University of Nevada, Reno

Ski Sense, a Method for Classifying Snow

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in
Mechanical Engineering

by

Florian M. McLelland

Dr. Floris van Breugel/Thesis Advisor

May, 2022

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

Florian M. McLelland

entitled

Ski Sense, a Method for Classifying Snow

be accepted in partial fulfillment of the
requirements for the degree of

Master of Science

Floris van Breugel
Advisor

Ryan Tung
Committee Member

Anne Nolin
Graduate School Representative

David W. Zeh, Ph.D., Dean
Graduate School

May, 2022

Abstract

The fields of snow science, snow sports, and surface classification rarely cross paths, however understanding the specific type of snow on a mountain has value to all of these industries. Understanding the structure, quality, and type of snow is crucial for assessing avalanche safety, interpreting satellite imagery and choosing the right pair of skis for your weekend trip. Generally, the type of snow is known before it is skied on and in the past, little research has been done on the classification of snow while skiing on it.

This work explores the information that can be extracted from strain signals coming from an alpine ski with the goal of classifying four snow types. Data was collected from powder, slushy, groomer, and icy snow. Using the strain information and ski boot angular velocity I was able to classify these snow types independently of the skiing style with a maximum success rate of 97% by implementing a Naïve Bayes classifier. Comparisons of classifier performance between strain gauges, indicates that the optimal placement of strain gauges is halfway between the binding and the tip/tail of the ski. Additionally, orienting the strain gauges parallel and perpendicular to the length of the ski provides different, but complimentary information. Results from the principal component cluster plots show greater separation of powder snow on perpendicular strain gauges and parallel strain gauges show more defined clusters for hard snow types. The ability to classify the snow being skied on opens the door for many applications in the snow sports industry, as well as the robotics industry, in the context of surface classification.

To arrive at the results, I have implemented numerical methods of data analysis including singular value decomposition and uniform manifold approximation and projection. These tools have allowed me to extract meaningful information from real world data. Additionally along the journey I attempted using other techniques such as Dynamic Mode Decomposition to recreate a model of the ski's bending dynamics, however, due to the limited number of sensors this approach was unsuccessful.

To my parents.

Acknowledgements

I would like to first acknowledge and thank my advisor Dr. Floris van Breugel for providing the opportunity to take on this project starting as an undergraduate student. Throughout my undergraduate and graduate career Dr. van Breugel has provided me with valuable feedback and insight which have improved my coding and data analysis skills as well as my public speaking and writing skills. Because of my time in his lab I feel much better prepared for giving presentations. When I first stepped into Dr. van Breugel's office I had no idea what kind of a project I would be working on as his interests included a wide range of topics. Skiing was never something I imagined researching, only an activity I enjoyed in my free time, so I would like to thank him for providing me with such a cool and unique project. Getting to the point of actually collecting data I went through many failed attempts and had to manage lots and lots of wire, which the end result reflects that I am not an EE major. I would also like to thank my lab mate and friend Austin Lopez who helped me grapple with the coding/hardware interface. My good friend Matt Heintzen deserves a thank you as well for collaborating with me on future work and improvement of the project.

I would also like to thank my parents for their unconditional love and support which helped me keep a positive spirit during the challenge of graduate school. And my girlfriend Claire Dwyer who was always there for me and who would adventure with me on the weekends to disconnect from the work.

I would also like to thank my lab mates and friends who gave their support and encouragement and especially David Stupski for ensuring I had a good work life balance. Additionally Dr. Ryan Tung helped me a lot with some hard engineering concepts I did not think I was capable of understanding. I also appreciate him taking the time to workshop ideas and about my project. I would also like to thank Dr. Anne Nolin for providing outside feedback and valuable insight into the snow science industry.

It would also not be a complete list of thank yous if I didn't include my cat Luna. Al-

though she does not have the capability to provide feedback or insight, her playful character was a huge help through my last year of grad school.

Contents

List of Tables

List of Figures

Nomenclature

Acronyms:

$SG_{F\parallel}$ – Front parallel strain gauge.

$SG_{F\perp}$ – Front perpendicular strain gauge.

$SG_{B\parallel}$ – Back parallel strain gauge.

$SG_{B\perp}$ – Back perpendicular strain gauge.

$SG_{T\parallel}$ – Tip parallel strain gauge.

ADC – Analog to Digital Converter.

IMU – Inertial Measurement Unit.

GPS – Global Positioning System.

DC – Direct Current.

I/O – Input / Output.

I2C – Inter-Integrated Circuit.

SPI – Serial Peripheral Interface.

FFT – Fast Fourier Transform.

DFT – Descrete Fourier Transform.

SVD – Singular Value Decomposition.

UMAP – Uniform Manifold Approximation and Projection.

CVXPY – Convex optimization package in python.

TV – Total Variance.

Units of Measurements:

Ω – Resistance in ohms.

V – Voltage.

m – Distance in meters.

s – Time in seconds.

Variables:

${}^{N}\vec{M}^{B/B_p}$ – Total moment applied to body B at point B_p in reference frame N.

$\vec{\bar{I}}^{B/B_p}$ – Inertia dyadic of body B at point B_p.

${}^{N}\vec{\alpha}^{B}$ – Angular acceleration of body B in reference frame N.

${}^{N}\vec{\omega}^{B}$ – Angular velocity of body B in reference frame N.

$\vec{r}^{B_c m/B_p}$ – Distance of point B_p to the center of gravity of body B.

m^{B} represents the mass of body B.

${}^{N}\vec{a}_{p}^{B}$ – Linear acceleration of point B_p in reference frame N.

Chapter 1

Introduction

For humans it is a mundane task to tell the difference between a soft furry coat and ridged corduroy pants. We can even do it with our eyes shut by simply touching each surface and noticing their distinctly different feel. For a robot this is a much more challenging task, because we as humans must come up with a way to not only teach the robot what is corduroy and what is fur, but must provide this machine with the tools, sensors, and brains to separate the two surfaces into distinct classes. One such tool already implemented is computer vision, which has been quite successful at classifying surfaces including differently machined steel surfaces with slight differences in roughness [20]. Other methods include machine learning techniques for computer vision in object detection which are quite fast and accurate [8].

I wanted to provide the machine with a different tool which uses a more touch-based approach. Using physical interaction to classify the difference between corduroy and fur can be better than computer vision because the surface can have characteristics which do not appear visually, which is a necessity for successful computer vision. This also brings up the challenge of the way in which the machine interacts with the surface, creating the potential for error due to subjectivity. If you slide your hand across the corduroy slowly it will have a much different feel than if that speed is increased, leading to a problem that

requires some self-awareness.

A strain gauge is a sensor that relies on the physical interaction between itself and the body it is adhered to. Many living organisms use strain sensors in order to feel how the world is interacting with their bodies, and better navigate the environment. Moths utilize strain like sensors embedded in their wings to aid in flight control maneuvers [9], [28]. I decided to take this approach of embedded strain sensors in the arena of alpine skiing, which is of great interest to the local Reno, NV area, as well as a subset of the global population.

1.1 Snow Sports Industry

In the snow sports industry, performance is highly valued and it comes from a combination of skill and equipment. When skiing on different types of snow, different skis perform better than others. A stiffer ski is more responsive and has less vibrations when sliding, which is optimal for harder packed snow, while a less stiff ski is more forgiving and can be optimal for softer snow [35].

This brings up the challenge for ski manufacturers and users to provide and choose the correct pair of skis for each specific snow type being skied. To address this issue a company named Verispellis has created a stiffness changing ski actuated by a thermal heating pad controlled by the user [34]. This is a novel invention, however it is open loop and requires user input. I set out to develop a self sensing ski that could one day decide on its own to change stiffness based on the surface being skied. Doing this would close the loop and not require the user to change the stiffness of the ski manually. Creating such a ski would have wide benefits to users of the ski industry and could impact the skiers ability to traverse difficult terrain, and to not be limited by the ski they chose to use that day.

The field of snow sports is large and athletes and hobbyists are always trying to improve their performance. The improvement of skiing comes from practiced technique and equipment quality, which is where the fields of engineering and snow sports cross. In the ski

manufacturing industry, design is very intricate and complex, however data collection about the skis themselves is not very common. Few studies have been done on the quantitative performance of skis using strain data.

When speaking to the CEO of Moment skis located in Reno, NV, he told me that skis are tested in a qualitative manner where professional athletes review the "feel" of the ski. They report to the manufacturer how much vibration occurs when carving on groomers and how flexible the skis are when maneuvering difficult terrain. This is hardly a scientific approach because every skier will have a slightly different conclusion of the ski's performance restricting generalizations for manufacturing and design of skis.

Improving this manufacturing process can be done by understanding the true elasticity parameters of the ski [12], but it is more informative to understand how the ski performs in "real-world" scenarios which it will be in during use. Investigations to this end have been made by a group from Kanazawa University, Japan, where a ski was outfitted with bending and torsional load cells to collect data about the ski's behavior when skied on real snow [14]. This group was able to recreate the skis shape during a turn which could be useful for understanding ski performance and improving ski design. This project produced valuable results relating ski design to ski performance, however was only skied in one style and on one type of snow. Having this information however allowed me to conclude that there is valuable information in the strain of a ski.

1.2 Snow Sensing Techniques

One common method for snow characterization is through the measurement of snow albedo. Snow albedo is a measurement of how reflective the snow surface is and plays a large role in climate temperature, because the solar radiation that is not reflected by the surface of the earth is absorbed by the surface causing a heating effect [17]. Understanding the snow's albedo is very important to snow studies and understanding the earth's ability to reflect

light and reduce warming due to solar radiation. Typically, data for snow albedo is gathered using satellite imagery [22]. One limitation with satellite imagery is on a cloudy day the albedo measured from a satellite image will be of the cloud cover rather than the snow. The work in this thesis could be useful for making localized measurements of snow-pack because cloud cover will not impede the measurements. Although the data collected in this thesis is not related to snow albedo, further work could be done to relate this data to snow albedo.

A major driving factor in snow albedo is snow grain size [3]. As grain size increases albedo decreases due to reduced scattering within the snow-pack [24]. Being able to relate the snow grain size to the type of snow being skied on could prove valuable for snow science as it could be further related to the albedo of the snow. Snow type also affects the albedo of the snow because of the variance in grain size. New snow, like powder, has small grain size but when it begins to melt the grain size increases therefore impacting the snow's albedo [15]. Therefore, measuring the snow-type with skis can provide information about the snow's ability to reflect light.

One study explained that snow albedo is not a function of density and is, as previously stated, driven by grain size [2]. The experiment done in, [2] was collecting albedo data of fresh snow and then collecting albedo data after the snow had been compacted by a snowmobile. The results displayed that the two snow-packs of different densities but similar grain sizes provided the same albedo. Here it is shown that albedo does not necessarily correlate to snow-pack density, however using the work developed in this Thesis denser snow can be differentiated from lighter snow.

For back country skiers, understanding the quality of the snow is very important because it could mean the difference between life or death [7]. Typically, snow-pack quality is measured by digging a pit about 1.5 meters deep and performing a snow stability test [26]. These stability tests are done in a few localized areas, lacking information of the entire slope. Although these tests are quite effective for avalanche safety, they lack the ability to provide information about the entire slope. Having a ski that can sense information about

the snow-pack could also benefit validation of snow pit tests.

My goal was to differentiate what snow the ski was travelling on, so I needed to find characteristics of the different snow which would be present in the measurable data. To the human it is easy to tell the difference between powder snow and groomer snow, one is soft and pillow-like and the other is hard and rigid. Initially, one may think to use a simple pressure-sinkage device which measures the pressure of soil or snow by pressing down on a circular disk and measuring the force [31]. This would give a clear indication between a groomer and powder snow-pack, however it would be difficult and impractical to implement on a moving ski in real-time. Another characteristic of snow that can be noticed when skiing is induced vibrations from lateral sliding across snow. Two researchers, Glenne and Foss, outfitted a set of skis with accelerometers to measure the vibrations on skis and snowboards. The response spectra they reported displayed larger magnitude at higher frequency when traversing hard snow, and lower magnitudes on the softer snow [13]. This confirmed my intuition and experience from skiing on groomer and powder snow, that hard pack snow will result in higher frequency vibrations.

The extraction of characteristic frequency from the snow-type becomes quite cumbersome as the signals become complex. The spectra of the signals coming off the ski will not always be straight forward, so a simple threshold classifier will likely fail. The singular value decomposition is a method of data analysis that can be very useful for large complex data sets. Many applications of SVD involve parameter and data reduction, one example is image reduction [10]. Being able to reduce data using SVD does not only imply less data, but can be used in a more clever way to extract dominant patterns existing in your data and therefore can be used for setting up the framework for classification.

The SVD method is more generally used in Principal Component Analysis (PCA) in the context of feature reduction and simple classification. The SVD is what PCA is built on and is widely used for parameter reduction and separating data sets into classes [16].

This work is motivated by the application to a broader problem in engineering, surface

classification. I hope this study is of value to future engineers and scientists facing task of surface classification in the fields of snow science and others.

1.3 Layout of Thesis

This thesis begins with an explanation of the hardware methods and sensors used to collect the data. This was a large portion of the project as I built all the circuits in house. In this chapter, I also explain the experimental setup and the locations where the data was collected. Chapter 3 discusses the modeling and how the system was observed. It discusses the simplifications made and the signals I had access to and which ones I used to develop the system model. Chapter 4 explains how the data is processed and analyzed to develop a classifier for snow-type. Chapter 5 displays the results and chapter 6 explains the insights and significance of the results. Chapter 7 is the conclusion which discusses the findings and future work.

Chapter 2

Methods

During the first two years of the project I spent time learning, designing and building circuits and programming microprocessors to collect data. This hardware design took many attempts and eventually came to a successful module with some limitations that arose during the experiments.

2.1 Sensors and Ski Setup

To collect the ski strain data, four strain gauges were oriented parallel to the length of the ski and two perpendicular to the length of the ski. I used the two orientations to capture different information about the ski's deflections. Strain gauges are placed according to Figure 2.1, displaying three strain gauges in front of the binding, including one perpendicular and two parallel sensors. Two strain gauges were placed at the back of the ski, one perpendicular and one parallel. Additionally, there was one strain gauge placed beside the binding.

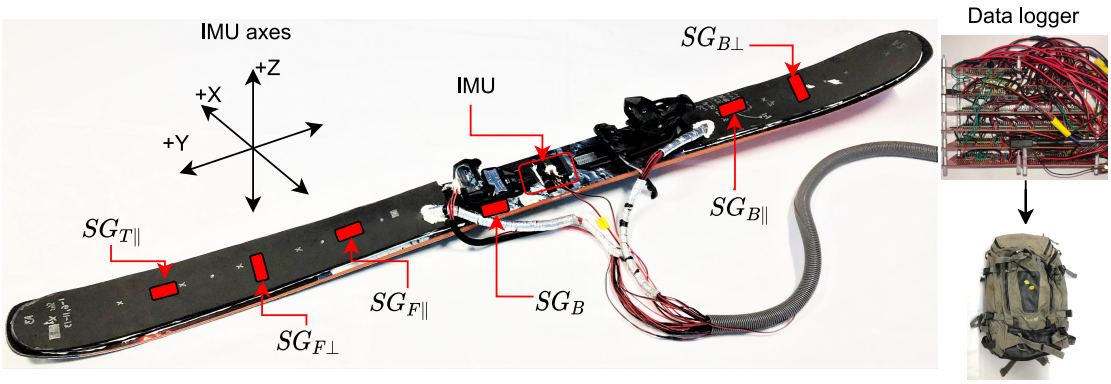

Figure 2.1: Physical ski used is a Moment Wildcat 108, outfitted with strain sensors on the front and back of the ski depicted in red. An IMU is located between the binding. All sensors are covered by closed cell foam padding and wires go to a data logger.

