Ao and Fregly
Journal of NeuroEngineering and Rehabilitation (2024) 21:194
https://doi.org/10.1186/s12984-024-01490-y

Journal of NeuroEngineering and Rehabilitation

RESEARCH

Open Access

Comparison of synergy extrapolation and static optimization for estimating multiple unmeasured muscle activations during walking

Di Ao[1*] and Benjamin J. Fregly[2]

Abstract

Background Calibrated electromyography (EMG)-driven musculoskeletal models can provide insight into internal quantities (e.g., muscle forces) that are difficult or impossible to measure experimentally. However, the need for EMG data from all involved muscles presents a significant barrier to the widespread application of EMG-driven modeling methods. Synergy extrapolation (SynX) is a computational method that can estimate a single missing EMG signal with reasonable accuracy during the EMG-driven model calibration process, yet its performance in estimating a larger number of missing EMG signals remains unknown.

Methods This study assessed the accuracy with which SynX can use eight measured EMG signals to estimate muscle activations and forces associated with eight missing EMG signals in the same leg during walking while simultaneously performing EMG-driven model calibration. Experimental gait data collected from two individuals post-stroke, including 16 channels of EMG data per leg, were used to calibrate an EMG-driven musculoskeletal model, providing "gold standard" muscle activations and forces for evaluation purposes. SynX was then used to predict the muscle activations and forces associated with the eight missing EMG signals while simultaneously calibrating EMG-driven model parameter values. Due to its widespread use, static optimization (SO) applied to a scaled generic musculoskeletal model was also utilized to estimate the same muscle activations and forces. Estimation accuracy for SynX and SO was evaluated using root mean square errors (RMSE) to quantify amplitude errors and correlation coefficient *r* values to quantify shape similarity, each calculated with respect to "gold standard" muscle activations and forces.

Results On average, compared to SO, SynX with simultaneous model calibration produced significantly more accurate amplitude and shape estimates for unmeasured muscle activations (RMSE 0.08 vs. 0.15, *r* value 0.55 vs. 0.12) and forces (RMSE 101.3 N vs. 174.4 N, *r* value 0.53 vs. 0.07). SynX yielded calibrated Hill-type muscle–tendon model parameter values for all muscles and activation dynamics model parameter values for measured muscles that were similar to "gold standard" calibrated model parameter values.

Conclusions These findings suggest that SynX could make it possible to calibrate EMG-driven musculoskeletal models for all important lower-extremity muscles with as few as eight carefully chosen EMG signals and eventually contribute to the design of personalized rehabilitation and surgical interventions for mobility impairments.

Keywords EMG-driven model, Synergy extrapolation, Static optimization, Model personalization, Muscle force, Muscle activation, Stroke

*Correspondence:
Di Ao
aodi@nwpu.edu.cn
Full list of author information is available at the end of the article

Background

Muscle forces are essential for maintaining body posture and engaging in functional activities. A comprehensive understanding of the forces generated by individual muscles is crucial for understanding the internal biomechanical mechanisms and motor control involved in human movement [1–3]. More importantly, this knowledge holds significant value in identifying musculoskeletal pathologies [4, 5] and neurological disorders [6, 7] as well as for designing effective rehabilitation or surgical interventions [8–10]. However, unlike joint moments, which can be measured in vivo either directly using dynamometers or indirectly using inverse dynamics, muscle forces cannot currently be measured easily in vivo, though ongoing research is seeking to address this limitation [11, 12]. Unfortunately, these research efforts have been hindered by technical challenges, high cost, and ethical considerations [11, 12], motivating the need for computational methods utilizing a musculoskeletal model to advance our understanding of muscle force generation during movement.

The two computational methods most commonly employed for estimating muscle activations and forces using a musculoskeletal model are electromyography (EMG)-driven modeling [6, 13–19] and static optimization (SO) [20–26]. Both methods employ a geometric model of the musculoskeletal system actuated by Hill-type muscle–tendon models [27], both utilize nonlinear optimization to resolve the "muscle redundancy problem" [28] (i.e., many more muscles than degrees of freedom (DOFs) in the skeleton, resulting in control indeterminacy), both require experimental joint kinematics and moments as inputs, and both find muscle activations and forces such that predicted net joint moments from a musculoskeletal model match experimental net joint moments calculated via inverse dynamics as closely as possible. However, the optimization problem formulations for these two methods are quite different, as outlined in Table S1 of the supplementary materials.

Differences in optimization problem formulation stemming from the use or non-use of experimental muscle excitations as inputs have important implications for the capabilities and limitations of EMG-driven modeling and SO. Since EMG-driven modeling uses experimental muscle excitations to constrain the time-varying shapes (and often amplitudes) of the predicted muscle excitations, model joint moments never match experimental joint moments perfectly. Consequently, EMG-driven modeling allows for calibration of musculoskeletal model parameter values when the optimization is performed over all time frames together. In contrast, SO finds muscle activations that make model joint moments match experimental joint moments perfectly. Consequently, there are no joint moment errors that can be used for calibrating musculoskeletal model parameter values. Furthermore, optimization of each time frame separately can sometimes produce muscle activation discontinuities between time frames [29, 30], while minimization of muscle activations with no constraints on the time-varying shapes of the predicted muscle activations produces the smallest possible muscle activations, resulting in minimum co-contraction solutions [27, 31] that may not be physiologically realistic for some subjects or movement conditions. Nonetheless, because of its simplicity and the ease with which it can be performed, SO remains the most commonly used computational method for estimating muscle activations and forces.

Although EMG-driven modeling possesses the advantages noted above and produces physiologically reasonable estimates of muscle activations and forces [18], missing EMG data from muscles that contribute significantly to a measured movement has limited the adoption of EMG-driven modeling for biomechanical research and routine clinical gait analysis. This issue is the result of two practical challenges. First, surface electrodes are incapable of measuring EMG signals from important deep muscles that contribute significantly to joint moments. Common examples are the iliacus and psoas muscles, strong hip flexor muscles that contribute significantly to walking function. While fine wire electrodes can capture EMG signals from deep muscles, their invasive nature, the need for specialized insertion skills, the substantial preparation time required for insertion, and the potential for discomfort and pain to the subject have limited their utilization. Furthermore, in certain scenarios, deep muscles may be inaccessible even with fine wire electrodes. For instance, the use of a fine wire electrode is contraindicated for safety reasons in subjects with a cancerous tumor near the muscle to be measured. Second, EMG systems possess a limited number of channels for collecting EMG data. Many EMG systems available in human movement labs provide support for 16 channels of data, which means only eight channels of EMG data can be collected per leg when measuring activities such as walking or running. However, EMG-driven lower extremity models require roughly 16 channels per leg to inform the model without omitting any important large muscles. These challenges are significant as the absence of EMG data from important muscles can have a negative impact on the reliability of force estimates for other muscles that span the same joints [19, 23]. To address the issue of missing EMG signals when performing EMG-driven modeling, researchers either exclude from the musculoskeletal model muscles with missing EMG data [22, 31], include such muscles in the model but assume that they generate only passive force [19], or include such muscles

and use SO to estimate the associated muscle activations [22, 23].

To provide a better alternative for addressing missing EMG signals, researchers have recently developed a modified EMG-driven modeling approach called "Synergy Extrapolation" (SynX) that uses muscle synergy concepts to estimate missing muscle excitation data [32–34]. The theoretical basis for SynX is that a large number (e.g., 8 or 16) of experimentally measured muscle excitations can be represented by a smaller number (e.g., 4 or 5) of muscle synergies composed of time-varying synergy excitations and associated time-invariant synergy vectors, where the weights in each synergy vector define how the associated synergy excitation contributes to all muscle excitations. The synergy excitations provide information about the timing of muscle contractions, while the synergy vectors provide information about the coordination of muscle contractions.

Based on this observation, the historical development of SynX followed a logical sequence of three studies. First, SynX was shown to work in theory for fitting eight missing muscle excitations using synergy excitations extracted from eight measured muscle excitations [32]. For this study, 16 muscle excitations per leg measured experimentally from three subjects during walking were split into two groups of eight "measured" and eight "missing" excitations, and synergy excitations calculated from the eight measured excitations were used to fit the eight missing excitations. This study only established the theoretical feasibility of SynX, since the fitting process required the use of the missing muscle excitations. Second, SynX was shown to work in practice for predicting a single missing muscle excitation if a musculoskeletal model with pre-calibrated parameter values was used in the process [33]. The same sets of 16 experimental muscle excitations were again split into two groups, where 15 muscle excitations were treated as "measured" and one muscle excitation at a time collected from a fine wire electrode was treated as "missing." A key limitation of this study was the need for a pre-existing calibrated musculoskeletal model before the missing muscle excitation could be predicted reliably, which necessitates a priori knowledge of the missing muscle excitation for initial model calibration. Third, SynX was shown to work in practice for predicting a single missing muscle excitation while simultaneously calibrating musculoskeletal model parameter values [34]. A multi-objective optimization problem was designed to predict one missing muscle excitation while simultaneously calibrating time-invariant musculoskeletal model parameter values and time-varying residual muscle activations needed to account for small errors in the measured muscle excitations. This study resolved the main limitation of the previous study by allowing EMG-driven model calibration and prediction of a single missing muscle excitation to be performed simultaneously. SynX has been used more recently to predict the activation of a single unmeasured upper-extremity muscle (e.g. biceps long head), achieving a Pearson's correlation coefficient of up to 0.99 with the same muscle activation calculated from experimental EMG data withheld for evaluation purposes [35]. The next logical study in this progression is to evaluate how well SynX works in practice for predicting multiple missing muscle excitations while simultaneously calibrating musculoskeletal model parameter values. If SynX can predict missing muscle excitations reliably using a low number of EMG signals collected using only surface electrodes, the applicability of EMG-driven modeling to research and clinical questions will be greatly expanded.

This study evaluated how well SynX can estimate muscle activations associated with eight channels of missing EMG data using synergy excitations associated with eight channels of measured EMG data while simultaneously calibrating musculoskeletal model parameter values. Experimental walking data collected from two subjects post-stroke were used for the evaluation. Time-varying quantities (muscle activations and forces along with net joint moments) and time-invariant model parameter values (activation dynamics and Hill-type muscle–tendon model parameter values) predicted by SynX were compared to "gold standard" results produced by EMG-driven model calibration using a complete set of EMG data where no EMG signals were regarded as missing. Time-varying quantities (muscle activations and forces) predicted by SO using a scaled generic musculoskeletal model were also compared to the "gold standard" results to determine which method provides the most reliable predictions. In addition, the reliability with which SynX and SO can predict muscle activations and forces when using pre-calibrated musculoskeletal models was evaluated to assess how model calibration affects muscle activation and force estimates from both methods.

Methods

Experimental data collection

Two previously published experimental walking datasets collected from hemiparetic subjects post-stroke were used for this study [17, 36]. One subject was high-functioning (S1, male, 1.70 m tall, mass 80.5 kg, age 79 years, right side hemiparesis, lower extremity Fugl-Meyer Motor Assessment score of 32 out of 34), while the other subject was low-functioning (S2, male, 1.83 m tall, mass 88.5 kg, age 62 years, right side hemiparesis, lower extremity Fugl-Meyer Motor Assessment score of 25 out of 34). After giving written informed consent, both subjects walked on a split-belt instrumented

treadmill (Bertec Corp., Columbus, OH, United States) at their self-selected speed (0.5 m/s for S1 and 0.35 m/s for S2) and fastest-comfortable speed (0.8 m/s for S1 and 0.65 m/s for S2) for one minute. All experimental procedures were approved by the University of Florida Health Science Center Institutional Review Board (IRB-01).

Sixteen channels of EMG data were collected from each leg of both subjects using both surface and fine wire electrodes (Motion Lab Systems, Baton Rouge, LA, United States). These extensive EMG data enabled every muscle in each leg of each subject's musculoskeletal model (see below) to have an associated experimental EMG signal, providing an opportunity to verify the reliability of muscle activations and forces estimated by SynX and SO. Measured EMG signals were expanded to muscles with similar anatomical function (e.g., semimembranosus and semitendinosus) before being used as inputs to the EMG-driven modeling process (Fig. 1 and Supplementary Table S2). Surface EMG data were collected and expanded to the subsequent superficial muscle groups: (1) *GlutMax*, expanded to gluteus maximus superior (*glmax1*), gluteus maximus middle (*glmax2*) and gluteus maximus inferior (*glmax3*); (2) *GlutMedMin*, expanded to gluteus medius anterior (*glmed1*), gluteus medius middle (*glmed2*), gluteus medius posterior (*glmed3*), gluteus minimus anterior (*glmin1*), gluteus minimus middle (*glmin2*), and gluteus minimus posterior (*glmin3*); (3) *SemiMembTen*, expanded to semimembranosus (*semimem*) and semitendinosus(*semiten*); (4) *RecFem*, used for rectus femoris (*recfem*); (5) *Bicfem*, expanded to biceps femoris long head (*bflh*) and biceps femoris short head (*bfsh*); (6) *VasMedInt*, expanded to vastus medialis (*vasmed*) and vastus intermedius (*vasint*); (7) *VasLat,* used for vastus lateralis (*vaslat*); (8) TibAnt, used for tibialis anterior (*tibant*);(9) *Peroneus*, expanded to peroneus brevis (*perbrev*) and peroneus long (*perlong*); (10) *Sol*, used for soleus (*soleus*). Additionally, fine-wire EMG data were collected and expanded to the following deep muscle groups: (1) *iliopsoas*, expanded to iliacus (*iliacus*) and psoas (*psoas*); (2) *Adductors*, expanded to adductor brevis (*addbrev*), adductor longus (*addlong*), adductor magnus distal (*addmagDist*), adductor magnus ischial (*addmagIsch*), adductor magnus middle (*addmagMid*), and adductor magnus proximal (*addmagProx*); (3) *Tibpost*, used for tibialis posterior (*tibpost*).

