A factorization-based algorithm to predict EMG data using only kinematics information

Marta Manzano[a] and Gil Serrancolí[a]*

[a]*Department of Mechanical Engineering, Universitat Politècnica de Catalunya, Barcelona, Catalonia, Spain*

Gil Serrancolí

Simulation and Movement Analysis Lab (SIMMA Lab)

Barcelona East School of Engineering (EEBE)

Universitat Politècnica de Catalunya

Av. Eduard Maristany 16

08019 Barcelona, Catalonia, Spain

A factorization-based algorithm to predict EMG data using only kinematics information

EMG analyses have several applications, such as identifying muscle excitation patterns during rehabilitation or training plans, or controlling EMG-driven devices. However, experimental measurements can be time consuming or difficult to obtain. This study presents a simple algorithm to predict EMG signals that can be applied in real time during running, given only the instantaneous vector of kinematics. We hypothesize that the factorization of the kinematics of the skeleton together with the EMG data of calibration subjects could be used to predict EMG data of another subject only using the kinematic information. The results showed that EMG signals of lower-limb muscles can be predicted accurately in less than a second using this method. Correlation coefficients between predicted and experimental EMG signals were higher than 0.7 in ten out of eleven muscles for most prediction trials and subjects, and their overall median value was higher than 0.8. These values confirm that this method could be used to accurately predict EMG signals in real time when only kinematics are measured.

Keywords: EMG prediction; muscle excitation; signal factorization; kinematic and muscle synergies

Introduction

The information provided by the measured EMG data can be crucial in different biomechanical contexts, such as athletic training [1] or rehabilitation programs [2], or to control a device driven by muscle excitations [3]. The measurement of EMG data provides information that helps in the diagnosis of different mobility impairments. In addition, it can be used to follow up rehabilitation programs in subjects who need to recover mobility, or to design subject-specific athletic training programs to enhance the performance of the muscles. The measurement of EMG data can be used to establish subject-specific rehabilitation programs for subjects with mobility impairments. However, the time and economic resources needed when performing an electromyography (EMG) analysis, as well as the acquisition of EMG data from a subject wearing a device, could become drawbacks. If the measurements are performed non-invasively, the electrodes, attached to the skin surface, might fall or reduce their contact with the movement of the patient, which can delay the process. If we could predict EMG data accurately using other available data (e.g. kinematics data) instead of measuring muscle activity experimentally, the process would be more efficient, less time consuming and more comfortable for both the injured subject and the clinician.

Research has been done over recent years in order to predict muscle activations. The delay between neural commands, represented by the electrical signal from the motoneurons (measured through EMG) and the subsequent muscle activations (result of the active contraction of muscle fibres) are usually modelled with a first-order ordinary differential equation [4,5]. On the one hand, in terms of musculoskeletal modelling, optimization techniques have been developed to estimate muscle activations and forces [6–8]. In this regard, computational algorithms based on multibody techniques are used to track or predict joint kinematics and dynamics. From the obtained muscle-tendon

kinematics and a muscle model, it is possible to estimate muscle activations and forces. Although consistent musculoskeletal kinematics and dynamics are obtained, these methods require scaling the models, calibrating the neuromusculoskeletal parameters, and applying inverse or forward dynamics analyses. The methods to calibrate subject-specific parameters are time consuming, and they usually cannot be used in real time due to the high computational cost. On the other hand, artificial neural networks have also been developed to predict EMG activity from kinematics data. For instance, Prentice et al. [9] developed a network to accurately predict EMG signals of eight lower-limb and trunk muscles from kinematics data (RMS > 0.9 for most muscles). However, a high number of strides are usually required to train the neural network (120 in the mentioned example).

Synergies between kinematics and EMG signals (normalized EMG) could be used to predict muscle activity using only kinematics data knowledge. The concept of muscle synergies has been used to analyze the muscle activation patterns when performing certain tasks [10–12], to control an EMG-driven device [13,14], or to analyze the differences in muscle activation patterns among different injured subjects [15,16]. The concept of muscle synergies is mathematically represented by the factorization of the matrix containing the envelope of EMG signals into two matrices of lower rank. This process is usually performed using a non-negative matrix factorization algorithm [17]. One matrix contains the time-dependent commands (with a lower number of commands than muscles) and the other contains the weights of each command to each muscle. Some studies have analyzed the synergies of EMG and kinematics data separately [18,19]. The mapping of EMG to kinematics data has been used to predict the kinematics, especially in upper limb movements [20,21]. Nevertheless, an integrated factorization of EMG and kinematics data have not been broadly explored. The possibility of predicting EMG data in real time using synergies obtained from the kinematics of the whole human body and

EMG signals of the lower limb has not been extensively studied yet. The use of human body kinematics to predict muscle activity could facilitate new technical developments using the predicted information but just measuring kinematics.

This study presents a method to predict the envelope of normalized EMG signals during running given only a vector of kinematics data at a certain time frame based on the optimization of synergy parameters obtained from another subject. We factorized scaled kinematics and normalized filtered EMG data of five subjects. After that, using the reference factorization data of this set of subjects, EMG data for other subjects were predicted. Each prediction is obtained in less than a second. Therefore, this algorithm also makes real-time use possible. In order to test the algorithm, two different sets of data were factorized to predict the EMG data a posteriori. The first one included three trials at 3 m/s from five different subjects, and the second group included three trials at 2 m/s, 3 m/s and 5 m/s from the same five subjects. The use of this method would speed up the evaluation of muscle excitations or would facilitate the control of certain EMG-driven devices, especially due to the current decreasing cost of motion capture systems.

