University of Nottingham

School of Mathematical Sciences

Cell Signalling

Jake Powell

A thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

One Day

Abstract

Original abstract: Intracellular Calcium oscillations is a versatile signalling mechanism responsible for many biological phenomena including immune responses and insulin secretion. There is now compelling evidence that whole-cell calcium oscillations are stochastic, arising from random molecular interactions at the subcellular level. This poses a significant challenge for modelling. Here, we utilise a probabilistic method that treats calcium oscillations as a point process. By employing an intensity function, we capture intrinsic cellular heterogeneity as well as inhomogeneous extracellular conditions, such as time-dependent stimulation. I will demonstrate how to simulate stochastic calcium spikes based on intensity functions. Furthermore, we will infer model parameters from real data using Bayesian methods, and utilise novel MCMC techniques to deal with priors on functions.

Acknowledgements

Thank some folk.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

Calcium (Ca^{2+}) is the alkaline earth metal, first discovered by Sir Humphry Davy in 1908 by electrolysis of its oxide. **?**, the name 'Calcium' is derived from the latin calx meaning lime as it was extracted by heating limestone.

Ca^{2+} is needed for the normal function of cells and their survival. The importance of Ca^{2+} in animal cells was discovered by accident in 1883 by Sidney Ringer **?**, predating its first isolation. In 1882, Ringer investigated the affects of sodium and potassium had on the contractions of isolated frog hearts, and found that the compounds effected the heart beating. **?** However, he discovered that the laboratory ran out of distilled water during the experiment and a technician had used pipe water instead. Hence, he repeated the experiment using distilled water but found the results differed from the original experiment, and he hypothesised the observed effects were "due to the inorganic constituents of the pipe water" **?**. That constituent was Ca^{2+}.

In the years since that discovery, it is known that in addition to influencing the heart beat, Ca^{2+} has 4 main biological roles :

- cofactor for enzymes or proteins,
- electrical (i.e. in the formation of action potentials in excitable cells),

- intracellular second messenger
- structural (i.e. in skeletal structures such as bones or shells).

In this thesis we consider Ca^{2+}'s role as an intracellular second messenger, or to be more precise we investigate how changes to the Ca^{2+} concentration inside a cell informs us of the applied stimulation. We begin by introducing intracellular Ca^{2+} signalling, before reviewing the literature of what mathematical models have been used to analyse the signals. Finally, we introduce the goals of our project.

1.0.1 Calcium Signalling

In its simplest form cells at rest have a Ca^{2+} concentration of $100nM$ but activates when the level rises to $\sim 1\mu M$. This elevation of Ca^{2+} can regulate many processes due to its versatility, i.e. speed, amplitude and spatio-temporal patterning.

Calcium takes part in a variety of biological functions such as cell maturation, **?**, steering bacterial infection Van Nhieu et al. (2003) and orchestrating fertilisation Santella et al. (2004); Whitaker (2006); Denninger et al. (2014). Along with these functions excessive Ca^{2+} can damage cells and even cause apoptosis **??**.

The overarching idea of the Ca^{2+} signalling network can be split into 4 compartments:

1. Signalling is triggered by a stimulus that generates various Ca^{2+}-mobilising signals.
2. The Ca^{2+}-mobilising signals activate the ON mechanism that feeds Ca^{2+} into the cytoplasm
3. Ca^{2+} functions as a messenger to simulate numerous Ca^{2+}-sensitive processes.
4. The OFF mechanism, composed of pumps and exchangers, remove Ca^{2+} from the cytoplasm to restore the resting state.

We will illustrate the process by considering Human embryonic kidney (HEK) cells exposed to the stimulant Carbohol, which induces Ca^{2+} oscillations driven by Ca^{2+}

release through the endoplasmic reticulum (ER) through inositol-1,4,5-triphosphate ($InsP_3$) receptors ($InsP_3R$), see figure –.

Carbohol acts as an agonist of the muscarinic M_3 receptor in the extracellular medium, which itself is coupled with a G-protein which induces the activation of Phospholipase C (PLC). The increase of PLC causes subsequent prodicution of IP_3 (and diacyl-glycerol DAG), via the hydrolysis of PLC and phosphatidylinositol 4,5-bisphosphate (PIP2). Therefore Carbohol has caused the activation of IP_3, a Ca^{2+}-mobilising signal.

