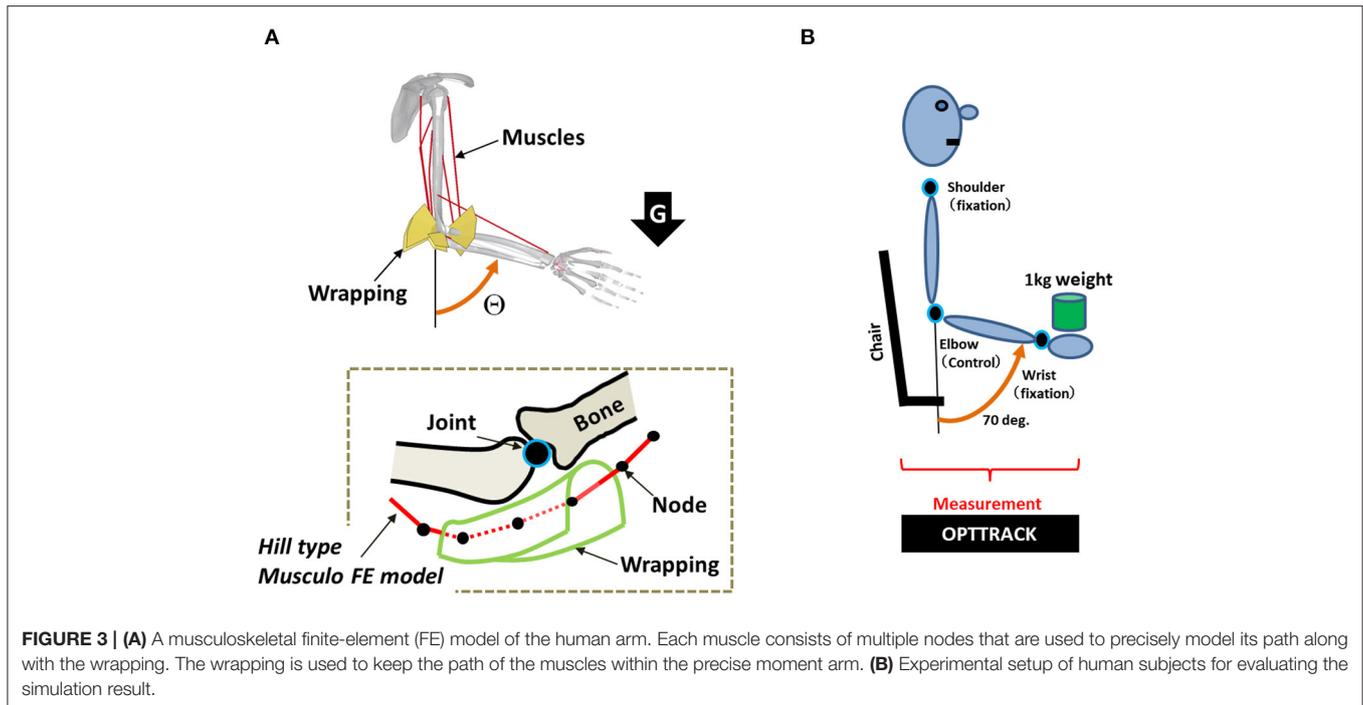
The SC-MS learned through the above process was recruited in the same time steps as the aforementioned learning time steps. Further information has been provided in detail in our previous study (Min et al., 2018).

Experimental Setup

To evaluate the proposed simulation model by comparison with the actions of four human subjects (four men, 40–44 years old) under the same conditions as those used in the simulation, we measured the loading responses of the study subjects, which is the same task as that in the simulation. All subjects were healthy and did not have any motor disorders. We assumed that these subjects have learned the MS recruited to achieve the aforementioned novel task throughout their whole life because the goal task of the MS described in section Learning and Recruitment Condition of an SC-MS can be achieved naturally by healthy subjects.

As shown in Figure 3B, in this experiment, the elbow joint angle was measured while the subject held a 1 kg load in his hand. To measure the responses to the loading condition through pure feedback control, the subjects were blindfolded with their eyes closed and were not informed about the timing of the loading. In addition, the distance between the initial falling point of the weight and the initial position of the hand was set close to zero. Furthermore, to approximate the novel condition as closely as possible, only data that were recorded during the first trial for each of the four subjects were used. The subject was instructed to try to recover as soon as possible the preloading posture set at 70°. All subjects were instructed to recover and maintain their preloading posture under this loading condition for 2.0 s. The shoulder and wrist joints were fixed during the measurement of the motion of the elbow joint. In this setting, we measured the positions of the shoulder, elbow, and wrist using OPTOTRAK 3020 (Northern Digital, Waterloo, Ontario, Canada), which is a three-dimensional position measurement device. We then used the measured positions of these three joints to calculate the angular movement of the elbow joint. The experimental setup has been described in detail in our previous study (Min et al., 2018).

All subjects provided written informed consent prior to their participation. The protocol was approved by the Tokyo Metropolitan Institute of Medical Science's ethics committees and was conducted in accordance with the ethical standards of the Declaration of Helsinki.



RESULTS

As mentioned in section Introduction, because of the dynamic modulation of an MS driven by the SC for novel feedback contexts, an SC-MS may be recruited in the CNS as a valuable learned MS. To validate this supposition, we tested the concept of T-SC in a model that tunes the SC-MS by gaining its signals to robustly recruit it in novel feedback contexts.

As shown in **Figure 1**, the T-SC is neurophysiologically achieved by tuning the SC-MS according to sensory feedback signals, which are generated in response to the context. Therefore, the T-SC may contribute to recruiting the SC-MS in the feedback control process. To validate this recruitment process, the simulation results of the SC-MS recruitment procedure in response to three novel sustained disturbances that did not involve the learning process of the SC are discussed in this section.

