AUTOMATED FEATURE EXTRACTION FROM LARGE CARDIAC ELECTROPHYSIOLOGICAL DATA SETS, AND A POPULATION DYNAMICS APPROACH TO THE DISTRIBUTION OF SPACE DEBRIS IN LOW-EARTH ORBIT

by

John Jurkiewicz

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
in Mathematics

at

The University of Wisconsin - Milwaukee
December 2022

ABSTRACT

AUTOMATED FEATURE EXTRACTION FROM LARGE CARDIAC ELECTROPHYSIOLOGICAL DATA SETS, AND A POPULATION DYNAMICS APPROACH TO THE DISTRIBUTION OF SPACE DEBRIS IN LOW-EARTH ORBIT

by

John Jurkiewicz

University of Wisconsin - Milwaukee, 2022
Under the Supervision of Professor Peter Hinow

We present two applications of mathematics to relevant real-world situations.

In the first chapter, we discuss an automated method for the extraction of useful data from large file-size readings of cardiac data. We begin by describing the history of electrophysiology and the background of the work's setting, wherein a new multi-electrode array-based application for the long-term recording of action potentials from electrogenic cells makes large-scale readings of relevant data possible, opening the way for exciting cardiac electrophysiology studies in health and disease. With hundreds of simultaneous electrode recordings being acquired over a period of days, the main challenge becomes achieving reliable signal identification and quantification. In the context of this method of data collection, we set out to develop an algorithm capable of automatically extracting regions of high-quality action potentials from terabyte size experimental results and to map the trains of action potentials into a low-dimensional feature space for analysis. We establish that our automatic segmentation algorithm finds regions of acceptable action potentials in large data sets of electrophysiological readings. We use spectral methods and support vector machines to classify our readings and to extract relevant features. We are able to show that action potentials from the same cell site can be recorded over days without detrimental effects to the cell membrane. The variability between measurements 24 h apart is comparable to the natural variability of the features at a single time point. this work contributes towards a non-invasive approach for cardiomyocyte functional maturation, as well as developmental, pathological and pharmacological studies. As the human-derived cardiac model tissue has the genetic makeup of its donor, a powerful tool for individual drug toxicity screening emerges.

In the second chapter we consider the population of objects, largely considered debris, in the region of outer space close to Earth. the presence of this debris in Earth's orbit poses a significant risk to human activity in outer space. This debris population continues to grow due to ground launches, loss of external parts from space ships, and uncontrollable collisions between objects. We examine the background of human space launch, the current methods of tracking objects, and modelling work done to date. We propose a diffusion-collision model for the evolution of debris density in Low-Earth Orbit (LEO) and its dependence on ground-launch policy, to arrive at a computationally feasible continuum-based model. We parametrize this model and test it against data from publicly available object catalogs to examine timescales for uncontrolled growth. Finally, we consider sensible launch policies and cleanup strategies and how they reduce the future risk of collisions with active satellites or space ships, along with considering extensions of the model to directly account for launch policy determination through minimization of certain functionals.

To my loving wife Gabrielle
for her unwavering support of my studies, even in the darkest times,
and to Beanie.
I miss you. I hope you’re proud of your daddy.
AMDG

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

AP	action potential
FP	field potential
MEA	multi-electrode array
hiPSC-CMs	human induced pluripotent stem cell derived cardiomyocytes
APD/APD$_m$	action potential duration (to m % level)
BCL	basic cycle length
DI	diastolic interval
DFT	discrete Fourier transform
SVM	support vector machine
LEO	Low-Earth Orbit
MEO	Medium-Earth Orbit
GEO	Geostationary Orbit
TLE	NORAD Two-Line Element Set

ACKNOWLEDGEMENTS

I wish to thank the reviewers of both manuscripts constituting this dissertation for their thorough and considerate comments that led to the full works below. I also want to thank Prof. Kerri Cahoy (MIT), Prof. Colin McInnes (University of Glasgow) and Dr. Jer-Chyi Liou (NASA) for their assistance in answering questions and sharing information.

Now on to the personal remarks.

I want to thank my wonderful advisor Peter Hinow for his constant teaching and consideration over the last 4 years of us working together. To say that the last four years have been quite possibly the darkest and most difficult of my life is an understatement. Through numerous personal emergencies and even the COVID-19 pandemic, all of which have negatively impacted my performance as a graduate student, he never once expressed dissatisfaction with my work or performance, which is the best support I could have asked for. I am forever in his debt for the lessons that he has taught me, and I doubt I can ever pay him back. I have him to thank for me being the mathematician I am today.

I want to thank my wonderful committee - Profs. Jeb Willenbring, Istvan Lauko, Gabriella Pinter, and Lijing Sun - for their service to me. Their comments during my proposal hearing shaped many aspects of my future research, and they have always been ready and willing to help with anything I have asked for.

I also wish to express my gratitude to my family: my parents, who through the entirety of my academic journey have always supported me, never pushing but always encouraging me to drive myself forward to the accomplishments they knew I could reach; my sisters Kathleen and Colleen, who have provided innumerable instances of assistance as I try to navigate the worlds of writing and academia; my brothers-in-law, the Matts, for always showing interest in my work; my nieces and nephews (Eva, Joey, Rosie, Walter, Maria and Tacy) for always believing in me; and my in-laws Susan and Paul, for their support and for always being patient with the time I've had to dedicate to my studies when we've gone to visit them out-of-state.

I owe a debt to all of my teachers who have helped me overcome my anxieties and driven me to explore my love of mathematics. In addition to the service of my wonderful committee members, many of whom have been my teachers in the past, I have no doubt that I would not be completing this dissertation wer it not for the encouragement of Bruce Wade, Eric Key, Robert Wood (of the Physics department) and my Calc II TA from spring of 2012, Lauren Williams (she had a way of making the class seem fun and low-stress which enabled me to explore my love of math at that early stage). Every person I have encountered during my education here has left a mark and spurred my love of learning. Though my schooling is coming to a close, the examples these individuals have given me has shown me that my education will never end.

And finally, though in no way least importantly, I want to thank my wife Gabrielle. She is my rock. Without her, I would not have weathered the pandemic with any shred of sanity intact. I know for a fact that I would definitely not have reached the successful end of my schooling here without the support she has shown me. Through all of the dark nights, when I doubted my own abilities and my future, and the horrible times when quitting was the only thing I wanted to do, she has been there, sitting with me in the dark and holding my hand, never letting me fall too far. I love you Gabby. I only wish I could show you even a fraction of the love you have shown to me. You have made me the luckiest man in the world. All of this is for you.

1 Automated Feature Extraction from Large Cardiac Electrophysiological Data Sets

In this chapter, we review some preliminaries from electrophysiology, signal processing and machine learning to understand the state of electrophysiological research at the time of completion of this work. We then propose an automated paradigm to efficiently analyze large sets of data obtained via microelectrode array (MEA) experimentation, partition the data into useful segments and extract features. The method relies on analysis of the time derivative of the electrical signal from the cardiac cell, called action potential. We further utilize this paradigm to extract features from a multitude of ethically-sourced human cardiac cells for purposes of statistical analysis. We apply the algorithm to extract data from the same cell site at different time-stamps to analyze the similarity for repeated physiological measurements. The method is capable of determining signal quality for virtually any type of experimental regimen for data collection, not exclusively MEA experimentation.

1.1 Preliminaries

1.1.1 Cardiac Electrophysiology

In the human circulatory system, the heart functions as an electromechanical pump powering the circulation of blood throughout the body. The cardiomyocytes - the excitable cells of the heart muscle - possess the ability to mechanically contract when the electrical potential of their environment changes. Various ion species (specifically K^+, Na^+, and Ca^{2+}) pass through the cell wall via ion channels, specialized proteins the manage the flux of these species. Crucially, these ion channels are voltage-gated, meaning the flux is dependent upon the potential difference between the cell interior and the cell exterior. Therefore, once the potential builds to a certain critical value, the ion channels open and allow a sudden influx of the ion species, quickly changing the internal voltage of the cell and triggering contraction of the cell. Once contraction occurs, the ions are allowed to leave the cell via the same ion channels and the process begins anew. Cells originating from different regions of the heart possess differing ion channels and thus different contractile behaviors; the electromechanical behavior of cardiomyocytes from the ventricles differs from that of myocytes in the ventricles. This distinction in behavior allows for the complex pumping behavior of the cardiac organ.

Measurement of the electrical behavior of the heart thus allows for insight into the mechanical behavior of the cell, and has had a long and storied history since the foundation of the field of electrophysiology (Hodgkin and Huxley, 1952; Harrington and Johnson, 1973). There are two types of electrical signal that are typically measured: **field potential (FP)** and **action potential (AP)**. Field potential is similar to a medical electrocardiogram; it is the electrical potential difference measured across the entire bulk of cardiac tissue, irrespective of how many cells constitute the tissue mass, and is thus something like an integral or aggregate measure of electrical behavior. Action potential is specific to one cell - it is the measure of the potential difference across the cell wall of a single cell. Due to the electromechanical behavior of cardiac cells, the action potential is the more important measure on a basic level, although field potential has been shown to be correlated strongly to the underlying action potential (Tertoolen et al., 2018).

Figure 1 demonstrates a typical action potential as a function of time. The **resting potential** is the baseline voltage the signal rests at between beats. When the voltage-gated ion channels trigger and allow for the influx of ion species, the transmembrane potential demonstrates a steep monotonic increase called **depolarization**. Upon reaching a peak voltage, the potential begins to monotonically decrease in the **repolarization phase**; cardiomyocytes from some areas of the heart exhibit a **plateau phase** during repolarization where the potential is essentially constant for a noticeable period of time. From this figure, certain data can be computed that encodes the morphology of the action potential.

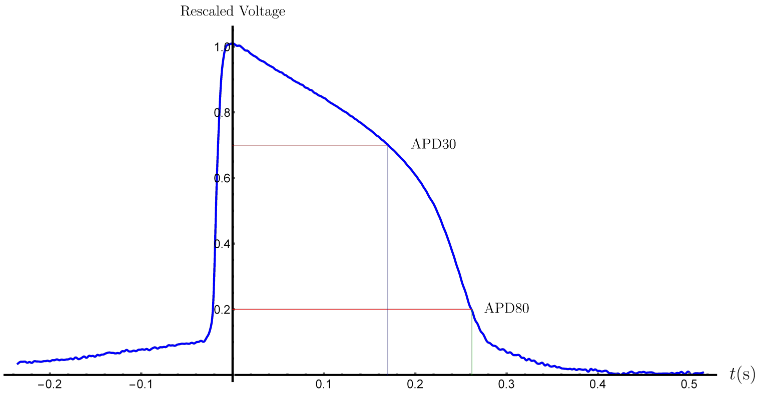

Figure 1: A typical action potential with voltage rescaled to rest at 0 and peak at 1 at time $t = 0$.

The **action potential duration** or **APD** is a measure of the temporal spread of the potential; typically it is reported as APD_n, which is the duration from peak voltage to n% repolarization (APD_{30} and APD_{80} are demonstrated in Figure 1). The comparative values of APD_{30} and APD_{80} give insight into the morphology of the potential wave - a ratio close to unity indicates a steep depolarization, while a lower ratio is indicative of a more gradual decay. Not shown in this figure is another important quantity, the **basic cycle lengthb (BCL)** which is the length of time between successive depolarizations - essentially the length of time between heartbeats.

Due to this close connection between electrical and mechanical behavior in the heart, the study of the electrical impulses can give valuable insight into the contractile behavior of cardiac tissue, as well as the emergence of arrhythmias (Anumonwo and Pandit, 2015; Qu et al., 2018).

Direct measurement of action potentials has proven to be time-consuming and technical, due to the need for intracellular access. Since the foundation of the field of electrophysiology with Hodgkin and Huxley, the voltage-clamp technique has been the gold standard (Fozzard and Beeler, Jr., 1975). Two electrodes are inserted into a single cell - a recording electrode and a current-bearing electrode. By varying the current applied to the cell, the transmembrane voltage can be kept at a fixed value and the complete behavior of the ion channels investigated directly. This allows for very detailed investigation of the action potential and enables precise control over the experimental parameters. However, the intrusive nature of the measurement means the cell is destroyed after measurement, so repeated experiments on the same specimen is not possible. Additionally, the technicality of the method means it is time-consuming and relegated to specialists.

From a reading of action potential, one can extract certain features of the action potential morphology. AP readings are analyzed by encoding the data stream as sequences of features (or biomarkers) that characterize the beat-by-beat behavior of the cells. These features as well as their counterparts in frequency space encode trends in the behavior of the cells. APs are typically encoded as action potential duration at $m\,\%$ (APD$_m$), the time from peak voltage until the cell has repolarized to $m\,\%$ of its resting potential. Additionally, differences and ratios of these features are commonly used (Britton et al., 2017). Using these features, the AP data can be projected into a low-dimensional feature space while still retaining the most relevant information for clinical purposes (Cantwell et al., 2019).

Mathematically, the behavior of action potentials is governed by ordinary differential equations, variations of the initial Hodgkin-Huxley model (Hodgkin and Huxley, 1952). By treating the cell wall as a capacitor and differentiating with respect to time, one obtains

$$\frac{\partial V}{\partial t} = \frac{1}{C_m}\left(\sum_k I_k\right)$$

Where C_m is the capacitance of the cell membrane, and I_k is the current or ion flux of the $k^t h$ ion species. The currents themselves are governed by time-dependent systems of non-linear differential equations as well. While the initial experiments of Hodgkin and Huxley focused on giant squid axons, the model is equally applicable to excitable cardiac cells, and subsequent refinements of the model have fine-tuned the currents to reflect the ion species present in the heart. This in turn leads to greater complexity of the model; the work of Grandi et al. (2010) accurately describes ventricular action potentials at the cost of utilizing a system of over 30 differential equations.

AP waves in cardiac tissue can also be measured indirectly, by means of recording field potential signals. Recent work has shown the feasibility of multi-electrode arrays (MEA) (Meyer et al., 2012) in the large-scale efficient recording of both AP and FP readings (Edwards et al., 2018; Zlochiver et al., 2019; Hayes et al., 2019; Balafkan et al., 2020).

1.1.2 Yamanaka Factors and hIPSCs

A determining factor in the applicability and efficiency of AP-related assays is the availability of ethically-sourced cardiac tissue for experimentation. The historical precedent in electrophysiology has been the use of animal cells cultured from non-human mammals (Mathur et al., 2016). Nonetheless, differences between nonhuman models and humans could contribute to clinical trial failures despite ideal nonhuman results (Li et al., 2020) - there are limited conclusions that can be drawn from comparisons to non-human tissue. Embryonic stem cells, while useful, carry significant ethical concerns. The pioneering work of Yamanaka (Takahashi and Yamanaka, 2006) has provided a new avenue of cardiac research by allowing for human cardiac tissue with little to no ethical concern. Yamanaka's work (which won the Nobel Prize) concerns the reprogramming of adult cells into stem cells - certain genetic factors (Yamanaka Factors) can be inserted into healthy adult tissue which reprograms the cells back to a pluripotent stem cell stage (called **human induced pluripotent stem cells (hIPCs)**), which can then be re-differentiated as needed. This enables large cell cultures of human tissue to be made available for research purposes.

