# Assembling and Testing Physically Based Models of Observed Complex Systems

**Henry D. I. Abarbanel**

**Department of Physics**

**and**

**Marine Physical Laboratory (Scripps Institution of Oceanography)**

**Center for Engineered Natural Intelligence**

**University of California, San Diego**

**habarbanel@ucsd.edu**

**with Daniel Rey, Nirag Kadakia, Daniel Breen, Sasha Shirman, Eve Armstrong, Dan Margoliash, Arij Daou, Gert Cauwenberghs, J. Wang, A. Aknin.**

**This talk is about the general principles of learning about model states and parameters from laboratory or field data. We want to transfer the information in the (noisy) data to a model (with model errors) of the physical or biological dynamical processes producing the measurements.**

**When unknown model parameters and unobserved state variables are estimated using the data, we can validate the model through prediction when new forcing is presented.**

**Using the same methods one can: design experiments; determine how many measurements are required; retrieve more information from existing measurements.**

**Examples will be discussed for**

**biophysical modeling of neurons and functional networks**
**neuromorphic engineering---neurons and functional networks on a chip**

## Some application areas for Data Assimilation:

Genetic regulatory networks

signal transduction pathways

systems biology; synthetic biology; Immunology

biophysical modeling of neurons and functional networks

neutrino astrophysics

coastal flows and transport of toxic constituents after storms

electrical and chemical engineering

identifying oil and gas reservoirs

hydrological models of streams and lakes

neuromorphic engineering---neurons and functional networks on a chip

numerical weather prediction

# Neuron Model

$$C\frac{dV(t)}{dt} = g_{Na}m(t)^3 h(t)(E_{Na} - V(t)) +$$
$$g_K n(t)^4 (E_K - V(t)) + g_L (E_L - V(t)) + \mathrm{I}_{applied}(t)$$

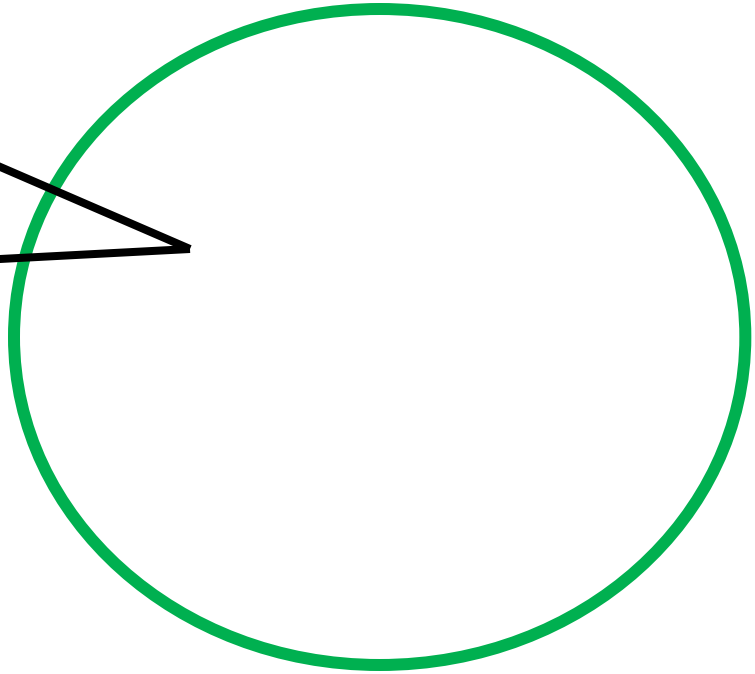
$$\frac{da(t)}{dt} = \frac{a_\infty (V(t) - a(t))}{\tau_x (V(t))} \qquad \mathrm{a}(t) = \{m(t), h(t), n(t)\}$$

$$a_\infty (V) = \frac{1}{2}\left(1 + \tanh\left(\frac{V - V_x}{dV_x}\right)\right)$$

$$\tau_x (V) = t_{1x} + t_{2x}\left(1 - \tanh^2\left(\frac{V - V_{xt}}{dV_{xt}}\right)\right)$$

**Inject current $I_{applied}(t)$**

**Measure Response Voltage V(t)**



**D = 4 L = 1 p = 20**

**Measure V(t) with selected $I_{applied}(t)$**

**Evaluate all parameters and all unobserved state variables a(t)**

# This is the challenge:

Using laboratory experiments on individual neurons and on collections of neurons. **Build biophysically based models of functional neural networks which matches experiments and predicts the response to new stimuli.**

**Our strategy is this**:

o create a model of the functional network of interest—e.g. song production network for song birds and, of course, the individual neurons in the network---what is a sensible model—we use Hodgkin-Huxley models.

o using the model itself, design experiments that stimulate all degrees of freedom of the neuron/network and measure enough quantities at each observation time---these are numerical simulations.

o Use the model along with data, voltage across membranes, and perhaps other measurements, to determine the unknown parameters in the model and the unobserved state variables in the model.

o “validate the neuron model”—via prediction—These validated neurons can can be used in network construction.

**One mainstream view of network modeling and operation is that details do not matter but some form of network "organization" or structure determines network operations.**

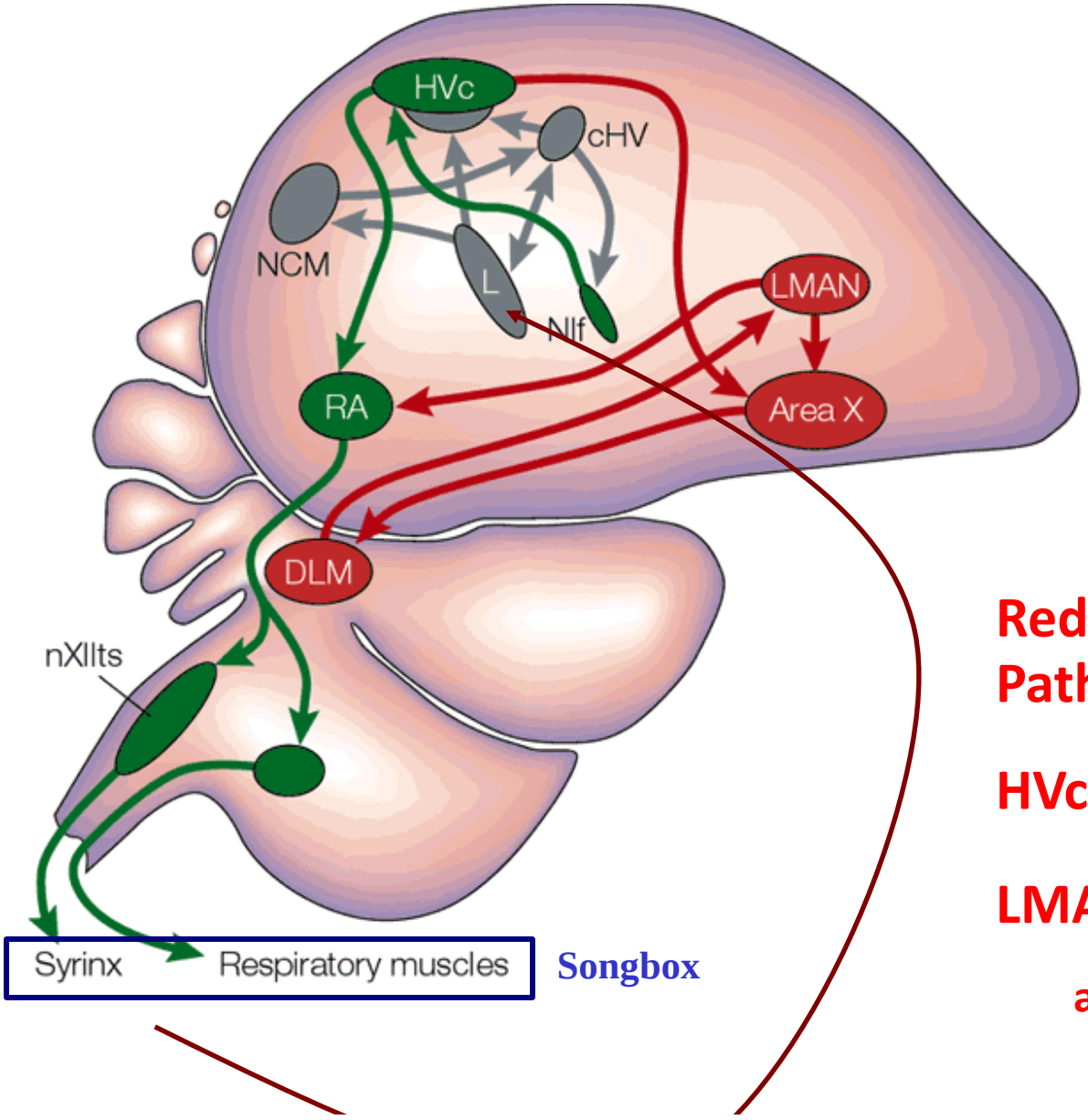
**Our use of data assimilation to design experiments, test and validate models of cells and systems points to advantages of other directions.**

**Why would we want the kind of detail of neural or cellular processes accurate modeling and careful data assimilation provides?**

- **use models of nerve cells (neurons) to compare healthy and diseased cells to provide biophysical targets for therapies**

- **use detailed models of regulatory networks for genetic action to design interventions**

- **use detailed, verified models of functional network connectivity and nodal performance to engineer functions into high-performance electronics---e.g. sequence generation and recognition with human accuracy but machine performance**