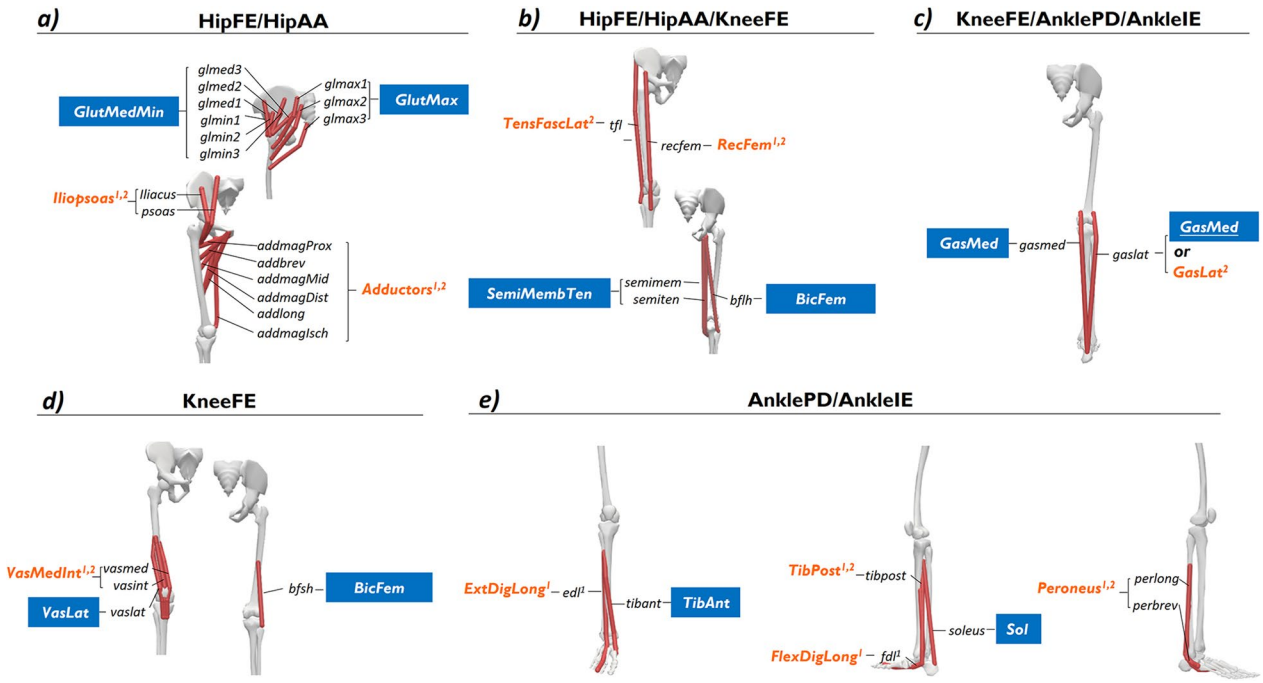

Fig. 1 Overview of experimental EMG channels treated as "measured" (blue boxes) and "unmeasured" (orange italicized text) when performing SynX and SO, along with associated muscles in the OpenSim model for each subject (black italicized text). A total of 16 channels of EMG data were collected from each leg. Each EMG signal was expanded to muscles in the OpenSim model with similar anatomical function before being used as inputs to the EMG-driven modeling process. Superscripts 1 and 2 indicate "unmeasured" EMG signals for subject S1 and S2, respectively. Muscles were categorized based on their actuated degrees of freedom (DOFs), which included: hip flexion/extension (HipFE), hip adduction/abduction (HipAA), hip internal/external rotation (HipRot), knee flexion/extension (KneeFE), ankle plantarflexion/dorsiflexion (AnklePD), and ankle inversion/eversion (AnkleIE)

Small differences existed in the EMG data collected from the two subjects. For the high-functioning subject (S1), only a single surface EMG signal (referred to as *GasMed)* was collected for both heads of gastrocnemius and expanded to medial gastrocnemius (*gasmed*) and lateral gastrocnemius (*gaslat*), and two fine-wire EMG signals (referred to as *ExtDigLong* and *FlexDigLong*) were recorded from extensor digitorum longus (*edl*) and flexor digitorum longus (*fdl*), respectively. For the low-functioning subject (S2), two surface EMG signals (referred as *GasMed* and *GasLat*) were recorded from medial gastrocnemius (*gasmed*) and lateral gastrocnemius (*gaslat*), respectively, and a fine-wire EMG signal (referred as *TensFascLat*) was recorded from tensor fasciae latae (*tfl*). Raw EMG data were collected at 1000 Hz and high-pass filtered at 40 Hz, demeaned, full-wave rectified, and low-pass filtered at 3.5/*tf* Hz, where where *tf* is the period of the gait cycle [17]. Processed EMG data were then normalized to the maximum values across all experimental gait cycles. The resulting processed EMG data will henceforth be referred to as "experimental muscle excitations" [17, 37].

A three-dimensional motion capture system (Vicon Corp., Oxford, United Kingdom) operating at 100 Hz was used to measure reflective surface marker trajectories, while two treadmill force plates (Bertec Corp., Columbus, OH, United States) recording at 1,000 Hz were used to measure ground reaction forces and moments. Raw motion capture and ground reaction data were low-pass filtered with a variable cut-off frequency of 7/*tf* Hz [38], where *tf* is the period of the gait cycle. Data from ten gait cycles (five cycles per speed) per leg were randomly chosen to calibrate the EMG-driven models and evaluate the accuracy of estimated muscle activations and forces. Following pre-processing, data from each gait cycle were resampled to 101 time points from heel-strike (0%) to subsequent heel-strike (100%) of the same foot. An extra 20 time frames, accounting for a maximum electromechanical delay of approximately 100 ms, were retained prior to the start of each gait cycle, yielding 121 time points for each of the 10 gait cycles.

Musculoskeletal model creation

A generic full-body OpenSim musculoskeletal model [39] was used as the starting point to create a personalized model of each subject. This generic model possessed 37 degrees of freedom (DOFs), 80 muscle–tendon actuators to control lower limb motion, and 17 ideal torque actuators to control the upper body motion. For each subject, a sequence of four analyses was performed using OpenSim 4.0 [40, 41] to prepare the model for EMG-driven modeling with SynX. First, OpenSim model scaling was performed using motion capture data collected while the subject stood in a static pose, allowing the generic model's anthropometry to match that of each subject more closely. Second, repeated OpenSim inverse kinematics (IK) analyses were performed within a nonlinear optimization to calibrate the locations and orientations of lower body joint centers and axes such that errors between model and experimental surface marker positions were minimized for isolated joint motion and walking trials [42]. The lower body DOFs affected by this calibration process were hip flexion/extension (*HipFE*), hip adduction/abduction (*HipAA*), hip internal/external rotation (*HipRot*), knee flexion/extension (*KneeFE*), ankle plantarflexion/dorsiflexion (*AnklePD*), and ankle inversion/eversion (AnkleIE). These six low-extremity DOFs were targeted because their associated experimental joint moments were needed for performing SynX and SO. Third, additional OpenSim IK analyses were performed using experimental marker data from the walking trials to obtain joint angle time histories. Fourth, OpenSim inverse dynamic (ID) analyses were performed using the previously calculated joint kinematics and the experimental ground reaction data from the walking trials to calculate experimental joint moments.

Muscle activation and force estimation

As illustrated in Fig. 2, both SynX and SO follow a four-phase process to generate accurate estimations of muscle activations and forces, which are outlined as follows: (1) In the case of SynX, the control inputs to the muscle–tendon models consist of muscle excitations, which involve processed experimental EMG data for "measured" muscles and SynX-estimated muscle excitations for unmeasured muscles. Muscle excitations are typically passed through an activation dynamics model to determine muscle activations. In the case of SO, the muscle–tendon models are directly controlled by muscle activations. Detailed information can be accessed in the sections titled "*Muscle activation estimation*" for SynX and "*Static Optimization Solution Process*" for SO; (2) Taking muscle activations as inputs, both approaches estimate muscle forces through a geometric musculoskeletal model driven by Hill-type muscle–tendon models. The physical musculoskeletal models provide muscle–tendon lengths based on joint kinematics, while the Hill-type models describe the physiology of muscle force generation. Additional insights can be found in the sections titled "*Muscle force estimation*" for SynX and "*Static Optimization Solution Process*" for SO; (3) Subsequently, the net joint moments are computed by aggregating the force contributions from all muscles across the specific DOF, which involves multiplying muscle forces by corresponding moment arms derived from the physical models based on joint angle histories. Detailed explanations are available

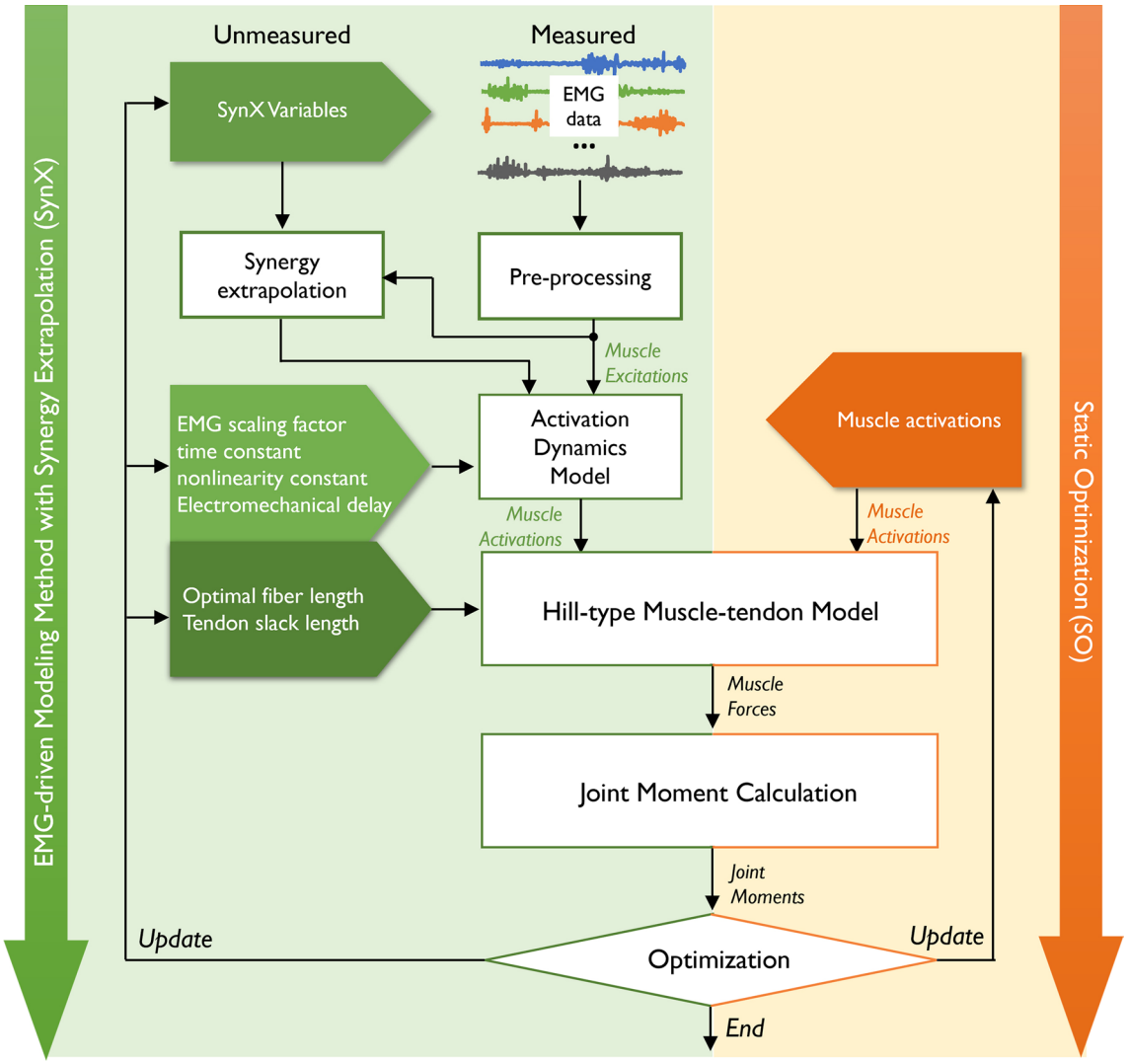

Fig. 2 Workflow for EMG-driven modeling with SynX (left panel with green background) and SO (right panel with orange background) as performed in this study. Both methods use experimental joint kinematics and moments as inputs and calculate muscle activations and forces such that predicted net joint moments from the musculoskeletal model closely match experimental net joint moments from inverse dynamics. However, notable differences exist in the optimization problem formulations for these two methods. For EMG-driven modeling with SynX, the design variables were time-invariant model parameter values and SynX variables, with the optimization problem being solved across all time frames together. In contrast, for SO, the design variables were time-varying muscle activations, typically utilizing model parameter values from scaled generic models or literature references, with the optimization problem being solved for each time frame separately. Muscle activations found by both approaches were used to estimate muscle forces and their respective contributions to net joint moments

in the sections titled "*Joint moment calculation*" for SynX and "*Static Optimization Solution Process*" for SO; (4) Last but not least, the estimated net joint moments are iteratively compared to the inverse dynamic joint moments through an optimization process. This iterative comparison leads to the estimation of a time-invariant muscle activation model, muscle–tendon model, and SynX-specific parameter values for SynX, as well as time-varying muscle activations for SO. Further details are provided in the sections titled "*EMG-driven model calibration with SynX*" for SynX and "*Static Optimization Solution Process*" for SO.

Synergy extrapolation solution process

SynX constructs unmeasured muscle excitations by multiplying time-varying muscle synergy excitations extracted from "measured" EMG signals by time-invariant synergy vector weights associated with "unmeasured" EMG signals. SynX was integrated into the framework of EMG-driven model calibration process, enabling the "unmeasured"

synergy vector weights to be determined as optimization design variables. This complex process involves using joint kinematics and associated musculoskeletal geometries (such as muscle–tendon lengths and moment arms) as inputs to sequentially estimate muscle activations, forces, and net joint moments. The estimated predicted net joint moments are iteratively compared to the inverse dynamic joint moments through an optimization process, resulting in the estimation of the time-varying muscle activations and forces, along with time-invariant musculoskeletal model and SynX-related parameter values. The SynX solution process involved four sequential steps as summarized below.