Much recent work has shown potential for human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) (Zwi et al., 2009) to allow for drug testing and disease modeling on viable human cardiac tissue in a controlled environment (Braam et al., 2010; Tertoolen et al., 2018; Edwards et al., 2018; Gintant et al., 2019; Kussauer et al., 2019). Further,the genetic makeup of hiPSC-CMs is identical to that of their donor which makes them a tool to test personalized drug responses (Strauss and Blinova, 2017). While

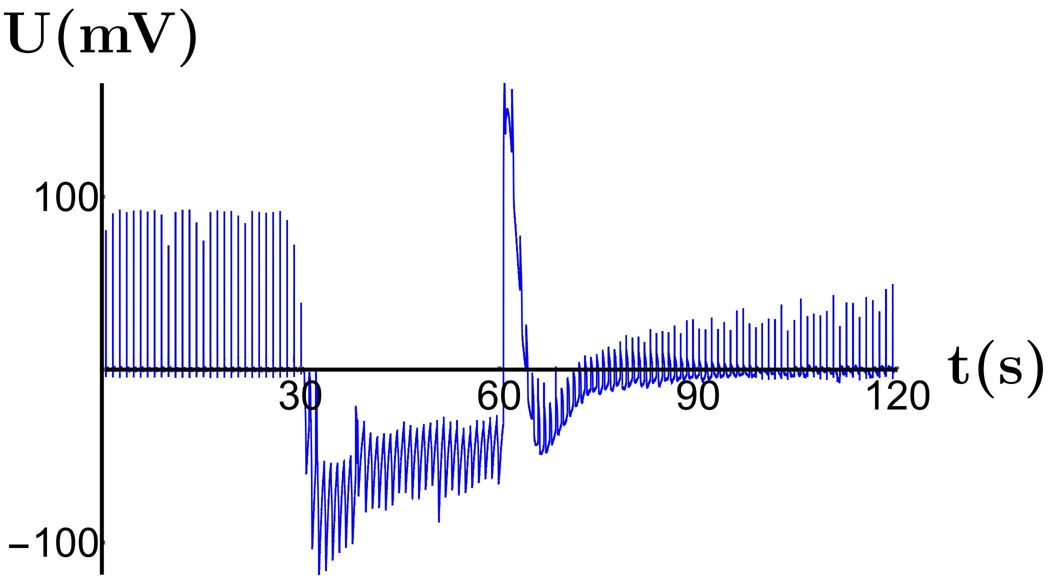

Figure 2: Raw data stream from a single electrode in an MEA. Prior to 30 s, the signal consists exclusively of FPs. From 30 s to 60 s, the electroporation stimulus is applied. From 60 s to 120 s, the signal includes APs that gradually degrade back to FP morphology.

there are concerns regarding the level of maturity of these hiPSC-CMs ranging from fetal to neonatal in development (Zhu et al., 2017), the possibility for safe human-tissue drug testing demands further research.

1.1.3 MEA Methods in Electrophysiology

A microelectrode array (MEA) consists of a normal experimental plate well wherein each well is augmented with an array of electrodes in the base of the plate. By reading potential differences between each electrode and reference electrodes in the side of the well, electrical measurements can be made regarding the material in the plate.

When cardiac cells are plated in the well, electrophysiological readings are possible. Typically, readings of FP are made due to the closed nature of the cells. However, MEA technology allows for the application of electrical stimuli to the cells to induce what is called **electroporation**, wherein the stimulus shocks the cell membrane and opens a hole in the cell wall, enabling access to the inside of the cell and thus allowing for transmembrane AP reading.

MEA technology has a number of advantages over the standard patch-clamp techniques of electrophysiology. Firstly, it has been shown (Edwards et al., 2018) that the electroporation stimulus is non-destructive - the cell remains alive during and after the application of stimulus, so repeated experiments and readings can be performed on the same cell site to measure time-dependent effects of treatment. Additionally, an MEA is highly parallel in its data collection - a standard plate well of 24 wells, each outfitted with 12 electrodes, creates 288 independent data streams from distinct cell sites simultaneously, collecting data much more efficiently than the single-cell patch-clamp technique, at the expense of voltage control and detailed knowledge of the makeup of the transmembrane current.

There are certian drawbacks to the MEA approach as well. Regardless of the experimental setup, access to AP data involves delivery of stimulating electrical impulses to the tissue, resulting in large regions of data without electrophysiological meaning. Further, electroporation does not offer permanent intracellular access, as the membrane nanopores close in time. This results in a degradation of the signal as the AP trace reverts to an FP morphology, see Figure 2. Electroporation may even fail entirely for a given cell and yield no usable AP data. The large volume of data generated thus requires a robust segmentation protocol to extract usable data. Manual segmentation is time-consuming and error prone for even modest data sets and becomes quickly infeasible as the volume of data grows. Therefore, the need for an automated segmentation method is imminent.

1.1.4 Signal Processing

We will consider a **finite discrete signal** to be a vector $\mathbf{U} = (x_i)_{i=1}^{N}$ in $\mathbb{R}^n$. In most practical applications, including the present consideration, these discrete signals arise as a consequence of evenly-spaced temporal sampling of a continuous function or time, and thus serve as a digital representation of that analog signal. For example, the vector $(1, 0, -1, 0, 1, 0, ...1)$ is an analog signal representing the continuous sinusoidal function

$f(t) = \cos(2\pi t)$.

The relative "size" of a signal or vector is measured via the **norm** of the vector. We consider two norms: the $\mathbf{l_2}$ **norm** of the vector $\mathbf{U}$, which measures the magnitude of the difference in the signal's deviation from 0 is defined by

$$||\mathbf{U}||_2 = \left(\sum_{i=1}^{N} |x_i|^2 \right)^{1/2}$$

and the $\mathbf{l_\infty}$ **norm**, measuring the greatest magnitude of the signal, is defined by

$$||\mathbf{U}||_\infty = \max_{1 \le i \le N} \{|x_i|\}$$

where the absolute values are understood to mean the modulus in the case of a complex-valued signal.

Given a discrete signal $\mathbf{U} = (x_i)_{i=1}^{n}$, we define its **discrete Fourier transform** (DFT) as the vector $\hat{\mathbf{U}} = (X_i)_{i=1}^{N}$ in $\mathbb{C}^N$ where (Oppenheim and Schafer, 1999)

$$X_k = \sum_{n=1}^{N} x_n e^{-2\pi i k n/N}, \quad k = 1, \ldots, N.$$

This is the discrete analog of the continuous Fourier transform. This creates a complex-valued vector: the argument $\frac{2\pi k}{N}$ represents the frequencies of even and odd sinusoids that sum to the signal, and the modulus $|X_k|$ indicates the relative contribution these sinusoids to the overall amplitude of the signal. In this way, the DFT also functions as a unitary (thus invertible) transformation $\mathfrak{F} : \mathbb{R}^N \to \mathbb{C}^N$. The **inverse DFT (iDFT)** is therefore the vector $(Y_i)_{i=1}^{n}$ defined by

$$Y_k = \frac{1}{N} \sum_{n=1}^{N} X_n e^{2\pi i k n/N}, \quad k = 1, \ldots, N.$$

Which recoups the original vector. Note the complex-conjugate symmetry in the DFT: $\overline{X}_k = X_{N-k}$. Thus the DFT is determined entirely by the constituent frequencies up to one-half the sampling frequency.

The **spectrum** of the discrete-time signal $\mathbf{U}$ is the set of all complex arguments appearing in the DFT. The **bandwidth** of the signal is the range of the spectrum, from lowest to highest frequency.

More generally, the z**-transform** of a signal $\mathbf{U} = (x_j)_{j=1}^{(}N)$ is the function $X(z)$ of the complex variable z defined by

$$X(z) = \sum_{j=1}^{N} x_j z^{-j}$$

Note that the DFT is vector formed by the evaluation of this function at the evenly-spaced frequencies $z_k = \frac{2\pi k}{N}$. This is an invertible transformation with the Laplace-like property that

$$\mathbf{U_1} * \mathbf{U_2} \mapsto X_1(z)X_2(z)$$

thus allowing for simplification of some properties in the frequency domain.

In realistic scenarios the measurement of a signal, either discrete or continuous, is affected by the presence of noise - thus the recorded signal $\tilde{\mathbf{U}}$ is $\tilde{\mathbf{U}}(t) = \mathbf{U}(t) + w(t)$ where $w(t)$ is an undesired function changing the recorded output from the ideal $\mathbf{U}$. The task of recovering the signal $\mathbf{U}$ from the recorded $\tilde{\mathbf{U}}$ then becomes the removal of w in such a way as to leave $\mathbf{U}$ as uncolored as possible.

A **filter** is a mathematical construct idealizing a realized circuit element which affects the amplitude of sinusoids in accordance with their frequencies. Thus, when applied to a signal, the filter will selectively attenuate some of the frequencies in the bandwidth of the signal according to its frequency response. The **passband** of the filter is the portion of the frequency domain that is unaffected by the filter's application, while the **stopband** is that portion which is attenuated essentially completely. Each filter is determined by a transfer fuction in the z-domain of the form

$$H(z) = \frac{b_0 + b_1 z^{-1} + ... + b_N z^{-N}}{a_0 + a_1 z^{-1} + ... + a_M z^{-M}}$$

Where the **order** of the filter is $\max(N, M)$. The filter's behavior is completely determined by the coefficients b_n, a_m. The transfer function encodes the effect of the filter on the frequency components of the signal. Thus in the frequency domain, the application of the filter amounts to multiplication of the original signal's z-transform with the transfer function. In the time domain, filter application amounts to recursive linear convolution of the signal and previous filtered values. An interesting effect of this convolutional application is that filters introduce a phase delay into the signal which is unavoidable for causal filters (i.e: filters depending only on past values of the signal). In applications where the signal is known in advance and not generated in real-time, an averaging process of forward and backward convolutions can be applied to negate the phase difference (as implemented in many signal processing software toolboxes).

Different choices of coefficients give differing behaviors for the filter on the passband and stopband. The specific choice of the **Butterworth filter** creates a filter which is completely flat (e.g. evenly applied) in both the passband and the stopband at the expense of a shallow transition between the bands, whereas the Chebyshev filters have steep transitions but features "ripples" of uneven application - either unwanted attentuation in the passband (Type I), or underattenuated frequencies in the stopband (Type II). (Oppenheim and Schafer, 1999)

1.1.5 Support Vector Machines

Suppose an n-tuple **x** in a subset $\Omega\mathbb{R}^n$ represents some collected data that is to be used in the analysis of a certain application. Suppose further there are two broad categories D_1, D_2 into which all possible data from this application fall - essentially a partition of $\Omega = D_1 \cup D_2$ into two disjoint classes according to the classification of the data. The task of separating Ω into this binary partition is the essence of the **binary classification problem**. In this work, due to the nature of the data generated and the needs of the below-mentioned applications, we focus exclusively on data that is **linearly separable** -

that is, the groupings of considered data in the classes D_1 and D_2 are separated by affine hyperplanes of the form $\mathbf{w} \cdot \mathbf{x} + b = 0$.

The automated application of the linearly separable binary classification problem is usually performed via the use of a classifying function. We assign any data in D_1 an arbitrary classification value of -1, and a value of $+1$ to the data in D_2. We classify an arbitrary datum $\mathbf{X}$ into either D_1 or D_2 by computing the classification value $y = \pm 1$ by

$$y(\mathbf{x}) = sgn(\mathbf{w} \cdot \mathbf{x})$$

Thus the task of classifying data is reformulated as the choice of $\mathbf{w}$ and b - respectively called the **weights** and the *bias* for the classifier.

Computing these hyperplanes is typically done via training - that is, utilizing a curated data set $(\mathbf{x}_i, y_i)$, called the **training set** where pre-classified data $\mathbf{x}_i$ are given with their classifications. While general training algorithms exist based on minimizing the error of the classification algorithm (adjusting the weights and bias in order to minimize the number of incorrect classifications of data in the training set), the **support vector machine** paradigm seeks to compute the hyperplane of maximal margin - the hyperplane with the greatest distance between it and the training subsets of D_1 and D_2. Our SVM is trained on this data via quadratic programming (Decoste and Schölkopf, 2002).

1.2 Statement of the Problem

Cardiac electrophysiological readings from an MEA at even reasonable sampling rates can easily create data streams into gigabyte range. Unfortunately, not all of the data generated, even for satisfactory experiments, is fully useful. Further, feature extraction from the useful data takes additional time and effort.

Therefore, we require an automated method to examine large data sets, extract the useful segments from each electrode, and compute classical features of action potential

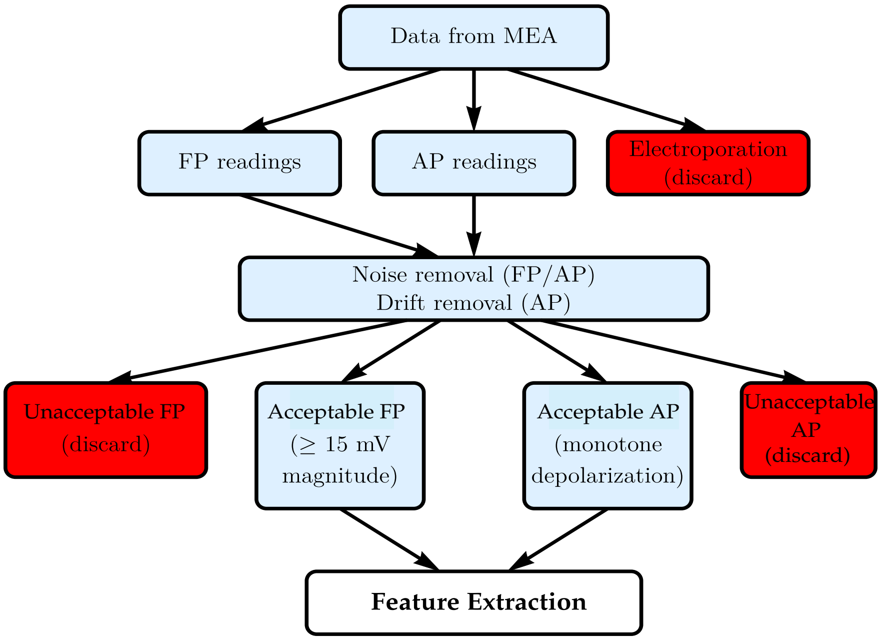

Figure 3: A flow chart showing the workflow for the segmentation algorithm.

for later analysis. Further, the presence of electrical artifacts from experimentation must be removed prior to any processing of the data

1.3 Data Segmentation and Feature Extraction

There is a natural division of labor for the processing algorithm to follow,as demonstrated in Figure 3. After data collection, regions of field potential and action potential have to be identified and extracted. After electrical artifacts (namely noise and baseline trend from electroporation stimulus) are removed, the data must be submitted in order to determine if it is acceptable for analysis or not. Once acceptable data is recovered, features can be extracted and analyzed.

