Chapter 32
A Statistical View on Calcium Oscillations

Jake Powell, Martin Falcke, Alexander Skupin, Tomas C. Bellamy, Theodore Kypraios, and Rüdiger Thul

Abstract Transient rises and falls of the intracellular calcium concentration have been observed in numerous cell types and under a plethora of conditions. There is now a growing body of evidence that these whole-cell calcium oscillations are stochastic, which poses a significant challenge for modelling. In this review, we take a closer look at recently developed statistical approaches to calcium oscillations. These models describe the timing of whole-cell calcium spikes, yet their parametrisations reflect subcellular processes. We show how non-stationary calcium spike sequences, which e.g. occur during slow depletion of intracellular calcium stores or in the presence of time-dependent stimulation, can be analysed with the help of so-called intensity functions. By utilising Bayesian concepts, we demonstrate how values of key parameters of the statistical model can be inferred from single cell calcium spike sequences and illustrate what information whole-cell statistical models can provide about the subcellular mechanistic processes that drive calcium oscillations. In particular, we find that the interspike interval distribution of HEK293 cells under constant stimulation is captured by a Gamma distribution.

J. Powell · T. Kypraios · R. Thul (✉)
Centre for Mathematical Medicine and Biology, School of Mathematical Sciences, University of Nottingham, Nottingham, UK
e-mail: pmxjp8@exmail.nottingham.ac.uk; theodore.kypraios@nottingham.ac.uk; ruediger.thul@nottingham.ac.uk

M. Falcke
Max Delbrück Centre for Molecular Medicine, Berlin, Germany

Department of Physics, Humboldt University, Berlin, Germany
e-mail: martin.falcke@mdc-berlin.de

A. Skupin
Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Belval, Luxembourg

National Biomedical Computation Resource, University California San Diego, La Jolla, CA, USA
e-mail: alexander.skupin@uni.lu

T. C. Bellamy
School of Life Sciences, University of Nottingham, Nottingham, UK
e-mail: tomas.bellamy@nottingham.ac.uk

32.1 Introduction

Calcium (Ca^{2+}) oscillations have long been recognised as a centrepiece in the world of intracellular Ca^{2+} signals [1–9]. Acting as a ubiquitous and versatile signalling mechanism, Ca^{2+} oscillations are responsible for inducing gene expression [10–12], controlling hormone secretion [13–17], orchestrating fertilisation [18–20] and steering bacterial invasion [21], to name but a few cellular functions. The notion of Ca^{2+} oscillations usually refers to transient increases in the whole-cell Ca^{2+} concentration that present themselves as a series of Ca^{2+} spikes. Since whole-cell calcium recordings yield averaged concentration values, it has often been assumed that mathematical models of intracellular Ca^{2+} oscillations can be directly based on the averaged Ca^{2+} concentration. To illustrate this concept, consider Ca^{2+} oscillations driven by Ca^{2+} release from the endoplasmic reticulum (ER) through inositol-1,4,5-trisphosphate ($InsP_3$) receptors ($InsP_3Rs$). In its simplest incarnation, these mathematical models assume that Ca^{2+} transport through all open $InsP_3Rs$ and the activity of all sarco-endoplasmic Ca^{2+} ATP (SERCA) pumps can be averaged across the cell to yield averaged Ca^{2+} release and resequestration, respectively. Since the activity of both $InsP_3Rs$ and SERCA pumps depends on the cytosolic Ca^{2+} concentration, these models implicitly assume that the gating of $InsP_3Rs$ and SERCA pumps is controlled by averaged Ca^{2+} concentration values.

This assumption may serve as a starting point to explore Ca^{2+} dynamics in systems for which detailed Ca^{2+} measurements are missing, and models based on averaged Ca^{2+} concentrations have been instrumental in furthering our understanding of Ca^{2+} oscillations [13, 16, 22–48]. However, the notion of mean Ca^{2+} values generally falls short of capturing the biology that underlies Ca^{2+} oscillations. The main reason for this is that $InsP_3Rs$ form clusters that are distributed throughout the cell at distances of 2–7µm [49–53]. This entails that the dynamics of $InsP_3Rs$ is controlled by the *local* Ca^{2+} concentration, not a global average. In other words, measuring the Ca^{2+} concentration across a cell, taking the spatial average and determining the gating of all $InsP_3Rs$ subject to the averaged Ca^{2+} concentration misrepresents the actual $InsP_3R$ dynamics. In addition, there are only a few tens of $InsP_3Rs$ per cluster [54–56]. Since binding of Ca^{2+} and $InsP_3$ to $InsP_3Rs$ is random and hence transitions between different states of the $InsP_3R$ occur stochastically, the relative fluctuation in the number of open $InsP_3Rs$ is considerable. This stochasticity might even be enhanced by the fact that at basal Ca^{2+} concentration, the actual numbers of Ca^{2+} ions in the vicinity of an $InsP_3$ is small [57–60]. Taken together, these observations strongly suggest that intracellular Ca^{2+} is a spatially extended stochastic medium, which prompts the question on how to best describe $InsP_3$ mediated Ca^{2+} oscillations mathematically.

One approach starts with the dynamics of single $InsP_3Rs$, groups them into clusters and then places the clusters into a three-dimensional representation of the

cytosol—see [61] for a recent perspective. In these models, InsP$_3$Rs are described by stochastic models known as Markov chains, which consist of different states of the InsP$_3$R such as open, closed and inhibited and contain rules for stochastic transitions between different states. Clusters of InsP$_3$Rs communicate with each other through Ca^{2+} diffusion. One advantage of such hierarchical modelling lies in its mechanistic interpretation. It allows questions to be answered about how Ca^{2+} oscillations are shaped by e.g. the distance between InsP$_3$R clusters, single channel current and Ca^{2+} buffers. However, these models require as input a significant number of parameters, such as gating constants for the InsP$_3$R, and are computationally expensive. In order to reduce the computational load, Langevin-type models have been put forward. In essence, they approximate the exact stochastic dynamics of the Markov chains.

In terms of modelling philosophy, the above approaches fall into the category of bottom-up techniques. At the other end of the spectrum lie so-called top-down methods. Here, we construct models that directly describe key properties of Ca^{2+} spikes such as amplitude and frequency without explicitly incorporated mechanistic details as e.g. the possible states of an InsP$_3$R. At first sight, this might appear less advantageous as different model behaviours cannot immediately be linked to specific molecular processes. However, there are distinct advantages. Firstly, the computational demand is significantly lower than with bottom-up approaches. This puts us in an ideal position to generate large numbers of realistic Ca^{2+} spike sequences, which in turn can serve as input to signalling cascades that decode Ca^{2+} spikes. Secondly, top-down models provide a powerful framework for fitting data and testing hypotheses on Ca^{2+} spike generation. Consequentially, we can use the knowledge gained from top-down models to improve bottom-up approaches, which in turn will advance our mechanistic understanding of Ca^{2+} oscillations.

In this review, we present the current state of statistical modelling of Ca^{2+} oscillations. The techniques that we employ are well established amongst statisticians, but are less familiar to modellers and experimentalists in the field of Ca^{2+} signalling. We therefore mainly focus on describing the underlying concepts and how they are related to the physiology of Ca^{2+} signalling. We discuss practical approaches for how to ascertain whether our statistical assumptions are consistent with measured Ca^{2+} spike sequences and what we can learn from our statistical analysis regarding the mechanisms that underlie Ca^{2+} spike generation.

32.2   Interspike Interval Statistics

We outlined in the introduction the mechanistic reasons for why Ca^{2+} oscillations are stochastic. At this point, one might argue—as is often done—that the molecular fluctuations present at InsP$_3$R clusters average out at the whole-cell level. In other words, since a cell can contain a large number of InsP$_3$R clusters, the stochastic contributions cancel. To test this hypothesis, Skupin et al [62] measured spontaneous Ca^{2+} oscillations in microglia, astrocytes and PLA cells as well as carbachol-induced oscillations in HEK293 cells. They found that their data is

consistent with stochastic whole-cell Ca^{2+} oscillations, which was also confirmed in later experiments [63]. This conclusion rests on results as shown in Fig. 32.1. We plot representative fluorescence traces for carbachol stimulated HEK293 cells in Fig. 32.1a, b. Cells were initially stimulated with $20\,\mu M$ carbachol before the solution was switched to $50\,\mu M$ carbachol. Figure 32.1a illustrates the well-known phenomenon of frequency encoding, by which the frequency of Ca^{2+} oscillations increases with an increase in stimulation strength. In Fig. 32.1c we plot the Ca^{2+} spike times for a larger number of cells. If Ca^{2+} oscillations were deterministic and governed by averaged Ca^{2+} concentrations, we would expect an almost constant spread of Ca^{2+} spike times, i.e. an almost constant value for the interspike interval (ISI), not the observed large variability, which is present at both stimulation strengths. Our argument for stochastic Ca^{2+} oscillations is further strengthened by the results shown in Fig. 32.1d. Here, each triangle corresponds to a sequence of Ca^{2+} spikes from one cell and denotes its mean μ and its standard deviation σ. We observe that the standard deviation is of the same magnitude as the mean, which is another strong indicator of stochastic behaviour. Importantly, similar results have been obtained for a number of additional cell types and under different conditions [61, 62], which lends even more support for the stochasticity of Ca^{2+} oscillations. Given the insights that $\sigma - \mu$ plots can provide into the nature of Ca^{2+} oscillations, we have recently released CaSiAn [64], a user friendly tool that allows for automatic ISI detection from fluorescence time course data and interactive investigation of the relationship between μ and σ.