Methods

Experimental data

The experimental data of this study consisted of EMG and kinematic running data (five steps starting from heel-strike) at four velocities (2, 3, 4, and 5 m/s). These data were extracted from the study of Hamner and Delp [22], available online, and they contain running trials of ten subjects (gender: all male, age: 29 ± 5 years, height: 1.77 ± 0.04 m, and 70.9 ± 7.0 kg). EMG data were available for 11 muscles of the right leg: biceps femoris long head (bflh), gastrocnemius lateralis (gaslat), gastrocnemius medialis (gasmed), gluteus maximus (glmax), gluteus medius (glmed), rectus femoris (rf),

semimembranosus (semimem), soleus (soleus), tibialis anterior (ta), vastus lateralis (vaslat) and vastus medialis (vasmed). These data were high-pass filtered at 10 Hz with a Butterworth filter of 3rd order, rectified and low-pass filtered at 6 Hz. After that, filtered EMG of each subject was normalized against the maximum EMG value for each muscle over the four trials at different running speeds. Kinematics data consisted of joint coordinates from a model with 29 degrees of freedom (dofs) [23], which was scaled specifically for each subject. Joint coordinates were obtained from marker trajectories using an inverse kinematics analysis in OpenSim [24] and then low-pass filtered at 6 Hz. Both EMG and kinematics were parameterized with b-splines (polynomial piecewise functions) using 20 nodes per second (similar to [25]). Velocities and accelerations of joint coordinates were computed from the b-spline representation of the joint coordinates. Finally, for all subjects, joint coordinates were demeaned (mean subtraction) and joint coordinates, velocities, and accelerations were scaled so that all values ranged between -1 and 1, ensuring that all types of variables played the same role during the optimization process. Following this approach, the scale factors for coordinates, velocities, and accelerations were 1, 10, and 100, respectively. This scaling process ensures that all kinematics and EMG data are ranged approximately between -1 and 1.

Factorization

In this phase (calibration phase), we factorized the matrices containing the experimental reference data (scaled) for three cycles (from heel strike to heel strike) of five subjects (calibration subjects). The concept of factorizing data consists of representing them with a lower number of degrees of freedom. In this study we factorized matrices containing, in columns, kinematics and EMG data. This process maps the EMG data with the kinematics. These matrices contain the scaled relative joint coordinates, all 29 scaled velocities and accelerations, and normalized EMG. Global joint coordinates (3

translations and 3 rotations of the pelvis with respect to the ground) were removed from the set to avoid having solutions dependent on the position of the subject with respect to the laboratory reference. These data were time normalized by each cycle to have 101 frames per cycle.

Each resultant matrix *M* of 92 columns (see Appendix A1) was factorized with time-dependent variables (Commands – *C*) and weights for each input variable to each command (Weights – *W*), such as follows:

$$M \sim C \cdot W \tag{1}$$

where *M* is the matrix (*nframes* x *nvar*) containing kinematics and EMG data, *nframes* is the number of time frames, and *nvar* is the number of variables factorized (92 in this case). Matrix *C* is *nframes* x *nsyn* and matrix *W* is *nsyn* x *nvar*, where *nsyn* is the number of synergies (i.e. the number of commands) used to factorize the matrix *M* (Figure 1). One matrix *C* was obtained for each subject, and a common matrix *W* was obtained for all five subjects simultaneously. Note that if *nsyn* is much lower than *nvar*, the factorization allows reconstructing the signals with a much lower number of elements than the ones of matrix *M* (*nframes x nvar*). Matrices *C* and *W* were computed by solving an optimization problem minimizing the following difference:

$$\min \sum_{s}^{ntrials} \sum_{j}^{nvar} \sum_{i}^{nframes} \left(M_{ijs} - C_{iks} \cdot W_{kj} \right)^2 \tag{2}$$

subject to boundary constraints of the weights *W* (the module of the weight vectors constrained to be equal to one). Note that since matrices *M* and *C* contain both positive and negative values, which can lead to negative values in the reconstructed EMG data, all negative predicted EMG values were truncated to zero.

In total, 40 sets of factorization variables were calculated for: a) 20 different number of synergies (*nsyn*), from 1 to 20, i.e., the optimization was run 20 times to obtain

combinations of *C* and *W*, with a number of synergies (commands) from 1 to 20; b) two sets of input trials, using either three trials simultaneously at 3 m/s (intermediate velocity) or three trials simultaneously at 2, 3 and 5 m/s; c) five subjects simultaneously. Case b) was included to check whether the information of trials at different speeds improved the predictions compared to results when using only trials at a single speed. For each factorization (a) and set of input trials (b), a single matrix *W* was estimated, so that the contribution to each command was the same independently of the subject and velocity, and 15 different matrices *C* were obtained, one for each velocity and subject.

In brief, the optimization problems solved during the factorization consist of minimizing Eq. (2) (objective function) while constraining the module of the weight vectors (rows of matrix *W*) to be equal to 1 (equality constraints). These optimization problems were solved using *fmincon* (with an interior-point algorithm) in 64-bit MATLAB (using a standard laptop IntelCore i7-6700HQ CPU 2.6GHz, 16 GB RAM).

Prediction of the EMG envelope

The prediction of the EMG envelope was performed for the other five subjects (prediction subjects) of the data set. The input data to the algorithm were only the experimental kinematics of the current frame for the prediction subject (v_{kin}). A combination of factorization terms obtained in the previous phase multiplied by *nsyn* scale factors *f* was searched to best fit the new kinematics vector v_{kin} (Figure 1). In this phase, an optimization problem was run frame by frame to find the proper scale factors *f*. For example, when predicting EMG data for a prediction subject, we scaled the previously obtained matrix *W* so that the current kinematics vector (single time frame) of the prediction subject fits one of the frames of the reconstructed kinematics data of one calibration subject (Figure 1, with Subject P – prediction – and Subjects C – calibration). In this prediction phase, the input data were the matrices of commands *C* and the weights *W* (with *nvar*

columns) obtained in the calibration phase. The first 81 ($nvar_{kin}$) columns of W contained information from the joint coordinates, velocities, and accelerations W_{kin}, and the last 11 columns ($nvar_{EMG}$) contained information from the EMG data W_{EMG}. This process consisted of scaling the weighting factors so that the commands of one data trial (C_{iks}) multiplied by these new scaled weights ($W_{kin_{kj}} \cdot f_k$) fit the kinematics data:

$$\min\left(\min_{i} \sum_{j}^{nvar_{kin}} \left(v_{kin\,j} - C_{iks} \cdot W_{kin_{kj}} \cdot f_k \right)^2 \right) \tag{3}$$

where v_{kin} is the vector containing joint angles, velocities, and accelerations at the current time frame and f_k are the optimized scale factors. For each set of factorization variables, the predictions were calculated frame by frame for the trials at 2, 3, 4, and 5 m/s. *fmincon* was also used to solve the optimization.

Note that this algorithm uses gradient-based methods, and Eq. (3) is not continuous. Therefore, to evaluate whether or not using the non-continuous cost function with *fmincon* leads to local minima, we also performed a test consisting of solving the optimization problem using a global optimization algorithm (*ga* in MATLAB) in one case. This test was carried out for the prediction of EMG of one subject using the factorization variables with five synergies, at 2, 3, 4, and 5 m/s.