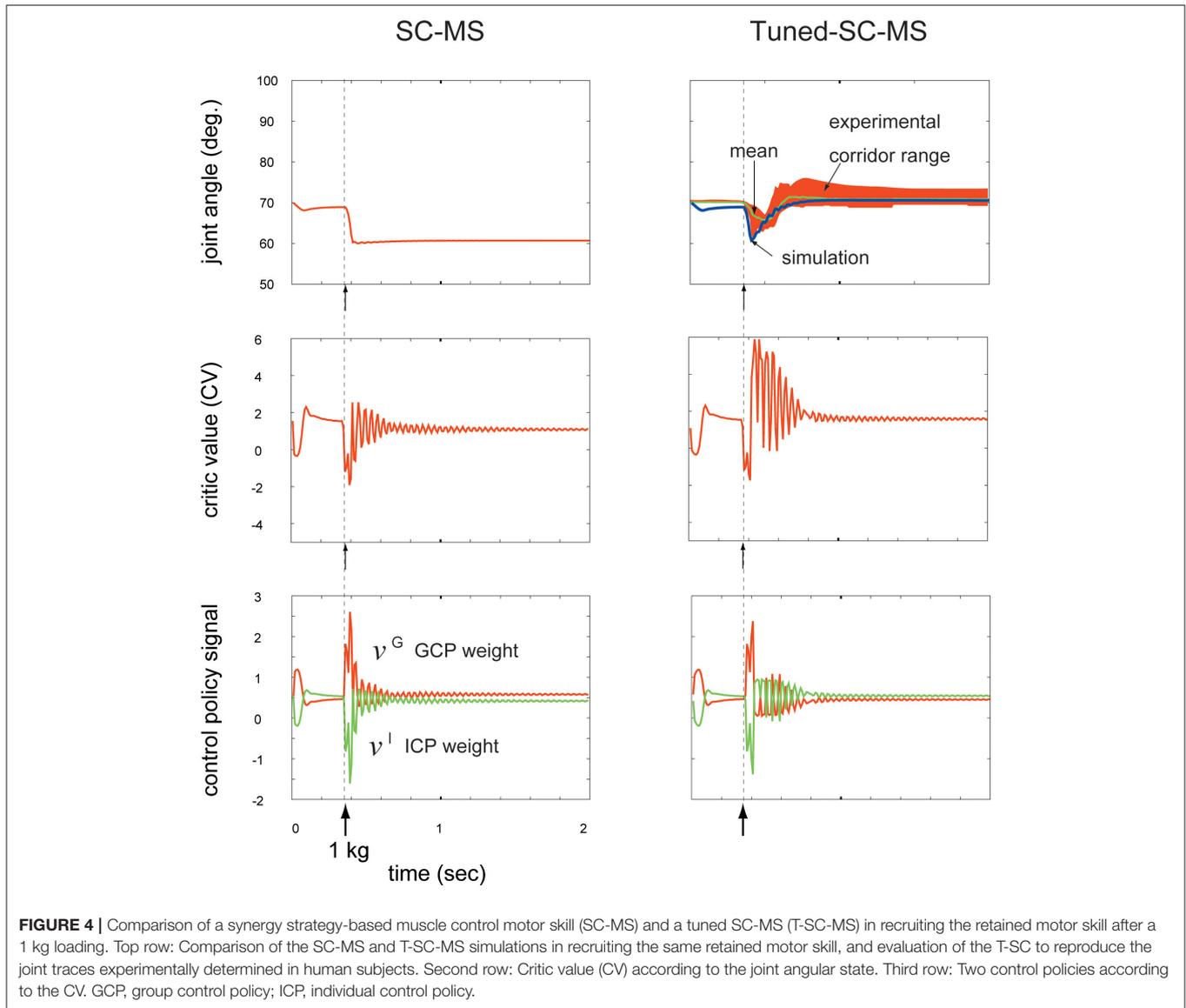
T-SC in a Novel Sustained Disturbance

Novel dynamic contexts in recruiting a learned MS are classified into transient and sustained disturbances. In our previous work (Min et al., 2018), we tested the motion control robustness of the SC-MS in these two types of dynamic contexts that did not involve SC-MS learning. In this test, the SC-MS demonstrated good recruitment against a transient disturbance, such as an impacting force, by recovering the pre-impacted context well. However, the SC-MS revealed the limitations of its recruitment in response to sustained disturbances; it only recovered to the point below the pre-disturbed point, as shown in **Figure 4**. As this difference in recruiting an SC-MS is attributed to the difference between their loading durations for an SC-MS, we hypothesized that the SC-MS needs to be tuned with accumulative gain signals

that consider the duration of the disturbance. To address this issue, we validated the T-SC for novel sustained disturbances as follows.

To validate the T-SC, we compared the performance of the SC-MS and the tuned SC-MS (T-SC-MS) in recovering the preloading posture against a sustained 1 kg loading. This sustained loading did not involve SC-MS learning. The results are shown in the top row of **Figure 4**. The left column in the top row of **Figure 4** demonstrates that the SC-MS found a new posture at 60° , which is below the preloading posture at 70° . By contrast, the T-SC-MS could successfully recover the preloading posture. This simulated recovering joint angular trace was within the real motion trace corridor range derived from the four experimental subjects using the same conditions as the simulation. This achievement of T-SC-MS was simulated through the following T-SC process.

As shown in **Figure 1**, the joint angular context may be afferently copied to the CNS as the contextual feedback signal and transferred to the BG via M1. According to this feedback signal, striosomal molecules functioning as adaptive critics in the BG (Houk et al., 1995) may estimate the CV as the evaluation of the recruitment of the SC-MS in recovering the preloading state. This is shown in the second row of **Figure 4**. After sustained 1 kg loading, the CV of the SC-MS decreases accordingly, and remained below, the CV of the preloading state according to maintaining the new posture below the preloading posture after 0.6 s. In comparison with SC-MS recruitment, the CV of the T-SC-MS also decreased during the undershooting, but it recovered to the level of the preloading state. Consequently, the T-SC-MS was more highly valued than the SC-MS in recruiting the learned MS to recover the preloading state after 0.4 s.



According to this CV-based evaluation of the SC-MS recruitment, the BG may optimally regulate the synergistic role redundancy between the GCP and ICP of individual muscles by using the rule based on Equation (1). As mentioned above, the CVs decreased from the preloading value because of their corresponding undershooting, shown in the top row of **Figure 4**. According to these CVs, as shown in both columns of the third row of **Figure 4**, the GCP increased whereas the ICP decreased because the GCP-driven group unit control is more effective than the ICP-driven individual muscle unit control in recovering the preloading context during undershooting. This synergy between the GCP and the ICP is differently regulated by SC-MS and T-SC-MS according to their CVs as follows.

After 0.4 s, SC-MS and T-SC-MS differently regulated the synergistic redundancy between the GCP and the ICP compared with before 0.4 s as follows. T-SC-MS started to recover the pre-disturbed CV from its lowest value after 0.4 s. According to this CV recovery, the GCP and ICP started to recover from

their highest and lowest values, respectively. After 0.9 s, the two control policies successfully recovered to their pre-disturbed values and were then kept stable at that state. In comparison with the T-SC-MS, the SC-MS maintained the new CV below the pre-disturbed CV after 0.7 s, thereby insufficiently recovering its two pre-disturbed control policies. This comparison is demonstrated in the third row of **Figure 4**. This superior achievement of T-SC-MS compared with SC-MS for the same novel disturbance is attributed to the following tuning process of an SC-MS.

As shown in **Figure 5A**, the SC-MS-produced signals were loaded with G^F and G^E , the agonistic and antagonistic signals of the TG, respectively, according to the recruitment process of a learned MS in **Figure 1**. These two TG signals were dynamically generated through the following feedback gain process of a learned MS in **Figure 5B**.