To appreciate the fact that μ and σ are of the same order of magnitude, we introduce a key concept for this review: the conditional Ca^{2+} spike intensity $q(t|s)$, $t > s$. Based on it, we obtain the conditional Ca^{2+} spike probability $q(t|s)dt$, which represents the probability to observe a Ca^{2+} spike in the time interval $[t, t + dt]$ given a Ca^{2+} spike at time s. In [62], the following ansatz was made:

$$q(t|s) = \begin{cases} 0, & s \leq t \leq T_r + s, \\ \lambda \left[1 - e^{-\xi(t-s-T_r)}\right], & T_r + s \leq t. \end{cases} \tag{32.1}$$

Here, T_r denotes the cellular refractory period. Numerous experiments have shown that there exists a minimal amount of time T_r after a Ca^{2+} spike before another Ca^{2+} spike can be triggered [62, 63, 66]. Therefore, the conditional intensity vanishes, i.e. $q = 0$, for a time T_r after the last Ca^{2+} spike. It is important to note that T_r is significantly longer than the recovery time of InsP$_3$Rs [63]. Once the refractory period has passed, the conditional intensity for a Ca^{2+} spike starts to increase at a rate ξ and eventually approaches an equilibrium value λ. This reflects the notion that a cell has to recover from the last Ca^{2+} spike. While ξ is a single number, it subsumes numerous recovery processes such as refilling of the ER or replenishment of InsP$_3$ following degradation by InsP$_3$-3-kinase and InsP$_3$-5-phosphatase. The values of T_r, ξ and λ can be directly inferred from Fig. 32.1d as outlined below. Due to the strong linear relationship between the mean and the standard deviation, we posit that
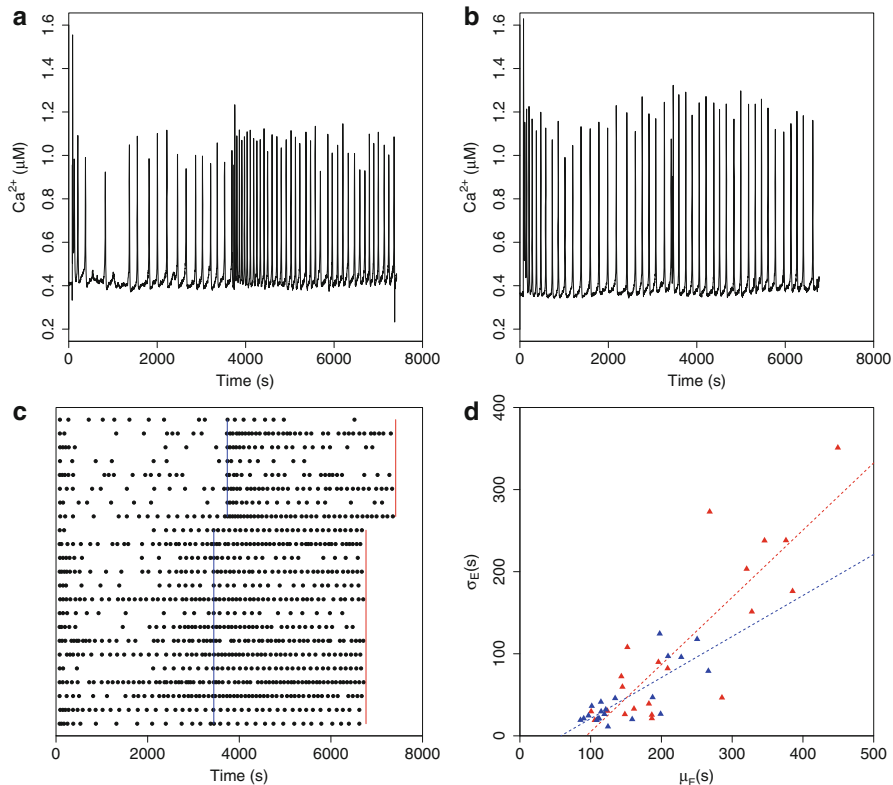
Fig. 32.1 (**a**, **b**) Fura-2 fluorescence intensity traces of two HEK293 cells stimulated first with 20 μM carbachol and then with 50 μM carbachol. The solution was exchanged at 3738s in (**a**) and at 3444s in (**b**). (**c**) Raster plot of Ca^{2+} spike times for the same stimulus protocol as in (**a**, **b**). The blue line indicates solution exchange and the red line denotes the end of the experiment. (**d**) Relationship between the standard deviation σ_E and the mean μ_E for the data shown in (**c**). Each triangle corresponds to data from one cell, and the line is the best linear fit. Red refers to 20 μM carbachol, and blue to 50 μM carbachol. (For details of the experiments see [65])

$$\sigma = \alpha(\mu - T_r)\,, \tag{32.2}$$

a relationship that has been shown to hold true for another 8 cell types and 10 conditions (see [61] for further discussion). When the standard deviation equals zero, successive Ca^{2+} spikes are separated by a constant period. Such Ca^{2+} spike sequences appear deterministic since there is no variation in the ISI, but the interpretation is different. The lack of ISI variability results from the fact that when the Ca^{2+} spike generation probability is high, i.e. λ is large, a Ca^{2+} spike is initiated as soon as the cell exits its refractory period. Therefore, the mean of the ISI distribution at a vanishing standard deviation equals T_r. This corresponds to the intersections of the red and blue lines with the x-axis in Fig. 32.1d, respectively. To

determine ξ and λ, we start from Eq. (32.1) and derive the ISI probability density $f(t, s)$, i.e. the probability density for Ca^{2+} spikes to occur at times t and s. This is equivalent to the probability of a Ca^{2+} spike at t given that the last spike occurred at s and no Ca^{2+} spike during the time $(s - t)$. Based on this interpretation of the ISI probability, we obtain

$$f(t, s) = q(t|s) \exp\left\{-\int_s^t q(u|s)du\right\} , \qquad (32.3)$$

where the exponential term corresponds to the absence of Ca^{2+} spikes between s and t. The mean μ and the standard deviation σ of the ISI distribution then follow from Eq. (32.3) as

$$\mu = \int_0^\infty tf(t, 0)dt , \qquad \sigma^2 = \int_0^\infty t^2 f(t, 0)dt - \mu^2 . \qquad (32.4)$$

For practical purposes, we can set $T_r = 0$ in the computation of μ and σ, since a constant T_r only shifts the mean and does not affect the standard deviation. To put it another way, we evaluate Eq. (32.4) for $T_r = 0$ and then add T_r to obtain the mean ISI μ. Next, we fit the equations in (32.4) to data such as shown in Fig. 32.1d to obtain cell specific values for ξ and λ. This is achieved in a two-step process. Firstly, we determine the experimental mean μ_E and standard deviation σ_E from individual Ca^{2+} spike sequences as shown in Fig. 32.1c. This gives one data point in Fig. 32.1d. Since μ and σ^2 in Eq. (32.4) depend on ξ and λ through $f(t, 0)$ via $q(t|s)$, we can perform a least square fit of Eq. (32.4) to the experimental data μ_E and σ_E to obtain single cell estimates for ξ and λ. Figure 32.2a, b display results for HEK293 cells stimulated with $30 \mu M$ carbachol. While the distribution for ξ exhibits a localised peak, the distribution of λ is much broader. A similar behaviour is observed for spontaneously spiking astrocytes as seen in Fig. 32.2c, d. A comparison of the Ca^{2+} spike rate λ reveals that it is almost an order of magnitude larger for HEK293 cells than for astrocytes, which might be attributed to the fact that the former is stimulated, but the latter is not. Intriguingly, the time scale for recovery ξ is almost 10-fold larger for HEK293 cells than for astrocytes, indicating that HEK293 cells recover more slowly than astrocytes after a Ca^{2+} spike. The existence of wide distributions for ξ and λ also points towards significant cell-to-cell variability, which provides another argument in favour of a statistical description of Ca^{2+} spikes.

It is now instructive to evaluate Eq. (32.4) for a constant conditional intensity function $q = r > 0$, which corresponds to a homogenous Poisson process. It emerges from the general form of the conditional intensity function in Eq. (32.1) in the limit of fast recovery, i.e. a large value of ξ. In this case, the integrals can be computed analytically and we obtain $\mu = \sigma = r$, which is consistent with the scaling in Eq. (32.2). This provides further intuition for the statement made above that stochastic effects need to be taken into account when the mean and the standard deviation are of similar magnitude.
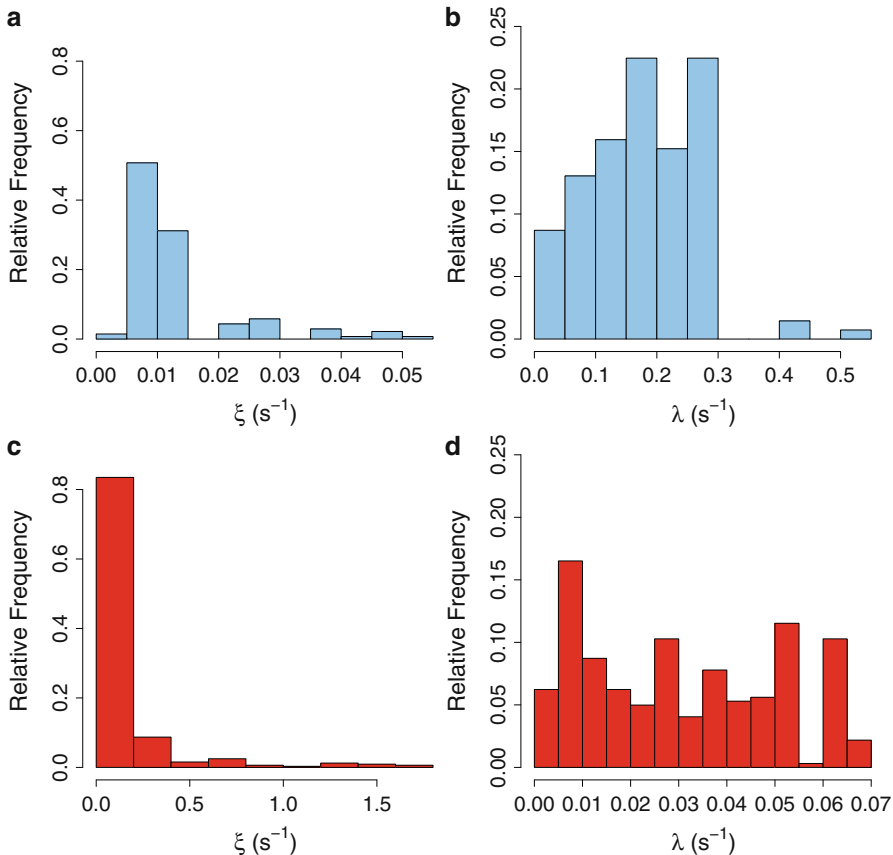
Fig. 32.2 Relative frequency for ξ (**a**, **c**) and λ (**b**, **d**) for HEK293 cells (top, blue) and astrocytes (bottom, red). HEK293 cells were stimulated with $30\,\mu$M carbachol, while Ca^{2+} spikes in astrocytes were spontaneous. $N = 138$ for HEK293 cells and $N = 321$ astrocytes. For experimental details, see [62]