An inertial measurement unit (IMU) was placed below the ski boot to measure angular velocity and linear acceleration applied by the skier. Having these measurements provided signals from the skier allowing us to separate the skier's input from the snow's input. Most of the ski's motion was caused by torques applied by the skier while the vibration was driven by the interaction with the snow surface. The IMU uses I2C communication to send its data to a separate teensy at 200Hz.

A GPS sensor was positioned in the skier's backpack, connected to an antenna placed on the skier's helmet. Having GPS data allowed us to examine the velocity of each trajectory. The GPS uses serial communication to send its data to the same teensy connected to the IMU at a rate of 10Hz. All sensors and sensor specifications are tabulated in Table 2.1. The GPS I used has a maximum sampling rate of 18Hz, however when sampling at this frequency some of the data was lost due to some form of overloading of the unit, leading me to choose a lower sampling rate.

Table 2.1: Sensors used on the ski in experiments include strain gauges, an IMU, and a GPS. Strain sensor sampling rate is dependent on response time which generally ranges upwards of 300KHz [6], therefore it's sample rate is dependent on the analog to digital conversion rate which is set to 1KHz. The IMU was programmed to sample at 200Hz and the GPS was set to sample at 10Hz. All three types of sensors sample at different rates requiring signal interpolation.

Sensors	Part Number	Sample Rate	Company
Strain gauge (120 Ω), ADC (24 bit)	RS pro: 632180	1KHz	Allied, Protocentral
IMU (6Dof)	LMS9DS1	200Hz	SparkFun
GPS	ZOE-M8Q	10Hz	SparkFun

Due to the symmetry of our experiments only the right ski needed to have sensors on it. I applied a silicone coating to each strain gauge and the IMU to waterproof the circuitry, providing impact protection, minimizing water entry points, and reducing temperature fluctuations. Wire leads from each sensor were routed inside the skier's pant leg into a backpack where the data logger was placed during experiments.

Many of the times we set out to do an experiment, unbeknownst to us the data simply did not collect and store on the memory chip. To avoid this issue I installed a set of LEDs which acted as an array of signals indicating that sensors were working and were in fact collecting data. Table 2.2 outlines the exact use of each LED. In the final data logger there are two yellow, one blue and two green LEDs. The yellow LEDs are on when strain data is recording (data must be post processed to check for corruption error). One green LED indicates when enough satellites have linked to the GPS and one green LED indicates when data is being recored to the IMU and when the IMU has made a connection to the microprocessor.

Table 2.2: The array of LEDs have specific functionality shown here. The two orange LEDs are on when the microprocessor is collecting and storing strain gauge data (it does not indicate corrupt or error data). The blue LED indicates when the GPS is recording data. The green LED for the GPS indicates if the module has connected to enough satellites to receive reliable position data. The green LED connected to the IMU is on when the sensor is connected and collecting data. The blue and yellow LEDs blink faintly when other parts of the code are running and it is not actively storing data to the micro-SD card. This is hard to notice.

LED	Orange	Blue	Green	
Quantity	2	1	2	
Sensor	Strain	GPS	IMU	GPS
When on	Recording	Recording	Connected/ Recording	Connected to satellite

The IMU uses I^2C, which is a form of digital communication, and in some cases the signals from an I^2C sensor can become corrupted and cause an interrupt in the data collection which is why I implemented the green LED indicator for the IMU. The Arduino function *Wire.available()* forces the program to stall until there is data available. Additionally I^2C communication is prone to interference which I faced due to the $>$ 1m wire connections, and to avoid corruption from this I implemented pull up resistors [19]. The details of how to use the data acquisition system and the LED notifiers is available in Appendix A.

Initially, when performing lab tests, I attempted using only one Teensy 3.5 microprocessor as it has 64 digital I/O pins which is sufficient for the number of data streams. When doing this, the performance of the data acquisition was very intermittent. Through many debugging attempts where I only connected one sensor type at a time I was able to achieve reliable data collection when either only strain gauges, the IMU, or the GPS were being sampled individually. This was due to the different forms of digital communication between sensors and the Teensy. The IMU uses I^2C while the ADC for the strain gauges uses SPI communication causing inconsistencies in the data collection. Although this is not the most robust form of trouble shooting, it is effective. If I were a seasoned electrical engineer I may have been able to more elegantly collect data from multiple communication forms. Figure 2.2 displays the number of embedded systems used, and it is shown that 4 Teensies are used in order to mitigate the communication error, two for 8 strain gauges, one for the GPS and

one for the IMU.

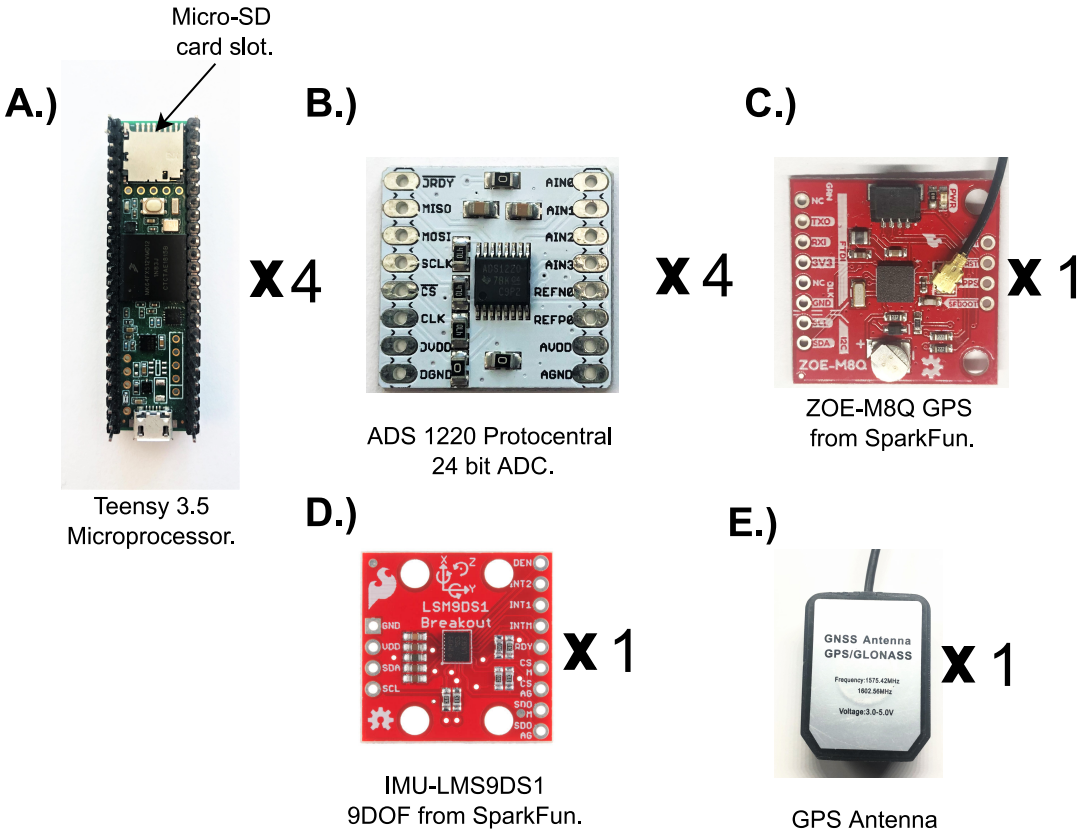

Figure 2.2: The sensors that are embedded systems and microprocessors are displayed. A is the Teensy 3.5 which has a micro-SD card slot where incoming data is stored. There were 3 Teensies used, 2 of which collected data from two ADCs each and one collected data from the IMU and the GPS. B shows the ADC used which has four single ended analog inputs which can be alternatively used as two differential analog inputs which is how I used them due to the functionality of a Wheatstone-bridge circuit Each ADC was gathering data from two strain gauges due to the amount of inputs on the chip. C is the GPS which uses the antenna shown in D. E is the 9 axis IMU, including linear acceleration, angular velocity and heading (which was not used). The image for E was available under the creative commons license from the SparkFun website [36].

A strain gauge under load only exhibits a small change in resistance making it difficult to measure voltage change across the gauge. To mitigate this issue I used a common Wheatstone-Bridge circuit detailed in Figure 2.3. The circuit is initially balanced in an unperturbed system, however a voltage is created when the strain gauge experiences stress due to actuation of the ski. Note that the strain gauges have a resting Resistance of 120Ω which is the same as the other three resistors in the circuit. In most engineering applications a minimum of a half bridge circuit is used due to it's ability to compensate for strain induced

by changes in environment temperature. In a half bridge, two strain gauges are placed on opposing sides of the circuit (replacing R_1 in Figure 2.3) so that the pattern is: resistor, strain gauge, resistor, strain gauge. This does two things: it allows the circuit to remain balanced under temperature induced strain, and it doubles the strain reading when the system is put under load. I was unable to use this method because the second strain gauge needs to be placed on the opposite side of the specimen, which would mean that gauge would be in contact with the snow. Another method for dealing with a temperature shift is to implement a load cell which is mounted to the specimen. I did not use this technique as it forces permanent holes to be drilled in the ski whereas adhering the gauges to the surfaces of the ski, mitigates this permanent damage.

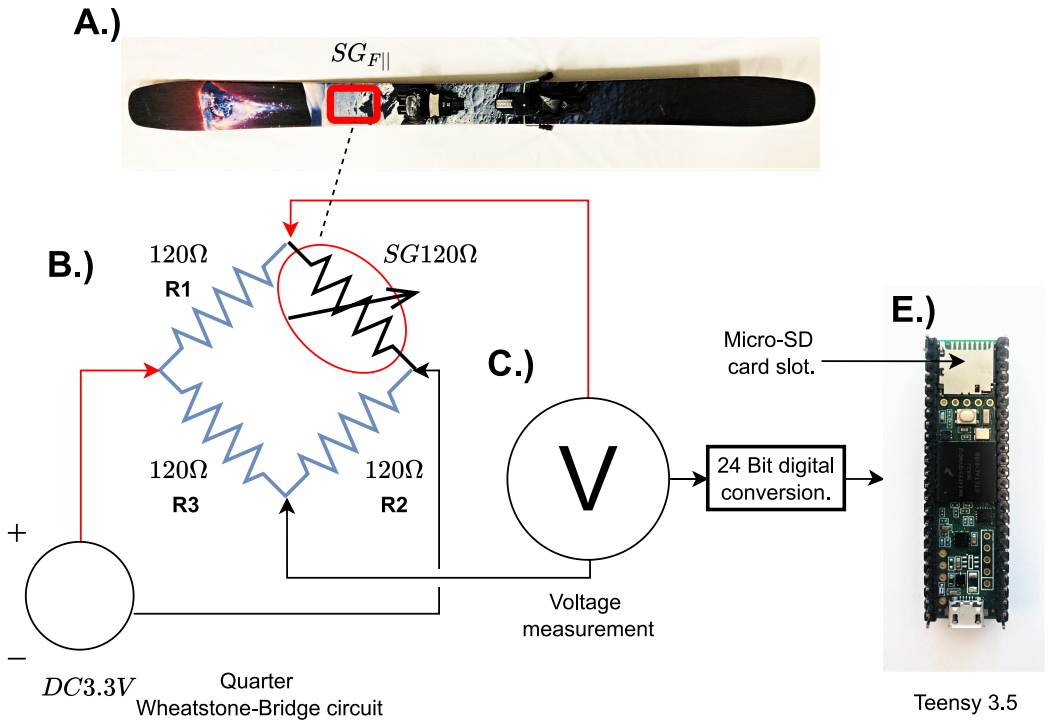

Figure 2.3: Each strain gauge is connected to a quarter Wheatstone-Bridge circuit consisting of three more resistors of the same resistance value. This creates a balanced circuit when at equilibrium, but when the strain gauge experiences stress and deforms, it's resistance value changes causing the circuit to become unbalanced creating a voltage that can be measured. This voltage signal is read by the ADC and stored on the micro-SD card.

The output voltage from the Wheatstone-Bridge circuit is connected to a 24-bit 1KHz analog to digital converter (ADC) by Protocentral which then sends the digital data to a

"teensy 3.5" via Serial Peripheral Interface (SPI).

Placing the strain gauge on the top surface of the ski was done to minimize temperature effects from the snow and to mitigate damage from sliding on the snow. As mentioned previously, a typical technique to mitigate temperature effects is to place a strain gauge on the opposite side of the specimen as seen in 2.4, however on a ski this would be the surface in contact with the snow. This would create a temperature differential contradicting the initial goal of the second gauge placement. I settled on one sensor placed on the top of the ski. Strain gauge temperature changes occur much slower than the ski bending modes and vibrations, therefore the temperature affect can be filtered out using a high pass filter.

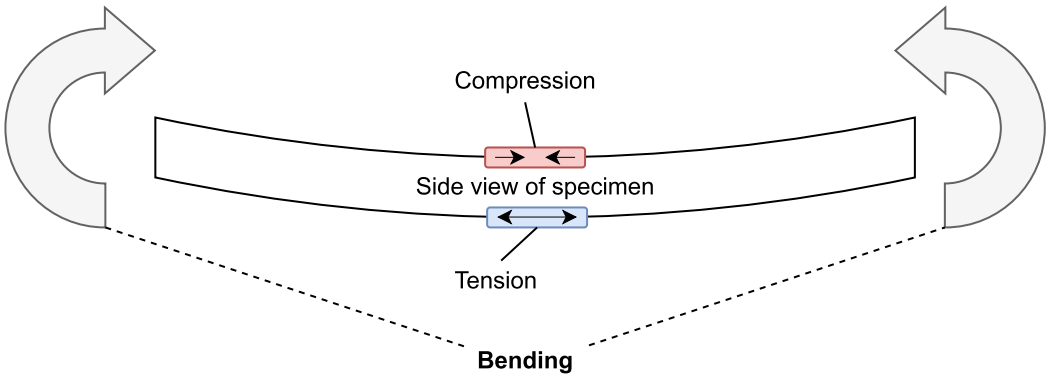

Figure 2.4: To mitigate thermal induced strain, sensors can be placed on either side of the specimen are wiring the senors on opposing sides of the bridge circuit. One sensor will be in tension and the other in compression during an actual system excitation.

Further justification for no temperature calibration is that the data was mean subtracted in order to ensure the magnitude of the strain signals are meaningful and can be compared to one another. Changes in strain gauge resistance creates a measurable signal. In order to justify my claim that strain readings can be mean subtracted and still provide meaningful results, the strain resistance must scale linearly with temperature. According to [4] the resistance of strain gauges scale linearly with temperature.