IP_3 then diffuses into the cell across the cell membrane, and binds to IP_3Rs situated on the ER. The ER is a compartment inside the cell that stores Ca^{2+}, on its membrane IP_3R are positioned in clusters with 5-30 IP_3R channels per cluster. **?**.

Describe the structure of IP_3R. The IP_3R opens in response to IP_3 and Ca^{2+} binding to activation sites. The activated channels release large amount of Ca^{2+} into the cytoplasm.

Some of the released Ca^{2+} binds to a receptors inactivation site which closes the channel. The channel cannot reopen for some time after inactivation (the channel is in a refractory state) **??**. After IP_3R closure cystolic Ca^{2+} is removed through plasma membrane ATPase (PCMA) and sarco-plasmic reticulum Ca^{2+}ATPase (SERCA) pumps out of the cell and into the ER respectively **?**. Cytosolic Ca^{2+} is also controlled by buffers - compounds that bind free Ca^{2+} such as ATP and calretinin - hence allowing Ca^{2+} concentration to remain in an operational range. Since Ca^{2+} is cytotoxic. The intracellular Ca^{2+} buffering capacity depends on the type of cell **???**

Repeated increases and decreases in Ca^{2+} caused by the opening/closing of IP_3R along with Ca^{2+} pump activity creates Ca^{2+} oscillations. Depending on the cell type the period of oscillations varies from seconds to minutes, i.e. 10ms spike duration in toadfish swimbladder muscle or 1 day for circadian rhythms **??** and oscillations range from regular spiking to bursting **?**.

Changes in cytosolic Ca^{2+} concentration generally doesn't occur uniformly across the cell, rather local Ca^{2+} changes can create Ca^{2+} events are often categorised into 3 groups: blips, puffs and waves. A Blip occurs when Ca^{2+} is released from a

single receptor, whereas a puff arises when a cluster of receptors is activated. Ca^{2+} can diffuse from one puff site to surrounding clusters activating the receptors and causing wave propagation. In astocytes at least three puff sites need to be activated synchronously to cause a wave **?**. Intracellular Ca^{2+} waves spread throughout the cell and can transfer information from one part of the cell to another.

1.0.1.1 Encoding and decoding Ca^{2+} Patterns

Cells use spatio-temporal patterns to transmit information (both within the same cell and to surrounding cells) **???** and consequently to initiate an appropriate physiological response. **??**. However, it is not yet fully understood how encoding and decoding mechanisms convey information. To date, it has been shown that Ca^{2+} responses can vary amongst different cell types **?**, and even in the same cell type **?**. Along with cell type, the agonist type **???** and its concentration **?** also affects Ca^{2+} response.

The relationship between Ca^{2+} oscillations and agonist was first studied in 1986 by Woods et al. **?** They found that Ca^{2+} oscillations depended on the hormone concentration used for simulation in single, isolated rat hepatocyte cells. From this, they hypothesised that only the frequency of Ca^{2+} oscillations were effected by the agonist and both amplitude and width of the oscillations were unaffected, which was confirmed in 1987 **?**.

Two hypothesis' for stimulus encoding followed from the fixed agonist experiments. Namely, amplitude and frequency modulation (AM and FM **?**. AM proposes that the concentration of agonist increases the amplitude of Ca^{2+} signals, whereas FM claims that the frequency of the Ca^{2+} signals increase with the stimulus strength. Recent studies have shown that Ca^{2+} spike times scales exponentially with the stimulus strength for FM K et al. (2014).

1.0.2 How do we generate spike sequences?

The aim of our work is to fit statistic models to real spike data. However, before any model fitting it is best to understand how we get spike data from individual cells.

1.0.3 Calcium Signalling Models

Calcium signalling is often modelled using either deterministic or stochastic models. In this section we provide a short overview of these models. For a more detailed view, we refer readers to **?????**.

1.0.3.1 Ca^{2+} oscillation modelling

We have shown that cytosolic Ca^{2+} concentration, $c(t)$) is controlled through IP_3R and SERCA/PMCA pumps. IP_3R controls Ca^{2+} release from the ER into the cytoplasm (J_{IP_3R}). Whereas SERCA/PMCA pumps transport Ca^{2+} from the cytoplasm back into the ER (J_{SERCA}) and outside the cell (J_{PMCA}) respectively. Combining the fluxes, we can mathematically describe the Ca^{2+} concentration by

$$\frac{dc}{dt} = J_{IP_3R} - J_{SERCA} - J_{PMCA}. \tag{1.1}$$

We could include more terms such as the flux into the cytoplasm from RyR, or leak from the internal stores into the cytosol. One of the first realistic models for Ca^{2+} oscillations was produced by De Young and Keizer (DYK) in 1992 **?**. The model became an archetype for more recent deterministic **???** and stochastic models **??**.

