MASTER THESIS

Multi-variable optimization problem for tailoring ankle-foot orthosis stiffness to end-users' needs

Emma A. Gille

FACULTY OF ENGINEERING TECHNOLOGY
DEPARTMENT OF BIOMECHANICAL ENGINEERING

Examination committee:
dr. E.H.F. van Asseldonk
dr. C. Bayón Calderón
dr.ir. F.J. Wouda

CENTRO DE AUTOMÁTICA Y ROBÓTICA

UNIVERSITY OF TWENTE.

Multi-variable optimization problem for tailoring ankle-foot orthosis stiffness to end-users' needs

Emma A. Gille[1]

***Abstract*—Ankle Foot Orthoses (AFOs) are devices commonly used to assist or rehabilitate gait, with stiffness being a key parameter influencing their effectiveness. However, the absence of clinical guidelines for tuning AFO stiffness often results in inconsistent outcomes and low user satisfaction. This may be due to two key challenges: the limited number of gait variables considered, which may not capture the full complexity of walking, and variability in how stiffness relates to gait variables across individuals.**

To address these challenges, we propose a user-tailored optimization framework that identifies and incorporates relevant gait variables based on individual needs and preferences. First, key performance variables across five gait domains were examined in five participants with cerebral palsy (CP) to assess their relevance for predicting stiffness levels, using SHapley Additive exPlanations (SHAP) to interpret feature importance. Subsequently, user and healthcare professional preferences were integrated into the stiffness optimization framework.

Our findings highlight the importance of multiple performance variables in capturing gait complexity and reveal that the most relevant variables differ between participants. Within the optimization framework, we identified a stiffness level that minimized the total error for each participant. The optimal stiffness varied between the participants, emphasizing the need for a personalized approach to stiffness optimization. Incorporating user and clinician preferences did not alter the optimal stiffness levels.

***Index Terms*—Cerebral palsy, ankle foot orthosis, stiffness optimization, performance variables, gait analysis, user & clinician preferences**

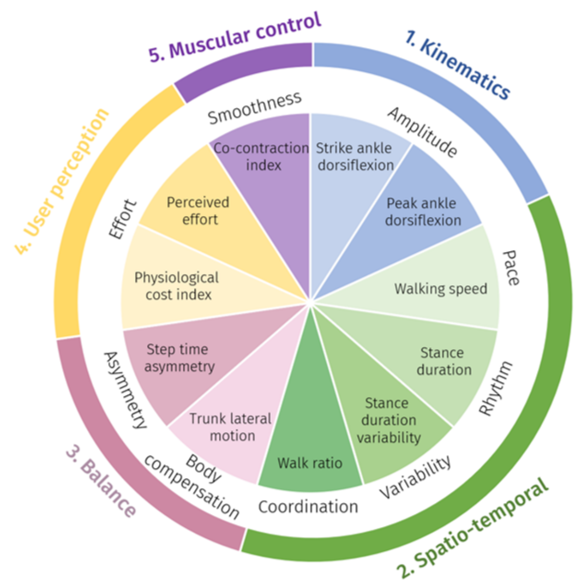


Fig. 1: Five domains that represent different aspects of gait. Several performance variables (pie chart at the center) were selected in this work to assess these gait aspects.

I. Introduction

Cerebral palsy (CP), with a prevalence of 1.7 to 3.1 per 1,000 live births [1], is a group of motor disorders caused by early damage to the developing brain [2]. It is a neurological condition that primarily affects movement, muscle tone, and posture, often resulting from brain injury or abnormal brain development during early childhood [3]. The severity and type of motor impairments vary depending on the location and extent of brain damage [3].

Walking is a fundamental aspect of daily life, playing a crucial role in children's development by enabling mobility, participation in activities, and social engagement [3]. For children with CP, improving mobility is a key goal in achieving greater independence and an active lifestyle [4].

Ankle foot orthoses (AFOs) are commonly used to assist or rehabilitate gait in people with motor or neurological disorders like CP [5]. These devices provide support to the ankle joint, correct gait deviations, and effectively reduce energy cost when walking. The AFO stiffness is a feature that highly influences the final user's walking performance [5]–[7]. Therefore, to maximize the benefits of an AFO, it is very important that this stiffness is tailored to the end-user's needs.

Optimal stiffness varies depending on the user's capacity [8], the task [9] and the terrain type [10], among other factors. Therefore, objective data of the performance of the user in all these different conditions would be useful to inform clinical-decision making in prescribing new AFOs [11]. However, clear clinical guidelines to determine a personalized value of the AFO's stiffness are lacking [11], and available experimental methods still face challenges like long measuring times and lack of inclusion of user preferences [5].

Human in the loop optimization (HILO) is a method aiming to optimize AFO stiffness based on an objective function that most often targets metabolic cost [12]. This target variable is computed indirectly from breath by breath measurements of carbon dioxide and oxygen. Nevertheless, metabolic cost requires a long measuring time, and even when a reduction in metabolic cost is achieved via HILO, this improvement is not reliably perceived by the user [13]. This can be explained by the fact that users prioritize several aspects of gait simultaneously, with metabolic cost not being the highest priority. Therefore, trying to optimize a single physiological measurement (i.e., metabolic cost) without considering other

[1]Department of Biomechanical Engineering, University of Twente, The Netherlands. Correspondence: e.a.gille@student.utwente.nl

variables and user preferences, may not yield the desired results in terms of user satisfaction and acceptance [13].

Gait domains, aside from effort, that may also be relevant for optimization include kinematics, spatio-temporal parameters, balance, user perception, and muscular control (Fig.1). Variables related to these domains are typically not used in combination during AFO stiffness optimization, limiting the understanding of gait performance [12]. When multiple domains are considered, the combination is often partial and constrained by predefined decision schemes [14]. These decision methods assess the performance of each target variable at different stiffness levels and rank optimal stiffness based on the first target variable. In case of a tie, they use subsequent variables one by one to make a decision. The hierarchy of target variables in these decision schemes is manually selected by researchers, potentially leading to high variability in the results between participants [14].

Tankink et al. [15] highlighted the importance of selecting the appropriate cost function (including the objective and variables), as different choices can highly influence the results and, in turn, affect the ability to design personalized features that suit the individual. This emphasizes that a "one-size-fits-all" approach may be suboptimal. Therefore, optimization of AFO stiffness should be improved by incorporating a tailored multi-variable hierarchy that aligns with individual user's priorities.

The aim of this thesis is to develop an alternative method that better captures the complexity of gait while incorporating user preferences and achieving faster results than HILO methods in selecting an optimal stiffness. This leads to the primary research question: **RQ: How can a multi-variable optimization model be designed to determine the optimal stiffness of an Ankle-Foot Orthosis, tailored to the specific needs of individual users?**

A key step in this process is analyzing which variables matter in predicting the stiffness, so the first sub-question is: **rq1**: Which performance variables are most relevant for predicting stiffness, and how does their sensitivity vary across users? If we find that across participants there are variables with weaker correlations to stiffness, those variables could be excluded as they may be less effective for optimization. It would reduce the number of required sensors and complexity of the protocol, and potentially increasing user comfort during the experiment. Focusing on the most relevant variables allows for a more efficient and personalized optimization process.

Beyond performance variables, the inclusion of user and clinician preferences is essential. A key challenge in the optimization process is the divergence between clinical functional objectives and patient preferences [11]. For instance, factors such as replicability of normal walking patterns and adaptability to walking speed have been shown to differ significantly in perceived importance between healthcare professionals and end-users [11]. This possible discrepancy raises the second sub-question: **rq2**: How will the optimal stiffness based on preferences of the user differ from that based on the preferences of the health care professional?

Addressing this question will refine the optimization model, ensuring a balance between clinical effectiveness and user

Fig. 2: Magnified view of the leaf spring-CAM mechanism in the inGAIT-VSO, highlighting the adjustable slider for stiffness adjustment. Figure taken from [17].

satisfaction.

By addressing the above-mentioned research questions, this thesis aims to develop a personalized multi-variable stiffness optimization framework that integrates both user and clinician preferences. The proposed optimization is tested on five pediatric participants with CP using the inGAIT-VSO device [16].

II. METHODS

This section outlines the steps involved in stiffness optimization. First, the inGAIT-VSO is introduced in section II-A. Next, the selection of performance variables used in the optimization is outlined in section II-B1, followed by the feature selection process for identifying key variables in section II-B2. Subsequently, the stiffness optimization process, including the incorporation of user and clinician preferences, is covered in sections II-C and II-D, respectively. Finally, the methods for studying this process with five participants diagnosed with CP are described in section II-E.

A. Variable stiffness orthosis

The inGAIT-VSO is a pediatric-focused Variable Stiffness Orthosis (VSO) designed for children with CP [9], [16], see Fig. 2. It features a non-actuated mechanism that provides a non-linear ankle angle-torque curve, active during both the stance and swing phases. The stiffness curve can be manually adjusted according to user/clinician preferences to impact the user's gait. The inGAIT-VSO is sensorized and allows for data capturing without needing sophisticated lab equipment. For a detailed description of the inGAIT-VSO, refer to [16].