Multiway ANOVA tests were used to analyse the influence of the velocities used in the calibration phase, the velocity of the prediction trial, and the subject, on the Pearson correlation coefficients and RMS values for muscle predictions. Differences were considered significant when p-value < 0.05.

Additionally, we also tested the algorithm with a set of data from another database. EMG data were predicted for two overground running trials at 5 and 7 m/s using data available from Dorn et al. [26]. The calibration data were commands and weight variables of running trials at 2, 3 and 5 m/s.

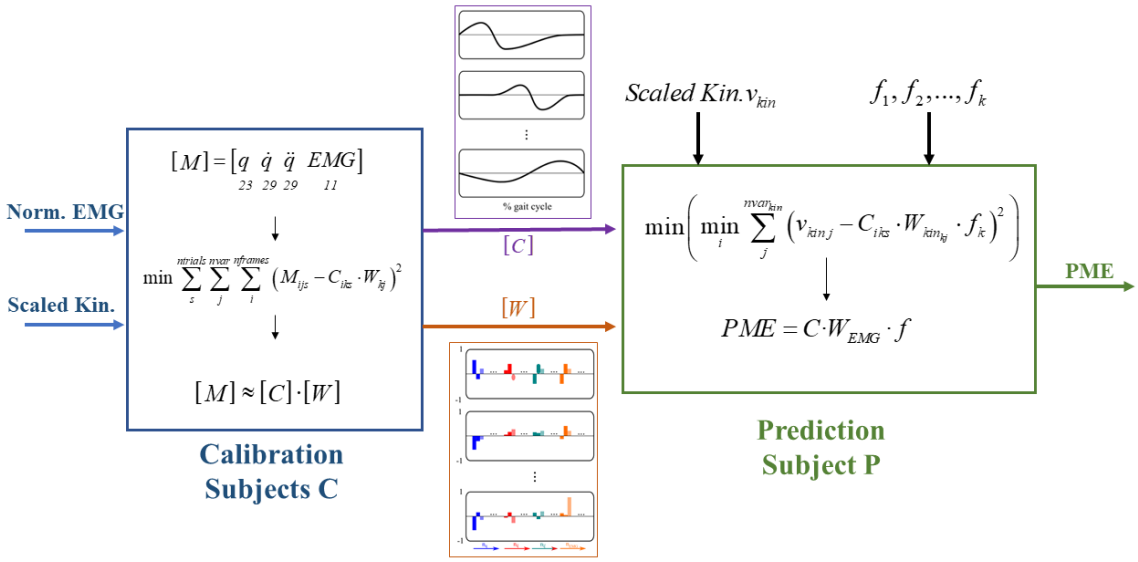

Figure 1. Process to predict muscle excitations (PME as normalized EMG) using three trials from five subjects in the calibration phase. *M* is the matrix containing, in columns, the scaled kinematics and normalized EMG of one of the calibration subjects (Subject C, calibration subject). The numbers below angular coordinates q, velocities $\dot{q}$, accelerations $\ddot{q}$, and normalized EMG stand for the number of columns. *C* and *W* are matrices containing the commands and the weights of the factorization, respectively. *i* is the time frame, *j* is the signal, *k* is the synergy and *s* is the trial. v_{kin} is the vector containing the kinematics data of Subject P (prediction subject) at one frame, and *f* is the vector of scale factors multiplying the matrix *W*. In the calibration phase, *C* and *W* are estimated from data of Subjects C and in the prediction phase *f* vector is estimated using *C*, *W* and the v_{kin} to predict normalized EMG signals.

Results

The estimated values of the kinematics (coordinates, velocities, and accelerations) and EMG signals after the factorization (calibration phase) reproduced all experimental data quite accurately, especially when using more than eight synergies ($r > 0.9$ when using trials at a single velocity, Figure 2a). The fact of using different velocities to factorize the signals slightly influenced the accuracy of the reconstruction (Figure 2b).

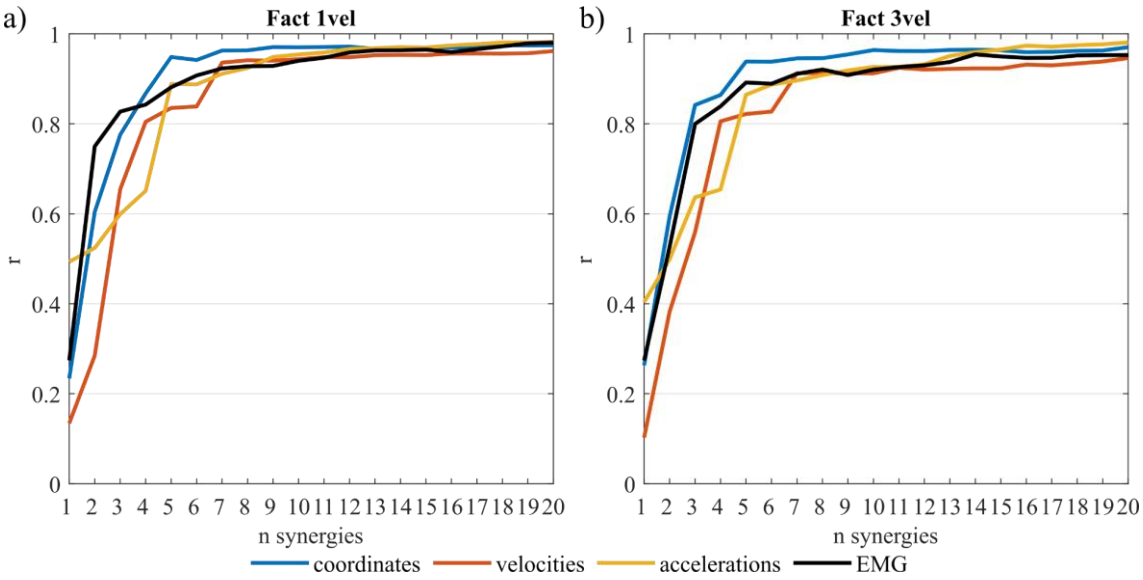

Figure 2. Median correlation coefficients between the experimental and the reconstructed data (angles, velocities, accelerations, and EMG signals) obtained at the calibration phase by each number of synergies for all subjects. Left: using three trials for each subject at 3 m/s. Right: using one trial at 2, 3 and 5 m/s for each subject.