Under the rule based on Equation (5), the three k -term components of the TG in the third graph in the top row of **Figure 5B** were determined according to the joint angle and

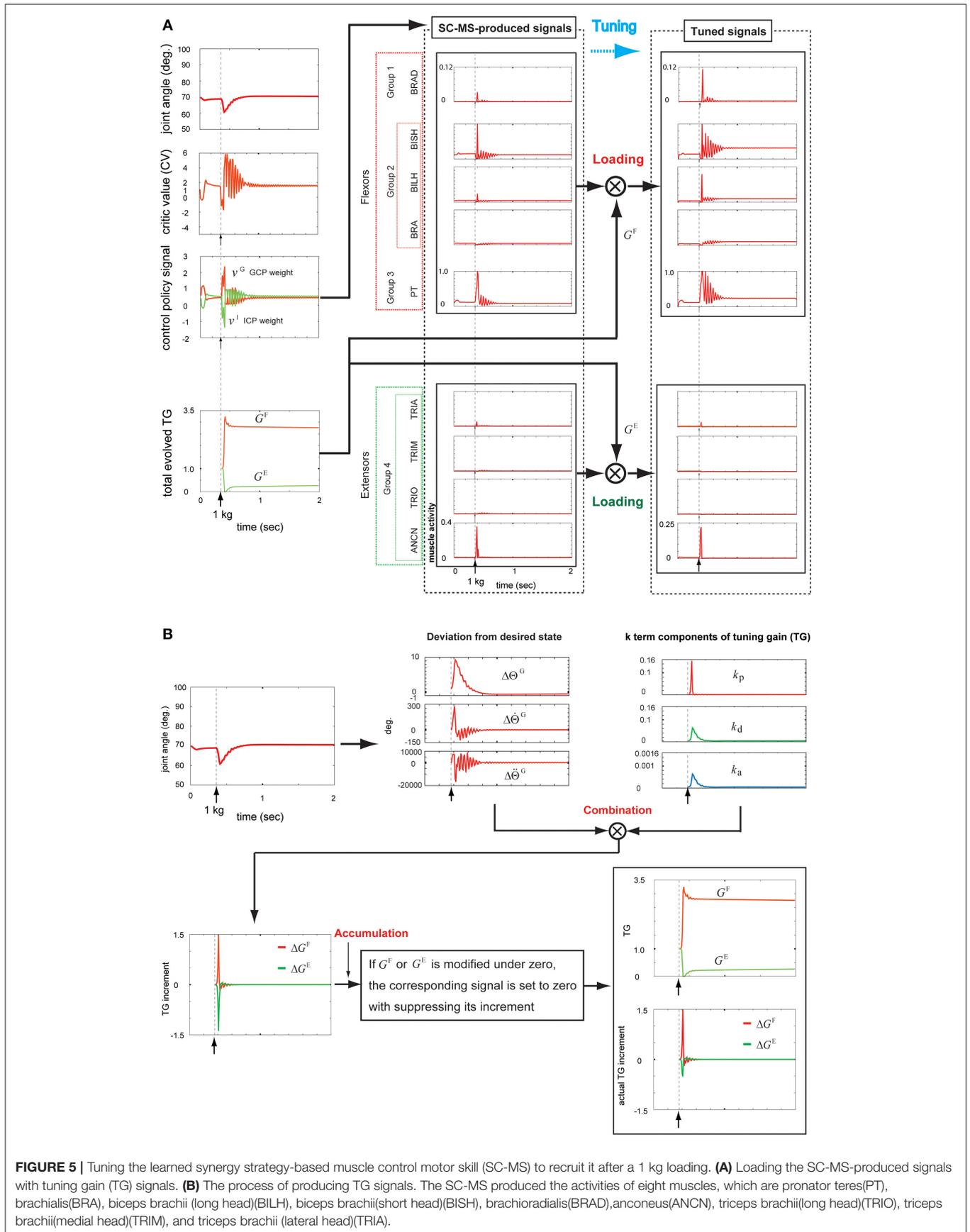


FIGURE 5 | Tuning the learned synergy strategy-based muscle control motor skill (SC-MS) to recruit it after a 1 kg loading. **(A)** Loading the SC-MS-produced signals with tuning gain (TG) signals. **(B)** The process of producing TG signals. The SC-MS produced the activities of eight muscles, which are pronator teres (PT), brachialis (BRA), biceps brachii (long head) (BILH), biceps brachii (short head) (BISH), brachioradialis (BRAD), anconeus (ANCN), triceps brachii (long head) (TRIO), triceps brachii (medial head) (TRIM), and triceps brachii (lateral head) (TRIA).

velocity deviations from the desired states, $\Delta\Theta_t^G$ and $\Delta\dot{\Theta}_t^G$, respectively, in the second graph in the top row of **Figure 5B**. As described in **Figure 2A**, the component k_p is designed to be sensitive to the increase in $\Delta\dot{\Theta}_t^G$, which is attributed to the high undershooting. Therefore its increase was faster and higher than both k_d and k_a that are designed to respond to the increase of $\|\Delta\Theta_t^G\|$, which was attributed to the undershooting or overshooting angle deviation (**Figure 2B**).

According to Equation (4), these three k -term components combined with their corresponding deviations from the desired states to produce three incremental components of G^F and G^E , as shown in the second and third graphs in the top row of **Figure 5B**. This combination produced the ΔG^F and ΔG^E , as shown in the first graph in the bottom row of **Figure 5B**. According to the rule of Equation (6), G^F and G^E must be above zero. Therefore, if either $G^F(t)$ or $G^E(t)$ is modified to be below zero, the corresponding signal is set to zero by suppressing its increment. Under this rule, ΔG^F and ΔG^E were regulated to modify G^F and G^E , respectively, as shown in the second graph in the bottom row of **Figure 5B**. The bottom part of this graph shows ΔG^F and ΔG^E , which actually contribute to modifying the G^F and G^E , respectively. The tracks of G^F or G^E were as follows.

After a sustained 1 kg loading, G^F increased from 1.0 to its peak value of 3.3 in response to the drop in the joint position, followed by a decrease in responding to the recovery of the preloading context, and finally remained at a stable value of 2.7 to maintain the preloading context. By contrast, G^E decreased under the same conditions from 1.0 to the lower value 0.0 in response to the decrease in the joint position before increasing, reflecting the preloading context recovery, and finally reached a stable value of 0.25 to retain the preloading posture. Hence, G^F increased and G^E decreased from 1.0 during the recovery of the preloading posture in response to a sustained disturbance. These TG signals, as shown in **Figure 5A**, contributed to additionally modulating the SC-MS through tuning its signals as follows.

The parameter G^F is reflected in the increased activities of agonists that were kept at higher values in the recovered preloading context compared with their preloading activities. Conversely, the parameter G^E decreased the antagonist activities and then maintained them at lower values in the recovered preloading context in comparison with their preloading activities.

The aforementioned results demonstrate that the SC-MS can be robustly recruited for a novel feedback context with additional modulation, which was achieved through tuning its signals.

T-SC in Further Novel Contexts

As shown in section T-SC in a Novel Sustained Disturbance, we verified that an SC-MS can be robustly recruited by tuning it for a novel sustained 1 kg loading, which did not involve learning of the SC-MS. In this section, we demonstrate the versatility of this recruitment process in further novel contexts.