B. Feature selection

1) Performance variables: To take into account different aspects of gait, we defined five domains: 1. Kinematic Properties, 2. Spatio-Temporal Characteristics, 3. Balance and Stability, 4. User Perception and Physiological Demand, and 5. Muscular Control (Fig. 1 & Table I). Within each domain, specific

TABLE I: Summary of performance variables and their corresponding domain, gait aspects, typical developing (TD) values, acceptable bounds (as obtained from the literature), and the computational interval, with "W.S." referring to walking segment and "W.T." to walking trial.

Domain	**Gait Aspect**	**Performance variables**	**TD values**	**Bounds**	**Interval**
Kinematic Properties	Amplitude	Strike ankle dorsiflexion (deg)	-2.15	[-5.9, 1.6]	Gait cycle
		Peak ankle dorsiflexion (deg)	12.5	[9.5, 15.5]	Gait cycle
Spatio-Temporal Characteristics	Pace	Walking speed (m/s)	≥ 1.2	[1, 1.4]	Gait cycle
	Rhythm	Stance duration (%)	57.97	[56, 59.9]	Gait cycle
	Variability	Stance duration variability (%)	≤ 1.93	[0, 1.93]	W.S.
	Coordination	Walk ratio (cm·min/step)	0.45	[0.38, 0.53]	Gait cycle
Balance and Stability	Body compensation	Trunk lateral motion (m/s^2)	≤ 1.47	[1.08, 1.86]	Gait cycle
	Asymmetry	Step time asymmetry (ms)	≤ 32	[25.6, 38.4]	W.S.
User Perception and Physiological Demand	Effort	Physiological cost index (beats/m)	≤ 0.39	[0.3, 0.48]	W.S.
		Perceived effort (Borg's scale)	≤ 12	[6, 12]	W.T.
Muscular Control	Smoothness	Co-contraction index (–)	26.5	[24, 29]	Gait cycle

performance variables were selected to comprehensively assess both functional and physiological aspects of user performance, with a particular focus on the ankle joint. The selection was based on previous literature [18] and prioritized variables that could be measured using wearable and onboard lab equipment, aiming to minimize duplication and redundancy within and between domains.

For the Kinematics domain, the angle between the foot and shank at heel strike, representing the ankle angle as a degree of dorsiflexion (DF), and the peak ankle DF, representing the maximum DF observed during the stance phase for each gait cycle, were computed (Table I). Within the Spatio-Temporal domain, walking speed, stance duration (the percentage of the gait cycle during which the foot is in contact with the ground), and walk ratio (defined as the ratio between step length and cadence) were computed every gait cycle as variables representative of pace, rhythm and coordination respectively. Additionally, the stance duration variability was determined every 20-meter walking segment by looking at the standard deviation of the percentage of stance phase durations computed along that segment. For the Balance and Stability domain, the trunk lateral motion was computed as a form of compensation mechanism every gait cycle, while step time asymmetry (STA) was computed for every 20-meter walking segment. Trunk lateral motion was derived from the root mean square of the lateral acceleration as done in [19]. The STA was computed using the formula presented in [20]:

$$STA = \frac{\sum_{i=2}^{n}(ST_i - ST_{i-1})}{2n} \tag{1}$$

where ST_i is the step time of the step i in milliseconds, and n is the total number of steps on the corresponding walking segment. For the domain User Perception and Physiological Demand, the Physiological Cost Index (PCI) [21], used to assess the energy cost of walking by relating changes in heart rate to walking speed, was calculated every walking segment as:

$$PCI = \frac{HR_{walk} - HR_{rest}}{v_{walk}} \tag{2}$$

where HR_{rest} and HR_{walk} are the heart rate measurements at rest and during walking, respectively, and v_{walk} represents walking speed. Moreover, the Borg's scale [22], a subjective measure of perceived exertion, was used to report user's perceived effort. Finally, for the Muscular Control domain, the co-contraction index (CCI) was computed during stance phase as presented in [23]:

$$CCI = \sum_{i=1}^{101} \frac{EMG_L(i)}{EMG_M(i)}(EMG_L(i) + EMG_M(i)) \tag{3}$$

where i represents the individual time points of stance phase (0–100%, or 101 data points), and $EMG_L(i)$ and $EMG_M(i)$ are the normalized activations of the less and more active muscles at point i, respectively. The CCI quantifies the simultaneous activation of antagonist muscle pairs, indicating the level of muscle co-contraction in the lower leg during walking. Specifically, the antagonist muscles used were the tibialis anterior (TA) and gastrocnemius medialis (GM).

Some performance variables were measured separately for the left and right sides of the body, while others were non-side-dependent. Specifically, trunk lateral motion, STA, PCI, and perceived effort were assessed as non-side-dependent values, whereas all other variables were calculated individually for both the left and right sides.

2) Classification model for feature importance: A classification model was developed using XGBoost, with performance variables as inputs and stiffness as the output. This model enhances the basic gradient boosting method by passing residuals from each decision tree to the next, hierarchically (see Fig. 3). It offers the benefit of being well-suited for high-dimensional datasets and includes inherent regularization to prevent overfitting [24]. The purpose of training this model and employing an explainability framework was to identify the variables that were significantly affected by stiffness. To train the XGBoost model, observations of each performance variable were obtained across a range of stiffness levels (see section II-E). Some variables could only be computed once per walking trial or segment, rather than for each gait cycle (Table I). To ensure consistency with other variables, these values were repeated to match the number of gait cycles within each respective trial or walking segment, resulting in an equal

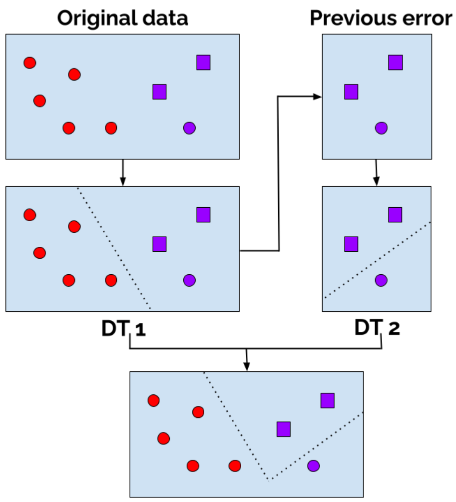


Fig. 3: Schematic illustration of functioning of XGBoost models. Original data has two classes circles and squares. The first decision tree (DT_1) gives an imperfect classification. The errors or residuals from DT_1 are fed as input to the next decision tree (DT_2). DT_1 and DT_2 are combined to give the final prediction

number of data points per variable, as required by SHAP for accurate feature importance assessment.

To quantify the contribution of each variable to the predictions, SHapley Additive exPlanations (SHAP) were employed [25]. SHAP values explain the contribution of each individual variable to the overall prediction. The SHAP value of a variable is calculated by assessing the effect of excluding that variable from the prediction process. If removing the variable has no effect on the result, its SHAP value will be 0, indicating the variable's irrelevance to the prediction. Different combinations of variables were tested to account for interactions between features, ensuring accurate calculation of SHAP values. The XGBoost model for classification and SHAP for result explanation were performed for each participant individually. Since SHAP values do not represent absolute quantities but rather indicate how much each variable deviates from the baseline prediction, they were normalized so that the total SHAP sum for each participant equaled 10.

We created two thresholds for the total SHAP value to categorize variable importance: a low-relevance threshold at 2.5% (0.25 out of 10) and a high-relevance threshold at 10% (1.0 out of 10). Variables with SHAP values below the low-relevance threshold were considered to have minimal influence on the model's prediction. Values between the two thresholds were interpreted as moderately important, while values above the high-relevance threshold indicated a strong contribution to stiffness prediction. These thresholds were selected to avoid overinterpreting minor fluctuations and to support clear groupings of variable influence.

C. *Stiffness Optimization*

After computing the SHAP values to assess the sensitivity of each variable to the stiffness prediction, the goal was to determine the optimal stiffness for each participant's performance. This optimization process has the main objective of minimizing the error between measured participant's data and reference values from typically developing (TD) children (see these TD values and their acceptable bounds in Table I). The TD values were obtained from the literature (i.e.,strike ankle DF, peak ankle DF, walking speed, stance duration, stance duration variability, and walk ratio [26], trunk lateral motion [19], STA [20], PCI [16], perceived effort [22], and CCI [27]). We tried to base our decisions primarily on data from TD children's gait. When pediatric data was unavailable, data from healthy adults was used.

For each gait cycle, the errors between the computed performance variables and the TD values were calculated using the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(f(X_{TD}, X_{CP}))^2} \tag{4}$$

where N is the number of gait cycles within a walking trial for a specific variable, X_{TD} is the TD value of that variable (from Table I), and X_{CP} represents the computed value of the same variable. RMSE was chosen because it penalizes larger errors more heavily, making it more sensitive to outliers. These errors were calculated for each gait cycle and performance variable. The error function $f(X_{TD}, X_{CP})$ was defined in three ways, based on the characteristics of the performance variables.