Equation (32.3) expresses the ISI distribution in terms of the conditional intensity function. It is often convenient to reverse the approach and start from an ISI distribution. Firstly, we obtain ISIs directly from experimental recordings, which inform us about the possible shapes of ISI distributions. Secondly, some ISI distributions that have been shown to capture experimental data cannot be derived from closed form intensity functions as e.g. in Eq. (32.1). A point in case is the Gamma distribution, which is consistent with Ca^{2+} oscillations in HEK293 cells [67] and also with voltage spikes in neurons [68, 69]. One common representation for the density of the Gamma distribution reads as

$$f_G(t, s) = \frac{\beta^\alpha}{\Gamma(\alpha)} (t - s)^{\alpha-1} e^{-\beta(t-s)} , \tag{32.5}$$

where α and β are called the shape parameter and rate, respectively, and Γ denotes the standard Gamma function. Suppose for a moment that the time between successive Ca^{2+} puffs follows a Poisson distribution with rate β. In contrast to Ca^{2+} spikes, Ca^{2+} puffs correspond to localised Ca^{2+} liberation through a cluster of InsP$_3$Rs. In addition to Ca^{2+} release through single InsP$_3$Rs, Ca^{2+} puffs are considered the basic building blocks in the hierarchy of Ca^{2+} signals [2, 61, 70]. A Gamma distribution where α is a positive integer returns the probability that α Ca^{2+} puffs have occurred for the first time. In other words, the Gamma distribution is a probability distribution for a combination of events to happen for the first time. This interpretation makes it an appealing candidate for Ca^{2+} spikes. The reason is that Ca^{2+} spikes are thought to form when a small number of Ca^{2+} puffs generates a region of elevated Ca^{2+} in the cell, which then initiates Ca^{2+} release throughout the cell. Alternatively, recent experiments in astrocytes suggest that the co-occurrence of a certain number of Ca^{2+} puffs is sufficient to trigger a Ca^{2+} spike [71]. This also fits well with a body of research that shows that Ca^{2+} puffs and Ca^{2+} spikes can be described as first-passage time problems [63, 72–76]. As an interesting observation, note that the mean ISI for Eq. (32.5) is α/β, so that the mean interpuff interval for α puffs is $1/\beta$, which is consistent with the mean interpuff time when puffs are described by a Poisson process with rate β. To relate a given ISI distribution to the conditional intensity function, we find that

$$q(t|s) = \frac{f(t, s)}{1 - \int_s^t f(u, s)\mathrm{d}u} , \tag{32.6}$$

which is equivalent to Eq. (32.3) as shown in Appendix 1. Equations (32.3) and (32.6) allow us to switch between conditional intensity functions and ISI distributions depending on what our modelling question requires.

32.3 Beyond Stationary Ca^{2+} Spike Sequences

The discussion so far assumed that successive Ca^{2+} spikes are independent and are described by the same statistics. The conditional intensity function $q(t|s)$ only depends on the time since the last spike $(t - s)$, but not on the absolute Ca^{2+} spike times t and s. Hence, the probability for two spikes to be separated by say 80s is the same irrespective of whether the first spike occurs 10s into the experiment or 1000s. The same holds true for the ISI density in Eq. (32.5) which only depends on the time difference $(t - s)$ between successive Ca^{2+} spikes. A consequence of the independence of Ca^{2+} spikes is that we can immediately write down the probability density for n Ca^{2+} spikes occurring at times $y_1, y_2, \ldots, y_n$. If we collect the Ca^{2+} spike times in a set $\mathbf{y} = \{y_1, \ldots y_n\}$ the probability density for the entire Ca^{2+} spike sequence is given by

$$p(\mathbf{y}) = f_1(y_1, 0) f(y_2, y_1) \cdots f(y_n, y_{n-1}) f_n(T, y_n) , \tag{32.7}$$

where $f_1(y_1, 0)$ denotes the probability density for the first spike to occur at y_1 and $f_n(T, y_n)$ is the probability that no spike happens after y_n until the end of the experiment at time T. The probability for a Ca^{2+} spike sequence factorises in the probabilities of individual and identical ISIs, which are properties often referred to as independence and stationarity, respectively. We separate out the contributions from f_1 and f_n since they do not correspond to ISI probabilities and hence are often modelled by different probability distributions, e.g. a Poisson distribution.

However, there are numerous reasons for why ISI probabilities do not remain constant over time and hence ISIs at different times of the experiment follow different probability distributions. For example, while the ER refills between Ca^{2+} spikes, the level of refilling can decrease as Ca^{2+} leaves the cell across the plasma membrane. In most experiments, $InsP_3$ is formed in response to activation of cell surface receptors, but the efficiency of this pathway may decrease over time. Both factors lower the propensity for the generation of Ca^{2+} spikes as the experiment progresses and introduces trends when plotting ISIs. When analysing Ca^{2+} spikes, we can remove trends and only consider Ca^{2+} spikes after initial transients. This presents a sensible approach when cells experience constant stimulation such as in step change experiments. However, under physiological conditions, hormones arrive in a time-dependent manner, so do neurotransmitters and paracrine signals. To mimic such an in vivo environment, cells need to be challenged with time-varying stimuli. As soon as we introduce an explicit time-dependence, ISI distributions are no longer stationary, but depend on the absolute time of the experiment.

This raises the question on how to mathematically describe the non-stationarity of Ca^{2+} spike sequences. One approach is to introduce an explicit time-dependence into the ISI distribution by making the parameters change over time. While conceptually appealing, the practicalities of this approach are limited. For instance, if we believe that the parameters change continuously over time, it is not apparent how to constrain the model best given that we sample the values of the parameters at only a few discrete time points, viz. the times of Ca^{2+} spikes. Another issue arises from the fact that the probability of a Ca^{2+} spike sequence does not necessarily factorise any more as in Eq. (32.7), but we need to consider the full multivariate probability $p(\mathbf{y}) = p(y_1, \ldots, y_n)$, which can pose significant challenges.

A more practical approach was put forward in [68]. At the heart of it lies a time transformation that maps the time of the original experiment, denoted by t, to a new time u via

$$u(t) = \int_0^t x(v)\mathrm{d}v \,, \tag{32.8}$$

where x is called the intensity function and relates to the level of Ca^{2+} spiking as we will illustrate below. As such, x is always strictly positive and hence associates each value of t with a unique values of u through Eq. (32.8). A consequence of this

mapping is that in the new time u, ISIs become independent [68]. This means that the probability density for a Ca^{2+} spike sequence factorises again and we have

$$p(\mathbf{y}|x) = g_1(u_1, 0|x)g(u_2, u_1|x) \cdots g(u_n, u_{n-1}|x)g_n(U, u_n|x), \qquad (32.9)$$

where $u_i = u(y_i)$, $U = u(T)$ and the dependence of $\mathbf{y}$ on the left hand side enters on the right hand side through u being a function of t. We explicitly include x to emphasise that the transformation depends on the intensity function. What makes Eq. (32.9) particularly useful is that the probability density g is related to the original ISI probability density f via

$$g(u_i, u_{i-1}|x) = x(y_i)f(u_i, u_{i-1}), \qquad (32.10)$$

which follows from the conservation of probability [67, 77]. We illustrate a practical calculation for Eq. (32.10) in Appendix 2.

At this point, it might appear that the intensity function is mathematically convenient, but detached from the actual biology. As it turns out, the contrary holds true. For the models of Ca^{2+} spiking considered here, $x(t)$ corresponds to the probability of Ca^{2+} spiking *independent* of the history of the Ca^{2+} spike sequence. Put differently, if there are N identical Ca^{2+} spiking cells, $Nx(t)$ is the expected number of Ca^{2+} spikes at time t. To illustrate this concept, we chose an intensity function (red line in Fig. 32.3), generated 10,000 Ca^{2+} spike sequences from it and binned them (light blue histogram). By using a large number of Ca^{2+} spike sequences, binning is equivalent to taking the average across all possible histories that led to a Ca^{2+} spike in the respective bin. The excellent agreement between the intensity function and the histogram confirms the above interpretation of $x(t)$. For the practicalities of generating the Ca^{2+} spike sequences, we refer the reader to Appendix 3.

32.4 Bayesian Computation

A main motivation for pursuing a statistical approach is to fit models of Ca^{2+} spike generation more easily to experimental data and hence learn more about the nature of Ca^{2+} oscillations. This requires us to infer the parameters of the ISI distribution, e.g. α and β for the Gamma distribution, and the time course of $x(t)$ from measured Ca^{2+} spike sequences. For ease of reference, we will call all unknowns of the model, i.e. the intensity function and the parameters of the ISI distribution, hyperparameters and denote them by θ.

There are a number of ways for achieving this goal. Here, we will make use of Bayesian inference, that addresses the following question: what does the data tell us about the parameters of the model? Expressed more formally, we are interested in $p(\theta|\mathbf{y})$, i.e. the probability distribution of the hyperparameters given a Ca^{2+} spike sequence. This probability is called the *posterior* distribution. The advantage of this
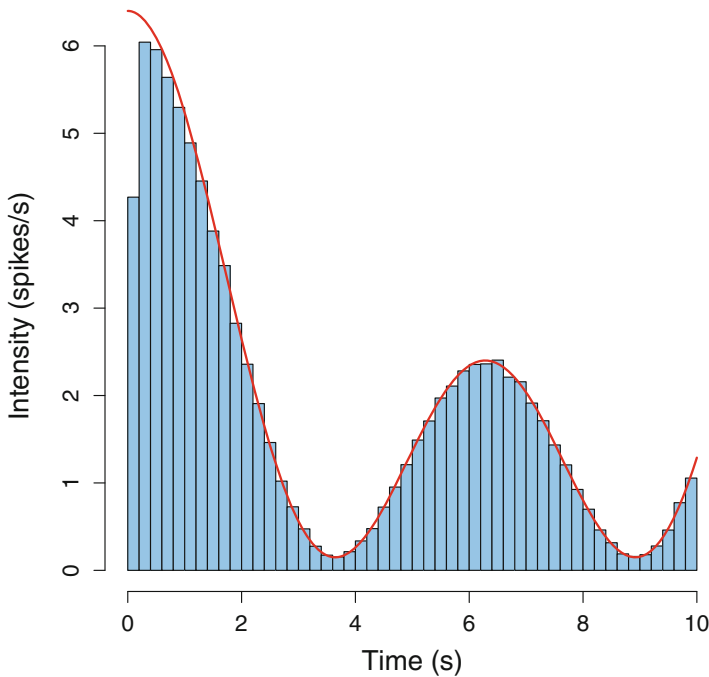
Fig. 32.3 Intensity function (red) and peristimulus-time histogram (blue) obtained from 10,000 Ca^{2+} spikes when the ISI distribution is given by a Gamma distribution. Parameter values are $x(t) = 2\cos(t) + 2\cos(0.5t) + 2.4$ and $\alpha = 6.2$ and $\beta = 6.2s$

approach is that we do not merely obtain a single value, but the full probability distribution for the parameters that are consistent with the data. This allows us to judge how well the model captures the data and what parameter values to use to describe the underlying biology. For instance, consider one of the hyperparameters, say θ_1. If the distribution for θ_1 is sharply peaked around a value θ_1^*, we can be confident that θ_1^* is a sensible estimate for θ_1. On the other hand, if the probability distribution is broad or exhibits multiple maxima, we are pressed much harder to interpret the results. It might also indicate that we based our original model on incorrect assumptions. In addition to these conceptual benefits, the posterior distribution is all we need to answer any question we have about the experiment. For instance, we can determine summary statistics such as mean and variance as well as the behaviour of functions that depend on hyperparameters.

