Multiscale Modeling

ML for Model Reduction

ML for Multiscale Modeling

Applications in Neuroscience & Software Dev.

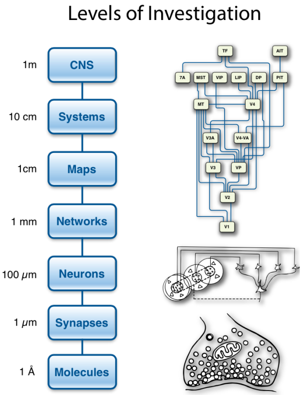
Machine Learning for Generalized Multiscale Modeling

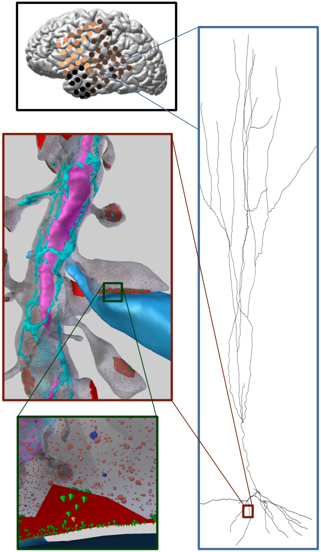
Oliver K. Ernst

CNL, Salk Institute
& UC San Diego Physics

October 3, 2017

Levels of Investigation in Neuroscience





Multiscale Modeling of Rxn-Diffusion Systems

Two problems for multiscale modeling:

1. **"How"** - Model reduction of a complex system
2. **"What"** - What accuracy to use to describe the system in practice - may change over time!

Particle Distribution

Particle Number	Uniform	Non-uniform
Low	Gridless SSA (Stochastic Sim. Algorithm)	Particle-Based (MCell)
↕	Gridless SSA	Gridded SSA
↕	Stochastic ODEs	Stochastic PDEs
Infinite	ODEs (Mass action)	PDEs (Finite elements)

Chemical master equation (CME):
$\dot{p} = Wp$
describes all dynamical regimes

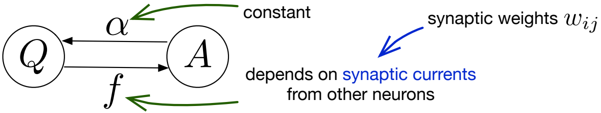
Neural Master Equation

Goal: analysis rather than simulation

Model	References	Number of Parameters	Sampling Possible?	Fit for Large N?	Direct Estimates of Pattern Probabilities?	Low-Dimensional Model of Entire Distribution?
Pairwise maximum entropy	Schneidman et al. (2006); Shlens et al. (2006)	$\sim N^2$	Yes	Difficult	Difficult	No
K-pairwise maximum entropy	Tkacik et al. (2013, 2014)	$\sim N^2$	Yes	Difficult	Difficult	No
Spatiotemporal maximum entropy	Marre et al. (2009); Nasser et al. (2013)	$\sim RN^2$	Yes	Difficult	Difficult	No
Semi-restricted Boltzmann Machine	Köster et al. (2014)	$\sim N^2$	Yes	Difficult	Difficult	No
Reliable interaction model	Ganmor et al. (2011)	Data dependent	No	Yes	Approximate	No
Generalized linear models	Pillow et al. (2008)	$\sim DN^2$	Yes	Difficult	No	No
Dichotomized gaussian	Amari et al. (2003); Macke et al. (2009)	$\sim N^2$	Yes	Yes	No	No
Cascaded logistic	Park et al. (2013)	$\sim N^2$	Yes	Yes	Yes	No
Population coupling	Okun et al. (2012, 2015)	$3N$	Yes	Yes	No	No
Population tracking	This study	N^2	Yes	Yes	Yes	Yes

Note: For the "Number of parameters" column, N indicates the number neurons considered, $\sim$ indicates "scales with," D indicates the number of coefficients per interaction term, and R indicates the number of timepoints across which temporal correlations are considered.

Two state model:



C. ODonnell, J. T. Goncalves, N. Whiteley, C. Portera-Cailliau, and T. J. Sejnowski. The population tracking model: A simple, scalable statistical model for neural population data. Neural Computation 2016.

Why is solving the CME hard?