For some performance variables, such as strike ankle DF, peak ankle DF, stance duration, walk ratio, and CCI, deviations in either direction of the TD value contributed to the error:

$$f(X_{TD}, X_{CP}) = X_{TD} - X_{CP} \tag{5}$$

For walking speed, exceeding the TD value was not penalized because higher speeds were considered beneficial. Consequently, only values lower than the TD value were penalized:

$$f(X_{TD}, X_{CP}) = \begin{cases} X_{TD} - X_{CP} & \text{if } X_{CP} < X_{TD} \\ 0 & \text{if } X_{CP} \geq X_{TD} \end{cases} \tag{6}$$

For variables where a reduction was generally beneficial, such as stance duration variability, trunk lateral motion, STA, PCI, and perceived effort, only values exceeding the TD value were penalized:

$$f(X_{TD}, X_{CP}) = \begin{cases} 0 & \text{if } X_{CP} \leq X_{TD} \\ X_{TD} - X_{CP} & \text{if } X_{CP} > X_{TD} \end{cases} \tag{7}$$

To ensure comparability between errors from different variables, we normalized the RMSE, making the errors unitless and standardized, see Equation 8. For this normalization, the bounds from Table I were used to account for natural gait variability, as they define the acceptable range considering TD values. The normalization ensures that the error magnitude is strongly influenced by the range of the bounds, which is intentional, as larger ranges typically lead to larger errors.

$$E_{norm} = \frac{RMSE}{b_{high} - b_{low}} \tag{8}$$

where E_{norm} is the normalized error for a specific variable, and b_{high} and b_{low} represent the upper and lower acceptable bounds for the TD values of the same variable, respectively.

Subsequently, the normalized errors were used in the cost function that aims to match the TD values and computed performance variables:

$$\min_{X_{CP}} \sum_{v=1}^{11} E_{norm} \tag{9}$$

where from (4) and (8)

$$E_{norm} = \frac{1}{b_{high} - b_{low}} \sqrt{\frac{1}{N} \sum_{n=1}^{N} (f(X_{TD}, X_{CP}))^2} \tag{10}$$

In these equations v indexes the performance variables. The normalized errors of all performance variables per trial were summed, and the stiffness that resulted in the lowest total error was identified.

The errors that fall within the bounds of Table I were considered acceptable, as they represent the natural variation observed in TD children. Using Equation 8, the error resulted in an acceptable value of 0.5 for the following variables: strike ankle DF, peak ankle DF, walking speed, stance duration, walk ratio, trunk lateral motion, STA, PCI, and CCI. For the remaining variables, the acceptable error was 0 since one of their bounds coincided with their TD values. The total acceptable error for each participant, referred to as the acceptable error line, was calculated by summing the acceptable errors for all variables included in their optimization.

D. User and clinician preferences

To assess user and clinician preferences for the optimization, a questionnaire was designed with two main components: Likert scale questions and a ranking task. For the first question, participants rated the importance of each considered domain (see Table I) using a Likert scale. Each domain received a score between 0.2 and 1 (in increments of 0.2), with higher values indicating greater importance. For the second question, participants selected their top three preferred domains through a ranking task. Bonus points were assigned based on rank: +1 for the first choice, +0.8 for the second, and +0.6 for the third. The full questionnaire is available in Appendix VI-A, and the completed responses can be found in Table III.

The questionnaire scores quantified the weight assigned to each domain, reflecting its relative preference. The total domain weight was calculated as follows:

$$W_{pref} = W_{likert} + W_{rank} \tag{11}$$

where W_{pref} is the total weight derived from the questionnaire, W_{likert} is the weight assigned based on the Likert scale responses, and W_{rank} is the weight derived from the ranking task.

To ensure a uniform application of weights, the total weight for each domain was divided by the number of variables within that domain, with the division accounting for variables measured on both sides (left and right) as individual entities:

$$W_{pref,v} = \frac{W_{pref}}{n_{domain}} \tag{12}$$

where $W_{pref,v}$ is the adjusted weight, accounting for variable quantity, and n_{domain} denotes the number of variables within a domain. These adjusted weights were then applied to the mean normalized errors of the performance variables, resulting in the following modification considering Equations 9 and 10:

$$\min_{X_{CP}} \sum_{v=1}^{11} \frac{W_{pref,v}}{b_{high} - b_{low}} \sqrt{\frac{1}{N} \sum_{n=1}^{N} f(X_{TD}, X_{CP}))^2} \tag{13}$$

where $W_{pref,v}$ is the adjusted preference weight given by (12).

E. Study

1) Participants: Five participants with CP took part in this study (weight 37 ± 19.3 kg, height 1.38 ± 0.17 m, age 9.4 ± 2.8 years-old) (Table II). Participants met the following inclusion criteria: aged 5–17 years, predominantly spastic uni- or bilateral CP, Gross Motor Function Classification System (GMFCS) [28] levels I–III, sufficient cognitive ability, and the capacity to complete the walking protocol. Flexible equinus deformities or drop-foot were preferred, with Ashworth scale scores ranging from 1 to 3. Exclusion criteria included leg surgery in the previous 6 months or botulinum toxin A injections in the previous 3 months, significant musculoskeletal deformities, unhealed skin lesions, gastrocnemius shortening $> 10°$, or visual impairments or behavioral issues that could hinder protocol performance. The Local Ethics Committee at Hospital Infantil Universitario Niño Jesús (HNJ) gave the approval (R-0064/23) and ensured the study was conducted in alignment with the Declaration of Helsinki. Participants and their families were informed accordingly, and parental consent was obtained before participation.

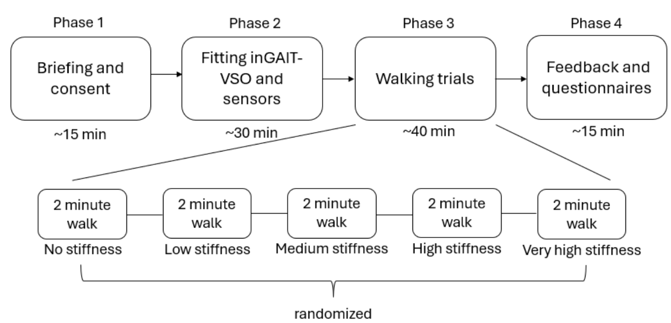


Fig. 4: Flowchart of the study layout.

2) Experimental protocol: The study involved a single session per participant of about 2h. Initially, participants were assisted in putting on the inGAIT-VSO bilaterally or on their affected side based on their clinical manifestation (diplegia or hemiparesis), and were given the opportunity to walk with the equipment and report any discomfort. After a brief familiarization of approximately one minute with the device, they performed five trials of the two-minute walking test (2mwt) [29], walking back and forth along a 20-meter straight

TABLE II: Participants characteristics.

Participant ID	Age (years-old)	Weight (kg)	Height (m)	Shoe size (EU)	Clinical manifestation	GMFCS	Sex
CP01	9	38	1.39	36	Hemiparesis (left)	I	Male
CP02	10	24	1.34	33	Spastic diplegia	III	Female
CP03	14	74	1.69	41	Hemiparesis (right)	I	Female
CP04	5	26	1.2	31	Hemiparesis (right)	I	Male
CP05	9	23	1.3	32	Spastic diplegia	II	Male

corridor. For each trial, the inGAIT-VSO was set to a different stiffness level: (1) No stiffness (spring disconnected), (2) Low stiffness (0.2 Nm/kg as peak restoring torque at 12 deg of DF), (3) Medium stiffness (0.4 Nm/kg at 12 deg of DF), (4) High stiffness (0.6 Nm/kg at 12 deg of DF), and (5) Very high stiffness (0.8 Nm/kg at 12 deg of DF). The value of 12 degrees of DF was chosen in accordance with previous literature [16], [30] as it is the approximate maximum value seen during stance in healthy gait. To minimize fatigue-related bias, the order of trials (i.e., stiffness levels) was randomized for each participant. Participants were allowed to rest as needed between trials.

After the five trials, participants and their physiotherapist completed the questionnaire presented in section II-D to assess their preferences. A detailed flowchart of the study protocol is provided in Fig. 4.

3) Data acquisition: During the 2mwt trials, a combination of sensors was used to compute the required performance variables listed in Table I. Magnetic encoders (AS5048b, AMS-OSRAM AG, Premstaetten, Austria) integrated in the inGAIT-VSO recorded the ankle angle in the sagittal plane at 100 Hz. Force-sensitive resistors (FSRs, FlexiForce A502, Tekscan Inc, MA, USA) embedded within the insoles were used to detect heel strike and toe-off events also at 100 Hz [16]. Two GaitUp sensors (Physilog, Lausanne, Switzerland) were attached to the shoes to monitor foot movements sampled at 128 Hz. Surface electromyography (EMG) electrodes (Trigno Delsys, Natick, MA, USA) were bilaterally placed on the TA, GM, soleus (SOL), and gastrocnemius lateralis (GL) muscles to register muscle activity with a sampling rate of 1926 Hz. An inertial measurement unit (IMU) was attached to the trunk of the participants at the sternum height, and recorded torso motion at 100 Hz. Finally, a smartwatch (Versa 3, Fitbit, San Francisco, USA) was used to measure heart rate at 1 Hz. Participants also reported their perceived effort after each trial using the Borg's scale [22].