1.3.1 Automated Classification for FP vs. AP

To classify portions of the signal as FP or AP, we look in the Fourier space, or frequency domain. The spectra of FP and AP signals are shown in Figure 5. We see from these

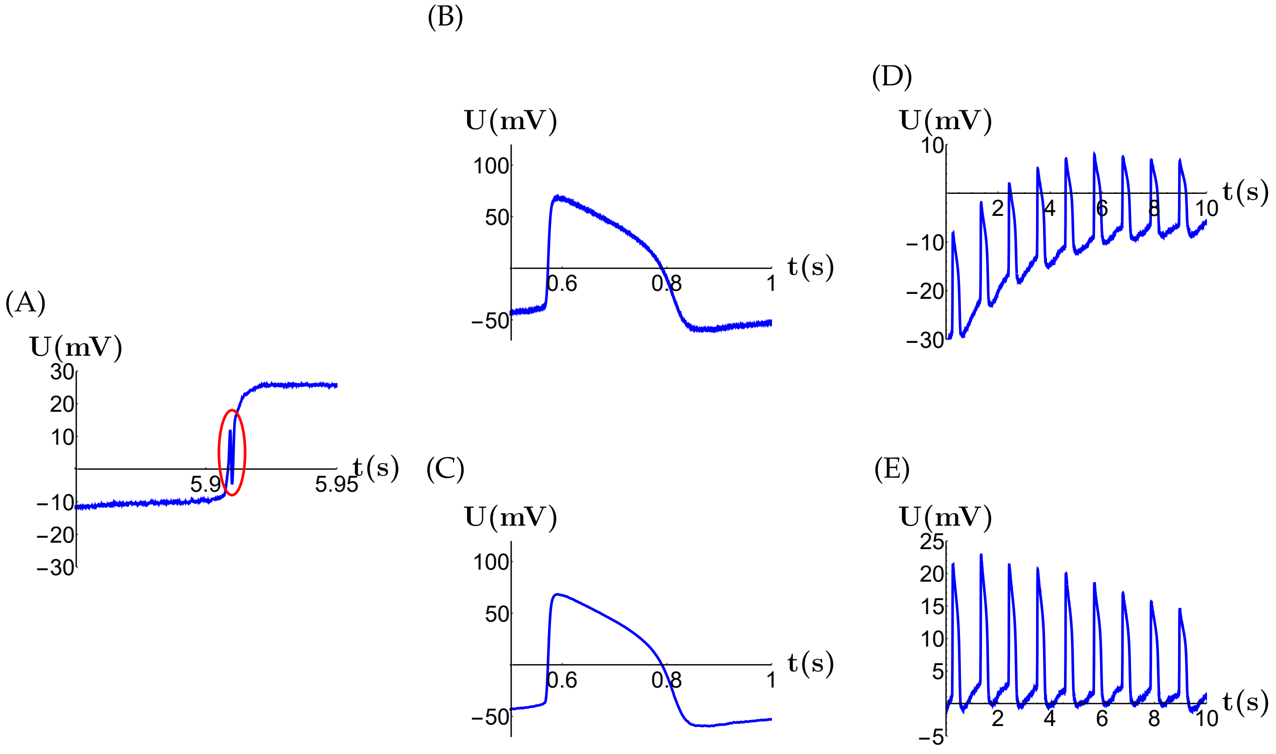

Figure 4: **(A)** Unacceptable AP trace featuring non-monotone depolarization. **(B) & (C)** AP trace before and after application of digital filter. **(D) & (E)** AP trace before and after removal of baseline trend.

figures that the bandwidths of the signals are distinct enough to allow a classification in the Fourier space - the bandwidth of FP is approximately 1300 Hz, while the bandwidth of AP is approxmiately 200 Hz.

Let $\mathbf{U}$ denote the voltage signal, $\hat{\mathbf{U}}$ its DFT, and $\hat{\mathbf{U}}_{\mathbf{200}}$ be the portion of the DFT with argument below 200 Hz. We encode $\hat{\mathbf{U}}$ by its L^2-mass (or energy) and the fraction of the mass below 200 Hz, i.e.

$$\mathbf{U} \mapsto \left(||\hat{\mathbf{U}}||_{\mathbf{2}}, \frac{||\hat{\mathbf{U}}_{\mathbf{200}}||_{\mathbf{2}}}{||\hat{\mathbf{U}}||_{\mathbf{2}}} \right). \tag{1}$$

The second component of this encoding allows for distincion based on bandwidth, while the first allows for a time-domain distinction as well (noting that the DFT is an isomorphism between sequence spaces, and that FP tends to have a lower l_2 mass than AP, see Figure 7); including both time-domain and frequency domain information allows for a more refined separation to be seen. We then seek to determine a method to classify this low-dimensional encoding as either belonging to FP or AP signals. We accomplish this

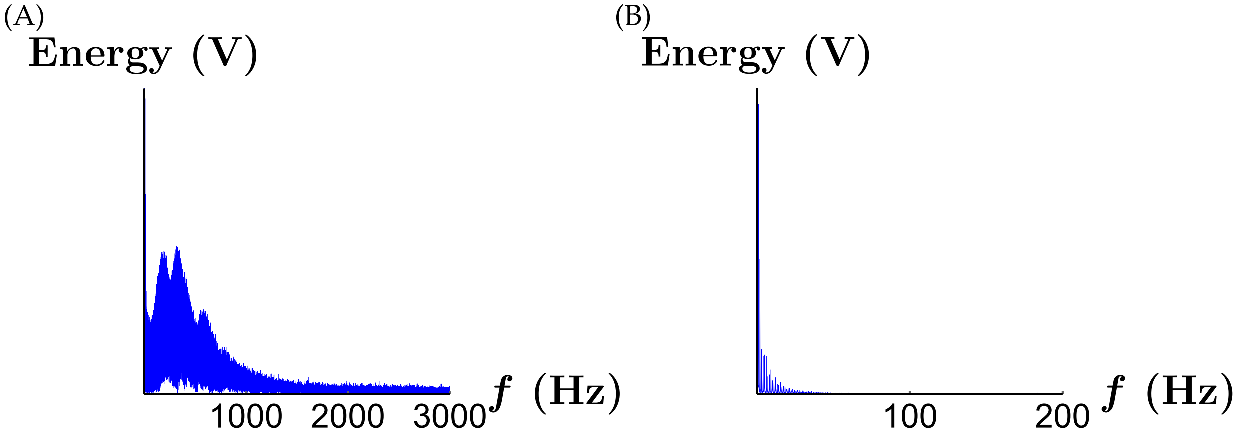

Figure 5: **(A)** FP spectrum. **(B)** AP bandwidth

with a SVM. We formulate a training set by encoding 494 signals: 255 FP and 235 AP. 60 AP training data were created *in silico* from the model due to Grandi et al. (2010) with randomized pacing protocols, while the remaining 175 were extracted by hand from collected data. The training set is presented in Figure 6(A), and this shows the separation that this two-dimensional encoding creates. The SVM trained on the training set computed the hyperplane of maximal margin demonstrated in Figure 6(A). Note that the margin is actually a conservative choice as it is determined by two isolated data points at the upper end of the FP cloud. As the training data is linearly separable, the SVM is able to achieve 100% accuracy in classification on the traning set (Fine, 1999).

The SVM approach also allows a more sequential view of the decay of AP signals to FP-type signals, as the feature point for a signal can be seen to move from one side of the hyperplane to the other as time increases and the nanopore closes, as in Figure 6.

1.3.2 Detection of Acceptable FP and AP signals

The next task, after extracting raw FP and AP data, is to determine if the extracted data is acceptable. For FP, the decision of acceptability, in our paradigm and experimental protocol, is somewhat trivial: acceptable FPs are those with a signal range of greater than 15 mV, simply to show that the cell's electrical activity is detectable and of sufficient level for experimental significance.

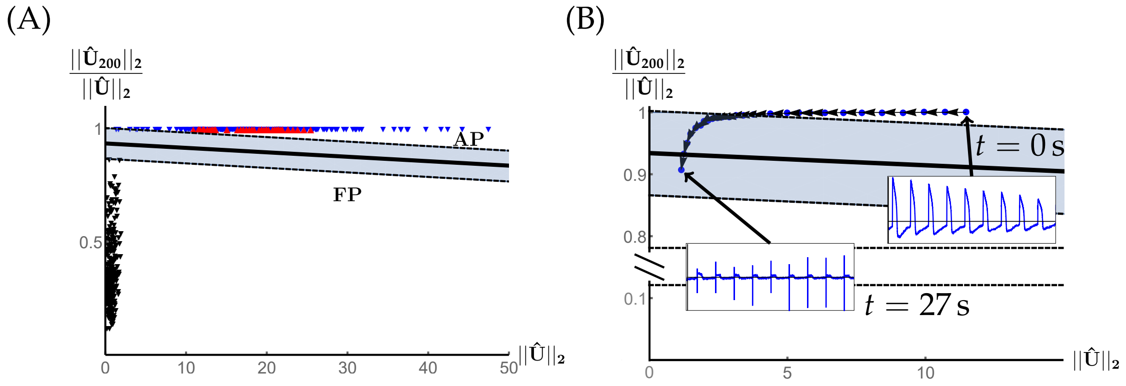

Figure 6: **(A)** The two frequency features from Equation (1). This choice allows linear classification of a signal morphology as a sequence of FPs or APs. *In vitro* AP recordings are marked by blue triangles, and *in silico* simulations are marked by red triangles. The bold line indicates the line of maximal margin obtained via SVM training. The region between the maximal margins on either side is shaded in gray. A signal with Fourier features in this region is discarded as neither AP nor FP. **(B)** Feature-space encoding of signal train in Figure 8(B). Each successive state represents an increment of $\Delta t = 1$ s in the signal. The two-dimensional encoding allows for a refined view of the signal degradation to FP-like behavior.

AP acceptability is more involved. Acceptability of AP readings is essentially in correspondence to the successful completion of electroporation. Therefore, an AP reading is considered acceptable when the electroporation was sufficient to yield a transmembrane potential reading of recognizably AP morphology, and it becomes unacceptable when the nanopore closes enough to introduce artifacts into the trace that distort this morphology. In particular, one looks at the depolarization phase. For a true AP reading, the depolarization phase consists of a steep monotonic increase in potential; as the nanopore closes, certain non-monotonicities become evident in the depolarization (see Figure 4). Once these non-monotonicities manifest, the signal is no longer considered acceptable, as it no longer represents a true AP. Thus, AP acceptability is determined by inspecting the time derivative of the AP.

To isolate non-monotonicity that is due specifically to nanopore closure and not any electrical artifacts from data collection, it is necessary to remove electrical noise and baseline trends in the APs prior to determining the data quality, see Figure 4 (B)-(E). This is done by applying filters to the data. Specifically, we use two Butterworth filters (Butterworth, 1930) applied sequentially: a low-pass filter of order 9 with cutoff frequency 225 Hz for noise removal, and a high-pass filter of order 2 with cutoff frequency of 0.5 Hz for removal of baseline trend.The orders for these filters was determined to maintain stability and filter the maximal amount of noise while not sacrificing computational performance. The filtered data are numerically differentiated using centered differences. We then locate positive maximums of the derivative, which are due to depolarization, and look for possible significant negative minimums in a neighborhood of a local maximum (a certain degree of negativity is expected in the neighborhood of the peak as repolarization takes hold, but our method seeks a significant negativity). In order to extract a signal of sufficient length to apply meaningful analysis, if any of the first 4 APs of the signal is found unacceptable the entire stream is discarded. For signals with more than 4 acceptable beats, the signal is then truncated at the earliest instance of non-monotonicity rather than after a certain period of time.

1.3.3 Feature Extraction

Once the signal has been processed, acceptable data streams are submitted for feature extraction. We compute three classical features of APs, namely basic cycle length (BCL), APD_{30} and APD_{80}. The BCL is computed using the same program that locates depolarization spikes from the segmentation step: depolarization maxima in the numerical derivative are located, and the time between successive deplarizations are then computed and counted as the BCL. Consequently, a signal containing m acceptable AP signals only yields $m-1$ BCL features; as such we only compute the remaining features for these $m-1$ beats to analyze the relationship between features. The APD_n is computed by isolating each individual pulse as above. Each pulse is shifted and re-scaled so that the maximum voltage occurs at time zero with magnitude 1, while the minimal value of the voltage for that pulse is shifted to a value of 0. The time it takes to decay by $n\%$ is then computed by finding the timestamp of the shifted signal sample X_i at which $|X_i - (1-n)|$ is closest to

0. We also compute the features in the Fourier space for these signals. Frequency spectra are extracted by taking a 15 s segment of each signal starting at $t = 0$ s for that signal and computing the DFT for that 15 s window. The spectrum of the DFT is then plotted against the frequency. The 15 s window moves in steps of 0.5 s.

1.3.4 Algorithm implementation

The algorithm is written in the Open Source language `Python` (Python Software Foundation, 2020), carefully documented and available at `https://github.com/jurkiew4/Action-Potential-Segmentation`. Semi-regular maintenance of this code is performed by the author.

1.4 Results

Initial validation tests of the algorithm were carried out by running the program on one full MEA plate of 288 electrodes. These results were consistent with by-hand segmentation and feature extraction but at a minor fraction of the time commitment, even when performed on a commercially available personal laptop running an Intel i7 processor with 8 GB RAM.

The first industrial application of the algorithm was performed by running the segmentation program on 864 candidate data streams from untreated cells (3 full MEA plates of 288 electrodes each). This yields 408 coupled pairs of acceptable processed AP and FP traces. The number of extracted data streams allows deep analysis of the underlying experimental protocol and its results. The duration of acceptable AP traces, and thus the duration of electroporation, were found to vary in length from 5 to 50 s, while on average the nanopores remain open for about 30 s. Representative coupled 15 s traces are shown in Figure 7.

The amount of data generated allows for an aggregate picture of the feature behavior for the entire bulk of cardiac tissue. Figure 9 shows the values of APD_{30} plotted against

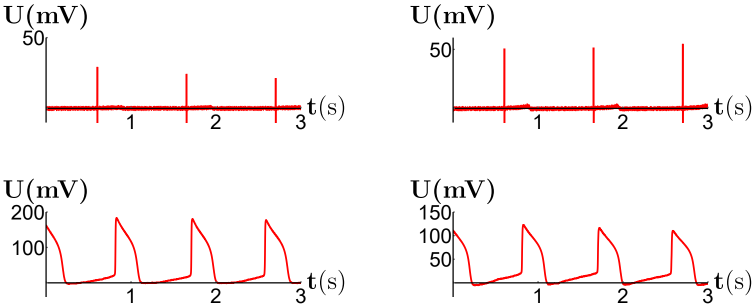

Figure 7: Coupled FP (top) and AP (bottom) readings from the same electrodes. Note the slightly higher beat rate of AP vs. FP, a result of the 1 Hz electroporation stimulus.