Step 1: Muscle activation estimation

For the first task of the SynX solution process, muscle activations were found for muscles with and without experimental EMG data. The transformation of "measured" muscle excitations into activations of all muscles served as the core of the SynX solution process [33, 34]. First, muscle excitations $e_m^{musc}(t)$ for each muscle with experimental EMG data were scaled using a muscle-specific scale factor ranging from 0.05 to 1. These EMG scaling factors were strategically incorporated into the muscle activation model calibration process to ensure that the magnitude of muscle excitations is appropriately normalized. This normalization is crucial for accurate muscle force estimations, as required by the Hill-type model in Eq. (6), which specifies that muscle activations must fall within the predefined range of 0 to 1. Acknowledging that actual maximum activation levels tend to surpass those observed experimentally during walking and recognizing that peak EMGs, even if identified from maximal voluntary contraction (MVC) trials, may not entirely capture the complete extent of muscle activation due to inherent limitations in capturing spatial variability [43], we chose to allow the optimization process to iteratively adjust the scaling factors, and this data-driven approach enabled us to determine the normalized muscle activations, unhindered by the constraints of empirical MVC-based normalization methods. Second, to center the data and eliminate any bias due to varying means in the original muscle excitations, the mean of each scaled measured muscle excitation was subtracted from the original dataset before conducting muscle synergy analysis (MSA) using principal component analysis (PCA) to extract a small number of muscle synergies, specifically five for the present study:

$$e_m^{musc}(t) = W_m(t)H_m + \mu_m + \varepsilon_m(t) \tag{1}$$

where $W_m(t)$ specifies the time-varying measured synergy excitations, H_m specifies the associated time-invariant measured synergy vector weights, μ_m stands for the average values of each measured muscle excitation, and $\varepsilon_m(t)$ stands for the decomposition residuals that could not be accounted for by $W_m(t)H_m + \mu_m$. Following MSA, both unmeasured muscle excitations $e_{SynX}^{musc}(t)$ and residual muscle excitations $e^{res}(t)$ added to the measured muscle excitations were constructed from the measured synergy excitations:

$$\begin{cases} e_{SynX}^{musc}(t) = W_m(t)H_{SynX} + \mu_{SynX} \\ e^{res}(t) = W_m(t)H_{res} + \mu_{res} \end{cases} \tag{2}$$

where H_{SynX} represents the unmeasured synergy vector weights,μ_{SynX} represents the average values of each unmeasured muscle excitation, H_{res} represents the residual synergy vector weights, and μ_{res} represents the average values of each residual muscle excitation. Henceforth, we denote the union of H_{SynX}, μ_{SynX}, H_{res} and μ_{res} as SynX design variables, which were all time-invariant and determined through an optimization process implemented within the EMG-driven model calibration process (Fig. 2). Once unmeasured and residual muscle excitations were constructed, two sets of muscle excitations were calculated when residual muscle excitations were and were not included:

$$\begin{cases} e^{musc}(t) = \left\{ e_m^{musc}(t), e_{SynX}^{musc}(t) \right\} \\ e_{res}^{musc}(t) = \left\{ e_m^{musc}(t) + e^{res}(t), e_{SynX}^{musc}(t) \right\} \end{cases} \tag{3}$$

where $e^{musc}(t)$ defines the muscle excitations without residual muscle excitations included, while $e_{res}^{musc}(t)$ defines the muscle excitations with residual muscle excitations included. Both $e^{musc}(t)$ and $e_{res}^{musc}(t)$ were utilized in subsequent steps to compute corresponding muscle activations denoted as $a^{musc}(t)$ and $a_{res}^{musc}(t)$, respectively. Third, neural activations $u^{musc}(t)$ were determined from constructed muscle excitations by employing a first-order ordinary differential equation for activation dynamics [44]:

$$\begin{aligned} \frac{du^{musc}(t)}{dt} &= (c_1 e^{musc}(t-d) + c_2)(e(t-d) - u^{musc}(t)) \\ c_1 &= 1/\tau_{act} - 1/\tau_{dact} \\ c_2 &= 1/\tau_{dact} \\ \tau_{dact} &= 4\tau_{act} \end{aligned} \tag{4}$$

where τ_{act} and τ_{dact} are activation and deactivation time constants, and d specifies the electromechanical time delay. Fourth, a nonlinear one-parameter transformation model was utilized to compute each associated muscle activation $a^{musc}(t)$ [45]:

$$a^{musc}(t) = (1 - c_3)u^{musc}(t) + c_3\left[\frac{g_1}{g_2(u^{musc}(t) + g_3)^{g_4} + g_5} + 1\right] \quad (5)$$

where c_3 is an activation nonlinearity constant that characterizes the curvature of the relationship of each muscle, and g_1 to g_5 are constant coefficients obtained by fitting published experimental data from isometric contractions [45]. Our EMG-driven modeling approach solves for muscle activations with (i.e., $a_{res}^{musc}(t)$) or without (i.e., $a^{musc}(t)$) residual excitations included over all time frames simultaneously by adjusting of the following design variables: SynX design variables, EMG scale factors, electromechanical time delays, activation time constants, and activation nonlinearity constants. Further details are provided in the section entitled "EMG-driven model calibration with SynX".

Step 2: Muscle force estimation

For the second step of the SynX solution process, muscle forces were estimated using the activations for measured and unmeasured muscles found in the first task. Taking the estimated muscle activations as inputs, our EMG-driven modeling process employed a Hill-type muscle tendon model with rigid tendon [17, 27, 46] to predict the force generated by a given muscle–tendon actuator m, which was formulated as (Fig. 2):

$$\begin{aligned} F^{musc}(t,\theta,\dot{\theta}) &= F_o^m \cdot \left[a^{musc}(t) \cdot f_l^a(\tilde{l}^{musc}(t,\theta)) \cdot f_v^a(\tilde{v}^{musc}(t,\theta,\dot{\theta})) + f_l^p(\tilde{l}^{musc}(t,\theta))\right] \cdot \cos\alpha \\ \tilde{l}^{musc}(t,\theta) &= \frac{l^{mt}(t,\theta) - l_s^t}{l_o^m} \\ \tilde{v}^{musc}(t,\theta,\dot{\theta}) &= \frac{v^{mt}(t,\theta,\dot{\theta})}{10 \cdot l_o^m} \end{aligned} \quad (6)$$

where $f_l^a(\tilde{l}^{musc}(t,\theta))$ and $f_v^a(\tilde{v}^{musc}(t,\theta,\dot{\theta}))$ describe the normalized active muscle force–length and force–velocity relationships, respectively, $f_l^p(\tilde{l}^{musc}(t,\theta))$ defines the normalized passive muscle force–length relationship, $\tilde{l}^{musc}(t,\theta)$ and $\tilde{v}^{musc}(t,\theta,\dot{\theta})$ denote the time-varying normalized muscle fiber length and velocity, respectively, $F^{musc}(t,\theta,\dot{\theta})$ and $a^{musc}(t)$ denote the muscle force and muscle activation generated by the muscle–tendon actuator at time t, F_o^m is the maximum isometric force, α is the pentation angle of the muscle (values of which were taken from literature [47]), l_o^m denotes optimal muscle fiber length, and l_s^t denotes tendon slack length. These values (apart from pennation angles) were calibrated through an optimization process. More details regarding the determination of l_o^m and l_s^t values for each muscle force estimation method can be found in the section entitled "EMG-driven model calibration with SynX".

Step 3: Joint moment calculation

For the third step of the SynX solution process, model net joint moments were calculated using the forces for measured and unmeasured muscles found in the second task. Once the muscle forces $F^{musc}(t,\theta)$ were estimated, their contributions to the net joint moment at each joint j were calculated using the corresponding muscle moment arms:

$$M^{joint}(t,\theta,\dot{\theta}) = \sum F^{musc}(t,\theta,\dot{\theta}) \cdot r^{musc}(t,\theta) \quad (7)$$

$$r^{musc}(t,\theta) = -\frac{\partial l^{mt}(t,\theta)}{\partial\theta} \quad (8)$$

where $M^{joint}(t,\theta,\dot{\theta})$ is net joint moment at time t produced by all spanning muscles spanning the joint, and $r^{musc}(t,\theta)$ is the muscle moment arm for muscle m at time t, which was defined as the negative of the partial derivative of muscle–tendon length $l^{mt}(t,\theta)$ with respect to generalized coordinate θ [48]. The negative sign in Eq. 8 was implemented for consistency with the OpenSim modeling environment. When utilizing SynX for estimating unmeasured muscle excitations, net joint moments were computed with ($M_{res}^{joint}(t,\theta,\dot{\theta})$) and without ($M^{joint}(t,\theta,\dot{\theta})$) inclusion of residual excitations in the measured muscle excitations, as stipulated by the cost function for EMG-driven model calibration.

Step 4: EMG-driven model calibration with SynX

For the fourth step of the SynX solution process, the first three tasks were performed iteratively within a nonlinear optimization that adjusted three categories of design variables (see Fig. 2): (1) SynX parameter values including synergy vector weights and average values associated with unmeasured muscle excitations as well as synergy vector weights and average values associated with residual muscle excitations; (2) activation dynamics model parameter values consisting of EMG scale factors, electromechanical delays, activation time constants, and activation nonlinearity constants; and (3) muscle–tendon model parameter values consisting of optimal muscle fiber lengths and tendon slack lengths. EMG-driven model calibration typically adjusts muscle forces by altering muscle–tendon model parameter values such that the differences

between model-predicted and inverse dynamic (ID) joint moments are minimized. However, to estimate unmeasured muscle excitations via SynX during EMG-driven model calibration, the primary cost function was formulated as a trade-off between minimizing net joint moment tracking errors and minimizing unmeasured and residual muscle activation magnitudes [34]:

$$\begin{aligned}\min J \triangleq & \sum\left(\frac{M_{res}^{joint}(t,\theta,\dot{\theta})-M^{ID}(t,\theta,\dot{\theta})}{MAD_1}\right)^2 \\ & +\sum\left(\frac{M^{joint}(t,\theta,\dot{\theta})-M^{ID}(t,\theta,\dot{\theta})}{MAD_2}\right)^2 \\ & +\sum\left(\frac{a_{SynX}^{musc}(t)}{MAD_3}\right)^2 \\ & +\sum\left(\frac{a_{res}(t)}{MAD_4}\right)^2\end{aligned} \tag{9}$$

where $M_{res}^{joint}(t,\theta,\dot{\theta})$ refers to model-predicted joint moments when residual muscle excitations are included in net joint moment calculations, $M^{joint}(t,\theta,\dot{\theta})$ refers to model-predicted joint moments when residual muscle excitations are excluded in net joint moment calculations, $M^{ID}(t,\theta,\dot{\theta})$ refers to inverse dynamic net joint moments obtained from OpenSim ID analyses, $a_{SynX}^{musc}(t)$ represents unmeasured muscle activations estimated by SynX, and $a_{res}(t)$ signifies residual muscle activations added to the measured muscle activations, which are equivalent to $a_{res}^{musc}(t)-a^{musc}(t)$. Normalization of all four cost function terms was achieved using a set of maximum allowable deviations (MAD), the values of which were determined by performing a sensitivity analysis as described in [34]. Further details regarding initial guesses, upper/lower bounds for design variables, additional inequality constraints, and penalty terms can be found in previously published studies [17, 33, 34]. All optimization procedures were performed using MATLAB's "*fmincon*" nonlinear optimization function with its sequential quadratic programming algorithm.

Static optimization solution process

The static optimization solution process involved determining muscle activations $a^{musc}(t)$ at each time instant t by performing an inverse dynamics-based optimization. In the standard SO approach, the muscle redundancy problem is addressed by minimizing the energetic cost represented by the sum of squares of muscle activations while ensuring that inverse dynamic joint moments are matched perfectly at the solution [20]:

$$\begin{aligned}&\textit{for}\ \text{time frame}\ t: \\ &\min J=\sum a^{musc}(t)^2 \\ &\textit{subject to} \\ &M^{joint}(t,\theta,\dot{\theta})-M^{ID}(t,\theta,\dot{\theta})=0 \\ &0\le a^{musc}(t)\le 1\end{aligned} \tag{10}$$

The net joint moments for SO were calculated by substituting the estimated muscle activations $a^{musc}(t)$ into the Hill-type muscle–tendon model and multiplying the resulting muscle forces by their corresponding moment arms, as depicted in Eqs. 6 through 8. In contrast to the EMG-driven modeling method, the muscle activations estimated for SO were used directly as design variables in the optimizations, which were solved individually for each time frame. Furthermore, model parameter values were taken from the scaled generic OpenSim model rather than being calibrated during the optimization process.

Synergy extrapolation and static optimization evaluation

Muscle selection heuristics

Given 16 measured muscle excitations for each leg of both subjects, we had to choose 8 muscle excitations to be treated as measured and 8 to be held back and treated as missing for SynX and SO evaluation purposes. A prior study [32] provided guidance for which eight muscles to select as measured and which eight to select as missing so as to maximize reconstruction accuracy for the eight missing muscle excitations. In that study, an investigation of all possible combinations of eight measured and eight missing EMG signals yielded the following muscle selection heuristic: (1) Choose muscles easily accessible by surface EMG electrodes; (2) Choose most frequently occurring muscle in the top 10% of muscle combinations that yielded the highest SynX accuracy from each primary lower extremity function group; (3) Choose two hip/knee biarticular muscles at minimum; (4) Choose remaining most frequent muscles to fill eight muscle combinations. Following this muscle selection heuristic, given a limited number of eight EMG channels, indicated that researchers should collect surface EMG data from commonly chosen uniarticular and biarticular flexor and extensor muscles from each major muscle group, as illustrated in Fig. 1. The selected uniarticular muscles included a hip extensor (*GlutMax*), a knee extensor (*VasLat* considered preferable over *VasMed*), an ankle plantarflexor (*Sol*), and an ankle dorsiflexor (*TibAnt*). Uniarticular hip flexors (*Iliacus and Psoas*) were omitted due to the difficulty in measuring these muscles with surface electrodes. The chosen biarticular muscles included a posterior thigh muscle (*SemiMembTen*, or *Bicfem*) and a posterior calf muscle (*GasMed* or *GasLat*).

Additionally, adding *GlutMedMin* to the list appeared to be a reasonable choice if one more muscle was needed.

Synergy extrapolation methodological choices

Implementation of SynX requires making several methodological choices that can impact the accuracy of estimated muscle activations and forces. Using the same two experimental data sets, previous studies investigated the influence of various methodological choices on SynX performance [33, 34], including EMG normalization methods, matrix decomposition algorithms, the number of muscle synergies, and assumptions regarding the variability of synergy vector weights across trials for the reconstruction of unmeasured and residual muscle excitations. Those studies systematically assessed the results for all possible methodological combinations and found that principal component analysis (PCA) with either five or six synergies consistently predicted unmeasured muscle excitations with reasonable accuracy. In contrast, non-negative matrix factorization (NMF) did not achieve acceptable prediction accuracy. Additionally, for the same number of synergies, employing trial-specific unmeasured synergy vector weights and speed-specific residual synergy vector weights resulted in optimal SynX performance for both subjects in terms of estimation accuracy and computational efficiency. Notably, EMG normalization had no significant impact on SynX performance. Thus, the key methodological choices for SynX were informed by insights from prior studies, as detailed in Table 1.

Optimization problems

The present study had three primary objectives that influenced the optimization problems that were formulated and solved. First, the study aimed to evaluate the performance of SynX when treating multiple channels of EMG data (i.e., eight) as "unmeasured." Second, the study sought to compare estimates of muscle activations and forces from SynX and SO with those from a "gold standard" reference. Third, the study aimed to analyze the accuracy of estimated unmeasured muscle activations and forces from both SynX and SO when using model parameter values associated with different levels of personalization.

