Comparison of Synergy Extrapolation and Static Optimization for Estimating Multiple Unmeasured Muscle Activations during Walking

Di, Ao

Institute of Medical Research,

Northwestern Polytechnical University,

Xi'an, Shaanxi, China

Phone: +1 (86) 18292866101

aodi@nwpu.edu.cn

Benjamin, J. Fregly

Department for Mechanical Engineering,

Rice University,

Houston, Texas, USA

Phone: +1 (713) 348-4906

Fax: (713) 348-5423

fregly@rice.edu

Corresponding author: Di, Ao

Abstract

Background: Calibrated electromyography (EMG)-driven musculoskeletal models can provide great 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 unclear.

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) 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, SynX 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) compared to SO. 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.

Key words: EMG-driven model, Synergy extrapolation, Static optimization, Model personalization, Muscle force, Muscle activation, Stroke

Background

Muscle forces are essential for maintaining body posture and engaging in functional activities. Knowledge of the forces exerted by individual muscles is crucial for understanding the internal biomechanical mechanisms and motor control involved in human movement [1–3]. More importantly, knowledge of muscle forces could be useful for 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* directly using dynamometers or indirectly using inverse dynamics, muscle forces cannot currently be measured easily *in vivo*, though ongoing research is seeking to develop new experimental methods that can measure muscle or tendon forces *in vivo* during human movement [11,12]. Unfortunately, these research efforts have been hindered by technical challenges, high cost, and ethical considerations[11,12], motivating the search for computational methods that can enhance our knowledge of muscle forces.

Musculoskeletal modeling enables computational estimation of unmeasurable or difficult to measure internal biomechanical quantities, such as muscle forces and joint contact forces, that influence human movement generation. The estimation process uses musculoskeletal computer models that represent the bones, muscles, joints, neural control, and external forces specific to the subject and task being modeled [13–15]. These computer models typically employ a geometric model of the musculoskeletal system actuated by Hill-type muscle-tendon models [16]. The control inputs to these muscle-tendon models are either muscle excitations, which are equivalent

to processed experimental electromyographic (EMG) data, or muscle activations, which are muscle excitations that have been time delayed and passed through an activation dynamics model[17]. In addition to estimating unmeasurable time-varying internal quantities (e.g., muscle activations and forces), musculoskeletal modeling can be used to estimate unmeasurable time-invariant model parameter values (e.g., optimal muscle fiber length, tendon slack length) that have a significant influence on muscle force generation [18].

The two computational methods most commonly employed for estimating muscle activations and forces using a musculoskeletal model are EMG-driven modeling [6,19–25] and static optimization (SO) [26–32]. Both methods utilize nonlinear optimization to resolve the "muscle redundancy problem" [33] (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 (Table 1). For EMG-driven modeling, the design variables are time-invariant model parameter values (i.e., EMG scale factors, electromechanical delays, activation dynamics parameter values, Hill-type muscle-tendon model parameter values), the cost function minimizes the sum of squares of errors between model and experimental joint moments, the constraints bound muscle activations to be less than or equal to one, and the optimization problem is solved over all time frames together. For SO, the design variables are time-varying muscle activations, the cost function typically minimizes the sum of squares of muscle activations [26,34], the constraints enforce no errors between model and experimental joint moments in addition to bounds on muscle activations, and the optimization problem is solved for each time frame separately.

These optimization problem formulation differences, which stem from using experimental muscle excitations as inputs for EMG-driven modeling but not SO, have important implications for the capabilities and limitations of both methods. 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, minimization of this joint moment mismatch allows for calibration of musculoskeletal model parameter values when the optimization is performed over all time frames together. In contrast, since SO finds muscle activations that make model joint moments match experimental joint moments perfectly, 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 [14,15], 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 implemented and performed, SO remains the most commonly used computational method for estimating muscle activations and forces.

Table 1: Comparison of optimization problem formulations and solutions for EMG-driven modeling and static optimization in their most fundamental forms.

	EMG-driven Modeling	**Static Optimization**
Design Variables	EMG scale factors Electromechanical delays Activation dynamics parameters Muscle-tendon model parameters	Muscle activations
Cost Function	Min sum of squares of errors between model and experimental joint moments	Min sum of squares of muscle activations
Constraints	Muscle activations ≤ 1	Muscle activations ≤ 1 Errors between model and experimental joint moments = 0
Experimental Inputs	Muscle excitations	Joint kinematics

	Joint kinematics Joint moments	Joint moments
Solution Process	All time frames together	Each time frame separately
Solution Quantities	Muscle excitations Muscle activations Muscle forces	Muscle activations Muscle forces

Although EMG-driven modeling possesses the advantages noted above and produces physiologically reasonable estimates of muscle activations and forces [24], missing EMG data from muscles that contribute significantly to a measured movement has limited the adoption of EMG-driven modeling for routine clinical gait analysis and biomechanical research. This issue is the result of two practical challenges. First, surface electrodes are incapable of measuring EMG signals from deep muscles. Despite their non-invasive nature and easy application, surface electrodes are unable to measure EMG signals from important deep muscles that contribute significantly to joint moments, such as the iliacus and psoas muscles during walking. 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 close to 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 [25,29]. To address the issue of missing EMG signals when performing EMG-driven modeling,

147 researchers either exclude muscles with missing EMG data from the musculoskeletal model
148 [28,35], include such muscles in the model but assume that they generate only passive force [25],
149 or include such muscles and use SO to estimate the associated muscle activations [28,29].

150 To provide a better alternative for addressing missing EMG signals, researchers have recently
151 developed a modified EMG-driven modeling approach called "Synergy Extrapolation" (SynX)
152 that uses muscle synergy concepts to estimate missing muscle excitation data [36–38]. The
153 theoretical basis for SynX is that a large number (e.g., 8 or 16) of experimentally measured muscle
154 excitations can be represented by a smaller number (e.g., 4 or 5) of muscle synergies composed of
155 time-varying synergy excitations and associated time-invariant synergy vectors, where the weights
156 in each synergy vector define how the associated synergy excitation contributes to all muscle
157 excitations. The synergy excitations provide information about the timing of muscle contractions,
158 while the synergy vectors provide information about the coordination of muscle contractions.
159 Given 16 experimental muscle excitations, if a lower dimensional set of 4 or 5 muscle synergies
160 are calculated using either all 16 excitations or a subset of 8 carefully selected excitations, the
161 resulting synergy excitations will be almost the same in both cases [36]. This observation
162 demonstrates that the time-varying synergy excitations extracted from the first 8 muscle
163 excitations can be used as basis functions for constructing the remaining 8 muscle excitations.

164 Based on this observation, the historical development of SynX followed a logical sequence of three
165 studies. First, SynX was shown to work in theory for fitting eight missing muscle excitations using
166 synergy excitations extracted from eight measured muscle excitations [36]. For this study, 16
167 muscle excitations per leg measured experimentally from three subjects during walking were split
168 into two groups of eight "measured" and eight "missing" excitations, and synergy excitations
169 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 [37]. 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 [38]. The same sets of 16 experimental muscle excitations were again split into groups of 15 "measured" muscle excitations and one "missing" fine wire muscle excitation. A multi-objective optimization problem was designed to predict the 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 [39]. 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,

194 the applicability of EMG-driven modeling to research and clinical questions will be greatly
195 expanded.

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

210 # Methods

211 ***Experimental Data Collection***

212 Two previously published experimental walking datasets collected from a high-functioning
213 hemiparetic subject (S1, male, 1.70 m tall, mass 80.5 kg, age 79 years, right side hemiparesis,
214 lower extremity Fugl-Meyer Motor Assessment score of 32 out of 34) and a low-functioning
215 hemiparetic subject post-stroke (S2, male, 1.83 m tall, mass 88.5 kg, age 62 years, right side
216 hemiparesis, lower extremity Fugl-Meyer Motor Assessment score of 25 out of 34) were used for

this study [23,40]. 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

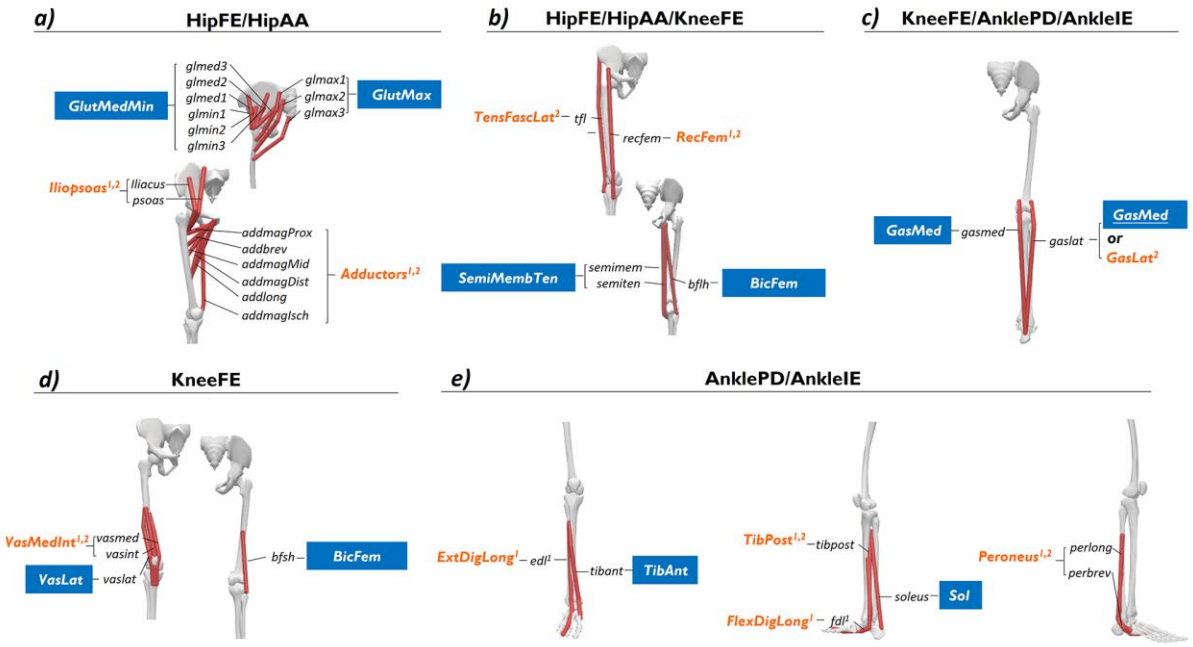

Figure 1 The assumption about "measured" and "unmeasured" EMG channels when performing SynX and SO as well as the associated muscles in the OpenSim model for each subject. The EMG channels assumed "measured" are denoted by blue boxes, while those assumed "unmeasured" are indicated by orange italic texts. The superscripts 1 and 2 represent the assumption of "unmeasured" EMG channels for subject S1 and S2, respectively. The muscles were categorized based on their actuating degrees of freedom (DOFs).