APD_{80} for each AP trace extracted via the segmenter program. We see a noticeable linear correlation between the APDs in seconds; fitting a regression line to this plot gives the relation $\text{APD}_{80} = 0.14044 + 0.68770 \cdot \text{APD}_{30}$, with $R^2 = .746346$ within the region of physiologically meaningful APDs.

In addition to the aggregate view, analyzing the dynamic (beat-to-beat) behavior of the AP signals manifest a few noticeable trends. In the Fourier space, we see the signal degradation to an FP-like trace, which sets in after between 4 and 50 beats. In the longer time series of the temporal features such as Figure 8(A), we observe increasing BCLs and increasing APDs - a greater time between beats and a correspondingly increased duration of beat. The BCL data tends to increase towards a value greater than 1 s; as the electroporation stimulus that opened the nanopores was applied at a frequency of 1 Hz Edwards et al. (2018), this indicates that the intrinsic beat rate of these cells is less than 1 Hz. Remarkably, even large fluctuations in BCL frequently do not manifest similar fluctuations in APD. This implies a largely uniform AP morphology in spontaneously beating cells.

In Figure 10(A) we show AP traces from the same electrode recorded 25 h apart. To analyze similarity and reproducibility of measurements across days, we run the segmentation algorithm on data sets from the same MEA plate at 3 different times, $t = 0$, 48 and 96 h, respectively. Figure 10(B)&(C) shows a strong coincidence of the APD_{80} and BCL

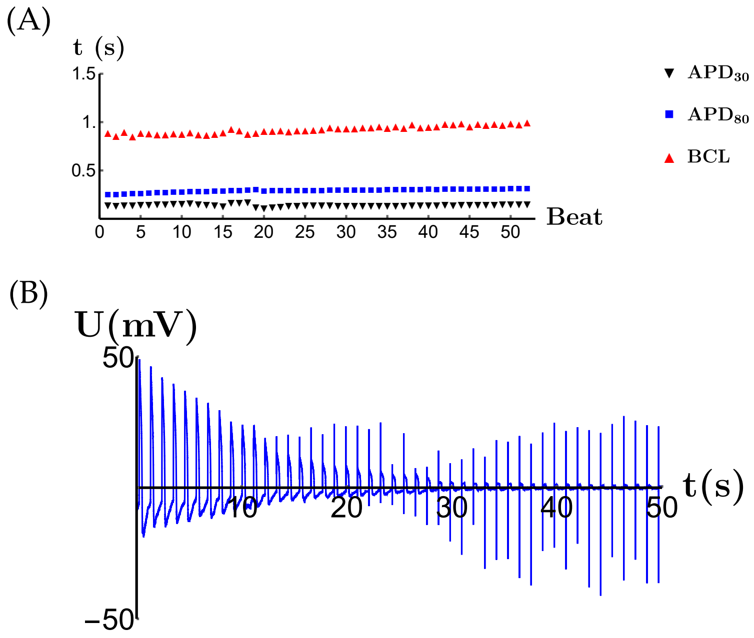

Figure 8: **(A)** Time series plot of APD_{30}, APD_{80} and BCL for a single data stream. **(B)** An AP signal chain of 50 s recorded after electroporation on an MEA. The unstable nanopores begin closing after approximately 13 s, resulting in a transitional signal morphology. After 30 s the nanopores are almost fully closed and the signal has reverted to an FP-like morphology.

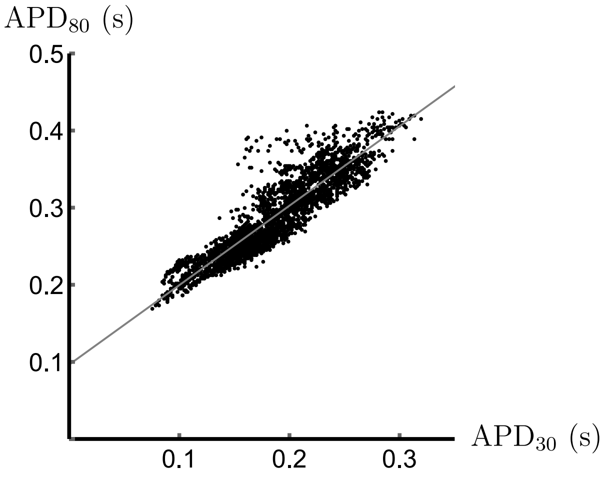

Figure 9: Scatter plot of APD_{30} vs. APD_{80} showing linear correlation.

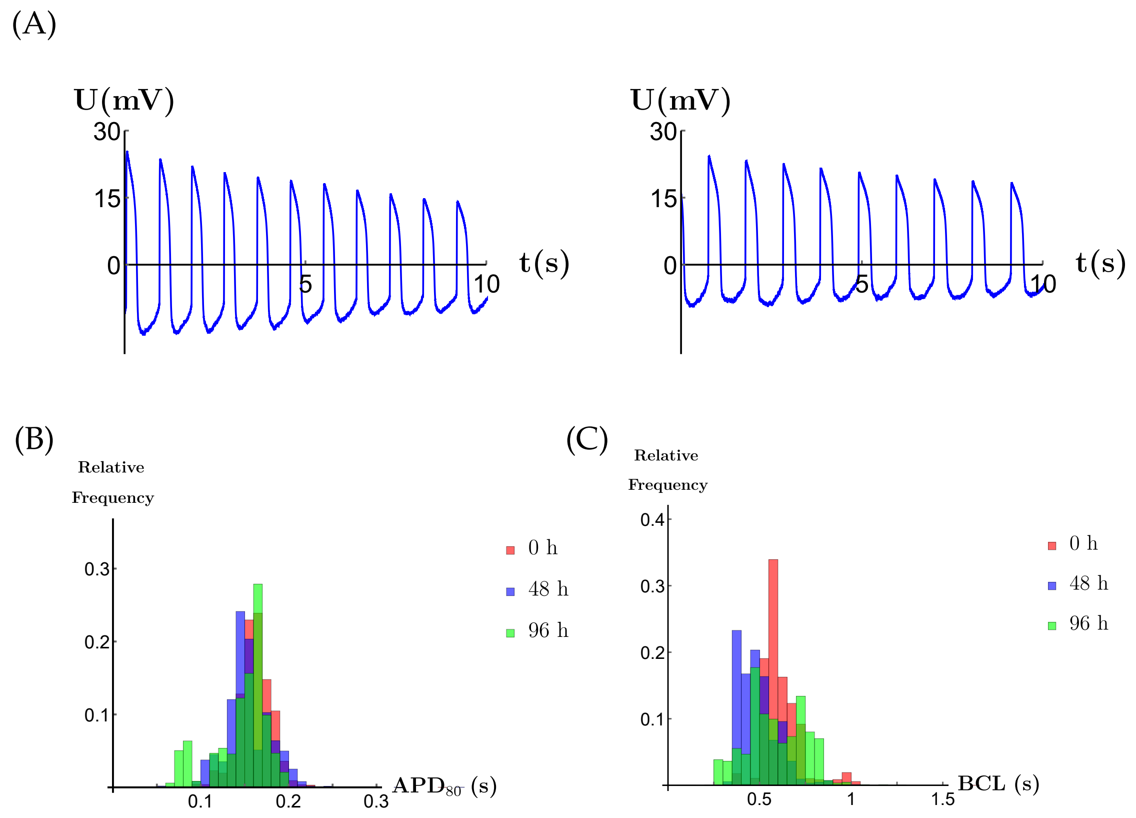

Figure 10: **(A)** AP traces from the same cell site taken 25 h apart, both recorded 10 s post-electroporation. The similarity in morphology and magnitude indicates consistency of measurements across multiple experiments. **(B) & (C)** Distributions of APD_{80} and BCL collected from the same MEA plates at 0, 48 and 96 h, respectively.

distributions. Treating each sequence of APDs from each day as a random sample from an unknown probability distribution, we form the empirical density function for each distribution. Measuring the pairwise distances between the empirical distributions on each day in the supremum norm we observe a maximal distance of 0.3, similar to the maximal supremum distance between distributions of different cell batches. Zhu et al. (2016) have shown that populations of stem cell-derived myocytes can exhibit significant variability in mean APD, and our results conform to these observations. Moreover, any day-to-day variance in APD readings is comparable to the natural variability of readings at a single time point. Therefore, the current experimental protocol allows for repeated comparable measurements on the same cell site.

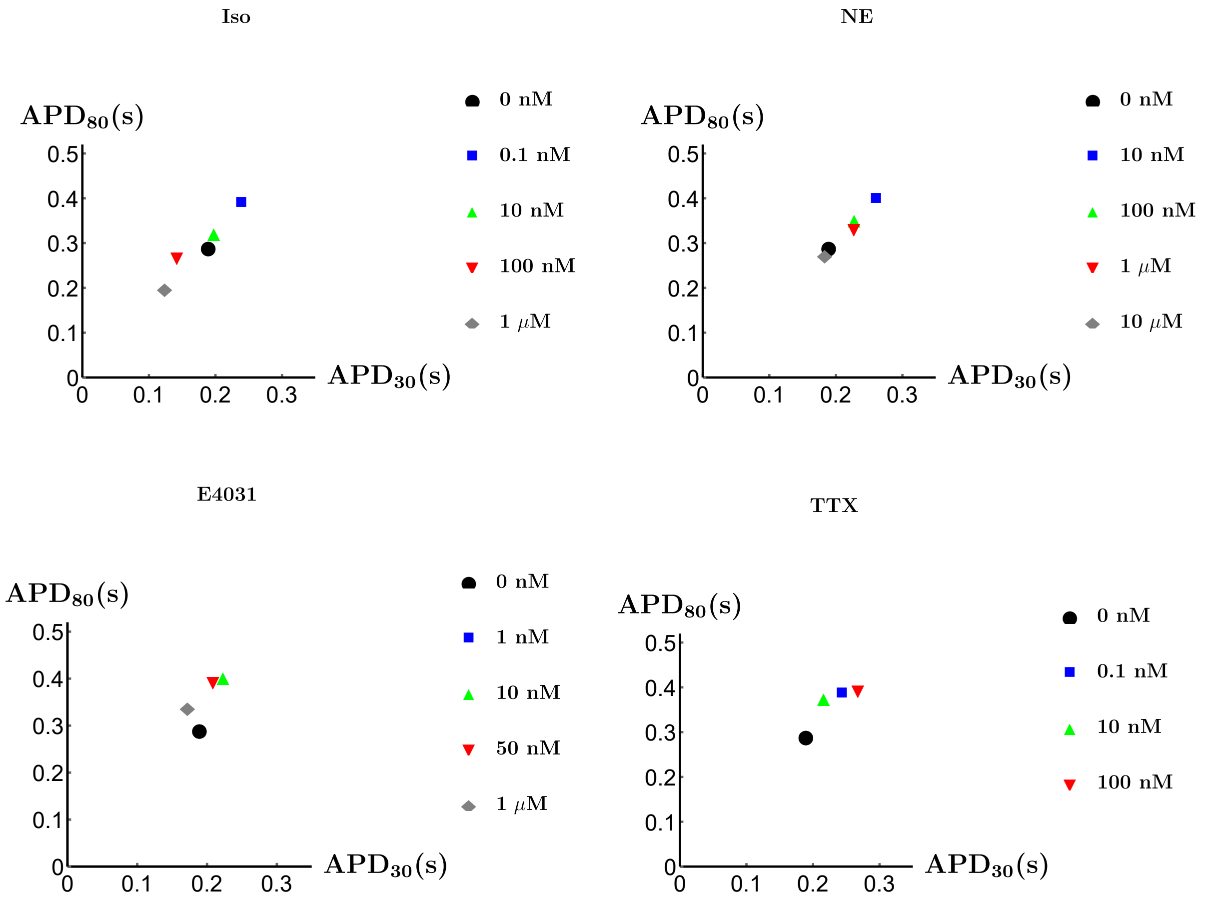

Figure 11: Plots of mean APD$_{30}$ vs. mean APD$_{80}$ in the presence of differing concentrations of (clockwise from top left): isoproteronol, norepinephrine, tetrodotoxin and E4031.

1.4.1 Application of Pharmaceutical Substances

hIPSC methods are used not only to generate cardiac tissue, but to utilize that cardiac tissue in the investigation of the cardiotoxic effects of pharmaceutical compounds (Edwards et al., 2018). Thus any segmentation algorithm must be robust to the effects of these compounds on the signal morphology and still extract meaningful information. Here we examine the performance of the segmentation algorithm on cardiomyocyte populations treated with four different compounds and analyze the extracted data.

Four compounds were applied to the cardiomyocytes in differeing concentrations: isoproteronol, norepinephrine (a neurotransmitter), E4031 (a HERG-channel blocker) and tetrodotoxin (or TTX - a sodium-blocking neurotoxin) (Edwards et al., 2018). A total of four differing concentrations (amounts differed by compound as well, and are displayed in Figure 11) were applied to the cell population. Applying the segmentation algorithm to the data sets collected from these populations yielded the data in Figure 11, which displays the orbit of the average APD_{30} vs. average APD_{80} in the presence of these concentrations. We observe a dose-dependent action of the compounds in accordance with the observations of Edwards et al. (2018).

The behavior of the isoproteronol and norepinephrine populations exhibits some interesting trends in that, while lower doses tend to increase both APDs from the control population averages, higher doses decrease both linearly. E4031 and TTX both exhibit the same dose-dependent action, but noticeably different than the first compounds: neither action is constrained to the line of best fit from figure 9. It appears that isoproteronol lengthens APD_{30} relative to APD_{80}, thus altering the morphology of the AP trace. TTX exhibits a similar greater lengthening of APD_{30}, but in a more chaotic manner relative to dose; further, no electrical activity was recorded from the cells with the highest TTX concentration applied, indicating complete cessation of cellular activity.

We thus see that our segmentation algorithm can be applied equally well in the case of chemically-altered cellular data, and aid in the analysis of the effects of these compounds on cellular electrical behavior. Coupled with the genetic identicality to the donor tissue exhibited by hIPSC-CMs, this segmenter has potential applications in the field of personalized medicine.