Two key challenges:

1. **System state space size explosion:**
   - # states in CME grows exponentially with # variables describing system
   - e.g. Gillespie SSA - but requires sufficient trajectories!
2. **Moment closure:**
   - Diff. eqs. for moments depend on higher order moments (analogous: BBGKY hierarchy)

Review: Johnson, T. et al. Model reduction for stochastic CaMKII reaction kinetics in synapses by graph-constrained correlation dynamics. *Phys. Biol.* **12** (2015) 045005

Short outline of the rest of the talk

Goal:

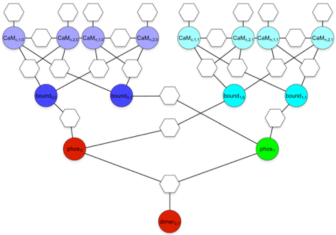
Develop ML approaches to address these problems!

Short outline of the rest of the talk:

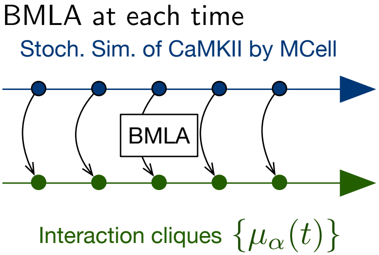
1. ML for model reduction
   - How: Moment closure + State space size explosion
2. Multiscale modeling extensions
   - What accuracy to use to describe the system in practice?
3. The future:
   - Applications to synaptic biochemistry?
   - Interfaces with library versions of MCell?
   - MCell/NEURON?

ML for Model Reduction: Graph Constrained Correlation Dynamics (GCCD)

$$\tilde{p}(s, t; \{\mu\}) = \frac{1}{\mathcal{Z}(\mu(t))} \exp\left[- \sum_{\alpha} \mu_{\alpha}(t) V_{\alpha}(s) \right] \qquad (1)$$

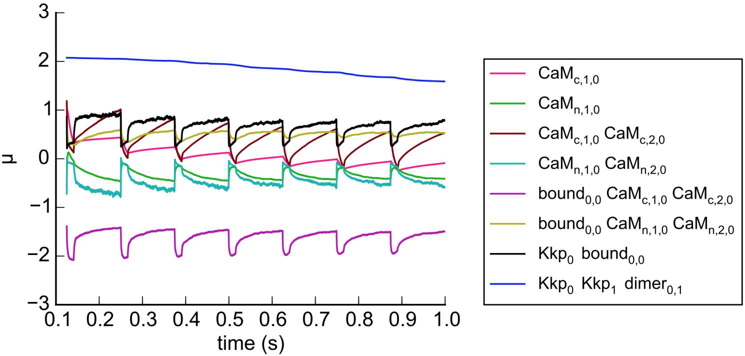


Example Markov Random Field (MRF) for CaMKII



Johnson, T. et al. Model reduction for stochastic CaMKII reaction kinetics in synapses by graph-constrained correlation dynamics. *Phys. Biol.* **12** (2015) 045005

Graph Constrained Correlation Dynamics (GCCD)



Reduced parameters μ obey some dynamical model

1. What are the time evolution operators? ML them!
2. Spatial dynamics

Johnson, T. et al. Model reduction for stochastic CaMKII reaction kinetics in synapses by graph-constrained correlation dynamics. *Phys. Biol.* **12** (2015) 045005

Define space of reduced dynamics

True (fine scale)	$p(n, \boldsymbol{\alpha}, \boldsymbol{x}, t) = n$ particles at $\boldsymbol{x}$ of species $\boldsymbol{\alpha}$
Reduced model	$\tilde{p}(n, \boldsymbol{\alpha}, \boldsymbol{x}, t) = \frac{1}{\mathcal{Z}(\{\nu\})} \exp\left[-\sum_{k=1}^{K} \sum_{\langle i \rangle} \nu_k(\boldsymbol{x}_{\langle i \rangle}, \boldsymbol{\alpha}_{\langle i \rangle}, t) \right]$

- $\nu_k(\boldsymbol{x}_{\langle i \rangle}, \boldsymbol{\alpha}_{\langle i \rangle}, t) =$ k-**particle interaction functions**
  - $\nu_1(\{\alpha\}, \{x\}, t) =$ single particle (self) interaction
  - $\nu_2(\{\alpha, \beta\}, \{x, y\}, t) =$ two particle interaction
  - ... **up to order** K
- Mean Field Theory - style, but slowly time-varying

Simplify: reduced dynamics for well-mixed systems

- Absence of space / well-mixed case:

True (fine scale)	$p(n,t)$ ← CME $\dot{p} = Wp$
Reduced model	$\tilde{p}(n,t) = \frac{1}{\mathcal{Z}(\{\nu\})} \exp\left[-\sum_{k=1}^{K} \binom{n}{k} \nu_k(t) \right]$ ← ?

