UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

DATA-DRIVEN GAIT ANALYSIS USING DYNAMIC MODE DECOMPOSITION FOR MONITORING CHEMOTHERAPY-INDUCED PERIPHERAL NEUROPATHY

A DISSERTATION
SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of
Doctor of Philosophy

By KANGJUN SEO

Norman, Oklahoma

2025

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/341604

DATA-DRIVEN GAIT ANALYSIS USING DYNAMIC MODE DECOMPOSITION FOR MONITORING CHEMOTHERAPY-INDUCED PERIPHERAL NEUROPATHY

A DISSERTATION APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr. Chad B. Roller, Chair

Dr. Hazem H. Refai, Co-Chair

Dr. Samuel Cheng

Dr. Choon Yik Tang

Dr. Curt Adams

© Copyright by KANGJUN SEO 2025
All Rights Reserved.

Acknowledgments

Portions of this dissertation are based on work completed in collaboration with Dr. Hazem H. Refai and Dr. Elizabeth S. Hile. I was responsible for the design of the study, data analysis, and manuscript preparation. Dr. Refai and Dr. Hile contributed through supervision and editorial feedback. In addition, Dr. Hile provided the dataset that served as the foundation for this work.

Table of Contents

Acknowledgments **iv**

List Of Tables **vii**

List Of Figures **viii**

Abstract **xii**

1 Introduction **1**

2 Literature Review **6**
2.1 Conventional Gait Metrics and Variability Analysis 6
2.2 Wearable and Sensor-Based Gait Analysis 7
2.3 Dynamic Mode Decomposition in Gait Studies 7
2.4 Baseline Gait Signatures with DMD 8
2.5 Novelty of the Present Approach 8

3 Data-Driven Modeling of Dynamical Systems **9**
3.1 Dynamical Systems 9
3.1.1 Classifications of Dynamical Systems 11
3.1.2 Linear Dynamics and Spectral Decomposition 12
3.2 Dynamic Mode Decomposition (DMD) 16
3.2.1 Operator-based Decomposition: Koopman Operator Theory 16
3.2.2 DMD Algorithm 18
3.3 DMD Variants 21
3.3.1 Augmented DMD with Hankel Matrix 21
3.3.2 DMD with Control 23

4 Plantar Pressure and Gait Cycle **25**
4.1 Introduction 25
4.2 Materials and Methods 27
4.2.1 Database 27
4.2.2 Data Structure 29
4.2.2.1 Plantar Pressure map 29
4.2.2.2 Temporal Signal of Plantar Pressure in a Gait Cycle: Ground Reaction Force 31
4.2.2.3 Spectral Features for Gait Cycle 32

4.3 Results . . . 34
4.3.1 Participants . . . 34
4.3.2 Data Processing . . . 34
4.3.3 DMD Reconstruction of a Gait Cycle . . . 35
4.3.4 Identifying Plantar Pressure Baseline . . . 36
4.3.5 Validation of Approach . . . 37
4.4 Discussion . . . 40
4.5 Conclusions . . . 42

5 Detection of Change in Gait Cycle 44
5.1 Introduction . . . 44
5.2 Methods . . . 47
5.2.1 Data Drift Detection . . . 47
5.2.2 Test statistic metrics . . . 48
5.2.3 Permutation Testing in Drift Detection . . . 49
5.3 Results . . . 51
5.3.1 Distribution of DMD spectral features . . . 51
5.3.2 Drift Detection via Permutation Testing . . . 51
5.3.3 Longitudinal Feature Drift . . . 54
5.3.4 Summary of Drift Detection Across Participants . . . 56
5.4 Discussion . . . 60
5.5 Conclusion . . . 62

6 Beyond the Floor: Toward Multisensor Fusion for Gait Analysis 63
6.1 Introduction . . . 63
6.2 Data Processing . . . 67
6.2.1 Data Structure . . . 68
6.2.1.1 Plantar Pressure . . . 68
6.2.1.2 Inertial Measurement Units . . . 68
6.2.1.3 Data Fusion . . . 68
6.2.2 Data Matrix . . . 69
6.3 DMD with control to fusion dataset . . . 70
6.3.1 Variability of the Gait Cycles . . . 74
6.4 System Identification Perspective . . . 75
6.5 Conclusion . . . 77

Reference List 81

List Of Tables

5.1 Kolmogorov-Smirnov metric, $\sup_x |E_b(x) - E_t(x)|$, where $E_b(x)$ and $E_t(x)$ are the empirical cumulative distribution function for the baseline visit (V_b) and the target visit labeled by V_t. In the 3rd through 6th columns, only drift-detected modes are presented. Here, f_α is the frequency, λ_α is the decay rate, $|A_\alpha|$ and $\angle A_\alpha$ are the magnitude and the phase of the initial condition of the DMD features ($\alpha = 1, 2, \cdots, 6$). In the last column, the overall drift status is presented. 55

5.2 Energy metric, $\sqrt{2\sum_x |E_b(x) - E_t(x)|^2}$. As in TABLE 5.1, only drift-detected modes are presented in the 3rd through 6th columns, and in the last column, the overall drift status is presented. Notice that the overall drift status and the modes of the DMD features are identical to the results in TABLE 5.1. . 56

5.3 Anderson-Darling, $\sum_x \left(\frac{|E_b(x)-E_t(x)|}{\sqrt{(m+n)D(x)(1-D(x))}} \right)^2$ with m and n are the number of footsteps in the baseline visit (V_b) and target visit (V_t), respectively and $D(x)$ is the joint ECDF of all footstep samples. As in TABLE 5.1 and TABLE 5.2, in the 3rd through 6th columns, only drift-detected modes are presented, and in the last column, the overall drift status is presented. Notice that the drift status for subject 1 is detected as drift from warning status and subject 6 from stable to warning status. 57

List Of Figures

1.1 Plantar pressure measurements from (a) cancer patient 15 and (b) cancer patient 26 before chemotherapy. Top row: Temporal transitions of plantar pressure distributions at key stance-phase events. Bottom row: Corresponding ground reaction forces acting on the center of pressure (black dot). Reproduced from (Seo et al., 2024), licensed under CC BY 4.0. 2

3.1 Schematic illustration of the Koopman operator confined in a finite m-dimensional invariant subspace spanned by the measurement functions $y \in \mathbb{R}^m$. In this case the Koopman operator is represented by a $m \times m$ matrix K_t. 18

4.1 Contact areas and ground reaction forces from (a-e) participant 1 and (f-j) participant 2. The upper rows show the spatial distribution of the contact area (shaded by gray) and the center of pressure (black dot) at various times in a gait cycle stance phase (Levine et al., 2012): (a,f) the first heel contact of the lead leg at $t = 0$, (b,g) foot-flat, (c,h) heel-off, (d,i) contralateral heel contact (Ben-Gal et al., 2020) and (e,j) toe-off of the lead leg are presented. Below each contact area, the value of the ground reaction force on the center of pressure is marked on the ground reaction force curve (solid line). Reproduced from Seo et al. (2024), licensed under CC BY 4.0. 29

4.2 Footsteps generated on the Tekscan Strideway pressure mapping device indicate the spatial trajectory of the walking movement. To show the whole footstep, the plantar pressures are integrated over time as labeled by the bar, $\overline{P}_s(x_c, y_c) = \int dt_c P_s(x_c, y_c, t_c)$. Note that the gait events of heel strike and toe-off always overlapped (temporally) a period of plantar contact from the contralateral foot. This defines the double support phase of gait. 31

4.3 The scatter plot of the spectral feature SF_α in 3-dimensional feature space defined by the decay rate γ, the frequency f, and the initial amplitude A, including all the gait cycles of a foot. The red squares represent the centroids of each SF_α points, which will be discussed in Result section. 33

4.4 The dynamic mode decomposition (DMD) reconstruction. Here, $\mathrm{DM}_\alpha(t) = (\lambda_\alpha^{1/\Delta t})^t w_\alpha$ with λ_α defined in Eq. (4.4). So, $b_\alpha \cdot \mathrm{DM}_\alpha(t) = (\lambda_\alpha^{1/\Delta t})^t b_\alpha$. The green circles represents the eigenvectors w_α and its complex conjugates at the initial time $t = 0$. Each α-th mode is filled by the blue and red colors. The squares reprents the α-th DMD mode in the state at $t = 0$ and the final state that are connected to the trajectory of the blue line. Reproduced from Seo et al. (2024), licensed under CC BY 4.0. 36

4.5 Gait baselines of (a) left foot and (b) right foot at normal walking pace and of (c) left foot and (d) right foot at fast walking pace for the participants before the chemotherapy treatment. Freq, GR, and Ini Cond denote frequency f_k, decay rate α_k, and initial condition $s_{0,k}$, respectively. Reproduced from Seo et al. (2024), licensed under CC BY 4.0. 38
4.6 Projection of the centroids of the dynamic mode decomposition (DMD) modes on the initial condition against frequency plane for each of 16 participants (a) at normal pace walks and (b) at fast walks. The color of the circles denotes the centroid of the k-th DMD mode (The conjugate modes are not displayed as they have negative frequencies against the same initial conditions). The participants are distinguished by circle sizes. Frequency and Ini Cond denote frequency f_k and initial condition $s_{0,k}$, respectively. Reproduced from Seo et al. (2024), licensed under CC BY 4.0. 39
4.7 Overview of the method for extracting gait features by the dynamic mode decomposition (DMD) on the plantar force input signal. The 2-dimensional data matrix is constructed by stacking the time-delayed coordinates, equivalent to propagation at a constant speed. The DMD modes, the eigenvalues of the Koopman operator, are complex functions flowing along their trajectories. The blue and red circles indicate the eigenvalues at each mode, while the empty green circles indicate all the eigenvalues for comparison. The black squares are the initial and final values that sum up the pair of complex conjugate modes (the red and blue circles). The eigenvalues' frequency, decay rate, and initial condition characterize the gait features. The centroids (red squares) from the multiple sample gait cycles identify the unique feature of the walking activity. Reproduced from Seo et al. (2024), licensed under CC BY 4.0. 42

5.1 Overview of the workflow. a. Spatiotemporal plantar pressure data was collected as participants walk over a Tekscan Strideway mat, capturing ground contact patterns. b. From the pressure data, ground reaction force (GRF) signals are computed for each gait cycle, reflecting the dynamic interaction between foot and floor. c. The GRF time series was embedded into a Hankel matrix, which captures time-shifted signal patterns to serve as input for DMD. d. The Hankel matrix was decomposed into dynamic modes, each characterized by a decay rate, frequency, and initial amplitude–capturing the underlying temporal dynamics of the gait. e. DMD feature distributions from follow-up visits are compared to each individual's baseline, enabling identification of deviations indicative of emerging gait abnormalities. f. Permutation-based tests (Kolmogorov–Smirnov, Anderson–Darling, and Energy distance) assess the significance of distributional changes, with False Discovery Rate control determining the final drift status as Stable, Warning, or Drift. 45

5.2 (a) The overlap between the distribution of the features derived from DMD eigenvalues, γ_α vs. f_α for $\alpha = 1, 2, \cdots, 6$ of the baseline footsteps (blue dots) at V_b (Visit1) and that of the target footsteps (red dots) at V_2 and (b) the overlap between the footsteps at V_b and footsteps (yellow dots) at V_3. For visual clarity, each distribution of the points is enclosed by the same colored lines, and the corresponding centroid of each DMD α^{th} mode is indicated by a cross with blue for V_b, red for V_2, and yellow for V_3, respectively. Notice that the target data at V_2 remains the same as the baseline data with slight shifts of the centroids. In contrast, the centroids at V_3 are shifted away from the baseline boundaries, indicating a change in the gait pattern. 52
5.3 The histograms show the drift status for the DMD features (frequency f_α, decay rate λ_α, magnitude of the initial condition $|A_\alpha|$ and its phase $\angle A_\alpha$ for $\alpha = 1, 2, \cdots, 6$) at (a) the target visit V_2 and (b) the target visit V_3 against the baseline at the baseline visit (V_1) for the same participant presented in Figure 5.2 using Kolmogorov-Smirnov metric, $\sup_x |E_b(x) - E_t(x)|$, where $E_b(x)$ and $E_t(x)$ are the empirical cumulative distribution function for V_1 and V_t for $t = 2, 3$. p-values are obtained by the number of times a metric value from a permutation exceeds the initial metric value. The drift status of each DMD feature is determined by the conditions mentioned in the text with the drift threshold $\theta_D = 0.05$ and the warning threshold $\theta_W = 0.1$. 59

6.1 Movements of various body parts during the gait cycle. (Top row) a. Double stance: Initial contact by the left foot (green) and toe-off by the right foot (yellow). b. Swing phase of the right foot (yellow) while the left foot (green) in midstance. c. Double stance: Initial contact by the right foot (yellow) and toe-off by the left foot (green). d. Swing phase of the left foot (green) while the right foot (yellow) in midstance. (Bottom row) Ground reaction force (GRF) and acceleration magnitude recorded by IMU 1, aligned with the corresponding gait phases, illustrating a strong correlation between the sginals. Yellow indicates right foot and limb and green indicates left foot and limb. 0 and 7 mark the plantar pressure contact areas, 1 and 5 are the IMUs mounted on the top of the feet, 2 and 6 are IMUs attached to the ankles, and 3 is lumbar-mounted IMU, and 4 is the sternum-moutned IMU. 65
6.2 Data matrix D of a combined dataset of IMU and Tekscan. The numbers in the bracket represents the IMU devices and Tekscan contact point as shown in Figure 6.1(b). For example, the first column labeled with FootPush presents D_1 data matrix including the signals of the magnitude of accelration (a), its components (a_x, a_y, a_z), the magnitude of angular velocity (ω), and its components (ω_x, ω_y,ω_z). The last signal in the first column labeled with grf is D_7 including GRF by a foot on the ground. 70
6.3 Eigenvalues with lowest residual error. 71
6.4 The first 3 low frequency and long-tail DMD modes. 72

6.5 Data matrix D of a combined dataset of IMU (blue lines) and Tekscan superimposed with DMD reconstructed signals $\hat{D}$ (red lines). 73
6.6 The amplitude fluctuation based on the DMD modes. 74
6.7 Frequency content derived from the poles of the system matrix A, where the GRF is used as the external input and inertial signals (acceleration and angular velocity) from IMU sensors are used as outputs. (a) Schematic of the system configuration, showing GRF as input (D_7 in Figure 6.1) and IMU responses as output. (b) Frequency content of acceleration components (a_x, a_y, a_z) from IMU device 1. (c) Frequency content of angular velocity components (ω_x, ω_y, ω_z) from IMU device 2. DBSCAN clustering was applied to identify dominant spectral patterns (colored points) and outliers (grey points). 77

Abstract

Seo, Kangjun Ph.D., University of Oklahoma, Aug, 2025. Data-driven Analysis Using Dynamic Mode Decomposition For Monitoring Chemotherapy-induced Peripheral Neuropathy

Chemotherapy-induced peripheral neuropathy (CIPN) is a prevalent adverse effect of cancer treatment, known to downgrade sensory and motor function and significantly increase fall risk among survivors. Subtle gait alterations—such as reduced step length, prolonged stride times, and increased double-support phases—often accompany CIPN. However, these changes also occur in other conditions, making early detection of CIPN challenging using conventional gait metrics. To address this challenge, this dissertation develops a data-driven framework that links an individual's unique walking patterns to the onset of treatment-induced abnormalities. Central to this approach is the application of dynamic mode decomposition (DMD) and machine learning models to kinetic and kinematic signals during the gait cycle. Particularly, DMD is well-suited for analyzing gait signals, as it provides a low-dimensional, equation-free representation of high-dimensional, time-resolved biomechanical data in an interpretable way. This data-driven gait modeling enables the identification of individual-specific dynamical gait features that are not captured by standard spatiotemporal gait parameters, facilitating early detection of abnormalities.

The dissertation comprises three interrelated studies. First, baseline gait features were established for cancer survivors prior to chemotherapy by decomposing their plantar pressure data into dominant DMD mode features—frequency, decay rate, and initial amplitude. These features were used in a machine learning model that accurately identified individuals with 86–89% accuracy, confirming the uniqueness and consistency of each subject's gait dynamics. Second, a longitudinal monitoring framework was developed to detect deviations from baseline using permutation-based statistical tests applied to the distributions of DMD features. This method detected early gait deviations in 87% of participants by their second or third treatment visits, enabling timely identification of emerging CIPN symptoms. Third, the framework was extended by integrating kinetic (plantar pressure) and kinematic (IMU-based

acceleration and angular velocity) datasets using various DMD variants and state-space system identification techniques. This multi-sensor approach reveals the impact of foot-ground interaction on whole-body dynamics, offering a more comprehensive understanding of gait adaptation under neuropathic stress.

Collectively, these contributions establish a robust, interpretable, and scalable framework for personalized gait analysis. By capturing individual characteristic gait baseline and detecting deviations in the dynamical features, this work offers preemptive, AI-driven diagnostic tools in healthcare, particularly for monitoring conditions like CIPN that impact mobility and quality of life.

Chapter 1

Introduction

Human walking is a well-practiced and highly repeatable motor process. The ground reaction force generated with each step, captured through plantar pressure, is sufficiently unique to identify individuals with high accuracy. This plantar pressure signature provides critical insights into the force distribution during walking, enabling evaluations of movement execution and early detection of abnormal gait patterns (Razak et al., 2012; Feng et al., 2011). In clinical contexts, especially among cancer survivors undergoing chemotherapy with neurotoxic effects, disruptions in the neuromuscular system can lead to increased imbalance and fall risk. Subtle changes in gait may serve as early indicators of such impairments, long before traditional clinical assessments detect changes.

Recent advancements in sensor technology have significantly improved the resolution and accessibility of plantar pressure measurements, supporting applications ranging from personalized footwear development (Shorten, 2009) to clinical foot health monitoring (Mann et al., 2016; Springer et al., 2016). Despite these advances, constructing a robust baseline model of an individual's gait remains a challenge due to the natural variability in footfalls. Identifying spatiotemporal features that consistently appear across steps is critical to establishing a reliable reference against which deviations can be measured.

Walking consists of cyclic movements, where each foot transitions from one position of support to the next, generating forces that maintain body propulsion and stability. The fundamental unit of this process is the gait cycle, which is typically divided into two main phases: stance and swing. The stance phase when the foot is in contact with the ground

comprises three main events: (1) heel strike, (2) midstance, and (3) toe-off. During the swing phase, the foot is airborne, and the ground reaction force drops to zero. Gait parameters such as displacement, speed, plantar pressure, impulse, and muscle activity are commonly expressed relative to this cycle. As illustrated in Figure 1.1, plantar pressure exhibits complex spatial and temporal dynamics throughout the gait cycle, complicating traditional analytical modeling.

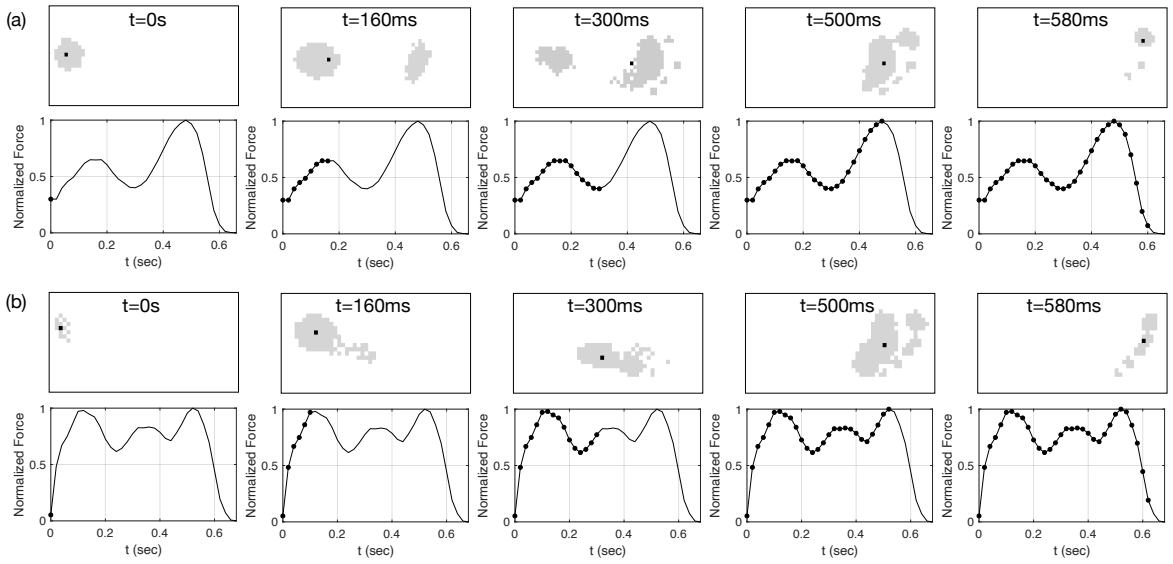


Figure 1.1: Plantar pressure measurements from (a) cancer patient 15 and (b) cancer patient 26 before chemotherapy. Top row: Temporal transitions of plantar pressure distributions at key stance-phase events. Bottom row: Corresponding ground reaction forces acting on the center of pressure (black dot). Reproduced from (Seo et al., 2024), licensed under CC BY 4.0.

With the proliferation of wearable gait sensing technologies, machine learning methods have become powerful tools for analyzing and tracking neuropathology (Harris et al., 2022). Plantar pressure datasets inherently encode nonlinear, high-dimensional dynamical behaviors in both spatial and temporal domains. A theoretical foundation for analyzing such dynamics was introduced in 1931 by Koopman (Koopman, 1931; Koopman and von Neumann, 1932), who demonstrated that nonlinear dynamical systems could be represented linearly in infinite-dimensional function spaces through spectral decomposition. Building upon this

foundation, Schmid and Sesterhenn developed the Dynamic Mode Decomposition (DMD) algorithm (Schmid and Sesterhenn, 2008; Schmid, 2010), initially for fluid dynamics, to identify coherent spatiotemporal structures from complex datasets.

DMD combines spatial dimensionality reduction with spectral decomposition, enabling a representation of system evolution through a set of spatial modes and corresponding temporal dynamics, characterized by frequency and decay/growth rate. DMD offers two primary advantages: (1) it is an equation-free, data-driven method that requires no knowledge of governing dynamics, and (2) it extracts low-rank, interpretable spatiotemporal patterns. However, classical DMD is limited in handling univariate time series or systems lacking spatial richness (Tu et al., 2014). To address this, the Hankel DMD method embeds delayed copies of the signal into a Hankel matrix (Arbabi and Mezić, 2017), enriching the data representation and enabling feature extraction even from scalar measurements. Variants of DMD, including Hankel DMD, have since found wide application across engineering, biomedical, and physical sciences (Berger et al., 2014; Brunton et al., 2016a; Kutz et al., 2016; Proctor et al., 2016; Parmar et al., 2022).

In this dissertation, we first employ Hankel DMD to extract individual-specific dynamics of gait features from plantar pressure signals recorded via a pressure-sensor-embedded walkway. The ground reaction force (GRF) time series that is spatially integrated from pressure maps is modeled through low-rank Koopman modes that capture dominant oscillatory and decaying behaviors over time. These modes are characterized by frequency, decay rate, and initial amplitude. By clustering multiple gait cycles from a given individual using these features, we construct a baseline gait signature unique to each person.

While this baseline analysis framework enables robust individual identification, it does not account for temporal changes in gait dynamics, particularly, those caused by certain conditions such as chemotherapy induced peripheral neuropathy (CIPN). Cancer survivors undergoing neurotoxic treatments often exhibit sensory deficits and reduced ankle strength

before traditional gait metrics like stride length reveal abnormalities (Winters-Stone et al., 2017). To capture such changes, we introduce a longitudinal gait monitoring framework based on repeated recordings across multiple clinical visits.

Using the same Hankel DMD framework, we track changes in spectral features over time by comparing the distribution of DMD mode parameters (frequency, decay rate, and amplitude) against each individual's baseline. Statistical tests, including Kolmogorov–Smirnov (KS), Anderson–Darling (AD), and Energy distance metrics, are used to assess distributional drifts, with false discovery rate (FDR) control applied to classify the changes as Stable, Warning, or Drift.