Another potential for error is that the current flowing through the Wheatstone circuit will cause the sensor to heat up inducing strain. If the sensor was not at its operating temperature, the strain signal would show a drift, however the data was taken from at least

three minutes into the run giving the sensor ample time to reach equilibrium temperature.

2.2 Data Logging

All data was stored on micro-SD cards located on the teensies in binary form, which were housed in a plastic box within the skier's backpack and powered by 5v battery packs. Analysis was done in post on a separate computer using custom software written in python. To align all of the data, the master teensy generated a random pulse signal with values ranging from zero to one, which was sent to the other teensies. Figure 2.5 shows the manual alignment process, where in A the pulse is not aligned between microprocessors and B shows the aligned signals. Once the signals were aligned, the IMU and GPS signals were interpolated to the 0.001s time step of the strain gauge.

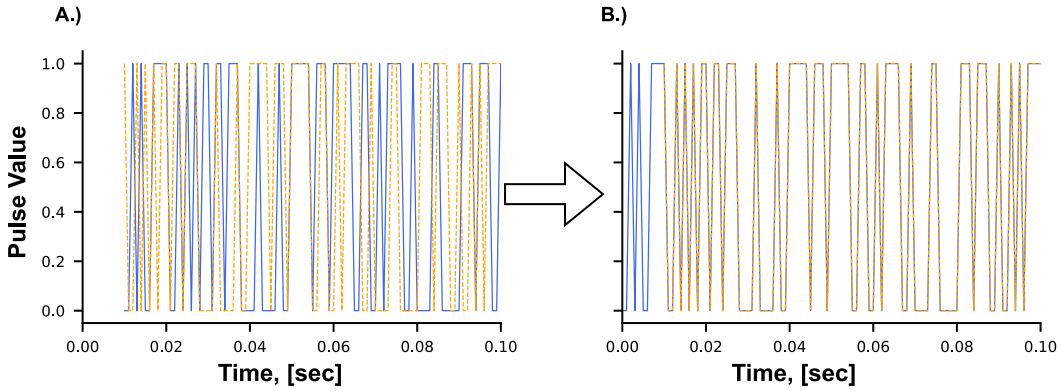

Figure 2.5: Random pulse sent from the Master teensy to all other microprocessors. A.) Pulse is not aligned. B.) the pulse has been time shifted and aligned. The blue pulse signal is sent from the master microprocessor which is then received at a later time point. The received signal is shown in yellow.

2.3 Snow, Terrain and Skier Trajectories

A set of Wildcat skis provided by Moment Skis, Reno, NV, was outfitted with an array of sensors and skied on various snow conditions at local mountain resort Mt. Rose, NV. The dimensions of the Wildcat skis are: length 180cm, tip and tail 141mm, waist 116 mm. The

four types of snow skied included: powder, slushy, groomer, and icy snow, which were all taken from the two different ski runs at Mt. Rose. Powder is very light and soft causing the ski to sink in while slushy snow is still soft however it has wet and heavy characteristics. Groomer snow is flat and very compact but still allowing ski-edge contact to be made. Icy is also flat and compact however maintaining edge contact is much more difficult. Each snow type was sampled on a separate day and was not sampled the same amount due to limited access to powder and icy snow. To minimize variability of data from terrain features such as moguls or bumps, only flat sections were skied. For consistency we recruited one skier whose height is 6' 1" and weight is 155lbs. Table 2 shows the details of each ski run and quantity of ski runs per snow type.

Table 2.3: All recorded ski runs are tabulated with the date, temperature, type of snow, type of turns and field notes about the snow's condition.

Location: Mt. Rose, NV				
Snow Type	Temp F°	Turning Style	Date	Notes
Groomer	25	Tight x3 Big x3	01/26/21	Smooth groomer.
Powder	20	Big Big	01/28/21	Untouched Snow. Skied out snow.
Icy	20	Tight Big	03/30/21	Sunny no wind. Groomed, but icy
Slushy	40	Tight x2 Big x2	04/01/21	Sunny and windy. Some 6" piles from previous skiers.

Powder snow is limited, because after many skiers have traversed the slope, it becomes uneven and subjects the ski to inputs characteristic to the terrain features rather than snow type. Two powder data sets exist, one pristine untouched data set and one skied out data set from later that day. Icy snow data is also limited due to the heating effect from the sun. Generally icy snow occurs when there are warmer afternoon temperatures but sub-zero night temperatures which cause ice layers to form.

To determine the impact of skiing style on our classification, tight and big ski turns were executed during data collection. Both left and right turns were used in the data analysis, ensuring both modes of actuation are included, because only the right ski had sensors on it. Although it is hard to visually notice the difference between fast and slow turns on the GPS trajectories in 2.6 B, looking at the smoothed angular velocities about the z-axis in 2.6 C, it can be seen that the blue trace has a lower frequency than the dark magenta trace.

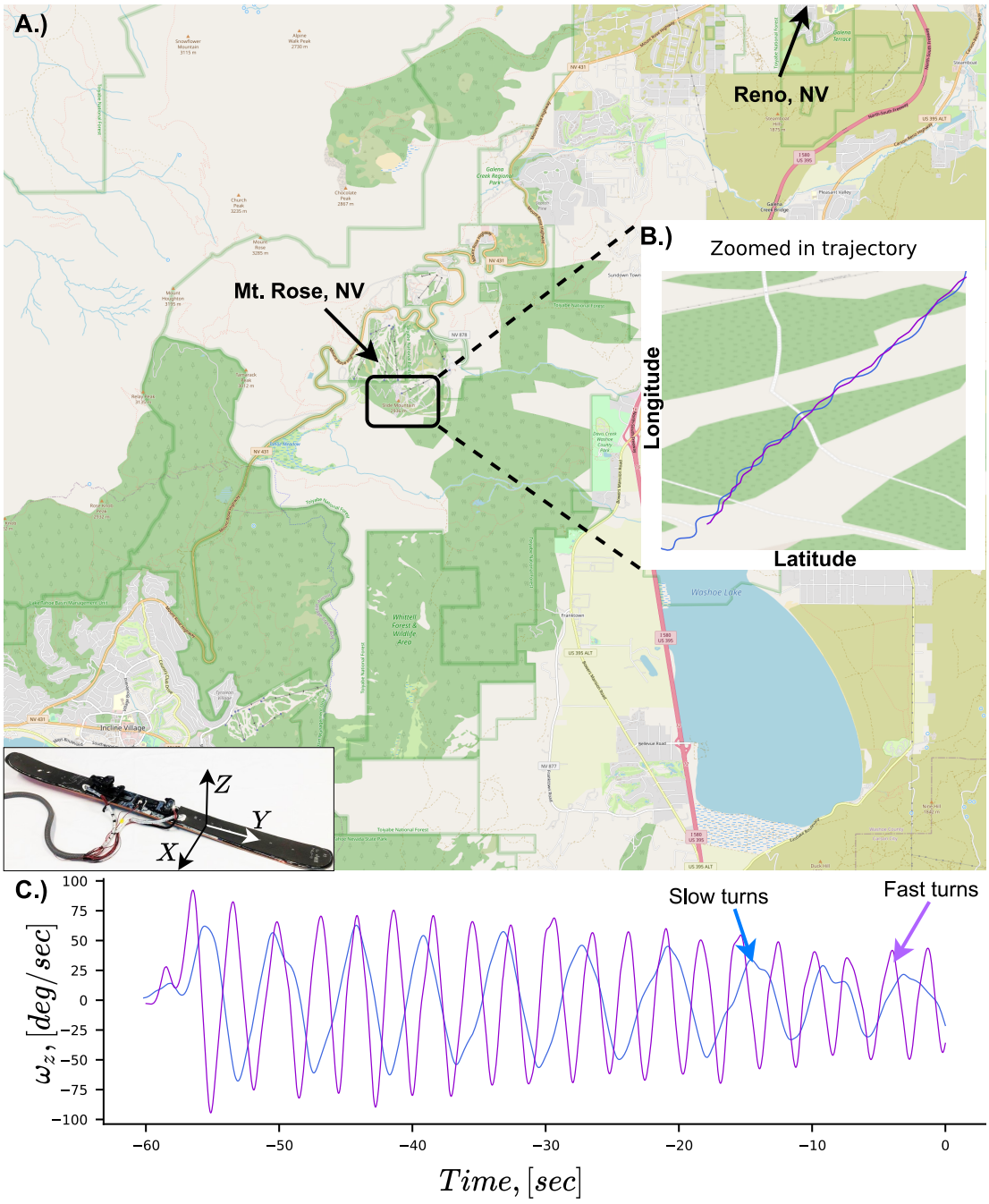

Figure 2.6: Example trajectories for fast and slow turning styles. A.) the zoomed out map of the Mt. Rose wilderness area including parts of Washoe valley, NV. B.) zoomed in GPS trajectory of the fast and slow turns. C.) angular velocities about the z-axis showing higher and lower frequencies.

Chapter 3

Modeling

Interactions between an alpine ski and snow surfaces are extremely complex due to the many non-linearities. The four common ski snow interactions are: smooth flat snow contact, self-excited vibration due to lateral sliding which occurs on harder snow, free vibrations when the ski is not in contact with the snow, and pure edge contact without sliding know as carving [29]. The non-continuous nature of these four modes is what creates a non-linear system. Due to this variety of modes, finding four analytical transfer functions is challenging, so instead I took the torque applied at the binding as the input and the strain at specific locations along the ski as the output signal to develop a "transfer function" between the ski and snow. I found that the angular velocity alone (which was directly measured) as a proxy for the input led to a much clearer classification of snow type.

3.1 Block Diagram

The entire skier-ski-snow system is more complex still because it includes a feedback system in which the skier has a desired trajectory and makes corrections to the actual trajectory. The complete system is detailed in Figure 3.1 and the red dashed outline highlights the subsystem we analyzed. In the complete block diagram, the skier applies forces and torques

to the ski-snow system to achieve a desired trajectory. A curvature of the ski forms as a result of the inputs which create an edge interface with the snow further leading to a true trajectory, which the skier can see and feel and adjust their inputs. Strain is a function of position along the ski.

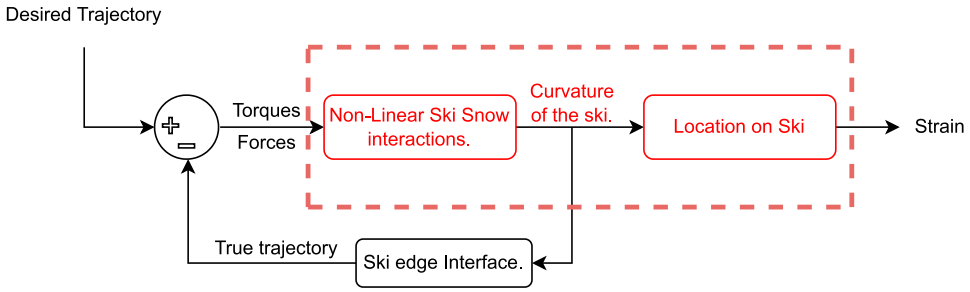

Figure 3.1: Skier-ski-snow block diagram. The skier applies torques to the ski in order to navigate the terrain and through the ski-snow dynamics the ski is deformed. The deformation causes strain which is picked up by the sensors.

The block diagram in Figure 3.2 shows the component of the skier-ski-snow system which I am analyzing. The input signal is only the torque from the skier because those are the dominant input signals. The output signals are the strain signals collected from the ski. This simple reduced block diagram generates the framework for analyzing the system in the general context of a transfer function.

Figure 3.2: Simplified block diagram of ski snow system, showing a transfer function between the skier's inputs and the measurable strain output.

The dynamics of skier actuation is very complex because there are many joints and linkages of the human body. Even simplifying the human body down to a combination of rigid links, rather than a flexible system, a complex intricate system remains. During a ski run, a skier moves his arms changing the system's inertia dyadic. This changing parameter can be mitigated because the torques are considered to be applied by the foot and I am only considering the ski and boot as one rigid body rather than the entire skier.

3.2 Torque vs. Angular Velocity

A skier applies toques and forces to the ski to navigate the terrain [5]. In classic skiing practice, the majority of the skier's weight is applied to the outside ski and a slight upward force is applied to the inside ski causing that ski to have less contact with the snow. Although forces are applied by the skier, torque is the dominant skier input because turning is primarily caused by parabolic edge-snow contact which is accomplished by torques applied by the skier.

The torque on the ski can be defined using the moment of effective force in equation 3.1 ${}^{N}\vec{M}^{B/B_p} \triangleq \vec{\vec{I}}^{B/B_p} \cdot {}^{N}\vec{\alpha}^{B} + {}^{N}\vec{\omega}^{B} \times (\vec{\vec{I}}^{B/B_p} \cdot {}^{N}\vec{\omega}^{B}) + \vec{r}^{B_c m/B_p} \times m^{B\,N}\vec{a}_p^{B}$, where I is the inertia dyadic which contains the products and moments of inertia of the ski and boot, ω is the angular velocity, α is the angular acceleration at the IMU, a is the linear acceleration at the IMU, r is the distance from the center of mass and m is the mass of the rigid body in question. Dimensions were taken from the ski's tip tail and binding locations to generate an approximate physical model of the ski in Solidworks 2020 computer aided design software to provide an estimate of the inertia dyadic. Dimensions of the skis tip and tail widths are 14.5cm, the waist is 12cm, the length is 180cm, the tip and tail thicknesses are 5.7mm and the waist thickness is 14.4mm. The model is shown in Figure 3.3. Table 3.1 displays the moments and products of inertia of the ski boot and ski modeled. Assuming uniform density throughout the ski, the material used in the model is ABS plastic, which is arbitrary as it only scales the inertia dyadic because of the assumption of uniform density. The complete torque equation for a three dimensional object was used because torque about one axis can cause rotation about another axis.

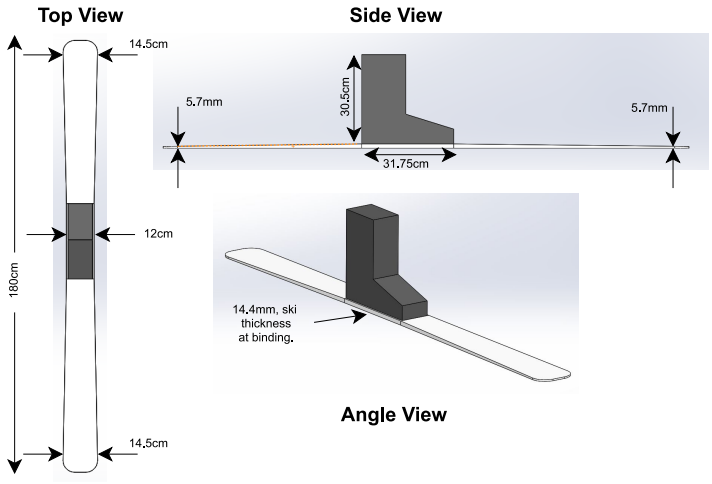

Figure 3.3: The Solidworks model of the ski is depicted with dimensions concurrent with the actual ski used for experiments. This model provided the inertia dyadic for calculations of the torques.