What are the time evolution operators of reduced model?

Define by **basis functions** $\{F_k\}$:

$$\boxed{\frac{d}{dt}\nu_k(t) = F_k\left(\{\nu_l\}_{l=1}^{K}\right)} \tag{2}$$

Variational problem for the time evolution operators

MaxEnt problem:

Minimize KL divergence between p and $\tilde{p}$

$$\text{Action: } S = \int_0^\infty dt \sum_{n=0}^{\infty} p \ln \frac{p}{\tilde{p}} \tag{3}$$

$$\text{Solve: } \frac{\delta S}{\delta F_k(\{\nu\})} = 0 \tag{4}$$

Solution:

$$\text{Gives } \{F_k\} \rightarrow \frac{d}{dt}\nu_k(t) = F_k\left(\{\nu_l\}_{l=1}^{K}\right) \rightarrow \{\nu_k\}$$

$$\tilde{p}(n,t) = \frac{1}{\mathcal{Z}(\{\nu\})} \exp\left[-\sum_{k=1}^{K} \binom{n}{k} \nu_k(t)\right]$$

which is at all times MaxEnt to true CME p

Boltzmann machine (BM) - like learning algorithm

Initialize:

Fully visible MRF with interaction cliques up to order K

Initial guess $F_k(\{\nu\})$

While not converged:

1. Solve for reduced dynamics $\frac{d}{dt}\nu_k(t) = F_k(\{\nu\})$ from random IC
2. **Boltzmann machine step:**

   Awake phase: Evaluate true moments $\left\langle \binom{n}{k'} \right\rangle_p$ by stoch. sim.

   Asleep phase: Evaluate $\left\langle \binom{n}{k'} \right\rangle_{\tilde{p}}$ by Gibbs sampling
3. **Evaluate the objective function:**

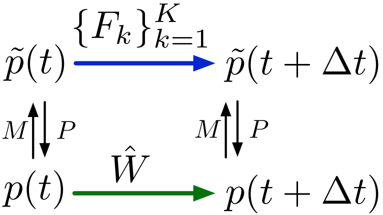
$$\frac{\delta S}{\delta F_k(\{\nu\})} = \sum_{k'=1}^{K} \int_0^\infty dt \left(\left\langle \binom{n}{k'} \right\rangle_p (t) - \left\langle \binom{n}{k'} \right\rangle_{\tilde{p}} (t) \right) \frac{\delta \nu_{k'}(t)}{\delta F_k(\{\nu\})} \tag{5}$$

$$\frac{d}{dt}\left(\frac{\delta \nu_{k'}(t)}{\delta F_k(\{\nu\})} \right) = \delta(t)\delta_{kk'}\delta(\{\nu\} - \{\nu_0\}) + \sum_{l=1}^{K} F_l(\{\nu_0\}) \frac{\partial}{\partial \nu_{l0}} \left(\frac{\delta \nu_{k'}(t)}{\delta F_k(\{\nu\})} \right) \tag{6}$$

4. **Gradient descent:**

$$F_k(\{\nu\}) \to F_k(\{\nu\}) - \lambda \frac{\delta S}{\delta F_k(\{\nu\})} \tag{7}$$

Multiscale modeling!



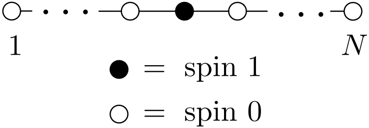
- How do you know which interactions to include?
- Which spatial correlations are important?

Shift the ML problem:

- **Previously:**
  Known: which correlations are important: $\nu_k \rightarrow \langle \binom{n}{k} \rangle$
  ML: what are the basis functions F_k (time evolution)?
- **Now:**
  Known: analytic F_k controlling different correlations
  ML: which are important to include?