The DYK model considers only two fluxes, one into the cytosol via the IP_3R channels and a leak from the internal store, and one out of the cytosol via a ATP-dependent pump (i.e. the SERCA pump). The SERCA pump is described by the Hill function of $c(t)$

$$J_{SERCA} = \frac{vc^2}{k^2 + c^2}, \tag{1.2}$$

where parameters v, k represent maximum Ca^{2+} uptake and pump activation constant respectively **?**.

Recall that studies have found that the IP_3R is activated when at least 3 of the 4 subunits are activated. Thus the DYK model assumes that an IP_3R consists of 3 identical independent subunits. Each subunit contains one IP_3R binding site, one Ca^{2+} activation site and one Ca^{2+} inhibition site. As such each subunit can be in one of 8 states (figure –), which can be specified by the binary triplet $ijk \in \{0, 1\}^3$,

where the indexes correspond to the IP_3R site, Ca^{2+} activation site, Ca^{2+} inhibition site in order. An index of 0 means the site is unbound whereas 1 represents the site been bound. A subunit is activated when it is in state 110 i.e. the IP_3R site and Ca^{2+} activation site is bound and the Ca^{2+} Inhibition site is unbound. Hence the the channel is open when all 3 subunits are activated. The probability that the subunit is in state ijk is denoted by x_{ijk}. By applying mass-action kinetics we retrieve eight ODEs for each subunit state, for example

$$\frac{dx_{000}}{dt} = b_1 x_{100} + b_5 x_{010} + b_4 x_{001} - x_{000}\left(a_1 I + a_4 c + a_5 c\right). \tag{1.3}$$

However, since we must be in one of the 8 states we can replace one of the ODEs with the condition $\sum x_{ijk} = 1$. IP_3 concentration can be fixed, however for more realistic results it is possible to include it's own dynamics, e.g **????**.

Thus, with the IP_3R subunit open probability of x_{110}, the outward flux of the DYK model can be computed by

$$J_{IP_3R} = c_1\left(v_1 x_{110}^3 + v_2\right)\left(c_{er} - c\right), \tag{1.4}$$

where the parameters c_1, c_{er}, v_1 and v_2 correspond to the ratio of ER to the cytosolic volume, Ca^{2+} concentration in the ER, maximal Ca^{2+} influx and Ca^{2+} leak respectively.

1.0.3.2 Ca^{2+} waves

Whilst the DYK improved understanding of Ca^{2+} signalling toolbox, it only considers temporal changes in the cytosolic Ca^{2+} concentration. It is know (see figure –) that Ca^{2+} patterns vary in both time and space. To incorporate this into (**??**) a diffusion term is added

$$\frac{\partial c}{\partial t} = D_c \frac{\partial^2 c}{\partial x^2} \tag{1.5}$$

1.0.3.3 Stochastic Models

CHAPTER 2

Model and Bayesian Inference

2.1 The Model

2.2 Simulating spike sequences

CHAPTER 3

Non-trivial Model Properties

3.1 ISI Distribution

3.1.1 Gamma distribution

Throughout we have assumed that the Gamma ISI distribution is as follows

$$p(y_i, y_{i-1}) = \frac{\gamma x(y_i)}{\Gamma(\gamma)} \big[\gamma X(y_{i-1}, y_i)\big]^{\gamma-1} \exp(-\gamma X(y_{i-1}, y_i)). \tag{3.1}$$

Which one can see a generalisation of a Gamma(γ, γ). We have used this distribution because it was used in . However, can we generalise further and consider a Gamma(α, β) ISI distribution? The answer is no. This is because the parameters x and β are non-identifiable.

Suppose our ISI interval is given by the following (α, β) Gamma distribution.

$$p(y_i, y_{i-1}) = \frac{\beta x(y_i)}{\Gamma(\alpha)} \big[\beta X(y_{i-1}, y_i)\big]^{\alpha-1} \exp(-\beta X(y_{i-1}, y_i)), \tag{3.2}$$

where

$$X(s,t) = \int_s^t x(u)du, \qquad \text{for any } s, t \in [0, T]. \tag{3.3}$$

Now we will show that this is equivalent to a (γ, γ) Gamma distribution by scaling the intensity function.