**Green: Motor Pathway**

**HVc→RA→**

**Respiration/Syrinx**

***Song Production***

**Red: Anterior Forebrain Pathway (AFP)**

**HVc→Area X→DLM→**

**LMAN→** **Area X**

**and HVC**

***Control and Song***

***Maintenance***

**Auditory Feedback**

**Isolated Neurons from the Avian Song System**

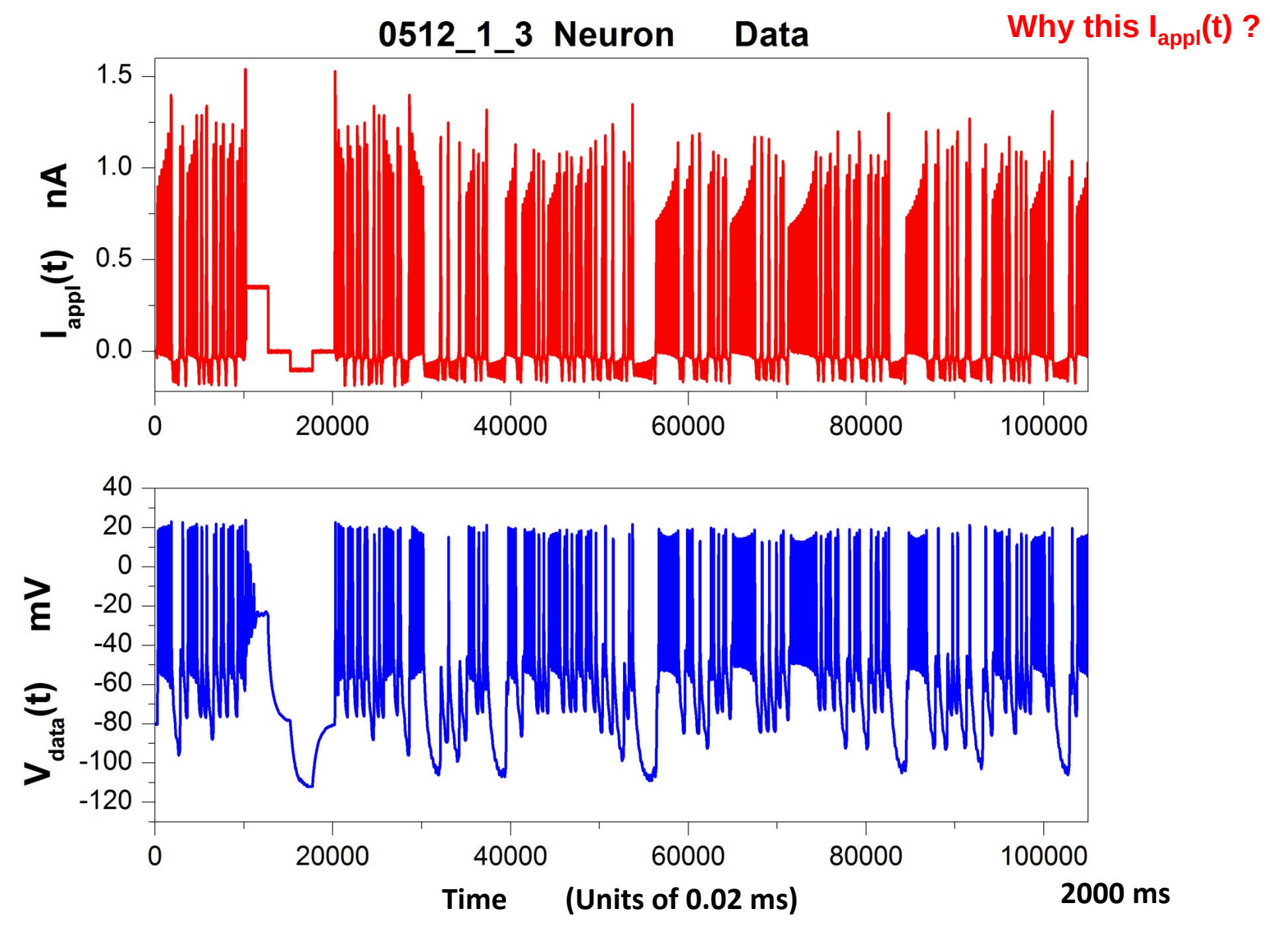
**On each neuron many different $I_{applied}(t)$ measurements in time "epochs" of 2-6 seconds**

**Membrane Voltage Observed**

**Sampling time 0.02 ms (50 kHz), 500-1500 ms of observations**

**Use all this to estimate the unknown parameters in the neuron and unmeasured state variables, then predict the response of the neuron to new stimuli (forcing).** **Model Validation**

**Neurobiological Laboratory Experiments**
**Margoliash Laboratory, UChicago**


0512_1_3 Neuron Data
Why this $I_{appl}(t)$ ?
$I_{appl}(t)$ nA
1.5
1.0
0.5
0.0
0
20000
40000
60000
80000
100000
$V_{data}(t)$ mV
40
20
0
-20
-40
-60
-80
-100
-120
Time (Units of 0.02 ms)
2000 ms

# Data Assimilation:

## Transfer of Information from Measurements to a Model of the Observations

We have

noisy measurements $y_k(t);\ k = 1, \ldots L;$ and

errors in the model $x_a(t+\Delta t)=f_a(x(t));$ a=1,…,D >>L,

uncertain initial conditions at the starting time $t_0$.

We make measurements in a time window $[t_0,T]$. We wish to incorporate the information in these measurements at $t_0, t_1, \ldots, t_N = T$ into our statistical estimate of the complete state of the model at T and into our statistical estimate of the model parameters.

**Statistical data assimilation is communication of information from measurements (transmitter) to a dynamical model (receiver).**

**At the end of an observation window $[t_0,T]$ we want the conditional probability distribution of the state x(T) of the system**

$$P(x(T)|\text{observations})$$

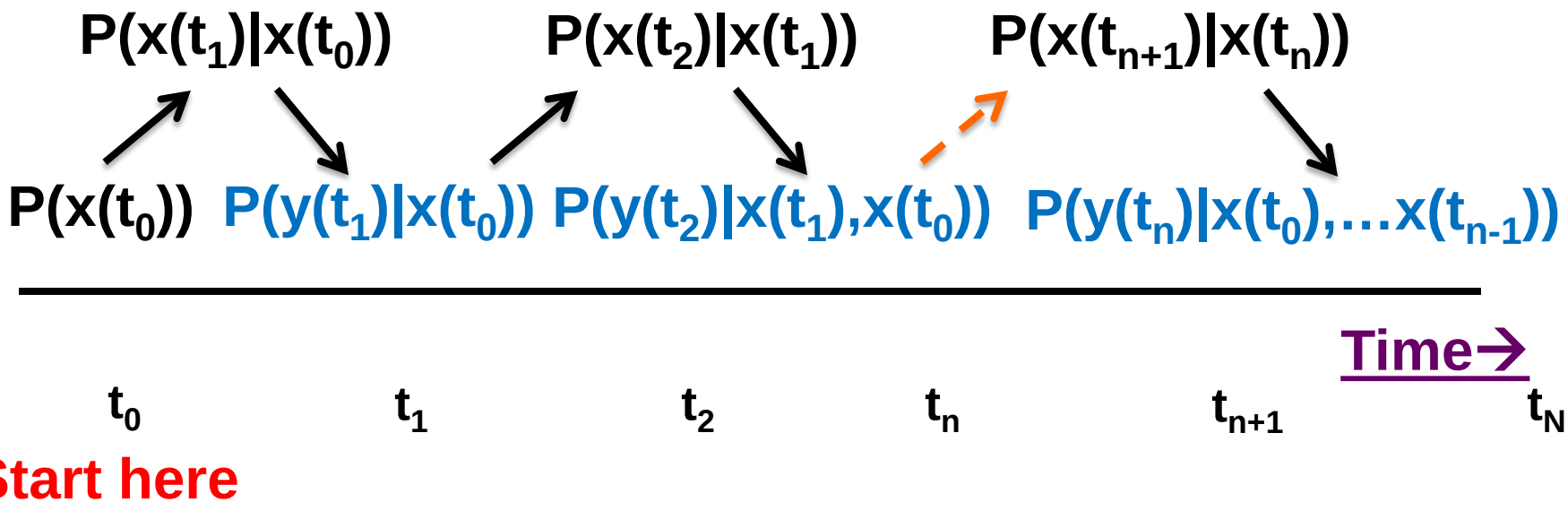
**given measurements during the window.**

**We then want to predict the future conditional probability distribution $P(x(t > T))$ for new forcing of the system.**

1st observation $y(t_1)$

2nd observation $y(t_2)$

$y(t_n)$ $y(t_{n+1})$



# Conditional probability of states X(N) given observations Y(N): *P(X(N)|Y(N))*

$$X(N) = \{x(t_0), x(t_1), \ldots, x(t_N)\} \qquad Y(N) = \{y(t_0), y(t_1), \ldots, y(t_N)\}$$

$$P(X(N)\,|\,Y(N)) = \prod_{k=0}^{N} P(y(k)\,|\,X(k), Y(k-1)) \prod_{k=0}^{N-1} P(x(n+1)\,|\,x(n))\; P(x(0))$$

Information Transfer  Dynamics of System

$$= e^{-A_0(X)}$$

Expected value of functions $G(X)$ on the path X(N),

$$\langle G(X) \rangle = \frac{\int dX\; e^{-A_0(X)} G(X)}{\int dX\; e^{-A_0(X)}}$$

*e.g.* G(X) = X; X - <X>; other moments

$$t = \{t_0, t_1, \ldots t_n, \ldots, t_N = T\}$$

$$y_l(t_n)$$

$$l = 1, 2, \ldots, L$$

**Data source:** **Transmitter**

$$L \le D$$

**Model:** **Receiver**

$$x_a(t_{n+1}) =$$

$$f_a(x(t_n));$$

$$a = 1, 2, \ldots, D$$

$$y_l(n) = h_l(x(n))$$

**Generalized synchronization of the transmitter and receiver**

In neurobiology: We have measurements of <u>voltage</u> across the cell membrane during the observation window

$Y=\{y(t_0) = V(t_0), y(t_1) = V(t_1)\ldots, y(T) = V(T)\}$ (curr

L = 1, D = 12, p = 50 Maximal conductances $g_{Na}$, $g_{SK}$, ….

We have a model (current conservation)

$CdV(t)/dt = I_{Na}(t)+I_K(t) +I_{Ca}(t)+I_{SK}(t)+\ldots.+I_{stimulus}(t)$

$$\frac{dx(t)}{dt} = F(x(t))$$

Measurements have errors, model has errors

$$E[G(X)|Y] = \langle G(X) \rangle = \frac{\int dX e^{-A_0(X)} G(X)}{\int dX e^{-A_0(X)}}$$

Path is through state variable+parameter space

X={x(N),x(N-1),...x(0)} including parameters

**Everything hinges on the structure of $A_0(X)$ in path space**

Two methods for evaluating these high dimensional integrals:

(1) Laplace's method (1774); seek minima of $A_0(X)$--there are multiple minima

$$\frac{\partial A_0(X)}{\partial X_\alpha} = 0 \quad \text{and} \quad \frac{\partial^2 A_0(X)}{\partial X_\alpha \partial X_\beta} > 0$$

(2) Monte Carlo searches of $\exp[-A_0(X)]$—distribution in path space.

First method seeks minima of $A_0(X)$

Second method samples near these minima.

We focus on the Laplace method. How can we find the minima $X^q$ q=0,1,….

$$\left.\frac{\partial A_0(X)}{\partial X_\alpha}\right|_{X^q} = 0 \quad q=0,1,... \quad A_0(X^0) < A(X^{q\neq 0})$$

Which minimum gives the biggest contribution to the integral ?
Smallest action $A_0(X^0)$ from path $X^0$

**Everything rests on the structure of $A_0(X)$ in path space.**

## Action for State and Parameter Estimation

Information Transfer

$$A_0(X) = -\sum_{n=0}^{N} \log\{CMI(y(n) \mid X(n), Y(n-1))\}$$

$$-\sum_{n=0}^{N-1} \log\{P(x(n+1) \mid x(n))\} - \log\{P(x(0))\}$$

Dynamics

Initial Condition

$A_0(X)$ is nonlinear in X and has multiple minima. Location and number of these minima depend on the model and on the number of measurements at each observation time in $[t_0,t_N]$.

