Markov Abstractions of Electrochemical Reaction-Diffusion in Synaptic Transmission for Neuromorphic Computing

Margot Wagner [1,2,3*], **Thomas M. Bartol** [3,4], **Terrence J. Sejnowski** [2,3,4,5], **and Gert Cauwenberghs** [1,2]

[1] *Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA*
[2] *Institute for Neural Computation, University of California, San Diego, La Jolla, CA, USA*
[3] *Computational Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA, USA*
[4] *Center for Theoretical Biological Physics, University of California, San Diego, La Jolla, CA, USA*
[5] *Division of Biological Sciences, University of California, San Diego, La Jolla, CA, USA*

Correspondence*:
Margot Wagner, Integrated Systems Neuroengineering Lab, University of California, San Diego, Department of Bioengineering, 9500 Gilman Dr, La Jolla, CA 92093
mwagner@ucsd.edu

ABSTRACT

Progress in computational neuroscience towards understanding brain function is challenged both by the complexity of molecular-scale electrochemical interactions at the level of individual neurons and synapses, and the dimensionality of network dynamics across the brain covering a vast range of spatial and temporal scales. Our work abstracts the highly detailed, biophysically realistic 3D reaction-diffusion model of a chemical synapse to a compact internal state space representation that maps onto parallel neuromorphic hardware for efficient emulation on very large scale, and offers near-equivalence in input-output dynamics while preserving biologically interpretable tunable parameters.

Keywords: neuromorphic, synapse, Markov chain, Monte Carlo, synaptic transmission, computational efficiency

1 INTRODUCTION

It has been known since the pioneering of computer architecture by John von Neumann that brains are far more effective and efficient in processing sensory information than digital computers, owing to the massively parallel distributed organization of neural circuits in the brain that tightly couple synaptic memory and computing at fine grain scale (von Neumann, 1958). Yet, modern day computers still follow the "von Neumann" architecture where computing and memory are kept separate, incurring severe penalties in computing bandwidth due to the bottleneck in data flow between fast centralized processing and slow vast memory. Moore's law relentless scaling of semiconductor technology, with a doubling of integration density

every two years, has allowed the von Neumann architecture to remain fundamentally unchanged since its advent. As the shrinking dimensions of transistors supporting Moore's law progression are approaching fundamental limits, it has become essential to consider alternative novel computing architectures to meet increasing computational needs in this age of the deep learning revolution, which itself is driven by advances rooted in a deeper understanding of brain function (Sejnowski, 2020). At the forefront of this movement are neuromorphic systems, introduced by Carver Mead (1990) as a solution to these limitations. Neuromorphic engineering looks towards human brains as inspiration for hardware systems due to their highly efficient computational nature. The human brain is regarded as the pinnacle of efficient computing, operating at an estimated rate of 10^{16} complex operations per second while consuming less than 20 W of power (Churchland and Sejnowski, 1992; Cauwenberghs, 2013). Therefore, neuromorphic engineering looks to mimic the function and organization of neural structures using hybrid analog and digital systems. There is significant overlap in the physics of computation between the two. Information is carried in the form of charge in which neuronal membrane dynamics are represented similarly in hardware using metal-oxide-semiconductor field-effect transistors (MOSFETs) (Mead, 1989). In the MOSFET sub-threshold region of operation, electrons and holes as the carriers of current between n- or p-type channels behave akin to ions flowing through ion channels that mediate current across the neuronal cell membrane. Fundamentally, these hardware systems share analogous properties to their biological counterparts including diffusion as the primary mechanism of carrier transport, charge stochasticity, and energy barriers modulated by gating voltage. Paired with Boltzmann distributions of charge, these systems are able to emulate current as an exponential function of the applied voltage, capturing the same biophysics underlying the neuronal dynamics (Mead, 1989; Broccard et al., 2017).

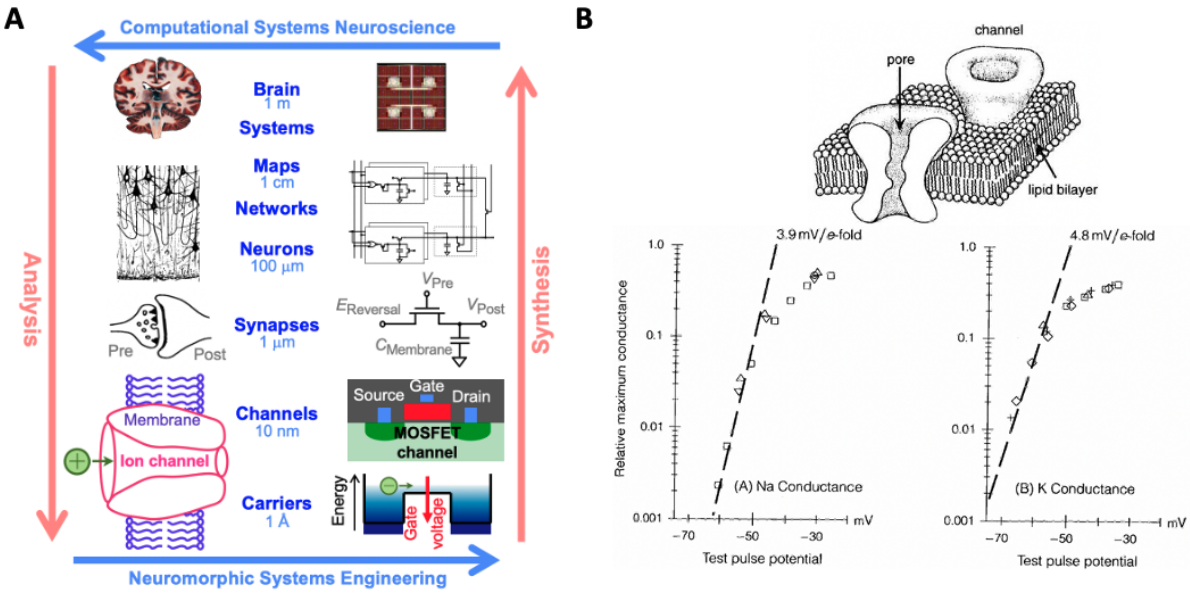


Figure 1. (A) Multiscale correspondence between the fields of computational neuroscience and neuromorphic engineering (Churchland and Sejnowski, 1992; Cauwenberghs, 2013) and (B) equivalence in Boltzmann thermodynamics between metal-oxide-semiconductor field-effect transistors (MOSFETs) (left) and ion channels resulting in current as an exponential function of applied voltage in both systems (right) (Hodgkin and Huxley, 1952; Mead, 1989).

Since the introduction of neuromorphic engineering, models of different complexity have been introduced for neurons, typically ranging from more detailed and realistic conductance-based Hodgkin-Huxley models

to simpler integrate-and-fire models allowing for better scalability. Synaptic connectivity between neurons is of primary concern in the field currently because synaptic strength and plasticity is fundamental to learning and memory in both biological and artificial representations of neural networks (Broccard et al., 2017; Indiveri et al., 2011). In neuromorphic architectures, synapses instantiate both computation and memory, and a new focus on compact electronic implementations of this computational memory has been emerging recently including the use of memristors (Boybat et al., 2018). Efficient representation of synapses is a crucial topic of concern as there are roughly 10^4 synapses for each neuron, totalling approximately 10^{16} in the human brain. They are diverse in nature and have highly complex temporal and spatial dynamics, which further complicates their representations (Broccard et al., 2017). Currently, there is a push for efficient synaptic models while maintaining the intricate dynamical behavior exhibited biophysiologically. Current models include time-multiplexing synapses, analog bistable synapses, and binary synapses to name a few, but the need for scalable and dynamically complex models of synaptic function and transmission is still existent and critical (Broccard et al., 2017; Bartolozzi and Indiveri, 2007).