Table 3.1: The moments and products are displayed in units of kg/m^2.

	I_x	I_y	I_z
I_x	0.77	0.01	0.00
I_y	0.01	0.16	-0.06
I_z	0.00	-0.06	0.62

$$
\begin{aligned}
{}^{N}\vec{M}^{B/B_p} = \vec{I}^{B/B_p} \cdot {}^{N}\vec{\alpha}^{B} + {}^{N}\vec{\omega}^{B} \times (\vec{I}^{B/B_p} \cdot {}^{N}\vec{\omega}^{B}) + \\
\vec{r}^{B_c m/B_p} \times m^{B}\, {}^{N}\vec{a}_{p}^{B}
\end{aligned}
\tag{3.1}
$$

Where:

${}^{N}\vec{M}^{B/B_p}$ represents the three dimensional torque applied to body B at point B_p in reference frame N.

$\vec{I}^{B/B_p}$ represents the inertia dyadic of body B at point B_p.

${}^{N}\vec{\alpha}^{B}$ represents the three dimensional angular acceleration of body B in reference frame N.

${}^{N}\vec{\omega}^{B}$ represents the three dimensional angular velocity of body B in reference frame N.

$\vec{r}^{B_c m/B_p}$ represents the three dimensional distance of point B_p to the center of gravity of body B.

m^{B} represents the mass of body B.

${}^{N}\vec{a}_{p}^{B}$ represents the three dimensional linear acceleration of point B_p in reference frame N.

Measuring the true skier input of torque is challenging, because most torque transducers use strain gauges to acquire this data, and typically this method only results in measurement of torque about a single axis [21]. Whereas this system requires all three axes due to the 6 degrees of freedom, causing a direct measurement of torque to require a multi-axis transducer. This can be avoided by using an IMU which provides angular velocity, which can be differentiated to provide the necessary angular acceleration for the torque equation. The angular acceleration was calculated using an estimate of the measured angular velocity using a Butter-worth Filter implemented in the Pynumdiff python package to get a solution for torque that contained less noise [33]. Figure 3.4 displays the calculated torques and the smoothed angular velocities in the same colors but dashed traces. The zoomed in plot displays a phase shift of approximately $\pi/2$ radians occurring between the angular velocity and torques. Because my analysis uses the frequency content the phase shift is not as crucial.

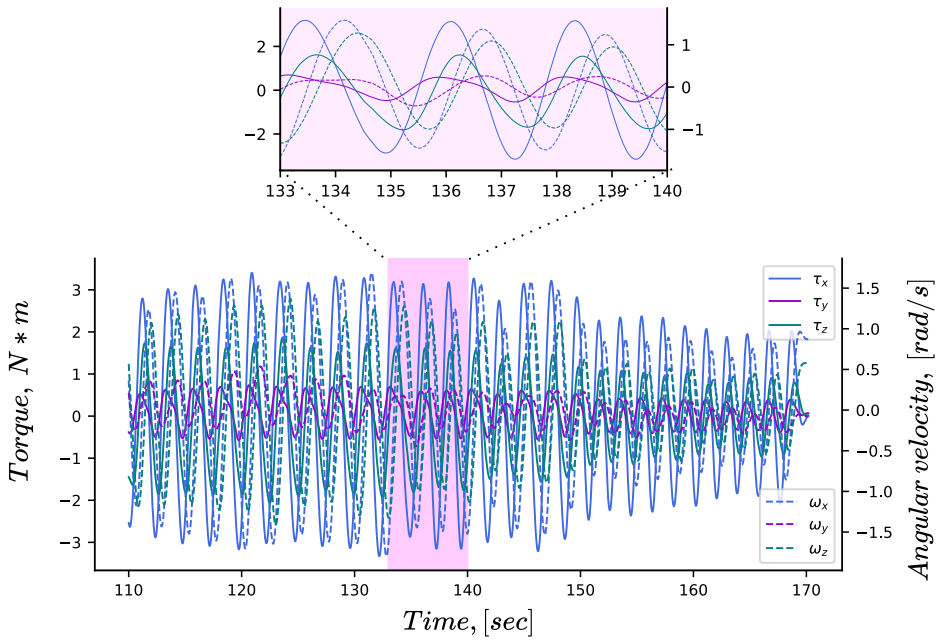

Figure 3.4: Smoothed torques plotted over smoothed angular velocities. Heavy smoothing is applied, making the units arbitrary. The angular velocity and torques have similar frequencies but display a time lag.

Although torque is indirectly available through equation 3.2, it is undesirable to use that calculated signal, because it contains multiplication and derivative terms which increase noise levels. In an attempt to avoid using the highly processed torque signal, an investigation into the details of the torque equation was performed. Additionally a comparison between the Discrete Fourier Transform (DFT) of the torque, angular velocity and angular acceleration was done.

The ski center of mass is very close to the location of the measurement allowing equation 3.1 to be simplified to the following: ${}^{N}\vec{M}^{B/B_p} = \vec{I}^{B/B_p} \cdot {}^{N}\vec{\alpha}^{B} + {}^{N}\vec{\omega}^{B} \times (\vec{I}^{B/B_p} \cdot {}^{N}\vec{\omega}^{B})$. This simplification still contains terms associated with the angular velocity. Additionally the equation becomes driven by the angular accelerations because the terms multiplied by the products of inertia become much smaller than the terms multiplied by the moments of inertia. The multiplication of the angular velocities also becomes much smaller than the terms including angular acceleration.

$$M_x = I_{xx} * \dot{\omega}_x + I_{xy} * \dot{\omega}_y + \underbrace{\omega_x \omega_y (I_{zz} - I_{yy})}_{\text{Multiplication of angular velocities}} + \underbrace{I_{yz}(\omega_y^2 - \omega_z^2) - \omega_x \omega_x I_{xy}}_{\text{Terms associated with products of inertia}} \tag{3.2}$$

A comparison of the DFTs of angular velocity and torque in Figure 3.5 shows that the two frequency signals have high similarity, while the frequency content of angular velocities do not. This further confirms that the input signal of torque is driven by the angular acceleration. This allows the analysis to use angular acceleration as a replacement for the torque input in the transfer function setting, however the angular acceleration still contains high noise amplification because it is the derivative of the angular velocity making less ideal to use than an unprocessed signal.

In order to relate the angular velocity (which is the measured unprocessed signal) the derivative can be taken. However as previously stated the time derivative amplifies noise due to the nature of the finite difference. An alternative approach is taken using the derivative definition in the Fourier domain where $DFT(\frac{d}{dt}f(t)) = \mathscr{F}DFT(f(t))$, where $\mathscr{F}$ is frequency array. Using this property, the DFT of angular velocity can be transformed into the DFT of angular acceleration by multiplication of the frequency array as seen in equation 3.3.

$$DFT(\dot{\omega}) = \mathscr{F}DFT(\omega) \tag{3.3}$$

Performing this multiplication scales the DFT of the angular velocity and still increases the high frequency noise so this multiplication term is dropped. Doing so can be justified because the DFT of angular velocity is linearly scaled by the multiplication of $\mathscr{F}$.

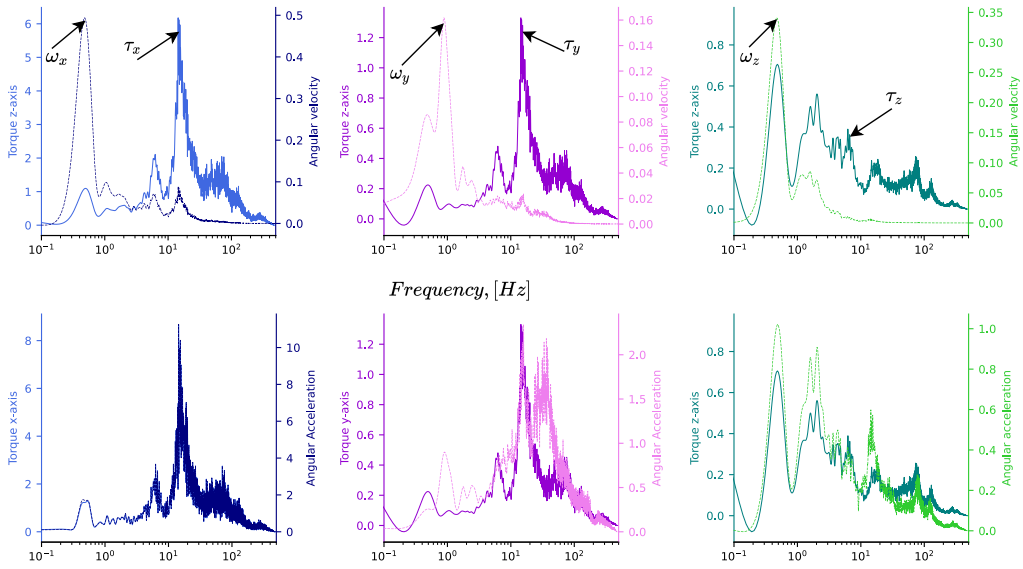

Figure 3.5: DFT of torque, angular velocity and angular acceleration. Signals have been slightly smoothed using the same smoothing parameters. Higher frequencies do not match because noise levels.

Using the angular velocity as the "input" signal for the ski-snow system allows the analysis to remain within the context of an engineering modeling context. Figure 3.6A shows the new block diagram using angular velocities as an input and keeping the discussion within the realm of transfer functions. Figure 3.6B shows the DFT ratio representative of the magnitude portion of a bode plot, which is commonly used for system analysis in the frequency domain.

3.3 DFT ratio

The transfer function defined in Figure 3.6 is not a true physical transfer function, because the input is not a force or torque. Therefore the transfer function in Figure 3.6 will be referred to as the "DFT ratio". These DFTs were solved using the SciPy python package [30]. The system model can be written in algebraic form using 3.4, where $\mathcal{A}$ is the $DFT(\omega)$, $\mathcal{B}$ is the $DFT(strain)$, and $\mathcal{X}$ is the DFT ratio which is being solved for.

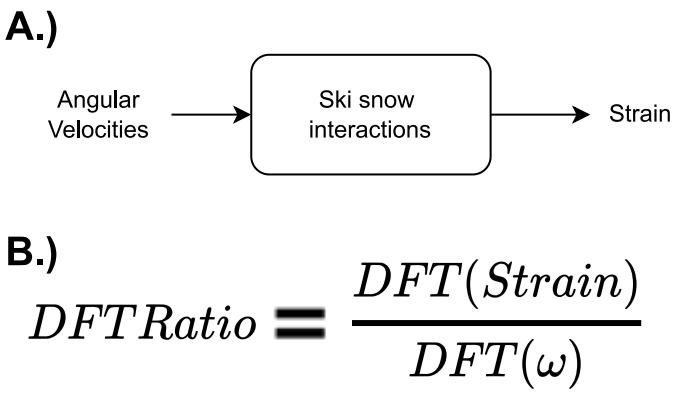

Figure 3.6: The system modeling approach has been simplified by using the angular velocities as an input into the system. A displays the model block diagram and B shows the relation to a transfer function through the use of DFTs.

$$\mathcal{A}\mathcal{X} - \mathcal{B} = 0 \tag{3.4}$$

Where:

$\mathcal{A}$ represents the DFT of the angular velocities.

$\mathcal{B}$ represents the DFT of the output strain.

$\mathcal{X}$ represents the DFT ratio.

Initially solving for $\mathcal{X}$ appears to be trivial and can be done using a naïve division of $\frac{\mathcal{B}}{\mathcal{A}}$. Performing the naïve division to solve for $\mathcal{X}$ on real data can lead to non-physical results which appear as jagged dips approaching zero caused by a small number being divided by a large number. Such dips are displayed in Figure 3.7A.

To mitigate these non-physical features, I implemented the python convex solver "CVXPY" which minimizes the loss function which I defined as equation 3.5 to find an optimal solution to the problem [23], [25]. The advantage of using the loss function is that constraints can be assigned like "Total Variance" which minimizes the difference between neighboring points.

$$\mathcal{L} = \sum_{n=1}^{n} (\mathcal{A}\mathcal{X} - \mathcal{B})^2 + \gamma TV(\mathcal{X}), \tag{3.5}$$

where:

$\mathcal{L}$ represents the loss function.

TV represents the total variance of the variable $\mathcal{X}$.

γ represents the weighting factor of how much the loss function values the total variance constraint.

To display the potential error arising from dividing DFTs (which often contain very small numbers), a simple mass spring damper simulation was subjected to a noisy chirp signal. The ratio of the input and output DFTs is displayed in Figure 3.7A where the jagged dips appear in teal and the true magnitude of the system Bode plot is displayed in black. The teal signal in Figure 3.7B was solved for by implementing the total variance constraint, leading to smoother and more accurate estimate of the true system response.

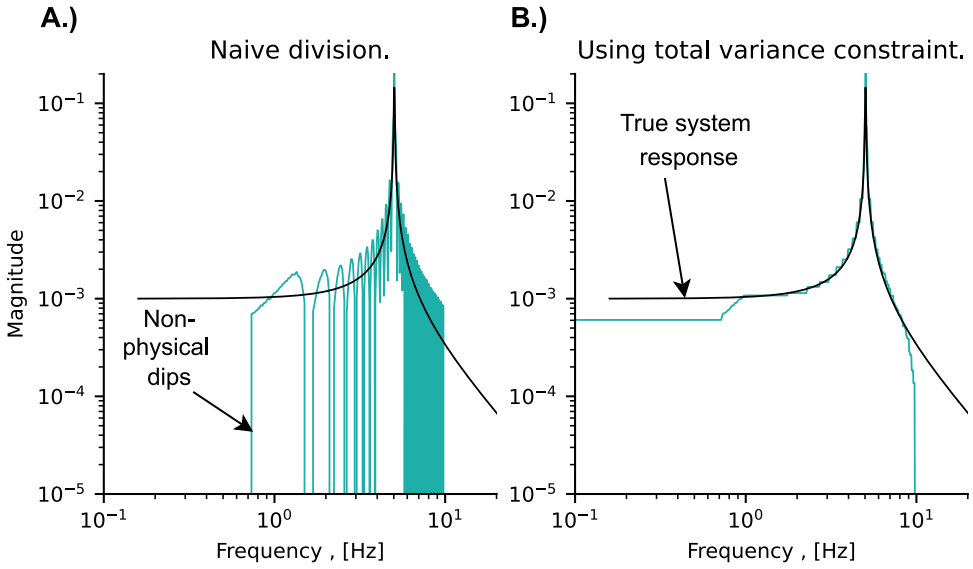

Figure 3.7: Non-physical dips in the transfer function are eradicated when constraining the total variance. Data is from a mass spring damper system subjected to a noisy chirp signal containing frequencies from 0.1-20Hz. The peak seen in both plots align at the resonant frequency of the system. The red trace represents the true no noise magnitude portion of the bode plot taken using classic systems engineering technique.

