Dissertations, Theses, and Masters Projects   Theses, Dissertations, & Master Projects

2012

Automated Reductions of Markov Chain Models of Calcium Release Site Models

Yan Hao
College of William & Mary - Arts & Sciences

Follow this and additional works at: https://scholarworks.wm.edu/etd

Part of the Applied Mathematics Commons, and the Biophysics Commons

Recommended Citation

Automated Reductions of Markov Chain Models of
Calcium Release Site Models

Yan Hao

Hebei, China

Master of Science, The College of William and Mary, 2010
Bachelor of Science, Tsinghua University, 2006

A Dissertation presented to the Graduate Faculty
of the College of William and Mary in Candidacy for the Degree of
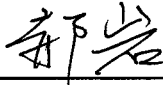Doctor of Philosophy

Department of Applied Science

The College of William and Mary
January, 2012

APPROVAL PAGE

This dissertation is submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

_____
Yan Hao

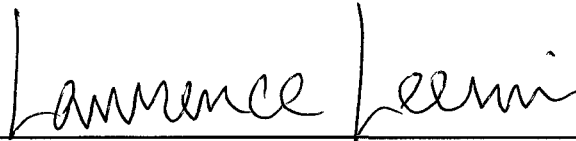Approved by the Committee, August, 2011

_____
Committee Chair
Associate Professor Gregory Smith, Applied Science
The College of William and Mary

_____
Associate Professor Peter Kemper, Computer Science
The College of William and Mary

_____
Professor Lawrence Leemis, Mathematics
The College of William and Mary

_____
Assistant Professor Leah Shaw, Applied Science
The College of William and Mary

ABSTRACT PAGE

Markov chain models have played an important role in understanding the relationship between single channel gating of intracellular calcium (Ca2+) channels, specifically 1,4,5-trisphosphate receptors (IP3Rs) and ryanodine receptors(RyRs), and the stochastic dynamics of Ca2+ release events, known as Ca2+ puffs and sparks. Mechanistic Ca2+ release site models are defined by the composition of single channel models whose transition probabilities depend on the local calcium concentration and thus the state of the other channels. Unfortunately, the large state space of such compositional models impedes simulation and computational analysis of the whole cell Ca2+ signaling in which the stochastic dynamics of localized Ca2+ release events play an important role. This dissertation introduces, implements and validates the application of several automated model reduction techniques that significantly reduce the computational cost of mechanistic compositionally defined Ca2+ release site models.

A common feature of Ca2+ channel models is the separation of time scales. For example, the well-known bell-shaped equilibrium open probability of IP3Rs can be reproduced by Markov Chain models that include transitions mediated by fast Ca2+ activation and slower Ca2+ inactivation. Chapter 2 introduces an automated model reduction technique that is based on fast/slow analysis that leverages these time scale differences. Rate constants in the single channel model are categorized as either fast or slow, groups of release site states that are connected by fast transitions are identified and lumped, and transition rates between reduced states are chosen consistent with the conditional probability distributions among states within each group. The fast/slow reduction approach is validated by the fact that puff/spark statistics can be efficiently computed from reduced Ca2+ release site models with small and transient error.

For Markov chain Ca2+ release site models without time-scale separation, the manner in which the full model states should be aggregated for optimal reduction is difficult to determine a priori. In Chapter 3, a genetic algorithm based approach that mimics the inheritance, mutation and selection processes of natural evolution is implemented to reduce these models. Given a full model of interest and target reduced model size, this genetic algorithm searches for set partitions, each corresponding to a potential scheme for state aggregation, that lead to reduced models that well-approximate the full model. A whole cell model with coupled local and global Ca2+ signaling is simplified by replacing a compositionally defined full Ca2+ release site model with a reduced model obtained through the genetic algorithm.

In Chapter 4, a Langevin formulation of Ca2+ release sites is introduced as an alternative model reduction technique that is applicable when the number of channels per Ca2+ release site is too large for the previously discussed reduction methods, but not so large that the stochasticity of Ca2+ release is negligible. The Langevin formulation for coupled intracellular Ca2+ channels results in stochastic differntial equations that well-approximate the corresponding Markov chain models when release sites possess as few as 20 channels, and the agreement improves as the number of channels per release site increases. Importantly, the computational time required by the Langevin approach does not increase with the size of Ca2+ release site.

Contents

List of Figures

List of Tables

Acknowledgments

Chapter 1

Introduction

Although widely recognized as an important component of all skeletal structures, calcium, as a divalent cation (Ca^{2+}), is an ubiquitous element of intracellular signaling and more importantly, it is a versatile second messenger which regulates many cellular activities [Carafoli et al., 2001]. To achieve this versatility, the Ca^{2+}-signaling system operates in many different ways. First of all, the time scale of Ca^{2+}-signals may vary over several orders of magnitude [Marchant and Parker, 2000]. Rapid transient changes in Ca^{2+} concentration ($[Ca^{2+}]$) control cell locomotion, neural transmission, hormonal secretion and muscle contraction directly. For example, Ca^{2+} triggers exocytosis within microseconds at the synaptic junction [Berridge et al., 2003]. Sustained $[Ca^{2+}]$ elevation plays pivotal roles in many vital cellular functions such as fertilization, gene expression and apoptosis where Ca^{2+} operates over minutes or even hours [Berridge et al., 1998]. The versatility of Ca^{2+}-signaling also lies in terms of amplitude and spacial patterns. For example, the Ca^{2+} signal that triggers muscle contraction are in forms of local $[Ca^{2+}]$ spikes while a Ca^{2+} wave propagates throughout the cell during the fertilization process [Berridge et al., 2000].

1.1 Calcium homeostasis

The Ca^{2+} signaling system consists of a large number of components sometimes called "The Ca^{2+}-signalling tool kit" including a variety of receptors, transducers, channels, buffers, effectors, enzymes, pumps and exchangers. Even though the mechanisms of these components are beyond the scope of this thesis, it is important to briefly review a few of them which will appear later in this report.

A common feature of the Ca^{2+}-signalling processes is a significant elevation in $[Ca^{2+}]$: Ca^{2+} signals usually initiate with a brief pulse of Ca^{2+} which raises the local cytosolic $[Ca^{2+}]$ level roughly 10 times or higher [Berridge et al., 2003]. This feature highly depends on the sustained macroscopic $[Ca^{2+}]$ gradients across cell surface and intracellular membranes. The membranes that surround cells and other intracellular organelles are semi-permeable which helps to maintain approximately 10^4 fold $[Ca^{2+}]$ gradients between the cytosol (approximately 100 μM) and the extracellular space, endoplasmic reticulum (ER), sarcoplasmic reticulum (SR) and mitochondria $[Ca^{2+}]$ [Cheng and Lederer, 2008]. Sustained high $[Ca^{2+}]$ in the cytosol is toxic and the low cytosolic $[Ca^{2+}]$ level is actively maintained by the Ca^{2+} homeostasis mechanisms including 1) the plasmalemmal Na^+/Ca^{2+} exchanger (NCX) and Ca^{2+}-ATPase (PMCA), which extrude cytosolic Ca^{2+} to the extracellular space. 2) the ER/SR Ca^{2+}-ATPase (SERCA) which transport Ca^{2+} back to ER/SR and 3) Ca^{2+} buffers which are Ca^{2+} binding proteins that sequester free Ca^{2+}, as illustrated in Fig. 1.1 [Cheng and Lederer, 2008, Berridge et al., 2003].
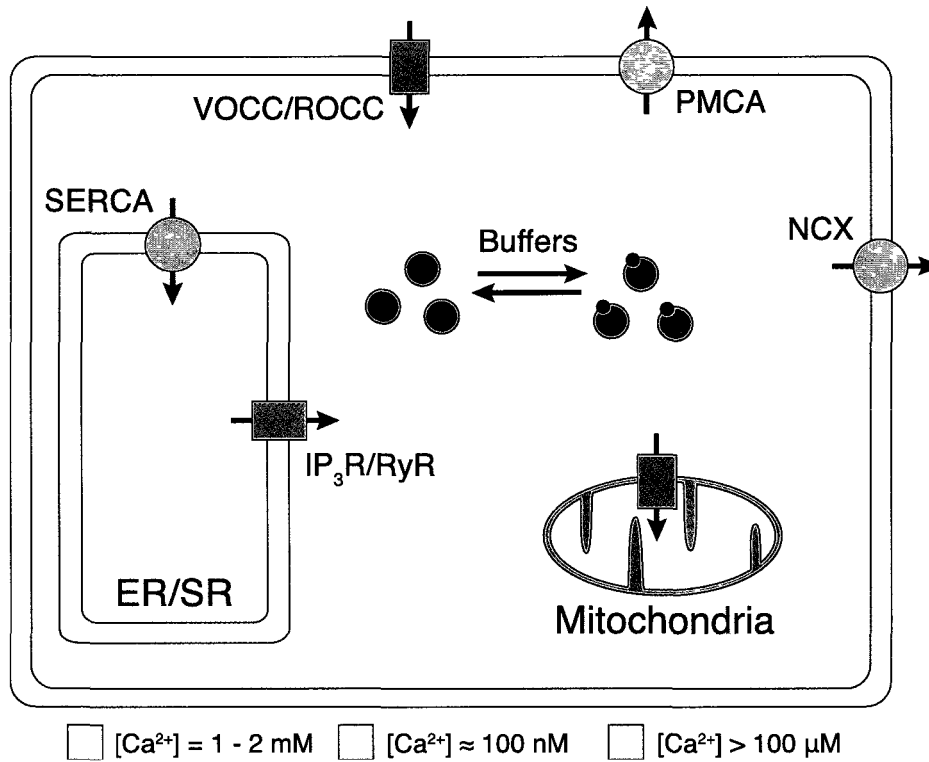
Figure 1.1: Schematic representation of intracellular Ca^{2+} signaling components. Calcium concentration gradients across the plasma membrane, the endoplasmic reticulum (ER), and the sarcoplasmic reticulum (SR) are maintained by plasma membrane Ca^{2+} ATPase (PMCA) and sarcoplasmic endoplasmic reticulum Ca^{2+} ATPase (SERCA) pumps, and other ion transport proteins such as the Na^+/Ca^{2+} exchanger (NCX). Uptake of Ca^{2+} into the mitochondrial matrix is mediated by a highly selective Ca^{2+} channel previously thought to be a uniporter [Kirichok et al., 2004]. Free cytosolic Ca^{2+} is buffered by a host of Ca^{2+} binding proteins and chelators. Release of Ca^{2+} from the ER or SR is mediated by two families of ROCCs, inositol 1,4,5-trisphosphate receptors (IP$_3$Rs) and ryanodine receptors (RyRs). The arrows through transport proteins and channels indicate the direction of Ca^{2+} movement. Adapted from [Groff, 2008]

3

The Ca^{2+}-signals can be generated by using both internal and external Ca^{2+} resources. Ca^{2+} entry from the extracellular space is mediated by Ca^{2+}-permeant channels which are gated by voltage, ligands, temperature and a variety of other mechanisms [Cheng and Lederer, 2008]. In contrast, the Ca^{2+} release from the intracellular Ca^{2+} reservoir, ER/SR, is mediated by only two families of Ca^{2+} channels, which are the modeling objects of this thesis, the ryanodine receptor (RyR) [Fill and Copello, 2002, Meissner, 1994] and the Inositol-1-4-5-trisphosphate receptor (IP_3R) [Berridge, 1993a, Foskett et al., 2007]. These Ca^{2+} channels are selectively permeable intermembrane protein tetramers which, when activated, allow Ca^{2+} flow passively down their concentration gradients. Though both families of Ca^{2+} channels have three major isoforms (type 1, 2, and 3) [Cheng and Lederer, 2008] and require Ca^{2+} binding on their cytosolic side to activate, they are distinct in several aspects. For example, RyRs are expressed primarily in excitable cells while IP_3Rs exist in many different cell types. More importantly, the opening of IP_3Rs requires co-agonist Inositol-1-4-5-trisphosphate (IP_3) in addition to Ca^{2+} binding [Bezprozvanny et al., 1991] while the absence of plant alkaloid ryanodine will lower the open probability of RyRs rather than prevent the opening of these channels.

1.2   Local calcium signaling

The distribution of IP_3Rs and RyRs on the ER/SR membranes are known to be in clusters, dubbed Ca^{2+} release sites/complexes. For example, in skeletal and cardiac myocytes, Ca^{2+} release sites are composed of a minimum of 10 and a maximum of more than 200 RyRs depending on cell type and species, and the distances between release sites are 5-20 times larger than the distance between neighboring channels [Franzini-Armstrong et al., 1999]. In the cortical regions of immature *Xenopus laevis*

oocytes, IP_3Rs are co-located in clusters of 5 to 50 with inter-cluster distance on the order of microns [Sun et al., 1998]. As mentioned in Sec. 1.1, IP_3Rs and RyRs are known to be activated by Ca^{2+}. A small increase in cytosolic $[Ca^{2+}]$ close to the Ca^{2+} channels may elevate the open probability of the channels and promote further release from the ER/SR, a phenomenon known as Ca^{2+}-induced Ca^{2+} release (CICR) [Bers, 2002].

The clustering of IP_3Rs and RyRs and the CICR phenomenon of results in strong interactions and cooperative gating within Ca^{2+} release sites and further generates a variety of intracellular Ca^{2+} signals. Three distinct modes of Ca^{2+} mobilization that have been observed via confocal microfluorimetry in oocytes, cardiomyocytes, and many other cell types: 1) Fundamental response: localized Ca^{2+} elevations caused by the activation of a single channel, also referred to as Ca^{2+} blips or quarks depending on whether the events are mediated by an IP_3R or RyR [Niggli, 1999, Berridge, 2006]; 2) Elementary response: Ca^{2+} elevations due to the activation of multiple IP_3Rs or RyRs associated with a single Ca^{2+} release site, also referred to as Ca^{2+} puffs and sparks [Cheng et al., 1993a, Cannell et al., 1995, Yao et al., 1995, Parker et al., 1996, Berridge, 2006]; and 3) Global response: global Ca^{2+} elevations such as oscillations and waves that involve multiple release sites [Cheng et al., 1996, Berridge, 2006].

1.3   Markov chain model of single calcium channels

Historically, the stochastic dynamics of single Ca^{2+} ion channel gating has been successfully modeled using Markov chains. More importantly, when Markov chain models of these intracellular Ca^{2+}-regulated Ca^{2+} channels are coupled via a mathematical representation of a Ca^{2+} microdomain, simulated Ca^{2+} release sites may exhibit the phenomenon of stochastic Ca^{2+} excitability where channels open and close in a con-

certed fashion reminiscent of Ca^{2+} puffs and sparks [Swillens et al., 1999]. Since the main focus of this thesis is to present novel techniques to reduce compositionally defined Markov chain Ca^{2+} release site models, it would be helpful to briefly review the basic mathematical definitions and properties of Markov chains.

1.3.1 Stochastic processes and Markov chains

To define a Markov chain, we must first give the definition of a stochastic process.

Definition Given a probability space $(\Omega, \mathcal{F}, P)$, a *stochastic process* (or random process) with state space $\mathcal{S}$ is a collection of $\mathcal{S}$-valued random variables on Ω indexed by a set T ("time") [Gardiner, 2009]. That is, a stochastic process X is a collection

$$\{X_t : t \in T\}$$

where each X_t is an F-valued random variable on Ω.

A Markov process, named after Andrey Markov, refers to an $\mathbb{R}^n$-valued stochastic process $X = \{X_t : t \in T\}$ on a probability space $(\Omega, \mathcal{F}, P)$ that satisfies the Markov property: the conditional probability distribution of future states of the process, given the present state and the past states, depend only upon the present state; that is, the past is irrelevant because it doesn't matter how the current state was obtained. The state space S of Markov processes can be continuous, when S is discrete (finite or countable), the Markov process is named a Markov chain [Norris, 1997].

1.3.2 Discrete-time Markov chains

A "discrete-time" stochastic process refers to systems where only the order of events is important. In these systems, the "time" set T takes the form of a set of integer

valued "steps." State changes randomly between "steps" which refer time as well as physical distance or any other discrete measurement.

When a Markov chain takes discrete values and is indexed by a discrete time, it is named a discrete-time Markov chain (DTMC), and the Markov property can be reformulated as follows:

$$\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_0 = x_0) = \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}). \qquad (1.1)$$

Let S_n ($n \in \mathbb{Z}$ and $n \geq 0$) represent the state of the DTMC at the nth step, the statistical information of a DTMC can be compactly collected by a matrix

$$\boldsymbol{P} = \{p_{ij}\}, \qquad (1.2)$$

where

$$p_{ij} = Pr\{S_{n+1} = j | S_n = i\} \qquad (1.3)$$

indicates the transition probability from state i to state j in a single step. Notice that $\boldsymbol{P}$ is a stochastic matrix, i.e.

$$\sum_{j=1}^{N} p_{ij} = 1, \qquad (1.4)$$

where N is the cardinality of the state space.

The mth power of $\boldsymbol{P}$, $\boldsymbol{P}^m$, is called the m-step transition probability matrix, where each element

$$p_{ij}^{(m)} = Pr\{S_{n+m} = j | S_n = i\}, \qquad (1.5)$$

indicates the probability that the system visits state j from state i after exactly m steps. Given any initial distribution of the Markov chain $\boldsymbol{\pi}(0)$, the probability that

7

this DTMC is in state i after m steps is consequently

$$\boldsymbol{\pi}(m) = \boldsymbol{\pi}(0)\boldsymbol{P}^m. \tag{1.6}$$

Definition A DTMC is called *aperiodic* if for any state $i \in \mathcal{S}$, the set $\{m \geq 0 : p_{ii}^{(m)} > 0\}$ has no common divisor other than 1

Definition A DTMC is called *irreducible* if there exists $m \in \mathbb{Z}^+$ such that $p_{ij}^{(m)} > 0$ for all $i, j \in \mathcal{S}$.

If a DTMC has both aperiodic and irreducible properties, it has a unique limiting distribution $\bar{\pi}$ which satisfies:

$$\bar{\pi} = \bar{\pi}P \quad \text{and} \quad \bar{\pi}\mathbf{e} = 1, \tag{1.7}$$

where $\boldsymbol{e}$ is a commensurate vector of ones.

1.3.3 Continuous-time Markov chains

When modeling the gating of Ca^{2+} ion channels, we more often use continuous-time Markov chains (CTMC), where the "time" set T is the set of all non-negative real numbers, for two main reasons. First, CTMCs monitor the time when the channel changes its states, and second the dwell times that these ion channels reside in each state are observed to follow exponential distributions [Bezprozvanny and Ehrlich, 1994].

Analogous to the P-matrices for DTMCs, CTMCs are often described by their corresponding infinitesimal generator matrix

$$\mathbf{Q} = \{q_{ij}\}, \tag{1.8}$$

8

where

$$
q_{ij} = \begin{cases} \lim\limits_{\Delta t \to 0} \dfrac{Pr\{S(t+\Delta t) = S_j | S(t) = S_i\}}{\Delta t} & (i \neq j) \\[2ex] -\sum\limits_{i \neq j} q_{ij} & (i = j) \end{cases} \tag{1.9}
$$

indicates the rate of transition from state i to state j per unit time. CTMCs also have a corresponding transition probability matrix

$$
\mathbf{P}(t) = \{p_{ij}(t)\}, \tag{1.10}
$$

where

$$
p_{ij}(t) = Pr\{S_{t_0+t} = j | S_{t_0} = i\}. \tag{1.11}
$$

Notice that $\mathbf{P}$ is a function of time describing the integrated change of rate given by matrix $\mathbf{Q}$. This relationship between the P-matrix and Q-matrix is elegantly described by the differential Chapman-Kolmogrov equation:

$$
\frac{d}{dt}\mathbf{P}(t) = \mathbf{P}(t)\mathbf{Q}. \tag{1.12}
$$

Similar to the stationary distribution of DTMCs, irreducible CTMCs also have a unique stationary distribution $\bar{\pi}$ which leading to a zero changing rate of $\pi(t)$, which satisfies:

$$
\bar{\pi}Q = \mathbf{0} \quad \text{and} \quad \bar{\pi}\mathbf{e} = 1, \tag{1.13}
$$

9
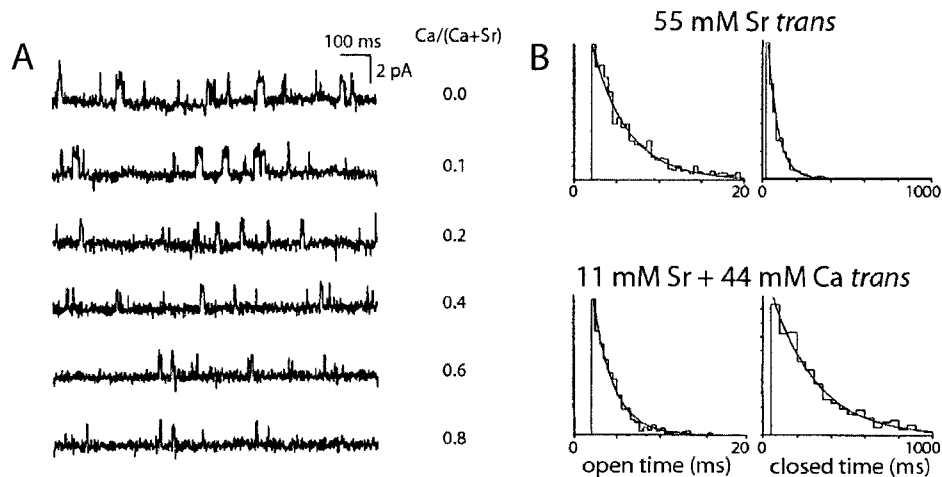
Figure 1.2: (A) Planar lipid bilayer recordings of reconstituted type I IP₃Rs using different mole fractions of $[Ca^{2+}]$ on the *trans* (luminal) side of channels demonstrate gating between two conductance levels, 0 pA (closed) and 2 pA (open). The $[Ca^{2+}]$ and $[IP_3]$ of the cis bath is 0.2 μM and 2 μM, respectively. (B) The distribution of open and closed dwell times are well fitted by exponential functions of the *trans* $[Ca^{2+}]$. Adapted from [Bezprozvanny and Ehrlich, 1994]

1.4 Markov chain models of single calcium channel gating

Because the conformational changes of IP₃Rs and RyRs during the gating processes are not yet detectable in vivo, our first understanding of the single Ca^{2+} channel gating kinetics is from the planar lipid bilayer current recordings of reconstituted single channels [Bezprozvanny et al., 1991,Bezprozvanny and Ehrlich, 1994]. Fig. 1.2A shows several sample planar lipid bilayer recordings of reconstituted type I IP₃Rs. The channel is closed most of the time and occasionally opens briefly resulting a 2 pA current deviation. The channel transfers stochastically between the open and closed state, and the dwell times of both states are exponentially distributed.

Starting from the 1990's the stochastic gating of the IP₃Rs and RyRs are success-fully modeled by Markov chains [Stewart, 1994,Colquhoun and Hawkes, 1995]. Here

10

we introduce the Markov chain models of single Ca^{2+} channels using the simplest two-state model with one closed state (C) and one open state (O):

$$\text{(closed) } C \underset{\beta}{\overset{\alpha}{\rightleftarrows}} O \text{ (open)}, \qquad (1.14)$$
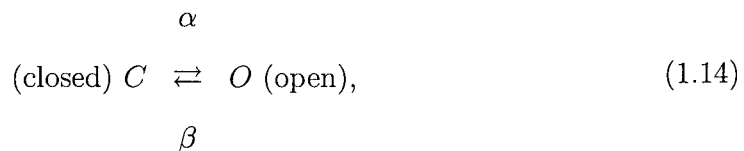
where α and β are the forward and backward transition rates, respectively, with units of time^{-1}. The transition rates of the single Ca^{2+} channels are often selected such that the model reproduces the experimental observations. Eq. 1.14 defines a CTMC, $X(t)$, with state space $S \in \{C, O\}$ and the dwell times in the open and closed states are exponentially distributed with expectations $1/\alpha$ and $1/\beta$ respectively. The corresponding infinitesimal generator matrix $\mathbf{Q}$ is given by

$$\mathbf{Q} = \{q_{ij}\} = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix} \qquad (1.15)$$

Markov chain models of IP$_3$R and RyR gating often include additional experimental observations other than simply opening and closing. For example, Ca^{2+}-dependent activation [Bezprozvanny et al., 1991, Keizer and Levine, 1996] can be included in the two-state model given in Eq. 1.15 by adding the [Ca^{2+}] into the forward transition rate

$$\text{(closed) } C \underset{k^-}{\overset{k^+c}{\rightleftarrows}} O \text{ (open)}, \qquad (1.16)$$
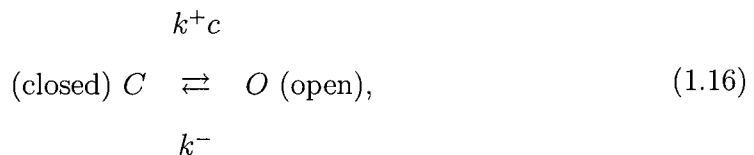
where c is the [Ca^{2+}] in the small cytosolic side space close to the channel and k^+ is the association rate constant with units conc^{-1}time^{-1}. Adding states is another

way to reflect the phenomena observed in experiments. For example, the three-state single channel models in Chap. 3 and Chap. 4 include an additional long-lasting closed state which can be reached from the open state to reflect the Ca^{2+} inactivation [Schiefer et al., 1995, Mak and Foskett, 1997]. It is important to be aware that conformational structures are also important aspects to be included in single Ca^{2+} channel models. These models are biophysically more realistic but often have a larger number of states and are computationally more expensive to study. The DeYoung-Keizer model [De Young and Keizer, 1992, Shuai et al., 2007], for example, includes 4 subunits and 9 states per unit which results in 495 total states.

1.5 Compositionally defined Markov chain calcium release site models

Clusters of Ca^{2+} channels, often called Ca^{2+} release sites or release complexes, can also be modeled by Markov chains [Ríos and Stern, 1997, Swillens et al., 1998, Stern et al., 1999, Shuai and Jung, 2002b, Rengifo et al., 2002, Hinch et al., 2004, Mazzag et al., 2005, DeRemigio and Smith, 2005, Nguyen et al., 2005, Huertas and Smith, 2007]. In these Ca^{2+} release site models intracellular Ca^{2+}-regulated Ca^{2+} channels are coupled via a mathematical representation of a Ca^{2+} microdomain, simulated Ca^{2+} release sites may exhibit the phenomenon of stochastic Ca^{2+} excitability where channels open and close in a concerted fashion reminiscent of Ca^{2+} puffs and sparks. These models usually are composed of several identical Ca^{2+} channels and include information such as the gating dynamics (state space) of the single channel, the geometry of the release site and the dynamics of cytosolic Ca^{2+} homeostasis.

Among all concurrent mathematical models of Ca^{2+} release sites, compositionally

defined Markov chain models contain the most detail of the physiological realism of channel interaction and collective gating. As the number of channels the release site possesses increases, however, the number of the states of release site models grows exponentially, which is the so-called compositional state-space explosion problem. Based on prior observation [Nguyen et al., 2005] that the spatial organization of Ca^{2+} release sites does not significantly affect its collective gating dynamics, instantaneous mean-field coupling assumption, where all channels are identical and indistinguishable [DeRemigio and Smith, 2005, Groff and Smith, 2008a, Hao et al., 2009], is employed to reduce the state space. Unfortunately, further reduction is still necessary for many Ca^{2+} release site models under the mean-field assumption to be computationally tractable.

Chapter 2

Reduction of calcium release site models via fast/slow analysis and iterative aggregation/disaggregation

2.1 Summary

Mathematical models of calcium release sites derived from Markov chain models of intracellular calcium channels exhibit collective gating reminiscent of the experimentally observed phenomenon of calcium puffs and sparks. Such models often take the form of stochastic automata networks in which the transition probabilities of each channel depend on the local calcium concentration and thus the state of the other channels. In order to overcome the state-space explosion that occurs in such compositionally defined calcium release site models, we have implemented several automated procedures for model reduction using fast/slow analysis. After catego-

rizing rate constants in the single channel model as either fast or slow, groups of states in the expanded release site model that are connected by fast transitions are lumped, and transition rates between reduced states are chosen consistent with the conditional probability distribution among states within each group. For small problems these conditional probability distributions can be numerically calculated from the full model without approximation. For large problems the conditional probability distributions can be approximated without the construction of the full model by assuming rapid mixing of states connected by fast transitions. Alternatively, iterative aggregation/disaggregation may be employed to obtain reduced calcium release site models in a memory-efficient fashion. Benchmarking of several different iterative aggregation/disaggregation-based fast/slow reduction schemes establishes the effectiveness of automated calcium release site reduction utilizing the Koury-McAllister-Stewart method.

Mathematical modeling has played an important role in understanding the relationship between single channel gating of intracellular calcium (Ca^{2+}) channels and the stochastic dynamics of Ca^{2+} release events known as Ca^{2+} puffs and sparks. Ca^{2+} release site models are defined by the composition of single channel models whose transition probabilities depend on the local calcium concentration and thus the state of the other channels. Because the large state space of such models impedes computational analysis of the dynamics of Ca^{2+} release sites, we implement and validate the application of several automated model reduction techniques that leverage separation of time scales, a common feature of single channel models of inositol 1,4,5-trisphosphate receptors and ryanodine receptors. The authors show for the first time that memory-efficient iterative aggregation/disaggregation-based numerical schemes are effective for fast/slow reduction of compositionally defined Ca^{2+} release site models.

The results of this chapter have been presented as "Reduction of calcium release site models via fast/slow analysis and iterative aggregation/disaggregation" in *Chaos* [Hao et al., 2009].

2.2   Introduction

Localized intracellular Ca^{2+} elevations known as puffs and sparks arise from the concerted gating of inositol 1,4,5-trisphosphate receptors (IP_3Rs) and ryanodine receptors (RyRs), intracellular Ca^{2+} channels that are clustered at release sites on the surface of the endoplasmic reticulum or sarcoplasmic reticulum [Cheng et al., 1993b, Cheng et al., 1996, Yao et al., 1995, Parker and Yao, 1996, Parker et al., 1996, Berridge, 1997a, Berridge, 1998].   When Markov chain models of these intracellular Ca^{2+}-regulated Ca^{2+} channels are coupled via a mathematical representation of a Ca^{2+} microdomain, simulated Ca^{2+} release sites may exhibit the phenomenon of "stochastic Ca^{2+} excitability" where channels open and close in a concerted fashion reminiscent of Ca^{2+} puffs and sparks [Swillens et al., 1999, Nguyen et al., 2005].   Detailed modeling and analysis of the stochastic dynamics of Ca^{2+} release has helped to develop our understanding of the relationship between single channel kinetics and emergent phenomena that lead to localized Ca^{2+} elevations such as Ca^{2+} puffs and sparks [Swillens et al., 1998, Swillens et al., 1999, Shuai and Jung, 2002b, Shuai and Jung, 2003, DeRemigio and Smith, 2005, Shuai et al., 2006, Shuai et al., 2007, Huertas and Smith, 2007, Groff and Smith, 2008b, Groff and Smith, 2008a, DeRemigio and Smith, 2008]. However, the state-space explosion that results when Ca^{2+} release site models are compositionally defined in terms of single channel models is a challenge to physiologically realistic modeling of the stochastic dynamics of Ca^{2+} release [Nguyen et al., 2005, DeRemigio et al., 2008].
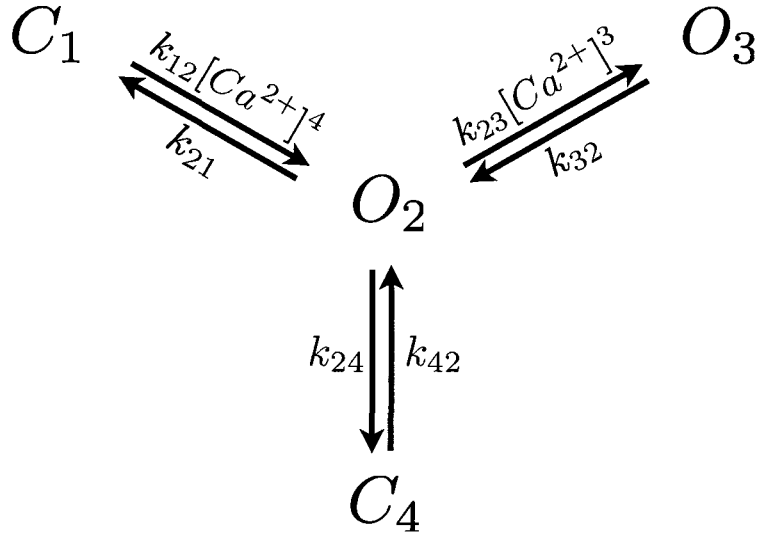
Figure 2.1: State-transition diagram for the Keizer-Levine RyR model [Keizer and Levine, 1996]. This model includes 2 closed (C_1 and C_4) and 2 open (O_2 and O_3) states. The $C_1 \rightarrow O_2$ and $O_2 \rightarrow O_3$ transitions involve binding of 4 and 3 Ca^{2+} ions, respectively, while the other transitions do not involve Ca^{2+}. Parameters as in [Keizer and Levine, 1996]: $k_{12} = 1500 \ \mu M^{-4}s^{-1}$; $k_{23} = 1500 \ \mu M^{-3}s^{-1}$; in s^{-1}: $k_{21} = 28.8$, $k_{32} = 385.9$, $k_{24} = 1.75$, $k_{42} = 0.1$.

Quasi-static approximation based on a separation of time scales is a well-established approach to reducing single channel models of Ca^{2+}-regulated Ca^{2+} channels. Ordinary differential equation (ODE) models of the dynamics of whole cell Ca^{2+} responses are often reduced through the observation that Ca^{2+} activation of IP_3Rs or RyRs is a faster process than Ca^{2+}-dependent or -independent inactivation. For example, the 4-state Keizer-Levine RyR model shown in Fig. 2.1 can be reduced to a 2-state model that can be represented by a single Hodgkin-Huxley-style gating variable in whole cell models of Ca^{2+} oscillations, because the $C_1 \leftrightarrow O_2$ and $O_2 \leftrightarrow O_3$ transitions are fast compared to the $O_2 \leftrightarrow C_4$ transitions [Keizer and Levine, 1996]. Similarly, the well-known 8-state DeYoung-Keizer IP_3R subunit model can be reduced to 2 states by assuming both IP_3 potentiation and Ca^{2+} activation are fast compared to Ca^{2+} inactivation [De Young and Keizer, 1992, Li and Rinzel, 1994].

The fast/slow analysis that occurs in many ODE models of intracellular Ca^{2+} responses is straightforward because the intracellular channels are coupled to the bulk cytosolic $[Ca^{2+}]$, the dynamics of which are assumed to be slow compared to the fast transitions within identified groups of states (e.g., C_1, O_2, and O_3 in Fig. 2.1). While fast/slow reduction can be applied to Markov chain models of Ca^{2+} release sites, the kinetics of domain Ca^{2+} near clusters of intracellular channels are considerably faster than the kinetics of bulk Ca^{2+} (milliseconds as opposed to seconds). Consequently, in the release site models that are the focus of this chapter, the domain $[Ca^{2+}]$ is assumed to be an instantaneous function of the number of open channels at a release site. That is, domain Ca^{2+} is not an environmental variable extrinsic to the Ca^{2+} release site model, but rather an intrinsic aspect of the model that is algebraically determined from the current release site state [Hinch et al., 2004, DeRemigio and Smith, 2005, Groff and Smith, 2008b, Groff and Smith, 2008a]. The focus of this chapter is the implementation and validation of automated fast/slow reduction procedures for this particular class of Ca^{2+} release site models, which are large structured time-homogeneous Markov chains.