4) Data processing: Data from the sensors of the inGAIT-VSO, EMG sensors, IMU sensor, and smartwatch were processed using MATLAB 2021b (MathWorks, Natick, MA, USA). Data from the GaitUp sensors were processed using GaitUp Lab (Physilog, Lausanne, Switzerland). EMG data were pre-processed to remove noise and artifacts. This involved band-pass filtering (30-300 Hz), full-wave rectification, and low-pass filtering (3 Hz). The resulting linear envelopes were normalized to the maximum activation observed in the no stiffness walking trial.

After pre-processing, all data were resampled to the frequency of the inGAIT-VSO and synchronized. The data were then segmented into gait cycles using heel strike events detected by the FSRs. Each gait cycle was linearly interpolated, resulting in 300 data points per cycle. Variables were calculated as described in section II-B1 with a computation interval specified in Table I.

III. Results

Data acquisition during the study sessions went smoothly for all participants, allowing the computation of performance variables for each trial. As CP03 exceeded the inGAIT-VSO's weight limit of 60 kg, this participant did not perform the high and very high stiffness conditions due to design limitations.

Performance variables that depended on the side where the inGAIT-VSO was worn (i.e., strike ankle DF and peak ankle DF) were computed bilaterally for participants wearing the device on both sides (CP02 and CP05), and unilaterally for the affected side of participants with hemiparesia (CP01, CP03, CP04).

A. Feature selection for stiffness prediction

Trained XGBoost models for each participant derived SHAP values representing the contribution of each variable to stiffness prediction (Fig. 5). The variables did not consistently hold the same level of influence across participants and stiffness levels. For CP01, the SHAP values indicated that peak ankle DF had a dominant influence on the model's predictions, with a SHAP value of 6.4, which is well above the high-relevance threshold for strong contribution. All other variables were below this threshold. In contrast, for CP02, CP03, CP04, and CP05, the influence was more evenly distributed. CP02, CP04, and CP05 each had at least four variables with SHAP values above the high-relevance threshold, while CP03 had three.

Interestingly, certain variables were important for stiffness prediction across multiple participants. Peak ankle DF and PCI had SHAP values above the high-relevance threshold in four participants, while stance duration variability and STA exceeded this threshold in three. On the other hand, some variables did not contribute meaningfully to the prediction for any participant, with SHAP values consistently below the low-relevance threshold. These included walking speed, stance duration (except on the right side for CP01), and walk ratio, all on both sides.

B. Stiffness Optimization

Stiffness optimization revealed a stiffness level that minimized the total error for each participant (Fig. 6). These optimal stiffness levels varied among participants, with no single level being optimal for all. Specifically, CP02 and CP03 showed clear minima at Medium stiffness, CP04 at High, and

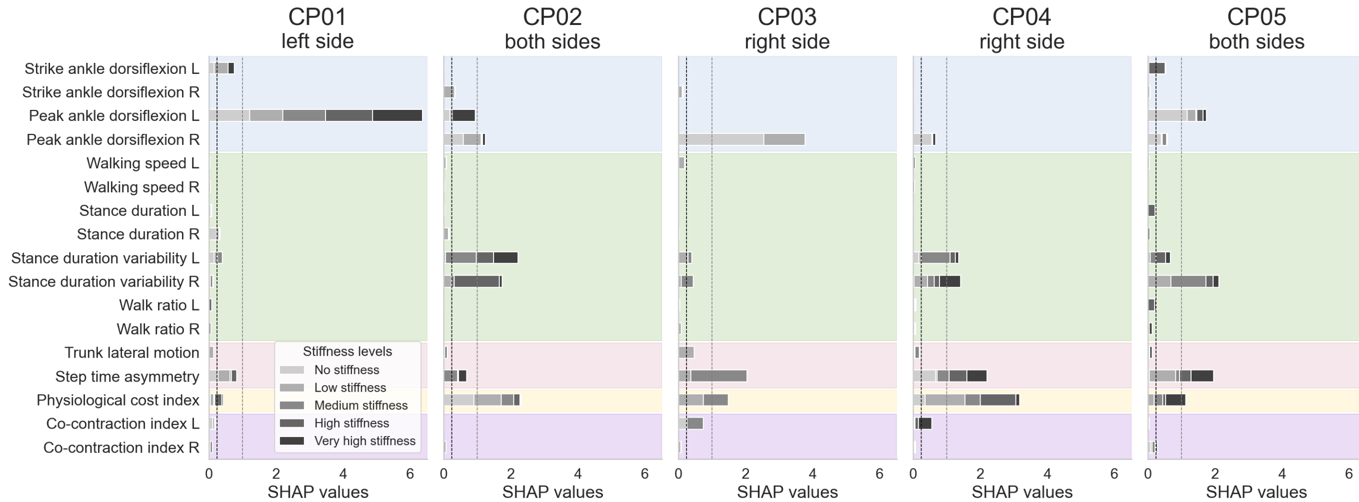


Fig. 5: SHAP values for all participants. Bars represent cumulative SHAP values per variable, indicating their impact on stiffness prediction — larger bars signify greater influence. Titles indicate whether the inGAIT-VSO was worn on the left, right, or both legs. "L" and "R" on the y-axis refer to the side where variables were computed. The black vertical dashed line marks the low-relevance threshold (2.5%), and the grey vertical dashed line marks the high-relevance threshold (10%).

CP05 at Low. For CP01, Low stiffness also resulted in the lowest error, though the difference between the lowest and second-lowest errors was small.

Notably, the magnitude of the total errors differed between participants. The errors for CP01 and CP03 were approximately 2 to 3 times higher than their acceptable error lines, while the errors for the other participants were 3 to 5 times higher than their acceptable error lines.

Large errors from some variables such as strike ankle DF, stance duration, walk ratio, or CCI are present in all error bars across participants. However, there are notable differences between participants. For instance, stance duration variability contributes more to the error for CP02, CP04, and CP05, while it is less influential for the other participants. Additionally, step time asymmetry resulted in higher errors for CP01, CP02, and CP04 compared to the other participants.

Interestingly, some variables that did not reflect a high SHAP value (e.g. stance duration or walk ratio, Fig. 5) did have an influence on the total error in the optimization (Fig. 6). Other variables, as peak ankle DF, stance duration variability, STA, or PCI, influenced both the total error and had high SHAP values in some participants.

C. User and Clinician Preferences

The questionnaire responses revealed differences in domain priorities between users and clinicians (Table III). Users consistently rated Balance and Stability as important, with all users giving it a 5 on the Likert scale and ranking it in their top three. In contrast, clinicians prioritized Kinematics, with four out of five rating it as a 5 and all ranking it in their top three.

Incorporating user and clinician preferences in the optimization problem influenced the magnitude of the total errors, as shown by the varying sizes of the error bars compared to the results without preferences (Fig. 7). However, neither user nor clinician preferences affected the optimal stiffness level or the overall optimization results for each participant. Notably, for all participants except CP02, incorporating user preferences resulted in larger error bars across all stiffness levels compared to clinician preferences.

IV. Discussion

In this thesis, an optimization framework was developed to determine the optimal stiffness of a variable stiffness orthosis (inGAIT-VSO). The framework was applied to five pediatric participants with CP. Within this optimization framework, a series of performance variables were selected to capture different aspects of gait. These variables were also evaluated to explore their relevance for stiffness prediction.

A. Feature selection for stiffness prediction

Key findings revealed participant-specific differences in the influence of performance variables on stiffness predictions (Fig. 5). For CP01, a single dominant variable (i.e., peak ankle DF) stood out. In contrast, the other participants had a broader set of variables that shared contributions for stiffness prediction. These variables spanned different gait domains, suggesting that excluding any of them could result in a loss of important information for characterizing gait. Interestingly, for all cases analyzed, at least one variable from each gait domain–except for the Muscular Control domain–exceeded the chosen high-relevance threshold of 10%. This suggests that the Muscular Control domain might have the least overall influence on stiffness prediction in this group of participants.