We validated this method using longitudinal plantar pressure data from 15 cancer survivors. Our model detected CIPN-related gait changes in 80% of participants by the second or third follow-up visit. This study may contribute to clinical gait analysis by:

1. Introducing a longitudinal gait monitoring framework based on DMD-derived dynamic features.

2. Applying permutation-based statistical tests with FDR control for distributional drift detection.

3. Demonstrating clinical utility using real-world longitudinal gait data.

4. Providing a foundation for future embedded gait monitoring systems.

Although plantar pressure data offers rich insights into foot-ground interactions, it only captures part of the full-body biomechanics involved in walking. Gait is a multimodal process, with contributions from both kinetic and kinematic variables. To develop a more comprehensive understanding of human gait, we extend our modeling approach to fuse plantar pressure data with signals from inertial measurement units (IMUs).

In our final study, we integrate ground reaction forces from the Tekscan Strideway with acceleration and angular velocity data from wearable IMUs attached to various body segments. This fusion enables a more complete representation of locomotor behavior, capturing both external contact forces and internal limb motions. To analyze this multimodal dataset, we implement advanced DMD algorithms including Optimized DMD, which enhances mode selection through gradient-based residual minimization, Sparsity-Promoting DMD, which enforces ℓ_1-regularization to isolate dominant modes, and DMD with Control (DMDc), which explicitly models external influences on system dynamics. In this formulation, acceleration and angular velocity are treated as intrinsic state variables, while GRF serves as the control input. This approach aligns with system identification theory, allowing us to construct interpretable input-output models in a state-space framework.

The remainder of this dissertation is organized as follows. Chapter 2 reviews related literature on plantar pressure analysis, DMD-based modeling, and gait sensing technologies. Chapter 3 presents the theoretical foundations of DMD and its variants. Chapter 4 introduces the baseline gait feature extraction framework using Hankel DMD. Chapter 5 extends this framework to a longitudinal monitoring, incorporating statistical drift detection. Chapter 6 concludes this dissertation by exploring multimodal gait modeling using sensor fusion and advanced DMD algorithms, including DMDc.

Chapter 2

Literature Review

Gait analysis has been widely studied across clinical and engineering domains, with a range of sensing modalities and analysis methods applied to detect neuromuscular and age-related gait impairments.

2.1 Conventional Gait Metrics and Variability Analysis

Researchers in Hsieh et al. (2019) applied statistical analyses to distinguish between cancer survivors with neuropathy, cancer survivors without neuropathy, and healthy control participants. The study considered gait speed and average step length and width, as well as their standard deviations, representing step length and step width variability, respectively. All participants completed two walking trials at a self-selected pace across a 6-meter Zeno™ pressure-sensitive walkway.

In Zhang et al. (2021), the researchers investigated fatigue-related gait instability in older adults to reduce fall risk. The study focused on detecting the effect of fatigue on plantar pressure variability using in-shoe pressure measurement systems consisting of 99 capacitive-type pressure sensors. Metrics such as peak pressure, pressure-time integrals, and median contact time were computed for each plantar region, along with measures of asymmetry and variability.

The study in Franco et al. (2019) examined plantar pressure variability across different age groups and between repeated walking sessions. Participants performed barefoot walking

trials at their preferred gait speed, during which average and peak pressures were analyzed in the forefoot (FF), midfoot (MF), and rearfoot (RF) regions. The data were averaged for each region and normalized to total foot pressure to account for differences in body mass and foot size.

2.2 Wearable and Sensor-Based Gait Analysis

In Gaßner et al. (2020), a sensor-based gait analysis approach was proposed to identify characteristic gait features in patients with Huntington's disease (HD). Gait data were collected using SHIMMER wearable sensors attached to the posterior aspect of both shoes. These sensors recorded tri-axial accelerometer and gyroscope signals during standardized walking tests. Gait parameters such as stride length, gait velocity, and stride, stance, and swing durations were computed on a per-stride basis, using methods developed in Barth et al. (2015); Rampp et al. (2015).

2.3 Dynamic Mode Decomposition in Gait Studies

In Boudali et al. (2017), the authors employed the Dynamic Mode Decomposition (DMD) method to analyze gait dynamics by mapping upper limb motion to contralateral lower limb motion, both with and without cane assistance. Dynamic variables for the upper limbs included angular positions and velocities of the shoulder, elbow, wrist, and club, while lower limb variables comprised contralateral hip, knee, and ankle angles. These measurements were derived using a motion capture system equipped with cameras.

2.4 Baseline Gait Signatures with DMD

In prior work (Seo et al., 2024), we demonstrated that DMD features extracted from pre-treatment plantar pressure signals exhibit consistent, individualized gait signatures. Using frequency, decay rate, and initial amplitude as modal descriptors, the study achieved ≥86% accuracy in subject identification. However, that study focused exclusively on baseline identification and did not address longitudinal gait changes or abnormalities resulting from treatment.

2.5 Novelty of the Present Approach

The studies above differ from the proposed approach in several key ways. Studies such as Razak et al. (2012) and Feng et al. (2011) utilize in-shoe pressure sensors, while Boudali et al. (2017) relies on video-based motion capture. In contrast, the proposed system uses a pressure mat, which, when properly calibrated, provides stable and repeatable measurements with lower variability. Prior studies often focus on averaging kinematic variables, thereby limiting the ability to capture temporal dynamics that DMD can reveal. Additionally, existing methods tend to analyze pressure at discrete points or regions on a foot template, whereas the proposed approach examines full pressure distributions over time at each sensor location.

Although Boudali et al. (2017) applied DMD to study joint kinematics extracted from video, such methods require preprocessing to obtain angular data and are subject to uncertainties related to recording and processing quality. In contrast, the pressure mat in our system captures direct interactions between the foot and surface. Therefore, any manifestation of chemotherapy-induced peripheral neuropathy (CIPN) is expected to influence the pressure patterns directly. Finally, Boudali et al. (2017) analyzes gait using only 12 spatial parameters, whereas the proposed method offers a significantly higher spatial resolution, providing a more detailed and robust representation of gait dynamics.

Chapter 3

Data-Driven Modeling of Dynamical Systems

3.1 Dynamical Systems

A dynamical system refers to a system whos state is uniquely described by a set of variables and whos behavior is governed by predefined rules. It can be represented by a set of the first order diffrential equations over either discrete time steps or a continuous time line. In general, the mathematical formulations for the continuous time dynamical system is given by a first-order differential equation,

$$\frac{dx(t)}{dt} = \dot{x}(t) = f(x(t)),\ t \in \mathbb{R},\ x \in \mathbb{R}^m, \tag{3.1}$$

where $x(t) = [x_1(t), x_2(t), \cdots, x_m(t)]^T$ is the state variable in the m-dimensional space at time t, and $f = [f_1(x(t)), f_2(x(t)), \cdots, f_m(x(t))]^T$ is a m-dimensional vector function of the state variable $x(t)$, called a propagator, governing the dynamics over time. Eq. (3.1) is a system of differential equations as

$$\begin{cases} \dot{x}_1(t) = f_1(x_1, x_2, \cdots, x_m) \\ \dot{x}_2(t) = f_2(x_1, x_2, \cdots, x_m) \\ \qquad\qquad \vdots \\ \dot{x}_m(t) = f_m(x_1, x_2, \cdots, x_m) \end{cases} \tag{3.2}$$

The main tasks can be considered in two-folds:

1. One is to solve for a state variable $x(t)$ given a rule $f(x)$. This can be done either analytically or numerically, depending on the complexity of the function $f(x)$. However, in most cases, the analytical approach is not attainable in high dimensional systems. So many numerical techniques have been developed to achive best approximate solutions using general-purpose solvers including ordinary differential euqtions (ODEs) (Ascher and Petzold, 1998; Hairer et al., 1993; Hairer and Wanner, 1996), simplectic integrators (Hairer et al., 2006; Sanz-Serna and Calvo, 1994), Krylov subspace methods (Saad, 2003; Gallopoulos et al., 1992), spectral decompositions (Trefethen, 2000; Boyd, 2001), and so forth.

2. Another is to find a rule $f(x)$ given a state variable $x(t)$. The main objectives are to infer best approximate function $f(x)$ governing the dynamics of the state variables $x(t)$ obtained by simulated or experiemental data. The main methods include derivative estimation (Chartrand, 2011; Savitzky and Golay, 1964), system identification techniques, state-space modeling (Ljung, 1999; Overschee and Moor, 1994), time series forcasting (Box et al., 2015; Kalman, 1960), machine learning (neural ODE) (Chen et al., 2018; Raissi et al., 2019), Koopman operator theory (Brunton et al., 2016b; Tu et al., 2014), and so forth. For these data-driven modeling, the state variable $x(t)$ are sampled to a time series x_k with a finite timestep Δt, or $x_k = x((k-1)\Delta t)$ with $k = 1, 2, \cdots, N, ,$ leading to discrete time dynamical systems.

The equivalent discrete-time dynamical system is given by

$$x_{k+1} = F_t(x_k), \tag{3.3}$$

where F_t is an iterative flow map of x_k, the state at time $(k-1)\Delta t$ onto x_{k+1}, the state at the advanced time step $k\Delta t$, or $F_t : x_k \rightarrow x_{k+1}$. Then, in the discrete time domain, the state can

be represented by a time series of $m \times N$ matrix representation, $[x_1, x_2, x_3, \cdots, x_N]$. Note that x_k is a m-dimensional column vector. So, the Eq. (3.3) can be cast in a matrix form

$$\begin{bmatrix} | & | & & | \\ x_2 & x_3 & \cdots & x_N \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ F_t(x_1) & F_t(x_2) & \cdots & F_t(x_{N-1}) \\ | & | & & | \end{bmatrix}. \tag{3.4}$$

Since the flow map F_t is related to the integral of f over the timestep,

$$F_t(x_k) = x_k + \int_{(k-1)\Delta t}^{k\Delta t} f(x(\tau))\, d\tau, \tag{3.5}$$

throughout this disseration, the propagator, f, and the flow map, F_t, are used interchangeably, unless otherwise specified.

3.1.1 Classifications of Dynamical Systems

In principle, the dynamical systems can be classified into an autonomous and non-autonomous system, depending on whether f, or F_t, depends on time explicitly or not.

- Autonomous system: $f = f(x(t))$, or $F_t = F_t(x_k)$.
- Non-autonomous system: $f = f(x(t), t)$, or $F_t = F_t(x_k, k)$.

In general, when a flow map involves the state variable at the time $k+1$ and k (or, a differential equation contains only a first order derivative), the system is referred to as a first-order system. Since the non-autonous systems and the higher-order systems can be converted into a autonomous first-order systems by expanding the dimension of the state variables, we refer a dynamical system to only autonomous first-order dynamical systems throughtout this dissertation, that is, $\dot{x}(t) = f(x)$, or $x_{k+1} = F_t(x_k)$.

Furthermore, the autonomous system can be categorized into a linear system and a nonlinear system based on how f, or F_t, depends on the state x:

- Linear system: $f(x) = Kx$, or $F_t(x) = Kx$ with K a $m \times m$ constant matrix.
- Nonlinear system: $f(x)$ is a nonlinear function of x. For example, $f(x) = (x + 1)x$, which may be considered as non-constant $K = x+1$. Nonlinear dynamical systems exhibit novel behavior where small changes in initial conditions can lead to unpredictable outcomes, resulting in bifurcations, chaos, or limit cycles and preventing analytical predictions in the long-term behavior (Strogatz, 2015).

3.1.2 Linear Dynamics and Spectral Decomposition

Particulary, a linear dynamical system is useful for the first task (finding x given f) as it can be analytically solvable in a closed-form solution. The equation (3.1), or (3.3), along with $f(x) = Kx$, becomes a systems of linear equations,

$$\dot{x}(t) = Kx(t),$$

leading to the state $x(t)$, or x_k, at any time $t \geq t_0$ with a given initial state $x(t_0) = x_1$ in terms of the propagator such that

$$x(t) = \exp(Kt)\, x(t_0), \tag{3.6}$$

where $\exp(Kt) = \sum_{l=0}^{\infty} 1/l!\,(Kt)^l = I + Kt + 1/2(Kt)^2 + \cdots$. Here, I is an identity matrix with the same dimension as K. Particulary, when K is invertible, $KK^{-1} = K^{-1}K = I$. So the state variable $x(t)$ can be expanded in a polynomial function of time t.

$$x(t) = \left(I + Kt + \frac{1}{2}KKt^2 + \frac{1}{6}KKKt^3 + \cdots \right) x(t_0). \tag{3.7}$$

We can easily check whether Eq. (3.6), or (3.7), is a solution of Eq. (3.1) by taking derivative of $\exp(Kt)$ with respect to time t, that is,

$$\begin{aligned}\frac{d}{dt}\exp(Kt) &= \frac{d}{dt}\left(I + Kt + 1/2(Kt)^2 + 1/6(Kt)^3 + \cdots\right)\\ &= K + K^2t + 1/2K^3t^2 + \cdots\\ &= K\left(I + Kt + 1/2K^2t^2 + \cdots\right)\\ &= K\exp(Kt),\end{aligned}$$

leading to

$$\frac{dx(t)}{dt} = \frac{d}{dt}\left[\exp(Kt)\,x(t_0)\right] = K\left[\exp(Kt)\,x(t_0)\right] = Kx(t).$$

Even though the differential equation is expressed in a simple form, anlyzing the state variable $x(t)$ is not easy to deal with as each component is interwined by other components:

$$\dot{x}(t) = K\,x(t) \leftrightarrow \begin{cases} \dot{x}_1(t) = K_{11}x_1(t) + K_{12}x_2(t) + \cdots + K_{1m}x_m(t)\\ \dot{x}_2(t) = K_{21}x_1(t) + K_{22}x_2(t) + \cdots + K_{2m}x_m(t)\\ \qquad\qquad\vdots\\ \dot{x}_m(t) = K_{m1}x_1(t) + K_{m2}x_2(t) + \cdots + K_{mm}x_m(t)\end{cases} \tag{3.8}$$

Among many techniques for the analysis, prediction, numerical simulation, estimation, and control of such solution, the simplest way to analyze the dynamical system using the linear propagator (or, flow map) is to use the spectral decomposition, which is a decomposition by eigenvalues and eigenvectors of K, whose eigendecompositon is given by

$$K = W\Lambda W^{-1} \text{ with } \Lambda = \begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m\end{bmatrix} \text{ and } W = \begin{bmatrix} | & | & & | \\ w_1 & w_2 & \cdots & w_m \\ | & | & & |\end{bmatrix}, \tag{3.9}$$

Here, Λ is diagonal matrix of the eigenvalues λ_α ($\alpha = 1, 2, \cdots, m$) and W is a $m \times m$ matrix whose columns are the linearly independent eigenvectors w_α (that is, $W\overline{W}^T = \overline{W}^T W = I$ with T indicating transpose and the bar notation indicating complex conjugation, leading to $W^{-1} = \overline{W}^T$) associated with eigenvalues. With the mulitplicative property of the eigenvalue matrix Λ, its power can be computed easily,

$$
\begin{aligned}
\Lambda^2 &= \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix} = \begin{bmatrix} \lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_m^2 \end{bmatrix} \\
\Lambda^3 &= \Lambda^2 \Lambda = \begin{bmatrix} \lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_m^2 \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix} = \begin{bmatrix} \lambda_1^3 & & & \\ & \lambda_2^3 & & \\ & & \ddots & \\ & & & \lambda_m^3 \end{bmatrix} \\
&\cdots \\
\Lambda^l &= \Lambda^{l-1} \Lambda = \begin{bmatrix} \lambda_1^l & & & \\ & \lambda_2^l & & \\ & & \ddots & \\ & & & \lambda_m^l \end{bmatrix}.
\end{aligned}
$$

Since $\exp(\Lambda t) = I + \Lambda t + \frac{1}{2}(\Lambda t)^2 + \frac{1}{6}(\Lambda t)^3 + \cdots$ is still a diagonal matrix, $\exp(Kt) = W \exp(\Lambda t) W^{-1}$, and the equation (3.6) becomes

$$
x(t) = W \exp(\Lambda t) W^{-1} x(t_0) \tag{3.10}
$$

for continuous time dynamics, and, for discrete time dynamics,

$$
\begin{aligned}
x_2 &= W \exp(\Lambda \Delta t) W^{-1} x_1 \\
x_3 &= W \exp(2\Lambda \Delta t) W^{-1} x_1 \\
&\vdots \\
x_{k+1} &= W \exp(k\Lambda \Delta t) W^{-1} x_1 \\
&\vdots \\
x_N &= W \exp((N-1)\Lambda \Delta t) W^{-1} x_1
\end{aligned} \tag{3.11}
$$

The matrix W^{-1} defines a transformation of x into intrinsic eigenvector coordinates $z = W^{-1}\,x$, where the dynamics become decoupled as

$$\dot{z}(t) = \Lambda\, z(t) \leftrightarrow \begin{cases} \dot{z}_1(t) = \lambda_1 z_1(t) \\ \dot{z}_2(t) = \lambda_2 z_2(t) \\ \vdots \\ \dot{z}_m(t) = \lambda_m z_m(t) \end{cases} \rightarrow \begin{cases} z_1(t) = \exp(\lambda_1 t)\, z_1(t_0) \\ z_2(t) = \exp(\lambda_2 t)\, z_2(t_0) \\ \vdots \\ z_m(t) = \exp(\lambda_m t)\, z_m(t_0) \end{cases} \tag{3.12}$$

for a continuous time dyanmical system, and

$$\begin{aligned} z_2 &= \exp(\Lambda \Delta t)\, z_1 \\ z_3 &= \exp(2\Lambda \Delta t)\, z_1 \\ &\vdots \\ z_N &= \exp((N-1)\Lambda \Delta t)\, z_1 \end{aligned} \tag{3.13}$$

for a discrete time dynamical system.

Linear dynamical systems are desirable since it is likely to transform the system into eigenvector coordinates, where the dynamics become decoupled, motivating many decomposition techniques such as Taylor expansion, Fourier transform, Wavelet transform, Singular value decomposition (SVD), and so forth. However, for nonlinear systems, no such closed-form solution or simple linear transformation of coordinates exist in general.

Real-world systems are in general nonlinear and show multi-scale behavior in both sapce and time. They exhibit chaotic behaviors with uncertainty in the equation of motion, in the parameters, and in the measurements of the system. Moreover, some systems are not specified by the basic equations of motion that make intractable to derive from the first principles.

3.2 Dynamic Mode Decomposition (DMD)

3.2.1 Operator-based Decomposition: Koopman Operator Theory

To address the challenges in nonlinear dyanmical systems, there are many techniques have been developed. In particular, there are two key approaches for the modern data-driven dynamical systems, Operator theoretic representations and Data-driven regression and machine learning. The former approach represent nonlinear dynamical systems in terms of infinite-dimensional but linear operators, such as the Koopman operator that advances measurement functions and Perron-Frobenius operator that advances probability densities and ensembles through the dynamics. The latter investigate systems in a data-driven model, such as the dynamic mode decomposition (DMD), the data-driven Koopman methods, the sparse identification of nonlinear dynamics (SINDy), and various machine learning algorithms.

The Koopman operator is a linear, infinite dimensional operator that represents the action of the nonlinear dynamical systems on the Hilbert space of measurement functions of the state (Rowley et al., 2009; Mezic, 2013). In contrast to the linearization techniques, the action of Koopman operator does not linearize the dynamics. Instead, it represents the measurement function of the dynamical system in terms of infinite-dimensional operator.

Mathematically, the Koopman operator $\mathcal{K}$ that is an infinite-dimensional linear operator acts on the Hilbert space $\mathcal{H}$ of scalar-valued meansurement functions of the state x, or, $g : \mathbb{C}^d \to \mathbb{C}$. The action the Koopman operator on the measurement function g equals to the composition of the function g with the flow map F_t such that $\mathcal{K}g = g \circ F_t$. This composition operation leads to the advancement of meansurement function along with the flow F_t as

$$\mathcal{K}g(x_k) = g(F_t(x_k)) = g(x_{k+1}). \tag{3.14}$$

The compositional action of the Koopman operator holds for all measurement functions g evaluated at any state x. Therefore, the mapping over timestep Δt is linear even though the underlying dynamics $F_t(x)$ is nonlinear. This is why the Koopman operator is fundamentally different from linearinzation techniques.

The heart of the Koopman opertor theory is the spectral decomposition of the measurement functions $g(x)$. Consider the eigenvalue decomposition of the Koopman operator

$$\mathcal{K}\phi(x) = \lambda\,\phi(x), \tag{3.15}$$

where λ is the eigenvalue of the Koopman operator $\mathcal{K}$ and $\phi(x)$ is the corresponding eigenfunction. Since the Koopman operator is linear operator, any measurement function $g(x)$ can be expressed in a superposition of the eigenfunctions.

$$g(x) = \sum_{\alpha=1}^{\infty} \phi_\alpha(x)\,\upsilon_\alpha, \tag{3.16}$$

where υ_α are the Koopman mode corresponding to the eigenfunctions $\phi_\alpha(x)$. Due to the orthonormality of the eigenfunctions $\langle \phi_\alpha, \phi_\beta \rangle = \delta_{\alpha,\beta}$ with $\langle x, y \rangle = \int dx\, y^*(x)x(x)$ being a projection operation, the Koopman mode is a projection onto the corresponding eigenfucntions, or, $\upsilon_\alpha = \langle g(x), \phi_\alpha(x) \rangle$. Then, the advancement of the measurement function $g(x)$ in time is simply expressed as a linear superposition multiplied by eigenvalues such that

$$g(x_k) = \sum_{\alpha=1}^{\infty} \phi_\alpha(x_k)\,\upsilon_\alpha \rightarrow g(x_{k+1}) = \sum_{\alpha=1}^{\infty} \lambda_\alpha\,\phi_\alpha(x_k)\,\upsilon_\alpha,$$

since $g(x_{k+1}) = \mathcal{K}g(x_k) = \sum_{\alpha=1}^{\infty} \mathcal{K}\left[\phi_\alpha(x_k)\right]\,\upsilon_\alpha = \sum_{\alpha=1}^{\infty} \left[\lambda_\alpha\,\phi_\alpha(x_k)\right]\,\upsilon_\alpha$.

Sometimes, it is possible to find the invariant subspace spanned by the m-dimensional measurement functions $\mathbf{g}(x) = \{g_1(x), g_2(x), \cdots, g_m(x)\}$. Then, the infinite dimensional

Koopman operator can be reduced to a m-dimensional operator, leading to a $m \times m$ matrix representaion, K_t such that

$$K_t\, g(x_k) = g(x_{k+1}). \tag{3.17}$$

The scematic process of the iterative operation of the Koopman operator on the measurement function $y_k = g(x_k)$ along with the flow map of the original nonlinear dynamical systems $x_{k+1} = F_t(x_k)$ is illustred in Figure 3.1.

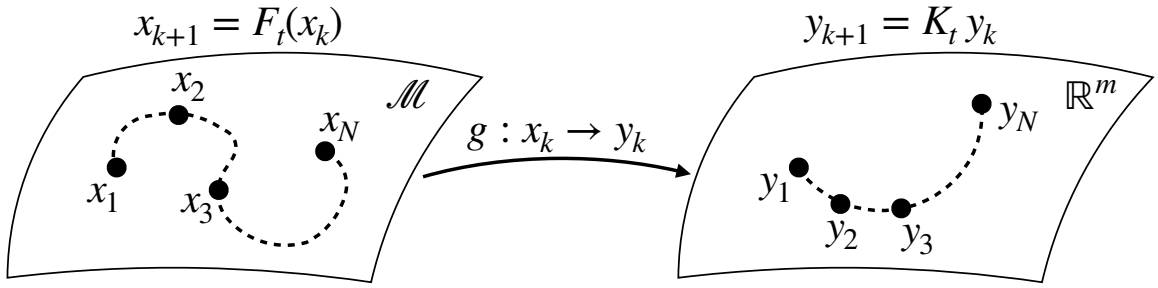


Figure 3.1: Schematic illustration of the Koopman operator confined in a finite m-dimensional invariant subspace spanned by the measurement functions $y \in \mathbb{R}^m$. In this case the Koopman operator is represented by a $m \times m$ matrix K_t.

However, the fact the Koopman operator is an infinite-dimensional operator is challenging to implement in numerical algorithm. Moreover, the quality of any finite-dimensional approximation depends on the choice of the measurement functions $g(x)$.