The remainder of this chapter is organized as follows. In Secs. 2.3 and 2.4 we motivate our model formulation and show a representative simulation of a Ca^{2+} release site composed of multiple Keizer-Levine RyRs interacting via a common domain $[Ca^{2+}]$. In Secs. 2.5 and 2.6 we demonstrate and validate fast/slow reduction of compositionally defined Ca^{2+} release site models. Importantly, the conditional probability distributions required for fast/slow reduction can be numerically approximated without the construction of the full model, resulting in a memory-efficient implementation. In Secs. 2.7 and 2.8 we show how iterative aggregation/disaggregation methods can be employed to obtain a reduced Ca^{2+} release site model, through exact calculation of the required conditional probability distributions. In Sec. 2.9 we show how a fast/slow

reduced Ca^{2+} release site model can be used to efficiently compute puff/spark statistics, such as the probability distribution of the time required to achieve a specified number of refractory channels after a step increase in $[Ca^{2+}]$. Sec. 2.10 discusses limitations and possible extensions this approach to reduction of Ca^{2+} release site models.

2.3 Model Formulation

Stochastic models of single channel gating often take the form of continuous-time discrete-state Markov chains (for review see [Colquhoun and Hawkes, 1995, Smith, 2002]). For example, Fig. 2.1 shows the state-transition diagram for the 4-state Keizer-Levine RyR that includes both fast Ca^{2+} activation and slower Ca^{2+}-independent inactivation [Keizer and Levine, 1996]. Under the assumption that domain $[Ca^{2+}]$ changes are fast compared to channel transitions, this single channel model is continuous-time Markov chain with infinitesimal generator matrix $Q = (q_{ij})$ given by

$$
Q = \begin{bmatrix}
\diamond & k_{12}c_\infty^4 & 0 & 0 \\
k_{21} & \diamond & k_{23}\left(c_\infty + c_*\right)^3 & k_{24} \\
0 & k_{32} & \diamond & 0 \\
0 & k_{42} & 0 & \diamond
\end{bmatrix}
\tag{2.1}
$$

where the states have been ordered C_1, O_2, O_3, C_4. The off-diagonal entries of the Q-matrix for this irreducible and time-homogeneous Markov chain are transition rates defined by

$$
q_{ij} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr\left[S(t + \Delta t) = j | S(t) = i\right]
\tag{2.2}
$$

A

B
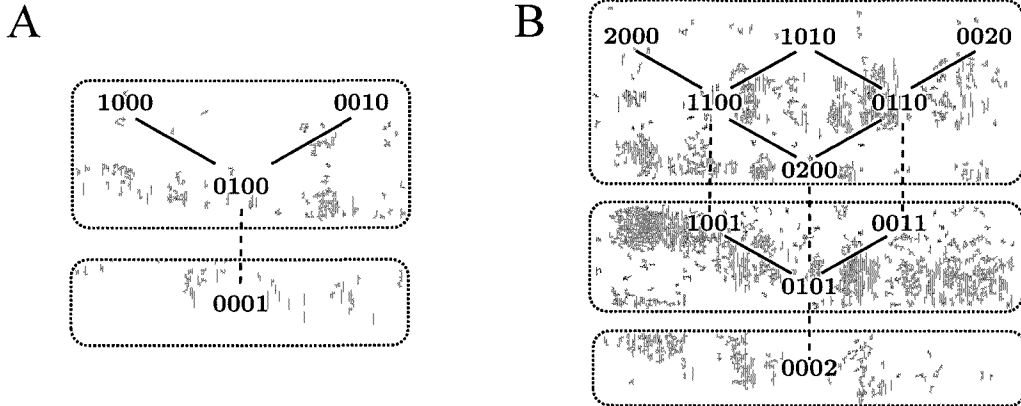
Figure 2.2: (A) Topology of the 4-state Keizer-Levine RyR model showing fast and slow transitions (solid and dotted lines, respectively). (B) Topology for the 10-state release site composed of 2 Keizer-Levine RyRs. Gray boxes indicate groups of states connected by fast transitions. The ordered M-tuples $(N_1, N_2, \ldots, N_M)$ satisfy $N_i \in \{0, \ldots, N\}$ and $\sum_i N_i = N$ where $N_i = n$ indicates n channels in state i.

where $i \neq j$ and $S(t) \in \{1, 2, 3, 4\}$ indicates the state of the stochastically gating channel at time t. The diamonds on the diagonal entries of the $\boldsymbol{Q}$-matrix indicate values leading to row sums of zero, $q_{ii} = -\sum_{j \neq i} q_{ij} < 0$. Note that the rate constants k_{24} (and k_{42}) for Ca^{2+}-independent inactivation of the RyR (and recovery from inactivation) have units of time^{-1}. The dissociation rate constants k_{21} and k_{32} also have units of time^{-1}. The association rate constants k_{23} and k_{12} have units of conc$^{-\eta}$ time^{-1} where $\eta = 3$ or 4 is the cooperativity for Ca^{2+} binding to the regulatory site of the channel. Consistent with the assumption of fast $[Ca^{2+}]$ changes, the background $[Ca^{2+}]$ denoted by c_∞ is used for the $C_1 \rightarrow O_2$ transition, while the concentration $c_\infty + c_*$ is used for the $O_2 \rightarrow O_3$ transition. The parameter c_* denotes the elevation over background Ca^{2+} experienced by the Ca^{2+} regulatory site of the channel when the channel is open.

Using the parameters of Fig. 2.1, $c_\infty = 0.1$ μM, and $c_* = 0.065$ μM, the equilibration rates for the three pairs of states in the Keizer-Levine RyR model are

20

$k_{12}c_\infty^4 + k_{21} = 28.9$ s^{-1} ($C_1 \leftrightarrow O_2$), $k_{23}\left(c_\infty + c_*\right)^3 + k_{32} = 393$ s^{-1} ($O_2 \leftrightarrow O_3$), and $k_{24} + k_{42} = 1.85$ s^{-1} ($O_2 \leftrightarrow C_4$). The solid lines of Fig. 2.2A correspond to the fast $C_1 \leftrightarrow O_2 \leftrightarrow O_3$ transitions in the Keizer-Levine RyR, while the dotted line corresponds to the slow $O_2 \leftrightarrow C_4$ transitions.

All of the statistical properties of the Keizer-Levine RyR can be calculated from its Q-matrix (Eq. 2.1). For example, the conditional probability of finding the channel in state j at time t provided it was in state i at time zero is

$$p_{ij}(t) = [e^{tQ}]_{ij} = \Pr\left[S\left(t\right) = j|S(0) = i\right], \tag{2.3}$$

where $t \geq 0$ and $[e^{tQ}]_{ij}$ indicates the element in the ith row and jth column of the matrix exponential. In fact, because the Markov chain is time homogeneous, $\Pr\left[S\left(t+s\right) = j|S(s) = i\right] = p_{ij}(t)$ for all $t \geq 0$ and $s \geq 0$.

The Ca^{2+} release site models that are the focus of this chapter involve N identical Keizer-Levine RyRs interacting via changes in local [Ca^{2+}] under the assumption of "instantaneous mean-field coupling" [Nguyen et al., 2005, DeRemigio and Smith, 2005, DeRemigio and Smith, 2008]. That is, we assume that the increase in local [Ca^{2+}] experienced by each channel is an instantaneous function of the number of open channels (N_O),

$$[\text{Ca}^{2+}](t) = c_\infty + c_* N_O(t). \tag{2.4}$$

Because identical channels coupled in this manner are indistinguishable, a release site composed of N M-state channels includes

$$\beta(N, M) = \binom{N + M - 1}{N} = \frac{(N + M - 1)!}{N!\,(M - 1)!} \tag{2.5}$$

distinct states. Each of the $\beta(N, M)$ states can be written as the ordered M-tuple

$(N_1, N_2, \ldots, N_M)$ where $N_i = n$ indicates n channels in state i, $N_i \in \{0, \ldots, N\}$, and $\sum_i N_i = N$. Fig. 2.2B uses this notation to illustrate the topology of a 10-state Ca^{2+} release site model composed of 2 coupled Keizer-Levine RyRs. In this case the states take the form $(N_{C_1}, N_{O_2}, N_{O_3}, N_{C_4})$ and, for example, the rate for the 2000 → 1100 transition is given by

$$N_{C_1} k_{12} \left(c_\infty + N_O c_* \right)^4 = 2k_{12} c_\infty^4$$

where $N_O = N_{O_2} + N_{O_3} = 0$ and $N_{C_1} = 2$ accounts for the fact that either one of the 2 channels can make a $C_1 \to O_2$ transition. Similarly, the rate for the 0110 → 0020 transition is given by

$$N_{O_2} k_{23} \left(c_\infty + N_O c_* \right)^3 = k_{23} \left(c_\infty + 2c_* \right)^3,$$

because $N_{O_2} = 1$ and $N_O = N_{O_2} + N_{O_3} = 2$. Consistent with Fig. 2.2A, the solid and dotted lines in Fig. 2.2B indicate those transitions associated with fast Ca^{2+}-dependent activation and slow Ca^{2+}-independent inactivation, respectively.

2.4   Representative calcium release site simulations

Fig. 2.3A shows the stochastic dynamics of a Ca^{2+} release site composed of eight identical Keizer-Levine RyRs coupled in the fashion described in Sec. 2.3. In each of the three simulations shown, the single channel model parameters follow Fig. 2.1, the background $[Ca^{2+}]$ is $c_\infty = 0.1$ μM, and simulations are performed using the exact numerical method attributed to Gillespie [Gillespie, 1977]. When the coupling strength c_* is relatively small (0.06 μM, top panel), increases in the number of open channels usually involve one or a few Ca^{2+} channels, reminiscent of the experimentally observed phenomena of Ca^{2+} blips and quarks [Sun et al., 1998, Niggli, 1999].
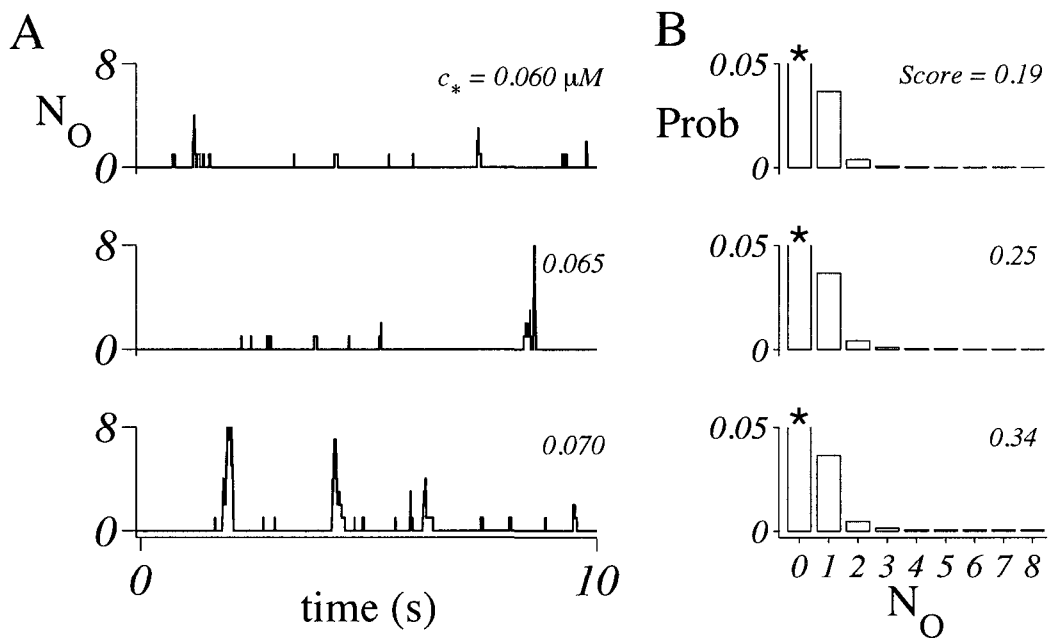
Figure 2.3: A: Representative Ca^{2+} release site simulations involving eight Keizer-Levine RyRs instantaneously coupled via a domain $[Ca^{2+}]$ given by $c = c_\infty + N_O c_*$ where $c_\infty = 0.1$ μM and $c_* = 0.06$ (top), 0.065 (middle), and 0.07 μM (bottom). When the coupling strength c_* is sufficiently large, the stochastic dynamics of the number of open channels at a release site (N_O) is reminiscent of puffs/sparks. B: Probability distribution of the number of open channels directly calculated from the generator matrix of the Ca^{2+} release site Markov chain models and the corresponding puff/spark *Score* of 0.19 (top), 0.25 (middle), and 0.34 (bottom). Asterisks indicate truncated bar for $Pr\,[N_O = 0] = 0.9576$ (top), 0.9561 (middle), and 0.9537 (bottom).

However, when the coupling strength is increased to $c_* = 0.065$ and 0.07 μM (middle and bottom panels), the stochastic dynamics of the number of open channels at a release site (N_O) becomes more robust and concerted. These events often involve a significant fraction of the channels at the release site. Event durations (100–300 ms) and inter-event intervals (20–50 s) are similar to the experimentally observed localized Ca^{2+} elevations known as Ca^{2+} puffs and sparks.

Fig. 2.3B shows the steady-state probability distribution of the number of open channels at these simulated Ca^{2+} release sites, that is, $Pr\,[N_O = n]$ where $n \in \{0, 1, \ldots, N\}$.

Note that these distributions are not estimated via Monte Carlo simulation, but rather directly calculated from the stationary distribution of the 165-state expanded Markov chain corresponding to 8 coupled Keizer-Levine RyRs ($165 = \beta(8, 4)$ in Eq. 2.5). That is, after constructing the Q matrix for the Ca^{2+} release site model, we numerically solve

$$\pi Q = 0 \quad \text{subject to} \quad \pi e = 1, \tag{2.6}$$

where Q is a 165×165 matrix, π is 1×165 row vector, and e is a 165×1 column vector of ones (see Sec. 2.11.1). Each element of the probability distribution of the number of open channels ($\Pr[N_O = n]$) is then constructed as the sum of the appropriate elements of π. Note that for the different values of the coupling strength used in Fig. 2.3B, only subtle differences in the probability distribution of N_O are visible. On the other hand, the presence or absence of puff/sparks in Ca^{2+} release site simulations such as Fig. 2.3A can be accessed from $\Pr[N_O = n]$ without recourse to Monte Carlo simulation using a response measure dubbed the puff/spark *Score* [Nguyen et al., 2005],

$$Score = \frac{\text{Var}[f_O]}{\text{E}[f_O]} = \frac{1}{N} \frac{\text{Var}[N_O]}{\text{E}[N_O]}, \tag{2.7}$$

where $f_O = N_O/N$ is the fraction of open channels. The puff/spark *Score* takes values between 0 and 1, and a *Score* of greater than approximately 0.25 indicates the presence of robust stochastic Ca^{2+} excitability (as in the middle and bottom panels of Fig. 2.3B).

In the Ca^{2+} release site model composed of 8 Keizer-Levine RyRs (Fig. 2.3), higher values of the Ca^{2+} coupling strength ($c_* > 0.1\ \mu M$) lead to sparks with physiologically unrealistic duration and ultimately a tonically active release site with low puff/spark *Score* ($c_* > 0.4\ \mu M$, not shown). Of course, release site simulations using a different number of channels (N) or a different single channel model lead to results distinct from

the representative simulations of Fig. 2.3. Such modeling has played an important role in understanding the relationship between the single channel gating of intracellular Ca^{2+} channels and the stochastic dynamics of Ca^{2+} puffs and sparks (for review see [Groff and Smith, 2008b]).

2.5 Fast/slow reduction for calcium release site models

In the context of ODE modeling of whole cell Ca^{2+} responses, the Keizer-Levine RyR model was reduced from 4 to 2 states by observing that transition rates between the dis-inactviated states (C_1, O_2, and O_3) are much faster than the transition rates to and from the inactivated state C_2 [Keizer and Levine, 1996]. Similarly, the 4-state Markov chain of Eq. 2.1 can be reduced to a 2-state model,

$$\text{(disinact) } C_1 \cup O_2 \cup O_3 \underset{\hat{q}_{21}}{\overset{\hat{q}_{12}}{\rightleftharpoons}} C_4 \text{ (inact)} \tag{2.8}$$
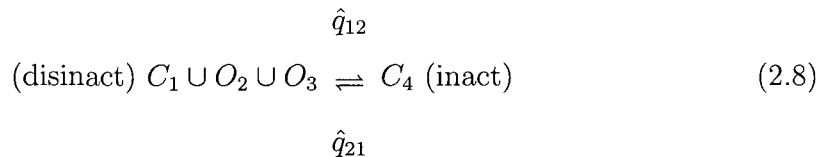
where $C_1 \cup O_2 \cup O_3$ indicates the dis-inactivated macrostate. While the transition rate from the inactivated state to the dis-inactivated macrostate in the reduced model can be "read off" the full model ($\hat{q}_{21} = q_{42}$, see Fig. 2.1), determining the transition rate from the dis-inactivated macrostate to the inactivated state ($\hat{q}_{12}$) requires an estimate of the steady-state conditional probability of being in state O_2 given that the channel is in $C_1 \cup O_2 \cup O_3$, because the product of this conditional probability and q_{24} gives rate of inactivation in the reduced model. Under the assumption of rapid mixing of dis-inactivated states, this conditional probability can be found using Hill's
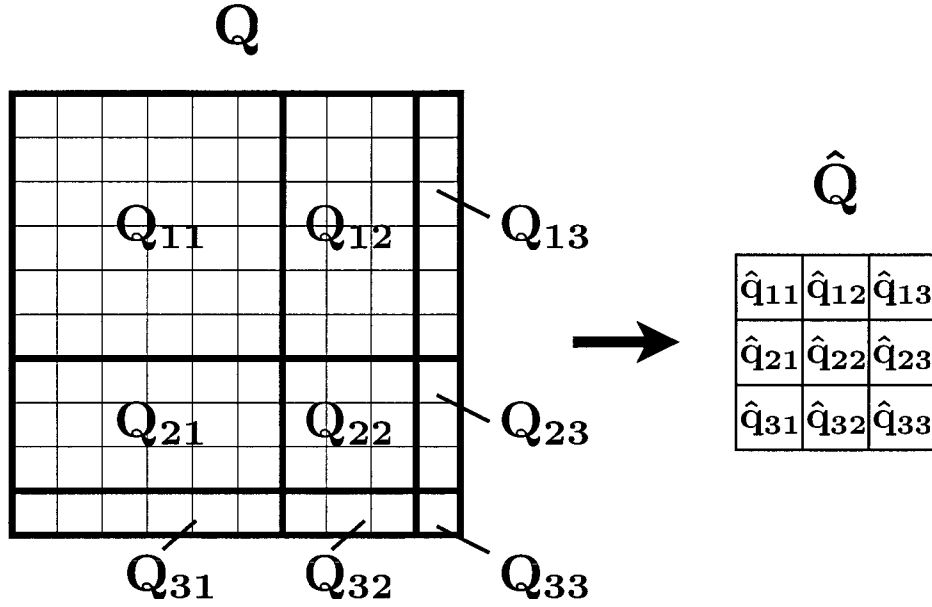
Figure 2.4: Partition and contraction of the 2-channel Keizer-Levine release site transition matrix. The transition matrix of the original model (left) is partitioned into blocks corresponds to the groups formed by classifying fast and slow transitions.

diagrammatic method [Hill, 1989] applied to the subgraph $C_1 \leftrightarrow O_2 \leftrightarrow O_3$ resulting in the expression

$$\Pr\left[O_2|C_1 \cup O_2 \cup O_3\right] = \frac{q_{12}q_{32}}{q_{21}q_{32} + q_{12}q_{32} + q_{12}q_{23}}.$$

Thus,

$$\hat{q}_{12} = q_{24}\frac{q_{12}q_{32}}{q_{21}q_{32} + q_{12}q_{32} + q_{12}q_{23}} \tag{2.9}$$

is the required transition rate for dis-inactivation in the reduced Keizer-Levine RyR (Eq. 2.8). In the reduced single channel model, the open probability conditioned on occupation of the dis-inactivated macrostate is

$$\Pr\left[O|C_1 \cup O_2 \cup O_3\right] = \frac{q_{12}q_{32} + q_{12}q_{23}}{q_{21}q_{32} + q_{12}q_{32} + q_{12}q_{23}}$$

where $O = O_2 \cup O_3$, while the open probability conditioned on occupation of the inactivated state is zero.

Fast/slow reduction for Ca^{2+} release sites composed of several channels can be illustrated by considering $N = 2$ Keizer-Levine RyRs coupled via a common domain $[Ca^{2+}]$. As discussed in Sec. 2.4, we assume $[Ca^{2+}](t) = c_\infty + c_* N_O(t)$ where $N_O(t)$ is the number of open channels (0, 1, or 2). Fig. 2.2B shows the transition state diagram for 2 coupled Keizer-Levine RyRs where each release site state is labeled by four digits $n_1 n_2 n_3 n_4$ with $n_i \in \{0, 1, 2\}$ and $\sum_i n_i = 2$. As mentioned above, the solid lines correspond to fast $C_1 \leftrightarrow O_2 \leftrightarrow O_3$ transitions, while the dotted lines correspond to slow $O_2 \leftrightarrow C_4$ transitions. The gray boxes of Fig. 2.2B indicate groups of states connected by fast transitions that are good candidates for lumping during a fast/slow reduction procedure that will result in a 3-state Ca^{2+} release site model.

As illustrated in Fig. 2.4, the fast/slow reduction procedure begins by constructing the Q-matrix for 2 coupled channels consistent with the partitioning in Fig. 2.2B. The resulting matrix takes the form

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \tag{2.10}$$

where block Q_{11} is 6×6, block Q_{22} is 3×3, and block Q_{33} is 1×1 (see Fig. 2.4). To perform the model reduction, we require an estimate of the conditional probability of being in the various substates of each block. Under the assumption of rapid mixing within lumped states, these conditional probability distributions are well-approximated by the solutions of the linear systems,

$$\hat{\pi}_i Q_{ii}^+ = 0 \quad \text{subject to} \quad \hat{\pi}_i e_i = 1, \tag{2.11}$$

27

where $i \in \{1, 2, 3\}$. In this expression, $\boldsymbol{Q}_{ii}^{+}$ is given by

$$\boldsymbol{Q}_{ii}^{+} = \boldsymbol{Q}_{ii} + \text{diag}\left(\sum_{j \neq i} \boldsymbol{Q}_{ij} \boldsymbol{e}_j\right), \tag{2.12}$$

where the sum is over two column vectors, the "diag" operation converts the resulting column vector into a diagonal matrix commensurate with $\boldsymbol{Q}_{ii}$, the unknowns $\hat{\boldsymbol{\pi}}_1$, $\hat{\boldsymbol{\pi}}_2$, and $\hat{\boldsymbol{\pi}}_3$ are 1×6, 1×3, and 1×1, respectively, and the $\boldsymbol{e}_j$ are commensurate column vectors of ones. The approximate conditional probability distributions $\hat{\boldsymbol{\pi}}_i$ are then used to calculate the transition rates between lumped states yielding the reduced model

$$\hat{\boldsymbol{Q}} = \begin{bmatrix} \hat{q}_{11} & \hat{q}_{12} & \hat{q}_{13} \\ \hat{q}_{21} & \hat{q}_{22} & \hat{q}_{23} \\ \hat{q}_{31} & \hat{q}_{32} & \hat{q}_{33} \end{bmatrix}, \tag{2.13}$$

where

$$\hat{q}_{ij} = \hat{\boldsymbol{\pi}}_i \boldsymbol{Q}_{ij} \boldsymbol{e}_j \tag{2.14}$$

for $i \neq j$ and $\hat{q}_{ii} = \sum_{j \neq i} -\hat{q}_{ij}$. Pseudocode for this fast/slow reduction procedure is presented in Algorithm 1.

Algorithm 1: Fast/slow reduction

---

require: $\hat{b}^2$ matrices $\{Q_{ij}\}$ where $\hat{b}$ is size of reduced model

for $i = 1, \ldots, \hat{b}$

$\quad Q_{ii}^+ \leftarrow Q_{ii} + \text{diag}\left\{\sum_{j \neq i} Q_{ij} e_j\right\}$

$\quad$ solve $\hat{\pi}_i Q_{ii}^+ = 0$ subject to $\hat{\pi}_i e_i = 1$

endfor

for $i = 1, \ldots, \hat{b}$

$\quad$ **for** $j = 1, \ldots, i-1, i+1, \ldots, \hat{b}$ **do**

$\quad\quad \hat{q}_{ij} \leftarrow \hat{\pi}_i Q_{ij} e_j$

$\quad$ **endfor**

$\quad \hat{q}_{ii} \leftarrow \sum_{j \neq i} -\hat{q}_{ij}$

endfor

return $\hat{Q} = (\hat{q}_{ij})$

2.6 Validation of fast/slow reduction for release sites

This section validates the numerical approach to fast/slow reduction outlined in Sec. 2.5 using a release site model composed of 8 4-state Keizer-Levine RyRs. Mean-field coupling of these channels leads to a 165×165 Q-matrix (cf. Eq. 2.10) that is partitioned into 81 blocks when states C_1, O_2, and O_3 are lumped. The 9 square blocks on the diagonal of the partitioned generator matrix are of size 45, 36, 28, 21, 15, 10, 6, 3, and 1 (see Sec. 2.11.2). The fast/slow reduction procedure outlined in Algorithm 1 leads to a reduced model specified by the 9×9 matrix $\hat{Q}$ (cf. Eq. 2.13).

Perhaps the most straightforward way to validate this approach is to compare
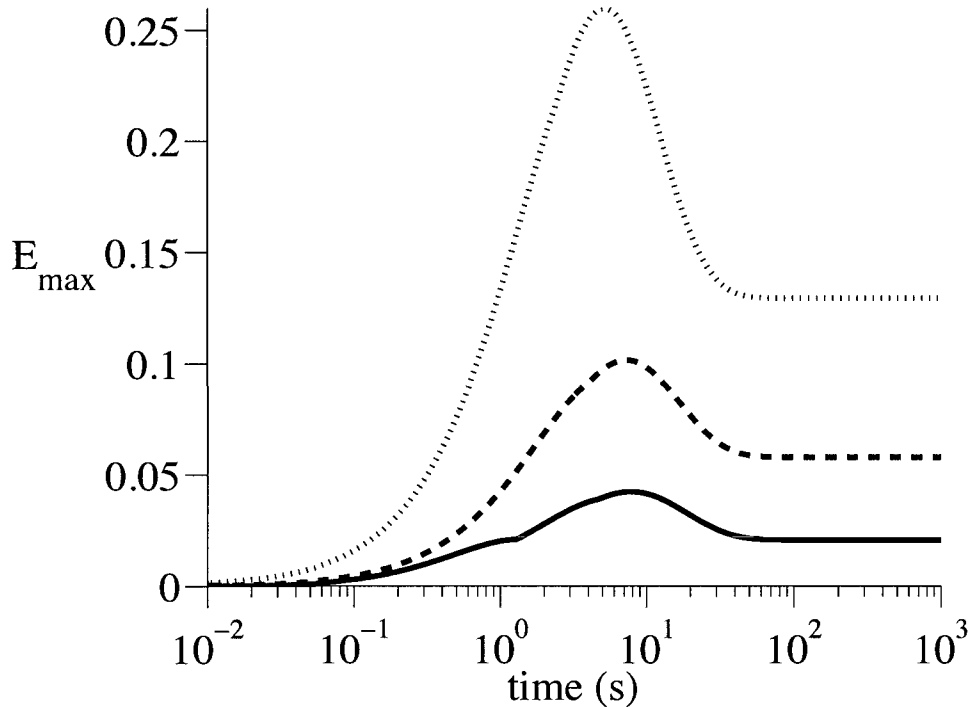
Figure 2.5: Error of fast/slow reduction (Algorithm 1) for a release site composed of eight 4-state Keizer-Levine RyRs quantified as in Eqs. 2.15–2.19. Solid, dashed and dotted lines use Ca^{2+} coupling strengths of $c_* = 0.06$, 0.065, and 0.07 μM, respectively (cf. Fig. 2.3). Background $[Ca^{2+}]$ is $c_\infty = 0.1$ μM and other parameters are as in Fig. 2.1.

the transition probability matrices of the reduced model ($\hat{P} = e^{t\hat{Q}}$) to the transition probability matrix of the full model ($P = e^{tQ}$) , see Eq. 2.3. Assuming the full and reduced models have b and $\hat{b}$ states, respectively, we write

$$\hat{E}(t) = \hat{P}(t) - UP(t)V \tag{2.15}$$

where V is a $b \times \hat{b}$ collector matrix [Nicola, 1998],

$$
V = \begin{bmatrix} e_1 & 0 & \cdots & 0 \\ 0 & e_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e_{\hat{b}} \end{bmatrix},
$$

and the e_i are column vectors of ones with lengths commensurate with Q_{ii}, and U is a $\hat{b} \times b$ distributor matrix given by

$$
U = \begin{bmatrix} \bar{\pi}_1 & 0 & \cdots & 0 \\ 0 & \bar{\pi}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \bar{\pi}_{\hat{b}} \end{bmatrix}. \tag{2.16}
$$

The exact conditional probability distributions $\bar{\pi}_i$ that compose U are row vectors given by

$$
\bar{\pi}_i = \frac{\pi_i}{\pi_i e_i}, \tag{2.17}
$$

where

$$
\pi = [\pi_1, \pi_2, \cdots, \pi_{\hat{b}}] \tag{2.18}
$$

is the conformally partitioned exact stationary distribution of the full model satisfying Eq. 2.6.

The solid line of Fig. 2.5 shows the maximum absolute error,

$$
E_{\max}(t) = \max_{ij} \left| \hat{E}_{ij}(t) \right|, \tag{2.19}
$$

for a 9-state fast/slow reduced Ca^{2+} release site model obtained by contracting a full

31
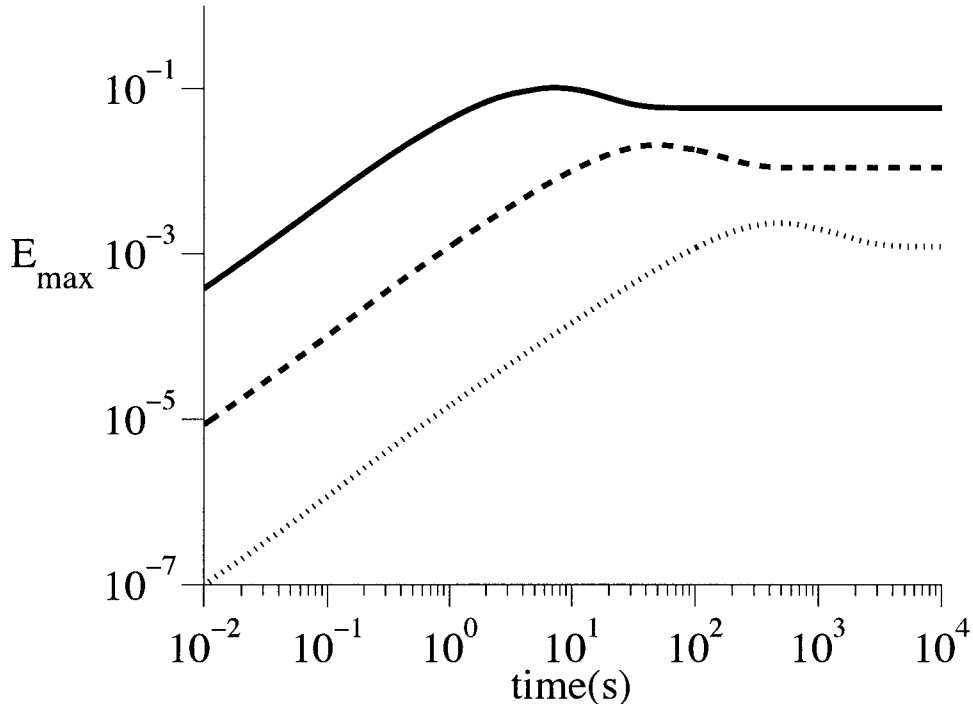
Figure 2.6: Logarithmic plot of the error of fast/slow reduction (Eq. 2.19) when Algorithm 1 is applied to a release sites composed of 8 4-state Keizer-Levine RyRs with parameters as in Fig. 2.5 (solid line). Dashed and dotted lines show the error of fast/slow reduction is decreased when slow transition rates are decreased by 10 and 100×, respectively. Other parameters as in Fig. 2.5.

model with eight 4-state Keizer-Levine RyRs and a coupling strength of $c_* = 0.06$ μM (as in the top panel of Fig. 2.3A). For small values of t both $\hat{\boldsymbol{P}}$ and $\boldsymbol{P}$ are approximated by identity matrices and consequently $E_{max}(t) \approx 0$. Note that $E_{max}(t)$ reaches a peak of 0.05 at $t \approx 10$ s and approaches a limiting value of 0.02 as $t \to \infty$, a value that corresponds to the maximum absolute error of the stationary distribution of the reduced model when compared to the contracted stationary distribution of the full model. (To see this, recall that the columns of $\lim_{t\to\infty} \boldsymbol{P}(t)$ are identical and each row is given by the elements of the stationary probability distribution for the full model that satisfies Eq. 2.6). The total absolute error of the stationary distribution

32
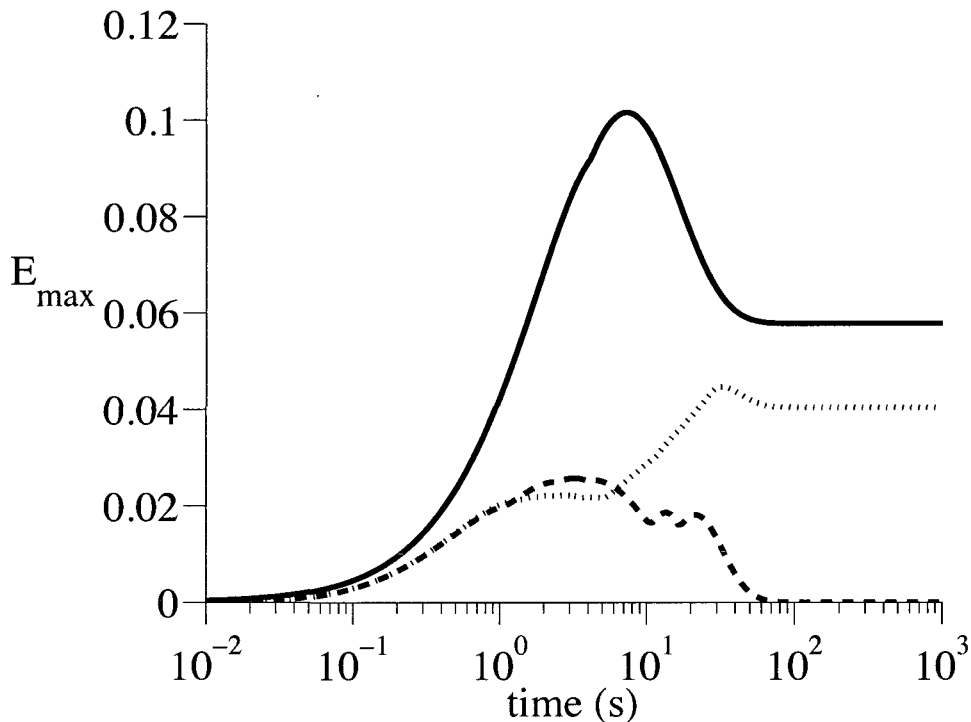
Figure 2.7: Error of fast/slow reduction (Eq. 2.19) when Algorithms 1 (solid line), 2 (dotted line), and 3 (dashed line) are applied to a release site composed of eight 4-state Keizer-Levine RyRs. Parameters: $c_\infty = 0.1$ μM, $c_* = 0.065$ μM, and as in Fig. 2.1.

of the fast/slow reduced model is $\sum_j |\hat{E}_{ij}(\infty)| \approx 0.047$.

