

Can Measured Synergy Excitations Accurately Construct Unmeasured Muscle Excitations?

(Winner of the 2015 SB3C Undergraduate Student Paper Competition in
Human Dynamics, Injury, Tissue Mechanics, and Cellular & Tissue Engineering)

Nicholas A. Bianco¹,Carolynn Patten², Benjamin J. Fregly^{1,3,4}

¹Dept. of Mechanical & Aerospace Engineering, University of Florida, Gainesville, FL, USA

²Neural Control of Movement Lab, Malcom Randall VA Medical Center and Department of
Physical Therapy, University of Florida, Gainesville, FL, USA

³Dept. of Biomedical Engineering, University of Florida, Gainesville, FL, USA

⁴Dept. of Orthopaedics & Rehabilitation, University of Florida, Gainesville, FL, USA

Corresponding author:

B.J. Fregly, Ph.D.

Dept. of Mechanical & Aerospace Engineering

231 MAE-A Building, P.O. Box 116250,

University of Florida

Gainesville, FL 32611-6250

Phone: (352) 392-8157

Fax: (352) 392-7303

E-mail: fregly@ufl.edu

Keywords: biomechanics, electromyography, EMG-driven, optimization, muscle synergy analysis, motor modules, musculoskeletal modeling, non-negative matrix factorization, synergy extrapolation, variance accounted for

ABSTRACT

Accurate prediction of muscle and joint contact forces during human movement could improve treatment planning for disorders such as osteoarthritis, stroke, Parkinson's disease, and cerebral palsy. Recent studies suggest that muscle synergies, a low-dimensional representation of a large set of muscle electromyographic (EMG) signals (henceforth called "muscle excitations"), may reduce the redundancy of muscle excitation solutions predicted by optimization methods. This study explores the feasibility of using muscle synergy information extracted from eight muscle EMG signals (henceforth called "included" muscle excitations) to accurately construct muscle excitations from up to 16 additional EMG signals (henceforth called "excluded" muscle excitations). Using treadmill walking data collected at multiple speeds from two subjects (one healthy, one post-stroke), we performed muscle synergy analysis on all possible subsets of eight included muscle excitations and evaluated how well the calculated time-varying synergy excitations could construct the remaining excluded muscle excitations (henceforth called "synergy extrapolation"). We found that some, but not all, eight-muscle subsets yielded synergy excitations that achieved $>90\%$ extrapolation variance account for (VAF). Using the top 10% of subsets, we developed muscle selection heuristics to identify included muscle combinations whose synergy excitations achieved high extrapolation accuracy. For 3, 4, and 5 synergies, these heuristics yielded extrapolation VAF values approximately 5% lower than corresponding reconstruction VAF values for each associated eight-muscle subset. These results suggest that synergy excitations obtained from experimentally measured muscle excitations can accurately construct unmeasured muscle excitations, which could help limit muscle excitations predicted by muscle force optimizations.

INTRODUCTION

The ability to determine muscle and joint contact forces reliably during walking could be useful for developing improved treatments for disorders such as osteoarthritis, stroke, Parkinson's disease, and cerebral palsy [1]. However, experimental approaches do not currently exist for measuring muscle and joint contact forces *in vivo* under normal conditions. Similarly, computational approaches for predicting these quantities are limited by the redundancy inherent in movement coordination (i.e., many more muscles actuating the skeleton than degrees of freedom in the skeleton) [2].

To address the indeterminacy issue, computational studies generally use either optimization or electromyographic (EMG) driven methods to estimate muscle and joint contact forces in musculoskeletal models. Optimization methods include static optimization, which is an inverse dynamic approach that utilizes prescribed kinematics and net joint moments to find muscle excitations, and dynamic optimization, which is a forward dynamic approach that solves for kinematics, net joint moments, and muscle excitations simultaneously [1,3–5]. While these methods can provide unique muscle and joint contact force solutions, they depend on subjective cost functions that make assumptions about what the body is maximizing or minimizing. Furthermore, predicted muscle excitations frequently do not agree well with EMG measurements. In contrast, EMG-driven methods use processed experimental EMG data to directly control muscle excitations in a musculoskeletal model [6–9]. Similar to static optimization, these methods require prescribed kinematics and net joint moments, but unlike optimization methods in general, they avoid the need for an assumed cost function. However, variations in the quality of EMG data, especially for muscles not accessible by surface electrodes, limit these methods [10,11]. Consequently, even EMG-driven methods often use

optimization methods to calculate unmeasured muscle excitations, which may be inconsistent with the actual muscle excitations [12].

Muscle synergy analysis has been proposed as a way to reduce the indeterminacy of muscle excitations predicted by optimization methods [13,14]. Muscle synergy analysis is a computational approach that uses non-negative matrix factorization to decompose a large number of experimentally measured EMG signals into a smaller number of independent time-varying “synergy excitations,” along with weights that describe how each synergy excitation contributes to each measured muscle excitation [15–17]. Typically, only three to five synergy excitations are needed to account for over 90% of the variability in up to 32 muscle excitations measured experimentally during walking [18–20]. To date, muscle synergy analysis has been used primarily in descriptive studies to analyze experimental EMG data [18,19]. For example, muscle synergy analysis has been applied to EMG data collected from individuals post-stroke or with Parkinson’s disease during walking revealing that, compared to healthy individuals, a reduced number of synergy excitations is often required to reach 90% variance accounted for (VAF) [20,21]. Recent studies have also incorporated muscle synergy concepts into muscle force optimizations [22–28]. One study improved prediction of knee contact forces by replacing a high-dimensional set of 44 independent muscle excitations with a lower-dimensional set of 5 synergy excitations [27]. The synergy excitations were extracted from 13 experimental EMG signals and were used to construct 44 interdependent muscle excitations. However, no evidence exists that synergy information obtained from a small number of experimentally measured EMG signals can accurately construct muscle excitations associated with unmeasured EMG signals.