(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). 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. Surface EMG data were collected from the following superficial muscle groups (figure 1): 1) *GlutMax*, including gluteus maximus superior (*glmax1*), gluteus maximus middle (*glmax2*) and gluteus maximus inferior (*glmax3*); 2) *GlutMedMin*, including 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*, including semimembranosus (*semimem*) and semitendinosus(*semiten*); 4) *RecFem*, including rectus femoris (*recfem*); 5) *Bicfem*, including biceps femoris long head (*bflh*) and biceps femoris short head (*bfsh*); 6) *VasMedInt*, including vastus medialis (*vasmed*) and vastus intermedius (*vasint*); 7) *VasLat*, including vastus lateralis (*vaslat*); 8) TibAnt, including tibialis anterior (*tibant*); 9) *Peroneus*, including peroneus brevis (perbrev) and peroneus long (perlong); 10) *Sol*, including soleus (soleus). Additionally, fine-wire EMG data were collected from the following deep muscle groups (Fig.2): 1) *iliopsoas*, including iliacus (*iliacus*) and psoas (*psoas*); 2) *Adductors*, including 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*, including tibialis posterior (*tibpost*). Small differences existed in the EMG data collect from the two subjects. For the high-functioning subject (S1), a surface EMG signal (referred as *GasMed)* was also collected and expanded to medial gastrocnemius (*gasmed*) and lateral gastrocnemius (*gaslat*), and two fine-wire EMG signals (referred 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 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 [23].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" [23,41].

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 1000 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 [42], where *tf* is the period of the gait cycle. Data from ten gait cycles (five cycles per speed) per leg were randomly chosen to simultaneously 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 [43] 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 were performed using OpenSim 4.0 [44,45] to prepare the model for EMG-driven modeling with SynX. First, OpenSim model scaling was performed so that the generic model's anthropometry would more closely match that of each subject. Second, repeated OpenSim inverse kinematics (IK) analyses within a

nonlinear optimization were performed 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 [46]. 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

SynX and SO were both utilized to estimate muscle activations and forces, and the resulting estimates from both methods were compared to a "gold standard" for evaluation purposes. As illustrated in figure 2, both approaches take joint kinematics and associated musculoskeletal geometries (i.e., muscle-tendon lengths and moment arms) as inputs to estimate muscle activations, muscle forces, and net joint moments. Subsequently, the estimated predicted net joint moments are iteratively compared to the inverse dynamic joint moments through an optimization process, leading to the estimation of the time-varying muscle activations and forces for SynX and SO, as well as time-invariant musculoskeletal model and SynX-related parameter values for SynX.

Synergy Extrapolation Solution Process

The SynX solution process involved four tasks as summarized below.

- *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 excitations from measured muscles into activations of all muscles itself involved four distinct steps [37,38]. First, muscle excitations

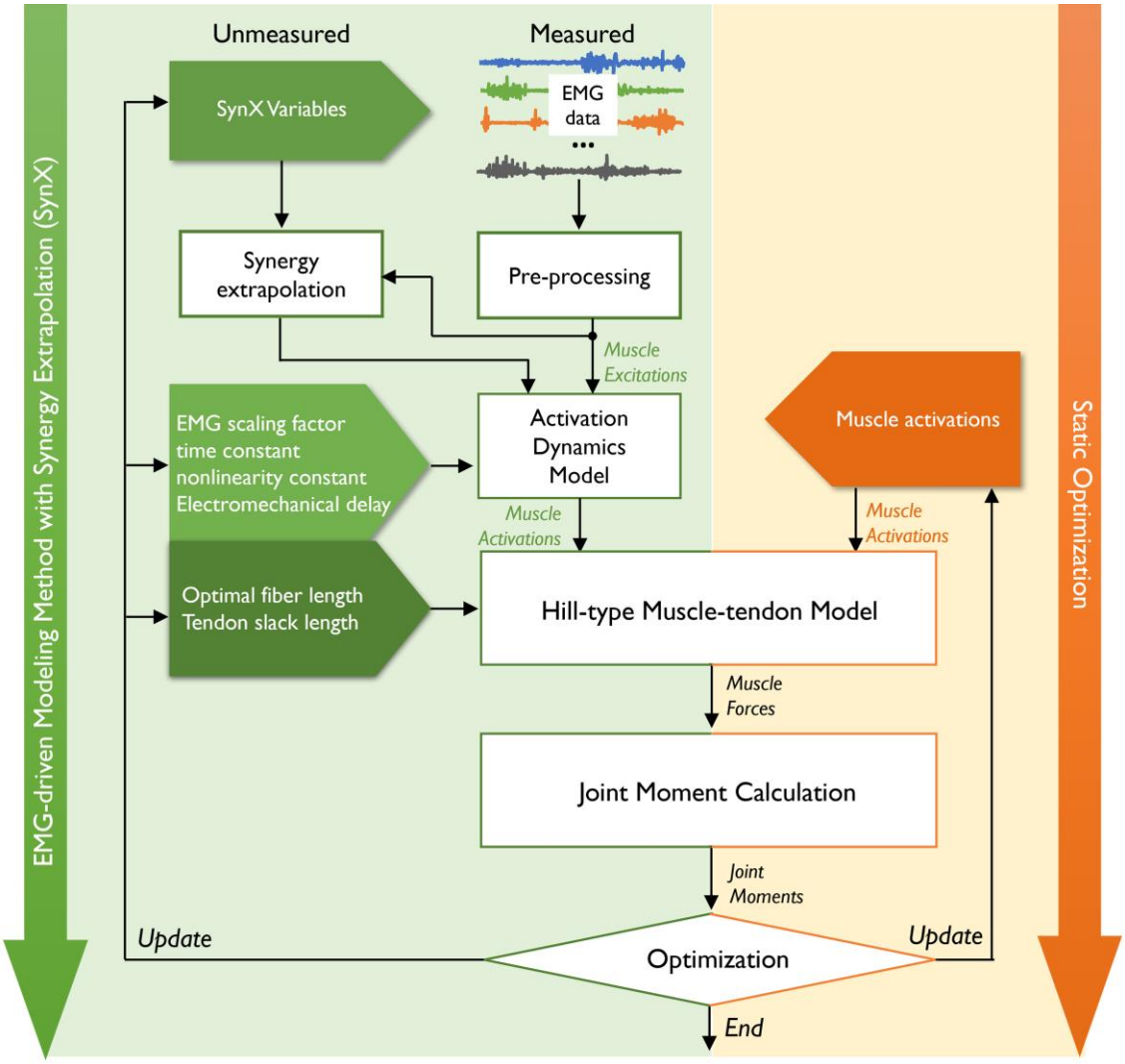

Figure 2 The workflow for EMG-driven modeling with SynX (left panel with a green background) and SO (right panel with an orange background) as performed in this study. Both methods employ experimental joint kinematics and moments as inputs and aim to determine muscle activations and forces in such a way that the predicted net joint moments from a musculoskeletal model closely match the experimental net joint moments calculated via inverse dynamics. However, there are notable differences in the optimization problem formulations for these two methods. In EMG-driven modeling with SynX, the design variables consist of time-invariant model parameter values and SynX variables, with the optimization problem being solved across all time frames together. Conversely, for SO, the design variables encompass time-varying muscle activations, typically utilizing model parameter values from scaled generic models or literature references, and the optimization problem is solved for each time frame separately. Subsequently, both techniques leverage the Hill-type muscle-tendon model to estimate muscle forces and their respective contributions to the joint moments.

300 $e_m^{musc}(t)$ for muscles with experimental EMG data were adjusted using a muscle-specific scale

301 factor ranging from 0.05 to 1, acknowledging that actual maximum activation levels tend to

302 surpass those observed experimentally during walking.

Second, muscle synergy analysis (MSA) was conducted on the scaled muscle excitations using principal component analysis (PCA) to extract a small number of muscle synergies, specifically six for the present study:

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

where $W_m(t)$ specifies the time-varying measured synergy excitations, H_m specifies the associated 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:

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

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 variables, which were all time-invariant and determined through an optimization process implemented within the EMG-driven model calibration process (Figure 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:

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

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 [47]:

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

where τ_{act} and τ_{dact} are activation and deactivation time constants. d specifies the electromechanical time delay.

Fourth, a nonlinear one-parameter transformation model was utilized to compute each associated muscle activation $a^{musc}(t)$ [48]:

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

where c_3 is an activation nonlinearity constant that characterizes the curvature of the relationship of each muscle. g_1 to g_5 are constant coefficients obtained by fitting published experimental data

from isometric contractions[48]. 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 the same set of design variables, encompassing SynX variables, EMG scale factors, electromechanical time delays, activation time constants, and activation nonlinearity constants, where further details are provided in section 2.3.1.4.

- *Muscle force estimation*

For the second task 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 [16,23,49] to predict the force generated by a given muscle-tendon actuator, m, which was formulated as (figure 2):

$$F^{musc}(t,\theta,\dot{\theta}) = F_o^m \cdot \left[a^{musc}(t) \cdot f_l^a(\tilde{l}^{musc}(t,\theta)) \cdot f_v^a(\tilde{v}^{musc}(t,\theta,\dot{\theta})) + f_l^p(\tilde{l}^{musc}(t,\theta)) \right] \cdot \cos\alpha$$

$$\tilde{l}^{musc}(t,\theta) = \frac{l^{mt}(t,\theta) - l_s^t}{l_o^m} \tag{6}$$

$$\tilde{v}^{musc}(t,\theta,\dot{\theta}) = \frac{v^{mt}(t,\theta,\dot{\theta})}{10 \cdot l_o^m}$$

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 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 [50]), l_o^m

355 denotes optimal muscle fiber length, and l_s^t denotes tendon slack length. These values (apart from
356 pennation angles, which were taken from the literature) were calibrated through an optimization
357 process or taken from the scaled OpenSim models. More details regarding the determination of
358 l_o^m and l_s^t values for each muscle force estimation method can be found below.

359 • *Joint moment calculation*

360 For the third task of the SynX solution process, model joint moments were calculated using the
361 forces for measured and unmeasured muscles found in the second task. Once the muscle forces
362 $F^{musc}(t,\theta)$ were estimated, their contributions to net joint moment at joint j were calculated using
363 the corresponding muscle moment arms:

364 $$M^{joint}(t,\theta,\dot{\theta}) = \sum F^{musc}(t,\theta,\dot{\theta}) \cdot r^{musc}(t,\theta) \quad (7)$$

365 $$r^{musc}(t,\theta) = -\frac{\partial l^{mt}(t,\theta)}{\partial \theta} \quad (8)$$

366 where $M^{joint}(t,\theta,\dot{\theta})$ is joint moment at time t, which is defined as the sum of contributions from
367 all spanning muscles $r^{musc}(t,\theta)$ is muscle moment arm for muscle m at time t, which was
368 defined as the negative of the partial derivative of muscle-tendon length $l^{mt}(t,\theta)$ with respect to
369 generalized coordinate θ [51]. The negative sign in Eq. 8 was implemented for consistency with
370 the OpenSim modeling environment. When utilizing SynX for estimating unmeasured muscle
371 excitations, net joint moments were computed with ($M_{res}^{joint}(t,\theta,\dot{\theta})$)and without ($M^{joint}(t,\theta,\dot{\theta})$)
372 incorporating residual excitations into the measured muscle excitations, as stipulated by the cost
373 function for EMG-driven model calibration.