## Standard model, Gaussian Error Action

If observations have Gaussian noise and models have Gaussian errors, action is

$$A_{\mathrm{GEA}}(X) = \sum_{n=0}^{N}\sum_{l=1}^{L}\frac{R_m(n,l)}{2}(x_l(n)-y_l(n))^2 +$$

$$\sum_{n=0}^{N-1}\sum_{a=1}^{D}\frac{R_f(a)}{2}(x_a(n+1)-f_a(x(n)))^2$$

This is not Gaussian, if *f(x)* is nonlinear. Finding the minima at any $R_m$ and $R_f$ is not hard (IPOPT), but finding the path with the smallest action—a challenge

Looking in continuous time shows the challenge clearly:

$$A_{GEA}(x(t),\, \dot{x}(t)) = \int_{t_0}^{t_f} dt\; L(x(t),\, \dot{x}(t))$$

$$= \int_{t_0}^{t_f} dt \left[ \sum_{l=1}^{L} \frac{R_m(t)}{2} (x_l(t) - y_l(t))^2 + \sum_{a=1}^{D} \frac{R_f(a)}{2} (\dot{x}_a(t) - F_a(x(t)))^2 \right]$$

$\delta A_0(X) = 0$ gives the Euler-Lagrange equations:

$$\left( \delta_{ab} \frac{d}{dt} + DF_{ab}(x(t)) \right) \left( \frac{dx_b(t)}{dt} - F_b(x(t)) \right) = \delta_{al} \frac{R_m}{R_f} (x_l(t) - y_l(t))$$

**'nudging'**

With the boundary conditions
$p_a(t_0) = p_a(t_f) = 0$.

$$p_a(t) = \frac{\partial L(x(t), \dot{x}(t))}{\partial \dot{x}_a(t)}$$

$$A_{\mathrm{GEA}}(X)=\sum_{n=0}^{N}\sum_{l=1}^{L}\frac{R_m(n,l)}{2}(x_l(n)-y_l(n))^2+\sum_{n=0}^{N-1}\sum_{a=1}^{D}\frac{R_f(a)}{2}(x_a(n+1)-f_a(x(n)))^2$$

To determine the paths for the lowest minimum action; find minimum for very small model error value $R_f$, then move slowly in $R_f$ to a larger value, then an even larger value. We call this **annealing**; distinct from standard simulated annealing.

If $R_f \rightarrow \infty$, model error is 0. We look at the opposite limit $R_f \rightarrow 0$, where model plays no role and dynamical phase space structure is absent.

At $R_f = 0$, minimum is degenerate at $x_l(t) = y_l(t)$; other unmeasured states undetermined. With $R_f = R_{f0}$ very small, choose $N_0$ initial starting paths with $x_l(t) = y_l(t)$, others chosen from a uniform distribution, this is a set of $X_0$ for numerical minimization using IPOPT. We call outcomes $X_1$.

Use these as initial starting paths with $R_f = \alpha R_{f0}$ ; $\alpha>1$, to arrive at $N_0$ paths $X_2$. Increase $R_f$ to $\alpha^2 R_{f0}$, … and continue using outcome paths as initial guesses for next optimizations, slowly increasing $R_f$ by powers of $\alpha$.

Plot $A_0(X^q)$ versus $\beta=\log_\alpha\left[R_f/R_{f0}\right]$. Action level plots.

## Simple Model Neuron NaKL

## $D+N_P = 4 + 19$, $L = 1$ Voltage

## Twin Experiment on NaKL Neuron

$$C\frac{dV(t)}{dt} = g_{Na}m(t)^3h(t)(E_{Na} - V(t)) +$$

$$g_K n(t)^4(E_K - V(t)) + g_L(E_L - V(t)) + \mathrm{I}_{applied}(t)$$

$$\frac{dx(t)}{dt} = \frac{x_\infty(V(t) - x(t))}{\tau_x(V(t))} \qquad x(t) = \{m(t), h(t), n(t)\}$$

$$x_\infty(V) = \frac{1}{2}\left(1 + \tanh\left(\frac{V - V_x}{dV_x}\right)\right)$$

$$\tau_x(V) = t_{1x} + t_{2x}\left(1 - \tanh^2\left(\frac{V - V_{xt}}{dV_{xt}}\right)\right)$$

**Generate data from NaKL equations**

**y(t) = x(t) + σN(0,1) noise**

**D = 4, L = 1**

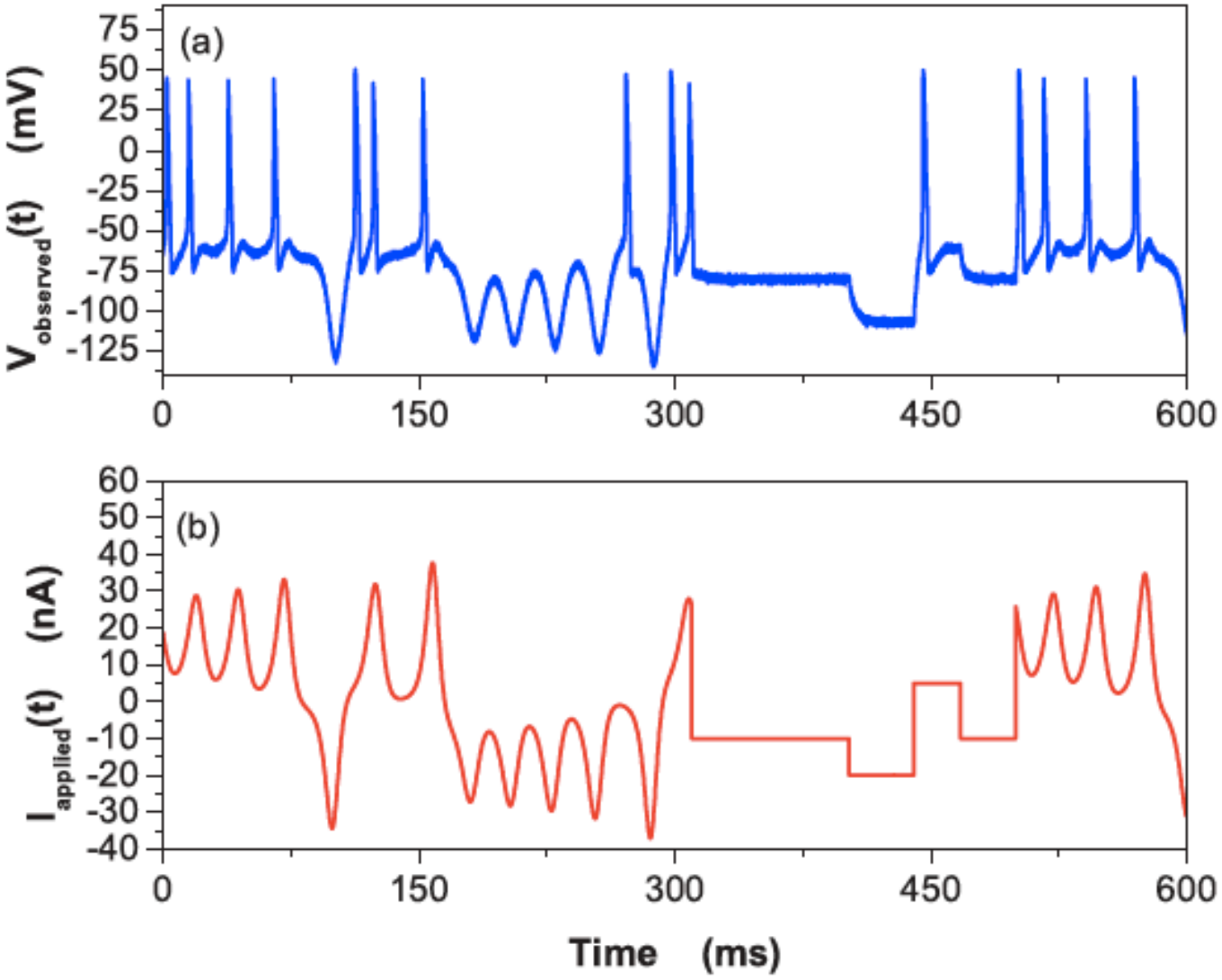
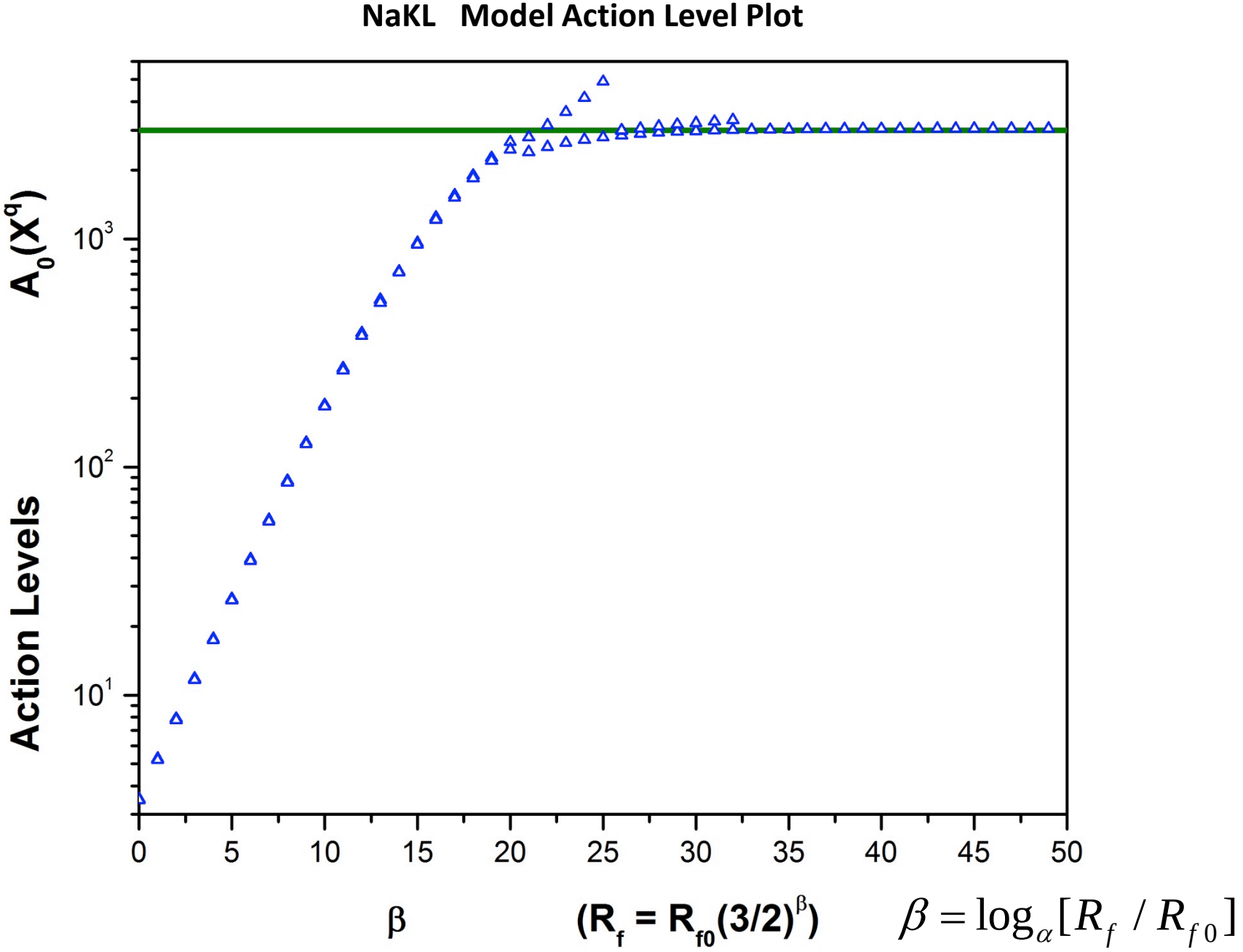