Modelling synapses is a challenging task due to their intricacy and sheer quantity. As noted above, there are an estimated 10^{16} synapses in the human brain. They vary in function and type, including both chemical and electrical synapses. They are incredibly variable, and exhibit behavior spanning multiple different time and spatial scales, as well as being highly stochastic in nature (van Rossum et al., 2003; Wang et al., 2020). Additionally, synaptic plasticity causes changes in strength over time, associated with learning and memory. Synaptic transmission involves a multitude of mechanisms and molecular components, making simulations including all components not readily scalable. In order to capture the sophisticated dynamics of synapses in a scalable manner, abstractions have to be made according to the research problem in question. The stochastic nature of synapses also makes large scale simulations more complicated as modelling stochastic processes is typically more computationally demanding. It has been shown in multiple instances that the noise present due to the stochastic variability in synapses is highly integral to synaptic transmission, so this becomes an important feature to maintain (Malagon et al., 2016). For example, Moezzi et al. (2014) proved that models including ion channel noise in calcium channels paired with the existence of a presynaptic mechanism causing random delays in synaptic vesicle availability best capture the interspike interval behavior of auditory nerve fiber models. Additionally, multiple experimental works have found the existence of presynaptic vesicles that are released into the synaptic cleft with some probability (Castillo and Katz, 1954; Korn and Faber, 1991). There are multiple similar conclusions found in modelling and experimental results as recently discussed by McDonnell et al. (2016).

Synapses form the connections between neurons and the strength of these connections changes over time, forming the basis of learning and memory in both biological and artificial neural networks. The computations involved in accurately modelling the biophysics of synapses are complex due to the highly nonlinear nature of their dynamics. Yet, most of the neural network models in use today abstract synaptic strength to a single scalar value, tuned to a specific task. The learning rule for updating synaptic strength is typically globally applied using abstractions of synaptic plasticity such as spike-time dependent plasticity, and its causal extensions for scalable real-time hardware implementation (Pedroni et al., 2019). Physical constraints and limitations in VLSI implementations restrict the functional form of synaptic representation. In turn, these abstractions restrict the potential computing power of neuromorphic systems and restrain achievable benchmarks in approaching the functional flexibility, resilience, and efficiency of neural computation in the biological brain. Our work addresses the need for a more biophysically realistic model of the synapse with biologically tunable parameters to represent synaptic plasticity, while offering a path towards efficient real-time implementation in neuromorphic hardware.

Synaptic transmission is dictated by a series of events initiated by presynaptic stimulation in the form of action potentials. An action potential causes membrane depolarization which leads to stochastic opening and closing of voltage-dependent calcium channels (VDCCs) lying on the presynaptic membrane and a resulting influx of calcium to the presynaptic terminal. Neurotransmitter release is modulated by calcium binding to calcium sensors near the neurotransmitter filled vesicles at the active zone, but calcium has other fates as it diffuses from the VDCCs. In addition to binding to the calcium sensors, it can bind to calbindin, which acts as a buffer, or it can be removed by plasma membrane calcium (PMCA) pumps. If enough calcium is able to bind to the calcium sensors, though, then neurotransmitters are released across the synaptic cleft and initiate downstream effects at the postsynaptic membrane. This process of synaptic transmission is the basis of learning and communication in the brain.

Abstracting this for computational efficiency, we created a series of Markov state transitions to fully realize the system with multiple internal states allowing for a biophysically tunable model of synaptic connectivity implementable in neuromorphic architectures. Markov models have a history of use as a stochastic discrete state alternative to Hodgkin-Huxley type formulations since their introduction (Hodgkin and Huxley, 1952; Armstrong, 1971; Colquhoun, 1973). Additional stochastic models have been introduced, including the Gillespie method (1977), which has been used to model neural channel noise (Gillespie, 1977; Skaugen and Walloe, 1979; Chow and White, 1996). Markov models have also found use in whole-cell models (Winslow et al., 1999). Further extensions utilize a particle model (Koch, 1999). The importance of the inclusion of stochasticity in ion channel behavior and synaptic transmission generally cannot be understated. Its inclusion has been demonstrated time and time again in experimental work and is thought to be integral in the form and function of synaptic transmission (McDonnell et al., 2016). This provides an additional complication in modelling synapses and has been handled in various different stages of transmission, including the stochastic models of vesicle release using probabilistically generated quantal components, stochastic models of transmitter diffusion, and stochastic models of receptors (Castillo and Katz, 1954; Bartol et al., 2015; van Rossum et al., 2003). These simulations are computationally expensive due to the high transition rates paired with the small number transitions necessitates a small time step. Specifically, Markov models have shown to be an effective method of modelling ion channels, but require high computational cost to effectively do so.

This paper looks to abstract the computational complex and nonlinear nature of synaptic transmission dynamics in a manner that is efficient and readily scalable for implementation in neuromorphic silicon very large-scale integrated (VLSI) circuits. This is done by introducing an efficient stochastic sampling scheme within a Markov chain representation of the components integral to stochastic presynaptic quantal transmission.

2 MATERIALS AND METHODS

2.1 Markov Chain Models

The cascade of events from the action potential stimulus input to the presynaptic neurotransmitter release output can be equivalently modelled as Markov chain to fully realize the system with multiple internal states instead of directly tracking all molecules and their kinetics in a computationally complex spatiotemporal 3D reaction-diffusion model. Each internal Markov state is assumed to be dependent solely on the state at the previous timepoint and is conditionally independent of all previous timepoints, simplifying simulations. Therefore, the fully biophysically complex system of synaptic transmission can be abstracted and sampled to create a Markov Chain Monte Carlo (MCMC) which answers the same question of neurotransmitter release utilizing tunable biophysical parameters while allowed scalability

for implementation in neuromorphic architectures. The MCMC was implemented in Python using state transition models creating more efficient simulation for internal states of the presynaptic terminal and neurotransmitter release due to stimulus input.

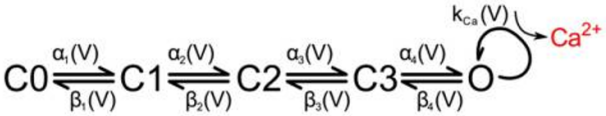


Figure 2. State diagram for voltage-dependent calcium channels and resulting calcium influx in the presynaptic membrane (Bartol et al., 2015).

For any given stimulus input, the VDCCs are assigned transition probabilities between states based on a five-state kinetic model (Figure 2) found experimentally and validated computationally with four closed states and a single open state (Bischofberger et al., 2002; Church and Stanley, 1996; Bartol et al., 2015). Prior to the stimulus, all VDCCs begin in a closed state, and the concentration of calcium in the presynaptic terminal is at steady-state. The transition probabilities are voltage dependent akin to a Hodgkin-Huxley model where the number of open VDCCs at any given moment is used to determine the number of calcium entering the presynaptic terminal based on experimental I-V curves and the resulting I-V equation found in (Bischofberger et al., 2002). Calcium influx is captured similarly by including transitions from the open VDCC state to the open VDCC state and an internal calcium generation. Using this, influx of calcium is modelled over the entire stimulus input due to VDCC opening.

