

Automated, predictive, and interpretable modeling of *C. elegans* nociceptive escape dynamics

Ilya Nemenman

Department of Physics, Department of Biology,
& Initiative in Theory and Modeling of Living Systems
Emory University

JAMES S.
MCDONNELL
FOUNDATION

nemenmanlab.org



Outline

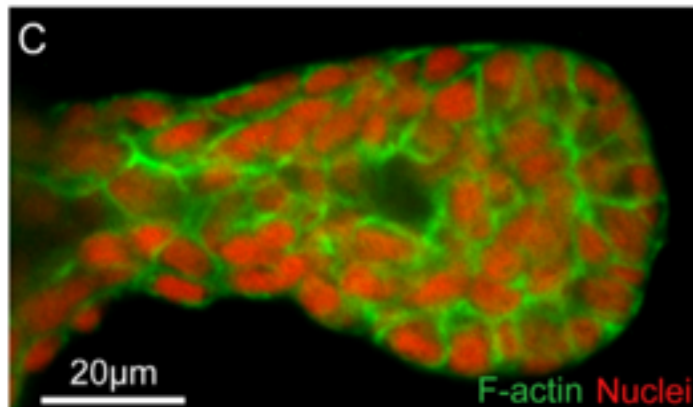
- What does my group do?
- Some philosophy
- Worms, take 1: statistics
- Worms, take 2: dynamics

What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**

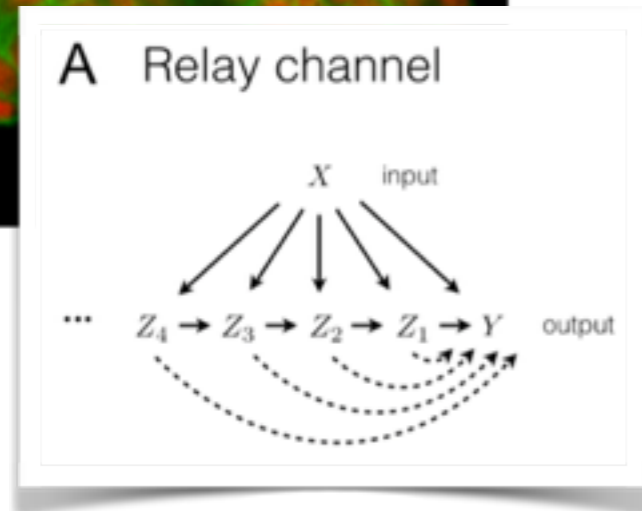
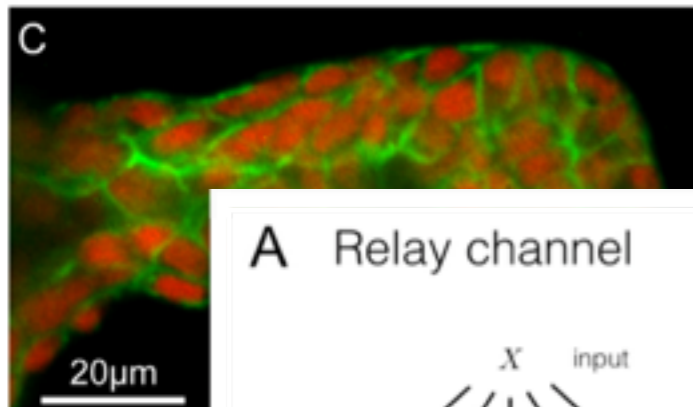
What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**



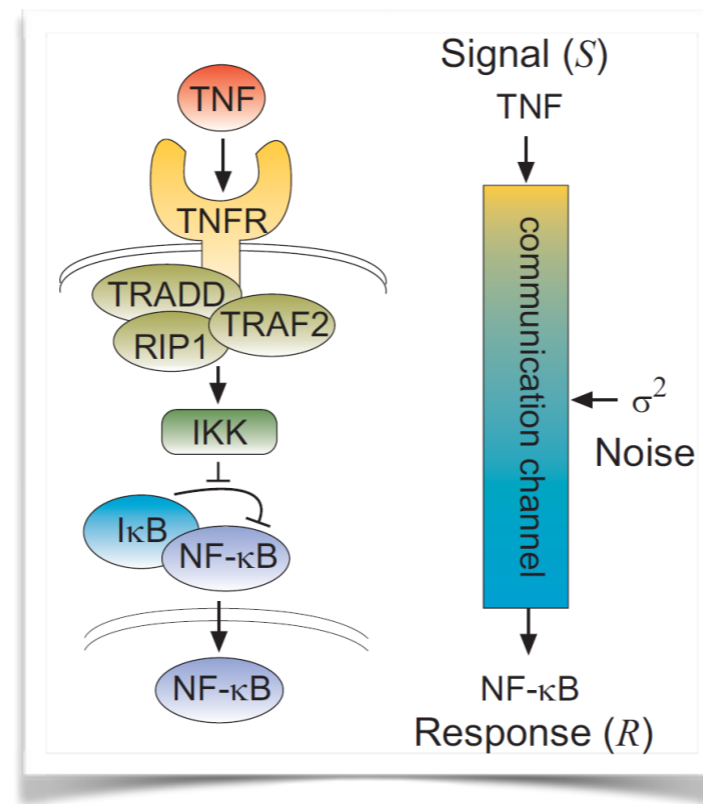
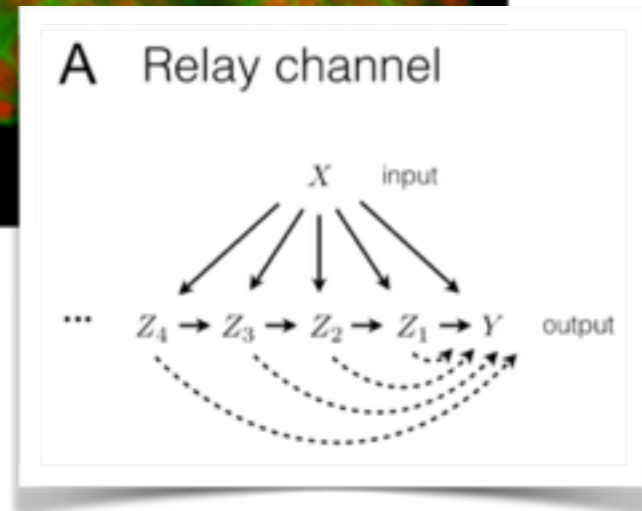
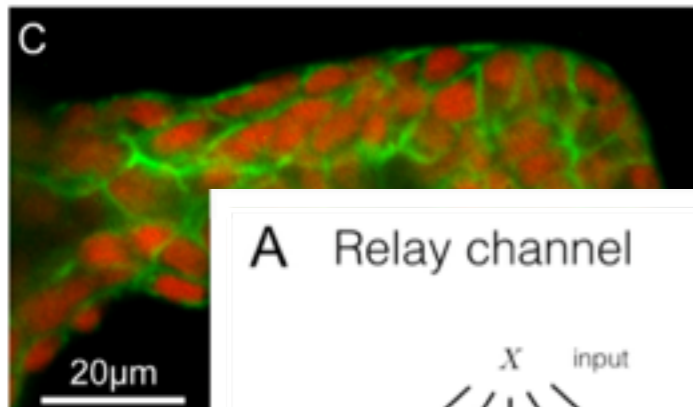
What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**



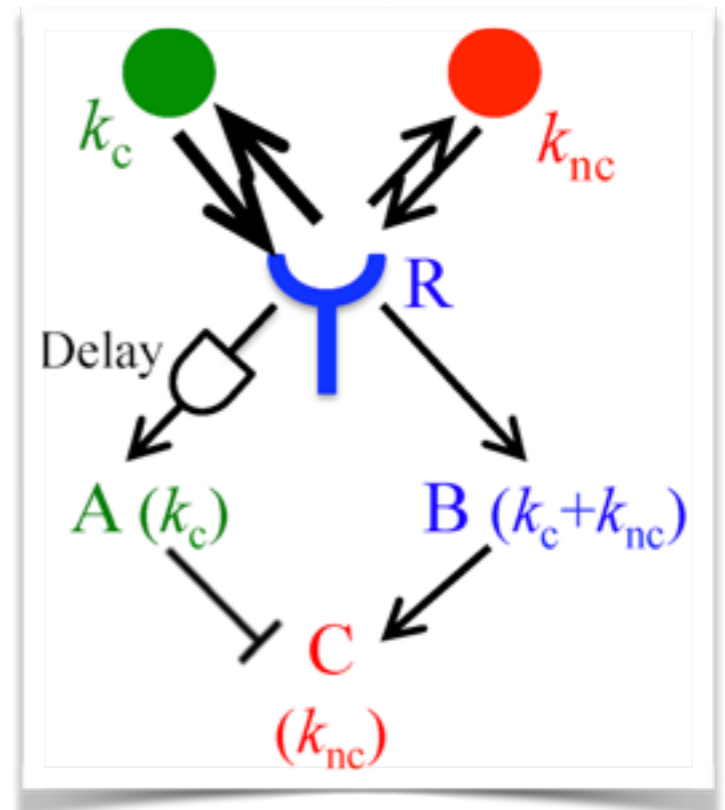
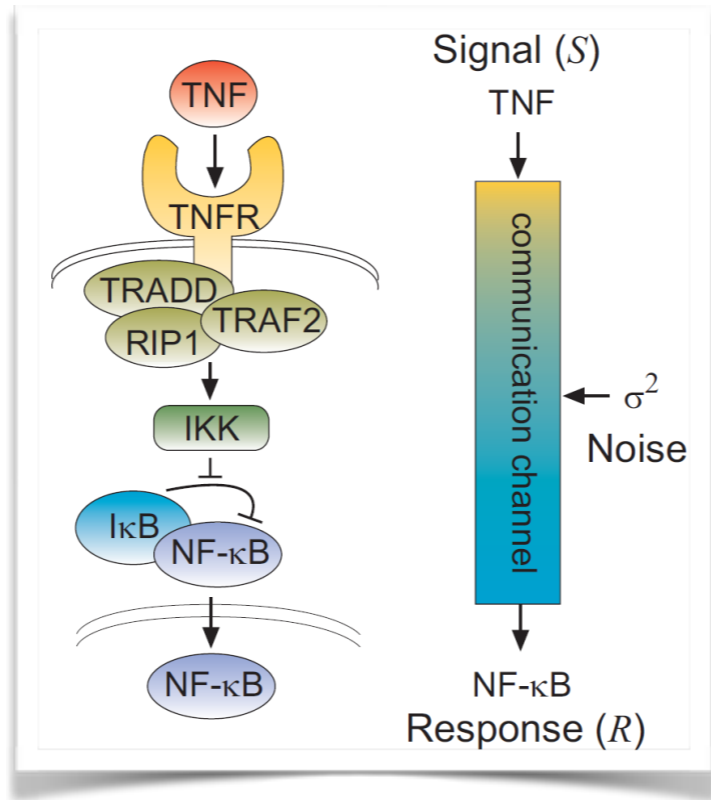
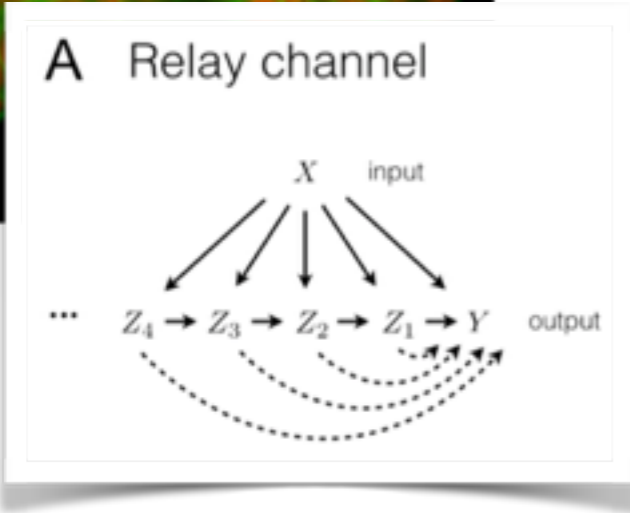
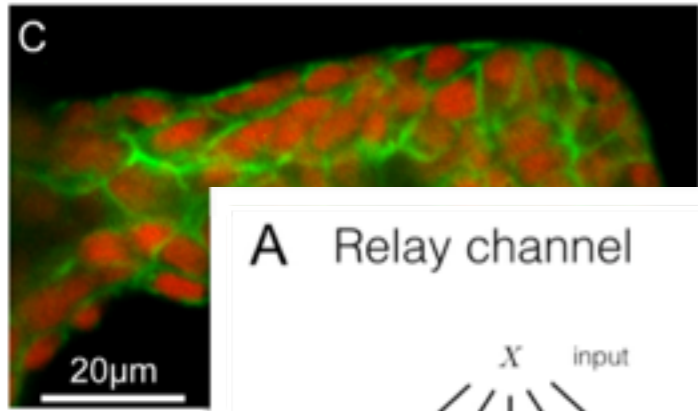
What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**



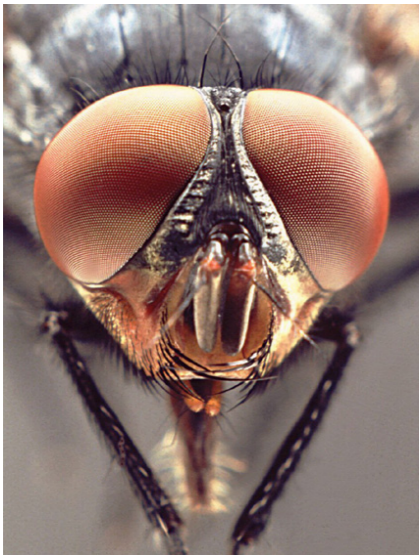
What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism** used by **biology to learn the environment**?



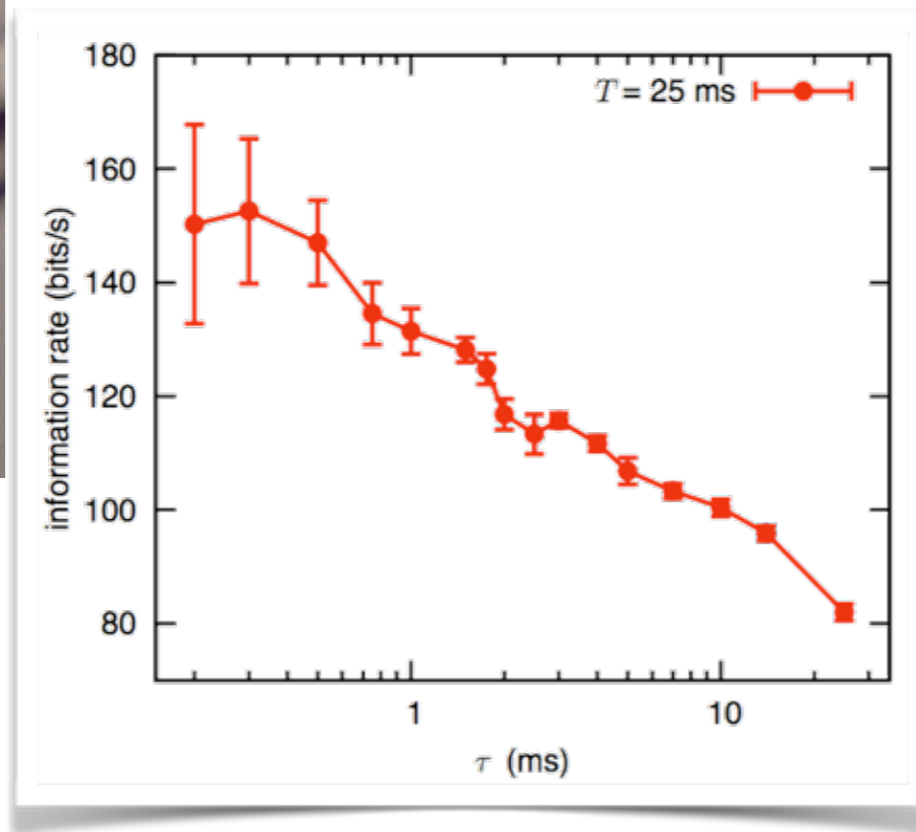
What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**