This study explored the feasibility of using synergy excitations extracted from eight surface EMG signals collected during walking to accurately construct muscle excitations derived from

surface and fine-wire EMG signals collected simultaneously from up to 16 other muscles. EMG data collected from one healthy subject (24 channels from one leg) and one subject post-stroke (16 channels from each leg) during treadmill walking were subdivided into multiple subsets of eight commonly collected surface EMG signals and all remaining surface and fine-wire EMG signals. Two novel measures were investigated to quantify the ability of synergy excitations extracted from each eight-muscle subset to construct the remaining muscle excitations not included in the subset, a process we call “synergy extrapolation.” The most appropriate measure was used to quantify how the choice of muscles in each eight-muscle subset affected construction accuracy for the remaining muscle excitations. Thus, the two main questions we addressed were: 1) Is synergy extrapolation using surface EMG data theoretically possible for walking, and if so, 2) Which muscles accessible by surface EMG provide the best information for synergy extrapolation?

METHODS

Experimental Data Collection and Reduction

Experimental EMG data were collected from one healthy subject and one subject post-stroke while walking at different speeds on an instrumented treadmill. Institutional review board approval and subject written informed consent were obtained prior to participation. Eleven treadmill walking trials ranging from 0.4 to 1.4 m/s were collected from the healthy subject, and seven trials ranging from 0.2 to 0.8 m/s were collected from the stroke subject. From each trial, ten consecutive gait cycles were chosen for analysis. A total of 24 EMG signals (17 surface, 7 fine-wire) were collected from lower extremity muscles of the healthy subject’s dominant leg, while a total of 16 EMG signals (12 surface, 4 fine-wire) were collected from each leg of the

stroke subject (Table 1). For the healthy subject, two recordings of rectus femoris activity were taken to capture fully its biarticular action. Whereas SENIAM guidelines state that the rectus femoris electrode should be placed 50% of the way between the superior aspect of the patella and the anterior superior iliac spine, we chose electrode placements located more proximally and distally compared to this location to capture both hip flexion and knee extension components, respectively. EMG signals were band-pass filtered (zero-lag fourth-order Butterworth) from 10 to 400 Hz for surface EMG and 20 to 400 Hz for fine-wire EMG to account for additional soft tissue motion artifacts. All EMG data were then demeaned, full-wave rectified, and low-pass filtered (zero-lag fourth-order Butterworth) at $7/(\text{cycle time})$ Hz (6 Hz for a typical cycle time of approximately 1.2 seconds) to account for speed-related changes in motion artifacts [29]. The resulting processed EMG signals defined experimental muscle excitation profiles. Muscle excitations in EMG-driven studies are typically normalized to the excitation level recorded during a maximum voluntary contraction (MVC) [6,8,30,31]. Alternatively, since MVC trials were not performed for either subject in this study, we normalized each muscle excitation by its maximum value observed across all walking trials [9].

Synergy Extrapolation Approach

In preparation for performing thousands of muscle synergy analyses, we grouped muscle excitations from both subjects into multiple pairs of “included” and “excluded” subsets. Muscle excitations placed in an “included” subset were used in an associated synergy analysis, while those placed in an “excluded” subset were omitted from that synergy analysis. Each “included” subset was composed of muscle excitations from eight muscles with surface EMG data, while the corresponding “excluded” subset was composed of the remaining 16 (dominant leg of

healthy subject) or 8 (each leg of stroke subject) muscle excitations from muscles with either surface or fine-wire EMG data. The total number of included subsets for each subject (24,310 for the healthy subject and 495 for each leg of the stroke subject) was determined by the number of possible combinations of eight muscles with surface EMG data. Eight muscle excitations were used in each included subset since this number is consistent with the number of experimental EMG signals commonly used in synergy analysis studies [20,25,32]. In practice, included muscle excitations would be those that can be measured easily using surface electrodes, while excluded muscle excitations would be those that cannot be measured easily due to limitations in number of available EMG channels or the need for fine-wire electrodes.

We performed muscle synergy analysis on each included muscle excitation subset using non-negative matrix factorization (NMF) [15–17]. Each synergy analysis produced a low-dimensional set of time-varying signals (henceforth called “synergy excitations”) and a corresponding set of weighting vectors (henceforth called “synergy vectors”) that specified how each synergy excitation contributed to each included muscle excitation [15,16]. The NMF optimization algorithm minimizes the following objective function $f(W, H_i)$,

$$f(W, H_i) = \frac{1}{2} \|WH_i - A_i\|_F^2 \quad (1)$$

where W is an n time points \times p synergies matrix containing synergy excitations (presumably representing both included and excluded muscle excitations) in columns, H_i is a p synergies \times m_i muscles matrix containing the synergy vectors associated with the included muscle excitations in rows, A_i is an n time points \times m_i muscles matrix containing the included (subscript i) muscle excitations in columns, and $\| \cdot \|_F$ represents the Frobenius norm [33]. Synergy vectors were allowed to vary between gait cycles, and therefore synergy analysis using Eq. (1) was performed

on each gait cycle separately. Prior to synergy analysis, we normalized each muscle excitation profile to have unit variance to allow the variations in each muscle to be considered with equal importance in the NMF solution approach [15]. Three to five synergies were extracted from each included muscle excitation subset, which is typical of other studies involving synergy analysis [18–20]. The resulting synergy excitations were normalized by a scaling factor to have a maximum value of one, consistent with the maximum value of muscle excitations used in muscle force optimizations. Corresponding synergy vectors were multiplied by the same scaling factor to keep reconstructed included muscle excitations unaltered.