To compute the posterior probability, we make use of Bayes' theorem, which states that

$$p(\theta|\mathbf{y}) = \frac{p(\mathbf{y}|\theta)p(\theta)}{p(\mathbf{y})} \ . \tag{32.11}$$

The right hand side contains the likelihood function $p(\mathbf{y}|\theta)$, the so-called prior $p(\theta)$ and the normalisation

$$p(\mathbf{y}) = \int p(\mathbf{y}|\theta) p(\theta) \mathrm{d}\theta . \qquad (32.12)$$

The conceptual appeal of Eq. (32.11) stems from its numerator. We already encountered an example for a likelihood function in Eq. (32.9). It represents how likely it is to observe what we have measured for a given θ and hence reflects our believes about the potential mechanisms that drive Ca^{2+} spike generation. The prior distribution allows us to provide sensible input for the hyperparameter values before we see the data. For instance, if we believe that some hyperparameter, say θ_1 again, has a value close to some θ_1^0, we pick a probability distribution that is centred around θ_1^0. On the other hand, if we are uncertain about possible values of θ_1, we choose a wider prior. Thinking about priors for hyperparameters that are numbers immediately leads us to probability distributions in the classical sense such as Poisson distributions or Gamma distributions. But what about a prior for the intensity function $x(t)$? To answer this question, it is helpful to return to the biological interpretation of $x(t)$, viz. the probability density for a Ca^{2+} spike at time t irrespective of the history of the Ca^{2+} spike sequence. If we challenge cells with a constant stimulus as in e.g. a step-change experiment, a reasonable assumption is that $x(t)$ remains constant over longer periods of time, but not necessarily at the same value for the entire experiment. For instance, as the experiment continues, Ca^{2+} spikes may become less frequent compared to the beginning of the experiment due to receptor desensitisation or changes to ER Ca^{2+} load. We can mimic this biological response by assuming an $x(t)$ that has a large constant value when the experiment is started and smaller constant value towards the end. In this particular illustration, we assumed that there are two different levels. To allow more flexibility, suppose that there are k levels and that the probability for having k levels is Poissonian with rate γ. If we further assume that each level h_i is drawn from a Gamma distribution with parameters a and b, we find that the prior for the intensity function is

$$p(x) = \mathrm{e}^{-\gamma} \frac{\gamma^k}{k!} \prod_{i=1}^{k} \frac{a^b}{\Gamma(a)} h_i^{a-1} \mathrm{e}^{-bh_i} . \qquad (32.13)$$

Because the number of changepoints is independent from the levels h_i, which again are independent from each other, $p(x)$ factorises into individual contributions [78]. In Fig. 32.4a we illustrate three candidates for such piecewise constant priors with different numbers of changepoints and different level values.

While piecewise linear intensity functions possess computational advantages— e.g. the integral in Eq. (32.22) in Appendix 2 can be computed analytically—one issue with them is that they are discontinuous, i.e. they have jumps. This might be undesirable in some situation, which leads us to priors for continuous functions. An example for this is when cells are challenged with a time-varying stimulus as e.g.
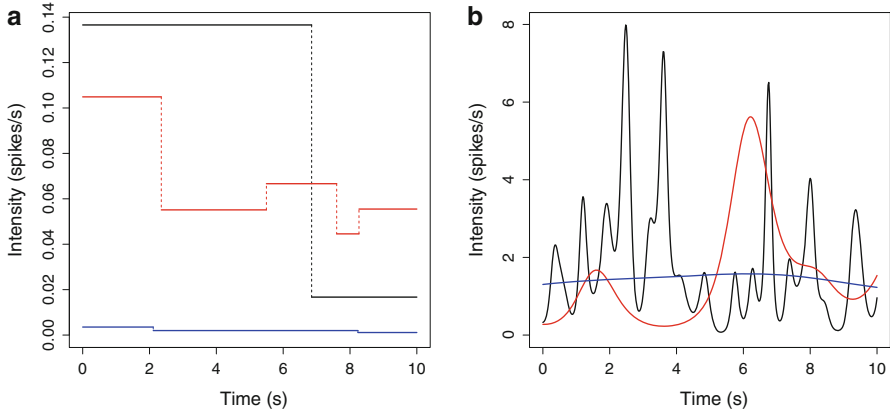
Fig. 32.4 Candidate intensity functions for (**a**) a piecewise linear prior and (**b**) a GP prior. The different colours indicate (**a**) different numbers of change points and different levels and (**b**) different values of κ. Here, blue corresponds to $\kappa = 5\,\mathrm{s}$, red to $\kappa = 1\,\mathrm{s}$ and black to $\kappa = 0.2\,\mathrm{s}$

in [67]. Since the stimulus changes smoothly over time, a reasonable assumption is that the intensity function inherits this smoothness. Among the different choices that can be made for continuous intensity functions, we here focus on so-called Gaussian processes (GPs). Consider the intensity function at some time point t. Instead of fixing a unique value $x = x(t)$, we prescribe a probability distribution $g(x)$. In other words, for a fixed time t there is a probability $g(x(t))\mathrm{d}x$ that the value of the intensity function lies in the interval $[x(t), x(t) + \mathrm{d}x]$. We here assume that the logarithm of the intensity function follows a Gaussian distribution of the form

$$f_{\mathrm{GP}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{(\mu - x)^2}{2\sigma^2} \right\}, \tag{32.14}$$

where μ denotes the mean and σ the standard deviation, respectively. GPs derive their name from the fact that we employ a Gaussian distribution. The reason for assuming that $\log x(t)$ rather than $x(t)$ itself follows a Gaussian distribution is that $x(t)$ is always positive, but a Gaussian distribution can yield negative values. By using the logarithm, we enforce the positiveness of the intensity function. To ensure that the intensity function is continuous, we need to guarantee that the values of x at two close-by time points t and s are not too far apart. This is achieved by imposing a correlation function

$$\Sigma(s, t) = \sigma^2 \exp\left\{ \frac{(s - t)^2}{\kappa^2} \right\}, \tag{32.15}$$

which we have chosen to be a squared exponential. Here, σ is the same as in Eq. (32.14) and κ measures how smooth the GP is. The larger κ the smoother

the intensity function. Figure 32.4b shows three different realisations of a GP for varying values of κ. Observe that all three functions are smooth, and that there are less wiggles for larger values of κ, which is consistent with the interpretation above. Since we have to discretise time for any practical computation, suppose that there are n time points, i.e. we represent the time of the experiments at n discrete time points ranging from $t_1 = 0$ to $t_n = T$, where T is the duration of the experiment. A practical representation for these time points are the values at which the Ca^{2+} concentration is measured and is determined by e.g. the frame rate of the microscope cameras. The prior for a GP is a multivariate Gaussian distribution and reads as

$$p(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left\{\frac{1}{2}(\mathbf{t} - \boldsymbol{\mu})\Sigma^{-1}(\mathbf{t} - \boldsymbol{\mu})\right\}, \qquad (32.16)$$

where $\mathbf{t}$ is a vector of length n representing the discretised time of the experiment, $\boldsymbol{\mu}$ is a vector of length n denoting the mean of the GP at each time point t_n, and Σ is given by Eq. (32.15).

Having introduced different priors for the hyperparameters of the model including the intensity function x, we can return to Eq. (32.11). While it is conceptually appealing and offers us a full picture of the parameters of the model as constrained by the experiment, it is computationally challenging. The reason is the integral in the denominator, which runs over the entire hyperparameter space. Since this can be high-dimensional, we require computationally efficient methods as a direct integration is often prohibitively expensive if not impossible. There are various methods for tackling this problem. For instance, instead of computing $p(\theta|\mathbf{y})$ directly including the integration of the denominator, we can determine the maximum of the distribution and its variance [79, 80]. This will provide us e.g. with the most likely intensity function that is consistent with the data as well as confidence intervals, see [67]. A different approach is to try and sample from $p(\theta|\mathbf{y})$ without having to explicitly compute it. The main idea is that if we can sample from a probability distribution, we know the possible values and the associated probabilities (values that are more likely than others are sampled more frequently) without having to determine a closed form solution. This often suffices for practical purposes. A technique that does this is known as Markov chain Monte Carlo [81].

32.5 Analysing Ca^{2+} Spike Sequences

Having introduced key concepts for a statistical analysis of Ca^{2+} spike sequences in the previous sections, we now apply them to different experiments. A crucial input to our model is the ISI probability density, see Eq. (32.9). However, we do not know *a priori* which distribution is most consistent with the data. To establish this, we can make use of the following transformation. Let the measured Ca^{2+} spike times be given by $y_1, \ldots y_n$ again. We define transformed ISIs by
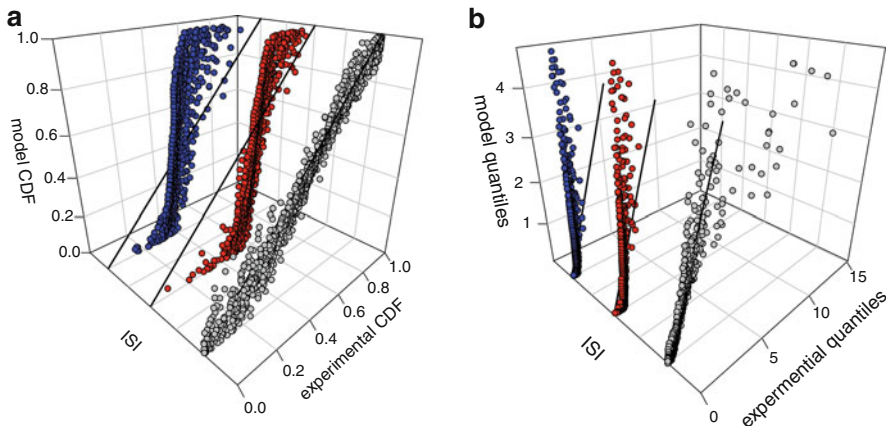
Fig. 32.5 K-S (**a**) and Q-Q (**b**) plots for data from 23 individual cells stimulated with carbachol cells. The initial concentration was $20\,\mu$M, which was increased to $50\,\mu$M. The ISI distributions are inverse Gaussian (blue), Poisson (red) and Gamma (grey). We used piecewise linear functions as prior for the intensity function

$$\tau_k = \int_{y_{k-1}}^{y_k} q(s|y_{k-1})\mathrm{d}s\,, \tag{32.17}$$

with q given by Eq. (32.6). It can now be shown that if the mechanisms that generate the observed Ca^{2+} spikes are consistent with the ISI distribution that we use in Eq. (32.6), the transformed ISIs τ_k are exponentially distributed with unit rate [82]. This leaves us with the task of comparing two probability distributions: the standard exponential distribution and the distribution of the transformed ISIs. The quantile-quantile (Q-Q) plot and the Kolmogorov Smirnov (K-S) plot are two powerful graphical approaches to examine differences between probability distributions. In Fig. 32.5 we show Q-Q and K-S plots for HEK293 cells in a multistep experiment. We tested three different ISI distributions: a Poisson, an inverse Gaussian and a Gamma distribution. Each cell was analysed individually and gave rise to a separate sequence of dots; no data assimilation was performed. For the Q-Q plot, we determine the quantiles of the transformed ISIs and plot them against the quantiles of the exponential distributions. For the K-S plot, we compute the cumulative distributions of the transformed ISIs and the exponential distribution, respectively, and plot them against each other. Identical probability distributions possess identical quantiles and identical cumulative distributions, respectively. Hence any deviation from a 45° straight line in the Q-Q and K-S plot points towards differences between the distributions and hence indicates that we need to improve our assumptions about the ISI distributions.