3.2.2 DMD Algorithm

Dynamic Mode Decomposition (DMD) is a numerical implementation of the Koopman eigenvalues λ_α and modes υ_α in a finite-dimensional approximation. The heart of DMD algorithm is to determine the Koopman eigenvalues and modes directly from the data.

The definition of prototype DMD algorithm is given by Tu et al. (2014): Given a dynamical system $x_{k+1} = F_t(x_k)$ and two sets of data matrices X and X',

$$X = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_{N-1} \\ | & | & & | \end{bmatrix}, \quad X' = \begin{bmatrix} | & | & & | \\ x'_1 & x'_2 & \cdots & x'_{N-1} \\ | & | & & | \end{bmatrix}$$

with x_k an initial state and $x'_k = x_{k+1}$ its corresponding output after evolution by time Δt, the DMD modes are defined as the eigenvectors of $K = X' X^{\dagger}$, where $\dagger$ denotes the Moore-Penrose pseudoinverse.

Let $g = [g_1, g_2, \cdots, g_m]^T$ be a vector-valued observables (measurement functions) with $g_j : \mathcal{M} \to \mathbb{C}$ for $j = 1, 2, \cdots, m$, and let

$$Y = \begin{bmatrix} g_1(x_1) & g_1(x_2) & \cdots & g_1(x_{N-1}) \\ g_2(x_1) & g_2(x_2) & \cdots & g_2(x_{N-1}) \\ \vdots & \vdots & & \vdots \\ g_m(x_1) & g_m(x_2) & \cdots & g_m(x_{N-1}) \end{bmatrix}, \quad Y' = \begin{bmatrix} g_1(x'_1) & g_1(x'_2) & \cdots & g_1(x'_{N-1}) \\ g_2(x'_1) & g_2(x'_2) & \cdots & g_2(x'_{N-1}) \\ \vdots & \vdots & & \vdots \\ g_m(x'_1) & g_m(x'_2) & \cdots & g_m(x'_{N-1}) \end{bmatrix} \tag{3.18}$$

be the matrices of the measurements on the observables $y_k = g(x_k)$. The columns of Y' are the observables evolved forward in time $y'_k = g(F(x_k)) = g(x'_k)$. Then, DMD algorithm on the data of observables approximates the Koopman operator by a matrix $K_Y = Y Y^{\dagger}$. The conditions for DMD to be successful are (1) the sufficiently large space of the observables so that the Koopman eigenfunctions ϕ_α belongs to the span of g and (2) the sufficiently rich data so that the eigenvectors of K_Y belongs to the range of the data X. Under these conditions satisfied, the Koopman eigenvalues are the DMD eigenvalues. Using eigenvalues $\Lambda = \{\lambda_\alpha\}$, eigenvectors $W = \{w_\alpha\}$, and projection coefficients,

$$b_\alpha = \left(W^{-1}\right)_\alpha y_1, \tag{3.19}$$

we can expand the snapshot y_k as

$$y_k = g(x_k) = \sum_{\alpha=1}^{\infty} \lambda_\alpha^{k-1} w_\alpha b_\alpha, \tag{3.20}$$

In the numerical approximation of the infinite expansion to the finite r modes, we get a DMD implementation as

$$y_k \simeq \sum_{\alpha=1}^{r} \lambda_\alpha^{k-1} w_\alpha b_\alpha. \tag{3.21}$$

It has been known that the dimensionality reduction is robust in the context of big data and data analysis by means of the signular value decomposition (SVD) (Hastie et al., 2009; Strang, 2016). The SVD-enhanced DMD algorithm is summarized as the following: Considering the data matrix $D = [g(x_1), g(x_2), \cdots, g(x_N)]$ defined in Eq. (3.18),

1. Define $Y = [D_1, D_2, \cdots, D_{N-1}]$ and $Y' = [D_2, D_3, \cdots, D_N]$

2. Compute the SVD of Y as $Y = USV^*$, where U is the left singular matrix, which is used as the basis to compute a matrix representation of DMD operator K, S are a diagonal singular value matrix, and V are the right singular matrix. The exact DMD operator K is the finite-dimensional operator that maps the columns of Y to Y', or $K = Y'Y^\dagger$.

3. Compute the matrix $A = U^*Y'VS^{-1}$, which is an approximate solution of K in the reduced dimensional subspace by U, in other words, $A = U^*KU$.

4. Decompose A into $A\,\eta_\alpha = \eta_\alpha \lambda_\alpha$ for $\alpha = 1, 2, \cdots, r$.

5. Then, the pairs $(\lambda_\alpha, \eta_\alpha)$ lead to the exact dynamic modes $w_\alpha = 1/\lambda_\alpha Y'VS^{-1}\eta_\alpha$ and the projected dynamic modes $w_\alpha = U\eta_\alpha$. Note that the exact dynamic modes become the projected dynamic modes when the column space of Y' lies in the range of Y.

The success of the Exact DMD method is attributed to the reduction of the dimensionality, which is determined by the truncation of the non-negative singular values of S. Because the truncation is equivalent to the low energy filtering in the Fourier transform, deciding how many singular values to keep in the practical computation is crucial. There have been many different criteria proposed for optimal truncation. The hard threshold techniques truncate at either the elbows in the singular value plot on a logarithmic scale or the singular value retaining 99% or 99.9% of the total variance in the data (Jolliffe, 2002; Gavish and Donoho, 2014; Shlens, 2014). The r number of the largest singular values are considered in the expansion in Eq. (3.21) discarding the rest of the small singular values. In this work, we choose the latter method to find the optimal r, or, $\|Y' - KY\|_2 \leq \epsilon = 10^{-6}N$, where $\|\cdot\|_2$ is the Euclidean 2-Norm.

3.3 DMD Variants

3.3.1 Augmented DMD with Hankel Matrix

When the data contains only the temporal dimension, the standard DMD method fails to obtain the spatial dynamic modes. To generalize the exact DMD method to the one-dimensional time-sequential data, (Arbabi and Mezić, 2017) proposed the algorithm to create artificial spatial dimensionality using the Hankel matrix, which involves stacking multiple time-shifted copies of the data.

Since the pressure signal in a gait cycle converges to zero at the end of each cycle, without loss of generality, the time sequence can be extended to $[D_1, D_2, \cdots, D_N, \cdots, D_{N+m-1}]$,

where $D_{N+1} = D_{N+2} = \cdots D_{N+m-1} = 0$ (zero-padding). Then, the Hankel matrix of the observable function g_1 is given by the $m \times N$ matrix with lower half-diagonal zeros,

$$D = \begin{bmatrix} D_1 & D_2 & \cdots & D_N \\ D_2 & D_3 & \cdots & D_{N+1} \\ D_3 & D_4 & \cdots & D_{N+2} \\ \vdots & \vdots & & \vdots \\ D_m & D_{m+1} & \cdots & D_{N+m-1} \end{bmatrix}. \tag{3.22}$$

This process ensures that the Hankel DMD models the dynamics of the entire signal.

Defining $Y = [D_0, D_1, \cdots, D_{N-2}]$ and $Y = [D_1, D_2, \cdots, D_{N-1}]$,

$$Y = \begin{bmatrix} D_1 & D_2 & \cdots & D_{N-1} \\ D_2 & D_3 & \cdots & D_N \\ D_3 & D_4 & \cdots & D_{N+1} \\ \vdots & \vdots & & \vdots \\ D_m & D_{m+1} & \cdots & D_{N+m-2} \end{bmatrix} \text{ and } Y' = \begin{bmatrix} D_2 & D_3 & \cdots & D_N \\ D_3 & D_4 & \cdots & D_{N+1} \\ D_4 & D_5 & \cdots & D_{N+2} \\ \vdots & \vdots & & \vdots \\ D_{m+1} & D_{m+2} & \cdots & D_{N+m-1} \end{bmatrix} \tag{3.23}$$

we obtain the eigenvalues $\{\lambda_\alpha\}$, the corresponding eigenvectors $\{w_\alpha\}$ of the Hankel DMD operator $K = Y'\, Y^\dagger$. Along with the initial projection coefficient b_α, Eq. (3.19), each snapshot D_k can be expanded as in the exact DMD expansion,

$$D_k = \sum_{\alpha=1}^{r} \lambda_\alpha^{k-1} w_\alpha b_\alpha. \tag{3.24}$$

Note that, since the first row describes the dynamics of the systems, we should consider the spectral modes of the systems in terms of only the first row of D. It's known that he optimal ratio of embedding to window length, m/N influences reconstruction accuracy, Koopman spectral convergence (Tu et al., 2014; Arbabi and Mezić, 2017; Brunton et al., 2017; Schmid, 2010; Towne et al., 2018). In this dissertation, we found that the optimized ratio m/N and

truncation r for the low-rank dataset with zero-padding can be obtained by minimizing the estimation error, $\|Y' - KY\| \leq \epsilon$ with a threshold $\epsilon = 10^{-6}(N + m - 1)$.

3.3.2 DMD with Control

We modeled the onset of the gait cycle in the double stance phase, where both feet are on the ground. One foot at toe-push-off pushes up its foot and the leg and the whole body forward. At the same time, the other foot at foot-flat supports the body against the ground, pivoting the swings of the other foot and leg. Thus, the Tekscan pressure signals play different roles in a gait cycle. As a consequence, the leading dynamical system consists of the IMU signals considered as output dynamics and the Tekscan pressure considered as input control. Then, the gait starts with a toe-push-off force signal as a system input. Figure 6.2 illustrates the signal of the leading dynamical system in a matrix form.

The dynamical system is formulated as

$$Y' = KY + B\Upsilon, \tag{3.25}$$

where $Y = [D_1, D_2, \cdots, D_{N-1}]$ is a $m \times N - 1$ data matrix including the initial value of the signals, $Y' = [D_2, D_3, \cdots, D_N]$ is one time-sample advanced signals with the same size of Y, K is a $m \times m$ matrix associated with time-evolution of the signals, and Υ is an input signal of the toe-push-off force signal, and B is weighted effects on the IMU and Tekscan sensors. The main task is to find the underlying dynamical meachanism characterized by K under the influence of the input Υ and B.

We employed the DMD with control algorithm (Heersink et al., 2017; Rosenfeld and Kamalapurkar, 2021; Mieg and Mönnigmann, 2025) for the unknown K and B, given collected

data signals D and input signal Υ. Using the truncated signular value decomposition (SVD) of the augmented input data matrix Y with Υ

$$\Omega = \begin{bmatrix} Y \\ \Upsilon \end{bmatrix} = USV^* = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} SV^* \tag{3.26}$$

and the truncated output matrix $Y' = \hat{U}\hat{S}\hat{V}^*$, we can compute the A in a reduced dimension corresponding to K

$$A = \hat{U}^* Y' V S^{-1} U_1^* \hat{U}. \tag{3.27}$$

With the eigenvalue decomposition of A such as $A\eta = \eta\Lambda$, the exact dynamic mode is given by

$$W = Y' V S^{-1} U_1^* \hat{U} \eta \tag{3.28}$$

corresponding to the eigenvalues Λ.

Chapter 4

Plantar Pressure and Gait Cycle

4.1 Introduction

Gait in healthy individuals is a well-practiced, highly repeatable, yet complex sensorimotor function (Clark, 2015). Slower gait speeds are linked to poorer health outcomes across diverse populations, including adults with cancer (Liu et al., 2019; Pamoukdjian et al., 2017; Luo et al., 2023). Emerging applications of artificial intelligence (AI) in gait analysis leverage clustering techniques to identify patterns in complex features derived from gait sensor data (Harris et al., 2022). Plantar pressure data obtained from sensor-embedded walkways provide rich, multidimensional information that can be examined for spatiotemporal consistency and variability both within and between individuals. Distinctive features within these pressure maps are often unique enough to enable high-accuracy individual identification (Liu et al., 2019), and sensitive enough to detect neuromotor impairments (Cao et al., 2022; Zou et al., 2023).

Cancer survivors face an elevated risk of falls (Huang et al., 2017), with neurotoxic chemotherapy being a known contributing factor (Bao et al., 2016; Winters-Stone et al., 2017). Alterations in plantar pressure distribution have been documented in individuals with diabetes, experimental neuropathy (Al-Angari et al., 2017; Eils et al., 2002), and in older adults (Zhang et al., 2021); however, little has been published regarding the spatial or temporal gait dynamics associated with chemotherapy-induced peripheral neuropathy (CIPN).

When pressure or force-sensing technologies are used to assess CIPN-related gait abnormalities, the primary metrics reported are typically gait speed, step length, and stride duration (Winters-Stone et al., 2017; Gilchrist and Tanner, 2016; Marshall et al., 2017; Hsieh et al., 2019). These studies consistently show that individuals with CIPN tend to take shorter steps, require more time per stride, and spend a greater proportion of time in the double-support phase (Winters-Stone et al., 2017; Zahiri et al., 2019). However, such gait features are not unique to CIPN and are also observed in conditions such as frailty and multimorbidity (Thaler-Kall et al., 2015). As a result, relying solely on these parameters is unlikely to distinguish CIPN from other cancer-related gait disturbances (Gaßner et al., 2020). Advanced gait classification systems for complex neurologic conditions increasingly incorporate multivariate modeling and sensor fusion techniques (Harris et al., 2022; Pulido-Valdeolivas et al., 2018).

Despite the potential, we found few examples in the literature of Hankel DMD being applied to plantar pressure data. Instead, most prior applications of DMD in gait research have focused on joint kinematics, such as angular position and velocity from motion capture, including comparisons between cane users and healthy controls (Boudali et al., 2017). Other studies have employed tri-axial accelerometers and gyroscopes embedded in footwear to identify gait features most correlated with disease progression, such as in Huntington's disease (Gaßner et al., 2020). Machine learning models trained on average per-stride values—such as gait speed, stride length, and various temporal metrics—have demonstrated clinical relevance, particularly for variability in stride timing (Barth et al., 2015; Rampp et al., 2015). These variability metrics could similarly be applied to pressure dynamics, which may reveal deeper complexity and individual variation in conditions like CIPN, especially given its heterogeneous presentation across patients and chemotherapeutic agents. Capturing this dynamic information allows pressure development during each step to be evaluated in terms of its spatial, temporal, and magnitude characteristics.

As an initial step toward automating CIPN detection, our objective is to develop a composite model that characterizes a cancer survivor's baseline gait prior to chemotherapy, using spatiotemporal features extracted from plantar pressure data. In this paper, we detail our processing pipeline for plantar pressure data collected from 16 female cancer survivors before the initiation of neurotoxic chemotherapy. We describe the steps taken to prepare the data for input into a DMD algorithm and outline how we used DMD to extract dynamic gait features. Our aim is to quantify the full-stance-phase pressure dynamics and to identify unique baseline gait patterns for each participant.

4.2 Materials and Methods

We performed developmental computational work using data from a study to determine the feasibility of longitudinal performance assessments in women who are starting a course of neurotoxic chemotherapy (taxane/platinum-based) as adjuvant treatment for breast or gynecologic cancer. The study received approval #11473 from the University Institutional Review Board, and all women signed consent prior to enrollment. The developmental work focused on establishing the dynamical property of the pressure data, the data-driven modeling of the plantar pressure patterns using the dynamic mode decomposition to define the features, and initiation of a machine learning algorithm to identify individual's characteristics associated with the features.

4.2.1 Database

For this computational work, we selected a database of plantar pressure files from 16 women collected by research assistants from the laboratory of E.H. and analyzed by K.S.. The data were from women with breast or gynecological cancer who could stand and walk without assistance and were not recovering from a recent lower limb injury or surgery. Women with

metastases to the nervous system or spine, or with progressive neurological conditions, were excluded. A few days before their first chemotherapy infusion, each woman walked at their self-selected usual pace for 4 to 8 total passes over a 6m×1m high-resolution Strideway force-sensor embedded ground floor mat manufactured by Tekscan Inc. with 4 sensors per cm^2 spatial resolution and 50Hz sampling rate. Each woman then repeated a series of 4 to 8 total passes but at their self-selected *fast but safe* pace. Standardized non-skid socks were provided for all walks.

The raw data format is 3000 frames of 2D spatial pressure snapshot with 128×895 unit sensors. The dimension of each unit cell is 5mm×5mm, covering the entire footstep by approximately 430 unit sensors. The sampling rate is 50Hz (Δt = 20ms). The stance phase in the average speed gait cycle takes about 40 to 50 frames, corresponding to 0.8 sec to 1 sec. Our pressure analyses are performed on the datasets with the pressure values calibrated to the natural force unit, Newton (N/m^2), or Pascal (Pa).

A participant made 4 to 8 passes in a walking measurement, generating 8 to 10 footsteps in a pass, where the participant walks in one direction. The datasets contain the intervals between the passes as the participant stepped off the walkway and turned around for the next pass. In the dataset, these appear as snapshots with no pressure (filled with zeros). We discarded the zero-pressure snapshots to condense all passes into a single gait dataset for a single individual without a pause. As the spatial dimensional limit of the setup environment, the walking direction in each pass is alternated with each pass (i.e., footfalls are aligned in the left direction in all even-numbered passes and then in the right direction in all odd passes). We flipped the data in every other pass row and column-wise to maintain the left foot in the higher columns and the right foot in the lower columns in each 2D snapshot.

4.2.2 Data Structure

The collected data forms a 3D tensor with snapshots of a 2D pressure map and 1D temporal frame. For most cases in this database, each foot (right or left) generated 4 to 5 footsteps on the ground sensor when walking at self-selected paced, yielding 25 to 36 sample footsteps per visit.

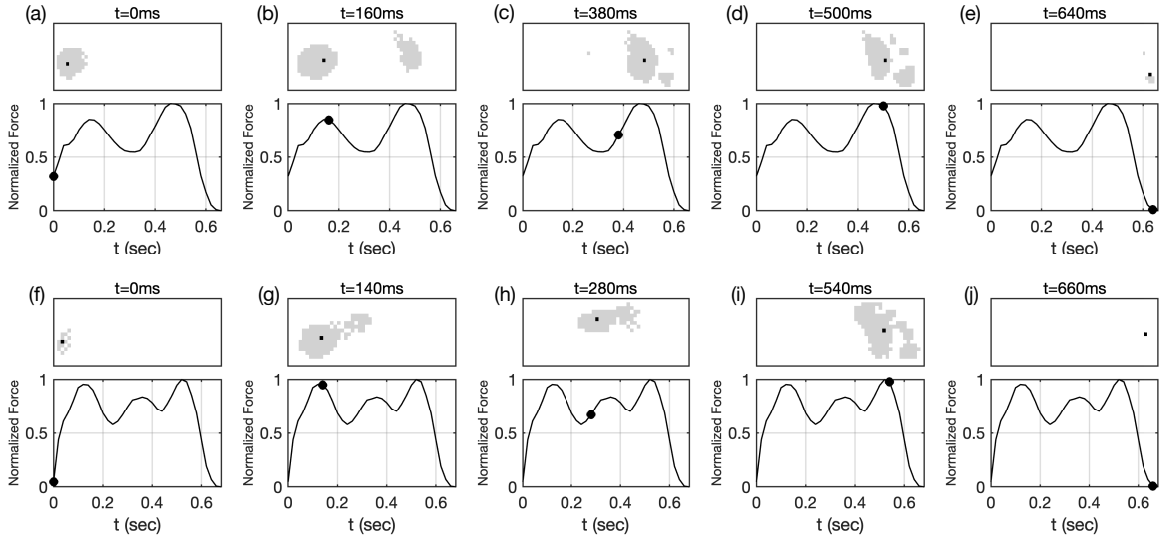


Figure 4.1: Contact areas and ground reaction forces from (a-e) participant 1 and (f-j) participant 2. The upper rows show the spatial distribution of the contact area (shaded by gray) and the center of pressure (black dot) at various times in a gait cycle stance phase (Levine et al., 2012): (a,f) the first heel contact of the lead leg at $t = 0$, (b,g) foot-flat, (c,h) heel-off, (d,i) contralateral heel contact (Ben-Gal et al., 2020) and (e,j) toe-off of the lead leg are presented. Below each contact area, the value of the ground reaction force on the center of pressure is marked on the ground reaction force curve (solid line). Reproduced from Seo et al. (2024), licensed under CC BY 4.0.

4.2.2.1 Plantar Pressure map

The pressure measured on the sensor is a reaction force of the ground against the contact force exerted by the respective part of the foot during the gait motion. The pressure at each time frame forms a spatial distribution, as shown in Figure 4.1. It shows the temporal evolution of the spatial pressure distribution and the corresponding center of pressure during each

gait cycle's stance phase. The value of the ground reaction force is obtained by integrating the pressure over the space at each time frame, resulting in a time series force signal acting on the center of pressure.

To describe and characterize the dynamical patterns of the plantar pressure in a gait cycle, we consider the plantar pressure as a spatial function $P_s(x, y, t)$ in time sequence, where x is the coordinate on the sensor along the walking direction, y is the coordinate perpendicular to x, $t = 0, \Delta t, 2\Delta t, \cdots, N\Delta t$ is the timesteps with an interval $\Delta t = 20$ millisecond, and the supscript $s = l, r$ indicates the left and right foot, respectively. Generally, the plantar pressure is a periodic function

$$P_s(x, y, t) = P_s(x + cL_x, y + cL_y, t + cT), \tag{4.1}$$

where $c = 1, 2, 3, \cdots, C$ indicates the c-th gait cycles of total C gait cycles, T is the temporal interval from a gait cycle to the next in the same side foot, and L_y and L_x are the transverse and logitudinal distances between the footsteps of each foot, respectively (See Figure 4.2). To clearly differentiate the footsteps between the gait cycles, the plantar pressure generated during the c-the gait cycle is labeled as

$$P_s(x + cL_x, y + cL_y, t + cT) = P_s(x_c, y_c, t_c), \tag{4.2}$$

in which t_c is truncated to the specific gait cycle, $(c-1)T < t_c \leq cT$, and x_c is a coordinate relative to cL_x, that is, $x_c = x - cL_x$. In the same fashion, $y_c = y - cL_y$. Usually, the stride speed is calculated by taking the ratio of the longitudinal distance L_y to the time difference between gait cycles, $v_c = L_y/(t_{c+1} - t_c)$, and the mean stride speed is given by $\overline{v} = \frac{1}{C}\sum_{c=1}^{C} v_c$.

Notice that, due to the inherent variability of human locomotion, the difference,

$$\delta_s(c, c') = \left\| P_s(x_c, y_c, t_c) - P_s(x_{c'}, y_{c'}, t_{c'}) \right\|,$$

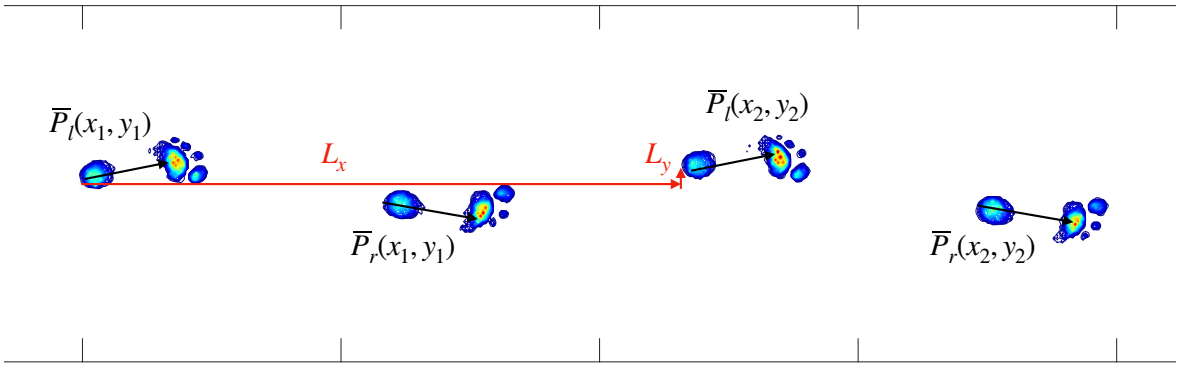


Figure 4.2: Footsteps generated on the Tekscan Strideway pressure mapping device indicate the spatial trajectory of the walking movement. To show the whole footstep, the plantar pressures are integrated over time as labeled by the bar, $\overline{P}_s(x_c, y_c) = \int dt_c P_s(x_c, y_c, t_c)$. Note that the gait events of heel strike and toe-off always overlapped (temporally) a period of plantar contact from the contralateral foot. This defines the double support phase of gait.

which reflects the local variability of the pressure distributions over the multiple gait cycles, is always present even in the well-controlled walking.

