

# Toward a remote assessment of walking bout and speed: application in patients with multiple sclerosis

Arash Atrsaei, Farzin Dadashi, Benoit Mariani, Roman Gonzenbach, Kamiar Aminian

**Abstract**—Gait speed as a powerful biomarker of mobility is mostly assessed in controlled environments, e.g. in the clinic. With wearable inertial sensors, gait speed can be estimated in an objective manner. However, most of the previous works have validated the gait speed estimation algorithms in clinical settings which can be different than the home assessments in which the patients demonstrate their actual performance. Moreover, to provide comfort for the users, devising an algorithm based on a single sensor setup is essential. To this end, the goal of this study was to develop and validate a new gait speed estimation method based on a machine learning approach to predict gait speed in both clinical and home assessments by a sensor on the lower back. Moreover, two methods were introduced to detect walking bouts during daily activities at home. We have validated the algorithms in 35 patients with multiple sclerosis as it often presents with mobility difficulties. Therefore, the robustness of the algorithm can be shown in an impaired or slow gait. Against silver standard multi-sensor references, we achieved a bias close to zero and a precision of 0.15 m/s for gait speed estimation. Furthermore, the proposed machine learning-based locomotion detection method had a median of 96.8% specificity, 93.0% sensitivity, 96.4% accuracy, and 78.6% F1-score in detecting walking bouts at home. The high performance of the proposed algorithm showed the feasibility of the unsupervised mobility assessment introduced in this study.

**Index Terms**—functional assessment, gait speed, inertial sensor, multiple sclerosis, unsupervised assessment, wearables

## I. INTRODUCTION

WALKING speed is a powerful clinical marker in predicting the functional status of the individuals [1]. It has been shown that gait speed can predict functional decline [2], is associated with functional status [3], and can track the stages of a mobility-related disease such as multiple sclerosis (MS) [4].

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 721577.

A. Atrsaei and K. Aminian are with the Laboratory of Movement Analysis and Measurement, École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland (e-mail: atrsaei.arash@epfl.ch, kamiar.aminian@epfl.ch).

F. Dadashi and B. Mariani are with Gait Up S.A., 1015 Lausanne, Switzerland (e-mail: farzin.dadashi@gaitup.ch, benoit.mariani@gaitup.ch)

R. Gonzenbach is with Department of Neurology, Center for Neurorehabilitation, 7317 Valens, Switzerland (e-mail: Roman.Gonzenbach@klinik-valens.ch)

Mobility impairments are a major concern in patients with MS. Due to progressive nature of the disease, these impairments can gradually lead to a decreased activity and quality of life [5]. Therefore, it is important to monitor these patients' mobility as it indicates advancing neurological problems [6].

Mobility is usually assessed by filling out questionnaires or a clinician's observation of the way the patient walks. For instance, in MS patients, the expanded disability status scale (EDSS) is a well-accepted clinical score that measures the impairments in these patients [7] in which a score higher than 4.0 is indicative of mobility impairments [8]. However, like other questionnaire-based methods, EDSS assessments can be subjective and may not report the actual performance of the patients during daily routines [9]. Furthermore, this scale is not sensitive enough to the changes in the severity of the disease [10].

To have a more objective assessment, the measurement of gait speed has been employed through timed walks in the clinic, e.g. 10-meter walk test (10MWT) [11], [12] by stop-watches, walkways, or sometimes inertial sensors. In patients with MS, It has been shown that gait speed was reduced compared to healthy controls [13]. Furthermore, gait speed significantly decreased in higher disability levels, i.e. with higher EDSS [14].

However, gait speed is often obtained in clinical environments under supervised conditions which may not reflect the performance of the patients in real-life settings [15]. Moreover, clinical assessments cannot continuously measure the amount of daily activities or more specifically amount of walking bouts. Body worn inertial measurement units (IMUs) coupled with the dedicated algorithms have the potential to overcome these challenges by estimating the walking speed during locomotion periods regardless of the measurement environment.

There are actually numerous previous works on the estimation of gait speed by IMUs on the lower limbs [16]–[20]. These methods mostly employed the zero-velocity update approach where the drift in the velocity signal caused by the integration of the acceleration is removed by detecting stance phases and updating the velocity to zero [16], [17]. Although these methods provide high accuracy due to the biomechanical nature of the lower limb during walking, a single sensor placed on trunk or wrist is preferable in free living context where the usability and comfort are better for the users [21], [22]. However, opposed to the lower-limb based methods, zero velocity update approach is challenging when

the sensor is placed on the trunk or the wrist because of the lack of motionless instances. Therefore, the methods based on trunk or wrist are mostly based on either biomechanical models [23], [24] or machine learning approaches [22], [25]–[33].

In biomechanical models, mostly the step length was modeled by an inverted pendulum and then by detecting the gait cycle events based on the peaks of the acceleration signal, the gait speed was estimated [23]. The model was further improved by considering the rolling of the ankle and the rotation of the trunk affecting the trajectory of the center of mass [24], [34]. As these biomechanical models are often validated in healthy individuals with unimpaired gait, extending them to individuals with mobility impairments can be challenging [27]. In machine learning approaches, several features were extracted mostly from the acceleration signal, and a prediction model was trained to estimate gait speed [22], [25]–[33], [35]. These features often fall into time-domain and frequency-domain categories which represent the intensity, periodicity, and posture of the movement [28]. Although in these studies, high accuracy was obtained to estimate the gait speed in supervised conditions such as laboratory environments, little is known about their performance in unsupervised daily activities and in domestic environments. For instance, in a study on MS patients to estimate the gait speed on the belt, time and frequency-domain features were extracted from both the vertical and horizontal components of the accelerometer signal on the lower back [27]. A model was trained based on a dataset of healthy adults and was tested in patients with MS in a six-minute walk test on the treadmill in the lab. The accuracy of the model was established and it was shown that the estimated gait speed had high correlations with EDSS.

To detect walking bouts based on a single IMU, the existing methods in the literature can be mostly divided into peak detection and machine learning methods. In peak detection algorithms, steps (and consequently walking bouts) can be detected for instance, from the peaks of pitch angular velocity signal of the foot [36], acceleration signal of the trunk [23], [37]–[41], or the spectrum of the accelerometer signal on the wrist [22]. As detecting those peaks require a threshold to be selected, the robustness of the algorithm is affected by inter-individual variability. Furthermore, the pattern of a signal from an individual with an impaired gait might differ from that of a healthy individual [42]. Enhancing peak detection by using advanced and complicated filters and smoothing functions [21], [42], [43] or using information from IMUs on another location [44] seems to solve the aforementioned issues. In machine learning algorithms, several features from acceleration or angular velocity signals can be mapped into the activity classes [45]–[47]. With reference systems such as video camera [47], a classifier can be trained to detect walking bouts.

However, lack of validation in daily activities and real-life conditions also exists in the literature when it comes to locomotion detection [42]. For instance, in [48], two commercial accelerometers were bench-marked in detecting walking periods in the 6-minute walk tests in the lab. In spite of obtaining a good accuracy in the lab for detecting the steps, the

performance of the system is not known during daily living activities. As it has been shown in [49], lower accuracy in detecting walking periods has been obtained in home settings compared to the laboratory-based tests. In another study, a reduction of almost 10% in the performance of the activity classification was observed when the algorithm was tested in real-life daily activities rather than laboratory data [45].

To this end, the goal of this study was to design a new method to estimate instantaneous gait speed and detect the walking bouts using a single IMU on the lower back. The main novelty of our method relied on a training phase in which a multi-sensor IMU setup was used to train the gait speed estimation and walking bout detection models. The algorithms were then tested and validated by a single IMU on the lower back both in supervised condition, i.e. at the clinic and unsupervised condition, i.e. at home. It should be mentioned that firstly, we have used a single IMU rather than a network of multiple IMUs on the body to be as less obtrusive as possible during daily activities and reduce the complexity associated with the sensor setup. Furthermore, we focused on the lower back sensor for its closeness to the center of mass compared to the wrist IMU and its ease of use compared to the foot placement.

## II. DATA COLLECTION

### A. Participants

The recruitment and clinical measurements were performed in Valens rehabilitation hospital, Valens, Switzerland. The study was approved by the ethical committee of St. Gallen Canton (ethics number: BASEC-ID 2017-01949). 35 participants with MS were recruited in this study. The demographic data was shown on Table I. The inclusion criteria for recruitment were confirmed diagnosis of MS by the criteria of McDonald 2017 [50], EDSS of 1.0 to 6.5, ability to use smartphones and sensors properly, and having enough space at home to perform 10MWT. Participants were excluded from the study if they had obvious cognitive deficits, or they were pregnant, breastfeeding, or desirous to be pregnant.

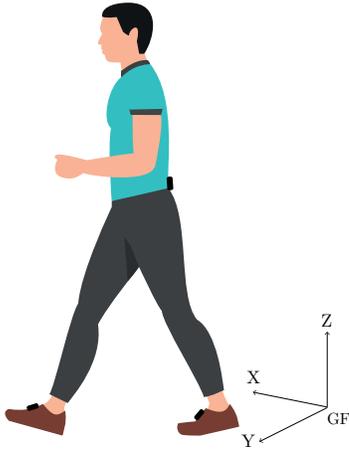
TABLE I: Demographic data of the participants

Number of participants (female)	35 (23)
Age	49.7 ± 13.0 yr
Height	172.4 ± 8.6 cm
Weight	72.4 ± 12.7 kg
EDSS	4.7 ± 1.0
min EDSS = 2.5, max EDSS = 6.5	

### B. Protocol and sensor setup

During the measurements, participants were equipped with three IMUs (Physilog 5<sup>®</sup>, Gait Up, CH), one on the waist and two on both feet (Fig. 1). It was instructed to attach the sensors by a rubber clip to the shoes and to the belt on the middle of the lower back. Each IMU included 3D accelerometer and gyroscope data recorded at a sampling rate of 128 Hz. The sensors on the feet were used to provide reference values

for training and validation of waist worn sensor algorithms (see section II-C). Participants were also given a smartphone that could provide the patients with the instructions about performing the tests. Moreover, the smartphone was used to connect to the IMUs by Bluetooth to start and stop the measurements and store the data. The data was then transferred to a secured server through internet by the same app from the smartphone.