For both, the K-S and the Q-Q plot, the data points deviate significantly from a straight line with slope 1 for the Poisson and the inverse Gaussian distribution. On the other hand, we observe a strong correlation between the 45° line and the data points for the Gamma distribution. This visual inspection is corroborated by
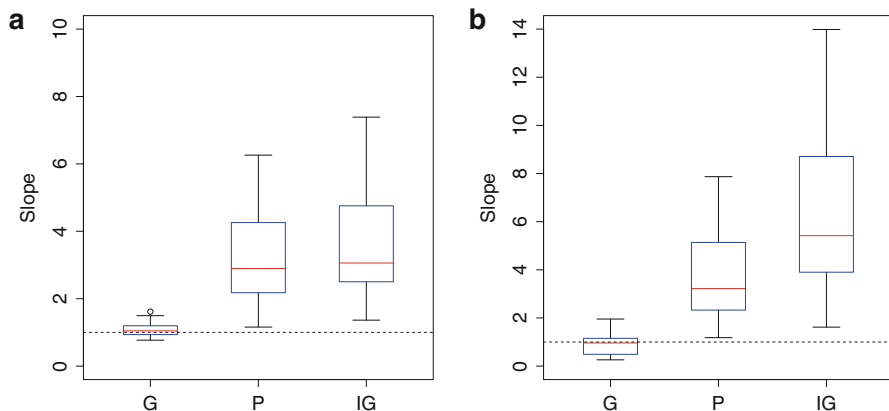
Fig. 32.6 Box and whisker plots for the data presented in Fig. 32.5 showing results for the (**a**) K-S plot and (**b**) Q-Q plot. The red line represents the median of the distribution, and the box delineates the range of data from the first to the third quartile. The whiskers indicate the spread of the data

the box-and-whisker plots in Fig. 32.6. Because we treated cells individually, we can determine the slope of a linear fit for each cell. The plots in Fig. 32.6 show the statistics for these slopes. The box represents the spread of data within the second and third quartile, and the red line indicates the median. The whiskers provide a measure for the overall spread of the data. Generally speaking, the smaller the box and the closer the whiskers to the box, the less spread is in the data. The Poisson and the inverse Gaussian distribution generally exhibit a larger spread than the Gamma distribution. Moreover, the median of the Gamma distribution is closer to one. In [67], we found that for HEK293 cells stimulated with $10\,\mu$M and $100\,\mu$M charbachol, respectively, the Gamma distribution worked best. These results and the new analysis presented here strongly suggest that the ISI statistics for Ca^{2+} spike sequences are captured by a Gamma distribution. A further argument to support this conclusion is that the data in [67] and [65] were acquired independently in different labs with different setups.

In order to obtain the results in Figs. 32.5 and 32.6 we had to estimate the intensity function $x(t)$ for each cell. Figure 32.7 displays $x(t)$ for two different cells. Since we analyse step change experiments, we first chose piecewise constant functions as a prior for x as discussed in Sect. 32.4. The mean intensity function is shown as a solid blue line, while the 95% confidence interval is represented by the shaded blue area. We clearly see an increase in the intensity function as the stimulus strength is stepped up. Moreover, during extended periods of time, the intensity function is almost constant, which has significant consequences for the interpretation of the mechanisms that drive Ca^{2+} spike generation as discussed below.

As pointed out in Sect. 32.4, piecewise constant functions are not the only possible prior. GPs constitute another possible class, and corresponding results are
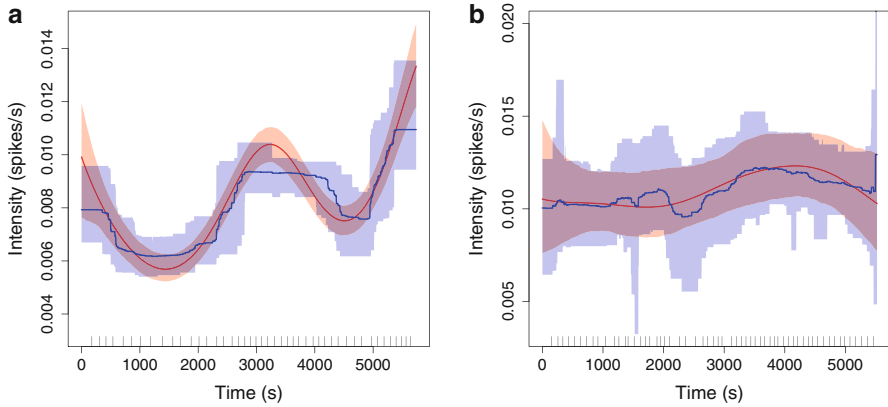
Fig. 32.7 Intensity functions (solid lines) and 95% confidence interval (shaded regions) for two cells in a multistep experiment. The initial stimulus was $20\,\mu$M carbachol and changed to $50\,\mu$M carbachol at $t = 2581$s (**a**) and $t = 2524$s (**b**). The prior for the intensity function is a GP (red) or piecewise constant (blue). The ticks along the x-axes indicate the Ca^{2+} spike times

shown in red in Fig. 32.7. Vitally, the intensity function obtained with a GP prior closely follows that derived from a piecewise constant prior. Given that the two priors represent significantly different functional forms of the intensity function, the consistency between the two approaches lends strong support for the validity of the derived intensity functions. Moreover, if we were to only use GPs as priors, a valid step is to interpolate the smooth prior with a piecewise linear function, which allows us to compute the mean ISI.

Having identified intensity functions that are consistent with measured Ca^{2+} spike sequences, we can now determine the conditional intensity functions $q(t|s) = q(y_i|y_{i-1})$. We start from Eq. (32.6), set $t = y_i$, $s = y_{i-1}$ and then replace $f(y_i, y_{i-1})$ with $g(u_i, u_{i-1}|x)$ from Eq. (32.10) by using Eq. (32.8) (see also Appendix 2). In other words, the conditional intensity function $q(t|s)$ is a highly nonlinear transformation of the estimated intensity function $x(t)$ given the Ca^{2+} spike times. In Fig. 32.8 we plot $q(t|s)$ for the data shown in Fig. 32.7. We notice that immediately after a Ca^{2+} spike $q(t|s)$ remains almost zero for some time before it increases. This indicates that during this period, Ca^{2+} spikes cannot occur, which is equivalent to saying that there is a refractory period. Importantly, we did not include an explicit refractory period in our model, i.e. we did not choose an ISI probability distribution that vanishes for a certain amount after the last Ca^{2+} spike. For instance, the Gamma distribution in Eq. (32.5) does not *per se* stay close to zero for small values of $(t - s)$. It only does so for certain values of α. Since the value of α is part of estimating $x(t)$, the vanishing of the conditional intensity function is an emergent result of the model. These findings are consistent with the presence of a refractory period T_r in the ansatz in Eq. (32.1). There, we chose the conditional intensity function and derived the ISI distribution, while for Fig. 32.8, we decided upon a certain ISI distribution and derived the conditional intensity function. The
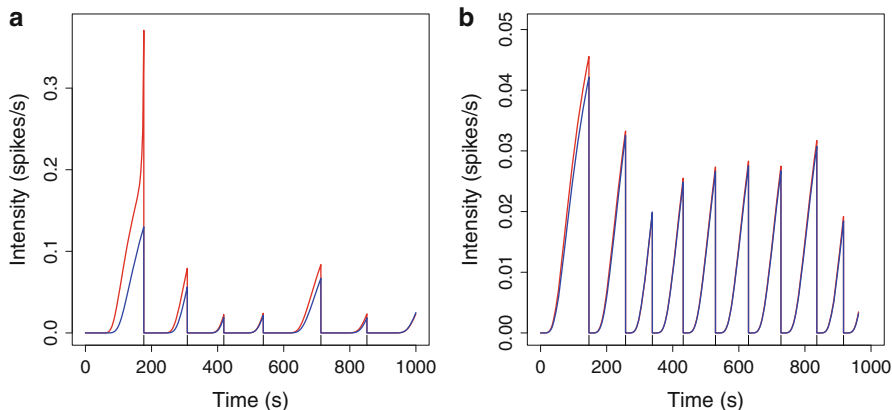
Fig. 32.8 Conditional intensity functions corresponding to the data in Fig. 32.7. The ticks along the x-axes indicate the Ca^{2+} spike times

agreement between the ansatz in Eq. (32.1) and the estimated conditional intensity function in Fig. 32.8 lends strong support to the former.

In addition to the conditional intensity function, we can also interrogate the ISI distribution. For a time-dependent intensity function $x(t)$, the corresponding ISI distribution is time-dependent as well, see Eqs. (32.21) and (32.22) in Appendix 2. However, when the intensity function is constant, this time-dependence is lost, and we can use the *same* ISI distribution for the entire period that x does not change. Inspecting Fig. 32.7a, we observe that the intensity function obtained with a PWL prior (blue line) is almost constant between 600s and 2000s, while a similar behaviour is seen in Fig. 32.7b during the first 2000s for the GP prior (red curve). Taking these values for x together with the inferred parameter values for the Gamma distribution, we now plot the corresponding ISI distributions in Fig. 32.9 based on Eq. (32.25). For the stronger stimulus ($50\,\mu M$, blue line), the ISI distribution is shifted towards the left compared to the weaker stimulus ($20\,\mu M$, red curve). In addition, the variance is more pronounced in the former than in the latter. To quantify this, we compute the mean μ and the standard deviation σ using the inferred intensity function x and the associated values of the Gamma distribution. We obtain $\mu = 159.07s$ and $\sigma = 22.50s$ for the small stimulus and $\mu = 96.73s$ and $\sigma = 32.24s$ for the stronger stimulus, respectively. We can compare this with the mean and standard deviation determined directly from the Ca^{2+} spike sequences shown in Fig. 32.7. We find $\mu_E = 166.57s$ and $\sigma_E = 24.416s$ at $20\,\mu M$ and $\mu_E = 96.31s$ and $\sigma_E = 27.03s$ at $50\,\mu M$, respectively. The good agreement between the experimentally determined statistics (μ_E, σ_E) and the estimated quantities (μ, σ) demonstrates the usefulness of the Bayesian inference approach that we have employed here.