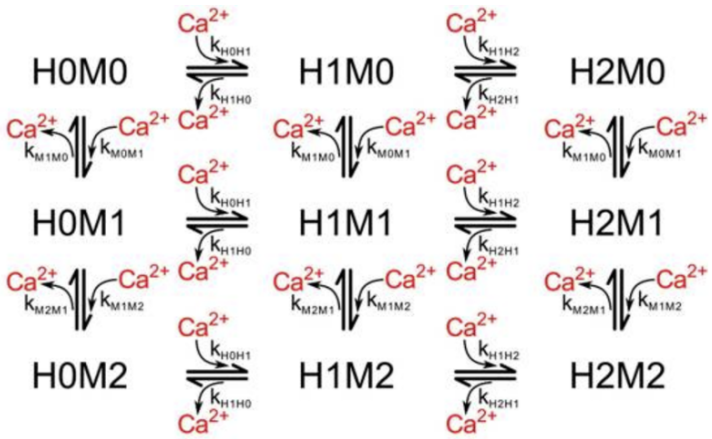


Figure 3. State diagram for calbindin binding (Bartol et al., 2015).

Once calcium has entered the presynaptic terminal, much of it binds to calbindin, which acts as a buffer and primarily modulates the amount of calcium that is able to reach the calcium sensors at the active zone. The state transitions are reversible first-order reactions, thus transition probabilities are dependent of the free calcium in the system and are updated as that amount changes over time. Calbindin has four binding sites, two high affinity sites and two medium affinity sites, leading to a nine-state calcium concentration-dependent kinetic model (Figure 3) (Nagerl et al., 2000). By modelling the binding and unbinding of calcium to calbindin as a loss or gain of free calcium respectively, calcium transients can also be elucidated.

Our Markov chain is a discrete-state chain in discrete time. Markov chains are modelled by a probability that the chain will move to another state given its current state and is conditionally independent of all previous timesteps. The probabilities are by nature only dependent on the current state of the Markov chain, thus:

$$P(X_{n+1} = x_j | X_n = x_i, X_{n-1}, ..., X_0) = P(X_{n+1} = x_j | X_n = x_i) = P_{ij} \tag{1}$$

For state transitions, the probability is the transition rate times the time step. In the case of calbindin transitions, this is further multiplied by the amount of free unbound calcium for forward reactions. In the case of a multi-state system, this gives a transition probability matrix for the likelihood of transition from a given state at the current time step to any other state at the next time step. This matrix is sparse, with nonzero probabilities only for adjacent states where transition is possible. In the case of the five-state VDCC system, this gives

$$P_{ij} = \begin{cases} k_{ij}\,\Delta t & j = i \pm 1 \\ 1 - \sum_{k=0}^{N} P_{ik} & j = i \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Typically Markov state transitions are modelled via a discrete inverse transform method, where given a random variable X the transition probabilities $P(X = x_i)$ describe a partition of unity (Figure 4). Therefore, we can generate a random number uniformly, $R \sim U(0, 1)$ and map it into discrete values of X. For example, in a two state system, $X_j = 0$ if $R \leq P_{i,0}$ or $X_j = 1$ if $P_{i,0} < R \leq P_{i,0} + P_{i,1} = 1$. This involves searching the state space for the next state given the current state for each molecule in the system at each time step, which can be a slow process for systems with a large number of states and molecules.

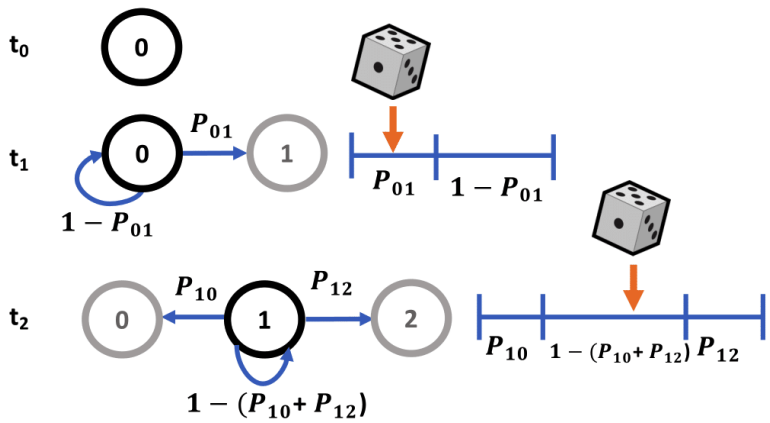


Figure 4. Markov sampling scheme for state transitions.

Here we have implemented an optimized MCMC sampling strategy involving sampling from a multinomial distribution, where $X \sim \text{Multi}(n, p_1, \ldots p_k)$ with the probability mass function for n experiments each with one of p possible outcomes:

$$f(x_1, \ldots, x_k; n, p_1, \ldots p_k) = \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} \ldots p_k^{x_k} \tag{3}$$

At each time point, the number of transitioning molecules of interest is n while the transition probabilities, whether voltage or concentration dependent, are p_k. This expedites computation by only requiring a single computation at each time, sampling the distribution of all n molecules at once.

2.2 MCell Models

MCell is a modelling software that uses spatially realistic 3D geometries and Monte Carlo reaction-diffusion modelling algorithms, which allows for biophysically realistic simulations of high complexity as it specifically tracks the state of every molecules in space and time (Bartol et al., 2015). Due to the accuracy and specificity, it provides a ground truth for biological simulations but does so at the cost of computational complexity.

To validate and compare our Markov models of synaptic transmission, we built a biophysically realistic stochastic 3D reaction-diffusion system with all major components for presynaptic vesicle release variability in response to a stimulus input (Figure 5). The model included realistic geometry for a CA3-CA1 *en passant* synapse with parameters set from experimental data (Nadkarni et al., 2010; Bartol et al., 2015). The CA3-CA1 synapse was chosen for the simulations as it is highly studied experimentally and is important for learning and memory. All kinetics and parameters match those used for the equivalent Markov models.

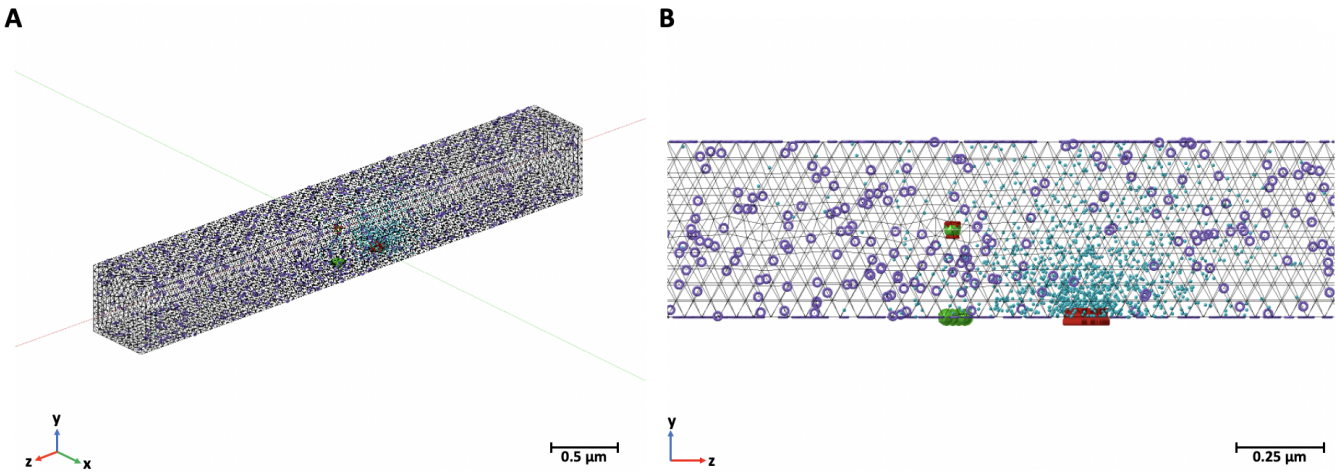


Figure 5. MCell model for synaptic transmission containing voltage-dependent calcium channels (red), calcium (blue), calcium sensors (green), and plasma membrane calcium pumps (purple). Calbindin not pictured here due to their large number.