**Fig. 1:** The location of inertial sensors on feet and belt, GF is the global frame in which the Z axis is the vertical axis, and X and Y axes lie on the horizontal plane

The data was collected in two phases. First in clinic, where the patients were asked to perform the 10MWT in three sessions with their preferred speed. During this test, patients had to walk for 10 meters, turn and walk back to their initial position. Patients were also instructed how to use the three sensors and the smartphone to become ready for the second phase happening in real life conditions. With the instructions given in clinic and also displayed on the phone, half of the patients (18 patients randomly chosen from the 35 patients) were asked to perform the 10MWT at home for 12 sessions, one session per week with their preferred speed. For each 10MWT session at home, the duration of the recorded signal as well as the 3D trajectory of the foot [16] were checked individually. If the duration of performing the test was too short, e.g. 2 or 3 seconds or patients had performed too many turns, that measurement was discarded due to its contrast to the protocol of the test. If possible, patients were requested to perform their usual daily activities while wearing the IMUs for 6 hours of measurement. They were allowed to go outside and perform their habitual activities.

### C. Reference values of the gait speed and walking bouts

Reference values for gait speed were obtained with a previously validated algorithm using IMUs on the feet of which the error was achieved as  $1.4 \pm 5.6$  cm/s in a study with healthy older and younger adults [16] or  $2.8 \pm 2.4$  cm/s in another study on patients with Parkinson disease [51]. These studies validated their method by an optical motion tracking system, i.e. gold standard method. The gait speed values obtained by this method were used for training and validation of our speed

estimation algorithm using waist sensor. In this way, we have validated our algorithm against a robust system which had been validated before against a gold standard system. We have defined such a reference as silver standard reference.

Throughout the measurements performed during daily activities at home, the aim was to determine the periods in which the participant was walking. Therefore, two classes were considered during daily activities: locomotion and non-locomotion. Foot worn IMUs were used to label the activities by applying a previously validated algorithm with an accuracy of 98% [36].

It should be noted that our reference for gait speed estimation and walking detection algorithms did not depend on sensor location on the foot. Firstly, because gait speed is determined in the global coordinate system [16], and in a global coordinate system, the placement and orientation of the attached sensor does not matter. Moreover, the reference algorithms based on the foot IMUs perform an automatic sensor to segment calibration during gait, making the reference values for locomotion periods and gait speed, independent of the sensor placement on the foot.

## III. GAIT SPEED ESTIMATION

### A. Waist vertical acceleration

To suppress the dependency of the algorithm to the orientation of the IMU attached to the belt, the vertical acceleration signal in the global frame (Z axis in Fig. 1) was chosen to extract the features [52]. In this way, a need for calibration of the sensor to anatomical frame could be avoided. Therefore, it will maintain reliability as the participants might wear the sensor in different ways throughout the day.

The accelerometer data was transformed from the IMU frame ( $a_s$ ) to the global frame ( $a_g$ ) by quaternions.

$$a_g = q \otimes [0 \quad a_s] \otimes q^* \quad (1)$$

In which  $q$  is the quaternion calculated by the fusion of the accelerometer and gyroscope data with a gradient descent algorithm [53],  $q^*$  is the quaternion conjugate, and  $\otimes$  operator is quaternion multiplication.

The vertical acceleration ( $a$ ) signal was then obtained by subtracting the gravity acceleration ( $g$ ) from the z axis component of  $a_g$ :

$$a = a_{g,z} - g \quad (2)$$

### B. Feature extraction

The vertical acceleration signal was low-pass filtered with a third-order Butterworth filter and a cut-off frequency of 3 Hz. The azimuth angle and angular velocity signals were also calculated by the method introduced by [54]. Next, the three signals (vertical acceleration, azimuth angle, and azimuth angular velocity) were divided into windows of 2 seconds with a one second overlap. For each window  $k$  a set of 18 features were extracted (Appendix I). 11 features were then selected out of 18 features by the backward elimination method according to [55]. All of these 11 features belonged to the vertical acceleration signal and reflect its statistical and frequency properties: Mean ( $x_{k,1}$ ), standard deviation ( $x_{k,2}$ ), maximum

$(x_{k,3})$ , minimum ( $x_{k,4}$ ), sum of absolute values ( $x_{k,5}$ ), sum of squared values ( $x_{k,6}$ ), maximum of the vertical velocity ( $x_{k,7}$ ), mean of the vertical velocity ( $x_{k,8}$ ), minimum of the vertical velocity ( $x_{k,9}$ ), dominant frequency ( $x_{k,10}$ ), and the amplitude of dominant frequency ( $x_{k,11}$ ) as defined in Appendix I.

The feature vector ( $x_k$ ) was then mapped into the instantaneous gait speed ( $V_k$ ):

$$V_k = f(x_k) + n \quad (3)$$

in which  $n$  is a white Gaussian noise with zero mean and covariance  $\sigma^2$ . Gaussian process regression (GPR) was used to build the regression model with MATLAB Regression Learner toolbox [56]. GPR was chosen because of its non-parametric characteristic which makes it a data-driven regression. Furthermore, being a stochastic model, GPR can estimate the uncertainty of the prediction [56], [57].

### C. Cross-validation

To evaluate the performance of the algorithm, leave-one subject-out strategy was used. For a given participant, the 10MWTs performed in the clinic by the rest of the participants were used as the training data for gait speed estimation of the 10MWT performed in the clinic by the given participant. The error was estimated as the root mean square (RMS), median, and inter-quantile range (IQR) of the difference of instantaneous gait speed estimated by the algorithm and reference value. If the participant had also performed the 10MWT and daily activities at home, the error for these two settings were also calculated. In the end, the median (bias) and IQR (precision) of each of the three parameters mentioned before were reported considering all the participants. Note that none of the measurements performed at home were used as the training dataset.

Bland-Altman plot [58] was also used to visualize the performance of the algorithm in each of the three settings, i.e. 10MWT at clinic, 10MWT at home, and daily activities at home. Furthermore, to investigate the degree of association between the reference and estimated values, the Pearson's correlation was calculated.

To demonstrate the impact of the severity of the disease on the estimated parameters, the patients were divided into two groups of mild ( $EDSS \leq 4.5$ ) and severe ( $EDSS \geq 5$ ) stage of the disease [59]. The gait speed obtained during the three assessment settings was compared between the two groups. Moreover, we compared the RMS error between the two groups, to investigate if the severity of the disease had an effect on the accuracy of the gait speed estimation algorithm. Wilcoxon rank sum test was used to show if there is a significant difference between the two groups. To assess the discriminative power and the effect size of the algorithm in predicting mild or severe stages of the disease, the Cohen's  $d$  value [60] was calculated for the gait speed and the 11 features during the 10MWT:

$$d = \frac{\sqrt{2}(\mu_{mild} - \mu_{severe})}{\sqrt{\sigma_{mild}^2 + \sigma_{severe}^2}} \quad (4)$$

in which  $\mu_{mild}$  and  $\sigma_{mild}$  are the mean and standard deviation values of a parameter from the mild group and  $\mu_{severe}$  and  $\sigma_{severe}$  belong to the severe group.

According to [60], a Cohen's  $d$  value of 0.8 and higher represents a large effect size while a value smaller than 0.4 shows a small effect size.

## IV. WALKING BOUT DETECTION

We have devised two methods to detect the walking bouts or locomotion periods: one based on a threshold on the gait speed developed from the previous section (GST Method) and the other by a machine learning approach (ML Method) based on the norm of the acceleration.

### A. GST method

We hypothesized that for gait speed values close to zero, the person is at rest and for higher speeds, the person is walking. Therefore, by applying a threshold ( $v_0$ ) on the gait speed obtained from the previous part, the periods in which the gait speed is higher than  $v_0$  are considered as locomotion while any other periods are considered as rest or non-locomotion.  $v_0$  was fixed as the optimal threshold from the receiver operating characteristic (ROC) curve in the range of 0 to 3 m/s.

### B. ML method

The gravity acceleration ( $g$ ) was first subtracted from the norm of the accelerometer data ( $a_s$ ).

$$a_n = \sqrt{a_{s,x}^2 + a_{s,y}^2 + a_{s,z}^2} - g \quad (5)$$

A third-order Butterworth filter with a cut-off frequency of 20 Hz was applied on the  $a_n$  signal which was divided further into windows of 1 second. With backward elimination method, 3 features were chosen to be extracted from each window  $k$ : Mean ( $x'_{k,1}$ ), standard deviation ( $x'_{k,2}$ ), and sum of absolute values ( $x'_{k,3}$ )

A naïve Bayes classifier was used to train and test the data. This method was chosen due to its better performance compared to other classification methods (Appendix II).

### C. Logical rules

To further improve the accuracy of the walking bout classification, two additional rules have been employed:

- 1) The non-locomotion periods that have less than 3 seconds duration and are between two locomotion periods, should be converted into locomotion.
- 2) The locomotion periods of less than 3 seconds should be converted into non-locomotion periods.

The biomechanical reason behind these two rules was that a walking period of less than 3 seconds is assumed to do not be a real walking bout [36]. Furthermore, a non-locomotion period of less than 3 seconds between two locomotion periods can be considered as a short negligible pause. These two rules were applied respectively after we had estimations from both GST and ML classification methods.

#### D. Cross-validation

We used again the leave-one subject-out strategy to train and test the classifiers for both of the ML and GST methods. For the ML method, for each patient, we used the remaining patients' daily activity data to train the classifier and the classifier was tested on the unseen patient. For the GST method, the threshold  $v_0$  was chosen from the training set and was tested on the unseen participant. The procedure was repeated as the same number as the participants.