1.5 Discussion and Conclusion

MEA technology for industrial-scale recording of coupled FP and AP data allows for sophisticated and extensive analysis of electrophysiological phenomena. Combined with the introduction of induced pluripotent stem cells by Yamanaka in 2006 (Takahashi and Yamanaka, 2006), this creates the potential for vast amounts of data to be collected in a

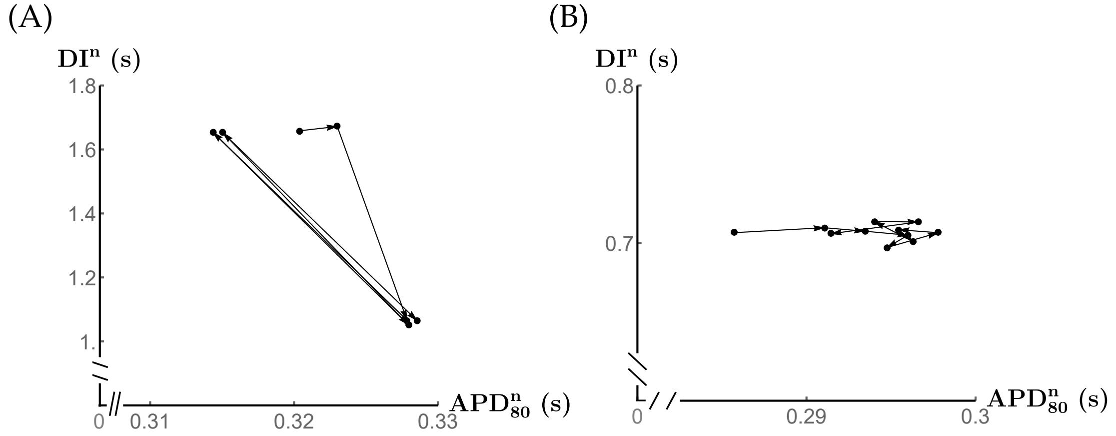

Figure 12: Dynamic plot of (APD_{80}, DI) from two electrode streams for sequences of seven and 11 beats, respectively. Panel (A) shows alternans while (B) shows strong variability in the APD tied to a constant DI. Note that both point sets are largely confined to one-dimensional subspaces.

short period of time.

To parse the significant amount of data generated by even a single MEA experiment and determine which readings are useful sound for analysis, we examine the signal in a 2-dimensional frequency-domain feature space and inspect the time derivative of the signal. Our algorithm efficiently locates and classifies AP and FP signals and determines the acceptability of the signal using only consumer-grade hardware. This allows for quick extraction of relevant data from an experiment without human intervention, freeing up time for experimental analysis. Our algorithm also efficiently addresses the random duration of intracellular access and the slow reversion of AP signals to FP morphology. This enables maximal information extraction for a given experiment. By analyzing the derivative to locate the onset of degradation, we eliminate dependence of the algorithm on a specific experimental procedure and make it applicable to non-electroporation based MEA methods, such as those proposed by Hayes et al. (2019). Further, this algorithm handles the introduction of pharmaceutical compounds and extracts meaningful data regarding the cardiac effects of these compounds as efficiently as it handles untreated cellular data.

MEA technology allows for repeated large-scale electrophysiological readings on the same cells at different timestamps, enabling detailed analysis of the effect of time on various physiological readings. We observe that the variability in readings taken from the same cell site at different days is comparable to the variation in different cell populations and thus is not an impediment to the reproducibility of experiments over timescales on the order of days. This long-scale consistency represents a significant strength of the MEA approach to electrophysiology and its applicability to clinical research.

The large amount of data that can be extracted from this experimental setup also allows for detailed analysis of the behavior between features for a given data stream, and the efficient extraction of these features makes feasible many exciting avenues for future research. One of the most fruitful of these relationships so far investigated for myocytes in the literature is the concept of cardiac restitution, the relationship between APD for a given beat to the diastolic interval (DI) of the preceding beat (where $\mathrm{DI} = \mathrm{BCL} - \mathrm{APD}$). Earlier work on this topic typically employ relationships of the form $\mathrm{APD}^{n+1} = F(\mathrm{DI}^n)$ which can be rewritten as $\mathrm{APD}^{n+1} = \mathrm{BCL} - F(\mathrm{APD}^n) := G(\mathrm{APD}^n)$ (Guevara et al., 1984; Kulkarni and Tolkacheva, 2015; Kesmia et al., 2019). These approaches are heavily dependent on the pacing protocols used to generate data, and require a steady-state response with fixed BCL to be achieved prior to analysis. As such, these models have limited applicability in the case of spontaneously-beating tissue such as is observed in cultures of hiPSC-CMs. To investigate the behavior of cardiac restitution in spontaneous-beating cells with variable BCL, we propose turning to the language of discrete dynamical systems. Consider

$$(\mathrm{APD}^{n+1}, \mathrm{DI}^{n+1}) = \Phi(\mathrm{APD}^n, \mathrm{DI}^n) \tag{2}$$

and observe these pairs from beat to beat; representative orbits from our data are shown in Figure 12. Figures of this type suggest the presence of an underlying dynamical system for the two-dimensional state space. Note the presence of some randomness in the dynamics of these pairs: while some cells like in Figure 12(A) exhibit an alternating long and short APD and corresponding DI (a condition known as **alternans**), others can exhibit extreme variability in APD with no corresponding change in DI. Therefore, a random dynamical system of the form

$$(\mathrm{APD}^{n+1}, \mathrm{DI}^{n+1}) = \Phi_\lambda(\mathrm{APD}^n, \mathrm{DI}^n) \tag{3}$$

Where $(\Phi_\lambda)_\lambda \in S)$ is a parametrized family of functions $\Phi_\lambda : [0,1] \to [0,1]$, (S, μ) is a probability space and λ chosen from S. Investigation of the statistics for the choice of λ, done via statistical analysis of the Poincaré plots for a given data set, would enable a stochastic picture of the spontaneous beat behavior of these hIPSC-CMs.

An additional approach to research on this topic may use deep neural networks to fit a function to these time series rather than attempting to explicitly determine the formula of the system. It may also be possible to treat each pair as a complex number $z^n = \mathrm{APD}^n + i\,\mathrm{DI}^n$ and exploit the field structure of $\mathbb{C}$ to apply a conformal neural network (Cantwell et al., 2019) instead of a traditional deep neural network.

A mention of the extensibility of this algorithm is also necessary. The particular data used in the construction and validation was in the HDF5 file format, but any file format for data is theoretically possible. Further, while the above implementation only extracts a small number of numerical features primarily from AP traces for analysis, the data arrays created by this method can be mined for any number of features, allowing user input as to which features to extract for a given application. In particular, FP duration is not computed by this program. This requires accurate identification of the repolarization wave in the FP complex. Automated detection of such waves is quite challenging, with several algorithms developed (Elgendi et al., 2015; Shang et al., 2019), but these come with computational overhead and as such are not included in the base implementation of our segmenter. MEA signals do not always manifest a recognizable repolarization wave, necessary to compute this quantity. This may be tied to cell linkage to the substrate in the MEA (Li et al., 2017) or perhaps to immaturity of cells.

Finally, the geometric construction of the MEA system allows the collection not only of temporal data, but also of spatial data for the bulk of tissue in the MEA well. Therfore, a full data set could be considered to have a graph structure on which a temporal dynamical system acts. With a finer spatial resolution (some MEA plate models can have up to 250 electrodes per plate) this could be used to generate a continuum model for hIPSC-CM tissue bulks as well.

2 A Population Dynamics Approach to the Distribution of Space Debris in Low-Earth Orbit

Keywords: Population dynamics, space debris, Kessler Syndrome, diffusion, Low-Earth Orbit

2.1 Introduction

2.1.1 Kepler's Laws for Orbital Bodies

Consider an object of mass m in orbit around a central spheroidal body of mass M. In the absence of other forces, the orbital body is bound by Newton's Law of Gravitation

$$\mathbf{F}(r) = \frac{GmM}{r^2}\hat{\mathbf{r}}$$

which governs its trajectory around the central body. This system possesses 6 degrees of freedom: the three spatial Cartesian position variables (r_x, r_y, r_z) and the three Cartesian velocity variables (v_x, v_y, v_Z).

In practice, the orbit of a body around a spherical central body is defined by 6 **Keplerian** or **orbital elements** that encode the position and velocity of the object more readily. The classical Keplerian elements are as follows.

Every ideal orbit occurs in a plane. From the central body, select a reference plane (typically the body's equatorial plane) and a reference direction (for earth, this is the prime meridian). The **inclination** i is the angle of the orbital plane relative to the reference plane, while the **right angle of the ascending node** Ω is the angle of one of the points of intersection between the orbital and reference planes, relative to the reference direction - specifically, it is the point at which the orbiting object crosses in an arbitrary but fixed "ascending" direction across the reference plane. The **argument of periapsis** ω is the angle from the reference angle to the point on the orbital ellipse closes to the central body. The **eccentricity** e is the same as the elliptic eccentricity of the orbit. The **mean motion** n of the object is the constant angular speed required to make one full rotation in the object's period; according to Kepler's 3rd Law, this is related to the **semi-major axis** a of the orbital ellipse by $n = \sqrt{\frac{\mu}{a^3}}$ where $\mu = GM$ is the product of the gravitational constant and the central mass. Finally, the **mean anomaly** M is the proportion of the full orbit from periapsis that has been complete at time of measurement, reported as an angle in a fictitious circular orbit. That is, if the object is three-quarters of the way through its orbital period in its elliptical orbit, the mean anomaly will be $M = \frac{3}{4} \times 2\pi = \frac{3\pi}{2}$ - i.e. the circular orbit a fictitious object travelling at angular speed n would complete in the same period as the object's elliptic orbit.

A seventh element is implicitly used in the translation of these elements into Cartesian coordinates: the **eccentric anomaly** E, which is defined as follows: suppose there exists a circle concentric with the orbital ellipse of radius a in the orbital plane, with x-axis passing through the central body and periapsis. For an object at any point P on the orbital ellipse, let P' be the point on this circle with the same x-coordinate in the orbital plain. E is then the central angle of the point P' on the circle of radius a. Kepler's Equation $M = E - e\sin(E)$ relates the mean anomaly to the eccentricity and the eccentric anomaly, allowing the 6-tuple $(i, \Omega, e, \omega, n, M)$ to encode the position and velocity of a body in Cartesian coordinates (Returns, 2022).

2.1.2 Simplified General Perturbations (SGP) Model

In practice, the orbits of objects in orbit around real central bodies do not conform identically to Kepler's Laws. Objects orbiting Earth, for example, are subject to perturbations in their orbits due to the oblateness of the Earth. Additionally, solar wind can influence the behavior of the orbits. It is thus necessary to incorporate these perturbations into any attempt to solve for the orbit of a real object with the precision necessary for modern applications.

The work of Brouwer (1959) and Lyddane (1963) laid the groundwork for what became the Simplified General Perturbations (SGP) Model, currently on SGP4 since 1981 (Hoots, 1981). The full model is incredibly complex and intensive, so we make no attempt to reproduce it in full detail here. Instead, we attempt a sketch of the work and refer the reader to the literature for more depth.

This work assumes the spherical potential present due to a spherical earth is perturbed by certain zonal harmonics representing the effects of gravity, atmosphere and solar wind. Further, the work uses what are known as Delaunay variables, which are related to the Keplerian elements:

$L = (\mu a)^{1/2}$	M
$G = L(1-e)^{1/2}$	ω
$H = G\cos(i)$	Ω

Table 1:

From the canonical Delaunay variables, the author performs two invertible conjugations of variables to arrive at the 6-tuple $(L'', G'', H'', M'', \omega'', Omega'')$, in terms of which the conjugated equations of motion can be solved exactly up to second order in a small parameter k_2. These six conjugate variables are taken to describe the actual motion of the object in orbit, while the original six Keplerian elements are called **osculating elements** in that they describe the Keplerian orbit an object with a certain instantaneous state would follow.

The work of Brouwer (1959) makes certain simplifying assumptions which render the model unstable in the vicinity of a critical inclination, around 63.43°. Lyddane (1963) refines the model by focusing on the **Poincaré variables**:

$x_1 = L$	$y_1 = i + \omega + \Omega$
$x_2 = (2(L - G))^{1/2} \cos(\omega + \Omega)$	$y_2 = -(2(L - G))^{1/2} \sin(\omega + \Omega)$
$x_3 = (2(G - H))^{1/2} \cos(\Omega)$	$y_3 = -(2(G - H))^{1/2} \sin(\Omega)$

Table 2:

Focusing on these transformed variables provides a more stable approach that still reduces to Brouwer's model for all inclinations and eccentricities.

2.1.3 Near-Earth Space

Near-Earth space is the terminology used for the portion of outer space directly adjacent to the planet Earth. Historically, there is a canonical separation of this region into three altitude regions for more precise definition: the Low Earth Orbit (LEO, with an altitude $\leq$ 2,000 km), the Medium Earth Orbit (MEO, altitude between 2,000 - 36,000 km), and the Geosynchronous Earth Orbit (GEO, altitude $\geq$ 36,000 km). In each of these three broad regions, the dominant forces and characteristics of existing objects differ substantially, and therefore each region requires different scientific treatment (Rossi, 2005). Additionally, the orbital periods accessible to a satellite in each region are substantially different, with larger periods accessible only at greater altitude ranges. This limits the type of satellite that can be placed into each orbit.

Of these three regions, the GEO region is the most resistant to analysis due to technological limitations - current space tracking technology is limited to observation of objects above a size of about 30 cm at that altitude. Further, debris in the GEO region is not as impacted by atmospheric drag or the gravitational effects of the Earth's oblateness, and therefore objects in this region do not exhibit decaying orbital patterns. Instead the gravitational effects of sun and moon as well as solar radiation are the dominant influences on the behavior of these objects. Objects in this region can access orbital periods equal to 24 hours, allowing for the satellite to remain essentially fixed above an equatorial point on Earth's surface - this is the desired orbit of communication satellites and other scientific satellites. As such, orbits of objects in this region are by necessity rather rather static and predictable. For this reason, and the low density of cataloged objects in GEO (541 active satellites as of February 2022 (Johnston, 2022)) the GEO region density is generally considered to be stationary in its activity over any realistic timescale.

The MEO region is of interest due to the presence of the constellation of GPS satellites in a 12-hour orbit, accessible only in this region.

The LEO space is the most interesting to analyze due to its proximity to earth and the sheer number of objects in orbit within it. Since the beginning of human space exploration in 1957, the LEO region has been inhabited by artificial objects. While the population of near-Earth space in the period immediately following this time was comprised of active man-made bodies, the abandonment of satellites, inter-object collisions or unitary explosions have led to an ever-increasing presence of so-called space debris - inactive, non-controllable entities evolving according to Keplerian laws. This filling population risks runaway collision of existing debris objects and poses an increasing danger not only to safe operation of active spacecraft, but to any object existing in near-Earth space. Indeed, in 2022 a discarded rocket stage collided with the Moon causing a noticeable crater (Space.com, 2022a). Investigation of the activity and evolution of this debris environment is therefore an important direction of research.