Both correlation coefficients and RMS difference values between the predictions of EMG data and the corresponding experimental values show that the best predictions were obtained using a low number of synergies (Figure 3). In terms of RMS values, the lowest median values for EMG data were obtained using from two to five synergies with both calibration sets (median RMS < 0.13). In terms of correlation coefficient values, the best predictions were obtained using between three and five synergies with the calibration trials at 3 m/s (median r > 0.8) and between four and five synergies with the calibration trials at three velocities (median r > 0.8). In 10 out of 11 muscles, the predictions were obtained with r > 0.7 when factorizing with data at 3 m/s (the exception was the ta) and in 9 out of 11 muscles when using calibration trials at three different velocities (the exceptions were ta and bflh). Figure 3 also shows that the higher is the number of synergies, the worse is the prediction of tibialis anterior (ta) and gluteus maximus (glmax), especially when using all calibration trials at 3 m/s.

Figure 4 shows an example of EMG predictions of a trial at 4 m/s. In that case, factorization data using five synergies from calibration trials at 3 m/s were used. Predictions of gaslat, gasmed, semimem, soleus, vaslat and vasmed are accurate in terms of normalized magnitude (mean RMS > 0.8) and shape (mean r > 0.8). Bflh, glmax and ta had mean r values between 0.7 and 0.8, and RMS < 0.2, and glmed and rf had mean r values ≥ 0.85, but with RMS values slightly higher (0.15 and 0.23, respectively). The magnitude of all predicted signals from this particular example is overall slightly lower than the experimental values. Except for the small peaks of gasmed at the flight phase, the timing of the peaks for all muscles was predicted precisely. In all muscles, the onset-offset periods were also predicted accurately.

The shape was better reproduced for the prediction trials at 3 and 4 m/s ($p < 0.01$), independently of the factorization used (Figure 5), whereas the magnitude was better predicted for prediction trials at 2 and 3 m/s ($p < 0.01$), also independently of the factorization used (Figure 5).

The algorithm could also predict the data of another database. For instance, when testing the algorithm with the second database [26] at 5 m/s and using five synergies, five muscles had r values higher than 0.75 (bflh, glmax, glmed, vaslat and vasmed), four muscles between 0.5 and 0.75 (gaslat and gasmed, semimem, and soleus), and tibant and recfem with r < 0.5. At 7 m/s (velocity higher than the used during the calibration), EMG of two muscles were predicted with r > 0.75 (gaslat and gasmed), two muscles between 0.5 and 0.75 (soleus and vasmed), four muscles with r between 0.4 and 0.5 (bflh, glmax, glmed, and vaslat), and the other muscles with r < 0.4 (recfem, semimem, and tibant).

The use of a gradient-based method to solve for the scale factors of the factorization variables did not seem to lead to differences compared to the use of a global optimization approach. The results showed that the correlation coefficients between the

EMG data obtained using *fmincon* and *ga* were 0.9909 ± 0.0031 at 2 m/s, 0.9949 ± 0.0015 at 3 m/s, 0.9960 ± 0.0012 at 4 m/s and 0.9970 ± 0.0010 at 5 m/s.

