

Modeling the Kinematics of Human Locomotion Over Continuously Varying Speeds and Inclines

Kyle R. Embry^{ID}, Student Member, IEEE, Dario J. Villarreal, Member, IEEE,
Rebecca L. Macaluso, and Robert D. Gregg^{ID}, Senior Member, IEEE

Abstract—Powered knee and ankle prostheses can perform a limited number of discrete ambulation tasks. This is largely due to their control architecture, which uses a finite-state machine to select among a set of task-specific controllers. A non-switching controller that supports a continuum of tasks is expected to better facilitate normative biomechanics. This paper introduces a predictive model that represents gait kinematics as a continuous function of gait cycle percentage, speed, and incline. The basis model consists of two parts: basis functions that produce kinematic trajectories over the gait cycle and task functions that smoothly alter the weight of basis functions in response to task. Kinematic data from 10 able-bodied subjects walking at 27 combinations of speed and incline generate training and validation data for this data-driven model. Convex optimization accurately fits the model to experimental data. Automated model order reduction improves predictive abilities by capturing only the most important kinematic changes due to walking tasks. Constraints on a range of motion and jerk ensure the safety and comfort of the user. This model produces a smooth continuum of trajectories over task, an impossibility for finite-state control algorithms. Random sub-sampling validation indicates that basis modeling predicts untrained kinematics more accurately than linear interpolation.

Index Terms—Human locomotion, optimization, predictive models, prosthetic limbs, robot control.

I. INTRODUCTION

BIOLOGICAL biped locomotion is capable of a diverse array of tasks and feats of agility unmatched by modern

Manuscript received April 9, 2018; revised September 25, 2018; accepted October 29, 2018. Date of publication November 5, 2018; date of current version December 6, 2018. This work was supported in part by the National Institute of Child Health & Human Development of the NIH under Award DP2HD080349 and Award R01HD094772 and in part by the NSF under Award CMMI-1652514. Robert D. Gregg, IV, Ph.D., holds a Career Award at the Scientific Interface from the Burroughs Wellcome Fund. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or NSF. (*Corresponding author: Robert D. Gregg.*)

K. R. Embry is with the Department of Mechanical Engineering, University of Texas at Dallas, Richardson, TX 75080 USA (e-mail: kyle.embry@utdallas.edu).

D. J. Villarreal is with the Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75205 USA.

R. L. Macaluso is with the Department of Bioengineering, University of Texas at Dallas, Richardson, TX 75080 USA.

R. D. Gregg is with the Departments of Bioengineering and Mechanical Engineering, University of Texas at Dallas, Richardson, TX 75080 USA (e-mail: rgregg@ieee.org).

Digital Object Identifier 10.1109/TNSRE.2018.2879570

prosthetic legs. The goal of powered knee and ankle prostheses is to improve transfemoral amputee quality of life by restoring normative biomechanics for as many locomotion tasks as possible. One example is the Vanderbilt Generation 3 powered knee and ankle prosthesis. This device has been shown to allow amputees to walk with gait kinematics that are similar to the normal gait kinematics of subjects with healthy limbs, and is showing promising results in terms of metabolic performance and minimizing back muscle activation related to the risk of low back pain [1].

Modern powered knee and ankle prostheses require complex control algorithms to provide amputees with intuitive control over their prostheses. A common control strategy, borrowed from gait analysis [2], is to create separate controllers for discrete locomotion activities [3]–[6] and phases of the gait cycle [7]–[9]. A high level controller selects among these controllers using classifier algorithms, which generally depend on input from inertial measurement units (IMU) or electromyography (EMG) sensors [10]–[13]. This finite-state machine (FSM) paradigm requires many controllers to approximate the continuum of all tasks and phases. As the number of controllers increases, so does the time and expertise needed to tune the system to a subject [7]. Therefore a new method capable of supporting a wide range of tasks and gait phases with a single controller is expected to reduce tuning time and promote normative biomechanics.

A different approach to restoring normative biomechanics without the drawbacks inherent to the FSM method is using normative kinematic trajectories as a virtual constraint [14]. Virtual constraints enslave the progression of actuated joints to the progression of a mechanical signal called a phase variable [15]–[18]. Our lab has used healthy human subject kinematic trajectories as the desired trajectories of prosthetic joints and a phase variable derived from the motion of the intact hip joint. This method was shown to improve amputee subjects' walking performance at discrete speed and incline combinations from 0.67 to 1.21 m/s and −2.5 to 9.0 deg, while reducing the necessary number of tuning parameters and therefore the configuration time [19]. We have also shown promising results that tracking normative trajectories can reduce compensations associated with passive prostheses, including vaulting and hip circumduction [20]. However, this approach is still limited by the use of an FSM to switch between locomotion tasks.

Other researchers in the fields of biped robotics and prosthetics have demonstrated the potential of replacing FSM algorithms with continuous parameterizations. Holgate *et al.* [21] proposed a mapping from percent gait and stride length to desired actuator position for an ankle prosthesis. This model represents a continuum of phase and walking speed, as stride length is related to walking speed. A more recent approach for controlling powered ankle prostheses modeled a surface representing the time evolution of selected locomotion variables, and regressed that surface over a velocity range to create a manifold [22]. This method creates realistic gait at many speeds in simulation, and provided more power at higher speeds in experiments with an ankle prosthesis. Some biped robots depend on a finite set of optimal gaits, each designed for a specific task, and either define non-periodic transitional gaits to guide the robot back to a preprogrammed periodic gait [23] or interpolate the optimal kinematics for all other tasks in between [24]–[26]. This results in robust gait and the ability to handle a variety of terrains, but requires a large and well-structured set of optimal gaits to interpolate between. The required density of training data is infeasible for human subject experiments.

This paper introduces a method to unify phase and task via a reduced-order data-driven model which will be referred to as the basis model (BM). The joint kinematics over many combinations of speed and incline are parameterized as a function of a phase variable and a novel addition, task variables. This model uses normative data from healthy human subjects performing a variety of walking speeds and inclines to determine optimal model parameters. The goal is to predict kinematic gait patterns for a range of walking speeds and inclines given relatively little training data. Our contributions over prior work [27] include an automated method to select a minimal set of bases (Section II-D.3). This step helps to create a simpler model that provides greater model generalization by avoiding overfitting to noisy data [28]. Other new contributions are range of motion constraints on the output, (Section II-D.4), minimum jerk regularization of the kinematic surface (Section II-D.5), a comparison of this method to another interpolative modeling method, and a more thorough cross-validation process (Section III-B).

The data used to train this model was collected via the procedure given in Section II-A. The model format and training method are described in detail in Section II-C. The predictive accuracy of the basis method, compared to linear interpolation, is discussed in Section II-F, and the numerical results are shown in Section III. Finally, the significance of this work and future plans to utilize the basis model to generate reference trajectories for a transfemoral powered prosthetic leg are discussed in Section V.