The dotted and dashed lines of Fig. 2.5 show $E_{\max}(t)$ for the fast/slow reduced model when the coupling strength is increased to $c_* = 0.065$ and 0.07 μM (as in the middle and bottom panels of Fig. 2.3A). Stochastic Ca^{2+} excitability is more pronounced and the puff/spark *Score* increases for these values of c_* (see Fig. 2.3B) and both the peak (0.10 and 0.26) and steady state (0.06 and 0.13) errors show a corresponding increase. Perhaps more importantly, Fig. 2.6 repeats this analysis using the standard value of the Ca^{2+} coupling strength ($c_* = 0.065$ μM) and modified parameter sets for the Keizer-Levine RyR model in which the rate of the slow transitions (k_{24}

and k_{42}) is decreased by 10 and 100× (dashed and dotted lines, respectively). Note that $E_{\max}(t)$ decreases as the separation of time scales between Ca^{2+}-dependent activation and Ca^{2+}-independent inactivation increases, thereby validating the fast/slow reduction procedure of Algorithm 1.

Because Fig. 2.5 indicates significant model reduction error, we considered alternative fast/slow reduction procedures that follow a solution method for nearly completely decomposable Markov chains presented in Stewart's monograph [Stewart, 1994, pages 285–294]. This approach is distinct from Algorithm 1 in that the diagonal elements of the diagonal blocks $\boldsymbol{Q}_{u}$ of the partitioned generator matrix are not adjusted to remove negative entries corresponding to slow transitions between lumped states (Eq. 2.11). Because the transition rates between macrostates are slow, this is a subtle difference. Nevertheless, Fig. 2.7 shows a decreased model reduction error using this modified fast/slow reduction procedure (Algorithm 2, dotted line) compared to the previously discussed method (Algorithm 1, solid line). Note that an important step in Algorithm 2 involves solving for the Perron vector of $\boldsymbol{P}_{u}$, a substochastic matrix given by $\boldsymbol{P}_{u} = \boldsymbol{I} + \boldsymbol{Q}_{u}/\delta$ for suitable δ. The Perron vector $\boldsymbol{u}_{i}$ solves $\boldsymbol{u}_{i}\boldsymbol{P}_{u} = \lambda\boldsymbol{u}_{i}$ subject to $\boldsymbol{u}_{i}\boldsymbol{e}_{i} = 1$ where λ is the spectral radius of $\boldsymbol{P}_{u}$ (see Sec. 2.11.1).

Algorithm 2: Modified fast/slow reduction

---

require: $\hat{b}^2$ matrices $\{Q_{ij}\}$

for $i = 1, \ldots, \hat{b}$

    $\delta \leftarrow \max_k |Q_{ii}(k, k)|$

    $P_{ii} \leftarrow I + Q_{ii}/\delta$

    $u_i \leftarrow$ the Perron vector of P_{ii}

endfor

for $i = 1, \ldots, \hat{b}$

    for $j = 1, \ldots, i - 1, i + 1, \ldots, \hat{b}$

        $\hat{q}_{ij} \leftarrow \hat{\pi}_i Q_{ij} e_j$

    endfor

    $\hat{q}_{ii} \leftarrow \sum_{j \neq i} -\hat{q}_{ij}$

endfor

return $\hat{Q} = (\hat{q}_{ij})$

Algorithm 3: Gold standard reduction with substantial

storage requirement

---

require: $\hat{b}^2$ matrices $\{Q_{ij}\}$

solve $\pi Q = 0$ subject to $\pi e = 1$ where $Q = (Q_{ij})$

for $i = 1, \ldots, \hat{b}$

    $\bar{\pi}_i \leftarrow \pi_i / \pi_i e_i$ where $\pi = [\pi_1, \pi_2, \ldots, \pi_{\hat{b}}]$

endfor

for $i = 1, \ldots, \hat{b}$

    for $j = 1, \ldots, i-1, i+1, \ldots, \hat{b}$

        $\hat{q}_{ij} \leftarrow \bar{\pi}_i Q_{ij} e_j$

    endfor

    $\hat{q}_{ii} \leftarrow \sum_{j \neq i} -\hat{q}_{ij}$

endfor

return $\hat{Q} = (\hat{q}_{ij})$

2.7 Reduction using correct conditional probability

As discussed in Sec. 2.6, the reduction error obtained using both the original and modified fast/slow reduction methods (Algorithms 1 and 2) is initially zero and asymptotically approaches a finite value as $t \to \infty$ (solid and dotted lines of Fig. 2.7). As expected, inspection of numerical results associated with Figs. 2.5–2.7 confirms that the reduction error is larger when the conditional probability distributions estimated in a block-by-block fashion by Algorithms 1 and 2 become less accurate (not shown). That is, the vector norms $\|\hat{\pi}_i - \bar{\pi}_i\|$—with $\hat{\pi}_i$ and $\bar{\pi}_i$ given by Eq. 2.11 and Eq. 2.17, respectively—are larger when Algorithms 1 and 2 are not performing well. Thus, the

error present in the fast/slow reduction approach is potentially avoidable, provided a better approximation of the conditional probability distributions can be obtained.

Eq. 2.5 indicates that a Ca^{2+} release site model composed of eight 4-state Keizer-Levine RyRs includes $\beta(8,4) = 165$ distinguishable states. For this relatively small release site model, the exact conditional probability distributions $\bar{\pi}_i$ can be calculated using Eqs. 2.17 and 2.18, because the numerical solution of the stationary distribution of the full problem is tractable (π, Eq. 2.6). In this case the rate constants for the reduced model are given by $\hat{q}_{ij} = \bar{\pi}_i \boldsymbol{Q}_{ij} \boldsymbol{e}_j$ for $i \neq j$ (cf. Eq. 2.14). For any given partitioning of states—i.e., the $\hat{b}^2$ matrices $\{\boldsymbol{Q}_{ij}\}$—the reduced model thus obtained will be referred to as the "gold standard," because the conditional probability distributions used to perform the reduction are exactly calculated. While this reduction may not be optimal, the fact that $\hat{\boldsymbol{P}}(\infty) = \boldsymbol{U}\boldsymbol{P}(\infty)\boldsymbol{V}$ (cf. Eq. 3.25) means that the error of the gold standard reduced model does at least approach zero as $t \to \infty$. The dashed lines of Fig. 2.7 show how this important feature of the gold standard reduced model (Algorithm 3) leads to finite integrated error, which is not a property of the other reductions. In addition, the peak value of $\hat{E}_{\max}$ obtained (0.03) is significantly smaller than the results of Algorithms 1 and 2 (0.10 and 0.05, respectively).

Because Algorithm 3 uses the exact conditional probability distributions $\bar{\pi}_i$ (Eq. 2.17), its reduction error—the dashed line of Fig. 2.7—indicates that the time scales of Ca^{2+}-dependent activation and Ca^{2+}-independent inactivation in the release site model are not completely separated. Fig. 2.8 shows that when this "gold standard" reduction procedure is repeated using modified parameter sets for the Keizer-Levine RyR model in which the rate of the slow transitions (k_{24} and k_{42}) is decreased by 10 and $100\times$, the peak error decreases from 0.03 to 5.7×10^{-3} and 6.6×10^{-4}, respectively (cf. Fig. 2.6).
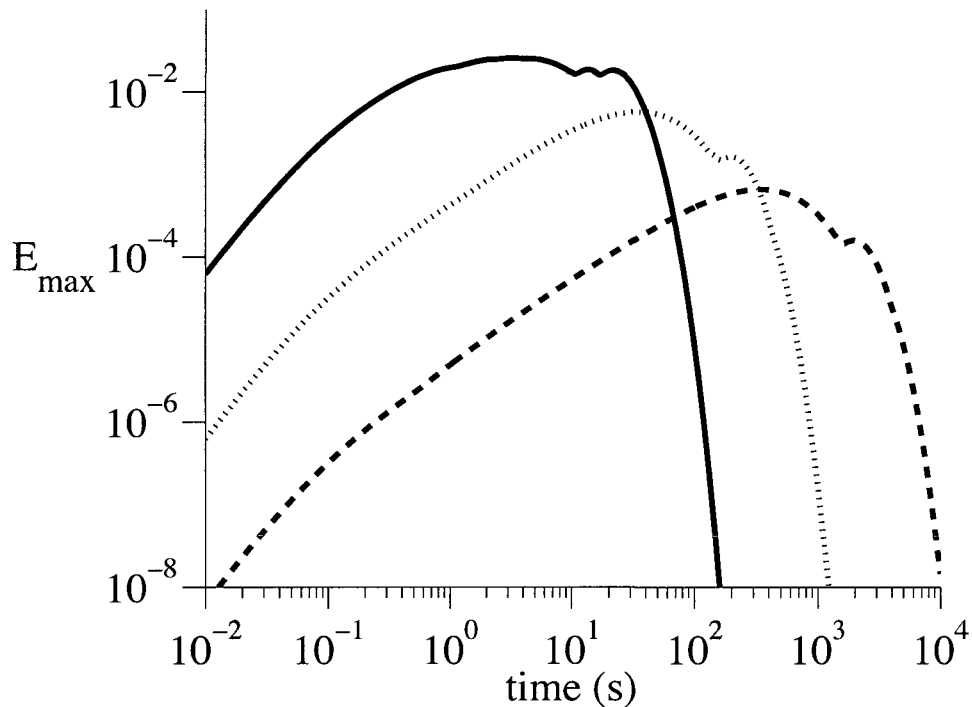
Figure 2.8: Logarithmic plot of the error of fast/slow reduction (Eq. 2.19) when Algorithm 3 is applied to a release sites composed of eight 4-state Keizer-Levine RyRs with parameters as in Fig. 2.5 (solid line). Dotted and dashed lines show a decreased error when the rate of slow transitions is decreased by 10 and 100×, respectively.

2.8 Iterative aggregation/disaggregation methods

Using Ca^{2+} release sites composed of a small number of channels, Sec. 2.7 showed that model reduction using exact conditional probability distributions ($\bar{\pi}$, Algorithm 3) is superior to fast/slow reduction procedures that use approximate conditional probability distributions ($\hat{\pi}$, Algorithms 1 and 2). On the other hand, the storage requirements of Algorithm 3 are far in excess of Algorithms 1 and 2. (Recall that Algorithm 3 solves for the full model stationary distribution (Eq. 2.6), Algorithm 1 sequentially solves for the stationary distributions of the various blocks of the partitioned generator matrix of the full model (Eq. 2.11), and Algorithm 2 sequentially solves for

38

the Perron vectors of $P_{ii} = I + Q_{ii}/\delta$.) Indeed, the substantial storage requirements of Algorithm 3 make it inappropriate as a fast/slow reduction procedure for Ca^{2+} release sites with a large number of states.

Iterative aggregation/disaggregation (IAD) methods are a well-known alternative to direct methods for calculating the stationary distribution of large Markov chains [Stewart, 1994]. Because these methods often perform well when a Markov chain is irreducible and nearly completely decomposable, we implemented a memory-efficient version of Algorithm 3 that solves for the stationary distribution of the full model using the Koury-McAllister-Stewart IAD method (see Algorithm 4) [Koury et al., 1984]. For comparison, we also implemented release site reduction procedures that utilize the IAD methods of Vantilborgh and Takahashi (algorithms not shown) [Cao and Stewart, 1985].

Table 2.1 shows the number of iterations required for convergence of the Koury-McAllister-Stewart and Vantilborgh algorithms for Ca^{2+} release sites composed of up to 80 4-state Keizer-Levine RyRs when $c_\infty = 0.1$ μM and $c_* = 0.06$ μM. The residuals given by $\|\pi Q\|_1$ calculated in a block-by-block fashion from $\{\pi_i\}$ and $\{Q_{ij}\}$ are also shown. Small residuals indicate convergence of the IAD methods to the correct stationary probability distribution π, yielding the exact conditional probability distributions $\bar{\pi}_i$ (Eq. 2.18), and a "gold standard" reduced model $\hat{Q} = (\hat{q}_{ij})$ where $\hat{q}_{ij} = \bar{\pi}_i Q_{ij} e_j$ for $i \neq j$. Our implementation of the Takahashi IAD method was less successful than the Koury-McAllister-Stewart and Vantilborgh methods and did not converge for $N \geq 30$ (not shown).

Table 2.1 shows that the number of iterations required for the Koury-McAllister-Stewart and Vantilborgh IAD methods first increases and then decreases as a function of N, presumably reflecting the fact that the Ca^{2+} release site dynamics change significantly when N is increased with fixed c_* (note that the puff/spark *Score* increases

and decreases in a similar fashion). In fact, for $N \geq 50$ the low puff/spark *Scores* in Table 2.1 reflect tonically active Ca^{2+} release sites.

To ensure that the success of model reduction using the Koury-McAllister-Stewart method for large N is not dependent on the release sites being tonically active, benchmark calculations were repeated using c_* values selected to ensure that the full model exhibited robust Ca^{2+} excitability (*Score* > 0.25). Using these parameters, Table 2.2 demonstrates successful release site reduction using the Koury-McAllister-Stewart method (Algorithm 4) with up to 80 Keizer-Levine RyRs. While the number of iterations required for convergence depends on the Ca^{2+} coupling strength, the residuals are consistently small.

In both Table 2.1 and 2.2, the N+1-state reduced Ca^{2+} release site models are contractions of full models with $\beta(N, 4)$ states (Eq. 2.5). The largest Ca^{2+} release site model successfully reduced using the Koury-McAllister-Stewart IAD method (see Algorithm 4) included $\beta(80, 4) = 91881$ states and $2 \cdot 3\beta(80, 3) = 531360$ transitions, where 3 corresponds to the number of edges in the state-transition diagram for the Keizer-Levine RyR (Fig. 2.2A) and $3\beta(80, 3)$ is the number of edges in state-transition diagram of the 80-RyR Ca^{2+} release site (cf. Fig. 2.2B).

Algorithm 4: Reduction using Koury-McAllister-Stewart IAD

require: $\hat{b}^2$ matrices $\{\boldsymbol{Q}_{ij}\}$ and *tolerance*

$\delta \leftarrow \max_i \max_k |\boldsymbol{Q}_{ii}(k,k)|$

for $i = 1, \ldots, \hat{b}$

    $\boldsymbol{P}_{ii} \leftarrow \boldsymbol{I} + \boldsymbol{Q}_{ii}/\delta$

    $\boldsymbol{y}_i^{(0)} \leftarrow$ row vector of $1/b$ commensurate with $\boldsymbol{P}_{ii}$

    for $j = 1, \ldots, i-1, i+1, \ldots, \hat{b}$

        $\boldsymbol{P}_{ij} \leftarrow \boldsymbol{Q}_{ij}/\delta$

    endfor

endfor

$\boldsymbol{y}^{(0)} \leftarrow \left[\boldsymbol{y}_1^{(0)}, \boldsymbol{y}_2^{(0)}, \ldots, \boldsymbol{y}_{\hat{b}}^{(0)} \right]$, $m \leftarrow 0$, *change* $\leftarrow \infty$

while *change* > *tolerance*

    $m \leftarrow m + 1$

    for $i = 1, \ldots, \hat{b}$ **do**

        $\tilde{\boldsymbol{y}}_i^{(m-1)} = \boldsymbol{y}_i^{(m-1)} / \|\boldsymbol{y}_i^{(m-1)}\|_1$

    endfor

    for $i = 1, \ldots, \hat{b}$ **do**

        for $j = 1, \ldots, \hat{b}$ **do**

            $\boldsymbol{A}^{(m-1)}(i,j) = \tilde{\boldsymbol{y}}_i^{(m-1)} \boldsymbol{P}_{ij} \boldsymbol{e}_j$

        endfor

    endfor

    solve $\boldsymbol{w}^{(m-1)}(\boldsymbol{A}^{(m-1)} - \boldsymbol{I}) = \boldsymbol{0}$ where $\|\boldsymbol{w}^{(m-1)}\|_1 = 1$

    $\boldsymbol{z}^{(m)} \leftarrow \left[w_1^{(m-1)} \tilde{\boldsymbol{y}}_1^{(m-1)}, \ldots, w_{\hat{b}}^{(m-1)} \tilde{\boldsymbol{y}}_{\hat{b}}^{(m-1)} \right]$

    for $j = \hat{b}, \hat{b} - 1, \ldots, 1$ **do**

        solve $\boldsymbol{y}_j^{(m)} = \boldsymbol{y}_j^{(m)} \boldsymbol{P}_{jj} + \sum_{i<j} \boldsymbol{z}_i^{(m)} \boldsymbol{P}_{ij} + \sum_{i>j} \boldsymbol{y}_i^{(m)} \boldsymbol{P}_{ij}$

    endfor

    change $\leftarrow \|\boldsymbol{y}^{(m)} - \boldsymbol{y}^{(m-1)}\|_1$

endwhile

return $\hat{\boldsymbol{Q}} \leftarrow \delta(\boldsymbol{A}^{(m-1)} - \boldsymbol{I})$ and $\boldsymbol{\pi} \leftarrow \boldsymbol{y}^{(m-1)}$

2.9 Example direct calculations using fast/slow reduction
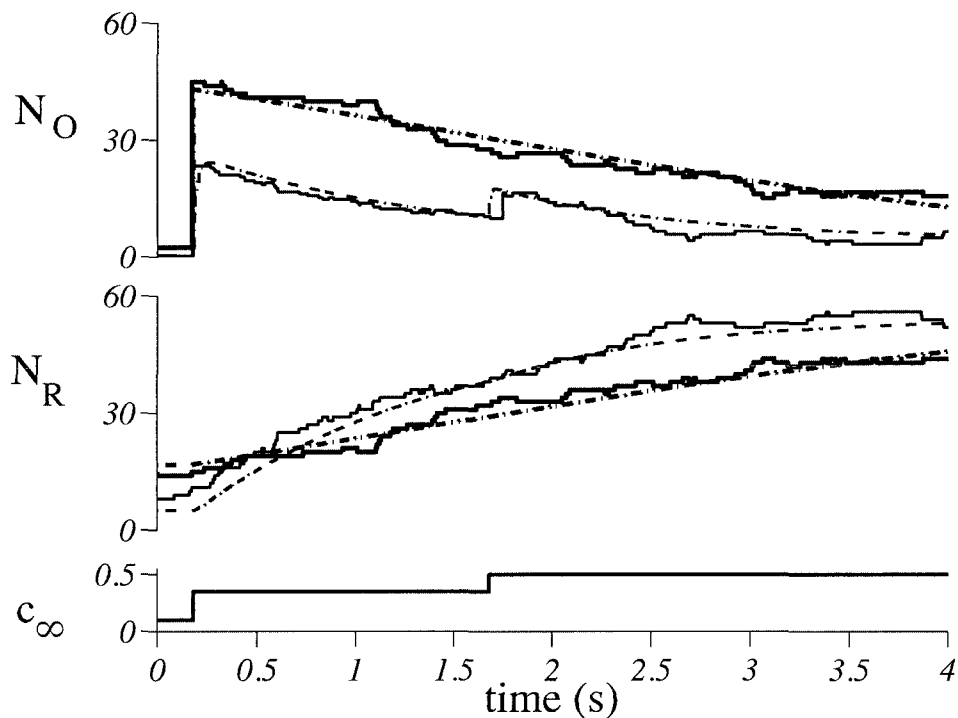
Figure 2.9: Monte Carlo simulation (solid lines) and direct calculation (broken lines) of the number of open (N_O) and refractory (N_R) channels in a stochastic simulation of a Ca^{2+} release site composed of 60 Keizer-Levine RyRs that are either independently gating (thin lines, $c_* = 0$) or coupled (thick lines, $c_* = 0.02$ μM) following an increase of the background $[Ca^{2+}]$ from $c_\infty = 0.1$ to 0.35 and subsequently to 0.5 μM (bottom panel).

As mentioned in Sec. 4.2, automated fast/slow reduction techniques are of interest because they may facilitate studies of Ca^{2+} release site dynamics that would otherwise be intractable due to the state-space explosion that occurs when multiple single channel models are coupled; below we illustrate this point. The thin solid lines of Fig. 2.9 show the number of open (N_O) and refractory (N_R) channels as a function

of time in a stochastic simulation of a Ca^{2+} release site composed of 60 independently gating Keizer-Levine RyRs ($c_* = 0$). While the background $[Ca^{2+}]$ is initially $c_\infty = 0.1\ \mu M$, this value is increased to 0.35 μM and 0.5 μM at times indicated in the lower panel. Note the increase in N_O upon the second step in $[Ca^{2+}]$ corresponds to the phenomenon of "Ca^{2+} adaptation" that is an important aspect of the chapter that introduced the RyR model used here (cf. Fig. 2C in [Keizer and Levine, 1996]). For comparision, the thick solid lines of Fig. 2.9 show results for a Ca^{2+} release site composed of 60 *coupled* Keizer-Levine RyRs ($c_* = 0.02\ \mu M$); interestingly, in this case adaptation is no longer observed.

More important to our present purposes are the broken lines of Fig. 2.9, which show exact results obtained from the probability distribution $\pi(t)$ directly calculated using matrix exponentials of fast/slow reduced release generator matrices, that is,

$$\boldsymbol{\pi}(t) = \begin{cases} \boldsymbol{\pi}_0 & t < t_1 \\ \boldsymbol{\pi}_0 e^{(t-t_1)\hat{\boldsymbol{Q}}_1} & t_1 \leq t < t_2 \\ \boldsymbol{\pi}_0 e^{(t_2-t_1)\hat{\boldsymbol{Q}}_1} e^{(t-t_2)\hat{\boldsymbol{Q}}_2} & t_2 \leq t, \end{cases} \tag{2.20}$$

where $\boldsymbol{\pi}_0 \hat{\boldsymbol{Q}}_0 = \boldsymbol{0}$ subject to $\boldsymbol{\pi}_0 e = 1$, and $\hat{\boldsymbol{Q}}_0$, $\hat{\boldsymbol{Q}}_1$, and $\hat{\boldsymbol{Q}}_2$ are generator matrices reduced from the full model evaluated with $c_\infty = 0.1$, 0.35, and 0.5 μM, respectively. While it is possible to obtain similar results by performing many Monte Carlo simulations and averaging, direct numerical calculation is computationally more efficient because the matrix exponential calculations of Eq. 2.20 use the 61-state reduced generator matrix ($0 \leq N_R \leq N$) as opposed to the 39711-state full model ($\beta(60,4)$ in Eq. 2.5).
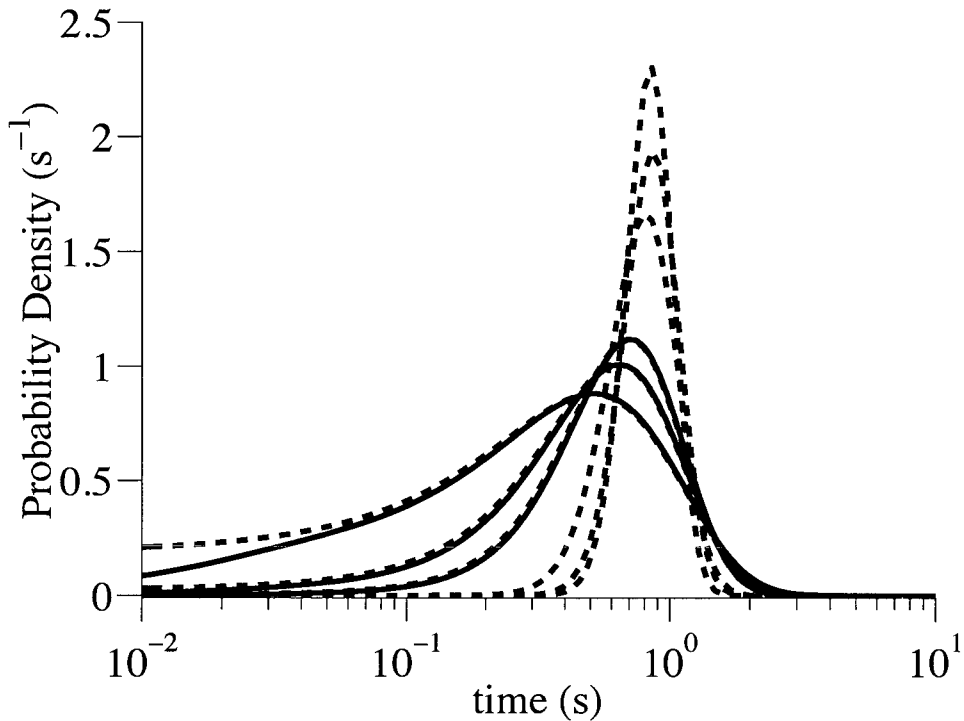
Figure 2.10: Directly calculated probability density of the time until the number of refractory channels (N_R) increases to $N/2$, half the total number of channels in the release site model. Solid lines show results obtained using the full model generator matrices for 8, 12, and 16 channels (sizes 165–969; see Eq. 2.5). Dashed lines show results obtained using the fast/slow reduced generator matrix for 8, 12, 16, 40, 60 and 80 channels (sizes 9–81). In all calculations the initial probability distribution is the stationary distribution for $c_\infty = 0.1$ μM; at time zero this background $[Ca^{2+}]$ is increased to $c_\infty = 0.35$ μM. The coupling strengths were chosen so that $c_* N = 0.52$ μM (e.g., in the eight channel case $c_* = 0.065$ μM).

Fig. 2.10 gives another example of how automated fast/slow reduction can be used in conjunction with matrix analytic formulas to probe the stochastic dynamics of Ca^{2+} release sites, the size of which would otherwise make direct numerical calculations unfeasible, and Monte Carlo simulation inefficient and unreliable. Using $N = 8$, 12, and 16 channels, the solid lines of Fig. 2.10 present direct calculations of the probability density of the time until the number of refractory channels (N_R) increases

44

to $N/2$, half the total number of channels in the release site model. These were calculated by permuting the generator matrix of the full model into the following form:

$$Q = \begin{pmatrix} Q_{aa} & Q_{ab} \\ Q_{ba} & Q_{bb} \end{pmatrix}, \tag{2.21}$$

where each partition contains rates for transitions between (or within) aggregate classes of states where $N_R < N/2$ (a) and $N_R \geq N/2$ (b). The probability distribution is given by [Ball and Geoffrey, 2000, Ball et al., 2000]

$$f(t) = -\phi_a e^{tQ_{aa}} Q_{aa} e_a > 0, \tag{2.22}$$

where e_a is a commensurate column vector of ones, ϕ_a is a row vector giving the initial probabilities of each state, and for simplicity we assume $\phi_a = \pi_a / \pi_a e_a$ where $\pi = (\pi_a \ \pi_b)$ is the stationary distribution solving $\pi Q = 0$. The dashed lines of Fig. 2.10 repeat these calculations using the generator matrix for the fast/slow reduced model. Not only does the agreement validate the reduction method, but perhaps more importantly, by using the fast/slow reduced generator matrix we are able to calculate the distributions for release sites composed of 40, 60 and 80 channels (dashed lines). Because the matrix exponential in Eq. 2.22 must be calculated for many different values of t, full model calculations are extremely time consuming if not impossible due to storage limitations. On the other hand, calculating the matrix exponentials in the reduced model case takes less than a second. While performing the model reduction using the IAD-based reduction method (Algorithm 4) is overhead, this step need be performed only once.

2.10 Discussion

We have implemented and validated several numerical procedures for reducing compositionally defined calcium release site models through fast/slow analysis. In all the approaches presented here, rate constants in the single channel model are categorized as either fast or slow, groups of states in the release site model that are connected by fast transitions are identified and lumped, and transition rates between reduced states are chosen consistent with exact or approximate conditional probability distributions among states within each group. For Ca^{2+} release site models that are small enough to allow direct calculation of the stationary distribution of the full model, Algorithm 3 is preferred in spite of its substantial storage requirements, because the exact conditional probability distributions result in a reduced model that is natural for the chosen partitioning of states. For release sites composed of many channels, the conditional probability distributions can be approximated without the construction of the full model by assuming rapid mixing of states connected by fast transitions (Algorithms 1 and 2). Alternatively, an iterative aggregation/disaggregation (IAD) method can be employed to obtain a reduced Ca^{2+} release site model in a memory-efficient fashion.

We compared the convergence properties of reduction algorithms using three IAD methods: Koury-McAllister-Stewart, Vantilborgh, and Takahashi [Koury et al., 1984, Cao and Stewart, 1985]. Our results suggest that Koury-McAllister-Stewart IAD-based reduction method is superior in the context of Ca^{2+} release site modeling (Algorithm 4). Calculations performed using Vantilborgh IAD required more iterations to converge than Koury-McAllister-Stewart, while those using the Takahashi method often did not converge (not shown). Note that memory-efficient implementation of model reduction using Algorithm 4 begins with enumeration of the state

space of a full Ca^{2+} release site model. This preliminary step must also be performed without excessive storage requirements (see Appendix B and Algorithms 5 and 6).
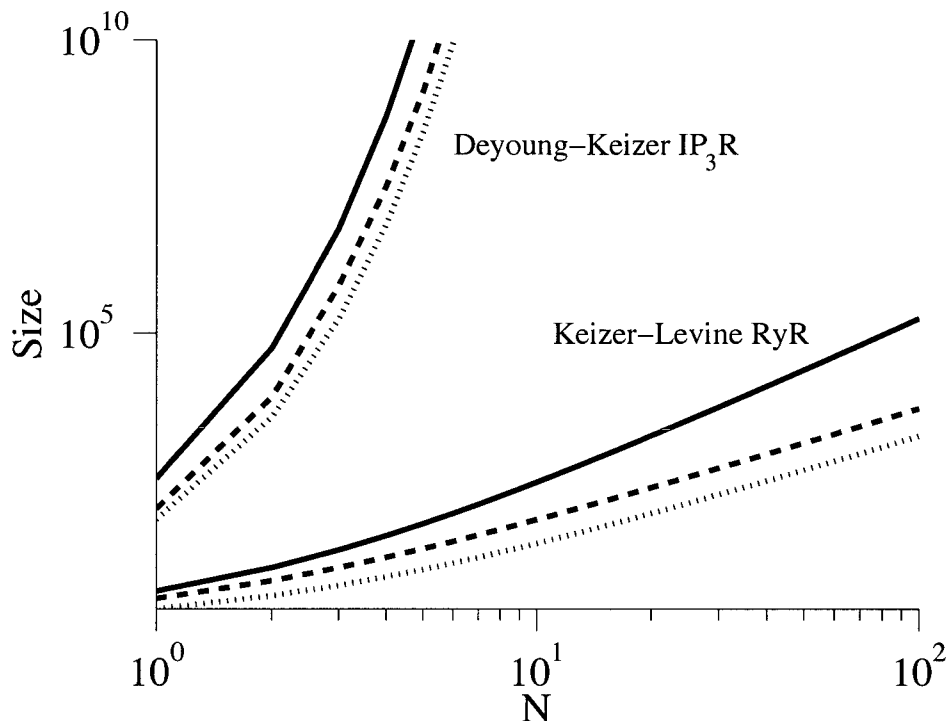
Figure 2.11: State space size (Eq. 2.5) of the full Ca^{2+} release site model (solid line), size of the largest (dashed line) and average (dotted line) diagonal block of the partitioned generator matrix for $1 \leq N \leq 100$ 4-state Keizer-Levine RyRs (see Fig. 2.4). The reduced release site model has $\hat{b} = N + 1$ states (not shown) because slow transitions in the Keizer-Levine RyR separate 2 groups of states (Fig. 2.2).

We were able to validate Algorithms 1–4 by confirming that the transition probability matrix of the reduced model well-approximates the corresponding contraction of the full model transition probability matrix, provided the separation of time scales between fast and slow processes is large enough (Figs. 2.6 and 2.8). As expected, both Algorithm 1 and 2 yield more error than the memory-inefficient reduction that uses the exact conditional probability distributions (Algorithm 3). Note that the Koury-McAllister-Stewart IAD-based Algorithm 4 produces the same reduced model

as Algorithm 3. The essential difference between Algorithms 3 and 4 is the numerical scheme used to calculate the exact conditional probability distributions. Because Algorithm 3 is not tractable for large Ca^{2+} release site models, we recommend Algorithm 4 to investigators interested in Ca^{2+} release site model reduction based on a separation of time scales.

It is important to note that while we have validated the four model reduction procedures presented here (Algorithms 1–4), the performance of a particular reduced model is a complicated matter that will depend on the single channel model used and, of course, the choice of parameters that influence the time scale separation of transitions identified as fast and slow.

While the error measure based on transition probability matrices (Eq. 3.25) is sufficient for our present purposes, we have not yet performed a detailed study of puff/spark duration and inter-event interval in full and reduced Ca^{2+} release site models. The extent to which model reduction may perturb measures of particular relevance to the stochastic dynamics of Ca^{2+} release is a question that deserves further consideration. Because puff/spark statistics are coarser measures of release site dynamics than the transition probability matrix itself, a reduced model could perform well with respect to the distribution of spark durations (for example), even when $\hat{E}_{max}(t)$ is not promising. While it is of some concern that $\hat{E}_{max}(t)$ often grows with the number of channels (Fig. 2.12), this doesn't adversely affect the reduced model probability densities of Fig. 2.10.

Although beyond the scope of this chapter, Algorithm 4 can be implemented in a distributed parallel fashion. Such implementation would likely be required to perform fast/slow reduction when Ca^{2+} release sites are composed of single channel models with many states. For example, a DeYoung-Keizer-like IP$_3$R model [De Young and Keizer, 1992] that includes 4 independent eight-state subunits—each with one bind-
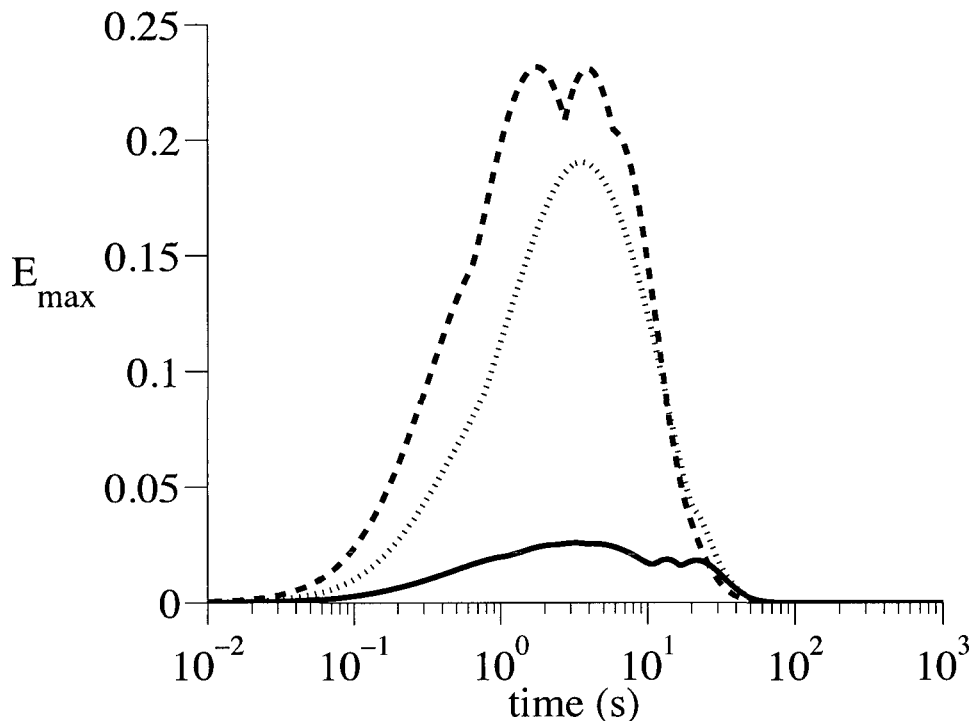
Figure 2.12: Reduction error of "gold standard" reduction procedure (Algorithm 3) for release sites composed of 8 (solid line), 12 (dotted line) and 16 (dashed line) four-state Keizer-Levine RyRs. $c_\infty = 0.1$ μM, $c_* = 0.065$ μM, and other parameters as in Fig. 2.1.

ing site for IP$_3$ and 2 binding sites for Ca^{2+}—results in a single channel model with $\beta(4,8) = 330$ distinguishable states (Eq. 2.5). Assuming fast IP$_3$-potentiation, fast Ca^{2+}-activation, and slow Ca^{2+}-inactivation, the topology of the fast and slow transitions results in 2 groups of 4 states for each subunit. This results in five groups with 35, 80, 100, 80, and 35 states for the single channel model, that is, $\beta(n_{\text{disinact}}, 4)\beta(n_{\text{inact}}, 4)$ for $n_{\text{inact}} = 0, 1, 2, 3, 4$ and $n_{\text{inact}} + n_{\text{disinact}} = 4$. Assuming a release site composed of N DeYoung-Keizer-like IP$_3$Rs, Fig. 2.11 shows the state space size of the full model (solid line), and the size of the largest (dashed line) and average (dotted line) diagonal block (cf. Eq. 2.10). Note that the limiting slopes for the DeYoung-Keizer IP$_3$R are

49

much greater than those observed for the Keizer-Levine RyR. For the Keizer-Levine RyR, the number of states in the full model is $\beta(N, 4) \sim N^3$ and the largest block size is $\beta(N, 3) \sim N^2$ (all channels in the largest group that includes 3 states; see Fig. 2.2). For the DeYoung-Keizer-like IP$_3$R, the number of states in the full model is $\mathcal{O}(N^{330})$ and the largest block size is $\mathcal{O}(N^{326})$ (326 = 330 states − 5 groups + 1).