Let $\gamma = \alpha$ and $\kappa = \frac{\beta}{\alpha}$, thus giving $\beta = \gamma\kappa$. Substitute into (2) gives

$$p(y_i, y_{i-1}) = \frac{\gamma\kappa x(y_i)}{\Gamma(\gamma)} \left[\gamma\kappa X(y_{i-1}, y_i)\right]^{\gamma-1} \exp(-\gamma\kappa X(y_{i-1}, y_i)). \tag{3.4}$$

Now define new intensity function by $\tilde{x} = \kappa x$, hence giving

$$\tilde{X}(s,t) = \int_s^t \tilde{x}(u)du = \int_s^t \kappa x(u)du = \kappa X(s,t), \qquad \text{for any } s, t \in [0, T]. \tag{3.5}$$

Put this into (3) gives

$$p(y_i, y_{i-1}) = \frac{\gamma\tilde{x}(y_i)}{\Gamma(\gamma)} \left[\gamma\tilde{X}(y_{i-1}, y_i)\right]^{\gamma-1} \exp(-\gamma\tilde{X}(y_{i-1}, y_i)), \tag{3.6}$$

which we see is a (γ, γ) Gamma distribution.

Hence, where we have introduced a two-variable Gamma distribution we cannot retrieve initial choices of variables. I.e β is replaced as a variable by the intensity function. The choice of having $\alpha = \beta$ is a sensible choice as this allows the expected number of spikes to not depend on the ISI parameters and only depends on x.

3.1.2 Inverse Gaussian distribution

Similarly we have not considered the full generalised Inverse Gaussian ISI distribution

$$p(y_i, y_{i-1}) = x(y_i) \left(\frac{\lambda}{2\pi X(y_{i-1}, y_i)^3}\right)^{0.5} \exp\left[-\frac{\lambda(X(y_{i-1}, y_i) - \mu)^2}{2\mu^2 X(y_{i-1}, y_i)}\right]. \tag{3.7}$$

However, we again get non-identifiable variables, in this case if we multiple all the parameters by a constant we get an identical distribution.

Suppose we have an Inverse Gaussian ISI distribution with parameters (λ, μ) and intensity function x. We want to show this is identical to an IG$((k\lambda, k\mu))$ with

intensity kx for any constant k. Firstly, note that the constant comes out the front of of the integral of the intensity function.

$$\int_s^t \kappa x(u)du = \kappa \int_s^t x(u)du = \kappa X(s,t), \qquad \text{for any } s,t \in [0,T]. \tag{3.8}$$

Hence, we get

$$p(y_i, y_{i-1}|kx, k\lambda, k\mu) = kx(y_i)\left(\frac{k\lambda}{2\pi(kX(y_{i-1},y_i))^3}\right)^{\frac{1}{2}} \exp\left[-\frac{k\lambda(kX(y_{i-1},y_i)-k\mu)^2}{2(k\mu)^2 kX(y_{i-1},y_i)}\right], \tag{3.9}$$

$$= kx(y_i)\left(\frac{1}{k^2}\frac{\lambda}{2\pi X(y_{i-1},y_i)^3}\right)^{\frac{1}{2}} \exp\left[-\frac{k\lambda(k(X(y_{i-1},y_i)-\mu))^2}{2k^3\mu^2 X(y_{i-1},y_i)}\right], \tag{3.10}$$

$$= \frac{k}{k}x(y_i)\left(\frac{\lambda}{2\pi X(y_{i-1},y_i)^3}\right)^{\frac{1}{2}} \exp\left[-\frac{k^3}{k^3}\frac{\lambda(X(y_{i-1},y_i)-\mu)^2}{2\mu^2 X(y_{i-1},y_i)}\right], \tag{3.11}$$

$$= x(y_i)\left(\frac{\lambda}{2\pi X(y_{i-1},y_i)^3}\right)^{\frac{1}{2}} \exp\left[-\frac{\lambda(X(y_{i-1},y_i)-\mu)^2}{2\mu^2 X(y_{i-1},y_i)}\right], \tag{3.12}$$

$$= p(y_i, y_{i-1}|x, \lambda, \mu) \qquad \text{as required.} \tag{3.13}$$

3.1.3 Log Normal distribution

Next we shall consider the full Log Normal distribution

$$p(y_i, y_{i-1}) = \frac{x(y_i)}{X(y_{i-1},y_i)\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\log X(y_{i-1},y_i)-\mu)^2}{2\sigma^2}\right\} \tag{3.14}$$