The classification results were compared to the reference system and the confusion matrix was obtained. Furthermore, specificity, sensitivity, accuracy, and F1-score were reported for both of the classification methods:

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (6)$$

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (8)$$

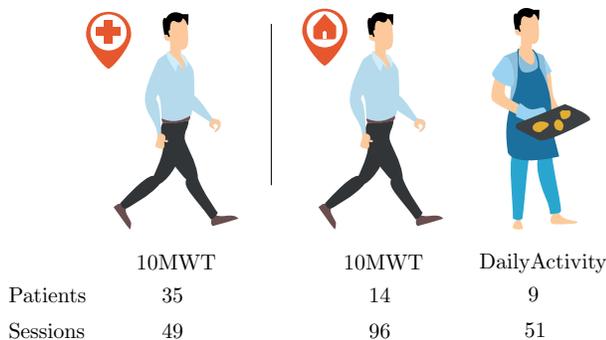
$$\text{F1-score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (9)$$

in which TP, TN, FP, and FN stand for true positive, true negative, false positive, and false negative, respectively.

### V. RESULTS

4 out of 18 patients left the study before performing the home assessments. Furthermore, out of 14 patients, only 9 of them performed the daily activity measurements.

For each of the settings, i.e. 10MWT at the clinic, 10MWT at home, and daily activities at home, patients did not perform all the sessions. Therefore, we analyzed the data of 49 sessions of 10MWT at the clinic, 96 sessions of 10MWT at home, and 51 sessions of daily activities for all the patients that remained in each setting (Fig. 2). During daily activities, over 300 hours of data were collected from the 9 patients that completed the measurements.



**Fig. 2:** The available data for analysis in home and clinical assessments for 10MWT and daily activities. For the 10MWT, there were 35 and 14 patients at the clinic and home, respectively. For the daily activities at home, data from 9 patients was available. Number of sessions shows how many measurements we had in each setting.

#### A. Gait speed estimation

The average gait speed for all the 10MWT trials at the clinic and home was shown in Fig. 3 for the 14 patients that could perform the test also at their home. It can be observed that except patients 11 and 15, all the participants had either a higher gait speed in the clinic, e.g. patients 7, 21, and 27 or a gait speed in the same range as their gait speed at home, e.g. patients 19, 25, and 33.

The vertical acceleration signal along with the corresponding estimated gait speed for the 10MWT at home and also a part of the daily activity is shown in Fig. 4 for participant number 15 (male, age=43, weight=84 kg, height=184 cm, EDSS=5.5) as an example. It can be observed that the estimated and reference values are in close agreement (Fig. 4b and Fig. 4d). Comparing Fig. 4b and Fig. 4d, gait speed has generally higher values during the 10MWT compared to daily activities. The same difference can also be observed in the intensity or the peaks of the vertical acceleration signals (Fig. 4a vs. Fig. 4c).

Stacking all the participants together, the median and IQR of the error metrics (i.e. RMS, median, IQR) were reported in Table II for three different settings: 10MWT in the clinic, 10MWT at home, and daily activity. For the daily activities, only the walking periods estimated by the reference system were considered. To have an overview of the range of the gait speeds in each setting, the IQR values of the gait speed ( $V_k$ ) were [0.21, 1.05], [0.24, 1.05], and [0.24, 0.80] m/s for 10MWT in the clinic, 10MWT at home, and daily activities at home, respectively.

**TABLE II:** The RMS, median (bias), and IQR (precision) of the error of the gait speed estimation in m/s. For each of these three parameters the median and [IQR] values were reported.

	RMS (m/s)	Median (m/s)	IQR (m/s)
10MWT at the clinic	0.10 [0.08 0.12]	-0.01 [-0.04 0.03]	0.10 [0.07 0.15]
10MWT at home	0.13 [0.11 0.14]	-0.02 [-0.06 0.04]	0.15 [0.13 0.18]
Daily activity	0.15 [0.14 0.18]	0.00 [-0.02 0.03]	0.15 [0.15 0.19]

The RMS error of the 10MWT at the clinic was 0.10 m/s and it was increased to 0.13 m/s at home. Furthermore, the RMS error of the predicted values during daily activities at home was 0.15 m/s which was the greatest of the three settings. In general, the IQR of RMS error varied between 0.08 and 0.18 m/s for all the settings and participants. The bias of the estimation error was -0.01, -0.02, and 0.00 m/s for the three settings, respectively. The precision was obtained as 0.10 m/s at the clinic, and 0.15 m/s at home.

The amount of error was illustrated by the Bland-Altman plot (Fig. 5), stacking all the participants and all the sessions together within each setting. The mean of the error was 0.00, 0.00, and 0.02 m/s for 10MWT at the clinic, 10MWT at home, and daily activities, respectively. Furthermore, the 95% confidence intervals were [-0.20, 0.21], [-0.20, 0.21], and [-0.23, 0.25] m/s, respectively.

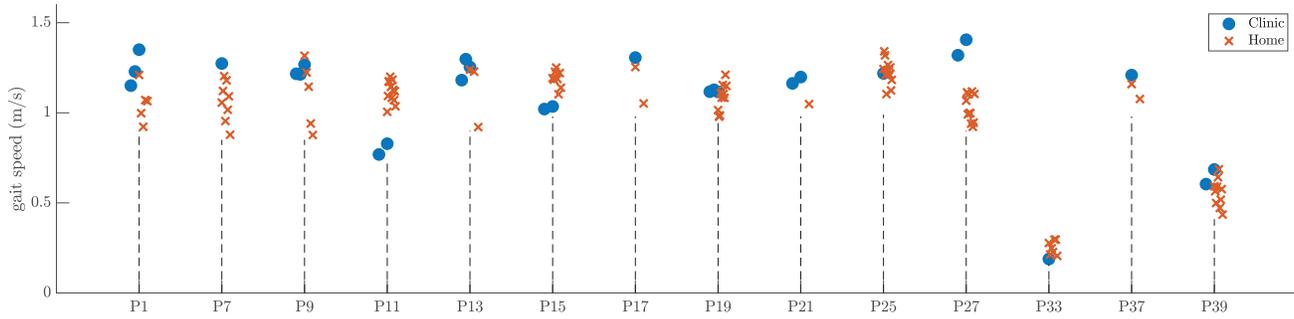


Fig. 3: Mean values of the gait speed for the 10MWT for participants that could perform the test both at the clinic and home

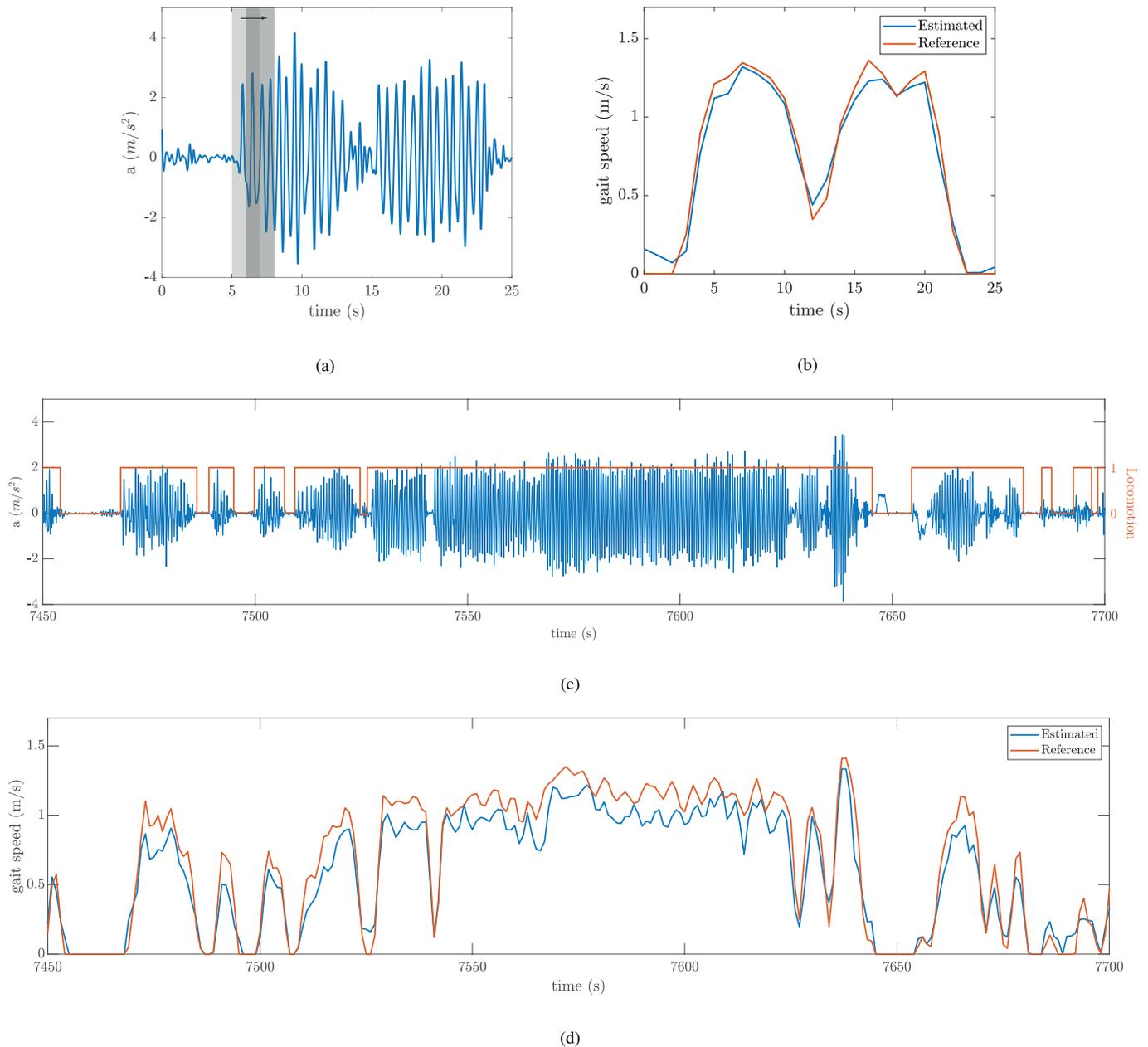


Fig. 4: Example of the signals for P15 (a) vertical acceleration during 10MWT at home along with two consecutive examples of windows used to extract features. Length of the windows is 2 seconds and they have one second overlap. (b) Estimated versus reference values of gait speed for 10MWT at home (c) Vertical acceleration signal during daily activities with the reference locomotion periods in red. (d) Estimated versus reference values of gait speed for daily activities

Pearson's correlation coefficients between the estimated values of the gait speed and the reference values were obtained as 0.96, 0.95, and 0.89 for 10MWT at the clinic, 10MWT at home, and daily activities, respectively. It can be noticed that high correlations exist between the predicted and reference values.