To address these primary objectives, we formulated six optimization problems to estimate unmeasured muscle activations and, for SynX, to calibrate model parameter values (Fig. 3). The first optimization problem (termed "*Params*") utilized all 16 channels of EMG data to calibrate each EMG-driven musculoskeletal model, providing "gold standard" muscle activations and forces for evaluation. The second optimization problem (termed "$SynX_{Unmeasured} + Params$") assessed the performance of SynX when multiple channels of EMG data (i.e., eight) were considered "unmeasured." This optimization problem calibrated EMG-driven models (including activation dynamics model, muscle–tendon model, and SynX variable values) for each leg of each subject while simultaneously estimating missing muscle excitations using SynX. The third optimization problem (termed "$SynX^{Params}_{Unmeasured}$") employed SynX to estimate the unmeasured muscle excitations within a well-calibrated EMG-driven model utilizing the model parameter values found in the "gold standard (*Params)*" optimization. The fourth optimization problem (termed "$SO^{Generic}_{All}$") used SO to estimate muscle activations for all muscles using muscle–tendon model parameter values taken from scaled generic OpenSim models, representing the most commonly formulated SO method. The fifth and sixth optimization problems (termed "SO^{Params}_{All}" and "$SO^{Params}_{Unmeasured}$") utilized SO to estimate muscle activations for all muscles and only unmeasured muscles, respectively, using model parameter values from the "gold standard" (*Params)* optimization. When performing the fourth and sixth optimizations $SynX^{Params}_{Unmeasured}$ and $SO^{Params}_{Unmeasured}$ to estimate only unmeasured muscle excitations/activations, the muscle activations of the measured muscles were taken from the "gold standard" (*Params)* optimization.

Table 1 Methodological choices for synergy extrapolation

Description	Methods (Abbreviations)
Matrix factorization algorithm	Principal component analysis (PCA)
EMG normalization method	Maximum value over all trial
Number of muscle synergies	5
Category of unmeasured synergy vector weights	Trial-specific
Category of residual synergy vector weights	Speed-specific
Number of missing EMGs	8
Number of measured EMGs	8

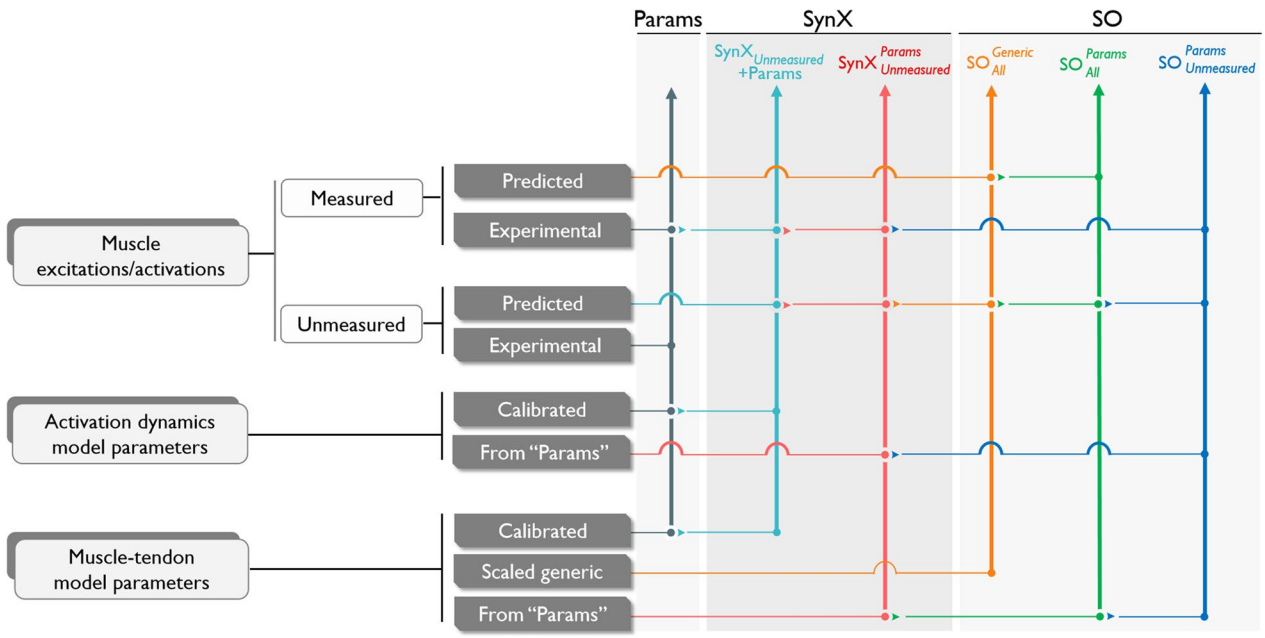

Fig. 3 Summary of six optimization problems investigated in this study. Two optimizations used SynX to predict unmeasured muscle excitations (*termed* $SynX_{Unmeasured} + Params$ and $SynX^{Params}_{Unmeasured}$), three optimizations used static optimization (SO) to predict unmeasured muscle activations (termed $SO^{Generic}_{All}$, SO^{Params}_{All} and $SO^{Params}_{Unmeasured}$), and one "gold standard" optimization used the complete set of EMG signals with no muscle excitations predicted by SynX or SO (termed *Params*). The calibration cases were named based on the prediction method for unmeasured muscle excitations or activations as well as the categories of design variables included in the optimization problem formulation. The subscripts indicate which set of muscle excitations or activations were predicted computationally, while the superscripts indicate which set of model parameters were employed for calculating muscle activations and forces. In each column of the optimizations, the arrows indicate whether each group of muscle excitations or activations were predicted or obtained experimentally as well as which values were used if model parameters were not calibrated through optimization. The term "Scaled Generic" denotes scaled generic model parameter values, while "From Params" refers to the model parameter values derived from the "gold standard" (*Params*) optimization

Evaluation metrics and statistical analyses

Several common evaluation metrics were used to evaluate the ability of SynX and SO to estimate muscle activations and forces for unmeasured muscles as well as net joint moments across all cases. First, root mean square errors (RMSEs) were computed to quantify magnitude errors between experimental (from "*Params*" case) and predicted (from two SynX and three SO cases) muscle activations and forces. Similarly, Pearson correlation coefficients (r) were computed to quantify shape similarity between experimental and predicted unmeasured muscle activations and forces. Correlations were categorized as weak ($r < 0.35$), moderate ($0.35 < r \leq 0.67$), strong ($0.67 < r \leq 0.9$), or very strong ($r \leq 0.9$) [49]. Furthermore, mean absolute errors (MAEs) between model and experimental net joint moments were also calculated for the "*Params*" case and the two SynX cases "$SynX_{Unmeasured} + Params$" and "$SynX^{Params}_{Unmeasured}$." Evaluation metrics, including RMSEs, r values, and MAE values, were calculated by concatenating the data across all calibration trials and legs of both subjects.

Multiple statistical analyses were performed to compare the evaluation metrics resulting from different SynX and SO cases. Paired t-tests were performed on RMSE and r values to identify significant differences in the accuracy of estimated unmeasured muscle activations between any two of the five SynX or SO cases. Paired t-tests were also used to identify significant differences in the accuracy of estimated muscle forces between any two of the five SynX or SO cases. In addition, paired t-tests were performed to compare joint moment matching errors (MAE values) between the "*Params*" case and the two SynX cases. All statistical analyses were performed in MATLAB with a significance level of $p < 0.05$.

Results

Muscle activations

Muscle activations for unmeasured muscles estimated using SynX and SO were compared with those produced by EMG-driven model calibration using a complete set of EMG data ("*Params*") (Figs. 4 and 5, Table 2). When SynX was used with simultaneous calibration of EMG-driven model parameter

values ("$SynX_{Unmeasured} + Params$"), activations for unmeasured muscles were estimated with low RMSE values (0.08 ± 0.06, maximum value ≤ 0.17) and moderate to strong correlation r values (0.55 ± 0.13, minimum value ≥ 0.38) across most muscles compared to "gold standard" muscle activations ("Params"). Among unmeasured muscles, SynX exhibited superior performance for superficial muscles (e.g. rectus femoris, lateral gastrocnemius and vastus intermedius) compared to deep muscles (e.g. iliacus, extensor digitorum longus and tibialis activations for hip adductors (RMSE = 0.01, $r \geq 0.43$) and flexor digitorum longus (RMSE = 0.05, $r = 0.92$), which are deep muscles, maintained comparable accuracy to that of superficial muscles. SynX with simultaneous calibration of EMG-driven model parameter values ("$SynX_{Unmeasured} + Params$") produced significantly more accurate predictions of unmeasured muscle activations compared to standard SO with scaled generic parameter values ("$SO_{All}^{Generic}$"). This finding was evident in terms of both magnitude (characterized by RMSE values, $p \leq 0.05$) and shape (characterized by correlation r values, $p \leq 0.05$) across unmeasured muscles and subjects (Figs. 4 and 5, Table 2). Even for muscles with relatively low estimation accuracy using both methods, such as iliacus, psoas, and extensor digitorum longus, SynX outperformed standard SO in reproducing the shape and magnitude of unmeasured muscle activations (Table 2). Moreover, SO ("$SO_{All}^{Generic}$") exhibited weak correlations ($r \leq 0.35$) in the muscle activation predictions for the majority of unmeasured muscles, apart from tibialis posterior ($r = 0.53$), extensor digitorum longus ($r = 0.48$), and flexor digitorum longus ($r = 0.89$). Notably, the SynX-based optimization generated smooth muscle activation profiles for all unmeasured and measured muscles, whereas SO generated muscle activation profiles with discontinuities and that generally underestimated the "gold standard" muscle activations (see Fig. 4).

SynX-based and SO-based methods were both sensitive to the level of musculoskeletal model personalization (Figs. 4 and 5, Table 2). For SynX, using a well-calibrated EMG-driven model ("$SynX_{Unmeasured}^{Params}$") yielded lower RMSE values for unmeasured muscle activations (0.05 ± 0.05) compared to when SynX variables and EMG-driven model parameter values were calibrated simultaneously ("$SynX_{Unmeasured} + Params$") ($0.08 \pm 0.06$). Unmeasured muscle activations produced by "$SynX_{Unmeasured}^{Params}$" exhibited strong or very strong correlations with those generated from by "*Params*" with the exception of the extensor digitorum longus ($r = 0.42$). For SO, well-calibrated model parameter values in "SO_{All}^{Params}" led to more accurate estimation of unmeasured muscle activations compared to using scaled generic model parameter values in "$SO_{All}^{Generic}$", although the difference was not substantial.

When a well-calibrated EMG-driven model was used to estimate only unmeasured muscle activations, SynX case "$SynX_{Unmeasured}^{Params}$" produced more accurate and reliable estimates compared to SO case "$SO_{Unmeasured}^{Params}$", as evidenced by lower RMSE values ($p \leq 0.05$) and higher correlation r values ($p \leq 0.05$) (Figs. 4 and 5, Table 2). Similar to all SO-estimated muscle activations, muscle activations estimated by SO case "$SO_{Unmeasured}^{Params}$"generally underestimated the "gold standard" muscle activations while demonstrating abrupt changes.

Muscle forces

In addition, muscle forces for unmeasured muscles estimated using SynX and SO were compared with those produced by EMG-driven model calibration using a complete set of EMG data ("*Params*") (Figs. 6 and 7, Table 3). When SynX was used with simultaneous calibration of EMG-driven model parameter values ("$SynX_{Unmeasured} + Params$"), forces for unmeasured muscleswere estimated with low RMSE values (101.3 ± 90.1) and moderate to strong correlation r values (0.53 ± 0.17) across most muscles compared to "gold standard" muscle forces ("*Params*") (Fig. 3 and Table 3). Among unmeasured muscles, SynX exhibited superior performance for vastus intermedius, vastus medialis, flexor digitorum longus, lateral gastrocnemius and a

(See figure on next page.)

Fig. 4 Average muscle activations for the "unmeasured" muscles (upper) and "measured" muscles (lower) across calibration trials, legs and subjects from "*Params*" optimization (blue solid lines), SynX-based optimizations ($SynX_{Unmeasured} + Params$: red solid lines and $SynX_{Unmeasured}^{Params}$: yellow solid lines) and SO-based optimizations ($SynX_{All}^{Generic}$: purple dashed lines, $SynX_{All}^{Params}$: green dashed lines and $SO_{Unmeasured}^{Params}$:grey dashed lines). Data are reported for a complete gait cycle, where 0% indicates initial heel-strike and 100% indicates subsequent heel-strike. For measured muscles, the curves associated with $SynX_{Unmeasured}^{Params}$ and $SO_{Unmeasured}^{Params}$ were underneath the curves associated with "*Params*" since the associated muscle activations were taken from the "*Params*" optimization rather than estimated

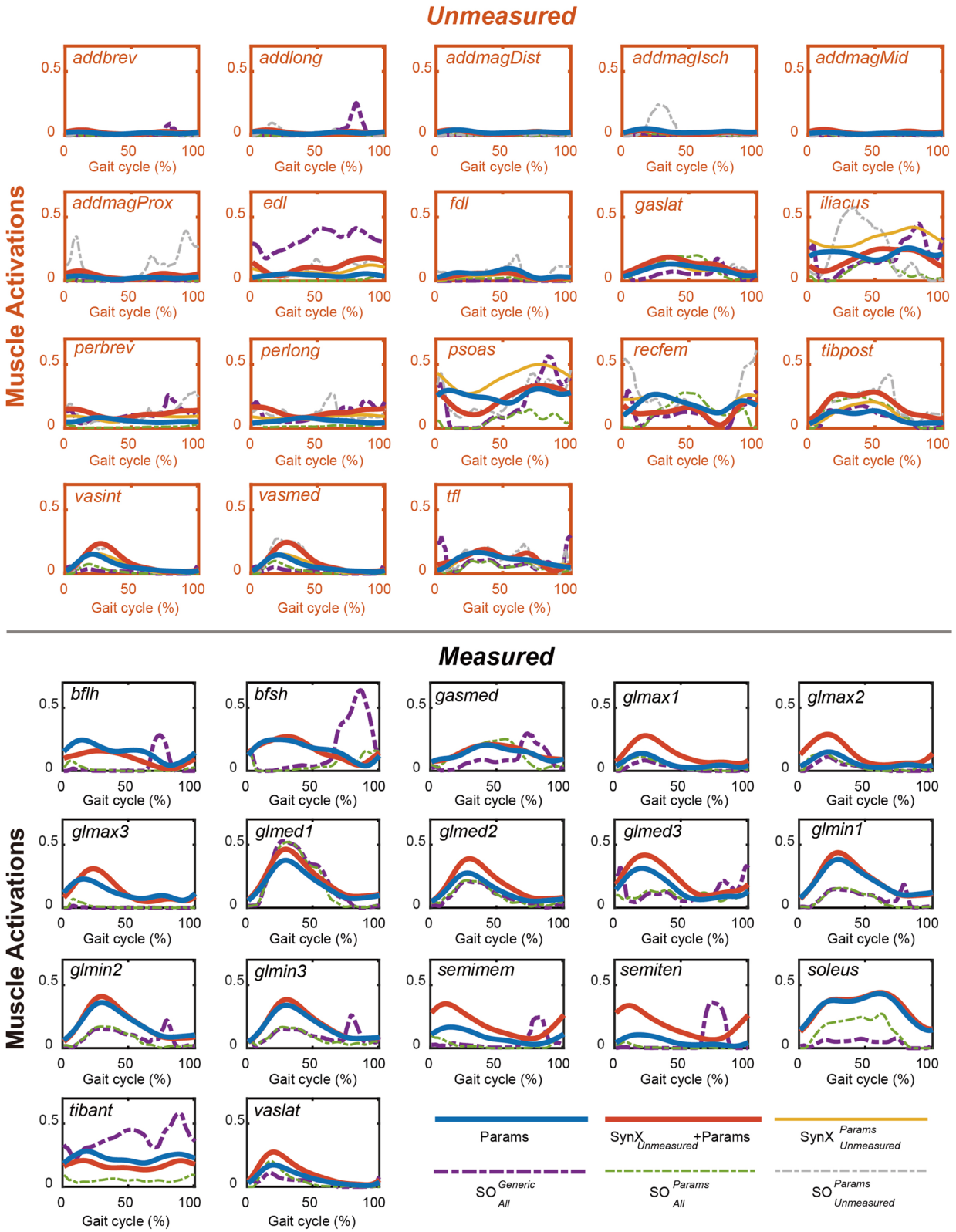

Fig. 4 (See legend on previous page.)