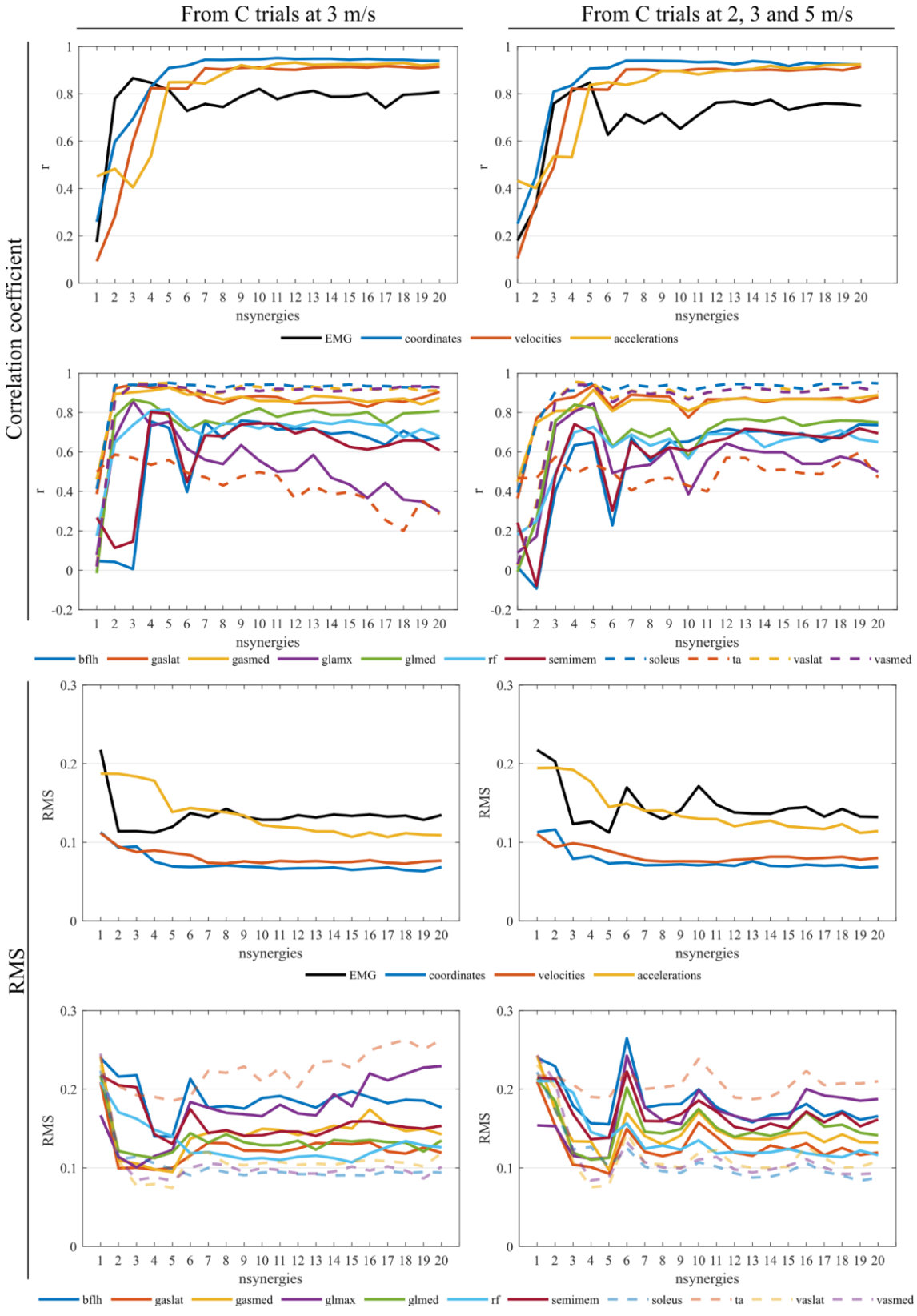

Figure 3. Median of correlation coefficients (first two rows) and root mean square errors (last two rows) between the EMG predictions and the normalized experimental EMG data for all prediction subjects. The left column contains the predictions using factorization data from a calibration C using three trials at 3 m/s for all calibration subjects, and the right column from calibration data using three trials at 3, 4 and 5 m/s for all calibration subjects. The first and third rows represent the median values for each type of variable, and the second and fourth rows, the EMG predictions for each muscle. All predictions are for all 20 factorization sets, according to the number of synergies used (from 1 to 20).

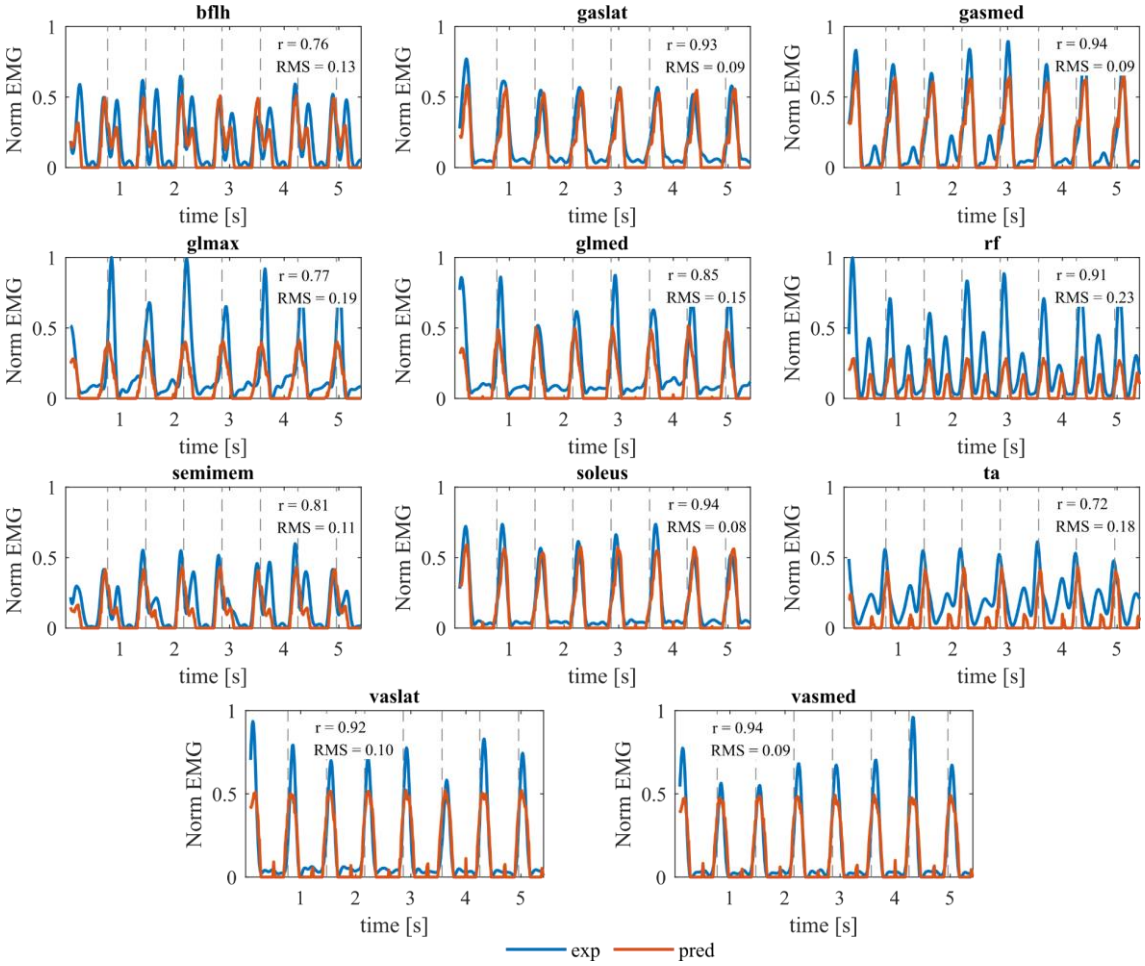

Figure 4. Example of normalized EMG predictions for one subject at 4 m/s obtained using the calibration data factorizing calibration trials (3 m/s) with five synergies. Vertical dashed lines indicate the heel-strike events.