Throughout this chapter we assume that the fast and slow transitions of the single channel model are identified by the modeler, and this specification is used to partition the full model generator matrix (cf. Fig. 2.2). While this makes sense given the likely prior understanding of time scales of single channel kinetics, this approach neglects the effect of [Ca^{2+}] changes on separation of time scales. That is, a Ca^{2+}-dependent transition such as $C_1 \rightarrow O_2$ or $O_2 \rightarrow O_3$ in the Keizer-Levine RyR may be slow or fast depending on $N_O(t)$. While the memory-efficient Algorithm 4 leads to the "gold standard" reduced model for any given partitioning, the approach to partitioning used here may not be optimal. In fact, when a 165-state release site is reduced to 9 states as in Figs. 2.5–2.8, there are $\beta(9, 165) \approx 3 \times 10^{14}$ possible partitioning schemes. Given the separation of time scales in the Keizer-Levine RyR, the chosen partitioning scheme is presumably among the best, but it is unclear how to demonstrate this without enumerating all the possibilities and comparing reduction errors. An important topic for future work is automated determination of the optimal partitioning of a full model generator matrix to achieve a target number of reduced model states. In cases where the reduction error is defined in terms of a puff/spark statistic of interest (e.g., spark duration), the optimal partitioning schemes would presumably be sensitive to the aggregate classes of states being lumped (e.g., closed vs. open) as well as separation of time scales [DeRemigio et al., 2008]. In future work we hope to combine the automated fast/slow reduction procedure presented here with whole cell modeling techniques that include a probability density-based description of

the local [Ca^{2+}] experienced by clusters of intracellular and plasma membrane Ca^{2+} channels [Williams et al., 2007, Williams et al., 2008].

Algorithm 5: $\mathbb{B}(nball, nbin)$

Recursive state space generation

---

require: $nball$, $nbin$

if $nbin = 1$ **return** $nball$

if $nball = 0$ **return** $1 \times nbin$ matrix of zeros

$\boldsymbol{B} \leftarrow \emptyset$ (an empty matrix)

for $\ell = nball, nball - 1, \ldots, 0$

    $\boldsymbol{B}_R = \mathbb{B}(nball - \ell, nbin - 1)$

    $\boldsymbol{B}_L \leftarrow$ column vector of ℓ's with same number

      of rows as $\boldsymbol{B}_R$

$$\boldsymbol{B} \leftarrow \left[\begin{array}{cc} \boldsymbol{B} \\ \hline \boldsymbol{B}_L & \boldsymbol{B}_R \end{array} \right]$$

endfor

return $\boldsymbol{B}$

2.11   Appendices

2.11.1   Implementation of fast/slow reduction procedures

Algorithms 1–6 were implemented in Matlab (The MathWorks, Inc.). Equations of the form $\boldsymbol{xA} = \boldsymbol{0}$ subject to $\boldsymbol{xe} = 1$ were solved by evaluating $\boldsymbol{x} \leftarrow (\ \boldsymbol{0}\ \ 1\ )/(\ \boldsymbol{A}\ \ \boldsymbol{e}\ )$ where the slash corresponds to Matlab's `mrdivide` command. When solving an equation of the form $\boldsymbol{xP} = \lambda \boldsymbol{x}$ subject to $\boldsymbol{xe} = 1$ we used Matlab's `eigs` command to find the eigenvector corresponding to the eigenvalue with largest real part and then

normalize the result. In our implementation of Algorithm 4, the aggregation and dis-aggregation steps were solved using `eigs` and `mldivide`, in spite of the fact that the aggregated system for nearly completely decomposable Markov chains is expected to be ill-conditioned [Stewart, 1994, pages 321–322]. It is possible that the inferior performance of the Takahashi method could be improved with a different implementation of these steps (Table 2.1).

2.11.2 Generation of state space and blocks of partitioned full model

Instantaneous mean-field coupling of N identical M-state channels yields a Ca^{2+} release site model with $\beta(N, M)$ states where

$$\beta(N, M) = \binom{N + M - 1}{N} = \frac{(N + M - 1)!}{N! \, (M - 1)!}.$$

Assuming transitions in the single channel model are labelled fast or slow in a manner that results in L groups of states of size m_1, m_2, ..., m_L with $\sum_{i=1}^{L} m_i = M$, the partitioned matrix corresponding to Eq. 2.10 will have $\beta(N, L)$ blocks, each of which can be labelled as $(n_1 n_2 \ldots n_L)$ indicating n_i channels in group i where $\sum_{i=1}^{L} n_i = N$. The diagonal block corresponding to macrostate $(n_1 n_2 \ldots n_L)$ is a square matrix of size

$$\prod_{i=1}^{L} \binom{n_i + m_i - 1}{n_i}.$$

An important aspect of the memory-efficient model reduction approach of Algorithm 4 is construction of the $\hat{b}^2$ input matrices $\{Q_{ij}\}$. To ensure that the storage requirements of specifying the full model are not limiting, it is helpful to construct the Q_{ij} independently. This was accomplished using a recursive function $\mathbb{B}(nball, nbin)$

that returns a matrix enumerating (in anti-lexicographical order) the number of ways that *nball* indistinguishable items can be arranged in *nbin* distinguishable locations (Algorithm 5). For example, the full state space for two four-state channels is the 10×4 matrix

$$
\mathbb{B}(2,4) =
\begin{bmatrix}
2 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 \\
& & \vdots & \\
0 & 0 & 0 & 2
\end{bmatrix} .
\tag{2.23}
$$

When the state space of the full model ($\mathbb{B}(N, M)$) is large, the state space of the reduced model ($\mathbb{B}(N, L)$) is constructed instead, where L is the number of groups of states separated by slow transitions ($L < M$). Denoting the rows of $\mathbb{B}(N, L)$ as $\{n_1 n_2 \ldots n_L\}$ where $n_1 = 0, 1, \ldots, m_1$; $n_2 = 0, 1, \ldots, m_2$; etc., the states in the full model that compose any particular lumped state $n_1 n_2 \ldots n_L$ can be enumerated as following:

$$
\begin{bmatrix}
\mathbb{B}_1(n_1, m_1) & \mathbb{B}_1(n_2, m_2) & \cdots & \mathbb{B}_1(n_L, m_L) \\
\mathbb{B}_1(n_1, m_1) & \mathbb{B}_1(n_2, m_2) & \cdots & \mathbb{B}_2(n_L, m_L) \\
\vdots & \vdots & & \vdots \\
\mathbb{B}_1(n_1, m_1) & \mathbb{B}_1(n_2, m_2) & \cdots & \mathbb{B}_K(n_L, m_L) \\
\mathbb{B}_1(n_1, m_1) & \mathbb{B}_2(n_2, m_2) & \cdots & \mathbb{B}_1(n_L, m_L) \\
\vdots & \vdots & & \vdots \\
\mathbb{B}_K(n_1, m_1) & \mathbb{B}_K(n_2, m_2) & \cdots & \mathbb{B}_K(n_L, m_L)
\end{bmatrix} ,
$$

where $\mathbb{B}_k(n_\ell, m_\ell)$ indicates the kth row of $\mathbb{B}(n_\ell, m_\ell)$ and an upper case K indicates the final row.

Algorithm 6: $\mathbb{R}(B)$

Determine transition rates for a given block of the full model

require: origin and destination states B

$n \leftarrow$ number of rows of B

$R \leftarrow n \times n$ matrix of zeros

for $i = 1, \ldots, n$

    for $j = i + 1, i + 2, \ldots, n$ **do**

        $\Delta \leftarrow (j$th row of $B) - (i$th row of $B)$

        if Δ contains exactly one -1 and one 1 **then**

            $R(i, j) \leftarrow$ index of the -1 in Δ

            $R(j, i) \leftarrow$ index of the 1 in Δ

        endif

    endfor

endfor

return R

With the subset of the full model state space corresponding to a particular lumped state $n_1 n_2 \ldots n_L$ available, it is possible to construct the blocks Q_{ij} of the partitioned full model without knowledge of the entire state space. This is accomplished using Algorithm 6, which takes as input a matrix B corresponding to a set of states and returns as output the matrix $\mathbb{R}(B) = R = (r_{k\ell})$, where the $r_{k\ell}$ are nonzero if and only if a transition is possible between states B_k and B_ℓ and, when a transition is possible, the origin and destination states of the one channel that changes state are $r_{k\ell}$ and $r_{\ell k}$, respectively. For example, focusing on the subsequence of rows of Eq. 2.23

corresponding zero inactivated channels ($N_{C_2} = 0$),

$$
B = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix},
$$

the function $\mathbb{R}(B)$ evaluates to

$$
\mathbb{R}(B) = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 2 & 0 & 2 & 1 & 1 & 0 \\ 3 & 3 & 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 & 2 & 0 \\ 0 & 3 & 2 & 3 & 0 & 2 \\ 0 & 0 & 3 & 0 & 3 & 0 \end{bmatrix}. \tag{2.24}
$$

The diagonal block of the full model corresponding to transitions within states of B is then given by

$$
Q_{BB} = \begin{bmatrix} \cdot & q_{12} & q_{13} & \cdot & \cdot & \cdot \\ q_{21} & \cdot & q_{23} & q_{12} & q_{13} & \cdot \\ q_{31} & q_{32} & \cdot & \cdot & q_{12} & q_{13} \\ \cdot & q_{21} & \cdot & \cdot & q_{23} & \cdot \\ \cdot & q_{31} & q_{21} & q_{32} & \cdot & q_{23} \\ \cdot & \cdot & q_{31} & \cdot & q_{32} & \cdot \end{bmatrix}
$$

where the dots indicate zero, the q_{ij} are the $i \to j$ transition rates of the single channel

55

model that either do not depend on $[\text{Ca}^{2+}]$ or are evaluated using N_O consistent with the relevant row of B, and the indices for these transition rates are chosen by reading off the elements of $\mathbb{R}(B)$ and $\mathbb{R}(B)^T$. Off-diagonal blocks of the full model corresponding to transitions between two groups of states (B_- and B_+) are found in a similar manner, beginning with the evaluation of

$$
\mathbb{R}\left(\begin{bmatrix} B_- \\ B_+ \end{bmatrix}\right) = \begin{bmatrix} R_{--} & R_{-+} \\ R_{+-} & R_{++} \end{bmatrix}
$$

using Algorithm 6. The matrices R_{-+} and R_{-+} provides indices of the single channel model transition rates needed to produce $Q_{B_- B_+}$ and $Q_{B_+ B_-}$.

	KMS		Vantilborgh		
N	Iter	Resid	Iter	Resid	*Score*
10	15	1.9e-10	15	7.4e-10	0.35
20	36	2.6e-9	85	2.9e-8	0.49
30	33	2.3e-9	99	5.6e-8	0.33
40	28	2.2e-9	73	3.8e-8	0.23
50	22	2.9e-9	59	4.5e-8	0.15
60	13	4.5e-10	14	3.4e-9	<0.01
70	5	1.2e-9	8	9.4e-10	<0.01
80	9	1.31e-9	15	1.88e-9	<0.01

Table 2.1: Benchmark calculations using two iterative aggregation/disaggregation algorithms: Koury-McAllister-Stewart (KMS) and Vantilborgh. The number of iterations (Iter) before convergence of the iteration vector (*tolerance* $= 10^{-8}$ in Algorithm 4), and the residual (Resid) of the calculated stationary distribution vector π given by $\|\pi Q\|_1$ are shown. Parameters: $c_\infty = 0.1$ μM, $c_* = 0.06$ μM and as in Fig. 2.1. Because the Ca^{2+} coupling strength is fixed, release sites with large N are tonically active resulting in low puff/spark *Score* (cf. Table 2.2).

	KMS			
N	Iter	Resid	c_*	*Score*S
30	44	1.7e-9	0.04	0.50
30	39	2.9e-9	0.05	0.41
30	33	2.3e-9	0.06	0.33
40	53	1.7e-9	0.03	0.51
40	46	1.6e-9	0.04	0.39
40	33	3.3e-9	0.05	0.30
50	57	3.3e-9	0.03	0.44
50	46	2.1e-9	0.04	0.30
50	28	4.4e-9	0.05	0.22
60	97	3.01e-9	0.02	0.52
60	64	3.54e-9	0.03	0.36
60	38	1.47e-9	0.04	0.23

Table 2.2: Benchmark calculations using c_* values chosen so that the puff/spark *Score* of the full model indicated robust Ca^{2+} excitability. See legend of Table 2.1.

Chapter 3

Reduction of calcium release site models using a genetic algorithm

3.1 Summary

Mathematical models of calcium release sites derived from Markov chain models of intracellular calcium channels exhibit collective gating reminiscent of the experimentally observed phenomenon of calcium puffs and sparks. Such models often take the form of stochastic automata networks in which the transition probabilities of each channel depend on the local calcium concentration and thus the state of the other channels. In prior work, to overcome the state-space explosion that occurs in such compositionally defined calcium release site models, we have implemented several automated procedures for model reduction based on fast/slow analysis where rate constants in release site models are categorized as either fast or slow, groups of states that are connected by fast transitions are lumped, and transition rates between reduced states are chosen consistent with the conditional probability distribution among states within each group. Here we implemented a genetic algorithm based approach to find reduced

models that produce moderate errors for problems without time-scale separation. Given a full model and a specified target size for a reduced model, this algorithm mutates and select members of a population of set partitions, each corresponding to a potential scheme for state aggregation, using a fitness function that favors partitions leading to reduced models approximate the full model on the behaviors of interest. With this genetic algorithm based approach, release site/complex models which are too large to benefit from computational approaches, moment closure approaches for example, are able to be reduced to reasonable sizes with very low reduction errors for a wide range of SR [Ca^{2+}] and a reasonable amount of computational effort.

3.2   Introduction

As a second messenger, calcium ions (Ca^{2+}) plays an important role in many physiological activities. Signaling occurs when the cell is stimulated to release calcium ions (Ca^{2+}) from the endoplasmic/sarcoplasmic reticulum (ER/SR), the intracellular Ca^{2+} reservoir, and/or when Ca^{2+} enters the cell through plasma membrane ion channels [Clapham, 1995]. The intracellular Ca^{2+} release which causes localized Ca^{2+} elevations known as puffs and sparks arise from concerted gating of clusters of inositol 1,4,5-trisphosphate receptors (IP3Rs) or ryanodine receptors (RyRs) on the surface of (ER/SR) [Berridge, 1997b, Cheng et al., 1996, Yao et al., 1995]. In cardiac myocyte excitation-contraction coupling (ECC), for example, the cell membrane depolarizes causing L-type Ca^{2+} channels to open and the Ca^{2+} influx further activates RyRs located on SR, known as Ca^{2+}-induced Ca^{2+} release (CICR) [Endo, 1977].

In the literature, the behavior of single IP3R/RyR channel gating is often modeled by continuous-time descrete-state Markov chans (CTMCs) [Colquhoun and Hawkes, 1995, Smith, 2002]. When Markov chain models of these channels are coupled via a

Ca^{2+} microdomain, by which the transition rates between the states of each channel become dependent to the states of other channels, the simulated Ca^{2+} channel clusters (release sites) may exhibit stochastic excitability that is reminiscent of Ca^{2+} puff/sparks [Nguyen et al., 2005]. However, the number of states possessed by these compositionally defined Ca^{2+} release site models increases exponentially as the number of channels increases. This combinatorial state-space explosion causes some modeling approaches to become intractable.

While the dynamics of any individual Ca^{2+} release site can in principle be obtained by Monte Carlo simulation regardless of model complexity, in practice these simulations are prohibitively computationally intensive due to large state spaces. Moreover, because cells usually possess a large number of release sites, compositionally defined Ca^{2+} release site models have often been excluded from multiscale whole cell simulations. On the other hand, many recently developed approaches that accelerate whole cell simulations, e.g., probability density and moment closure approaches [Williams et al., 2007, Williams et al., 2008], require release sites to be as compact as possible while retaining the physiological realism of collective channel gating. For these reasons, we developed several automated approaches based on fast/slow analysis [Hao et al., 2009] to reduce Markov chain Ca^{2+} release site models where the rate constants in release site models are categorized as either fast or slow, groups of states that are connected by fast transitions are lumped so that the full model is compressed into a tractable size while the physiological gating and interaction properties of the channels are preserved. However, when the time-scale separation between transition rates that is necessary for fast/slow analysis is absent, the manner in which the full model states should be partitioned and aggregated for optimal reduction is difficult to determine a priori. Naively enumerating all partitions for a Markov chain Ca^{2+} release site models and choosing the one with the least error is not possible because

the number of possible partitions is too large. For example, a release site model composed of merely 5 three-state Ca^{2+} channels (15 states) can be partitioned in 10^{10} distinct ways. In this chapter we discuss the implementation of a genetic algorithm that is able to automatically and rapidly select partition schemes that reduce the corresponding Markov chain model reduction error.

Developed in the 1970s by John Holland [Holland, 1975], genetic algorithms are widely used as computational schemes to find exact or approximate solutions for optimization and search problems. Genetic algorithms have been applied to various aspects of biological research, e.g. the profiling of the gene expression in bacteria [DiGesú et al., 2005, To and Vohradsky, 2007] and phylogenetic analysis of proteins [Hill et al., 2005]. However, the application of genetic algorithms in the context of the automated reduction of Ca^{2+} release site models is novel. In our implementation, a population of set partitions are randomly generated, each corresponding to a potential scheme for state aggregation. The program 'evolves' this population by selecting the partitions that lead to reduced models that approximate the full model behavior.

Unlike the fast/slow analysis of Chap. 2 that assumes fixed ER/SR $[Ca^{2+}]$ and instantaneous coupling between the channels [Hao et al., 2009], we motivate a whole cell homeostasis formulation which takes both local and global Ca^{2+}signaling into consideration and the reduced models selected by the genetic algorithm generate small error for a wide range of ER/SR $[Ca^{2+}]$.

The remainder of this paper is organized as follows. In Section 3.3 we motivate the model reduction process by partitioning a minimal whole cell model of Ca^{2+} homeostasis where bidirectional influence of local and global Ca^{2+} signals are modeled. In section 3.4 we introduce genetic algorithms and detail their implementation in the model reduction context. In Section 3.5, we demonstrate that the reduced model approximates the full model with regard to several important steady-state responses

observed in the minimal whole cell environment. To show that the reduction technique is applicable to more realistic Ca^{2+} release site models, we also present Ca^{2+} release site reduction results using a single channel model that includes both cytosolic and luminal Ca^{2+} regulation.

3.3 Model formulation

3.3.1 A minimal whole cell model

We will demonstrate and validate our Ca^{2+} release site model reduction approach using a whole cell model of a quiescent cytosolic environment that takes Ca^{2+} homeostasis into account (Fig. 3.1). Similar to previous work by Hartman and colleagues [Hartman et al., 2010], in this minimal whole cell model both local and global Ca^{2+} responses to the stochastic gating of Ca^{2+} channels are considered and release and reuptake fluxes are balanced. Fig. 3.1 shows the components and fluxes of the model. A large number of heterogeneous local Ca^{2+} signals associated with a large number of Ca^{2+} release sites are coupled to the bulk cytosolic and ER/SR $[Ca^{2+}]$. Each Ca^{2+} release site is composed of 10–30 Ca^{2+} channels. In this formulation, release sites may experience different "domain" $[Ca^{2+}]$, but all channels in a given release site experience the same local cytosolic and luminal $[Ca^{2+}]$. Consistent with prior work by Hinch and colleagues [Hinch et al., 2004, Hinch et al., 2006, Greenstein et al., 2006], when the number of open channels in a Ca^{2+} release site changes, the local $[Ca^{2+}]$ is assumed to rapidly reach a new equilibrium in the spatially restricted domain. The change in the balance of the leak and reuptake by endo(sarco)plasmic reticulum Ca^{2+}-ATPase (SERCA) pumps caused by this change in domain $[Ca^{2+}]$ will influence the bulk Ca^{2+} concentrations and further affect the puff/spark dynamics.
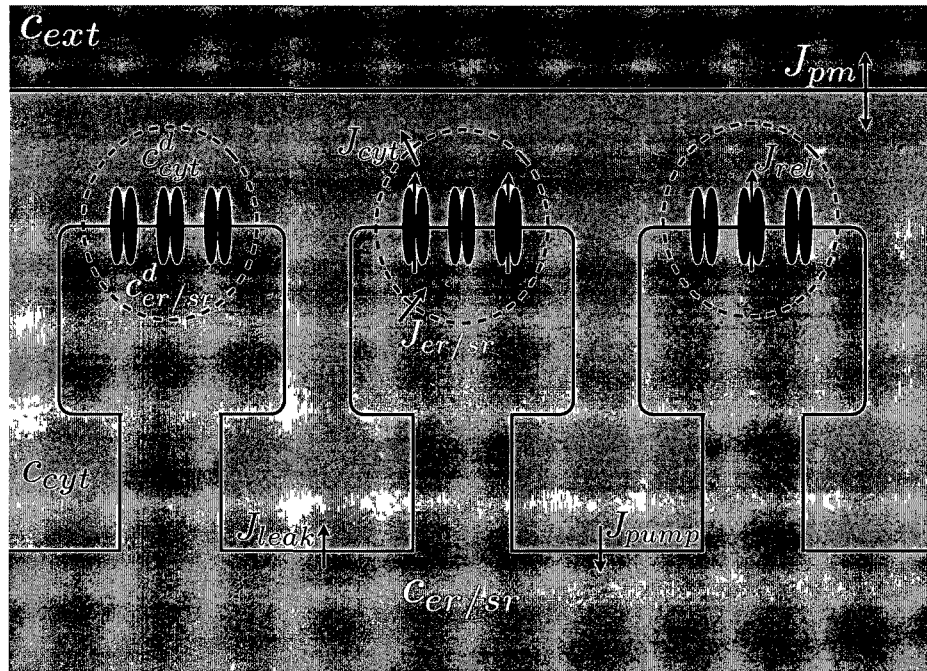
Figure 3.1: Diagram of model components and fluxes. The bulk endoplasmic/sarcoplasmic reticulum $[Ca^{2+}]$ is represented by $c_{er/sr}$, the bulk cytosolic and external $[Ca^{2+}]$ is c_{cyt} and c_{ext} respectively. Ca^{2+} channels locate on the ER/SR membrane forming release sites. The domain $[Ca^{2+}]$ (c_{cyt}^d and $c_{er/sr}^d$) are rapidly changed by the release currents (J_{rel}) when the number of open channels changes. Other fluxes considered in this model are: diffusion from cytosolic domain to the bulk cytosol (J_{cyt}), diffusion from the bulk ER/SR to the luminal side domains ($J_{er/sr}$), a passive leak from the ER/SR to the cytosol (J_{leak}), the SERCA pump flux that resequesters Ca^{2+} in to the ER/SR (J_{pump}) and fluxes across the plasma membrane (J_{pm}).

3.3.2 Steady-state of domain concentration

Fig. 3.1 illustrates the fluxes in this whole cell formulation. The domain [Ca^{2+}] for each release site c_{cyt}^d and $c_{er/sr}^d$ are coupled to each other via the release flux J_{rel} when one or more channels are open. As in mentioned in Sec. 3.3.1, the domain [Ca^{2+}] associated with each release site is distinct and all domains are coupled to the bulk cytosolic and luminal compartments via the fluxes J_{cyt} and $J_{er/sr}$. Under these assumptions, the domain fluxes are given by:

$$J_{rel}^n = \nu_{rel}\gamma_n(c_{er/sr}^{d,n} - c_{cyt}^{d,n}), \tag{3.1}$$

$$J_{cyt}^n = \nu_{cyt}(c_{cyt}^{d,n} - c_{cyt}), \tag{3.2}$$

$$J_{er/sr}^n = \nu_{er/sr}(c_{er/sr} - c_{er/sr}^{d,n}), \tag{3.3}$$

where ν_{rel} is the maximum release rate through a release site, c_{cyt} and $c_{er/sr}$ are the bulk cytosolic and ER/SR concentrations, and $\gamma_n = n/N$ is the fraction of open channels at an N-channel release site. The rate constants ν_{cyt} and $\nu_{er/sr}$ determine the time required for the decay and refilling of the cytosolic and luminal microdomains, respectively [Mazzag et al., 2005, Huertas and Smith, 2007].

Because the dynamics of domain Ca^{2+} is fast compared to the stochastic gating of Ca^{2+} channels (3.3.1), the domain fluxes associated with each release site must balance for any specific release site:

$$J_{rel}^n = J_{cyt}^n = J_{er/sr}^n. \tag{3.4}$$

The domain [Ca^{2+}] of any release site with n channels open can be obtained directly by solving Eq. 3.4 as a function of the bulk cytosolic and luminal [Ca^{2+}] (c_{cyt} and

$c_{er/sr}$), that is,

$$c_{cyt}^{d,n} = \frac{\nu_{cyt}}{\nu_{cyt} + \tilde{\nu}_{er/sr}} c_{cyt} + \frac{\tilde{\nu}_{er/sr}}{\nu_{cyt} + \tilde{\nu}_{er/sr}} c_{er/sr} \tag{3.5}$$

$$c_{er/sr}^{d,n} = \frac{\tilde{\nu}_{cyt}}{\tilde{\nu}_{cyt} + \nu_{er/sr}} c_{cyt} + \frac{\nu_{er/sr}}{\tilde{\nu}_{cyt} + \nu_{er/sr}} c_{er/sr} \tag{3.6}$$

where

$$\tilde{\nu}_{cyt} = \frac{\gamma_n \nu_{rel} \nu_{cyt}}{\gamma_n \nu_{rel} + \nu_{cyt}}, \quad \text{and} \quad \tilde{\nu}_{er/sr} = \frac{\gamma_n \nu_{rel} \nu_{er/sr}}{\gamma_n \nu_{rel} + \nu_{er/sr}}. \tag{3.7}$$

Notice that for a release site with N channels, the number of open channels takes integer values from 0 to N. Consequently, there are $N + 1$ pairs of cytosolic and luminal domain [Ca^{2+}] values for any given values of the bulk concentration ($c_{cyt} and c_{er}$).

3.3.3 Concentration balance equations for the bulk cytosol and ER

As shown in Fig. 3.1, the bulk cytosolic and luminal [Ca^{2+}] are both influenced by the Ca^{2+} fluxes to and from their associated microdomains, J_{cyt}^n and J_{er}^n. The bulk concentrations also interact via a SERCA pump flux that takes the form:

$$J_{pump} = \frac{\nu_{pump} c_{cyt}^2}{\nu_{pump}^2 + c_{cyt}^2} \tag{3.8}$$

and a passive leak from the ER/SR to the cytosol,

$$J_{leak} = \nu_{leak}(c_{er/sr} - c_{cyt}). \tag{3.9}$$

Following previous work [Hartman et al., 2010], our model formulation assumes a permeabilized cell and the plasma membrane flux J_{pm} is

$$J_{pm} = k_{pm}(c_{ext} - c_{cyt}), \tag{3.10}$$

where k_{pm} is chosen large enough so that the bulk cytosolic [Ca^{2+}] is "clamped" to the extracellular bath ($c_{ext} = 0.1\mu$ M).

Now that all Ca^{2+} fluxes are defined, the concentration balance equations for the bulk cytosolic and ER compartments are given by:

$$\frac{dc_{cyt}}{dt} = J_{cyt}^T + J_{leak} - J_{pump} + J_{pm}, \tag{3.11}$$

$$\frac{dc_{er/sr}}{dt} = \frac{1}{\lambda_{er/sr}} \left(J_{er/sr}^T - J_{leak} + J_{pump} \right), \tag{3.12}$$

where $\lambda_{er/sr} = V_{er/sr}/V_{cyt}$, V_{cyt} and $V_{er/sr}$ are the effective cytosolic and ER/SR volumes, i.e., taking Ca^{2+} buffering into account. J_{cyt}^T and $J_{er/sr}^T$ are the sums of fluxes over all release sites. Notice that under the fast domain Ca^{2+} assumption, there are only $N + 1$ pairs of possible domain [Ca^{2+}] values (Sec 3.3.2) and consequently J_{cyt}^T and $J_{er/sr}^T$ can be expressed as,

$$J_{cyt}^T = \sum_{n=0}^{N} f_n v_{cyt}^T \left(c_{cyt}^{d,n} - c_{cyt} \right) \tag{3.13}$$

$$J_{er/sr}^T = \sum_{n=0}^{N} f_n v_{er/sr}^T \left(c_{er/sr} - c_{er/sr}^{d,n} \right) \tag{3.14}$$

where f_n is the fraction of release sites with n open channels.

3.3.4 Markov chain model of single channel gating

The stochastic gating of single channels is studied by a continuous-time discrete-state Markov chain model. For simplicity, this single channel model has only three states including both Ca^{2+} activation and Ca^{2+} inactivation. The transition diagram of this model is given by

$$
\begin{array}{ccccc}
k_a^+(c_{cyt}^d)^2 & & k_b^+(c_{cyt}^d)^2 & & \\
C & \rightleftharpoons & O & \rightleftharpoons & R \\
k_a^- & & k_b^- & &
\end{array}
\tag{3.15}
$$
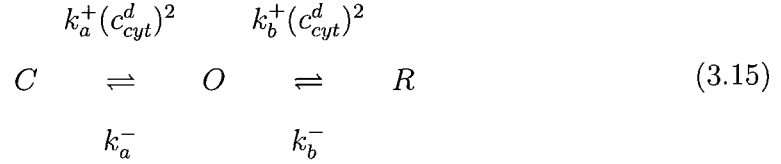
In this transition diagram $k_i^+(c_{cyt}^d)^2$ and k_i^-, where $i \in \{a, b\}$, are transition rates with units of reciprocal time. k_i^+ is an association rate constant with units of $conc^{-\eta}$ $time^{-1}$ where η is the cooperativity of Ca^{2+} binding, and c_{cyt}^d is the domain $[Ca^{2+}]$ experienced by the release site on the cytosol side. Under the assumption that the formation and collapse of local Ca^{2+} is fast compared to channel gating, when the local Ca^{2+} concentrations are specified, the transition-state diagram Eq. 3.15 defines a continuous time Markov chain with infinitesimal generator matrix $\boldsymbol{Q} = (q_{ij})$ given by:

$$
\boldsymbol{Q} =
\begin{bmatrix}
\diamond & k_a^+(c_{cyt}^d)^2 & 0 \\
k_a^- & \diamond & k_b^+(c_{cyt}^d)^2 \\
0 & k_b^- & \diamond
\end{bmatrix}
\tag{3.16}
$$

The off-diagonal entries of the Q-matrix for this irreducible and time-homogeneous Markov chain are transition rate from state i to state j, defined by

$$
q_{ij} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} Pr[S(t + \Delta t) = j | S(t) = i],
\tag{3.17}
$$

where $i \neq j$ and the diamonds ($\diamond$) on the diagonal entries are negative values leading to row sums of zero.

All of the statistical properties of the Ca^{2+} channel can be calculated from its Q-matrix (Eq. 3.16). Importantly, the time evolution of the probability distribution over all three state of this model can be calculated by solving the ordinary differential equation (ODE) system:

$$\frac{d\pi}{dt} = \pi Q,$$ (3.18)

where $\pi(t) = (\pi_C, \pi_O, \pi_R)$ is a row vector containing the probability of finding the channel in each state at time t, given the initial condition $\pi(0)$. Notice that the limiting probability distribution π_s of Markov chains (the steady state of Eq. 3.18) does not depend on the initial condition $\pi(0)$, and can be obtained by solving

$$\pi_s Q = 0 \quad \text{subject to} \quad \pi_s \mathbf{e} = 1,$$ (3.19)

where $\mathbf{e}$ is a commensurate column vector of ones.

3.3.5 Compositionally defined Ca^{2+} release site models

The Ca^{2+} release site models that are used to demonstrate the implementation of the reduction approach involve N identical Ca^{2+} channels. These channels interact via cytosolic local $[Ca^{2+}]$ changes under the "fast domain mean-field coupling" assumption [Nguyen et al., 2005, DeRemigio and Smith, 2005] illustrated in Fig. 3.1, which means these identical channels are indistinguishable. In general, a release site composed of NM-state channels includes

$$\beta(N, M) = \binom{N + M - 1}{N} = \frac{(N + M - 1)!}{N! \, (M - 1)!}$$ (3.20)

distinct states. In this 3-state single channel model case, the N-channel release site has $\beta(N, 3) = (N + 2)(N + 1)/2$ states, each can be written in the form of ordered 3-

tuple (N_C, N_O, N_R), where $N_i = k, (i \in \{C, O, R\})$ indicates k channels are in state i, and $\sum_i N_i = N$. With this notation, the states of any release site is a well ordered set and can be conveniently ranked anti-lexicographically. Fig. 3.2 enumerated all the states and illustrate the topology of the 3-state single channel model (Fig. 3.2A) and a release site composed of 2 3-state channels (Fig. 3.2B).
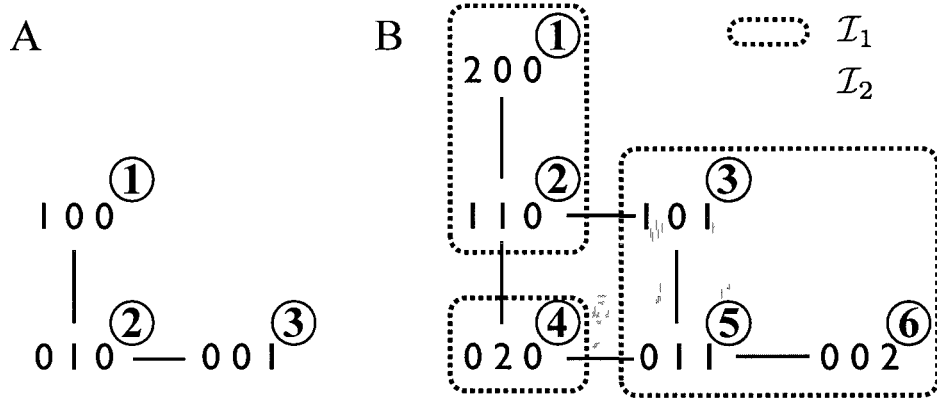
Figure 3.2: (A) The tuple representation of the three-state single channel model in Eq. 3.15. States C, O, R are represented by $(100), (010), (001)$ respectively. (B) The topology and connectivity of a release site composed of two three-state channels in the tuple representation. The 6 states CC, CO, CR, OO, OR, RR are represented by $(200), (110), (101), (020), (011), (002)$ respectively. The ranks of the states are labeled in circles. Dashed line boxes and grey boxes represent two sample 3-partitions of the 2-channel release site $\mathcal{I}_1$ and $\mathcal{I}_2$ in 3.4.2.