Some patterns emerged regarding which variables were classified as important. For instance, peak ankle DF and PCI exceeded the high-relevance threshold for four out of five participants, consistently playing an important role in stiffness prediction. Similarly, STA and stance duration variability

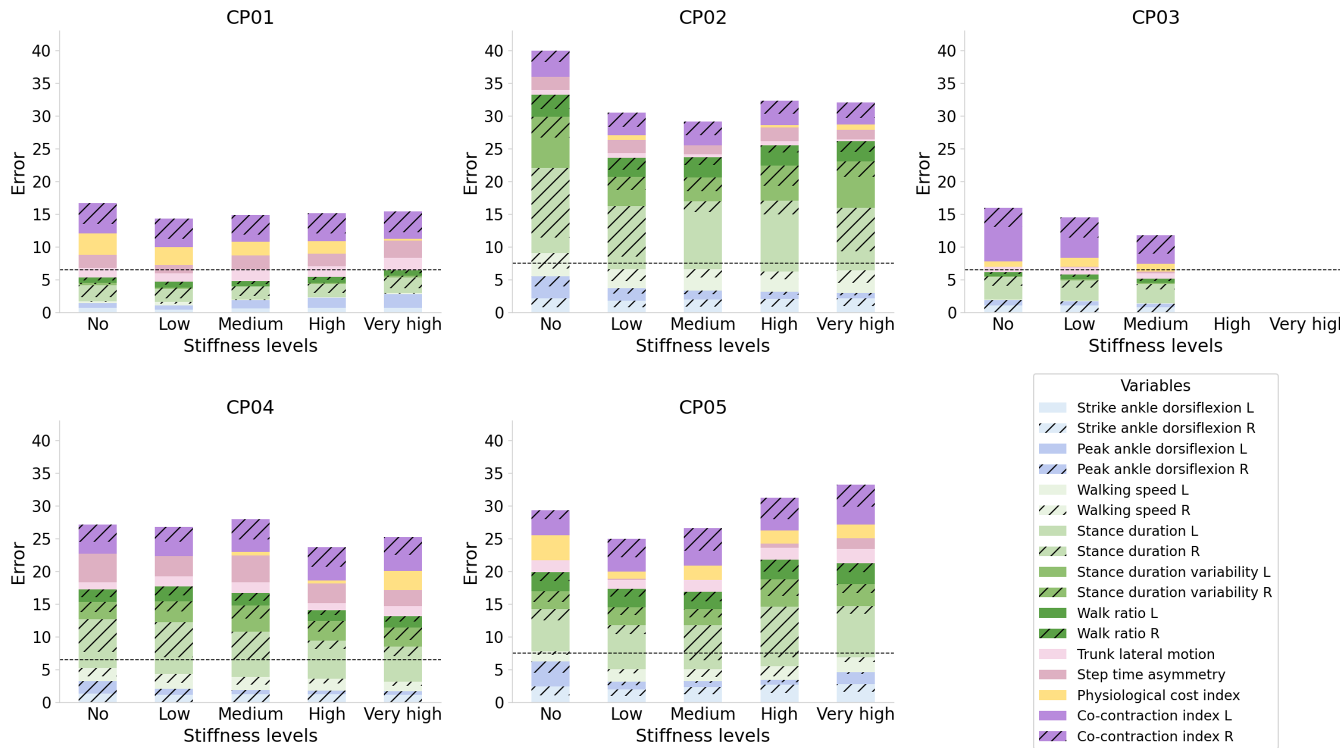


Fig. 6: Total errors across stiffness levels for all participants. Dashed horizontal lines indicate the acceptable error lines: 6.5 for CP01, CP03, and CP04, and 7.5 for CP02 and CP05, reflecting differences in the number of performance variables.

surpassed the high-relevance threshold for three out of five participants.

Contrarily, walking speed and walk ratio consistently fell below the low-relevance threshold across all participants, indicating minimal influence on stiffness prediction. Stance duration also showed limited importance, exceeding the low-relevance threshold only once. This suggests that these variables contributed little to stiffness prediction in this group. For walking speed, this low relevance could be due to some participants not fully understanding the instruction to walk as fast as possible, while others may have been affected by fatigue or varying motivation.

While some common patterns were visible, individual variation across participants remained. This underscores the risk of relying on a limited set of variables or gait domains as done in other studies [12], [14], as they may not fully capture the complexity of individual gait patterns. Our findings support the need for customized cost functions and argue against "one-size-fits-all" approaches in optimization frameworks.

B. *Stiffness Optimization*

We identified the stiffness level that minimized total error for each participant (Fig. 6), however, it is difficult to assess whether the optimization was truly successful, as there is no defined ground truth to compare with. Stiffness tuning in clinical practice is traditionally based on empirical methods such as manual adjustment, patient feedback, and observational gait assessment [31]. These methods are subjective, relying heavily on the clinician's experience and interpretation, which complicates the establishment of a reliable ground truth. This subjectivity also highlights the importance of developing an objective optimization framework that evaluates stiffness levels based on quantifiable gait outcomes rather than solely on visual assessment or patient-reported feedback.

Some correlations have been observed between the computed errors and the level of GMFCS, with larger errors corresponding to higher levels of GMFCS. Participants CP01 and CP03, both classified as GMFCS I, had total errors roughly two to three times their respective acceptable error lines (see Fig. 6). In contrast, CP02 and CP05 classified as GMFCS Level III and II respectively, showed total errors at least four to six times higher than the acceptable error line. For CP04 (GMFCS I), the large errors observed may be attributed to the participant's young age (only 5 years old).

An overlap between feature importance and optimization results would suggest that the features influencing the prediction also produce errors in the optimization. However, a lack of overlap is not necessarily negative: SHAP reflects how much a variable changes across stiffness levels, whereas stiffness optimization focuses on the size of the error of that variable. For example, a variable with a consistently high error might strongly affect optimization results but have little influence on SHAP as it does not vary much between stiffness levels.

Initially, we considered excluding the variables that consistently fell below the low-relevance threshold (SHAP $<$ 2.5%) from the optimization. That would maybe allow to

TABLE III: User and clinician responses to questions 1 and 2 of the preference questionnaire. Question 1 used a Likert scale, where 5 indicated highest importance and 1 indicated lowest importance. Question 2 was a ranking task where 1 = most important, 2 = second, and 3 = third. Unranked domains were not in the top three.

	User preferences									
	Likert scale					**Ranking task**				
Participant ID	Kinematics	Spatio-Temporal	Balance	User Perception	Muscular Control	Kinematics	Spatio-Temporal	Balance	User Perception	Muscular Control
CP01	4	5	5	4	4		1	2		3
CP02	5	3	5	4	4	2		1		3
CP03	5	5	5	5	5			1	3	2
CP04	4	4	5	5	5		1	2	3	
CP05	3	4	5	5	5			3	2	1
	Clinician preferences									
	Likert scale					**Ranking task**				
Participant ID	Kinematics	Spatio-Temporal	Balance	User Perception	Muscular Control	Kinematics	Spatio-Temporal	Balance	User Perception	Muscular Control
CP01	5	4	4	4	4	1	2		3	
CP02	5	5	3	4	4	1	2			3
CP03	5	5	5	5	5	1		3	2	
CP04	4	3	4	5	5	3			2	1
CP05	5	4	3	5	4	2	1		3	

reduce computation time and perform future experiments with a simplified protocol (less sensors). However, SHAP results varied across participants, and only walking speed and walk ratio consistently fell below this threshold. As keeping these variables would not eliminate any of the sensors used and did not affect computation time, we finally retained all of them. For each participant, excluding the low-relevance variables from the stiffness optimization was performed and showed no effect on the lowest total errors achieved (see Appendix VI-B).

The optimization framework relies on TD values and bounds from literature, which significantly influence the optimization. In cases where pediatric gait data were unavailable, we selected the most relevant alternatives, including data from healthy adults, ensuring that each variable was supported by one reliable source and was appropriate within the context of our study. While gait characteristics can vary slightly across age groups, a sensitivity analysis (Appendix VI-C) shows that the optimization is robust to small variations in TD values and bounds, indicating that minor differences in literature-based inputs do not substantially affect the results.

C. Comparability across participants and domains

The optimization results are not directly comparable across participants due to differences in the number of variables included. Participants CP01, CP03, and CP04 wore the AFO on only one side, meaning that Kinematic variables were calculated unilaterally. The remaining participants wore the AFO on both sides, resulting in bilaterally computed Kinematic variables. We considered averaging the contributions of bilateral side-dependent variables to allow comparisons across participants. However, we did not do so because that would have diminished the ability to evaluate each side independently and would have required making assumptions about the relative importance of each side.

Similarly, we chose not to normalize stiffness optimization based on the number of variables per domain, allowing domains with more variables to have a greater impact on the optimization. This decision stems from the fact that variable importance across domains is not directly comparable; for example, the four variables from the Spatio-Temporal domain do not collectively hold the same significance as the single variable from the Muscular Control domain. Since each variable was carefully selected to minimize redundancy, we treated them as equally important, basing stiffness optimization on performance variables rather than gait domains.

D. User and clinician preferences

As previously reported in [11], our findings also revealed a divergence between user and clinician preferences (Table III). Users consistently rated Balance and Stability as highly important, whereas clinicians prioritized Kinematics.

With the inclusion of weights reflecting user and clinician preferences in the optimization, the optimal stiffness levels did not change for any of the participants (Fig. 7). The weights only altered the magnitude of the errors, not the optimization results. This may be due to the weights assigned to the questionnaire responses. To explore the influence of weight selection, two alternative weighting approaches were tested: (1) only using responses from the Likert scale without considering the ranking task, and (2) emphasizing the high-value responses of both the Likert scale and the top-ranked domain (see Appendix VI-D for details).

While some differences between approaches were observed (Fig. 10), they had no effect on the final optimization outcomes. This suggests that the choice of weighting approach had minimal influence on the optimization results, indicating that the stiffness optimization is primarily driven by performance variables rather than preferences. Future work could explore additional methods to incorporate user and clinician preferences, potentially leading to higher levels of user satisfaction and acceptance.

E. Study limitations

The inclusion of only five participants limits the generalizability of the findings. However, even within this small group, the results consistently pointed to the need for individualized stiffness optimization. Future work will aim to validate and strengthen these findings.