Recruitment in Undershooting Attributed to a Novel Sustained 2 kg Loading

To examine the recruitment of a learned SC-MS by tuning it in an additional severe undershooting context, we simulated the recruitment of a T-SC-MS for sustained 2 kg loading, which is

two times the weight of the 1 kg loading used in section T-SC in a Novel Sustained Disturbance. The simulation of this recruitment process is shown in **Figures 6A,B**. The top graph of the first column in **Figure 6A** shows the joint angular trace during the recruitment process of the SC-MS for this disturbance as follows.

After loading with 2 kg, the joint angular trace dropped below the preloading position to about 62° and then increased to be maintained at about 65° for a short time. However, the joint angle declined again to about 55° , but finally recovered to about 75° near the goal state and then was kept stable at that state. Two decreases in angle value and some overshoot during the recruitment of the SC-MS showed an incomplete recovery for 2 kg loading in comparison with the process for 1 kg. This difference is attributed to further severe disturbances over a 1 kg loading. This movement could be achieved through the following recruitment processes of the T-SC-MS.

As mentioned in section T-SC in a Novel Sustained Disturbance, the parameter CV evaluates the recruitment of an SC-MS to achieve the goal state. The second graph of the first column in **Figure 6A** shows the CV as the evaluation of the recruitment of T-SC-MS for a 2 kg loading. Owing to the aforementioned severe decrease, the CV for a 2 kg loading decreased further than the CV for a 1 kg loading during undershooting. As shown in the third graph of the first column of **Figure 6A**, this decrease in the CV increased the GCP more than the decrease in the CV under the 1 kg loading. Accordingly, the ICP was suppressed further than that under the 1 kg loading. This CV-driven synergy between two control policies regulates the control redundancy of individual muscle units. Through this regulation, the muscle activities are produced, as shown in the second column of **Figure 6A**. These signals were loaded with G^F and G^E according to the recruitment process of a learned MS in **Figure 1**. These two TG signals were dynamically produced through the following feedback gain process of a learned MS in **Figure 6B**.

Under the rule based on Equation (5), the three k -term components of the TG shown in the third graph in the top row of **Figure 6B** were determined by the joint angle and velocity deviations from the desired states, $\Delta\Theta_t^G$ and $\Delta\dot{\Theta}_t^G$, respectively, in the second graph in the top row of **Figure 6B**. After loading with 2 kg, as mentioned above, the joint movement developed in two downward steps. In the first step, the component k_p drastically increased to respond to the increase of $\Delta\dot{\Theta}_t^G$, which was attributed to the high-speed downward motion. Concurrently, both k_d and k_a increased by less than k_p because they responded to $\|\Delta\Theta_t^G\|$, the increase in which was less than the increase in $\Delta\dot{\Theta}_t^G$. During this step, the response traces of the three k -term components were similar to those observed with a 1 kg loading. In the second step, all three k -term components increased substantially in response to the large increases in both $\|\Delta\Theta_t^G\|$ and $\Delta\dot{\Theta}_t^G$. This is attributed to the feedback context, in which the joint angle state was far from its preloading state with a high downward speed. After this second drop, the joint angular state mostly recovered by 0.8 s to its preloading goal state before slowly reaching the preloading state. In response to this recovery, the k_p drastically decreased in response to the decrease in $\Delta\dot{\Theta}_t^G$ in

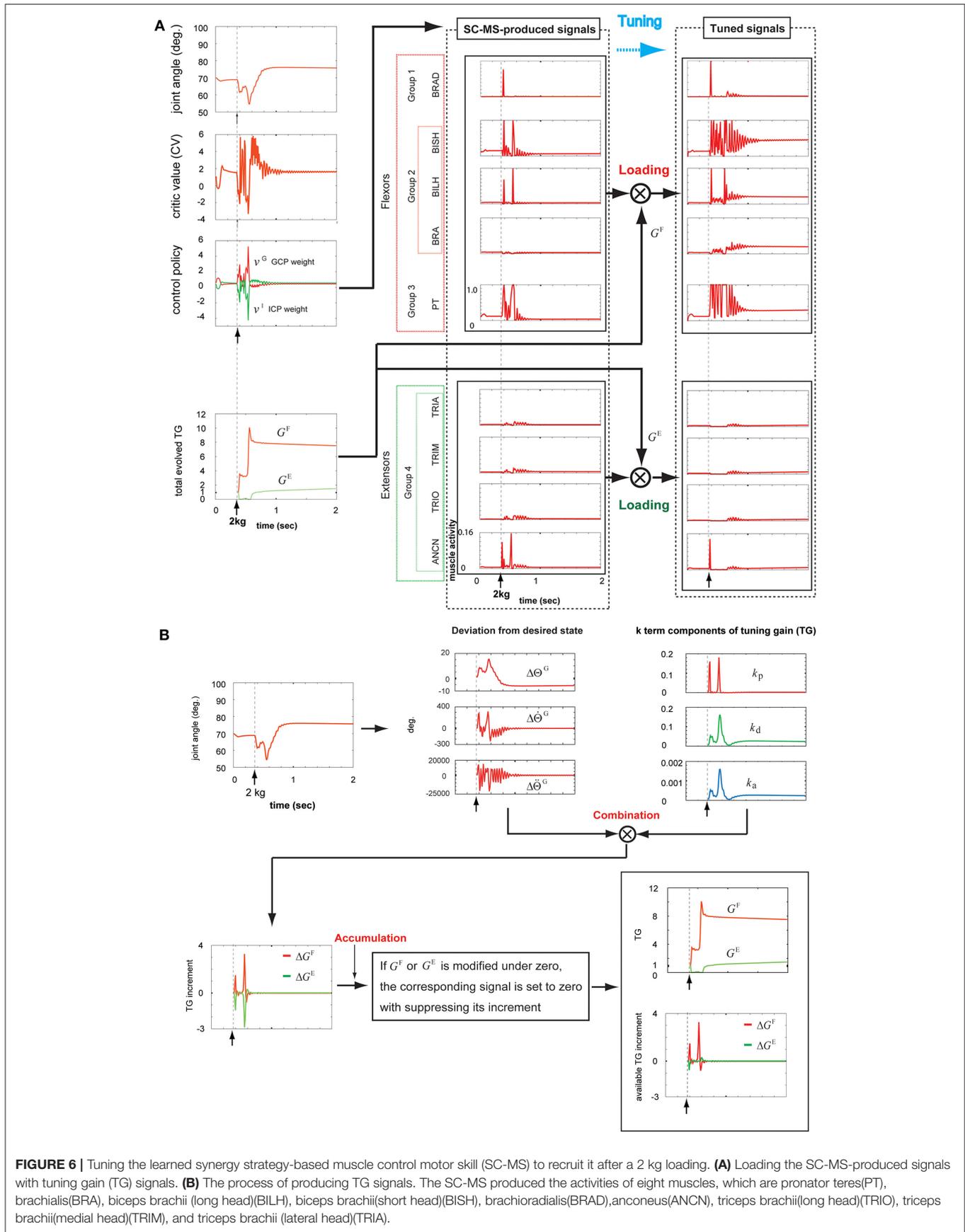


FIGURE 6 | Tuning the learned synergy strategy-based muscle control motor skill (SC-MS) to recruit it after a 2 kg loading. **(A)** Loading the SC-MS-produced signals with tuning gain (TG) signals. **(B)** The process of producing TG signals. The SC-MS produced the activities of eight muscles, which are pronator teres(PT), brachialis(BRA), biceps brachii (long head)(BILH), biceps brachii(short head)(BISH), brachioradialis(BRAD),anconeus(ANCN), triceps brachii(long head)(TRIO), triceps brachii(medial head)(TRIM), and triceps brachii (lateral head)(TRIA).