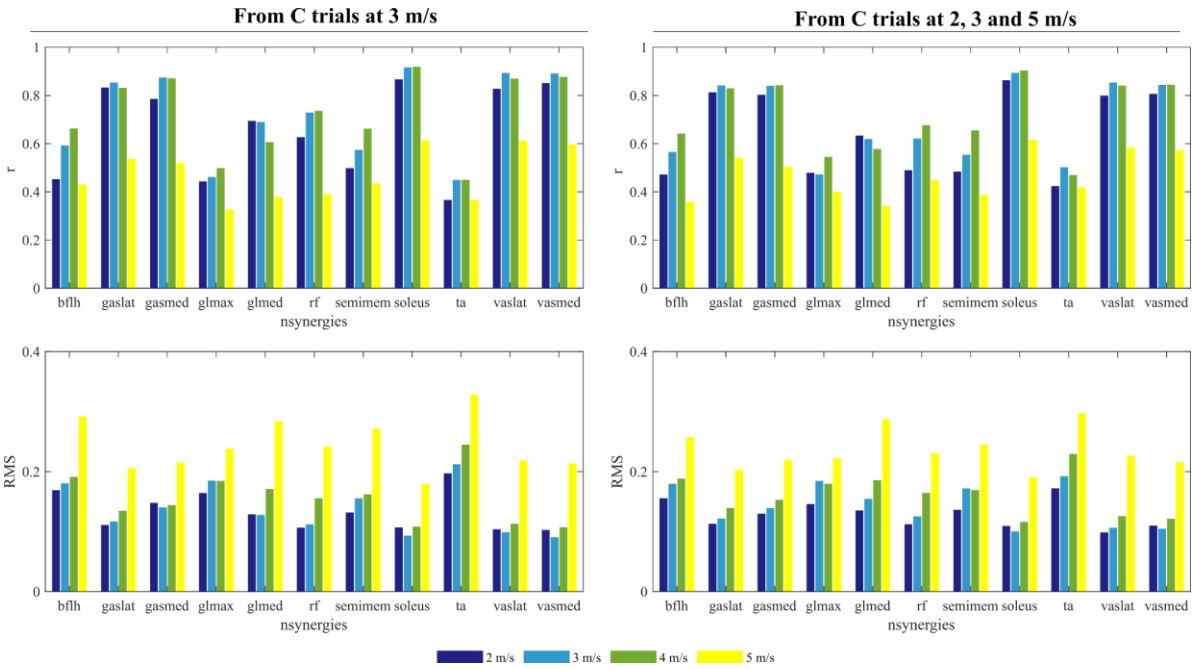

Figure 5. Mean of correlation coefficients (first row) and RMS values (second row) between EMG predictions and normalized experimental EMG data over all 21 factorization sets. The predictions are obtained using the factorization from calibration trials at 3 m/s (left column) and from calibration trials at 2, 3 and 5 m/s (right column).

Discussion

The goal of this study was to investigate the accuracy of the predictions of EMG data during running when using only experimental kinematic data. This method consists of factorizing, in the first place, the matrix of kinematics (23 angular coordinates, and 29 velocities and accelerations of the skeletal model) and normalized EMG data (11 muscles) into two matrices of lower rank, containing the time-dependent commands and the non-time-dependent weights of each signal to each command. Secondly, it predicts muscle excitations (represented with the normalized EMG in this study) by giving only the vector of the experimental kinematic data at one frame. The predictions took 0.287 ± 0.068 seconds using factorized data of trials at 3 m/s, and 0.304 ± 0.070 seconds using factorized data at three different trials (test using five synergies, i.e. five commands). If the

kinematics data could be obtained at least every 0.3 seconds, this method could predict the EMG data at this frequency.

The algorithm optimized scale factors of the weights so that the reconstructed kinematics at that frame matched with one time frame of the reference kinematics. Then, with the reference factorization terms and the new scale factors, we predicted the normalized EMG signals for the 11 muscles. The results showed that we can obtain accurate EMG predictions using this method (correlation coefficients higher than 0.7 for 10 of the 11 muscles). These predictions are better than the ones typically obtained using musculoskeletal systems [27–29] and similar to the ones obtained with complex neural networks [9].

We analyzed whether the results were influenced by the number of synergies and the velocities used to factorize data. We obtained a higher accuracy in the reproduction of kinematics and EMG data from intrasubject factorization analysis as we increased the number of synergies (using commands and weights of the same subject). However, no improvement in EMG predictions was observed when increasing the number of synergies (using commands and weights from the other subjects) above five. The best overall EMG predictions were obtained using four and five synergies. This conclusion is very close to the five muscle synergies used to account for the variability of muscle excitation patterns in other studies [30,31]. The fact of obtaining a similar optimum number of synergies when factorizing kinematics and EMG simultaneously to other studies factorizing only EMG data was not obvious a priori, since vector weights could have contained only non-zero kinematic weights. The presence of such a high number of variables during the factorization (81 kinematic variables and 11 EMG signals) gives more information to the system, establishing how the different angular coordinates, velocities and accelerations

might affect the EMG data. In the end, this allows to create a more accurate algorithm that will be able to predict more effectively the muscle excitation patterns of a subject.

Muscle prediction accuracy varies from muscle to muscle. This could be related to the number of joints the muscle is spanning. Although it is considered that the CNS controls the movement synergistically, it seems obvious that EMG data of one muscle would be more related to the coordinates, velocities, and accelerations of the joint/s that the muscle is spanning. Therefore, the fact that ta (monoarticular muscle) has the lowest prediction accuracy could be explained by the amount of directly related kinematic data included in the factorization matrix (lower than in biarticular muscles).

Since running speed may affect the pattern of muscle synergies [32], we used two sets of factorization terms. We factorized the signals using only three trials at 3 m/s per subject, and also using three trials at 2, 3 and 5 m/s per subject. The type of factorization had no significant effect on the prediction results. A possible explanation could be that the scale factors optimized during the prediction phase allow a similar kinematics data tracking, independently of the velocity. This would be the case when having commands with similar shapes and different magnitudes across velocities. Overall, the trials with the best shape prediction were the ones at 3 and 4 m/s, and the trials with the best magnitude prediction were the ones at 2 and 3 m/s, independently of the set of factorization terms.

Additionally, we also tested the capability of the algorithm to predict data from another source [26]. The predicted data were overall accurate for the trial at 5 m/s (velocity used during the calibration data). At 7 m/s, the predictions were still accurate for some muscles (especially gaslat and gasmed). Despite the protocol probably being different (difference in sensor location) and despite the calibration data being measured while running on a treadmill and the prediction data while running overground, we could closely identify the onset-offset patterns.

We used a gradient-based method to predict EMG data and a non-continuous function, which is likely to fall in local minima. Nevertheless, we also proved that a global optimization algorithm (which took around 4 seconds per frame) led to very similar results.

The presented approach has some limitations. Firstly, the analysis was performed in running data only. It is known that synergy components are task-specific [33,34]. Therefore, before applying this method to another type of movement, running style or to a completely different running velocity, reference kinematics and EMG should be collected and factorized first. Secondly, the experimental data came from healthy subjects; future lines of research will be focused on the analysis of the capability to predict muscle excitations in injured subjects. Our results suggest that the information of kinematics and EMG synergies from five subjects simultaneously can yield to the accurate prediction of EMG data in a different subject. Since we expect to have more variability of kinematics and EMG data in injured subjects than in healthy subjects, the optimization of scale factors of factorization parameters may play an important role to obtain accurate predictions. Thirdly, the non-negative factorization of EMG signals has been related to the pattern used by the central nervous system to activate the muscles (with values of factorization variables ranging from 0 – not activated – to 1 – fully activated). In our study, we factorized the data with variables that can have negative and positive values (since kinematics data could be positive or negative), whose physical meaning has a difficult interpretation. However, the results show that the algorithm can be used to predict EMG data accurately. Fourthly, reliable surface EMG data for some muscles could be challenging to obtain due to crosstalk; therefore, the predictions of their signals are expected to be poor.