- *Optimization problem formulation*

For the fourth task of the SynX solution process, the first three tasks were performed iteratively within a nonlinear optimization that adjusts three categories of design variables (see figure 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; 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 joint moment tracking errors and minimizing unmeasured and residual muscle activation magnitudes [38]:

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

where $M_{res}^{joint}(t,\theta,\dot{\theta})$ refers to model-predicted joint moments when residual muscle excitations are included in joint moment calculations, $M^{ID}(t,\theta,\dot{\theta})$ refers to inverse dynamic 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 [38]. Further details regarding initial guesses, upper/lower bounds for design variables, additional inequality constraints, and penalty terms can be found in previously published studies [23,37,38]. All optimization procedures were performed using MATLAB's "*fmincon*" 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 [26]:

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

The net joint moments for SO were estimated by multiplying the muscle forces obtained through substituting muscle activations $a^{musc}(t)$ into the Hill-type muscle-tendon model with the corresponding muscle moment arms, as depicted in Eqs. 6 through 8. In contrast to the EMG-

410 driven modeling method, the muscle activations estimated for SO were used directly as design
411 variables in the optimizations, which were solved individually for each time frame. Furthermore,
412 model parameter values were taken from scaled generic models or literature references rather than
413 being calibrated during the optimization process.

414 ***Synergy Extrapolation and Static Optimization Evaluation***

415 *Muscle selection heuristics*

416 Given 16 measured muscle excitations for each leg of both subjects, we had to decide which 8
417 muscle excitations would be treated as measured and which 8 would be held back and treated as
418 missing for SynX and SO evaluation purposes. A prior study [36] provided guidance for which
419 eight muscles to select as measured and which eight to select as missing so as to maximize
420 reconstruction accuracy for the eight missing muscle excitations. In that study, an investigation of
421 all possible combinations of eight measured and eight missing EMG signals yielded the following
422 muscle selection heuristic: 1) Choose muscles easily accessible by surface EMG electrodes; 2)
423 Choose most frequently occurring muscle in the top 10% of muscle combinations that yielded the
424 highest SynX accuracy from each primary lower extremity function group; 3) Choose two hip/knee
425 biarticular muscles at minimum; 4) Choose remaining most frequent muscles to fill eight muscle
426 combinations. Following the observation of SynX performance, our muscle selection heuristic,
427 given a limited number of eight EMG channels, indicated that researchers should collect surface
428 EMG data from commonly chosen uniarticular and biarticular flexor and extensor muscles from
429 each major muscle group, as illustrated in figure 1. The selected uniarticular muscles included a
430 hip extensor (*GlutMax*), a knee extensor (*VasLat* considered preferable over *VasMed*), an ankle
431 plantarflexor (*Sol*), and an ankle dorsiflexor (*TibAnt*). Uniarticular hip flexor (*Iliopsoas*) was
432 omitted due to the difficulty in measuring these muscles reliably 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. Even the collection of EMG data from less commonly chosen muscles spanning all three joints (*Adductors*, *tfl*, and *Peroneus*) may facilitate to improve the estimation accuracy to some extent due to the unique stabilizing roles they play in the frontal plane, they were excluded from the "measured" muscles that was attributed to the difficulty in reliable surface EMG measurement of these muscles.

Synergy extrapolation methodological choices

Table 2 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

Implementation of SynX requires making several methodological choices that can impact the accuracy of estimated muscle activations and forces. Previous studies investigated the influence of various methodological factors on SynX performance [37,38], 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. We 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.

449 In contrast, non-negative matrix factorization (NMF) did not achieve acceptable prediction
450 accuracy. Additionally, for the same number of synergies, employing trial-specific unmeasured
451 synergy vector weights and speed-specific residual synergy vector weights resulted in optimal
452 SynX performance for both subjects in terms of estimation accuracy and computational efficiency.
453 Notably, EMG normalization had no significant impact on SynX performance. Thus, the key
454 methodological choices for SynX in this study were informed by insights from prior research, as
455 detailed in Table 1.

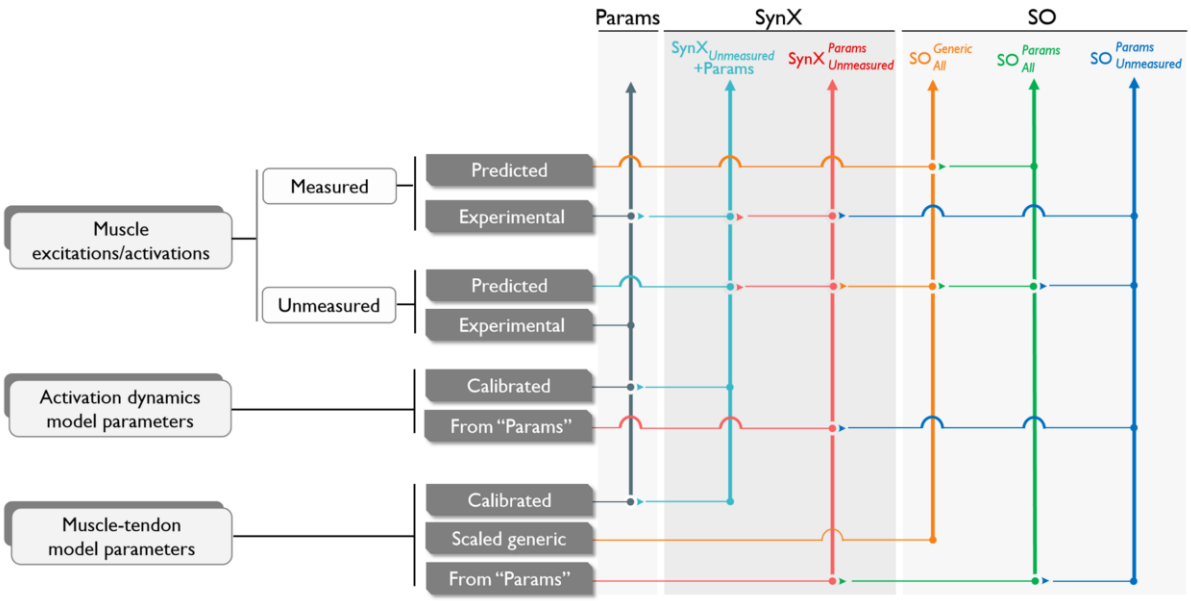

Figure 3 Summary of six optimizations performed in this study, which included two optimizations using SynX to predict unmeasured muscle excitations (*termed* $SynX_{Unmeasured} + Params$ and $SynX^{Params}_{Unmeasured}$), three optimizations using static optimization (SO) to predict unmeasured muscle activations (termed $SO^{Generic}_{All}$, SO^{Params}_{All} and $SO^{Params}_{Unmeasured}$) and one "gold standard" optimization using 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 during model calibration. In each column of the optimizations, the arrows indicate whether each group of muscle excitations or activations were predicted or obtained experimentally. Moreover, the arrows indicate which values were used if the model parameters were not calibrated through optimization. The term "Scaled Generic" denotes the scaled generic model parameter values, while "From Params" refers to the model parameter values derived from the "gold standard (*Params*)" optimization.

Optimization problems

In this study, three primary objectives were pursued. Firstly, the study aimed to evaluate the performance of SynX when treating multiple channels of EMGs (i.e., eight) as "unmeasured". Secondly, it sought to compare the estimates of muscle activations and forces from SynX with those from SO. Thirdly, the study also aimed to analyze the accuracy of estimated unmeasured muscle activations and forces for 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, calibrate model parameter values (figure 3). The first optimization problem utilized all 16 channels of EMG data to calibrate each EMG-driven musculoskeletal model, providing "gold standard" muscle activations and forces for evaluation (termed "$Params$"). The second optimization problem assessed the performance of SynX when multiple channels of EMG data (i.e., eight) were considered "unmeasured". This optimization problem calibrated EMG-driven models for each leg of each subject while simultaneously estimating missing muscle excitations using SynX, where activation dynamics model, muscle-tendon model, and SynX parameter values were calibrated concurrently (termed "$SynX_{Unmeasured} + Params$"). The third optimization problem used SO to estimate muscle activations for all muscles using muscle-tendon model parameter values taken from scaled generic OpenSim models (termed "$SO_{All}^{Generic}$"), representing the most commonly formulated SO method. The accuracy of estimated muscle activations and forces was further quantitatively compared to those from the optimization "$SynX_{Unmeasured} + Params$" to assess the estimation performance of both SynX and SO. The fourth optimization problem employed SynX to estimate the unmeasured muscle excitations within a well-calibrated EMG-driven model (termed "$SynX_{Unmeasured}^{Params}$"), utilizing the

model parameter values found in the "gold standard (*Params*)" optimization. The fifth and sixth optimization problems utilized SO to estimate muscle activations for all muscles (termed "SO_{All}^{Params}") and only unmeasured muscles (termed "$SO_{Unmeasured}^{Params}$") using the model parameter values from the "gold standard (*Params*)" optimization, rather than scaled generic values. when performing the fourth and sixth optimizations of $SynX_{Unmeasured}^{Params}$ and $SO_{Unmeasured}^{Params}$ to estimate only unmeasured muscle excitations/activations, the muscle activations of the measured muscles were determined from the "gold standard (*Params*)" optimization.

Evaluation Metrics and Statistical Analyses

Several common evaluation metrics were utilized to evaluate the ability of SynX and SO to estimate muscle activations and muscle forces for unmeasured muscles and 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 interpreted based on [52], categorized as weak ($r < 0.35$), moderate ($0.35 < r \leqslant 0.67$), strong ($0.67 < r \leqslant 0.9$), or very strong ($r \leqslant 0.9$). 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 "$SynX_{Unmeasured}^{Params}$." 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 also 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 (optimization problem "*Params*"). This comparison was conducted to assess the accuracy of estimated muscle activations (figure 4 and 5, table 3). Initially, during the simultaneous calibration of EMG-driven model parameters, SynX effectively estimated unmeasured muscle activations, demonstrating low RMSE values (≤ 0.17, $= 0.08 \pm 0.06$) and moderate or strong correlation r values (≥ 0.38, $= 0.55 \pm 0.13$) across most muscles for optimization "$SynX_{Unmeasured} + Params$" (figure 3 and table 3). Notably, among these unmeasured muscles, SynX exhibited superior performance for the superficial muscles (e.g. rectus femoris, lateral gastrocnemius and vastus intermedius) compared to the deep-located muscles (e.g. iliacus, extensor digitorum longus and tibialis posterior) in terms of both shape and magnitude. However, the estimates for adductor muscles (RMSE = 0.01, $r \geq 0.43$) and flexor digitorum longus (RMSE = 0.05, $r = 0.92$) that typically rely on fine-wire electrodes for EMG collection remained reasonably accurate.