downward speed, and both k_d and k_a concurrently decreased in response to the decrease in $\|\Delta\Theta_t^G\|$. After 0.8 s, k_d and k_a increased slightly by about 1.0 s in response to the slight overshooting of $\|\Delta\Theta_t^G\|$ and then decreased quite slowly in response to the quite slow decrease of $\|\Delta\Theta_t^G\|$ to zero. Under the rule of Equation (4), these three k -term components combined with the corresponding deviations from the desired state, which are shown in the second graph in the top row of **Figure 6B**. These combinations produced the ΔG^F and ΔG^E , as shown in the first graph in the bottom row of **Figure 6B**. According to the rule of Equation (6), ΔG^F and ΔG^E were regulated to modify G^F and G^E , respectively, as shown in the second graph in the bottom row of **Figure 6B**. The bottom part of this graph shows ΔG^F and ΔG^E , which actually contribute to modifying the G^F and G^E , respectively. The tracks of the TGs and their contribution to tuning the SC-MS were as follows.

G^F increased in response to the two drops in joint angle value and then slowly decreased to a stable level. To assist G^F , the parameter G^E was suppressed during the first drop, but increased substantially during the second drop before slowly dropping to a stable level. These two agonist and antagonist TG signals were loaded onto the SC-MS-produced muscle activities, as shown in the second and third columns of **Figure 6A**. Through this loading, they were tuned to the optimal muscle activities for recruiting the SC-MS under a 2 kg loading.

Recruitment in Overshooting Attributed to a Novel Sustained-1 kg Loading

The overshooting during the SC-MS learning process is transiently driven by incorrectly controlling the joint and is eventually suppressed by gravity. Therefore, the SC-MS learned to control the overshooting with very little extensor activation, which functions as the agonist for overshooting. Because of this learning condition of the SC-MS, the overshooting driven by the sustained negative disturbance on the hand is further severe novel disturbance in recruiting the SC-MS than the undershooting driven by the sustained positive disturbances such as 1 kg or 2 kg loading. Therefore, to recruit the SC-MS during the sustained negative disturbance-driven overshooting, the SC-MS needs to be tuned more than the sustained positive disturbance-driven undershooting. By simulating the recruitment of the T-SC-MS during the overshooting driven by a sustained negative disturbance, we tested the tuning process to robustly recruit the SC-MS in an entirely novel context as follows.

After loading a -1 kg weight on the simulated hand, the joint angular trace was raised to about 98° and then decreased to about 66° (top graph, first column of **Figure 7A**). Finally, the joint angular trace overshot by about 76° and then stably recovered to the preloading state. According to this contextual joint angular trace, the CV was determined as shown in the second graph of the first column of **Figure 7A**. Further, gravity, which reflects the movement, needs to be considered in determining the CV. As mentioned in the first paragraph of this section, it is comparatively easy for an SC-MS to suppress the incorrect control-driven transient overshooting because of gravity during its learning process. Therefore, even if the overshooting attributed to the sustained negative disturbance on the hand is a further severe context for the SC-MS, the CV

for recruiting the SC-MS during the overshooting is less than that during the undershooting [second graph, second column of **Figure 4** (first column of 5A), and first column of 6A]. Because of this evaluation of the CV, as shown in the third graph in the first column of **Figure 7A**, the change in the GCP and ICP weights in response to this negative sustained disturbance was suppressed to a small range in comparison with its response to positive sustained disturbances such as a 1 or 2 kg loading. This process was regulated using Equation (1). As mentioned above, this response is attributed to the learning condition of SC-MS, in which the transient overshooting driven by incorrect control is controlled by a small amount of activity of the extensors because gravity contributes to the recovery of the preloading state from its overshooting state. Because of this learning condition of SC-MS, the extensors functioning as agonists for negative sustained disturbance need to be loaded more than the flexors functioning as agonists for positive sustained disturbance. To process this additional modulation, the SC-MS-produced muscle activities shown in the second column of **Figure 7A** were loaded with the antagonistic and agonistic TGs, G^F and G^E , respectively. These TGs were produced by the following process.

As shown in the first graph in the bottom row of **Figure 7B**, the three incremental components of G^F and G^E were produced by combining the three k -term components of the TG (third graph, top row of **Figure 7B**) and their corresponding deviations from the desired state (second graph, top row of **Figure 7B**). These three components of G^F and G^E were summed to produce ΔG^F and ΔG^E , respectively, which were accumulated to produce G^F and G^E as shown in the second graph in the bottom row of **Figure 7B**. The precise process was achieved as follows.

In agreement with the rules shown in **Figures 2A,B**, during the initial overshooting shown in the first graph in the top row of **Figure 7B**, the k_p component was almost suppressed in response to the high overshooting but the k_d and k_a components drastically increased in response to the large increase in $\|\Delta\Theta_t^G\|$. After the joint movement started to recover to its preloading state at about 0.57 s, k_p drastically increased in response to the large increase in $\Delta\Theta_t^G$; this increase was attributed to the high-speed downward motion, and both the k_d and k_a components decreased in response to the decrease in $\|\Delta\Theta_t^G\|$. Thereafter, the joint angle finally recovered to its preloading state via the slight undershooting and the second overshooting, which was less pronounced. In response to this recovery, k_p decreased to almost zero, whereas k_d and k_a stably decreased to their preloading values via their transient increase, as shown in the third graph in the top row of **Figure 7B**. According to Equation (4), these three k -term components combined with their corresponding deviations of desired states to produce the ΔG^F and ΔG^E , as shown in the first graph in the bottom row of **Figure 7B**. Under the rule of Equation (6), the ΔG^F and ΔG^E were regulated to modify the G^F and G^E , respectively, as shown in the second graph in the bottom row of **Figure 7B**. The bottom part of this graph shows the ΔG^F and ΔG^E , which actually contribute to modifying the G^F and G^E , respectively. The tracks of the TGs and their contribution to tuning the SC-MS were as follows.

G^E substantially increased during the overshooting and then stabilized at a lower level owing to the recovery of the preloading state. To assist G^E , G^F was completely suppressed during the

overshoot and then increased for a very short time before decreasing slowly to a stable level.

Because of this modulation of G^F , as shown in the top of the second and third columns of **Figure 7A**, the SC-MS-produced signals of flexors were dynamically unloaded in response to the overshooting attributed to a negative sustained disturbance because they were antagonists for the negative disturbance. In comparison with the flexors, the extensors were highly loaded with G^E to function as agonists against the overshooting (the bottom part of the second and third columns, **Figure 7A**). Through this tuning process, the handicap in recruiting an SC-MS under overshooting conditions attributed to sustained negative disturbances, which was mentioned in the first paragraph of this section, could be overcome.