Several algorithms are available to factor matrix A_i into W and H_i . In our study, we used a combination of the default algorithms included with the MATLAB `nnmf` function (The MathWorks Inc., Natick, MA). The multiplicative update algorithm was used to generate initial guesses for W and H_i , and these initial guesses were then used in the alternating least squares algorithm, which is faster and allows for more evaluations of the objective function. To avoid finding local minima, we performed each synergy analysis ten times using ten different initial guesses for the multiplicative update algorithm, taking the solution with lowest final cost function value as our best estimate of the global minimum. For a detailed explanation of both algorithms, refer to Berry *et al.* 2007 [33].

We evaluated how accurately the synergy excitations W obtained from each included muscle excitation subset A_i were able to construct each corresponding excluded muscle excitation subset A_e – a process we call “synergy extrapolation.” Synergy extrapolation assumes that the synergy excitations obtained from the included muscle excitations apply equally well to the excluded muscle excitations, though the additional synergy vector weights H_e for the excluded muscle excitations remain unknown. How to find these unknown weights in practice remains a

topic of ongoing research. For the present study, our goal was to determine whether synergy extrapolation could work in theory by taking advantage of our knowledge of the excluded muscle excitations in each included-excluded subset pair. Specifically, given W as found using Eq. (1) and A_e from the same subset pair, we calculated H_e using linear least-squares regression, as indicated by the following matrix equations,

$$\begin{aligned} W^T W H_e &= W^T A_e \\ H_e &= (W^T W)^{-1} W^T A_e \end{aligned} \tag{2}$$

where W is the previously found n time points $\times p$ synergies matrix containing synergy excitations in columns, H_e is a p synergies $\times m_e$ muscles matrix containing the synergy vectors associated with the excluded muscle excitations, and A_e is an n time points $\times m_e$ muscles matrix containing the excluded (subscript e) muscle excitations in columns. In our study, we solved Eq. (2) using the MATLAB backslash operator, which did not force the recovered synergy vector weights to be positive. Once H_e was calculated, we reconstructed the corresponding excluded muscle excitations using matrix multiplication,

$$\tilde{A}_e = W H_e \tag{3}$$

where \tilde{A}_e is the synergy-based approximation of A_e . We performed this synergy extrapolation approach for every included-excluded muscle subset pair, thereby providing a comprehensive set of extrapolation solutions. Variance accounted for (VAF), a measure of goodness-of-fit that accounts for magnitude and shape, was utilized to evaluate all extrapolation solutions for excluded muscle excitations [15]. Each available data set (healthy subject leg, stroke subject non-paretic and paretic leg) was analyzed separately. An overview of our entire computational approach is provided in Fig. 1.

Synergy Extrapolation Evaluation

VAF values calculated for reconstructed excluded muscle excitations through synergy extrapolation were averaged across excluded muscles and gait cycles to create a single Extrapolation VAF for each excluded muscle subset. We sorted included muscle subsets based on the Extrapolation VAF of their corresponding excluded muscle subset. The 10% of included muscle subsets that resulted in excluded muscle subsets with the highest Extrapolation VAF were identified as the Top 10% of Subsets. Extrapolation ability was measured two different ways: 1) the number of subsets that exceeded 90% Extrapolation VAF, and 2) the Mean Extrapolation VAF for the Top 10% of Subsets at each walking speed. We used these two measures to define “acceptable” muscle subsets which were then compared to determine which included muscle subsets produced the most accurate extrapolation results.

Using the measure that best described extrapolation ability, we developed four different heuristics for selecting included muscle subsets likely to extrapolate with the highest accuracy possible. Since these heuristics were based on the results of the preceding synergy analysis, we describe the four heuristics in detail within the Results section and describe here only the rationale for their development. To define the number of synergies for developing heuristics, we determined how many synergies were needed such that Extrapolation VAF exceeded 90% on average for all three data sets. For each data set, we calculated the frequency with which each included muscle appeared in an acceptable subset across all trials. Next, we developed heuristics based on overall muscle frequency and importance of having muscles in different functional groups. Included muscle subsets selected from these heuristics were evaluated based on their ability to construct accurately excluded muscle excitations using the presented synergy extrapolation method, noting that under normal conditions, the excluded muscle excitations

would not be known. In addition, the individual reconstruction VAF values for included muscle excitations chosen by each heuristic were calculated and averaged together for three to five extracted synergy excitations. Reconstruction VAF results for included muscles were compared to Extrapolation VAF results for excluded muscles to determine the number of synergy excitations required to achieve a desired extrapolation accuracy.

RESULTS

Overall, synergy excitations extracted from subsets of eight included muscle excitations derived from surface EMG data were able to construct the remaining excluded muscle excitations derived from surface and fine-wire EMG data with high accuracy, as demonstrated by both measures of extrapolation ability. The number of eight-muscle subsets able to achieve 90% Extrapolation VAF increased as the number of synergies was increased (Table 2). Acceptable combinations were found for most walking speeds when using four and five synergy excitations, while almost no acceptable combinations were found when using three synergy excitations. While variations were also observed across walking speeds, these variations were found to be sensitive to small reductions in the VAF cutoff value (Fig. 2). For example, for the paretic leg of the stroke subject analyzed using five synergies, three walking speeds had virtually no combinations that achieved the 90% VAF cutoff value. However, when the VAF cutoff was reduced to 85%, more than 200 muscle combinations met the VAF requirement for all walking speeds. Furthermore, reducing the VAF cutoff to 80% allowed over 90% of muscle combinations to meet the cutoff for every trial in all three experimental data sets. Variations across speeds were also observed in the Mean Extrapolation VAF for the Top 10% of Subsets, although not as acutely (Fig. 3). While this measure is contrary to cutoffs traditionally used in synergy analysis,

it eliminates the subjectivity of a 90% VAF value and may be a better indicator of extrapolation ability. Consequently, we chose the Mean Extrapolation VAF for Top 10% of Subsets as the most appropriate measure for identifying heuristics that produce high Extrapolation VAF.