4.2.2.2 Temporal Signal of Plantar Pressure in a Gait Cycle: Ground Reaction Force

In this work, we focus on understanding the essential dynamical characteristics of each survivor's gait cycle and identifying their personalized baseline. To that end, we integrate the spatial components of the plantar pressure of each footstep as

$$GRF_s(t_c) = \int dx_c\, dy_c\, P_s(x_c, y_c, t_c), \tag{4.3}$$

where $GRF_s(t_c)$ describes the temporal development of the plantar force at the c-th gait cycle of each foot, called a ground reaction force (GRF), meaning the temporal profile of the total contact force between the foot and the ground. The unit is the standard force unit, Newton (N) or Pascal multiplied by the contact area (Pa·m^2) at each timestep $t = (k-1)\Delta t$. Therefore, the strength of the spatially integrated pressure $GRF_s(t_c)$ results from the correlation of the ground reaction force with the instantaneous contact area.

4.2.2.3 Spectral Features for Gait Cycle

In general, the DMD eigenvalues $\{\lambda_\alpha\}$ and the corresponding eigenvectors $\{w_\alpha\}$ are complex numbers that consist of the real and imaginary parts. Thus, we can rewrite the eigenvalues as

$$\lambda_\alpha = \exp(\gamma_\alpha \Delta t)\, \exp(i2\pi f_\alpha \Delta t),\ \alpha = 1, 2, \cdots, r \tag{4.4}$$

where $\exp(\gamma_\alpha \Delta t)$ is the magnitude of the eigenvalue, $|\lambda_\alpha|$, and $\exp(i2\pi f_\alpha \Delta t)$ is the rotation relative to the real. The time evolution of the signal can be considered as a discretized temporal signal sequence with a sampling rate $1/\Delta t$.

$$D = [D_1, D_2, \cdots, D_N], \tag{4.5}$$

where the k-th element is the GRF at instant $t = (k-1)\Delta t$. Applying the Hankel DMD to the time-series data, the GRF can be reconstructed by the DMD eigenvalues and the DMD eigenvectors $\hat{D} = [\hat{D}_1, \hat{D}_2, \cdots, \hat{D}_N]$ as

$$GRF(t = (k-1)\Delta t) \simeq \hat{D}_k = \sum_{\alpha=1}^{r} \left[e^{(\gamma_\alpha + i2\pi f_\alpha)\Delta t} \right]^{k-1} w_\alpha b_\alpha. \tag{4.6}$$

Now, we can extract the features of the GRF time-series in terms of the decay (or growth) rate

$$\gamma_\alpha = \frac{\mathrm{Re}\ln(\lambda_\alpha)}{\Delta t}, \tag{4.7}$$

and the frequency,

$$f_\alpha = \frac{\mathrm{Im}\ln(\lambda_\alpha)}{2\pi\, \Delta t}, \tag{4.8}$$

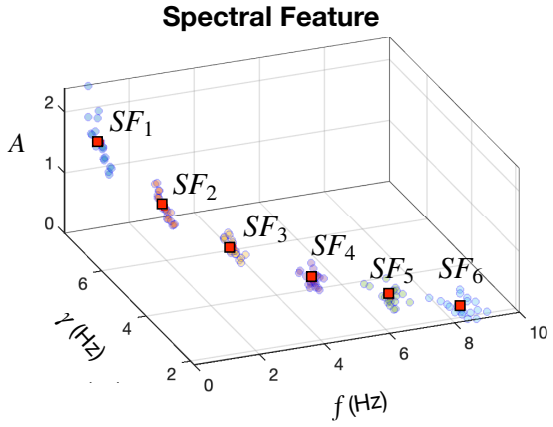


Figure 4.3: The scatter plot of the spectral feature SF_α in 3-dimensional feature space defined by the decay rate γ, the frequency f, and the initial amplitude A, including all the gait cycles of a foot. The red squares represent the centroids of each SF_α points, which will be discussed in Result section.

where we used $x = e^{\ln(x)}$ and $\ln(\lambda) = (\gamma + i2\pi f)\,\Delta t$. The unit of the decay rate γ_α and the frequency f_α are the inverse of second, $\sec^{-1}$, or Hz. And the initial amplitude is defined as

$$A_\alpha = w_\alpha b_\alpha. \tag{4.9}$$

Since the Hankel Matrix is constructed with the time-shifted copies of the measured data, the eigenvectors are not independently determined. Therefore, we can define the spectral features of the dynamical systems in terms of the set of triplets

$$SF_\alpha = (\gamma_\alpha, f_\alpha, A_\alpha),\ \alpha = 1, 2, \cdots, r. \tag{4.10}$$

Figure 4.3 presents the distribution of the spectral features for all the gait cycles from a foot. The clear separation between the sepctral feature at different modes α enables us to analyze each mode of DMD spectral feature independently, providing the deeper understanding how each mode is associated with specific participant's unique walking pattern. The distribution

of the points within each mode (for example, blue circles labled with SF_1) will be discussed in the next chapter regarding to the detection of health condition based on the gait analysis.

The growth rate refers to the positive γ_α, while the decay rate refers to the negative γ_α. Generally, the positive γ_α generates an unstable diverging sequence in time series, while the negative γ_α generates a stable converging sequence. Notably, the plantar pressure signal for each footstep converges to zero at the end of the gait cycle. So, throughout this dissertation, we will use the decay rate terminology.

4.3 Results

This section presents the results of applying our DMD framework to the plantar pressure data in the cancer survivor database. We apply the DMD framework to each individual.

4.3.1 Participants

Participants were 16 women with breast ($n = 3$), ovarian ($n = 6$) or uterine ($n = 7$) cancers ranging in age from 35 to 80 years (mean 62.9±12.6 yrs). They were 6% Black non-Latina, 13% White Latina, 81% White non-Latina. Seven women (44%) reported a diagnosis of diabetes or pre-diabetes, and five women (31%) reported being diagnosed with peripheral neuropathy in the past, including one woman without diabetes but with recurrent cancer. Four women (25%) reported falling in the past year. Self-selected usual gait speeds ranged from 0.71 to 1.17 m/s (mean 0.96±0.15 m/s).

4.3.2 Data Processing

We integrated the spatiotemporal data over the time dimension, producing multiple footsteps in a 2D matrix, and then cropped the region to include each footstep. Once the region was

determined, we trimmed the duration of the gait cycle, producing the time of initial foot contact at $t = 0$. For the analysis of every foot (e.g., right or left), the size of the time frames for each step was extended to 100 frames with zero-padding on the measured data. Notice that this process ensures the convergence of the DMD analysis by enclosing all the eigenvalues within a unit circle in the complex plane.

The temporal pressure profile in each gait cycle was obtained by integrating over the spatial dimension, numerically equivalent to row- and column-wise summation in the 3D tensor. Notice that in Eq. (4.3), we denote it as the sequence in time series $GRF_s(t_c) = D = [D_1, D_2, \cdots, D_{100}]$ with $t = (k-1)\Delta t$ ($k = 1, 2, 3, \cdots, 100$) for the c-th gait cycle of the left ($s = l$) and the right ($s = r$) foots, respectively.

4.3.3 DMD Reconstruction of a Gait Cycle

We obtained the spectral feature of the gait cycle, $SF_\alpha = (\gamma_\alpha, f_\alpha, A_\alpha)$, for all the footstep c with a optimal truncation $r = 12$ in a given tolerance $\epsilon = 10^{-6}N$ with $N = 100$. So the DMD reconstructed GRF, $\hat{D}$, can be expanded with the parameters as, for each gait cycle,

$$\hat{D}_k = \sum_{\alpha=1}^{6} \exp(\gamma_\alpha t)\, \exp(i2\pi f_\alpha t)\, A_\alpha + (\text{complex conjugate}), \quad t = (k-1)\Delta t$$

Figure 4.4 illustrates the reconstruction mechanism in the DMD framework. The eigenvectors w_α and the complex conjugates w_α^* are plotted in the complex plane with the blue and red filled circles, respectively (top row). As time progresses, the DMD modes $e^{\gamma_\alpha t} \times e^{i2\pi f_\alpha t}$ move following the trajectory indicated by the blue lines.

The bottom row of Fig. 4.4 shows how the reconstructed signal depends on the number of modes included in the mode expansion, showing that the the shape of the reconstructed signal approaches to that of the original signal with the more modes added. However, the

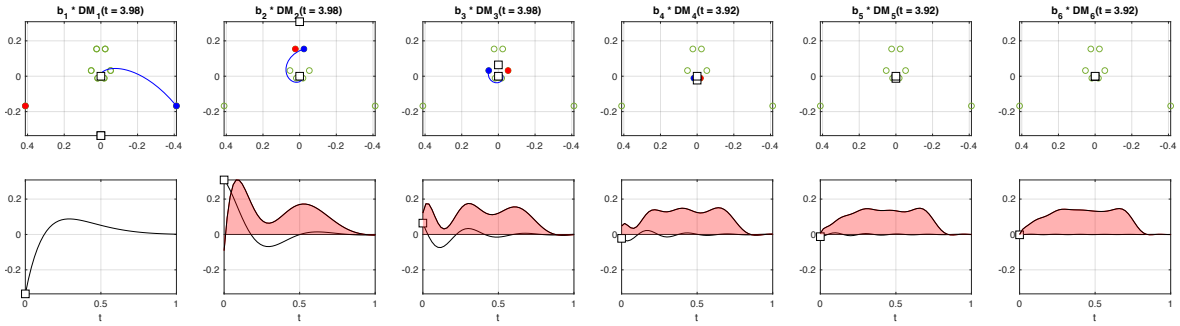


Figure 4.4: The dynamic mode decomposition (DMD) reconstruction. Here, $\mathrm{DM}_\alpha(t) = (\lambda_\alpha^{1/\Delta t})^t\, w_\alpha$ with λ_α defined in Eq. (4.4). So, $b_\alpha \cdot \mathrm{DM}_\alpha(t) = (\lambda_\alpha^{1/\Delta t})^t\, b_\alpha$. The green circles represents the eigenvectors w_α and its complex conjugates at the initial time $t = 0$. Each α-th mode is filled by the blue and red colors. The squares reprents the α-th DMD mode in the state at $t = 0$ and the final state that are connected to the trajectory of the blue line. Reproduced from Seo et al. (2024), licensed under CC BY 4.0.

first two dominant modes already form the temporal profile with peaks at heel-strike and toe-off moments in the gait cycle. Thus, those dominant modes, corresponding to low frequency with larger decay rate, inform us the general walking behavior across all the participants. As we increase the number of modes, the details of the signal, such as the locations of the peaks and the relative heights between them, become closer to the measured data. It implies that thes high frequency with low decay modes are key elements for identifying individual's unique walking pattern, playing an important role in establishing their baseline based on DMD spectral features.

4.3.4 Identifying Plantar Pressure Baseline

The advantage of the DMD framework in this temporal pressure signal is that the small number r of the DMD modes can capture its dynamical patterns in the 3-dimensional feature space of SF_α. Each mode α is separated from the other, allowing for the analysis in the machine-learning algorithms. Taking left footsteps and right footsteps into separate consideration, we can construct the feature space for each side with a collection of the parameter

triplets $SF_{\alpha,s,c}$ with $s = l, r$ and $c = 1, 2, 3, \cdots, C$, where c is indices for the multiple gait cycles and C is the total number of the left and right footsteps, respectively.

We employ the supervised clustering algorithm to find the centroid,

$$\overline{SF}_{\alpha,s} = \frac{1}{C} \sum_{c=1}^{C} SF_{\alpha,s,c}, \tag{4.11}$$

for each DMD mode α and each side of the footstep. Figure 4.5 presents the dynamic modes as the feature points, including all steps before chemotherapy treatment at both normal and fast paces. The figure shows the DMD modes are clustered at each mode and separated from other modes, leading to well-defined centroids. The general patterns that the centroids increase in the projection on (γ,A)-plane and decrease in the projection on (γ, f)-plane were observed across all the participants. However, the specific values of the centroids for each individual in both left and right feet and at different walking paces are distinguishable, enabling one to take the set of centroids as a baseline of each gait motion.

4.3.5 Validation of Approach

Projecting the centroids of DMD modes on the plane consisting of the initial amplitude, A_α and frequency, f_α results in regressions that may be distinguishable per participant. (Figure. 4.6) To validate the DMD mode-based personal identification, we blindly select a patient x and calculate the distance between the centroids of the participants $\left\{\overline{SF}_{\alpha,s}(x_i)\right\}$ labeled by x_i and the randomly selected person labeled by x, whose DMD modes are identified by $\{SF_{\alpha,s,c}(x) | \, \alpha = 1, \cdots, r, \, c = 1, \cdots, C\}$. Here, $s = \{\text{left, right}\}$, c represents the s-side

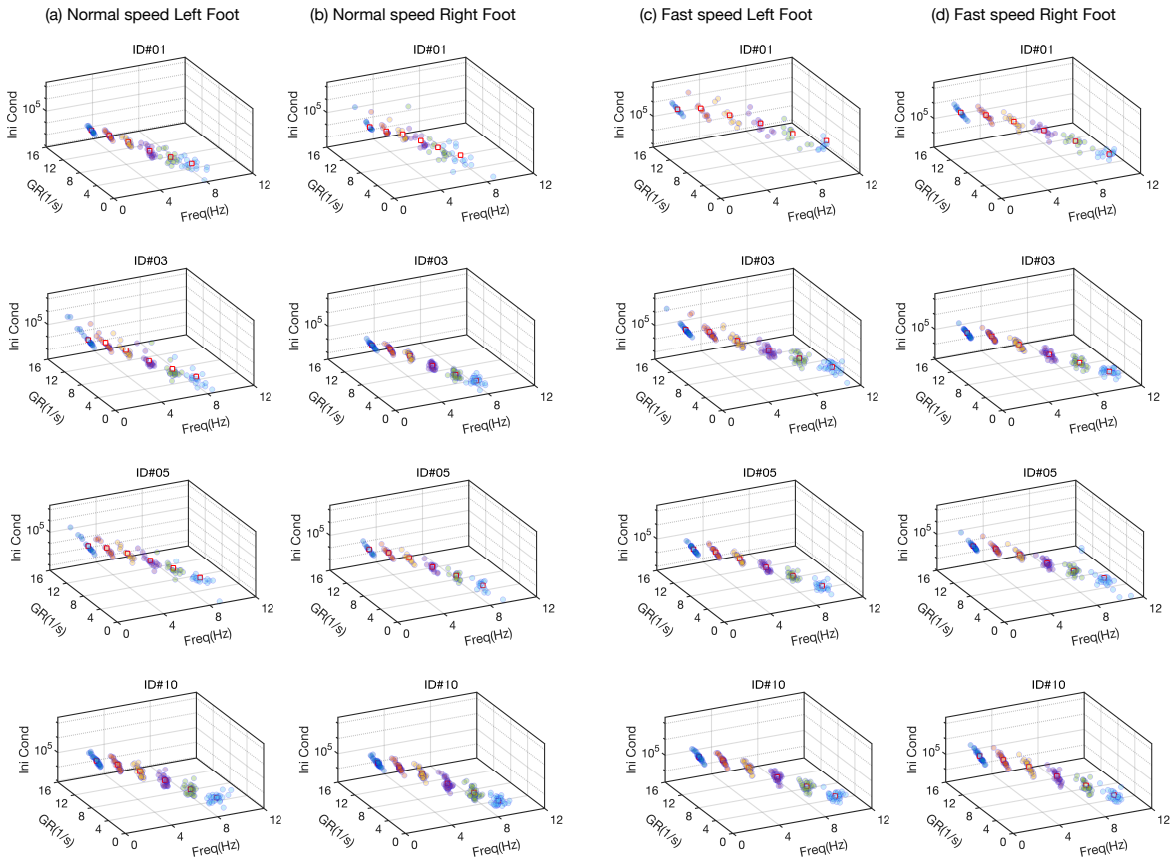


Figure 4.5: Gait baselines of (a) left foot and (b) right foot at normal walking pace and of (c) left foot and (d) right foot at fast walking pace for the participants before the chemotherapy treatment. Freq, GR, and Ini Cond denote frequency f_k, decay rate α_k, and initial condition $s_{0,k}$, respectively. Reproduced from Seo et al. (2024), licensed under CC BY 4.0.

step of the participant x and C is the total number of the s-side steps that are randomly selected. The cost function is defined as a set of Mahalanobis distances as $L_s(x_1, \cdots, x_{10}\,|x) = \{l_s(x_1\,|x), l_s(x_2\,|x), \cdots, l_s(x_{10}\,|x)\}$, where

$$l_s(x_i\,|x) = \sum_{\alpha=1}^{r} \left\| \frac{1}{C_s} \sum_{c=1}^{C_s} S_\alpha(x_i)^T \cdot \left(SF_{\alpha,s,c}(x) - \overline{SF}_{\alpha,s,c}(x_i) \right) \right\|_2 \tag{4.12}$$

with a variance

$$S_\alpha(x_i) = \left[\frac{1}{C-1} \sum_{c=1}^{C} \left(SF_{\alpha,s,c}(x_i) - \overline{SF}_{\alpha,s}(x_i) \right)^2 \right]^{-1/2}. \tag{4.13}$$

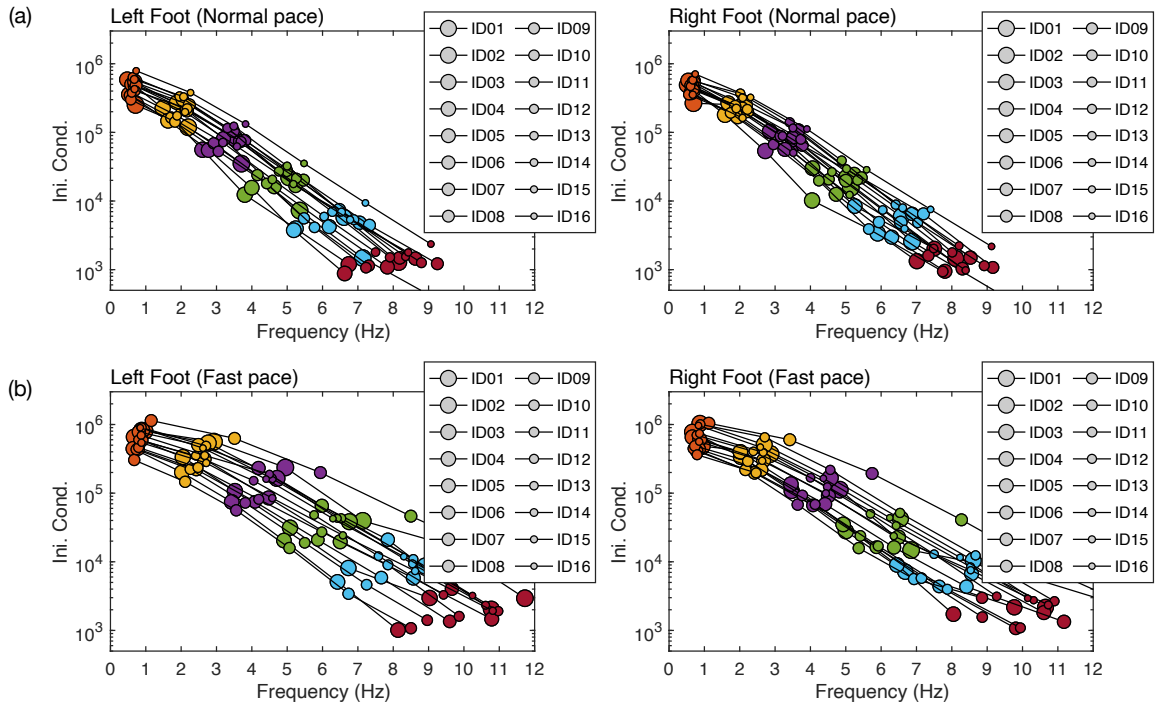


Figure 4.6: Projection of the centroids of the dynamic mode decomposition (DMD) modes on the initial condition against frequency plane for each of 16 participants (a) at normal pace walks and (b) at fast walks. The color of the circles denotes the centroid of the k-th DMD mode (The conjugate modes are not displayed as they have negative frequencies against the same initial conditions). The participants are distinguished by circle sizes. Frequency and Ini Cond denote frequency f_k and initial condition $s_{0,k}$, respectively. Reproduced from Seo et al. (2024), licensed under CC BY 4.0.

We estimated the person's ID, $\hat{x} = \arg\min L_s(x_1, \cdots, x_{10}\,|x)$. We repeated the 100, 000 random selections of 30% sample steps ($C_s = 0.3C$) for the person's ID x under a uniform distribution. The identification accuracies were 89.89% and 88.10% for the left and right footstep samples, respectively, at a self-paced *normal* speed. The identification accuracies for the left and right footstep samples at a self-paced *fast* speed were 86.51% and 88.52%. The high accuracies of the identification rates imply that the centroids of the DMD modes are likely to represent the individual's unique walking baseline, providing the framework for future studies on the health information associated with gait variability away from the baseline.

4.4 Discussion

In this study, we present a novel framework for analyzing gait as a complex spatiotemporal dynamical system using a data-driven model derived from high-resolution plantar pressure sensors embedded in the Tekscan Strideway walkway. We employed dynamic mode decomposition (DMD), a basis modal decomposition technique for infinite-dimensional linear dynamical systems. Unlike conventional methods such as proper orthogonal decomposition (POD) or Fourier transforms, DMD captures key dynamical features using low-rank modal basis functions that are interpretable in physical terms—frequency, decay rate, and initial condition. While developing a universal model of time-evolving plantar pressure distributions remains intractable, we demonstrate that DMD offers a viable, equation-free approach for extracting spectral properties from pressure data with high fidelity.

Because plantar pressure signals are inherently the result of intricate spatial and temporal dynamics, our approach offers deeper insights into the underlying biomechanical processes than traditional gait parameters like step length or width. We propose that DMD-based modeling of pressure dynamics holds promise for capturing clinically relevant indicators of current and future health status, particularly in cancer survivors at risk for chemotherapy-induced peripheral neuropathy (CIPN) at the onset of neurotoxic treatment. In this initial work, we show that a small set of DMD mode parameters can represent an individual's baseline gait pattern. When gait is sufficiently stable, these parameters may characterize both natural step-to-step variability and pathological deviations. However, expanding this analysis to a larger cohort and incorporating additional spatial degrees of freedom will be necessary to develop a higher-order model capable of tracking the onset, progression, and potential recovery from chemotherapy-related neuromuscular toxicity.

We also applied clustering algorithms to group DMD modes and identify representative centroids. These centroids—and the distribution and relationships among contributing

modes—may serve as personalized markers of gait patterns at critical points in the cancer care trajectory. Among the six dominant truncated modes (and their complex conjugates), the mode with the lowest frequency and highest decay rate—along with its conjugate—was consistently associated with stance phase dynamics, as the center of pressure moved from heel to forefoot. This mode exhibited minimal frequency variation but showed substantial variation in initial condition and decay rate across steps, suggesting potential for use in developing a gait asymmetry index. Later stance phase features, including peak pressure timing and magnitude, were primarily captured in the three lowest-frequency modes with the slowest decay rates. Early stance (heel-strike) appeared in higher-frequency modes with faster decay. Although these are preliminary observations, they demonstrate the model's sensitivity to biomechanical events like heel strike and toe-off, and hint at its utility in capturing subtle changes due to chemotherapy-related sensory or strength loss, which are clinically important due to their implications for fall risk.

Despite its simplicity, clustering of DMD modes enabled the development of individualized regression profiles that may serve as a basis for detecting deviations from baseline. Prior work by Feng et al. (2011) demonstrated the use of support vector machines (SVM) on plantar pressure data for subject identification, achieving a recognition rate of 96%—insensitive to weight but dependent on gait speed. In contrast, our approach links identification to interpretable dynamical variables, including gait speed and pressure amplitude. Notably, our cancer survivor dataset includes a greater number of steps and passes per participant compared to previous studies that used only two 6-meter walking trials (Hsieh et al., 2019).

We also acknowledge the importance of age as a factor in future between-subject comparisons, particularly in relation to step-to-step variability in plantar pressure patterns (Franco et al., 2019). A key strength of this work lies in the novel application of a data-driven DMD approach to high-resolution pressure sensor data, offering new possibilities for individualized gait analysis and monitoring in clinical populations.