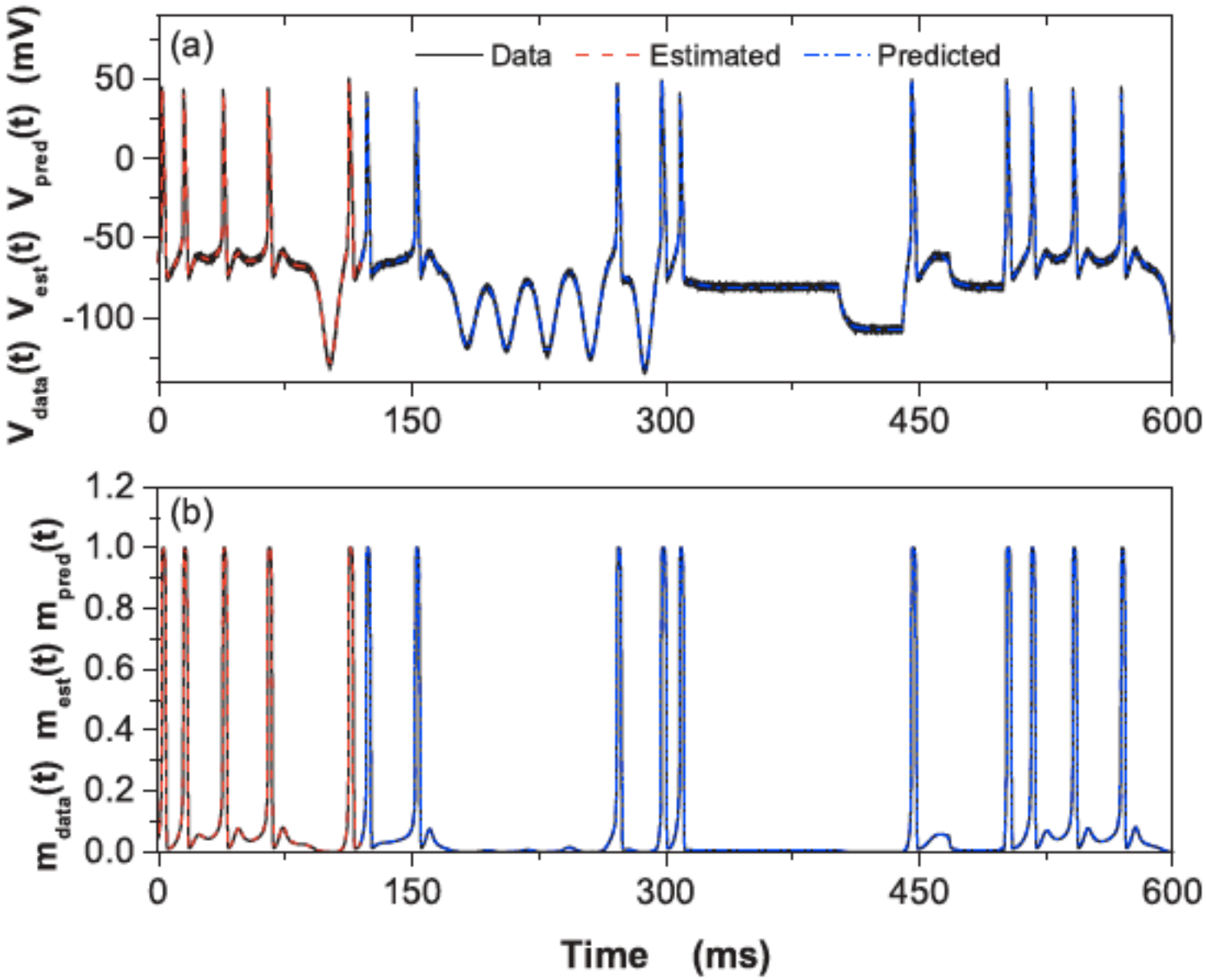

FIG. 8. (Color online) (a) Voltage response of the standard Hodgkin-Huxley neuron model with $Na^+$, $K^+$, and leak channels, our NaKL model, in response to the applied (injected) stimulus current shown in the (b).

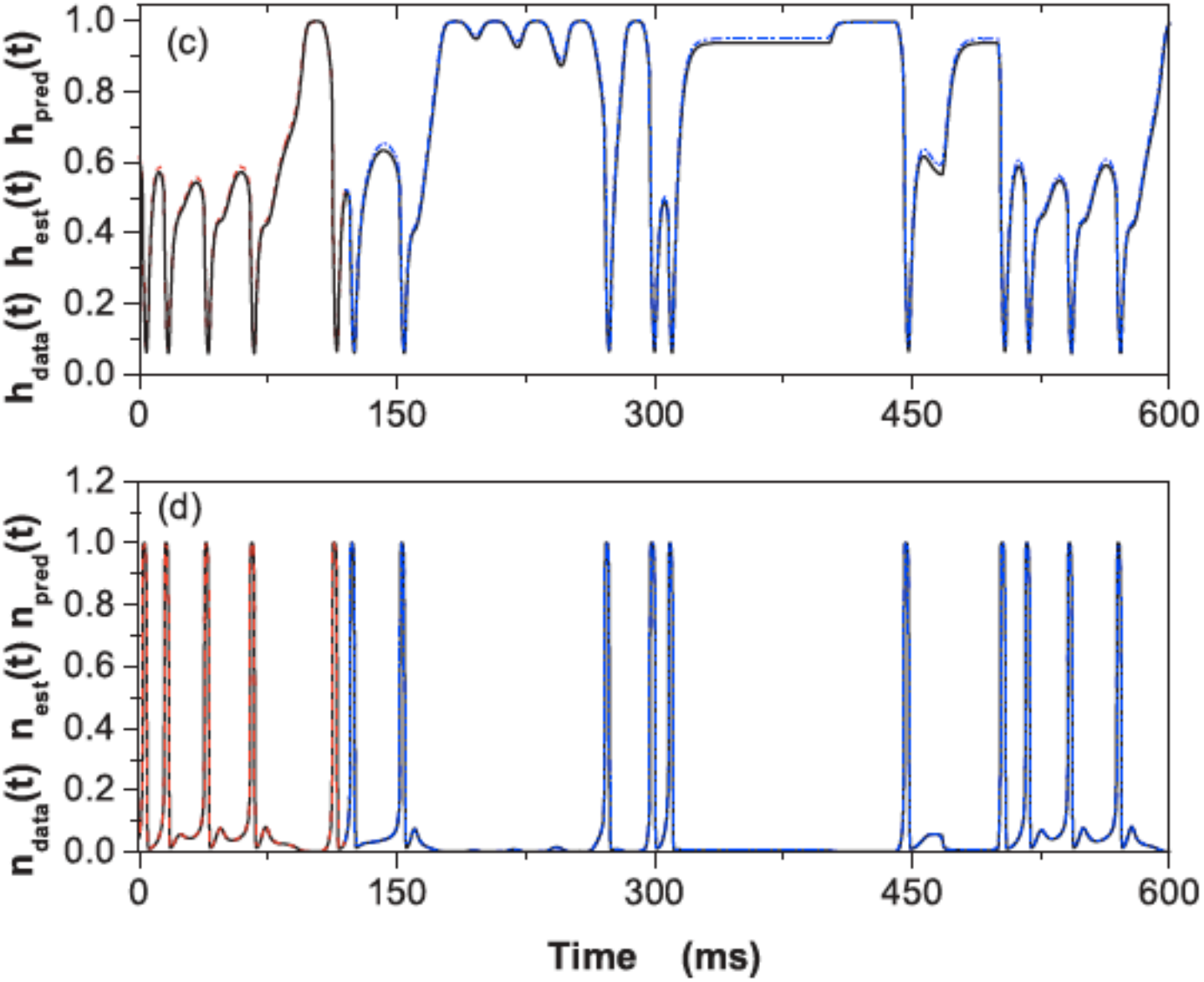
## Annealing in Model Error Accuracy

Paths giving minima of the action depend on the number of measurements L.

For the Standard Model GEA, when action levels are independent of $R_f$, the action level is dictated by statistics of measurement error term---a consistency check on the action level evaluations.


NaKL Model Action Level Plot
$A_0(X^q)$
Action Levels
$10^3$
$10^2$
$10^1$
0
5
10
15
20
25
30
35
40
45
50
β
$(R_f = R_{f0}(3/2)^{\beta})$
$\beta = \log_{\alpha}[R_f / R_{f0}]$


(a)
Data
Estimated
Predicted
$V_{data}(t)$ $V_{est}(t)$ $V_{pred}(t)$ (mV)
50
0
-50
-100
0
150
300
450
600
(b)
$m_{data}(t)$ $m_{est}(t)$ $m_{pred}(t)$
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
150
300
450
600
Time (ms)


(c)
(d)
$h_{data}(t)$ $h_{est}(t)$ $h_{pred}(t)$
$n_{data}(t)$ $n_{est}(t)$ $n_{pred}(t)$
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
150
300
450
600
Time (ms)

## NaKL Neuron Twin Experiment

| Parameters | Known | Estimated | LB | UB |
|---|---|---|---|---|
| $g_{Na}$ | 120.0 | 108.4 | 50.0 | 200.0 |
| $E_{Na}$ | 50.0 | 49.98 | 0.0 | 100.0 |
| $g_K$ | 20.0 | 21.11 | 5.0 | 40.0 |
| $E_K$ | −77.0 | −77.09 | −100.0 | −50.0 |
| $g_L$ | 0.3 | 0.3028 | 0.1 | 1.0 |
| $E_L$ | −54.0 | −54.05 | −60.0 | −50.0 |
| C | 0.8 | 0.81 | 0.5 | 1.5 |
| $V_m$ | −40.0 | −40.24 | −60.0 | −30.0 |
| $dV_m$ | 0.0667 | 0.0669 | 0.01 | 0.1 |
| $\tau_{m0}$ | 0.1 | 0.0949 | 0.05 | 0.25 |
| $\tau_{m1}$ | 0.4 | 0.4120 | 0.1 | 1.0 |
| $V_h$ | −60.0 | −59.43 | −70.0 | −40.0 |
| $dV_h$ | −0.0667 | −0.0702 | −0.1 | −0.01 |
| $\tau_{h0}$ | 1.0 | 1.0321 | 0.1 | 5.0 |
| $\tau_{h1}$ | 7.0 | 7.76 | 1.0 | 15.0 |
| $V_n$ | −55.0 | −54.52 | −70.0 | −40.0 |
| $dV_n$ | 0.0333 | 0.0328 | 0.01 | 0.1 |
| $\tau_{n0}$ | 1.0 | 1.06 | 0.1 | 5.0 |
| $\tau_{n1}$ | 5.0 | 4.97 | 2.0 | 12.0 |