Muscle selection heuristics developed using muscle frequency in the Top 10% of Subsets across all walking trials (Table 3) led to reliable identification of eight-muscle subsets with high extrapolation accuracy (Table 4). Muscle combinations selected by all four heuristics achieved average Extrapolation VAF of 86.9%, 89.7%, and 91.9% for three, four, and five synergy excitations, respectively, with only small variations observed between heuristics (Table 5). These combinations achieved average reconstruction VAF of 91.8%, 95.7%, and 97.8% for three, four, and five synergy excitations, respectively. Thus, average reconstruction VAF values were greater than Extrapolation VAF values by an average of 5.5%. The combinations of eight muscles selected by heuristics represented all primary lower extremity muscle functions during walking.

DISCUSSION

This study addressed two primary questions: 1) Is synergy extrapolation – using synergy excitations obtained from one set of muscle excitations to construct another set of muscle excitations – theoretically possible for walking, and 2) if so, which muscles accessible by surface EMG provide the most information for synergy extrapolation? Our results clearly demonstrate that numerous subsets of eight muscle excitations derived from surface EMG data (i.e., the “included” muscles) during walking are able to accurately construct between eight and 16 additional muscle excitations derived from surface and fine-wire EMG data collected simultaneously (i.e., the “excluded” muscles), indicating that synergy extrapolation is theoretically possible for walking. However, many included muscle subsets did not yield high

Extrapolation VAF values, indicating that the choice of included muscles for calculating the necessary synergy excitations is important. We also developed heuristics for selecting combinations of 8 commonly collected surface EMG signals that provided high extrapolation accuracy. Based on these results, synergy extrapolation may be useful for limiting predicted muscle excitations in muscle force optimizations, since the problem of finding a time-varying muscle excitation for each muscle can be reduced to the problem of finding only 3, 4, or 5 time-varying synergy excitations with associated synergy vector weights. It may also be useful for limiting predicted muscle excitations in EMG-driven models with missing EMG signals, since the problem of finding a time-varying muscle excitation for each unmeasured muscle is reduced to the problem of finding only 3, 4, or 5 synergy vector weights. Furthermore, because our results were consistent between our two subjects, our findings may be equally applicable to healthy individuals as well as individuals with neurological impairments.

While our study demonstrates that synergy extrapolation is theoretically possible, it does not provide practical information regarding how to find the unknown synergy vector weights. In real life conditions, some muscles will have available surface (and possibly fine-wire) EMG data while many others will not. To evaluate whether or not synergy extrapolation could work on a theoretical basis, we divided our large number of available EMG signals into “included” and “excluded” categories, where the “included” muscles represented those from which experimental EMG could be collected – thus synergy analysis could be performed, while the “excluded” muscles represented those from which experimental EMG data would be more difficult to collect. We took advantage of our knowledge of excluded muscle excitations to determine how well they could be constructed using synergy excitations obtained from the included muscles. Of course, finding the unknown synergy vector weights for the excluded muscles this way would

not be possible in practice. Determining how to find these unknown weights remains a topic of ongoing research.

At a minimum, by showing that synergy extrapolation can work for walking, we have demonstrated that it is reasonable to reduce the indeterminacy in muscle force optimization problems by using a small number of synergy excitations to construct all muscle excitations [27]. For example, consider a static optimization problem for walking where a 5 degree-of-freedom leg model is controlled by 50 muscle-tendon actuators and the gait cycle is divided into 100 time points. Theoretically we would have $5 \times 100 = 500$ inverse dynamic moments to be matched by $50 \times 100 = 5,000$ unknown muscle excitation values (ignoring activation dynamics), which is underdetermined by a factor of 10. If we now model all muscle excitations using 5 synergies, we would have $5 \times 100 = 500$ unknown synergy excitation values and $5 \times 50 = 250$ unknown synergy vector weights for a total of 750 unknowns. For this situation, the problem would be underdetermined by only a factor of 1.5, making the design space much smaller and the solution much closer to unique. However, because the unknown synergy vector weights apply to all time frames, the optimization problem must now be solved over all time frames simultaneously, increasing the computational cost of the solution process significantly (e.g., [27,34]). If the unknown synergy excitations could be obtained from experimentally measured muscle excitations, they could serve as subject-specific basis functions to constrain and simplify the construction of all muscle excitations [26]. In this case, only the synergy vector weights would be unknown, and the problem formulation could theoretically become overdetermined.

While synergy-based muscle force optimizations seem promising, technical challenges remain with implementing such approaches. When synergy solutions are calculated, a constraint must be placed on the synergy excitations W or the synergy vectors H such that the product WH

is unique. Some investigators require that the maximum value of each synergy vector be one [35,36], while others require that the maximum value of each synergy excitation be one [37]. The MATLAB nnmf algorithm normalizes the magnitude of each synergy vector to be one, which means that the amplitude of the corresponding synergy excitations can vary significantly as a function of the number of muscle excitations included in the analysis. In a muscle force optimization problem, the simplest approach is to add a constraint requiring the magnitude of each synergy vector to be one, as done in a recent study [28]. Forcing all predicted synergy vectors or synergy excitations to have a maximum value of one would be more difficult to implement, since maximum value constraints are not differentiable. When experimental EMG data are available from muscles with high Extrapolation VAF, how to use the calculated synergy excitations also requires further investigation. No reliable method currently exists for normalizing experimental muscle excitations, and the shapes of calculated synergy excitations are affected by how the experimental muscle excitations are normalized. Thus, it is not clear how synergy excitations calculated from experimental muscle excitations should be used to construct muscle excitations predicted by muscle force optimizations.