The discrete limit & Ising models

- Continuous → discrete



$$\tilde{p}(\{s_i\}) = \frac{1}{Z} \exp\left[h \sum_{i=1}^{N} s_i + J \sum_{i=1}^{N-1} s_i s_{i+1} \right]$$

- MaxEnt consistent with moments
  - $\sum_{i=1}^{N} s_i$ = ave. # of particles
  - $\sum_{i=1}^{N-1} s_i s_{i+1}$ = ave. # of nearest neighbors

Basis functions:

Analytically find:
$$\frac{dh}{dt} = F_h(h, J)$$
$$\frac{dJ}{dt} = F_J(h, J)$$

Analytic basis functions

- Partition function in 1D is analytically accessible
  Solves inverse Ising problem:

$$\text{At large } N : \ln \mathcal{Z} \approx \lambda_+^N \Rightarrow \begin{cases} \left\langle \sum_{i=1}^N s_i \right\rangle (t) = \partial_h \ln \mathcal{Z} \\ \left\langle \sum_{i=1}^{N-1} s_i s_{i+1} \right\rangle (t) = \partial_J \ln \mathcal{Z} \end{cases} \tag{8}$$

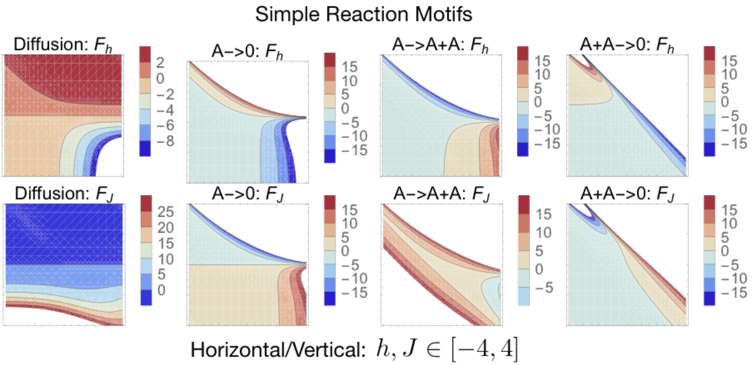
Solution for basis functions:

$$\begin{pmatrix} F_h(h,J) \\ F_J(h,J) \end{pmatrix} = \begin{pmatrix} \partial_h^2 \ln \mathcal{Z} & \partial_h \partial_J \ln \mathcal{Z} \\ \partial_h \partial_J \ln \mathcal{Z} & \partial_J^2 \ln \mathcal{Z} \end{pmatrix}^{-1} \begin{pmatrix} \frac{d}{dt} \left\langle \sum_{i=1}^N s_i \right\rangle \\ \frac{d}{dt} \left\langle \sum_{i=1}^{N-1} s_i s_{i+1} \right\rangle \end{pmatrix} \tag{9}$$

RHS $\frac{d}{dt} \langle \dots \rangle$ given by CME

- Moment closure: $\frac{d}{dt} \langle \dots \rangle$ may not close
  But: all moments can be related to h, J since $\mathcal{Z}$ is analytically accessible!

Analytic solutions for simple systems



1. **General reversible process:** $A \leftrightarrow B + C$.
2. **Substrate Enzyme Product (SEP):** $E + S \leftrightarrow ES \rightarrow E + P$.

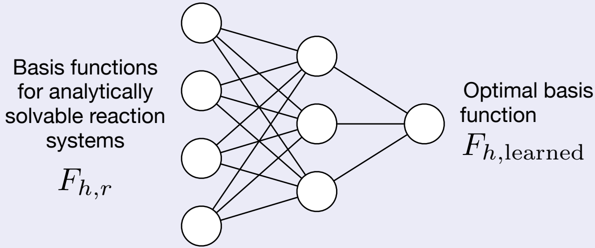
Linearity of Basis Functions → ML for Multiscale

- Linearity in the CME ⇒ linearity in the basis functions
  $\dot{p} = \sum_r W_r p \Rightarrow F_h = \sum_r F_{h,r}$

What accuracy should be used to describe system ?