## Lorenz96 Model D = 5

$$\frac{dx_a(t)}{dt} = x_{a-1}(t)(x_{a+1}(t) - x_{a-2}(t)) - x_a(t) - f$$

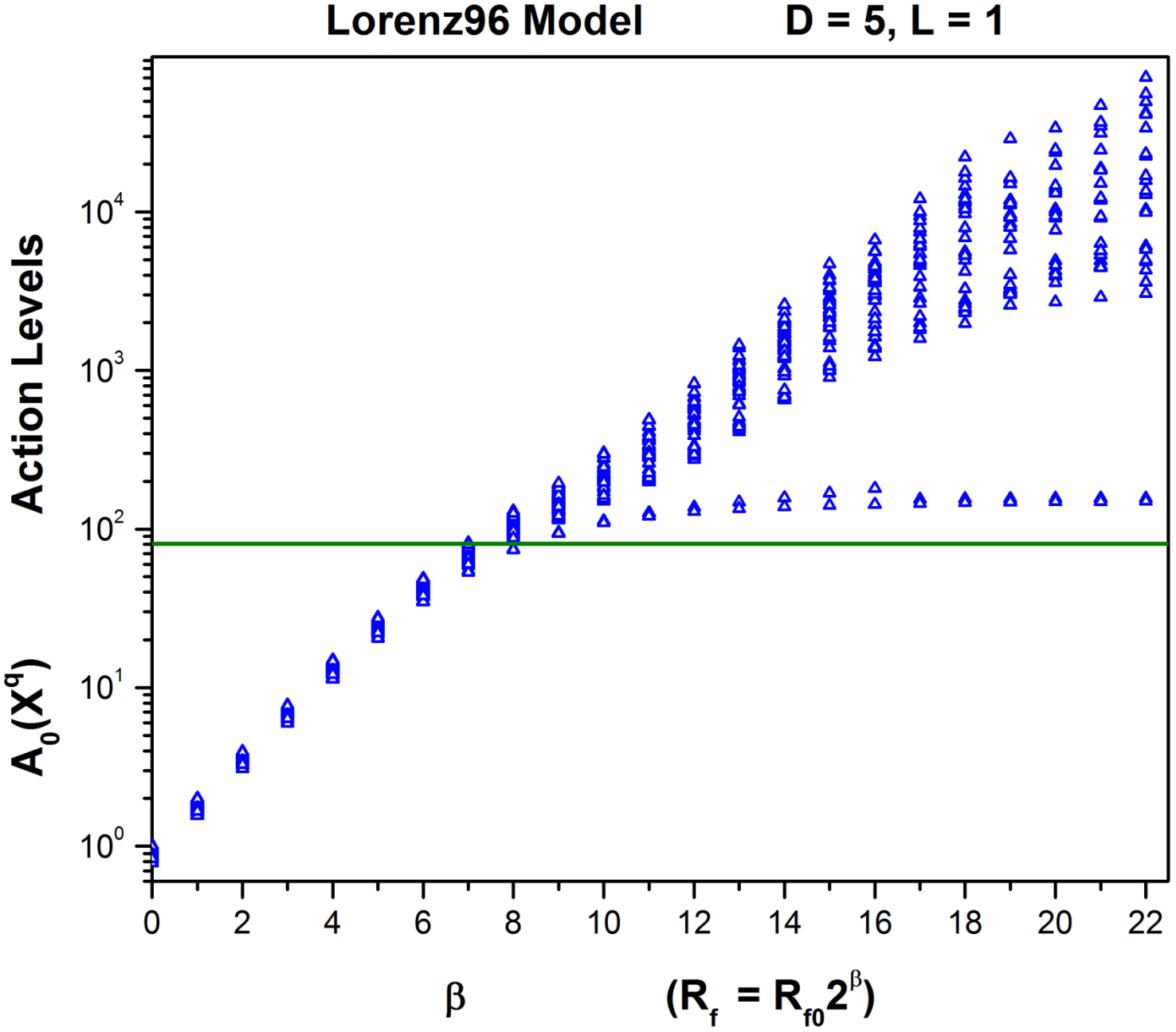
$$a = 1, 2, ..., D;$$

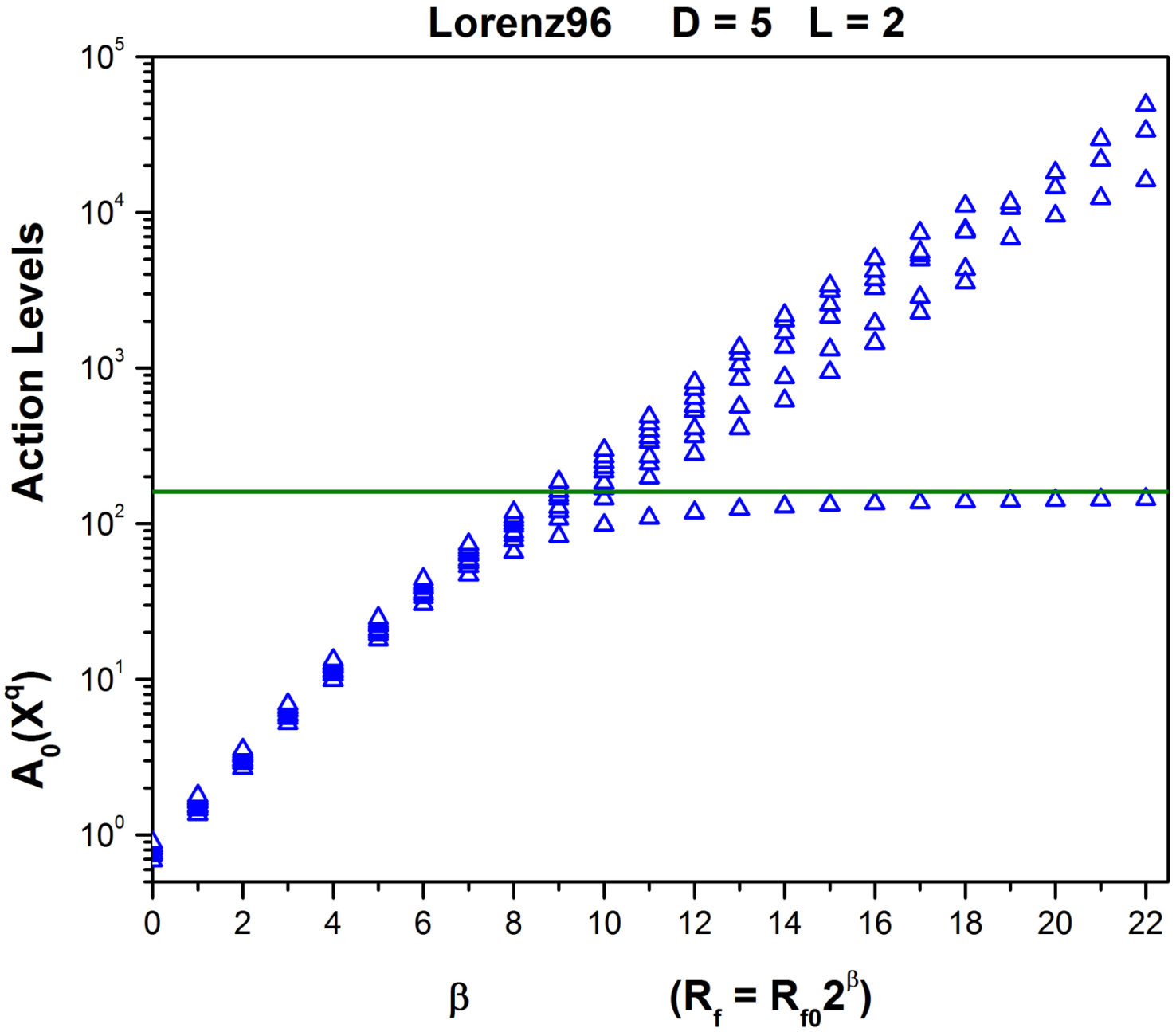
$$x_{-1}(t) = x_{D-1}(t);\ x_0(t) = x_D(t);\ x_{D+1}(t) = x_1(t).$$

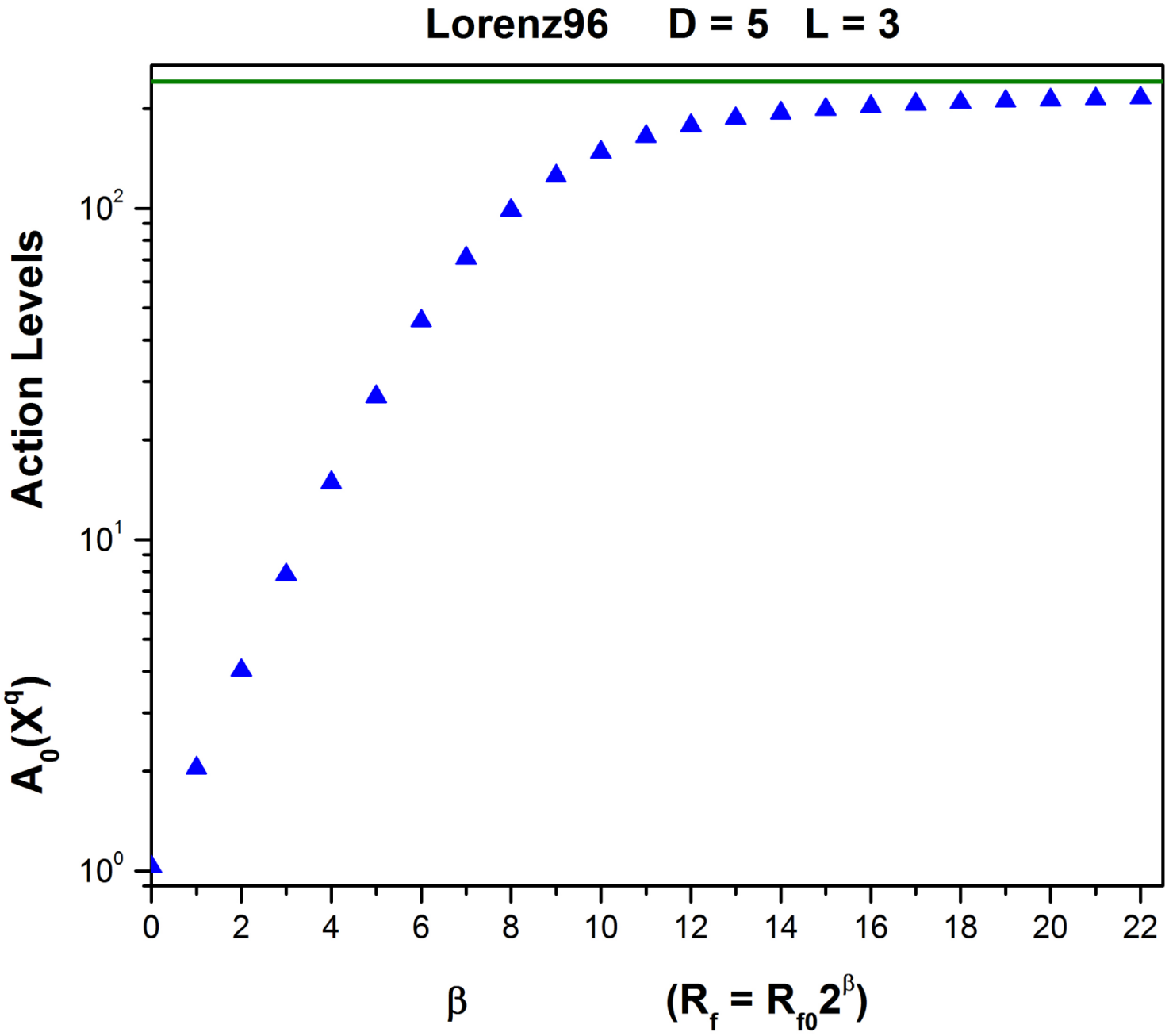
f is a fixed parameter f = 8.17. Solutions are chaotic.