Despite these challenges, at least three recent studies have attempted to use synergy-based methods to estimate muscle forces using static optimization. Sartori *et al.* (2013) [26] constructed 34 muscle excitations using 5 “excitation primitives” fitted to 5 synergy excitations obtained from 16 experimentally measured muscle EMG signals. Each excitation primitive was a task-generic Gaussian-shaped impulsive curve representing data collected from two subjects performing four different gait tasks. Using this low-dimensional representation, the authors were able to match experimental hip, knee and ankle moments and experimental muscle excitations reasonably well for both subjects over all four gait tasks. Walter *et al.* (2014) [27] constructed 44

muscle excitations using 5 synergy excitations extracted from 13 experimental muscle EMG signals collected from a subject implanted with a force-measuring knee replacement. They showed that use of 5 synergy excitations instead of 44 independent muscle excitations improved the accuracy of knee contact force predictions. However, muscle excitations predicted by the 5 synergy excitations were only slightly closer to the 13 experimental muscle excitations than were those predicted by the 44 independent muscle excitations. Serrancolí *et al.* (2016) [34] found 44 muscle excitations that could be closely fit by 5 synergy excitations derived from 10 experimental muscle EMG signals collected from a different instrumented knee subject. Neuromusculoskeletal model parameter values were calibrated using a two-level optimization approach, and the calibrated model predicted knee contact and leg muscle forces for walking trials withheld from the calibration process. Keeping predicted muscle excitations close to an experimental synergy solution resulted in accurate knee contact force predictions only when model parameter values were well calibrated. The results of the present study provide an experimental justification for using synergy-based methods to construct muscle excitations, including those of deep muscles requiring fine-wire EMG, in these recent muscle force optimization studies.

We defined two measures, 90% Extrapolation VAF and Mean Extrapolation VAF for the Top 10% of Subsets, to evaluate the ability of synergy excitations derived from included muscle subsets to construct muscle excitations not included in the synergy analysis. In this way, we hoped to eliminate the potential bias of a single measure in selecting “acceptable” muscle subsets. We averaged the individual muscle VAF values into a single Extrapolation VAF to facilitate interpretation of extrapolation ability for any muscle combination. The 90% Extrapolation VAF measure was defined based on typical VAF cutoffs used in synergy analyses

reported in the literature [18–20]. This measure was able to identify subsets that extrapolated well, but it omitted subsets that might be deemed acceptable if a small reduction in cutoff was allowed. This observation motivated us to use the Mean Extrapolation VAF for the Top 10% of Subsets measure for the remainder of the study. This measure provided finer granularity and showed clearly that a large number of subsets were on the cusp of achieving 90% Extrapolation VAF.

The four muscle selection heuristics in Table 4 were defined in an attempt to choose, in a methodological manner, subsets of muscle excitations that contain the most information (i.e., that extrapolate well). Heuristics were chosen to reflect obvious interpretations of the muscle frequency results in Table 3. All heuristics performed well, identifying muscle combinations representative of all primary lower extremity functions (e.g., a hip extensor, a hip adductor, a knee flexor, a knee extensor, either head of biarticular gastrocnemius, soleus, tibialis anterior, and peroneus longus). Reconstruction VAF values were greater than Extrapolation VAF values by 5.5%, on average, indicating that knowledge of reconstruction VAF may be useful for estimating the level of Extrapolation VAF and that Extrapolation VAF can be expected to be slightly lower than reconstruction VAF.

Though our muscle selection heuristics were generated using EMG data from only two subjects, the available EMG data were extensive, and our findings provide unique information about which muscle EMG signals should be prioritized to maximize information content (Table 5). Overall, our muscle selection heuristics suggest that given a limited number of EMG channels, researchers should collect surface EMG data from muscles that would commonly be selected – uniarticular and biarticular flexor and extensor muscles from each major muscle group. For uniarticular muscles, these include a hip extensor (GlutMax), a knee extensor (VasLat

over VasMed), an ankle plantarflexor (Sol), and an ankle dorsiflexor (TibAnt). No uniarticular hip flexor (Iliopsoas) or uniarticular knee flexor (BiFemShort) were included as these muscles are difficult to measure with surface electrodes. For biarticular muscles, selected muscles include a posterior thigh muscle (SemiMemb, SemiTend, or BiFemLong – no clear preference), possibly an anterior thigh muscle (RF), and a posterior calf muscle (GasMed or GasLat). Our muscle selection heuristics also suggest that researchers should consider collecting surface EMG data from some muscles that are not commonly selected - frontal plane muscles (AddLong or AddMag, TFL, and PerLong) spanning all three joints. Identification of these additional muscles is not surprising given the unique stabilizing roles they play in the frontal plane. Other frontal plane stabilizers (GlutMed, Grac) were not identified, possibly due to the difficulty in measuring these muscles reliably with surface EMG. However, from Table 2, if we added one more muscle to the list, it appears that GlutMed would be a reasonable choice. Interestingly, VasLat appears to be preferable over VasMed, while GasMed and GasLat, or SemiMem and BiFemLong, appear in the same list, suggesting that EMG data from these muscles may contain more independent information than commonly believed. While these muscle selection heuristics are useful, we emphasize that they should be viewed with caution given the limited number of subjects analyzed.

Our methodology was motivated by the potential utility of muscle synergies in predictive modeling applications. Rather than calculating a VAF value for the entire constructed muscle set, VAF values were calculated for each individual muscle and then averaged. In this way, a poorly constructed muscle would have greater influence on the measures of extrapolation ability. This behavior was desired since a poorly constructed muscle excitation can significantly influence a predictive walking optimization, especially if the muscle is a major contributor to a particular

lower extremity joint moment. Since muscle excitations in walking simulations are constrained to be between 0 and 1, synergy excitations obtained from synergy analysis of included muscle subsets were normalized to be between 0 and 1 prior to extrapolation. Interestingly, when synergy vectors were normalized to be between 0 and 1 instead, with the corresponding synergy excitations scaled accordingly, synergy vectors found by synergy extrapolation were less consistent with those found directly from synergy analysis of the excluded muscle excitations.