However as in the above two cases we get non-identifiable variables. In this case, if we multiply the intensity by a factor of κ this is identical to subtracting $\log\kappa$ to μ. Suppose we have an Log Normal ISI distribution with parameters (μ, σ) and intensity function κx. We want to show this is identical to an LN$((\mu + \log\kappa, \sigma))$ with intensity κx for any constant κ.

$$p(y_i, y_{i-1}|\kappa x, \mu, \sigma) = \frac{\kappa x(y_i)}{\kappa X(y_{i-1}, y_i)\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\log \kappa X(y_{i-1}, y_i) - \mu)^2}{2\sigma^2}\right\}, \quad (3.15)$$

$$= \frac{x(y_i)}{X(y_{i-1}, y_i)\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\log\ X(y_{i-1}, y_i) + \log\kappa - \mu)^2}{2\sigma^2}\right\} \quad (3.16)$$

$$= \frac{x(y_i)}{X(y_{i-1}, y_i)\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\log\ X(y_{i-1}, y_i) - (\mu - \log\kappa))^2}{2\sigma^2}\right\} \quad (3.17)$$

$$= p(y_i, y_{i-1}|x, \mu - \log\kappa, \sigma) \qquad \text{as required.} \quad (3.18)$$

3.1.4 Weibull distribution

Next we shall consider the Weibull distribution

$$p(y_i, y_{i-1}|k, \lambda) = \frac{x(y_i)k}{\lambda}\left(\frac{X(y_{i-1}, y_i)}{\lambda}\right)^{k-1} \exp\left\{-\left(\frac{X(y_{i-1}, y_i)}{\lambda}\right)^k\right\} \quad (3.19)$$

However again we get non-identifiable variables. In this case, if we multiply the intensity by a factor of m this is identical to dividing λ by m.

Suppose we have an Weibull ISI distribution with parameters (k, λ) and intensity function mx. We want to show this is identical to an Weibull$(k, \frac{\lambda}{m})$ with intensity x for any constant m.

$$p(y_i, y_{i-1}|mx, k, \lambda) = \frac{mx(y_i)k}{\lambda}\left(\frac{mX(y_{i-1}, y_i)}{\lambda}\right)^{k-1} \exp\left\{-\left(\frac{mX(y_{i-1}, y_i)}{\lambda}\right)^k\right\} \quad (3.20)$$

$$= \frac{x(y_i)k}{\lambda/m}\left(\frac{X(y_{i-1}, y_i)}{\lambda/m}\right)^{k-1} \exp\left\{-\left(\frac{X(y_{i-1}, y_i)}{\lambda/m}\right)^k\right\} \quad (3.21)$$

$$= p(y_i, y_{i-1}|x, k, \lambda/m) \qquad \text{as required.} \quad (3.22)$$

3.2 Speed of the GP

In this section we discuss how a naive implementation of the GP prior in chapter 1 will lead to a computationally expensive algorithm. We will explain different methods to improve the algorithm, and justify our implementation.

In this framework we assume a priori that the intensity function $x(t)$ comes from a GP. To allow for computations we discretise time, often set to the frame rate of experimental data. Thus, the prior reduced to a multivariate normal distribution, over a large number of indexes, often greater then 1000 steps. To sample from the posterior intensity we utilise an under-relaxed method, where in each iteration we propose a candidate intensity x_{can} depending on the current value x_{cur} by

$$x_{\text{can}} = \sqrt{1-\omega^2}\log\left(x_{\text{cur}}\right) + \omega\nu, \qquad \nu \sim \mathcal{N}\left(\mathbf{0}, \Sigma\right)$$

Thus, in each iteration we must generate a draw (ν) from the multivariate normal distribution. This requires the inversion and determinant of the covariance matrix to be calculated, whose dimension is the discretisation chosen for the intensity, thus we need to be able to invert 1000×1000 matrices, quickly. There exists different methods to compute the matrix decomposition such as via eigenvalues?? or a choleski decomposition. However, these methods tend to be time consuming once the dimension increases.

We have considered two different methods to improve the time taken, namely by implementing a projection when computing likelihoods and by expressing the GP by its spectral decomposition and simulating by a fast fourier transform (FFT).

3.2.1 Projection

Explain how projection works.