II. METHODS

A. Experimental Protocol

The experimental protocol was approved by the Institutional Review Board at the University of Texas at Dallas (UTD). All 10 subjects (5 female) were able-bodied and provided written informed consent. The subject were mean age 23 years

($\sigma = 2.8$ years), mean height 170 cm ($\sigma = 8.2$ cm), and mean weight 64 kg ($\sigma = 7.7$ kg). This selection of subjects represents a population of young, active, and healthy walkers. This should produce the desired kinematics of otherwise healthy amputee community ambulators who can fully utilize the potential capabilities of an advanced powered knee and ankle prosthesis like the 2nd generation UTD LEG [29]. However, it should be noted that the emphasis of this paper is on the solution method, not the specific kinematic data. The only assumptions made about the training data kinematics is that they are periodic and continuously adapt to changes in speed and incline (within the bounds tested). Researchers with different patient populations of interest are encouraged to replicate this optimization method with different training data.

All tests consisted of subjects walking at a steady speed and incline on a Bertec instrumented split-belt treadmill for one minute. For each test the subject walked at a constant speed of 0.8 m/s, 1.0 m/s, or 1.2 m/s and a constant ground slope ranging from -10 degrees to +10 degrees at 2.5 degree increments. All subjects walked at every combination of speed and incline, resulting in 27 different tasks with unique identifiers, χ_j with $j = 1, 2, \dots, 27$. A 10-camera Vicon motion capture system recorded the subjects' kinematics at 100 Hz. The order of trials was randomized and subjects took breaks to prevent fatigue. A supplementary data set from these experiments is available for download [30].

B. Data Processing

Kinematic data were filtered using Vicon's Plug-In Gait pipeline [31] which fills gaps and smooths trajectories with a Woltring filter (smoothing factor of ten). The Dynamic Plug-in Gait Model was applied to calculate joint angles [31]. Positive joint angles represent hip flexion, knee flexion, and ankle dorsiflexion throughout this paper.

Kinematic data for each subject were separated into gait cycles and interpolated to contain 150 points per stride. The number of data points was chosen to slightly exceed the number of recorded frames for any stride. All gait cycles begin at heel strike (1) and end just before the next heel strike (150). Heel strike was defined as when the force plate measurement crossed 10% of the subject's weight, or as the time of maximal anterior position of the heel marker if the subject stepped on the wrong force plate. For each speed and incline combination, the intra-subject mean of each joint angle was calculated across the phase dimension using MATLAB. Outliers, defined as a trajectory that is more than three standard deviations away from the mean trajectory at any point in the stride, were removed. After outlier rejection, the inter-subject mean and standard deviation were calculated across the phase dimension.

C. Gait Model

The basis model represents joint kinematics as a function of phase and task. Phase is measured by a phase variable, $\phi \in \{\mathbb{R} | 0 \leq \phi < 1, \phi > 0\}$, which is a cyclic and monotonic scalar that increases from 0 to 1 once per stride. Each independent dimension of the ambulation task is measured by

a task variable, which are concatenated into a task vector χ . In our case, $\chi = (v, \alpha)$, where v is the subject's forward speed linearly mapped from a range of 0.6 m/s to 1.4 m/s to a range of 0 to 1, and α is the ground slope, linearly mapped from -10 degrees to 10 degrees to a range of 0 to 1. The methods in this paper can be expanded to include more dimensions of χ , like a sit/stand variable or a flat/stairs variable, but those transitions are out of the scope of this paper.

Gait kinematics are modeled as the weighted summation of N basis functions of phase, $b_k(\varphi)$. The weight of each basis function changes for each unique task, as determined by the task functions $c_k(\chi)$. This yields the following separable expression for the joint angle q of the hip, knee, or ankle:

$$q(\varphi, \chi) = \sum_{k=1}^N b_k(\varphi) c_k(\chi), \quad (1)$$

where the total number of basis functions is N , indexed by k .

The basis functions b_k are finite Fourier series of degree F :

$$b_k(\varphi) = \beta_{00k} + \sum_{i=1}^F (\beta_{1ik} \cos(i\varphi) + \beta_{2ik} \sin(i\varphi)), \quad (2)$$

where the $\beta_{tik} \in \mathbb{R}$ terms are coefficients that will be solved for with convex optimization. The t index denotes the type of function (constant, cosine, or sine) that the coefficient pairs with, i indexes the order of the terms, and the k index shows what basis function b_k the terms belong to. This general form is widely used in the interpolation of periodic functions. The order $F = 10$ is selected to ensure our model can match the significant frequency content of human kinematics, as in [32].

The scalar task functions c_k describe how the basis function's weight continuously changes with respect to changing tasks variables. We use Bernstein basis polynomials as task functions because of their frequent use to parameterize geometric shapes on finite intervals [33]. This also means that our resulting model is a Bezier curve, which will produce closed-form solutions for derivatives and guaranteed boundedness. All task functions have the format:

$$c_k(\chi) = \binom{\gamma}{\kappa} f(\chi)^\kappa (1 - f(\chi))^{\gamma - \kappa}, \quad (3)$$

where $\binom{\gamma}{\kappa}$ is a binomial coefficient. The $f(\chi)$, γ and κ terms will vary depending on whether a given task function is meant to model changes due to speed, incline, or constant components, as shown by the following sets:

$$K_0 = \{1\}, \quad K_1 = \{2, 3, 4\}, \quad K_2 = \{5, \dots, 8\}, \quad (4)$$

where all task functions c_k with $k \in K_0$ are constant functions, $k \in K_1$ are speed functions, and $k \in K_2$ are incline functions. There are three speed terms because we elected to use a 2nd order Bernstein function for speed, the highest order appropriate for only three recorded speeds. A third order Bernstein basis (four terms) was selected for inclines due to the relative complexity of kinematic changes due to incline. The other parameters of (3) take the following values:

$$\begin{aligned} k \in K_0 &\implies f(\chi) = 0, \quad \gamma = 0, \quad \kappa = 0, \\ k \in K_1 &\implies f(\chi) = v, \quad \gamma = 2, \quad \kappa = 0, 1, 2, \\ k \in K_2 &\implies f(\chi) = \alpha, \quad \gamma = 3, \quad \kappa = 0, 1, 2, 3. \end{aligned}$$

D. Objectives and Constraints of the Basis Model

Convex optimization is used to solve for the parameters β that determine the shape of the basis functions. The basis model has several objectives and constraints, and each will be discussed separately.

1) Fitting to Training Data: Our main goal is that the basis model (1) can approximate average human kinematics for a range of tasks. To do this we solve for β in b_k that satisfy:

$$\bar{\mathbf{d}}_{\varphi_i \chi_j} \approx \sum_{k=1}^N b_k(\varphi_i) c_k(\chi_j), \quad (5)$$

where $\bar{\mathbf{d}}_{\varphi_i \chi_j}$ is the inter-subject mean of recorded kinematics of a given joint at discrete phase point φ_i and discrete task vector χ_j .