A second limitation is the variation in computation intervals across variables. Three of the eleven variables were extracted

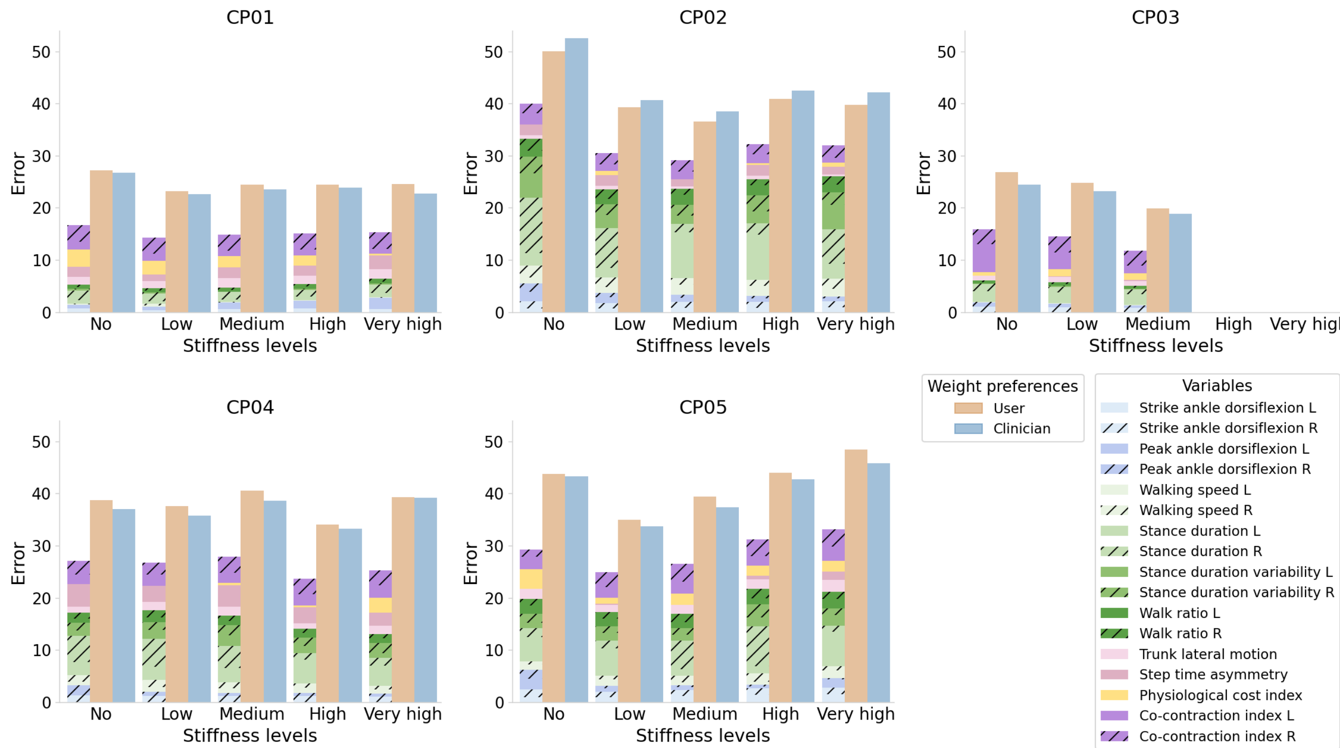


Fig. 7: Optimization incorporating user and clinician preferences for all participants. For each stiffness level, three bars are shown (from left to right): the original optimization, optimization including user preferences, and optimization including clinician preferences.

for each walking segment, and one was calculated only once per trial, resulting in repeated values across gait cycles to match the number of data points of other variables. This repetition could potentially reduce variability within each stiffness level, making differences between stiffness levels appear smaller or larger than they actually are. Furthermore, variables with repeated values may exhibit a lower RMSE due to their limited variation, this can reduce their influence in the stiffness optimization.

The Borg scale, used to estimate perceived effort, was excluded from both feature importance and stiffness optimization due to bias in registering this variable. Some participants repeated the same value across trials, while others appeared to adjust their ratings based on the (randomized) trial order, suggesting they compared each new rating to the previous one rather than reflecting actual perceived effort. Additionally, because only one Borg value was recorded per trial, any variation could disproportionately impact predictions, making it an overly dominant variable for feature importance.

Lastly, the study focused on kinematic and temporal domains but excluded kinetics, as measuring them would have required assessment of forces along the 20-meter corridor. The ankle torque measured with inGAIT-VSO was also excluded to avoid bias in SHAP values, as it directly relates to the AFO stiffness. Including kinetics in future research could provide deeper insights.

V. Conclusion

Within this thesis, a multi-variable optimization framework was developed to identify the optimal stiffness in the inGAIT-VSO for children with CP. To achieve this, we first aimed to identify relevant optimization variables and observe how they varied across users. Our findings show that multiple variables were classified as important, supporting the hypothesis that relying on a single variable is insufficient to capture the complexity of gait. Additionally, the variability in relevant variables among participants highlights the need for individualized stiffness optimization.

Within the optimization framework, we identified an optimal stiffness level for each participant that minimized total error. These levels varied among participants, emphasizing the importance of a personalized approach to stiffness optimization. Incorporating user and clinician preferences into the framework did not change the optimal stiffness levels, suggesting that the optimization is primarily driven by performance variables rather than preferences. Future work may explore additional methods to incorporate user and clinician preferences, potentially leading to higher levels of user satisfaction and acceptance.

Funding

The inGAIT project is funded by the Dutch Research Council NWO (Veni-TTW-2020, ref. number 18079).

ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisors for their guidance and support throughout this project. I am also very thankful to my colleagues at the Center of Automation and Robotics (CAR) in Madrid for their collaboration and friendship. Finally, I am grateful to my family — and especially to Rik and Sarah — for their constant support and for always inspiring me to embrace new opportunities.

REFERENCES

[1] K. Vitrikas, H. Dalton, and D. Breish, "Cerebral palsy: An overview," 2020.

[2] I. Novak, C. Morgan, L. Adde, J. Blackman, R. N. Boyd, J. Brunstrom-Hernandez, G. Cioni, D. Damiano, J. Darrah, A. C. Eliasson, L. S. D. Vries, C. Einspieler, M. Fahey, D. Fehlings, D. M. Ferriero, L. Fetters, S. Fiori, H. Forssberg, A. M. Gordon, S. Greaves, A. Guzzetta, M. Hadders-Algra, R. Harbourne, A. Kakooza-Mwesige, P. Karlsson, L. Krumlinde-Sundholm, B. Latal, A. Loughran-Fowlds, N. Maitre, S. McIntyre, G. Noritz, L. Pennington, D. M. Romeo, R. Shepherd, A. J. Spittle, M. Thornton, J. Valentine, K. Walker, R. White, and N. Badawi, "Early, accurate diagnosis and early intervention in cerebral palsy: Advances in diagnosis and treatment," 2017.

[3] S. Armand, G. Decoulon, and A. Bonnefoy-Mazure, "Gait analysis in children with cerebral palsy," *EFORT Open Reviews*, vol. 1, 2016.

[4] B. C. Conner, N. M. Remec, C. M. Michaels, C. W. Wallace, E. Andrisevic, and Z. F. Lerner, "Relationship between ankle function and walking ability for children and young adults with cerebral palsy: A systematic review of deficits and targeted interventions," 2022.

[5] D. Totah, M. Menon, C. Jones-Hershinow, K. Barton, and D. H. Gates, "The impact of ankle-foot orthosis stiffness on gait: a systematic literature review," *Gait & posture*, vol. 69, pp. 101–111, 2019.

[6] J. S. Lora-Millan, M. Nabipour, E. H. F. van Asseldonk, and C. Bayón, "Advances on mechanical designs for ankle-foot orthoses," *submitted to Frontiers in Bioengineering and Biotechnology*, 2023.

[7] C. Bayón, "Moving forward: The importance of tailored orthotic management in children with cerebral palsy," *Developmental Medicine & Child Neurology*, vol. 00, jun 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1111/dmcn.15684

[8] N. F. Waterval, F. Nollet, J. Harlaar, and M.-A. Brehm, "Modifying ankle foot orthosis stiffness in patients with calf muscle weakness: gait responses on group and individual level," *Journal of NeuroEngineering and Rehabilitation*, vol. 16, pp. 1–9, 2019.

[9] N. V. Crey, D. J. Lam, E. A. Bywater, M. Shepherd, and E. J. Rouse, "The Variable Stiffness Orthosis: Customizable Mechanics for Assistance and Rehabilitation," *TechRxiv*, vol. 27-11, nov 2024.

[10] M. K. Shepherd and E. J. Rouse, "The vspa foot: A quasi-passive ankle-foot prosthesis with continuously variable stiffness," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 12, pp. 2375–2386, 2017.

[11] C. Bayón, M. v. Hoorn, A. Barrientos, E. Rocon, J. P. Trost, and E. H. v. Asseldonk, "Perspectives on ankle-foot technology for improving gait performance of children with cerebral palsy in daily-life: requirements, needs and wishes," *Journal of neuroengineering and rehabilitation*, vol. 20, no. 1, p. 44, 2023.