Chapter 4

Data Analysis

4.1 Raw Data

Although the data was taken from relatively consistent slopes, real world conditions change rapidly making the signals even richer. To maintain the signal's richness I analyzed the data from each ski run in two turn segments so I could analyze the classifier's performance when providing it with only a fraction of the information existing in the entire run. A minimum of two turns was necessary to ensure all bending modes of the ski were included in the data. Figure 4.1A shows the raw strain signal from the $SG_{B\|}$ in blue and the smoothed z-axis angular velocity in red. Using the smoothed z-axis angular velocity trace, I extracted segments of approximately two-turns using the point of zero deg/sec as an indicator for transitions between left and right turns. The zoomed in two turn segment is displayed in Figure 4.1B, where high and low frequency signals are apparent.

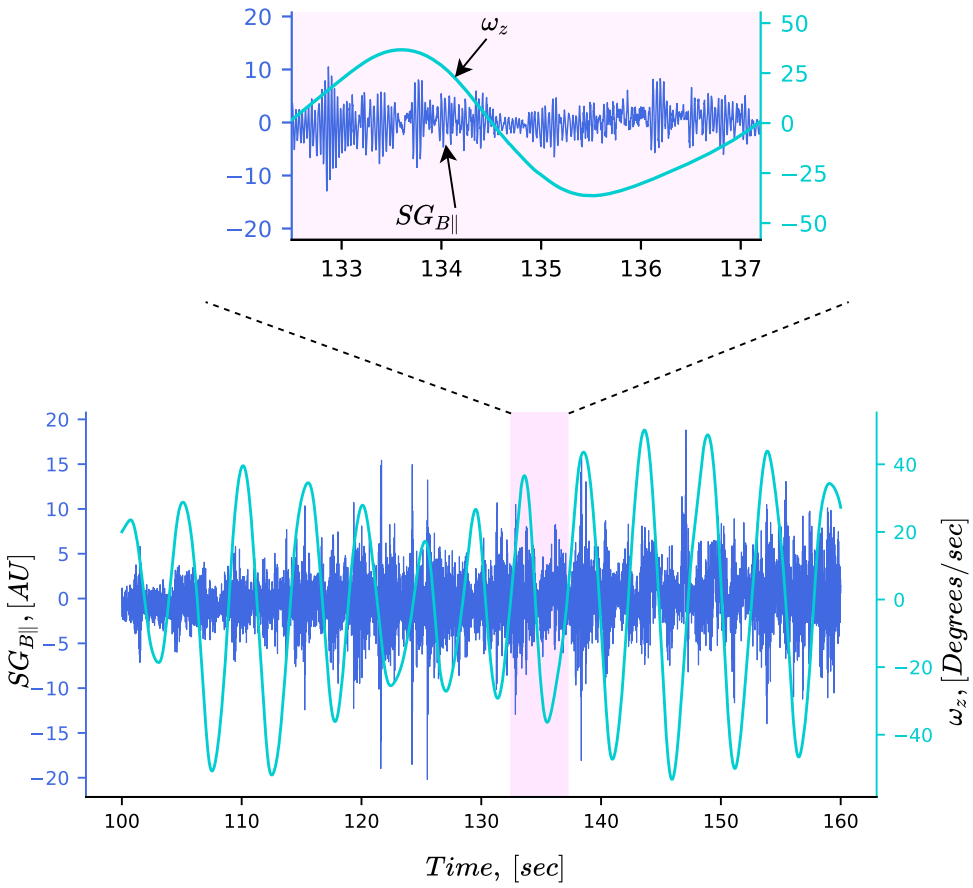

Figure 4.1: Full ski run displaying smoothed angular velocity about the z-axis and strain from the back parallel strain gauge. High and low frequency content is visible in the strain signal. A two-turn segment is taken to show that it holds a large amount of information with the intention of only needing to use a small portion of the signal to classify the snow.

4.2 Frequency Content

To numerically analyze these high and low frequencies observed in the zoomed in portion of Figure 4.1 (top), the DFT of the signals was taken providing the frequency content within the signal. Having the frequency content allowed me to visualize the dominant frequencies created by the ski-snow interactions as well as the frequencies caused by the skier seen in Figure 4.2. A high frequency spike is apparent in the groomer data shown in Figure 4.2A which does not exist in Figure 4.2B. These high frequency vibrations are common to feel when skiing groomer snow and are non existent on soft powder snow, which aligns with the

frequency content of the two signals. The low frequency magnitude spike in the angular velocity and strain follows our intuition that the skier imparts some signal onto the ski, and provides us with evidence that there is a relationship between the skier and the strain signal.

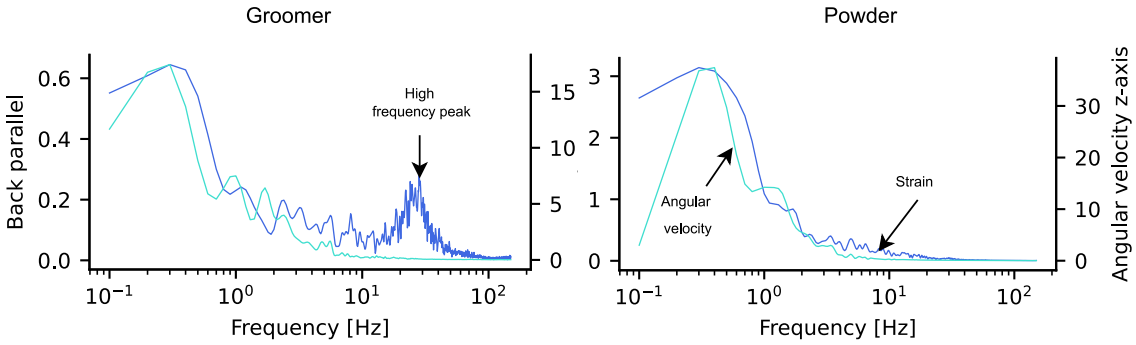

Figure 4.2: Comparison of a groomer and powder DFT. The groomer DFT has a high frequency spike and the powder does not. Both groomer and powder have a shared low frequency.

Motivated by the apparent differences in the $DFT(\omega)$ and $DFT(strain)$ the modeling approach developed in chapter 3 was used to create a DFT ratio, allowing the frequency content of both the input (angular velocity) and output (strain) to be used. This ratio is similar to finding a transfer function of a system's inputs to outputs, however because angular velocity is not a true input, it is referred to as the DFT ratio.

4.3 Creating a Matrix of DFT Ratios

Once all the DFTs had been created for each strain gauge and each axis of angular velocity, I solved for the respective DFT ratios. The DFT ratios were solved based on the flowchart in Figure 4.3A where each angular velocity component creates a ratio with each individual strain gauge. Figure 4.3B shows the large matrix of appended DFT ratios containing all the information of the skier-ski-snow interaction. The transpose of this matrix shown in Figure 4.3C is done to setup the data in the format needed to perform a Singular Value Decomposition which is explained in the next section.

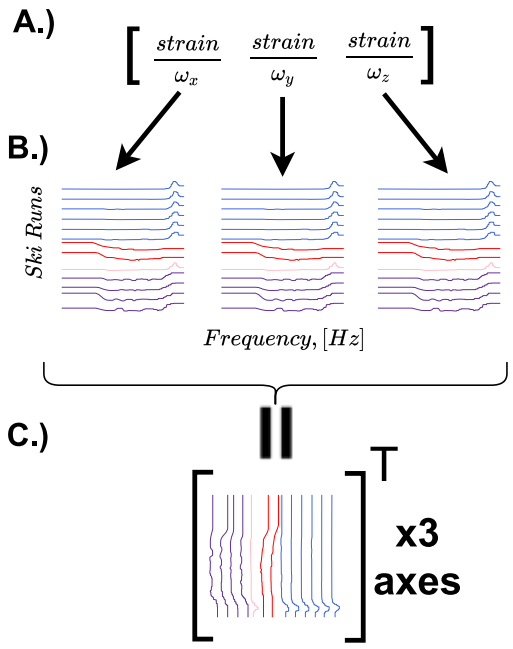

Figure 4.3: A.) DFT ratios are created for each angular velocity axis and individually for each strain gauge, they are then appended horizontally in B.) which is transposed creating the DFT ratio matrix in C.).

When generating these DFT ratios using the total variance constraint a flattening effect occurs which can be tuned using a hyper parameter γ acting as a weight telling the convex solver how much importance to place on minimizing the total variance. The higher the value of γ the flatter the DFT ratios become. Figure 4.4 shows that when a large value for γ is used, the DFT ratio becomes unusable as it is predominantly a flat line. When a value of 0.5 is used the DFT ratio provides peaks and valleys correlating to the type of snow skied.

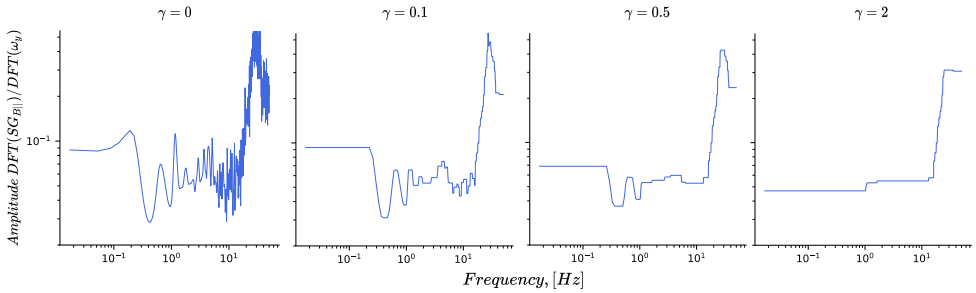

Figure 4.4: As γ increases the DFT ratio becomes flatter.

4.4 Singular Value Decomposition and Projection

To develop a classifier I implemented the Singular Value Decomposition (SVD), which is what extracts the dominant coherent structures of a matrix. In this case the coherent structures correspond to dominant shapes of the DFT ratio plots. A shape with a high frequency peak will most likely be associated with harder snow, whereas a DFT ratio with a lower frequency peak or no peak will correlate to softer snow.

The SVD performs an eigen decomposition of the original non-symmetric DFT ratio matrix, creating three new matrices U, Σ and V. U is an NxM matrix whose columns are the left singular vectors of AA^T (where A is the original matrix), Σ is an MxM matrix consisting of the singular values of A on the diagonals, and V is an MxM matrix whose columns are the right singular values of A^TA [11]. Equation 4.1 shows the relationship between the original DFT ratio matrix A and the decomposed matrices U, Σ and V:

$$A = U\Sigma V^T. \tag{4.1}$$

The large DFT ratio matrix is shown in Figure 4.5A which is decomposed into the matrices U, Σ and V are shown in 4.5. U represents the transformed set of bases where the coherent structures exist, Σ contains the singular values on the diagonal which are the weights of the new features existing in the rows of V. Figure 4.5C displays the singular values which drop off very fast. The diminishing singular values can be used to determine how many rows of the V matrix are needed to recreate the data. For visualization I chose to keep the top five features, which are still too many to visualize because humans can only intuitively perceive three dimensions. To visualize all five features on a two-dimensional plane I used Scikit-Learn's [32] Uniform Manifold Approximation and Projection (UMAP) which is a technique that preserves the Euclidean distance in the higher dimensional space [27]. This projection is shown in Fig. 4.5D-E which is where clusters begin to form according to snow

type. To summarize, each DFT ratio is broken down into its principal components using SVD, and the top five components are then projected down to two dimensions using UMAP leading to a single data point which appears in a cluster plot.

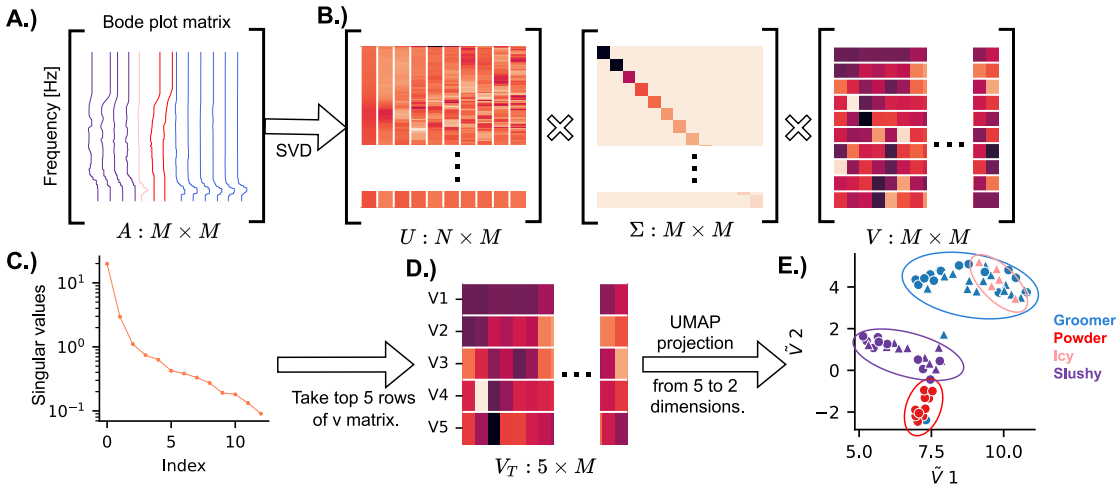

Figure 4.5: SVD analysis done on ski-snow transfer function. A.) Is the horizontally stacked transfer functions in the frequency domain. B.) Is the 3 matrices that the SVD creates, U is the new set of bases that the data exists in, S is a diagonal matrix of singular values, and V is a matrix of features in the SVD space. C.) is a log10 scaled plot of the singular values showing that they drop in value very quickly, D.) is the top 5 rows of the V matrix and E.) is the projection from 5D space to 2D space while preserving the Euclidian distance that exists in the higher dimension.

4.5 Classification

Although visually we can see that there are four distinct groups of data that have formed according to snow type in Figure 4.5E, we wanted to quantify the separation of the clusters in a statistically rigorous manner by training a classifier through supervised learning. A Gaussian Naïve Bayes classifier was trained with 70% randomly selected data from the top five V rows and then tested with the remaining 30% using the Scikit-learn library [18]. The Gaussian Naïve Bayes classifier provided a more parsimonious decision boundary compared to other methods such as the Support Vector Machine.

The Gaussian Naïve Bayes classifier was trained using different combinations of strain gauges. This was done by appending the DFT ratios together based on what combination

of strain gauge was used and taking the SVD. The rows of the V matrix from the SVD operation were used as features. Two feature decisions are made here. One is the combination of strain gauges, and the other is the quantity of rows used to train the classifier. This described workflow is displayed in Figure 4.6.

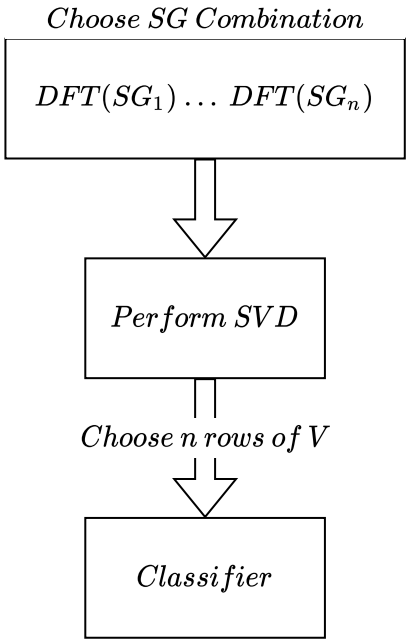

Figure 4.6: Classifier workflow is displayed. First the combination of strain gauges is displayed, next the SVD is taken of the appended DFT ratios, and lastly the desired number of V rows are provided to the classifier.