While VAF is a useful and commonly accepted measure of muscle excitation reconstruction quality, how well a reconstructed muscle excitation approximates the original muscle excitation may not be fully captured in a single percentage value. To illustrate this issue, we took sample EMG data from the paretic leg of the stroke subject walking at 0.7 m/s and applied our synergy extrapolation procedure. For this example, a single included muscle subset was selected using our muscle selection heuristics. Fig. 4 shows the synergy-constructed muscle excitations using 3, 4, and 5 synergies for muscles in the included and excluded subset pair, and VAF values corresponding to each approximation are provided in Table 6. As the number of synergies used in the analyses was increased, VAF values increased and all muscles reached 90% VAF at 5 synergies. The included muscle excitations reached higher VAF values and had superior muscle reconstruction quality in general, matching the shape and magnitude of the original muscle excitations at 5 synergies. For the excluded muscle subset, some muscle excitations were not approximated quite as well, failing to match well the shape and magnitude of the original muscle excitation. Notably, both the iliopsoas and rectus femoris in the excluded muscle subset were barely able to achieve 90% VAF at 5 synergies and were only slightly improved from their 3 and 4 synergy reconstructions. Underperforming excluded muscle excitation reconstructions such as these suggest that additional synergies extracted from this included muscle subset may not

improve construction quality further. Furthermore, to represent these excluded muscle excitations accurately in a predictive modeling simulation, it may be necessary to include these muscles in the experimental EMG recordings.

One unexpected observation was that for a given number of synergies, the non-paretic leg of the stroke subject had a slightly higher VAF than did the paretic leg (Table 5). This observation is contrary to results reported by Clark *et al.* 2010 [20]. However, when we performed synergy analysis using the same eight muscles used in that study, the VAF for the two legs became nearly equal, though the paretic leg still had higher variability. Furthermore, methodological differences exist between how we performed our synergy analyses and how Clark *et al.* performed theirs. The fact that our stroke subject was high functioning may have also affected the comparison between non-paretic and paretic legs. Importantly, as other studies have suggested, the number and choice of muscles had a significant influence on the synergy solutions in the present study [38]. Given the known heterogeneity in this clinical population, analysis of EMG data from a single stroke subject is a significant limitation.

In conclusion, this study has demonstrated that for both a healthy subject and a stroke subject, synergy excitations extracted from eight muscle excitations derived from commonly measured surface EMG signals during walking can be used to construct all remaining muscle excitations from the same data set with good accuracy. Even though the extrapolation process was more challenging for the healthy subject data set (16 excluded muscle excitations to be constructed by synergy extrapolation versus only eight for each leg of the stroke subject), Extrapolation VAF was similar for both subjects, suggesting that proper selection of included muscles, and use of a sufficient number of included muscles, are the most critical issues. These findings support the use of synergy excitations derived from experimental EMG signals to

construct muscle excitations for missing EMG signals as part of a muscle force optimization. However, construction of the missing EMG signals is most likely to work well if the collected EMG signals follow our proposed muscle selection heuristics.

Acknowledgments: This study was funded by NSF REU fellowship CBET 1430584 and NSF grant CBET 1159735.

Nomenclature:

AddLong = adductor longus
AddMag = adductor magnus
BiFemLong = biceps femoris long head
BiFemShort = biceps femoris short head
EMG = electromyography
GasLat = lateral gastrocnemius
GasMed = medial gastrocnemius
GlutMax = gluteus maximus
GlutMed = gluteus medius
Grac = gracilis
PerLong = peroneus longus
RF = rectus femoris
RF2 = rectus femoris (second recording location)
SemiMemb = semimembranosus
SemiTend = semitendinosus
Sol = soleus
TFL = tensor fasciae latae
TibAnt = tibialis anterior
VAF = variance accounted for
VasLat = vastus lateralis
VasMed = vastus medialis

REFERENCES

- [1] Erdemir, A., McLean, S., Herzog, W., and van den Bogert, A. J., 2007, “Model-based estimation of muscle forces exerted during movements,” *Clin. Biomech.*, **22**(2), pp. 131–154.
- [2] Crowninshield, R. D., and Brand, R. A., 1981, “A physiologically based criterion of muscle force prediction in locomotion,” *J. Biomech.*, **14**(11), pp. 793–801.