**'Twin Experiments' Use to test methods of data assimilation; use to design experiments.**

**Generate data with known model; add noise to model output; present l = 1,2,…,L < D time series to assimilation procedure.**


Lorenz96 Model
D = 5, L = 1
Action Levels
$A_0(X^q)$
$10^0$
$10^1$
$10^2$
$10^3$
$10^4$
0
2
4
6
8
10
12
14
16
18
20
22
β
$(R_f = R_{f0}2^{\beta})$


Lorenz96 D = 5 L = 2
Action Levels
$A_0(X^q)$
$10^5$
$10^4$
$10^3$
$10^2$
$10^1$
$10^0$
0
2
4
6
8
10
12
14
16
18
20
22
β
$(R_f = R_{f0}2^{\beta})$


Lorenz96 D = 5 L = 3
Action Levels
$A_0(X^q)$
$10^2$
$10^1$
$10^0$
0
2
4
6
8
10
12
14
16
18
20
22
β
$(R_f = R_{f0}2^{\beta})$

## Back to Song System Nucleus HVC

## Interneurons, L = 1  Voltage

$$C\frac{dV(t)}{dt} = g_{Na}m(t)^3 h(t)(E_{Na} - V(t)) +$$

$$g_K n(t)^4 (E_K - V(t)) + g_L (E_L - V(t)) +$$

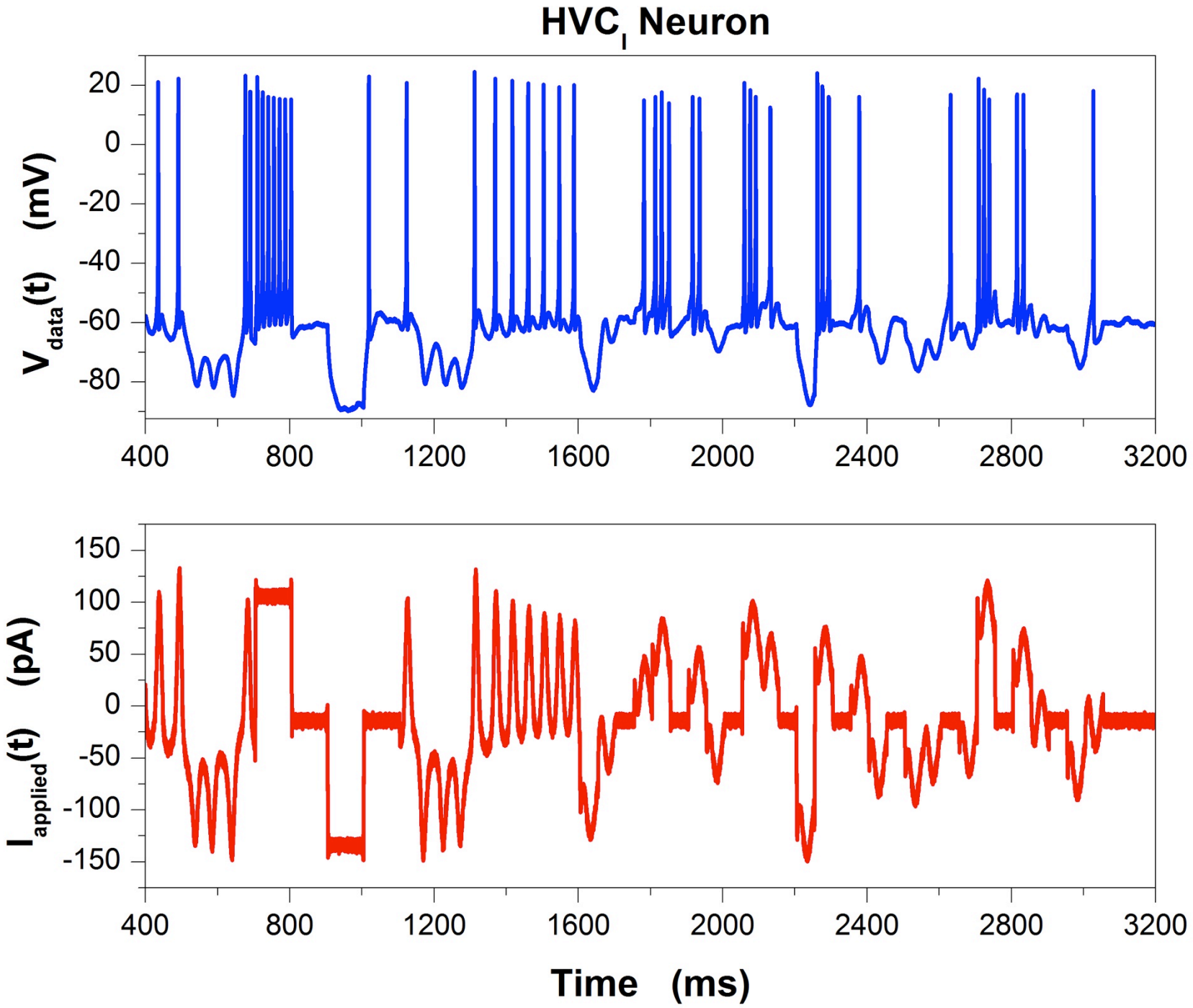
$$g_{Ca} a(t) b(t) V(t) \frac{[Ca^{2+}]_{ext} - [Ca^{2+}](t) e^{-V(t)/V_T}}{1 - e^{-V(t)/V_T}}$$

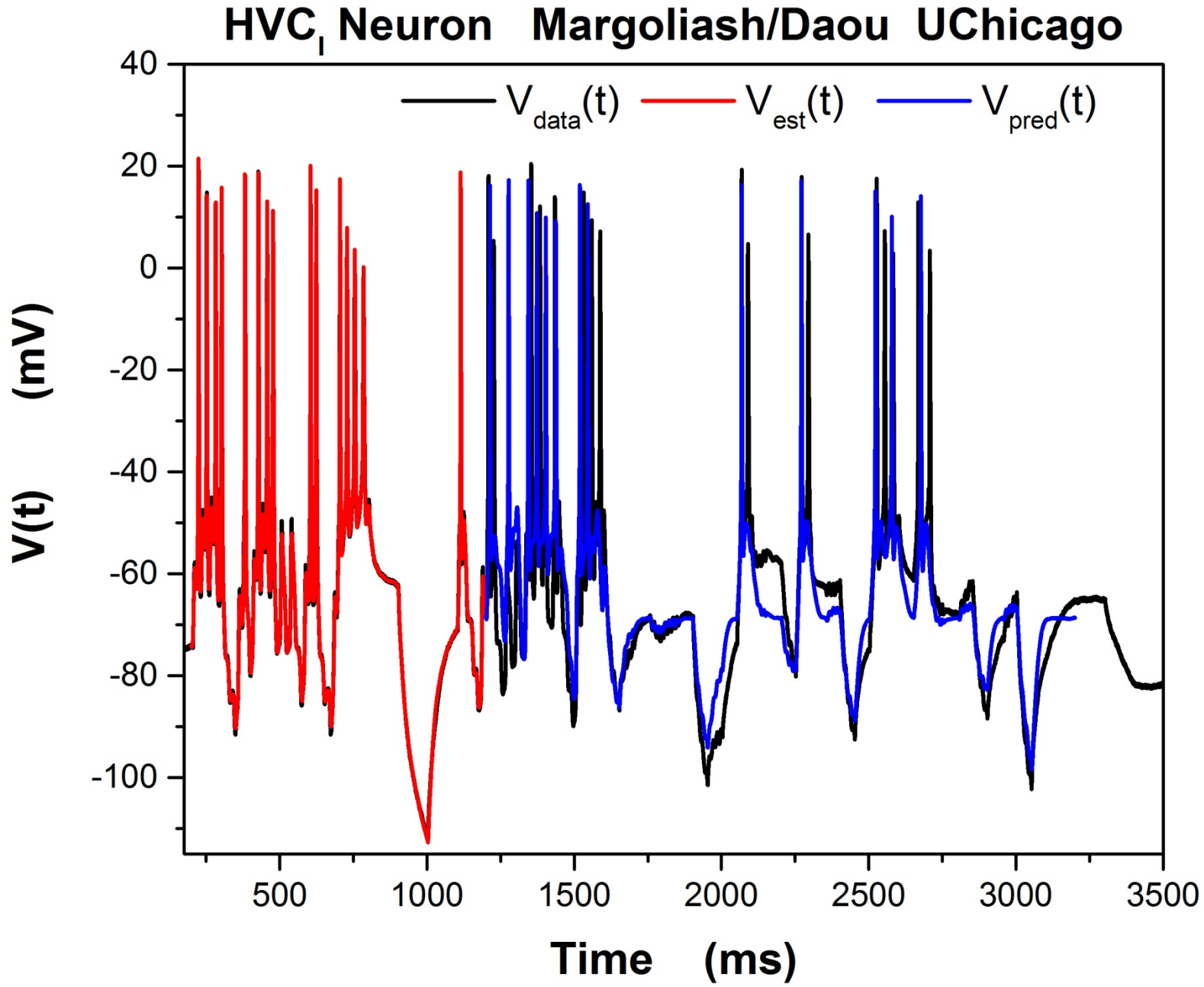
$$+ \text{ other currents} + \mathrm{I}_{applied}(t)$$