3.2.2 Spectral representation

What is spectral density function? Define/explain. In the case of a GP with 0 mean and squared exponential covariance function (write equation?) the spectral density is

given by

$$S(s) = \sigma^2 \left(2\pi\ell^2\right)^{1/2} \exp\left(-2\pi^2\ell^2 s^2\right)$$

However, not all kernels have a closed form for the spectral density, as such we can numerically calculate S via a FFT. Explain more.

Since A zero-mean GP is a real-values 1D-IV stationary stochastic process we can use the following theorem to express the gaussian process as an infinite series. *To every real-valued 1D-IV stationary stochastic process $f_0(t)$ with mean value equal to zero and two-sided power spectral density function $S_{f_0,f_0}(\omega)$, two mutually orthogonal real processes $u(\omega)$ and $v(\omega)$ with orthogonal increments $du(\omega)$ and $dv(\omega)$ can be assigned such that*

$$f_0(t) = \int_0^\infty [\cos(\omega t)du(\omega) + \sin(\omega t)dv(\omega)]$$

By the infinite series representation we can simulate the GP by the following series as $N \to \infty$

$$f(t) = \sqrt{2}\sum_{n=0}^{N-1} A_n \cos(\omega_n t + \Phi_n)$$

where

$$A_n = \left(2S(\omega_n)\Delta\omega\right)^{1/2}, \qquad n = 0, 1, 2, \ldots, N-1$$

$$\omega_n = n\Delta\omega, \qquad n = 0, 1, 2, \ldots, N-1$$

$$\Delta\omega = \omega_c/N$$

and

$$A_0 = 0 \qquad \text{or} \qquad S(\omega_0 = 0) = 0$$

In equation - ω_c represents an upper cutoff frequency for which the spectral density is assumed to be zero for any larger frequency. To estimate ω_c we use the following

Algorithm 1: Simulating Gaussian Process via spectral decomposition.

Input: N, δt, $k(t_1, t_2)$

Output: Draw from GP

Set $t = 0$, $y_{\text{cur}} = 0$, $\mathbf{y} = \emptyset$ and $h = T/K$;

while $t < T$ **do**

  Draw $a \sim U(0, 1)$ and set $C = 0$;

  while $C < a$ **do**

    $t = t + h$;

    if $t > T$ **then**

      STOP;

    Numerically calculate $C = \int_{y_{\text{cur}}}^{t} p(y_{\text{cur}}, u|x(u), \theta)du$;

  Add t to the set of spike times $\mathbf{y}$, and set $y_{\text{cur}} = t$;

Return spike times $\mathbf{y}$;

criteria

$$\int_0^{\omega_c} S(\omega)d\omega = (1 - \epsilon) \int_0^{\infty} S(\omega)d\omega$$

To further improve the computation time we can rewrite this infinite series to a format which allows for the use of fast Fourier transform. Rewriting we get

$$f(p\Delta t) = \Re \left(\sum_{n=0}^{M-1} B_n \exp \left[i(n\Delta\omega)(p\Delta t) \right] \right)$$

and B_n The limitations of sampling from the GP in this method is that the step size, and number of steps is no longer independent. However, by choosing the value of M in a smart way we can still maintain the correct step size. Via

3.3 Estimating Length Scale

CHAPTER 4

Refractory Period

4.0.1 How it affects the model

We will consider having a Gamma ISI throughout, although the same concept has been used for Poisson, inverse Gaussian and Weibull. Previously, without the inclusion of $T_{\min}$ we have the following ISI distribution.

$$p(y_i, y_{i-1}|x) = \frac{\gamma x(y_i)}{\Gamma(\gamma)} \big[\gamma X(y_{i-1}, y_i)\big]^{\gamma-1} \mathrm{e}^{-\gamma X(y_{i-1}, y_i)}, \tag{4.1}$$

where

$$X(s,t) = \int_s^t x(u)du, \qquad \text{for any } s,t \in [0,T], \tag{4.2}$$

where $\gamma > 0$ is the ISI parameter and Γ is the Gamma function.