To write a standard form convex optimization problem that will satisfy (5), we define the column vector x :

$$\begin{aligned} x_k &= [\beta_{00k} \ \beta_{11k} \ \beta_{21k} \ \dots \ \beta_{1Fk} \ \beta_{2Fk}]^T \in \mathbb{R}^{1+2F}, \\ x &= [x_1 \ x_2 \ \dots \ x_N]^T \in \mathbb{R}^{N(1+2F)}. \end{aligned}$$

A matrix A_i stores the phase dependent terms from (2):

$$\begin{aligned} a_i &= [1 \ \cos(\varphi_i) \ \sin(\varphi_i) \ \dots \ \cos(F\varphi_i) \ \sin(F\varphi_i)] \\ &\in \mathbb{R}^{1 \times (1+2F)}, \\ A_i &= \begin{bmatrix} a_i & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & a_i & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & a_i \end{bmatrix} \in \mathbb{R}^{N \times N(1+2F)}, \end{aligned}$$

which yields $A_i x_k = b_k(\varphi_i)$ from (2). Similarly the task functions from (3) are stored in a row vector:

$$y^j = [c_1(\chi_j) \ c_2(\chi_j) \ \dots \ c_N(\chi_j)] \in \mathbb{R}^{1 \times N},$$

which gives the equivalent statements:

$$y^j A_i x = \sum_{k=1}^N b_k(\varphi_i) c_k(\chi_j).$$

This matrix format makes it explicit that the basis model is linear with respect to x . Minimizing any norm of $\bar{\mathbf{d}}_{\varphi_i \chi_j} - y^j A_i x$ will constitute a convex objective function, guaranteeing the existence of a global minimum. However, one degree of error has different effects at different points in the gait cycle. To represent this, our objective function will divide the norm of $\bar{\mathbf{d}}_{\varphi_i \chi_j} - y^j A_i x$ by the standard error of the data at that phase point and task, $\text{SE}(\bar{\mathbf{d}}_{\varphi_i \chi_j})$. This follows the assumption that points in the gait cycle with high variance do not need to be followed as accurately.

2) Testing Against Validation Data: To avoid overfitting to the specific tasks recorded in our experiment, and to test the feasibility of recording fewer trials in future experiments, we employed a form of repeated random sub-sampling validation called Monte Carlo cross-validation [34]. Cross validation, in general, consists of breaking a data set into two groups, a training set and a validation set. The model parameters are solved in order to approximate the training set, and the model's performance is tested on the remaining data in the validation set. Monte Carlo cross-validation starts with the

data being randomly divided into training and validation sets in z unique ways. Each random instance of cross-validation will be called a test split. In our case, the model parameters are optimized with respect to the training data from tasks χ_j with $j \in H_h$, where H_h is a subset of the tasks recorded in experiments, and $h = 1, 2, \dots, z$ indexes through the unique randomized sets of all test splits. The performance of the model is tested against the validation set H_h^C .

A risk with data-driven models is the potential need for exorbitant amounts of data to produce accurate results. For example, if we added other task variables to our task vector (stairs, running, transitions between) it would become a combinatorial problem to measure a fine grid of every combination of tasks as was done for this experiment. Because of this, it is important that our model can be trained using sparse data in the task space. To test this ability, we created the random training sets H_h using Latin Hypercube Sampling (LHS), a common technique for the design of sparse experiments [35]. LHS discretizes the independent variables of an experiment into n bins, and chooses n points to sample such that each bin of each independent variable is sampled exactly once. To simulate LHS for our current experiment, we take nine unique samples from the available 27 as training data for each test split, such that each incline is represented once and each speed is represented exactly three times. We will generate $z = 250$ unique test splits with the LHS method to test if our modeling method can make predictions from sparse data.

3) Automated Model Order Reduction: The basis model has a total of $N = 8$ basis functions and task functions (4). It would be advantageous to reduce this number if possible. Low rank models which select only the most important model features to explain the response are more likely to provide predictive accuracy by avoiding fitting to model noise, and are simpler to interpret [36]. To facilitate the use of sparsity inducing norms to create lower order models, we introduce a selector function $g(K)$ which returns a vector similar to y^j where only task functions with $k \in K$ are nonzero:

$$g_k(K) = \begin{cases} c_k(\chi_j) & \text{if } k \in K, \\ 0 & \text{otherwise,} \end{cases} \quad \forall k = 1, 2, \dots, N. \quad (6)$$

It is important to note that making $c_k = 0$ also eliminates the contribution of b_k towards the full model (1). During the optimization process, it is safe to assume $b_k = 0$ if $k \notin K$. The g function is used to calculate the constant terms of lower-order basis models, denoted as:

$$\Lambda_{ij}^K = g(K)A_i \in \mathbb{R}^{N(1+2F)},$$

which leads to the lower order model:

$$\Lambda_{ij}^K x = \sum_{k \in K} b_k(\varphi_i) c_k(\chi_j). \quad (7)$$

We will use convex, sparsity-inducing norms on $\Lambda_{ij}^K x$ to reduce the model and use only the most important bases, a set that is defined as $K^* \subseteq \{K_0 \cup K_1 \cup K_2\}$.

4) Range of Motion Limits: The basis model will calculate desired joint trajectories for a prosthetic leg, so bounding the model's output is a matter of safety for the user and machine. The magnitude of each task function is known before solving for β , and we can rely on this fact to bound the range of motion of each type of task variable (constant, speed, or incline) over a grid of input values indexed by m :

$$R_i^{\max} = \max_m \Lambda_{im}^1 x + \max_m \sum_{k \in K_1^*} \Lambda_{im}^k x + \max_m \sum_{k \in K_2^*} \Lambda_{im}^k x,$$

$$R_i^{\min} = \min_m \Lambda_{im}^1 x + \min_m \sum_{k \in K_1^*} \Lambda_{im}^k x + \min_m \sum_{k \in K_2^*} \Lambda_{im}^k x,$$

$$\forall i = 1, 2, \dots, 150 \text{ and } \forall m = 1, 2, \dots, 100,$$

where $K_i^* = K_i \cap K^*$ is the set of each type of task function that was determined to be significant, and i indexes over all discrete points of phase. The grid of checkpoints are combinations of 100 speeds and 100 inclines uniformly spanning the domain. The goal is to build a dense enough grid that it is unlikely for the function to leave the range of motion bounds between grid points. R_i^{\max} gives the maximum position of the model at discrete phase φ_i while R_i^{\min} gives the minimum. Both can be constrained over the range of i to limit the range of motion at all phase points.