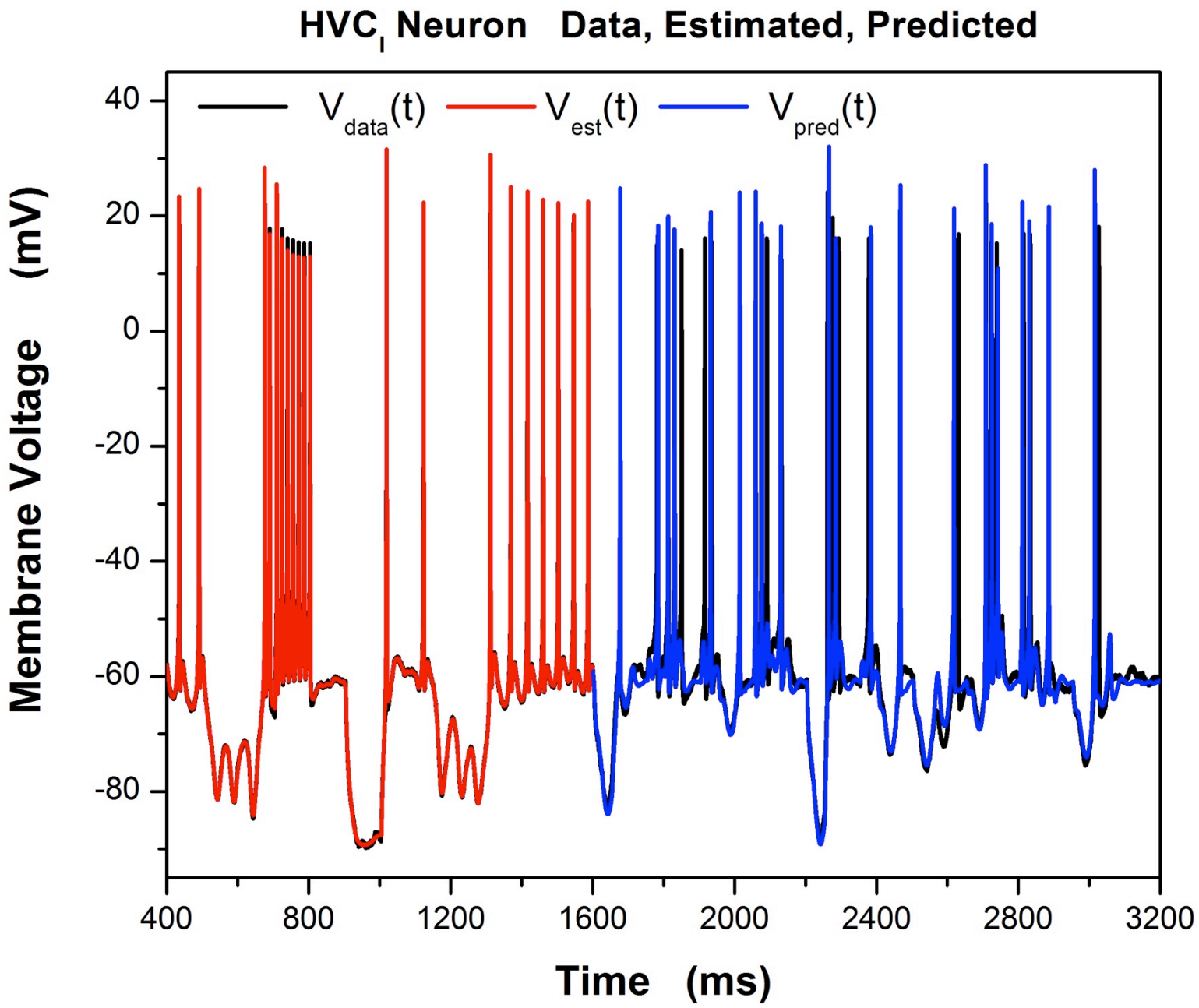
$$\frac{dx(t)}{dt} = \frac{x_\infty (V(t) - x(t))}{\tau_x (V(t))} \qquad x(t) = \{m(t), h(t), n(t), a(t), b(t)\}$$

$$x_\infty (V) = \frac{1}{2}\left(1 + \tanh(\frac{V - V_x}{dV_x})\right)$$

$$\tau_x (V) = t_{1x} + t_{2x}(1 - \tanh^2(\frac{V - V_{xt}}{dV_{xt}}))$$


HVC$_I$ Neuron
$V_{data}(t)$ (mV)
20
0
-20
-40
-60
-80
400
800
1200
1600
2000
2400
2800
3200
$I_{applied}(t)$ (pA)
150
100
50
0
-50
-100
-150
Time (ms)


HVC$_I$ Neuron Margoliash/Daou UChicago
$V_{data}(t)$
$V_{est}(t)$
$V_{pred}(t)$
V(t) (mV)
40
20
0
-20
-40
-60
-80
-100
500
1000
1500
2000
2500
3000
3500
Time (ms)

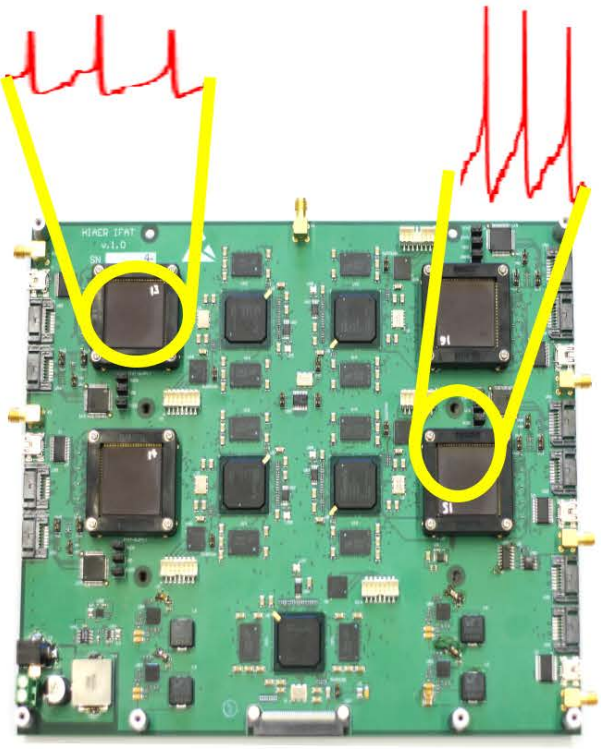

HVC$_I$ Neuron Data, Estimated, Predicted
$V_{data}(t)$
$V_{est}(t)$
$V_{pred}(t)$
Membrane Voltage (mV)
40
20
0
-20
-40
-60
-80
400
800
1200
1600
2000
2400
2800
3200
Time (ms)

# VLSI Neuromorphic Chip

- Test parameters on the chip to check quality of fabrication versus design

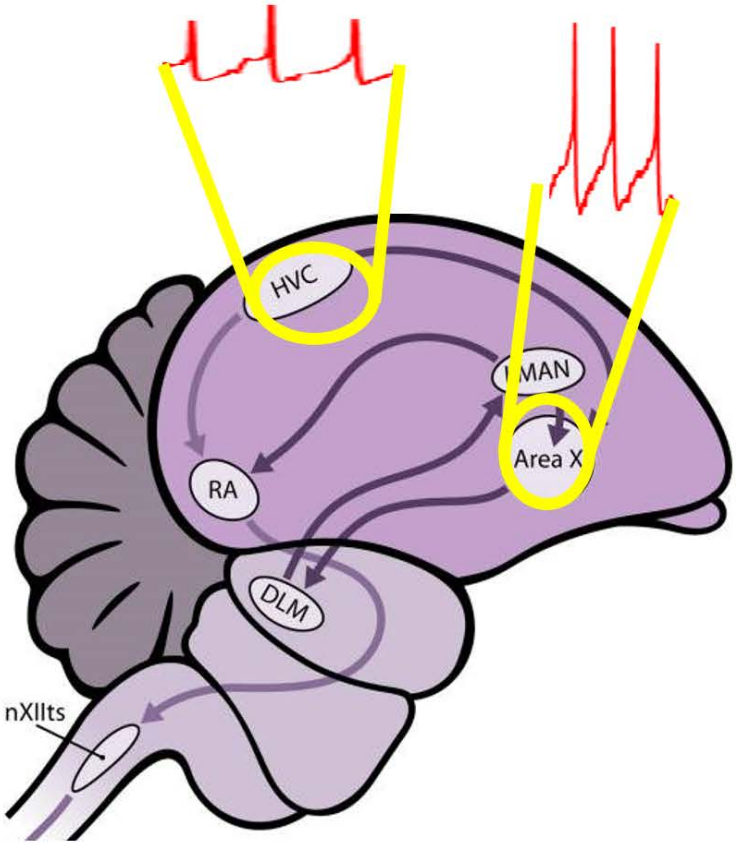
- Use a twin experiment to test method of data assimilation: generate data from the VLSI chip. Use “voltages” on chip neurons as measured quantities to estimate parameters known from first step.

- Use voltage data from biological neuron to readjust chip parameters and state variables to those for the data, then predict voltage response to new current stimulation.



Neuromorphic Chip

Data Assimilation




Brain

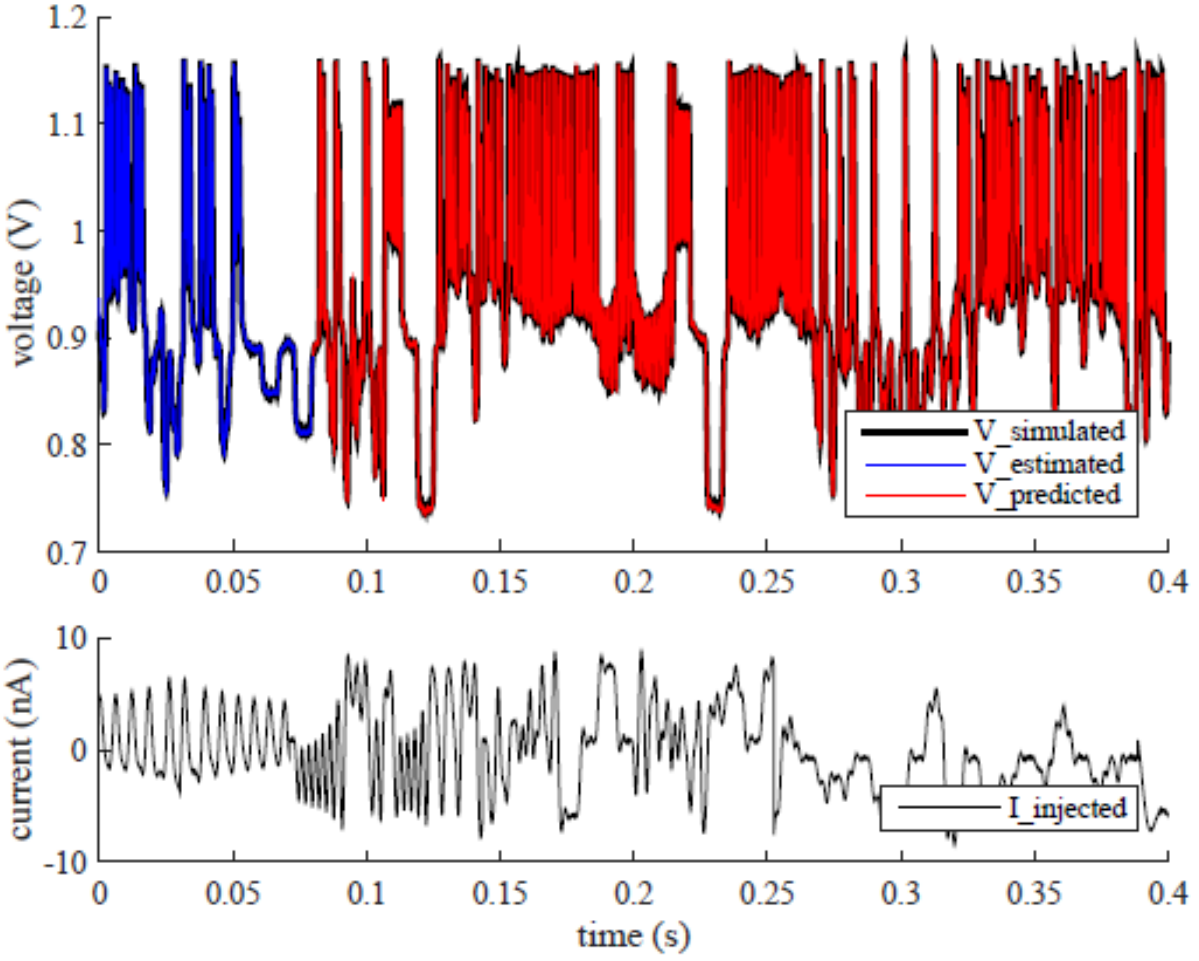



Fig. 3. Estimated (blue), predicted (red), and simulated (black) data of the NeuroDyn model in software. The prediction is obtained by integrating the equations of motion forward using the estimated parameters and configuration of the system at the end of the estimation window. The injected stimulating current protocol (black) is also displayed.