Our study in this chapter has focused on processing the temporal characteristics of a plantar pressure signal. It does not include the spatial pressure distribution that is embedded in the 2D plantar pressure signal; if processed, it will improve the accuracy of the proposed method. Furthermore, we developed the method using data from a small sample of only 16 patients. Accordingly, the results should be interpreted with caution and will require confirmation with larger datasets.

4.5 Conclusions

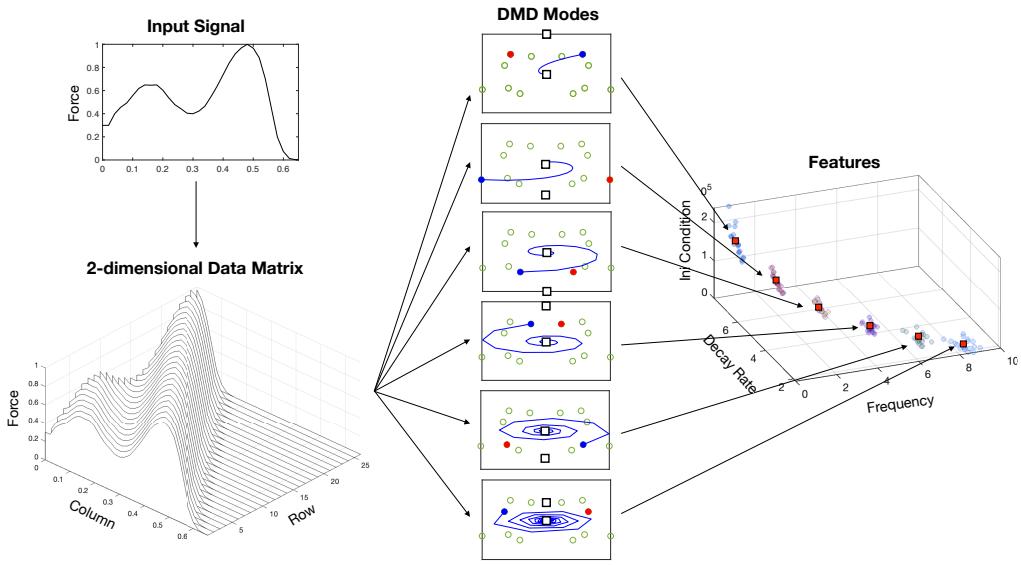


Figure 4.7: Overview of the method for extracting gait features by the dynamic mode decomposition (DMD) on the plantar force input signal. The 2-dimensional data matrix is constructed by stacking the time-delayed coordinates, equivalent to propagation at a constant speed. The DMD modes, the eigenvalues of the Koopman operator, are complex functions flowing along their trajectories. The blue and red circles indicate the eigenvalues at each mode, while the empty green circles indicate all the eigenvalues for comparison. The black squares are the initial and final values that sum up the pair of complex conjugate modes (the red and blue circles). The eigenvalues' frequency, decay rate, and initial condition characterize the gait features. The centroids (red squares) from the multiple sample gait cycles identify the unique feature of the walking activity. Reproduced from Seo et al. (2024), licensed under CC BY 4.0.

This chapter presented our application of Dynamic Mode Decomposition (DMD) to spatiotemporal plantar pressure data collected from women undergoing treatment for breast and gynecological cancers as they walked over a high-resolution sensor-embedded walkway. From these data, we extracted low-rank DMD modes that characterize pressure dynamics during key phases of the stance cycle—phases known to vary among individuals with peripheral neuropathies. This data-driven model captures the temporal evolution of plantar forces unique to each individual's gait. For each step, we quantified features based on the DMD mode parameters: frequency, decay (or growth) rate, and initial condition. We then clustered the steps for each participant to determine representative (central) DMD parameters for each foot. Using this approach, we achieved identification accuracy exceeding 86% across our sample. An overview of the methodology is illustrated in Figure 4.7.

Chapter 5

Detection of Change in Gait Cycle

5.1 Introduction

The human gait, characterized by cyclic phases of stance and swing, typically demonstrates consistent patterns in healthy individuals. However, subtle deviations from these patterns often precede clinically observable mobility defects, indicating neuromuscular pathologies. Among cancer survivors receiving neurotoxic chemotherapy, sensory damage in the feet and reduced ankle strength due to chemotherapy-induced peripheral neuropathy (CIPN) increase fall risk, often before obvious changes in conventional gait metrics like stride length or walking speed (Winters-Stone et al., 2017).

Pressure-sensor walkways offer detailed spatiotemporal representations of footfalls, enabling the detection of subtle early indicators of gait abnormalities. The preceding chapter demonstrated the effectiveness of Dynamic Mode Decomposition (DMD) in analyzing the temporal characteristics of plantar pressure patterns, achieving high-accuracy individual identification based on distinctive spectral features of gait cycles (Seo et al., 2024). However, that framework did not explore how these baseline signatures may change over time or capture systematic longitudinal variations.

This chapter addresses the existing gap by introducing a longitudinal monitoring framework that leverages DMD-derived gait features, characterized by three spectral components; frequency, decay rate, and initial amplitude. By analyzing plantar pressure data from cancer

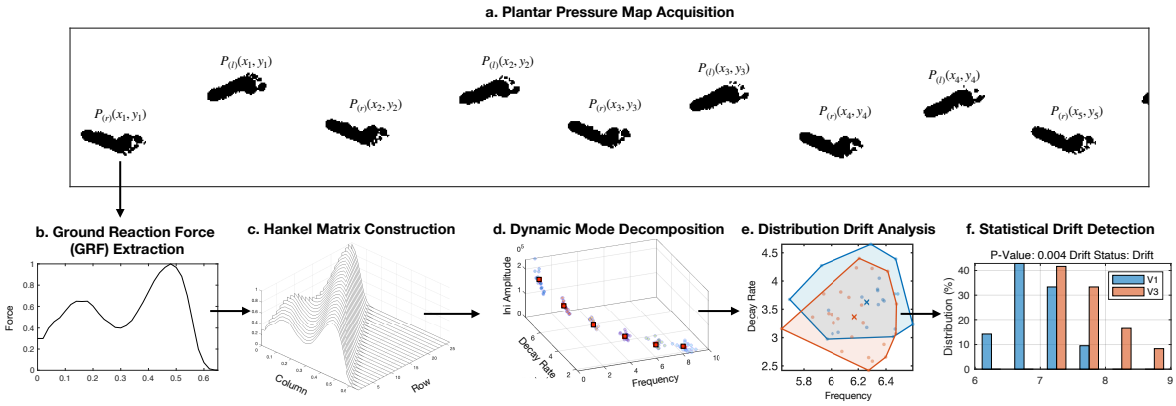


Figure 5.1: Overview of the workflow. a. Spatiotemporal plantar pressure data was collected as participants walk over a Tekscan Strideway mat, capturing ground contact patterns. b. From the pressure data, ground reaction force (GRF) signals are computed for each gait cycle, reflecting the dynamic interaction between foot and floor. c. The GRF time series was embedded into a Hankel matrix, which captures time-shifted signal patterns to serve as input for DMD. d. The Hankel matrix was decomposed into dynamic modes, each characterized by a decay rate, frequency, and initial amplitude–capturing the underlying temporal dynamics of the gait. e. DMD feature distributions from follow-up visits are compared to each individual's baseline, enabling identification of deviations indicative of emerging gait abnormalities. f. Permutation-based tests (Kolmogorov–Smirnov, Anderson–Darling, and Energy distance) assess the significance of distributional changes, with False Discovery Rate control determining the final drift status as Stable, Warning, or Drift.

patients across multiple clinical visits during their chemotherapy treatment, we aim to detect the distributional drift of the DMD spectral feature points relative to each individual's baseline. To detect the these changes, We employed a permutation-based statistical tests, including Kolmogorov-Smirnov (KS), Anderson-Darling (AD), and Energy distance tests. Then, the outcomes were classified using a False Discovery Rate (FDR) approach into three categories: Stable, Warning, or Drift.

Figure 5.1 illustrates an overview of the proposed framework. Panels (a) through (d) show the baseline identification algorithm, consistent with the approach described in the previous chapter. Each spatially integrated pressure signal from the spatiotemporal plantar pressure map, the ground reaction force (GRF), collected across multiple walking passes on the strideway, was transformed into DMD modes. This process produced sets of modal

spectral feature points within a feature space defined by frequency, decay rate, and initial amplitude. As previously discussed, the distribution of these feature points for each DMD mode, along with their central tendency, plays a crucial role in detecting deviations from baseline and assessing treatment-related changes. Panels (e) and (f) of Fig. 5.1 illustrate this procedure. After establishing an individual's baseline distribution using the pre-treatment dataset, we evaluated distributional drift by comparing it with post-treatment datasets collected at later clinical visits.

We validated our approach using longitudinal plantar pressure data from 15 cancer survivors, successfully identifying CIPN-related gait changes in 80% of participants by their second or third follow-up visits. Our algorithm was careful not to raise false alarms when a patient's walking pattern stayed normal.

This study contributes to the existing literature by:

1. Introducing a novel longitudinal monitoring framework based on DMD-derived dynamic gait features.

2. Utilizing a robust multi-metric, permutation-based statistical approach with False Discovery Rate control.

3. Demonstrating comprehensive validation through real-world clinical data from cancer survivors.

4. Establishing a foundation for future applications in real-time, embedded gait monitoring systems for early clinical intervention.

The remainder of this chapter is structured as follows: Section 2 summarizes relevant work in gait analysis and drift detection methodologies. Section 3 details the dataset and DMD-based feature extraction algorithm. Section 4 describes the proposed drift detection framework and presents our experimental results. Section 5 discusses clinical implications before concluding.

5.2 Methods

In the context of longitudinal gait analysis, detecting significant changes in a patient's gait pattern over time is essential for identifying the onset or progression of neurotoxic effects, associated with chemotherapy-induced peripheral neuropathy (CIPN). Given the variability inherent in biological systems and human movement, we utilized a statistical distributional drift detection framework to assess whether the spatiotemporal dynamics of plantar pressure data have shifted meaningfully relative to a pre-treatment baseline.

5.2.1 Data Drift Detection

In our study, we employed MATLAB's `detectdrift` function, part of the Statistics and Machine Learning Toolbox (The MathWorks, Inc., 2024b), to evaluate distributional shifts in gait dynamics following cancer therapy. The `detectdrift` function provides a framework to monitor potential distributional changes in a streaming or temporally segmented dataset. It works by comparing distributions across sliding or fixed windows using a permutation-based hypothesis testing approach. Internally, it performs statistical two-sample tests such as the Kolmogorov–Smirnov (KS) test, energy distance, or Anderson-Darling (AD) test. The function outputs drift decision flags, p-values, and a drift status indicator for each segment.

This capability is particularly useful in our pipeline after constructing a dynamic modal feature space using Dynamic Mode Decomposition (DMD). Once baseline gait features are established, we apply `detectdrift` to subsequent post-treatment gait sessions to detect abnormality onset. The permutation testing nature of the method ensures that the drift detection is robust to small sample sizes and non-parametric variations.

5.2.2 Test statistic metrics

Let $\mathcal{D}_0 = \{SF^{(0)}(t_c)\}_{c=1}^{C_0}$ define the baseline distribution of DMD spectral features collected prior to chemotherapy. Then, $\mathcal{D}_t = \{SF^{(t)}(t_c)\}_{c=1}^{C_t}$ represents the same patient's gait data at a subsequent visit labeled by t. Here, C_0 and C_t are the total numbers of the steps associated with the gait cycle before and after the treatments, respectively. The main task is to test the null hypothesis

$$H_0 : \mathcal{D}_t \sim \mathcal{D}_0 \tag{5.1}$$

against the alternative

$$H_1 : \mathcal{D}_t \nsim \mathcal{D}_0, \tag{5.2}$$

where the distribution of the DMD features drift signficantly from that of baseline features.

In this dissertation, we evaluate H_0 using non-parametric thwo-sample tests. These methods are robust to unknown data distribution and particularly suitable for moderate-sized, high-dimensional samples like those derived from DMD.

- Kolmogorov-Smirnov (KS) Test: Compares emprirical cumulative distribution function (ECDFs) of two samples in one dimension, whose test statistic is

  $$d = \sup_x ||F_0(x) - F_t(x)||,$$

  where F_0 and F_t are ECDFs of $\mathcal{D}_0$ and $\mathcal{D}_t$. KS is applied separately to each feature dimension (Massey, 1951). The MATLAB built-in function `kstest` computes the critical value using an approximate formula or by interpolatiion in table, returning a test decision for the null hypothesis at the 5% significance level with 1 for the rejection or 0 otherwise, p-value of the hypothesis test, and the approximate critical value of the test.

- Anderson-Darling (AD) Test: Extends the KS test by weighting descrepancies in the tails more heavily.

$$d = \sum_x \left(\frac{\|F_0(x) - F_t(x)\|}{\sqrt{(C_0 + C_t)D(x)(1 - D(x))}} \right)^2$$

with $D(x)$ being the joint ECDF of all data. This improves sensitivity in clinical contexts where abnormal gait changes often menifest in rare step behaviors. The MATLAB built-in function `adtest` decides to reject or not reject the hypothesis based on comparing the p-value for the hypothesis test with the specified significance level, returning the test decision, p-value, and the critical value of the test.

- Energy Distance Test: Employed for multivariate data that quantifies the difference between two distributions using pairwise distance (Székely and Rizzo, 2013)

$$d = \sqrt{2 \sum_x \|F_0(x) - F_t(x)\|^2}$$

5.2.3 Permutation Testing in Drift Detection

The `detectdrift` function in MATLAB employs *permutation testing* (Good, 2000) as its core methodology to determine whether a given variable has experienced statistical drift between the baseline and target datasets. Permutation testing is a non-parametric hypothesis testing technique used to evaluate the significance of a test statistic by simulating its distribution under the null hypothesis.

Specifically, `detectdrift` simulates the null hypothesis (no drift) by repeatedly permuting the observations of the selected variable across the two groups and recomputes the test statistic for each permutation. This process yields an empirical distribution of the test statistic under the assumption that the two groups originate from the same distribution.

If the original, unpermuted test statistic lies well within this empirical distribution, then it is not considered statistically significant, and the algorithm fails to reject the null hypothesis,

indicating no evidence of drift. Conversely, if the original test statistic is an outlier relative to the permutation-based null distribution, `detectdrift` rejects the null hypothesis and concludes that drift has occurred in that variable.

Let $[L_{\text{CI}}, U_{\text{CI}}]$ denote the lower and upper bounds of the confidence interval (CI) for the p-value, and let θ_D and θ_W denote the drift and warning thresholds, respectively, where typically $\theta_D < \theta_W$ (for example, $\theta_D = 0.01$, $\theta_W = 0.05$). Then, the decision logic is as follows:

- Drift: Detected if $U_{\text{CI}} < \theta_D$. This implies strong evidence that the variable has changed distribution between the baseline and target data.

- Warning: Detected if the confidence interval partially overlaps with the warning region, i.e.,

$$\theta_D < L_{\text{CI}} < \theta_W \quad \text{or} \quad \theta_D < U_{\text{CI}} < \theta_W.$$

  This suggests a possible drift but with uncertainty, warranting further observation.

- Stable: Assigned if $L_{\text{CI}} > \theta_W$, indicating strong evidence that no significant drift has occurred.

This tiered classification provides an interpretation of drift detection results, which is essential in real-world applications where hard decision boundaries may not capture gradual or emerging changes. In our analysis, these classifications help prioritize which gait variables or modes require closer monitoring over time. Thus, this method is especially suited to the analysis of high-dimensional, nonstationary biological data such as gait dynamics, where model-free statistical rigor is desired. It supports our framework by quantifying shifts in gait features across sessions with high sensitivity and low model bias.

5.3 Results

This section presents the outcomes of our proposed method for detecting gait abnormalities through shifts in the distribution of DMD mode spectral features over multiple clinic visits. Unlike our previous framework (Seo et al., 2024), which used centroids of the samples of DMD feature points in feature space, particularly for individual identification, the current framework focuses on detecting longitudinal distributional shifts in these DMD modal spectral feature points using statistical hypothesis testing.

5.3.1 Distribution of DMD spectral features

We used the same dataset described in Seo et al. (2024) from 15 cancer survivors across multiple clinical visits. Per visit, each participant completed 4-–8 passes over the strideway in alternating walking directions, generating approximately 8—10 footsteps per pass. To ensure consistent left/right foot orientation, pressure maps from alternating directions were flipped as necessary, aligning left-foot contact with the upper half of the frame and right-foot contact with the lower half.

For each footstep, we extracted 12 DMD modes spectral features, $FS_{\alpha,c} = (\gamma_{\alpha,c}, f_{\alpha,c}, A_{\alpha,c})$ from the dataset from the subsequent clinical visits. The number of the modes $r = 12$ was selected to ensure reconstruction error below a tolerance at $\epsilon = 10^{-6}$ with $N = 100$ time samples. Separate distributions of the DMD modal feature points in feature space were constructed for left and right feet. Figure 5.2 illustrates a representative distribution of these features in the 2D space of decay rate γ vs. frequency f.

5.3.2 Drift Detection via Permutation Testing

To assess whether gait dynamics changed over time, we compared the distribution of each DMD feature between a participant's baseline visit and their follow-up visits. For each mode

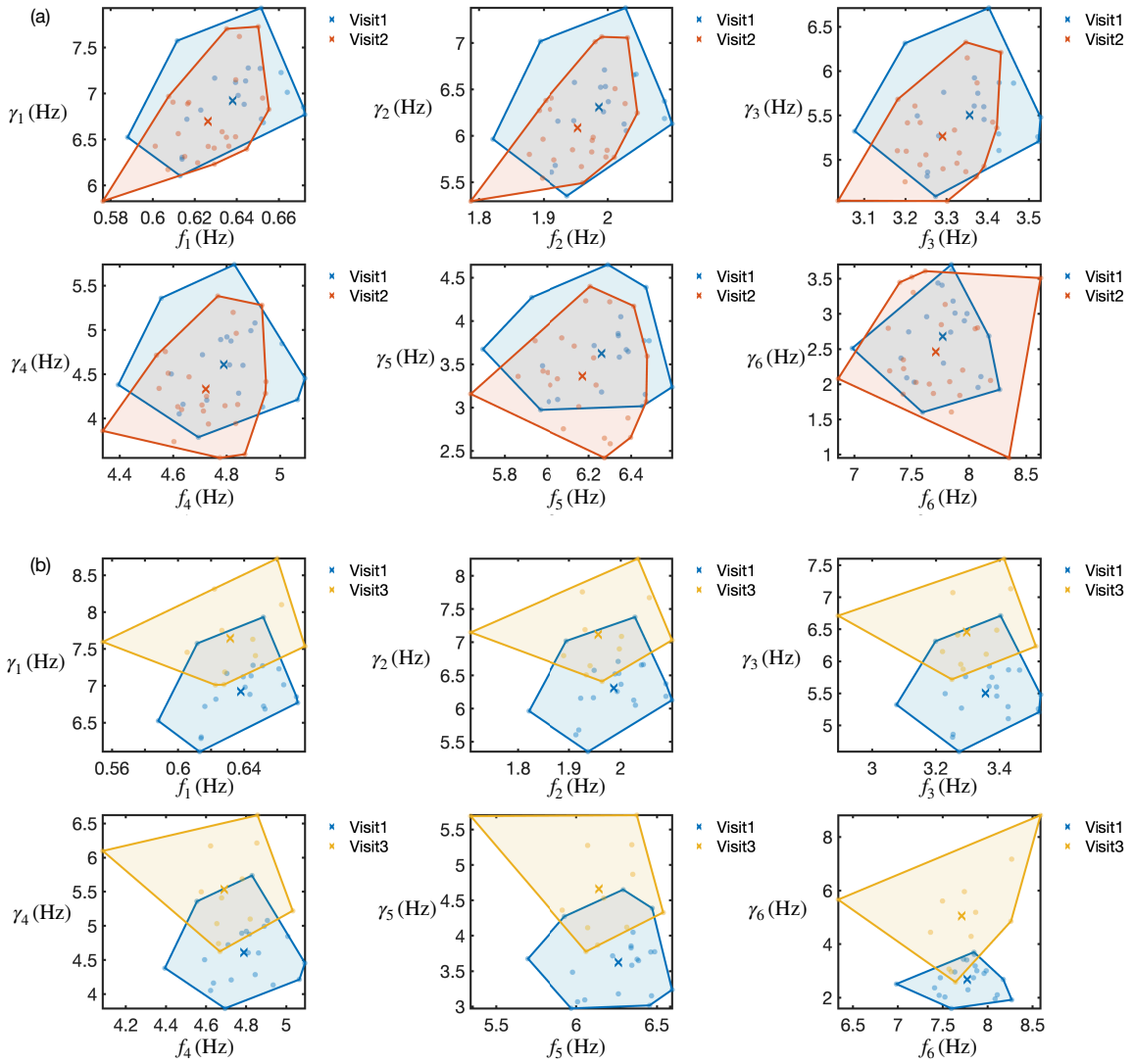


Figure 5.2: (a) The overlap between the distribution of the features derived from DMD eigenvalues, γ_α vs. f_α for $\alpha = 1, 2, \cdots, 6$ of the baseline footsteps (blue dots) at V_b (Visit1) and that of the target footsteps (red dots) at V_2 and (b) the overlap between the footsteps at V_b and footsteps (yellow dots) at V_3. For visual clarity, each distribution of the points is enclosed by the same colored lines, and the corresponding centroid of each DMD α^{th} mode is indicated by a cross with blue for V_b, red for V_2, and yellow for V_3, respectively. Notice that the target data at V_2 remains the same as the baseline data with slight shifts of the centroids. In contrast, the centroids at V_3 are shifted away from the baseline boundaries, indicating a change in the gait pattern.

k and each feature component $(\lambda_k, f_k, |s_{0k}|, \angle s_{0k})$, we conducted a permutation test to evaluate the null hypothesis: the feature distributions in the baseline and target visits are identical.

We used permutation testing to determine the drift status for each variable (4 features per mode for 6 modes) in the baseline data (obtained during the first visit) and its counterpart in the target data (obtained during subsequent visits). A permutation test is a nonparametric statistical significance test in which the distribution of a metric (test statistic) under a null hypothesis (no drift) is obtained by computing the values of that metric under all possible rearrangements of a variable in Baseline and Target. Trying all possible variable permutations might be infeasible depending on the number of variables and footsteps. Therefore, a sufficient number of permutations are required to obtain a reasonable metric estimate for the selected variable. Under the null hypothesis, many values of the metric recorded during permutation testing can be as extreme as the initial test statistic. This suggests sufficiently high confidence that the footsteps of the specified variable in the baseline and target data come from the same distribution. Therefore, no evidence of drift is found, and it fails to reject the null hypothesis.

If the initial test statistic is identified as an outlier, it rejects the null hypothesis. This suggests sufficiently high confidence that the footsteps of the specified variable in the baseline and target data come from different distributions. Therefore, drift is detected.

The steps in permutation testing are the following:

- Compute an initial metric value from the original data for the DMD k^{th} mode of C_0 footsteps in the baseline data and C_t footsteps in the target data.

- Permute the footsteps of the k^{th} mode of the baseline data and the target data. Then, separate them into two vectors with sizes C_0 and C_t, respectively. Then, compute the same metric value. Repeat this process to obtain a distribution of the specified metric.

- Fitting x, the number of times a metric value obtained from a permutation is greater than the value of the initial metric value, to a binomial distribution, estimate the 95%

confidence interval in a range of a lower bound L_{CI} and an upper bound U_{CI} for the p-value: $p = x/N_{\text{perm}}$, where N_{perm} is the number of permutations.

- With the drift threshold θ_D and the warning threshold $\theta_W > \theta_D$, determine the drift status based on the following conditions: (1) Drift status when $U_{\text{CI}} < \theta_D$, (2) Stable status when Lower $L_{\text{CI}} > \theta_W$, and (3) Warning status for otherwise.

We used the (1) Kolmogorov-Smirnov metric, (2) Anderson-Darling metric, and (3) Energy metric as a test statistic in permutation testing for detecting drift between the baseline and target visit. The following tables summarize the 15 participants' DMD feature drift results.