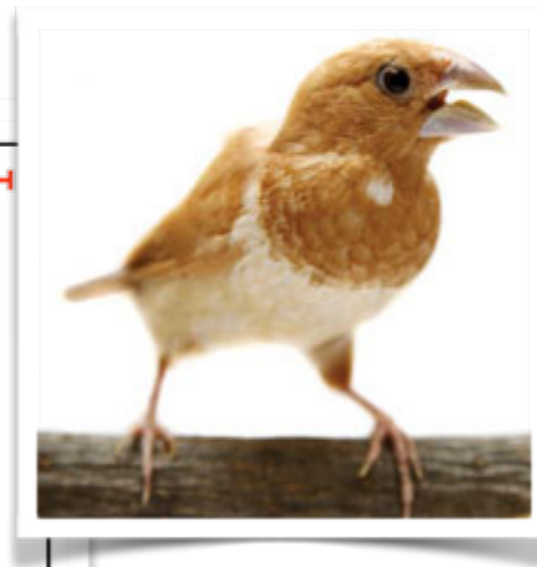
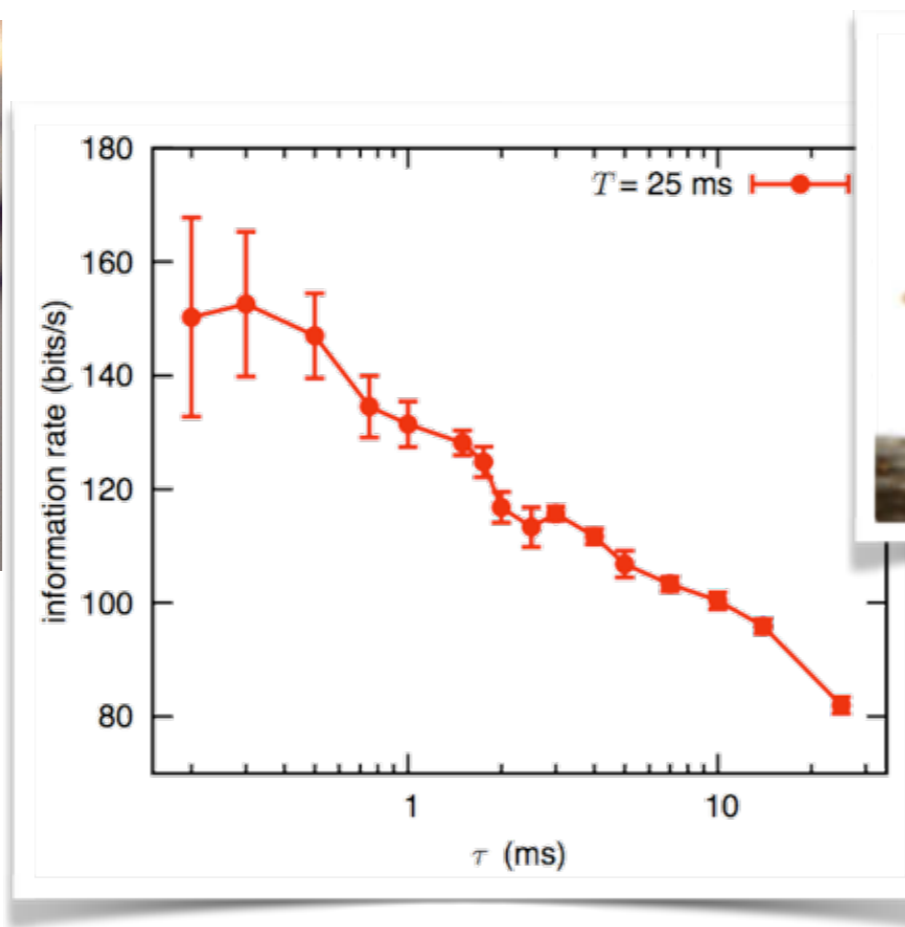
What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**



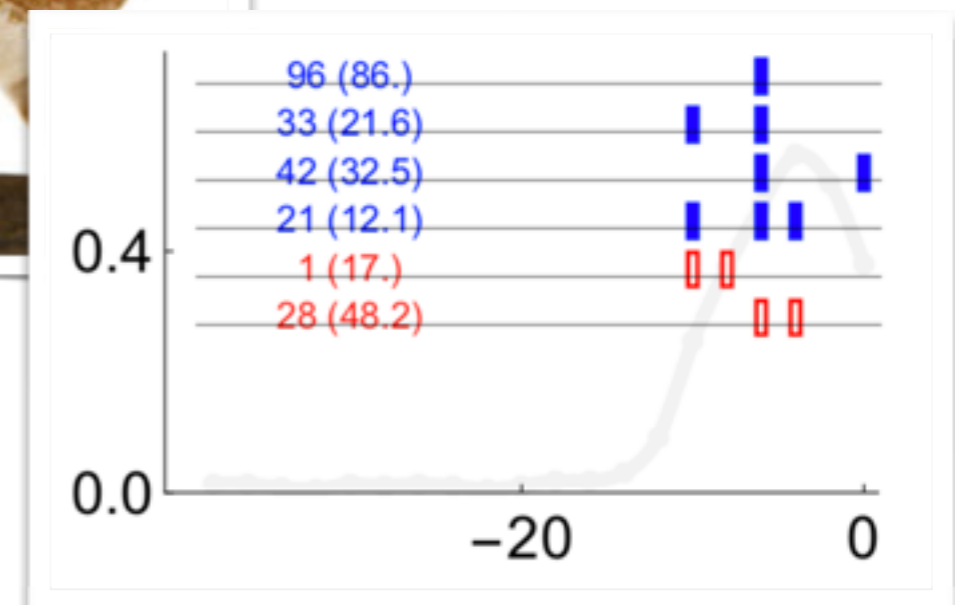
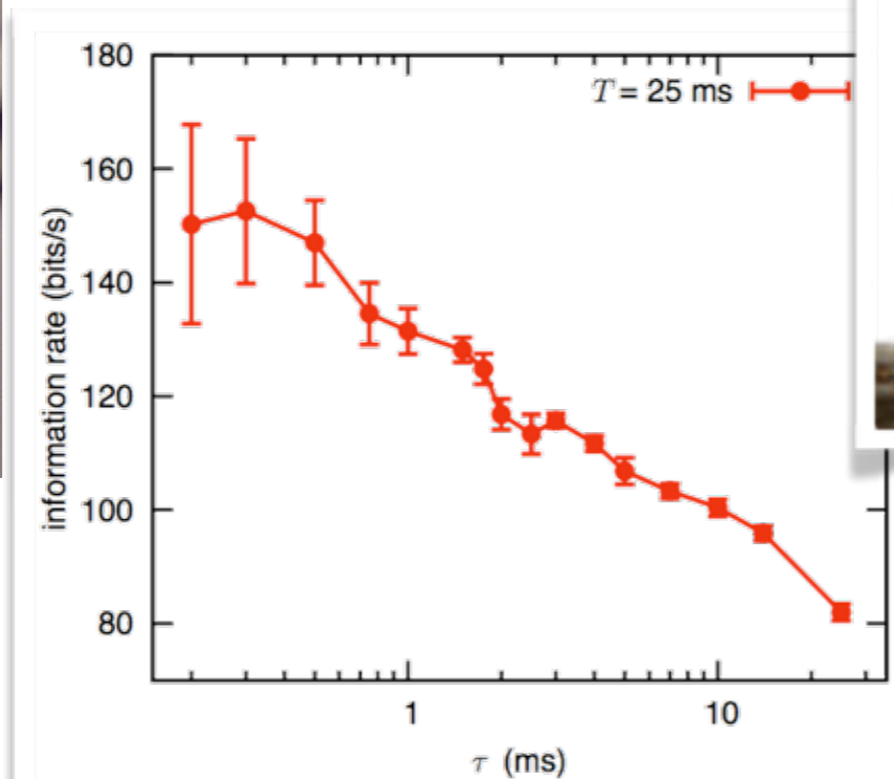
What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**



What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**



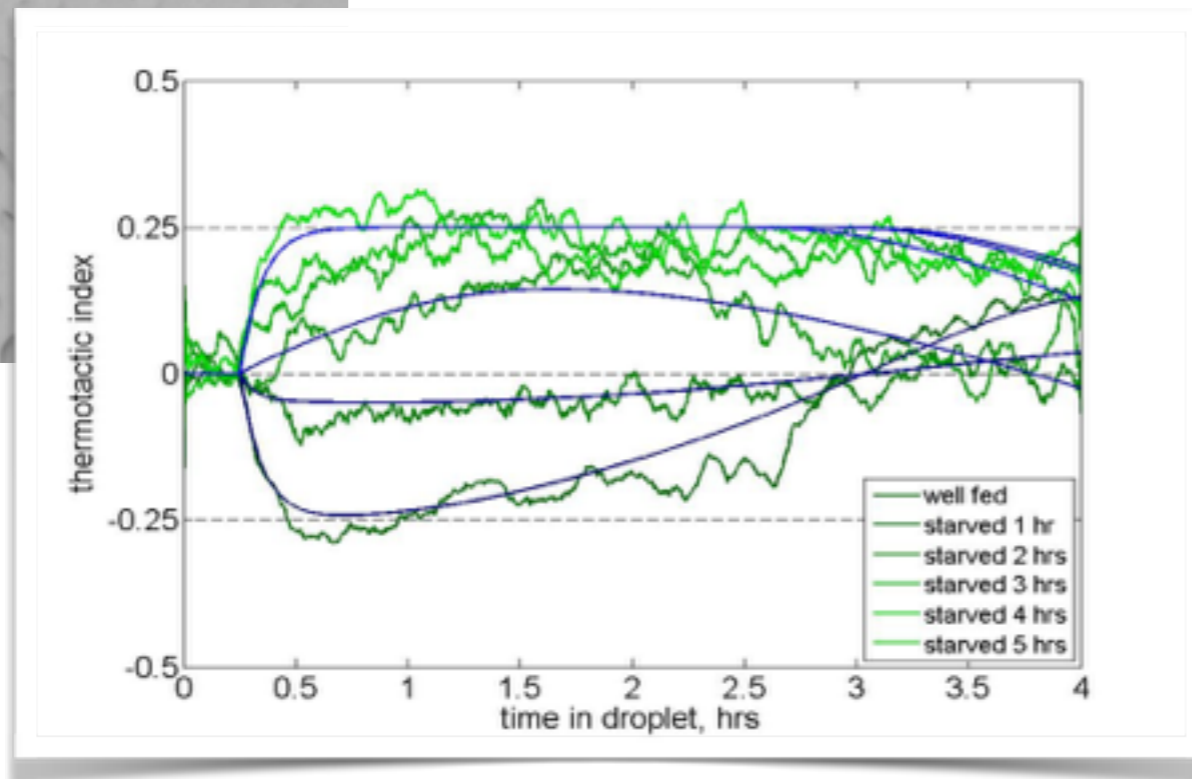
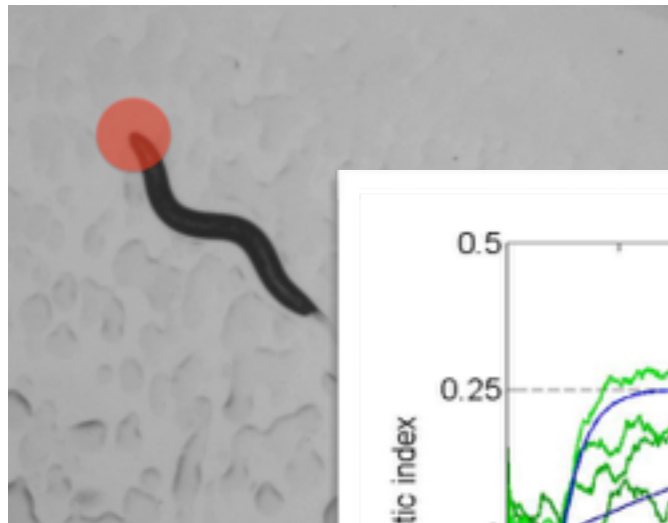
What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**



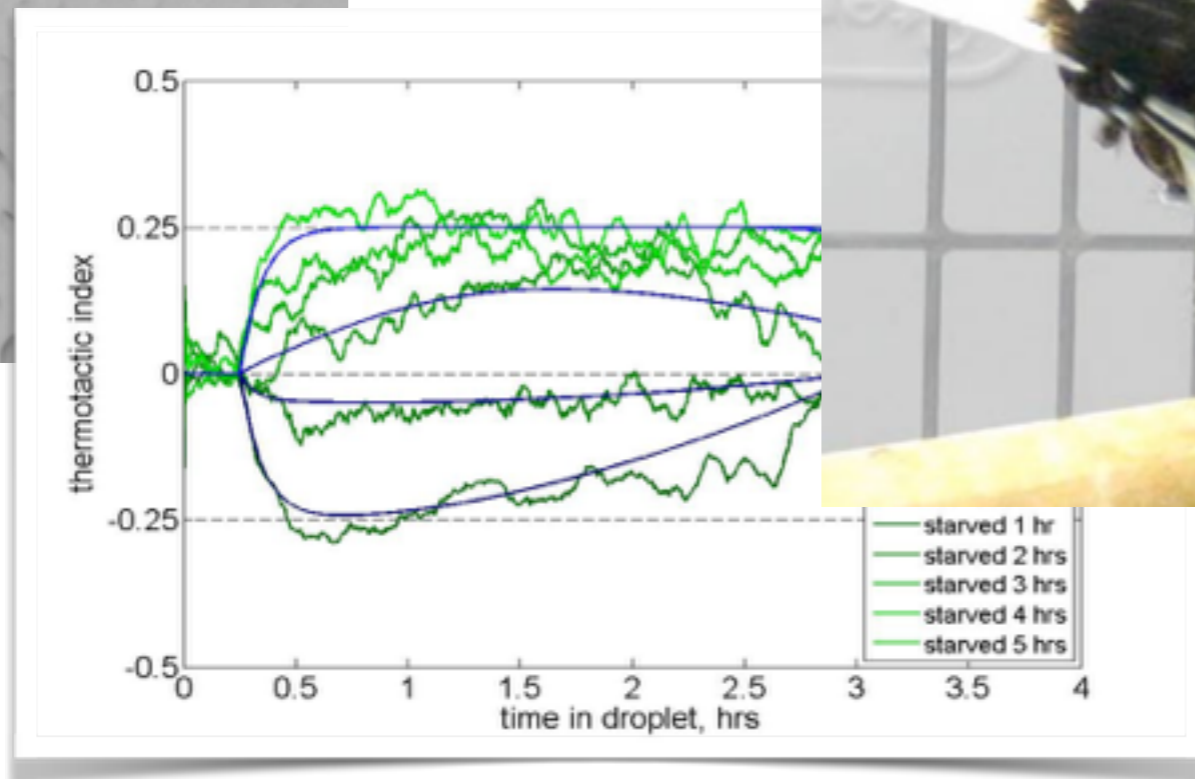
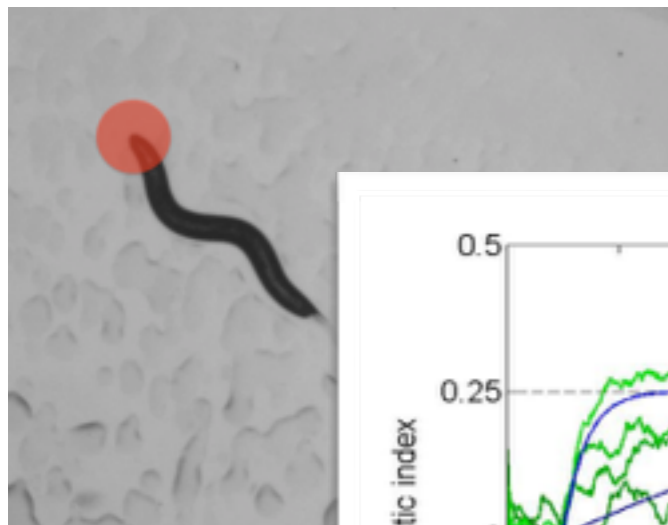
What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**