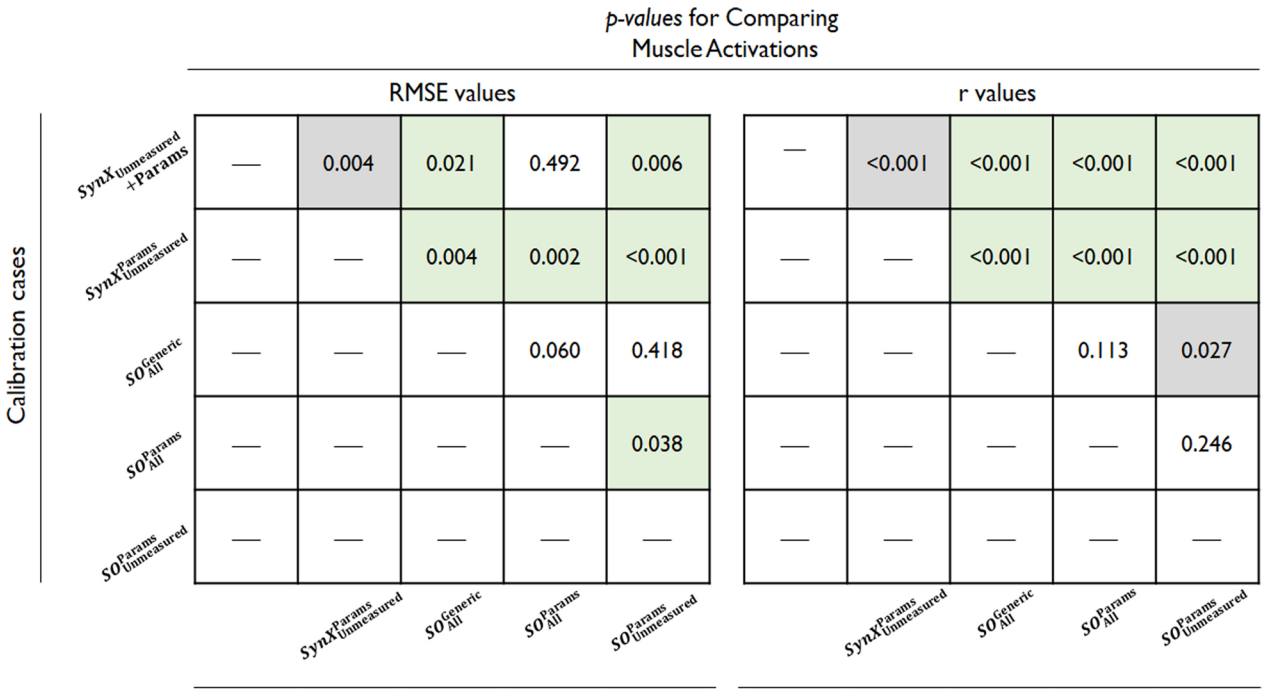

Fig. 5 *p-values* obtained from paired *t-tests* used to compare the estimation accuracy of muscle activations, as indicated by RMSE values (left) and *r* values (right), between different optimizations. Initially, RMSE and *r* values for muscle activations were calculated between the experimental ("*Params*") optimization and each of the other optimizations, where results across all calibration trials, legs, and subjects being concatenated and displayed in Table 3. Subsequently, the RMSE and *r* values for muscle activations from each optimization were compared to the results from every other optimization to determine the statistical significance of the differences in estimation accuracy between each pair of optimizations. All statistical analyses were performed in MATLAB, and the level of statistical significance was set to $p < 0.05$. A box with green background indicates that the estimation performance for the y-axis optimization was significantly better (lower RMSE values or higher r values) than for the x-axis optimization, while a box with grey background indicates that the estimation performance for the y-axis optimization was significantly worse (higher RMSE values or lower r values) than for the x-axis optimization

majority of adductor muscles in terms of both shape and magnitude.

RMSE values (101.3 ± 90.1) for SynX case "$SynX_{Unmeasured} + Params$"were significantly smaller ($p = 0.028$) than those for standard SO case "$SO^{Generic}_{All}$" (174.4 ± 174.4). Furthermore, correlation r values for muscle forces between SynX cases full EMG-driven model calibration ("*Params*" case) were moderately strong or better across all unmeasured muscles. Conversely, for SO, the correlation was generally weak for most muscles, except for moderate correlations observed for rectus femoris ($r = 0.42$), lateral gastrocnemius ($r = 0.51$), tibialis posterior ($r = 0.48$), and extensor digitorum longus ($r = 0.58$) (Table 4).

Model personalization had a considerable influence on the accuracy of estimated muscle forces for both SynX and SO (Figs. 6 and 7, Table 3). SynX demonstrated improved estimation of muscle force shapes ($p \leq 0.05$) and magnitudes ($p \leq 0.05$) when simultaneous calibration of SynX variables and EMG-driven model parameters ("$SynX_{Unmeasured} + Params$") was replaced with a well-calibrated EMG-driven model ("$SynX^{Params}_{Unmeasured}$"). Similarly, SO benefited from using well-calibrated model parameter values in achieving more accurate estimation of unmeasured muscle forces, leading to significantly different correlation r values between cases "$SO^{Generic}_{All}$" and "SO^{Params}_{All}", while RMSE values remained statistically comparable between them.

When model parameter values were taken from a full EMG-driven calibration ("*Params*"), SynX ("$SynX^{Params}_{Unmeasured}$") predicted unmeasured muscle forces more reliably than did SO ("$SO^{Params}_{Unmeasured}$"(, as evidenced by significantly lower RMSE values ($p \leq 0.05$) and higher correlation r values ($p \leq 0.05$).

Model parameters

In general, for measured muscles, the four activation dynamics model parameters and two Hill-type muscle–tendon model parameters showed a high degree of similarity between the "$SynX_{Unmeasured} + Params$ "and "*Params*" cases (Fig. 8, left panel). In contrast, for unmeasured muscles, only the two Hill-type muscle–tendon model parameters showed a high degree of

Table 2 Root mean square error (RMSE) values and Pearson correlation coefficient (*r*) values for estimated muscle activations calculated between the experimental ("Params") case and each calibration case

Unmeasured Muscles	Calibration cases									
	$SynX_{Unmeasured}$ +*Params*		$SynX^{Params}_{Unmeasured}$		$SO^{Generic}_{All}$		SO^{Params}_{All}		$SO^{Params}_{Unmeasured}$	
	RMSE	*r*	RMSE	*r*	RMSE	*r*	RMSE	*r*	RMSE	*r*
Adductor brevis	0.02	0.54	0.01	0.69	0.04	−0.17	0.02	−0.03	0.02	0.37
Adductor longus	0.02	0.60	0.01	0.83	0.11	−0.13	0.02	0.12	0.04	0.37
Adductor magnus distal	0.02	0.53	0.01	0.84	0.03	0.01	0.03	−0.07	0.03	−0.11
Adductor magnus ischial	0.02	0.55	0.01	0.89	0.04	0.23	0.04	−0.08	0.11	0.32
Adductor magnus middle	0.02	0.43	0.01	0.63	0.02	−0.17	0.02	−0.15	0.02	0.26
Adductor magnus proximal	0.04	0.58	0.01	0.76	0.03	−0.14	0.04	−0.21	0.19	0.47
Iliacus	0.17	0.43	0.17	0.86	0.24	0.05	0.21	0.08	0.33	0.04
Psoas	0.14	0.45	0.17	0.78	0.30	0.02	0.25	0.07	0.16	0.50
Rectus femoris	0.09	0.63	0.06	0.76	0.21	0.35	0.14	0.30	0.21	0.11
Tensor fasciae latae[2]	0.16	0.66	0.04	0.98	0.29	−0.09	0.24	−0.15	0.15	0.71
Vastus medialis	0.09	0.65	0.02	0.92	0.07	0.13	0.05	0.79	0.10	0.61
Vastus intermedius	0.08	0.67	0.02	0.92	0.08	0.12	0.06	0.81	0.08	0.59
Lateral gastrocnemius[2]	0.09	0.62	0.04	0.79	0.09	0.18	0.07	0.48	0.08	0.67
Tibialis posterior	0.15	0.53	0.08	0.58	0.11	0.53	0.13	0.54	0.20	0.16
Peroneus brevis	0.08	0.41	0.05	0.74	0.15	−0.31	0.07	0.29	0.11	0.25
Peroneus longus	0.09	0.42	0.05	0.79	0.13	0.13	0.06	0.43	0.14	0.25
Extensor digitorum longus[1]	0.17	0.38	0.08	0.42	0.62	0.48	0.08	0.09	0.15	0.21
Flexor digitorum longus[1]	0.05	0.92	0.09	0.82	0.14	0.89	0.11	0.95	0.11	0.38
Mean ± standard deviation	0.08 ± 0.06	0.55 ± 0.13	0.05 ± 0.05	0.78 ± 0.14	0.15 ± 0.15	0.12 ± 0.30	0.09 ± 0.07	0.24 ± 0.36	0.12 ± 0.08	0.34 ± 0.22

The RMSE and *r* values were calculated when the data across all calibration trials and subjects were concatenated

[1] Indicates the EMG signals were assumed unmeasured only for S1 and [2]indicates the EMG signals were assumed unmeasured only for S2

(See figure on next page.)

Fig. 6 Average muscle forces for the "unmeasured" muscles (upper) and "measured" muscles (lower) across calibration trials, legs and subjects from "*Params*" optimization (blue solid lines), SynX-based optimizations ($SynX_{Unmeasured} + Params$: red solid lines and $SynX^{Params}_{Unmeasured}$: yellow solid lines) and SO-based optimizations ($SynX^{Generic}_{All}$: purple dashed lines, $SynX^{Params}_{All}$: green dashed lines and $SO^{Params}_{Unmeasured}$:grey dashed lines). Data are reported for a complete gait cycle, where 0% indicates initial heel-strike and 100% indicates subsequent heel-strike. For measured muscles, the curves associated with $SynX^{Params}_{Unmeasured}$ and $SO^{Params}_{Unmeasured}$ were underneath the curves associated with "*Params*" since the associated muscle activations were taken from the "*Params*" optimization rather than estimated

similarity between the "$SynX_{Unmeasured} + Params$"and "*Params*" cases (Fig. 8, right panel). The four activation dynamics model parameters exhibited substantial differences for unmeasured muscles between SynX case "$SynX_{Unmeasured} + Params$" and the full EMG-driven model calibration case "*Params*".

Discussion

This study evaluated the ability of synergy extrapolation (SynX) to perform concurrent estimation of a large number of unmeasured muscle excitations and parameter values in an EMG-driven model. The approach was evaluated using gait datasets collected from two post-stroke subjects performing treadmill walking at self-selected and fastest-comfortable speeds. EMG signals measured from eight muscles bilaterally were treated as "unmeasured" and estimated using synergy information extracted from EMG signals measured from another eight muscles bilaterally that were treated as "measured." The muscle activations, forces, and model parameter values for the unmeasured muscles were quantitatively compared to "gold standard" values obtained when all 16 channels of EMG data were used to calibrate an EMG-driven musculoskeletal model for each leg of each subject. The results revealed that the estimated unmeasured muscle activations and forces were reasonably accurate and reliable in term of both shape and magnitude (Figs. 4 and 6, Tables 2 and 3). Moreover, Hill-type muscle–tendon model parameter values for both unmeasured and measured muscles, including optimal fiber length and tendon slack length, exhibited a high level of agreement with "gold standard" model parameter values (Fig. 8). When SO estimates of unmeasured muscle activations and forces were compared with SynX estimates, the SynX results were more accurate and realistic than those from SO (Figs. 4 and 6, Tables 2 and 3), which contained abrupt changes and tended to underestimate the unmeasured muscle quantities. When the sensitivity of estimated unmeasured muscle activations and forces to the level of model personalization was investigated, both SynX and SO generated more accurate estimates when utilizing well-calibrated muscle–tendon model parameters. However, SynX demonstrated superior performance to SO in estimating unmeasured muscle activations and forces when employing model parameter values from full EMG-driven model calibration.

Several factors help explain why SynX demonstrated superior performance over SO. First, by utilizing measured synergy excitations as building blocks, SynX reduced the problem of finding unknown time-varying muscle excitations to identifying a small number of unmeasured synergy vector weights. This simplification led to a substantial reduction in the search space for the optimization in comparison with SO-based approaches [33]. Second, unlike SO-based approaches, which solve for muscle activations one time frame at a time, the dependence between time frames in weighted synergy excitations resulted in continuous muscle activations, improving the physiological plausibility of muscle excitation estimates. Third, the time-invariance of unmeasured and residual synergy vector weights enabled a single-layer optimization process, simultaneously achieving EMG-driven model personalization and muscle activation estimation, which enhanced the accuracy of muscle force estimation. Fourth, calibration of synergy-structured residual muscle excitations was integrated into SynX to enhance the accuracy of predicted unmeasured muscle excitations. Unlike SO, SynX introduced residual muscle excitations to prevent missing muscle excitations from compensating for errors in measured muscle excitations, improving solution accuracy as a consequence [34]. Fifth, SynX-based methods minimized muscle activations only for unmeasured muscles whereas SO-based methods minimized muscle activations, and thus co-activation between agonist and antagonist muscles, for all muscles [24, 50]. Last, the SynX-based methods leveraged the concept of muscle synergies, making the method more physiological reasonable [51, 52].