Now with the inclusion of $T_{\min}$ the ISI distribution changes to

$$p(y_i, y_{i-1}|x) = \begin{cases} 0 & \text{if } y_i - y_{i-1} < T_{\min} \\ \frac{\gamma x(y_i)}{\Gamma(\gamma)} \big[\gamma X(y_{i-1}, y_i)\big]^{\gamma-1} \mathrm{e}^{-\gamma X(y_{i-1}, y_i)} & \text{o/w} \end{cases} \tag{4.3}$$

However, we can rewite this equation and alter the definition of X to account for

the refractory period. Thus,

$$p(y_i, y_{i-1}|x) = \frac{\gamma x(y_i)}{\Gamma(\gamma)}\left[\gamma \tilde{X}(y_{i-1}, y_i)\right]^{\gamma-1} \mathrm{e}^{-\gamma \tilde{X}(y_{i-1}, y_i)}, \tag{4.4}$$

where

$$\tilde{X}(s,t) = \int_{s+T_{\min}}^{t} x(u)du, \qquad \text{for any } s,t \in [0,T], \tag{4.5}$$

Next we want to calculate the likelihood for an entire spike sequence. In general we have:

$$p(\mathbf{y}|x) = p_1(y_1|x)p_T(T, y_N|x)\prod_{i=2}^{N} p(y_i, y_{i-1}|x). \tag{4.6}$$

Above we have dealt with the ISI distribution, however we need to decide how $T_{\min}$ affects p_1 and p_T. We assume that $T_{\min}$ has no effect on p_1, since we do not know the time of the last spike before the initial spike. Hence, the refractory period may already of occurred. Thus,

$$p_1(y_1|x) = x(y_1)e^{-X(0,y_1)} \tag{4.7}$$

$T_{\min}$ however does effect p_T. If the time difference between the last spike and end time is less then the refractory period then p_T can be disregarded as it has no effect on the likelihood. If the difference is longer then the refractory period then we have

$$p_T(T, y_N|x) = e^{-\tilde{X}(y_N, T)}, \tag{4.8}$$

Thus when calculating the likelihood for x we do not need to change the formula, we only need to change our calculation of X by including $T_{\min}$ to the bottom of the integrand. However we will need to distinguish the case whether to include a term for p_T or not.

4.0.2 Estimating $T_{\min}$

We estimate $T_{\min}$ using MCMC, where if we are estimating other parameters ($x(t)$/ISI parameters) we use Gibbs sampling and sample from $T_{\min}$ separately.

$T_{\min}$ only affects the likelihood via $X(y_{i-1}, y_i)$ hence we get the marginal for each ISI distribution reduces to :

Distribution	log marginal
Poisson	$-X(y_{i-1}, y_i)$
Gamma	$(\gamma - 1) \log X(y_{i-1}, y_i) - \gamma X(y_{i-1}, y_i)$
Inverse Gaussian	1
Log Normal	1
Weibull	

CHAPTER 5

Application to Real Data

words

5.1 Section

words

And a formula too (a complicated one... Ohm's law)

$$V = \rho \frac{l}{a} I \,. \tag{5.1}$$

And also a more complicated one to show the maths font:

$$Q = \mathcal{L} + \sum_{i=1}^{N} \prod_{j=1}^{N} a_{ij} \,. \tag{5.2}$$

CHAPTER 6

Conclusion

And a formula too (a complicated one... Ohm's law)

$$V = \rho \frac{l}{a} I \,. \tag{6.1}$$

And also a more complicated one to show the maths font:

$$Q = \mathcal{L} + \sum_{i=1}^{N} \prod_{j=1}^{N} a_{ij} \,. \tag{6.2}$$

Bibliography

Denninger, P., Bleckmann, A., Lausser, A., Vogler, F., Ott, T., Ehrhardt, D. W., Frommer, W. B., Sprunck, S., Dresselhaus, T., and Grossmann, G. (2014). Male–female communication triggers calcium signatures during fertilization in arabidopsis. *Nature Communications*, 5(1).

K, T., SC, T., G, M., VL, P., A, M., AP, T., A, S., CW, T., and M, F. (2014). Reliable encoding of stimulus intensities within random sequences of intracellular Ca^{2+} spikes. *Science Signaling*, 7(331):ra59–ra59.

Santella, L., Lim, D., and Moccia, F. (2004). Calcium and fertilization: the beginning of life. *Trends in Biochemical Sciences*, 29(8):400–408.

Van Nhieu, G. T., Clair, C., Bruzzone, R., Mesnil, M., Sansonetti, P., and Combettes, L. (2003). Connexin-dependent inter-cellular communication increases invasion and dissemination of shigella in epithelial cells. *Nature Cell Biology*, 5(8):720–726.

Whitaker, M. (2006). Calcium at fertilization and in early development. *Physiological Reviews*, 86(1):25–88.