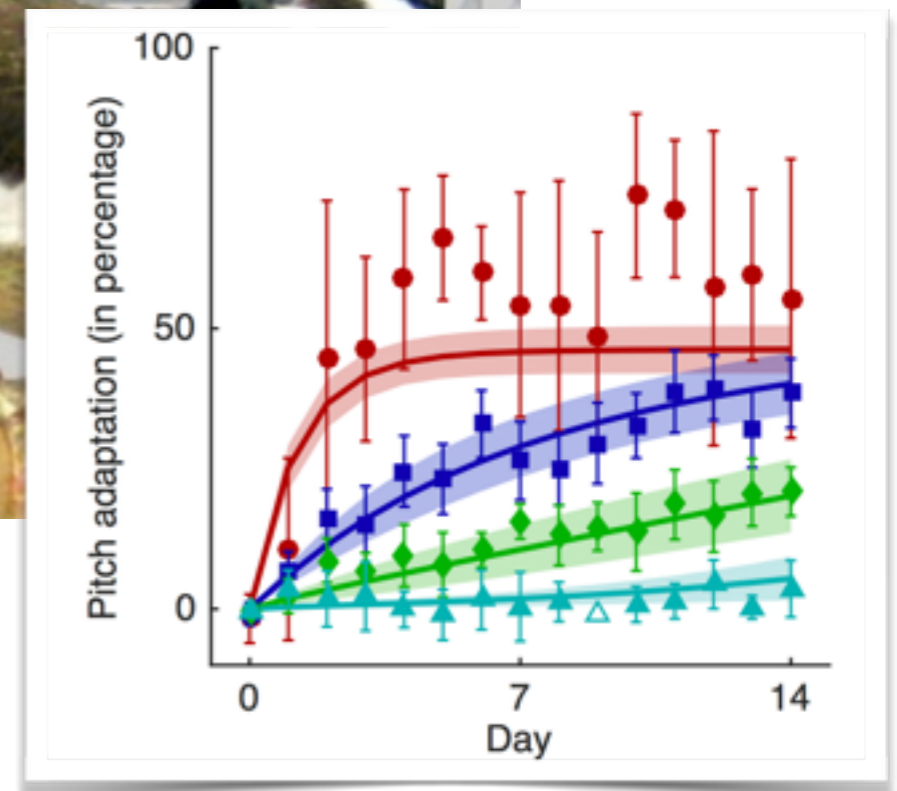
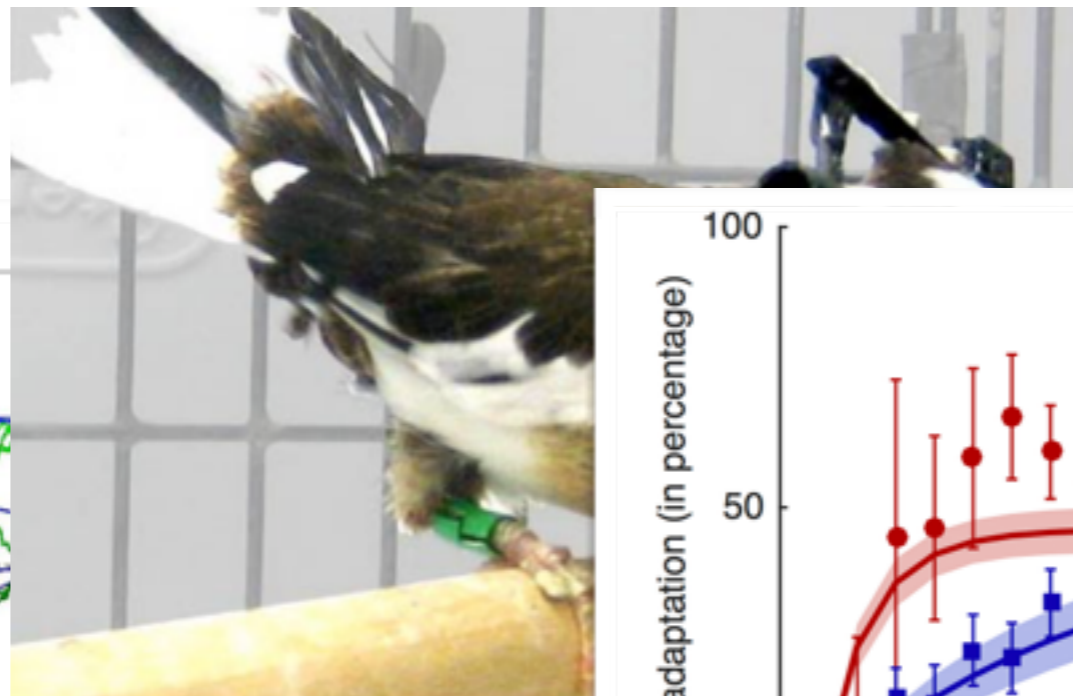
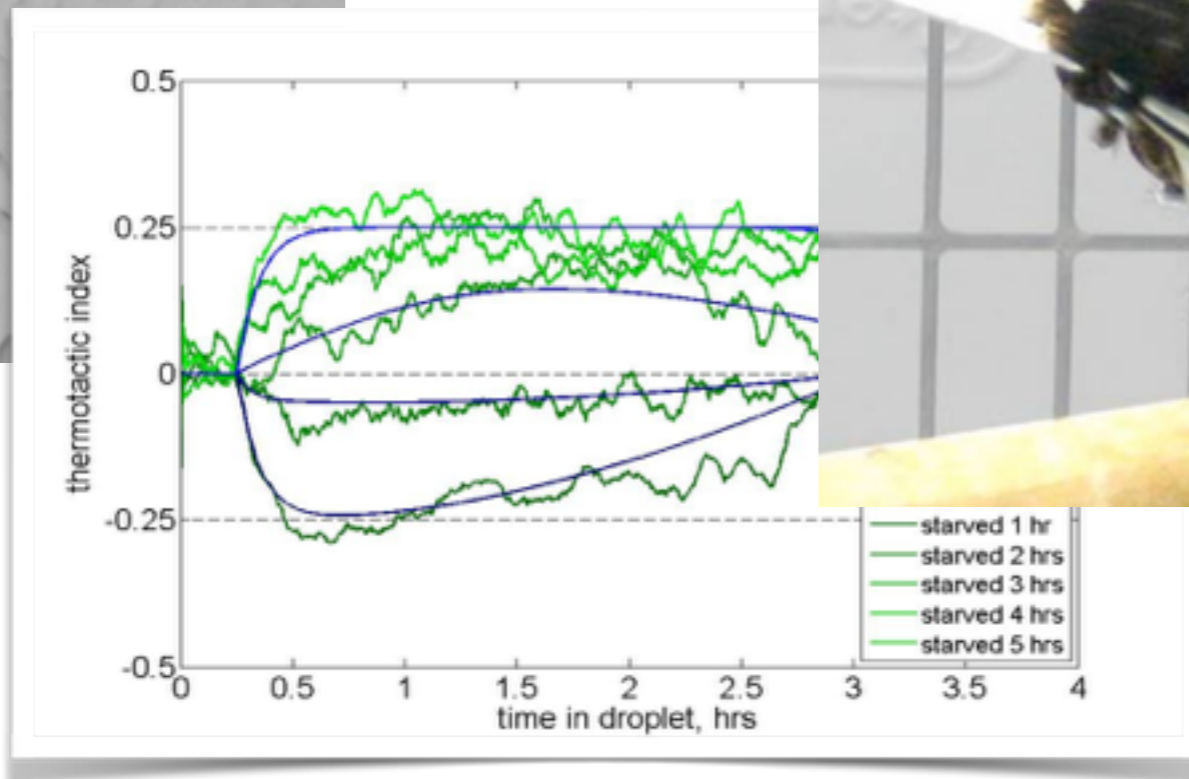
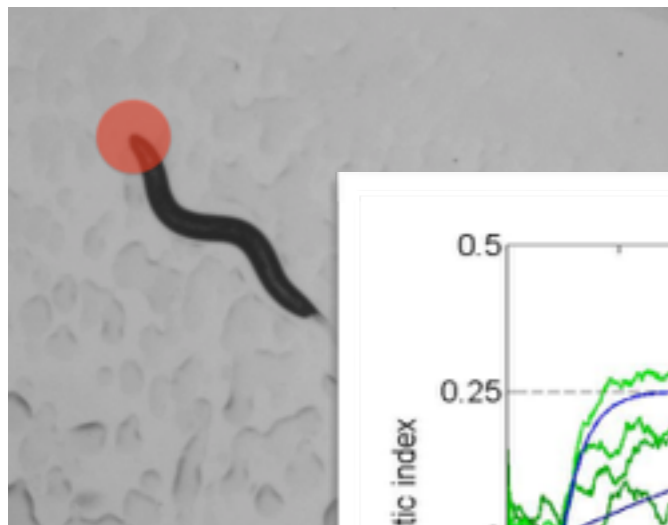
What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism used by biology to learn the environment?**



What do we study?

- How biology processes information and learns?
 - Neurons, individual cells, brains, populations...
 - What are the limits to our information processing capabilities?
 - What are the **computational** and biophysical **mechanism** used by **biology to learn the environment**?



Why do we study this?

Why do we study this?

“Studying string theory cannot be more exciting than studying the brain that can study string theory.”

Thanks

- Bryan Daniels (Arizona State)
- Will Ryu (Toronto)
- KaWai (George) Leung (data science somewhere in Boston)
- Aliya Mohammadi (Toronto)

Successes of q-bio

- We know that circuits (neural, genetic, signaling) can be understood mathematically.
- Noise and individual variability are not just “...it happens,” but can also be modeled.
- We (often) understand why the circuits are the way they are from control-theoretic, information-theoretic, evolutionary, and optimization principles.
- There is probably no “paradigm change”, no “new physics” at the small circuits level: putting enough experiments, stat mech, and nonlinear dynamics together, we can understand every circuit.
- Measuring 1000s of variables is relatively easier than knowing what to do with the measurements.
- And yet, **more is different**, and we don't know how to deal with it.

How do we model this?



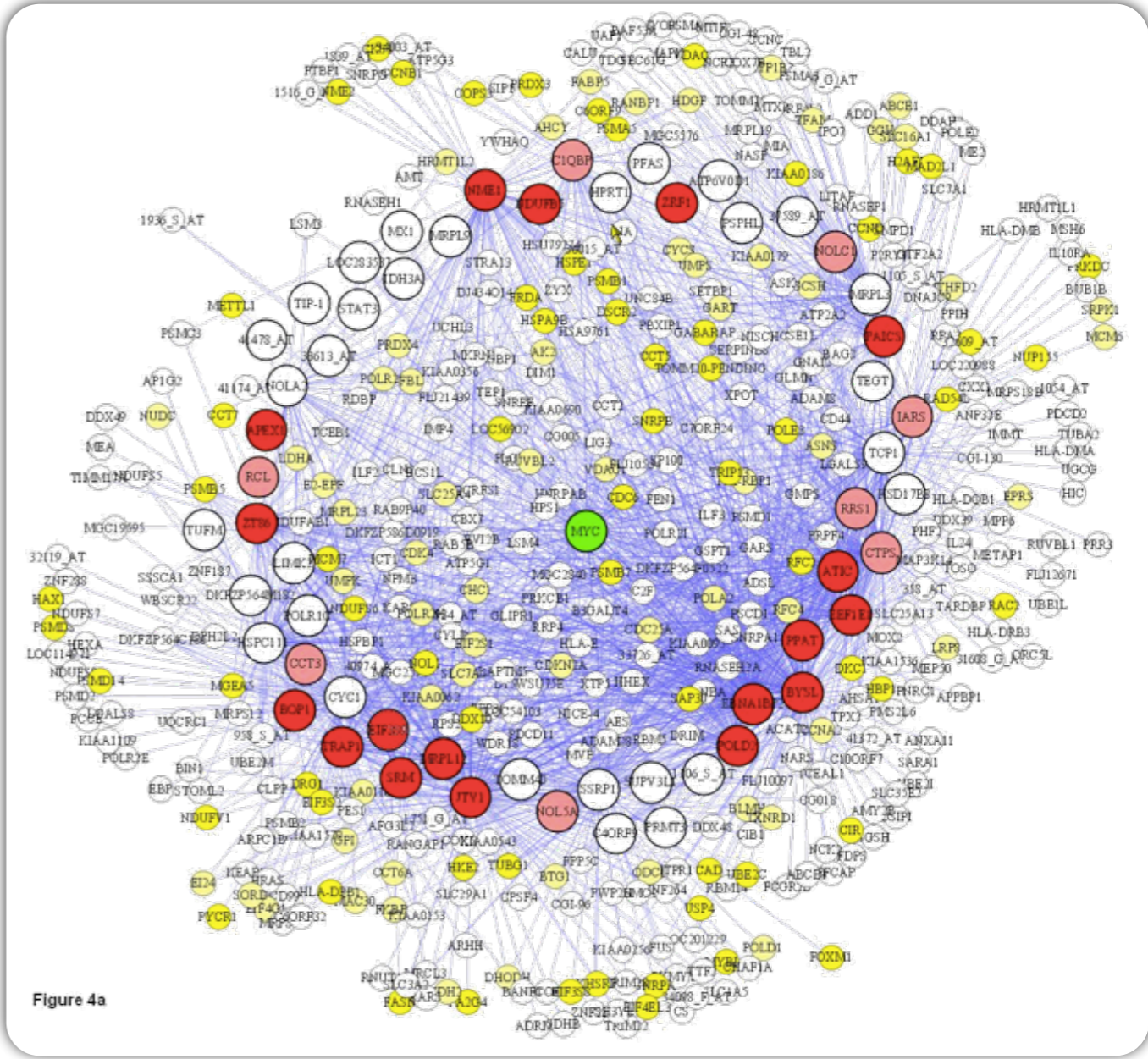
How do we model this?



How do we go from circuits to organisms?

A culture's icons are a window onto its soul. Few would disagree that, in the culture of molecular biology that dominated much of the life sciences for the last third of the 20th century, the dominant icon was the double helix. In the present, post-modern, 'systems biology' era, however, it is, arguably, the hairball.

A.D. Lander. BMC Biology 2010, 8:40



Margolin et al., 2006

Of exactitude in science

...In that Empire, the craft of Cartography attained such Perfection that the Map of a Single province covered the space of an entire City, and the Map of the Empire itself an entire Province. In the course of Time, these Extensive maps were found somehow wanting, and so the College of Cartographers evolved a Map of the Empire that was of the same Scale as the Empire and that coincided with it point for point. Less attentive to the Study of Cartography, succeeding Generations came to judge a map of such Magnitude cumbersome, and, not without Irreverence, they abandoned it to the Rigours of sun and Rain. In the western Deserts, tattered Fragments of the Map are still to be found, Sheltering an occasional Beast or beggar; in the whole Nation, no other relic is left of the Discipline of Geography.

From *Travels of Praiseworthy Men* (1658) by J. A. Suarez Miranda (a fictional reference).
By Jorge Luis Borges and Adolfo Bioy Casares.
English translation quoted from J. L. Borges, *A Universal History of Infamy*,
Penguin Books, London, 1975.

At CNS 2014

At CNS 2014

- A common opinion: “**The final** theory of biological systems will be a large multiscale computational model. We need more experimental data to specify its details.”

At CNS 2014

- A common opinion: “**The final** theory of biological systems will be a large multiscale computational model. We need more experimental data to specify its details.”
- There’s something wrong with this statement, in my opinion.
 - What is the “**final**” theory?
 - Do we need the theory of “**everything**” in any biological (or physical) system?
 - What’s the final theory of my coffee mug?

At CNS 2014

- A common opinion: “**The final** theory of biological systems will be a large multiscale computational model. We need more experimental data to specify its details.”
- There’s something wrong with this statement, in my opinion.
 - What is the “**final**” theory?
 - Do we need the theory of “**everything**” in any biological (or physical) system?
 - What’s the final theory of my coffee mug?
- “Don’t model bulldozers with quarks.” (Goldenfeld and Kadanoff, 1999)
 - Theories must lose details and must be developed to explain **limited** sets of phenomena.

At CNS 2014

- A common opinion: “**The final** theory of biological systems will be a large multiscale computational model. We need more experimental data to specify its details.”
- There’s something wrong with this statement, in my opinion.
 - What is the “**final**” theory?
 - Do we need the theory of “**everything**” in any biological (or physical) system?
 - What’s the final theory of my coffee mug?
- “Don’t model bulldozers with quarks.” (Goldenfeld and Kadanoff, 1999)
 - Theories must lose details and must be developed to explain **limited** sets of phenomena.
- Otherwise: The best material model of a cat is another, or preferably the same, cat. (*Philosophy of Science*, Wiener and Rosenblueth, 1945)

Without loosing details, biology ceases to be science

- Western scientific tradition:
 - There are laws (of nature, of god, whatever). A rock is a rock everywhere. It falls the same in Pisa and in Atlanta.
 - There're causes and there are effects.
 - There is “useless information” (Oscar Wilde).
 - But this belief requires closing one's eyes to minor discrepancies
 - Two balls dropped from the Leaning Tower didn't actually land simultaneously.
 - “If we had the STM in the 1920s, there wouldn't be the Debye theory of solids.” (H. Levine)
- Non-western tradition, e.g., buddhism
 - Pratityasamutpada: dependent origination: “*Pratitya samutpada* is sometimes called the teaching of cause and effect, but that can be misleading, because we usually think of cause and effect as separate entities, with cause always preceding effect, and one cause leading to one effect. According to the teaching of Interdependent Co-Arising, cause and effect co-arise (*samutpada*) and everything is a result of multiple causes and conditions...” — Thich Nhat Hanh

So...

- Are there *phenomenological, coarse-grained*, and yet **functionally accurate representations of activity of (some) biological systems**, or are we forever doomed to every **detail mattering**?
 - Such models would not answer *every* question, but specific questions on coarse scales.
- How can we learn such phenomenological models directly from data?

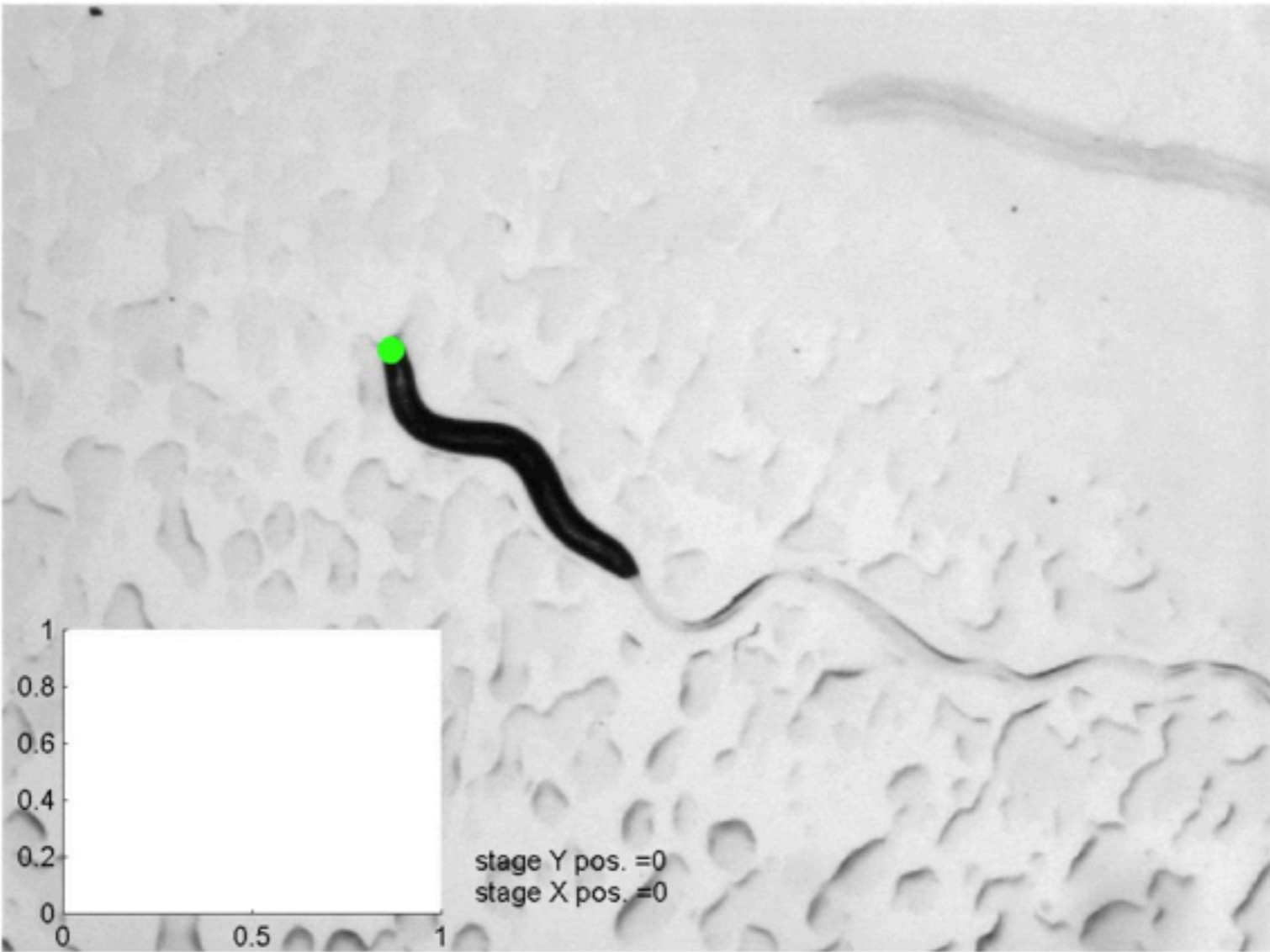
In this talk

- We will study one data set of nociceptive escape response in the worm.
- We will first build a purely phenomenological model of this process, designed specifically to answer behavioral questions.
- We will then build a dynamical model of the same data using an automated inference system.