523

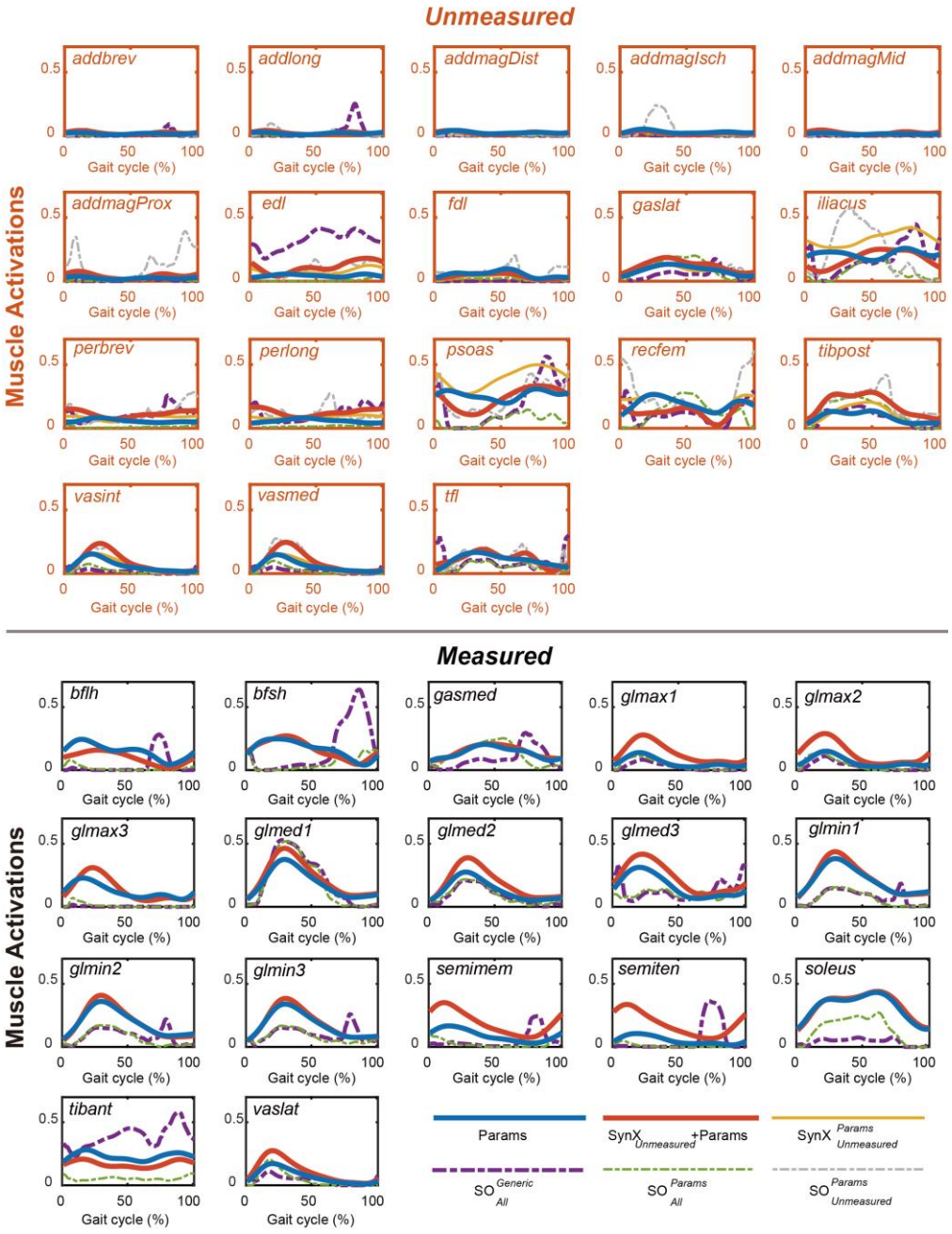

Figure 4 Average muscle activations for the "unmeasured" muscles (upper) and the "measured" muscles (lower) across calibration trials, legs and subjects from "*Params*" optimization (blue solid curves), SynX-based optimizations ($SynX_{Unmeasured} + Params$:red solid curves and $SynX^{Params}_{Unmeasured}$: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

524 Second, SynX combined with model optimization "$SynX_{Unmeasured} + Params$" produced significantly

525 more accurate predictions of unmeasured muscle activations compared to the standard SO used

526 within optimization "$SO_{All}^{Generic}$". This finding was evident in terms of both magnitude (characterized

527 by RMSE values, $p \leq 0.05$) and shape (characterized by correlation r values, $p \leq 0.05$) across

528 unmeasured muscles and subjects (figure 4 and 5, table 3). Even for muscles with relatively low

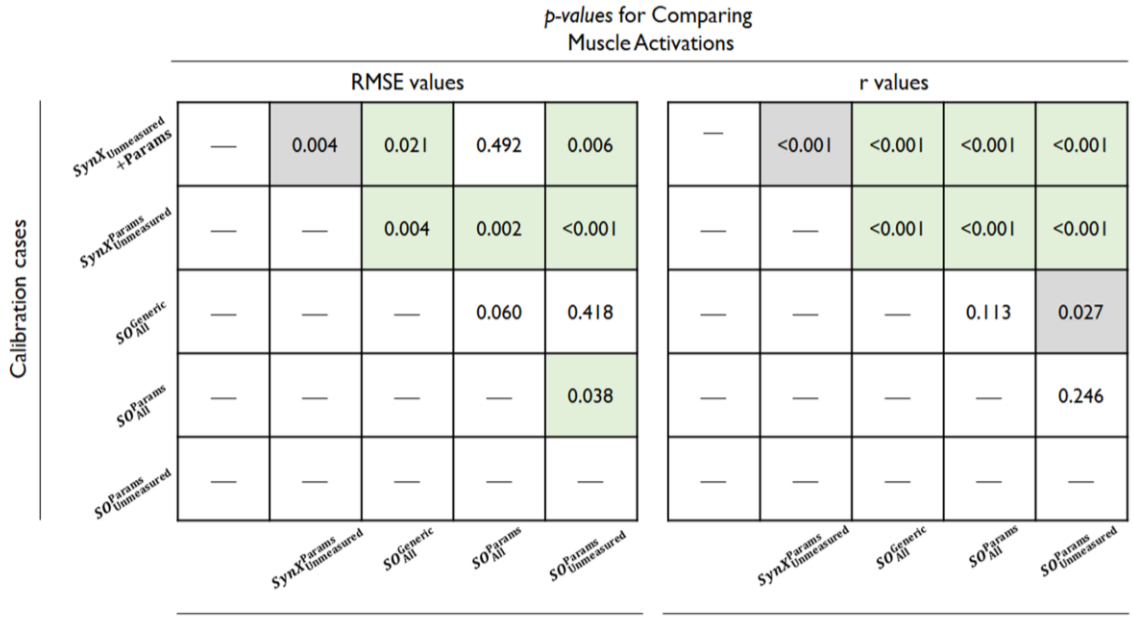

Figure 5 *p-values* obtained from paired *t-test* used to compare the estimation accuracy of muscle activations, as indicated by RMSE values (left) and *r* values (right), between different optimizations. Initially, RMSE and *r* values were calculated between the experimental ("*Params*" optimization) and estimated muscle activations from various optimizations, with the results across all calibration trials, legs, and subjects being concatenated and displayed in table 3. Subsequently, the RMSE and *r* values from each optimization were individually 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 at $p < 0.05$. 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 it 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 it for the x-axis optimization.

529 estimation accuracy using both methods, such as iliacus, psoas and extensor digitorum longus,

530 SynX outperformed SO in reproducing the shape and magnitude of unmeasured muscle activations

531 (table 3). Moreover, SO exhibited weak correlation ($r \leq 0.35$) in the muscle activation predictions

for the majority of unmeasured muscles within optimization "$SO_{All}^{Generic}$", 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, whereas SO exhibited discontinuities and underestimated muscle activations, featuring abrupt changes (see figure 3).

Third, both SynX-based and SO-based methods were sensitive to the level of musculoskeletal model personalization (figure 4 and 5, table 3). For SynX, employing a well-calibrated EMG-driven model for optimization "$SynX_{Unmeasured}^{Params}$" resulted in lower RMSE for the estimation of unmeasured muscle activations (RMSE = 0.05 ± 0.05) compared to optimization "$SynX_{Unmeasured} + Params$", when SynX variables and EMG-driven model parameter values were calibrated simultaneously. The estimated unmeasured muscle activations from optimization "$SynX_{Unmeasured}^{Params}$" exhibited strong or very strong correlations with those generated from optimization "$Params$", with the exception of the extensor digitorum longus (r = 0.42). For SO, well-calibrated model parameter values in optimization "SO_{All}^{Params}" led to more accurate estimation of unmeasured muscle activations compared to using scaled generic model parameter values in optimization "$SO_{All}^{Generic}$", although the difference was not substantial.

Last, with the use of a well-calibrated EMG-driven model to estimate unmeasured muscle activations only, SynX in optimization case "$SynX_{Unmeasured}^{Params}$" demonstrated more accurate and reliable estimates compared to SO in optimization "$SO_{Unmeasured}^{Params}$". This was evidenced by remarkably lower RMSE values ($p \leq 0.05$) and higher correlation r values ($p \leq 0.05$) (figure 4 and

552 5, table 3). Similar to all SO-estimated muscle activations, the estimates obtained from SO in
553 optimization "$SO^{Params}_{Unmeasured}$" showed general underestimation and abrupt changes.

554 ***Muscle Forces***

555 The study quantitatively evaluated the magnitude and shape similarity of muscle forces estimated
556 using various optimization methods and those estimated from EMG-driven modeling calibration
557 using a full set of EMGs (optimization "*Params*") through RMSE and correlation *r* measurements.

558 First, SynX provided reasonably accurate and reliable estimation of muscle forces that closely
559 matched those obtained from "*Params*" optimization for both subjects, as shown in figure 6 and 7,
560 table 4. In terms of unmeasured muscle forces, the RMSE values using SynX for optimization
561 "$SynX_{Unmeasured} + Params$" (= 101.3 ± 0.13) were significantly smaller (p = 0.028) than those using
562 standard SO for optimization "$SO^{Generic}_{All}$" (= 174.4 ± 174.4). Furthermore, the correlation *r* values
563 between the muscle forces estimated by SynX and those provided by the full EMG-driven model
564 calibration was moderate or higher across all unmeasured muscles. Conversely, for SO, the
565 correlation was generally weak for most muscles, except for moderate correlations observed for
566 rectus femoris (r = 0.42), lateral gastrocnemius (r = 0.51), tibialis posterior (r = 0.48), and extensor
567 digitorum longus (r = 0.58) (table 4).