3.4 Reduction technique

Our basic strategy of reducing Ca^{2+} release site models to a smaller model with pre-determined size $\hat{b}$ includes three major steps:

Step 1. Partition the full model into $\hat{b}$ groups.

Step 2. Lump the states within each group.

Step 3. Find proper transition rates between groups.

In previous work [Hao et al., 2009], *Step* 1 was achieved automatically based on the separation of time scales. In this paper, we employ a genetic algorithm to search for k-partition schemes that generate small reduction errors for general Ca^{2+} release site models, especially those without time-scale difference. By "k-partition" we mean a partition that divides the states of a release site model into k groups. *Step* 2 and *Step* 3 in this reduction technique are carried on from [Hao et al., 2009]. Even though the genetic algorithm based technique can be used to reduce any Ca^{2+} release site models to any pre-determined size, an example of partitioning the 6-state release site model in Fig. 3.2B into 3 groups will be described in detail hereafter for the sake of simplicity.

3.4.1 Conventional genetic algorithm

Genetic algorithms are probabilistic search algorithms that were introduced by John Holland in the 1970s [Holland, 1975]. They have been used to find exact or approximate solutions to optimization and search problems with objective functions that are discontinuous, nonlinear, difficult to calculate, etc. [Davis, 1991, Michalewicz, 1994] based on the mechanics of natural selection. These algorithms manipulate a population of solutions to the objective function and implement a "survival of the fittest" strategy in their search for better solutions. The general methodology of genetic algorithms is displayed in a flowchart in Fig. 3.3.

The algorithm starts with **Initialization** where a number of *"individuals"* (solution candidates) are randomly generated to form an initial *"population."* The size N_p of the *population* is usually kept as a constant throughout the entire search procedure. Then, this *population* goes through the **Evaluation** procedure where each *individual*
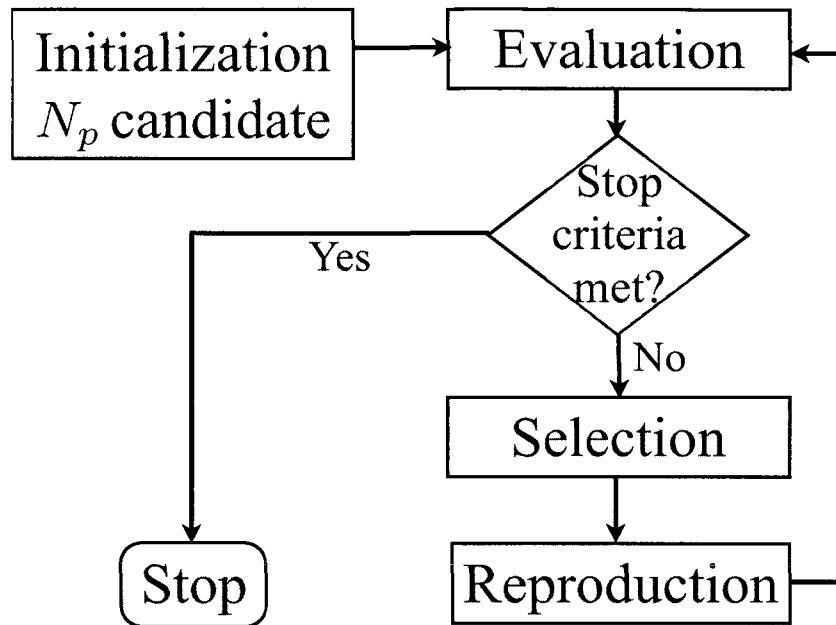
Figure 3.3: A simplified flow chart of the general procedures of genetic algorithms. The program start with the **Initialization** subroutine then loop through **Evaluation**, **Selection** and **Reproduction** till the stop criteria is met.

will have his "fitness" evaluated by the objective function. After each round of evaluation the program checks whether the termination criteria is satisfied. The program usually ends if a solution satisfies the minimum criteria or a predetermined number of generations is reached. If none of the termination criteria is satisfied, the program will move on to the **Selection** process, which is usually stochastic and designed so that the *individual* with better "fitness" have higher probability to be selected compared to those who are less fit. Only a fraction of the current population (N_s *individuals*, where $N_s < N_p$) can survive and enter the **Reproduction** process as the "*parent*" solutions. For each "*child*" (a new solution candidate) to be produced, one or more "*parent*" solutions are selected, recombined (crossover) and/or varied (mutation). The **Reproduction** process continues till N_p *individuals* are generated thus a new generation of *population* is formed. The new generation will then go through *Eval-*

71

uation process to have their "fitness" evaluated and the entire program continues untill the one or more of the termination criteria are satisfied.

3.4.2 Implementation of genetic algorithm to partition states

Initialization

Our purpose of using genetic algorithms is to find partition schemes of full Ca^{2+} release site models so that the resulting reduced models better approximate the full models. In this context, each "Individual" $(\mathcal{I}_i)$ is a set partition scheme which divides the $\beta(N, M)$ states that the full model possesses into $\hat{b}$ groups. To make physical sense, the requirement of the partition process is that each group must be connected within, that is, there is a path from any state to any other state. In Fig. 3.2 gives two sample valid 3-partitions which divides the 6-state release site model into 3 groups by dashed circles and grey boxes:

$$\mathcal{I}_1 = (\{1, 2\}, \{3, 5, 6\}, \{4\})$$

$$\mathcal{I}_2 = (\{1\}, \{2\}, \{3, 4, 5, 6\})$$

In the *Initialization* process, N_p distinct 3-partitions are randomly generated.

Evaluation

In the *Evaluation* process each of the N_p "Individuals" (partition schemes) must be applied to the full model and have their corresponding reduced model compared to the full model. The "fitness" of each "Individual" is then assigned in a manner that favors the ones that produce less error.

We demonstrate this procedure in more detail with

$$\mathcal{I}_1 = (\{1,2\}, \{3,5,6\}, \{4\})$$

given in 3.4.2. First, the generator matrix associated with the two-channel release site model Q is permuted to the order of $\mathcal{I}$. The permutation for $\mathcal{I}_1$ is shown in Fig. 3.4. The new generator matrix $\tilde{Q}$ is then partitioned into a $\hat{b} \times \hat{b}$ block matrix ($\hat{b}$ = 3 for $\mathcal{I}_1$) following the scheme given by each $\mathcal{I}$. The stationary distribution $\tilde{\pi}$ of $\tilde{Q}$ is conformally partitioned as

$$\tilde{\pi} = [\tilde{\pi}_1, \tilde{\pi}_2, \dots, \tilde{\pi}_{\hat{b}}]. \tag{3.21}$$

The generator matrix $\hat{Q}$ of the target reduced Ca^{2+} release site model is a $\hat{b} \times \hat{b}$ matrix

$$\hat{Q} = \begin{bmatrix} \hat{q}_{11} & \hat{q}_{12} & \cdots & \hat{q}_{1\hat{b}} \\ \hat{q}_{21} & \hat{q}_{22} & \cdots & \hat{q}_{2\hat{b}} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{q}_{\hat{b}1} & \hat{q}_{\hat{b}2} & \cdots & \hat{q}_{\hat{b}\hat{b}} \end{bmatrix}, \tag{3.22}$$

where

$$\hat{q}_{ij} = \bar{\pi}_i \tilde{Q}_{ij} e_j \tag{3.23}$$

for $i \neq j$ and $\hat{q}_{ii} = \sum_{j \neq i} -\hat{q}_{ij}$. $\bar{\pi}_i$ is the conditional probability distribution of the states within group i:

$$\bar{\pi}_i = \frac{\tilde{\pi}_i}{\tilde{\pi}_i e_i}, \tag{3.24}$$

and e_i are commensurate column vectors of ones.

When the reduced matrix is generated, the transition probability matrix (jump

73
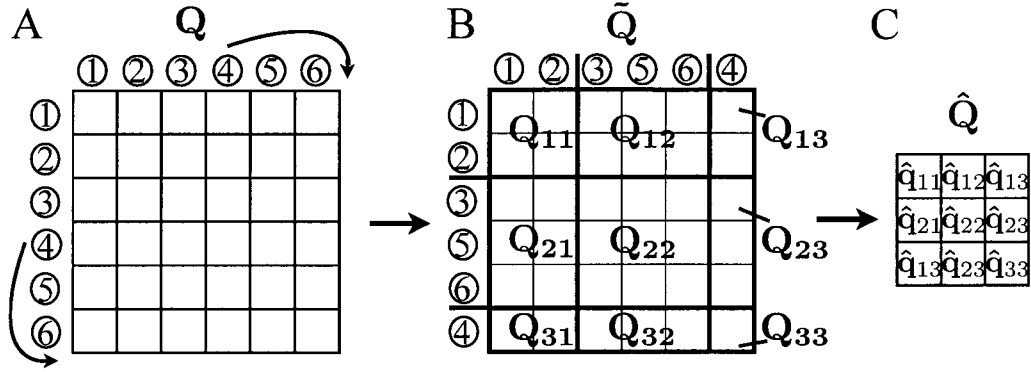
Figure 3.4: Permutation of states and partition structure for two three-state channels following partition scheme $\mathcal{I}_1 = (\{1,2\},\{3,5,6\},\{4\})$. (A) The rows and columns of the expanded generator matrix $Q^{(2)}$ are both permuted following the order given by the grouping scheme in $\mathcal{I}_1$. (B) The block structure given by the thicker lines shows the the partitioning of the generator matrix following $\mathcal{I}_1$. (C) The corresponding reduced matrix calculated as Eq. 3.23

matrix) of their corresponding reduced model ($\hat{\mathbf{P}} = e^{t\hat{\mathbf{Q}}}$) is compared to the transition probability matrix of the full model ($\mathbf{P} = e^{t\mathbf{Q}}$). Assuming the full model has b states, we write

$$\hat{E}(t) = \hat{P}(t) - UP(t)V \tag{3.25}$$

where V is a $b \times \hat{b}$ collector matrix [Nicola, 1998]

$$V = \begin{bmatrix} \mathbf{e}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{e}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{e}_{\hat{b}} \end{bmatrix},$$

the $\mathbf{e}_i$ are column vectors of ones with lengths commensurate with Q_{ii}, and U is a

$\hat{b} \times b$ distributor matrix given by

$$
U = \begin{bmatrix}
\bar{\pi}_1 & 0 & \cdots & 0 \\
0 & \bar{\pi}_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \bar{\pi}_{\hat{b}}
\end{bmatrix}. \tag{3.26}
$$

Notice that, $\bar{\pi}_i$ is the exact conditional probability distribution of the states within group i, calculated from the stationary distribution of the full model. The transition probabilities of the reduced model and the full model agree with each other exactly in the limit. As shown in Fig. 3.5, the maximum difference on transition probabilities falls below 10^{-9} within 1 second.
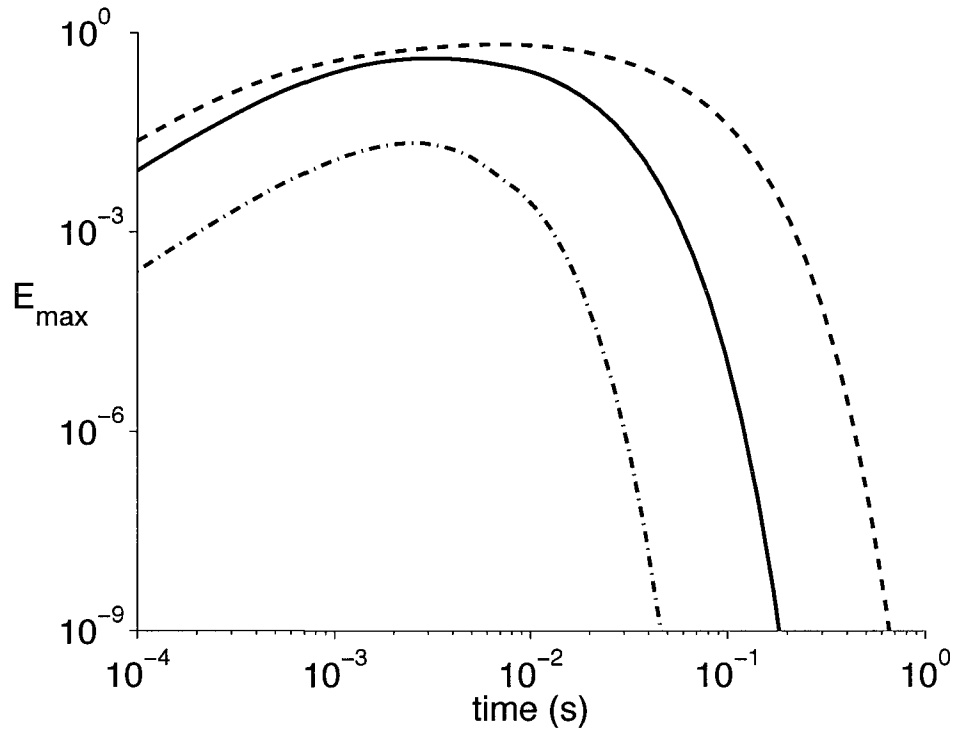
Figure 3.5: The maximum (E_{max}) of the transition probability matrix $\hat{E}(t)$ as a function of time from the reduction of two three-state Ca^{2+} channels (Eq. 3.15) following the partition scheme $\mathcal{I}_1$ when the ER/SR $[Ca^{2+}]$ ($c_{er/sr}$) is 100 μM (dot-dashed line), to 600 μM (solid line), and 1100 μM (dashed line). Parameters: $k_a^+ = 4.5\ \mu M^{-\eta}$ ms^{-1}, $k_b^+ = 0.2\ \mu M^{-\eta}\ ms^{-1}, k_a^- = k_b^- = 500\ ms^{-1}$, $c_{cyt} = 0.1\ \mu$M, $\eta = 2$. Cytosolic side domain $[Ca^{2+}]$ is calculated from Eq. 3.5.

The $b \times b$ matrix $\hat{\boldsymbol{E}}(t)$ is cumbersome to use for evaluation, we consequently define $E_{\mathrm{max}}(t) = \max_{ij} \left| \hat{\boldsymbol{E}}_{ij}(t) \right|$, the element of $\hat{\boldsymbol{E}}(t)$ with largest absolute value at time t. Note that E_{max} is a function of both time and $c_{er/sr}$ because the transition rates of the full Ca^{2+} release site model are functions of the luminal $[Ca^{2+}]$ (Sec. 3.3.1) and consequently Fig. 3.5 plots the reduction error E_{max} as a function of $c_{er/sr}$ assuming the same partition scheme $\mathcal{I}$ and reduction procedure. For example, Fig. 3.5 shows $E_{\mathrm{max}}(t; c_{er/sr})$ for Ca^{2+} release site model in Fig. 3.2(B) reduced to a 3-state model following the partition scheme given by $\mathcal{I}_1$. As validated in [Hao et al., 2009], the reduced model well approximates the full model when E_{max} is small. The maximum transition error is in the range $100 - 2000 \ \mu M$ as the luminal $[Ca^{2+}]$ is raised from $100 \ \mu M$ (dot-dash-dot line) to $1100 \ \mu M$ (dashed line).
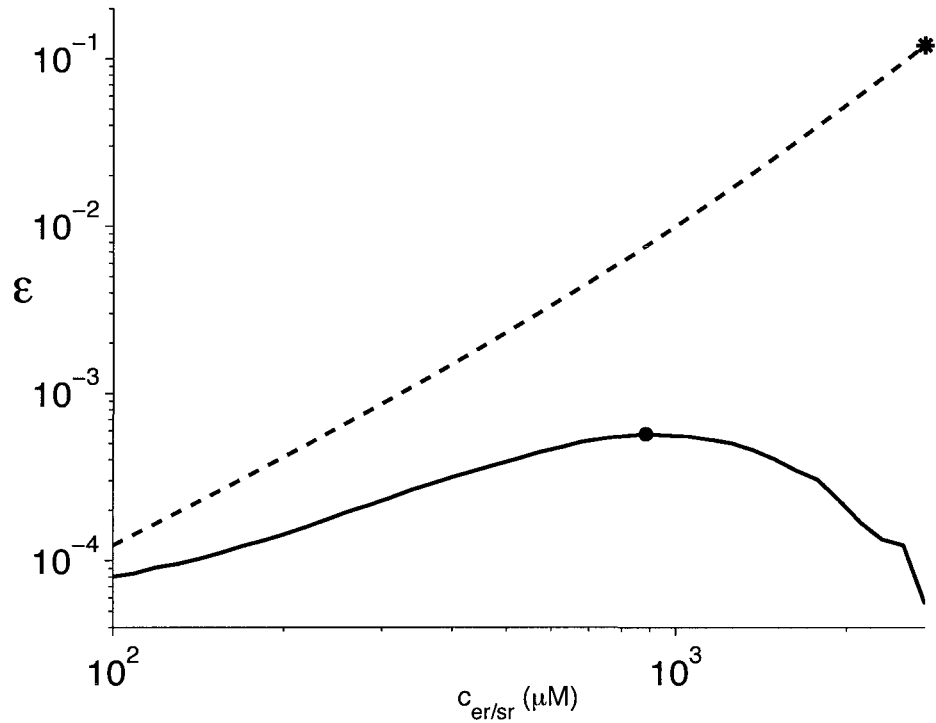
Figure 3.6: The integrated reduction error $\mathcal{E}$ as a function of $c_{er/sr}$ $(100 - 2000\ \mu\mathrm{M})$. The reduction errors associated with partition scheme $\mathcal{I}_1$ and $\mathcal{I}_2$ are shown by the dashed and solid line, respectively. The star and dot indicate the maximum values. Parameters are as in Fig. 3.5.

To have the reduced model applicable in the whole cell simulation described in Sec 3.3.1, the objective we want to achieve through the genetic algorithm is to pick partitions that produce small reduction errors for all possible ER/SR $[Ca^{2+}]$ values at all times. Consequently, we define

$$\mathcal{E}(c_{er/sr}) = \int E_{\max}(t, c_{er/sr})dt, \tag{3.27}$$

the area under each curve in Fig. 3.5. Then, for any partition scheme $\mathcal{I}$, the integrated error $\mathcal{E}$ can be calculated as a function of $c_{er/sr}$ and the maximum $\mathcal{E}(c_{er/sr})$ selected as the global reduction error of scheme $\mathcal{I}$. In Fig. 3.6 the dashed and solid lines show the integrated error $\mathcal{E}$ associated with $\mathcal{I}_1$ and $\mathcal{I}_2$, respectively, as a function of $c_{er/sr}$ (150 − 2000 μM). Because partitions that result in lower reduction errors are preferred, the "fitness" of a given partition scheme $\mathcal{F}$ is defined by

$$\mathcal{F}_{\mathcal{I}} = \frac{1}{\max_{c_{er/sr}} \mathcal{E}_{\mathcal{I}}(c_{er/sr})}. \tag{3.28}$$

As shown in Fig. 3.6, when the full model is partitioned and lumped following $\mathcal{I}_1$ (dashed line), the maximum possible error (star) generated by the reduced model $\mathcal{E}_1$ is approximately 210 times larger than the maximum error (dot) generated by using $\mathcal{I}_2$. The "fitness" of $\mathcal{I}_1$ is consequently 210 times less than the fitness of $\mathcal{I}_2$.

Selection and Reproduction

This section introduces how the genetic algorithm implementation forms the "next generation" from the current "population." The conventional **reproduction** process in genetic algorithms usually consists of **selection**, **crossover** and **mutation**. However, when implementing the genetic algorithm to optimize set partitions, the

popular *crossover* procedure, which requires a pair of "parents" to generate one "child," was time-consuming because of the combinatorial bookkeeping required to ensure that every group of states in the "child" set partition was a connected component. Because the goal of our work is to reduce computational cost of Ca^{2+} release site models, we avoided a *crossover* process in our genetic algorithm implementation and incorporated the *selection* process into the *mutation* process for simplicity and computational effectiveness.

We start with building a discrete probability distribution used in the selection of the next generation of set partitions. The probability mass function (PMF), which indicates the probability $\mathcal{P}_i$ for each "individual," $\mathcal{I}_i$, to be selected, is given as a member of the subsequent generation

$$\mathcal{P}_i = \frac{\mathcal{F}_i}{\sum_i \mathcal{F}_i} \qquad (1 \leq i \leq N_p), \tag{3.29}$$

that is, the probabilities of selection is proportional to the "fitness," $\mathcal{F}_i$. To generate each "child," we start with randomly select a "parent" ($\mathcal{I}$) from the current "population" following the corresponding PMF. For example, the probability of $\mathcal{I}_1$ being selected is 210 times less than the probability that $\mathcal{I}_2$ is selected.

After a parent $\mathcal{I}$ is selected, a *mutation* process begins by randomly selecting and joining a pair of groups of states in the parent partition scheme, and then randomly splitting this aggregated group of states into two new connected groups. Notice that the two groups initially selected is a valid pair only when they are originally connected to each other in the full model, i.e., at least one state in the first group is connected by a transition to at least one state in the second group. As an example, for $\mathcal{I}_1$ in Fig. 3.2 all three groups are connected in the full model so any two groups can be

picked, while for $\mathcal{I}_2$ groups $\{1\}$ and $\{3,4,5,6\}$ are not connected so they are not a valid pair.

In our implementation of the genetic algorithm, we allow multiple mutations of this kind for each partition scheme, selected for the subsequent generation. A geometric distribution is assigned to the number of mutations N_m,

$$Pr(N_m = k) = (1-p)^{k-1}p, \qquad (3.30)$$

where $k = 1, 2, \ldots$ and $p = 0.8$.

This **selection** and **mutation** process continues until the new generation of size N_p is generated, their fitnesses determined as this new generation is used in another cycle of **evaluation**. The genetic algorithm continues until a set partition is found that has a "fitness" $\mathcal{F} \geq 1000$ (reduction error $\mathcal{E}$ less than 0.1% for all luminal $[Ca^{2+}]$) or the algorithm is terminated by reaching the maximum number of allowed iterations (2000).

3.5 Results

In this section we first validate the genetic algorithm implemented in Sec 3.4.2 by showing that the algorithm converges and produces set partition schemes that generate small reduction error. To further demonstrate that the genetic algorithm can be applied to general Ca^{2+} channels, we use this approach to reduce a release site that is composed of several four-state channels (Fig. 3.8) that are activated by cytosolic Ca^{2+} and the activation affinity is regulated by the luminal $[Ca^{2+}]$. We then integrate the reduced Ca^{2+} release site model into the whole cell model simulation and show that the reduced model is a good approximation to the full model.

3.5.1 Reducing Ca^{2+} release site models that are composed of 3-state channels

Fig. 3.7 shows an example of the convergence of the genetic algorithm. We applied the genetic algorithm to reduce a Ca^{2+} release site that is composed of 10 3-state channels (66 states) to a 11-state model. The population size $N_p = 10$ and the reduction error $\mathcal{E}$ was measured for 50 log-spaced $c_{er/sr}$ values range from 100μM to 2000μM. Each column of stars is the 10 individuals of a "generation". The black stars indicate the individual (partition) that produces the smallest error in its generation. The criteria that ends the program is set to be $\mathcal{E} < 10^{-3}$ or 2000 generations generated, whichever satisfied first. In this specific reduction experiment, the program did generate 2000 generations and the minimum $\mathcal{E}$ was 0.0043.
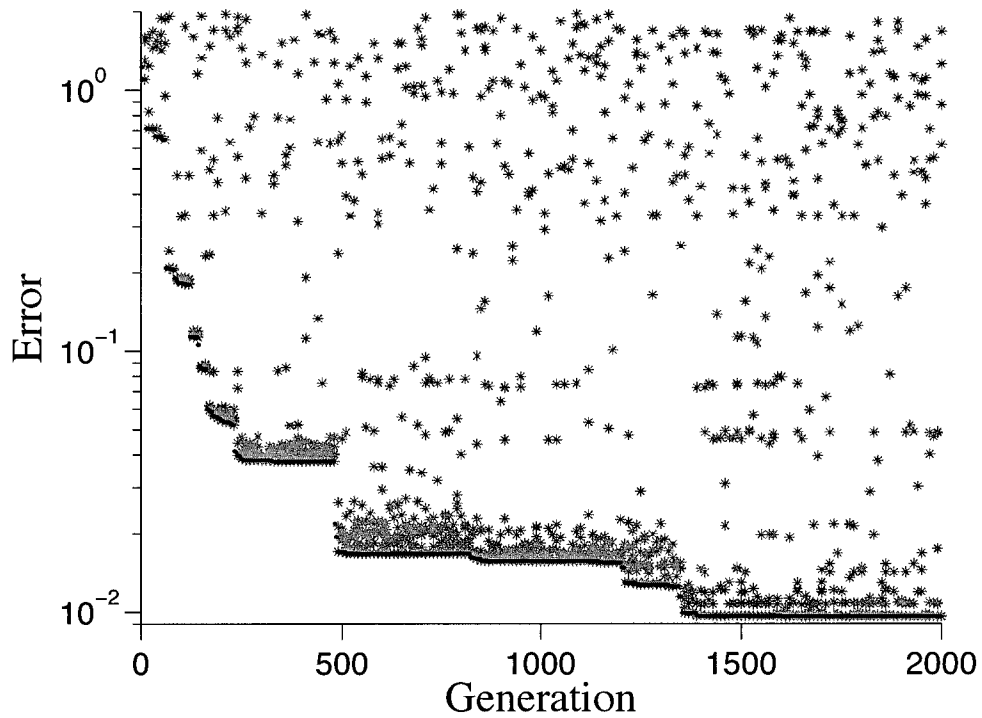
Figure 3.7: A sample evolution record from the genetic algorithm. A Ca^{2+} release site composed of 10 three-state channels is designated to be reduced to a 11-state model. One of every 10 generations is plotted. Each column of stars indicates a generation of 10 individuals and the one that produces the least error is indicated by the black star. Parameters are as in Fig. 3.5.

3.5.2 Reducing Ca²⁺ release site models that are composed of 4-state channels with luminal regulation

In this section we demonstrate that reduced Ca^{2+} release models can replace the full model in the whole cell Ca^{2+} homeostasis model described in Sec. 3.3.1 with good accuracy. To validate that the reduction procedure fits a wide variety of models, we introduce a 4-state model (Fig. 3.8) which is activated by cytosolic Ca^{2+} and the activation affinity is regulated by the luminal $[Ca^{2+}]$.

The 4-state Ca^{2+} channel model is assumed to have a regular or "unsensitized" mode (states C_u, O_u) in which the activation dissociation constant ($K_a = \sqrt{k_a^-/k_a^+}$), is higher than the activation dissociation constant ($K_d = \sqrt{k_d^-/k_d^+}$) of the "sensitized" mode (C_s, O_s). We also assume that the channel is more likely to be in the "sensitized" mode when the ER/SR $[Ca^{2+}]$ is high. The transition rates are $k_i^+(c_{cyt}^d)^2$, $k_j^+ c_{er/sr}^d$ and k_i^-, where $i, j \in \{a, ..., e\}$, with units of reciprocal time. k_i^+ is an association rate constant with units of conc$^{-\eta}$ time^{-1} where η is the cooperativity of Ca^{2+} binding while c_{cyt}^d and $c_{er/sr}^d$ are the domain $[Ca^{2+}]$ experienced by the release site on the cytosol and ER/SR side respectively. Notice that we assume that the Ca^{2+} binding cooperativity ($\eta = 1$) of the channel sensitization (luminal regulation) process is different from the binding cooperativity of the activation process ($\eta = 2$).

$$O_u \underset{k_c^-}{\overset{k_c^+ c_{er}^d}{\rightleftharpoons}} O_s$$

$$C_u \underset{k_e^-}{\overset{k_e^+ c_{er}^d}{\rightleftharpoons}} C_s$$
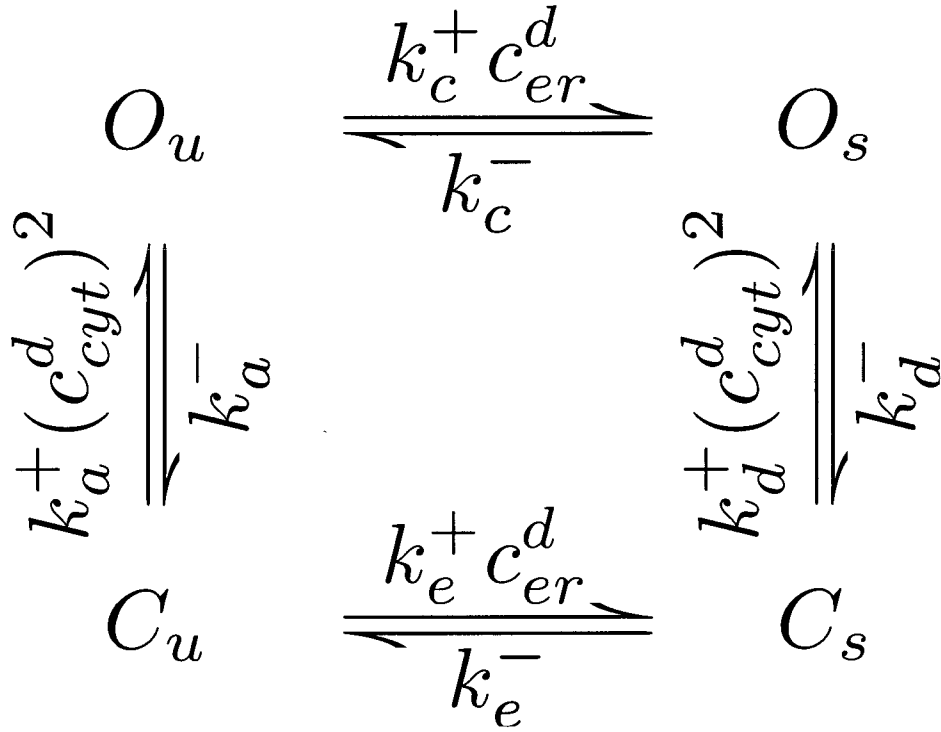
Figure 3.8: Transition diagram of the four-state Ca^{2+} channel model. The channel is activated by cytosolic Ca^{2+} (transitions $C_u \rightarrow O_u$ and $C_s \rightarrow O_s$) and is "sensitized" by ER/SR Ca^{2+} (transitions $C_u \rightarrow C_s$ and $O_u \rightarrow O_s$). Parameters: $k_a^+ = k_d^+ = 4.5 \ \mu M^{-2} ms^{-1}$, $k_c^+ = k_e^+ = 1 \ \mu M^{-1} ms^{-1}$, $k_a^- = k_b^- = 500 \ ms^{-1}$, $c_{cyt} = 0.1 \ \mu M$.

An important motivation of using this 4-state model is that luminal regulation of RyRs is observed in many experiments [Shannon et al., 2004, Stevens et al., 2009] but the detailed mechanism is yet not clear. In this paper, we are interested in how the "sensitization" of the activation of each individual Ca^{2+} channels affects the cooperative gating of the Ca^{2+} release site. Consequently we experiment on different sensitized activation rates as well as the dissociation constant $K_c = k_c^-/k_c^+$ (Fig. 3.8) of the sensitization process. On the other hand, the parameters of the regular or "unsensitized" Ca^{2+} activation were picked consistent with the parameters in [Hartman et al., 2010], where many puff/spark statistics of a group of 10 2-state Ca^{2+}-activated channels were studied.

We assumed that the number of Ca^{2+} release sites is large enough that the distribution of release site states can be well approximated by solving $\pi(t)$ from Eq. 3.18 instead of using Monte Carlo simulation. However, simply substituting π_n for f_n in Eqs. 3.13 and 3.14 will fail because the "fast domain" assumption is a singular limit of the ODE system. Consequently, instead of using Eqs. 3.11 and 3.12 we consider the total cytosolic ($\hat{c}_{cyt}$) and ER/SR $[Ca^{2+}]$ ($\hat{c}_{er/sr}$), which are sums of the bulk and domain concentrations weighted by effective volume ratios,

$$\hat{c}_{cyt} = c_{cyt} + \Lambda_{cyt}^d \bar{c}_{cyt}^d \tag{3.31}$$

$$\hat{c}_{er/sr} = c_{er/sr} + \frac{\Lambda_{sr}^d}{\lambda_{sr}} \bar{c}_{er/sr}^d, \tag{3.32}$$

where, $\bar{c}_{cyt}^d$ and $\bar{c}_{er/sr}^d$ are the given by

$$\bar{c}_{cyt}^d = \sum_{n=0}^{N} \pi_n c_{cyt}^{d,n} \tag{3.33}$$

$$\bar{c}_{er/sr}^d = \sum_{n=0}^{N} \pi_n c_{er/sr}^{d,n}, \tag{3.34}$$

which are the mean values of the cytosolic and SR domain Ca^{2+} concentrations. The effective volume ratios in Eqs. 3.31 and 3.32 are given by

$$\Lambda^d_{cyt} = \frac{V^{d,T}_{cyt}}{V_{cyt}}, \tag{3.35}$$

$$\Lambda^d_{sr} = \frac{V^{d,T}_{sr}}{V_{cyt}}, \tag{3.36}$$

where $V^{d,T}_{cyt}$ and $V^{d,T}_{sr}$ are the effective volumes of the aggregated cytosolic and SR domains, respectively. The equations that balance $\hat{c}_{cyt}$ and $[Ca^{2+}]$ $\hat{c}_{er/sr}$ are given by:

$$\frac{d\hat{c}_{cyt}}{dt} = J^T_{rel} + J_{leak} - J_{pump} + J_{pm}, \tag{3.37}$$

$$\frac{d\hat{c}_{er/sr}}{dt} = \frac{1}{\lambda_{er/sr}} \left(-J^T_{rel} - J_{leak} + J_{pump} \right). \tag{3.38}$$

The total release flux J^T_{rel} is given by

$$J^T_{rel} = \sum_{n=0}^{N} \pi_n \gamma_n v^T_{rel} \left(c^{d,n}_{er/sr} - c^{d,n}_{cyt} \right), \tag{3.39}$$

where $\gamma_n = n/N$, $c^{d,n}_{cyt}$ and $c^{d,n}_{er/sr}$ are given by Eqs. 3.5–3.7, π_n is the probability that a randomly sampled release site has n open channels, and can be found from $\boldsymbol{\pi} = (\pi_0, \pi_1, \cdots, \pi_N)$ by integrating Eq. 3.18.
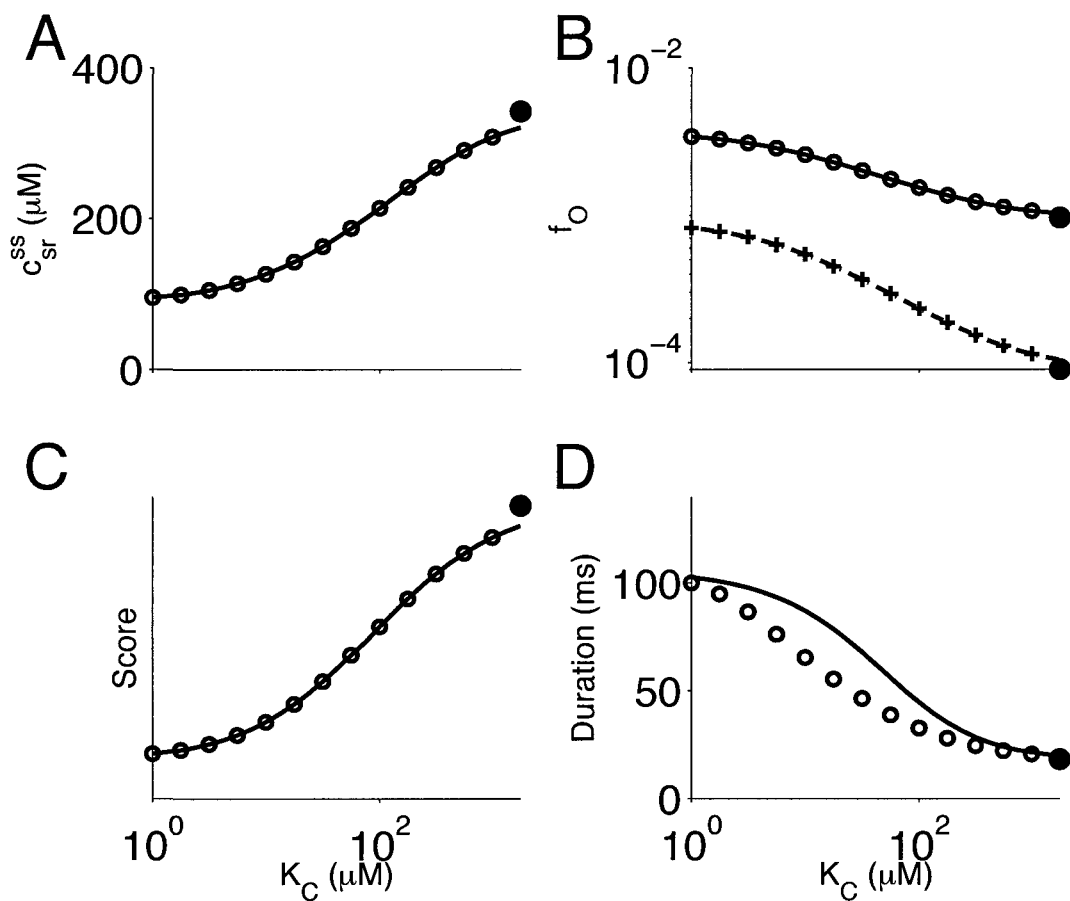
Figure 3.9: Effects of luminal regulation calculated from release sites composed of 10 luminal regulated Ca^{2+} channels. Results from the full release site model and reduced model are shown by lines and empty circles (and crosses in B), respectively. The filled circles show corresponding results from the release site that composed of 10 two-state Ca^{2+} activated channel without luminal regulation. (A) Steady state ER/SR $[Ca^{2+}]$ as a function of K_C. (B) Steady state open probability (dashed line) and the fraction of open channels (solid line) as a function of K_C. (C) Spark scores as a function of K_C. (D) Spark durations as a function of K_C.