Chapter 5

Results

5.1 Clustering of Strain and Angular Velocity Separately

To provide further justification for using the top five rows of the V matrix from the SVD calculation, I plotted the top five rows in a pair plot, which plots each five rows against the each other. When looking at the rows plotted against each other in 5.1, it can be seen that there is variance which provides meaningful information and therefore the higher dimensions of V should be used in the classification. When looking at V_1 plotted against V_3 the orange powder cluster of data is much more separated from the rest of the data, justifying it's use to the classifier. Going into more detail Figure 5.1 provides comparisons between the rows of the V matrix coming from an SVD of the DFT ratio of $SG_{B||}$ and $\omega_{x,y,z}$. Figure 5.1 A,C,D, and H show distinct separation between hard and softer snows, while E,F, and G separate powder and slushy snows. I and J display the most mixing of snow types. Figure 5.1 B and C show a group of icy points separated from the main cluster of points, however it is mixed with some groomer data points. The red lines and circle indicate where the separation and grouping is occurring.

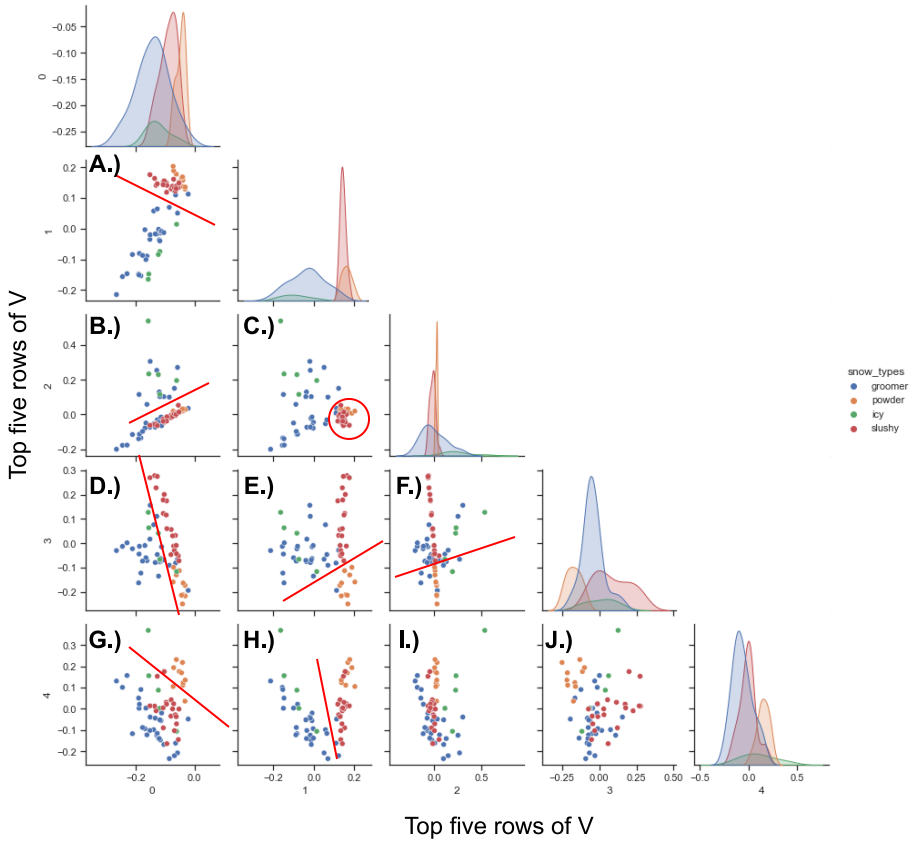

Figure 5.1: This figure provides comparisons between the rows of the V matrix coming from an SVD of the DFT ratio of $SG_{B||}$ and $\omega_{x,y,z}$. A,C,D, and H show distinct separation between hard and softer snows, while E,F, and G separate powder and slushy snows. I and J display the most mixing of snow types. B and C show a group of icy points separated from the main cluster of points, however it is mixed with some groomer data points. The red lines and circle indicate where the separation and grouping is occurring.

To better understand where the strain signals were coming from I performed the SVD and UMAP analysis to the DFT of the angular velocities and strain outputs separately. I studied the relationship between the two forms of actuation by putting the DFT of the raw angular velocities through the SVD UMAP analysis to see how the skier's actuation style (Tight or Big) would cluster when plotting the projected principal components. Figure 5.2A shows fast turns clustering at the top right side of the plot and slow turns on the bottom left, however no groups form due to snow-type.

Next I analyzed the frequency content of only the strain signals to see how the clusters compare to the angular velocity clusters. This comparison is shown in Figure 5.2 B, where

the clusters previously defined by turning style have now disappeared and transformed into clusters according to snow type. Clusters beginning to form due to snow type motivates the idea that the strain signal is impacted by the snow more than the skier's actuation style. To relate the two signals (angular velocity and strain), I used the DFT ratio motivated by its similarity to a transfer function. Figure 5.2C shows a cluster plot of the DFT ratio. This cluster plot shows a visual improvement from the previous plot, providing motivation to use the DFT ratios for the classifier.

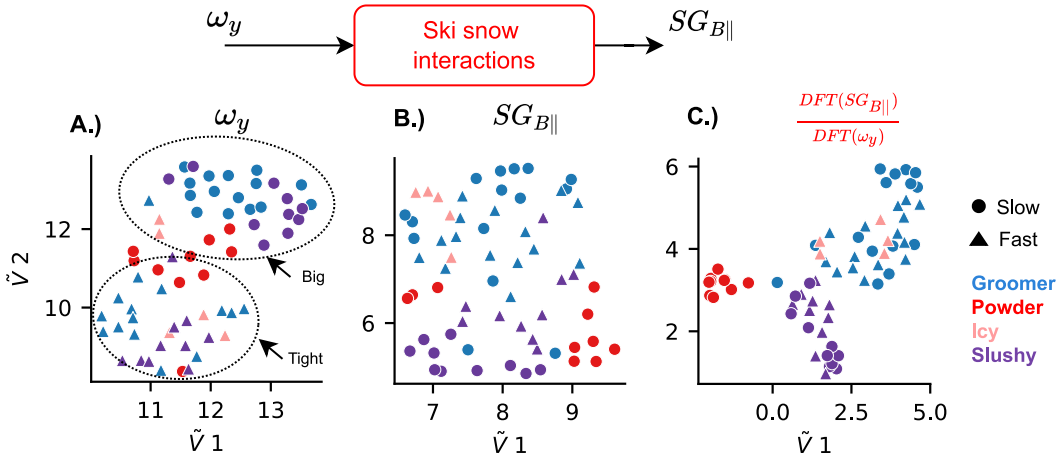

Figure 5.2: A.) Scatter plot showing separation of angular velocities based on turning speed. Xs represent fast turns and dots represent slow turns. B.) Shows that when just the strain signal is used, the turning styles mix snow type is the only clusters that happen. C.) Using UMAP to project the top five features the clusters are much more apparent.

5.2 Clustering Using Individual Strain Gauges

Figure 5.3 displays cluster plots from all the strain gauges using the DFT ratio analysis. This figure shows how the quality of clusters relates to the location on the ski. The strain gauge next to the binding and the front parallel strain gauge do not display clusters according to snow type while the gauges further towards the tip and tail of the ski form much more apparent clusters. Additionally, strain gauges oriented perpendicular to the ski separate out the powder snow much more than the parallel gauges. $SG_{B||}$ provides distinct separation between the softer snows and the harder snows, which is seen from the purple (slushy)

and powder (red) snow clustering at the bottom left of the plot while the other snows are located at the top right. These are all qualitative results, which we quantified using the Naïve Bayes classification algorithm.

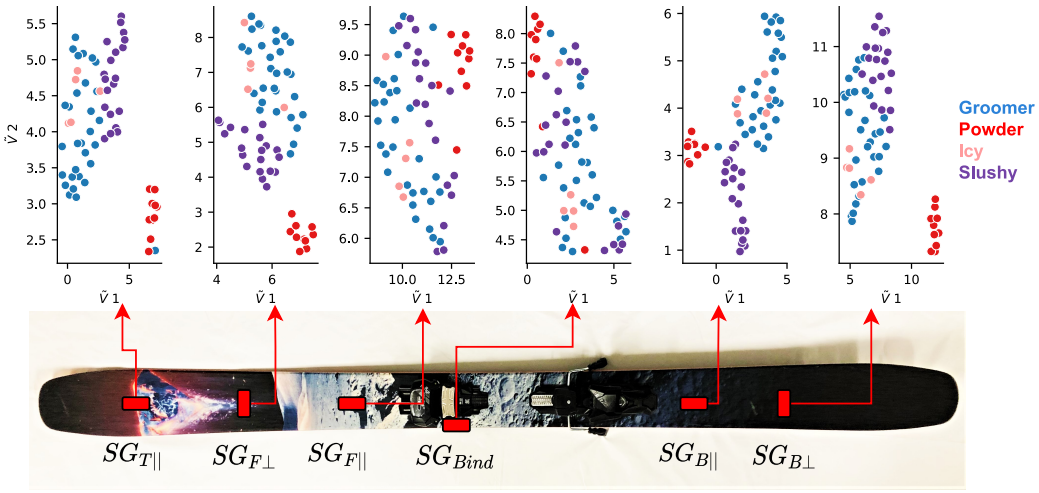

Figure 5.3: Clusters of "V" rows of SVD projected into a two-dimensional UMAP space. The strain gauges closer to the binding cluster poorly compared to the strain gauges closer to the tip which are most likely more affected by the snow and not the skier's input.

5.3 What do the DFT ratios tell us?

Plotting the actual DFT ratios from all the ski runs, provides insight into what the SVD is using to form the groups and what coherent shapes exist in the data. Figure 5.4 displays the DFT ratios from every ski run and all strain gauges. Each plot from Figure 5.4 is a ski run and each trace corresponds to a different strain gauge. All groomer DFT ratios in blue display a high frequency peak as well as the Icy run. The slushy and powder ratios have a low frequency plateau and a dip at around 0.5Hz. The slushy ratios have a more pronounced peak at 5Hz. The parallel strain gauges are plotted with solid lines, SG_C is plotted with triangles and the perpendicular strain gauges are plotted with circles. SG_C apears in the top section of each plot and across all snow types and ski runs it exhibits a peak at 5Hz. Both perpendicular strain gauges appear in the bottom portion of the plots due to a lower output amplitude. This lower amplitude in output is mot likely because

these two gauges are oriented perpendicular to the dominant bending axis. Because SG_C generally has a similar shape across all snow types, it will be more difficult for the SVD to pick out the differences.

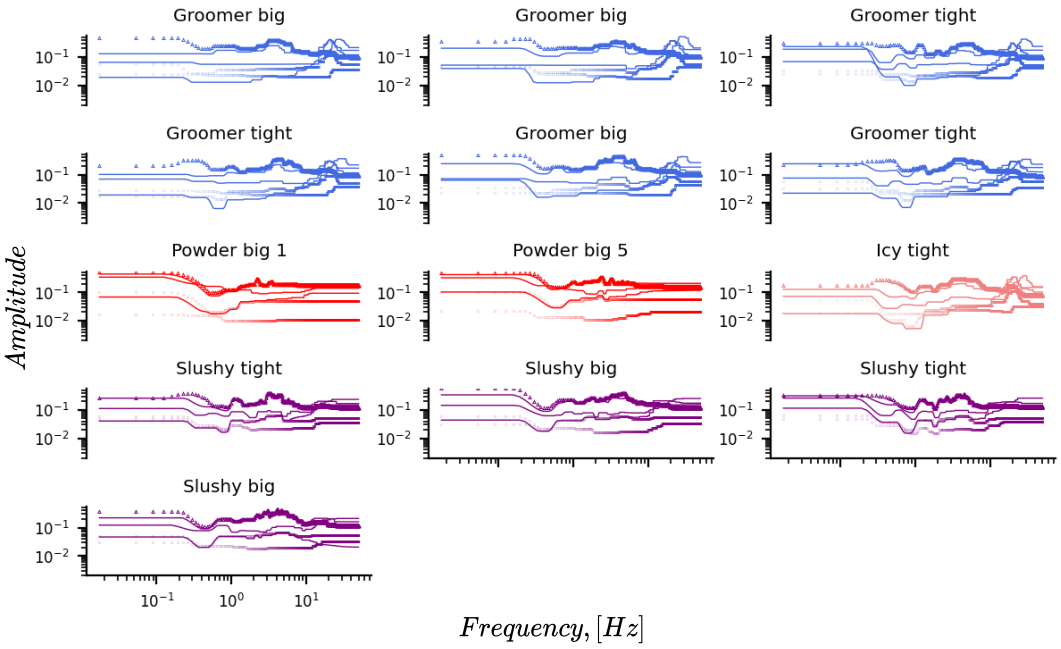

Figure 5.4: All groomer DFT ratios in blue display a high frequency peak as well as the Icy run. The slushy and powder ratios have a low frequency plateau and a dip at around 0.5Hz. The slushy ratios have a more pronounced peak at 5Hz. The parallel strain gauges are plotted with solid lines, SG_C is plotted with triangles and the perpendicular strain gauges are plotted with circles.

5.4 Naïve Bayes Classification

Based on the evidence in the previous section that strain gauges further from the binding perform better at separating snow types into groups than others and perpendicular strain gauges separate powder from harder snows, I took advantage of this by using different combinations of strain gauges for the Gaussian Naïve Bayes classifier. Using more strain gauges introduces more features to the classifier, potentially causing performance decrease after an optimal number of features. This is called *The Curse of Dimensionality* which states that arbitrarily increasing features leads to decreased classifier performance after a certain number of features are used [1]. Implementing a Gaussian Naïve Bayes classifier I

iterated through all the features at my disposal in order to find the peak performance of the classifier.

Performance of this classifier is displayed in Figure 5.5, where the maximum performance is shown along with the corresponding number of features used. When combining $SG_{T||}$ and $SG_{B\perp}$ the peak performance is 97.4% and the number of features used is 6. The combination of $SG_{F\perp}$ and $SG_{B||}$ provides a performance of 94.9% at 8 features. The combination of all strain gauges provides a performance of 97.4% with 10 features. Lastly using only the strain gauge at the binding, the classifier's maximum performance is 90% at 5 features. In each combination of sensors there is an optimal number of features, which if increased the performance decreases. To bring back the ideas relating to a physical transfer function I performed the classification using torque as input and then using frequency $* \omega$ as input to analyze the effect of the noise amplification due to processing.

The middle 4 plots display combinations of strain gauges and use frequency $* \omega$ as the input for the DFT ratio. Using these ratios results in reduced performance namely for SG_C with 90% accuracy at 23 features. The combination of $SG_{T||}$ and $SG_{B\perp}$ shows 92.3% peak performance at 8 features which is a lower maximum performance requiring 2 more features to achieve it. Using the combination of $SG_{F\perp}$ and $SG_{B||}$ is better than the pure angular velocity ratio, however, it uses 11 featuress to reach a maximum performance of 97.4%. When combining all strain gauges the classifier's performance drops to 93.2%.