The MCell model includes the canonical presynaptic geometry for an average terminal as a rectangular box measuring 0.5μm by 0.5μm by 4μm containing the calbindin buffer, steady-state calcium concentration, PMCA pumps, VDCCs and calcium sensors modulating neurotransmitter release (Nadkarni et al., 2010). The detailed kinetics of these systems have been discussed in detail elsewhere and are based on experimental data (Bartol et al., 2015). The active zone is based on that of an average presynaptic active zone with only one being including and containing seven docked neurotransmitter vesicle release sites. The VDCCs of type P/Q and are stationed at a biophysically realistic distance from the active zone, also located on the membrane. PMCAs are homogenously placed across the presynaptic membrane while calbindin are placed in the volume. This is a flexible architecture that can respond to any stimulus input and allows for monitoring of that states of each molecule in the system. To compare with the Markov models, we used the same single action potential stimulus.

3 RESULTS

3.1 Voltage-Dependent Calcium Channels

The efficient Markov chain implementation has strong agreement with the full MCell model in terms of the internal state transients in response to an external stimulus. Figure 6a shows the fraction of open VDCCs over time in response to the action potential, which controls the amount of calcium influx to the system. Transients for all states can be found in the supplemental material. As expected, the number of closed VDCCs (state C0) decreases over the duration of the stimulus. The internal states (C1-C3) subsequently increase and decrease as the membrane voltage increases and the forward rates for the VDCCs increase, leading to an exponential increase in the open VDCCs while the membrane depolarizes. At the maximum membrane potential, almost all VDCCs are in the open state. As the membrane repolarizes, the reverse reaction rates constants increase, and the VDCCs close. This leads to another increase and decrease in the internal VDCC states as the receptors go from their open to resting state (C0).

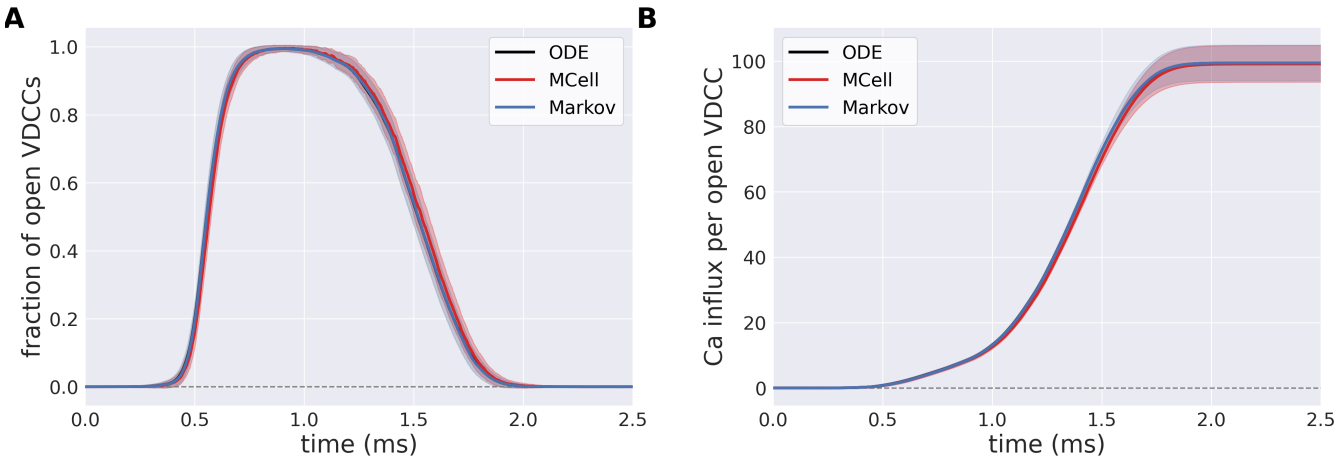


Figure 6. (A) Fraction of open voltage-dependent calcium channels (VDCCs) and (B) calcium influx through open VDCCs, in response to action potential stimulus for stochastic MCell, Markov, and deterministic ODE forward Euler simulation.

In its open state, VDCCs allow for the probabilistic influx of calcium through the channels into the presynaptic bouton. This is exemplified in Figure 6b, where there is an increase in the calcium influx through the open VDCCs over the course of the stimulus. Again, there is strong agreement between the more computationally complex MCell model and the computational efficient Markov equivalent model.

3.2 Calbindin Buffer

Simulations of homogeneous calcium and calbindin were run using both the MCell, Markov and deterministic simulation schemes. In the presence of calcium, the forward binding reaction is heavily favored and this is highlighted in Figure 7 (left) where free calcium exponentially decreases. A similar transient is apparent for the unbound state of calbindin, as it quickly transitions to different stages of high and medium binding. While over the course of the simulation, all the free calcium is removed from the system, calbindin states reach a new steady-state where there are still unbound calbindin. Similarly, the fully bound state, H2M2, rapidly increases and reaches a new steady state that is still only 1% of all calbindin. This is due to the high concentration of calbindin in the presynaptic bouton. Even once all the calcium is in a bound state, there is plenty of unbound or partially bound calbindin remaining in the system. Calbindin acts as a strong buffer allowing for calcium storage and asynchronous neurotransmitter

release, so this and slow unbinding rates becomes an important feature of calbindin. The rapid extent to which calcium binds to calbindin shows the impact of buffering on calcium's ability to diffuse and bind to the calcium sensors in the active zone. The inclusion of calbindin at such high concentrations becomes a key feature of maintaining the relatively low release rates of neurotransmitters even in the presence of a stimulus.