5) Jerk Minimization: It is important that the output of the basis model resembles natural human motion, and minimizing the jerk of the kinematic trajectories is an effective way to do so [37], [38]. Jerk-minimization will smooth the model output, and is very simple to implement as a convex objective. The jerk at each point in the model can be written as:

$$J_{ij}^K = \frac{\partial^3}{\partial \varphi^3} \Lambda_{ij}^K,$$

$$\text{vec}(J^K x) = [J_{11}^K x \dots J_{L1}^K x \dots J_{1M}^K x \dots J_{LM}^K x].$$

Minimizing a norm of $\text{vec}(J^K x)$ will reduce joint jerk.

E. Solving for Optimal Model Parameters

Our optimization is a two-step process. Step 1 uses all basis functions and reduces the model dimensionality using group L_1 regularization [39], as follows:

$$\begin{aligned} & \underset{x, \rho}{\text{minimize}} \rho + \lambda \Omega(x), \\ & \text{such that } -\rho \text{SE}(\mathbf{d}_{\varphi_i \chi_j}) \leq \bar{\mathbf{d}}_{\varphi_i \chi_j} - y^j A_i x \leq \rho \text{SE}(\mathbf{d}_{\varphi_i \chi_j}), \\ & \Omega(x) := \sum_{k=1}^N \max_{i,j} |\Lambda_{ij}^k x|, \\ & \forall i = 1, 2, \dots, 150 \text{ and } \forall j = 1, 2, \dots, 27, \end{aligned} \quad (8)$$

where ρ is a bound on the absolute error between the experimental data $\bar{\mathbf{d}}_{\varphi_i \chi_j}$ and the model's prediction $y^j A_i x$, divided by $\text{SE}(\mathbf{d}_{\varphi_i \chi_j})$. The function Ω penalizes the number of nonzero basis functions used. The L_∞ norm (max of absolute values) is taken of each low order model $\Lambda_{ij}^k x$ to determine its size and the L_1 norm (summation of strictly positive elements inside the Ω function) is taken across these groups. Minimizing the L_1 norm encourages sparsity in the residual [40]. In our case, this will encourage a sparse set of basis functions/task

functions. The scalar λ is a weighting term, which weights the relative importance of our two objectives: fitting the data accurately (ρ), and using a minimal set of bases (Ω).

The optimization problem (8) has a guaranteed global minimum because it is a convex optimization problem. Each element of $\bar{\mathbf{d}}_{\varphi_i \chi_j} - y^j A_i x$ is affine with respect to vector x . This means the upper bound $\rho \text{SE}(\mathbf{d}_{\varphi_i \chi_j})$ and lower bound $-\rho \text{SE}(\mathbf{d}_{\varphi_i \chi_j})$ form affine constraints. The regularizing function Ω is convex, because $\Lambda_{ij}^k x$ is affine with respect to x , the L_∞ norm of an affine expression is always convex, and the sum of convex functions is convex [40]. Similarly, the objective function $\rho + \gamma \Omega(x)$ is the sum of an affine expression and a convex function, making it a convex function. As (8) has a convex objective and affine constraints, it is a standard convex optimization problem.

To determine the desired order of the model, we perform a Principle Component Analysis (PCA) of the complete data set. We define the desired order of the data to be the number of eigenvalues greater than three (a threshold that separated the very large and very small eigenvalues) plus one to account for a constant offset (PCA is based on zero mean data, in contrast to our method). This results in a desired order of four for the hip and five for the knee and ankle. We then iteratively weight λ with the bisection method until only the desired number of terms are sufficiently large (> 1).

The second optimization step is smoothly fitting the model to data, using only the basis functions deemed significant by the previous step, denoted K^* . This problem is formulated as follows:

$$\begin{aligned} & \underset{x}{\text{minimize}} \quad \rho + \delta \|\text{vec}(J^{K^*} x)\|_2, \\ & \text{such that } -\rho \text{SE}(\mathbf{d}_{\varphi_i \chi_j}) \leq \bar{\mathbf{d}}_{\varphi_i \chi_j} - \Lambda_{ij}^{K^*} x \leq \rho \text{SE}(\mathbf{d}_{\varphi_i \chi_j}), \\ & R_1 < R_i^{\min}, \\ & R_2 > R_i^{\max}, \\ & \forall i = 1, 2, \dots, 150 \quad \text{and } \forall j \in H_h, \end{aligned} \quad (9)$$

where the term δ is the relative weight between our two objectives, fitting the data accurately (ρ) and smoothing the model ($\|\text{vec}(J^{K^*} x)\|_2$). After some iterations we settled on a value of $\delta = 1e-5$, but the solution is not highly sensitive to differences in this value. The lower and upper range of motion bounds of each joint, R_1 and R_2 respectively, are based on values from [41].

The convexity analysis of (9) is similar to (8). We drop the basis regularization objective in favor of a smoothing regularization term $\|\text{vec}(J^{K^*} x)\|_2$. The L_2 norm of linear $J^{K^*} x$ is convex, so the objective function $\rho + \delta \|\text{vec}(J^{K^*} x)\|_2$ is convex. The minimization across multiple linear functions is convex, and the positive sum of convex functions is convex, so the lower bound constraint $R_1 < R_i^{\min}$ is valid. Similarly, the maximization of multiple linear functions is concave, and the positive sum of concave functions is concave, making the upper bound $R_2 > R_i^{\max}$ likewise a valid constraint.

F. Determining Predictive Accuracy

It is important to test if the basis model is capable of predicting kinematics for completely new tasks. To quantify

predictive accuracy, we calculate the error between the predicted kinematic value, $q(\varphi_i, \chi_j)$, and the real value found via experiments, $\bar{\mathbf{d}}_{\varphi_i \chi_j}$, defined as:

$$G_{hj} = \max_i \left\{ \frac{\bar{\mathbf{d}}_{\varphi_i \chi_j} - q(\varphi_i, \chi_j)}{\text{SE}(\mathbf{d}_{\varphi_i \chi_j})} \right\}, \quad (10)$$

$$\forall i = 1, 2, \dots, 150, \quad \forall j \in H_h, \quad \text{and } \forall h = 1, 2, \dots, z.$$

where taking the maximum over i gives the value of the largest error encountered in a gait cycle. This evaluation is repeated for all validation test splits $j \in H_h$ and stored in a vector of mean errors, e_μ , and a vector of maximum errors e_m :

$$(e_\mu)_h = \underset{j}{\text{mean}} G_{hj}, \quad (e_m)_h = \underset{j}{\max} G_{hj}.$$

We will compare the mean, max, and distribution of e_μ and e_m between the two methods compared in this paper: basis modeling (BM) as described in (1) and linear interpolation (LN) of kinematics with respect to phase, speed, and incline. Due to the extreme inaccuracy of linear extrapolation of gaits, our linear interpolation model uses linear interpolation for task vectors in the interior of the convex hull of the training data, and an FSM for task vectors on the boundary or outside the convex hull. The FSM uses the nearest training data to predict extrapolated data, as defined by Delaunay Triangulation [42].