After dividing the patients between mild and severe group, their gait speed during the three settings was compared between the two groups (Table III). Furthermore, the RMS error was also compared. The Wilcoxon rank sum test showed that the gait speed of the mild group was significantly higher than the severe group for both of the 10MWTs. However, the severe group had a significantly higher gait speed during daily activities at home. Furthermore, the gait speed obtained during the 10MWTs showed a large effect size while the effect size for the daily activities was small.

However, according to Table III with regard to the RMS error of the gait speed estimation, Wilcoxon rank sum test showed no significant difference between the mild and severe group. To evaluate the discriminative power of the features between mild and severe groups, the Cohen's *d* value was calculated for each feature and was shown in Fig. 6. It can be seen that 7 out of 11 features had a large effect size. Features 5, and 11 ( $x_{k,5}$ , and  $x_{k,11}$ ) had the highest effect size, while features 8 and 10 ( $x_{k,8}$  and  $x_{k,10}$ ) were the lowest among all the features.

### B. Walking bout detection

For the GST method, the median (IQR) values for  $v_0$  were obtained as 0.12 ([0.11, 0.12]) m/s. The performance metrics of the detection algorithm for all the subjects were summarized in Table IV.

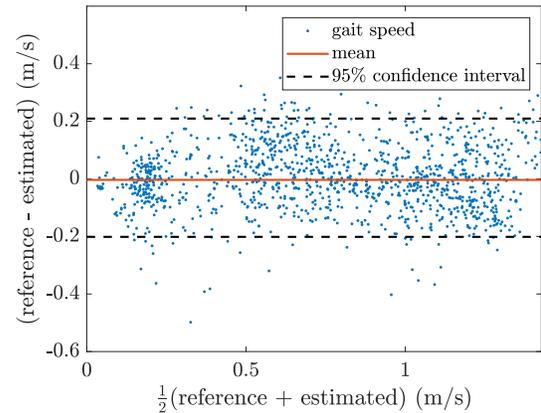
The confusion matrix where the predicted classes of activity (locomotion and non-locomotion) are shown in hours in Table V for both the reference labels and the predicted labels, demonstrate that out of almost 22.5 hours of locomotion 21.5 hours were detected correctly. However, in addition to those, there were about 25.1 hours that were falsely marked as locomotion.

For the ML method, the performance metrics as well as the confusion matrix are given in Tables IV and V, respectively. It can be observed that except the sensitivity, the performance metrics for the ML method were higher than the GST method.

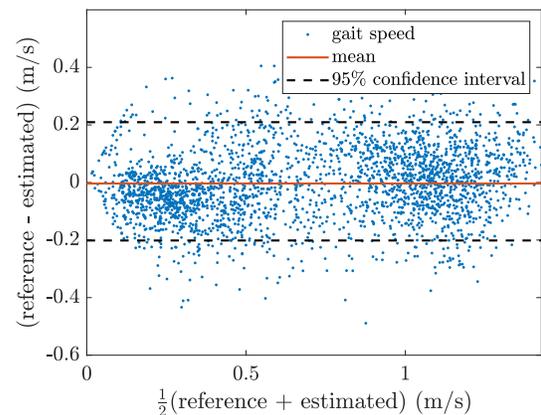
## VI. DISCUSSION

The designed algorithm based on a single IMU on the lower back was able to detect walking bouts and predict gait speed in slow and impaired gait such as in MS patients.

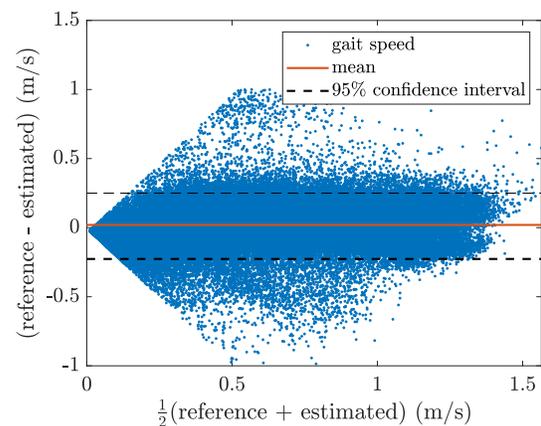
We chose the waist location for the IMU, rather than foot or wrist, due to several reasons. As movements during gait concern mostly the lower-limbs, foot or shank can be considered as the most accurate location to attach an IMU for gait assessment. For temporal parameters of gait, e.g. gait cycle time, movements of foot are characterized with features that make the extraction of these parameters more accurate and straight-forward than other locations [61]. For spatial parameters of gait, e.g. stride length, the existence of



(a)



(b)

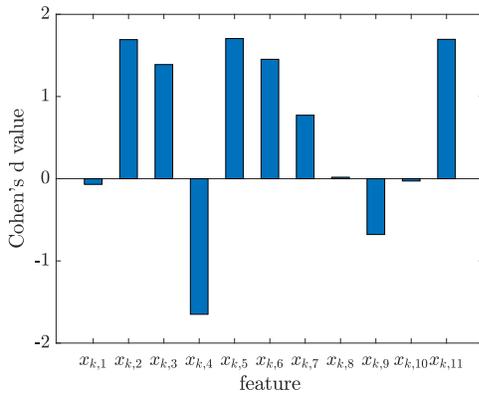


(c)

Fig. 5: Bland-Altman plot representing the estimation error of the gait speed in three settings: (a) 10MWT at the clinic, (b) 10MWT at home, (c) Daily activity

**TABLE III:** Comparison of the gait speed estimated by the lower back IMU and RMS error of gait speed estimation between patients with mild ( $EDSS \leq 4.5$ ) and severe ( $EDSS \geq 5$ ) stage of the disease, p-value of less than 0.05 was considered as significant

	Gait speed (mild) in m/s		Gait speed (severe) in m/s		Gait speed comp.	Effect size	RMSE comp.
	Median	IQR	Median	IQR	p-value	Cohen's d	p-value
10MWT at the clinic	1.20	[1.14 1.27]	0.83	[0.58 1.02]	<0.001*	1.76	0.31
10MWT at home	1.10	[1.00 1.21]	1.07	[0.50 1.17]	0.007*	1.05	0.52
Daily activity	0.50	[0.25 0.78]	0.52	[0.24 0.86]	<0.001*	-0.16	0.90



**Fig. 6:** The Cohen's d value for each of 11 features presented in III-B, the higher the absolute values, the higher the discriminative power of the features between the mild and severe group.

**TABLE IV:** The performance metrics (in %) of the locomotion detection algorithm for the GST and ML methods

	GST method		ML method	
	Median	IQR	Median	IQR
Specificity	89.7	[85.3 96.7]	96.8	[94.1 98.8]
Sensitivity	94.7	[93.0 95.5]	93.0	[89.2 95.7]
Accuracy	90.4	[86.7 96.7]	96.4	[94.2 98.6]
F1-score	60.4	[58.4 64.4]	78.6	[74.2 80.5]

**TABLE V:** Confusion matrix of the activity classification for the GST and ML methods in hours

		Predicted				Ref.
		Locomotion		Non-locomotion		
		GST	ML	GST	ML	
Ref.	Locomotion	21.5	20.5	1.0	2.0	22.5
	Non-locomotion	25.2	9.2	253.0	269.0	278.2
Predicted		46.7	29.7	254.0	271.0	300.7

zero-velocity update method allows the extraction of these parameters accurately without the drift problem [16]. Therefore, this IMU position can be considered as the most popular placement during lab assessments [62]. However, we need to attach the IMU either by straps around a bare foot or via rubber clips or straps to the shoes. We cannot expect the users to wear their shoes all the time during daily activities. Using straps around their foot might become uncomfortable after some time. Moreover, based on our direct observation by working with patients, sometimes patients are reluctant to

wear the IMU on the foot outside the laboratory environment, especially if they intend to go outdoors. Because the sensor might resemble a GPS tracker. To this end, we may sacrifice the accuracy and the number of extracted gait parameters for the sake of comfort of the users and consider wrist or waist placements during daily activities. Wrist IMU has the advantage of being embedded in a smart-watch and might have the least obtrusiveness; however, independent movements of wrist during gait can make the development of the algorithm more challenging [28], [46]. Therefore, the waist placement seems to be a compromise between comfort and accuracy. Because the IMU can be attached easily by a rubber clip to the belt and due to its closeness to the center of mass [21], [63], gait pattern can be captured more accurately than wrist location. Moreover, by using a single IMU on the waist, more information regarding the mobility of the patients can be extracted. For instance, gait asymmetry [38] or postural transitions (sit-to-stands or stand-to-sits) [52] can be evaluated which might not be possible with wrist IMU for the former and foot IMU for the latter.

In order to estimate the performance of the algorithm, a previously validated wearable system using sensors on feet was used, offering the possibility to validate the new algorithms in supervised and unsupervised setting.

Gait speed was estimated from the vertical acceleration in the global frame making the algorithm independent of the sensor placement and orientation on the lower back as in practice the sensor placements were done by the patients themselves in unsupervised settings.