Fig. 3.9 shows a comparison of 20 numerical calculations of the stationary dynamics of a Ca^{2+} release site composed of 10 4-state RyRs (286 states, lines) and the corresponding reduced 34 state model (circles and crosses) using different values of the disassociation rate of sensitization K_C. The filled circles show the results of a release site composed of 10 2-state RyRs. When the disassociation rate of sensitization K_C is high enough, the sensitized states are rarely visited and consequently the 4-state model results should approach the 2-state model results. As shown in Fig. 3.9 the 4-state model well approximates the 2-state model when K_C is approximately 1000 μM.

In Fig. 3.9, panel A shows that decreasing the K_C will decrease the bulk SR $[Ca^{2+}]$ and the results calculated from the reduced Ca^{2+} release site model are close approximations to the full model. Fig. 3.9B shows the open probability of a single 4-state channel (dashed line) as a function of K_C. The solid line in Fig. 3.9B shows the fraction of open channels f_O of the 10-channel release site, where

$$f_O = E[N_O]/N, \qquad (3.40)$$

and

$$E[N_O] = \sum_{n=0}^{N} n\pi_n \qquad (3.41)$$

is the average number of open channels per release site. The reduced model gives good approximation to both parameters of the full model (empty circles and crosses). As K_C decreases, both parameters increase which indicates adding the sensitized the states increase the open probability of the channels which further causes a lower steady state SR $[Ca^{2+}]$ which is consistent with Fig. 3.9A. In prior work [Nguyen

et al., 2005] we defined the puff/spark *Score* as

$$Score = \frac{\text{Var}[f_O]}{\text{E}[f_O]} = \frac{1}{N}\frac{\text{Var}[N_O]}{\text{E}[N_O]} \tag{3.42}$$

from which the presence or absence of puff/spark can be assessed. This measure ranges between 0 and 1, and values that are larger than 0.2 indicate presence of robust Ca^{2+} puffs/sparks. Fig. 3.9C, shows the *Scores* of the full model and the reduced model as a function of K_C. The reduced model *Scores* give a close approximation to the full model results.

Notice that the *Score* values are above 0.35 for all K_C values, indicating robust Ca^{2+} puff/sparks present in both the full and reduced model. We further studied the mean duration of spontaneous Ca^{2+} puff/sparks occurring as a function of K_C in the whole cell formulation, shown in Fig. 3.9D. We assume that a transition from $N_O = 4$ to $N_O = 5$ is considered to initialize a puff/spark and a transition from $N_O = 1$ to $N_O = 0$ (all channels closed) terminates the puff/spark. The mean puff/spark duration was calculated using the matrix analytic method described in [Groff and Smith, 2008a]. As K_C decreases, or as the channels are more likely to be in the sensitized, the puff/spark duration increases indicating the luminal regulation of the channel might lead to longer puff/sparks, which is consistent with experimental observations [Gyorke and Gyorke, 1998, Stevens et al., 2009]. Compared to the Ca^{2+} release site model composed of 10 2-state channels (filled circle in Fig. 3.9D), the average puff/spark duration of a release site composed of the same number of 4-state channels can be up to four times (when $K_C = 1$) longer. While in prior work [Hao et al., 2009], similar comparison to a Ca^{2+} release site composed of Keizer-Levine model [Keizer and Levine, 1996] and its corresponding reduced model gave good agreement, in this new study the reduced model tends to slightly underestimate the

puff/spark durations.

3.6   Discussion

We have implemented and validated a novel genetic algorithm based searching technique to find reduced models that produce moderate errors for Ca^{2+} release sites without time-scale separation. Given a full model and the designated size of the reduced model, this algorithm samples and evolves a population of set partitions, each corresponding to a potential scheme for state aggregation, toward the partitions that lead to reduced models which approximate the full model on the behaviors of interest. A Ca^{2+} release site composed of 10 four-state channels that are activated by the cytosolic Ca^{2+} and regulated by luminal Ca^{2+} is reduced by this technique and the steady state responses of the reduced model well approximates the full model in the minimal whole cell homeostasis environment proposed in Sec 3.3.1 (Fig. 3.9).

When a Ca^{2+} release site model is reduced, the resulting models usually have significantly fewer states, which is inevitably accompanied by losing some transition information and different state aggregation schemes may preserve different information. A main benefit from using genetic algorithms is that the **evaluation** function is flexible enough to pick state aggregation schemes that maximize any information that is of specific interest to the user. In this report, for example, we are interested in how luminal regulation affects the spark behavior of the Ca^{2+} release site. The **evaluation** function is consequently designed to assign higher fitness to the partitions which generate small errors in a wide range of ER/SR $[Ca^{2+}]$. As another example, if the spark frequency is crucial in some study, we can conveniently edit the **evaluation** function to calculate the spark frequency of each reduced model generated from partition $\mathcal{I}$ and assign higher *fitness* to the ones that better approximate the full model

spark frequency.

Although this genetic algorithm based reduction technique is very flexible, it is important to note that this procedure is more time-consuming compared to the fast/slow reduction technique [Hao et al., 2009], because for each *Individual* $\mathcal{I}$, we reduce the full model following the aggregation scheme, calculating the reduction error and then assign fitness. Because we use 10 as the population size and 2000 generations are generated untill the program terminates, the total time consumed is approximately 20000 fold that of the fast/slow procedure. Fortunately, for any specific objective assigned to the reduced model, the reduction procedures need to execute only once and the reduced release site model are potentially able to save significantly more time in the whole cell simulations.

An interesting phenomenon we observed when reducing Ca^{2+} release site models with the genetic algorithm based approach is that the state aggregation schemes that result in small reduction errors tend to be "heavy headed." That is, these "good" partitions usually feature one large group which collects more than 50% of the states while other groups contain significantly less (some times only one or two) states. Moreover, the states aggregated in the small groups are highly likely to be the states that are visited less often in the full model and this phenomenon exclusively exists in all Ca^{2+} release site reduction procedures. This observation is a good reason for the fact that the generator matrix associated with the reduced model $\hat{Q}$ are very ill conditioned. On the other hand, his observation suggests that we can generate a biased initial population to accelerate the evolution procedure.

Chapter 4

Langevin description of the stochastic dynamics of calcium release sites

4.1 Summary

Compositionally defined Markov chain models have been used to study the relationship between single channel gating of intracellular calcium (Ca^{2+}) channels and the stochastic dynamics of Ca^{2+} "puffs" and "sparks," intracellular Ca^{2+} release events that arise from the cooperative activity of clusters of Ca^{2+} channels. In such models, the transition probabilities of individual channels depend on the local Ca^{2+} concentration and thus the state of the other channels. Consequently, Markov chain models of Ca^{2+} release sites often possess intractably large state spaces that impede computational analysis. To overcome this difficulty, we derived a general Langevin formulation for the stochastic dynamics of Ca^{2+} release sites composed of a large number of intracellular Ca^{2+} channels. We validate this Langevin formulation by comparison to

Markov chain simulations and perform benchmark simulations that demonstrate its computational efficiency for single channel models with 2 and more states and release sites composed of 20 to 200 channels.

4.2 Introduction

Stochastic dynamics are becoming an increasingly important factor to be incorporated into the study of genetics, computational cell biology and system biology. For example, expression and nonexpression of cell surface pili during an infection of the urinary tract by *E. coli* occurs in a stochastic fashion [Low et al., 2001]. In cardiac myocytes, membrane depolarization during the action potential causes L-type Ca^{2+} channels to open, and Ca^{2+} current through these channels causes the release of a larger amount of Ca^{2+} from the sarcoplasmic reticulum, a process known as Ca^{2+}-induced Ca^{2+} release (CICR) [Beuckelmann and Wier, 1988]. Because different groups of RyRs experience different local Ca^{2+} concentrations, they stochastically gate in a manner that depends on whether nearby sarcolemmal Ca^{2+} channels have recently been open or closed [Stern, 1992]. In cell signal transduction which is restricted in a small volume, stochastic molecular fluctuations inevitably arise because of small molecular numbers. In this case, stochastic modeling approaches are required to faithfully reproduce detailed fluctuations in the number of molecules since even a change by a few molecules could produce a substantial change in concentration.

In the context of stochastic modeling of Ca^{2+} release sites, that is, clusters of Ca^{2+} release channels, the stochastic signaling dynamics are often modeled by Markov chain models because the distributions of the open/close dwell time follow exponential or Erlang distributions. For a single channel, Markov chain models specify each functional state and the transition rates between states are selected such that the

model reproduces experimental single channel records. Ideally, Ca^{2+} release sites are compositionally defined by specifying how multiple individual channels interact and thereby lead to concerted stochastic phenomena such as stochastic Ca^{2+} excitabiliy, known as puffs and sparks [Berridge, 1993b, Cheng et al., 1996, Nguyen et al., 2005].

In mechanistic models of cell signal transduction, gene regulatory networks and calcium signaling, etc., there is often a compositional structure whereby a number of same or similar components/modules are combined to create a larger system. These systems can usually be idealized as stochastic automata networks (SANs) where each component has many functional states and the transitions rates between them depend on the states of other components [Nguyen et al., 2005]. However, the compositional nature of signaling complexes results in a combinatorial state-space explosion that is an important practical consideration that can cause modeling approaches with state-dependent computational efficiency to become intractable.

While the dynamics of any individual Ca^{2+} release site can theoretically be obtained by Monte Carlo simulation regardless of model complexity, these simulations can be computationally intensive. In prior work, Smith and collaborators have developed and validated modeling approaches to accelerating multiscale calculations of calcium signaling in cardiac myocytes. For example, probability density and moment closure approaches [Williams et al., 2007, Williams et al., 2008], which assume large number of Ca^{2+} release sites, small number of states per release site and fast subspace $[Ca^{2+}]$ dynamics, can be approximately 1000 times faster than Monte Carlo simulations. Some approaches, on the other hand, focus on automated reduction of Ca^{2+} release site models where the release sites are compressed into a tractable size while the physiological gating and interaction properties of the channels are preserved [LaMar et al., 2011, Hao et al., 2009]. In this chapter we present a Langevin formulation of the stochastic dynamics of Ca^{2+} release sites that is an alternative model reduction

technique in the context of multiscale models of Ca^{2+} handling [Hinch et al., 2006] applicable when a cell has a very large number of Ca^{2+} release sites, but a moderate number of channels per release site. We validate our Langevin formulation of compositionally defined Ca^{2+} release sites by comparison of full and reduced signaling complexes that demonstrates that the statistic of puffs and sparks are indeed captured by the reduced model.

In the context of calcium signaling, prior work utilizing Langevin formulation has largely focused on the case where noise terms are added to a "common pool" compartmental model [Shuai and Jung, 2002a] and the intracellular calcium channels are assumed to stochastically gate in an independent fashion while coupled to the bulk cytosolic $[Ca^{2+}]$. In this prior work the model formulation accounts for the finite number of calcium release units in the cell, but the properties of the stochastic driving force in the Langevin equation is best interpreted as representing molecular fluctuations due to the finite number of calcium release sites when each release site is composed of a single channel (initially a continuous ordinary differential equation model).

The Langevin formulation presented here is distinguished from this previous work in the following ways. First, the channels in each Ca^{2+} release site are coupled, i.e. the opening or closing of the channels affects the transition rates of its neighboring channels. Second, prior work by Jung and collaborators essentially added a noise term to a single Hodgin-Huxley-style gating variable equation that represents the slow Ca^{2+} inactivation of an intracellular Ca^{2+} channel, that is derived under the assumption that Ca^{2+} activation of intracellular Ca^{2+} channels is a much faster process than Ca^{2+} inactivation [De Young and Keizer, 1992, Li and Rinzel, 1994]. In our approach, noise terms corresponding to every channel transition are added and no quasistatic assumption is made. Third, the implementation of the Langevin descrip-

tion is automated in the manner that any Markov model of a single Ca^{2+} channel can be transformed to the corresponding Langevin equations in an automated fashion. Our Langevin formulation is applicable and efficient in situations where the number of release sites in a cell is large enough that molecular fluctuations are negligible, while the number of channels per release site is moderate.

As an application of this approach to model reduction, we use a Langevin description of 20 and 60 release sites to address specific questions about the manner in which Ca^{2+} inactivation of single channel models influences spark/puff statistics. We focus on how correlations in spark amplitude, duration, and inter-event interval associated with successive release events can carry a signature of the presence and time-constants of inactivation processes occurring at the molecular level.

4.3 Model Formulation

4.3.1 Conventional Markov chain model with Ca^{2+} activation

To introduce the model formation, consider a minimal two-state model for ryanodine/inositol 1,4,5-trisphosphate receptor (RyR/IP3R) where only Ca^{2+} activation presents. A Markov chain description of such stochastic two-state single-channel gating is diagrammed by the following transition-state diagram:

$$\text{(closed) } C \underset{k^-}{\overset{k^+ c^\eta}{\rightleftharpoons}} O \text{ (open)} \tag{4.1}$$
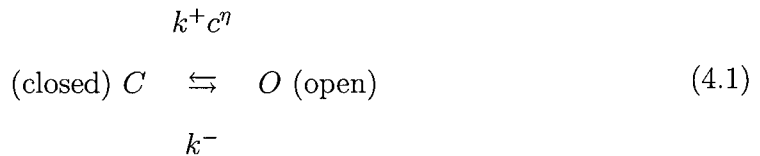
where $k^+ c^\eta$ and k^- are transition rates with units of reciprocal time, k^+ is an association rate constant with units of $conc^{-\eta}$ $time^{-1}$, η is the cooperativity of Ca^{2+} binding,

and c is the local $[\text{Ca}^{2+}]$ that may be either constant or a function of time. We make the assumption of instantaneous mean-field coupling [DeRemigio and Smith, 2005], and further assume that the local $[\text{Ca}^{2+}]$ experienced by all channels of a release site is
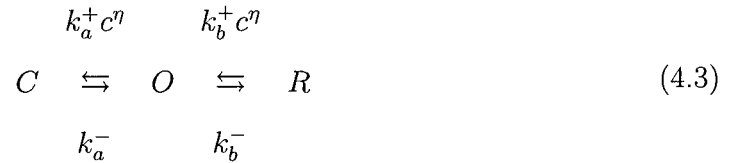
$$c = c_\infty + N_O c_*, \tag{4.2}$$

where N_O is the number of open channels, c_∞ is the fixed background $[\text{Ca}^{2+}]$ and c_* is the $[\text{Ca}^{2+}]$ elevation caused by each opening channel.

A minimal model which considers both Ca^{2+} activation and inactivation is also used below to evaluate the Langevin model. The transition diagram of this 3-state model is

$$
\begin{array}{ccccc}
 & k_a^+ c^\eta & & k_b^+ c^\eta & \\
C & \rightleftharpoons & O & \rightleftharpoons & R \\
 & k_a^- & & k_b^- &
\end{array}
\tag{4.3}
$$

where R indicates the long-lasting closed (refractory) state entered via Ca^{2+} inactivation. We assume that η, the cooperativity of Ca^{2+} binding, is the same for both the activation and inactivation process. Under the assumption of instantaneous mean-field coupling Eq. 4.2, the domain $[\text{Ca}^{2+}]$ for the $O \rightarrow R$ transition is greater than the $[\text{Ca}^{2+}]$ for the $C \rightarrow O$ transition due to the self-induced increase in local Ca^{2+} experienced by the channel when open.

The transition-state diagram Eq. 4.3 becames

$$
\begin{array}{ccccc}
 & k_a^+ c_\infty^\eta & & k_b^+ (c_\infty + c_*)^\eta & \\
C & \rightleftharpoons & O & \rightleftharpoons & R \\
 & k_a^- & & k_b^- &
\end{array}
\tag{4.4}
$$
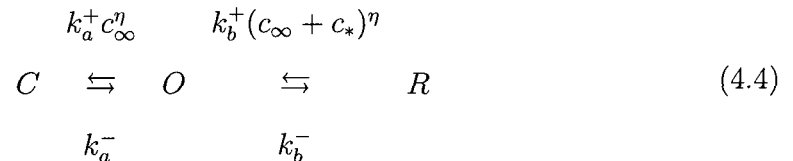
and the corresponding infinitesimal generator matrix is

$$Q = (q_{ij}) = \begin{pmatrix} \Diamond & k_a^+ c_\infty^\eta & 0 \\ k_a^- & \Diamond & k_b^+ (c_\infty + c_*)^\eta \\ 0 & -k_b^- & \Diamond \end{pmatrix} \qquad (4.5)$$

where the diagonal elements ($\Diamond$) are such that the row sum of the matrix Q are zero.

A vector π is the stationary distribution of a Markov chain if it satisfies

$$\pi Q = 0 \qquad \text{subject to} \qquad \pi e = 1 \qquad (4.6)$$

where πe is an inner product and e is a commensurate column vector of ones. The stationary distribution of Markov chain models can be most conveniently either solved analytically or numerically depending on the size of the model. For example, for the two-state single channel model (Eq. 4.1), we can analytically solve Eq. 4.6 to find

$$\pi_C = \frac{k^+ c^\eta}{k^+ c^\eta + k^-} \qquad \text{and} \qquad \pi_O = \frac{k^-}{k^+ c^\eta + k^-}.$$

If a collection of N such Ca^{2+} channels are co-localized at a release site and coupled via local $[Ca^{2+}]$, the transition diagram for the release site as a collective entity is

$$
\begin{array}{ccccccccc}
 & Nk^+ c_0^\eta & & (N-1)k^+ c_1^\eta & & 2k^+ c_{N-2}^\eta & & k^+ c_{N-1}^\eta & \\
0 & \rightleftharpoons & 1 & \rightleftharpoons & \ldots & \rightleftharpoons & N-1 & \rightleftharpoons & N \\
 & k^- & & 2k^- & & (N-1)k^- & & Nk^- &
\end{array}
\qquad (4.7)
$$
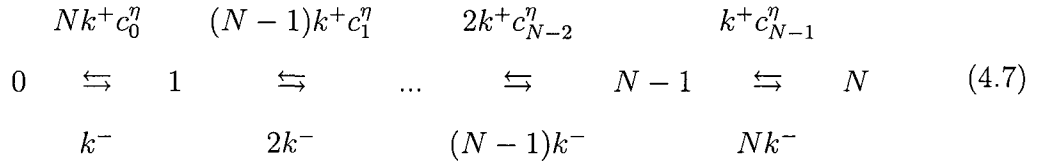
where the states of the system, $\{0, 1, \ldots, N\}$, indicate the number of open channels, and the local $[Ca^{2+}]$ used in each Ca^{2+}-mediated transition is $c_n = c_\infty + N_O c_*$ (Eq. 4.2).

99

In general, the transition between two neighboring states takes the form:

$$n \quad \underset{nk^-}{\overset{(N-n)k^+ c_n^\eta}{\rightleftharpoons}} \quad n+1 \qquad (4.8)$$
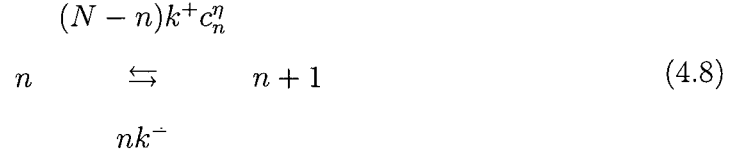
where the number of open channels n is a integer between 0 and N.

The absence or presence of the localized Ca^{2+} elevations known as puffs and sparks is one of the most important properties of Ca^{2+} release sites. In prior work we have defined a response measure dubbed the puff/spark *Score*, which is defined as the index of dispersion of the fraction of open channels [Nguyen et al., 2005]:

$$Score = \frac{\mathrm{Var}[f_O]}{\mathrm{E}[f_O]} = \frac{1}{N} \frac{\mathrm{Var}[N_O]}{\mathrm{E}[N_O]} \qquad (4.9)$$

This measure ranges between 0 and 1, and values that are close to or larger than 0.25 indicates presence of robust Ca^{2+} puffs/sparks.

4.3.2 The Langevin description of Ca^{2+} release sites composed of two-state channenls

When the number of channels, N, is large, Eq. 4.8 indicate that the number of open channels is changing at rate

$$\frac{dN_O}{dt} = \alpha(N - N_O) - \beta N_O, \qquad (4.10)$$

where the backward transition rate $\beta = k^-$ is a constant while the forward transition rate $\alpha = k^+(c_\infty + N_O c_*)^\eta$ is a function of N_O. By dividing both sides by N, the number of open channels, we can derive an ordinary differential equation (ODE) for

the fraction of open channels $f_O(t)$ taking values between 0 and 1,

$$\frac{df_O}{dt} = \alpha(1 - f_O) - \beta f_O. \tag{4.11}$$

In the Langevin description, fluctuations in the fraction of open channels due to the finite value of N are described using a stochastic ODE that takes the form

$$\frac{df_O}{dt} = g(f_O) + \xi(t). \tag{4.12}$$

In this equation, $g(f_O) = \alpha(1 - f_O) - \beta f_O$ corresponds to the deterministic dynamics previously described. These deterministic dynamics are supplemented with a rapidly varying forcing term, $\xi(t)$, which is a random variable parameterized by time. A solution of Eq. 4.12, $f_O(t)$, must satisfy the equation for a particular instantiantion of $\xi(t)$. Alternatively, if the statistics of $\xi(t)$ are given, $f_O(t)$ is a new random variable formally defined by Eq. 4.12 that has a probability density function that can be either calculated analytically or integrated numerically from the Fokker-Plank equation that corresponds to Eq. 4.12.

In order to use a Langevin equation of the form of Eq. 4.12 to simulate a large number of ion channels, we must make an appropriate choice for both the deterministic function $g(f_O)$ as well as the statistics of the random variable $\xi(t)$. Recalling the average rate equation for the dynamics of the open fraction of channels, Eq. 4.11, we write

$$\frac{df_O}{dt} = \alpha(1 - f_O) - \beta f_O + \xi(t). \tag{4.13}$$

An appropriate choice for ξ is a fluctuating function of time that has zero mean,

$$\langle \xi(t) \rangle = 0,$$

and an autocorrelation function given by

$$\langle \xi(t)\xi(t') \rangle = \gamma(f_O)\,\delta(t - t'),$$

where $\langle \rangle$ indicates a "trial" or "ensemble average." δ is the Dirac delta function and $\gamma(f_O)$ is a "two-time" covariance that depends on the system state, f_O. It can be shown [Keizer, 1987] that the variance $\gamma(f_O)$ is inversely proportional to N and proportional to the sum of the rates of both the $O \rightarrow C$ and $C \rightarrow O$ transitions, that is,

$$\gamma(f_O) = \frac{\alpha(1 - f_O) + \beta f_O}{N}. \tag{4.14}$$

An appropriate choice for $\xi(t)$ is thus $\xi(t) = \sqrt{\gamma(f_O)}\Delta B(t)$ where the ΔB are the increments of a Wiener process (discussed further in section 4.4.1).

4.3.3 The Langevin description of Ca^{2+} release sites that consist of channels with 3 or more states

When a release site consists of Ca^{2+} channels with N states ($N \geq 3$), the corresponding Langevin equation is

$$\frac{d\boldsymbol{f}}{dt} = Q\boldsymbol{f} + \boldsymbol{\xi}(t). \tag{4.15}$$

In Eq. 4.15 the state of the release site is represented by the $N \times 1$ column vector $\boldsymbol{f}$ where each element f_i indicates the fraction of the channels that are in state i. By conservation of probability we have $\mathbf{e}^T\boldsymbol{f} = 1$ where $\mathbf{e}$ is a $N \times 1$ column vector of ones, and the $N \times 1$ column vectors $\boldsymbol{\xi}(t)$ is the rapidly varying force with mean zero

$$\langle \boldsymbol{\xi}(t) \rangle = 0$$

and two-time covariance

$$\langle \boldsymbol{\xi}(t)\boldsymbol{\xi}^T(t')\rangle = \Gamma\left(\boldsymbol{f}\right)\delta\left(t - t'\right) \tag{4.16}$$

where $\Gamma = (\gamma_{ij})$ so that written in scalar form we have $\langle \xi_i(t)\xi_j(t')\rangle = \gamma_{ij}\left(\boldsymbol{f}\right)\delta\left(t - t'\right)$.

Because the Ca^{2+} release site models are composed of identical single channel models whose transitions take the form of (Eq. 4.5), the off-diagonal elements of the two-time covariance matrix $\Gamma = (\gamma_{i,j})$ take the form

$$\gamma_{ij} = -\frac{q_{ij}f_i + q_{ji}f_j}{N} \qquad (i \neq j) \tag{4.17}$$

$$\gamma_{ii} = -\sum_{j\neq i}\gamma_{ij} \tag{4.18}$$

where the diagonals are non-negative ($\gamma_{ii} \geq 0$) and the off-diagonals are non-positive ($\gamma_{ij} \leq 0, i \neq j$). In Matlab notation we can write the matrix Γ as

$$\Gamma = \bar{\Gamma} - \mathrm{diag}\left(\bar{\Gamma}\mathbf{e}\right)$$

where

$$\bar{\Gamma} = -\left(D_f\bar{Q} + \bar{Q}^T D_f\right).$$

The diagonal entries of matrix D_f are the values of $\boldsymbol{f}$, and $\bar{Q}$ is the generator matrix Q with the diagonal zeroed out: $\bar{Q} = Q - \mathrm{diag}\left[\mathrm{diag}\left(Q\right)\right]$. The simulation of this random vector $\boldsymbol{\xi}(t)$ will be discussed in Sec. 4.4.2.

4.4 Results

4.4.1 Simulation of the Langevin equations for two-state channels

While the trajectory of a real Wiener process, $B(t)$, is differentiable nowhere, it can be simulated by numerically integrating a piecewise constant approximation to the Wiener increment $\Delta B(t)$ where $\langle \Delta B(t) \rangle = 0$ and

$$\langle \Delta B(t) \Delta B(t') \rangle = \begin{cases} 1/\Delta t & t \in [t, t + \Delta t] \\ 0 & \text{otherwise} \end{cases}$$

The smaller Δt is, the closer our simulation approximates a real Wiener process.

The dynamical part of the SDE is integrated by the forward Newton's method where the time step is also chosen as Δt so that both the deterministic and stochastic part of the SDE are updated for each iteration. Consequently, the general formula for generating a random trajectory of f_O is:

$$f_O(i+1) = \Delta t \left(\alpha \left(1 - f_O(i) \right) - \beta f_O(i) + \sqrt{\frac{\alpha \left(1 - f_O(i) \right) + \beta f_O(i)}{N}} \eta \right), \quad (4.19)$$

where N is the total number of channels simulated and η is a Gaussian random variable with mean 0 and variance 1.

The Langevin equations of Ca^{2+} release sites that consist of 20 to 200 two-state channels (Eq. 4.1) are simulated and when the coupling strength c_* is such that the *Score* is optimized, robust puffs/sparks are observed (Fig. 4.1 A) as seen in the Monte Carlo simulations of corresponding Markov chain models (Fig. 4.1 B). The distributions of the number of open channels are compared to the steady state distributions of

the corresponding Markov chain model and show agreement for as few as 20 channels (Fig. 4.1 C). This agreement improves slightly as the number of channels simulated grows and all Langevin results more or less overestimate the probability of fewer open channels.
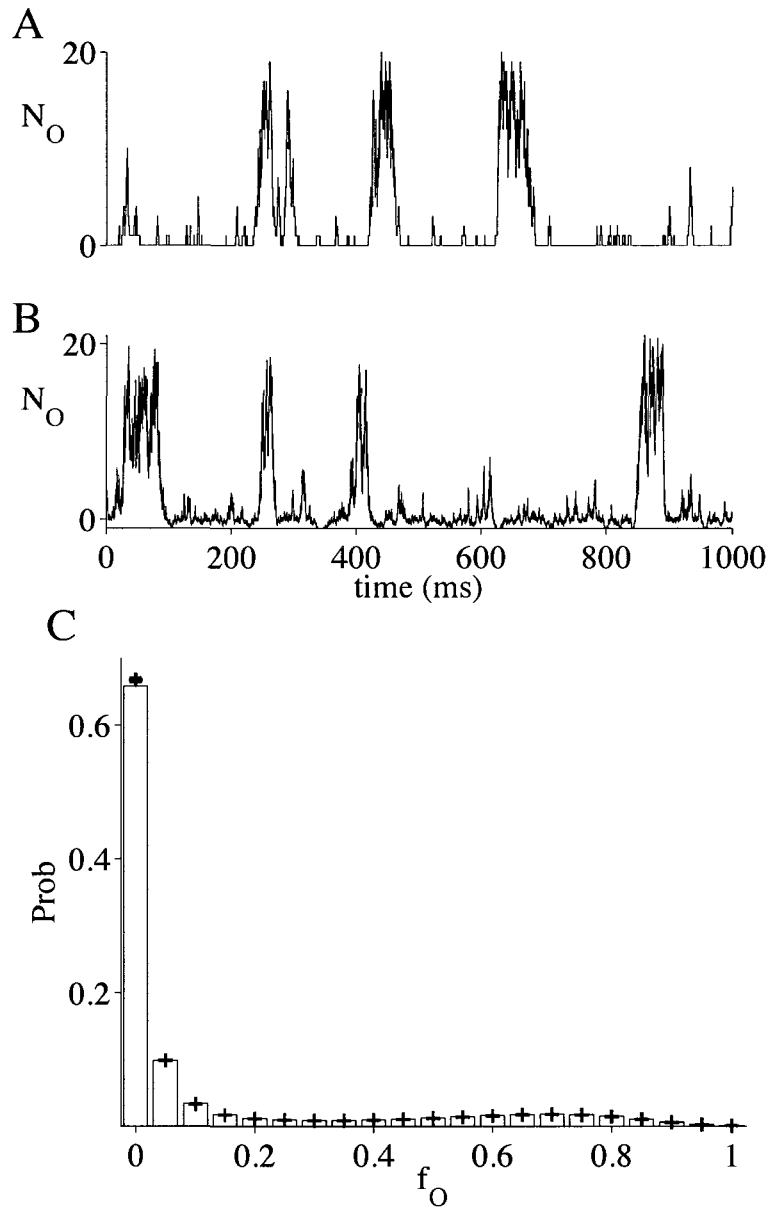
Figure 4.1: (A) Example Monte Carlo simulation of a Markov chain model of 20 two-state Ca^{2+} channels (Eq. 4.1). (B) Example simulation of the corresponding Langevin equation. (C) Stationary distribution of the Markov chain model (histogram) and probability distribution of the fraction of open channels for 50 Langevin simulations (mean values are shown by crosses and error bars indicate standard deviation) for 20 two-state Ca^{2+} channels. Parameters: $k^+ = 1.5\ \mu M^{-\eta}ms^{-1}$, $k^- = 0.5\ ms^{-1}$, $c_* = 0.06\ \mu M$, $\eta = 2$.
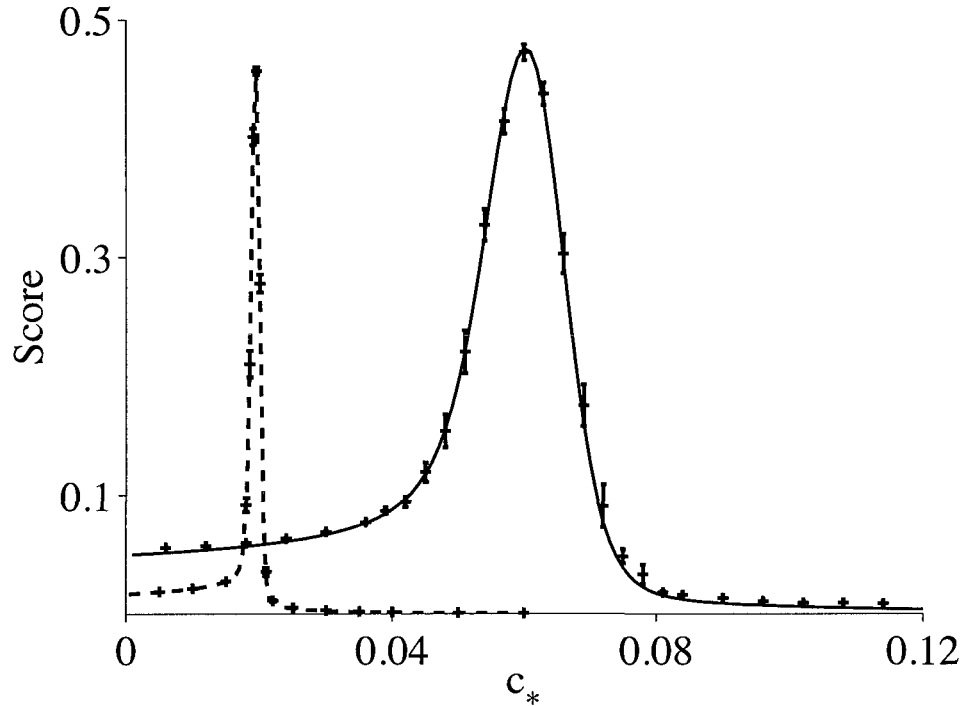
Figure 4.2: The Ca^{2+} puff/spark *Score* as a function of the coupling strength (c_*) when the number of two-state Ca^{2+}-activated channels (Eq. 4.1) is $N = 20$ (solid line) and $N = 60$ (dotted line) calculated from the Markov chain model stationary distributions. For comparison the corresponding Langevin equations were simulated and the dependence of the puff/spark *Score* on c_* was calculated using the mean and variance of the fraction of open channels for this stationary stochastic process. Error bars indicate mean and 95% confidence interval of *Scores* for 10 trials of the Langevin simulation using integration time step 10^{-4} second, each of which is 10 seconds in duration. Parameters are as in Fig. 4.1

Three possible mechanisms that contribute to puff and spark termination observed in prior works include: stochastic attrition, ryanodine/IP_3 receptor inactivation and luminal regulation [Stern and Cheng, 2004, Groff and Smith, 2008a]. Because inactivation and luminal regulation are not present in the two-state Ca^{2+} channel model (Eq. 4.1), the termination of the puffs/sparks in Fig. 4.1 are due to stochastic attrition. In prior work utilizing Markov chain models, it was found that for a given number of channels N, there often exists a range of coupling strengths (c_*) that gives rise to puff/sparks terminated by stochastic attrition. Furthermore, the range for c_* that leads to puff/sparks narrows as N increases [DeRemigio and Smith, 2005]. To determine whether the number of two-state channels at a Ca^{2+} release site and coupling strength affect the dynamics of stochastic attrition in the Langevin simulations is similar to the Markov chain models, the *Score* of release sites composed of 20 and 60 channels was calculated using both approaches and plotted as a function of coupling strength c_*. Fig. 4.2 shows that the *Scores* of the Langevin model well-approximate the *Scores* of the Markov chain model, thereby validating the Langevin formulation and our numerical implementation of the stochastic model.
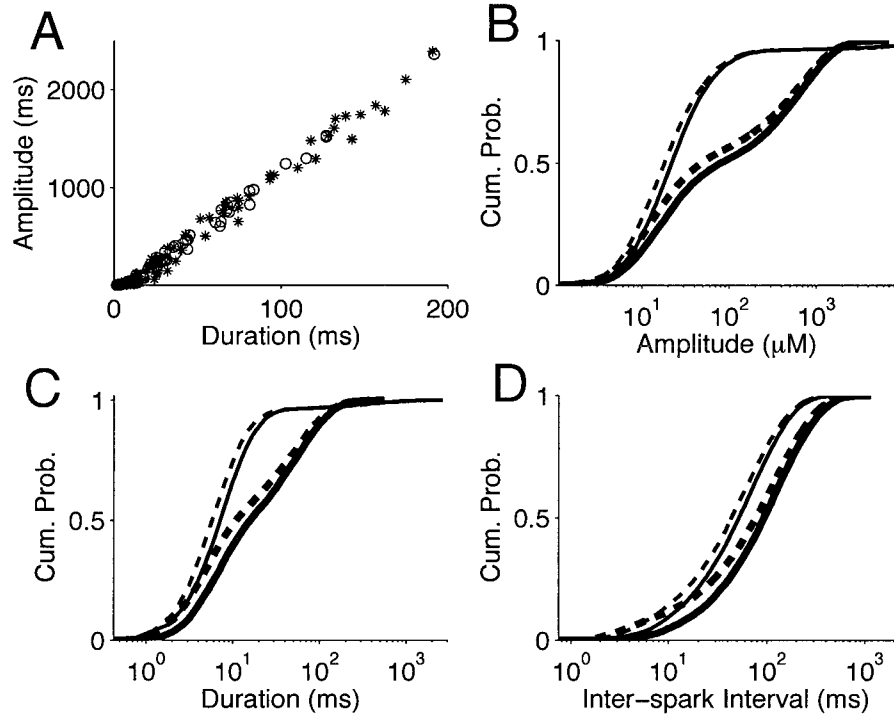
Figure 4.3: (A) Linear relationship between puff/spark amplitude and puff/spark duration is shown using the Langevin equation (circles) and Markov chain models (stars) of 20 two-sate (Eq. 4.1) Ca^{2+} channels. (B, C, D) The cumulative probability distribution of the puff/spark amplitude (B), duration (C) and inter-event intervals (D) observed in the Monte carlo simulations of the Markov chain (solid lines) and Langevin (dashed lines) models. Thick and thin lines correspond to N=20 and 60 two-state Ca^{2+} channels, respectively. Coupling strengths c_* are selected such that the puff/spark *Scores* are high (see Fig. 4.2). $c_* = 0.06$ μM and 0.0194 μM for the 20-channel and 60-channel release sites, respectively, . Other parameters as in Fig. 4.1.