568 Second, model personalization had considerable influence on the accuracy of estimating muscle
569 forces for both SynX and SO, as detailed in figure 6 and 7, and table 4. SynX demonstrated notably
570 improved estimation accuracy in terms of both shape ($p \leq 0.05$) and magnitude ($p \leq 0.05$) when
571 incorporating a well-calibrated EMG-driven model for optimization "$SynX^{Params}_{Unmeasured}$", as compared
572 to simultaneous calibration of SynX variables and EMG-driven model parameters for optimization

573

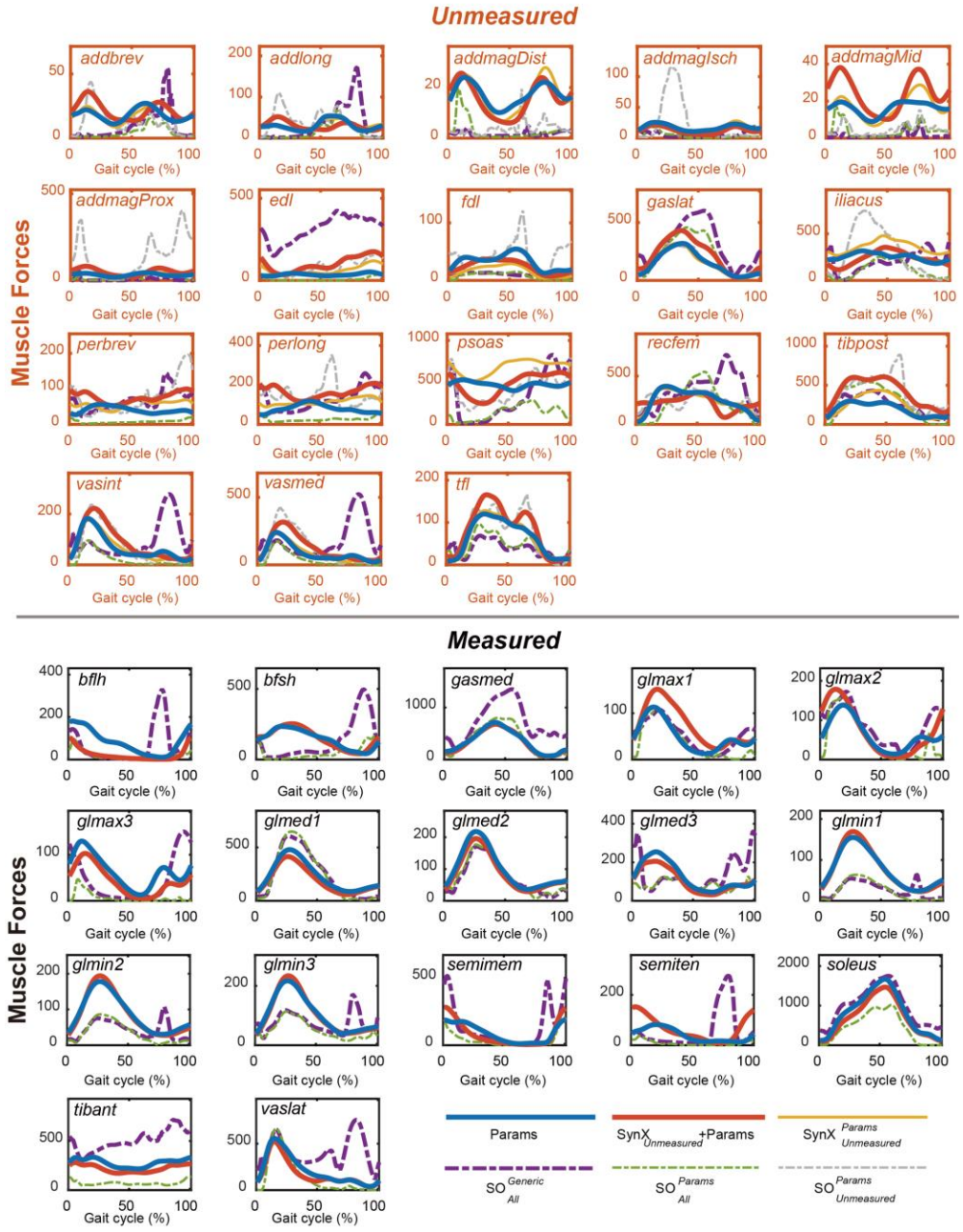

Figure 6 Average muscle forces for the "unmeasured" muscles (upper) and the "measured" muscles (lower) across calibration trials, legs and subjects from "*Params*" optimization (blue solid curves), SynX-based optimizations ($SynX_{Unmeasured} + Params$:red solid curves and $SynX^{Params}_{Unmeasured}$: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

574 "$SynX_{Unmeasured} + Params$". Similarly, SO benefited from well-calibrated model parameter values in

575 achieving more accurate estimation of unmeasured muscle forces, leading to significantly different

576 correlation *r* values between optimizations "$SO_{All}^{Generic}$" and "SO_{All}^{Params}", while the RMSE values

577 remained statistically comparable between them.

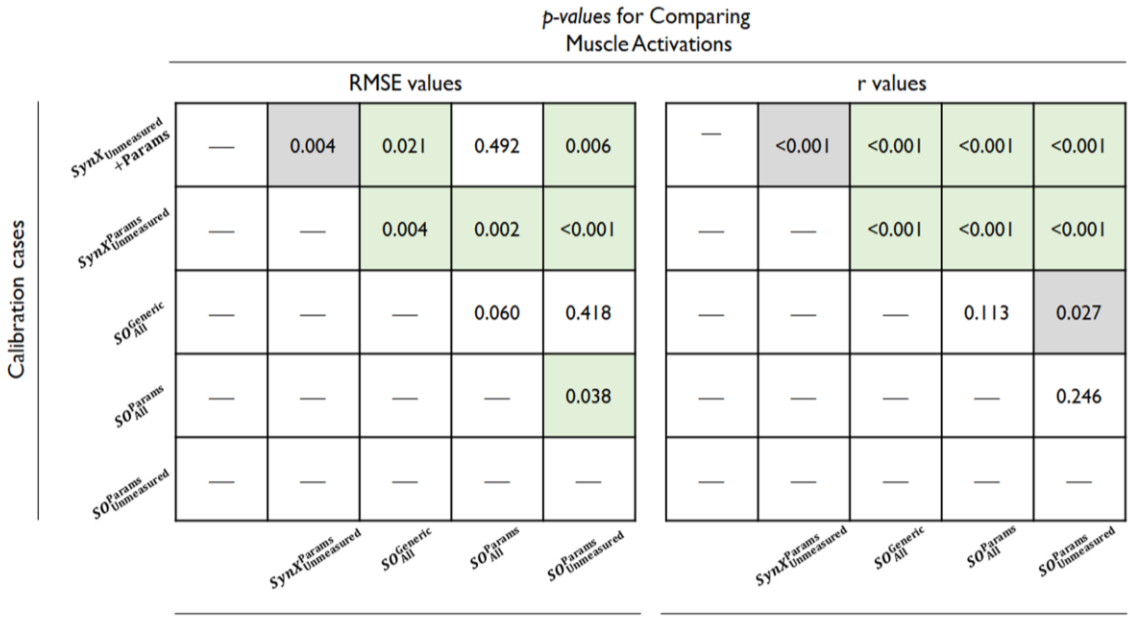

Figure 7 ***p-values*** **obtained from paired** ***t-test*** **used to compare the estimation accuracy of muscle forces,** as indicated by RMSE values (left) and *r* values (right), between different optimizations. Initially, RMSE and *r* values were calculated between the experimental ("*Params*" optimization) and estimated muscle activations from various optimizations, with the results across all calibration trials, legs, and subjects being concatenated and displayed in table 4. Subsequently, the RMSE and *r* values from each optimization were individually 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 at $p < 0.05$. 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 it 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 it for the x-axis optimization.

578 Finally, with model parameter values determined through a full EMG-driven calibration, "*Params*",

579 SynX predicted unmeasured muscle forces more accurately and reliably within optimization

580 "$SynX_{Unmeasured}^{Params}$" than SO within optimization "$SO_{Unmeasured}^{Params}$". This was evidenced by significantly

581 lower RMSE values ($p \leq 0.05$) and higher correlation *r* values ($p \leq 0.05$).

582 ***Joint Moments***

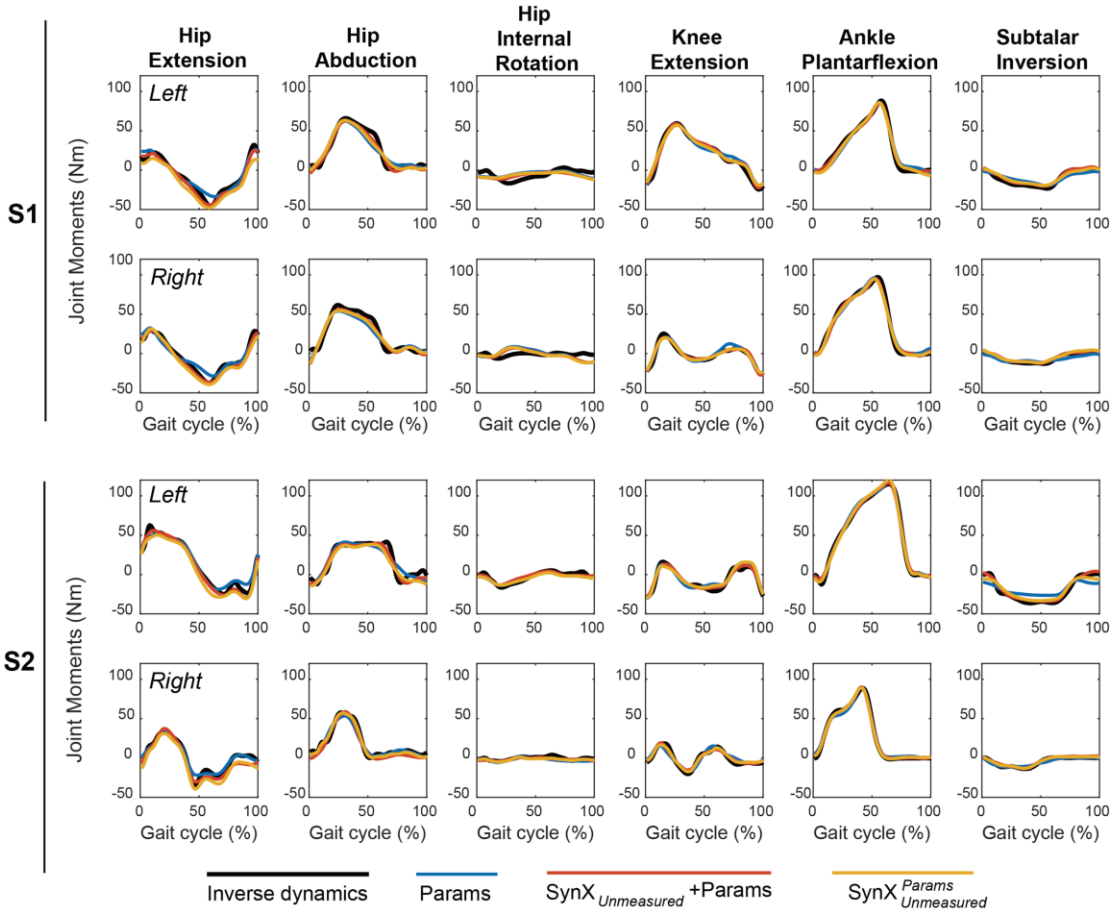

Figure 8 Average joint moments across calibration trials from "*Params*" optimization (blue solid curves) and SynX-based optimizations ($SynX_{Unmeasured} + Params$: red solid curves and $SynX^{Params}_{Unmeasured}$: yellow solid curves). Data are reported for the complete gait cycle, where 0% indicates initial heel-strike and 100% indicates subsequent heel-strike.