ML linear combinations of known basis functions

$$F_{h,\text{learned}} = \sum_r \theta_{h,r} F_{h,r} \text{ and similarly for } J \tag{10}$$



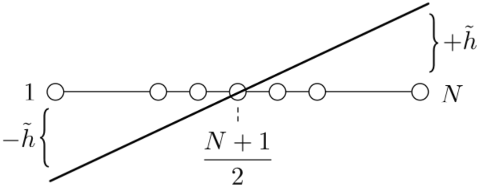
Train from stoch. sim. (e.g. MCell) to learn polynomial

1D → 2D → 3D

- Ising model: homogenous h, J

  Control: $\left\langle \sum_{i=1}^{N} s_i \right\rangle = \#\text{ particles}$ $\quad \left\langle \sum_{i=1}^{N-1} s_i s_{i+1} \right\rangle = \text{NN}$

- **Arbitrary spatial correlations:** Linear perturbations:



Control spatial moments:

$$\left\langle \sum_{i=1}^{N} \frac{i-\eta}{\zeta} s_i \right\rangle$$

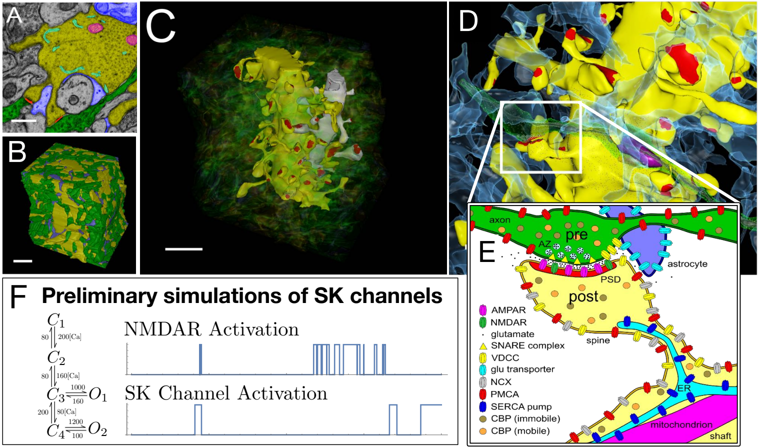
$$\left\langle \sum_{i=1}^{N-1} \frac{i+1/2-\eta}{\zeta} s_i s_{i+1} \right\rangle$$

- $\mathcal{Z}$ is analytically accessible $\Rightarrow$ analytic basis functions

Complex Spatial Correlations in 3D

Approximate using **piecewise linear functions** on mesh
Analogous to basis ("hat" functions) in **Finite Elements**

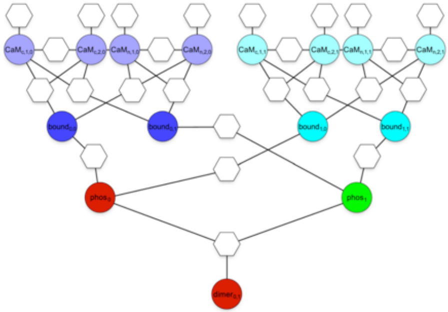
SK channels

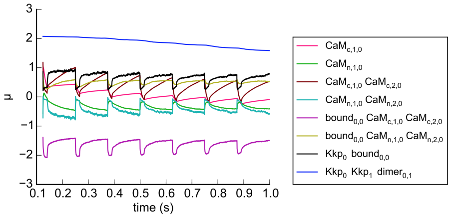


Bartol, T. et al. Computational reconstitution of spine calcium transients from individual proteins. *Frontiers in Synaptic Neuroscience* (2015)

Hirschberg, B. et al, Gating of recombinant small-conductance Ca-activated K channels by calcium. *J. Gen. Physiol* (1998)

Calmodulin/CaMKII complex

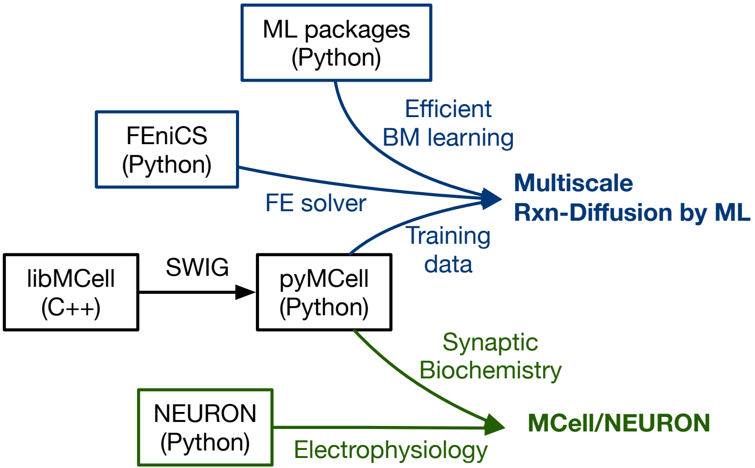




- Reduced models for simulations scaling up to electrophysiology
- PSD-CaMKII interactions?