Fig. 32.9 ISI probability density $f_G(t, 0|x)$ for the data shown in Fig. 32.7a (red) and Fig. 32.7b (blue) for t between approximately 600s and 2000s (red) and t between 1s and 2000s (blue)
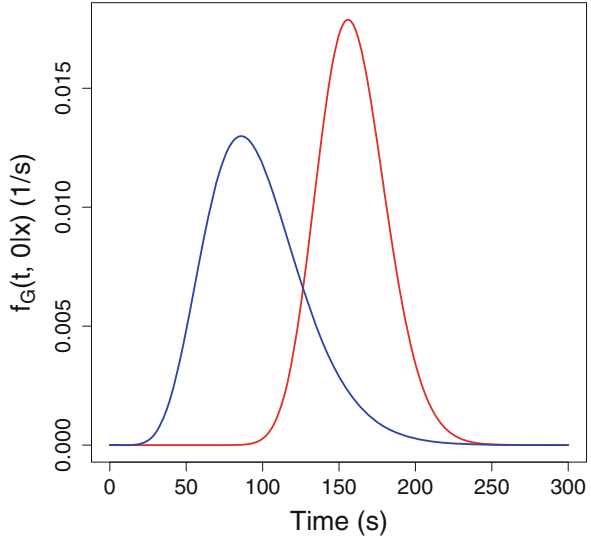
32.6   Concluding Discussion

Ca^{2+} spikes constitute a well-established mode of intracellular Ca^{2+} signalling across a large number of cell types. We can now draw on a substantial body of experimental measurements that have identified and characterised the cellular components that drive Ca^{2+} oscillations. Despite these successes, central questions remain open. Amongst them is a seemingly innocuous query: given a stimulus, can we predict the sequence of Ca^{2+} spikes? Since there is wide-ranging consensus that information about the stimulus is encoded in the properties of Ca^{2+} spike sequences, answering this question is critical for our understanding of intracellular Ca^{2+} signalling.

Addressing this issue from a modelling perspective is challenging for two reasons. The generation of Ca^{2+} spikes is firstly stochastic and secondly driven by the interaction of spatially distributed clusters of $InsP_3$ receptors. One avenue to make progress is to simulate partial differential equations for the intracellular calcium concentration (see [61] for a recent perspective). Here, we have reviewed a different framework that is a conceptual antipode to the first approach. While partial differential equations rely on mechanistic details and build oscillations from the bottom-up, the statistical ideas in this review aim at describing Ca^{2+} spikes directly at the cell level.

One advantage of a statistical model lies in its computational demands. It is considerably cheaper to generate Ca^{2+} spike sequences from a statistical model than to solve partial differential equations. This is particularly useful when studying cell populations, where intercellular heterogeneity calls for multiple Ca^{2+} spike sequences with different parameter values. But statistical models may also help

conceptually. Ultimately, Ca^{2+} dependent signalling is driven in many instances by the sequence of Ca^{2+} spikes. Hence, the properties of Ca^{2+} spikes such as their ISI distribution are of central interest. Since it is conceivable that different microscopic models based on detailed molecular processes all lead to the same cellular behaviour, statistical frameworks are ideally placed to capture this common identity of Ca^{2+} spiking.

The first step in our statistical analysis is to derive a model for whole-cell Ca^{2+} spiking. Since Ca^{2+} spikes are stochastic, we can express their occurrence most naturally in the language of probabilities. A core ingredient is the ISI distribution $f(t, s)$, or equivalently the conditional intensity function $q(t|s)$. It is worth noting that both depend on only two times, i.e. we assume that the generation of a Ca^{2+} spike only depends on the history since the last Ca^{2+} spike. This independence of successive Ca^{2+} spikes has been shown for astrocytes, PLA cells and HEK293 cells in [83] for constant stimulation. In general, however, this might be too strong an assumption. In particular, when cells are challenged with continuously changing stimuli—in order to mimic a more realistic cellular environment—correlations within the signal might be inherited by the Ca^{2+} spikes. One possibility is to generalise the conditional intensity function. Instead of depending only on the last time t, it now relies on the entire Ca^{2+} spike history H_t, i.e. $q = q(t|H_t)$. This, however, does not necessarily lead to a mathematically tractable problem. A more practical approach is the introduction of an intensity function. Essentially, it transforms the original Ca^{2+} spike times in such a way that they become independent. As a consequence, we can use the original ISI distributions $f(t, s)$ or conditional intensity functions $q(t|s)$, which entails that the parameters of the model are those of the ISI distribution and the intensity function, respectively.

This leaves us with the task of finding the parameter values given the Ca^{2+} spike sequences. The last condition is of particular relevance to the current approach. Our goal is to derive a model that is constrained by experimental data and whose parameter values can be sensibly estimated. We have achieved this by employing Bayesian ideas, which allow us to determine the probability distribution of the parameters given the Ca^{2+} spike sequences, i.e. $p(\theta|\mathbf{y})$. This is a distinct advantage of the Bayesian framework. Other methods, such as maximum likelihood estimators, also provide information about the parameters of the model. However, they only deliver *one* set of parameter values associated with a standard error, not entire probability distributions. Moreover, these approaches come with numerical challenges and are hard to pursue in higher dimensions.

Since the ISI distribution is a core component of the model, we first ascertained if our choices are consistent with the recorded Ca^{2+} spike sequences. As Figs. 32.5 and 32.6 illustrate, the Gamma distribution captures the data well, while the inverse Gaussian and the Poisson distribution fail to do so. It is worth noting that the data analysed here were obtained in different experiments than those used in [67], yet both data sets lead to the same conclusion: Ca^{2+} spikes are well described by a Gamma distribution. This might point to the mechanisms that generate Ca^{2+} spikes. Since the Gamma distribution returns the probability of the first time that

α events have occurred (see Eq. (32.5)), it is consistent with the interpretation that the formation of a critical nucleus of elevated Ca^{2+}, driven by the occurrence of a certain number of Ca^{2+} puffs, underlies the generation of a Ca^{2+} spike. At the moment, we cannot rule out that other probability distributions that we have not tested yet describe Ca^{2+} spikes equally well or even better. The advantage of our Bayesian modelling framework is that it works for *any* probability distribution, which allows us to test more candidate distributions in the future. Moreover, we have two complementary tests at our disposal, the Q-Q and the K-S plot. While both approaches check whether two probability distributions coincide, the K-S plot is more sensitive towards the centre of the distribution, while the Q-Q plot focusses on the tails.

When testing for the most likely ISI distribution, we had to estimate the intensity function $x(t)$ at the same time, since the ISI distribution explicitly depends on $x(t)$ (see Eqs. (32.21) and (32.22)). Following on from our results so far, we focussed on intensity functions obtained for the Gamma distribution. The intensity function is central to our understanding of Ca^{2+} spike generation. An almost constant intensity function indicates that Ca^{2+} spikes originate from *stationary* dynamics. This means that the ISI distribution is identical for each recorded Ca^{2+} spike time, which allows us to compute the mean and the variance of a Ca^{2+} spike sequence (see Fig. 32.9). From a biological perspective, this corresponds to a cell with no explicitly time-dependent processes such as a continuous depletion of the ER or an accumulating degree of receptor desensitisation. In the latter, this does not mean that receptor desensitisation does not occur, but that the fraction of desensitised receptors across the cell stays constant. A change of experimental conditions is directly translated into changes of the intensity functions. For instance, recent experiments with sinusoidal stimuli led to intensity functions that reflect the rises and falls of the applied agonist [67]. Moreover, since intensity functions are estimated from Ca^{2+} spike sequences of individual cells, they mirror the variability of observed responses. A key line of research is therefore to quantify and classify such diverse intensity functions.

Using both the ISI distribution and the intensity function we can compute the conditional intensity function $q(t|s)$, which corresponds to the probability of a Ca^{2+} spike at time t given that a Ca^{2+} spike occurred at time s. The shape of q allows us to discuss potential mechanisms that are involved in Ca^{2+} spike generation. For instance, as Fig. 32.8 illustrates, the conditional intensity function remains small after a Ca^{2+} spike before it increases. This period of low Ca^{2+} spiking probability is consistent with the observations of refractoriness. Plots like Fig. 32.8 also allow us to estimate the range of refractory periods, which we can then compare to the refractory period obtained from plots of the mean ISI against the ISI standard deviation as seen in Fig. 32.1d. In addition, we can compare the rise time of the conditional probability function with known timescales of e.g. ER refilling or $InsP_3$ receptor recovery to ascertain whether any of the molecular timescales match the cellular time scale, or whether we deal with an emergent timescale.

One motivation for pursuing a statistical approach is to obtain *distributions* for the parameter values that govern Ca^{2+} spike generation. The reason for why

distributions exist in the first place—and not a single parameter value only—lies in the inherent single-cell variability. Consider e.g. the variation of ξ and λ shown in Fig. 32.2. The recovery from global cellular inhibition is controlled by ξ. This involves inter alia resequestration of Ca^{2+} from the cytosol to the ER via SERCA pumps or recovery of $InsP_3$ levels. Since expression levels of SERCA pumps can vary amongst cells, recovery proceeds at different speed in different cells, which is captured by the variability of ξ. As for λ it controls the asymptotic Ca^{2+} spike rate. As Ca^{2+} spikes are believed to occur via the formation of a critical nucleus of elevated Ca^{2+} and subsequent propagation of a Ca^{2+} wave, the spatial distributions of $InsP_3Rs$ and SERCA pumps are crucial. These distributions vary significantly between cells, which directly impacts on the spread of λ.

As with all modelling approaches, the methodology presented here is not without its caveats. In its current form, we only consider Ca^{2+} spike times and leave aside other Ca^{2+} spike properties such as amplitude, duration, shape, or baseline Ca^{2+} concentration levels. However, these characteristics have been shown to impact on a number of Ca^{2+} dependent processes [4]. An interesting point in this respect is a potential interplay between release amplitude, release duration and the absolute refractory period. Our results in [65] suggest that the absolute refractory period is not controlled by cell variability and hence that there is only one value for all cells. We can further test this hypothesis by extending the model for a Ca^{2+} spike as in Eq. (32.25) in Appendix 3 to explicitly include a distributed refractory period. The advantage of the Bayesian approach is that the estimation process remains conceptually the same, but we need to estimate additional parameters. As stated above, one incentive for the current work is to relate the estimated parameter values to biophysical processes. Care needs to be taken here as different processes can potentially give rise to the same whole cell parameter values that we infer. Hence further tests are needed to discriminate between different alternatives. A consequence of this consideration is that cells might employ a number of different strategies to generate the same whole cell signal, and it will be fascinating to tease apart the advantages and disadvantages of specific routes to global Ca^{2+} signals.