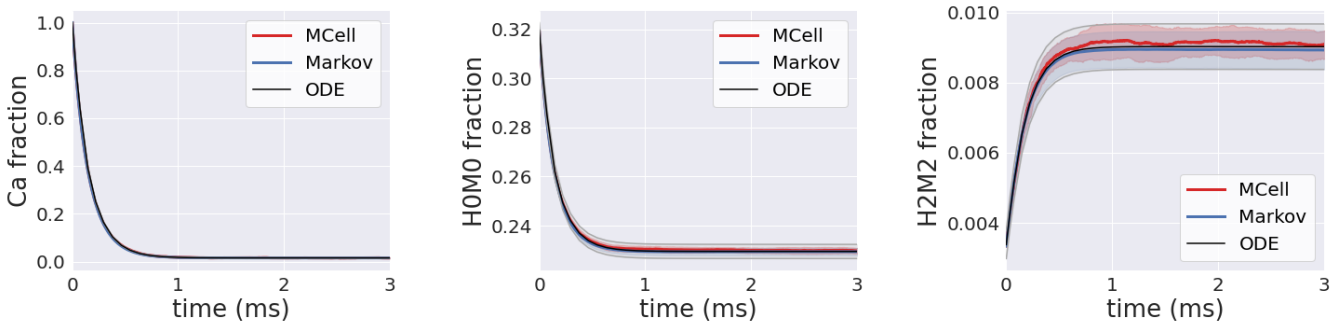


Figure 7. Transients for homogeneous calcium-calbindin buffer binding in the presynaptic bouton for free calcium *(left)*, the unbound calbindin state, H0M0 *(middle)*, and the fully bound calbindin state, H2M2 *(right)*.

3.3 Benchmarks

Runtime and total floating point operations were used as metrics for comparison between the simulation methods (Table 1). Here we compare MCell, the standard Markov model, and the multinomial Markov stochastic models. The deterministic Euler solution is included as well for a standard comparison. For the VDCC simulations, the optimized multinomial sampling Monte Carlo Markov chain model has a runtime on the order for the forward Euler deterministic solution. The MCell and standard Markov stochastic models exemplify a runtime an order of magnitude higher. The number of operations is also much high for the standard Markov model as compared to the multinomial Markov model.

Table 1. Benchmarks for different simulation types for both voltage-dependent calcium channel and calbindin binding simulations.

VDCC	**Runtime**	**Total Floating Point Operations (MFLOPs)**
MCell	139.20 ± 7.00 s	--
Standard Markov	109.16 ± 0.93 s	70.10
Multinomial Markov	9.40 ± 0.08 s	12.60
ODE (Euler)	10.05 ± 0.13 s	7.73
Calbindin	**Runtime**	**Total Floating Point Operations (MFLOPs)**
MCell	29.34 ± 0.39 s	--
Standard Markov	4.10E+4 ± 0.00 s	46.10
Multinomial Markov	2.37 ± 0.02 s	27.20
ODE (Euler)	0.72 ± 0.01 s	1.60

For the calbindin model, the multinomial Markov method is again an order of magnitude faster than the MCell model although it is also an order of magnitude slower than the deterministic model. The standard Markov model is four orders of magnitude slower than the multinomial model due to the high density of calbindin molecules in the simulation. This highlights the importance of our multinomial sampling scheme for simulations with large numbers of molecules and multiple states. The operations are also fewer for the multinomial case than the standard case. Overall, the multinomial Markov simulation provides a computationally efficient alternative to stochastic MCell simulations while maintaining the biological accuracy.

3.4 Complexity Analysis

MCell uses a scheduler, which allows for only making changes to the scheduled particles, though in the worst-case, this still scales with the total number of particles in the simulation, n. At each time step, a particle can transition to any of its adjacent or branched states, b, which is similarly described by a fan-out factor in electronic implementation. Therefore, the overall time complexity for MCell is $O(bnt)$. The classical Markov representation tracks every particle, regardless of if its transitions at each time point or not as well as sampling from the neighboring states for transitions at each time point, leading to a similar $O(bnt)$ time complexity, or at best $O(nt \log_2 b)$ for sampling implemented with efficient search algorithms. Both the multinomial Markov model and the Euler ODE implementation describe the system in terms of the state concentrations, leading to a dependence on the total number of states, $s \geq b$, rather than the number of particles. Due to efficient sampling methods, the multinomial Markov method is independent of the number of adjacent states, leading to a time complexity for both the multinomial Markov and Euler ODE methods of $O(bst)$.

Table 2. Space and time complexity for the various simulation strategies.

Complexity	Time Complexity	Space Complexity
MCell	$O(b\ n\ t)$	$O(b\ n \log_2 s)$
Standard Markov	$O(b\ n\ t)$	$O(b\ n \log_2 s)$
Multinomial Markov	$O(b\ s\ t)$	$O(b\ s \log_2 n)$
ODE (Euler)	$O(b\ s\ t)$	$O(b\ s \log_2 n)$

n (total number of molecules), s (total number of states), b (number of branching states), t (number of timepoints)

With regards to space complexity, the traditional Markov as well as MCell representation stores the state in bits for each particle as well as the branched states from the current states, leading to complexity of $O(bn \log_2 s)$ The efficient Markov model and ODE solution both simply store the number of molecules represented by bits in each state at each timepoint as well as the branched states resulting in a space complexity of $O(bs \log_2 n)$. There exists a trade off here between the number of particles in each state compared to the number of states where one is stored directly and one is stored as an index. Thus, for simulations where the number of states is less than the number of particles, the multinomial Markov model is an efficient representation of the system, which is typically the case for biochemical simulation. MCell is more efficient with large state-space systems, but the number of states could be sparsified in a multinational Markov representation by implementing dynamic instantiation and annihilation of states. Additionally, unseen or rarely seen states could be ignored by truncating based on probability of a particle being in

that state. This would functionally decrease the number of states in the system allowing for use of the multinomial Markov simulation method.

4 DISCUSSION

The goal of this work was to create a more computational efficient model of biologically realistic synaptic transmission for use in large scale neuromorphic systems. We created a multinomial Markov Chain Monte Carlo (MCMC) sampling strategy for capturing the internal states of vital molecules in the system in response to stimulus where transition probabilities could be voltage- or concentration-dependent and the next timestep could be predicted solely using the current timestep. This scheme was implemented to capture the dynamics of the stochastic opening and closing of voltage-dependent calcium channels (VDCCs) through multiple internal states as well as the resulting calcium influx into the presynaptic bouton through the open VDCCs. Once calcium has entered the presynaptic terminal, we also simulated calcium binding to the calbindin buffer which modulates calcium levels in the bouton, directly impacting the amount of calcium that reaches the calcium sensors in the active zone. This amount impacts the neurotransmitter release from the presynaptic side and the resulting effects on the postsynaptic side.

All simulations were modelled using the optimized multinomial Markov sampling method as well as a typical Markov sampling method and compared to highly detailed 3D geometric stochastic reaction-diffusion simulations done using MCell. The Markov simulations show agreement with the MCell simulations for the system dynamics including the number of open VDCCs and calcium influx in response to an action potential stimulus as well as the binding of calcium to the calbindin buffer. Differences are observed from the deterministic solution to the stochastic simulations implying the importance of stochasticity in these simulations to capture more biologically-realistic systems.

Exemplified by runtime and total number of operations, the optimized multinomial MCMC method of simulations was shown to be more efficient than the standard Markov model while being faster than the MCell equivalents. This is hopeful for scaling these biologically-realistic models to large-scale systems while maintaining biological tunability.

Next steps involve capturing modelling the remaining kinetics in a similar fashion including the binding and removal of calcium by the plasma membrane calcium (PMCA) pumps as well as binding to the calcium sensors. In addition, to capture the diffusion of calcium through the presynaptic terminal but specifically to the calcium sensors at the active zone, a diffusive kernel must be included to the system. Separation of space-time variables for the particles paired with the diffusive kernel provides a method to implement such a model utilizing an operator splitting approach. Separation of variables naturally has finite elements in space, which works to quantize the complex synaptic space and quantify boundary conditions allowing for a finite element approach where eigenvalues will give decay rates for transitions between spatially separated Markov graphs. The transition rates will be due to the difference between spatial nodes and the Markov reaction transitions.