We also compared the Langevin and Markov chain models with respect to three additional puff/spark statistics. As demonstrated diagrammatically in Fig. 4.4, we define the *Duration* (D) of a Ca^{2+} release event to be the random period of time between event initiation (after a release site makes a transition from $N_O = 0 \to 1$ open channels, shown by black dots in Fig. 4.4) and termination (when all channels close via a final $N_O = 1 \to 0$ transition). The *Amplitude* (A) of a Ca^{2+} release event is the integrated area under the $N_O(t)$. In confocal microscopy experiments, puff/spark amplitude is usually defined as the normalized fluorescence $\Delta F/F_0$, where F_0 denotes the resting fluorescence [Song et al., 1997]. Assuming a linear relationship between N_O and the Ca^{2+} flux, this definition of puff/spark amplitude reflects the cumulative amount of Ca^{2+} released from the ER/SR during an event.

To compare these puff/spark statistics in the Markov chain and Langevin models, 20 and 60 Ca^{2+} channels are simulated until 10^5 large events (where the number of open channels is transiently greater or equal to five, $N_O \geq 5$) are observed. The coupling strength (c_*) used is such that the puff/spark *Sores* of the release sites are optimized.

In Fig. 4.3 (A), the amplitudes of Ca^{2+} release events from the 20-channel release site are plotted as a function of the duration. Fig. 4.3 (B, C, D) shows the cumulative distribution functions (CDFs) for the puff/spark amplitudes (A), puff/spark durations (D) and puff/spark intervals (I) respectively. Consistent with Fig. 4.1, all three CDFs obtained from the Langevin model sit above the Monte Carlo result, indicating that more small events take place in Langevin simulations. In spite of this slight over estimate of the number of small events, the Langevin results generally agree with the Monte Carlo simulation results for as few as 20 channels (thick lines). Furthermore, this agreement is improved when the release site is composed of 60 channels (thin lines). Note that, as the number of channels simulated increases, the number of

110
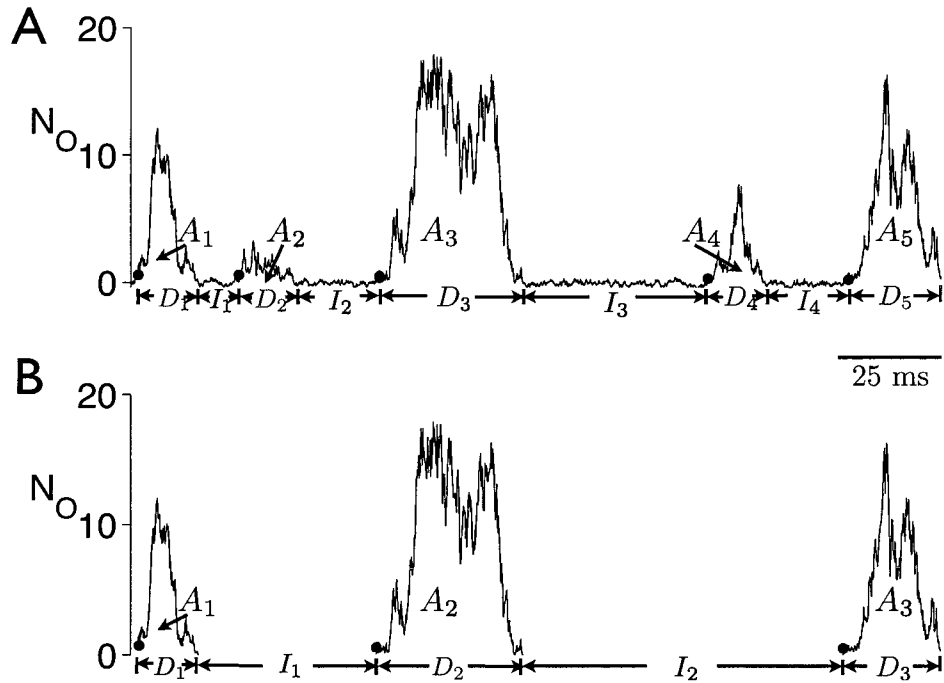
Figure 4.4: (A) a sample simulation of 20 three-state Ca^{2+} channels (Eq. 4.4) using Langevin description that includes five Ca^{2+} release events with amplitudes: $A_1 = 66.35$ ms, $A_2 = 15.05$ ms, $A_3 = 315.98$ ms, $A_4 = 28.79$ ms, and $A_5 = 130.55$ ms. (B) When threshold of puff/sparks is set to be 30 ms, only three events whose amplitudes are above the threshold are selected for further study.

states visited during Monte Carlo simulations using Gillespie algorithm also increases, consequently the run time for simulating the Makov chain model grows dramatically. In contrast, the computation cost of simulating the Langevin model doesn't change. For the simulations of 20 channels shown in Fig. 4.3, the Langevin simulation takes approximately twice as long as the Monte Carlo simulations. When 60 channels are simulated, the Langevin approach takes about 70% of the time that was consumed by Monte Carlo simulations.

4.4.2 Simulating Langevin equations of Ca²⁺ release sites that are composed of channels with 3 or more states

Similar to simulating the Langevin equation for two-state channels, when a Ca²⁺ release site is composed of $m(m > 2)$ states, we can produce a piecewise constant approximation to the rapidly fluctuating force $\boldsymbol{\xi}(t)$, which is a $m \times 1$ vector, with specified two-time covariance consistent with Eq. 4.16,

$$\langle \boldsymbol{\xi}(t)\boldsymbol{\xi}^T(t') \rangle = \begin{cases} \Gamma/\Delta t & t' \in [t, t + \Delta t] \\ 0 & \text{otherwise,} \end{cases} \tag{4.20}$$

where Γ is a $m \times m$ covariance matrix and $\boldsymbol{\xi}\boldsymbol{\xi}^T$ is a outer product. This is done by peforming a Cholesky factorization at each time step

$$\Gamma = A^T A,$$

where Γ (and thus A) are functions of the system state $\boldsymbol{f}(t)$ through Eqs. 4.17 and 4.18 and we have dropped the index i for clarity. The values for the piecewise constant fluctuating force are given by

$$\boldsymbol{\xi} = A^T \Delta \boldsymbol{B} \tag{4.21}$$

where the elements of $\Delta \boldsymbol{B}$ are independent and identically distributed normal random variables with mean zero and variance $1/\Delta t$. (If Γ is $m \times m$, there are of course m elements of $\Delta \boldsymbol{B}$ produced for each time step.)

Fig. 4.5 shows simulation trajectories for Ca²⁺ release sites composed of 20 three-state channels (Eq. 4.4). The parameters in these simulations are selected such that robust puff/spark events are observed (*Score* = 0.38 and 0.42 for panels A and B, respectively) when the Ca²⁺ inactivation affinity is sufficiently high (i.e., the dissoci-

ation constant $K_b = (k_b^-/k_b^+)^{1/\eta} = 0.58~\mu M$), but are tonically open ($Score < 0.1$) when the channels include low-affinity Ca^{2+} inactivation ($K_b = 5.8\mu M$). In prior work, Groff and Smith [Groff and Smith, 2008a] showed that high-affinity Ca^{2+} inactivation may facilitate puff/spark termination in two distinct ways depending on the Ca^{2+} inactivation rates. Fig. 4.5 A and B reproduce these puff/spark dynamics with two different termination mechanisms (see legend). Note that the variance of the number of refractory channels (N_R) is small when the inactivation rate is slower and sparks are terminating by stochastic attrition on a back ground of refractory channels ($Var[N_R] = 1.0$ in B). When sparks terminate due to the recruitment of refractory channels during each puff/spark event the variance of the number of refractory channels (N_R) is large ($Var[N_R] = 2.8$ in A).

Fig. 4.5A illustrates how the accumulation of refractory channels during puff/sparks can contribute to the termination of Ca^{2+} release events. Note that the number of refractory channels (N_R) increases significantly during each puff/spark event and decreases during the inter-event intervals. Moreover, N_R usually reaches its local maximum value, usually larger than the average number of refractory channels ($E[N_R]$), at the end of each puff/spark. In Fig. 4.5B, the inactivation and de-inactivation rate is reduced by 10-fold of those used in Fig. 4.5A without changing the dissociation constant for Ca^{2+} inactivation K_b. In this case, the number of refractory channels is relatively constant throughout the entire simulation and thus the puff/spark events do not appear to terminate by the accumulation of refractory channels. Groff and Smith referred to this phenomenon as puff/spark terminating by stochastic attrition as a background of refractory channels.
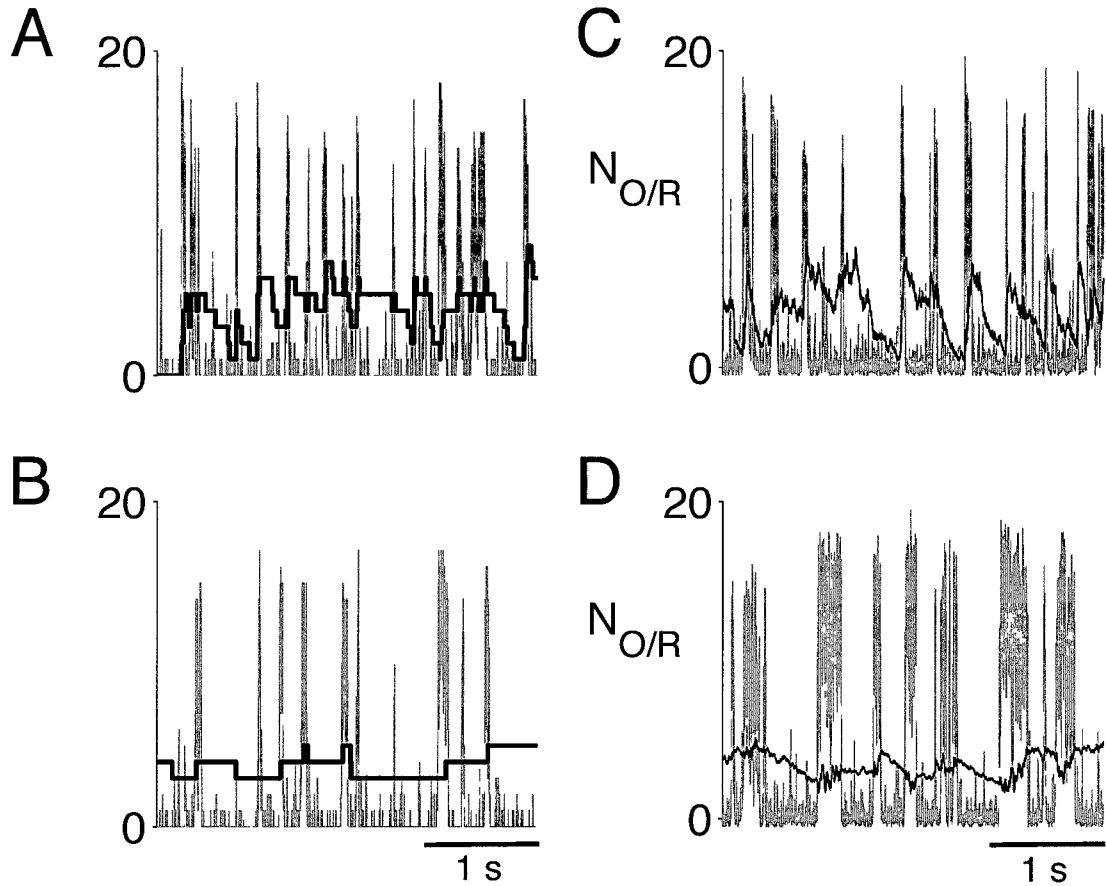
Figure 4.5: (A, C) Example Langevin and Monte Carlo simulation trajectory of 20 three-state channels (Eq. 4.4) showing puff/spark termination facilitated by the recruitment of refractory channels (black line) during puff/spark events. Parameters: $\eta = 2$, $k_a^+ = 1.5\ \mu\text{M}^{-\eta}\text{ms}^{-1}$, $k_b^+ = 0.015\ \mu\text{M}^{-\eta}\text{ms}^{-1}$, $k_a^- = 0.5\ \text{ms}^{-1}$, $k_b^- = 0.005$ ms^{-1}, $c_\infty = 0.05\ \mu\text{M}$, $c_* = 0.075\ \mu\text{M}$. (B, D) Example Langevin and Monte Carlo simulation when the Ca^{2+} inactivation rate is reduced 10-fold ($k_b^+ = 0.0015\ \mu\text{M}^{-2}\text{ms}^{-1}$ and $k_b^- = 0.0005\ \text{ms}^{-1}$). The number of refractory channels (black line) is nearly constant and puff/spark terminate without additional recruitment of refractory channels.

114
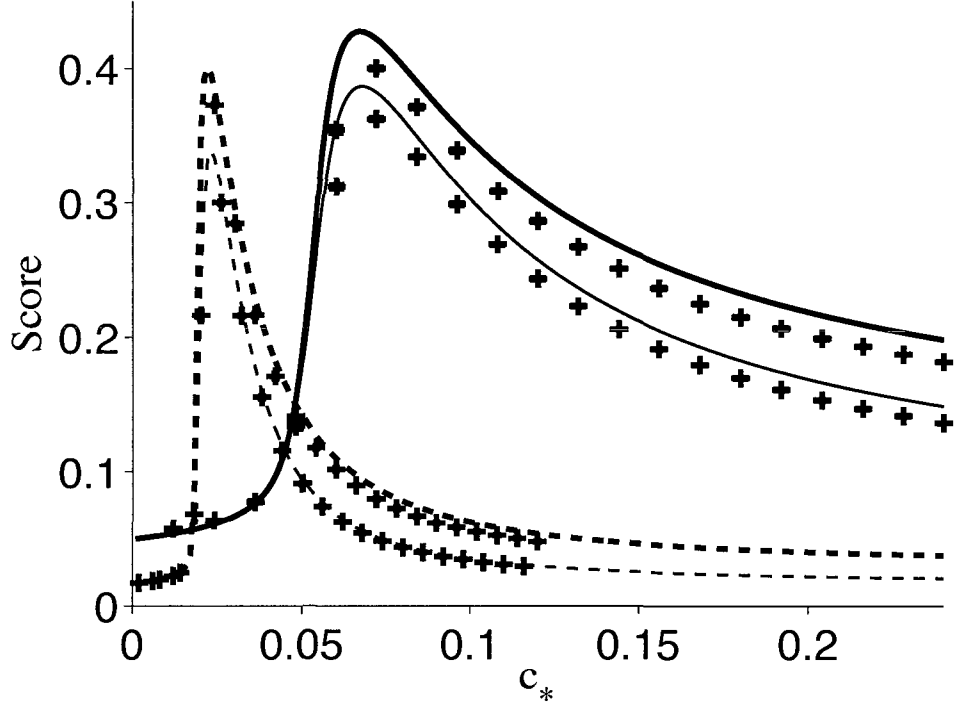
Figure 4.6: The Ca^{2+} puff/spark *Score* as a function of the coupling strength (c_*) when the number of three-state Ca^{2+}channels (Eq. 4.4) is $N = 20$ (solid lines) and $N = 60$ (dashed lines) are calculated using the stationary distribution of the Markov chain model. While the dissociation constant for Ca^{2+} inactivationis held constant $(K_b = 5.8 \ \mu M)$, two sets of inactivation rates are used (thin lines: $k_b^+ = 0.015 \ \mu M^{-2}ms^{-1}, k_b^- = 0.005 \ ms^{-1}$; thick lines: $k_b^+ = 0.0015 \ \mu M^{-2}ms^{-1}, k_b^- = 0.0005 \ ms^{-1}$). The corresponding Langevin simulation are calculated from the average probability distribution of f_O (10 trials of 10 seconds duration, crosses). Other parameters are as in Fig. 4.5

Fig. 4.6 shows how the stochastic excitability for Ca^{2+} release sites is affected by the coupling strength (c_*) when inactivation is present (Eq. 4.4) This can be compared to Fig. 4.2 where inactivation is absent. The lines are calculated directly from the stationary distribution of the Markov chain models. Each cross, on the other hand, is the average *Score* of 10 independent simulations at a same c_* value and the corresponding 95% confidence interval is given by the error bars. In the 20 channel simulations, Langevin results slightly underestimate the *Scores*, but the approximation improves as the simulated release sites are composed more channels. For both 20 (solid) and 60 (dashed) three-state channels, both methods show that a range of coupling strength (c_*) that gives rise to robust puff/sparks similar to the two-state case. Both methods also agree that the range for c_* that leads to stochastic excitability narrows as N increases.

4.5 Using the Langevin model to probe puff/spark statistics

The distribution of individual puff/spark durations, amplitudes and inter-event intervals can be measured in living cells using confocal fluorescence microscopy [Parker and Wier, 1997, Klein et al., 1999, Smith and Parker, 2009]. Although less commonly measured in experiments, we were interested in the relationship between successive puff/spark amplitudes and how puff/sparks and inter-puff/spark intervals may be correlated with each other, and how these correlations depend on single channel properties, such as Ca^{2+} inactivation (Eq. 4.4). Markov chain modeling of Ca^{2+}, these relationships are not easily obtained through direct analysis of Q matrices; consequently, Monte Carlo simulations are required and thus can be extremely time-

consuming when release sites are composed of large number of channels. We have found that our Langevin formulation (Section 4.3.3) reduces the computational cost of Ca^{2+} release site simulations to a degree that allows analyze of correlation in successive puff/spark properties.

In Sec. 4.4.1, we defined duration (D), amplitude (A) and inter-event intervals (I) for Ca^{2+} release events. Because small Ca^{2+} release events (blips) are often not detectable in confocal micro-fluorescence experiments, we define an amplitude "*threshold*", A_θ, and events whose amplitudes superthreshold $(A > A_\theta)$ are preserved in the analysis of correlations. Fig. 4.4A shows a simulation trajectory of a release site composed of 20 Ca^{2+} channels with five Ca^{2+} release events. If A_θ is set to be 30 ms, only three events are selected for analysis and the corresponding puff/spark durations, amplitudes and inter-puff/spark intervals are demonstrated in Fig. 4.4B.

Because amplitude threshold A_θ is to some extent arbitrary, Pearson correlation coefficients are calculated using a range of A_θ values. Fig. 4.7A shows the relationship between successive puff/spark amplitudes. When inactivation and de-inactivation rates are set to be the standard parameters (as in Fig. 4.5A), for Ca^{2+} release sites that are composed of 20 (solid line), 60 (dotted line), 80 (not shown) and 100 (not shown) channels, a small negative correlation between successive puff/spark amplitudes is observed that indicates event-to-event alternation in puff/spark amplitudes. This negative correlation reaches its minimum value when A_θ is about 15 ms for 20 channels and 20 ms for 60 channels.

When the Ca^{2+} inactivation/de-inactivation rates in the single channel model are increased by 10 times, the negative correlation for the 20-channel release site became barely observable but the negative correlation for the 60 channel release site is stronger (data not shown). Consequently, we did parameter studies on the inactivation and de-inactivation rates since, together with the number of channels, they seem to affect
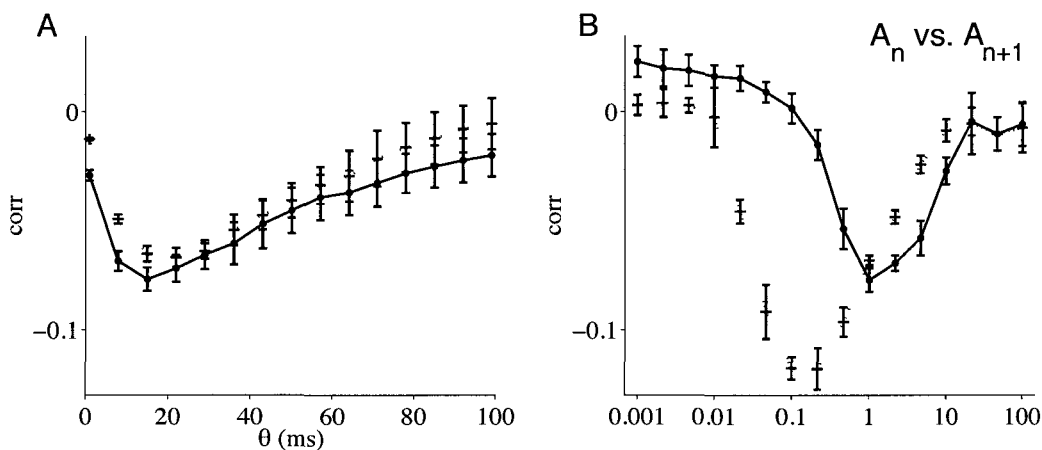
A

0

corr

−0.1

0   20   40   60   80   100
θ (ms)

B A_n vs. A_{n+1}

0

corr

−0.1

0.001 0.01 0.1   1   10   100

Figure 4.7: (A) Correlation of successive puff/spark amplitudes of release sites consist of 20 (solid) and 60 (dashed) 3-state Ca^{2+} channels as a function of selected amplitude threshold. Parameters are as in Fig. 4.5. (B) The strongest negative correlation for 20 (solid) and 60 channels (dotted) when the inactivation/de-inactication rate are changed to different folds of the parameters in Fig. 4.5.

the successive puff/spark amplitude correlations. To keep the Ca^{2+} affinity (K_b) fixed, the inactivation and de-inactivation rates are either reduced or increased by the same factor. For each pair of inactivation/de-inactivation rates studied, 10 independent Langevin simulations are carried out and 10^6 Ca^{2+}release events are collected from each individual simulation. The amplitude correlation of each parameter averaged through the ten trials is shown in Fig. 4.7B by dots and crosses for 20 and 60 channels respectively, and the standard deviations are shown by error bars.

Simulations of both 20 and 60 channels show that the negative correlations between successive puff/spark amplitudes are significantly weakened when the inactivation rates are reduced by 100 fold compared to the standard parameters. Recalling Fig. 4.5B, when inactivation rate is slow enough that the number of refractory channels stays almost constant, the Ca^{2+} release site approximates a two-state release site that is composed of fewer channels. On the other hand, when the inactivation rates

118

are large, the negative correlation is also weakened due to the fast recovery from the refractory state. The speculated reason for this negative correlation is consequently high affinity, moderately slow inactivation since when inactivation presents, a comparatively larger fraction of channels are inactivated by the end of each large puff/spark event (shown in Fig. 4.5A), thus fewer channels are available to participate in the next Ca^{2+} release event and consequently the next puff/spark is very likely to have a comparatively small amplitude. Similar study was carried out for Ca^{2+}channels without inactivation (data not shown) and the fact that no obvious negative correlation is found further supported our speculation.

Interestingly, as the release site under simulation is composed of a larger number of channels, the strongest negative correlations are reached at a slower inactivation rate. For instance, while the negative correlation for successive puff/spark amplitudes is the strongest at the standard parameter for 20 channels, the inactivation rate has to be reduced by 10 fold and 100 fold for 60 channels (Fig. 4.5B) and 80 channels (data not shown) respectively, to reach their strongest negative correlation. The strength of the maximum negative correlation, on the other hand, is shown to be slightly larger for larger release sites.

When the relationship between puff/spark amplitudes and their preceding inter-event intervals are studied in the same manner, positive correlations are shown for 20 and 60 channels when standard parameters are used (Fig. 4.8A). The explanation to this observation is also accumulated refractory channels terminating puff/sparks (as in Fig. 4.5A). In this case, longer inter-event intervals indicates more recover time for the channels that are inactivated by the end of the preceding puff/spark allowing more channels to be available for release events, consequently the following event amplitudes are likely to be larger compared to that of those who follow brief quiescent periods. Similar to the previous study, the inactivation/de-inactivation rates must be
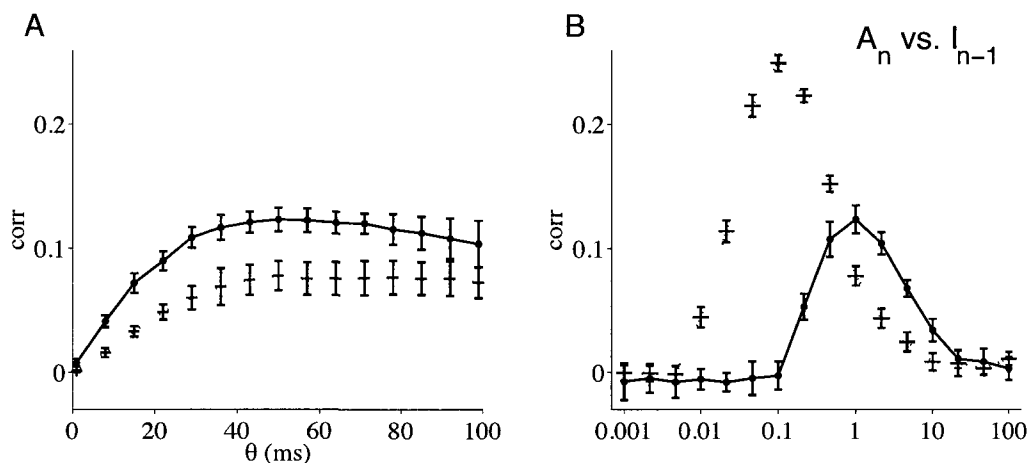
Figure 4.8: Correlation of puff/spark amplitudes (upper) and the previous quiescent duration of release sites consist of 20 (solid) and 60 (dotted) 3-state Ca^{2+} channels as a function of selected amplitude threshold. The change of the correlation as a function of the inactivation rate when 10% of the events are selected as puff/sparks,

slow enough such that the channels spend adequate time in the refractory state before recovery, yet not too slow so that the number of refractory channels increase as more channels are open. Actually, when the inactivation is either reduced or increased by 100 fold, the positive correlations fade out for both 20 and 60 channels (Fig. 4.8B).

On the other hand, the amplitudes of puff/sparks does not appear to affect the interval length before the next Ca^{2+} release events. In Fig. 4.9 the puff/spark amplitudes and their following inter-event intervals are also shown to have only very weak positive correlations (less than 10%) when standard parameters are used for 60 channels (dotted line in Fig. 4.9A) and this positive correlation only exists when the inactivation/de-inactivation rates are moderately slower than activation rate (dashed line in Fig. 4.8B) which are results consistent with the two results discussed above. However, the positive correlation is barely shown for 20 channels (solid lines). It was reported in literature that no significant correlation between peak spark amplitude and the immediate following interval in rat ventricular myocytes [Parker and Wier,
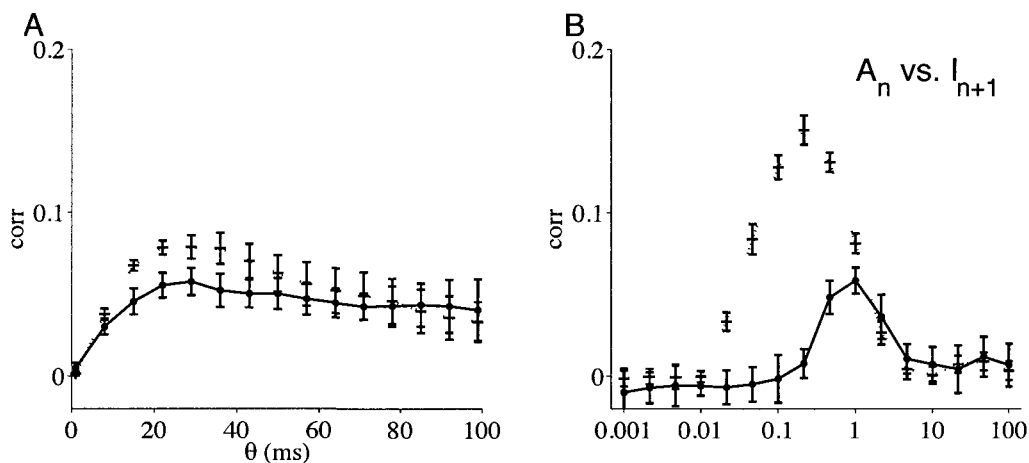
Figure 4.9: Correlation of puff/spark amplitudes (upper) and the following quiescent duration of release sites consist of 20 (solid) and 60 (dotted) 3-state Ca^{2+} channels as a function of selected amplitude threshold. The change of the correlation as a function of the inactivation rate when 10% of the events are selected as puff/sparks.

1997] Spark amplitudes and their preceding inter-event intervals are also reported to lack significant correlation in the study of frog muscle fibres [Klein et al., 1999].

4.6 Discussion

We have presented a Langevin formulation for the stochastic dynamics of calcium release sites. Using a two-state model with Ca^{2+}activation only and a three-state Ca^{2+} channel model with both Ca^{2+} activation and inactivation, we demonstrated that the Langevin description is a good substitution to models of the corresponding compositionally defined Markov chain Ca^{2+} release site models. The Langevin simulation results for the Ca^{2+} release sites that are composed of 20 – 150 channels are almost identical to the corresponding Monte Carlo simulation results on the stationary distributions of the fraction of open channels, the puff/spark scores and the distributions of puff/spark statistics.

121

As an application demonstrating the computational efficiency of this approach, using Ca^{2+} release sites that are composed of 20 or 60 three-state Ca^{2+} channels, we studied whether Ca^{2+} inactivation affects the correlations between the successive puff/spark amplitudes and inter-event intervals, which are computationally expensive to get statistics if otherwise simulated by Monte Carlo simulations. We observed some interesting correlations exist among puff/spark amplitude and inter-event interval lengths and that the existence of these correlations are highly dependent to Ca^{2+} inactivation rates.

4.6.1 Spark amplitude and inter-event interval correlations

We started with studying the correlations between successive puff/spark amplitudes of 20 and 60 three-state Ca^{2+} under the standard parameters (as in Fig. 4.5A) and discovered small negative correlations between successive puff/spark amplitudes. Then how the inactivation/de-inactivation rates affects this correlation is studied and negative correlations are shown only when the inactivation/de-inactivation are approximately 20 to 500 times slower as compared to the activation rates for 20 channels and 50 to 2000 times slower for 60-channel release sites.

The reason for this negative correlation is very likely to be that a larger fraction of channels are inactivated by the end of large release events causing less channels are available for the following puff/spark. To further verify our speculation, we studied the correlations between puff/spark amplitude and their preceding and following inter-event interval lengths. In the parameter range where negative correlations between successive puff/spark amplitudes are found, puff/spark amplitudes are shown to be positively affected by their preceding inter-event interval but are not significantly correlated to their following inter-event interval.

Even though many experimental studies are done to study the distribution of Ca^{2+} puff/spark amplitudes, rinsing times and durations [Cheng et al., 1999, Rios et al., 2001, Shen et al., 2004] and some of them study correlations between the peak amplitudes of Ca^{2+} release events and their corresponding rising time [Lacampagne et al., 2000], we didn't find data discussing the correlation between successive spark amplitudes. Yet the few experimental data which were collected toward the study of how Ca^{2+} puff/spark amplitudes interfere with the intervals preceding or following them showed no significant correlations between these statistics.