[12] M. A. Diaz, M. Voss, A. Dillen, B. Tassignon, L. Flynn, J. Geeroms, R. Meeusen, T. Verstraten, J. Babič, P. Beckerle *et al.*, "Human-in-the-loop optimization of wearable robotic devices to improve human–robot interaction: A systematic review," *IEEE Transactions on Cybernetics*, vol. 53, no. 12, pp. 7483–7496, 2022.

[13] R. L. Medrano, G. C. Thomas, and E. J. Rouse, "Can humans perceive the metabolic benefit provided by augmentative exoskeletons?" *Journal of neuroengineering and rehabilitation*, vol. 19, no. 1, p. 26, 2022.

[14] N. F. Waterval, M.-A. Brehm, J. Harlaar, and F. Nollet, "Individual stiffness optimization of dorsal leaf spring ankle–foot orthoses in people with calf muscle weakness is superior to standard bodyweight-based recommendations," *Journal of neuroengineering and rehabilitation*, vol. 18, no. 1, p. 97, 2021.

[15] T. Tankink, J. M. Hijmans, R. Carloni, and H. Houdijk, "Human-in-the-loop optimization of rocker shoes via different cost functions during walking," *Journal of biomechanics*, vol. 166, p. 112028, 2024.

[16] L. van Noort, N. V. Crey, E. J. Rouse, I. Martínez-Caballero, E. H. van Asseldonk, and C. Bayón, "A usability study on the ingait-vso: effects of a variable-stiffness ankle-foot orthosis on the walking performance of children with cerebral palsy," *Journal of neuroengineering and rehabilitation*, vol. 21, 12 2024. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/39090725/

[17] van Noort, "Sensor framework for pediatric cerebral palsy ankle-foot orthosis: Development and technical validation," Master's thesis, University of Twente, 2024.

[18] L. Carcreff, C. N. Gerber, A. Paraschiv-Ionescu, G. De Coulon, C. J. Newman, K. Aminian, and S. Armand, "Comparison of gait characteristics between clinical and daily life settings in children with cerebral palsy," *Scientific Reports*, vol. 10, no. 1, pp. 1–11, 2020.

[19] M. B. Speedtsberg, S. B. Christensen, J. Stenum, T. Kallemose, J. Bencke, D. J. Curtis, and B. R. Jensen, "Local dynamic stability during treadmill walking can detect children with developmental coordination disorder," *Gait and Posture*, vol. 59, 2018.

[20] P. Natarajan, R. D. Fonseka, L. Sy, R. J. Mobbs, and M. Maharaj, "Proposed objective scoring algorithm for clinical evaluation of walking asymmetry in lumbar disc herniation, based on relevant gait metrics from wearable devices: The gait symmetry index (gsitm) – observational study," *Brain and Spine*, vol. 2, 2022.

[21] P. Butler, M. Engelbrecht, R. E. Major, J. H. Tait, J. Stallard, and J. H. Patrick, "Physiological Cost Index of Walking for Normal Children and Its Use As an Indicator of Physical Handicap," *Developmental Medicine & Child Neurology*, vol. 26, no. 5, pp. 607–612, 1984. [Online]. Available: http://doi.wiley.com/10.1111/j.1469-8749.1984.tb04499.x

[22] G. Borg, "Psychophysical scaling with applications in physical work and the perception of exertion," *Scandinavian Journal of Work, Environment and Health*, vol. 16, pp. 55–58, 1990.

[23] B. A. Knarr, J. A. Zeni Jr, and J. S. Higginson, "Comparison of electromyography and joint moment as indicators of co-contraction," *Journal of Electromyography and Kinesiology*, vol. 22, no. 4, pp. 607–611, 2012.

[24] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[25] S. M. Lundberg, G. G. Erion, and S.-I. Lee, "Consistent individualized feature attribution for tree ensembles," *arXiv preprint arXiv:1802.03888*, 2018.

[26] BTS Bioengineering, "Bts bioengineering: Instrumentation for biomechanics and movement analysis," 2023, accessed: 2023-03-04. [Online]. Available: https://www.btsbioengineering.com

[27] Y. Fang and Z. F. Lerner, "How adaptive ankle exoskeleton assistance affects stability during perturbed and unperturbed walking in the elderly," *Annals of Biomedical Engineering*, vol. 51, 2023.

[28] R. Palisano, P. Rosenbaum, S. Walter, D. Russell, E. Wood, and B. Galuppi, "Development and reliability of a system to classify gross motor function in children with cerebral palsy," *Developmental medicine & child neurology*, vol. 39, no. 4, pp. 214–223, 1997.

[29] T. W. Pin and H. L. Choi, "Reliability, validity, and norms of the 2-min walk test in children with and without neuromuscular disorders aged 6-12," *Disability and rehabilitation*, vol. 40, pp. 1266–1272, 5 2018. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/28637155/

[30] E. Owen, "The importance of being earnest about shank and thigh kinematics especially when using ankle-foot orthoses," 2010.

[31] T. Kobayashi, A. K. Leung, and S. W. Hutchins, "Techniques to measure rigidity of ankle-foot orthosis: A review," 2011.

VI. APPENDIX

A. Preference questionnaire

The questionnaire used to assess user preferences is shown here. A similar questionnaire was provided to the clinician, with the questions tailored to the clinician's perspective on the user. The questionnaire is presented on the following two pages (pages 14–15).

PARTICIPANT ID: DATE:

We are going to ask you some questions about how you would like inGAIT to help you. If you don’t understand something, feel free to ask us.

Exercise 1: Indicate how much you would like inGAIT to help you with the following:

1. I would like inGAIT to help me move my **foot up** and **down easily when I walk**

Not important	Little important	Normal	Important	Very important

2. I would like inGAIT-VSO to help me take **bigger steps** and **walk faster**

Not important	Little important	Normal	Important	Very important

3. I would like inGAIT-VSO to help me walk steadily, **without moving** my **body** too much **from side to side**

Not important	Little important	Normal	Important	Very important

4. I would like inGAIT to help me walk for a **longer** time **without getting tired**

Not important	Little important	Normal	Important	Very important

5. I would like my muscles to be **relaxed** and **tension-free** when I walk with the inGAIT

Not important	Little important	Normal	Important	Very important

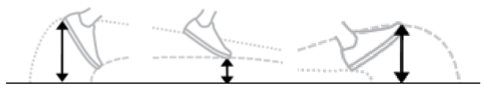
Exercise 2: Choose the most important phrases/the ones you like the most from the following box:

Write a **1** next to the one you find **most important**

Write a **2** next to the one you find **second most important**

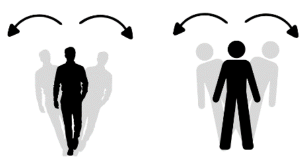
Write a **3** next to the one you find **third most important**

☐ I would like inGAIT to help me move my **foot up** and **down easily when I walk**



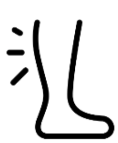
☐ I would like inGAIT to help me take **bigger steps** and **walk faster**

☐ I would like inGAIT to help me walk steadily, **without moving** my **body** too much **from side to side**



☐ I would like inGAIT to help me walk for a **longer** time **without getting tired**

☐ I would like my muscles to be **relaxed** and **tension-free** when I walk with the inGAIT



B. *Optimization excluding low-relevance variables*

Variables with low relevance (SHAP $\leq$ 0.5) were excluded from the optimization (see Fig. 8). This exclusion did not change the optimal stiffness level for any of the participants.

C. *Sensitivity analysis*

To evaluate the effect of slight variations in the TD values and bounds of all variables, perturbations of -10% and +10% were applied to the TD values and, separately, to the range of the bounds. Error calculation and normalization were performed as described in section II-C. The sensitivity bars are present but remain relatively small (Fig. 9), indicating that the introduced perturbations lead to only minor changes in the optimization outcomes.

D. *Preference weight testing*

TABLE IV: Overview of the three preference weighting approaches and their assigned weights: original approach, (1) Likert scale only, and (2) high-value responses. The upper part of the table shows the weights for the first question (Likert scale) of the questionnaire, while the lower part shows the weights for the second question (ranking task).

Likert scale

Response	original	(1) Likert scale only	(2) high-value responses
5	1	1	1
4	0.8	0.8	0.5
3	0.6	0.6	-
2	0.4	0.4	-
1	0.2	0.2	-

Ranking task

Response	original	(1) Likert scale only	(2) high-value responses
Top 1	1	-	1
Top 2	0.8	-	-
Top 3	0.6	-	-

The three weighting approaches, applied to both user and clinician preferences as outlined in Table IV, are illustrated in Fig. 10. The overall optimization results, and specifically the optimal stiffness levels, remained unchanged for all participants.