The distributions of each test split vector, e_μ and e_m , are compared between the basis modeling method and linear interpolation method using the Wilcoxon signed-rank test. A non-parametric test was used because the Anderson-Darling test rejected the null hypotheses that e_μ or e_m came from normal distributions [43].

III. RESULTS

A. Parameterizing Experimental Data

In this section, we will model hip, knee, and ankle kinematics and evaluate the ability of the basis model to predict untrained task kinematics. To solve this optimization problem, we used CVX, a package for specifying and solving convex programs [44], [45]. Percent gait acts as the phase variable of (1), which will be defined:

$$\varphi = \frac{t}{T}, \quad \varphi_i = \frac{t_i}{T},$$

where t and t_i are continuous and discrete time, respectively, from the beginning of a gait cycle, and T is the total time of a gait cycle.

After order reduction, the most important task functions are plotted in Fig. 1. The order of the legend gives the relative importance of each function. Each of these task functions are paired with a basis function, see Fig. 2.

B. Evaluating the Model of Experimental Data

Next, we compare the error in our prediction of all untrained tasks to the errors that would result from the use of linear interpolation. Table I shows the mean of the mean summary statistic vector (\bar{e}_μ) for all test splits, and the max of the max summary statistic vector ($\vee e_m$) for all test splits for both the basis modeling method (1) and linear interpolation.

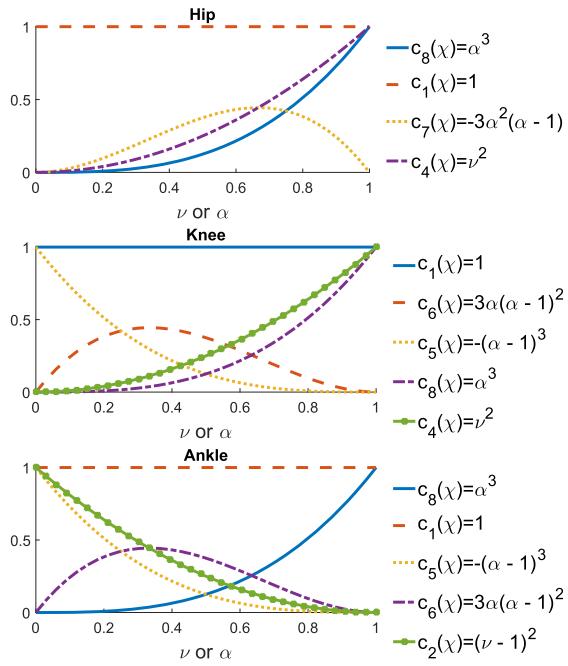


Fig. 1. The most important task functions for each joint, determined by group L_1 regularization. The legend gives the order of importance.

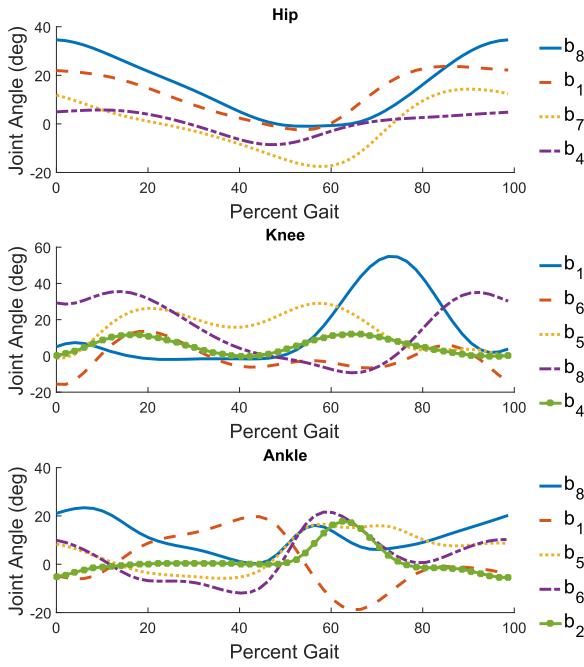


Fig. 2. The matching basis functions for each task function in **Fig. 1**.

Table I shows that the basis model (BM) and the linear interpolation (LN) method have similar hip error means, but the basis model has lower knee and ankle fitting in terms of \bar{e}_μ and $\vee e_m$. A one-sided Wilcoxon signed-rank test shows the e_μ and e_m distributions of test split results have lower medians for the basis model on all three joints with the p -values given in the third row of **Table I**. This shows that the basis model is repeatedly outperforming linear interpolation in terms of accurately predicting hip, knee, and ankle kinematics.

TABLE I
SUMMARY STATISTIC DISTRIBUTION MEANS, MAXIMUMS, AND
WILCOXON SIGNED RANK RESULTS AMONG TEST
SPLITS FOR ALL JOINTS

	Hip		Knee		Ankle	
	\bar{e}_μ	$\vee e_m$	\bar{e}_μ	$\vee e_m$	\bar{e}_μ	$\vee e_m$
BM	0.492	2.10	0.881	3.41	0.856	2.62
LN	0.518	1.59	1.09	5.83	0.994	3.14
p	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$	$\ll 0.05$

p is the p-value test result of a one sided Wilcoxon signed rank test that the median of a given summary statistic distribution resulting from the BM method is less than the median of the same distribution resulting from the LN method.

Fig. 3 and **Fig. 4** show example kinematic surfaces made from the basis method or linear interpolation, respectively. Both use the same random test split for training. In both figures, model outputs are represented as a blue-green surface, and the data used to train the model is shown in dashed blue lines. Solid red lines show the untrained data, which is what was used to calculate validation error. Note that the basis model is C^∞ smooth, while the linear interpolation surface has derivative discontinuities at training points and when extrapolating using an FSM.

IV. DISCUSSION

A. Advantages of Basis Modeling

Any powered prosthesis with a position tracking controller will require reference trajectories to follow. There are two basic data-driven approaches to providing those trajectories: an FSM that provides one trajectory for each supported task, or continuous prediction of gait based on the proximity of the current task to known data. The latter has several advantages that we believe are beneficial for emulating natural gait with a prosthesis. Continuous prediction provides a continuum of possible gaits, providing the user with fine control of their gait. Previous work showed continuous prediction over the gait cycle reduces configuration time of a powered leg to individual users [19], and we expect doing the same for both phase and task will have even greater benefits. We have also shown that continuous prediction of gait is more accurate than an FSM at predicting untrained kinematics, even when the FSM is emulated with perfect classification accuracy [27].