Getting back to the worms



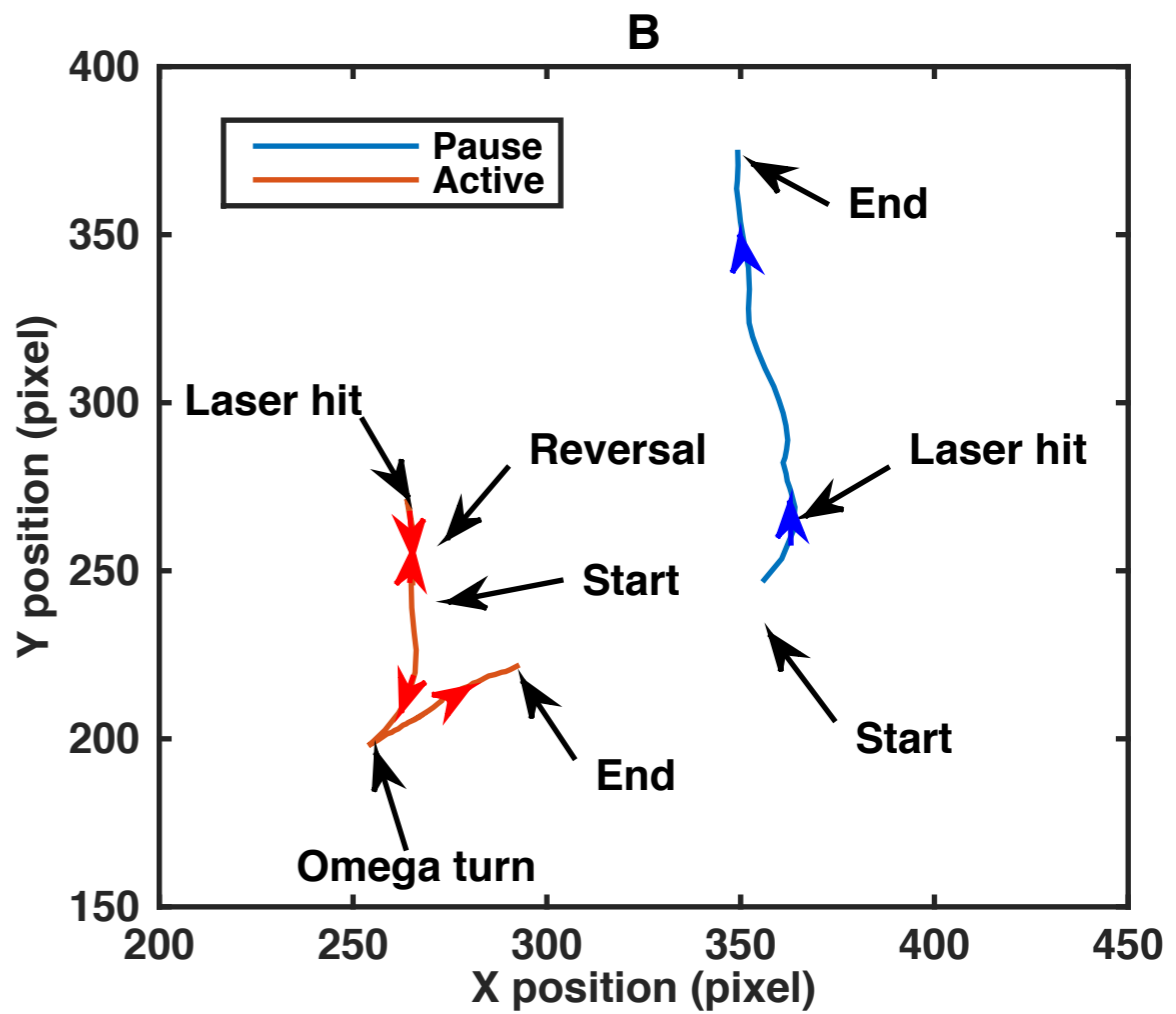
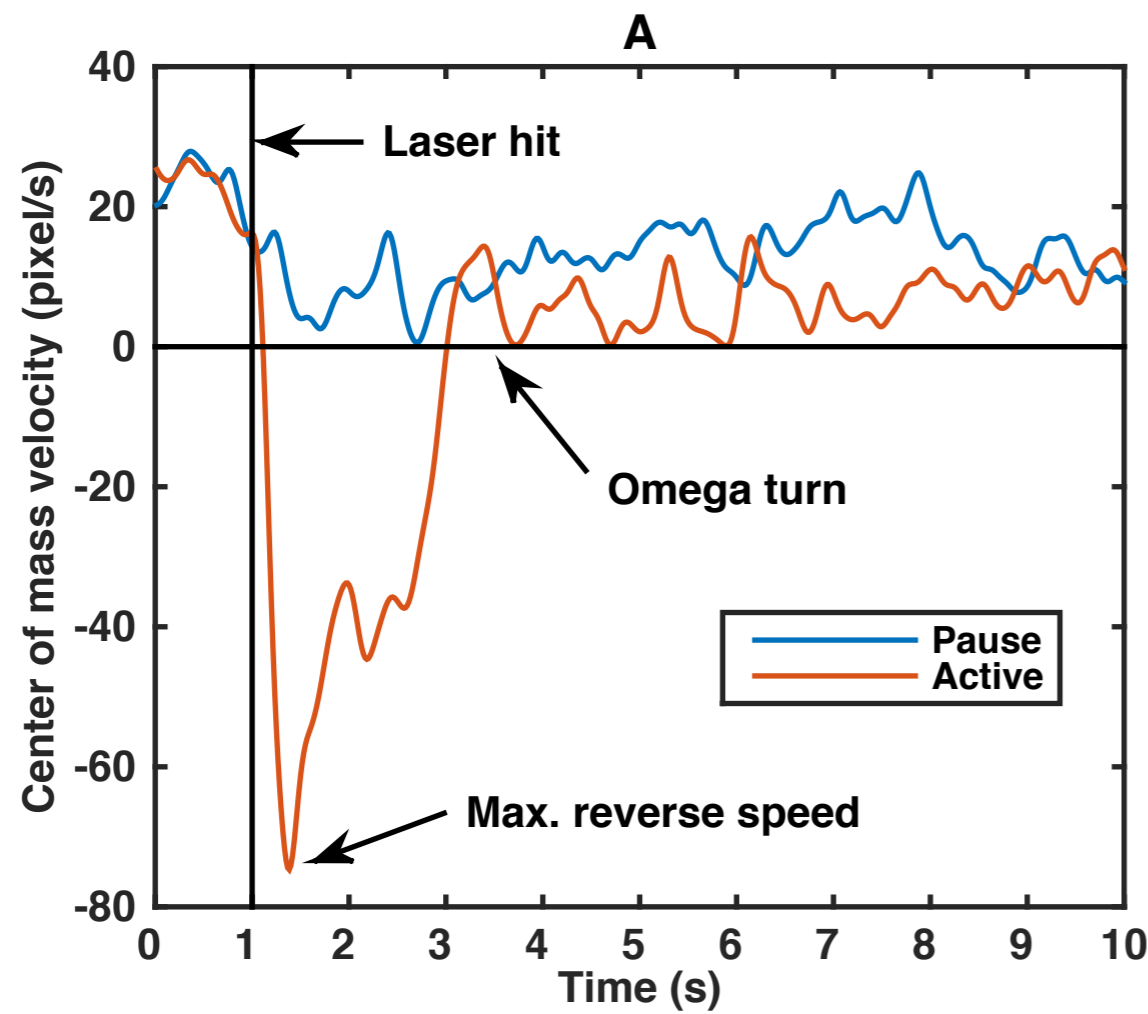
Getting back to the worms



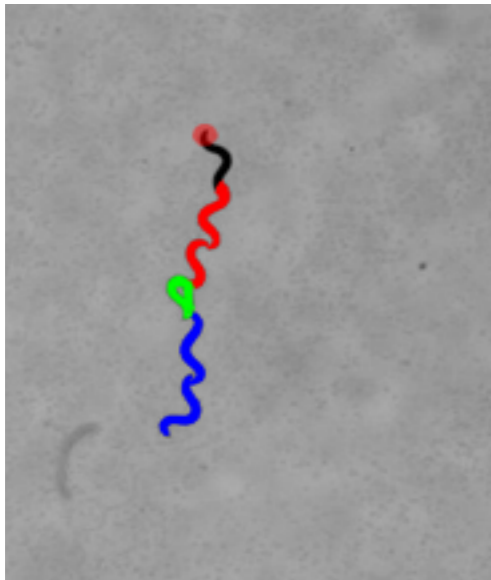
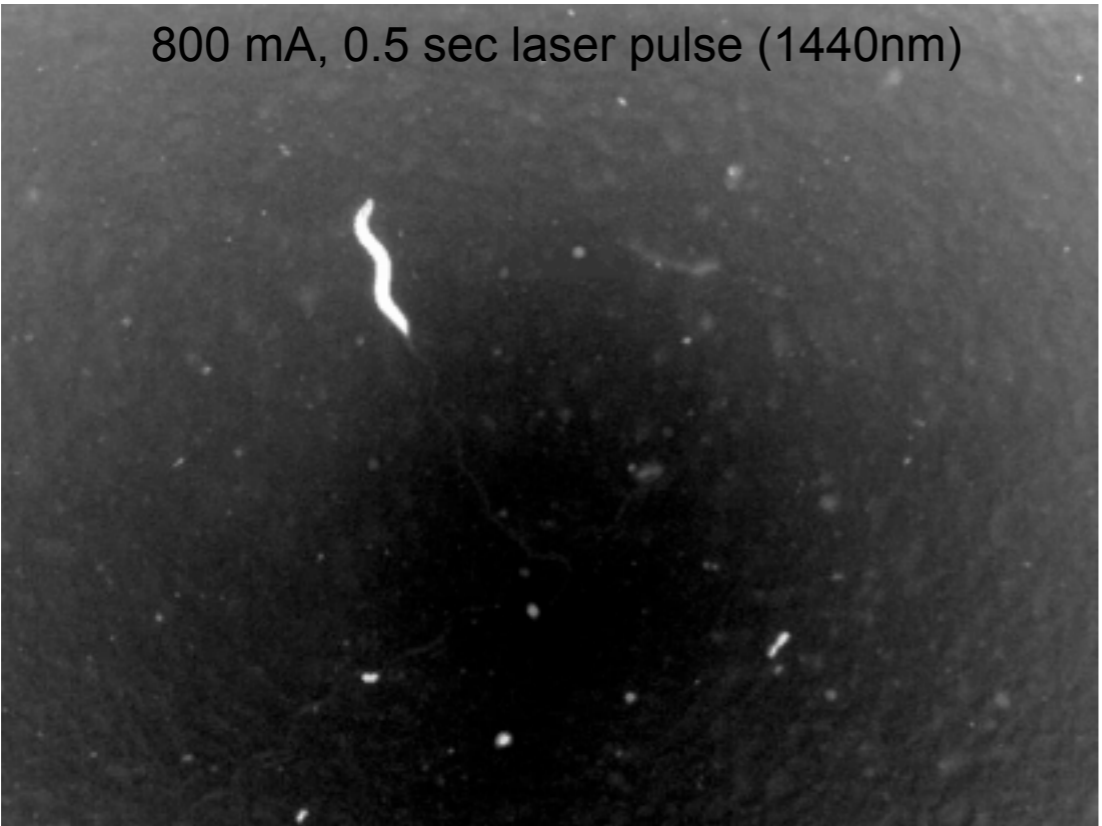
Questions

- How does the sensory information influence the behavior?
 - Can we predict the behavior from the sensory signals?
 - **Can we postdict the sensory signals from the behavior?**
- Nociception
 - **Can't ask an animal** "Could you please rate the pain you experience on the scale of 1-10?"
 - How do we test **analgesics** using animal models?
- How to distinguish **analgesic effects** of drugs from effects on motor behavior?
 - An old joke: how do you prove that a roach hears with its legs?
- How to model complex behaviors?
 - Detailed, mechanistic modeling, from molecules to neurons to behavior over long time scales is *unlikely to work*.
 - Need phenomenological models.