In conclusion, with the proposed algorithm, we can obtain accurate EMG predictions of most muscles with only the knowledge of the vector of kinematics at a certain time frame. Since the results are obtained in less than one second, this represents an efficient method to be used while acquiring kinematics data at 0.3 seconds per frame. This could be applied for different biomechanical purposes, such as in motion analysis to perform a follow-up of a rehabilitation treatment or an athletic training program, or to control EMG-driven devices.

Acknowledgments

G. Serrancolí acknowledges the support from the Serra Húnter Program.

Declaration of interest

None

References

1. Chimera NJ, Swanik KA, Swanik CB, Straub SJ. Effects of Plyometric Training on Muscle-Activation Strategies and Performance in Female Athletes. *Journal of Athletic Training* 2004; **39**(1):24–31.
2. Suetta C, Aagaard P, Rosted A, et al. Training-induced changes in muscle CSA, muscle strength, EMG, and rate of force development in elderly subjects after long-term unilateral disuse. *Journal of Applied Physiology* 2004; **97**(5):1954–1961. doi:10.1152/japplphysiol.01307.2003.
3. Andreasen DS, Allen SK, Backus DA. Exoskeleton with EMG based active assistance for rehabilitation. *Proceedings of the 2005 IEEE 9th International Conference on Rehabilitation Robotics* 2005; **2005**:333–336. doi:10.1109/ICORR.2005.1501113.
4. Buchanan TS, Lloyd DG, Manal K, Besier TF. Neuromusculoskeletal modeling: Estimation of muscle forces and joint moments and movements from measurements of neural command. *Journal of Applied Biomechanics* 2004;

20(4):367–395.

5. Winters JM, Stark L. Estimated mechanical properties of synergistic muscles involved in movements of a variety of human joints. *Journal of Biomechanics* 1988; **21**(12):1027–1041. doi:10.1016/0021-9290(88)90249-7.

6. Campen A Van, Pipeleers G, De Groote F, Jonkers I, De Schutter J. A new method for estimating subject-specific muscle – tendon parameters of the knee joint actuators : a simulation study. *International Journal for Numerical Methods in Biomedical Engineering* 2014; **30**(10):969–987. doi:10.1002/cnm.

7. Prilutsky BI, Zatsiorsky VM. Optimization-based models of muscle coordination. *Exercise and Sport Sciences Reviews* 2002; **30**(1):32–38. doi:10.1097/00003677-200201000-00007.

8. Erdemir A, McLean S, Herzog W, van den Bogert AJ. Model-based estimation of muscle forces exerted during movements. *Clinical Biomechanics* 2007; **22**(2):131–154. doi:10.1016/j.clinbiomech.2006.09.005.

9. Prentice S, Patla A, Stacey D. Artificial neural network model for the generation of muscle activation patterns for human locomotion. *Journal of Electromyography and Kinesiology* 2001; **11**(1):19–30.

10. Turpin NA, Guével A, Dossat A, Durand S, Hug F. Recognition of muscle functional organisation in rowing by synergy identification. *Computer Methods in Biomechanics and Biomedical Engineering* 2010; **13**(sup1):141–142. doi:10.1080/10255842.2010.495876.

11. Chvatal SA, Ting LH. Common muscle synergies for balance and walking. *Frontiers in Computational Neuroscience* 2013; **7**(May):1–14. doi:10.3389/fncom.2013.00048.

12. Ivanenko YP, Poppele RE, Lacquaniti F. Motor control programs and walking. *Neuroscientist* 2006; **12**(4):339–348. doi:10.1177/1073858406287987.

13. Godfrey SB, Ajoudani A, Catalano M, Grioli G, Bicchi A. A synergy-driven approach to a myoelectric hand. *IEEE International Conference on Rehabilitation Robotics* 2013. doi:10.1109/ICORR.2013.6650377.

14. Scheme E, Englehart K. Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use.

The Journal of Rehabilitation Research and Development 2011; **48**(6):643. doi:10.1682/jrrd.2010.09.0177.

15. Cheung VCK, Piron L, Agostini M, Silvoni S, Turolla A, Bizzi E. Stability of muscle synergies for voluntary actions after cortical stroke in humans. *Proceedings of the National Academy of Sciences* 2009; **106**(46):19563–19568. doi:10.1073/pnas.0910114106.

16. Serrancolí G, Monllau JC, Font-Llagunes JM. Analysis of muscle synergies and activation-deactivation patterns in subjects with anterior cruciate ligament deficiency during walking. *Clinical Biomechanics* 2016; **31**:65–73. doi:10.1016/j.clinbiomech.2015.09.019.

17. Tresch MC, Cheung VCK, d'Avella A. Matrix Factorization Algorithms for the Identification of Muscle Synergies: Evaluation on Simulated and Experimental Data Sets. *Journal of Neurophysiology* 2006; **95**(4):2199–2212. doi:10.1152/jn.00222.2005.

18. Lambert-Shirzad N, Machiel Van Der Loos HF. On identifying kinematic and muscle synergies: A comparison of matrix factorization methods using experimental data from the healthy population. *Journal of Neurophysiology* 2017; **117**(1):290–302. doi:10.1152/jn.00435.2016.

19. Takiyama K, Yokoyama H, Kaneko N, Nakazawa K. Speed-dependent and mode-dependent modulations of spatiotem-poral modules in human locomotion extracted via tensor decom-position. *Scientific Reports* 2020; **10**(1):1–15. doi:10.1038/s41598-020-57513-w.

20. Jiang N, Vujaklija I, Rehbaum H, Graimann B, Farina D. Is accurate mapping of EMG signals on kinematics needed for precise online myoelectric control? *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 2014; **22**(3):549–558. doi:10.1109/TNSRE.2013.2287383.

21. Muceli S, Jiang N, Farina D. Extracting signals robust to electrode number and shift for online simultaneous and proportional myoelectric control by factorization algorithms. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 2014; **22**(3):623–633. doi:10.1109/TNSRE.2013.2282898.

22. Hamner SR, Delp SL. Muscle contributions to fore-aft and vertical body mass center accelerations over a range of running speeds. *Journal of Biomechanics* 2012:1–8. doi:10.1016/j.jbiomech.2012.11.024.