DISCUSSION

A learned MS can potentially be used for effective motor control in a novel context. In addressing this issue, we hypothesized that an MS can be retained through learning it in the CNS and then recruiting it. Through the simulation using the proposed neurophysiological computational model, we have shown that the MS might be retained through learning the muscle synergy to achieve its task and recruited through dynamically tuning it according to novel feedback contexts. In this tuning, the learned muscle synergy, termed SC-MS, produces the muscle control signals through its dynamic modulation according to the feedback context and these signals are additionally loaded with tuning signals, termed TG signals, which are dynamically modulated according to the feedback context. Through this dynamic modulation, a skilled MS might be recruited in a variety of conditions besides those experienced during motor learning. Furthermore, this involvement of the muscles' synergy with a skilled MS demonstrates that it might subserve the learning and retaining of an MS in the CNS.

Dynamic Modulation of an MS According to the Feedback Context

To recruit a learned MS for novel contexts, we assumed that a learned MS is dynamically modulated for the feedback context. In addressing this issue, we used the concept of SC (Min et al., 2018), which dynamically regulates the redundant functional roles of individual muscles according to consecutive feedback contexts. As described in Equations (1) and (2), this SC-driven regulation contributes to the dynamic modulation of an MS for the feedback context. Consequently, this modulation contributes to robust recruitment of an MS in various novel feedback contexts that did not involve the learning of an MS, as shown in **Figures 4, 5A, 6A, 7A**. These results show that the SC may be an optimal strategy to learn an MS and to recruit it.

Robust Recruitment of a Learned MS Through Tuning It According to the Feedback Context and Its Implications

Even if an SC-MS is modulated according to the feedback context as mentioned in the above subsection, this modulation is learned

in the dynamics under the learning context of SC-MS. Because of this learning condition, to robustly recruit an SC-MS in a novel context, it needs to be additionally tuned. To validate this tuning, we hypothesized that a muscle loading tuner may operate in the CNS to tune the SC-MS through dynamically loading its muscle control signals according to the feedback context. This hypothesis was validated with the simulation results shown in **Figures 5A, 6A, 7A**, in which the SC-MS could be successively recruited through dynamically loading its muscle control signals according to the feedback context under three different novel sustained disturbances. This recruitment may involve the rapid adaptation of motion control to novel contexts without learning a new MS for them. If this rapid adaptation is impaired, the normal motion control in novel dynamic contexts may be seriously disturbed. To test the potential clinical implications, this hypothesis needs to be further studied in neurophysiology. Through this study, the proposed model may provide a new clinical view of motion control disorders attributed to cortico-BG loop-related CNS diseases in pathophysiology and therapeutics/rehabilitation. Furthermore, through the transcortical circuit, the recruitment-produced muscle control signals may be transferred to the cerebellum as a correction signal to train a neural network, on which a feedforward motor command is generated in the cerebellum (Kawato et al., 1987; Kawato, 1990; Kambara et al., 2009). Therefore, the T-SC may involve robust feedforward motion control in novel contexts.

Previous studies, such as proportional integral derivative control (Petkos and Vijayakumar, 2007) and optimal feedback control (Todorov and Jordan, 2002; Liu and Todorov, 2007) in modeling the feedback control process, focused only on the correction of the motor control error but did not address the contribution of a learned MS to feedback control. Our new approach to recruitment of a learned MS in novel contexts may offer a new viewpoint for this previously unaddressed feedback control issue.

Recruiting a Learned MS via the Cortico-Basal Ganglia Loop

The BG contributes to “stabilization augmentation” by facilitating an optimal activity that fits the desired situation and context while suppressing other ongoing CNS activities that would interfere with the desired behavior (Mink, 1996). Furthermore, Turner and Anderson (1997) showed that movement-related changes in pallidal discharge to specific parameters of movement are discharge of neurons in the skeletomotor portions of both pallidal segmentations. This BG response is demonstrated by encoding the combination of the sensory and contextual state through the sensory feedback process, which may involve online motion control with the selective facilitation and suppression of different frontal thalamocortical circuits (Turner and Anderson, 1997). As the BG reinforces a new MS through reinforcement learning and retains it subsequently (Lehéricy et al., 2005), this online motion control role of the BG may involve the recruitment of a learned MS retained in the BG, which is dynamically modulated by the selective facilitation and suppression of different frontal

thalamocortical circuits. Based on the aforementioned previous studies, this cortico-BG scheme may be a common framework for the learning and recruitment of an MS in the CNS. Therefore, the T-SC-driven recruitment of a learned MS through the cortico-BG loop may involve different kinds of motion control, which need to respond to various sensory feedback contexts via the M1 from different sensory areas, including the somatosensory cortex and the visual cortex. This hypothesis may be reasonable, even if it has recently been demonstrated that the roles of neural structures differ between different tasks (Paparella et al., 2020).

Muscle Control Scheme of the Corticospinal Tract in Recruiting an SC-MS

The experimental evidence introduced in section Recruiting a Learned MS via the Cortico-Basal Ganglia Loop shows that the BG may retain a learned MS and involve the recruitment of it to control movement according to the feedback context. Based on this concept, to recruit an SC-MS according to the feedback context, we assumed that the BG may dynamically modulate an SC-MS with the synergistic combination of two control policies of the SC, GCP, and ICP, which is driven by a combination of their inhibition and disinhibition. As shown in **Figure 1**, this synergistic combination of GCP and ICP in the BG produces muscle control signals through the corticospinal tract. As outlined in section Introduction, GCP-driven signals may function as group unit control signals that are decoded into synergistic combinations of MPs (Bizzi et al., 1991; d'Avella et al., 2003) retained in the spinal cord because the group units produce the contraction sets of muscles termed MPs in processing the SC. Furthermore, as the ICP-driven signals serve as the control signals for individual muscle units, they may be directly copied from the corticomotor neurons among the CSTs to MNs. Therefore, the ICP-driven signals sculpt GCP-driven signals through their synergistic combination to optimally modulate an SC-MS according to the feedback context. This recruitment of an SC-MS may support the concept introduced in section Introduction that muscle activities are produced by combining two pathways of MNs (Rathelot and Strick, 2009).

Evaluating the Proposed Model in Comparison With Human Subjects

Evaluating the proposed computational model in comparison with human subjects, the two disadvantages of the computational model were as follows.