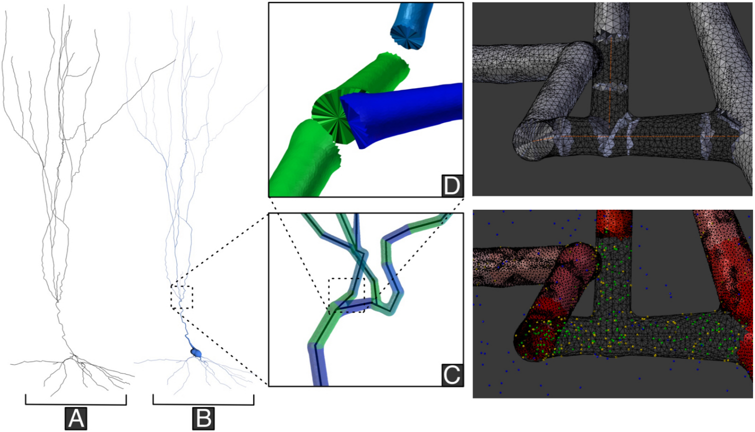
Johnson et al. Model reduction for stochastic CaMKII reaction kinetics in synapses by graph-constrained correlation dynamics. *Physical biology* 2015.

Interface with library versions of MCell: libMCell & pyMCell

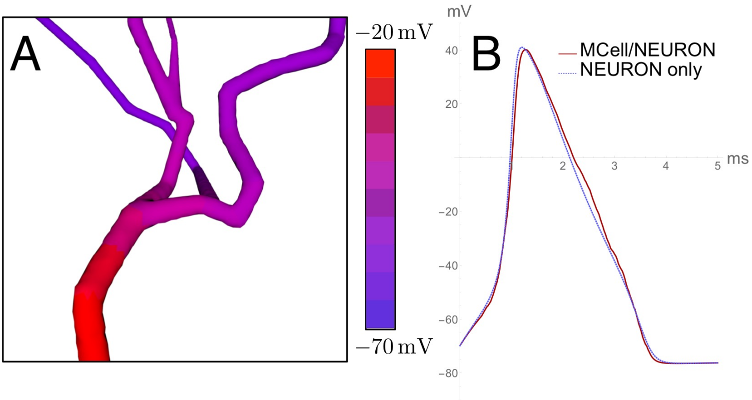


libMCell/pyMCell: Thanks Jacob, Bob, Evan, Tom et al.!

Multi-physics! MCell/NEURON



Multi-physics! MCell/NEURON



C

Closed-loop **"checkpointing"**	Closed-loop **Preliminary Python**	Open-loop **MCell only**
414.2 s = previous implementation without Python	1.6 s = proposed implementation, without parallel processing	0.4 s = maximum speed possible

Review

- **“How to do model reduction”:**
  → Boltzmann machine learning algorithms that *approximate time evolution operators* of a reduced model
- **“What accuracy to use to describe the system in practice”:**
  → ML algorithms to determine optimal scales from stoch. sim. data (e.g. MCell)
- **Future:**
  - 1D → 2D → 3D via Finite Elements
  - Tailor methods to real synaptic biochemistry: SK channels & CaMKII reduced model
  - Software packages interfacing with libMCell/pyMCell

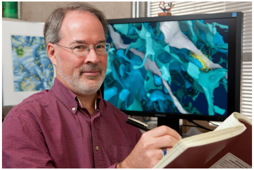
Thanks

Eric Mjolsness

UCIrvine

Terrence Sejnowski

salk
Where cures begin.



Tom Bartol

CNL

UC San Diego

MMBioS

E. Mjolsness. Time-ordered product expansions for computational stochastic system biology. *Physical Biology* 2013.

Johnson et al. Model reduction for stochastic CaMKII reaction kinetics in synapses by graph-constrained correlation dynamics. *Physical biology* 2015.