Comparing vertical acceleration signal and its corresponding gait speed revealed the relationship between gait speed and the signal from which the features were derived (Fig. 4). Higher gait speeds had generally higher intensity in the vertical acceleration obtained by the belt IMU. Gait speed was decreased in the middle of the 10MWT (Fig. 4b) which shows the moment that the patient reached the marked position (10-meter mark) and had to turn and go back to their original position. One interesting observation in the vertical acceleration signal during daily activities occurred around the time 7650 s (Fig. 4c). A pattern of increase, constant, and decreasing positive signal can be seen before the time 7650 s followed by the same pattern in the negative direction after that time. The motion that can be attributed to this pattern is probably being in an elevator in which the acceleration increases to a constant value and decreases back to zero. Please also note that during the first bump (the beginning of the elevator motion), gait speed was estimated as zero despite vertical acceleration signal had

some non-zero values when the lift had an acceleration. This shows the efficiency of the frequency-domain features, i.e.  $x_{k,10}$  and  $x_{k,11}$ , as these two features take into account the periodicity of gait. During the second bump (at the end of the elevator motion), the patient had a few steps with low gait speeds inside the elevator.

Using silver standard reference systems, the accuracy of the gait speed estimation algorithm was assessed by several means. The RMS, bias, and precision of the error shown in Table II, indicated that our algorithm could achieve with only a slight decrease in the performance when being validated in home assessments. The RMS error of the estimation was 0.10 m/s in the clinic, and 0.13 and 0.15 m/s at home in 10MWT and daily activities, respectively. Compared to other algorithms validated in supervised conditions [27], where the RMS error of the gait speed estimation during treadmill walking in MS patients was 0.12 m/s, our algorithm demonstrated slightly better accuracy. But compared to the reference [26], we achieved a lower accuracy as the RMS error of their gait speed estimation was 0.07 m/s. In this study, a model was trained based on a large group of older adults. Gait speed was obtained by an IMU on the lower back in straight walking tests in the clinic. The authors showed that adding demographic data and anthropometry can reduce the estimation error of gait speed. In [28], an RMS error of 0.05 m/s was achieved based on a personalized algorithm in healthy subjects and in conditions where only walking rather than other activities was included in the protocol. Compared to another study based on an accelerometer on the lower back in which the gait speed in MS patients was overestimated by a bias of 0.12 m/s [64], we achieved a bias of almost zero in all the three settings (Table II).

This was also shown in Bland-Altman plot (Fig. 5). However, during daily activities, the estimation error was higher for gait speeds lower than 1 m/s (Fig. 5c). In our opinion, there are two potential explanations for the large errors around the gait speed of 0.5 m/s. One reason can be a lower accuracy of our algorithm during slower walking speeds. While this reason is unlikely as we do not see these large errors around lower gait speeds during clinical assessments, it can be avoided by employing two regression models, one for lower gait speeds, and one for higher gait speeds. For instance, in another study, a low speed-specific model was used for gait speeds of less than 1 m/s to have a more accurate gait speed estimation [26]. The second reason which is more related to the complex context of the domestic environment, can be due to the slower gait speeds during very short walking bouts, e.g. less than 10 seconds, or activities that involve other tasks rather than gait. During daily activity monitoring, it has been shown in the literature that short walking bouts are generally accompanied by other tasks while longer walking bouts are generally single-task and occur more outdoors [65]. Furthermore, shorter walking bouts can have slower gait speeds than longer walking bouts [66]. Therefore, we believe that those large errors around 0.5 m/s can be more related to very short walking bouts that included shuffling or even were wrongly detected as gait. These tasks might potentially affect the accuracy of our silver standard reference as well as the predicted values by the IMU on

the belt. To avoid such errors, a constraint on the minimum duration of a walking bout, e.g. 10 or 15 seconds can be set to have a steady-state gait. The funnel shape in Fig. 5c is because of the instances in which either the reference or the estimated value was close to zero but their difference was slightly higher than zero. Thus, two lines with slopes of  $\pm 2$  appeared in this plot. The percentage of the reference and estimated values that were close to zero were 1.5% and 0.2% in Fig. 5a, 1.1% and 1.4% in Fig. 5b, and 5.0% and 1.2% in Fig. 5c which are in the same range as our reference system [36].

The correlation analysis between the estimated values and the reference values showed that although there was an excellent agreement between those two, this association decreased in home assessments. This can be due to the different context of the environment at home compared to the clinic, e.g. many obstacles, slopes, turnings in real-life settings.

Several comparisons were performed between the patients with mild ( $EDSS \leq 4.5$ ) and severe ( $EDSS \geq 5$ ) stages of the disease. Firstly, we noticed that the mild group had an estimated gait speed higher than the severe group during the 10MWTs (Table III). This states that our method had the potential of predicting the progression of the disease in an objective manner. Decreasing gait speed with the EDSS has also been observed in [14], [27], [31]. The patients in the severe stage of the disease had a higher mean of gait speed during daily activities. This can be because of the small sample size of this setting. Furthermore, large effect size values were observed for gait speed obtained during the 10MWTs performed in clinic and at home, but not for the daily activities. This might be again due to the small number of patients (9) for the daily activity measurements. Analyzing the discriminative power of each feature between mild and severe group revealed that features  $x_{k,5}$  (sum of absolute values) and  $x_{k,11}$  (amplitude of the dominant frequency) had the highest ability in differentiating between the two groups with  $x_{k,5}$  being the highest (Fig. 6). As defined in the Appendix, this feature was the sum of absolute values which can be somehow considered as the absolute value of the vertical velocity of the center of mass. Therefore, the gait speed extracted by the IMU along with the features have the potential to help the clinicians monitor the progress of the disease.

Some of the features in Fig. 6, e.g.  $x_{k,4}$  and  $x_{k,9}$ , had negative Cohen's d values. According to Eq. 4, a negative Cohen's d value means a higher mean of a feature from the severe group. Referring to Appendix I, these two features ( $x_{k,4}$  and  $x_{k,9}$ ) belong to the min value of the acceleration and its integration signals. As it can be seen in Fig. 4, the local minimums of the acceleration signal are negative. Therefore, the severe group of the patients have higher  $x_{k,4}$  value and consequently, a lower absolute value than the mild group which is expected as the intensity of the signal is lower for the severe group.

Another interesting observation in Table III was that for the 10MWT, the severe group had on average higher gait speed in the clinic compared to home. This is in contrast to the moderate group and also to a recent study on patients with Parkinson's disease (PD) that participants had generally higher gait speeds in the 10MWT performed in the clinic [67]. While

going into the details of this hot topic of “clinical vs. home assessment” is out of the scope of the current study, there are some explanations for this seemingly contradictory behavior. The main reason can be due to the fact that individuals behave differently in different settings due to several factors [15]. Other than sensorimotor system, psychological factors can impact our mobility. For instance, the white-coat effect in which patients perform worse in the clinic is one of these psychological factors [15]. Sometimes, the presence of an observer might induce stress in the patients causing them to not present their actual capacity. While in the aforementioned study on PD patients [67], the general trend was a higher gait speed in the clinic, there were a few patients having faster gait speeds at home. It is worth mentioning that in our study, two of the patients, i.e. P11 (EDSS=6.5) and P15 (EDSS=5.5) specifically, were among the severe group that their higher gait speed at home was evident from Fig. 3. Another probable explanation for this difference can be that the severe group had benefited more from their rehabilitation in the clinic compared to the moderate group. Because home assessments were performed after the clinical assessments in the rehabilitation center. However, more evidence with a larger dataset is required to confirm this reasoning. Whatever the reasons are for this observation, our method could detect these subtle changes in an objective manner. Furthermore, an unsupervised assessment at home can question the measurements performed in the clinic and help the clinicians have a better insight into the actual capacity of the patients.

For the locomotion detection during daily activities, we introduced two new methods. Both methods showed a specificity, sensitivity, and accuracy of higher than 90% with the ML method being superior in F1-score (78.6% versus 60.4% in GST method). Detecting walking bouts during daily activities can help the clinicians have an objective and more accurate estimation of the activity status of the patients at home. Furthermore, by having the duration of each walking bout, gait speed can be estimated in walking bouts with a specific duration such as the duration that it takes for the patients to perform the 10MWT. When optimizing by the ROC curve, the GST method had higher sensitivity and lower specificity compared to the ML method. One explanation can be that some movements of the lower back can induce a speed for the center of mass; however, these movements might not be actually a walking bout. Therefore, in GST method the predicted walking bouts were almost double the ML method. On the other hand, the GST method depends on the threshold that is being used. For instance, to have a specificity as the same as the ML method, the sensitivity of the GST method drops to 83%.

There are very few previous studies on walking bout detection that validated their algorithms in real-life daily activities. We have achieved comparable sensitivity (93.0% by ML method versus 87.1% in [46] and 94.0% in [42]) and specificity (96.8% by ML method versus 96.7% in [46] and 97.0% in [42]) compared to two studies one on 37 older adults [46] and the other on 15 children with cerebral palsy [42]. Regarding the F1-score, we achieved slightly higher performance (78.6% versus 74.9%) compared to the study conducted in

reference [46]. However compared to another study with 20 older adults [47], their proposed method achieved a higher F1-score of 87.1%. One explanation can be using a gold standard system, i.e. camera in their study as the reference system rather than a silver standard reference system as was the case in our study. Having a more accurate reference system can lead to a more accurate classifier. Another explanation for their superior results can be their choice of features as they have used numerous time-domain features from both the accelerometer and gyroscope signal from all the three axes. Although, using features from all the three axes of the sensors may provide additional information about the biomechanics of the movement, it makes the performance of the algorithm dependent to sensor placement changes. Using the norm of the accelerometer or the vertical acceleration in the global frame (as was the case in our study or in [46]) can prevent such dependencies.