Using the true input of calculated torque the results in the bottom portion of Figure 5.5, and offers the worst performance overall. The features needed to reach the maximum performances have increased and the maximum performances have significantly dropped.The low performance when using the DFT(torque) ratio is due to the amplified noise that results from the time derivative and high level of processing in the torque calculation. Using the DFT(angular velocity) provides the best performance because it is an unprocessed signal. The ratio of DFT(strain) and frequency times DFT(angular velocity) stills provides high quality classification results, and the slight decrease in performance is likely due to high

frequency noise being amplified by the multiplication of frequency.

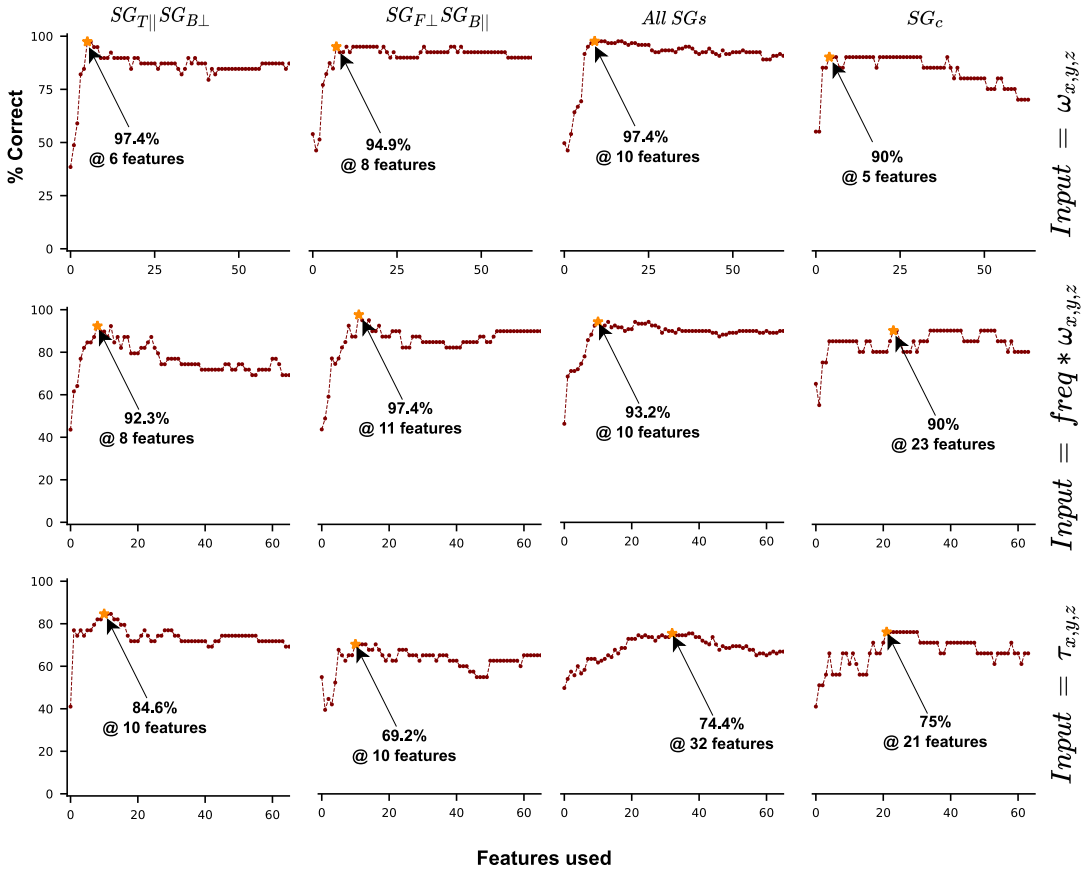

Figure 5.5: Percent accuracy of the Naïve Bayes classification using different combinations of strain gauges. Top: input is angular velocity, middle: input is frequency times angular velocity and bottom: input is toque calculartion.

Chapter 6

Discussion

Overall the SVD analysis creates visually distinct groups of the four snow-types. However, the icy data falls into the same category as the groomer data, on all strain gauges. $SG_{T||}$ has a cluster of icy data points that is surrounded by groomer but still independent, providing valuable information when classifying icy snow. The mixing of icy and groomer data, is caused by the fact that the icy data was also taken on groomed runs making the two data-sets inherently similar. Figure 6.1C and D show the DFT ratios for groomer and icy snow, which both share a similar overall shape with a high frequency peak, providing a reason for the mixing in the cluster plots.

Understanding where and how the clusters form can be done by inspecting the dominant shapes of the DFT ratios associated with each snow type. Looking at DFT ratios from $SG_{B||}$ in Figure 6.1C and D there is a high frequency peak which does not appear in the softer snows. Figure 6.1A does not have the low frequency dip or a high frequency peak distinguishing the powder snow from the other softer snow and the other two hard snows. Comparing the icy and groomer DFT ratios confirms that both snow types provide similar DFT ratios leading to close proximity clustering.

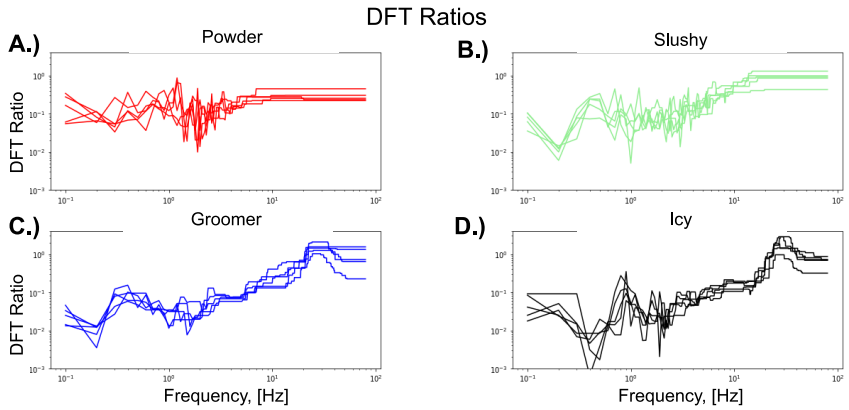

Figure 6.1: Each snow type exhibits a unique shape of DFT ratio. These unique shapes are what the SVD extracts and ultimately lead to the clusters forming.

As the location of the strain gauge gets closer to the binding the clustering quality decreases, as seen in $SG_{F||}$ and SG_C. The input onto the ski from the skier is closest to the binding and therefore has the potential to overwhelm the signal from the snow which is much more apparent in strain gauges closer to the tip and tail. The decrease in cluster quality could also be due to the increase in stiffness of the "beam" closer to the binding leading to smaller strain magnitudes. This conclusion provides evidence that future placement of strain gauges should be closer to the edges of the ski length. Figure 6.2 shows the percentage correct for each class of snow-type using 5 rows from the "V" matrix and cycling through 500 different random subsets of training and testing data. Powder snow depicted in red is performs very well which aligns with the cluster plots shown earlier in the results section. The icy class performs the worst overall most likely due to the fact that it is very similar to groomed snow. This result leads me to believe that the icy data is what is driving the classifier performance down when evaluating its correctness overall.

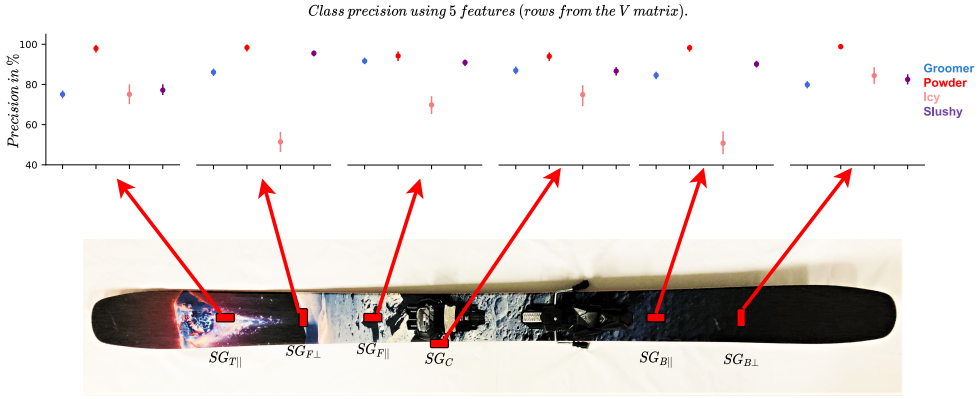

Figure 6.2: Classifier results on a per class basis are displayed with a 99% confidence interval. 500 different random training and test sets of data were used to achieve this statistical significance. 5 rows from the SVD "V" matrix were used as features.

The perpendicular strain gauges display greater separation of powder clusters from the other three harder snow types. One difference between the powder and hard snows is edge contact which could be causing a characteristic difference in the data due to torque about the y-axis. Fig. 6.3 A.) displays pure edge contact created by a hard snow surface resulting in a torque about the ski's y-axis. In Fig. 6.3 B.) however, the snow deforms to the ski and creates a layer of even support to balance the skier's force leading to minimal y-axis torque.

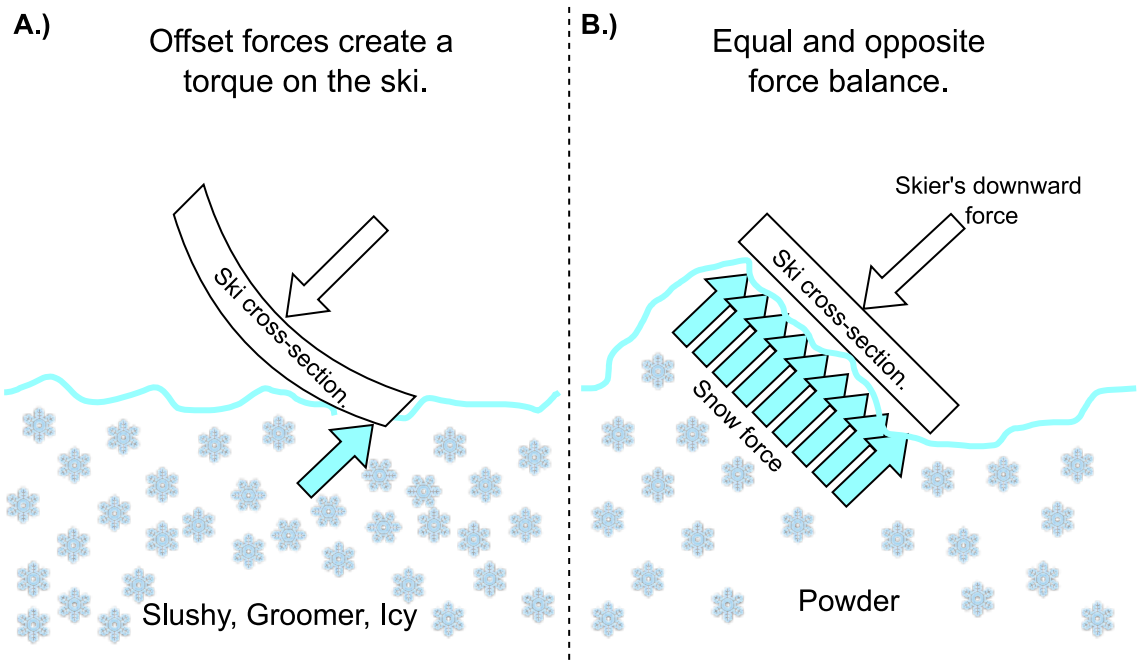

Figure 6.3: A.) Displays the ski making contact with a hard snow surface resulting in a torque being applied about its y-axis. In B.) a force balance is shown, because soft snow like powder will support the ski causing less deflection about the y-axis.

Using only two strain sensors, would make the data acquisition, much more manageable and compact, making it easier to collect more data on many different types of skis. Both perpendicular strain gauges show this hard/soft separation which means that a perpendicular strain gauge adds valuable information about the snow, however it does not make an impact whether it is placed on the front or the rear of the ski. This conclusion gives insight into future ski sensor design and placement, allowing future work to take advantage of the benefits of perpendicular strain gauges. In future work we would place two sensors only. Both sensors would be placed halfway between the binding and the tip of the ski, and would make use of both perpendicular and parallel orientations.

I wanted a more data driven approach to why the strain gauge right next to the binding did not cluster. To do this I performed a linear regression using the angular velocities to explain the strain gauge signals. Linear regression was performed on two turn segments of DFT data using a linear equation: $DFT(SG_C) \approx A * DFT(\omega)$ where A is a linear multiplier.

From this analysis I plotted the R^2 value of each regression in Figure 6.4. The plot of SG_C shows higher R^2 values signifying that this strain signal is more strongly related to the angular velocity signal, and therfore more snow-specific information.

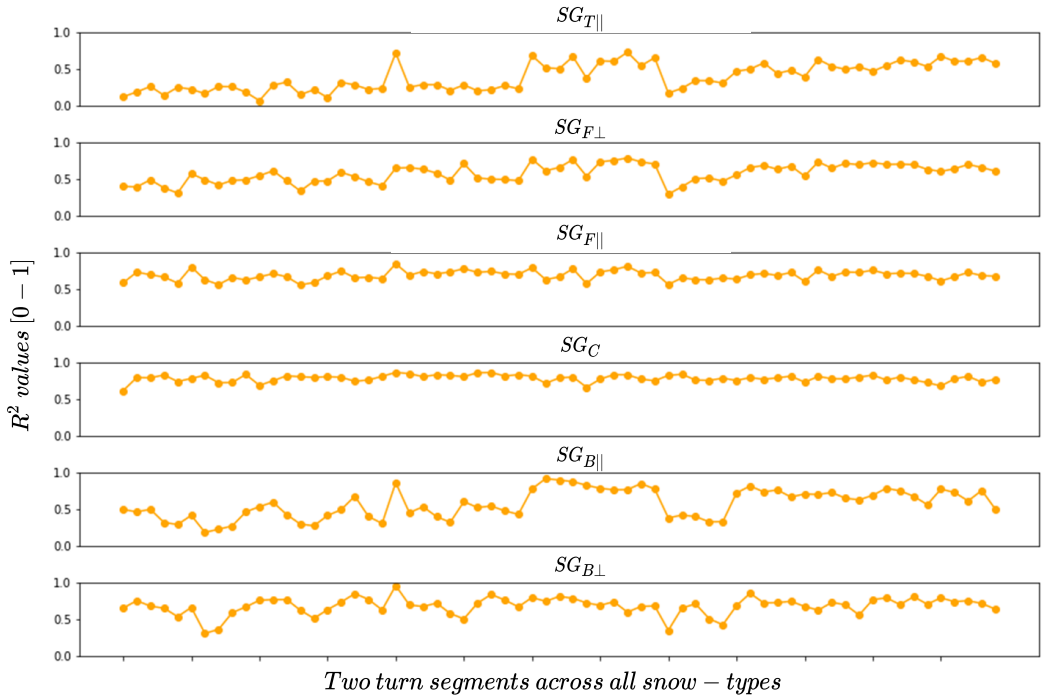

Figure 6.4: The R^2 values across all snow types are plotted and display a value close to one for SG_C. The other strain gauges are not well described by the angular velocity.