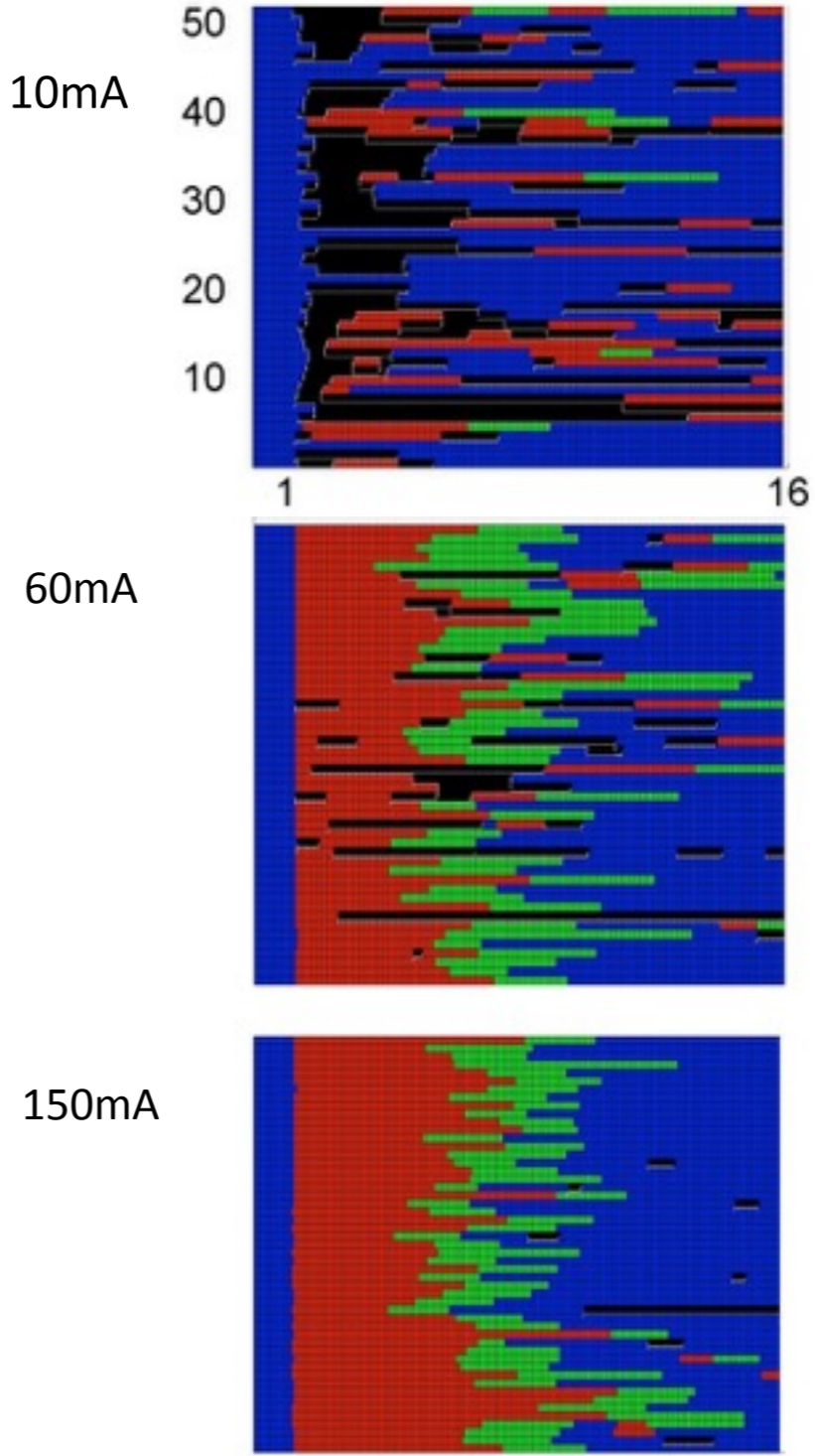
Typical center of mass trajectories



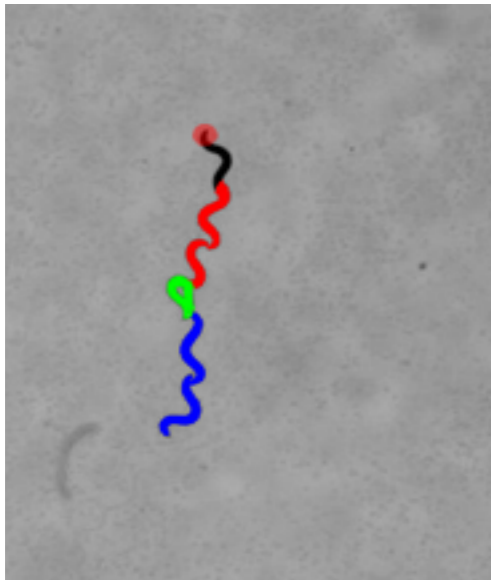
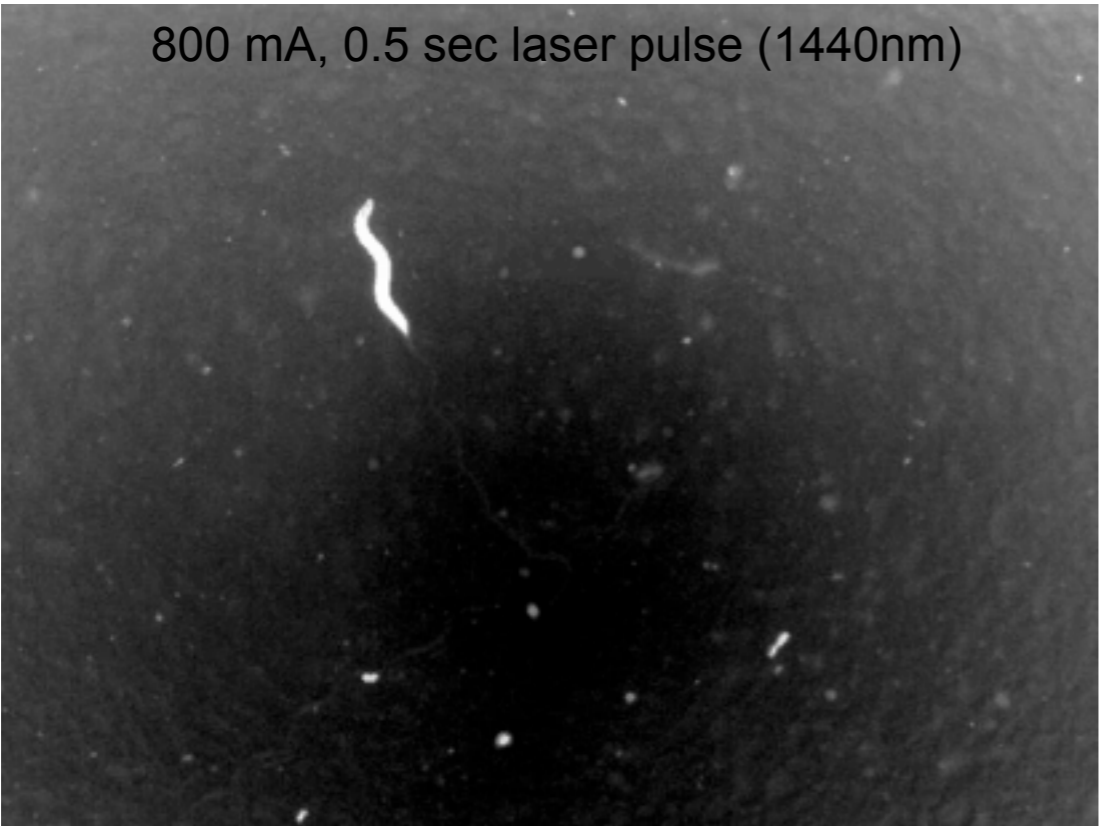
Properties of the escape behavior: Statistics of states depends on stimuli



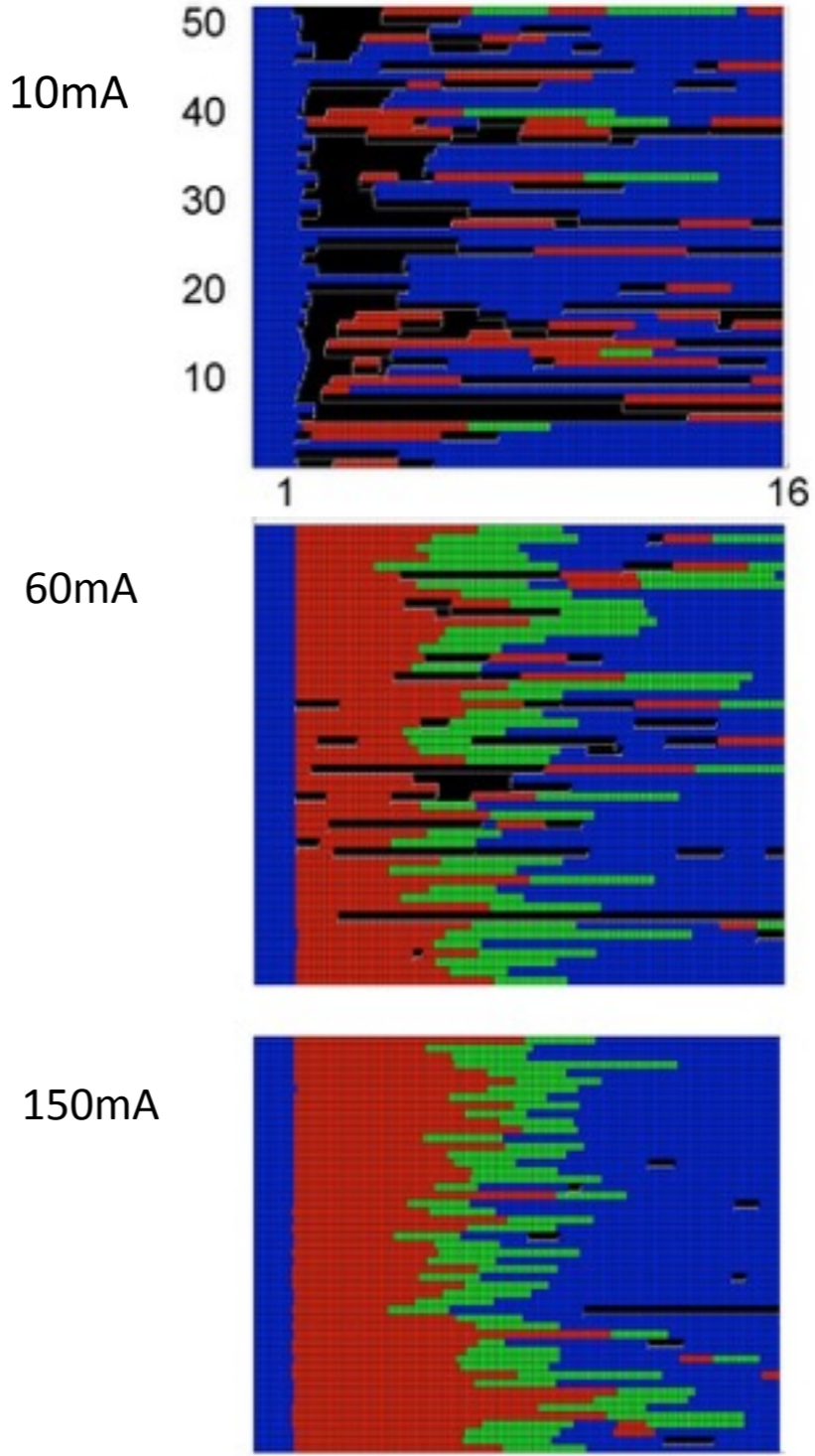
- REVERSE
- FORWARD
- PAUSE



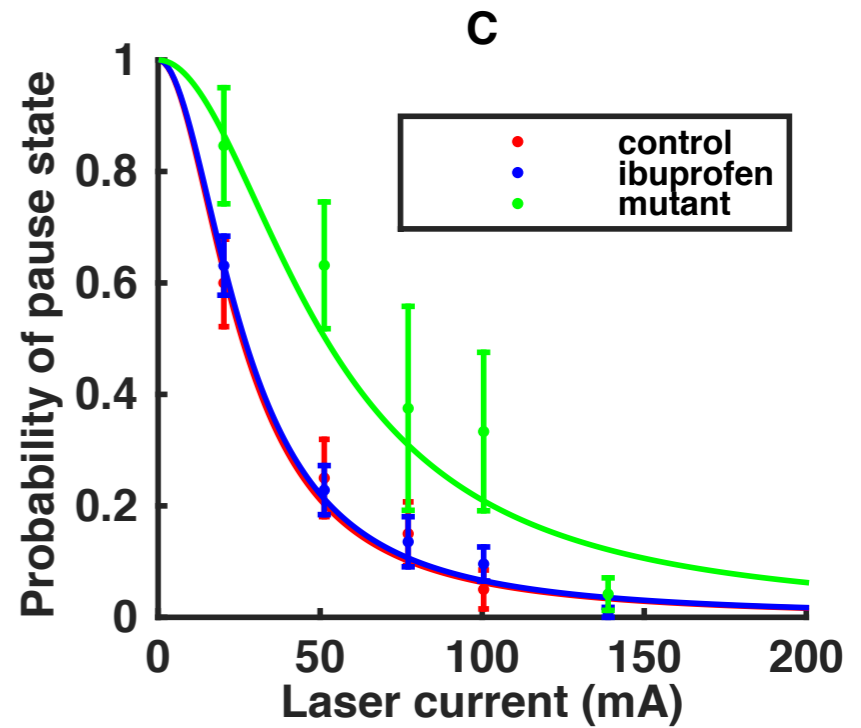
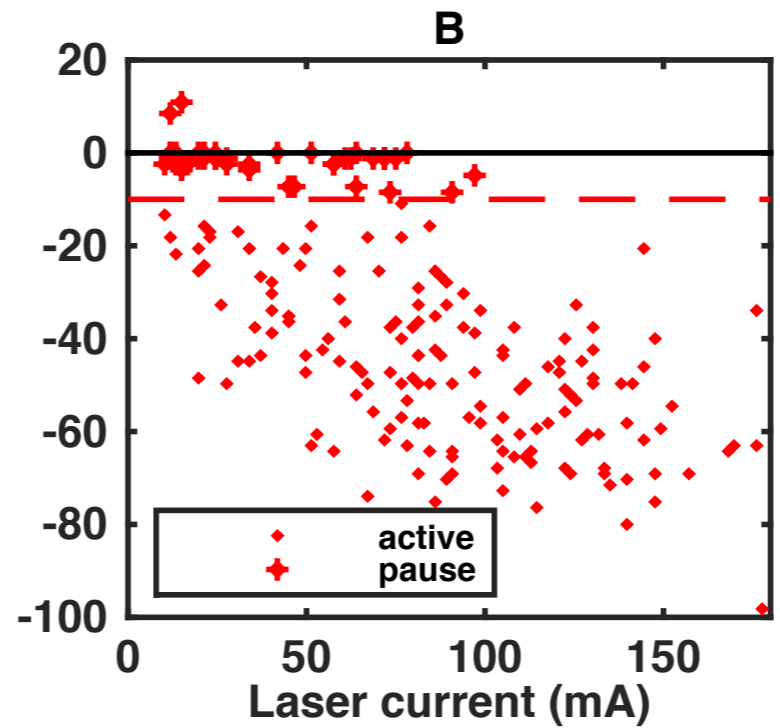
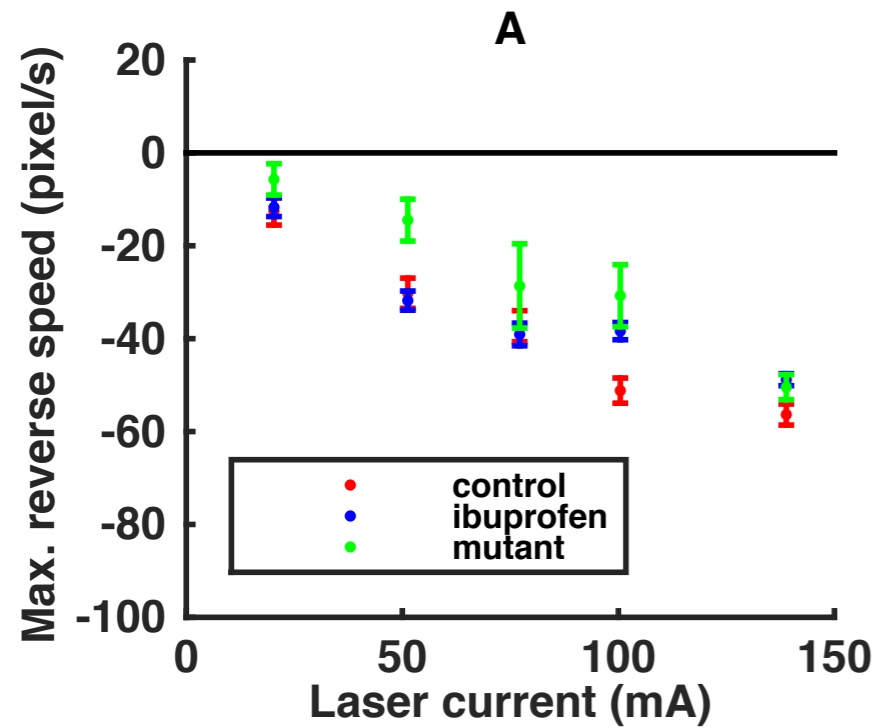
Properties of the escape behavior: Statistics of states depends on stimuli



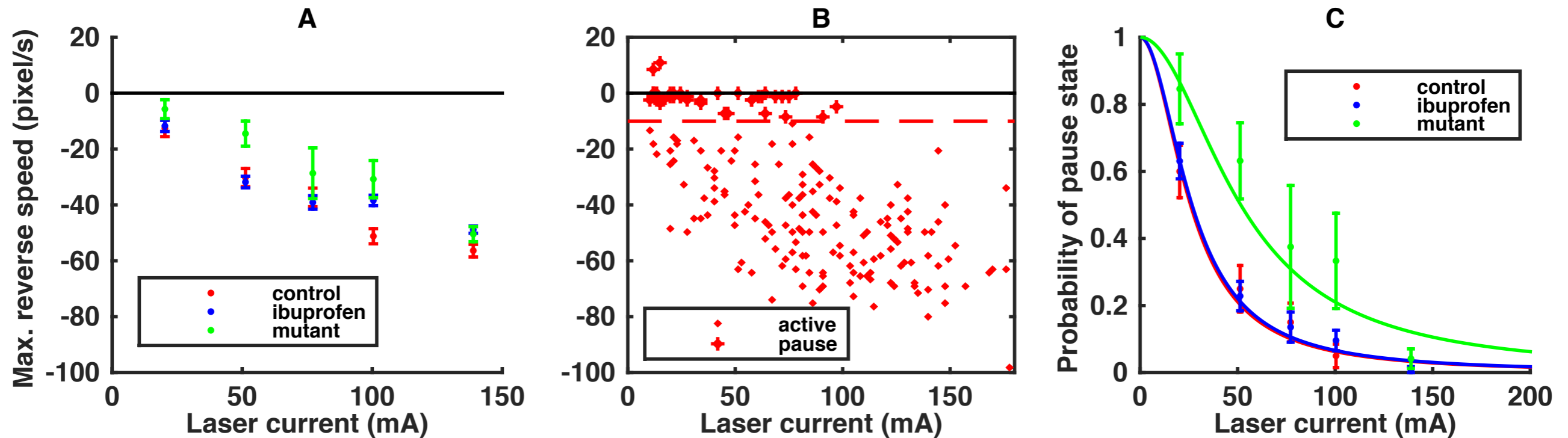
- REVERSE
- FORWARD
- PAUSE



Simple statistical characterizations are insufficient: Need whole behavior quantification



Simple statistical characterizations are insufficient: Need whole behavior quantification



- Max escape velocity shows little predictive power.
- Analysis made complicated by having paused and active escaping worms.
- Simple statistics show little power to distinguish effects of perturbations, which are visually significant.