The novelty of this work was mainly the method used to train the models to estimate gait speed and detect walking bouts by a single IMU on the lower back. We have used some of the features stated in the literature while we introduced a novel method to map these features into gait speed and state of activity. This method which is the training by a silver standard reference system, i.e. IMUs on the feet, allowed the estimation of instantaneous gait speed as well as state of activity (locomotion or non-locomotion) by a single IMU on the belt. This new method can have two potential applications: the first one which was shown in this paper, is training of the estimation model in the lab with a multi-sensor system. By having this model, the clinicians can ask the patients to wear only the lower back IMU which is more comfortable and easier to use (due to the reasons mentioned throughout the paper) during assessments at home or outside clinic. The other potential application which can be a future study, is a personalized approach, meaning that patients can be equipped initially with 3 IMUs on the feet and the lower back, and be asked to perform their daily activities while the sensors are recording the data. Once enough walking bouts were measured and a model with a clinically meaningful accuracy was trained, patients can detach the feet sensors and continue their daily activities with only the IMU on the belt. To better clarify the contribution of this study compared to the previous studies with similar machine learning-based algorithms or approaches, we see two main distinctions:

- Our proposed method and algorithms are capable of estimating the instantaneous gait speed rather than a single value gait speed over a trial of walking test as was the case in [25]–[27], [35]. This way of gait speed estimation allows the estimation of gait variability during walking tests as well as gait speed distribution during daily activities. Opposed to the studies in [25]–[27], [35], we validated our algorithm also during daily activities and walking tests performed in the patients’ home. Furthermore, our gait speed estimation approach was based on the vertical acceleration signal in the global frame making the algorithm robust to sensor placement changes [52].

- Compared to the recent study introduced by Soltani *et al.* [28] which trained their single-IMU-based gait speed estimation model by a Global Navigation Satellite System (GNSS) system, we used IMUs on the feet to train the model. Using GNSS signal might face problems related to the strength of the signals received from the satellites. Specially, when the patients are indoors, GNSS can have a weak signal acquisition. Furthermore, GNSS might raise privacy concerns due to tracking the location of the participants. In this regard, an IMU is preferable than the GNSS system.

The first limitation of our study was the small sample size of participants in the daily activity setting. Nevertheless, the dataset was large enough to perform the validation, as we had over 300 hours of daily activity measurements in which 22 hours were locomotion periods. Due to limitations in the materials and equipment as well as the long duration of the home assessments (12 weeks), half of the patients (18) were given the equipment for the home assessments. From these 18 participants, 4 of them quit the study due to personal and technical issues: one of them got a new job and was not willing to continue the study. Another participant was at a 4-week holiday and decided to stop participating in the study. One of the patients reported a long waiting time for the sensors to be connected to the smartphone by Bluetooth. Therefore, they sent back the system to be repaired. The fourth patient completely forgot to perform the home assessments. Even though text messages had to be sent as reminders to the patients that did not perform the tests on time, sending reminders to this specific patient was overlooked. Furthermore, 5 out of 14 patients did not perform the daily activity part. The fact that this part was not mandatory and patients were asked to perform it in case of possibility can explain the reason.

Among the remaining patients that performed the measurements, some of the sessions were missed or excluded. Around 25% of these sessions belonged to the data that was recorded by the sensors but the protocol of the test was not respected, and 75% of these sessions belonged to missing data. While the exact reason for the missing data is not known to us individually for each session, we assume it can be due to the difficulty of the participants to work with a smartphone other than their own phone, slow connection of Bluetooth, and disruption in data connection to transfer data.

While using a single sensor setup (as was the main goal of this study) rather than a complex multi-sensor system can overcome some of the issues mentioned above, designing a more autonomous system can be helpful. For instance, the IMUs used in this study were being charged through USB cable. A smart charging dock station can be designed instead, to transfer the data from the IMUs memory card to the server automatically while the IMUs are being charged. Moreover, using a faster mean of communication, e.g. WiFi, rather than Bluetooth could improve the connection between IMUs and smartphone. Another problem was that some of the patients did not respect the protocol of the test during the unsupervised assessment. An automated algorithm (e.g. checking the traversed distance or the azimuth angle of the trunk) embedded on the sensors can notify the user im-

mediately if the walking test was not performed correctly. In this case, the user can repeat the test and data loss would be reduced. Another solution can be implementing the 2-minute [68] or 6-minute [69] walk tests rather than 10MWT, as these tests are easier to perform in practice and have a less strict protocol. However, one should be careful about comparing the results of these tests between the patients. Because due to the protocol of the 2-minute or 6-minute walk tests, the effect of a complex environment at home such as obstacles cannot be controlled. This is why we had chosen the 10MWT as this test has a more strict protocol and takes less time than the 2-minute or 6-minute walk tests. Finally, recently, there are more and more studies showing the relationship between a walking test performed in the clinic and measurements performed during daily activities [21], [70]. By knowing these relations, clinicians can estimate patients' capacity from daily living assessments without actually performing the clinical gait tests.

While studying the usability of the system was out of scope of the current study, an impartial study can be helpful to investigate the usability and acceptability of such a system among the patients. Because in the end, if we want these wearables to be integrated into clinical assessments of the patients, their usability will have an important role [71].

One aspect of limitation was that our dataset could have been divided into three sets of training, development, and test. To have a more reliable machine learning model, a separate development set is needed to tune the hyperparameters. However, due to the small number of participant, we only had two training and test sets.

Another aspect of limitation can be using silver standard methods to validate the algorithms, therefore making the accuracy highly dependent on such methods. To have a more reliable references, walkways during clinical assessment and annotated videos during daily activities can be used to validate the gait speed estimation and walking bout detection methods, respectively. Finally, we observed relatively large errors around slower gait speeds during daily activities. While we hypothesize these errors can be due to the gait speeds during short walking bouts, more evidence with a larger dataset is required to confirm this assumption and investigate the effect of bout length on the performance of our methods.

Future work will be focused on the comparison of the gait speed between the clinical and home settings. This would help us have a deeper understanding of the degree of overlap between supervised and unsupervised assessments.

## VII. CONCLUSION

This study introduced Gaussian process regression method to estimate the gait speed in clinical as well as home environments using a single IMU at waist. Based on vertical acceleration in global frame, the method was independent on sensor placement on waist. The robustness of the method was guaranteed by dual validation in both supervised gait test and unsupervised daily activities at home. We built a model based on a multi-sensor setup in the clinic, and the model was tested with a simpler sensor setup, i.e. a single lower back

IMU at home. Moreover, two methods to detect the walking bouts were presented in which the machine learning-based approach seemed to perform better. This validation study can provide clinicians with an objective tool to assess the mobility performance of the patients in both clinic and home in an unsupervised manner. Now that an unsupervised assessment of mobility in MS patients seemed feasible, future research is now possible to compare the patients' mobility between supervised and unsupervised settings.

### APPENDIX I LIST OF FEATURES

List of the complete features before selection is given below with their definition as well as the references to the literature from which the features were selected. If a feature was selected based on our own hypothesis, we have mentioned the reason behind choosing that feature. 13 of these features belonged to the low-pass filtered vertical acceleration signal ( $a_f$ ). Denoting  $a_f$  at each sample  $i$  by  $a_{f,i}$ , and assuming within each window there are  $N$  samples, the features for each window  $k$  at time  $t_k$  are:

- Mean [27], [35]

$$x_{k,1} = \frac{1}{N} \sum_{i=1}^N a_{f,i}$$

- Standard deviation [35]

$$x_{k,2} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N |a_{f,i} - x_{k,1}|^2}$$

- Maximum, as a representative of the intensity of the movement

$$x_{k,3} = \max \{a_{f,i}\}$$

- Minimum, as a representative of the intensity of the movement

$$x_{k,4} = \min \{a_{f,i}\}$$

- Sum of absolute values [35]

$$x_{k,5} = \sum_{i=1}^N |a_{f,i}|$$

- Sum of squared values [35]

$$x_{k,6} = \sum_{i=1}^N a_{f,i}^2$$

- Maximum of the vertical velocity, as a representative of vertical oscillation due to inverted pendulum behaviour of gait

$$x_{k,7} = \max \left\{ \int a_f(i) di \right\}$$

- Mean of the vertical velocity, as a representative of vertical oscillation due to inverted pendulum behaviour of gait

$$x_{k,8} = \text{mean} \left\{ \int a_f(i) di \right\}$$

- Minimum of the vertical velocity, as a representative of vertical oscillation due to inverted pendulum behaviour of gait

$$x_{k,9} = \min \left\{ \int a_f(i) di \right\}$$

- First dominant frequency, as a representative of cadence

$$x_{k,10} = f_{\max,1}$$

- Amplitude of the first dominant frequency, as a representative of cadence

$$x_{k,11} = A_{f_{\max,1}}$$

- Second dominant frequency, as a representative of cadence

$$x_{k,12} = f_{\max,2}$$

- Amplitude of the second dominant frequency, as a representative of cadence

$$x_{k,13} = A_{f_{\max,2}}$$

Dominant frequency and its amplitude were obtained by computing the fast Fourier transform (FFT) of the signal throughout the window.

The remaining 5 features were from the azimuth angular velocity ( $\omega$ ) and azimuth angle ( $\theta$ ). Azimuth angular velocity and rotation were used as they are indicative of turns and people might have different behaviors during turnings [37].

- Mean of the azimuth angular velocity

$$x_{k,14} = \frac{1}{N} \sum_{i=1}^N \omega_i$$

- Standard deviation of the azimuth angular velocity

$$x_{k,15} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N |\omega_i - x_{k,14}|^2}$$

- Maximum of the azimuth angular velocity

$$x_{k,16} = \max \{\omega_i\}$$

- Minimum of the azimuth angular velocity

$$x_{k,17} = \min \{\omega_i\}$$

- Range of the azimuth angle

$$x_{k,18} = \max \{\theta_i\} - \min \{\theta_i\}$$

### APPENDIX II APPLIED CLASSIFICATION METHODS

We have applied several classification methods and we chose the naïve Bayes classifier based on its accuracy and computation time.

#### ACKNOWLEDGMENT

The authors would like to thank Sakura Nussbaum and Gaëlle Prigent for their support in data processing and providing the hardware and equipment. We would also like to thank European Union's Horizon2020 research and innovation programme for providing the funding under the Marie Skłodowska-Curie grant agreement No. 721577.