Upon inclusion of these elements, the entire process from stimulus to neurotransmitter release can be modelled as a series of Markov chains leading to powerful implications for synaptic transmission modelling. The while synapse can be included as well with the inclusion of a diffusive kernel across the synaptic cleft as well as downstream effects on the postsynaptic side, of which many mirror similar kinetics and dynamics as the presynaptic side leading to a natural extension of this modelling framework. The resulting system would be a biologically tunable model of synaptic transmission for any stimulus input in a highly efficient manner. This opens the door for large-scale implementations of synaptic transmission and learning readily implementable into neuromorphic architectures with biological realism never before seen.

Through the utilization of Markov-based abstractions applied to biophysically realistic 3D reaction-diffusion models of a chemical synaptic, we have created a compact and efficient internal state space representation of transmission. This is in response to the challenge presented by the high dimensionality and complex nature of molecular-scale interactions in synapses and across scales making implementation in very large-scale systems previously unattainable. The model is directly amenable to efficient emulation in parallel neuromorphic hardware systems while maintaining biophysically relevant and interpretable parameters that are readily tunable. This opens the door towards neuromorphic circuits and systems on very large scale that strike a greater balance between integration density and biophysical accuracy in modelling neural function at the whole-brain level.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

FUNDING

This work was supported by a NSF Graduate Research Fellowship to MW, the National Science Foundation, and the Office of Naval Research.

REFERENCES

Armstrong, C. M. (1971). Interaction of tetraethylammonium ion derivatives with the potassium channels of giant axons. *Journal of General Physiology* 58, 413–37. doi:10.1085/jgp.58.4.413

Bartol, T. M., Keller, D. X., Kinney, J. P., Bajaj, C. L., Harris, K. M., Sejnowski, T. J., et al. (2015). Computational reconstitution of spine calcium transients from individual proteins. *Frontiers in Synaptic Neuroscience* 7. doi:10.3389/fnsyn.2015.00017

Bartolozzi, C. and Indiveri, G. (2007). Synaptic dynamics in analog VLSI. *Neural Computation* 19, 2581–2603. doi:10.1162/neco.2007.19.10.2581

Bischofberger, J., Geiger, J. R. P., and Jonas, P. (2002). Timing and efficacy of Ca^{2+} channel activation in hippocampal mossy fiber boutons. *The Journal of Neuroscience* 22. doi:10.1523/JNEUROSCI.22-24-10593.2002

Boybat, I., Gallo, M. L., Nandakumar, S. R., Moraitis, T., Parnell, T., Tuma, T., et al. (2018). Neuromorphic computing with multi-memristive synapses. *Nature Communications* 9. doi:10.1038/s41467-018-04933-y

Broccard, F. D., Joshi, S., Wang, J., and Cauwenberghs, G. (2017). Neuromorphic neural interfaces: from neurophysiological inspiration to biohybrid coupling with nervous systems. *Journal of Neural Engineering* 14. doi:10.1088/1741-2552/aa67a9

Castillo, J. and Katz, B. (1954). Quantal components of the end-plate potential. *The Journal of Physiology* 124, 560–573. doi:10.1113/jphysiol.1954.sp005129

Cauwenberghs, G. (2013). Reverse engineering the cognitive brain. *PNAS* 110, 15512–15513. doi:10.1073/pnas.1313114110

Chow, C. C. and White, J. A. (1996). Spontaneous action potentials due to channel fluctuation. *Biophysical Journal* 71, 3013–3021. doi:10.1016/S0006-3495(96)79494-8

Church, P. and Stanley, E. (1996). Single L-type calcium channel conductance with physiological levels of calcium in chick ciliary ganglion neurons. *The Journal of Physiology* 496. doi:10.1113/jphysiol.1996.sp021665

Churchland, P. S. and Sejnowski, T. J. (1992). *The Computational Brain* (The MIT Press)

Colquhoun, D. (1973). *The relation between classical and cooperative models for drug action*. Drug receptors (Macmillan Press)

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* 81, 2340–2361. doi:10.1021/j100540a008

Hodgkin, A. and Huxley, A. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology* 117, 500–544. doi:10.1113/jphysiol.1952.sp004764

Indiveri, G., Linares-Barranco, B., Hamilton, T. J., van Schaik, A., Etienne-Cummings, R., Delbruck, T., et al. (2011). Neuromorphic silicon neuron circuits. *Frontiers in Neuroscience* 5. doi:10.3389/fnins.2011.00073

Koch, C. (1999). *Biophysics of Computation* (Oxford University Press)

Korn, H. and Faber, D. (1991). Quantal analysis and synaptic efficacy in the CNS. *Trends in Neuroscience* 14, 439–45. doi:10.1016/0166-2236(91)90042-s

Malagon, G., Miki, T., Llano, I., Neher, E., and Marty, A. (2016). Counting vesicular release events reveals binomial release statistics at single glutamatergic synapses. *The Journal of Neuroscience* 36, 4010–4025. doi:10.1523/JNEUROSCI.4352-15.2016

McDonnell, M. D., Goldwyn, J. H., and Lindner, B. (2016). Neuronal stochastic variability: Influences on spiking dynamics and network activity. *Frontiers in Computational Neuroscience* 10. doi:10.3389/fncom.2016.00038

Mead, C. (1989). *Analog VLSI and Neural Systems* (Addison-Wesley)

Mead, C. (1990). Neuromorphic electronic systems. In *Proceedings of the IEEE*. vol. 78, 1629–1636

Moezzi, B., Iannella, N., and McDonnell, M. D. (2014). Modeling the influence of short term depression in vesicle release and stochastic calcium channel gating on auditory nerve spontaneous firing statistics. *Frontiers in Computational Neuroscience* 8. doi:10.3389/fncom.2014.00163

Nadkarni, S., Bartol, T. M., Stevens, C. F., Sejnowski, T. J., and Levine, H. (2010). Modelling vesicular release at hippocampal synapses. *PLoS Computational Biology* 6. doi:10.1371/journal.pcbi.1000983

Nagerl, U. B., Novo, D., Mody, I., and Vergara, J. L. (2000). Binding kinetics of calbindin-D28k determined by flash photolysis of caged Ca^{2+}. *Biophysical Journal* 79, 3009–3018. doi:10.1016/S0006-3495(00)76537-4

Pedroni, B. U., Joshi, S., Deiss, S. R., Sheik, S., Detorakis, G., Paul, S., et al. (2019). Memory-efficient synaptic connectivity for spike-timing-dependent plasticity. *Frontiers in Neuroscience* 13. doi:10.3389/fnins.2019.00357

Sejnowski, T. J. (2020). *The Deep Learning Revolution* (The MIT Press)