583 Compared with the MAE values between model-predicted and experimental ID joint moments
584 from "*Params*" optimization, the MAE values for SynX-based optimization, with simultaneous
585 calibration of model parameters for optimization "$SynX_{Unmeasured} + Params$", were consistently lower
586 across all DoFs. On average, the MAE values were lower by 1.54 Nm for *HipFE*, 1.74 Nm for
587 *HipAA*, 0.37 Nm for *HipRot*, 1.88 Nm for KneeFE, 2.05 Nm for *AnklePD* and 2.27 Nm *AnkleIE*
588 across the legs of both subjects (see figure 8 and table 5). When the EMG-driven model parameter
589 values were fixed at the values determined from "*Params*" optimizations, the SynX-estimated joint

moments for optimization "$SynX^{Params}_{Unmeasured}$" showed significantly lower MAE values only for all DoFs ($p \leq 0.05$), except for *HipRot* ($p = 0.078$). It is important to note that inherent to the formulation of optimization, the joint moment matching errors were exceptionally small (MAEs ≤ 0.001) for all three SO-based optimizations, "$SO^{Generic}_{All}$", "SO^{Params}_{All}", and "$SO^{Params}_{Unmeasured}$".

Model Parameters

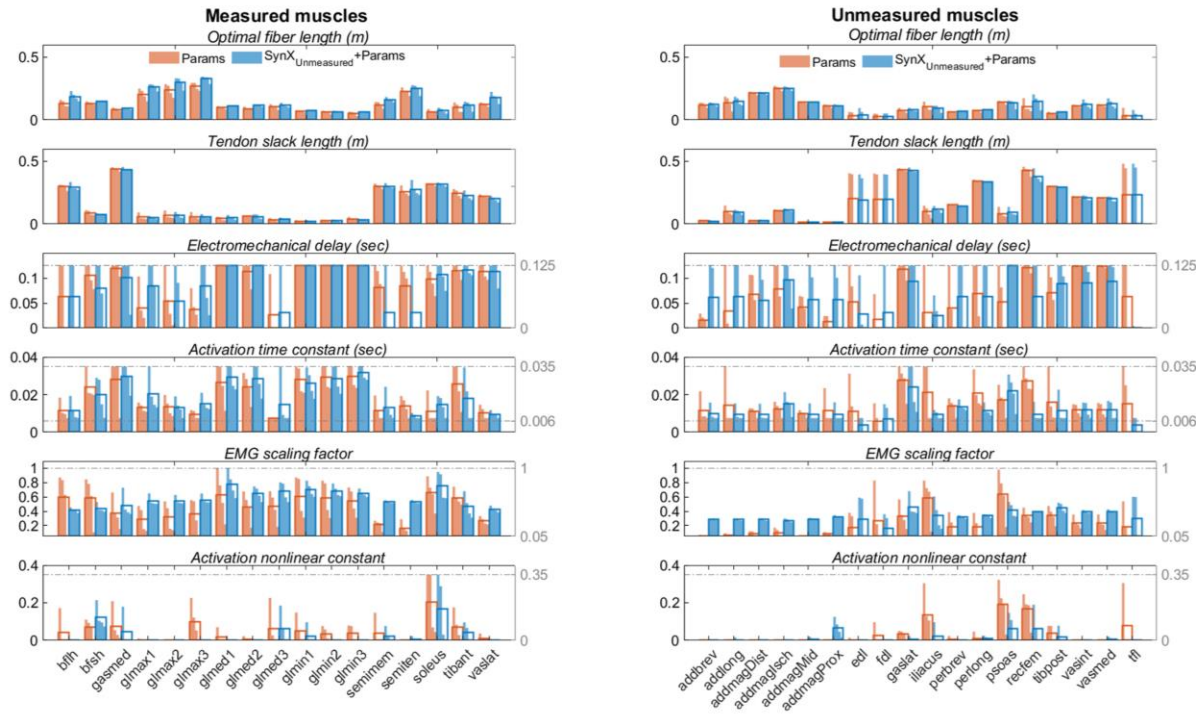

Figure 9 EMG-driven model parameters of two legs of both subjects from "gold standard (*Params*)" optimization (in blue) and "$SynX_{Unmeasured} + Params$" optimization (in orange). The upper and lower bounds for each of the four activation dynamics model parameters during optimization have been indicated 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.

In general, the four activation dynamic model parameters and the two Hill-type muscle-tendon model parameters from optimization showed a high degree of similarity between "$SynX_{Unmeasured} + Params$" optimization and "$Params$" optimization for the measured muscles (refer to figure 9, left panel). Conversely, for the unmeasured muscles, when simultaneously tuning SynX

599 variables, "$SynX_{Unmeasured} + Params$" still maintained the pattern defined by the parameter magnitudes
600 of the optimal fiber length and tendon slack length for each model parameter (refer to figure 9,
601 right panel). However, substantial discrepancies in the four activation dynamic model parameters
602 were observed for the unmeasured muscles between the SynX approach for "$SynX_{Unmeasured} + Params$"
603 and the full EMG-driven model calibration for "$Params$".

604 ## Discussion

605 This study extended the capability of synergy extrapolation (SynX) to enable the concurrent
606 estimation of a large number of unmeasured muscle excitations and calibration of an EMG-driven
607 model. The approach was developed and evaluated using gait datasets collected from two post-
608 stroke subjects performing treadmill walking at self-selected and fastest-comfortable speeds. EMG
609 signals measured bilaterally from eight muscles were treated as "unmeasured" and estimated using
610 the synergy information extracted from another eight muscles treated as "measured." The muscle
611 activations, forces, and model parameter values for the unmeasured muscles were quantitatively
612 compared to "gold standard" values obtained when all 16 channels of EMG data were used to
613 calibrate an EMG-driven musculoskeletal model for each leg of each subject. The results revealed
614 that the estimated unmeasured muscle activations and forces were reasonably accurate and reliable
615 in term of both shaped and magnitude (figures 4 and 6, Tables 3 and 4). Moreover, Hill-type
616 muscle-tendon model parameter values for both unmeasured and measured muscles, including
617 optimal fiber length and tendon slack length, exhibited a high level of agreement with the "gold
618 standard" model parameter values (figure 9). When SO estimates of unmeasured muscle
619 activations and forces were compared with SynX estimates, the SynX results were more accurate
620 and realistic than those from SO (figures 4 and 6, tables 3 and 4), which contained abrupt changes
621 and tended to underestimate the unmeasured muscle quantities. When the sensitivity of estimated

622 unmeasured muscle activations and forces to the level of model personalization was investigated,
623 both SynX and SO generated substantially more accurate estimates when utilizing well-calibrated
624 muscle-tendon parameters. However, SynX demonstrated superior performance to SO in
625 estimating unmeasured muscle activations and forces when employing model parameter values
626 from full EMG-driven model calibration.

627 SynX has demonstrated superior performance over SO for estimating muscle activations and forces
628 for several important reasons. First, by utilizing measured synergy excitations as building blocks,
629 it reduced the problem of finding unknown time-varying muscle excitations to identifying a small
630 number of unmeasured synergy vector weights. This led to a substantial reduction in the search
631 space for the optimization in comparison with SO-based approaches[37]. Second, unlike SO-based
632 approach, which solved a time frame of muscle activation at a time, the inherent constraints of
633 dependence between time frames in weighted synergy excitations resulted in smooth and
634 continuous estimated muscle activations, improving the physiological plausibility of the estimates.
635 Third, the time-invariance of unmeasured and residual synergy vector weights enabled a single-
636 layer optimization process, simultaneously achieving EMG-driven model personalization and
637 muscle activation estimation, which enhanced the accuracy of muscle force estimation. Fourth,
638 calibration of synergy-structured residual muscle excitations was integrated into SynX to enhance
639 the accuracy of predicted unmeasured muscle excitations. Unlike SO, where unmeasured muscle
640 activations, as design variables, inclined to deviate from experimental muscle excitations during
641 iterative adjustments for minimizing joint moment matching errors, SynX introduced residual
642 muscle excitations to account for joint moment matching errors, preventing predicted missing
643 muscle excitations from excessively compensating for joint moment prediction inaccuracy through
644 optimization. The addition of residual muscle excitations in turn prevented inaccuracy as a
645 consequence[38]. Fifth, the SynX-based methods did not require assumptions in the optimization

process, whereas SO approach led to an underestimation of muscle activations by minimizing co-activation between agonist and antagonist muscles concurrently[30,53]. Last, it has been theorized that muscle synergies are generated by the central nervous system to efficiently regulate the control of highly redundant musculoskeletal systems [54,55]. The SynX-based approach leveraged the concept of muscle synergy, making the method more physiological reasonable.

SynX offers benefits over other computational methods for the estimating missing EMG signals within [25,56–58] or outside [59–61] the context of musculoskeletal modeling. Below are some representative approaches that offer great insights for us to develop our method. First, one such method utilizes Gaussian process regression models to describe the synergistic relationship between a subset of muscles, which enables the estimation of unmeasured muscle excitations using information provided by a subset of measured muscle excitations [60]. However, the muscle excitations associated with "unmeasured" muscles must be initially known for conducting the required model training process, rendering this method infeasible when the "unmeasured" muscle excitations are truly unmeasurable due to experimental constraints or safety considerations. Second, an alternative approach employs low-dimension sets of impulsive excitation primitives to estimate unmeasured muscle excitations [25,56,57]. 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. Meanwhile, the primitive-driven excitations for measured muscles are minimally adjusted to improve joint moment estimation in EMG-assisted models. However, these adjustments masked the omission of active force generating properties for unmeasured muscles (i.e., iliacus and psoas), resulting in noticeable hip joint moment prediction errors. Furthermore, none of these studies

669 evaluated the accuracy of predicted unmeasured muscle excitations due to the lack of
670 corresponding experimental EMG data. Third, hybrid EMG-informed models that incorporates SO
671 to determine unmeasured muscle activations have been developed [28,29]. Satori et al. also
672 allowed minimal adjustments of measured muscle activations while predicting missing EMG
673 signals (e.g., from iliacus and psoas) using SO[28]. However, none of these methods have provided
674 evidence that estimation of unmeasured muscle activations was reliable and in reasonable
675 agreement with experimental measurements. Furthermore, due to the nature of SO, the resulting
676 muscle activations might exhibit unrealistic discontinuities. Last, as another well-established
677 approach within OpenSim, the computed muscle control (CMC) algorithm solves a static
678 optimization to determine muscle excitations necessary for achieving the desired accelerations for
679 tracking experimental motion, providing more accurate joint moments compared to SO [58,62–
680 64]. However, it has been observed that CMC may be less robust and computationally efficient
681 when estimating muscle function in human locomotion. All in all, EMG-driven modeling method
682 with SynX provides an enhanced approach for estimating unmeasured muscle excitations, forces
683 and joint moments in an efficient manner, without the requirement for prior knowledge of the
684 "unmeasured" muscle excitations during the model training phase.