**TABLE VI:** The employed machine learning methods along with their accuracy and processing time for ML method to detect the walking bouts

Method	Accuracy(%)	Training and test time (s)
Decision tree	95.0	48
LDA	94.6	12
Logistic regression	94.2	35
Naïve Bayes	94.7	16
kNN ( $k = 10$ )	95.0	219

## REFERENCES

- [1] S. Fritz and M. Lusardi, "White paper: "walking speed: the sixth vital sign";," *Journal of geriatric physical therapy*, vol. 32, no. 2, pp. 2–5, 2009.
- [2] J. S. Brach, J. M. VanSwearingen, A. B. Newman, and A. M. Kriska, "Identifying early decline of physical function in community-dwelling older women: performance-based and self-report measures," *Physical Therapy*, vol. 82, no. 4, pp. 320–328, 2002.
- [3] J. L. Purser, M. Weinberger, H. J. Cohen, C. F. Pieper, M. C. Morey, T. Li, G. R. Williams, and P. Lapuerta, "Walking speed predicts health status and hospital costs for frail elderly male veterans," *Journal of rehabilitation research & development*, vol. 42, no. 4, 2005.
- [4] J. A. Cohen, A. V. Krishnan, A. D. Goodman, J. Potts, P. Wang, E. Havrdova, C. Polman, and R. A. Rudick, "The clinical meaning of walking speed as measured by the timed 25-foot walk in patients with multiple sclerosis," *JAMA neurology*, vol. 71, no. 11, pp. 1386–1393, 2014.
- [5] N. G. LaRocca, "Impact of walking impairment in multiple sclerosis," *The Patient: Patient-Centered Outcomes Research*, vol. 4, no. 3, pp. 189–201, 2011.
- [6] J. Kesselring, "Disease progression in multiple sclerosis. i. impaired mobility and its impact on limitations of activities and social participation," *Eur Neurol Rev*, vol. 5, pp. 56–60, 2010.
- [7] J. F. Kurtzke, "Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (edss)," *Neurology*, vol. 33, no. 11, pp. 1444–1444, 1983.
- [8] J. Freeman, D. Langdon, J. Hobart, and A. Thompson, "The impact of inpatient rehabilitation on progressive multiple sclerosis," *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, vol. 42, no. 2, pp. 236–244, 1997.
- [9] H. Inojosa, D. Schriefer, A. Klöditz, K. Trentzsch, and T. Ziemssen, "Balance testing in multiple sclerosis—improving neurological assessment with static posturography?" *Frontiers in Neurology*, vol. 11, 2020.
- [10] A. Vienne-Jumeau, F. Quijoux, P.-P. Vidal, and D. Ricard, "Wearable inertial sensors provide reliable biomarkers of disease severity in multiple sclerosis: A systematic review and meta-analysis," *Annals of Physical and Rehabilitation Medicine*, 2019.
- [11] F. Bethoux and S. Bennett, "Introduction: enhancing mobility in multiple sclerosis," *International journal of MS care*, vol. 13, no. 1, pp. 1–3, 2011.
- [12] R. W. Motl, M. D. Goldman, and R. H. Benedict, "Walking impairment in patients with multiple sclerosis: exercise training as a treatment option," *Neuropsychiatric disease and treatment*, vol. 6, p. 767, 2010.
- [13] C. L. Martin, B. A. Phillips, T. Kilpatrick, H. Butzkueven, N. Tubridy, E. McDonald, and M. Galea, "Gait and balance impairment in early multiple sclerosis in the absence of clinical disability," *Multiple Sclerosis Journal*, vol. 12, no. 5, pp. 620–628, 2006.
- [14] J. L. Preiningerova, K. Novotna, J. Ruz, L. Sucha, E. Ruzicka, and E. Havrdova, "Spatial and temporal characteristics of gait as outcome measures in multiple sclerosis (edss 0 to 6.5)," *Journal of neuroengineering and rehabilitation*, vol. 12, no. 1, p. 14, 2015.
- [15] E. Warmerdam, J. M. Hausdorff, A. Atrsaei, Y. Zhou, A. Mirelman, K. Aminian, A. J. Espay, C. Hansen, L. J. Evers, A. Keller *et al.*, "Long-term unsupervised mobility assessment in movement disorders," *The Lancet Neurology*, 2020.
- [16] B. Mariani, C. Hoskovec, S. Rochat, C. Büla, J. Penders, and K. Aminian, "3d gait assessment in young and elderly subjects using foot-worn inertial sensors," *Journal of biomechanics*, vol. 43, no. 15, pp. 2999–3006, 2010.
- [17] A. M. Sabatini, C. Martelloni, S. Scapellato, and F. Cavallo, "Assessment of walking features from foot inertial sensing," *IEEE Transactions on biomedical engineering*, vol. 52, no. 3, pp. 486–494, 2005.
- [18] K. Aminian, B. Najafi, C. Büla, P.-F. Leyvraz, and P. Robert, "Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes," *Journal of biomechanics*, vol. 35, no. 5, pp. 689–699, 2002.
- [19] A. Rampp, J. Barth, S. Schülein, K.-G. Gaßmann, J. Klucken, and B. M. Eskofier, "Inertial sensor-based stride parameter calculation from gait sequences in geriatric patients," *IEEE transactions on biomedical engineering*, vol. 62, no. 4, pp. 1089–1097, 2014.
- [20] Y. Moon, R. S. McGinnis, K. Seagers, R. W. Motl, N. Sheth, J. A. Wright Jr, R. Ghaffari, and J. J. Sosnoff, "Monitoring gait in multiple sclerosis with novel wearable motion sensors," *PLoS one*, vol. 12, no. 2, p. e0171346, 2017.
- [21] F. A. Storm, K. Nair, A. J. Clarke, J. M. Van der Meulen, and C. Mazza, "Free-living and laboratory gait characteristics in patients with multiple sclerosis," *PLoS one*, vol. 13, no. 5, 2018.
- [22] B. Fasel, C. Duc, F. Dadashi, F. Bardyn, M. Savary, P.-A. Farine, and K. Aminian, "A wrist sensor and algorithm to determine instantaneous walking cadence and speed in daily life walking," *Medical & biological engineering & computing*, vol. 55, no. 10, pp. 1773–1785, 2017.
- [23] W. Zijlstra and A. L. Hof, "Assessment of spatio-temporal gait parameters from trunk accelerations during human walking," *Gait & posture*, vol. 18, no. 2, pp. 1–10, 2003.
- [24] J.-S. Hu, K.-C. Sun, and C.-Y. Cheng, "A kinematic human-walking model for the normal-gait-speed estimation using tri-axial acceleration signals at waist location," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 8, pp. 2271–2279, 2013.
- [25] S. Zihajehzadeh and E. J. Park, "Experimental evaluation of regression model-based walking speed estimation using lower body-mounted imu," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 243–246.
- [26] S. Byun, H. J. Lee, J. W. Han, J. S. Kim, E. Choi, and K. W. Kim, "Walking-speed estimation using a single inertial measurement unit for the older adults," *PLoS one*, vol. 14, no. 12, 2019.
- [27] R. S. McGinnis, N. Mahadevan, Y. Moon, K. Seagers, N. Sheth, J. A. Wright Jr, S. DiCristofaro, I. Silva, E. Jortberg, M. Ceruolo *et al.*, "A machine learning approach for gait speed estimation using skin-mounted wearable sensors: From healthy controls to individuals with multiple sclerosis," *PLoS one*, vol. 12, no. 6, 2017.
- [28] A. Soltani, H. Dejnabadi, M. Savary, and K. Aminian, "Real-world gait speed estimation using wrist sensor: A personalized approach," *IEEE journal of biomedical and health informatics*, 2019.
- [29] H. Vathsangam, A. Emken, D. Spruijt-Metz, and G. S. Sukhatme, "Toward free-living walking speed estimation using gaussian process-based regression with on-body accelerometers and gyroscopes," in *2010 4th International Conference on Pervasive Computing Technologies for Healthcare*. IEEE, 2010, pp. 1–8.
- [30] L. Shammas, T. Zentek, B. von Haaren, S. Schlesinger, S. Hey, and A. Rashid, "Home-based system for physical activity monitoring in patients with multiple sclerosis (pilot study)," *Biomedical engineering online*, vol. 13, no. 1, p. 10, 2014.
- [31] A. Supratak, G. Datta, A. R. Gafson, R. Nicholas, Y. Guo, and P. M. Matthews, "Remote monitoring in the home validates clinical gait measures for multiple sclerosis," *Frontiers in neurology*, vol. 9, p. 561, 2018.
- [32] A. M. Keppler, T. Nuritidinow, A. Mueller, H. Hoefling, M. Schieker, I. Clay, W. Böcker, and J. Fürmetz, "Validity of accelerometry in step detection and gait speed measurement in orthogeriatric patients," *PLoS one*, vol. 14, no. 8, 2019.
- [33] M. Schimpl, C. Lederer, and M. Daumer, "Development and validation of a new method to measure walking speed in free-living environments using the actibelt@ platform," *PLoS one*, vol. 6, no. 8, 2011.
- [34] D. A. Neumann, *Kinesiology of the musculoskeletal system-e-book: foundations for rehabilitation*. Elsevier Health Sciences, 2013.
- [35] S. Zihajehzadeh and E. J. Park, "Regression model-based walking speed estimation using wrist-worn inertial sensor," *PLoS one*, vol. 11, no. 10, p. e0165211, 2016.
- [36] C. M. el Achkar, C. Lenoble-Hoskovec, A. Paraschiv-Ionescu, K. Major, C. Büla, and K. Aminian, "Instrumented shoes for activity classification in the elderly," *Gait & posture*, vol. 44, pp. 12–17, 2016.
- [37] M. H. Pham, M. Elshehabi, L. Haertner, S. Del Din, K. Srulijes, T. Heger, M. Synofzik, M. A. Hobert, G. S. Faber, C. Hansen *et al.*, "Validation of a step detection algorithm during straight walking and turning in patients with parkinson's disease and older adults using an