The preceding discussion illustrates that advanced statistical modelling can provide valuable insights into the dynamics of Ca^{2+} spiking. Vitally, our approach works for cells that are dynamically stimulated with agonist time courses that mimic physiological conditions in vivo, which allows us to model cellular Ca^{2+} spiking in a realistic environment. By inferring parameter values from single cell measurements, we can determine their ISI distribution, which is a central ingredient to modelling Ca^{2+} spikes. Moreover, it allows us to compute statistical properties such as means and variances, which in turn quantify stochastic Ca^{2+} spikes. In addition, we showed how the statistical model allows us to infer potential mechanisms of Ca^{2+} spike generation. This connects the statistical approach with the mechanistic framework of simulating partial differential equations for Ca^{2+} signalling. In the future, it will be desirable to see these two complementary techniques working hand in hand, which has the potential to significantly enhance our understanding of the Ca^{2+} signalling toolkit.

Appendix 1

We here show the equivalence of Eqs. (32.3) and (32.6). For this, it is convenient to introduce

$$F(t, s) = 1 - \int_s^t f(u, s)\mathrm{d}u .$$
(32.18)

The right hand side of Eq. (32.6) can be written as a full derivative in the form

$$q(t|s) = -\frac{\mathrm{d}}{\mathrm{d}t} \ln F(t, s)$$
(32.19)

Multiplying through by (-1) and integrating both sides with respect to t yields

$$-\int_s^t q(u|s)\mathrm{d}u = \ln F(t, s) ,$$
(32.20)

noting that $\ln F(s, s) = 0$. When we exponentiate both sides and use the fact that $F(t, s) = f(t, s)/q(t|s)$ as per Eq. (32.6) we arrive at Eq. (32.3).

Appendix 2

Here, we demonstrate how to practically apply Eq. (32.10) when f is given by the density for the Gamma distribution as in Eq. (32.5). We obtain

$$g_G(u_i, u_{i-1}|x) = x(y_i)\frac{\beta^\alpha}{\Gamma(\alpha)} X_{i,i-1}^{\alpha-1}\mathrm{e}^{-\beta X_{i,i-1}} ,$$
(32.21)

where

$$X_{i,i-1} = \int_{y_{i-1}}^{y_i} x(v)\mathrm{d}v ,$$
(32.22)

since the difference $(t - s)$ in the transformed time u is given by

$$u(t) - u(s) = u(y_i) - u(y_{i-1}) = \int_0^{y_i} x(v)\mathrm{d}v - \int_0^{y_{i-1}} x(v)\mathrm{d}v = \int_{y_{i-1}}^{y_i} x(v)\mathrm{d}v .$$
(32.23)

A common choice for f_1 and f_n is a Poisson distribution, which leads to

$$g_1(u_1, 0|x) = x(y_1)\mathrm{e}^{-X_{1,0}} , \qquad g_n(U, u_n|x) = \mathrm{e}^{-X_{n+1,n}} ,$$
(32.24)

in the transformed time, where we have set $y_0 = 0$ and $y_{n+1} = T$.

Appendix 3

To generate the Ca^{2+} spikes that underlie the histogram in Fig. 32.3 we use inverse sampling [84]. Since $x(t)$ is non-constant, we need to use the time-dependent ISI density $g(u_i, u_{i-1}|x)$ from Eq. (32.10). For ease of presentation, we rewrite Eq. (32.10) in terms of the non-transformed Ca^{2+} spike times y_i using the results from Appendix 2 as

$$f_G(y_i, y_{i-1}|x) = x(y_i)\frac{\beta^\alpha}{\Gamma(\alpha)}\left[\int_{y_{i-1}}^{y_i} x(v)\mathrm{d}v\right]^{\alpha-1}\exp\left\{-\beta\int_{y_{i-1}}^{y_i} x(v)\mathrm{d}v\right\},$$

$$(32.25)$$

which we use in the definition of the cumulative probability function

$$F_G(t, y_{i-1}|x) = \int_{y_{i-1}}^t f_G(s, y_{i-1}|x)\mathrm{d}s. \tag{32.26}$$

Suppose now that the last Ca^{2+} spike occurred at time y_{i-1}. We find the next Ca^{2+} spike time as $y_i = y_{i-1} + \Delta$, where

$$\Delta = \inf\{t|F_G(t, y_{i-1}|x) > \omega\}, \tag{32.27}$$

and ω is a random number that is uniformly distributed between 0 and 1. To put it another way, we need to integrate the ISI probability density f_G from y_{i-1} until we obtain a value of ω for the integral and then add the corresponding upper bound of the integral to the previous Ca^{2+} spike time y_{i-1}.

References

1. Berridge MJ, Galione A (1988) Cytosolic calcium oscillators. FASEB J 2:3074–3082
2. Berridge MJ, Lipp P, Bootman MD (2000) The versatility and universality of calcium signalling. Nat Rev Mol Cell Biol 1:11–21
3. Parekh AB (2011) Decoding cytosolic Ca^{2+} oscillations. Trends Biochem Sci 36:78–87
4. Dupont G, Combettes L (2016) Fine tuning of cytosolic Ca^{2+} oscillations. F1000Research 5. https://f1000research.com/articles/5-2036/v1
5. Thul R, Bellamy TC, Roderick HL, Bootman MD, Coombes S (2008) Calcium oscillations. Adv Exp Med Biol 641:1–27
6. Dupont G, Combettes L, Bird GS, Putney JW (2011) Calcium oscillations. Cold Spring Harb Perspect Biol 3:pii:a004226
7. Dupont G, Falcke M, Kirk V, Sneyd J (2016) Models of calcium signalling. Interdisciplinary Applied Mathematics, vol 43. Springer, Cham
8. Schuster S, Marhl M, Höfer T (2002) Modelling of simple and complex calcium oscillations. From single-cell responses to intercellular signalling. Eur J Biochem 269:1333–1355
9. Uhlén P, Fritz N (2010) Biochemistry of calcium oscillations. Biochem Biophys Res Commun 396:28–32

10. Dolmetsch RE, Xu K, Lewis RS (1998) Calcium oscillations increase the efficiency and specificity of gene expression. Nature 392:933–936

11. Noren DP, Chou WH, Lee SH, Qutub AA, Warmflash A, Wagner DS et al (2016) Endothelial cells decode VEGF-mediated Ca^{2+} signaling patterns to produce distinct functional responses. Sci Signal 9:ra20–ra20

12. Di Capite J, Ng SW, Parekh AB (2009) Decoding of cytoplasmic Ca^{2+} oscillations through the spatial signature drives gene expression. Curr Biol 19:853–858

13. Nunemaker CS, Bertram R, Sherman A, Tsaneva-Atanasova K, Daniel CR, Satin LS (2006) Glucose modulates $[Ca^{2+}]_i$ oscillations in pancreatic islets via ionic and glycolytic mechanisms. Biophys J 91:2082–2096

14. Tse A, Tse FW, Almers W, Hille B (1993) Rhythmic exocytosis stimulated by GnRH-induced calcium oscillations in rat gonadotropes. Science (New York, NY). 260:82–84

15. Colsoul B, Schraenen A, Lemaire K, Quintens R, Van Lommel L, Segal A et al (2010) Loss of high-frequency glucose-induced Ca^{2+} oscillations in pancreatic islets correlates with impaired glucose tolerance in Trpm5$^{-/-}$ mice. Proc Natl Acad Sci U S A 107:5208–5213

16. Tsaneva-Atanasova K, Yule DI, Sneyd J (2005) Calcium oscillations in a triplet of pancreatic acinar cells. Biophys J 88:1535–1551

17. Tse A, Lee AK, Tse FW (2012) Ca^{2+} signaling and exocytosis in pituitary corticotropes. Cell Calcium 51:253–259

18. Santella L, Lim D, Moccia F (2004) Calcium and fertilization: the beginning of life. Trends Biochem Sci 29:400–408

19. Whitaker M (2006) Calcium at fertilization and in early development. Physiol Rev 86:25–88

20. Denninger P, Bleckmann A, Lausser A, Vogler F, Ott T, Ehrhardt DW et al (2014) Male-female communication triggers calcium signatures during fertilization in arabidopsis. Nat Commun 5:4645

21. Van Nhieu GT, Clair C, Bruzzone R, Mesnil M, Sansonetti P, Combettes L (2003) Connexin-dependent inter-cellular communication increases invasion and dissemination of Shigella in epithelial cells. Nat Cell Biol 5:720–726

22. Ullah G, Jung P, Machaca K (2007) Modeling Ca^{2+} signaling differentiation during oocyte maturation. Cell Calcium 42:556–564

23. Ullah G, Jung P (2006) Modeling the statistics of elementary calcium release events. Biophys J 90:3485–3495

24. Shuai JW, Jung P (2003) Optimal ion channel clustering for intracellular calcium signaling. Proc Natl Acad Sci U S A 100:506–510

25. Shuai JW, Jung P (2002) Stochastic properties of Ca^{2+} release of inositol 1,4,5-trisphosphate receptor clusters. Biophys J 83:87–97

26. Shuai J, Jung P (2002) Optimal intracellular calcium signaling. Phys Rev Lett 88:068102

27. De Young GW, Keizer J (1992) A single-pool inositol 1,4,5-trisphosphate-receptor-based model for agonist-stimulated oscillations in Ca^{2+} concentration. Proc Natl Acad Sci U S A 89:9895–9899

28. Li Y, Rinzel J (1994) Equations for InsP$_3$ receptor-mediated $[Ca^{2+}]_i$ oscillations derived from a detailed kinetic model: a Hodgkin-Huxley like formalism. J Theor Biol 166:461–473

29. Tang Y, Stephenson J, Othmer H (1996) Simplification and analysis of models of calcium dynamics based on IP$_3$-sensitive calcium channel kinetics. Biophys J 70:246–263

30. Gaspers LD, Bartlett PJ, Politi A, Burnett P, Metzger W, Johnson J et al (2014) Hormone-induced calcium oscillations depend on cross-coupling with inositol 1,4,5-trisphosphate oscillations. Cell Rep 9:1209–1218