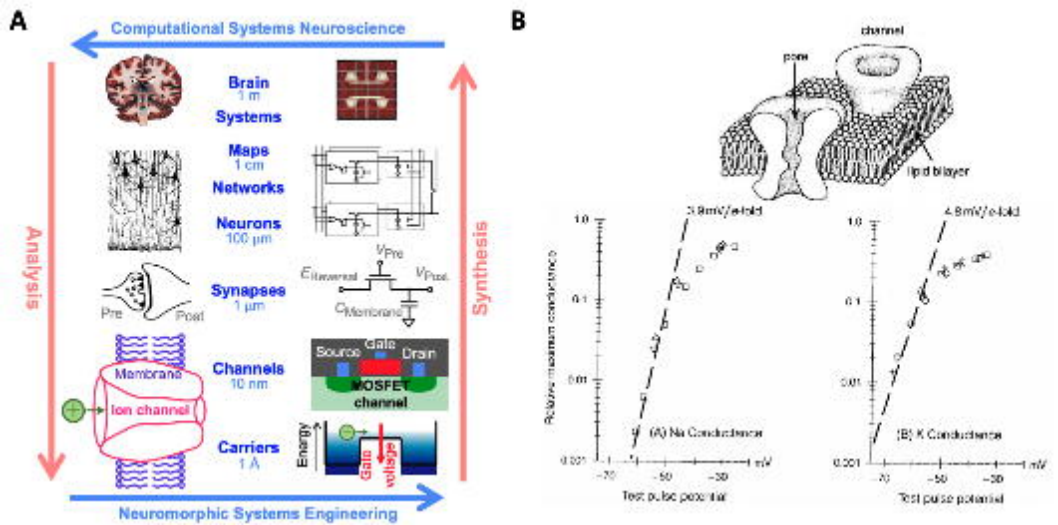
Skaugen, E. and Walloe, L. (1979). Firing behavior in a stochastic nerve membrane model based upon the Hodgkin-Huxley equations. *Acta Physiologica Scandinavica* 107, 343–363. doi:10.1111/j.1748-1716.1979.tb06486.x

van Rossum, M. C., O'Brien, B., and Smith, R. (2003). Effects of noise on the spike timing precision of retinal ganglion cells. *The American Physiological Society* 89, 2406–2419. doi:10.1152/jn.01106.2002

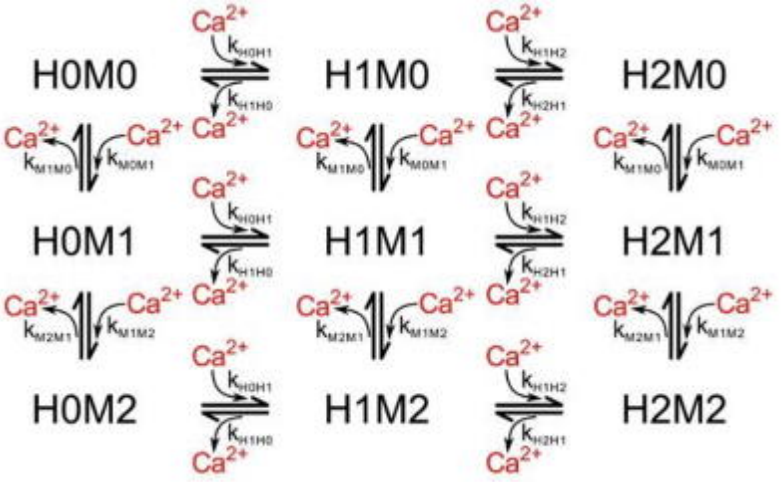
von Neumann, J. (1958). *The Computer and the Brain* (Yale University Press)

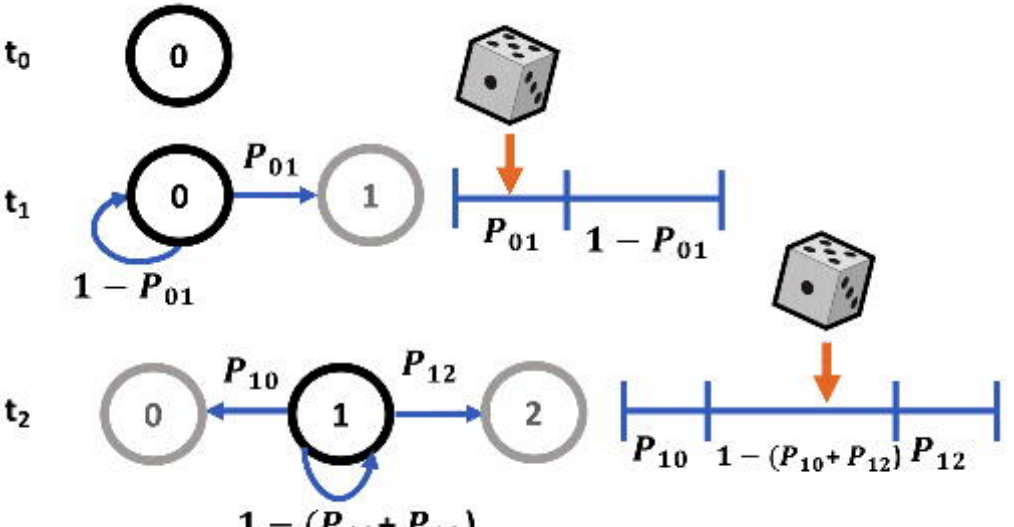
Wang, J., Cauwenberghs, G., and Broccard, F. D. (2020). Neuromorphic dynamical synapses with reconfigurable voltage-gated kinetics. *IEEE Transactions on Biomedical Engineering* 67, 1831–1840. doi:10.1109/TBME.2019.2948809

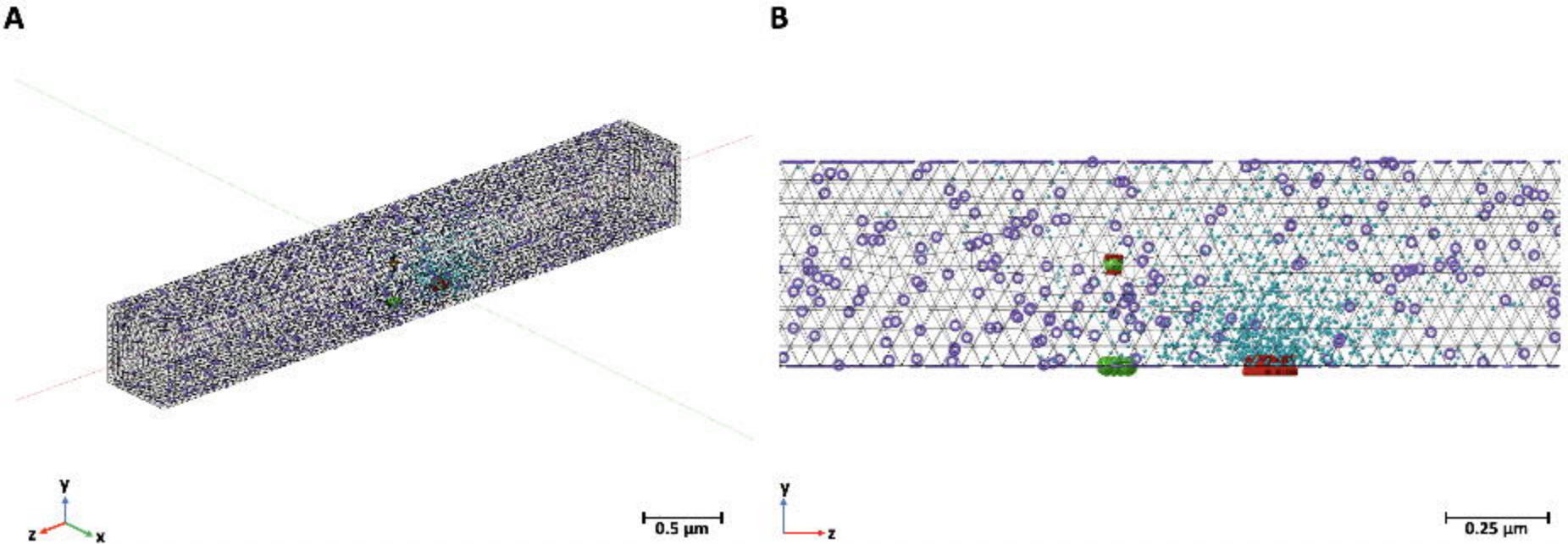
Winslow, R. L., Rice, J., Jafri, S., Marbán, E., and O'Rourke, B. (1999). Mechanisms of altered excitation-contraction coupling in canine tachycardia-induced heart failure, ii. *Circulation Research* 84, 571–586. doi:10.1161/01.RES.84.5.571