In this study, an SC-MS was learned only through one learning experience of a particular task, which was to move the hand to its goal within a limited joint angular range as described in section Learning and Recruitment Condition of an SC-MS. While learning the SC-MS, no disturbances were involved (Min et al., 2018), as described in section Learning and Recruitment Condition of an SC-MS. Therefore, the recruitment of an SC-MS under sustained disturbance was simulated as a pure novel recruitment, as described in section Results. To evaluate this simulation in comparison with human subjects, as shown in the top row of **Figure 4**, we approximated a pure novel recruitment as closely as possible using only those data that were recorded

during the first trial for each of the four subjects. However, the subjects have experienced and learned various tasks during their whole life and thereby have experienced various tasks under various sustained disturbances. Therefore, the sustained 1 kg loading on the hand is not a pure novel context for these subjects. Consequently, this should be taken into account when evaluating the simulation results through a comparison with the subjects' movements. Because of the disadvantage attributed to pure novel recruitment, an SC-MS is even more difficult to recruit under novel sustained disturbances, such as a sustained 1 kg loading, than the subjects. Considering this disadvantage, we may evaluate that an SC-MS can be robustly recruited through the proposed recruitment model termed T-SC.

As mentioned in section Introduction, innate and learned MSs are recruited in the CNS for effective and fast motion control in response to novel external disturbances. To validate this, the recruitment of a learned MS in a pure feedback control process is the most optimal task because the pure feedback control, which is not involved in the prediction of any disturbance, may need the most effective and fast response to the feedback context. Therefore, as described in section Results, novel recruitment with T-SC was simulated in pure feedback control. To evaluate this simulation by comparison with human subjects' movements, as shown in the top row of **Figure 4**, we approximated this pure feedback control process as closely as possible, as described in section Experimental Setup, through an experimental setting in which the subjects were blindfolded and not informed regarding the timing of the loading. To avoid the weight being misloaded on the subjects' hands, as shown in **Figure 3B**, the distance between the initial falling point of the weight and the initial position of the hand was set close to zero. Further, we instructed the subjects not to predict the timing of the loading weight. However, even though this instruction was given to the subjects, they might instinctively have some preliminary joint stiffness by co-contraction of both agonists and antagonists in preparation for the incoming disturbance before loading. Because of this, as shown in the top row of the right column of **Figure 4**, the mean joint angular trace of the subjects after loading undershot was slower than the simulating joint angular trace. In evaluating the simulation results in the top row of **Figure 4**, we considered that the simulation model was disadvantaged in responding to a disturbance in comparison with human subjects.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

ETHICS STATEMENT

All subjects were provided written informed consent prior to their participation. The protocol was approved by the ethics committees of the Tokyo Metropolitan Institute of Medical Science, and it was conducted in accordance with the ethical standards of the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

KM conceived and designed the theoretical and computational model and the experiments. KM and JL conducted the experiments. KM and SK analyzed the data. KM, SK, and JL wrote the manuscript. All authors read and approved the version to be published and agreed to be accountable for all aspects of the work.

FUNDING

This research was supported by the Tokyo Metropolitan Institute of Medical Science and grants-in-aid from the Ministry of

Education, Culture, Sports, Science and Technology in Japan (<http://www.mext.go.jp/>) (20K12638) to KM. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We thank Dr. Masami Iwamoto for the valuable advice on the motion control mechanisms of the musculoskeletal system. We also thank Dr. Yasuharu Koike for permitting us to use the measuring devices in his laboratory and Dr. Duk Shin for his experimental assistance.

REFERENCES

- Amis, A. A., Dowson, D., and Wright, V. (1979). Muscle strengths and musculoskeletal geometry of the upper limb. *Eng. Med.* 8, 41–48. doi: 10.1243/EMED_JOUR_1979_008_010_02
- Amundsen Huffmaster, S. L., Van Acker, G. M. III, Luchies, C. W., and Cheney, P. D. (2018). Muscle synergies obtained from comprehensive mapping of the cortical forelimb representation using stimulus triggered averaging of EMG activity. *J. Neurosci.* 38, 8759–8771. doi: 10.1523/JNEUROSCI.2519-17.2018
- Barroso, F. O., Torricelli, D., Moreno, J. C., Taylor, J., Gomez-Soriano, J., Bravo-Esteban, E., et al. (2014). Shared muscle synergies in human walking and cycling. *J. Neurophysiol.* 112, 1984–1998. doi: 10.1152/jn.00220.2014
- Barto, A. G. (1995). “Adaptive critics and the basal ganglia,” in *Models of Information Processing in the Basal Ganglia*, eds J. C. Houk, J. L. Davis, and D. G. Beiser (Cambridge, MA: MIT Press), 215–232.
- Bizzzi, E., Mussa-Ivaldi, F. A., and Giszter, S. (1991). Computations underlying the execution of movement: a biological perspective. *Science* 253, 287–291. doi: 10.1126/science.1857964
- d’Avella, A., and Pai, D. K. (2010). Modularity for sensorimotor control evidence and a new prediction. *J. Mot. Behav.* 42, 361–369. doi: 10.1080/00222895.2010.526453
- d’Avella, A., Saltiel, P., and Bizzzi, E. (2003). Combinations of muscle synergies in the construction of a natural motor behavior. *Nat. Neurosci.* 6, 300–308. doi: 10.1038/nn1010
- De Marchis, C., Di Somma, J., Zych, M., Conforto, S., and Severini, G. (2018). Consistent visuomotor adaptations and generalizations can be achieved through different rotations of robust motor modules. *Sci. Rep.* 8:12657. doi: 10.1101/362228
- Doya, K. (2000). Reinforcement learning in continuous time and space. *Neural Comput.* 12, 219–245. doi: 10.1162/089976600300015961
- Doya, K. (2007). Reinforcement learning: computational theory and biological mechanisms. *HFSP J.* 1, 30–40. doi: 10.2976/1.2732246/10.2976/1
- Doya, K. (2008). Modulators of decision making. *Nat. Neurosci.* 11, 410–416. doi: 10.1038/nn2077
- Giszter, S. F., Mussa-Ivaldi, F. A., and Bizzzi, E. (1993). Convergent force fields organized in the frog’s spinal cord. *J. Neurosci.* 13, 467–491. doi: 10.1523/JNEUROSCI.13-02-00467.1993
- Hada, M., Yamada, D., and Tsuji, T. (2007). “An analysis of equivalent impedance characteristics by modeling the human musculoskeletal structure as a multibody system,” in *Proceeding of the ECCOMAS Thematic Conference on Multibody Dynamics*, BMI (Milan), 1–20.
- Herter, T. M., Korbelt, T., and Scott, S. H. (2009). Comparison of neural responses in primary motor cortex to transient and continuous loads during posture. *J. Neurophysiol.* 101, 150–163. doi: 10.1152/jn.90230.2008
- Hikosaka, O., Takikawa, Y., and Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.* 80, 954–978. doi: 10.1152/physrev.2000.80.3.953
- Hill, A. V., and Sec, R. C. (1938). The heat of shortening and the dynamic constants of muscle. *Proc. Roy. Soc.* 126B, 136–195.
- Houk, J. C., Adams, J. L., and Barto, A. G. (1995). “A model of how the basal ganglia generate and use neural signals that predict reinforcement,” in *Models of Information Processing in the Basal Ganglia*, eds J. C. Houk, J. L. Davis, and D. G. Beiser (Cambridge, MA: MIT Press), 249–270.
- Ito, M., and Doya, K. (2011). Multiple representations and algorithms for reinforcement learning in the cortico-basal ganglia circuit. *Curr. Opin. Neurobiol.* 21, 368–373. doi: 10.1016/j.conb.2011.04.001
- Kambara, H., Kim, K., Shin, D., Sato, M., and Koike, Y. (2009). Learning and generation of goal-directed arm reaching from scratch. *Neural Netw.* 22, 348–361. doi: 10.1016/j.neunet.2008.11.004
- Kawato, M. (1990). Feedback-error-learning neural network for supervised motor learning. *Adv. Neural Comput.* 365–372. doi: 10.1016/B978-0-444-88400-8.50047-9
- Kawato, M., Furukawa, K., and Suzuki, R. (1987). A hierarchical neural-network model for control and learning of voluntary movement. *Biol. Cybern.* 57, 169–185. doi: 10.1007/BF00364149
- Kibushi, B., Hagio, S., Moritani, T., and Kouzaki, M. (2018). Lower local dynamic stability and invariable orbital stability in the activation of muscle synergies in response to accelerated walking. *Speeds. Front. Hum. Neurosci.* 12:485. doi: 10.3389/fnhum.2018.00485
- Lehéricy, S., Benali, H., Van de Moortele, P. F., Péligrini-Issac, M., Waechter, T., Ugurbil, K., et al. (2005). Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12566–12571. doi: 10.1073/pnas.0502762102
- Liu, D., and Todorov, E. (2007). Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *J. Neurosci.* 27, 9354–9368. doi: 10.1523/JNEUROSCI.1110-06.2007
- London, B. M., and Miller, L. E. (2013). Responses of somatosensory area 2 neurons to actively and passively generated limb movements. *J. Neurophysiol.* 109, 1505–1513. doi: 10.1152/jn.00372.2012
- Min, K., Iwamoto, M., Kakei, S., and Kimpara, H. (2018). Muscle synergy-driven robust motion control. *Neural Comput.* 30, 1104–1131. doi: 10.1162/neco_a_01063
- Mink, J. W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* 50, 381–425. doi: 10.1016/S0301-0082(96)00042-1
- Murray, W. M., Delp, S. L., and Buchanan, T. S. (1995). Variation of muscle moment arms with elbow and forearm position. *J. Biomech.* 28, 513–525. doi: 10.1016/0021-9290(94)00114-J
- Nambu, A., Tokuno, H., and Takada, M. (2002). Functional significance of the corticostriatal-thalamo-pallidal ‘hyperdirect’ pathway. *Neurosci. Res.* 43, 111–117. doi: 10.1016/S0168-0102(02)00027-5
- Nashed, J. Y., Kurtzer, I. L., and Scott, S. H. (2015). Context-dependent inhibition of unloaded muscles during the long-latency epoch. *J. Neurophysiol.* 113, 192–202. doi: 10.1152/jn.00339.2014
- Neumann, D. A. (2002). *Kinesiology of the Musculoskeletal System: Foundations for Physical Rehabilitation*. Maryland Heights, MO: Mosby.
- Paparella, G., Rocchi, L., Bologna, M., Berardelli, A., and Rothwell, J. (2020). Differential effects of motor skill acquisition on the primary motor and sensory cortices in healthy humans. *J. Physiol.* 598, 4031–4045. doi: 10.1113/JP279966