31. Politi A, Gaspers LD, Thomas J, Höfer T (2006) Models of IP$_3$ and Ca^{2+} oscillations: frequency encoding and identification of underlying feedbacks. Biophys J 90:3120–3133

32. Sun CH, Wacquier B, Aguilar DI, Carayol N, Denis K, Boucherie S et al (2017) The Shigella type III effector IpgD recodes Ca^{2+} signals during invasion of epithelial cells. EMBO J 36(17):2567–2580

33. Bessonnard S, De Mot L, Gonze D, Barriol M, Dennis C, Goldbeter A et al (2014) Gata6, Nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network. Development (Cambridge, England) 141:3637–3648

34. De Caluwé J, Dupont G (2013) The progression towards Alzheimer's disease described as a bistable switch arising from the positive loop between amyloids and Ca^{2+}. J Theor Biol 331:12–18

35. Dupont G, Lokenye EFL, Challiss RAJ (2011) A model for Ca^{2+} oscillations stimulated by the type 5 metabotropic glutamate receptor: an unusual mechanism based on repetitive, reversible phosphorylation of the receptor. Biochimie 93:2132–2138

36. Dupont G, Abou-Lovergne A, Combettes L (2008) Stochastic aspects of oscillatory Ca^{2+} dynamics in hepatocytes. Biophys J 95:2193–2202

37. Swillens S, Champeil P, Combettes L, Dupont G (1998) Stochastic simulation of a single inositol 1,4,5-trisphosphate-sensitive Ca^{2+} channel reveals repetitive openings during 'blip-like' Ca^{2+} transients. Cell Calcium. 23:291–302

38. Tran Van Nhieu G, Kai Liu B, Zhang J, Pierre F, Prigent S, Sansonetti P et al (2013) Actin-based confinement of calcium responses during Shigella invasion. Nat Commun 4:1567

39. Salazar C, Zaccaria Politi A, Höfer T (2008) Decoding of calcium oscillations by phosphorylation cycles: analytic results. Biophys J 94:1203–1215

40. Kummer U, Olsen LF, Dixon CJ, Green AK, Bornberg-Bauer E, Baier G (2000) Switching from simple to complex oscillations in calcium signaling. Biophys J 79:1188–1195

41. Kummer U, Krajnc B, Pahle J, Green AK, Dixon CJ, Marhl M (2005) Transition from stochastic to deterministic behavior in calcium oscillations. Biophys J 89:1603–1611

42. Sneyd J, Han JM, Wang L, Chen J, Yang X, Tanimura A et al (2017) On the dynamical structure of calcium oscillations. Proc Natl Acad Sci U S A 114:1456–1461

43. Cao P, Tan X, Donovan G, Sanderson MJ, Sneyd J (2014) A deterministic model predicts the properties of stochastic calcium oscillations in airway smooth muscle cells. PLoS Comput Biol 10:e1003783

44. Wang IY, Bai Y, Sanderson MJ, Sneyd J (2010) A mathematical analysis of agonist- and KCl-induced Ca^{2+} oscillations in mouse airway smooth muscle cells. Biophys J 98:1170–1181

45. Higgins ER, Cannell MB, Sneyd J (2006) A buffering SERCA pump in models of calcium dynamics. Biophys J 91:151–163

46. Sneyd J, Tsaneva-Atanasova K, Yule DI, Thompson JL, Shuttleworth TJ (2004) Control of calcium oscillations by membrane fluxes. Proc Natl Acad Sci U S A 101:1392–1396

47. Sneyd J, Tsaneva-Atanasova K, Reznikov V, Bai Y, Sanderson MJ, Yule DI (2006) A method for determining the dependence of calcium oscillations on inositol trisphosphate oscillations. Proc Natl Acad Sci USA 103:1675–1680

48. Nowacki J, Mazlan S, Osinga HM, Tsaneva-Atanasova K (2010) The role of large-conductance Calcium-activated K^+ (BK) channels in shaping bursting oscillations of a somatotroph cell model. Physica D Nonlinear Phenomena 239:485–493

49. Bootman M, Niggli E, Berridge M, Lipp P (1997) Imaging the hierarchical Ca^{2+} signalling system in HeLa cells. J Physiol 499 (Pt 2):307–314

50. Smith IF, Wiltgen SM, Parker I (2009) Localization of puff sites adjacent to the plasma membrane: functional and spatial characterization of Ca^{2+} signaling in SH-SY5Y cells utilizing membrane-permeant caged IP_3. Cell Calcium 45:65–76

51. Taufiq-Ur-Rahman, Skupin A, Falcke M, Taylor CW (2009) Clustering of $InsP_3$ receptors by $InsP_3$ retunes their regulation by $InsP_3$ and Ca^{2+}. Nature 458:655–659

52. Keebler MV, Taylor CW (2017) Endogenous signalling pathways and caged IP_3 evoke Ca^{2+} puffs at the same abundant immobile intracellular sites. J Cell Sci 130:3728–3739

53. Konieczny V, Keebler MV, Taylor CW (2012) Spatial organization of intracellular Ca^{2+} signals. Semin Cell Dev Biol 23:172–180

54. Bruno L, Solovey G, Ventura AC, Dargan S, Dawson SP (2010) Quantifying calcium fluxes underlying calcium puffs in Xenopus laevis oocytes. Cell Calcium 47:273–286

55. Shuai J, Pearson JE, Foskett JK, Mak DOD, Parker I (2007) A kinetic model of single and clustered IP_3 receptors in the absence of Ca^{2+} feedback. Biophys J 93:1151–1162

56. Smith IF, Parker I (2009) Imaging the quantal substructure of single IP_3R channel activity during Ca^{2+} puffs in intact mammalian cells. Proc Natl Acad Sci U S A 106:6404–6409

57. Flegg MB, Rüdiger S, Erban R (2013) Diffusive spatio-temporal noise in a first-passage time model for intracellular calcium release. J Chem Phys 138:154103
58. Dobramysl U, Rüdiger S, Erban R (2016) Particle-based multiscale modeling of calcium puff dynamics. Multiscale Model Simul 14:997–1016
59. Weinberg SH, Smith GD (2014) The influence of Ca^{2+} buffers on free $[Ca^{2+}]$ fluctuations and the effective volume of Ca^{2+} microdomains. Biophys J 106:2693–2709
60. Wieder N, Fink R, von Wegner F (2015) Exact stochastic simulation of a calcium microdomain reveals the impact of Ca^{2+} fluctuations on IP_3R gating. Biophys J 108:557–567
61. Falcke M, Moein M, Tilunaite A, Thul R, Skupin A (2018) On the phase space structure of IP_3 induced Ca^{2+} signalling and concepts for predictive modeling. Chaos (Woodbury, NY) 28:045115
62. Skupin A, Kettenmann H, Winkler U, Wartenberg M, Sauer H, Tovey SC et al (2008) How does intracellular Ca^{2+} oscillate: by chance or by the clock? Biophys J 94:2404–2411
63. Thurley K, Smith IF, Tovey SC, Taylor CW, Parker I, Falcke M (2011) Timescales of IP_3-evoked Ca^{2+} spikes emerge from Ca^{2+} puffs only at the cellular level. Biophys J 101:2638–2644
64. Moein M, Grzyb K, Gonçalves Martins T, Komoto S, Peri F, Crawford A et al (2018) CaSiAn: a calcium signaling analyzer tool. Bioinformatics (Oxford, England) 1:11
65. Thurley K, Tovey SC, Moenke G, Prince VL, Meena A, Thomas J et al (2014) Reliable encoding of stimulus intensities within random sequences of intracellular Ca^{2+} spikes. Sci Signal 7:ra59
66. Rooney TA, Sass EJ, Thomas AP (1990) Agonist-induced cytosolic calcium oscillations originate from a specific locus in single hepatocytes. J Biol Chem 265:10792–10796
67. Tilunaite A, Croft W, Russell N, Bellamy TC, Thul R (2017) A Bayesian approach to modelling heterogeneous calcium responses in cell populations. PLoS Comput Biol 13:e1005794
68. Barbieri R, Quirk MC, Frank LM, Wilson MA, Brown EN (2001) Construction and analysis of non-Poisson stimulus-response models of neural spiking activity. J Neurosci Methods 105:25–37
69. Miura K, Tsubo Y, Okada M, Fukai T (2007) Balanced excitatory and inhibitory inputs to cortical neurons decouple firing irregularity from rate modulations. J Neurosci 27:13802–13812
70. Thurley K, Skupin A, Thul R, Falcke M (2012) Fundamental properties of Ca^{2+} signals. Biochim Biophys Acta 1820:1185–1194
71. Croft W, Reusch K, Tilunaite A, Russel N, Thul R, Bellamy TC (2016) Probabilistic encoding of stimulus strength in astrocyte global calcium signals. Glia 64:537–552
72. Thul R, Falcke M (2006) Frequency of elemental events of intracellular Ca^{2+} dynamics. Phys Rev E 73(6 Pt 1):061923
73. Thul R, Falcke M (2007) Waiting time distributions for clusters of complex molecules. Europhys Lett 79:38003
74. Thul R, Thurley K, Falcke M (2009) Toward a predictive model of Ca^{2+} puffs. Chaos (Woodbury, NY) 19:037108
75. Thurley K, Falcke M (2011) Derivation of Ca^{2+} signals from puff properties reveals that pathway function is robust against cell variability but sensitive for control. Proc Natl Acad Sci U S A 108:427–432
76. Moenke G, Falcke M, Thurley K (2012) Hierarchic stochastic modelling applied to intracellular Ca^{2+} signals. PLoS One 7:e51178
77. Papoulis A, Pillai SU (2002) Probability, random variables and stochastic processes, 4th edn. McGraw Hill, Boston
78. Ross SM (2003) Introduction to probability models, 8th edn. Academic, Amsterdam
79. Cunningham JP, Shenoy KV, Sahani M (2008) Fast gaussian process methods for point process intensity estimation. In: Proceedings of the 25th International Conference on Machine Learning, Stanford University, Palo Alto, pp 192–199

80. Cunningham JP, Yu B, Shenoy K, Sahani M (2008) Inferring neural firing rates from spike trains using gaussian processes. In: Platt JC, Koller D, Singer Y, Roweis S (eds) Advances in neural information processing systems, vol 20. MIT Press. Cambridge, MA, pp 329–336
81. Robert CP, Casella G (2005) Monte Carlo statistical methods. Springer Texts in Statistics. Springer, New York
82. Brown EN, Barbieri R, Ventura V, Kass RE, Frank LM (2002) The time-rescaling theorem and its application to neural spike train data analysis. Neural Comput 14:325–346
83. Skupin A, Falcke M (2010) Statistical analysis of calcium oscillations. Eu Phys J Spec Top 187:231–240
84. Devroye L (1986) Non-uniform random variate generation. Springer, New York