5.3.3 Longitudinal Feature Drift

To illustrate the statistical change over time for the particular subject, figure 5.2 shows the projection of the DMD features of the footstep samples onto the DMD eigenvalues, frequency f_α and decay rate λ_α for $\alpha = 1, 2, \cdots, 6$, for the baseline visit V_b (labeled by Visit1) and the target visit (V_t) at (a) $t = 2$ (Visit2) and (b) $t = 3$ (Visit3). At visit 3, significant centroids and feature point distribution changes occurred and are shown. On the other hand, at visit 2, a minor shift between centroids and major feature points overlaps, which could be attributed to individual variability. To quantify the drift in the data distribution, we employed the detect drift built-in MATLAB function, which determines the presence of the drift between two distributions based on the given drift threshold $\theta_D = 0.05$, the warning threshold $\theta_W = 0.1$ and p-values using the permutation testing.

Figure 5.3 (a) shows the histograms of the 6 modes of the DMD features, frequencies f_α, decay rates λ_α, and the magnitudes $|A_\alpha|$ and the phase $\angle A_\alpha$ of the initial conditions. We used a 95% confidence interval with a warning threshold θ_W and a drift threshold θ_D of 0.1 and 0.05, respectively. For the overall decision on the presence of the drift, we utilized the FDR method for multiple hypothesis tests, based on the Bejamini-Hocchberg procedure, to compute the

Table 5.1: Kolmogorov-Smirnov metric, $\sup_x |E_b(x) - E_t(x)|$, where $E_b(x)$ and $E_t(x)$ are the empirical cumulative distribution function for the baseline visit (V_b) and the target visit labeled by V_t. In the 3rd through 6th columns, only drift-detected modes are presented. Here, f_α is the frequency, λ_α is the decay rate, $|A_\alpha|$ and $\angle A_\alpha$ are the magnitude and the phase of the initial condition of the DMD features ($\alpha = 1, 2, \cdots, 6$). In the last column, the overall drift status is presented.

Subject	V_t	f_α	λ_α	$\|A_\alpha\|$	$\angle A_\alpha$	Overall
1	t =2		k =3,4,5,6	k =2,3,5		W
2	t =3					S
3	t =2	k =1,2,3,4,5,6	k =1	k =6	k =1	D
4	t =2	k =1,2,3,4,5,6			k =1,2,3	D
5	t =2	k =1,2,3,4,5,6	k =4		k =2,3,5	D
6						S
7	t =2	k =1,2,3,4,5,6			k =3,4,6	D
8	t =2	k =1,2,3,4,5,6			k =3,4,6	D
9	t =2	k =1,2,3,4,5,6	k =1,2,3,4,5,6	k =4,5	k =1,2,3,4	D
10	t =3		k =1,2,3,4,5,6	k =1,2,3,4,5,6	k =1,2	D
11	t =2	k =1,2,3,4,5,6				D
12	t =3	k =1,2,3	k =1,2,3,4,5,6			D
13	t =2	k =1,2,3,4,5,6			k =3,4,5	D
14	t =2	k =1,2,3,4,5,6			k =3,4,5	D
15	t =2	k =1,2,3,4,5,6			k =1	D

FDR. Using 24 variables (f_α, λ_α, $|A_\alpha|$ and $\angle A_\alpha$ for $\alpha = 1, 2, \cdots, 6$), the algorithm ranks the p-values corresponding to the variables, divides the ranks by the number of variables $N_\alpha = 24$ to obtain $q = [1/N_\alpha, 2/N_\alpha, \cdots, k/N_\alpha]$, modifies the warning and drift thresholds for each sorted p-value by multiplying the initial warning and drift threshold values by the corresponding q value, then checks if any sorted p-values are smaller than the corresponding modified warning and drift thresholds to determine the drift status.

Table 5.2: Energy metric, $\sqrt{2\sum_x |E_b(x) - E_t(x)|^2}$. As in TABLE 5.1, only drift-detected modes are presented in the 3rd through 6th columns, and in the last column, the overall drift status is presented. Notice that the overall drift status and the modes of the DMD features are identical to the results in TABLE 5.1.

Subject	V_t	f_α	λ_α	$\lvert A_\alpha \rvert$	$\angle A_\alpha$	Overall
1	t =2		k =1,2,3,4,5,6	k =2,3,5,6		D
2	t =3	k =2				S
3	t =2	k =1,2,3,4,5,6	k =1,2,3,4,6	k =6		D
4	t =2	k =1,2,3,4,5,6	k =1		k =1,2,3	D
5	t =2	k =1,2,3,4,5	k =4	k =4	k =1,2,4	D
6	t =2					S
7	t =2	k =1,2,3,4,5,6		k =3	k =2,3,4,6	D
8	t =2	k =1,2,3,4,5,6		k =3	k =2,3,4	D
9	t =2	k =1,2,3,4,5,6	k =1,2,3,4,5,6	k =5	k =1,2,3,4	D
10	t =3		k =1,2,3,4,5,6	k =1,2,3,4,5,6	k =1,2	D
11	t =2	k =1,2,3,4,5,6				D
12	t =3	k =1,2,3,4	k =1,2,3,4,5,6			D
13	t =2	k =1,2,3,4,5,6			k =3,4,5	D
14	t =2	k =1,2,3,4,5,6			k =3,4,5	D
15	t =2	k =1,2,3,4,5,6	k =1,2		k =1	D

5.3.4 Summary of Drift Detection Across Participants

TABLE 5.1 present the results for all participants who demonstrated walking disorder due to chemotherapy-induced neuropathy using the Kolmogorov-Smirnov metric $\sup_x ||E_b(x) - E_t(x)||$, where $E_b(x)$ and $E_t(x)$ are the empirical cumulative distribution functions (ECDF) of the baseline visit (V_b) and the target visit (V_t) over the common domain, respectively. We presented the drift-detected visit associated with the onset of the disorder for each patient. The drift-detected mode numbers k of the DMD features (frequency f_α, decay rate λ_α,

Table 5.3: Anderson-Darling, $\sum_x \left(\frac{|E_b(x)-E_t(x)|}{\sqrt{(m+n)D(x)(1-D(x))}} \right)^2$ with m and n are the number of footsteps in the baseline visit (V_b) and target visit (V_t), respectively and $D(x)$ is the joint ECDF of all footstep samples. As in TABLE 5.1 and TABLE 5.2, in the 3rd through 6th columns, only drift-detected modes are presented, and in the last column, the overall drift status is presented. Notice that the drift status for subject 1 is detected as drift from warning status and subject 6 from stable to warning status.

Subject	V_t	f_α	λ_α	$\|A_\alpha\|$	$\angle A_\alpha$	Overall
1	t =2		k =1,2,3,4,5,6	k =1,2,3,4,5,6		D
2	t =3				k =2,5	W
3	t =2	k =1,2,3,4,5,6	k =1,3,4,6	k =6	k =1	D
4	t =2	k =1,2,3,4,5,6	k =1		k =1,2,3	D
5	t =2	k =1,2,3,4,5	k =4	k =4	k =2,3,5	D
6	t =3				k =1	W
7	t =2	k =1,2,3,4,5,6		k =3	k =2,3,4,6	D
8	t =2	k =1,2,3,4,5,6		k =3	k =2,3,4,6	D
9	t =2	k =1,2,3,4,5,6	k =1,2,3,4,5,6	k =4,5	k =1,2,3,4,6	D
10	t =3		k =1,2,3,4,5,6	k =1,2,3,4,5,6	k =1,2	D
11	t =2	k =1,2,3,4,5,6				D
12	t =3	k =1,2,3	k =1,2,3,4,5,6			D
13	t =2	k =1,2,3,4,5,6			k =3,4,5	D
14	t =2	k =1,2,3,4,5,6		k =1,4,5,6	k =2,5,6	D
15	t =2	k =1,2,3,4,5,6	k =1,2		k =1	D

magnitude of the initial condition $|A_\alpha|$ and phase of the initial amplitude $\angle A_\alpha$) were shifted from the baseline established at visit 1. Notice that most of the overall decisions are based on the shifts in the frequency modes, while the decay rate and the magnitude of the initial conditions are the second decision-makers.

TABLE 5.2 and TABLE 5.3 are the results using the Energy metric $\sqrt{2\sum_x |E_b(x) - E_t(x)|^2}$ and the Angerson-Darling metric $\sum_x \left(\frac{|E_b(x)-E_t(x)|}{\sqrt{(C_b+C_t)D(x)(1-D(x))}} \right)^2$ with C_b and C_t are the number of

footsteps at baseline and target visits, respectively. Here, $D(x)$ is the joint ECDF of all data. Notice that the results with the Energy metric are identical in the overall drift status and the drift-detected DMD modes α. On the other hand, the Anderson-Darling metric and while using the same thresholds $\theta_D = 0.05$ and $\theta_W = 0.1$, the drift statuses for subject 1 and subject 6 have changed labels from the warning to the drift and from the stable to the warning status, respectively. It is due to the fact that the Anderson-Darling metric, as compared to the Kolmogorov-Smirnov and the energy metrics, places more weight on observations at the tails of the distribution, thus lowering p-values.

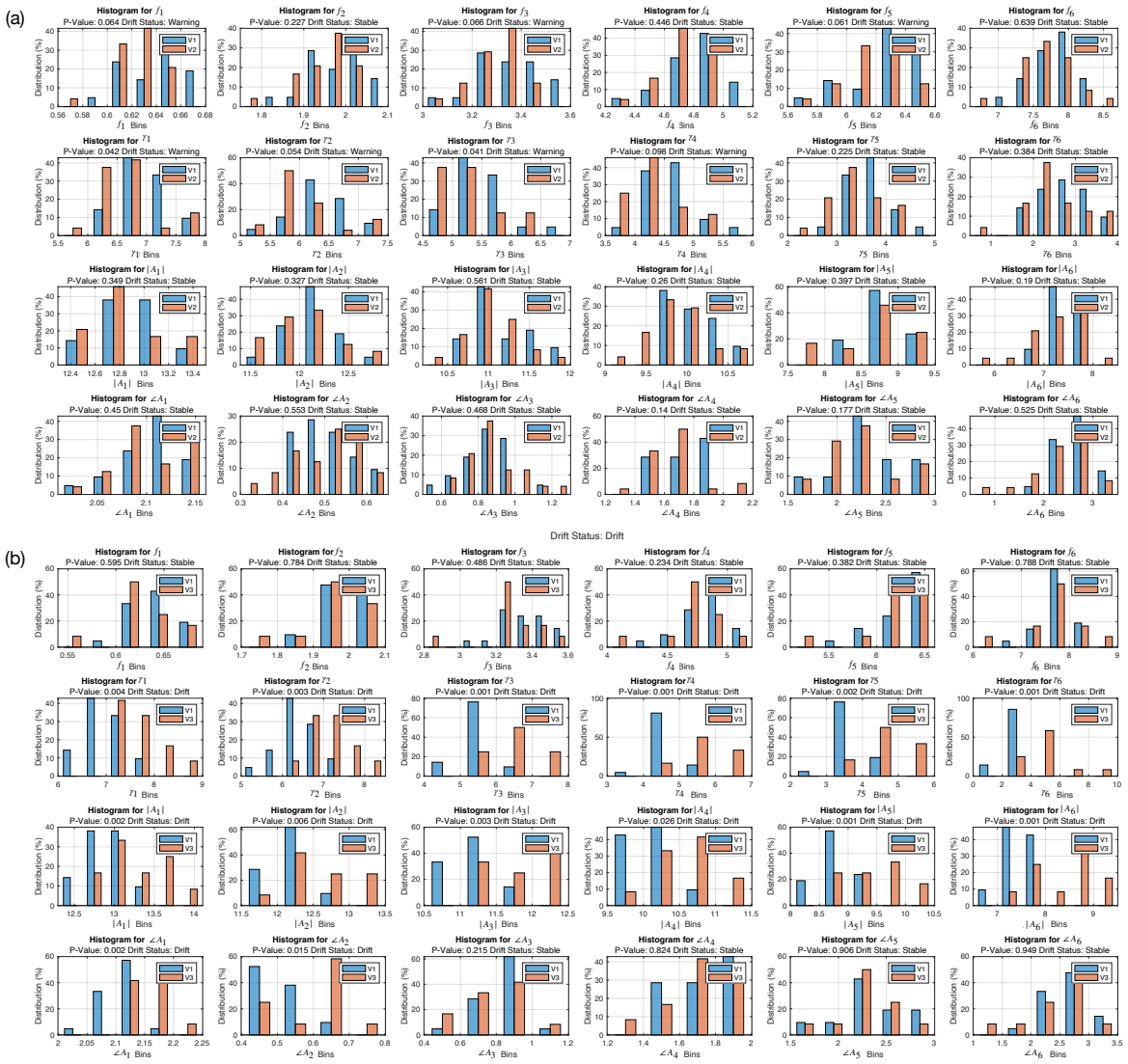


Figure 5.3: The histograms show the drift status for the DMD features (frequency f_α, decay rate λ_α, magnitude of the initial condition $|A_\alpha|$ and its phase $\angle A_\alpha$ for $\alpha = 1, 2, \cdots, 6$) at (a) the target visit V_2 and (b) the target visit V_3 against the baseline at the baseline visit (V_1) for the same participant presented in Figure 5.2 using Kolmogorov-Smirnov metric, $\sup_x |E_b(x) - E_t(x)|$, where $E_b(x)$ and $E_t(x)$ are the empirical cumulative distribution function for V_1 and V_t for $t = 2, 3$. p-values are obtained by the number of times a metric value from a permutation exceeds the initial metric value. The drift status of each DMD feature is determined by the conditions mentioned in the text with the drift threshold $\theta_D = 0.05$ and the warning threshold $\theta_W = 0.1$.

5.4 Discussion

This study extends our previous work on DMD-based gait signatures (Seo et al., 2024) by introducing a novel longitudinal framework capable of detecting emerging gait abnormalities through distributional drift in plantar-pressure features.

By comparing post-treatment gait dynamics to individualized pre-treatment baselines, the proposed permutation-based detection algorithm identified significant deviations in 12 out of 15 cancer patients (80% sensitivity). Importantly, the method maintained high specificity, correctly identifying the three participants whose gait remained stable across clinic visits.

Chemotherapy-induced peripheral neuropathy (CIPN) often manifests as subtle sensory degradation before any observable change in conventional gait metrics such as stride length or cadence. Because of this, CIPN is frequently underdiagnosed. The current approach offers a sensitive and interpretable alternative: instead of focusing on averaged spatiotemporal metrics, it detects foot-loading patterns during the stance phase of the gait cycle. The physical interpretability of DMD-derived features (decay rate, frequency, and initial amplitude) enhances their clinical utility. For example, changes in modal decay rates may reflect altered temporal coordination or fatigue-related instability, while frequency shifts may signal disruptions in motor control. The low-dimensional feature representation also enables robust detection with a limited number of gait cycles (typically 40–50 steps), making the method suitable for outpatient settings.

Although this study utilized the Tekscan Strideway mat, the framework is applicable to the sensing modality. Because it relies on time-resolved plantar-pressure images, it is readily adaptable to other platforms, including in-shoe sensors or wearable pressure-mapping systems. Furthermore, the algorithm's computational efficiency makes it feasible for real-time feedback applications. A single–session gait assessment could immediately flag patients

showing early signs of CIPN-related gait disorder, enabling timely referrals to physical therapy or adjustments in chemotherapy regimens.

Several limitations merit discussion. First, the sample was restricted to 15 female participants from a single clinical site, potentially limiting generalizability across sexes, pathologies, and healthcare environments. Additionally, gait assessments were performed under controlled indoor conditions, which may not capture the irregular terrains encountered in daily life. Second, while the current method distinguishes between Stable, Warning, and Drift statuses, it does not quantify the severity of gait abnormality. Future work will focus on mapping drift magnitude to clinical CIPN severity scales and linking feature changes to fall risk or functional decline. Third, although the reconstruction accuracy of the DMD model was high with 12 modes, higher-order pathological phenomena may require expanded modal resolution or multiresolution DMD techniques.

In addition to distributional drift detection, various anomaly detection techniques may enhance the interpretation of spatiotemporal gait features derived from DMD (Chandola et al., 2009). For more deeper understanding, machine learning-based approaches like one-class support vector machines (SVMs) (Schölkopf et al., 2001) or isolation forests (Liu et al., 2008) offer flexible, non-parametric alternatives, enabling learning individualized gait boundaries and flagging outlier steps. These models are particularly suited for detecting subtle deviations that may not immediately shift global distributions but still indicate early dysfunction.

Time-series-specific methods such as dynamic time warping (DTW) (Berndt and Clifford, 1994) or change point detection (CPD) (Truong et al., 2020) can also be leveraged to identify anomalies in gait temporal evolution across visits. Additionally, local density-based techniques such as the Local Outlier Factor (LOF) (Breunig et al., 2000) are valuable for capturing step-to-step variability, which may be especially relevant for detecting asymmetry or irregular transitions during stance and push-off phases. Integrating these anomaly detection

frameworks with DMD mode analysis offers a promising direction for early identification of clinically meaningful gait changes—particularly in longitudinal studies of cancer survivors undergoing neurotoxic chemotherapy.

5.5 Conclusion

This study introduced a novel, data-driven framework for detecting gait abnormalities through dynamic analysis of plantar pressure signals using Dynamic Mode Decomposition (DMD). By transforming each footstep into a compact set of modal features—-frequency, decay rate, and initial amplitude—-we enabled individualized, interpretable tracking of gait patterns over time.

To detect deviations from an individual's pre-treatment gait baseline, we employed a permutation-based drift detection approach using multiple statistical tests (Kolmogorov–Smirnov, Anderson–Darling, and Energy metrics), with False Discovery Rate (FDR) control to ensure robustness against multiple comparisons. The framework achieved 80% sensitivity in identifying chemotherapy-induced gait changes among cancer survivors.

Compared to traditional gait metrics, this method provides physical insight and earlier detection by identifying subtle redistributions in plantar pressure during stance. It operates on low-dimensional feature spaces and requires a relatively small sample size of gait cycles, making it suitable for outpatient monitoring or integration into wearable devices.

Beyond the scope of chemotherapy-induced peripheral neuropathy (CIPN), the proposed framework is broadly applicable to other progressive conditions affecting gait, such as diabetic neuropathy, Parkinsonian gait, and post-operative rehabilitation. Its computationally efficient design supports deployment in both clinical walkways and mobile, real-world monitoring environments.

Chapter 6

Beyond the Floor: Toward Multisensor Fusion for Gait Analysis

6.1 Introduction

Human gait is a sequential, alternating movement of the lower limbs, involving both legs and feet in the forward propulsion of the body's center of gravity. Plantar pressure represents the spatiotemporal distribution of contact forces between the foot and the ground. These pressure maps form a unique biomechanical signature that can identify individuals with high accuracy (Razak et al., 2012) and detect early signs of gait abnormalities (Feng et al., 2011). In parallel, the consistent patterns in body segment acceleration and orientation—captured through inertial measurements—are fundamental to maintaining stable and functional gait.

Variability in gait patterns often reflects differences in limb motion, walking speed, or asymmetric force distribution. In clinical populations, particularly adults undergoing neurotoxic chemotherapy, slower walking speeds and disrupted gait patterns are linked to an increased risk of falls. Gait, therefore, serves as a rich physiological signal closely tied to one's health status.

This chapter marks the final chapter of this dissertation and serves as a bridge to future work. Here, I reflect on an ongoing but unfinished project that extends the current pressure-based analysis by integrating data from inertial measurement units (IMUs). This fusion of Tekscan and IMU data holds promise for a more comprehensive and dynamic model of

human gait. While the analysis is still in progress, this sensor fusion framework represents a significant step forward—offering a more holistic and precise view of gait behavior that can ultimately enhance early detection of dysfunction and support longitudinal monitoring in clinical populations. This direction, while still under development, opens new possibilities beyond the floor.

In recent years, many technological advancements have emerged for monitoring gait patterns in various aspects with highly accurate and affordable manners, including from developing personalized footwear (Shorten, 2009) to assessing foot health (Mann et al., 2016; Springer et al., 2016). Plantar pressure measurements from sensor-embedded walkways allow spatiotemporal examination for consistencies and variations with high accuracy. Inertial measurement unit (IMU) devices measure acceleration, angular velocity, and magnetic field to monitor the movements of different body segments. However, there's been no analysis algorithm developed to model a comprehensive gait patterns using the multimodal data incorporating the plantar pressure and the resultant kinematic variables such as the acceleration and angular velocity at the same time.

The gait cycle consists of two major phases: the stance phase and the swing phase. The stance phase consists of a sequence of events as shown in Figure 1.1: (a) heel-strike, (b) foot-flat, (c) heel-off of the lead leg, (d) contralateral heel-strike, and (e) toe-push-off of the lead leg. Followed by toe-push-off, the lead foot enters the swing phase with no contact force on the ground.

The double stance is defined as the duration when both feet are in contact with the ground. The lead foot is in heel-strike to foot-flat, and the contralateral foot is in toe-push-off. The gait parameters are divided into two categories: kinematic variables, such as stride length and speed, and kinetic variables, such as plantar pressure and spatially integrated impulse (GRF). Figure 6.1(a) through (d) illustrates the sequential gait cycles, starting with a double stance with one foot in thrust (toe-push-off) and another in the pivot (heel-strike). Then, the

pushed-up foot begins to swing about the pelvis while the torso swings forward about the foot in contact with the ground. Figure 6.1(e) plots the spatially integrated plantar pressure (GRF) of both feet, superimposed with the magnitude of the acceleration of the body segment (label 1). The plantar pressures from both feet overlap during the double stance. Then, the swing foot's acceleration and the grounded foot's plantar pressure show similar traces.

This chapter presents a framework potential to provide a unified perspective for a unique gait feature in terms of kinematic (GRF) and kinetic variables (accelerations and angular velocities).

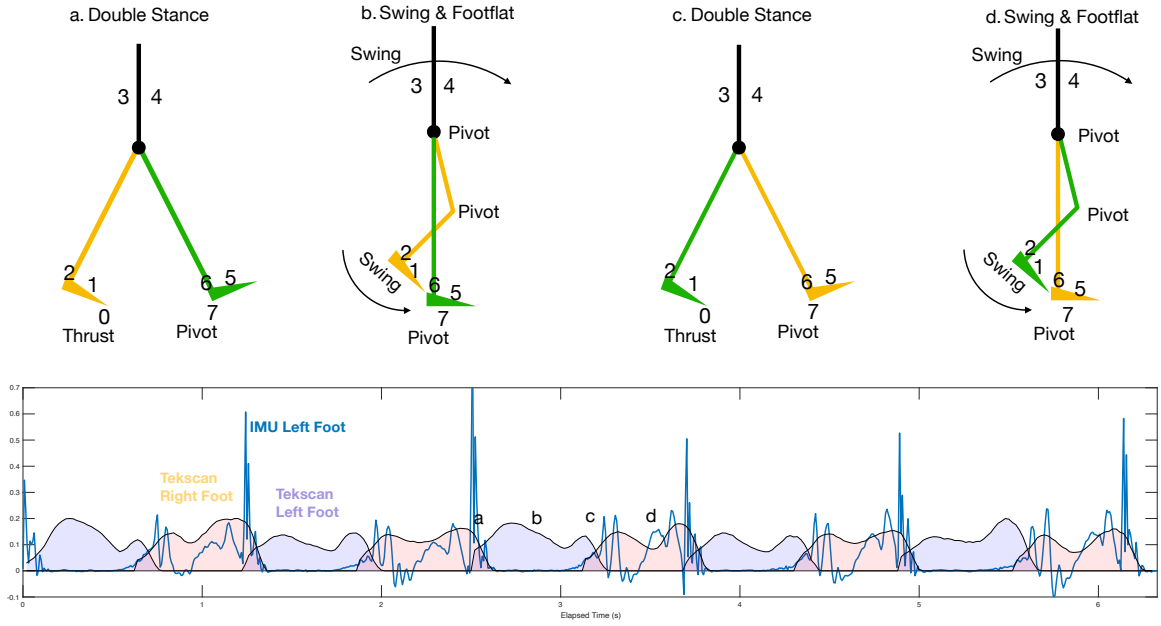


Figure 6.1: Movements of various body parts during the gait cycle. (Top row) a. Double stance: Initial contact by the left foot (green) and toe-off by the right foot (yellow). b. Swing phase of the right foot (yellow) while the left foot (green) in midstance. c. Double stance: Initial contact by the right foot (yellow) and toe-off by the left foot (green). d. Swing phase of the left foot (green) while the right foot (yellow) in midstance. (Bottom row) Ground reaction force (GRF) and acceleration magnitude recorded by IMU 1, aligned with the corresponding gait phases, illustrating a strong correlation between the sginals. Yellow indicates right foot and limb and green indicates left foot and limb. 0 and 7 mark the plantar pressure contact areas, 1 and 5 are the IMUs mounted on the top of the feet, 2 and 6 are IMUs attached to the ankles, and 3 is lumbar-mounted IMU, and 4 is the sternum-moutned IMU.