- [3] Anderson, F. C., and Pandy, M. G., 2001, "Static and dynamic optimization solutions for gait are practically equivalent," *J. Biomech.*, **34**(2), pp. 153–161.
- [4] Neptune, R. R., Sasaki, K., and Kautz, S. A., 2008, "The effect of walking speed on muscle function and mechanical energetics," *Gait Posture*, **28**(1), pp. 135–143.
- [5] Pandy, M. G., Lin, Y. C., and Kim, H. J., 2010, "Muscle coordination of mediolateral balance in normal walking," *J. Biomech.*, **43**(11), pp. 2055–2064.
- [6] Lloyd, D. G., and Besier, T. F., 2003, "An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo," *J. Biomech.*, **36**(6), pp. 765–776.
- [7] Doorenbosch, C. A. M., and Harlaar, J., 2003, "A clinically applicable EMG-force model to quantify active stabilization of the knee after a lesion of the anterior cruciate ligament," *Clin. Biomech.*, **18**(2), pp. 142–149.
- [8] Buchanan, T. S., Lloyd, D. G., Manal, K., and Besier, T. F., 2004, "Neuromusculoskeletal modeling: estimation of muscle forces and joint moments and movements from measurements of neural command," *J. Appl. Biomech.*, **20**(4), pp. 367–395.
- [9] Sartori, M., Reggiani, M., Farina, D., and Lloyd, D. G., 2012, "EMG-driven forward-dynamic estimation of muscle force and joint moment about multiple degrees of freedom in the human lower extremity," *PLoS One*, **7**(12), p. e52618.
- [10] De Luca, C. J., 1997, "The use of surface electromyography in biomechanics," *J. Appl. Biomech.*, **13**(2), pp. 135–163.
- [11] Disselhorst-Klug, C., Schmitz-Rode, T., and Rau, G., 2009, "Surface electromyography and muscle force: Limits in sEMG-force relationship and new approaches for applications," *Clin. Biomech.*, **24**(3), pp. 225–235.
- [12] Sartori, M., Farina, D., and Lloyd, D. G., 2014, "Hybrid neuromusculoskeletal modeling best tracks joint moments using a balance between muscle excitations derived from electromyograms and optimization," *J. Biomech.*, **47**(15), pp. 3613–3621.
- [13] Ting, L. H., Chvatal, S. A., Safavynia, S. A., and McKay, J. L., 2012, "Review and perspective: neuromechanical considerations for predicting muscle activation patterns for movement," *Int. j. numer. method. biomed. eng.*, **28**(10), pp. 1003–1014.
- [14] Fregly, B. J., Besier, T. F., Lloyd, D. G., Delp, S. L., Banks, S. A., Pandy, M. G., and D'Lima, D. D., 2012, "Grand challenge competition to predict in vivo knee loads," *J. Orthop. Res.*, **30**(4), pp. 503–513.
- [15] Ting, L. H., and Chvatal, S. A., 2010, "Decomposing muscle activity in motor tasks: methods and interpretation," *Mot. Control Theor. Exp. Appl.*
- [16] Lee, D. D., and Seung, H. S., 1999, "Learning the parts of objects by non-negative matrix factorization," *Nature*, **401**, pp. 788–791.
- [17] Tresch, M. C., Saltiel, P., and Bizzi, E., 1999, "The construction of movement by the spinal cord," *Nat. Neurosci.*, **2**(2), pp. 162–167.
- [18] Ivanenko, Y. P., Poppele, R. E., and Lacquaniti, F., 2004, "Five basic muscle activation patterns account for muscle activity during human locomotion," *J. Physiol.*, **556**(Pt. 1), pp. 267–282.
- [19] Cappellini, G., Ivanenko, Y. P., Poppele, R. E., and Lacquaniti, F., 2006, "Motor patterns

- in human walking and running,” *J. Neurophysiol.*, **95**(6), pp. 3426–3437.
- [20] Clark, D. J., Ting, L. H., Zajac, F. E., Neptune, R. R., and Kautz, S. A., 2010, “Merging of healthy motor modules predicts reduced locomotor performance and muscle coordination complexity post-stroke,” *J. Neurophysiol.*, **103**(2), pp. 844–857.
 - [21] Rodriguez, K. L., Roemmich, R. T., Cam, B., Fregly, B. J., and Hass, C. J., 2013, “Persons with Parkinson’s disease exhibit decreased neuromuscular complexity during gait,” *Clin. Neurophysiol.*, **124**(7), pp. 1390–1397.
 - [22] Allen, J. L., and Neptune, R. R., 2012, “Three-dimensional modular control of human walking,” *J. Biomech.*, **45**(12), pp. 2157–2163.
 - [23] Allen, J. L., Kautz, S. A., and Neptune, R. R., 2013, “The influence of merged muscle excitation modules on post-stroke hemiparetic walking performance,” *Clin. Biomech.*, **28**(6), pp. 697–704.
 - [24] McGowan, C. P., Neptune, R. R., Clark, D. J., and Kautz, S. A., 2011, “Modular control of human walking: Adaptations to altered mechanical demands,” **43**(3), pp. 1–15.
 - [25] Neptune, R. R., Clark, D. J., and Kautz, S. A., 2009, “Modular control of human walking: A simulation study,” *J. Biomech.*, **42**(9), pp. 1282–1287.
 - [26] Sartori, M., Gizzi, L., Lloyd, D. G., and Farina, D., 2013, “A musculoskeletal model of human locomotion driven by a low dimensional set of impulsive excitation primitives,” *Front. Comput. Neurosci.*, **7**, p. 79.
 - [27] Walter, J. P., Kinney, A. L., Banks, S. A., D’Lima, D. D., Besier, T. F., Lloyd, D. G., and Fregly, B. J., 2014, “Muscle synergies may improve optimization prediction of knee contact forces during walking,” *J. Biomech. Eng.*, **136**(2), p. 21031.
 - [28] Meyer, A. J., Eskinazi, I., Jackson, J. N., Rao, A. V., Patten, C., and Fregly, B. J., 2016, “Muscle Synergies Facilitate Computational Prediction of Subject-Specific Walking Motions,” *Front. Bioeng. Biotechnol.*, **4**(77).
 - [29] Hug, F., 2011, “Can muscle coordination be precisely studied by surface electromyography?,” *J. Electromyogr. Kinesiol.*, **21**(1), pp. 1–12.
 - [30] Jonkers, I., Spaepen, a., Papaioannou, G., and Stewart, C., 2002, “An EMG-based, muscle driven forward simulation of single support phase of gait,” *J. Biomech.*, **35**(5), pp. 609–619.
 - [31] Shao, Q., Bassett, D. N., Manal, K., and Buchanan, T. S., 2009, “An EMG-driven model to estimate muscle forces and joint moments in stroke patients,” *Comput. Biol. Med.*, **39**(12), pp. 1083–1088.
 - [32] De Groote, F., Jonkers, I., and Duysens, J., 2014, “Task constraints and minimization of muscle effort result in a small number of muscle synergies during gait,” *Front. Comput. Neurosci.*, **8**, p. 115.
 - [33] Berry, M. W., Browne, M., Langville, A. N., Pauca, V. P., and Plemmons, R. J., 2007, “Algorithms and applications for approximate nonnegative matrix factorization,” *Comput. Stat. Data Anal.*, **52**(1), pp. 155–173.
 - [34] Serrancolí, G., Kinney, A. L., Fregly, B. J., and Font-Llagunes, J. M., 2016, “Neuromusculoskeletal model calibration significantly affects predicted knee contact