Modeling whole behavior: Bayes to the rescue

$$P(I|\mathbf{v}) = \frac{P(\mathbf{v}|I)P(I)}{P(\mathbf{v})} = \frac{1}{Z}P(\mathbf{v}|I)P(I)$$

$$P(\mathbf{v}|I) = P(\mathbf{v}|s = p, I)P(s = p|I) + P(\mathbf{v}|s = a, I)P(s = a|I).$$

$$P(s = a|I, \mathcal{I}_0) = 1 - P(s = p|I, \mathcal{I}_0) = \frac{(I/\mathcal{I}_0)^2}{1 + (I/\mathcal{I}_0)^2}$$

Modeling whole behavior: Templated escape response

$$P(\mathbf{v}|a, I) = \frac{1}{(2\pi)^{\frac{T}{2}} |\Sigma_a|^{\frac{1}{2}}} \times \exp \left[-\frac{1}{2} (\mathbf{v} - f_{\mathcal{I}_1, \mathcal{I}_2}(I) \mathbf{u}_a)^T \Sigma_a^{-1} (\mathbf{v} - f_{\mathcal{I}_1, \mathcal{I}_2}(I) \mathbf{u}_a) \right]$$

$$f(I) \equiv f_{\mathcal{I}_1, \mathcal{I}_2}(I) = \mathcal{I}_1 + \frac{I}{1 + I/\mathcal{I}_2}$$

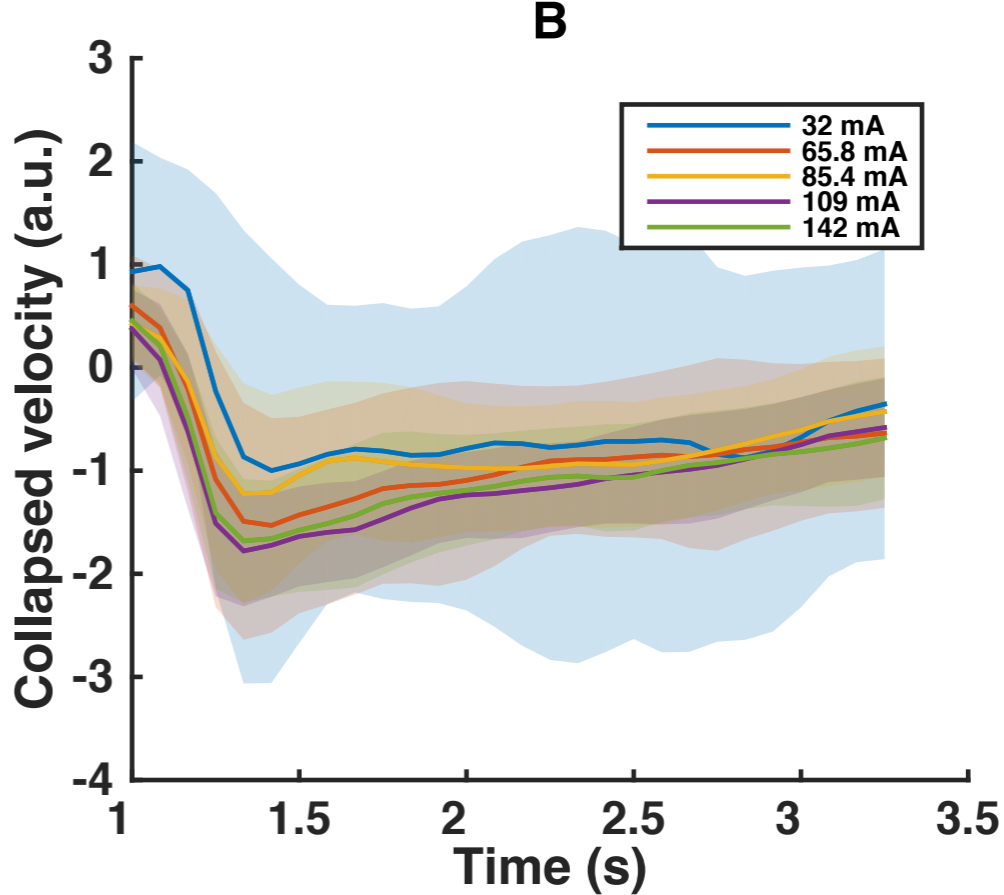
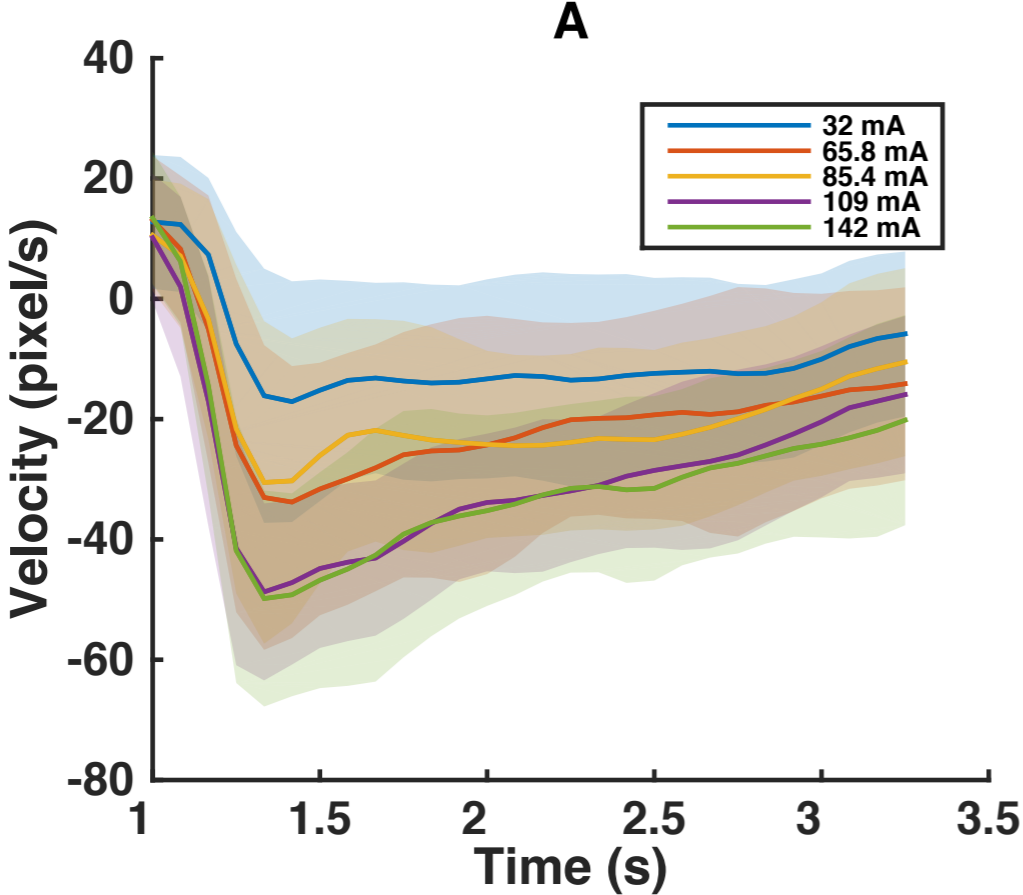
Modeling whole behavior: Templated escape response

$$P(\mathbf{v}|a, I) = \frac{1}{(2\pi)^{\frac{T}{2}} |\Sigma_a|^{\frac{1}{2}}} \times \exp \left[-\frac{1}{2} (\mathbf{v} - f_{\mathcal{I}_1, \mathcal{I}_2}(I) \mathbf{u}_a)^T \Sigma_a^{-1} (\mathbf{v} - f_{\mathcal{I}_1, \mathcal{I}_2}(I) \mathbf{u}_a) \right]$$

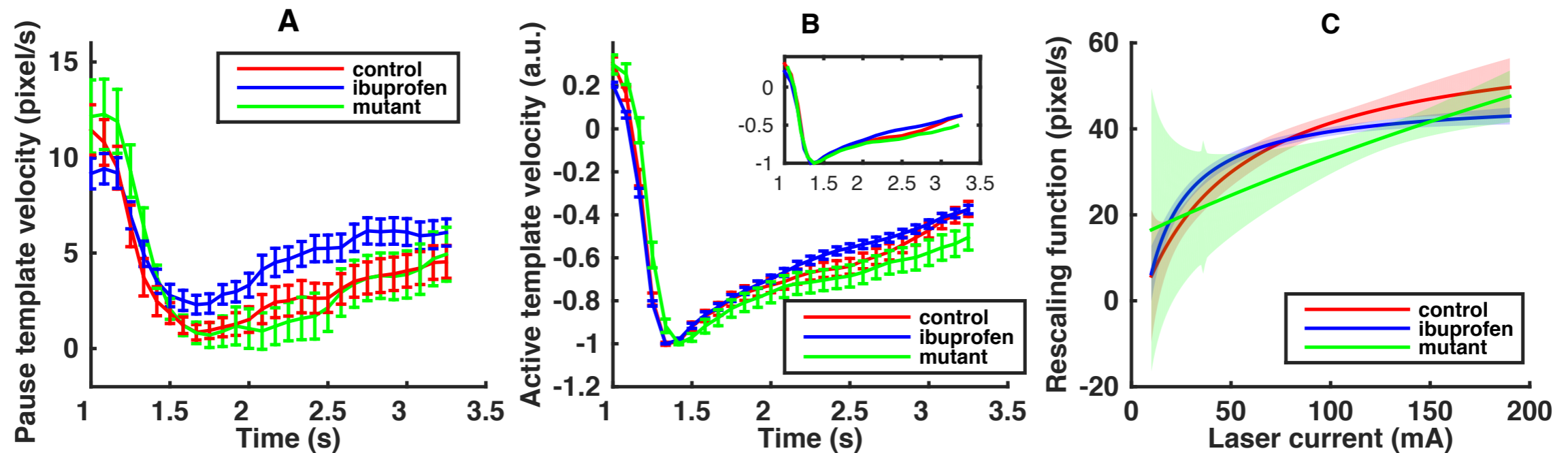
$$f(I) \equiv f_{\mathcal{I}_1, \mathcal{I}_2}(I) = \mathcal{I}_1 + \frac{I}{1 + I/\mathcal{I}_2}$$

- Similar templating (but without rescaling) for pausing.

How does the templating assumption work?

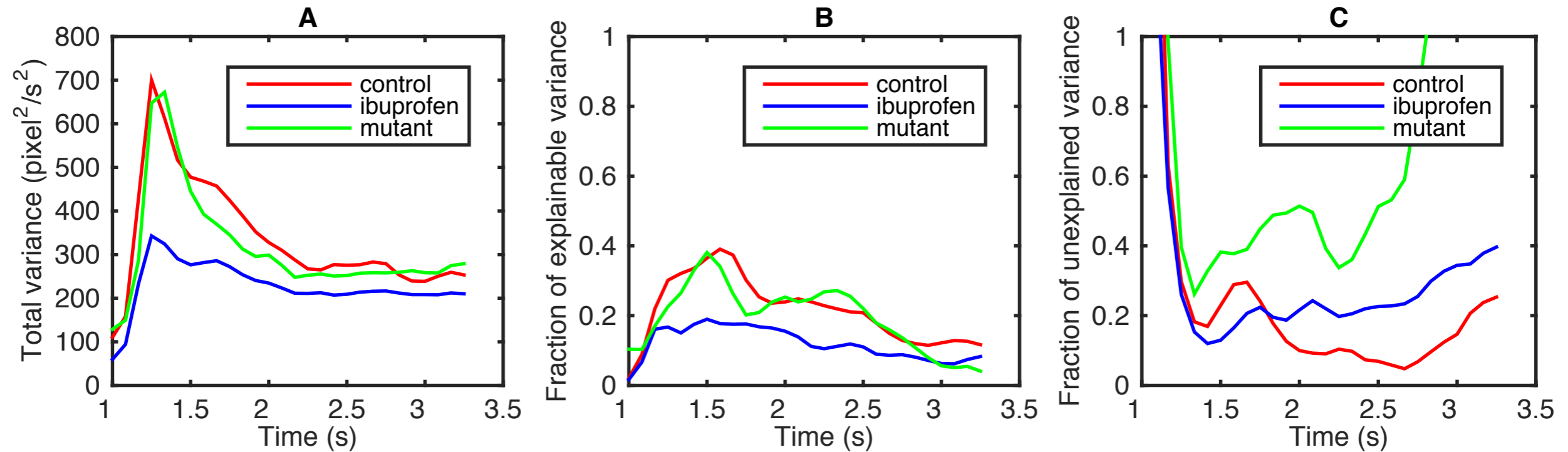


The templates and the scaling functions



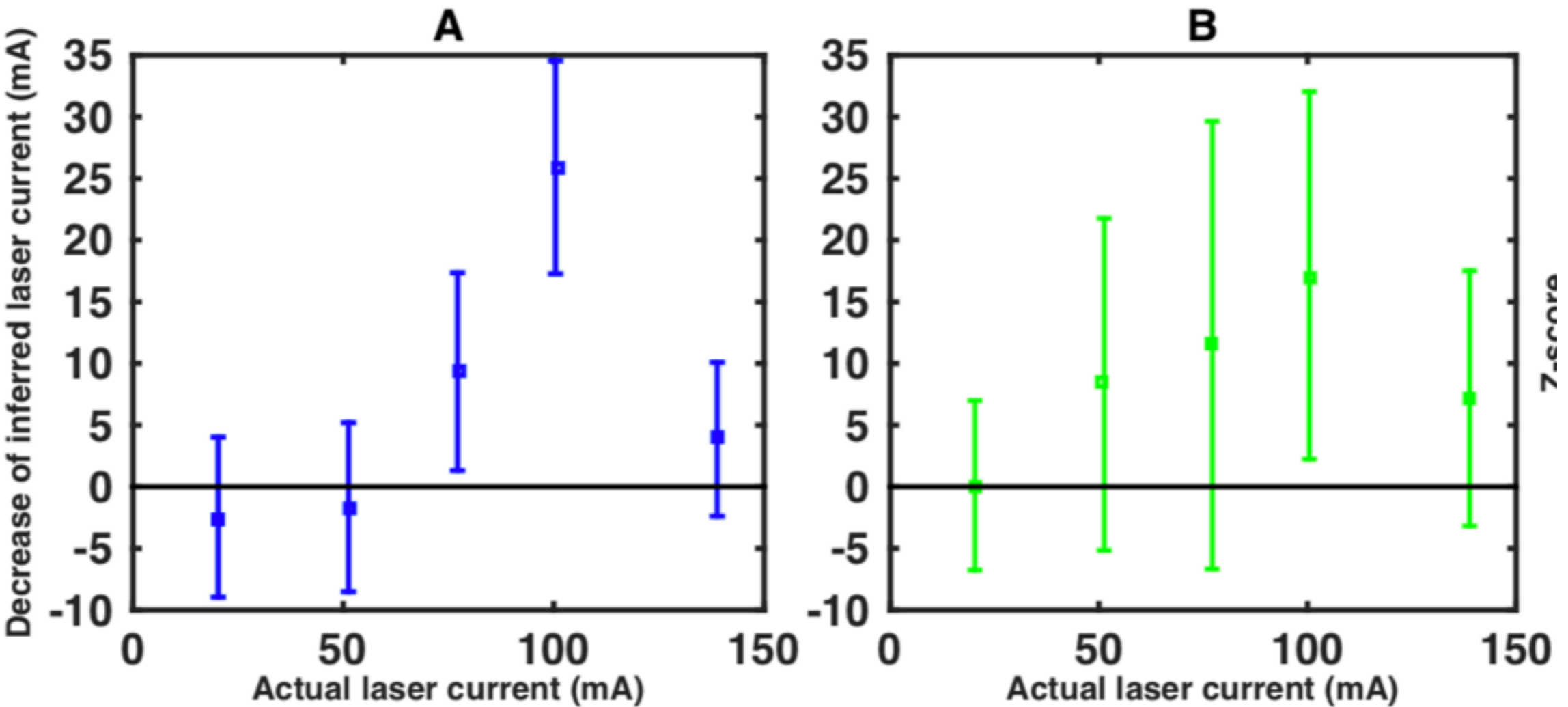
- Ibuprofen treatment does not change the template, and only changes parameters of the scaling function.
- Mutant changes the template and changes the shape of the scaling function!

Explanatory power of the template model



- The model has a much worse explanatory power for the mutant. Mutant changes the worm behavior **qualitatively**.
- Conclusion: **mutant is not an analgesic/sensory mutant only.**

Can quantify analgesic effects



Summary I:

- Can relate stimulus to behavior in a quantitative model.
 - Can quantify analgesic effects.
 - Can say if the effect is more than analgesic.
-
- But: can we automate the modeling?

Can we infer the equations governing the worm's dynamics?



Can we infer the equations governing the worm's dynamics?



First steps for dynamics (and keep the worm in mind)

- Assume that dynamics of cellular networks is **temporally local**, given by ordinary differential equations.
 - Do not fit curves; **fit dynamics**.
- Neglect stochasticity, and spatial structure for now

$$\begin{cases} \frac{dx_1}{dt} = f_1(x_1, x_2, \dots, x_n) \\ \dots \\ \frac{dx_n}{dt} = f_n(x_1, x_2, \dots, x_n) \end{cases}$$

First steps for dynamics (and keep the worm in mind)

- Assume that dynamics of cellular networks is **temporally local**, given by ordinary differential equations.
 - Do not fit curves; **fit dynamics**.

- Neglect stochasticity, and spatial structure for now

$$\begin{cases} \frac{dx_1}{dt} = f_1(x_1, x_2, \dots, x_n) \\ \dots \\ \frac{dx_n}{dt} = f_n(x_1, x_2, \dots, x_n) \end{cases}$$

- Can we automatically fit these functions f_i from data?
 - How do we enumerate the set of all possible multivariate functions?
 - How do we search through this list?
 - How do we not overfit?