2.1.4 The Kessler Syndrome

Kessler and Cour-Palais (1978) realized the inherent risks involved with the space debris population and proposed the **Kessler Syndrome**: the uncontrolled growth of the debris population if left to evolve naturally.

There is a tremendous risk involved with the runaway growth of space debris and its impact on future space missions. For example, take NASA's collision model for the number of new objects with mass greater than m, denoted $N(m)$ between two objects of masses M_1 and M_2 (Pardini et al., 1998):

$$N(m) = 0.4396(\frac{M_e}{m})^{.7496}$$

where M_e is the total mass of ejected material (dependent on the kinetic energy of the impacting mass - if kinetic energy is sufficiently high to destroy both objects totally, this is taken to be $M_1 + M_2$). A catastrophic collision between a 10kg and a 5kg mass would then create ≈ 3 objects greater than 1kg in mass, and almost 600 greater than 1g. As the likelihood of a collision between two particles is quadratic in the numerical density of the population, each collision dramatically increases the risk of future collisions.

Even singular objects can pose a danger to the future space environment. If the object is capable of exploding, NASA's empirical model for the explosions considered to be of "low intensity" (Pardini et al., 1998) is a piecewise exponential law:

$$N(m) = \begin{cases} 0.171Me^{-0.6502\sqrt{m}} & m \geq 1.936\text{kg} \\ 0.869Me^{-1.8202\sqrt{m}} & m < 1.936\text{kg} \end{cases}$$

where M is the exploding mass. While these occur rarely, their impact on the debris population is enormous.

High-velocity collisions ($\approx$ 10 km s^{-1}) such as those predicted by the Kessler syndrome have occurred in reality, and occasionally resulted in the destruction of active satellites, for example the Iridium 33-Kosmos 2251 collision in 2009. Unitary satellite breakup or destruction has also occurred. Examples are the aftermath of the intentional destruction of the Fengyun 1C weather satellite by a Chinese anti-satellite missile in 2007, or the destruction of Kosmos-1408 in a Russian anti-satellite weapon test in November 2021 (Carnegie Endowment for International Peace, 2022).

2.1.5 To a Population-Dynamics Approach

Active monitoring and forecasting of the population of space debris is an important point of action for space-faring nations. At the level of tracking individual objects, the gold standard in analysis of artificial satellites has been the Simplified General Perturbations model (SGP, currently on SGP4) (Hoots, 1981). This model, described in Section 2.1.2, provides an effective means of orbit determination for individual orbits over short timescales, and hence ideal for short-term policy and individual tracking. Nevertheless, this method suffers from certain shortcomings. This model requires input data to be provided in the format of NORAD two-line element (TLE) sets, which is a very particular and obfuscating data format to encode and decode orbital state vectors. Secondly, this model is a single-particle model, which requires independent application to each cataloged object in orbit. Therefore, propagating the entire population of orbital debris potentially requires tens of thousands of separate applications. Thirdly, over long timescales the predicted orbit can vary significantly from reality (Riesing and Kahoy, 2015). Finally, the model itself does not directly include the presence of inter-object collisions or single-object explosions, which, though rare, cannot be neglected.

There have been many attempts to incorporate collisions and explosions into applications of the SGP4 model, or similar. Nikolaev et al. (2012), for example, utilizes a modeling approach, computing the trajectory of every known object of size $\geq$10 cm. In the event of a detected collision, well-known empirical distributions for number and size of particles created in collisions are used to update the catalog with new objects, whose trajectories are subsequently propagated along with the original objects as an enlarged cohort, to account for population growth. NASA's benchmark is the LEGEND model (Liou et al., 2004). The LEGEND approach is similar to that of Nikolaev et al. (2012), but more all-encompassing - the entire space up to GEO is simulated in this model. Additionally, proprietary methods for orbital propagation are used in place of the standard SGP4. This has been the main tool used by NASA to study the near-earth environment since completion in 2004. While very accurate, these methods still require propagation of thousands of individual orbits and sophisticated collision detection. Hence, an analytical population-dynamics model can possibly incorporate the key features of orbital evolution, collisions and explosions combined with significant computational savings.

Analytical models at the population level were first introduced by McInnes (1993), see also (Zhang et al., 2019) for a more recent contribution. Although space debris is by definition non-biological material, the mathematical modeling approach to its behavior shares a surprising similarity to that of spatially structured biological populations that are subject to birth and death processes, as well as inter-species interactions. The model of McInnes (1993) includes a convection term and a growth term due to binary collisions. In addition, space debris is subject to the effect of solar wind which is included as a diffusion term. In this work, we propose a spherically symmetric diffusion model with interaction term for space debris population evolution. We examine public databases of existing TLEs to parametrize and initialize the model for the subsequent numerical simulations. The diffusion parameter varies with altitude, thereby incorporating the effects of both atmospheric drag, gravitational perturbations and solar wind. The use of a novel analytical model allows for significant computational savings and avoidance of numerical instability while capturing the important mechanisms of population growth and providing good approximation to debris activity. Moreover, we allow for explicit control at the level of launch policy and potential cleaning mechanisms. This provides the framework for testing of different strategies to mitigate the growth of the population density. Our work therefore has the potential to inform future space law and policy (Gast, 2022).

2.2 Introducing a spatial population model for LEO

We begin by stating the underlying assumptions on the structure of orbital debris. It is a natural assumption supported by empirical observations, as in Figures 13-15, that the orbital population is spherically symmetric. Thus let $u(r, t)$ denote the density of orbital objects at time t and a distance of r from the center of the Earth. Assuming the earth to be perfectly spherical, the total population size is obtained by integration,

$$U(t) = 4\pi \int_{r_E+200}^{r_E+2000} u(r,t) r^2 \, \mathrm{d}r, \tag{4}$$

where $r_E = 6378$ km is the equatorial radius of the Earth.

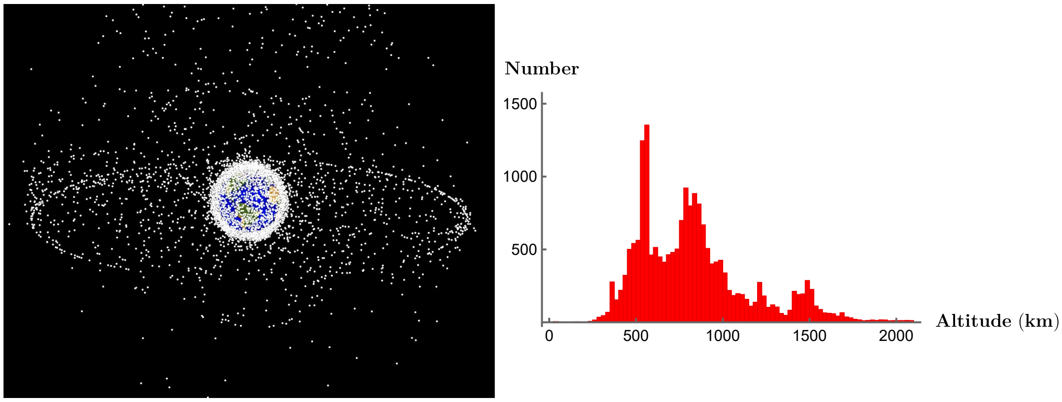

Figure 13: **(Left)** A computer generated image of space debris as seen from high Earth orbit. Clearly visible are the LEO cloud and the GEO ring. From NASA's Orbital Debris Program Office; available at `orbitaldebris.jsc.nasa.gov/photo-gallery`. **(Right)** Histogram snapshot of spatial density of orbital population in LEO as of March 22nd, 2022. Histogram generated by obtaining TLEs of all catalogued objects, and extracting altitudes by use of pyorbital. This is distinct from the density functions used in Equation 5.

LEO space is affected by several factors, namely gravitational effects from the Earth-Moon-Sun system, solar wind, and, in lower regions, atmospheric friction. To account for these factors on the behavior of the LEO population, we postulate that the spatial density $u(r,t)$ evolves in time according to a diffusion equation with source and collision term Φ,

$$\frac{\partial}{\partial t} u(r,t) = \frac{1}{r^2} \frac{\partial}{\partial r} \left(D(r) r^2 \frac{\partial}{\partial r} u(r,t) \right) + \Phi(u,r). \tag{5}$$

The diffusion term in this equation represents the combined effects of atmospheric drag, a multi-body gravitational system and solar wind. In order to capture the altitude dependence of these perturbations, we select a piecewise radially dependent diffusion parame-

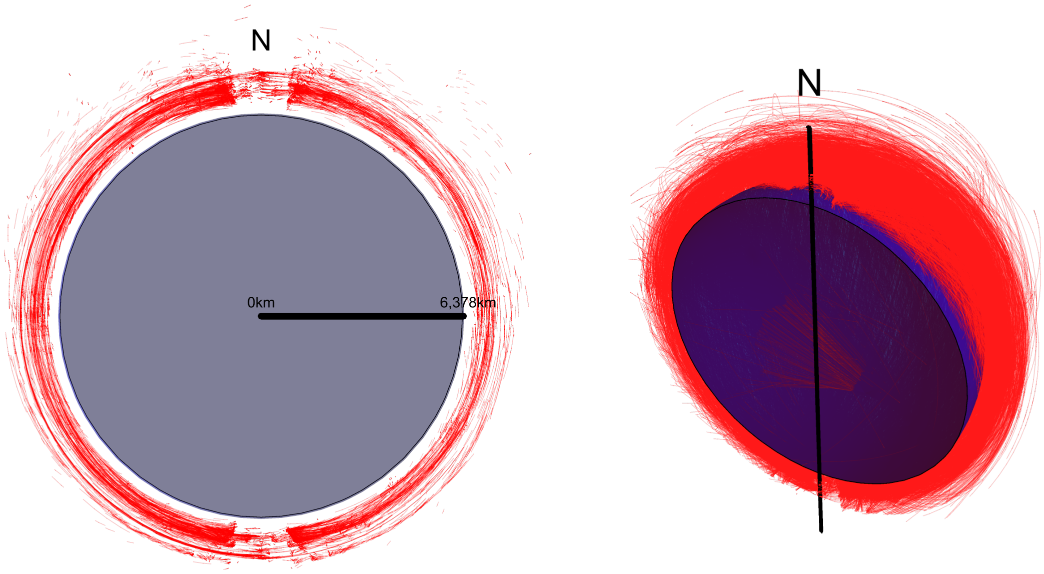

Figure 14: Cross-sectional views of the distribution of satellite orbits in LEO as of March 2nd, 2022. Aside from minor conical areas near the poles, orbital ellipses cover the volume of near-Earth space essentially uniformly.

ter

$$D(r) = \begin{cases} \alpha \exp(-\lambda(r - r_E)), & r < r_E + 1000 \\ \xi, & r \geq r_E + 1000 \end{cases} . \tag{6}$$

Here α and ξ are setting the diffusion rate in the near and the far field, respectively, and λ sets the spatial scale. The choice of these values is discussed in detail in Section 2.3. The functional form is inspired by the presence of atmospheric drag which decreases exponentially with altitude. Above a certain altitude only the solar wind effects the trajectories. In accordance with the observations of (Lemaitre and Hubaux, 2013), we choose a critical altitude of 1000 km as atmospheric drag works on a timescale of 10^3-10^4 years at this altitude.

The term Φ represents a "birth and death process" for the debris population. Debris is created from two sources namely collisions between present objects and deposition of new objects by ground launch. We set

$$\Phi(u, r) = C(u) + \Delta(r, t) \tag{7}$$

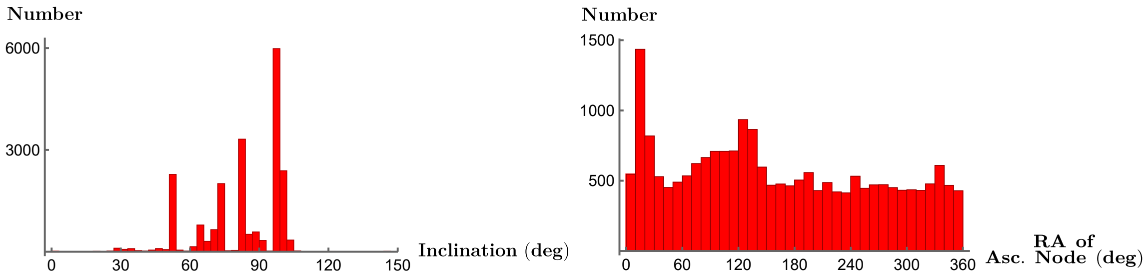

Figure 15: Histograms of **(Left)** orbital inclination and **(Right)** right ascension (longitude) of ascending node (the longitude at which the orbit intersects Earth's orbital plane) as of March 2nd, 2022. These together show that spherical symmetry is a reasonable assumption to model the dynamics of orbital debris.

where $C(u)$ is the rate of object creation due to binary collisions between orbital objects, and $\Delta(r,t)$ is the rate of deposition of objects into orbit. We are at the moment ignoring explosions as a source but leave the possibility that these be incorporated as stochastic impulses in the future. In McInnes (1993), a quadratic collision term of Smoluchowski type is proposed for binary collisions of space debris subject to certain averaging assumptions,

$$C(u) = \frac{\partial u(r,t)}{\partial t}|_{collision} = \frac{1}{\sqrt{2}}\beta\gamma\sqrt{\frac{GM}{r+r_E}}u^2(r,t) \tag{8}$$

where G is the gravitational constant, M is the mass of Earth, β is the average number of new particles created in a binary collision, and γ is the mean cross-sectional area of space debris. We assume that the deposition rate is separable, that is

$$\Delta(r,t) = R(r)T(t). \tag{9}$$

The idea is that there are "natural" locations for deposition of working objects and that merely the amount varies over time. The functional form is allowed to vary when we consider different strategies to mitigate debris growth, but this separable form is the standard for business-as-usual computations.

The model is completed by a homogeneous Dirichlet boundary condition at the lowest altitude as objects that enter the denser layers of the atmosphere burn up (although occasionally large pieces of space debris have reached the surface of the Earth) and exit the population. At the upper altitude we impose a homogeneous Neumann boundary condition since there is no debris migrating from the MEO region. Thus we have

$$u(r_E + 200, t) = 0, \quad \frac{\partial}{\partial r} u(r_E + 2000, t) = 0 \tag{10}$$

for all $t \geq 0$. Finally, we will discuss in Section 2.3 the selection of the initial condition

$$u(r, 0) = u_0(r), \quad r \geq r_E + 200. \tag{11}$$

2.3 Parametrization of the model

To validate the model and aid in setting certain parameter values, we compare our model's predictions to those of NASA's LEGEND model in the case of continued deposition with no mitigation (NASA, 2022). As LEGEND is only presented to the public in a limited manner and the available predictions are only in terms of the total population, these comparisons were done by integrating the density at each time stamp and comparing the evolution of the total population to the predictions of LEGEND over a timeframe of 100 years.