Compared to other computational methods, SynX offers benefits for estimating missing EMG signals within [19, 53–55] or outside [56–58] the context of musculoskeletal modeling. One published computational method used Gaussian process regression models to describe the synergistic relationship between a subset of muscles, which enabled estimation of unmeasured muscle excitations using information provided by a subset of measured muscle excitations [57]. However, the muscle excitations

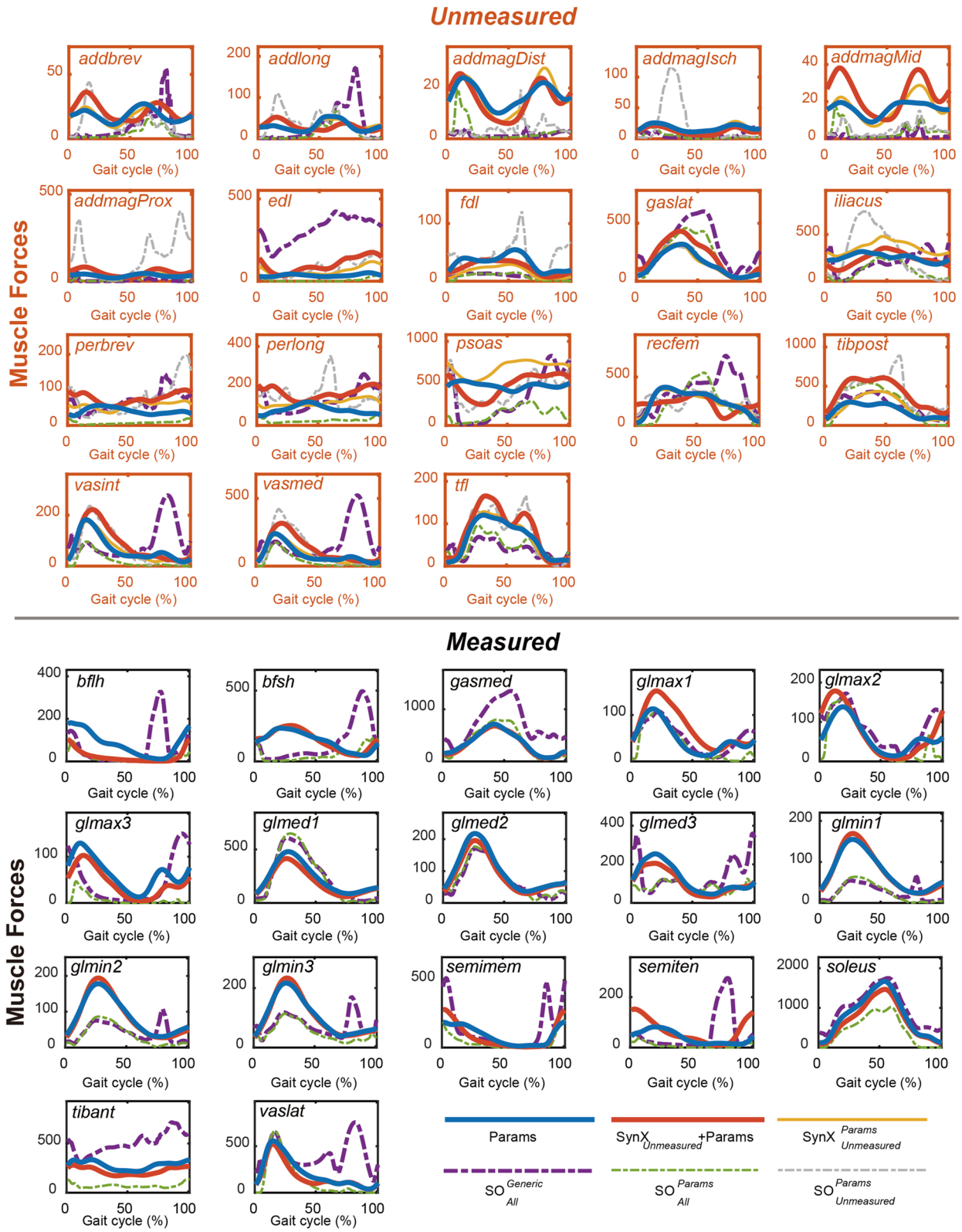

Fig. 6 (See legend on previous page.)

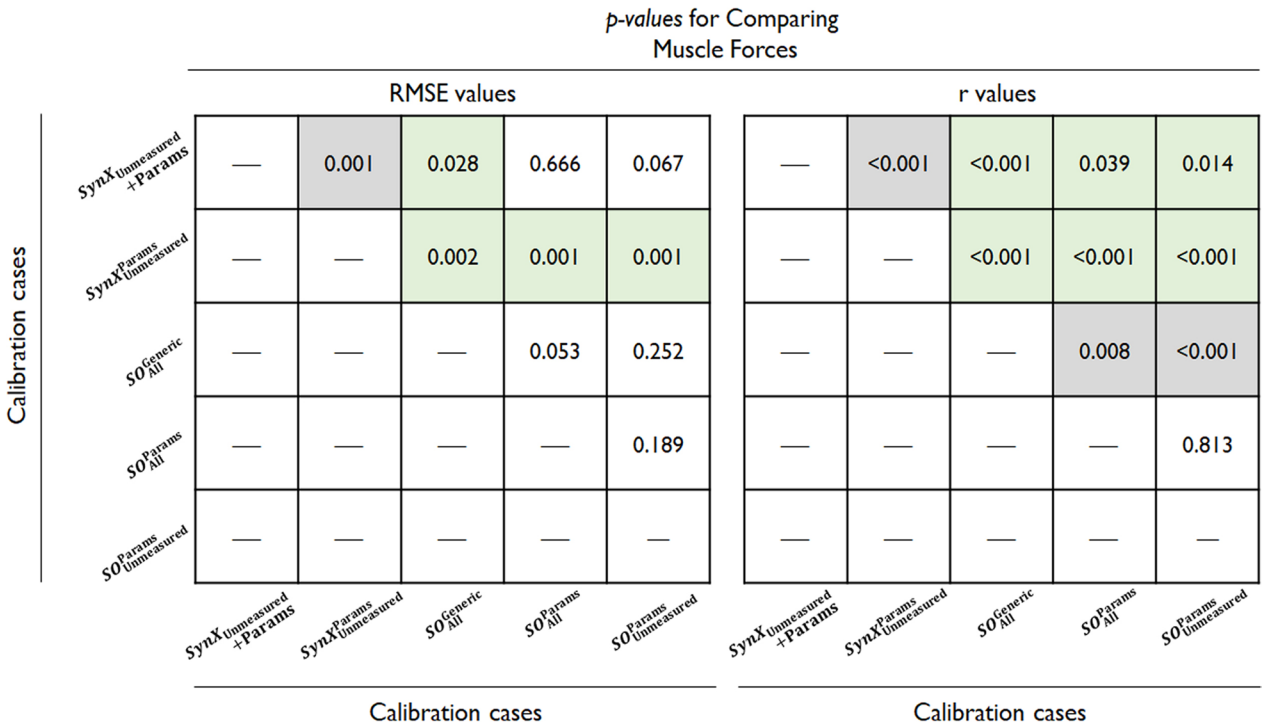

Fig. 7 p-*values* obtained from paired *t-tests* used to compare the estimation accuracy of muscle forces, as indicated by RMSE values (left) and *r* values (right), between different optimizations. Initially, RMSE and *r* values for muscle forces were calculated between the experimental ("*Params*") optimization and each of the other optimizations, where results across all calibration trials, legs, and subjects being concatenated and displayed in Table 3. Subsequently, the RMSE and *r* values for muscle forces from each optimization were compared to the results from every other optimization to determine the statistical significance of the differences in estimation accuracy between each pair of optimizations. All statistical analyses were performed in MATLAB, and the level of statistical significance was set to $p < 0.05$. A box with green background indicates that the estimation performance for the y-axis optimization was significantly better (lower RMSE values or higher r values) than for the x-axis optimization, while a box with grey background indicates that the estimation performance for the y-axis optimization was significantly worse (higher RMSE values or lower r values) than for the x-axis optimization

associated with "unmeasured" muscles must be known for conducting the required model training process, rendering this method infeasible when the "unmeasured" muscle excitations are truly unmeasured due to experimental or safety considerations. A second published computational method employed low-dimensional sets of impulsive excitation primitives to estimate unmeasured muscle excitations [19, 53, 54]. Each muscle is assigned to a module by evaluating associated weighting factors for the excitation primitives derived from measured muscle excitations. Muscles without EMG signals are assumed to belong to the same module as measured muscles that share the same innervation and contribute to the same mechanical action. The primitive-driven excitations for measured muscles are then minimally adjusted to improve joint moment estimation in EMG-assisted models. However, these adjustments mask the omission of active force generating properties for some unmeasured muscles (i.e., iliacus and psoas), resulting in noticeable hip joint moment prediction errors. A third published computational method used hybrid EMG-informed models that incorporate SO to determine unmeasured muscle activations [22, 23]. That study also allowed minimal adjustments of measured muscle activations while predicting missing EMG signals (e.g., from iliacus and psoas) using SO [22]. A final published computational method –the computed muscle control (CMC) algorithm within OpenSim – solves for the muscle excitations needed to achieve the desired accelerations for tracking an experimental motion [55, 59–61]. However, CMC has been observed to generate different muscle activation and force solutions depending on the point in time at which the simulation is started [55]. None of these studies provided evidence that estimation of unmeasured muscle activations was reliable and in reasonable agreement with experimental measurements. Furthermore, due to the nature of SO, the resulting muscle activations are likely to exhibit unrealistic discontinuities. All in all, EMG-driven modeling method with SynX provides an enhanced approach for estimating unmeasured muscle activations and forces in an efficient manner without requiring a priori knowledge of the "unmeasured" muscle excitations during a model training phase.

Table 3 Root mean square error (RMSE) values and Pearson correlation coefficient (*r*) values for estimated muscle forces calculated between the experimental ("Params") case and each calibration case

Unmeasured muscles	Calibration cases									
	$SynX_{Unmeasured}$ + *Params*		$SynX^{Params}_{Unmeasured}$		$SO^{Generic}_{All}$		SO^{Params}_{All}		$SO^{Params}_{Unmeasured}$	
	RMSE	*r*	RMSE	*r*	RMSE	*r*	RMSE	*r*	RMSE	*r*
Adductor brevis	14.4	0.35	8.1	0.71	24.1	−0.13	16.3	0.36	15.6	0.29
Adductor longus	23.8	0.47	11.4	0.87	74.7	−0.01	25.3	0.67	43.3	0.49
Adductor magnus distal	10.9	0.43	7.1	0.76	18.8	−0.11	19.4	−0.14	18.3	−0.19
Adductor magnus ischial	12.8	0.36	7.1	0.82	20.1	−0.05	21.4	−0.17	51.6	0.16
Adductor magnus middle	17.3	0.46	9.0	0.57	17.2	−0.12	16.8	−0.08	14.2	0.19
Adductor magnus proximal	39.2	0.42	12.5	0.69	34.4	0.03	42.7	−0.02	186.7	0.32
Iliacus	158.9	0.36	134.5	0.85	235.5	0.04	197.8	0.31	348.6	0.30
Psoas	273.6	0.40	249.4	0.83	530.9	−0.11	429.4	−0.05	285.9	0.37
Rectus femoris	136.8	0.68	58.3	0.94	350.8	0.42	181.0	0.70	156.2	0.47
Tensor fasciae latae[2]	115.1	0.66	20.7	0.99	126.8	0.31	142.9	0.13	83.3	0.82
Vastus medialis	115.4	0.60	31.5	0.92	290.1	−0.05	68.6	0.81	119.9	0.71
Vastus intermedius	74.4	0.66	23.8	0.91	154.4	−0.04	58.7	0.84	65.8	0.69
Lateral gastrocnemius[2]	162.6	0.86	67.7	0.94	153.9	0.51	144.9	0.73	95.8	0.89
Tibialis posterior	320.0	0.44	166.1	0.56	214.2	0.48	293.6	0.54	424.0	0.14
Peroneus brevis	50.2	0.48	31.9	0.77	82.0	−0.24	47.4	0.35	74.7	0.31
Peroneus longus	109.9	0.49	63.4	0.80	126.5	0.21	81.9	0.50	170.9	0.29
Extensor digitorum longus[1]	156.1	0.48	71.3	0.57	611.8	0.57	64.7	0.38	128.2	0.47
Flexor digitorum longus[1]	31.2	0.91	51.1	0.81	73.1	−0.45	62.3	0.94	63.3	0.39
Mean ± standard deviation	101.3 ± 90.1	0.53 ± 0.17	56.9 ± 65.4	0.79 ± 0.13	174.4 ± 174.4	0.07 ± 0.28	106.4 ± 110.9	0.38 ± 0.36	130.4 ± 116.8	0.39 ± 0.26

The RMSE and *r* values were calculated when the data across all calibration trials and subjects were concatenated

[1] Indicates the EMG signals were assumed unmeasured only for S1 and [2]indicates the EMG signals were assumed unmeasured only for S2