There are many options for how to construct a continuous model of human kinematics. The simplest conceptually is linear interpolation, and this has been used in some cases for biped robot gait [46]. The basis modeling method has several advantages over linear interpolation. Through convex optimization, we can force the basis model to fit experimental data while simultaneously minimizing jerk and constraining range of motion. Jerk-minimized kinematic trajectories have been shown to closely mimic natural human movements [37], [38]. The form of the basis model is also C^∞ smooth, regardless of solved parameters, see **Fig. 3**. The basis model also benefits from automated order reduction and a simple closed form expression. Our statistical analysis has shown that basis modeling more accurately predicts untrained kinematics given

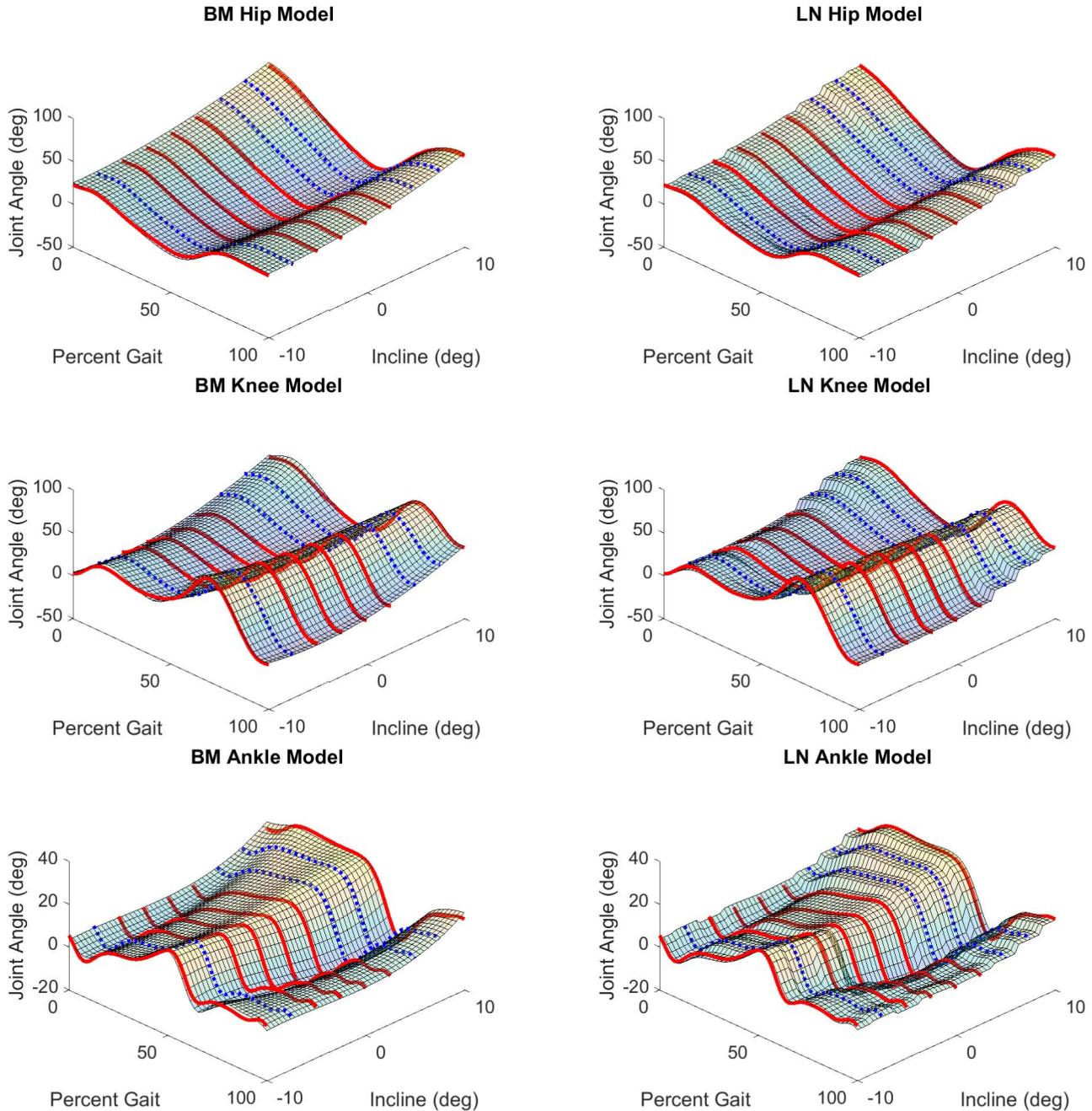


Fig. 3. Basis model joint surfaces generated for a random test split. The blue dashed mean kinematic trajectories were used for training, solid red is for validation. The true surface is four dimensional, so this is the 3D projection onto one speed, 1.0 m/s.

sparse irregular training data. Performance on sparse data is important because it allows us to design future experiments that will make better models with fewer gait experiments. This aspect will be vital when recording data on impaired populations that cannot be expected to participate in lengthy gait trials. Irregular data is also important as it allows us to utilize experimental design techniques like LHS, or possibly to combine several publicly available data sets of kinematics sampled on different (or irregular) intervals. This second possibility is of interest but is outside the scope of this paper.

Fig. 4. Linear interpolation model joint surfaces generated for a the same test split as Fig. 3, projected onto one speed, 1.0 m/s. The dashed blue mean kinematic trajectories were used for training, solid red is for validation.

B. Interpretation of Trends

Basis and task functions lend themselves well to an intuitive understanding of how task affects gait. The basis function $b_k(\phi)$ shows what kinematic trend will become more apparent as $c_k(\chi)$ increases. For example, take the ankle basis $b_8(\phi)$ in Fig. 2, which contributes additional dorsiflexion around heel strike. As $c_8(\chi) = \alpha^3$ increases (*i.e.*, as the incline of walking increases from -10 to $+10$ degrees), the contribution of $b_8(\phi)$ will increase. Keeping these two terms in mind, it can be deduced that as incline increases, so will

ankle dorsiflexion at heel strike. The intuitive behavior of $b_8(\phi)c_8(\chi)$ is a well documented feature of inclined walking, also shown in [47, Fig. 2]. A similar example can be seen on the knee joint $b_5(\phi)$, which will shrink as incline increases because it is paired with knee $c_5(\chi) = -(\alpha - 1)^3$, see Fig. 1 and 2. This would imply that the knee angle during late stance will be shallower for high inclines, matching again a discrete trend highlighted in [47, Fig. 2]. Part of the power of basis modeling is how these well-established trends emerge naturally from the solution process. The researcher does not have to explicitly define what trends in gait data are important and add terms intended to capture those phenomena.

V. CONCLUSION

This paper develops a method to model human locomotion as a function of phase, speed, and incline. This model will be used as the reference trajectory generator of a virtually constrained powered knee and ankle prosthesis to achieve a continuum of possible gaits over a range of tasks. Various convex optimization techniques are used to improve the accuracy of the model and deliver several useful properties. Model parameters are optimized to match a set of able-bodied human kinematic data recorded at UTD. Structured sparsity inducing optimization methods ensure that the order of the model is minimized, which helps the model to match important trends in the data while maintaining simplicity. Kinematic jerk minimization helps the model output to mimic natural human motion, and preserves the operational life of actuators. Range of motion constraints respect anatomical constraints and will promote safety in future prosthetics applications.