685 This study quantified the impact of model personalization, specifically focusing on muscle-tendon
686 parameters, on the estimation of muscle activations and forces in both EMG-driven modeling with
687 SynX and SO. In the case of SynX, the tracking errors between the estimated and experimental
688 estimates were remarkably reduced when muscle-tendon parameter values were personalized to a
689 suitable level for optimization "$SynX^{Params}_{Unmeasured}$". Meanwhile, the mean correlations between the
690 estimated and experimental values were also substantially increased, moving from moderate to
691 strong. Additionally, the matching errors of joint moments during optimization "$SynX^{Params}_{Unmeasured}$"

stayed closer to those obtained from full EMG-driven model optimization "$Params$", as opposed to calibrating model parameters concurrently within optimization "$SynX_{Unmeasured} + Params$". In the case of SO, consistent with previous studies [65], personalization of muscle-tendon parameters showed noticeable improvements in estimating of muscle activations and forces in terms of both shape and amplitude for optimization "SO_{All}^{Params}", with statistically significant enhancement observed only for the shape of muscle forces. In light of these observations, enhancing the level of model personalization generally improved the accuracy of estimation. However, there were significant variations in the degree of improvement among different 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 accurate shape. When conducting simulations using a scaled generic model, SynX successfully replicated muscle activations with the correct amplitude and shape, which SO did not achieve this.

Joint moment matching errors differ among optimizations using different approaches (figure 8 and table 5). First, the inverse dynamics (ID) and estimated joint moments exhibited a much closer agreement in the SO-based optimizations, "$SO_{All}^{Generic}$", "SO_{All}^{Params}", and "$SO_{Unmeasured}^{Params}$" than in the EMG-driven modeling optimizations, "$Params$", "$SynX_{Unmeasured} + Params$", and "$SynX_{Unmeasured}^{Params}$". The optimization formulation used by SO in equation (10) resulted in extremely small joint moment matching errors. However, the additional constraints within EMG-modeling methods, including muscle activation-contraction dynamics and the dependency between time frames of EMG signals, limited the torque-generating capacity of muscles, thereby preventing the reproduction of joint moments. Second, the joint moment matching errors, arranged in descending order, for optimizations associated with the EMG-driven modeling method are "$Params$",

715 "$SynX^{Params}_{Unmeasured}$" and "$SynX_{Unmeasured} + Params$". This observation was attributed to the increasing
716 degrees of freedom in the optimization, determined by additional SynX variables for
717 "$SynX^{Params}_{Unmeasured}$", and also additional model parameters for "$SynX_{Unmeasured} + Params$", which enabled
718 the optimizer to reduce the joint moment matching errors. This can also explain why the joint
719 moment matching errors were smaller when estimating 8 channels of unmeasured EMG signals in
720 this study compared to when estimating EMGs for only the iliacus and psoas in [38]. Last,
721 regardless of the level of model personalization and the number of channels of muscle activations
722 to estimate, SO consistently found the solutions of muscle activations at each time frame to almost
723 perfectly match the ID joint moments, although occasionally requiring a small amount of reserve
724 actuator torque due to model inadequacies. Consequently, static optimization does not possess the
725 joint moment matching errors needed to calibrate muscle-tendon model parameters.

726 This study also examined the impact of incorporating the SynX process in the EMG-driven
727 modeling framework on the calibrated model parameter values. It was observed that Hill-type
728 muscle-tendon model parameter values, specifically optimal fiber length and tendon slack length,
729 from optimization "$SynX_{Unmeasured} + Params$" closely approximated the "true" values obtained from
730 full EMG-driven model optimization "$Params$", as depicted in figure 9. However, the activation
731 dynamics model parameters, including electromechanical delay, activation time constant, EMG
732 scale factor, and activation nonlinear constant, were only reproduced with reasonable similarity
733 for the measured muscles. When SynX was used to estimate a large number of missing EMGs
734 during EMG-driven model calibration, it introduced additional flexibility through SynX variables
735 into the optimization problem. This, in turn, had a cascading effect on the calibrated model
736 parameter values for all muscles. Thus, beyond the primary cost terms specified in equation (9),
737 penalty terms, functioning as "soft constraints" were incorporated to restrict deviations of model

parameter values from the initial model or a designated reference value[23,37,38]. The objective was to minimize the impact of SynX on model parameter values. The strategy proved effective for the Hill-type muscle-tendon models, which are inherently highly nonlinear. Nevertheless, it remained challenging to maintain the values of the activation dynamics model parameters through the utilization of penalty terms. Typically, the transformations from muscle excitations to muscle activations were determined by activation dynamics model parameters, such as time shifts typically defined by electromechanical delay and amplitude scaling dictated by EMG scale factors. Within SynX, however, unmeasured muscle excitations were constructed using linear combinations of measured synergy excitations, which could already account for these potential transformations. Consequently, numerous combinations of model parameters values and SynX variables could result in identical muscle activations. Thus, without compromising the accuracy of muscle activations, the activation model parameter values might approach the designated values in the penalty terms, allowing SynX variables to adjust to provide required muscle activations.

The inclusion of the SynX in the EMG-driven model calibration process had minimal impact on the estimation of measured muscle activations and forces, as depicted in figures 4 and 6, respectively. The muscle activations and forces estimated by SynX for measured muscles from optimization " $SynX_{Unmeasured} + Params$ " remained closely aligned with those from calibration " $Params$ ", in contrast to the results from the commonly formulated " $SO_{All}^{Generic}$ " optimization. While the most significant deviations in the SynX-estimated muscle activations occurred for semimembranosus and semitendinosus (figure 4), and the greatest discrepancies in the SynX-estimated muscle forces were observed for the muscles spanning the hip joint, such as biceps femoris long head and gluteus maximus superior (figure 6), they still exhibit a similar shape to the experimental curves. These observations were largely attributed to the optimization formulation

761 for the SynX-incorporated EMG-driven calibration, resulting in minimal changes in the calibrated
762 activation dynamics model and Hill-type muscle-tendon model parameter values (figure 9).

763 Several methodological choices for SynX were necessary to consider, as they could potentially
764 impact SynX performance, as indicated in table 2. A series of previously published studies from
765 the author have extensively investigated various methodological choices, with the goal of
766 identifying an optimal combination that could yield the most reliable and accurate estimation of
767 unmeasured muscle activations [36–38]. Initially, principal component analysis (PCA) provided
768 more accurate, reliable, and efficient estimates of unmeasured muscle excitations compared to
769 non-negative matrix factorization (NMF), due to the non-negativity constraints for NMF and extra
770 design variables for PCA, both of which could result in a more restricted feasible search space for
771 NMF in comparison to PCA [37,38]. Additionally, PCA was particularly beneficial in our
772 framework because it permitted residual excitations to be both positive and negative, which could
773 be beneficial for achieving lower joint moment errors. Second, by comparing the results of five
774 different EMG normalization methods that were performed either within individual trials or across
775 all trials, we observed that EMG normalization does not have a significant influence on the SynX
776 performance[37]. As a result, the measured muscle excitations were normalized to their maximum
777 values across all trials before MSA to facilitate easy implementation. Furthermore, as the number
778 of synergies increased, the performance of SynX exhibited non-monotonic behavior, with five and
779 six synergies generally providing the best SynX performance and outcomes for EMG-driven
780 model calibration[38]. Hence, when an increasing number of muscles were treated “unmeasured”
781 in this study, five synergies were selected for generating the results in this study, considering the
782 reduction of computational costs. Last, according to the assumptions about the variability of
783 synergy vector weight across walking trials, we categorized them associated with unmeasured and
784 residual muscle excitations as trial-specific, speed-specific, and subject-specific, respectively,

while different concatenation strategies were used to extract corresponding synergy excitations. it was indicated that with an equal number of synergies, the trial-specific unmeasured synergy vector weights and speed-specific residual synergy vector weights produced the best SynX performance for the majority of subjects[38]. This insight shed lights on the categorization strategy of synergy vector weight across walking trials within this study.

A reasonable choice of a neural control strategy is essential for producing physiologically realistic predictive simulations of walking [66]. To date, prevailing predictive simulation studies have explored the optimality of neural control principles underlying human gait, and commonly reported that minimizing the sum of squares of muscle activations in the cost function, a typical practice in SO, can result in a human-like walking pattern [26,66,67]. However, the comparative results between the estimated muscle activations and forces from both SynX and SO in this study have raised a pivotal question: If the prevalent neural control strategy of minimizing the sum of squares of muscle activations fails to accurately estimate muscle activations when the joint kinematics and moments are known a priori from experimental walking data, how can it provide reliable estimates in the predictive simulations of walking when the joint kinematics and moments are unknown a priori? These findings may also yield valuable insights into the potential benefits of 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 preliminary verified for only one experimental scenario thus far [68], the results of the present study endorse 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 two post-stroke subjects, as these experimental datasets provided EMG signals for every 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 developed the framework using walking data with two representative speeds. It would be valuable to investigate its applicability for various dynamic movement conditions and experimental scenarios, including stair climbing and running. Third, to enhance computational efficiency, we integrated a rigid tendon model into our Hill-type muscle-tendon models. Research studies have indicated that 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 stroke subjects walked at slow speeds, it suggests that use of a rigid tendon model was appropriate. However, it would be worthwhile to expand our approach to include compliant tendons in our Hill-type muscle-tendon models, enabling the applications to tasks involving fast movements. Last, we analyzed the impact of personalizing muscle-tendon parameter values on SynX performance. However, various other aspects of model personalization, including skeletal geometries, muscle kinematics, and other physiological properties that contribute to muscle force generation, may also impact muscle force estimates. Future work should therefore aim to extend the methods to investigate whether SynX performance was sensitive to these aspects of model personalization. The author has recently developed an EMG-driven modeling method that can personalize muscle wrapping surface parameters [71]. Therefore, one of the forthcoming research directions would focus on examining the influence of personalizing muscle-tendon pathway on SynX-based estimates.