23. Hamner SR, Seth A, Delp SL. Muscle contributions to propulsion and support during running. *Journal of Biomechanics* 2010; **43**(14):2709–2716. doi:10.1016/j.jbiomech.2010.06.025.

24. Delp SL, Anderson FC, Arnold AS, et al. OpenSim: Open-source software to create and analyze dynamic simulations of movement. *IEEE Transactions on Biomedical Engineering* 2007; **54**(11):1940–1950. doi:10.1109/TBME.2007.901024.

25. Reinbolt JA, Haftka RT, Chmielewski TL, Fregly BJ. A computational framework to predict post-treatment outcome for gait-related disorders. *Medical Engineering and Physics* 2008; **30**(4):434–443. doi:10.1016/j.medengphy.2007.05.005.

26. Dorn TW, Schache AG, Pandy MG. Muscular strategy shift in human running: Dependence of running speed on hip and ankle muscle performance. *Journal of Experimental Biology* 2012; **215**(11):1944–1956. doi:10.1242/jeb.064527.

27. Serrancoli G, Kinney AL, Fregly BJ, Font-Llagunes JM. Neuromusculoskeletal Model Calibration Significantly Affects Predicted Knee Contact Forces for Walking. *Journal of Biomechanical Engineering* 2016; **138**(8):081001. doi:10.1115/1.4033673.

28. Serrancolí G, Kinney AL, Fregly BJ. Influence of musculoskeletal model parameter values on prediction of accurate knee contact forces during walking. *Medical Engineering and Physics* 2020; **85**:35–47. doi:10.1016/j.medengphy.2020.09.004.

29. Eskinazi I, Fregly BJ. A computational framework for simultaneous estimation of muscle and joint contact forces and body motion using optimization and surrogate modeling. *Medical Engineering and Physics* 2018; **54**:56–64. doi:10.1016/j.medengphy.2018.02.002.

30. Cappellini G, Ivanenko YP, Poppele RE, Lacquaniti F. Motor patterns in human walking and running. *Journal of Neurophysiology* 2006; **95**(6):3426–3437.

doi:10.1152/jn.00081.2006.

31. Ivanenko YP, Poppele RE, Lacquaniti F. Five basic muscle activation patterns account for muscle activity during human locomotion. 2004; **1**:267–282. doi:10.1113/jphysiol.2003.057174.

32. Kibushi B, Hagio S, Moritani T, Kouzaki M. Speed-Dependent Modulation of Muscle Activity Based on Muscle Synergies during Treadmill Walking. 2018; **12**(January):1–13. doi:10.3389/fnhum.2018.00004.

33. Cruz Ruiz AL, Pontonnier C, Sorel A, Dumont G. Identifying representative muscle synergies in overhead football throws. *Computer Methods in Biomechanics and Biomedical Engineering* 2015; **18**(sup1):1918–1919. doi:10.1080/10255842.2015.1070581.

34. Nazifi MM, Yoon HU, Beschorner K, Hur P. Shared and Task-Specific Muscle Synergies during Normal Walking and Slipping. 2017; **11**(February):1–14. doi:10.3389/fnhum.2017.00040.

Appendix A1. Data to be factorized

The data to be factorized contained 92 columns, corresponding to scaled joint coordinates (23), joint angular velocities (29), joint angular accelerations (29), and normalized filtered EMG of 11 muscles (11).

Table A1. Order and names of the degrees of freedom of joint coordinates, velocities, and accelerations, and names of the muscles of EMG data. Vel stands for angular velocity and acc for angular acceleration. Tx, ty, and tz are the three translational degrees of freedom in the space of the pelvis with respect to the ground. Muscle abbreviations are described in the Methods section.

Coordinates	**Velocities**	**Accelerations**	**Norm. EMG**
1 R Hip flexion angle	24 Pelvis tilt vel	53 Pelvis tilt acc	82 bflh
2 R Hip adduction angle	25 Pelvis list vel	54 Pelvis list acc	83 gaslat
3 R Hip rotation angle	26 Pelvis rotation vel	55 Pelvis rotation acc	84 gasmed
4 R Knee flexion angle	27 Pelvis tx vel	56 Pelvis tx acc	85 glmax
5 R Ankle angle	28 Pelvis ty vel	57 Pelvis ty acc	86 glmed
6 L Hip flexion angle	29 Pelvis tz vel	58 Pelvis tz acc	87 rf
7 L Hip adduction angle	30 R Hip flexion vel	59 R Hip flexion acc	88 semimem
8 L Hip rotation angle	31 R Hip adduction vel	60 R Hip adduction acc	89 soleus
9 L Knee flexion angle	32 R Hip rotation vel	61 R Hip rotation acc	90 ta
10 L Ankle angle	33 R Knee flexion vel	62 R Knee flexion acc	91 vaslat
11 Lumbar extension angle	34 R Ankle vel	63 R Ankle acc	92 vasmed
12 Lumbar bending angle	35 L Hip flexion vel	64 L Hip flexion acc	
13 Lumbar rotation angle	36 L Hip adduction vel	65 L Hip adduction acc	
14 R arm flexion angle	37 L Hip rotation vel	66 L Hip rotation acc	
15 R arm adduction angle	38 L Knee flexion vel	67 L Knee flexion acc	
16 R arm rotation angle	39 L Ankle angle vel	68 L Ankle angle acc	
17 R elbow flexion angle	40 Lumbar extension vel	69 Lumbar extension acc	
18 R forearm rotation angle	41 Lumbar bending vel	70 Lumbar bending acc	
19 L arm flexion angle	42 Lumbar rotation vel	71 Lumbar rotation acc	
20 L arm adduction angle	43 R arm flexion vel	72 R arm flexion acc	
21 L arm rotation angle	44 R arm adduction vel	73 R arm adduction acc	
22 L elbow flexion angle	45 R arm rotation vel	74 R arm rotation acc	
23 L forearm rotation angle	46 R elbow flexion vel	75 R elbow flexion acc	
	47 R forearm rotation vel	76 R forearm rotation acc	
	48 L arm flexion vel	77 L arm flexion acc	
	49 L arm adduction vel	78 L arm adduction acc	
	50 L arm rotation vel	79 L arm rotation acc	
	51 L elbow flexion vel	80 L elbow flexion acc	
	52 L forearm rotation vel	81 L forearm rotation acc	