The most important conclusion of this paper is that it is possible to accurately parameterize the human gait cycle as a continuous function of only phase, speed, and incline. This function is represented well by a separable model with basis functions that parameterize kinematics as a function of phase, and task functions that change the relative weight of basis functions depending on the current task. The basis modeling technique was shown to be more accurate than linear interpolation on sparse training data by Monte Carlo cross validation. This paper has shown that, when using the basis modeling method, very fine grids of task data (as collected for this paper) are unnecessary for producing satisfactory interpolation accuracy. This is an important conclusion, because when scaling the model up to a larger task space (*e.g.*, stair climbing) it will be infeasible to sample on the fine grid necessary for accurate linear interpolation. The basis model also provides direct intuition into how human gait is affected by changing tasks.

The next step in this research is developing sensing algorithms for powered prostheses to measure task in real time. This will rely heavily on existing techniques for IMU estimation of walking speed [48]–[50] and incline [51]–[55]. Given measurable tasks and phase variables [56], the following step is to implement this algorithm on the second-generation UTD leg [29].

ACKNOWLEDGMENT

The authors would like to thank Dr. Stephen Boyd for enlightening conversations about the potential of convex

optimization to solve for gait basis functions, and Sanyukta Bihari for her contributions to the collection of this experimental data.

REFERENCES

- [1] C. Jayaraman *et al.*, “Impact of powered knee-ankle prosthesis on low back muscle mechanics in transfemoral amputees: A case series,” *Frontiers Neurosci.*, vol. 12, p. 134, Mar. 2018.
- [2] J. Perry, J. Burnfield, and L. Cabico, *Gait Analysis: Normal and Pathological Function*, 2nd ed. Thorofare, NJ, USA: Slack, 2010.
- [3] A. J. Young, A. M. Simon, N. P. Fey, and L. J. Hargrove, “Intent recognition in a powered lower limb prosthesis using time history information,” *Ann. Biomed. Eng.*, vol. 42, no. 3, pp. 631–641, 2014.
- [4] R. Jiménez-Fabián and O. Verlinden, “Review of control algorithms for robotic ankle systems in lower-limb orthoses, prostheses, and exoskeletons,” *Med. Eng. Phys.*, vol. 34, no. 4, pp. 397–408, May 2012.
- [5] M. R. Tucker *et al.*, “Control strategies for active lower extremity prosthetics and orthotics: A review,” *J. Neuroeng. Rehabil.*, vol. 12, no. 1, p. 1, 2015.
- [6] B. E. Lawson, J. Mitchell, D. Truex, A. Shultz, E. Ledoux, and M. Goldfarb, “A robotic leg prosthesis: Design, control, and implementation,” *IEEE Robot. Autom. Mag.*, vol. 21, no. 4, pp. 70–81, Dec. 2014.
- [7] A. M. Simon *et al.*, “Configuring a powered knee and ankle prosthesis for transfemoral amputees within five specific ambulation modes,” *PLoS ONE*, vol. 9, no. 6, p. e99387, 2014.
- [8] M. F. Eilenberg, H. Geyer, and H. Herr, “Control of a powered ankle-foot prosthesis based on a neuromuscular model,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 2, pp. 164–173, Apr. 2010.
- [9] A. H. Shultz, B. E. Lawson, and M. Goldfarb, “Running with a powered knee and ankle prosthesis,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 3, pp. 403–412, May 2015.
- [10] H. L. Bartlett and M. Goldfarb, “A phase variable approach for IMU-based locomotion activity recognition,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 6, pp. 1330–1338, Jun. 2017.
- [11] R. Stolyarov, G. Burnett, and H. Herr, “Translational motion tracking of leg joints for enhanced prediction of walking tasks,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 4, pp. 763–769, Apr. 2017.
- [12] H. Huang, F. Zhang, L. J. Hargrove, Z. Dou, D. R. Rogers, and K. B. Englehart, “Continuous locomotion-mode identification for prosthetic legs based on neuromuscular-mechanical fusion,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2867–2875, Oct. 2011.
- [13] A. J. Young, T. A. Kuiken, and L. J. Hargrove, “Analysis of using EMG and mechanical sensors to enhance intent recognition in powered lower limb prostheses,” *J. Neural Eng.*, vol. 11, no. 5, p. 056021, 2014.
- [14] R. D. Gregg, T. Lenzi, L. J. Hargrove, and J. W. Sensinger, “Virtual constraint control of a powered prosthetic leg: From simulation to experiments with transfemoral amputees,” *IEEE Trans. Robot.*, vol. 30, no. 6, pp. 1455–1471, Dec. 2014.
- [15] E. R. Westervelt, J. W. Grizzle, C. Chevallereau, J. H. Choi, and B. Morris, *Feedback Control of Dynamic Bipedal Robot Locomotion*. New York, NY, USA: CRC Press, 2007.
- [16] K. Sreenath, H.-W. Park, I. Pouliquen, and J. W. Grizzle, “A compliant hybrid zero dynamics controller for stable, efficient and fast bipedal walking on MABEL,” *Int. J. Robot. Res.*, vol. 30, no. 9, pp. 1170–1193, Sep. 2010.
- [17] A. Ramezani, J. W. Hurst, K. A. Hamed, and J. W. Grizzle, “Performance analysis and feedback control of ATRIAS, a three-dimensional bipedal robot,” *ASME J. Dyn. Syst., Meas., Control*, vol. 136, no. 2, p. 021012, Dec. 2013.
- [18] M. J. Powell, H. Zhao, and A. D. Ames, “Motion primitives for human-inspired bipedal robotic locomotion: Walking and stair climbing,” in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2012, pp. 543–549.
- [19] D. Quintero, D. J. Villarreal, D. J. Lambert, S. Kapp, and R. D. Gregg, “Continuous-phase control of a powered knee-ankle prosthesis: Amputee experiments across speeds and inclines,” *IEEE Trans. Robot.*, vol. 34, no. 3, pp. 686–701, Jun. 2018.
- [20] S. Rezazadeh, D. Quintero, N. Divekar, E. Reznick, L. Gray, and R. D. Gregg, (Nov. 2018). “A phase variable approach for improved volitional and rhythmic control of a powered knee-ankle prosthesis.” [Online]. Available: <https://arxiv.org/abs/1811.05414> and <http://adsabs.harvard.edu/abs/2018arXiv181105414R>
- [21] M. A. Holgate, T. G. Sugar, and A. W. Böhler, “A novel control algorithm for wearable robotics using phase plane invariants,” in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2009, pp. 3845–3850.
- [22] N. Dhir, H. Dallali, E. M. Ficanha, G. A. Ribeiro, and M. Rastgaard, “Locomotion envelopes for adaptive control of powered ankle prostheses,” in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2018, pp. 1488–1495.