- inertial measurement unit at the lower back,” *Frontiers in neurology*, vol. 8, p. 457, 2017.
- [38] S. Del Din, A. Godfrey, and L. Rochester, “Validation of an accelerometer to quantify a comprehensive battery of gait characteristics in healthy older adults and parkinson’s disease: toward clinical and at home use,” *IEEE journal of biomedical and health informatics*, vol. 20, no. 3, pp. 838–847, 2015.
- [39] F. Massé, R. R. Gonzenbach, A. Arami, A. Paraschiv-Ionescu, A. R. Luft, and K. Aminian, “Improving activity recognition using a wearable barometric pressure sensor in mobility-impaired stroke patients,” *Journal of neuroengineering and rehabilitation*, vol. 12, no. 1, p. 72, 2015.
- [40] A. Godfrey, A. Bourke, G. O’laighin, P. Van De Ven, and J. Nelson, “Activity classification using a single chest mounted tri-axial accelerometer,” *Medical engineering & physics*, vol. 33, no. 9, pp. 1127–1135, 2011.
- [41] B. Dijkstra, W. Zijlstra, E. Scherder, and Y. Kamsma, “Detection of walking periods and number of steps in older adults and patients with parkinson’s disease: accuracy of a pedometer and an accelerometry-based method,” *Age and ageing*, vol. 37, no. 4, pp. 436–441, 2008.
- [42] A. Paraschiv-Ionescu, C. J. Newman, L. Carreff, C. N. Gerber, S. Armand, and K. Aminian, “Locomotion and cadence detection using a single trunk-fixed accelerometer: validity for children with cerebral palsy in daily life-like conditions,” *Journal of neuroengineering and rehabilitation*, vol. 16, no. 1, p. 24, 2019.
- [43] A. Hickey, S. Del Din, L. Rochester, and A. Godfrey, “Detecting free-living steps and walking bouts: validating an algorithm for macro gait analysis,” *Physiological measurement*, vol. 38, no. 1, p. N1, 2016.
- [44] D. Trojaniello, A. Cereatti, E. Pelosin, L. Avanzino, A. Mirelman, J. M. Hausdorff, and U. Della Croce, “Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors: application to elderly, hemiparetic, parkinsonian and choreic gait,” *Journal of neuroengineering and rehabilitation*, vol. 11, no. 1, p. 152, 2014.
- [45] I. C. Gyllensten and A. G. Bonomi, “Identifying types of physical activity with a single accelerometer: evaluating laboratory-trained algorithms in daily life,” *IEEE transactions on biomedical engineering*, vol. 58, no. 9, pp. 2656–2663, 2011.
- [46] A. Soltani, A. Paraschiv-Ionescu, H. Dejnabadi, P. Marques-Vidal, and K. Aminian, “Real-world gait bout detection using a wrist sensor: An unsupervised real-life validation,” *IEEE Access*, vol. 8, pp. 102 883–102 896, 2020.
- [47] M. Awais, L. Chiari, E. A. F. Ihlen, J. L. Helbostad, and L. Palmerini, “Physical activity classification for elderly people in free-living conditions,” *IEEE journal of biomedical and health informatics*, vol. 23, no. 1, pp. 197–207, 2018.
- [48] B. M. Sandroff, R. W. Motl, L. A. Pilutti, Y. C. Learmonth, I. Ensari, D. Dlugonski, R. E. Klaren, S. Balantrapu, and B. J. Riskin, “Accuracy of stepwatch™ and actigraph accelerometers for measuring steps taken among persons with multiple sclerosis,” *PLoS one*, vol. 9, no. 4, 2014.
- [49] B. Dijkstra, Y. P. Kamsma, and W. Zijlstra, “Detection of gait and postures using a miniaturized triaxial accelerometer-based system: accuracy in patients with mild to moderate parkinson’s disease,” *Archives of physical medicine and rehabilitation*, vol. 91, no. 8, pp. 1272–1277, 2010.
- [50] A. J. Thompson, B. L. Banwell, F. Barkhof, W. M. Carroll, T. Coetzee, G. Comi, J. Correale, F. Fazekas, M. Filippi, M. S. Freedman *et al.*, “Diagnosis of multiple sclerosis: 2017 revisions of the mcdonald criteria,” *The Lancet Neurology*, vol. 17, no. 2, pp. 162–173, 2018.
- [51] B. Mariani, M. C. Jiménez, F. J. Vingerhoets, and K. Aminian, “On-shoe wearable sensors for gait and turning assessment of patients with parkinson’s disease,” *IEEE transactions on biomedical engineering*, vol. 60, no. 1, pp. 155–158, 2012.
- [52] A. Atrsaei, F. Dadashi, C. Hansen, E. Warmerdam, B. Mariani, W. Maetzler, and K. Aminian, “Postural transitions detection and characterization in healthy and patient populations using a single waist sensor,” *Journal of NeuroEngineering and Rehabilitation*, vol. 17, no. 1, pp. 1–14, 2020.
- [53] S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, “Estimation of imu and marg orientation using a gradient descent algorithm,” in *2011 IEEE international conference on rehabilitation robotics*. IEEE, 2011, pp. 1–7.
- [54] M. El-Gohary, S. Pearson, J. McNames, M. Mancini, F. Horak, S. Mellone, and L. Chiari, “Continuous monitoring of turning in patients with movement disability,” *Sensors*, vol. 14, no. 1, pp. 356–369, 2014.
- [55] F. Dadashi, G. P. Millet, and K. Aminian, “Estimation of front-crawl energy expenditure using wearable inertial measurement units,” *IEEE Sensors Journal*, vol. 14, no. 4, pp. 1020–1027, 2013.
- [56] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.
- [57] F. Dadashi, G. Millet, and K. Aminian, “Gaussian process framework for pervasive estimation of swimming velocity with body-worn imu,” *Electronics letters*, vol. 49, no. 1, pp. 44–45, 2013.
- [58] J. M. Bland and D. G. Altman, “Applying the right statistics: analyses of measurement studies,” *Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology*, vol. 22, no. 1, pp. 85–93, 2003.
- [59] L. Angelini, I. Carpinella, D. Cattaneo, M. Ferrarin, E. Gervasoni, B. Sharrack, D. Paling, K. P. S. Nair, and C. Mazzà, “Is a wearable sensor-based characterisation of gait robust enough to overcome differences between measurement protocols? a multi-centric pragmatic study in patients with multiple sclerosis,” *Sensors*, vol. 20, no. 1, p. 79, 2020.
- [60] D. K. Lee, “Alternatives to p value: confidence interval and effect size,” *Korean journal of anesthesiology*, vol. 69, no. 6, p. 555, 2016.
- [61] B. Mariani, H. Rouhani, X. Crevoisier, and K. Aminian, “Quantitative estimation of foot-flat and stance phase of gait using foot-worn inertial sensors,” *Gait & posture*, vol. 37, no. 2, pp. 229–234, 2013.
- [62] M. Zrenner, A. Küderle, N. Roth, U. Jensen, B. Dümmler, and B. M. Eskofier, “Does the position of foot-mounted imu sensors influence the accuracy of spatio-temporal parameters in endurance running?” *Sensors*, vol. 20, no. 19, p. 5705, 2020.
- [63] C.-C. Yang and Y.-L. Hsu, “A review of accelerometry-based wearable motion detectors for physical activity monitoring,” *Sensors*, vol. 10, no. 8, pp. 7772–7788, 2010.
- [64] R. W. Motl, M. Weikert, Y. Suh, J. J. Sosnoff, J. Pula, C. Soaz, M. Schimpl, C. Lederer, and M. Daumer, “Accuracy of the actibelt® accelerometer for measuring walking speed in a controlled environment among persons with multiple sclerosis,” *Gait & posture*, vol. 35, no. 2, pp. 192–196, 2012.
- [65] J. M. Van Ancum, K. S. van Schooten, N. H. Jonkman, B. Huijben, R. C. Van Lummel, C. G. Meskers, A. B. Maier, and M. Pijnappels, “Gait speed assessed by a 4-m walk test is not representative of daily-life gait speed in community-dwelling adults,” *Maturitas*, vol. 121, pp. 28–34, 2019.
- [66] S. Del Din, A. Godfrey, B. Galna, S. Lord, and L. Rochester, “Free-living gait characteristics in ageing and parkinson’s disease: impact of environment and ambulatory bout length,” *Journal of neuroengineering and rehabilitation*, vol. 13, no. 1, p. 46, 2016.
- [67] H. Gaßner, P. Sanders, A. Dietrich, F. Marxreiter, B. M. Eskofier, J. Winkler, and J. Klucken, “Clinical relevance of standardized mobile gait tests. reliability analysis between gait recordings at hospital and home in parkinson’s disease: A pilot study,” *Journal of Parkinson’s disease*, no. Preprint, pp. 1–11, 2020.
- [68] D. Gijbels, B. Eijnde, and P. Feys, “Comparison of the 2-and 6-minute walk test in multiple sclerosis,” *Multiple Sclerosis Journal*, vol. 17, no. 10, pp. 1269–1272, 2011.
- [69] F. A. Storm, A. Cesareo, G. Reni, and E. Biffi, “Wearable inertial sensors to assess gait during the 6-minute walk test: a systematic review,” *Sensors*, vol. 20, no. 9, p. 2660, 2020.
- [70] A. Atrsaei, M. F. Corrà, F. Dadashi, N. Vila-Chã, L. Maia, B. Mariani, W. Maetzler, and K. Aminian, “Gait speed in clinical and daily living assessments in parkinson’s disease patients: performance versus capacity,” *npj Parkinson’s Disease*, vol. 7, no. 1, pp. 1–11, 2021.
- [71] E. Chiauzzi, C. Rodarte, and P. DasMahapatra, “Patient-centered activity monitoring in the self-management of chronic health conditions,” *BMC medicine*, vol. 13, no. 1, pp. 1–6, 2015.