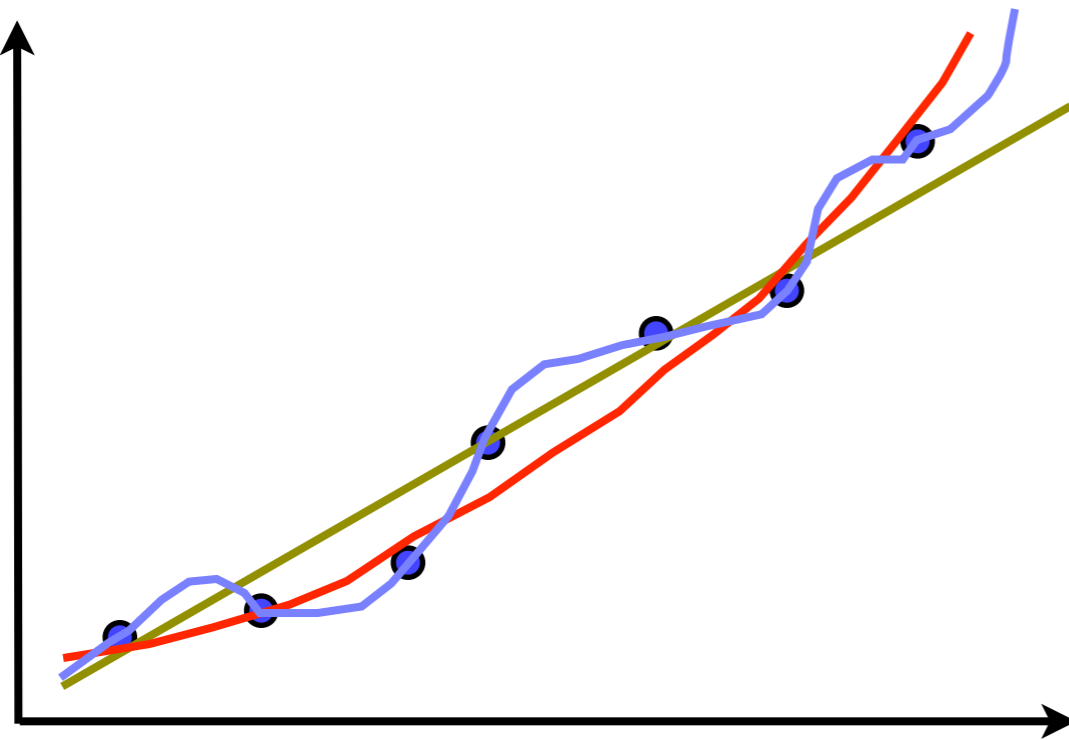
The field (and we) has worked on this problem for a while

- Bottom-up methods, reducing a known microscopic, mechanistic network (not the easiest path).
- Top-down methods that build **phenomenological models** from data directly, and without reconstructing a mechanistic network as an intermediate step
 - Crutchfield and McNamara, *Compl Syst* 1987
 - Voit et al, *Theor Biol Med Model* 2006
 - Lillacci and Khammash, *PLoS CB* 2010
 - Munsky, et al., *MSB* 2009, *Science* 2013
 - *Lipson et al., Science* 2009, *Phys Biol* 2011 - *EUREQA*
 - *Brunton et al., 2016 - SINDY*
 - AutomatedStatistician.com

Can we avoid the exhaustive search and the need for the correct basis?

Can we avoid the exhaustive search and the need for the correct basis?

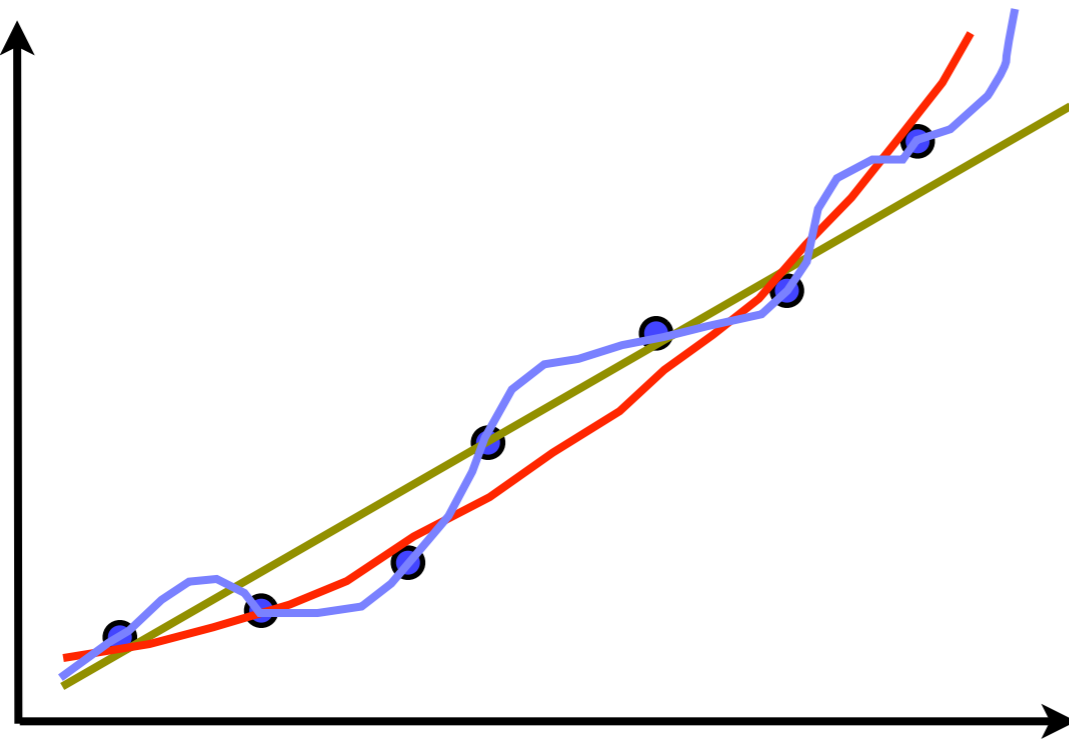
- Do not need an exhaustive search or exact fits when fitting dimensional curves with progressively increasing complexity



$$y_K(x) = \sum_{k=1}^K A_k x^k + \text{noise}$$

Can we avoid the exhaustive search and the need for the correct basis?

- Do not need an exhaustive search or exact fits when fitting dimensional curves with progressively increasing complexity

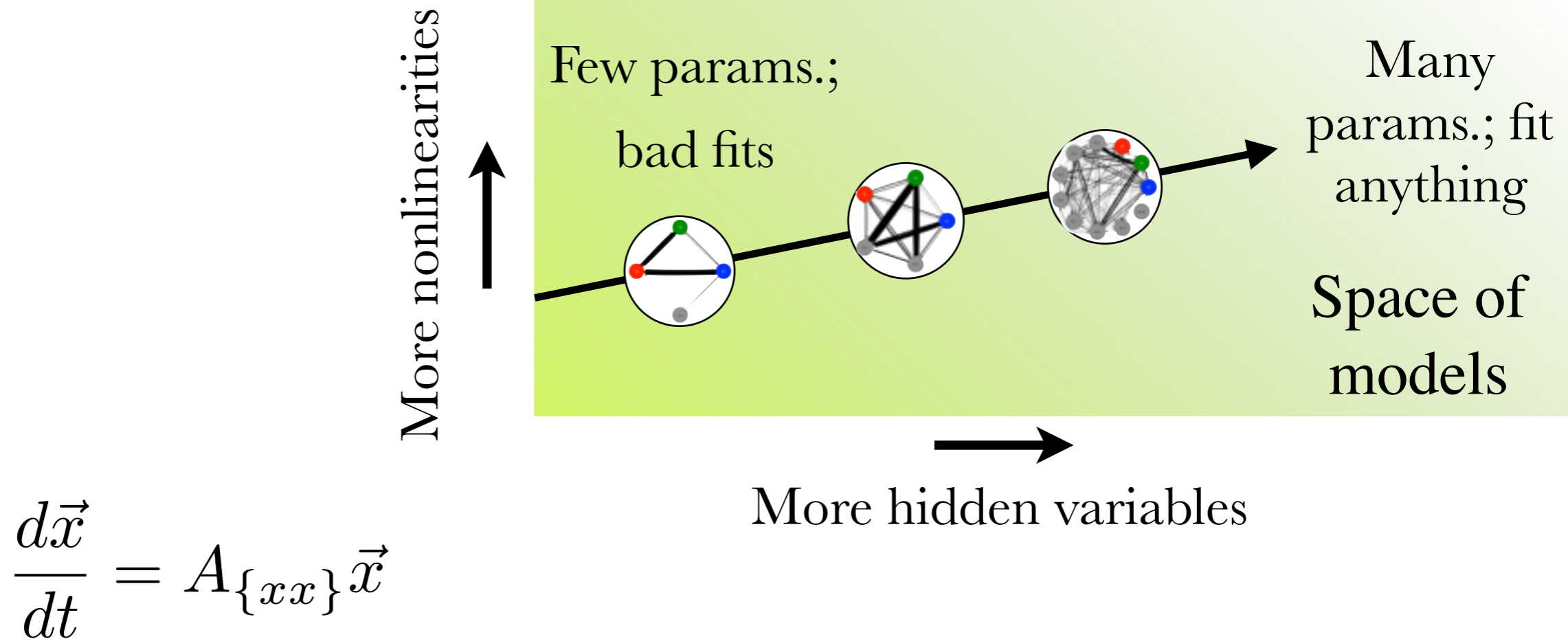


$$y_K(x) = \sum_{k=1}^K A_k x^k + \text{noise}$$

- Use **nested, complete** model families, e.g., Taylor series.
- Use Bayesian model selection to limit the complexity of the search space (the value of maximum K).

Schwartz, *Ann Stat* 1978; MacKay, *Neural Comp*, 1992
Balasubramanian, *Neural Comp*, 1996; Nemenman, *Neural Comp*, 2005

Why is fitting dynamics so hard?

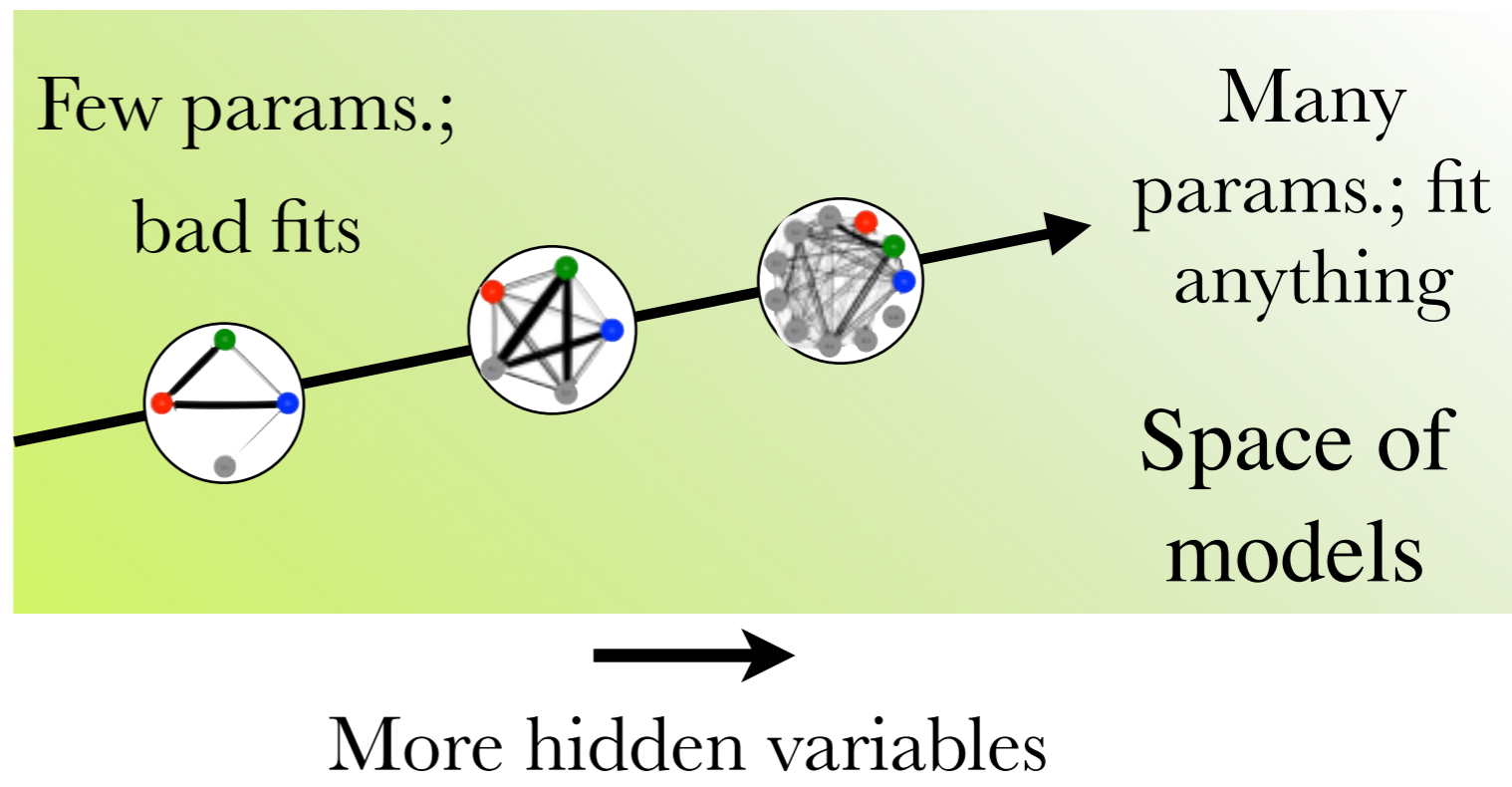


Why is fitting dynamics so hard?

$$\frac{d\vec{x}}{dt} = A_{\{xx\}}\vec{x} + A_{\{xx\}}^{(2)}\vec{x} \odot \vec{x} + \dots + A_{\{xx\}}^{(K)}\vec{x} \odot \dots \odot \vec{x}$$

More nonlinearities

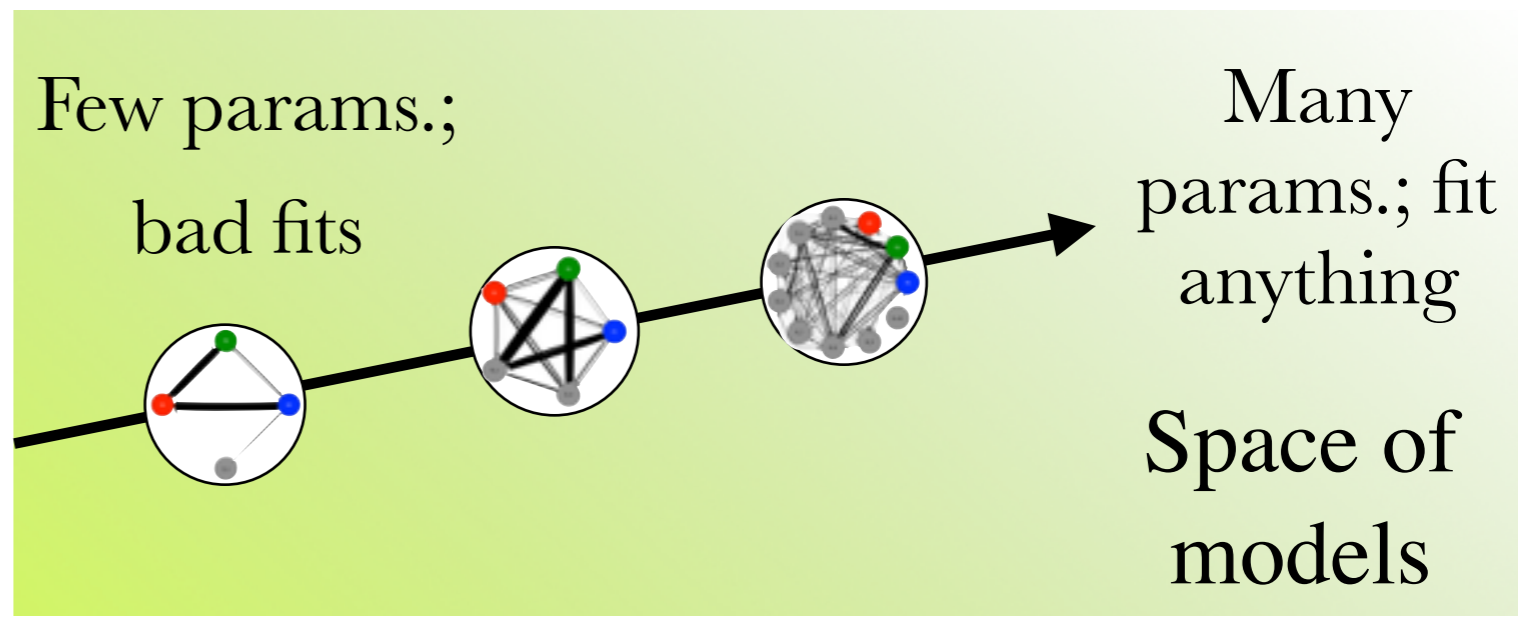
$$\frac{d\vec{x}}{dt} = A_{\{xx\}}\vec{x}$$



Why is fitting dynamics so hard?

$$\frac{d\vec{x}}{dt} = A_{\{xx\}}\vec{x} + A_{\{xx\}}^{(2)}\vec{x} \odot \vec{x} + \dots + A_{\{xx\}}^{(K)}\vec{x} \odot \dots \odot \vec{x}$$