- [23] C. Liu, C. G. Atkeson, and J. Su, "Biped walking control using a trajectory library," *Robotica*, vol. 31, no. 2, pp. 311–322, 2013.
- [24] R. Hartley, X. Da, and J. W. Grizzle, "Stabilization of 3D underactuated biped robots: Using posture adjustment and gait libraries to reject velocity disturbances," in *Proc. IEEE Conf. Control Technol. Appl.*, Aug. 2017, pp. 1262–1269.
- [25] X. Da and J. Grizzle. (Nov. 2017). "Combining trajectory optimization, supervised machine learning, and model structure for mitigating the curse of dimensionality in the control of bipedal robots." [Online]. Available: <https://arxiv.org/abs/1711.02223>
- [26] Q. Nguyen *et al.*, "Dynamic walking on randomly-varying discrete terrain with one-step preview," in *Robotics: Science and Systems*. Cambridge, MA, USA: RSS Foundation, 2017. [Online]. Available: <http://www.roboticsproceedings.org/>
- [27] K. R. Embry, D. J. Villarreal, and R. D. Gregg, "A unified parameterization of human gait across ambulation modes," in *Proc. IEEE Int. Conf. Eng. Med. Biol. Soc.*, Aug. 2016, pp. 2179–2183.
- [28] M. L. Bermingham *et al.*, "Application of high-dimensional feature selection: Evaluation for genomic prediction in man," *Sci. Rep.*, vol. 5, May 2015, Art. no. 10312.
- [29] T. Elery, S. Rezazadeh, C. Nesler, J. Doan, H. Zhu, and R. Gregg, "Design and benchtop validation of a powered knee-ankle prosthesis with high-torque, low-impedance actuators," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2018, pp. 2788–2795.
- [30] K. Embry, D. Villarreal, R. Macaluso, and R. Gregg, "The effect of walking incline and speed on human leg kinematics, kinetics, and EMG," *IEEE Dataport*, 2018, doi: [10.21227/gk32-e868](https://doi.org/10.21227/gk32-e868).
- [31] Vicon Motion Systems Ltd. (2018). *Vicon Nexus Reference Guide*. [Online]. Available: <https://docs.vicon.com/display/Nexus26/Vicon+Nexus+Reference+Guide>
- [32] D. Quintero, A. E. Martin, and R. D. Gregg, "Toward unified control of a powered prosthetic leg: A simulation study," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 1, pp. 305–312, Jan. 2018.
- [33] R. T. Farouki, "The Bernstein polynomial basis: A centennial retrospective," *Comput. Aided Geometric Des.*, vol. 29, no. 6, pp. 379–419, 2012.
- [34] M. Kuhn and K. Johnson, *Applied Predictive Modeling*, vol. 26. New York, NY, USA: Springer, 2013.
- [35] R. L. Iman, *Encyclopedia of Quantitative Risk Analysis and Assessment*. Hoboken, NJ, USA: Wiley, 2008.
- [36] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*, vol. 112. New York, NY, USA: Springer, 2013.
- [37] T. Lenzi, L. Hargrove, and J. Sensinger, "Speed-adaptation mechanism: Robotic prostheses can actively regulate joint torque," *IEEE Robot. Autom. Mag.*, vol. 21, no. 4, pp. 94–107, Dec. 2014.
- [38] T. Flash and N. Hogan, "The coordination of arm movements: An experimentally confirmed mathematical model," *J. Neurosci.*, vol. 5, no. 7, pp. 1688–1703, Jul. 1985.
- [39] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Structured sparsity through convex optimization," *Statist. Sci.*, vol. 27, no. 4, pp. 450–468, 2012.
- [40] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.
- [41] *Range of Joint Motion Evaluation Chart*, Dept. Social Health Services, Washington, DC, USA, 2014.
- [42] MathWorks. (2017). *ScatteredInterpolant*. [Online]. Available: <https://www.mathworks.com/help/matlab/ref/scatteredinterpolant.html>
- [43] E. Whitley and J. Ball, "Statistics review 6: Nonparametric methods," *Crit. Care*, vol. 6, no. 6, p. 509, 2002.
- [44] M. Grant and S. Boyd. (Mar. 2014). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1*. [Online]. Available: <http://cvxr.com/cvx>
- [45] M. C. Grant and S. P. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control (Lecture Notes in Control and Information Sciences)*, V. D. Blondel, S. P. Boyd, and H. Kimura, Eds. London, U.K.: Springer-Verlag, 2008, pp. 95–110. [Online]. Available: http://stanford.edu/~boyd/graph_dcp.html
- [46] X. Da, O. Harib, R. Hartley, B. Griffin, and J. W. Grizzle, "From 2D design of underactuated bipedal gaits to 3D implementation: Walking with speed tracking," *IEEE Access*, vol. 4, pp. 3469–3478, 2016.
- [47] A. Leroux, J. Fung, and H. Barbeau, "Postural adaptation to walking on inclined surfaces: I. Normal strategies," *Gait Posture*, vol. 15, no. 1, pp. 64–74, 2002.
- [48] S. Yang and Q. Li, "Inertial sensor-based methods in walking speed estimation: A systematic review," *Sensors*, vol. 12, no. 5, pp. 6102–6116, 2012.
- [49] J. Rueterborries, E. G. Spaich, B. Larsen, and O. K. Andersen, "Methods for gait event detection and analysis in ambulatory systems," *Med. Eng. Phys.*, vol. 32, no. 6, pp. 545–552, Jul. 2010.
- [50] Q. Li, M. Young, V. Naing, and J. M. Donelan, "Walking speed estimation using a shank-mounted inertial measurement unit," *J. Biomech.*, vol. 43, no. 8, pp. 1640–1643, 2010.
- [51] B. Chen and Q. Wang, "Combining human volitional control with intrinsic controller on robotic prosthesis: A case study on adaptive slope walking," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 4777–4780.
- [52] A. M. Sabatini, C. Martelloni, S. Scapellato, and F. Cavallo, "Assessment of walking features from foot inertial sensing," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 3, pp. 486–494, Mar. 2005.
- [53] W. Svensson and U. Holmberg, "An autonomous control system for a prosthetic foot ankle," in *Proc. Int. Fed. Autom. Control, 4th IFAC Symp. Mech. Syst. (IFAC)*, 2006, pp. 856–861.
- [54] F. Sup, H. A. Varol, and M. Goldfarb, "Upslope walking with a powered knee and ankle prosthesis: Initial results with an amputee subject," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 1, pp. 71–78, Feb. 2011.
- [55] H. Rehbinder and X. Hu, "Drift-free attitude estimation for accelerated rigid bodies," *Automatica*, vol. 40, no. 4, pp. 653–659, 2004.
- [56] D. J. Villarreal, H. A. Poonawala, and R. D. Gregg, "A robust parameterization of human gait patterns across phase-shifting perturbations," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 3, pp. 265–278, Mar. 2017.