Conclusions

In conclusion, this study demonstrated a significant advancement over previous research by highlighting the capability of SynX to reproduce a large number of 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 process had minimal impact on the calibrated Hill-type muscle-tendon model parameter values for all muscles and activation dynamics model parameter values for the 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 in this regard. 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 muscle impairments in situations where difficulties arise in collecting EMG signals from all important contributing muscles.

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 subject provided written informed consent before participation.

Consent for publication

Not applicable.

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*).

Competing interests

The authors declare that they have no competing interests.

Authors' 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.

Acknowledgements

Not applicable.

List of abbreviations

Abbreviation	*Definition*
EMG	electromyography
SynX	synergy extrapolation
SO	static optimization
RMSE	root mean square errors

DOFs	degrees of freedom
IK	inverse kinematics
ID	inverse dynamic
MSA	muscle synergy analysis
PCA	principal component analysis
NMF	non-negative matrix factorization

883 # Larger Tables

Table 3 Pearson correlation coefficient r values and root mean square error (RMSE) values were calculated between the experimental ("Params" calibration) and estimated muscle activations for different calibration cases. The RMSE and *r* values for each muscle were calculated when the data across all calibration trials, legs and subjects were concatenated.

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**

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

884

Table 4 Pearson correlation coefficient r values and root mean square error (RMSE) values were calculated between the experimental ("Params" calibration) and estimated muscle forces for different calibration cases. The r and RMSE values were calculated when the data across all calibration trials and subjects were concatenated.

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**

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

885

886

887

888

Table 5 Mean absolute error (MAE) values calculated between joint moments found from inverse dynamics and either "Params" calibration or SynX-based calibrations including $SynX_{Unmeasured} + Params$ and $SynX^{Params}_{Unmeasured}$.

DOFs	Calibration Case	Subjects			
		S1		S2	
		Left	Right	Left	Right
HipFE	*Params*	7.45	6.49	5.81	4.90
	$SynX_{Unmeasured} + Params$	4.73	3.96	3.97	5.84
	$SynX^{Params}_{Unmeasured}$	8.25	5.55	6.84	7.53
HipAA	*Params*	7.31	7.23	7.17	5.51
	$SynX_{Unmeasured} + Params$	4.88	5.65	4.85	4.90
	$SynX^{Params}_{Unmeasured}$	5.79	5.80	6.42	2.77
HipRot	*Params*	5.82	5.64	2.53	2.19
	$SynX_{Unmeasured} + Params$	5.61	4.81	2.55	1.75
	$SynX^{Params}_{Unmeasured}$	5.98	5.22	2.87	1.46
KneeFE	*Params*	6.33	4.26	4.90	4.70
	$SynX_{Unmeasured} + Params$	3.98	2.68	3.08	2.92
	$SynX^{Params}_{Unmeasured}$	3.94	2.87	3.88	2.02
AnklePD	*Params*	6.88	5.68	5.04	4.85
	$SynX_{Unmeasured} + Params$	4.51	4.11	3.14	2.50
	$SynX^{Params}_{Unmeasured}$	5.78	4.63	4.14	1.77
AnkleIE	*Params*	4.89	2.82	8.19	1.75
	$SynX_{Unmeasured} + Params$	2.50	1.89	2.73	1.46
	$SynX^{Params}_{Unmeasured}$	2.90	2.13	4.22	0.95

889

890

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. 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.

14. 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.

15. Vilimek M. Musculotendon forces derived by different muscle models.

16. Zajac FE. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. Crit Rev Biomed Eng. 1989;17:359–411.

17. 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.

18. Modenese L, Ceseracciu E, Reggiani M, Lloyd DG. Estimation of musculotendon parameters for scaled and subject specific musculoskeletal models using an optimization technique. J Biomech. 2016;49:141–8.

19. 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.

20. 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. 2003;

21. 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.

22. 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.

23. Meyer AJ, Patten C, Fregly BJ. Lower extremity EMG-driven modeling of walking with automated adjustment of musculoskeletal geometry. PloS One. 2017;12:e0179698.

24. 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.

25. 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.

26. Anderson FC, Pandy MG. Static and dynamic optimization solutions for gait are practically equivalent. J Biomech. 2001;34:153–61.

27. 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.

960 28. Sartori M, Farina D, Lloyd DG. Hybrid neuromusculoskeletal modeling to best track joint
961 moments using a balance between muscle excitations derived from electromyograms and
962 optimization. J Biomech. 2014;47:3613–21.

963 29. Zonnino A. Model-based estimation of individual muscle force based on measurements of
964 muscle activity in forearm muscles.

965 30. Kian A, Pizzolato C, Halaki M, Ginn K, Lloyd D, Reed D, et al. Static optimization
966 underestimates antagonist muscle activity at the glenohumeral joint: A musculoskeletal modeling
967 study. J Biomech. 2019;97:109348.

968 31. Shourijeh MS, Mehrabi N, McPhee J. Forward Static Optimization in Dynamic Simulation of
969 Human Musculoskeletal Systems: A Proof-of-Concept Study. J Comput Nonlinear Dyn.
970 2017;12:051005.

971 32. Michaud F, Lamas M, Lugrís U, Cuadrado J. A fair and EMG-validated comparison of
972 recruitment criteria, musculotendon models and muscle coordination strategies, for the inverse-
973 dynamics based optimization of muscle forces during gait. J NeuroEngineering Rehabil.
974 2021;18:17.

975 33. Bernstein N. The coordination and regulation of movements. No Title. 1967;

976 34. Ackermann M, van den Bogert AJ. Optimality principles for model-based prediction of
977 human gait. J Biomech. 2010;43:1055–60.

978 35. Sartori M, Gizzi L, Lloyd DG, Farina D. A musculoskeletal model of human locomotion
979 driven by a low dimensional set of impulsive excitation primitives. Front Comput Neurosci
980 [Internet]. 2013 [cited 2024 Jan 8];7. Available from:
981 http://journal.frontiersin.org/article/10.3389/fncom.2013.00079/abstract

982 36. Bianco NA, Patten C, Fregly BJ. Can Measured Synergy Excitations Accurately Construct
983 Unmeasured Muscle Excitations? J Biomech Eng. 2018;140:011011.

984 37. Ao D, Shourijeh MS, Patten C, Fregly BJ. Evaluation of Synergy Extrapolation for
985 Predicting Unmeasured Muscle Excitations from Measured Muscle Synergies. Front Comput
986 Neurosci. 2020;14:588943.

987 38. Ao D, Vega MM, Shourijeh MS, Patten C, Fregly BJ. EMG-driven musculoskeletal model
988 calibration with estimation of unmeasured muscle excitations via synergy extrapolation. Front
989 Bioeng Biotechnol. 2022;10:962959.

990 39. Tahmid S, Font-Llagunes JM, Yang J. Upper Extremity Muscle Activation Pattern Prediction
991 Through Synergy Extrapolation and Electromyography-Driven Modeling. J Biomech Eng.
992 2024;146:011005.

993 40. Fregly BJ. How Well Do Commonly Used Co-contraction Indices Approximate Lower Limb
994 Joint Stiffness Trends During Gait for Individuals Post-stroke? Front Bioeng Biotechnol. 2021;8.

995 41. McLean SG, Huang X, Van Den Bogert AJ. Association between lower extremity posture at
996 contact and peak knee valgus moment during sidestepping: implications for ACL injury. Clin
997 Biomech. 2005;20:863–70.

998 42. Hug F. Can muscle coordination be precisely studied by surface electromyography? J
999 Electromyogr Kinesiol. 2011;21:1–12.

1000 43. Rajagopal A, Dembia CL, DeMers MS, Delp DD, Hicks JL, Delp SL. Full-Body
1001 Musculoskeletal Model for Muscle-Driven Simulation of Human Gait. IEEE Trans Biomed Eng.
1002 2016;63:2068–79.

1003 44. Delp SL, Anderson FC, Arnold AS, Loan P, Habib A, John CT, et al. OpenSim: Open-
1004 Source Software to Create and Analyze Dynamic Simulations of Movement. IEEE Trans
1005 Biomed Eng. 2007;54:1940–50.

1006 45. 2018 AJ Seth OpenSim Simulating musculoskeletal dynamics and neuromuscular control to
1007 study human and animal movement.pdf.

1008 46. Reinbolt JA, Schutte JF, Fregly BJ, Koh BI, Haftka RT, George AD, et al. Determination of
1009 patient-specific multi-joint kinematic models through two-level optimization. J Biomech.
1010 2005;38:621–6.

1011 47. He J, Levine WS, Loeb GE. Feedback gains for correcting small perturbations to standing
1012 posture. IEEE; 1989. p. 518–26.

1013 48. Manal K, Buchanan TS. A one-parameter neural activation to muscle activation model:
1014 estimating isometric joint moments from electromyograms. J Biomech. 2003;36:1197–202.

1015 49. Hill AV. The heat of shortening and the dynamic constants of muscle. Proc R Soc Lond Ser
1016 B-Biol Sci. 1938;126:136–95.

1017 50. Serrancolí G, Kinney AL, Fregly BJ, Font-Llagunes JM. Neuromusculoskeletal Model
1018 Calibration Significantly Affects Predicted Knee Contact Forces for Walking. J Biomech Eng.
1019 2016;138:081001.

1020 51. An K, Takahashi K, Harrigan T, Chao E. Determination of muscle orientations and moment
1021 arms. J Biomech Eng. 1984;106:280–2.

1022 52. Taylor R. Interpretation of the correlation coefficient: a basic review. J Diagn Med Sonogr.
1023 1990;6:35–9.

1024 53. Shourijeh MS, Fregly BJ. Muscle Synergies Modify Optimization Estimates of Joint
1025 Stiffness During Walking. J Biomech Eng. 2020;142:011011.

1026 54. Tresch MC, Cheung VCK, d'Avella A. Matrix Factorization Algorithms for the
1027 Identification of Muscle Synergies: Evaluation on Simulated and Experimental Data Sets. J
1028 Neurophysiol. 2006;95:2199–212.

1029 55. Torres-Oviedo G, Ting LH. Muscle Synergies Characterizing Human Postural Responses. J
1030 Neurophysiol. 2007;98:2144–56.

56. 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 [Internet]. 2013 [cited 2023 May 10];7. Available from: http://journal.frontiersin.org/article/10.3389/fncom.2013.00079/abstract

57. Neptune RR, Clark DJ, Kautz SA. Modular control of human walking: A simulation study. J Biomech. 2009;42:1282–7.

58. 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.

59. Ajiboye AB, Weir RF. Muscle synergies as a predictive framework for the EMG patterns of new hand postures. J Neural Eng. 2009;6:036004.

60. 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.

61. 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.

62. Thelen DG, Anderson FC, Delp SL. Generating dynamic simulations of movement using computed muscle control. J Biomech. 2003;36:321–8.

63. 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.

64. 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.

65. 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.

66. 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.

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 [Internet]. 2016 [cited 2023 May 10];4. Available from: http://journal.frontiersin.org/article/10.3389/fbioe.2016.00077/full

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. 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.