4.6.2 Promise and limitations of the Langevin implementation

In prior work, Langevin equations are utilized to reduce the computation cost of a specific cluster of unregulated Li-Rinzel IP3 receptors where all channels are independent to each other [Shuai and Jung, 2002a]. Different from their work, which was restricted to a 2-state model, we aimed at developing an automated program which is able to transform Ca^{2+} release sites with arbitrary number of states into Langevin formulations. Luminal and cytosolic Ca^{2+} regulation, diffusions and buffering can be easily integrated into our program because the deterministic part of these stochastic differential equations are solved as dynamical systems:

$$\frac{dc_{myo}}{dt} = \beta_{myo} \left[v_{rel} \, \gamma_n \left(c_{sr} - c_{myo} \right) - v_{myo} \left(c_{myo} - c_{myo}^\infty \right) \right] \qquad (4.22)$$

$$\frac{dc_{sr}}{dt} = \frac{\beta_{sr}}{\lambda} \left[v_{rel} \, \gamma_n \left(c_{sr} + c_{myo} \right) + v_{sr} \left(c_{sr}^\infty - c_{sr} \right) \right], \qquad (4.23)$$

where Eq. 4.22 and Eq. 4.23 describe the Ca^{2+} change in the myocyte and in SR respectively, λ is the SR to myoplasmic domain volume ratio, $\beta_{myo} \left(c_{myo} \right)$ and $\beta_{sr} \left(c_{sr} \right)$

are functions of the domain concentrations accounting for rapid buffering.

With a three-state Ca^{2+} channel model, we validated that when the size of a Ca^{2+} release site is fairly large ($N \geq 20$), the Langevin formulation approximates Markov chain results very well and the results usually improve as the number of channels models increases. More importantly, the main advantage of the Langevin formulation is that the computation cost does not increase as the number of channels modeled increases.

However, the Langevin equations are currently solved by forward Newton's method with Cholvesky factorization finding the square root of the covariance matrix. These numerical methods, especially `Chokove` command in Matlab, are not very time efficient. We are still searching for numerical solutions of the Langevin formulations so that the computational efficiency is further improved.

4.7 Apendices

4.7.1 Producing Gaussian random variates with specified covariance matrix

An $n \times n$ real symmetric matrix M is "positive semi-definite" if $z^T M z \geq 0$ for all non-zero vectors z with real entries ($z \in R^n$). A covariance matrix Σ is square, symmetric, and positive semi-definite and therefore has the Cholesky-like decomposition

$$\Sigma = A^T A$$

and the random variables y can be calculated as

$$y = A^T x$$

124

where the x_i are independent and identically distributed normal random variables that satisfy the mean and covariance of $\boldsymbol{x}$ are $\langle \boldsymbol{x} \rangle = \boldsymbol{0}$ and $\langle \boldsymbol{x}\boldsymbol{x}^T \rangle = I$. Consequently, the mean and covariance of $\boldsymbol{y}$ are $\langle \boldsymbol{y} \rangle = \langle A^T\boldsymbol{x} \rangle = A^T\langle \boldsymbol{x} \rangle = A^T\boldsymbol{0} = \boldsymbol{0}$ and $\langle \boldsymbol{y}\boldsymbol{y}^T \rangle = \Sigma$ because

$$\langle \boldsymbol{y}\boldsymbol{y}^T \rangle = \langle A^T\boldsymbol{x}\boldsymbol{x}^T A \rangle = A^T\langle \boldsymbol{x}\boldsymbol{x}^T \rangle A = A^T I A = A^T A = \Sigma$$

4.7.2 Fokker-Planck descriptions for 2-state Ca^{2+} channels

From "Noise in spatially extended systems" [García-Ojalvo and Sancho, 1999], the Langevin equation for the zero-dimensional system

$$\frac{dx}{dt} = f(x) + g(x)\eta(t)$$

where

$$\langle \eta(t) \rangle = 0$$

and

$$\langle \eta(t)\eta(t') \rangle = 2D\delta\left(t - t'\right)$$

has the corresponding Fokker-Planck equation:

$$\frac{\partial}{\partial t}P(x,t) = -\frac{\partial}{\partial x}f(x)P(x,t) + D\frac{\partial^2}{\partial x^2}g^2(x)P(x,t).$$

Assuming natural boundary conditions (See Section 4.7.4), the steady-state probability distribution is given by

$$P_\infty(x) = \frac{A}{g(x)^2}\exp\left\{\int \frac{f(x)}{Dg^2(x)}dx\right\},$$

where A is a normalization constant.

4.7.3 Steady-state probability distribution for two-state channel

Assuming a two-state channel with constant transition rates the Langevin equation is

$$\frac{dx}{dt} = \underbrace{-\alpha x + \beta(1-x)}_{f(x)} + \underbrace{\sqrt{\frac{\alpha x + \beta(1-x)}{N}}}_{g(x)}\,\eta(t),$$

where

$$\langle \eta(t)\rangle = 0,$$

and

$$\langle \eta(t)\eta(t')\rangle = \delta\left(t - t'\right).$$

The corresponding Fokker-Planck equation is

$$\frac{\partial}{\partial t}P(x,t) = -\frac{\partial}{\partial x}f(x)P(x,t) + D\frac{\partial^2}{\partial x^2}g^2(x)P(x,t).$$

Assuming natural boundary conditions, we can define the stochastic potential as

$$U(x) = -\int \frac{f(x)}{g^2(x)}dx. \tag{4.24}$$

The steady state distribution is given by

$$P_\infty(x) = \frac{A}{g(x)^2}\exp\left\{-2U(x)\right\}.$$

When α, β are constants, we can calculate $U(x)$ analytically:

$$-\frac{U(x)}{N} = \int \frac{f(x)}{g^2(x)}dx = \int \frac{-\alpha x + \beta(1-x)}{\alpha x + \beta(1-x)}dx = \int \frac{\beta - (\alpha + \beta)x}{\beta + (\alpha - \beta)x}dx$$

Let $a = -(\alpha + \beta)$, $b = d = \beta$, $c = \alpha - \beta$, using the integral table formula,

$$\int \frac{ax+b}{cx+d}dx = \frac{(bc-ad)\log(cx+d)+acx}{c^2}$$

we find

$$-\frac{U(x)}{N} = \frac{2\alpha\beta}{(\alpha-\beta)^2}\log\left[\alpha x + \beta(1-x)\right] - \frac{\alpha+\beta}{\alpha-\beta}x.$$

The steady state distribution is given by

$$P_\infty(x) = A\left[\alpha x + \beta(1-x)\right]^{\frac{4N\alpha\beta}{(\alpha-\beta)^2}-1}\exp\left\{-2N\frac{\alpha+\beta}{\alpha-\beta}x\right\}.$$

If the transition diagram of a two-state channel with Ca^{2+} activation is given by, Eq. 4.1 then

$$\begin{cases} \alpha &=& k^- \\ \beta &=& k^+(c_\infty + rx)^2 \end{cases}$$

where r is the maximum (i.e. when all channels open) Ca^{2+} release caused by channel gating and x is the fraction of open channels. In this case α became a function of x and consequently the integral table formula is not applicable any more.

substitute α and β by their value:

$$\begin{aligned} U &=& -N\int \frac{k^+(c_\infty+rx)^2 - (k^- + k^+(c_\infty+rx)^2)x}{k^+(c_\infty+rx)^2 + (k^- - k^+(c_\infty+rx)^2)x}dx \\ &=& -N\int 1 - \frac{2k^-x}{k^+c_\infty^2 + (2k^+c_\infty r + k^- - k^+c_\infty^2)x + (k^+r^2 - 2k^+c_\infty r)x^2 - k^+r^2x^3}dx \end{aligned}$$

Theoretically, the integrand is is always integrable since it is a rational function. However, there is not a trivial formula for the cubic function above. Consequently, we integrated the integrand in Eq. 4.24 numerically.
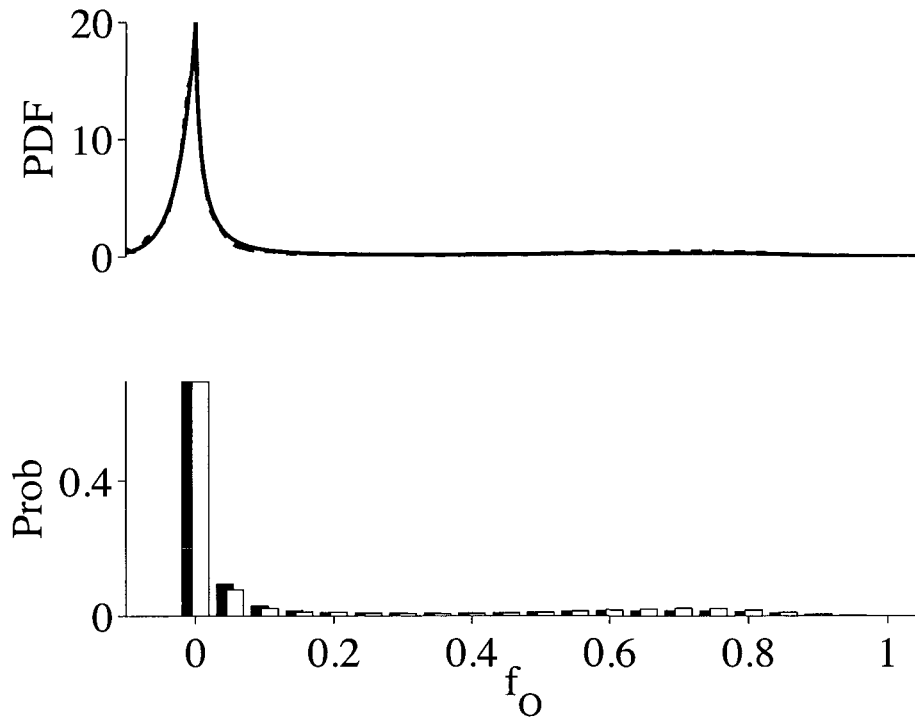
Figure 4.10: The steady state probability distribution of a 20-channel Ca^{2+} release site numerically integrated from the Fokker Plank equation is shown with the solid curve and discretized follow the number of open channels (white histogram). The same distribution statistically collected from a sample simulation is plotted with the dashed line. The corresponding Markov chain model stationary distribution is shown with black histograms. Parameters are as in Fig. 4.1

4.7.4 Boundary condition

In our simulations, free boundary condition is used in match with the assumption that is necessary for analytically solving the steady state solutions of Langevin equations. However, since open fractions of a release site have to be within $[0, 1]$ to make physical sense, the values of f_O which go beyond this interval are collected as either of the boundary value.

Chapter 5

Summary of results

This dissertation proposed several techniques to reduce the computational cost of compositionally defined Markov chain Ca^{2+} release site models, which is an important step toward mechanistic modeling that traverses the molecular and cellular levels of Ca^{2+} signaling dynamics. Compositionally defined Markov chain Ca^{2+} release site models reminiscent the physiological realism of interacting Ca^{2+} channels and collective gating. Unfortunately, the large state space of such compositional models impedes simulation and computational analysis in multiscale simulations of the whole cell Ca^{2+} signaling in which the stochastic dynamics of localized Ca^{2+} release events play an important role. In this dissertation we validated that our novel techniques are able to produce reduced models that have significantly fewer states while maximizing the preservation of the collective gating properties of computationally defined Markov chain Ca^{2+} release sites. We also utilize these reduction approaches to calculate some puff/spark statistics that are otherwise expensive to collect from experiments or traditional simulation methods.

Chapter 2 introduces several automated model reduction techniques based on fast/slow analysis that leverages these time scale differences that are often exhibited

by ligand-gated intracellular channels. Rate constants in the single channel model are categorized as either fast or slow, groups of release site states that are connected by fast transitions are identified and lumped, and transition rates between reduced states are chosen consistent with the conditional probability distributions among states within each group. The fast/slow reduction approach is validated by the fact that puff/spark statistics can be efficiently computed from reduced Ca^{2+} release site models with small and transient reduction error. For release site composed of many channels, the conditional probability distributions can be approximated without the construction of the full model by assuming a rapid mixing of states connected by fast transitions. Alternatively, an IAD method can be employed to obtain a reduced Ca^{2+} release site model in a memory-efficient fashion.

In Chapter 3, a genetic algorithm based approach is implemented to find reduced models with moderate errors in the absence of time-scale separation. Given a full model of interest and target reduced model size, this genetic algorithm searches for set partitions, each corresponding to a potential scheme for state aggregation, that lead to reduced models that well-approximate the full model mimicking the inheritance, mutation, and selection processes of natural evolution. Compared to the fast/slow based approach (Chapter 2), the genetic algorithm based approach is more expensive computationally but more flexible and realistic. The *Evaluation* subroutine can be easily modified so that the resulting reduced model is able to approximate the full model on different aspects and fit different model formulations.

While the techniques presented in Chapters 2 and 3 significantly reduce the size of mechanistic Ca^{2+} release site models, the stationary distribution of the full model is required for the implementation of both approaches. In practice, this limits the size of full models that can be targeted for reduction and, for release sites composed of identical channels, the number of channels per release site is constrained. In Chapter 4, a

130

Langevin formulation of the Ca^{2+} release sites is introduced as an alternative model reduction technique that is applicable in situations in which the number of channels per Ca^{2+} release site is too large for the previously discussed reduction methods, but not so large that the stochastic dynamics of Ca^{2+} release can be neglected. The Langevin formulation for coupled intracellular Ca^{2+} channels results in stochastic differential equations that well-approximate the corresponding Markov chain models when release sites possess as few as 20 channels, and the agreement improves as the number of channels per release site increases. Importantly, the computational time required by the Langevin approach does not increase with the size of Ca^{2+} release sites. With this approach we are able to study and report the correlations between puff/spark statistics theoretically for the first time.

Bibliography

[Ball and Geoffrey, 2000] Ball, F. and Geoffrey, Y. (2000). Superposition of Spatially Interacting Aggregated Continuous Time Markov Chains. *Methodology of Computing in Applied Probability*, 2(1):93–116.

[Ball et al., 2000] Ball, F. G., Milne, R. K., and Yeo, G. F. (2000). Stochastic models for systems of interacting ion channels. *IMA journal of mathematics applied in medicine and biology*, 17(3):263–293.

[Berridge, 1993a] Berridge, M. J. (1993a). Inositol trisphosphate and Ca^{2+} signaling. *Nature*, 361(6410):315–25.

[Berridge, 1993b] Berridge, M. J. (1993b). Inositol trisphosphate and Ca^{2+} signaling. *Nature*, 361(6410):315–25.

[Berridge, 1997a] Berridge, M. J. (1997a). Elementary and global aspects of Ca^{2+} signalling. *J Physiol (London)*, 499(Pt 2):291–306.

[Berridge, 1997b] Berridge, M. J. (1997b). Elementary and global aspects of calcium signalling. *J Physiol (Lond)*, 499 (Pt 2):291–306.

[Berridge, 1998] Berridge, M. J. (1998). Neuronal Ca^{2+} signaling. *Neuron*, 21(1):13–26.

[Berridge, 2006] Berridge, M. J. (2006). Calcium microdomains: organization and function. *Cell Calcium*, 40(5-6):405–12.

[Berridge et al., 1998] Berridge, M. J., Bootman, M. D., and Lipp, P. (1998). Calcium-a life and death signal. *Nature*, 395:645–648.

[Berridge et al., 2003] Berridge, M. J., Bootman, M. D., and Roderick, H. L. (2003). Calcium signalling: dynamics, homeostasis and remodelling. *Molecular cell biology*, 4:517–529.

[Berridge et al., 2000] Berridge, M. J., Lipp, P., and Bootman, M. D. (2000). The versatility and universality of Calcium signaling. *Molecular cell biology*, 1(6703):11–21.

[Bers, 2002] Bers, D. M. (2002). Cardiac excitation-contraction coupling. *Nature*, 415(6868):198–205.

[Beuckelmann and Wier, 1988] Beuckelmann, D. J. and Wier, W. G. (1988). Mechanism of release of calcium from sarcoplasmic reticulum of guinea-pig cardiac cells. *J Physiol (London)*, 405:233–255.

[Bezprozvanny and Ehrlich, 1994] Bezprozvanny, I. and Ehrlich, B. E. (1994). Inositol (1,4,5)-trisphosphate (IP$_3$)-gated Ca^{2+} channels from cerebellum: conduction properties for divalent cations and regulation by intraluminal Ca^{2+}. *J Gen Physiol*, 104(5):821–56.

[Bezprozvanny et al., 1991] Bezprozvanny, I., Watras, J., and Ehrlich, B. E. (1991). Bell-shaped Ca^{2+}-response curves of Ins(1,4,5)P3- and Ca^{2+}-gated channels from endoplasmic reticulum of cerebellum. *Nature*, 351(6329):751–4.

133

[Cannell et al., 1995] Cannell, M. B., Cheng, H., and Lederer, W. J. (1995). The control of calcium release in heart muscle. *Science*, 268(5213):1045–9.

[Cao and Stewart, 1985] Cao, W. and Stewart, W. (1985). Iterative aggregation/disaggregation techniques for nearly uncoupled Markov chains. *Journal of the Association for Computing Machinery*, 32(3):702–719.

[Carafoli et al., 2001] Carafoli, E., Santella, L., Brance, D., and Brisi, M. (2001). Generation, control, and processing of cellular calcium signals. *Crit. Rev. Biochem. Mol. Biol.*, 36:107–260.

[Cheng et al., 1996] Cheng, H., Lederer, M. R., Lederer, W. J., and Cannell, M. B. (1996). Ca^{2+} sparks and $[Ca^{2+}]_i$ waves in cardiac myocytes. *Am J Physiol*, 270(1 Pt 1):C148–59.

[Cheng and Lederer, 2008] Cheng, H. and Lederer, W. J. (2008). Calcium sparks. *Physiol. Rev.*, 88(4):1491–1545.

[Cheng et al., 1993a] Cheng, H., Lederer, W. J., and Cannell, M. B. (1993a). Calcium sparks: elementary events underlying excitation-contraction coupling in heart muscle. *Science*, 262(5134):740–4.

[Cheng et al., 1993b] Cheng, H., Lederer, W. J., and Cannell, M. B. (1993b). Calcium sparks: elementary events underlying excitation-contraction coupling in heart muscle. *Science*, 262(5134):740–4.

[Cheng et al., 1999] Cheng, H., Song, L. S., Shirokova, N., González, A., Lakatta, E. G., Ríos, E., and Stern, M. D. (1999). Amplitude distribution of calcium sparks in confocal images: theory and studies with an automatic detection method. *Biophysical Journal*, 76(2):606–17.

134

[Clapham, 1995] Clapham, D. E. (1995). Calcium signaling. *Cell*, 80(2):259–268.

[Colquhoun and Hawkes, 1995] Colquhoun, D. and Hawkes, A. (1995). A Q-matrix cookbook: how to write only one program to calculate the sigle-channel and macro-scopic predictions for any kinetic mechanism. In Sakmann, B. and Neher, E., editors, *Single-Channel Recording*, pages 589–633. Plenum Press, New York.

[Davis, 1991] Davis, L. (1991). *Handbook Of Genetic Algorithms*. Van Nostrand Reingold, New York.

[De Young and Keizer, 1992] De Young, G. W. and Keizer, J. (1992). A single-pool inositol 1,4,5-trisphosphate-receptor-based model for agonist-stimulated os-cillations in Ca^{2+} concentration. *Proc Natl Acad Sci USA*, 89(20):9895–9.

[DeRemigio et al., 2008] DeRemigio, H., Kemper, P., LaMar, M. D., and Smith, G. D. (2008). Markov chain models of coupled intracellular calcium channels: Kronecker structured representations and benchmark stationary distribution cal-culations. *Physical Biology*, 5(3):36003.

[DeRemigio and Smith, 2005] DeRemigio, H. and Smith, G. D. (2005). The dynamics of stochastic attrition viewed as an absorption time on a terminating Markov chain. *Cell Calcium*, 38(2):73–86.

[DeRemigio and Smith, 2008] DeRemigio, H. and Smith, G. D. (2008). Calcium re-lease site ultrastructure and the dynamics of puffs and sparks. *Math. Med. and Biol.*, 25(1):65–85.

[DiGesú et al., 2005] DiGesú, V., Giancarlo, R., LoBosco, G., Raimondi, A., and Scaturro, D. (2005). Genclust: a genetic algorithm for clustering gene expression data. *BMC Bioinformatics*, 6:289.

[Endo, 1977] Endo, M. (1977). Calcium release from the sarcoplasmic reticulum. *Phys Rev*, 57(1):71–108.

[Fill and Copello, 2002] Fill, M. and Copello, J. A. (2002). Ryanodine receptor calcium release channels. *Physiol Rev*, 82(4):893–922.

[Foskett et al., 2007] Foskett, J. K., White, C., Cheung, K. H., and Mak, D. O. (2007). Inositol trisphosphate receptor Ca^{2+} release channels. *Physiol Rev*, 87(2):593–658.

[Franzini-Armstrong et al., 1999] Franzini-Armstrong, C., Protasi, F., and Ramesh, V. (1999). Shape, size, and distribution of Ca^{2+} release units and couplons in skeletal and cardiac muscles. *Biophys J*, 77(3):1528–1539.

[García-Ojalvo and Sancho, 1999] García-Ojalvo, J. and Sancho, J. (1999). *Noise in spatially extended systems*. Springer.

[Gardiner, 2009] Gardiner, C. (2009). *Stochastic Methods*. Springer-Verlag, Berlin Heidelberg, Germany.

[Gillespie, 1977] Gillespie, D. (1977). Exact stochastic simulation of coupled chemical reactions. *J Phys Chem*, 81:2340–61.

[Greenstein et al., 2006] Greenstein, J. L., Hinch, R., and Winslow, R. L. (2006). Mechanisms of excitation-contraction coupling in an integrative model of the cardiac ventricular myocyte. *Biophys J*, 90(1):77–91.

[Groff, 2008] Groff, J. R. (2008). *Markov chain models of calcium puffs and sparks*. PhD thesis, The College of William and Mary, Williamsburg, VA.

[Groff and Smith, 2008a] Groff, J. R. and Smith, G. D. (2008a). Calcium-dependent inactivation and the dynamics of calcium puffs and sparks. *J. Theor. Biol.*, 253(3):483–99.

[Groff and Smith, 2008b] Groff, J. R. and Smith, G. D. (2008b). Ryanodine receptor allosteric coupling and the dynamics of calcium sparks. *Biophys. J.*, 95(1):135–154

[Gyorke and Gyorke, 1998] Gyorke, I. and Gyorke, S. (1998). Regulation of the cardiac ryanodine receptor channel by luminal ca2 involves luminal ca2+ sensing sites. *Biophys J*, 75:2801–2810.

[Hao et al., 2009] Hao, Y., Kemper, P., and Smith, G. D. (2009). Reduction of calcium release site models via fast/slow analysis and iterative aggregation/disaggregation. *Chaos. Focus issue: Intracellular Calcium Dynamics—A Change of Modeling Paradigm?*, 5(3):36003.

[Hartman et al., 2010] Hartman, J. A., Sobie, E. A., and Smith, G. D. (2010). Calcium sparks and homeostasis in a minimal model of local and global calcium responses in quiescent ventricular myocytes. *AJP: Heart and Circulatory Physiology*. doi:10.1152/ajpheart.00293.2010.

[Hill et al., 2005] Hill, T., Lundgren, A., Fredriksson, R., and Schioth, H. (2005). Genetic algorithm for large-scale maximum parsimony phylogenetic analysis of proteins. *Biochim Biophys Acta*, 1725(1):19–29.

[Hill, 1989] Hill, T. L. (1989). *Free Energy Transduction in Biology and Biochemical Cycle Kinetics.* Springer-Verlag, New York.

[Hinch et al., 2004] Hinch, R., Greenstein, J. L., Tanskanen, A. J., Xu, L., and Winslow, R. L. (2004). A simplified local control model of calcium-induced calcium release in cardiac ventricular myocytes. *Biophysical Journal*, 87(6):3723–36.

[Hinch et al., 2006] Hinch, R., Greenstein, J. L., and Winslow, R. L. (2006). Multiscale models of local control of calcium induced calcium release. *Prog Biophys Mol Biol*, 90(1-3):136–50.

[Holland, 1975] Holland, J. H. (1975). *Adaptation in natural and artificial systems*. The U. of Michigan Press.

[Huertas and Smith, 2007] Huertas, M. A. and Smith, G. D. (2007). The dynamics of luminal depletion and the stochastic gating of Ca^{2+}-activated Ca^{2+} channels and release sites. *J. Theor. Biol.*, 246(2):332–354.

[Keizer and Levine, 1996] Keizer, J. and Levine, L. (1996). Ryanodine receptor adaptation and Ca^{2+}(-)induced Ca^{2+} release-dependent Ca^{2+} oscillations. *Biophys J*, 71(6):3477–3487.

[Keizer, 1987] Keizer, J. E. (1987). *Statistical Thermodynamics of Nonequilibrium Processes*. Springer Verlag, Berlin.

[Kirichok et al., 2004] Kirichok, Y., Krapivinsky, G., and Clapham, D. E. (2004). The mitochondrial calcium uniporter is a highly selective ion channel. *Nature*, 427(6972):360–4.

[Klein et al., 1999] Klein, M. G., Lacampagne, A., and Schneider, M. F. (1999). A repetitive mode of activation of discrete ca2+ release events (ca2+ sparks) in frog skeletal muscle fibres. *J Physiol*, 515(2):391–411.

[Koury et al., 1984] Koury, J., McAllister, D., and Stewart, W. (1984). Iterative methods for computing stationary distributions of nearly completely decomposable Markov chains. *SIAM Journal on Algebraic and Discrete Methods*, 5(2):164–186.

[Lacampagne et al., 2000] Lacampagne, A., Klein, M. G., Ward, C. W., and Schneider, M. F. (2000). Two mechanisms for termination of individual ca2+ sparks in skeletal muscle. *Proc Natl Acad Sci USA*, 97(14):7823–7828.

[LaMar et al., 2011] LaMar, M. D., Kemper, P., and Smith, G. D. (2011). Reduction of calcium release site models via moment fitting of phase-type distributions. *Phys. Biol.*, 8:026015.

[Li and Rinzel, 1994] Li, Y. X. and Rinzel, J. (1994). Equations for IP$_3$R-mediated [Ca^{2+}]$_i$ oscillations derived from a detailed kinetic model: a Hodgkin-Huxley like formalism. *J Theor Biol*, 166(4):461–73.

[Low et al., 2001] Low, D. A., Weyand, N. J., and Mahan, M. J. (2001). Roles of dna adenine methylation in regulating bacterial gene expression and virulence. *Infection and Immunity*, 69:7197–1204.

[Mak and Foskett, 1997] Mak, D. O. and Foskett, J. K. (1997). Single-channel kinetics, inactivation, and spatial distribution of inositol trisphosphate (IP$_3$) receptors in Xenopus oocyte nucleus. *J Gen Physiol*, 109(5):571–87.

[Marchant and Parker, 2000] Marchant, J. S. and Parker, I. (2000). Functional interactions in Ca^{2+} signaling over different time and distance scales. *J Gen Physiol*, 116(5):691–6.

[Mazzag et al., 2005] Mazzag, B., Tignanelli, C., and Smith, G. D. (2005). The effect of residual Ca^{2+} on the stochastic gating of Ca^{2+}-regulated Ca^{2+} channels. *J. Theor. Biol.*, 235(1):121–150.

[Meissner, 1994] Meissner, G. (1994). Ryanodine receptor/Ca^{2+} release channels and their regulation by endogenous effectors. *Annu. Rev. Physiol*, 56:485–508.

[Michalewicz, 1994] Michalewicz, Z. (1994). *Genetic Algorithms + Data Structures = Evolution Programs*. Springer, New York.

[Nguyen et al., 2005] Nguyen, V., Mathias, R., and Smith, G. D. (2005). A stochastic automata network descriptor for markov chain models of instantaneously-coupled intracellular Ca^{2+} channels. *Bull. Math. Biol.*, 67(3):393–432.

[Nicola, 1998] Nicola, V. (1998). Lumping in markov reward processes. Technical report, RC14719, IBM Thomas Watson Research Centre, PO Box 704, Yorktown Heights, NY 10598.

[Niggli, 1999] Niggli, E. (1999). Localized intracellular Ca^{2+} signaling in muscle: Ca^{2+} sparks and Ca^{2+} quarks. *Annu Rev Physiol*, 61:311–35.

[Norris, 1997] Norris, J. (1997). *Markov chains*. Cambridge University Press, Cambridge.

[Parker et al., 1996] Parker, I., Choi, J., and Yao, Y. (1996). Elementary events of IP_3-induced Ca^{2+} liberation in Xenopus oocytes: hot spots, puffs and blips. *Cell Calcium*, 20(2):105–21.

[Parker and Wier, 1997] Parker, I. and Wier, W. G. (1997). Variability in frequency and characteristics of Ca^{2+} sparks at different release sites in rat ventricular myocytes. *J Physiol*, 505(Pt 2):337–44.

[Parker and Yao, 1996] Parker, I. and Yao, Y. (1996). Ca^{2+} transients associated with openings of inositol trisphosphate-gated channels in Xenopus oocytes. *J Physiol*, 491(Pt 3):663–8.

[Rengifo et al., 2002] Rengifo, J., Rosales, R., Gonzalez, A., Cheng, H., Stern, M. D., and Rios, E. (2002). Intracellular Ca^{2+} release as irreversible Markov process. *Biophys J*, 83(5):2511–2521.

[Rios et al., 2001] Rios, E., Shirokova, N., Kirsch, W. G., Pizarro, G., Stern, M. D., Cheng, H., and Gonzalez, A. (2001). A preferred amplitude of Ca^{2+} sparks in skeletal muscle. *Biophys J*, 80(1):169–83.

[Ríos and Stern, 1997] Ríos, E. and Stern, M. D. (1997). Calcium in close quarters: microdomain feedback in excitation-contraction coupling and other cell biological phenomena. *Annual review of biophysics and biomolecular structure*, 26:47–82.

[Schiefer et al., 1995] Schiefer, A., Meissner, G., and Isenberg, G. (1995). Ca^{2+} activation and Ca^{2+} inactivation of canine reconstituted cardiac sarcoplasmic reticulum Ca^{2+}-release channels. *J Physiol (Lond)*, 489 (Pt 2):337–48.

[Shannon et al., 2004] Shannon, T. R., Wang, F., Puglisi, J., Weber, C., and Bers, D. M. (2004). A mathematical treatment of integrated Ca^{2+} dynamics within the ventricular myocyte. *Biophysical Journal*, 87(5):3351–71.

[Shen et al., 2004] Shen, J., Wang, S., Song, L., Han, T., and Cheng, H. (2004). Polymorphism of ca2+ sparks evoked from in-focus ca2+ release units in cardiac myocytes. *Biophys. J.*, 86:182–190.

[Shuai and Jung, 2002a] Shuai, J. and Jung, P. (2002a). Stochastic properties of ca2+ release of inositol 1,4,5-trisphosphate receptor clusters. *Biophys J*, 83:87–97.

[Shuai et al., 2007] Shuai, J., Pearson, J. E., Foskett, J. K., Mak, D. O., and Parker, I. (2007). A kinetic model of single and clustered IP$_3$ receptors in the absence of Ca^{2+} feedback. *Biophys. J.*, 93(4):1151–1162.

[Shuai et al., 2006] Shuai, J., Rose, H. J., and Parker, I. (2006). The number and spatial distribution of IP$_3$ receptors underlying calcium puffs in xenopus oocytes. *Biophys. J.*, 91(11):4033–4044.

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

[Shuai and Jung, 2003] Shuai, J. W. and Jung, P. (2003). Optimal ion channel clustering for intracellular calcium signaling. *Proc Natl Acad Sci USA*, 100(2):506–510.

[Smith, 2002] Smith, G. D. (2002). Modeling the stochastic gating of ion channels. In Fall, C., Marland, E., Wagner, J., and Tyson, J., editors, *Computational Cell Biology*, pages 291–325. Springer-Verlag.

[Smith and Parker, 2009] Smith, I. F. and Parker, I. (2009). Imaging the quantal substructure of single ip3r channel activity during ca puff in intact mammalian cells. *PNAS*, 106:6404–6409.

[Song et al., 1997] Song, L. S., Stern, M. D., Lakatta, E. G., and Cheng, H. (1997). Partial depletion of sarcoplasmic reticulum Ca^{2+} does not prevent Ca^{2+} sparks in rat ventricular myocytes. *J Physiol*, 505(Pt 3):665–75.

[Stern, 1992] Stern, M. D. (1992). Theory of excitation-contraction coupling in cardiac muscle. *Biophys J*, 63(2):497–517.

[Stern and Cheng, 2004] Stern, M. D. and Cheng, H. (2004). Putting out the fire: what terminates calcium-induced calcium release in cardiac muscle? *Cell Calcium*, 35(6):591–601.

[Stern et al., 1999] Stern, M. D., Song, L. S., Cheng, H., Sham, J. S., Yang, H. T., Boheler, K. R., and Rios, E. (1999). Local control models of cardiac excitation-contraction coupling. a possible role for allosteric interactions between ryanodine receptors. *J Gen Physiol*, 113(3):469–89.

[Stevens et al., 2009] Stevens, S. C., Terentyev, D., Kalyanasundaram, A., Periasamy, M., and Györke, S. (2009). Intra-sarcoplasmic reticulum Ca^{2+} oscillations are driven by dynamic regulation of ryanodine receptor function by luminal Ca^{2+} in cardiomyocytes. *J Physiol (Lond)*, pages 4863–4872.

[Stewart, 1994] Stewart, W. (1994). *Introduction to the Numerical Solution of Markov Chains*. Princeton University Press, Princeton.

[Sun et al., 1998] Sun, X. P., Callamaras, N., Marchant, J. S., and Parker, I. (1998). A continuum of IP_3-mediated elementary Ca^{2+} signalling events in Xenopus oocytes. *J Physiol*, 509(Pt 1):67–80.

[Swillens et al., 1998] Swillens, S., Champeil, P., Combettes, L., and 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(5):291–302.

[Swillens et al., 1999] Swillens, S., Dupont, G., Combettes, L., and Champeil, P. (1999). From Ca^{2+} blips to Ca^{2+} puffs: theoretical analysis of the requirements for interchannel communication. *Proc Natl Acad Sci U S A*, 96(24):13750–13755.

[To and Vohradsky, 2007] To, C. and Vohradsky, J. (2007). A parallel genetic algorithm for single class pattern classification and its application for gene expression profiling in streptomyces coelicolo. *BMC Genomics*, 8:49.

[Williams et al., 2007] Williams, G. S. B., Huertas, M. A., Sobie, E. A., Jafri, M. S., and Smith, G. D. (2007). A probability density approach to modeling local control of Ca^{2+}-induced Ca^{2+} release in cardiac myocytes. *Biophys. J.*, 92(7):2311–2328.

[Williams et al., 2008] Williams, G. S. B., Huertas, M. A., Sobie, E. A., Jafri, M. S., and Smith, G. D. (2008). Moment closure for local control models of Ca^{2+}-induced Ca^{2+} release in cardiac myocytes. *Biophys. J.*, 95(4):1689–1703.

[Yao et al., 1995] Yao, Y., Choi, J., and Parker, I. (1995). Quantal puffs of intracellular Ca^{2+} evoked by inositol trisphosphate in Xenopus oocytes. *J Physiol*, 482(Pt 3):533–53.

Vita

Yan Hao was born August 25, 1984 in Hebei, China. From 1996-2002 she attended Shijiazhuang No. 2 Middle School, and was admitted to Tsinghua University in June, 2002. She was rewarded a scholarship for academic excellence for her undergraduate study and received a Bachelor of Science in Applied Mathematics in May 2006. She began graduate studies of mathematical biology in The Department of Applied Science at The College of William & Mary in August 2006. She received her Master of Science in Applied Science in January 2009 and continued her studies of mathematical biology in pursuit of a Doctorate of Philosophy.