Test parameters on the chip to check quality of fabrication versus design

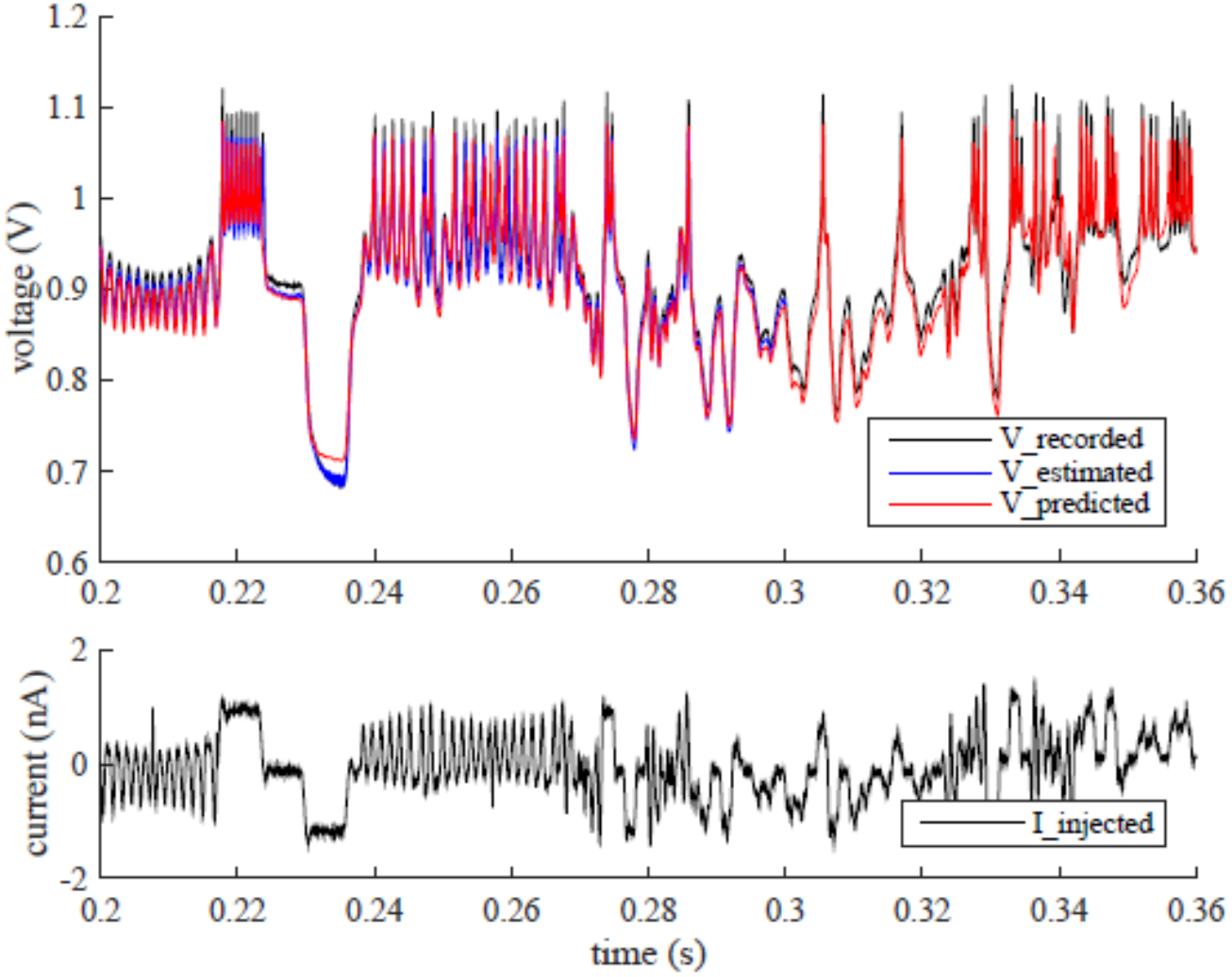



Fig. 4. Estimated (blue), predicted (red), and measured (black) data recorded from the NeuroDyn chip. The estimation and prediction were obtained, and the curent injection was applied, under identical conditions as in Fig. 3.

Use a twin experiment to test method of data assimilation: generate data from the VLSI chip. Use “voltages” on chip neurons as measured quantities to estimate parameters known from first step.

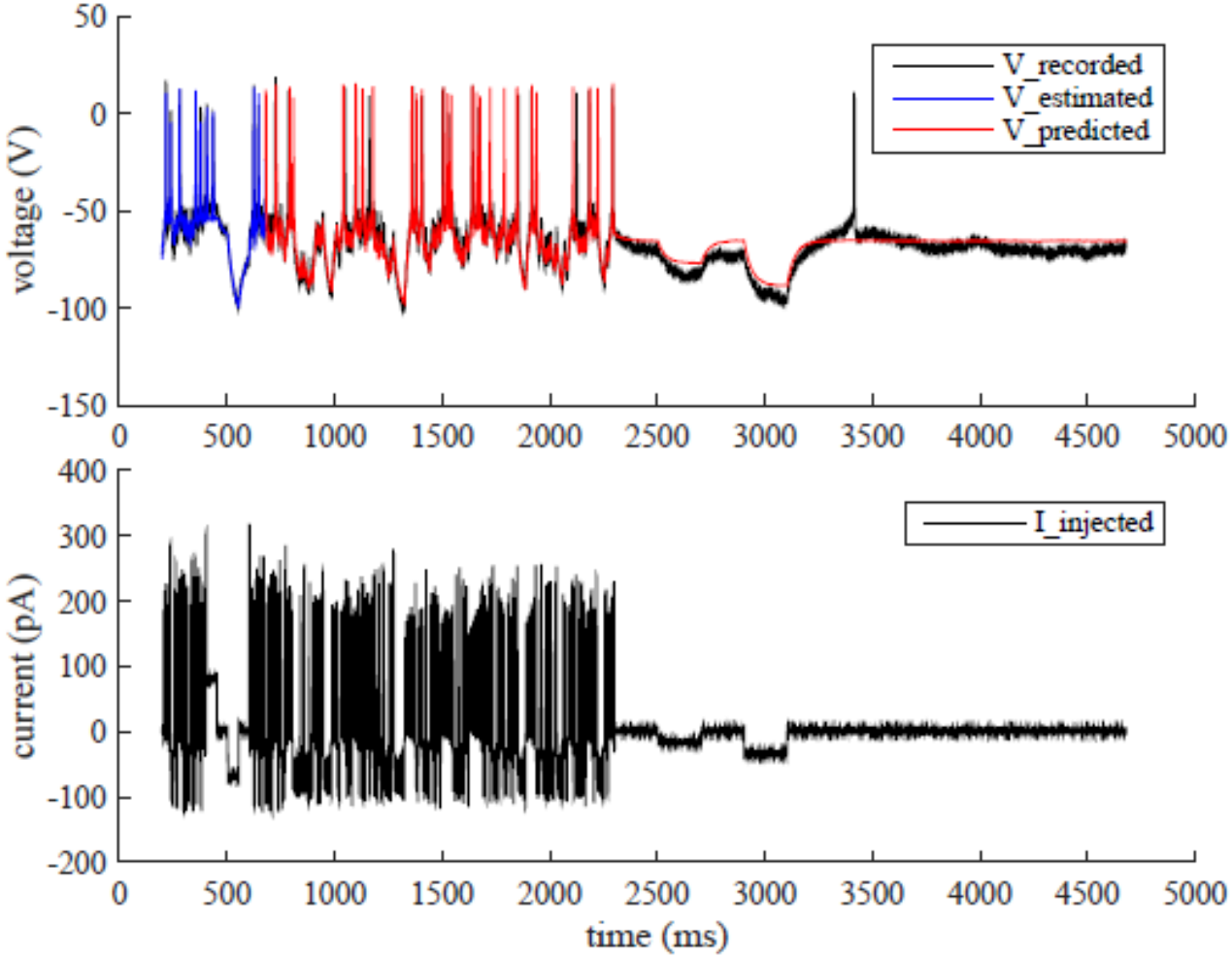



Fig. 5. Estimated (blue), predicted (red), and electrophysiological (black) data recorded *in vivo* from an HVC interneuron ($HVC_I$) [5] and instantiated onto the NeuroDyn hardware model.

Use voltage data from biological neuron to readjust chip parameters and state variables to those for the data, then predict voltage response to new current stimulation.

# So, what did we learn ?

1. Make a model—no algorithms for this, use your best knowledge of the (bio)physics.

2. Make a big model—experiments will prune the model

3. Do twin experiments to determine how many measurements you need to get the “global” minimum of $A_0(X)$—annealing.

4. Use twin experiments to design laboratory experiments

5. Do experiments to determine consistency of model with data.

6. Use the completed model and estimated x(T), via probability distribution or dx/dt = F(x(t)), to predict, for t > T. This validates (or not) the model.

## So, what did we learn ?

7. Use Laplace method + computable corrections to determine consistency of numerical methods.

8. If there are not enough measurements at each observation time, (a) get more; (b) use waveform information via time delays.

9. Using Data Assimilation, one can (a) test new fabrications of VLSI neurons, (b) test DA methods on verified VLSI chip, (c) complete neuron model on chip from biological data; predict response to new forcing.

# Unfinished Business

Measurements for Networks of Neurons---extracellular potentials? Other technology

Computational capability for the future

Port network models to VLSI

Use principles of network functions to solve similar problems in other space and time domains.