Free parameters and initial conditions are established using real-world orbital data in the TLE format obtained from Space-Track (US Space Command, 2010). As such, the data do not describe Keplerian orbits and are specialized for use with the simplified perturbations models (Hoots, 1981). This necessitates the use of specialized software to decode the TLE data into Keplerian elements that can be used to compute position vectors. We use the open-source Python distribution pyorbital to extract the osculating orbital data (`pypi.org/project/pyorbital`). The altitude at any given epoch is computed from the most recent TLE preceding the chosen epoch for each cataloged object to obtain the cumulative population function. This function is then differentiated and divided by $4\pi r^2$, as r ranges from $r_e + 200$ to $r_e + 2000$, to obtain the initial condition $u_0(r)$ for Equation (11).

To select the values of diffusion parameters α and ξ for Equation (6) we use variability present in the available TLE databases, as this variability is due to the effects of the same perturbing forces. In the same manner as Riesing and Kahoy (2015), we take the two most recent TLEs for every tracked object currently in LEO; the older TLE for each satellite is propagated to the epoch of the newer TLE, and the error in altitude (propagated to observed) is recorded, assuming that the more recent TLE gives the correct osculating value. We gather these errors into groups based on altitude and time difference in epoch and compute the standard deviation of each bin. For each group, we use the formula $\sigma = \sqrt{2Dt}$ to gather diffusion rates. Finally at each altitude we take the average value of these diffusion rates to model the "true" diffusivity. Choosing $\xi = 10^{-4}$ km^2 d^{-1} in keeping with the observations of Lemaitre and Hubaux (2013) and the cataloged data, we fit the above piecewise exponential function to this scatter plot via nonlinear regression as implemented in Wolfram Mathematica, giving $\alpha = 0.5783$ km^2 d^{-1} and thus forcing $\lambda = 0.0086$ km^{-1}. A plot of the diffusion rate vs. the altitude is shown in Figure 16.

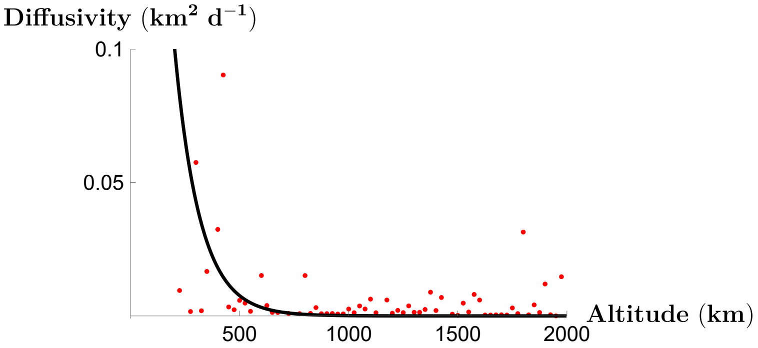

Figure 16: Diffusivity function for Equation 6 for individual orbits at different altitudes.

The parameters β and γ in the collision term (8) represent the number of new objects created in a binary collision and the average cross-sectional area of a piece of space debris, respectively. We set $\beta = 2000$, taking the Iridium-Cosmos collision of 2009 as a representative (ApolloSat.com, 2022). In Kessler and Cour-Palais (1978), a range of poten-

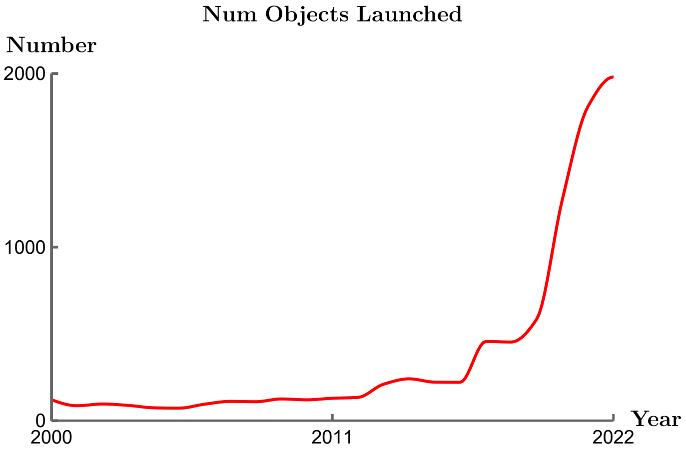

Figure 17: Number of objects launched into LEO per year since 2000. Note the uptick around 2019, when Starlink Satellites were first launched (Space.com, 2022b).

tial values for cross-sectional areas was obtained from radar data and is presented as 2 - 12 m^2. Therefore, we expect this parameter to be approximately in this range. In practice, we set this parameter value to ensure a good fit in the L_1 and L_∞ norms of our model to accepted standard predictions, namely those presented in LEGEND. This agreement does require the presence of a birth term as described in section 7 The parameters used in Equations (5)-(11) are collected in Table 3.

α	0.5783 km^2 d^{-1}
ξ	10^{-4} km^2 d^{-1}
λ	0.0086 km^{-1}
β	2000
γ	1.01×10^{-8} km^2

Table 3: Collected parameters for our model Equations (5)-(11).

The space and time components of the deposition function from Equation (9) are possibly the most difficult objects to estimate for our model. First we begin with

$$R(r) = \sum_{k=1}^{n} a_k \exp\left(-\frac{(r-r_k)^2}{\sigma_k^2}\right),$$

where the r_k are the typical orbit locations and the σ_k are the widths of those regions. This is motivated by inspection of the densities from 2014 through 2022; there appear to be certain "natural" altitudes at which most objects are deposited. Our choice of values was made to give $R(r)$ a total mass of 1, and is given in Table 4.

To set the $T(t)$ factor, we use the catalog of total number of objects launched for the last several years, maintained by the UN Office for Outer Space Affairs (accessible at `https://www.unoosa.org/oosa/osoindex/search-ng.jspx`). We interpolate this list of values via cubic spline interpolation. Additionally, we make the simplifying assumption of uniform deposition over the year - that is, if n objects are deposited during a year, the effective rate is $\frac{n}{365}$ per day. This choice results in an agreement between the propagated and observed densities from 2014 through 2020, see Section 2.4.

k	1	2	3	4
a_k (km^{-3})	$5.599 \cdot 10^{-16}$	$1.39 \cdot 10^{-11}$	$8.39 \cdot 10^{-12}$	$1.39 \cdot 10^{-11}$
r_k (km)	200	500	700	850
σ_k (km)	7.07	20.09	100	9.98

Table 4: Peak locations, widths and strengths of the deposition function R.

2.4 Simulation and comparison with data

Note that, due to the underlying simplifying assumptions inherent in a population-based model, we interpret any instances of blow-up in finite time as indicators of catastrophic, uncontrolled growth rather than true physical singularities.

To validate the model and parametrization, we compare predictions to those made by the LEGEND model. LEGEND was formulated in 2004 and does not include recent notable events such as the Kosmos-1408 explosion or the recent uptick in payload launches, so we propagate the 2020 density to 2120, assuming a launch policy similar to that in 2004, and compare with LEGEND. Figure 18 shows the similarities in forecasted populations; our model deviates from the LEGEND predictions by approximately 10% in the L_∞ norm.

The red curves in Figure 19 show the evolution over the past 8 years of radial space debris density, as obtained from Space-Track.com (US Space Command, 2010). Note the

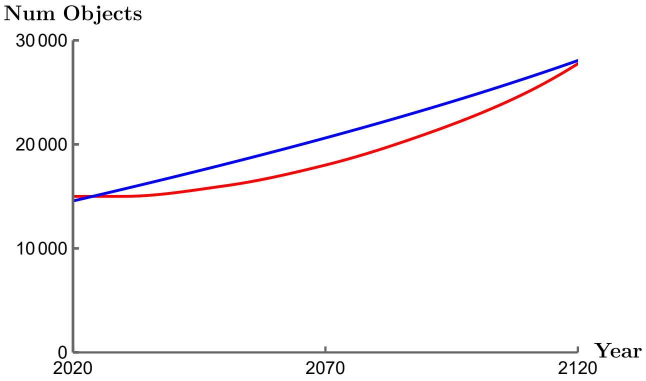

Figure 18: Comparison between forecasted total debris populations given by our model (blue) and LEGEND (red) - LEGEND graph reproduced from NASA (NASA, 2022).

steady, slow growth prior to 2022.

We implement numerical simulations of the model (5)-(11) using a custom driver program implementing the forward-Euler method. Figure 20 shows the evolution of the population (with no further deposition of objects) starting from two different epochs, 2020 and 2022, respectively. The presence of the noticeable peak around 450 km from the 2022 epoch is due to the Russian test of their anti-satellite weapon in November 2021, creating upwards of 1500 traceable pieces of debris in this altitude range (Carnegie Endowment for International Peace, 2022). While such unitary and exceptional events are difficult to incorporate into a deterministic model, it is clear that the presence of this debris causes significant hazard in form of much more pronounced growth of the population.

To analyze the effect of space policy on the debris population, we first assume that ground policy does not account for the removal of space debris. That is, $\Delta(r,t)$ is assumed to be always positive, and has the form specified above in Section 2.3. Comparisons of the propagated and observed populations for 2016, 2018 and 2020 are shown in Figure 19. Figure 21 show the impact of continued deposition at the then-current rate on the growth of the 2022 population. The slow growth of the 2022 population under deposition is replaced with almost immediate catastrophic growth - the density increases almost 10-fold in approximately 40 years.

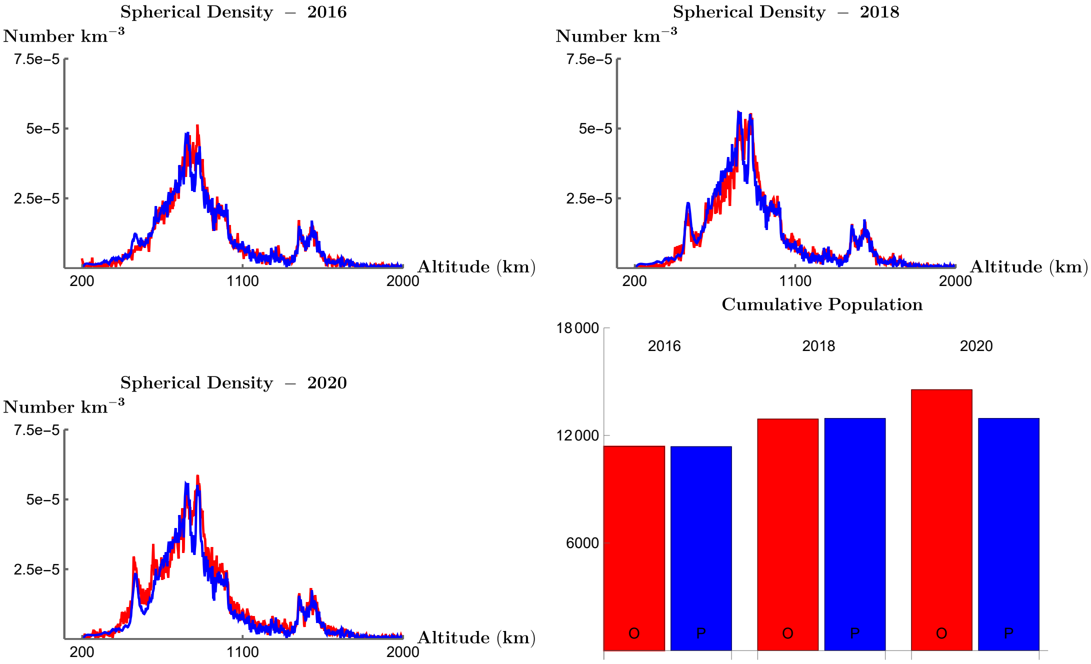

Figure 19: Comparison of observed (red) population and propagation of 2014 (blue) populations under the action of ground launches in 2016, 2018 and 2020 respectively, along with the population growth (lower right panel).

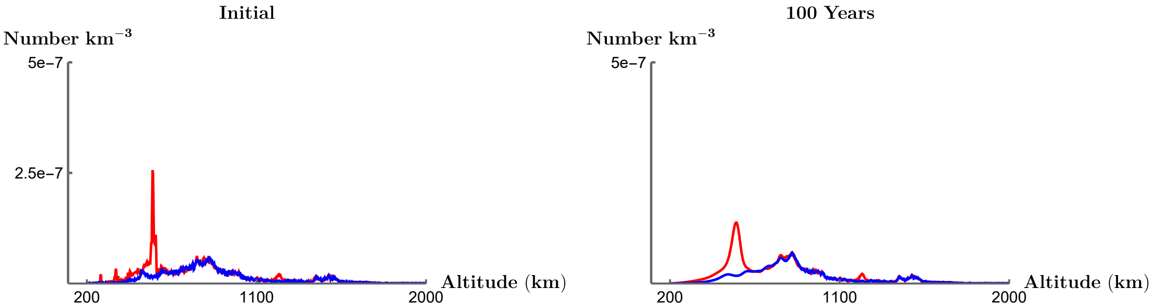

Figure 20: Initial densities (left) in 2020 (blue) and 2022 (red), along with densities propagated 100 years (right) in the event of no further deposition. Note that the Kosmos-1408 debris spike at 500 km altitude in the 2022 population sustains itself through the next 100 years and contributes to overall population growth, while both densities diffuse toward the lower altitude boundary, in keeping with Lemaitre and Hubaux (2013) and McInnes (1993).

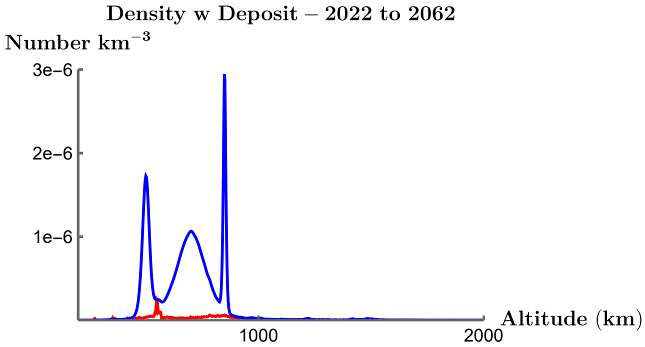

Figure 21: Evolution of the 2022 population under current depostion practices. The initial condition is shown in red while the propagated population density is in blue. Notice the tremendous growth relative to the no-deposition condition.

2.5 Discussion and conclusion

Model (5)-(11) is a first step towards an accurate continuum model and is amenable to efficient numerical implementation. Further, the inclusion of the $\Delta(r,t)$ term allows for in-depth analysis of the effect of ground launch policy on future evolution. In particular, we see that continuing on the current course of launching will cause cascading growth to occur at a sooner epoch than if the debris population were left to evolve on its own. As only a few examples have demonstrated, an important part of future research is a careful determination of the diffusion parameter and, once the technology becomes widely available, the debris removal rate.