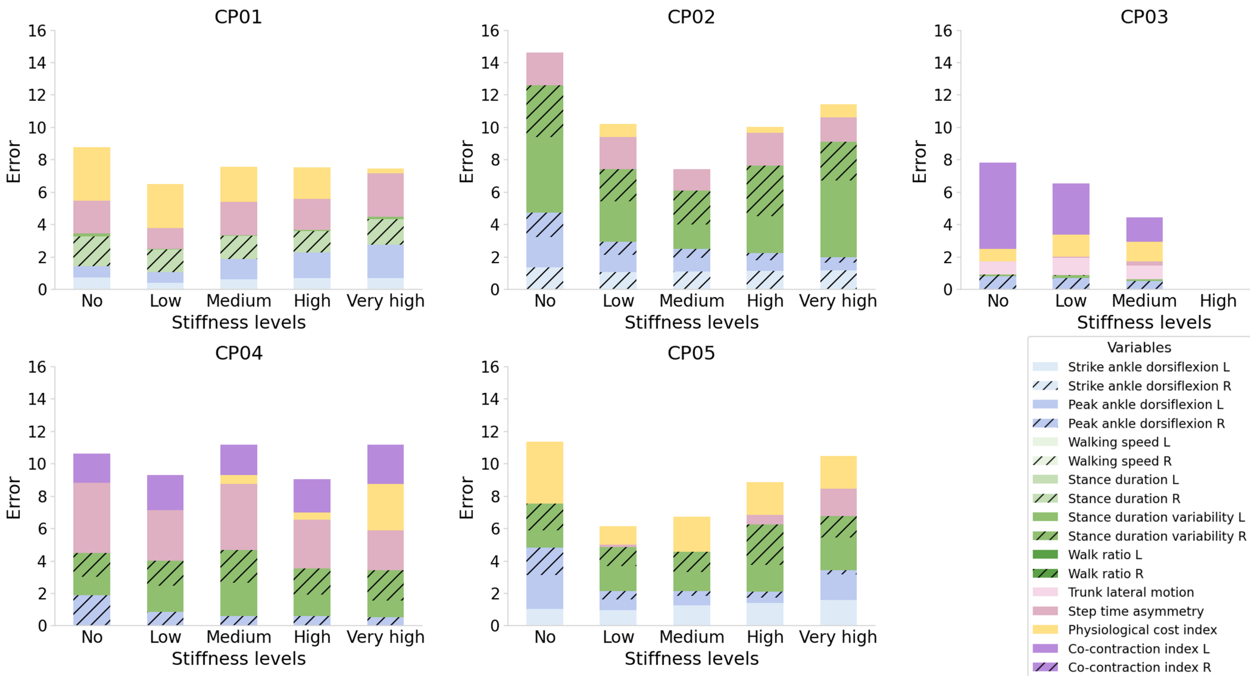


Fig. 8: Optimization excluding variables that fell below the low-relevance threshold (SHAP $< 2.5\%$) in the feature importance for stiffness prediction.

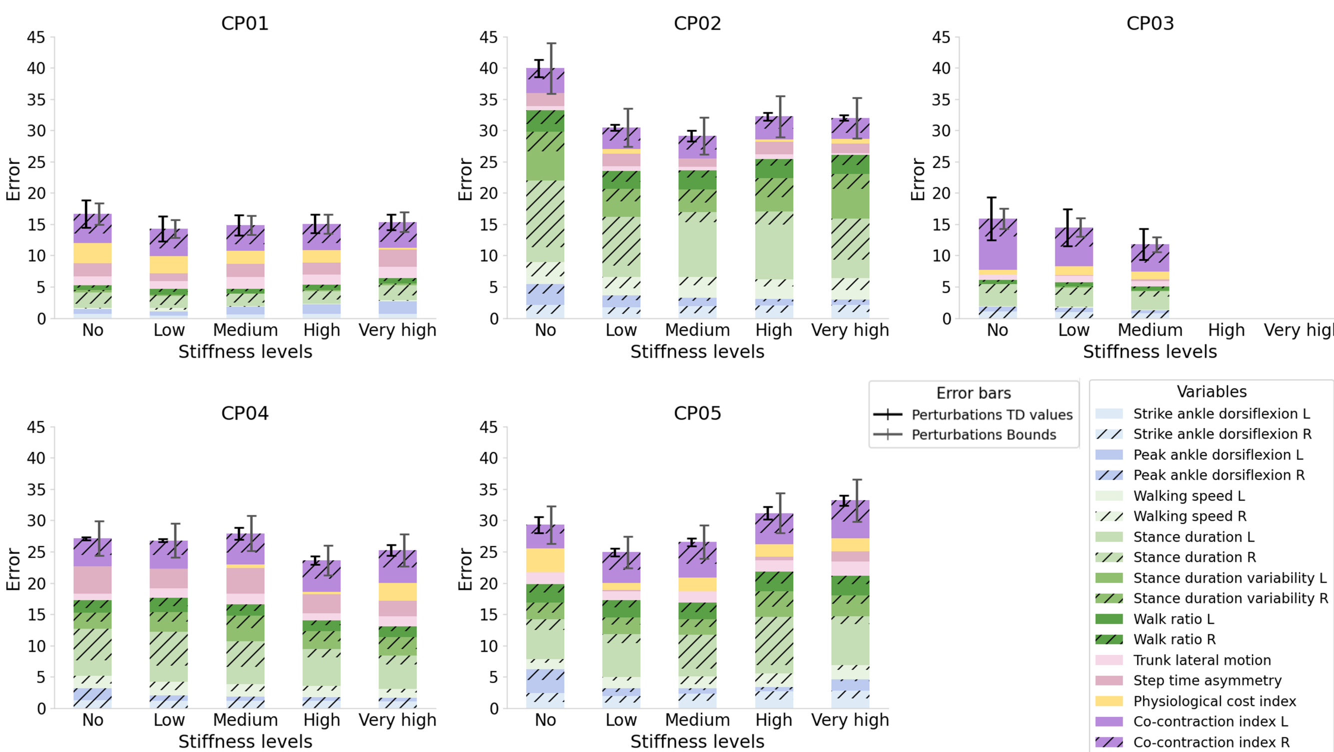


Fig. 9: Optimization including sensitivity analysis. Black error bars represent $\pm 10\%$ perturbations in TD values, and grey error bars represent $\pm 10\%$ perturbations in the range of the bounds.

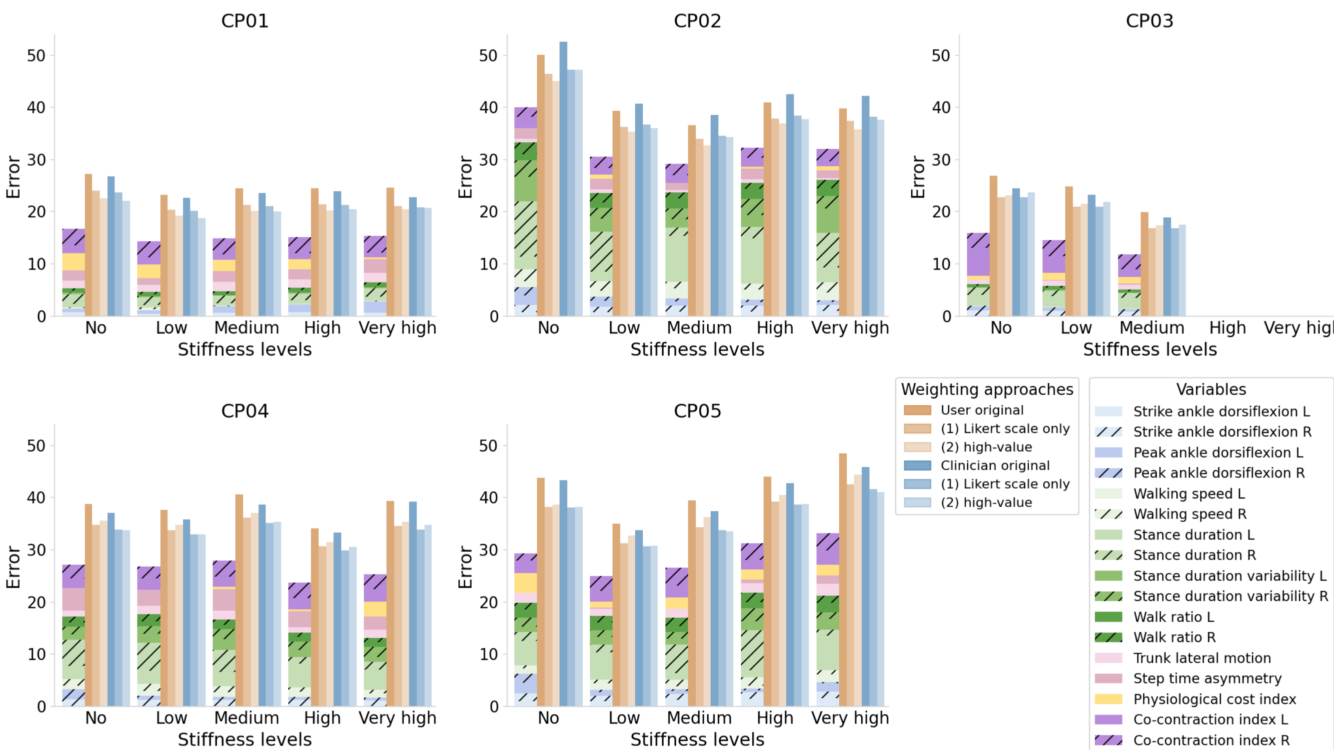


Fig. 10: Optimization incorporating user and clinician preferences for all participants. For each stiffness level, three grouped bars are shown (from left to right): the original optimization, optimization including user preferences, and optimization including clinician preferences. Shading within the user and clinician bars indicates different preference weighting approaches.