The DMD features dimensionality-reduction techniques combined with a spectral decomposition. The DMD uses the spectral decomposition of a highly correlated spatiotemporal dynamical system into a set of coherent spatial and temporal modes similar to other decomposition algorithms, such as singular value decomposition (SVD). In contrast to the SVD decomposition, utilizing a DMD method is beneficial as it offers a way to interpret the underlying law governing the dynamics and to predict future dynamical states. However, the DMD has a limitation in applying to low-rank systems (Tu et al., 2014). The Hankel DMD algorithm was introduced to remedy the issue by augmenting data matrix with time-delayed coordinates (Hankel matrix) (Arbabi and Mezić, 2017). Among many variants of DMD, Optimized DMD is developed to extract a subset of desired modes with the largest amplitudes for non-convex optimization techniques using a gradient-based search algorithm to reduce the least squares residual. The sparsity-promoting DMD incorporated the optimized DMD with compressive sensing to achieve a desirable number of modes by regularizing the non-convex optimization with the ℓ_1 norm penalty. DMD with control (DMDc) was introduced to produce an input-output model of the data from complex, high-dimensional systems that require applied external control.

To that end, we employed the aforementioned various DMD algorithms to extract the characteristic feature of the gait cycle using a combined dataset comprising GRF measurements and inertial signals, specifically, acceleration and angular velocity, recorded from various body spots (See Figure 6.1). By leveraging this fused dataset, we explore the potential data-driven modeling approaches that describe the relationship between kinetic and kinematic variables. In particular, we demonstrate that the gait cycle can be effectively modeled using DMDc, where the acceleration and angular velocity data are treated as system inputs and the plantar force as external control.

Moreover, therough the comparison of alternative optimization strategies, we propose a method for refining the DMD modes to better distinguish between intrinsic and extrinsic

features of gait. As complimentaty approach, we also examine the gait system from a control theoretic perspective by applying system identification techniques to model the body's movement using state space framework.

The remainder of the paper is structured as follows: First, we reviewed the work on gait analysis based on plantar pressure data from the walkway and acceleration and angular velocity data from IMU. In section 3, we presented the method to fuse the data sets from the pressure map and IMU signals and the algorithms to incorporate the data with DMD with control and optimization techniques. In section 4, we presented the implementation results for intrinsic and extrinsic gait features in terms of the DMD modes and control.

6.2 Data Processing

In this section, we examine the details of the data collected from the plantar pressure gound-based sensor (Tekscan) and the inertial measurement units (IMU) tied on crucial body parts. These two types of measurements offer the kinematic and the kinetic variables associated with a human gait motion. Thus, analysing the combined data provides the comprehensive understanding of the individual's walking mechanism. We employed the dynamic mode decomposition with control (DMDc) algorithm to investigate the spectral properties revealing the interplay between the kinetic input and the resultant kinematic outputs. Then, the projection parameters are optimized by means of the optimal closed-form scheme and the sparsity-promoted DMD algorithm.

6.2.1 Data Structure

6.2.1.1 Plantar Pressure

The pressure measured on the Tekscan ground sensor is a reaction force of the ground against the contact force exerted by the part of the foot during the gait motion. Each sensor unit captures the time-varying spatial distribution of the plantar pressure. The ground reaction force (GRF) integrated over the activated units corresponds to the force on the center of pressure. The trace of the GRF reflects the characteristics of the gait cycle's stance phase, which consists of the heel-strike, foot-flat, heel-off, contra-lateral heel-strike, and toe-push-off of the lead leg.

6.2.1.2 Inertial Measurement Units

The IMU measures kinematic variables such as accelerations, angular velocities, and magnetic fields in 3-dimensional coordinates in their inertial (sensor) frames. Particularly, the signals of the accelerations and angular velocities describe the movements of the feet and ankles, along with the lumbar and sternum, during the swing phase of the gait cycles pivoted by one foot on the ground. Since the IMU accelerations are measured in their inertial frames, each component describes the yaw, roll, and pitch in the x-, y-, and z-directions. At the same time, the IMU angular velocity components inform the device's orientation in their inertial frame. Thus, the location, linear velocity, and orientation of the body parts require the coordinate transformation to the earth reference frame. The IMU magnetic-field signals play a role in directing the geological orientations.

6.2.1.3 Data Fusion

To fully understand the underlying walking mechanism, we synchronized the IMU acceleration and angular velocity signals with the spatially integrated pressure signal from the

Tekscan mat. Since the Tekscan has a lower sampling rate (50 Hz) than that of the IMU (120 Hz), the pressure signal is oversampled using linear interpolation. We organized the IMU acceleration and the angular signals and the Tekscan pressure signals into a matrix form. The IMU sensors on the foot and ankle in swing are labeled with D_1 and D_2, the foot and ankle on the ground with D_6 and D_5, and the lumbar and sternum with D_3 and D_4, and the pressure signal with D_7, respectively. The partial pressure signal of the foot at toe-pushing off was labeled with D_0. Note that D_i for $i = 1, \cdots, 6$ contains 3-dimensional x-, y-, and z-components in device's inertial frame and D_7 contains a single component along the z-direction perpendicular to the ground.

6.2.2 Data Matrix

The combination of the IMU signals with the Tekscan pressure signals can be synchronized by matching the timestamps provided by the headers within ± 6 ms synchronization error. The initial timing of each gait cycle was determined at the sample point of the first non-zero pressure signal. After the time-synchroization, the Tekscan pressure signals were oversampled at the sampling rate of 128 Hz by linear interpolation up to two sample points. The ending timing for the input data matrix X was chosen at the last sample point of the oversampled non-zero pressure signals, corresponding to the duration of the stance phase of a gait cycle.

This window slice contains the double stance with both feet in contact with the ground: the leading foot in heel-strike and the lagging one in toe-push-off. As shown in Figure 6.2(a), the pressure signal from the lagging foot at the toe-push-off phase generates the swings of the torso about the foot on the ground and decays away, corresponding to the control input Υ. Once the force disappeared, its foot and leg begin to swing with foot-flat of the leading foot still in contact with the ground. The swing can be considered as the motion of a double pendulum with a pivot at the pelvis. After a gait cycle, the leading and lagging feet switch

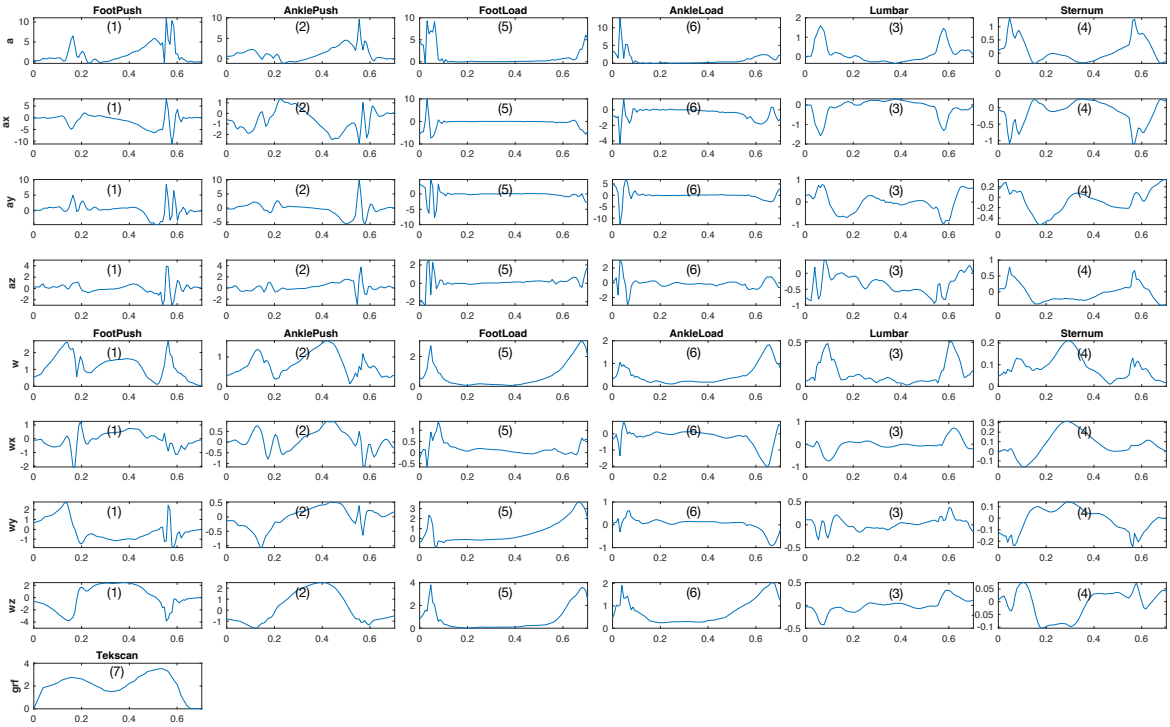


Figure 6.2: Data matrix D of a combined dataset of IMU and Tekscan. The numbers in the bracket represents the IMU devices and Tekscan contact point as shown in Figure 6.1(b). For example, the first column labeled with FootPush presents D_1 data matrix including the signals of the magnitude of accelration (a), its components (a_x, a_y, a_z), the magnitude of angular velocity (ω), and its components (ω_x, ω_y,ω_z). The last signal in the first column labeled with grf is D_7 including GRF by a foot on the ground.

roles, leading to the start of the next double stance. Figure 6.2(b) shows that the accelerations and angular velocities of the swing body parts (D_1 through D_6) reveal the pivoting plantar pressure by the leading foot (D_7). Figure 6.2 shows the data matrix D as a $49 \times N$ matrix form, where N is the number temporal samples.

6.3 DMD with control to fusion dataset

We computed the unique set of the eigenvalues λ and the corresponding DMD modes w for each participant by choosing the one with the minimum residues between D's and $\hat{D}$'s across all the gait cycles in each trial. Here, D denotes the fusion data of the IMU accelerations and

the angular velocities. $\hat{D}$ is the corresponding reconstructed signal based on the eigenvalues, the DMD modes, and the amplitudes. Figure 6.3(a) illustrates the eigenvalues of the time-evolution matrix K with the truncation values r for the toe-push-off force from the left foot and from the right foot, respectively. Figure 6.3(b) shows the corresponding spectra of the eigenvalues. The frequency and the decay rate are equivalent to the imaginary and real parts, respectively.

$$f_\alpha = \frac{T_s}{2\pi}\mathrm{Im}\ln\lambda_\alpha, \quad \gamma_\alpha = T_s\,\mathrm{Re}\ln\lambda_\alpha, \tag{6.1}$$

where $T_s = 128$ Hz is the sampling rate of the combined dataset. The indicated truncation ranks r for the left toe-push-off and right toe-push-off cases resulted from the optimizations of the residue error $\|D - \hat{D}\|$, which is in order of $O(1)$ in the normalized unit. This included optimizing both the truncation value r and the amplitudes $\hat{\alpha}$ using the standard DMD, optimal DMD, and sparsity-promoting DMD methods. The gray circles indicate (a) the eigenvalues and (b) the frequencies and decay rates of all the gait cycles during the trial. The size of each eigenvalue λ_i represents the magnitude of the corresponding optimized amplitude $\hat{\alpha}_i$ for $i = 1, \cdots, r$.

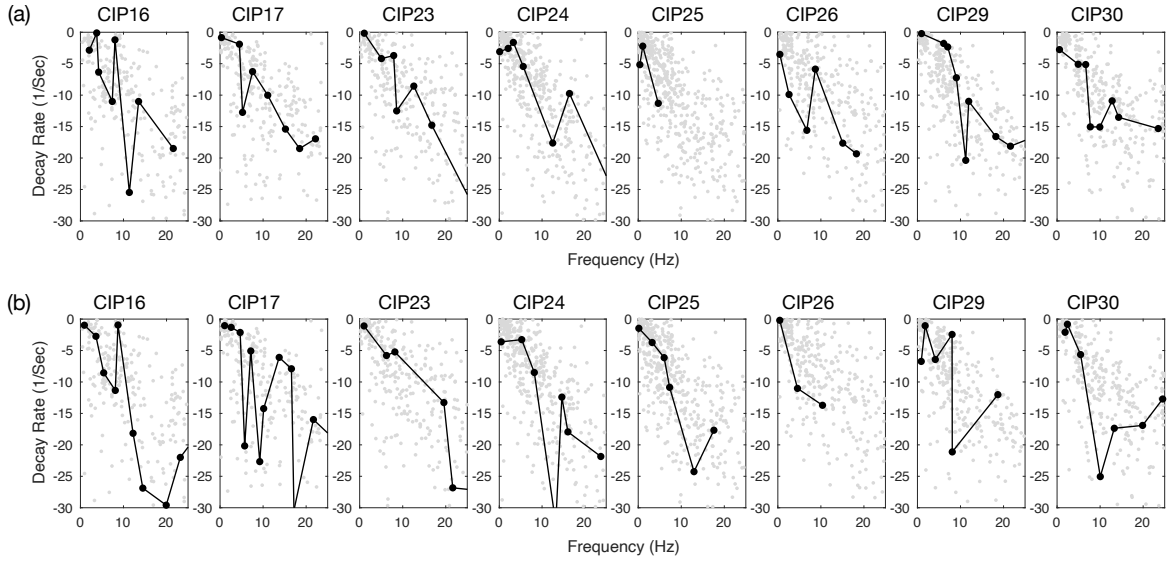


Figure 6.3: Eigenvalues with lowest residual error.

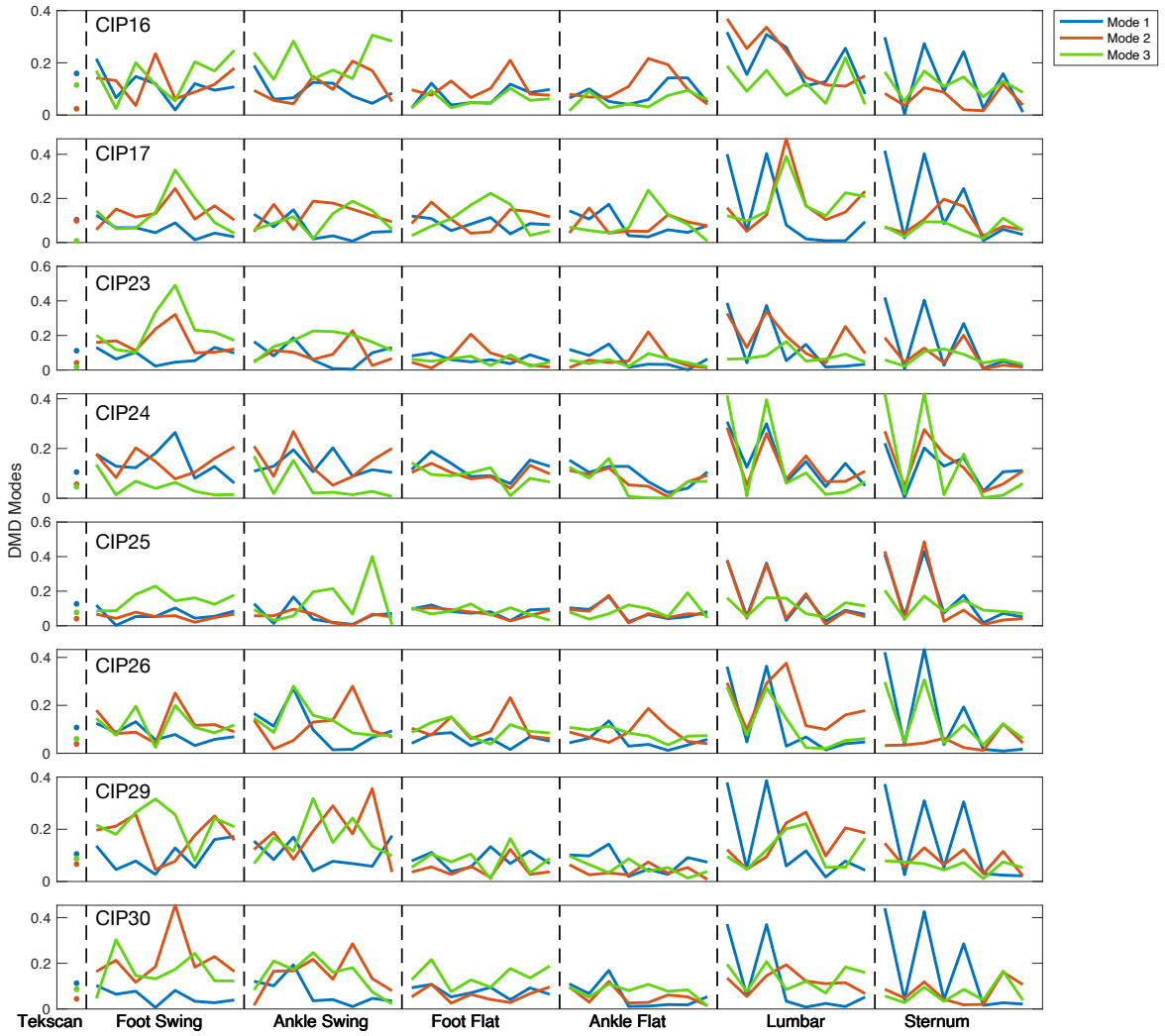


Figure 6.4: The first 3 low frequency and long-tail DMD modes.

Figure 6.6 shows the first four DMD modes associated with the eigenvalues (frequencies and decay rates) in Figure 6.3. The first red circle is the magnitude of the DMD mode at the Tekscan pressure signal from a leading foot on the ground D_7. The horizontal axis represents the locations of the IMU sensors and the vertical axis represents the magnitude of the DMD modes by denoted by the black circles.

For each IMU sensor, we included the magnitude of the acceleration $|a| = \sqrt{a_x^2 + a_y^2 + a_z^2}|$, and its components, a_x,a_y,a_z, and the magnitude of the angular velocity $|\omega| = \sqrt{\omega_x^2 + \omega_y^2 + \omega_z}$

and its components ω_x, ω_y, ω_z. Note that the components x, y, z are measured in the devive's inertial frame, corresponding to its yaw, roll, and pitch motions. Meanwhile, the pressure signal from the Tekscan is a measure of the force perpendicular to the ground. Notice that the DMD modes at the leading foot on the ground is suppressed since the 4th mode, while the DMD modes at the swinging body parts remains at the finite values. We can see the DMD reconstruction of the combined dataset in Figure 6.5. Each resultant reconstructed signal presents a good match with the measured signal during the swing phase, where the IMU sensors and Tekscan measure the smoothe motions from the moving parts of the body and limbs. In contrast, at the heel-strike and the toe-push-off events, the DMD reconstruction can't capture the fast fluctuating motion of the sensors as it requires DMD modes with higher frequency which has been truncated.

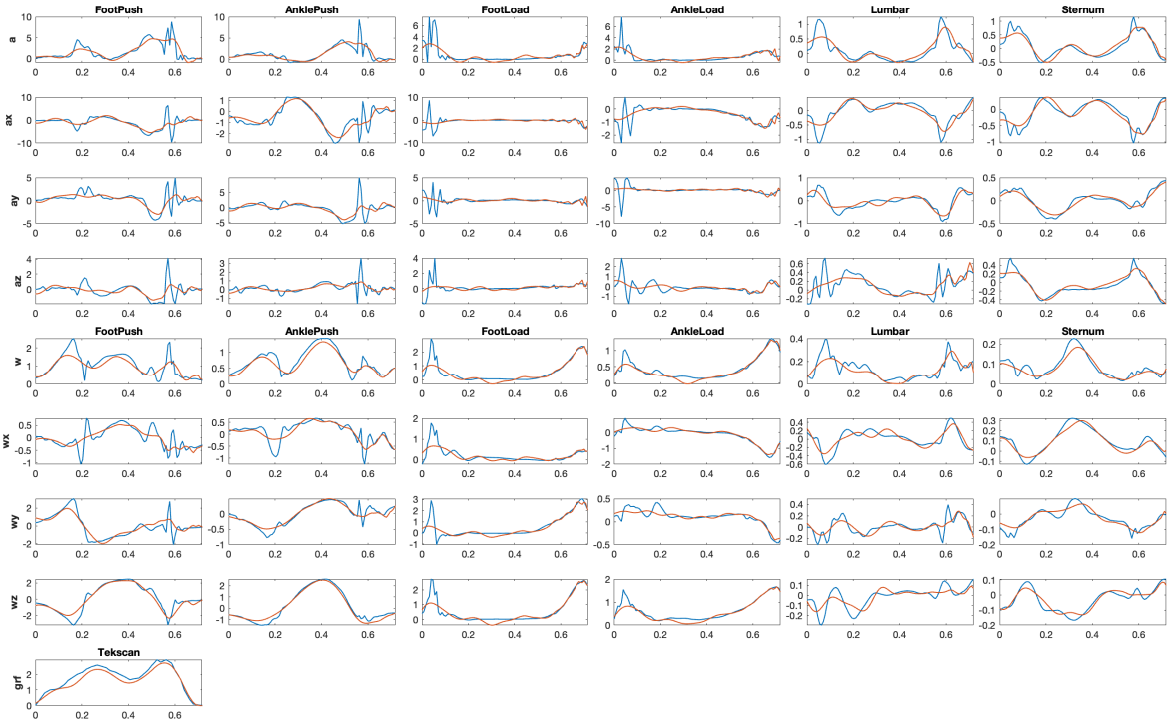


Figure 6.5: Data matrix D of a combined dataset of IMU (blue lines) and Tekscan superimposed with DMD reconstructed signals $\hat{D}$ (red lines).

6.3.1 Variability of the Gait Cycles

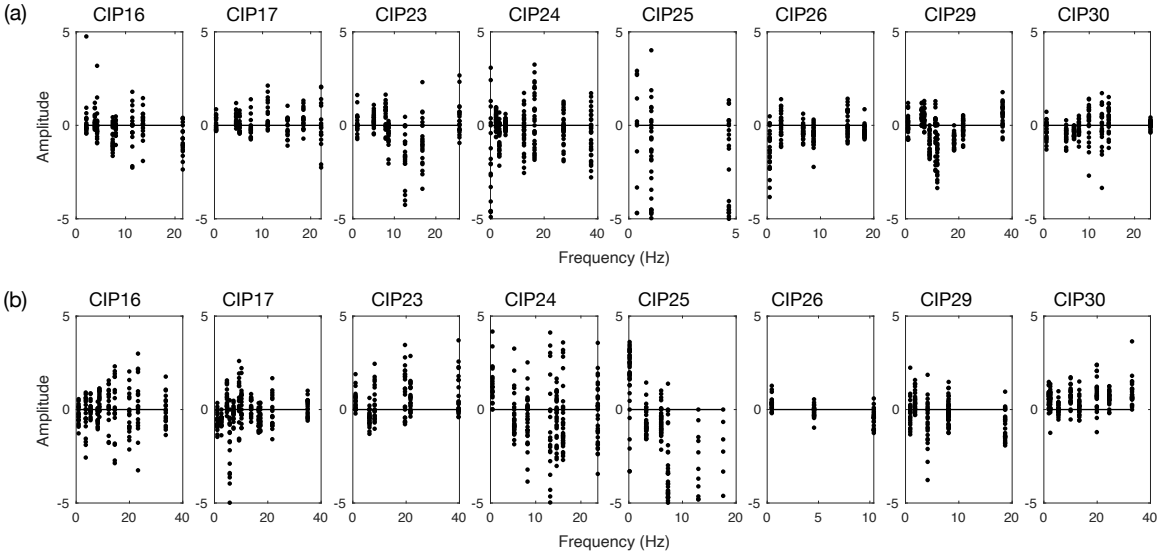


Figure 6.6: The amplitude fluctuation based on the DMD modes.