The performances of the Gaussian Naïve Bayes classifier are summarized in Table 6.1. I hypothesized that SG_C would have the poorest performance because visually the projected cluster plots did not form coherent clusters. Although SG_C displays the lowest performance, it is still relatively successful in classifying the test data, which signifies that even if visually the scatter plots do not form clusters according to snow-type, the combination of V rows provides the classifier with the necessary information for successful classification.

The combination of $SG_{T||}$ and $SG_{B\perp}$ performed the best. Although the combination of all strain gauges performs with the same maximum percentage, 10 features are used instead of 6. Requiring more features for classification, signifies that the lower 6 dimensions of V when all strain gauges are used provide less information than the combination of $SG_{T||}$

and $SG_{B\perp}$. Using less features is desirable because it leads to a more robust classification process, and reduces the risk of over fitting the data. Combining two perpendicular strain gauges results in a decrease in performance to 87.2%. However combining two parallel strain gauges results in 97.4% success rate only using 5 features.

Table 6.1: Summary of the classifier performances based on strain gauge combination. DFT ratio is: $\frac{DFT(SG)}{DFT(\omega_{x,y,z})}$.

Strain gauge combination	$SG_{T\|\|}$ $SG_{B\perp}$	$SG_{F\perp}$ $SG_{B\|\|}$	$SG_{B\|\|}$ $SG_{B\perp}$	All strain agauges	SG_C	$SG_{T\|\|}$ $SG_{B\|\|}$	$SG_{F\perp}$ $SG_{B\perp}$	$SG_{F\|\|}$ $SG_{B\|\|}$	$SG_{T\|\|}$ $SG_{F\|\|}$
Maximum performance	97.4%	94.9%	94.9%	97.4%	90%	97.4%	87.2%	94.9%	87.2%
Numbers of features at maximum performance	6	8	5	10	5	5	5	8	21

The results presented in this thesis show a clear method for classifying at least three of the snow-type: Powder Slushy and Groomer. Being able to perform this classification provides a method of localized snow classification which could be used to quantify snow grain size and therefore snow albedo providing relevance to the snow science community. Collecting this data would have to be part of a citizen science project where a modular strain sensing device would be placed on the skis of participants. Doing this would provide a more objective data set.

Additionally this work lays down the stepping stones for a closed loop ski that would change stiffness based on the snow-type it was skiing on. Having an adapting ski is beneficial to all skiers as well as manufacturers. This data can also be useful for ski manufacturers as it provides quantitative feedback on how much and at what frequency a ski is vibrating rather than relying on the current opinion of pro skiers.

On a broader scale this work provides insight into how strain information can be used to classify surfaces. Using DFT ratios of measurable signals to classify surfaces could also be

implemented in the self driving cars to provide a feedback system that tunes the suspension system making the car safer and more reliable on different surfaces. In order to make an effective prosthetic limb, the product must be as bio-mimetic as possible. To do this the leg would need to know what surface it is on in order to flex the correct muscles to achieve optimal balance and actuation.

Chapter 7

Conclusion

Humans and animals are very good at classifying surfaces by touch and sight, machines on the other hand, have a much harder time performing such a task. I set out to investigate this problem using strain information from a ski to classify four different snow-types on an alpine slope. The strain gauges placed on the ski provided signals directly related to the deformation of the ski which was caused by the different snow-types and skier torques. The skier's input is not a characteristic of the snow, requiring us to decouple the skier's input from the snow's impact on ski bending and vibration. I accomplished the separation of inputs by relating their frequency content similar to a transfer function analysis.

I developed a classifier using this relationship between the angular velocity and strain frequency content and applied a Singular Value Decomposition to the entire data set. Using the first 5 principal components from the SVD we were able to find groups forming according to snow type. The quality of the formed groups increased with distance from the binding and perpendicular strain gauges pulled the powder snow cluster further away from the harder snows. In future work and or design of ski sensing equipment, strain gauges would be placed further from the binding and only two gauges total would be needed. One gauge would be parallel and one would be perpendicular. To summarize, we have collected strain and IMU data from an alpine ski and successfully been able to classify the type of snow

being skied, independently of skiing style. This work lays out the preliminary results that prove the viability of using strain information to classify surfaces, and must be further investigated and validated with different skiers and terrain.

7.1 Future Work and Improvements

Although the work in this thesis led to concrete results for classifying snow based on frequency content of an actuated ski, there are many things that would be improved for future work. Some improvements are based on the findings of this work, and others are based on reflection and new knowledge coming from experience building hardware and performing experiments.

In future experiments I would choose ADCs with lower sampling rates. The current setup of 1KHz is impressive and reliable, however I believe a sampling rate of 200 would be sufficient as it captures the range of frequencies observed in the data collected. Having a strain gauge sampling rate of 200Hz would decrease the load on the microprocessors allowing a more compact setup, because more strain gauges could be connected to the same microprocessor.

Additionally, a 24-bit converter is necessary even though it is generally better to have more accuracy which is achieved by a higher conversion bit number. This high conversion number puts an unnecessary load on the microprocessors.

Future work also includes more validation of the results. Actuation of the ski in a different style other than fast or slow turning style may lead to different results because the actuation of the ski is based on the way ski is skied and the snow. This would be achieved by more data collections and having a wide variety of skier's with different skill levels and body types. A less experienced skier will have a different weight distribution and will likely not have consistent turns compared to an experienced skier. A taller or shorter skier as well as different weights will cause the ski to act differently.

Many times a skier will notice distinct sounds based on the snow-type they are skiing, especially when skiing over icy snow. To address the issue of icy snow mixing in with groomed snow the skis could be equipped with a microphone, which would pick up on the distinct sound difference present when skiing over icy snow. Using the frequency content of the sound signal coming from the microphone may increase the ability to split the icy and groomer snow-types.

In order to use this technique in practice further development of the hardware and on board software is needed. As stated an application for this algorithm is to classify surfaces. However to make this a system for live feedback the classifier would need to be on-board the ski or device doing the classification.

Bibliography

[1] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, no. 1, pp. 55–63, 1968. DOI: 10.1109/TIT.1968.1054102.

[2] C. F. Bohren and R. L. Beschta, "Snowpack albedo and snow density," *Cold Regions Science and Technology*, vol. 1, no. 1, pp. 47–50, 1979.

[3] S. G. Warren and W. J. Wiscombe, "A model for the spectral albedo of snow. ii: Snow containing atmospheric aerosols," *Journal of the Atmospheric Sciences*, vol. 37, no. 12, pp. 2734–2745, 1980.

[4] P. Cappa, G. De Rita, K. McConnell, and L. Zachary, "Using strain gages to measure both strain and temperature," *Experimental mechanics*, vol. 32, no. 3, pp. 230–233, 1992.

[5] E. Müller, "Analysis of the biomechanical characteristics of different swinging techniques in alpine skiing," *Journal of sports sciences*, vol. 12, no. 3, pp. 261–278, 1994.

[6] K. Ueda and A. Umeda, "Dynamic response of strain gages up to 300 khz," *Experimental mechanics*, vol. 38, no. 2, pp. 93–98, 1998.

[7] C. E. Page, D. Atkins, L. W. Shockley, and M. Yaron, "Avalanche deaths in the united states: A 45-year analysis," *Wilderness & Environmental Medicine*, vol. 10, no. 3, pp. 146–151, 1999.

[8] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, Ieee, vol. 1, 2001, pp. I–I.

[9] A. Sherman and M. H. Dickinson, "A comparison of visual and haltere-mediated equilibrium reflexes in the fruit fly drosophila melanogaster," *Journal of experimental biology*, vol. 206, no. 2, pp. 295–302, 2003.

[10] A. G. Akritas and G. I. Malaschonok, "Applications of singular-value decomposition (svd)," *Mathematics and computers in simulation*, vol. 67, no. 1-2, pp. 15–31, 2004.

[11] K. Baker, "Singular value decomposition tutorial," *The Ohio State University*, vol. 24, 2005.

[12] F. Rainer, W. Nachbauer, K. Schindelwig, and P. Kaps, "On the measurement of the stiffness of skis," *Science and Skiing III; Müller, E., Bacharach, D., Klika, R., Eds*, pp. 136–147, 2005.

[13] G. C. Foss and B. Glenne, "Reducing on-snow vibrations of skis and snowboards," *Sound and Vibration*, vol. 41, no. 12, pp. 22–27, 2007.

[14] T. Yoneyama, N. Scott, H. Kagawa, and K. Osada, "Ski deflection measurement during skiing and estimation of ski direction and edge angle," *Sports engineering*, vol. 11, no. 1, pp. 3–13, 2008.

[15] J. Dozier, R. O. Green, A. W. Nolin, and T. H. Painter, "Interpretation of snow properties from imaging spectrometry," *Remote Sensing of Environment*, vol. 113, S25–S37, 2009.

[16] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[17] Z. Wang and X. Zeng, "Evaluation of snow albedo in land models for weather and climate studies," *Journal of Applied Meteorology and Climatology*, vol. 49, no. 3, pp. 363–380, 2010.

[18] F. Pedregosa, G. Varoquaux, A. Gramfort, *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[19] Y. G. Guo, D. Z. Ren, and L. P. Xu, "The application of i2c bus digital sensor in airflow measurements," in *Applied Mechanics and Materials*, Trans Tech Publ, vol. 130, 2012, pp. 108–111.

[20] P. Morala-Argüello, J. Barreiro, and E. Alegre, "A evaluation of surface roughness classes by computer vision using wavelet transform in the frequency domain," *The International Journal of Advanced Manufacturing Technology*, vol. 59, no. 1, pp. 213–220, 2012.

[21] M. H. Muftah, S. M. Haris, K. Petroczki, and E. A. Khidir, "An improved strain gauge-based dynamic torque measurement method," *International Journal of Circuits, Systems and Signal Processing*, vol. 7, no. 1, pp. 66–73, 2013.

[22] F. A. Al-Wassai and N. Kalyankar, "Major limitations of satellite images," *arXiv preprint arXiv:1307.2434*, 2013.

[23] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.

[24] A. C. Adolph, M. R. Albert, J. Lazarcik, J. E. Dibb, J. M. Amante, and A. Price, "Dominance of grain size impacts on seasonal snow albedo at open sites in new hampshire," *Journal of Geophysical Research: Atmospheres*, vol. 122, no. 1, pp. 121–139, 2017.

[25] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, "A rewriting system for convex optimization problems," *Journal of Control and Decision*, vol. 5, no. 1, pp. 42–60, 2018.

[26] *Avalanche.org " snowpit*, Feb. 2018. [Online]. Available: https://avalanche.org/avalanche-encyclopedia/snowpit/.

[27] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426*, 2018.

[28] T. L. Mohren, T. L. Daniel, S. L. Brunton, and B. W. Brunton, "Neural-inspired sensors enable sparse, efficient classification of spatiotemporal data," *Proceedings of the National Academy of Sciences*, vol. 115, no. 42, pp. 10 564–10 569, 2018.

[29] R. C. Reid, P. Haugen, M. Gilgien, R. W. Kipp, and G. A. Smith, "Alpine ski motion characteristics in slalom," *Frontiers in Sports and Active Living*, vol. 2, p. 25, 2020.

[30] P. Virtanen, R. Gommers, T. E. Oliphant, *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020. DOI: 10.1038/s41592-019-0686-2.

[31] J. Mähönen, N. Lintzén, and J. Casselgren, "Portable bevameter for measuring snow properties in field," *Cold Regions Science and Technology*, vol. 182, p. 103 195, 2021.

[32] T. Sainburg, L. McInnes, and T. Q. Gentner, "Parametric umap embeddings for representation and semisupervised learning," *Neural Computation*, vol. 33, no. 11, pp. 2881–2907, 2021.

[33] F. Van Breugel, Y. Liu, B. W. Brunton, and J. N. Kutz, "Pynumdiff: A python package for numerical differentiation of noisy time-series data," *Journal of Open Source Software*, vol. 7, no. 71, p. 4078, 2022.

[34] *Home*. [Online]. Available: https://www.verispellis.com/.

[35] *Skis*. [Online]. Available: https://www.mechanicsofsport.com/skiing/equipment/skis.html.

[36] SparkFun, *Sparkfun 9dof imu breakout - lsm9ds1*, [Online; accessed April 9, 2022]. [Online]. Available: https://www.sparkfun.com/products/13284.

Appendix A

User Manual for the Data Logger

This appendix is a set of instructions on how to opperate the data logger propperly. It also contains information on what sensors are active and which connectors are active. This should be used to ensure successful data collection.

The data acquisition unit for the ski sense project collects data using 5 teensy 3.5 microprocessors. Each teensy is powered independently by a small battery pack. Three teensies record strain data, one teensy records GPS data, one teensy records IMU data, and all teensies collect pulse data. Each teensy records data to a separate designated onboard micro SD card. The duties of each teesny and some definitions are outlined here.

A.1 Teensy 1

1. Location: Second from the top (there is only one that is second from the top).
2. Functionality: Not functioning as a data collector, however it powers two ADCs which send data to other teensies.
3. Slots labeled 7-12 connect to this teensy.
4. LEDs: One yellow LED that does not always turn on. Not necessary to have working since this teensy does not collect data.

A.2 Teensy 2

1. Location: The protoboard at this location is labeled with a silver "#2". Fifth protoboard down top of board. The protoboard should be oriented in such a way that the column letters are right side up starting with j i h g f ....
2. Functionality: Records data from four strain gauges, and pulse.

3. Slots: 13,14,15,C.
4. LEDs: One yellow. This LED blinks when recording data.
5. Strain Gauges plugged in (in this order): 3,4,6,C.

A.3 Teensy 3

1. Location: The protoboard at this location is labeled with a silver "#3". Fifth protoboard down bottom of board. The protoboard should be oriented in such a way that the column letters are right side up starting with j i h g f .... This is the other teensy at the same level as teensy 2.
2. Functionality: Records data from four strain gauges, and pulse.
3. Slots: 1-6. Only 1-4 are used.
4. LEDs: One yellow. This LED blinks when recording data.
5. Strain Gauges Plugged in (in this order): 7,9,12,15.

A.4 Teensy 4 (Master)

1. Location: The very top teensy.
2. Functionality: Records data from one IMU, three accelerometers, and pulse.
3. Output: This teensy creates the pulse.
4. LEDs: One green. This LED blinks when recording.

A.5 Teensy 5

1. Location: The very bottom teensy.
2. Functionality: Records GPS, and pulse data.
3. LEDs: One green, one red, and one blue. Green signifies a connection with satellites. Red signifies no connection. Blue blinks when data is recorded.

Table A.1: The condition of all strain gauges and their conctors are tabulated. This table should be considered when setting up experiments.

Number	Strain Gauge Functionality	Slot Functionality	Notes
1, 3-6	Yes	Yes	
2	No	Yes	Data comes in but is a constant value, suggesting that the strain gauge is disconnected.
7-12	Yes	No	Slots connected to Teensy 1
13,14	Yes	Yes	Teensy 1 and 3 must be plugged in
15,C	Yes	Yes	