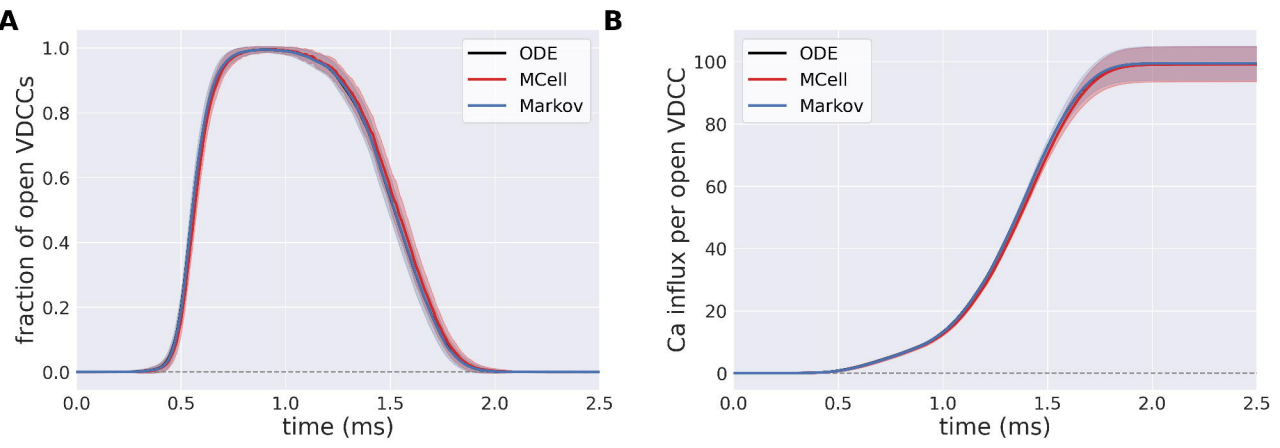

A
Computational Systems Neuroscience
Analysis
Synthesis
Brain
1 m
Systems
Maps
1 cm
Networks
Neurons
100 μm
Synapses
1 μm
Pre
Post
Channels
10 nm
Membrane
Ion channel
Carriers
1 Å
Source
Gate
Drain
MOSFET
channel
Energy
Gain
voltage
Neuromorphic Systems Engineering
B
channel
pore
lipid bilayer
(A) Na Conductance
(B) K Conductance
Test pulse potential

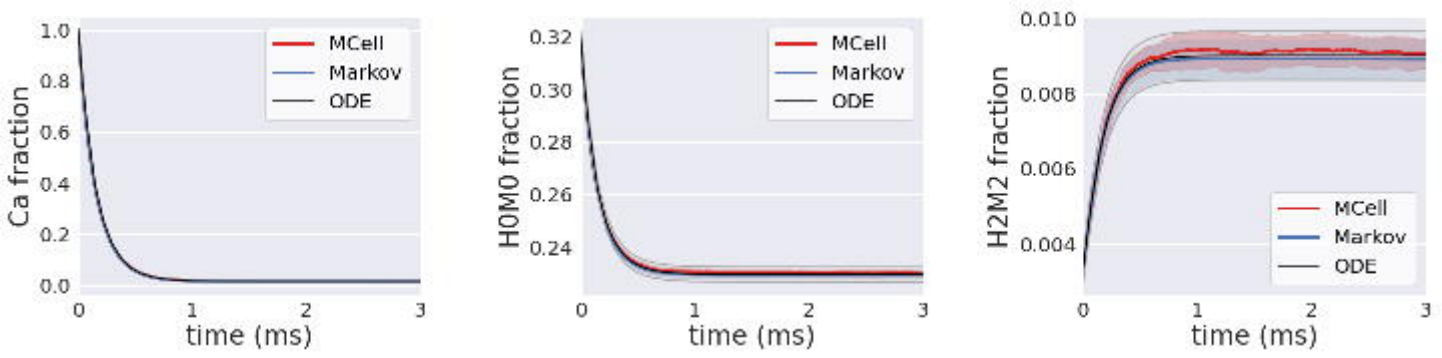
C0
C1
C2
C3
O
α1(V)
β1(V)
α2(V)
β2(V)
α3(V)
β3(V)
α4(V)
β4(V)
kCa(V)
Ca2+


H0M0
H1M0
H2M0
H0M1
H1M1
H2M1
H0M2
H1M2
H2M2
Ca2+
kH0H1
kH1H0
kH1H2
kH2H1
kM1M0
kM0M1
kM2M1
kM1M2


t_0
0
t_1
0
P_{01}
1
$1 - P_{01}$
P_{01}
$1 - P_{01}$
t_2
0
P_{10}
1
P_{12}
2
$1 - (P_{10} + P_{12})$
P_{10}
$1 - (P_{10} + P_{12})$
P_{12}


A
y
z
x
0.5 µm
B
y
z
0.25 µm


A
fraction of open VDCCs
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.5
1.0
1.5
2.0
2.5
time (ms)
ODE
MCell
Markov
B
Ca influx per open VDCC
100
80
60
40
20
0
0.0
0.5
1.0
1.5
2.0
2.5
time (ms)
ODE
MCell
Markov


Ca fraction
time (ms)
MCell
Markov
ODE
H0M0 fraction
time (ms)
MCell
Markov
ODE
H2M2 fraction
time (ms)
MCell
Markov
ODE

VDCC	Runtime	Total Floating Point Operations (MFLOPs)
MCell	139.20 ± 7.00 s	--
Standard Markov	109.16 ± 0.93 s	70.10
Multinomial Markov	9.40 ± 0.08 s	12.60
ODE (Euler)	10.05 ± 0.13 s	7.73
Calbindin	**Runtime**	**Total Floating Point Operations (MFLOPs)**
MCell	29.34 ± 0.39 s	--
Standard Markov	4.10E+4 ± 0.00 s	46.10
Multinomial Markov	2.37 ± 0.02 s	27.20
ODE (Euler)	0.72 ± 0.01 s	1.60

Complexity	Time Complexity	Space Complexity
MCell	$O(b\ n\ t)$	$O(b\ n \log_2 s)$
Standard Markov	$O(b\ n\ t)$	$O(b\ n \log_2 s)$
Multinomial Markov	$O(b\ s\ t)$	$O(b\ s \log_2 n)$
ODE (Euler)	$O(b\ s\ t)$	$O(b\ s \log_2 n)$

n (total number of molecules), s (total number of states), b (number of branching states), t (number of timepoints)