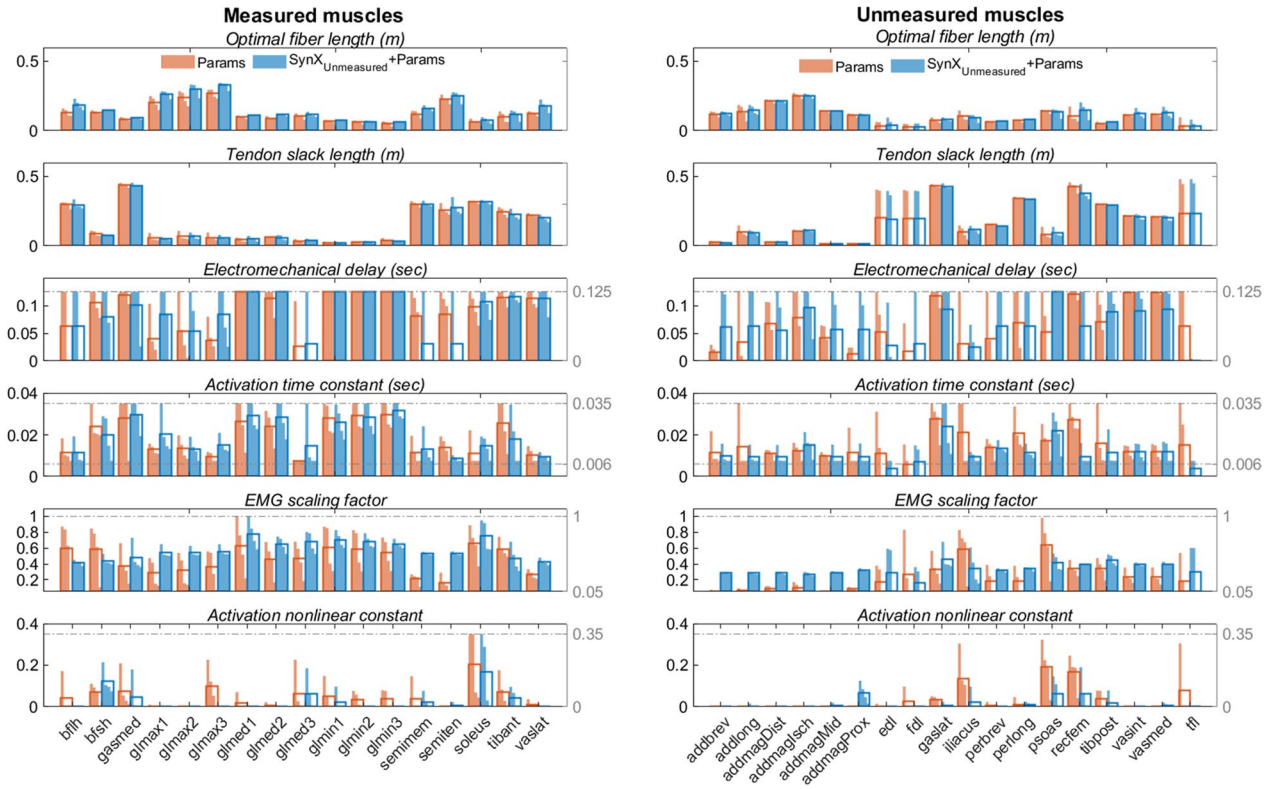

Fig. 8 EMG-driven model parameter values for two legs of both subjects from "$SynX_{Unmeasured} + Params$" optimization (in orange) and "gold standard (*Params*)" optimization (blue bars). The upper and lower bounds during optimization for each of the four activation dynamics model parameters have been indicated by grey dash-sot lines, where the upper and lower bounds for the scaling factors of optimal fiber lengths and tendon slack lengths were [0.6, 1.4] for all muscles

Personalized muscle–tendon model parameters have the potential to significantly enhance the accuracy of estimating muscle activations and forces, utilizing both SynX and SO approaches. For SynX, tracking errors between the estimated and experimental estimates of muscle activations were reduced substantially when muscle–tendon model parameter values were pre-personalized for case "$SynX^{Params}_{Unmeasured}$". Consequently, mean correlations between estimated and experimental muscle activations were also substantially increased, moving from moderate to strong. Additionally, compared to case "$SynX_{Unmeasured} + Params$", net joint moment matching errors for case "$SynX^{Params}_{Unmeasured}$" were closer to those obtained from full EMG-driven model calibration case "*Params*". For SO, consistent with previous studies [62], personalization of muscle–tendon model parameters produced noticeable improvements in estimated muscle activations and forces in terms of both shape and amplitude for case "SO^{Params}_{All}", with statistically significant improvements observed for only the shape of muscle forces. These observations suggest that enhancing the level of model personalization generally improves the accuracy of muscle activation estimation, consistent with a previous study demonstrating that model personalization also improves the accuracy of knee contact force estimation [47]. Despite these observations, significant variations existed in the degree of improvement among different optimization approaches. In scenarios where a well-calibrated musculoskeletal model is available, SynX has the ability to predict muscle activations for muscles lacking EMG data with reasonable amplitude and shape, whereas SO can predict unmeasured muscle activations with reasonable amplitude but not shape. Even though SynX started with the same scaled generic model as used by SO, it predicted unmeasured muscle activations with reasonably accurate amplitude and shape, which SO did not achieve.

Net joint moment matching errors were also found to differ among the different optimization approaches (Figure S3 and Table S3). First, inverse dynamics (ID) and estimated net joint moments exhibited much closer agreement in the SO-based optimizations ("$SO^{Generic}_{All}$", "SO^{Params}_{All}", and "$SO^{Params}_{Unmeasured}$") than in the EMG-driven modeling optimizations ("*Params*", "$SynX_{Unmeasured} + Params$", and "$SynX^{Params}_{Unmeasured}$"). The optimization formulation used by SO in Eq. (10) resulted

in negligible joint moment matching errors. However, the additional constraints within EMG-modeling methods, including muscle activation dynamics and dependency of EMG signals between time frames, limited the torque-generating capacity of muscles, thereby preventing more precise reproduction of joint moments. Second, joint moment matching errors, arranged in descending order, for optimizations associated with the EMG-driven modeling method are "*Params*", "$SynX^{Params}_{Unmeasured}$", and "$SynX_{Unmeasured} + Params$". This observation can be explained by the increasing number of design variables, and thus the increasing freedom, in the optimization problem formulations. Specifically, "*Params*" uses the fewest design variables, "$SynX^{Params}_{Unmeasured}$" adds SynX design variables, and "$SynX_{Unmeasured} + Params$" further adds EMG-driven model design variables, with each addition enabling the optimizer to reduce the joint moment matching errors further. This observation also helps explain why joint moment matching errors were smaller when estimating 8 channels of unmeasured EMG signals in the present study compared to estimating only two channels of unmeasured EMG signals (i.e., iliacus and psoas) in a previous study that used the same datasets [34]. Last, regardless of the numbers of muscle activations to estimate, SO consistently finds muscle activation solutions that match ID joint moments almost perfectly at each time frame, although occasionally a small amount of reserve actuator torque is needed due to model inadequacies. Consequently, static optimization does not possess the joint moment matching errors needed to calibrate muscle–tendon model parameter values.

Incorporation of the SynX process into an EMG-driven modeling framework had minimal impact on some calibrated model parameter values but a larger impact on others. Specifically, Hill-type muscle–tendon model parameter values, specifically optimal fiber length and tendon slack length, predicted by case "$SynX_{Unmeasured} + Params$" closely approximated the "gold standard" values obtained from full EMG-driven model calibration case "*Params*", as depicted in Fig. 8. This finding can be explained by the fact that Hill-type muscle–tendon model parameter values directly affect calculated net joint moments, which are the primary error terms used for EMG-driven model calibration. In contrast, activation dynamics model parameter values, including electromechanical delay, activation time constant, EMG scale factor, and activation nonlinear constant, predicted by case "$SynX_{Unmeasured} + Params$" showed only reasonable similarity to "gold standard" values for the measured muscles. This finding is understandable since conversion of muscle activations back to muscle excitations does not affect the joint moment matching errors that drive the SynX calibration process, and no data are available to the calibration process that make the conversion process unique.

Several methodological choices needed to be made to perform SynX, as indicated in Table 2, and these choices could potentially impact SynX performance. A series of previously published studies from the authors have extensively investigated various methodological choices, with the goal of identifying an optimal combination that could yield the most reliable and accurate estimation of unmeasured muscle activations [32–34]. Initially, principal component analysis (PCA) provided more accurate, reliable, and efficient estimates of unmeasured muscle excitations compared to non-negative matrix factorization (NMF). The non-negativity constraints for NMF and extra design variables for PCA both result in a more restricted feasible search space for NMF in comparison to PCA [33, 34]. Additionally, PCA was particularly beneficial in this computational framework since it permitted residual excitations to be both positive and negative, which could be beneficial for achieving lower joint moment errors. Second, by comparing the results of five different EMG normalization methods that were performed either within individual trials or across all trials, we observed that EMG normalization does not have a significant influence on the SynX performance [33]. As a result, measured muscle excitations were normalized to their maximum values across all trials before MSA to facilitate easy implementation. Furthermore, as the number of synergies increased, the performance of SynX exhibited non-monotonic behavior, with five or six synergies generally providing the best SynX performance and outcomes for EMG-driven model calibration [34]. Hence, when the present study treated a large number of muscles as "unmeasured," five synergies were selected for generating the results. Last, based on assumptions about how synergy vectors vary across walking cycles, we categorized synergy vectors associated with unmeasured and residual muscle excitations as trial-specific, speed-specific, or subject-specific, respectively, while different concatenation strategies were used to extract corresponding synergy excitations. We found that for equal numbers of synergies, trial-specific unmeasured synergy vector weights and speed-specific residual synergy vector weights produced the best SynX performance [34]. This observation informed the way synergy vector weights were allowed to vary across walking cycles analyzed in the present study.

A reasonable choice of a neural control strategy is essential for producing physiologically realistic predictive simulations of walking [63]. Presently, there is a perspective suggesting that human locomotion control could potentially be perceived as a hierarchical

structure with two layers. The lower layer is believed to create fundamental motor patterns, while the upper layer is thought to convey commands to the lower layer to adjust these patterns[64]. Specifically, the lower layer is proposed to involve two control mechanisms: reflexes and central pattern generators (CPGs), which may represent feedback and feedforward control, respectively [65]. In the context of typical voluntary movements like walking, it is hypothesized that the central nervous system's involvement in the upper layer may primarily pertain to task initiation. Following initiation, a cluster of motoneurons in the spinal cord is theorized to maintain continuity by generating predetermined activation patterns.[66]. Muscle synergies, where a single pathway simultaneously activates multiple muscles, have also been proposed as a lower layer control mechanism that reduces the degrees of freedom for complex control tasks [65]. This underscores the potential advantages of modeling muscle activations using the synergy concept over SO-based approach, aligning with the true motor control scheme for walking. However, given the consensus that humans use minimum effort to conduct well-practiced motor tasks, such as walking, SO hypothesis remains the primary and easily implemented neural control principle when performing optimal control analysis, with the standard practice of minimizing the sum of squares of muscle activations [20, 63, 67]. The comparison of estimated muscle activations and forces from both SynX and SO reported in this study raise a pivotal question: If the prevalent neural control strategy of minimizing the sum of squares of muscle activations fails to estimate muscle activations reliably when joint kinematics and moments are known a priori from experimental walking data, how can it provide reliable estimates in predictive simulations of walking when joint kinematics and moments are unknown a priori? Our findings also suggest the potential benefits of using muscle synergies for predicting walking motion with musculoskeletal models. While the reliability of a synergy-based neural control strategy for generating predictive simulations of walking has been verified for only one experimental scenario thus far [68], the results of the present study support further exploration of a synergy-based neural control strategy for generating predictive simulations of walking.

This study exhibited several limitations which may provide insights for future research endeavors. First, this study validated the effectiveness of our EMG-driven modeling framework incorporating SynX by analyzing gait datasets from only two subjects post-stroke. These experimental datasets were chosen since they provided EMG signals for every major muscle in our musculoskeletal model, enabling the evaluation of estimation accuracy. Further investigation is necessary to investigate diverse subject populations with larger sample sizes. Second, we evaluated SynX using only walking data and only two representative speeds. It would be valuable to evaluate SynX further using a broader set of dynamic movement conditions and experimental scenarios, including stair climbing, running, and upper extremity tasks [35]. Third, SynX has been evaluated thus far using only Hill-type muscle–tendon models with a rigid tendon. This choice was made to enhance computational efficiency and since rigid and compliant tendon models produce almost identical muscle force estimates for slow movements like walking at a healthy speed, but different muscle force estimates for faster movements such as running [69, 70]. As both of our subjects walked at slow speeds, use of a rigid tendon model was likely appropriate. However, it would be worthwhile to expand our approach to include compliant tendon models in our Hill-type muscle–tendon models, enabling applications involving fast movements. Fourth, it is unknown how well SynX would work when used to calibrate only a single-joint model. It would be interesting to perform SynX-based estimates of muscle activations and forces for single-joint model calibration using all eight EMG electrodes dedicated to muscles spanning the joint of interest (e.g., the knee) and then compare the results with those obtained from multi-joint model calibration employing all eight EMG electrodes placed across all joints in each leg, as proposed in this study. In the case of the knee, the accuracy of the two approaches could be evaluated objectively using in vivo measurements of knee contact forces available from the "Knee Grand Challenge" datasets [71]. This evaluation could offer valuable insights for researchers focusing on applications for a single joint. Last, we analyzed the impact of personalizing only activation dynamics and muscle–tendon model parameter values on SynX performance. However, personalization of other model aspects, including skeletal geometries (as recently investigated by the authors [72]), muscle kinematics, and other physiological properties that contribute to muscle force generation, may also impact muscle force estimates. Future work should investigate whether SynX performance is sensitive to these additional aspects of model personalization.

Conclusions

In conclusion, this study demonstrated a significant advancement over previous research by highlighting the capability of SynX to reproduce a large number of muscle activations associated with unmeasured muscle excitations while simultaneously calibrating EMG-driven model parameter values. Notably, the estimation accuracy of muscle activations and forces

in terms of shape and amplitude for the unmeasured muscles was significantly higher than that of the standard SO approach. The incorporation of SynX into an EMG-driven model calibration process had minimal impact on the calibrated Hill-type muscle–tendon model parameter values for all muscles and activation dynamics model parameter values for measured muscles. Additionally, when integrated with well-calibrated musculoskeletal models, both SynX and SO produced substantially more accurate estimates of unmeasured muscle activations and forces, with SynX demonstrating superior performance over SO. The findings suggest that SynX could effectively address the practical challenge of collecting a full set of EMG signals for EMG-driven modeling calibration in the lower extremity during walking, with significant implications for personalized treatments for movement impairments in situations where difficulties arise in collecting EMG signals from all important contributing muscles.

Abbreviations

EMG	Electromyography
SynX	Synergy extrapolation
SO	Static optimization
RMSE	Root mean square errors
DOFs	Degrees of freedom
IK	Inverse kinematics
ID	Inverse dynamics
MSA	Muscle synergy analysis
PCA	Principal component analysis
NMF	Non-negative matrix factorization

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12984-024-01490-y.

Additional file1.

Author contributions

B.J.F. designed and performed the experiments; D.A. wrote the programs; D.A. analyzed the data, prepared figures, and drafted the manuscript; D.A. wrote the manuscript; D.A. and B.J.F. revised the manuscript; D.A. and B.J.F. approved the final version of the manuscript.

Funding

This work was funded by the Cancer Prevention Research Institute of Texas (CPRIT) under grant RR170026 to B.J.F., the National Institutes of Health under grant R01 EB030520 to B.J.F., and the National Natural Science Foundation (NSF) of China under grant 32301090 to D.A.

Availability of data and materials

The SynX EMG-driven modeling process presented in this study is freely available within the Muscle–tendon Model Personalization Tool provided with the open-source Neuromusculoskeletal Modeling Pipeline software (https://nmsm.rice.edu).