- forces for walking,” *J. Biomech. Eng.*, **138**(81001).
- [35] Chvatal, S. A., and Ting, L. H., 2013, “Common muscle synergies for balance and walking,” *Front. Comput. Neurosci.*, **7**, p. 48.
 - [36] Routson, R. L., Clark, D. J., Bowden, M. G., Kautz, S. a, and Neptune, R. R., 2013, “The influence of locomotor rehabilitation on module quality and post-stroke hemiparetic walking performance.,” *Gait Posture*, **38**(3), pp. 511–517.
 - [37] Gonzalez-Vargas, J., Sartori, M., Dosen, S., Torricelli, D., Pons, J. L., and Farina, D., 2015, “A predictive model of muscle excitations based on muscle modularity for a large repertoire of human locomotion conditions,” *Front. Comput. Neurosci.*, **9**(114).
 - [38] Steele, K. M., Tresch, M. C., and Perreault, E. J., 2013, “The number and choice of muscles impact the results of muscle synergy analyses,” *Front. Comput. Neurosci.*, **7**, p. 105.

FIGURE CAPTIONS

Figure 1. Overview of “synergy extrapolation” approach. All possible combinations of eight surface EMG measurements were identified and labeled as the “included” muscle subsets for each full data set (healthy, stroke non-paretic leg, stroke paretic leg). Synergy analysis was performed on each included subset to determine a set of three to five time-varying “synergy excitations” that reconstructed the eight surface EMG signals. The subset of remaining muscle excitations not included in the synergy analysis was labeled as the “excluded” muscle subset and was fitted with the extracted synergy excitations through linear least-squares regression. The constructed excluded muscle excitations were compared to the experimental excluded muscle excitations using variance accounted for (VAF) calculations. VAF values were averaged across muscles and gait cycles to create a single Extrapolation VAF value for each included-excluded muscle subset pair. This synergy extrapolation approach was repeated for each walking speed of all three data sets.

Figure 2. Percentage of included muscle combinations that achieved a specified Extrapolation VAF value (80%, 85%, or 90%) for the healthy subject (top), the non-paretic leg of the stroke subject (middle), and the paretic leg of the stroke subject (bottom) across all walking speeds. Extrapolation VAF results were generated using five synergy excitations. Extrapolation ability varied across walking speeds, but no consistent trend appeared across the three data sets. The percentage of acceptable muscle combinations changed significantly with relatively small changes in VAF cutoff value. For these results in absolute numbers instead of percentage values, refer to Table 2.

Figure 3. Mean Extrapolation VAF for the Top 10% of Combinations averaged across walking trials (black circles) for the healthy subject (left), the non-paretic leg of the stroke subject (middle), and the paretic leg of the stroke subject (right) using three to five synergies. Error bars representing plus or minus one standard deviation demonstrate the variability across walking trials for each data set. Reference lines represent VAF cutoff values for typical synergy analyses (90% and 95%). Using five synergies, all three data sets achieved on average a Mean Extrapolation VAF of 90% or greater across all walking speeds. The paretic leg data set exhibited a reduced extrapolation ability with more variability across walking speeds compared to the non-paretic leg and the healthy data sets.

Figure 4. Comparison of reconstructed and experimental (black) muscle excitations for (a) eight included muscle excitations and (b) eight excluded muscle excitations using three (red), four (green), and five (blue) synergies. Sample EMG data were taken from the paretic leg of the stroke subject walking at 0.7 m/s. The included muscle excitations were selected based on the previously described muscle selection heuristics and were reconstructed using the synergy excitations and synergy vectors obtained from synergy analysis. The excluded muscle excitations were reconstructed using the synergy excitations obtained from the included muscles and our linear least-squares implementation of synergy extrapolation. VAF values for each reconstructed muscle excitation are provided in Table 6.

TABLE CAPTIONS

Table 1. Experimental surface and fine-wire EMG measurements recorded for this study. EMG measurements from two different locations of the rectus femoris were included since both contained unique signal characteristics.

Table 2. Number of included muscle combinations that were able to achieve 90% Extrapolation VAF for three to five synergy excitations. Analyses using four and five synergy excitations produced favorable extrapolation results for most walking speeds, while almost no acceptable combinations were found when using three synergy excitations.

Table 3. Frequency with which each muscle appeared in the Top 10% of Combinations across all walking trials. Percent values were calculated for each data set separately (healthy, stroke non-paretic leg, stroke paretic leg). Muscles were separated into groups based on their primary lower extremity functions: hip uniarticular muscles, hip/knee biarticular muscles, knee uniarticular muscles, knee/ankle biarticular muscles, and ankle uniarticular muscles.

Table 4. Muscle selection heuristics developed based on Table 3 to identify muscle combinations that extrapolate accurately.

Table 5. Muscle combinations selected from heuristics defined in Table 4. For these muscle combinations, the Extrapolation VAF exceeded 90% across all data sets and heuristics when using five synergy excitations. These muscle combinations also performed well using three and four synergies. Reconstruction VAF values were on average greater than Extrapolation VAF

values by 5.5%. Muscle combinations generally followed a similar distribution across groups representing all primary lower extremity functions.

Table 6. VAF values for sample muscle excitation reconstructions presented in Fig. 4 using 3, 4, and 5 synergy excitations. For 5 synergies, all muscle excitations, included and excluded, were reconstructed with at least 90% VAF.