- Petkos, G., and Vijayakumar, S. (2007). "Load estimation and control using learned dynamics models," in *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Piscataway, NJ: IEEE), 1527–1532.
- Pruszynski, J. A., Kurtzer, I., Nashed, J. Y., Omrani, M., Brouwer, B., and Scott, S. H. (2011). Primary motor cortex underlies multi-joint integration for fast feedback control. *Nature* 478, 387–390. doi: 10.1038/nature10436
- Rathelot, J. A., and Strick, P. L. (2009). Subdivisions of primary motor cortex based on cortico-motoneuronal cells. *Proc. Natl. Acad. Sci. U.S.A.* 106, 918–923. doi: 10.1073/pnas.0808362106
- Safavynia, S. A., and Ting, L. H. (2012). Task-level feedback can explain temporal recruitment of spatially fixed muscles synergy throughout postural perturbations. *J. Neurophysiol.* 107, 159–177. doi: 10.1152/jn.00653.2011
- Selen, L. P., Shadlen, M. N., and Wolpert, D. M. (2012). Deliberation in the motor system: reflex gains track evolving evidence leading to a decision. *J. Neurosci.* 32, 2276–2286. doi: 10.1523/JNEUROSCI.5273-11.2012
- Shinnamon, H. M. (1993). Preoptic and hypothalamic neurons and initiation of locomotion in the anesthetized rat. *Prog. Neurobiol.* 41, 323–344. doi: 10.1016/0301-0082(93)90003-B
- Spraker, M. B., Yu, H., Corcos, D. M., and Vaillancourt, D. E. (2007). Role of individual basal ganglia nuclei in force amplitude generation. *J. Neurophysiol.* 98, 821–834. doi: 10.1152/jn.00239.2007
- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning*. Cambridge, MA: MIT Press, 152–153.
- Suzuki, T., Kinugasa, R., and Fukushima, S. (2017). Activation of plantar flexor muscles is constrained by multiple muscle synergies rather than joint torques. *PLoS ONE* 12:e0187587. doi: 10.1371/journal.pone.0187587
- Thelen, D. G. (2003). Adjustment of muscle mechanics model parameters to simulate dynamic contractions in older adults. *J. Biomech. Eng.* 125, 70–77. doi: 10.1115/1.1531112
- Ting, L. H., and Macpherson, J. M. (2012). A limited set of muscle synergies for force control during a postural task. *J. Neurophysiol.* 93, 609–613. doi: 10.1152/jn.00681.2004
- Todorov, E., and Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* 5, 1226–1235. doi: 10.1038/nn963
- Toma, S., and Santello, M. (2019). Motor modules account for active perception of force. *Sci. Rep.* 9:8983. doi: 10.1038/s41598-019-45480-w
- Torres-Oviedo, G., Macpherson, J. M., and Ting, L. H. (2006). Muscle synergy organization is robust across a variety of postural perturbations. *J. Neurophysiol.* 96, 1530–1546. doi: 10.1152/jn.00810.2005
- Tresch, M. C., Philippe, S., and Emilio, B. (1999). The construction of movement by the spinal cord. *Nat. Neurosci.* 2, 162–167. doi: 10.1038/5721
- Turner, R. S., and Anderson, M. E. (1997). Pallidal discharge related to the kinematics of reaching movements in two dimensions. *J. Neurophysiol.* 77, 1051–1074. doi: 10.1152/jn.1997.77.3.1051
- Zajac, F. E., Topp, F. L., and Stevenson, P. J. (1985). "A dimensionless musculotendon model," in *Proceedings of the IEEE/Eight Annual Conference of the Engineering in Medicine and Biology Society* (Piscataway, NJ: IEEE), 601–604.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Min, Lee and Kakei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.