Of particular interest to scientists and policymakers are strategies for reduction of debris input and the active removal of space debris. Such strategies are a subject of active research. Klima et al. (2018), for example, approach the topic from a game theoretic perspective. As debris removal reduces the risk to all active satellites, every stakeholder may delay their action leading to a strategic dilemma. Especially due to the cost incurred in their use, such removal strategies, while physically possible as of writing, are still very much experimental. Still, despite these shortcomings, these new technologies represent the best feasible means to avoid the Kessler syndrome. A responsible stewardship of near-earth space is clearly a task for public and private institutions alike that benefit from human space activity (Gast, 2022).

The means of these removal strategies bears some mention. Bonnal et al. (2020) describe an orbiting laser station whose purpose is to improve knowledge about the trajectories of debris objects. There is also the possibility to impart velocity changes to objects using this orbital laser. Takahashi et al. (2018) propose a concept for space debris removal by a bi-directional plasma plume ejected from a satellite. One of the beams impacts the debris object while the other keeps the cleaning satellite in place. This has been demonstrated in a laboratory experiment. All of these removal techniques act rather locally, as opposed to the global removal considered in this paper, but could be incorporated into further refinements of the model as time-varying velocity fields in addition to the passive removal of the air drag diffusivity currently present.

There exist a number of directions for future refinement and expansion of the model. For example, Figure 20 demonstrates that, despite the high accuracy of the propagation model in accounting for collision events, certain events (such as unitary breakup or the Russian anti-satellite test) can occur with no advance warning and are not directly related to the population at any time. This hampers the ability to account for such events in a purely deterministic partial differential equation model. Therefore, the inclusion of stochastic impulses into $\Delta(r, t)$ could further enhance the accuracy and predictive power of the model. While the statistics of such events are at present not widely understood, this presents a very interesting and fresh area for investigation.

Because of the importance of policy to the overall evolution of the population, choosing an effective policy is a key strategic decision. Mathematically, we can formalize this choice into choosing a policy Δ that minimizes a functional

$$I(\Delta) = \int_{t_0}^{t_1} \int_{r_E+200}^{r_E+2000} F(r, t, u, u_r, u_t, \Delta) r^2 \, \mathrm{d}r \, \mathrm{d}t,$$

where the debris populations $u(r, t)$ is constrained by the model (5) - (11). It is likely, although unverified at writing, that the value of Δ that minimizes risk would be identically 0 or a negative value, so a more interesting path might be to derive a “safe” domain in which risk, while still present, is acceptable in some sense.

3 Conclusion

The above works, while distinct in nature, share a common theme of the application of mathematics to relevant industrial and scientific problems. Indeed the work of Chapter 2 shows the applicability of methods from population dynamics, typically dealing with living agents, to the evolution of a population of “dead” objects that still follow population-based laws. We remain confident that such outside applications of mathematics have a continued future in the realm of industry and science.

We thank the reader for their time and attention.

References

Anumonwo, J. and Pandit, S. (2015). Ionic mechanisms of arrhythmogenesis. *Trends Cardiovas. Med.*, **25**:487–496.

ApolloSat.com (2022). U.S. Satellite Destroyed in Space Collision. Access at `https://apollosat.com/space-collision/`.

Balafkan, N., Mostafavi, S., Schubert, M., Siller, R., Liang, K. X., Sullivan, G., and Bindoff, L. A. (2020). A method for differentiating human induced pluripotent stem cells toward functional cardiomyocytes in 96-well microplates. *Sci. Rep.*, **10**:18498.

Bonnal, C., McKnight, D., Phipps, C., Dupont, C., Missonnier, S., Lequette, L., Merle, M., and Rommelaere, S. (2020). Just in time collision avoidance - A review. *Acta Astronaut.*, 170:637–651.

Braam, S. R., Tertoolen, L. G. J., van de Stolpe, A., Meyer, T., Passier, R., and Mummery, C. L. (2010). Prediction of drug-induced cardiotoxicity using human embryonic stem cell-derived cardiomyocytes. *Stem Cell Res.*, **4**:107–116.

Britton, O. J., Bueno-Orovio, A., Virág, L., Varró, A., and Rodriguez, B. (2017). The electrogenic Na^+/K^+ pump is a key determinant of repolarization abnormality susceptibility in human ventricular cardiomyocytes: A population-based simulation study. *Front. Physiol.*, **8**:278.

Brouwer, D. (1959). Solution of the problem of artificial satellite theory without drag. *The Astronomical Journal*, 64:378–396.

Butterworth, S. (1930). On the theory of filter amplifiers. *Experimental Wireless and the Wireless Engineer*, **7**:536–541.

Cantwell, C. D., Mohamied, Y., Tzortzis, K. N., Garasto, S., Houston, C., Chowdhury, R. A., Ng, F. S., Bharath, A. A., and Peters, N. S. (2019). Rethinking multiscale cardiac electrophysiology with machine learning and predictive modelling. *Comput. Biol. Med.*, **104**:339–351.

Carnegie Endowment for International Peace (2022). The dangerous fallout of russia's anti-satellite missile test. Access at `https://carnegieendowment.org`.

Decoste, D. and Schölkopf, B. (2002). Training invariant support vector machines. *Mach. Learn.*, **46**:161–190.

Edwards, S. L., Zlochiver, V., Conrad, D. B., Vaidyanathan, R., Valiquette, A. M., and Joshi-Mukherjee, R. (2018). A multiwell cardiac μGMEA platform for action potential recordings from human iPSC-derived cardiomyocyte constructs. *Stem Cell Rep.*, **11**:522–536. Available at `www.cell.com/stem-cell-reports`.

Elgendi, M., Eskofier, B., and Abbot, D. (2015). Fast T wave detection calibrated by clinical knowledge with annotation of P and T waves. *Sensors*, **5**:17693–17714.

Fine, T. L. (1999). *Feedforward Neural Network Methodology*. Springer-Verlag, New York.

Fozzard, H. and Beeler, Jr., G. W. (1975). The voltage clamp and cardiac electrophysiology. *Circ. Res.*, **37**:403–413.

Gast, R. (2022). Ein All für alle (Die Zeit, 07-23-2022, in German).

Gintant, G., Burridge, P., Gepstein, L., Harding, S., Herron, T., Hong, C., Jalife, J., and Wu, J. C. (2019). Use of human induced pluripotent stem cell–derived cardiomyocytes in preclinical cancer drug cardiotoxicity testing: A scientific statement from the American Heart Association. *Circ. Res.*, **125**:e75–e92.

Grandi, E., Pasqualini, F. S., and Bers, D. M. (2010). A novel computational model of the human ventricular action potential and Ca transient. *J. Mol. Cell. Cardiol.*, **48**:112–121.

Guevara, M. R., Ward, G., Shrier, A., and Glass, L. (1984). Electrical alternans and period doubling bifurcations. *IEEE Computers in Cardiology*, pages 167–170.

Harrington, L. and Johnson, E. A. (1973). Voltage clamp of cardiac muscle in a double sucrose gap. *Biophys. J.*, **13**:626–647.

Hayes, H. B., Nicolini, A. M., Arrowood, C. A., Chvatal, S. A., Wolfson, D. W., Cho, H. C., Sullivan, D. D., Chal, J., Fermini, B., Clements, M., Ross, J. D., and Millard, D. C. (2019). Novel method for action potential measurements from intact cardiac monolayers with multiwell microelectrode array technology. *Sci. Rep.*, **9**:11893.

Hodgkin, A. L. and Huxley, A. F. (1952.). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. (London)*, **117**:500–544.

Hoots, F. R. (1981). Reformulation of the Brouwer geopotential theory for improved computational efficiency. *Celestial Mechanics*, 24:367–375.

Johnston, E. (2022). List of satellites in geostationary orbit. Access at `https://www.satsig.net/sslist.htm`.

Kesmia, M., Boughaba, S., and Jacquir, S. (2019). Nonlinear dynamics of two-dimensional cardiac action potential duration mapping model with memory. *J. Math Biol.*, **78**:1529–1552.

Kessler, D. J. and Cour-Palais, B. G. (1978). Collision frequency of artificial satellites: The creation of a debris belt. *J. Geophys. Res.*, 83:2637–2646.

Klima, R., Bloembergen, D., Savani, R., Tuyls, K., Wittig, A., Sapera, A., and Izzo, D. (2018). Space debris removal: Learning to cooperate and the price of anarchy. *Front. Robot. AI*, 5:54.

Kulkarni, K. and Tolkacheva, E. (2015). Real-time feedback based control of cardiac restitution using optical mapping. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pages 5920–5923.

Kussauer, S., David, R., and Lemcke, H. (2019). hiPSCs derived cardiac cells for drug and toxicity screening and disease modeling: What micro- electrode-array analyses can tell us. *Cells*, **2019**:1331.

Lemaitre, A. and Hubaux, C. (2013). Space debris long term dynamics. In Celletti, A., Locatelli, U., Ruggeri, T., and Strickland, E., editors, *Mathematical Models and Methods for Planet Earth*, volume 6 of *INdAM*, pages 113–121. Springer, Cham, Switzerland.

Li, J., Hua, Y., Miyagawa, S., Zhang, J., Li, L., Liu, L., and Sawa, Y. (2020). hiPSC-derived cardiac tissue for disease modeling and drug discovery. *Int. J. Mol. Sci.*, **2021** 21(23):8893.

Li, J., Minami, I., Shiozaki, M., Yu, L., Yajima, S., Miyagawa, S., Shiba, Y., Morone, N., Fukushima, S., Yoshioka, M., Li, S., Qiao, J., Li, X., Wang, L., Kotera, H., Nakatsuji, N., Sawa, Y., Chen, Y., and Liu, L. (2017). Human pluripotent stem cell-derived cardiac tissue-like constructs for repairing the infarcted myocardium. *Stem Cell Reports*, **9**:1546–1559.

Liou, J.-C., Hall, D. T., Krisko, P. H., and Opiela, J. N. (2004). LEGEND – a three-dimensional LEO-to-GEO debris evolutionary model. *Adv. Space Res.*, 34:981–986.

Lyddane, R. H. (1963). Small eccentricities or inclinations in the brouwer theory of the artificial satellite. *The Astronomical Journal*, 68:555–558.

Mathur, A., Ma, Z., Loskill, P., Jeeawoody, S., and Healy, K. (2016). In vitro cardiac tissue models: Current status and future prospects. *Adv. Drug Deliver. Rev.*, **96**:203–213.

McInnes, C. R. (1993). An analytical model for the catastrophic production of orbital debris. *ESA J.*, 17:293–305.

Meyer, T., Boven, K.-H., Günther, E., and Fejtl, M. (2012). Micro-electrode arrays in cardiac safety pharmacology. *Drug Safety*, **27**:763–772.

NASA (2022). Legend: 3d/od evolutionary model. `https://orbitaldebris.jsc.nasa.gov/modeling/legend.html`, Accessed 11-20-2022.

Nikolaev, S., Phillion, D., Springer, H. K., deVries, W., Jiang, M., Pertica, A., Henderson, J., Horsley, M., and Olivier, S. (2012). Brute force modeling of the Kessler syndrome. In *Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS), Maui, HI, September 2012*.

Oppenheim, A. V. and Schafer, R. W. (1999). *Discrete-Time Signal Processing*. Prentice-Hall, Upper Saddle River, NJ, 2nd edition.

Pardini, C., Anselmo, L., Rossi, A., Cordelli, A., and Farinella, P. (1998). A new orbital debris reference model. *The Journal of the Astronomical Sciences*, 46:249–265.

Python Software Foundation (1995-2020). The Python Language. Available at `www.python.org`.

Qu, Y., Page, G., Abi-Gerges, N., Miller, P., Ghetti, A., and Vargas, H. (2018). Action potential recording and pro-arrhythmia risk analysis in human ventricular trabeculae. *Frontiers Physiol.*, **8**:1108.

Returns, A. (2022). The Six Orbital Elements. Access at `https://www.astronomicalreturns.com/p/section-43-six-orbital-elements.html`.

Riesing, K. and Kahoy, K. (2015). Orbit determination from two-line element sets for ISS-deployed CubeSats. In *29th Annual AIAA/USA Conference on Small Satellites*.

Rossi, A. (2005). Population models of space debris. In Knežević, Z. and Milani, A., editors, *Dynamics of Populations of Planetary Systems*, volume 197 of *Proceedings IAU Colloquium*, pages 427–438. International Astronomical Union.

Shang, H., Wei, S., Liu, F., Wei, D., Chen, L., and Liu, C. (2019). An improved sliding window area method for T wave detection. *Comput. Math. Meth. M.*, **2019**:3130527.

Space.com (2022a). A rogue 3-ton piece rocket debris just collided with the moon. Access at `https://www.space.com/rogue-rocket-stage-hit-moon-today`.

Space.com (2022b). Starlink satellites: Everything you need to know about the controversial internet megaconstellation. Access at `https://www.space.com/spacex-starlink-satellites.html`.

Strauss, D. G. and Blinova, K. (2017). Clinical trials in a dish. *Trends Pharmacol Sci.*, **38**:4–7.

Takahashi, K., Charles, C., Boswell, R. W., and Ando, A. (2018). Demonstrating a new technology for space debris removal using a bidirectional plasma thruster. *Sci. Rep.*, 8:14417.

Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, **126**:663–676.

Tertoolen, L. G. J., Braam, S. R., van Meer, B. J., Passier, R., and Mummery, C. L. (2018). Interpretation of field potentials measured on a multi electrode array in pharmacological toxicity screening on primary and human pluripotent stem cell-derived cardiomyocytes. *Biochem. Biophys. Res. Commun.*, **497**:1135–1141.

US Space Command (2010). Space-track.org. `Space-Track.org`, Accessed 08-08-2022.

Zhang, B., Wang, Z., and Zhang, Y. (2019). Discrete evolution model based on mean spatial density for space debris environment. *Astrophys. Space Sci.*, 364:70.

Zhu, H., Scharnhorst, K. S., Stieg, A. Z., Gimzewski, J. K., Minami, I., Nakatsuji, N., Nakano, H., and Nakano, A. (2017). Two dimensional electrophysiological characterization of human pluripotent stem cell-derived cardiomyocyte system. *Sci. Rep.*, **7**:43210.

Zhu, R., Millrod, M. A., Zambidis, E. T., and Tung, L. (2016). Variability of action potentials within and among cardiac cell clusters derived from human embryonic stem cells. *Sci. Rep.*, **6**:18544.

Zlochiver, V., Kroboth, S. L., Beal, C. R., Cook, J. A., and Joshi-Mukherjee, R. (2019). Human iPSC-derived cardiomyocyte networks on multiwell micro-electrode arrays for recurrent action potential recordings. *J. Vis. Exp.*, **149**:e59906. Available at `www.jove.com/v/59906`.

Zwi, L., Caspi, O., Arbel, G., Huber, I., Gepstein, A., Park, I.-H., and Gepstein, L. (2009). Cardiomyocyte differentiation of human induced pluripotent stem cells. *Circulation*, **120**:1513–1523.