Figure 6.6(a) shows the magnitudes of the optimal amplitudes $\hat{\alpha}$ for corresponding system parameters, eigenvalues and the DMD modes. The gray colored circles represent the amplitudes all the gait cycles during the trial, which were the projections by the optimal DMD method to the system parameters (eigenvalues and the DMD modes). Figure 6.6(b) shows the effects of the weighted contributions to each body parts B from the toe-push-off control force Υ. Note that the control effects are inherent feature of the participant unless the external and environmental influences are small enough to maintain the states, meanwhile the values of the amplitudes depend on the specific circumstances and gait cycles. This implies the excessive variations in either the system parameters (eigenvalues and DMD modes) or the control effect indicates the severe change in the health condition.

6.4 System Identification Perspective

To complement our data-driven modeling of gait using DMD, we explore the use of state-space estimation as a control-theoretic framework for capturing the dynamics of human movement. This approach is particularly well-suited for modeling sensor-fused data, where inertial measurements (acceleration and angular velocity) represent internal kinematic states, and ground reaction forces (GRFs) measured via plantar pressure sensors represent external kinetic outputs.

A state-space model describes a dynamical system through a set of first-order difference equations that govern the evolution of internal (latent) state variables over time, along with equations linking those states to observed outputs. In discrete time, the system is typically formulated as

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \\ y_k &= Cx_k + Du_k + v_k, \end{aligned} \tag{6.2}$$

where x_k is the state vector at time $t = (k-1)\Delta t$ (latent gait dynamics), u_k is the input vector (GRF), y_k is the output vector (acceleration and angular velocity from IMUs), A, B, C, D are the system matrices to be estimated, and w_k and v_k are process and measurement noise, respectively.

State-space models provide a compact and interpretable representation of time-evolving systems, and they are well supported by identification and control theory (Ljung, 1999; Overschee and Moor, 2012). In the context of gait, this allows us to reconstruct unmeasured internal dynamics, track transitions between gait phases, and potentially predict future instability. To estimate the system matrices from data, we employ subspace-based identification methods (the N4SID algorithm) using the MATLAB System Identification Toolbox (The MathWorks, Inc., 2023). These methods require time-aligned input-output datasets and produce a model that best explains the observed dynamics in terms of latent states. This framework

allows us to explore the structure and dimensionality of gait control mechanisms, and to distinguish between intrinsic body-driven processes and external ground reaction responses.

In control theory and dynamical systems, poles represent the eigenvalues of the system matrix A in a discrete-time or continuous-time state-space model (Chen, 1998; Hespanha, 2009),

$$|\lambda I - A| = 0.$$

DMD eigenvalues are closely related to poles of the estimated state-space system. The poles magnitudes and angles describe frequency content of the gait dynamics that are corresponding to the frequency in DMD spectral features.

Figure 6.7 presents the frequency content extracted from the poles of the system matrix A, where the ground reaction force (GRF) is treated as the external input and the inertial measurements, acceleration and angular velocity, from the IMU devices serve as the system output. Figure 6.7(a) provides a schematic diagram of the system architecture, in which GRF data (D_7 in Figure 6.1) acts as the input channel. Figures 6.7(b) and (c) display the corresponding kinematic responses recorded from IMU device 1 and device 2, respectively.

To identify structure in the frequency content of the poles, we applied the DBSCAN clustering algorithm (Ester et al., 1996) using MATLAB's Statistics and Machine Learning Toolbox (The MathWorks, Inc., 2024a). Clustered frequency components are shown in color, while outliers are represented in grey. Notably, the extracted frequency clusters align well with those obtained through DMDc approach, supporting the consistency and relevance of DMD features in capturing the underlying gait dynamics.

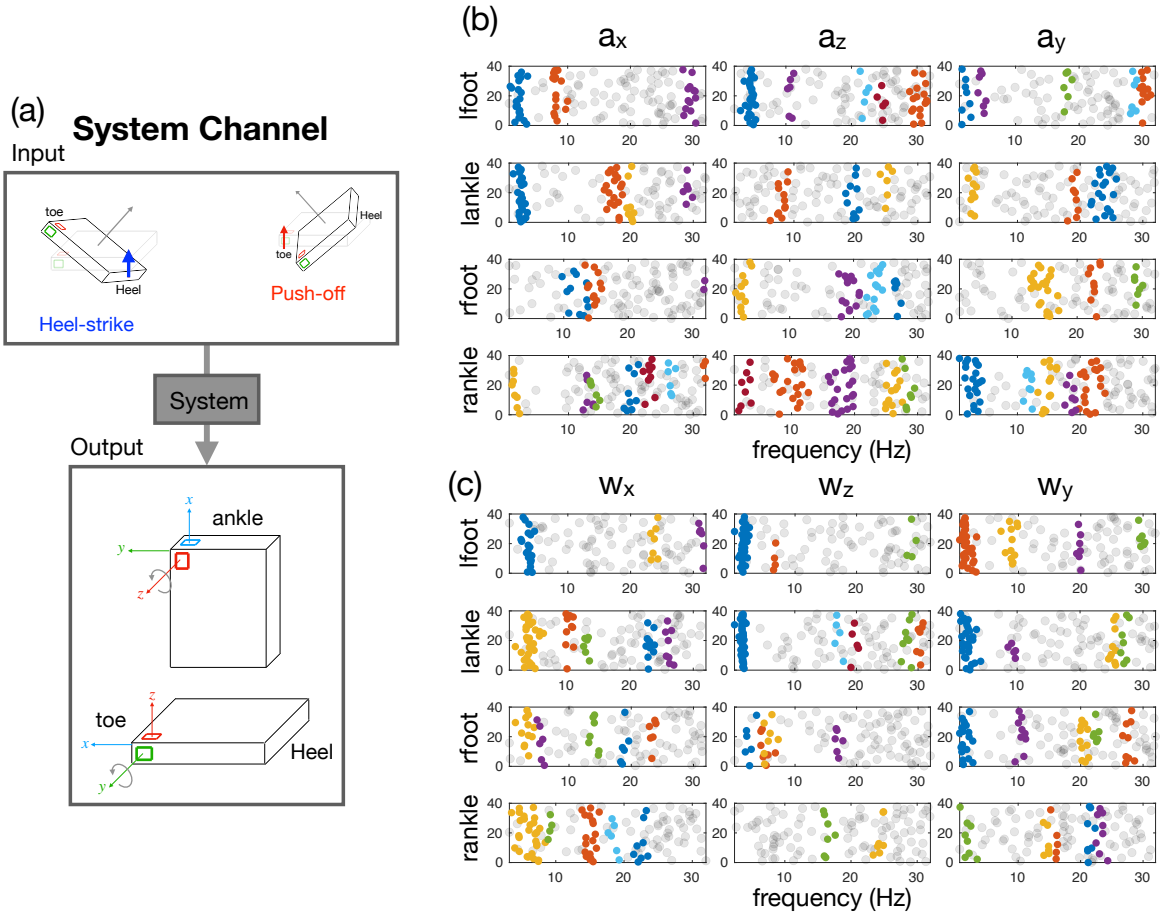


Figure 6.7: Frequency content derived from the poles of the system matrix A, where the GRF is used as the external input and inertial signals (acceleration and angular velocity) from IMU sensors are used as outputs. (a) Schematic of the system configuration, showing GRF as input (D_7 in Figure 6.1) and IMU responses as output. (b) Frequency content of acceleration components (a_x, a_y, a_z) from IMU device 1. (c) Frequency content of angular velocity components (ω_x, ω_y, ω_z) from IMU device 2. DBSCAN clustering was applied to identify dominant spectral patterns (colored points) and outliers (grey points).

6.5 Conclusion

In this work, we presented a comprehensive framework for gait analysis using multimodal measurements, integrating kinematic variables from portable inertial measurement units (IMUs) and kinetic variables from a ground reacted plantar pressure sensor (Tekscan Strideway).

While each dynamical variable such as acceleration, angular velocity, or GRF can be independently studied to provide insight into gait behaviors, such isolated perspectives may fail to capture the inherently coupled nature of human locomotion. The gait cycle arises from the tightly interwoven dynamics of multiple body segments, requiring a unified modeling framework capable of representing this complexity.

To this end, we employed combined datasets from IMUs and Tekscan strideway to construct data-driven dynamical models. By interpreting the double stance phase as the initiation of body movement, we modeled toe-push-off from the heel-strke as a control input, and characterized the resulting swing of the leading foot as system outputs. In earlier chapters, we demonstrated that applying DMD to spatially integrated plantar pressure signals (GRF) successfully extracted low-rank gait features and detected abnormal walking patterns. Building on that, we adopted DMDc to incorporate external control forces. This approach enabled us to capture both the transient and steady-state responses of the system.

Crucially, this study extends prior work by integrating spatial complexity directly from real multimodal sensor placements, rather than artificially augmenting spatial dimensions. As a result, DMD modes effectively characterized the dynamic interplay between IMU-recorded kinematics and Tekscan-measured kinetics under varying control inputs.

In parallel, we constructed a state-space model of the gait system using system identification techniques. Here, the GRF served as the external input, and IMU signals (acceleration and angular velocity) were treated as outputs. The estimated system matrix A provided a linear approximation of gait dynamics, and its eigenvalues (poles) revealed the dominant frequency content of the system response. We estimated these models and visualized the spectral content of the poles. Cluster analysis of these pole frequencies using the DBSCAN algorithm revealed patterns that strongly aligned with the DMDc-derived spectral components, reinforcing the validity of both modeling approaches. This convergence suggests that state-space estimation offers a complementary perspective on the gait dynamics encoded

in DMD modes, further strengthening the foundation for individualized gait modeling and potential clinical applications.

This final chapter presents an evolving yet promising direction in gait analysis by integrating kinetic and kinematic measurements through sensor fusion. By combining Tekscan-measured plantar pressure data with IMU-recoreded motion signals, we explored a unified modeling framework using DMDc and state-space estimation. These techniques not only preserved the interpretability of low-rank features but also revealed new insights into the interactions between foot-ground kinetics and limb kinematics. Although the analysis remains ongoing, preliminary findings suggest that this integrated approach may provide potentials for detecting subtle changes in gait patterns, with relevance for early diagnosis and individualized tracking of conditions.

Taken together, the work presented throughout this dissertation, from the identification of individual gait signatures to the detection of abnormality and now the fusion of multimodal data, lays the groundwork for a new class of data-driven clinical tools. As this research progresses, we anticipate richer models capable of informing rehabilitation, monitoring treatment effects, and enhancing our fundamental understanding of human locomotion. This chapter, though concluding the current dissertation, serves as a beginning to a broader, interdisciplinary journey in biomedical dynamics and personalized movement analytics.

Reference List

Al-Angari, H. M., *et al.*, 2017: Novel dynamic peak and distribution plantar pressure measures on diabetic patients during walking. *Gait & Posture*, **51**, 261–267.

Arbabi, H., and I. Mezić, 2017: Ergodic theory, dynamic mode decomposition and computation of spectral properties of the koopman operator. *SIAM Journal on Applied Dynamical Systems*, **16 (4)**, 2096–2126.

Ascher, U. M., and L. R. Petzold, 1998: *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. SIAM.

Bao, T., *et al.*, 2016: Long-term chemotherapy-induced peripheral neuropathy among breast cancer survivors: prevalence, risk factors, and fall risk. *Breast Cancer Research and Treatment*, **159 (2)**, 327–333.

Barth, J., *et al.*, 2015: Stride segmentation during free walk movements using multi-dimensional subsequence dynamic time warping on inertial sensor data. *Sensors*, **15**, 6419.

Ben-Gal, O., *et al.*, 2020: Using the loading response peak for defining gait cycle timing: A novel solution for the double-belt problem. *Journal of Biomechanics*, **110**, 109 963.

Berger, E., M. Sastuba, D. Vogt, B. Jung, , and H. B. Amor, 2014: Dynamic mode decomposition for perturbation estimation in human robot interaction. *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 593.

Berndt, D. J., and J. Clifford, 1994: Using dynamic time warping to find patterns in time series. *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining (AAAI Workshop)*, 359–370.

Boudali, A. M., *et al.*, 2017: Human locomotion analysis: Identifying a dynamic mapping between upper and lower limb joints using the Koopman operator. *Annu Int Conf IEEE Eng Med Biol Soc*, **1**, 88.

Box, G. E. P., *et al.*, 2015: *Time Series Analysis: Forecasting and Control*. 5th ed., Wiley.

Boyd, J. P., 2001: *Chebyshev and Fourier Spectral Methods*. 2nd ed., Dover Publications.

Breunig, M. M., *et al.*, 2000: Lof: Identifying density-based local outliers. *ACM SIGMOD Record*, **29 (2)**, 93–104.

Brunton, B. W., *et al.*, 2016a: Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *Journal of Neuroscience Methods*, **258**, 1.

Brunton, S. L., J. L. Proctor, and N. J. Kutz, 2016b: Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control. *PLoS ONE*, **11 (2)**, e0150 171.

Brunton, S. L., *et al.*, 2017: Chaos as an intermittently forced linear system. *Nature Communications*, **8**, 19–21.

Cao, Z., *et al.*, 2022: Characteristics of plantar pressure distribution in diabetes with or without diabetic peripheral neuropathy and peripheral arterial disease. *Journal of Healthcare Engineering*, **2022**, 2437 831.

Chandola, V., A. Banerjee, and V. Kumar, 2009: Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, **41 (3)**, 1–58.

Chartrand, R., 2011: Numerical differentiation of noisy, nonsmooth data. *ISRN Applied Mathematics*, **2011**, 1–11.

Chen, C. T., 1998: *Linear System Theory and Design*. Oxford University Press.

Chen, R. T. Q., Y. Rubanova, J. Bettencourt, and D. Duvenaud, 2018: Neural ordinary differential equations. *Advances in Neural Information Processing Systems (NeurIPS)*, **31**.

Clark, D. J., 2015: Automaticity of walking: functional significance, mechanisms, measurement and rehabilitation strategies. *Front. Hum. Neurosci.*, **9**, 246.

Eils, E., *et al.*, 2002: Modified pressure distribution patterns in walking following reduction of plantar sensation. *Journal of Biomechanics*, **35 (10)**, 1307.

Ester, M., H. Kriegel, J. Sander, and X. Xu, 1996: A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD)*, 226–231.

Feng, Y., Y. Ge, and Q. Song, 2011: A human identification method based on dynamic plantar pressure distribution. *2011 IEEE International Conference on Information and Automation*, **2011**, 329.

Franco, P. S., *et al.*, 2019: Within and between-days repeatability and variability of plantar pressure measurement during walking in children, adults and older adults. *Advances in Rheumatology*, **1**, 88.

Gallopoulos, E., Y. Saad, and G. W. Stewart, 1992: Efficient solution of parabolic equations by krylov approximation methods. *SIAM Journal on Scientific and Statistical Computing*, **13 (5)**, 1236–1264.

Gaßner, H., *et al.*, 2020: Gait variability as digital biomarker of disease severity in Huntington's disease. *Journal of Neurology*, **267**, 1594.

Gavish, M., and D. L. Donoho, 2014: The optimal hard threshold for singular values is $4/\sqrt{3}$. *IEEE Transactions on Information Theory*, **60 (8)**, 5040–5053.

Gilchrist, L., and L. Tanner, 2016: Gait patterns in children with cancer and vincristine neuropathy. *Pediatric Physical Therapy*, **28 (1)**, 16–22.

Good, P. I., 2000: *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. 2nd ed., Springer.

Hairer, E., C. Lubich, and C. Wanner, 2006: *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. 2nd ed., Springer.

Hairer, E., S. P. Nørsett, and G. Wanner, 1993: *Solving Ordinary Differential Equations I: Nonstiff Problems*. 2nd ed., Springer.

Hairer, E., and G. Wanner, 1996: *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*. 2nd ed., Springer.

Harris, E. J., I. H. Khoo, and E. Demircan, 2022: A survey of human gait-based artificial intelligence applications. *Frontiers in Robotics and AI*, **8**, 749 274.

Hastie, T., R. Tibshirani, and J. Friedman, 2009: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed., Springer.

Heersink, B., M. A. Warren, and H. Hoffmann, 2017: Network dynamic mode decomposition with control. *arXiv preprint*, arXiv:1709.02883.

Hespanha, J. P., 2009: *Linear Systems Theory*. Princeton University Press.

Hsieh, K. L., L. Trinh, and J. J. Sosnoff, 2019: Gait variability is altered in cancer survivors with self-reported neuropathy. *Gait & Posture*, **72**, 206.

Huang, M. H., J. Blackwood, M. Godoshian, and L. Pfalzer, 2017: Prevalence of self-reported falls, balance or walking problems in older cancer survivors from surveillance, epidemiology and end results-medicare health outcomes survey. *Journal of Geriatric Oncology*, **8 (4)**, 255–261.

Jolliffe, I. T., 2002: *Principal Component Analysis*. Springer.

Kalman, R. E., 1960: *A new approach to linear filtering and prediction problems*, Vol. 82. 35–45 pp.

Koopman, B. O., 1931: Hamiltonian systems and transformation in Hilbert space. *Proceedings of the National Academy of Sciences of the USA*, **17**, 315.

Koopman, B. O., and J. von Neumann, 1932: Dynamical systems of continuous spectra. *Proceedings of the National Academy of Sciences of the USA*, **18**, 255.

Kutz, J. N., X. Fu, and S. L. Brunton, 2016: Multiresolution dynamic mode decomposition. *SIAM Journal on Applied Dynamical Systems*, **15**, 713.

Levine, D., J. Richards, and M. W. Whittle, 2012: *Whittle's Gait Analysis*. 4th ed., Elsevier, London, UK.

Liu, F. T., K. M. Ting, and Z. H. Zhou, 2008: Isolation forest. *Proceedings of the 2008 IEEE International Conference on Data Mining*, IEEE, 413–422.

Liu, M. A., *et al.*, 2019: Gait speed, grip strength, and clinical outcomes in older patients with hematologic malignancies. *Blood*, **134 (4)**, 374–382.

Ljung, L., 1999: *System Identification: Theory for the User*. 2nd ed., Prentice Hall.

Luo, J., *et al.*, 2023: Optimal objective measurement of physical function and its predictive capacity for mortality among community-dwelling older women. *Geriatrics and Gerontology International*.

Mann, R., *et al.*, 2016: Plantar pressure measurements and running-related injury: A systematic review of methods and possible associations. *Gait & Posture*, **47**, 1.

Marshall, T. F., *et al.*, 2017: Chemotherapy-induced-peripheral neuropathy, gait and fall risk in older adults following cancer treatment. *Journal of Cancer Research and Practice*, **4 (4)**, 134–138.

Massey, F. J., 1951: The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, **46 (253)**, 68–78.

Mezic, I., 2013: Analysis of fluid flows via spectral properties of the Koopman operator. *Annual Review of Fluid Mechanics*, **45**, 357–378.

Mieg, L., and M. Mönnigmann, 2025: Optimal mode decomposition for control. *arXiv preprint*, arXiv:2504.08545.

Overschee, P. V., and B. D. Moor, 1994: *N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems*. Springer.

Overschee, P. V., and B. D. Moor, 2012: *Subspace Methods for System Identification*. Springer Science & Business Media.

Pamoukdjian, F., *et al.*, 2017: Slow gait speed is an independent predictor of early death in older cancer outpatients: Results from a prospective cohort study. *Journal of Nutrition, Health and Aging*, **21 (2)**, 202–206.

Parmar, N., H. Refai, and T. Runolfsson, 2022: A Survey on the Methods and Results of Data-Driven Koopman Analysis in the Visualization of Dynamical Systems. *IEEE Transactions on Big Data*, **8**, 723.

Proctor, J. L., S. L. Brunton, and J. N. Kutz, 2016: Dynamic mode decomposition with control. *SIAM Journal on Applied Dynamical Systems*, **15**, 142.

Pulido-Valdeolivas, I., *et al.*, 2018: Gait phenotypes in pediatric hereditary spastic paraplegia revealed by dynamic time warping analysis and random forests. *PLoS One*, **13 (3)**, e0192 345.

Raissi, M., P. Perdikaris, and G. E. Karniadakis, 2019: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, **378**, 686–707.

Rampp, A., *et al.*, 2015: Inertial sensor-based stride parameter calculation from gait sequences in geriatric patients. *IEEE Trans Biomed Eng.*, **62**, 1089.

Razak, A. H. A., A. Zayegh, R. K. Begg, , and Y. Wahab, 2012: Foot Plantar Pressure Measurement System: A Review. *Footwear Science*, **12**, 9884.

Rosenfeld, J. A., and R. Kamalapurkar, 2021: Dynamic mode decomposition with control liouville operators. *arXiv preprint*, arXiv:2101.02620.

Rowley, C. W., *et al.*, 2009: Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics*, **645**, 115–127.

Saad, Y., 2003: *Iterative Methods for Sparse Linear Systems*. 2nd ed., SIAM.

Sanz-Serna, J. M., and M. P. Calvo, 1994: *Numerical Hamiltonian Problems*. Chapman and Hall.

Savitzky, A., and M. J. E. Golay, 1964: Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **36 (8)**, 1627–1639.

Schmid, P. J., 2010: Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, **656**, 5.

Schmid, P. J., and J. Sesterhenn, 2008: Dynamic mode decomposition of numerical and experimental data. *61st Annual Meeting of the APS Division of Fluid Dynamics*.

Schölkopf, B., *et al.*, 2001: Estimating the support of a high-dimensional distribution. *Neural Computation*, **13 (7)**, 1443–1471.

Seo, K., H. H. Refai, and E. S. Hile, 2024: Application of Dynamic Mode Decomposition to Characterize Temporal Evolution of Plantar Pressures from Walkway Sensor Data in Women with Cancer. *Sensors*, **24**, 1.

Shlens, J., 2014: A tutorial on principal component analysis. ArXiv:1404.1100 [cs.LG].

Shorten, M., 2009: Plantar pressure distribution and footwear design. *Footwear Science*, **1**, 88.

Springer, S., U. Gottlieb, and M. Lozin, 2016: Spatiotemporal Gait Parameters as Predictors of Lower-Limb Overuse Injuries in Military Training. *Scientific World Journal*, **2016**, 5939 164.

Strang, G., 2016: *Introduction to Linear Algebra*. 5th ed., Wellesley-Cambridge Press.

Strogatz, S. H., 2015: *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. 2nd ed., Westview Press.

Székely, G. J., and M. L. Rizzo, 2013: Energy statistics: A class of statistics based on distances. *Journal of Statistical Planning and Inference*, **143 (8)**, 1249–1272.

Thaler-Kall, K., *et al.*, 2015: Description of spatio-temporal gait parameters in elderly people and their association with history of falls: results of the population-based cross-sectional kora-age study. *BMC Geriatrics*, **15**, 32.

The MathWorks, Inc., 2023: *System Identification Toolbox User's Guide*. Natick, Massachusetts, The MathWorks, Inc.

The MathWorks, Inc., 2024a: *DBSCAN Clustering - Statistics and Machine Learning Toolbox*. MathWorks.

The MathWorks, Inc., 2024b: *detectdrift (Statistics and Machine Learning Toolbox)*. The MathWorks, Inc., https://www.mathworks.com/help/stats/detectdrift.html.

Towne, A., O. T. Schmidt, and T. Colonius, 2018: Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *52nd AIAA Aerospace Sciences Meeting*.

Trefethen, L. N., 2000: *Spectral Methods in MATLAB*. SIAM.

Truong, C., L. Oudre, and N. Vayatis, 2020: Selective review of offline change point detection methods. *Signal Processing*, **167**, 107 299.

Tu, J. H., *et al.*, 2014: On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, **1(2)**, 391–421.

Winters-Stone, K. M., *et al.*, 2017: Falls, functioning, and disability among women with persistent symptoms of chemotherapy-induced peripheral neuropathy. *Journal of Clinical Oncology*, **35 (23)**, 2604–2612.

Zahiri, M., *et al.*, 2019: Using wearables to screen motor performance deterioration because of cancer and chemotherapy-induced peripheral neuropathy (cipn) in adults—toward an early diagnosis of cipn. *Journal of Geriatric Oncology*, **10 (6)**, 960–967.

Zhang, G., *et al.*, 2021: Plantar pressure variability and asymmetry in elderly performing 60-minute treadmill brisk-walking: Paving the way towards fatigue-induced instability assessment using wearable in-shoe pressure sensors. *Sensors*, **2021**, 3217.

Zou, Y. F., *et al.*, 2023: Clinical utility of plantar pressure measurements as screening in patients with parkinson disease with and without freezing of gait history. *Archives of Physical Medicine and Rehabilitation*, **104 (7)**, 1091–1098.