Declarations

Ethics approval and consent to participate

All experimental procedures were performed in accordance with Declaration of Helsinki and approved by the University of Florida Health Science Center Institutional Review Board (IRB-01), and the subjects provided written informed consent before participation.

Consent for publication

Not applicable.

Competing interest

The authors declare no competing interests.

Author details

[1]Institute of Medical Research, Northwestern Polytechnical University, Xi'an, Shaanxi, China. [2]Department for Mechanical Engineering, Rice University, Houston, TX, USA.

Received: 12 February 2024 Accepted: 15 October 2024
Published online: 01 November 2024

References

1. Pandy MG, Lin Y-C, Kim HJ. Muscle coordination of mediolateral balance in normal walking. J Biomech. 2010;43:2055–64.
2. Mathieu E, Crémoux S, Duvivier D, Amarantini D, Pudlo P. Biomechanical modeling for the estimation of muscle forces: toward a common language in biomechanics, medical engineering, and neurosciences. J NeuroEngineering Rehabil. 2023;20:130.
3. Berger DJ, d'Avella A. Effective force control by muscle synergies. Front Comput Neurosci. 2014;8:46.
4. Kim HJ, Fernandez JW, Akbarshahi M, Walter JP, Fregly BJ, Pandy MG. Evaluation of predicted knee-joint muscle forces during gait using an instrumented knee implant. J Orthop Res. 2009;27:1326–31.
5. Marra MA, Vanheule V, Fluit R, Koopman BH, Rasmussen J, Verdonschot N, et al. A subject-specific musculoskeletal modeling framework to predict in vivo mechanics of total knee arthroplasty. J Biomech Eng. 2015;137:020904.
6. Shao Q, Bassett DN, Manal K, Buchanan TS. An EMG-driven model to estimate muscle forces and joint moments in stroke patients. Comput Biol Med. 2009;39:1083–8.
7. Cano-de-la-Cuerda R, Pérez-de-Heredia M, Miangolarra-Page JC, Muñoz-Hellín E, Fernández-de-las-Peñas C. Is there muscular weakness in Parkinson's disease? Am J Phys Med Rehabil. 2010;89:70–6.
8. Trinler U, Hollands K, Jones R, Baker R. A systematic review of approaches to modelling lower limb muscle forces during gait: applicability to clinical gait analyses. Gait Posture. 2018;61:353–61.
9. Fregly BJ. A conceptual blueprint for making neuromusculoskeletal models clinically useful. Appl Sci. 2021;11:2037.
10. Vega MM, Li G, Shourijeh MS, Ao D, Weinschenk RC, Patten C, et al. Computational evaluation of psoas muscle influence on walking function following internal hemipelvectomy with reconstruction. Front Bioeng Biotechnol. 2022;10:1748.
11. Martin JA, Brandon SC, Keuler EM, Hermus JR, Ehlers AC, Segalman DJ, et al. Gauging force by tapping tendons. Nat Commun. 2018;9:1592.
12. Keuler EM, Loegering IF, Martin JA, Roth JD, Thelen DG. Shear wave predictions of achilles tendon loading during human walking. Sci Rep. 2019;9:13419.
13. Buchanan TS, Lloyd DG, Manal K, Besier TF. Neuromusculoskeletal modeling: estimation of muscle forces and joint moments and movements from measurements of neural command. J Appl Biomech. 2004;20:367–95.
14. Heine R, Manal K, Buchanan TS. Using hill-type muscle models and EMG data in a forward dynamics analysis of joint moment: evaluation of critical parameters. J Mech Med Biol. 2003;3(2):169–86.
15. Hoang HX, Diamond LE, Lloyd DG, Pizzolato C. A calibrated EMG-informed neuromusculoskeletal model can appropriately account for

muscle co-contraction in the estimation of hip joint contact forces in people with hip osteoarthritis. J Biomech. 2019;83:134–42.
16. Lloyd DG, Besier TF. An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo. J Biomech. 2003;36:765–76.
17. Meyer AJ, Patten C, Fregly BJ. Lower extremity EMG-driven modeling of walking with automated adjustment of musculoskeletal geometry. PLoS One. 2017;12: e0179698.
18. Sartori M, Reggiani M, Farina D, Lloyd DG. EMG-driven forward-dynamic estimation of muscle force and joint moment about multiple degrees of freedom in the human lower extremity. Gribble PL, editor. PLoS ONE. 2012;7: e52618.
19. Pizzolato C, Lloyd DG, Sartori M, Ceseracciu E, Besier TF, Fregly BJ, et al. CEINMS: a toolbox to investigate the influence of different neural control solutions on the prediction of muscle excitation and joint moments during dynamic motor tasks. J Biomech. 2015;48:3929–36.
20. Anderson FC, Pandy MG. Static and dynamic optimization solutions for gait are practically equivalent. J Biomech. 2001;34:153–61.
21. Heintz S, Gutierrez-Farewik EM. Static optimization of muscle forces during gait in comparison to EMG-to-force processing approach. Gait Posture. 2007;26:279–88.
22. Sartori M, Farina D, Lloyd DG. Hybrid neuromusculoskeletal modeling to best track joint moments using a balance between muscle excitations derived from electromyograms and optimization. J Biomech. 2014;47:3613–21.
23. Zonnino A. Model-based estimation of individual muscle force based on measurements of muscle activity in forearm muscles. IEEE Trans Biomed Eng. 2020;67:134–45.
24. Kian A, Pizzolato C, Halaki M, Ginn K, Lloyd D, Reed D, et al. Static optimization underestimates antagonist muscle activity at the glenohumeral joint: a musculoskeletal modeling study. J Biomech. 2019;97: 109348.
25. Shourijeh MS, Mehrabi N, McPhee J. Forward static optimization in dynamic simulation of human musculoskeletal systems: a proof-of-concept study. J Comput Nonlinear Dyn. 2017;12: 051005.
26. Michaud F, Lamas M, Lugrís U, Cuadrado J. A fair and EMG-validated comparison of recruitment criteria, musculotendon models and muscle coordination strategies, for the inverse-dynamics based optimization of muscle forces during gait. J NeuroEngineering Rehabil. 2021;18:17.
27. Zajac FE. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. Crit Rev Biomed Eng. 1989;17:359–411.
28. Bernstein N. The coordination and regulation of movements. 1967.
29. Schellenberg F, Oberhofer K, Taylor WR, Lorenzetti S. Review of modelling techniques for in vivo muscle force estimation in the lower extremities during strength training. Comput Math Methods Med. 2015;2015:1–12.
30. Vilimek M. Musculotendon forces derived by different muscle models. Acta Bioeng Biomech. 2007;9:41–7.
31. Sartori M, Gizzi L, Lloyd DG, Farina D. A musculoskeletal model of human locomotion driven by a low dimensional set of impulsive excitation primitives. Front Comput Neurosci. 2013. https://doi.org/10.3389/fncom.2013.00079.
32. Bianco NA, Patten C, Fregly BJ. Can measured synergy excitations accurately construct unmeasured muscle excitations? J Biomech Eng. 2018;140: 011011.
33. Ao D, Shourijeh MS, Patten C, Fregly BJ. Evaluation of synergy extrapolation for predicting unmeasured muscle excitations from measured muscle synergies. Front Comput Neurosci. 2020;14: 588943.
34. Ao D, Vega MM, Shourijeh MS, Patten C, Fregly BJ. EMG-driven musculoskeletal model calibration with estimation of unmeasured muscle excitations via synergy extrapolation. Front Bioeng Biotechnol. 2022;10: 962959.
35. Tahmid S, Font-Llagunes JM, Yang J. Upper extremity muscle activation pattern prediction through synergy extrapolation and electromyography-driven modeling. J Biomech Eng. 2024;146: 011005.
36. Fregly BJ. How well do commonly used co-contraction indices approximate lower limb joint stiffness trends during gait for individuals post-stroke? Front Bioeng Biotechnol. 2021. https://doi.org/10.3389/fbioe.2020.588908.
37. McLean SG, Huang X, Van Den Bogert AJ. Association between lower extremity posture at contact and peak knee valgus moment during sidestepping: implications for ACL injury. Clin Biomech. 2005;20:863–70.
38. Hug F. Can muscle coordination be precisely studied by surface electromyography? J Electromyogr Kinesiol. 2011;21:1–12.
39. Rajagopal A, Dembia CL, DeMers MS, Delp DD, Hicks JL, Delp SL. Full-body musculoskeletal model for muscle-driven simulation of human gait. IEEE Trans Biomed Eng. 2016;63:2068–79.
40. Delp SL, Anderson FC, Arnold AS, Loan P, Habib A, John CT, et al. OpenSim: open-source software to create and analyze dynamic simulations of movement. IEEE Trans Biomed Eng. 2007;54:1940–50.
41. Seth A, et al. OpenSim Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement. PLoS Comput Biol. 2018;14: e1006223.
42. Reinbolt JA, Schutte JF, Fregly BJ, Koh BI, Haftka RT, George AD, et al. Determination of patient-specific multi-joint kinematic models through two-level optimization. J Biomech. 2005;38:621–6.
43. Cronin NJ. Spatial variability of muscle activity during human walking: the effects of different EMG normalization approaches. Neuroscience. 2015;300:19–28.
44. He J, Levine WS, Loeb GE. Feedback gains for correcting small perturbations to standing posture. IEEE; 1989. p. 518–26.
45. Manal K, Buchanan TS. A one-parameter neural activation to muscle activation model: estimating isometric joint moments from electromyograms. J Biomech. 2003;36:1197–202.
46. Hill AV. The heat of shortening and the dynamic constants of muscle. Proc R Soc Lond Ser B-Biol Sci. 1938;126:136–95.
47. Serrancolí G, Kinney AL, Fregly BJ, Font-Llagunes JM. Neuromusculoskeletal model calibration significantly affects predicted knee contact forces for walking. J Biomech Eng. 2016;138: 081001.
48. An K, Takahashi K, Harrigan T, Chao E. Determination of muscle orientations and moment arms. J Biomech Eng. 1984;106:280–2.
49. Taylor R. Interpretation of the correlation coefficient: a basic review. J Diagn Med Sonogr. 1990;6:35–9.
50. Shourijeh MS, Fregly BJ. Muscle synergies modify optimization estimates of joint stiffness during walking. J Biomech Eng. 2020;142: 011011.
51. Tresch MC, Cheung VCK, d'Avella A. Matrix factorization algorithms for the identification of muscle synergies: evaluation on simulated and experimental data sets. J Neurophysiol. 2006;95:2199–212.
52. Torres-Oviedo G, Ting LH. Muscle synergies characterizing human postural responses. J Neurophysiol. 2007;98:2144–56.
53. Sartori M, Gizzi L, Lloyd DG, Farina D. A musculoskeletal model of human locomotion driven by a low dimensional set of impulsive excitation primitives. Front Comput Neurosci. 2013;7.
54. Neptune RR, Clark DJ, Kautz SA. Modular control of human walking: a simulation study. J Biomech. 2009;42:1282–7.
55. Trinler U, Leboeuf F, Hollands K, Jones R, Baker R. Estimation of muscle activation during different walking speeds with two mathematical approaches compared to surface EMG. Gait Posture. 2018;64:266–73.
56. Ajiboye AB, Weir RF. Muscle synergies as a predictive framework for the EMG patterns of new hand postures. J Neural Eng. 2009;6: 036004.
57. Gurchiek RD, Ursiny AT, McGinnis RS. A gaussian process model of muscle synergy functions for estimating unmeasured muscle excitations using a measured subset. IEEE Trans Neural Syst Rehabil Eng. 2020;28:2478–87.
58. Gonzalez-Vargas J, Sartori M, Dosen S, Torricelli D, Pons JL, Farina D. A predictive model of muscle excitations based on muscle modularity for a large repertoire of human locomotion conditions. Front Comput Neurosci. 2015;9:114.
59. Thelen DG, Anderson FC, Delp SL. Generating dynamic simulations of movement using computed muscle control. J Biomech. 2003;36:321–8.
60. Thelen DG, Anderson FC. Using computed muscle control to generate forward dynamic simulations of human walking from experimental data. J Biomech. 2006;39:1107–15.
61. Roelker SA, Caruthers EJ, Hall RK, Pelz NC, Chaudhari AMW, Siston RA. Effects of optimization technique on simulated muscle activations and forces. J Appl Biomech. 2020;36:259–78.
62. Davico G, Lloyd DG, Carty CP, Killen BA, Devaprakash D, Pizzolato C. Multi-level personalization of neuromusculoskeletal models to estimate physiologically plausible knee joint contact forces in children. Biomech Model Mechanobiol. 2022;21:1873–86.

63. De Groote F, Falisse A. Perspective on musculoskeletal modelling and predictive simulations of human movement to assess the neuromechanics of gait. Proc R Soc B. 2021;288:20202432.
64. Bersani A, Davico G, Viceconti M. Modeling human suboptimal control: a review. J Appl Biomech. 2023;39:294–303.
65. Song S, Kidziński Ł, Peng XB, Ong C, Hicks J, Levine S, et al. Deep reinforcement learning for modeling human locomotion control in neuromechanical simulation. J NeuroEngineering Rehabil. 2021;18:126.
66. Latash ML. Fundamentals of motor control. Academic Press; 2012.
67. Falisse A, Serrancolí G, Dembia CL, Gillis J, Jonkers I, De Groote F. Rapid predictive simulations with complex musculoskeletal models suggest that diverse healthy and pathological human gaits can emerge from similar control strategies. J R Soc Interface. 2019;16:20190402.
68. Meyer AJ, Eskinazi I, Jackson JN, Rao AV, Patten C, Fregly BJ. Muscle synergies facilitate computational prediction of subject-specific walking motions. Front Bioeng Biotechnol. 2016. https://doi.org/10.3389/fbioe.2016.00077.
69. De Groote F, Kinney AL, Rao AV, Fregly BJ. Evaluation of direct collocation optimal control problem formulations for solving the muscle redundancy problem. Ann Biomed Eng. 2016;44:2922–36.
70. Domire ZJ, Challis JH. The influence of an elastic tendon on the force producing capabilities of a muscle during dynamic movements. Comput Methods Biomech Biomed Engin. 2007;10:337–41.
71. Fregly BJ, Besier TF, Lloyd DG, Delp SL, Banks SA, Pandy MG, et al. Grand challenge competition to predict in vivo knee loads: grand challenge competition. J Orthop Res. 2012;30:503–13.
72. Ao D, Li G, Shourijeh MS, Patten C, Fregly BJ. EMG-Driven Musculoskeletal Model Calibration With Wrapping Surface Personalization. IEEE Trans Neural Syst Rehabil Eng. 2023;31:4235–44.

Publisher's Note