More nonlinearities



$$\frac{d\vec{x}}{dt} = A_{\{xx\}}\vec{x}$$

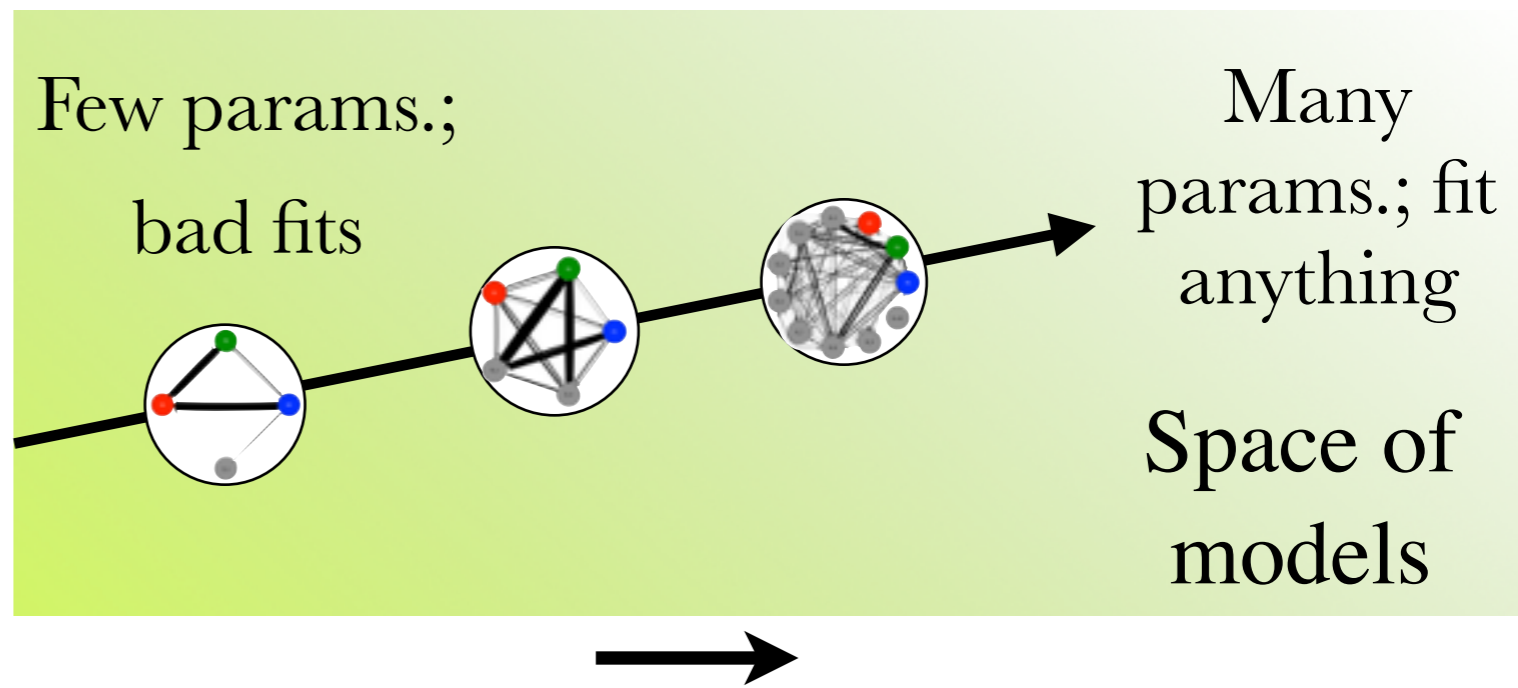
More hidden variables

$$\begin{cases} \frac{d\vec{x}}{dt} = A_{\{xx\}}\vec{x} + B_{\{x\}1}\xi_1 + \dots + B_{\{x\}K}\xi_K \\ \frac{d\xi_1}{dt} = A_{1\{x\}}\vec{x} + B_{11}\xi_1 + \dots + B_{1K}\xi_K \\ \dots \\ \frac{d\xi_K}{dt} = A_{K\{x\}}\vec{x} + B_{1K}\xi_1 + \dots + B_{KK}\xi_K \end{cases}$$

Why is fitting dynamics so hard?

$$\frac{d\vec{x}}{dt} = A_{\{xx\}}\vec{x} + A_{\{xx\}}^{(2)}\vec{x} \odot \vec{x} + \dots + A_{\{xx\}}^{(K)}\vec{x} \odot \dots \odot \vec{x}$$

More nonlinearities



$$\frac{d\vec{x}}{dt} = A_{\{xx\}}\vec{x}$$

More hidden variables

$$\begin{cases} \frac{d\vec{x}}{dt} = A_{\{xx\}}\vec{x} + B_{\{x\}1}\xi_1 + \dots + B_{\{x\}K}\xi_K \\ \frac{d\xi_1}{dt} = A_{1\{x\}}\vec{x} + B_{11}\xi_1 + \dots + B_{1K}\xi_K \\ \dots \\ \frac{d\xi_K}{dt} = A_{K\{x\}}\vec{x} + B_{1K}\xi_1 + \dots + B_{KK}\xi_K \end{cases}$$

- Hidden degrees of freedom and nonlinearities breaks nestedness -- no consistency.
- Choose any (reasonable) **complete** path through the model space
 - Good choice — good fits with little data; Bad choice — not worse than exhaustive search.

Two types of model families

- Both nested and complete.
- Account for nonlinearities **and** hidden variables.
- Biologically reasonable.

Daniels and Nemenman, Nature Comm 2015; PLoS ONE 2015

Two types of model families

- Both nested and complete.
- Account for nonlinearities **and** hidden variables.
- Biologically reasonable.

Sigmoidal
recurrent
networks

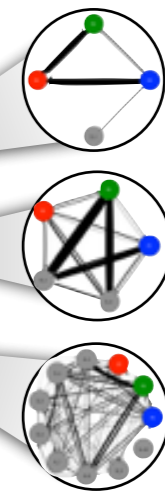
Daniels and Beer,
arXiv 2010

Degradation Interactions Input

▼ ▼ ▼

$$\frac{dx_i}{dt} = -x_i/\tau_i + \sum_{j=1}^J W_{ij} \xi(x_j + \theta_j) + \sum_{k=1}^K V_{ik} I_k$$

with $\xi(y) = 1/(1 + e^{-y})$



Daniels and Nemenman, Nature Comm 2015; PLoS ONE 2015

Two types of model families

- Both nested and complete.
- Account for nonlinearities **and** hidden variables.
- Biologically reasonable.

Sigmoidal recurrent networks

Daniels and Beer, arXiv 2010

Degradation
Interactions
Input

$$\frac{dx_i}{dt} = -x_i/\tau_i + \sum_{j=1}^J W_{ij} \xi(x_j + \theta_j) + \sum_{k=1}^K V_{ik} I_k$$

with $\xi(y) = 1/(1 + e^{-y})$

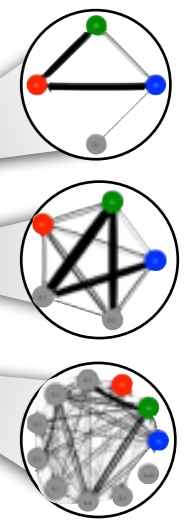


S-systems

Savageau et al., 1976-...

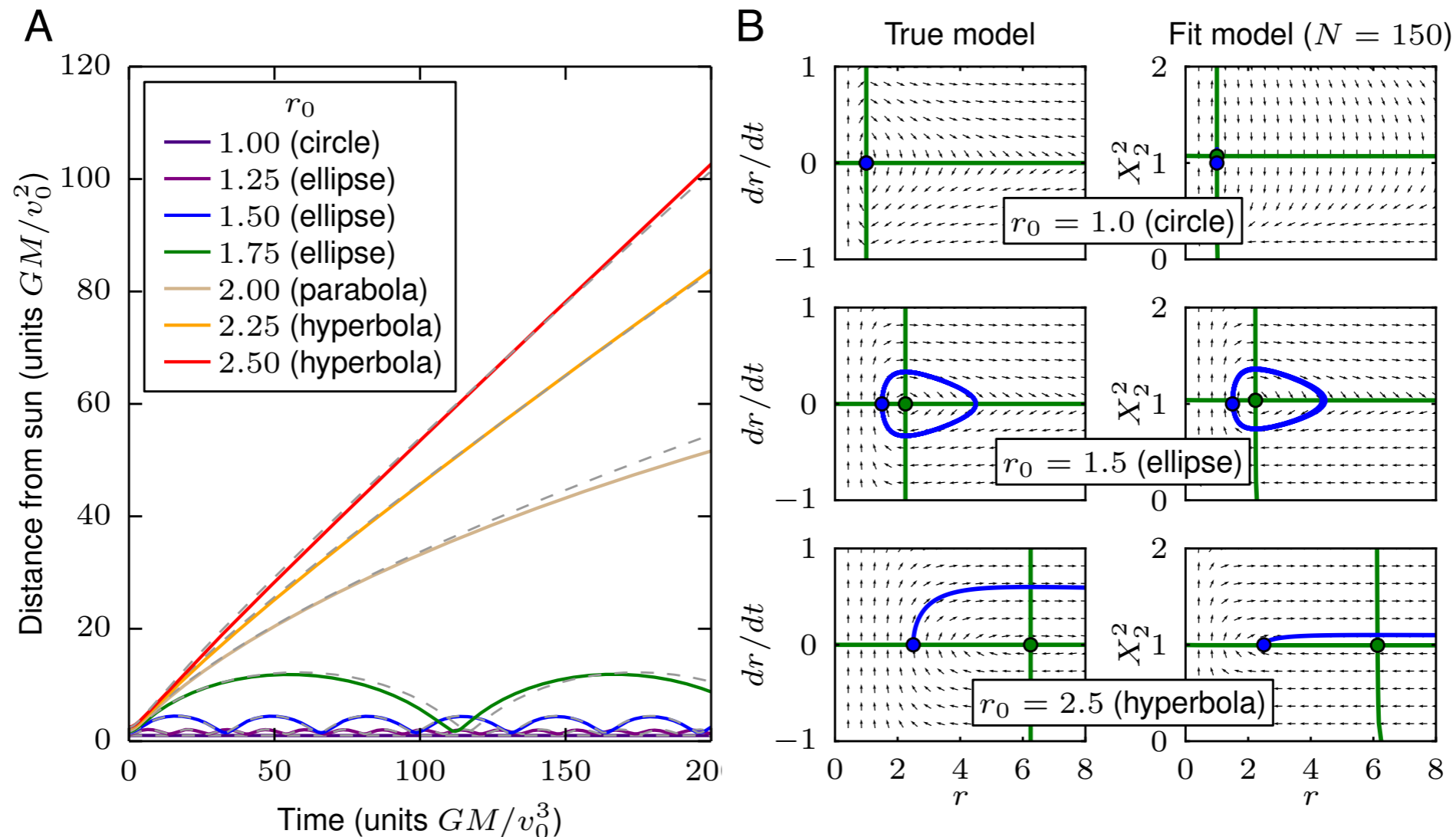
Interactions and input dependence

$$\frac{dx_i}{dt} = A_i \prod_j x_j^{\alpha_{ij}} \prod_k I_k^{a_{ik}} - B_i \prod_j x_j^{\beta_{ij}} \prod_k I_k^{b_{ik}}$$



Daniels and Nemenman, Nature Comm 2015; PLoS ONE 2015

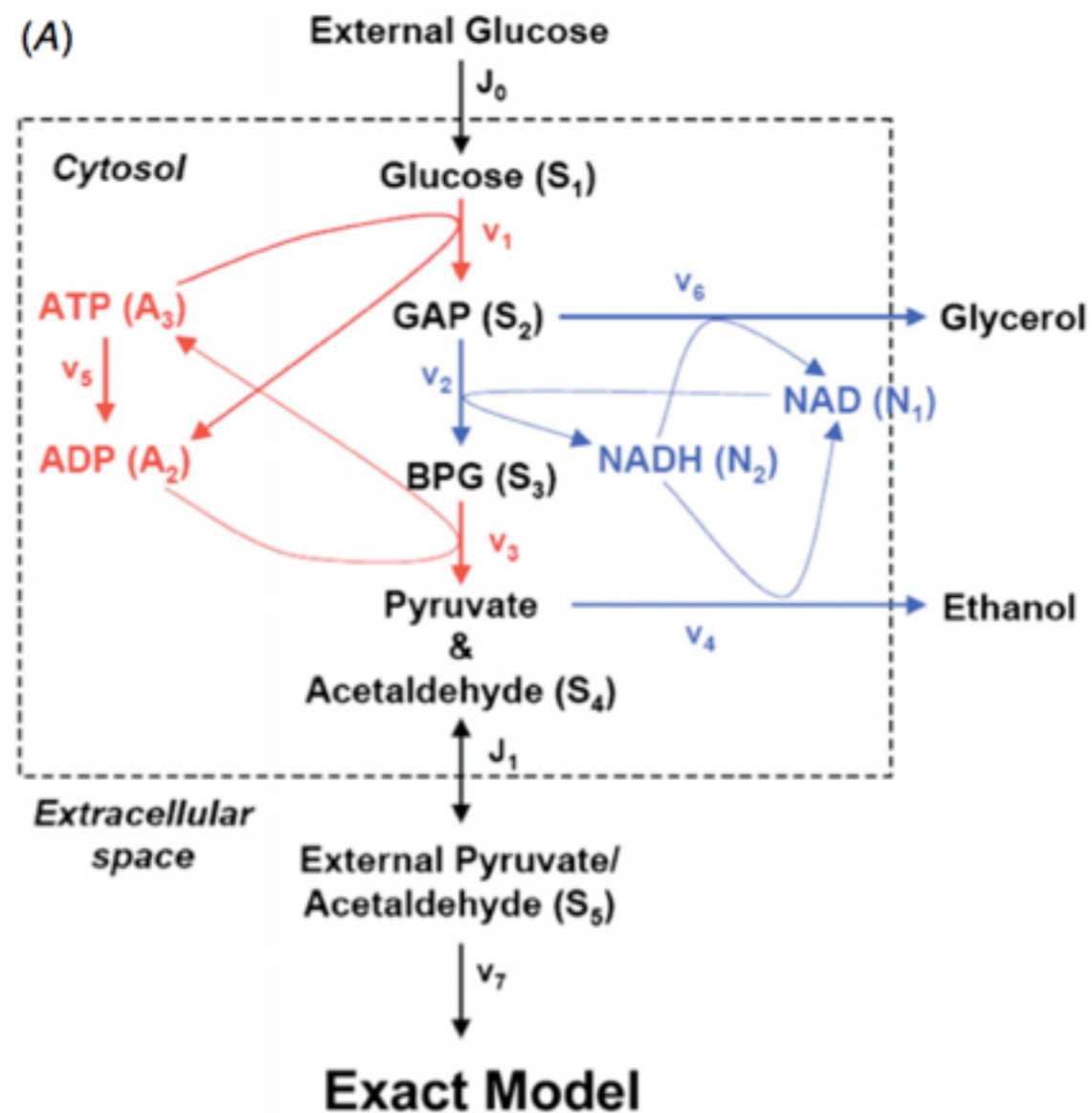
Finding laws that we already know: An automated Sir Isaac (*Sirlsaac on GitHub*)



- Finds the hidden variable needed to account for the Newton's laws.
- Accounts for different classes of trajectories.

Daniels and Nemenman, Nature Comm 2015, PLoS ONE 2015

Test Model: Yeast Glycolytic Oscillator

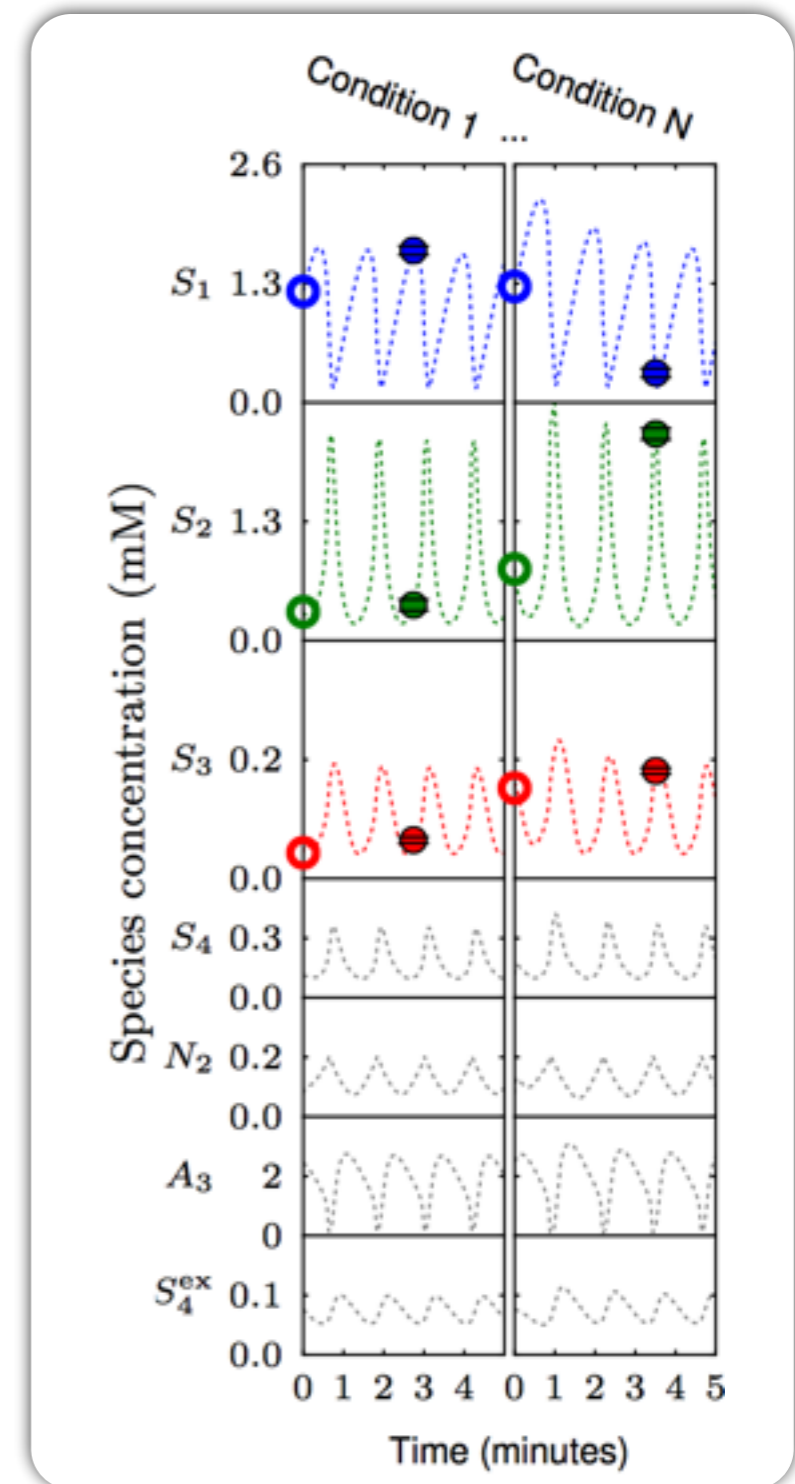


- 7 species, 28 parameters
- Complex rational dynamical laws

Ruoff et al., 2003

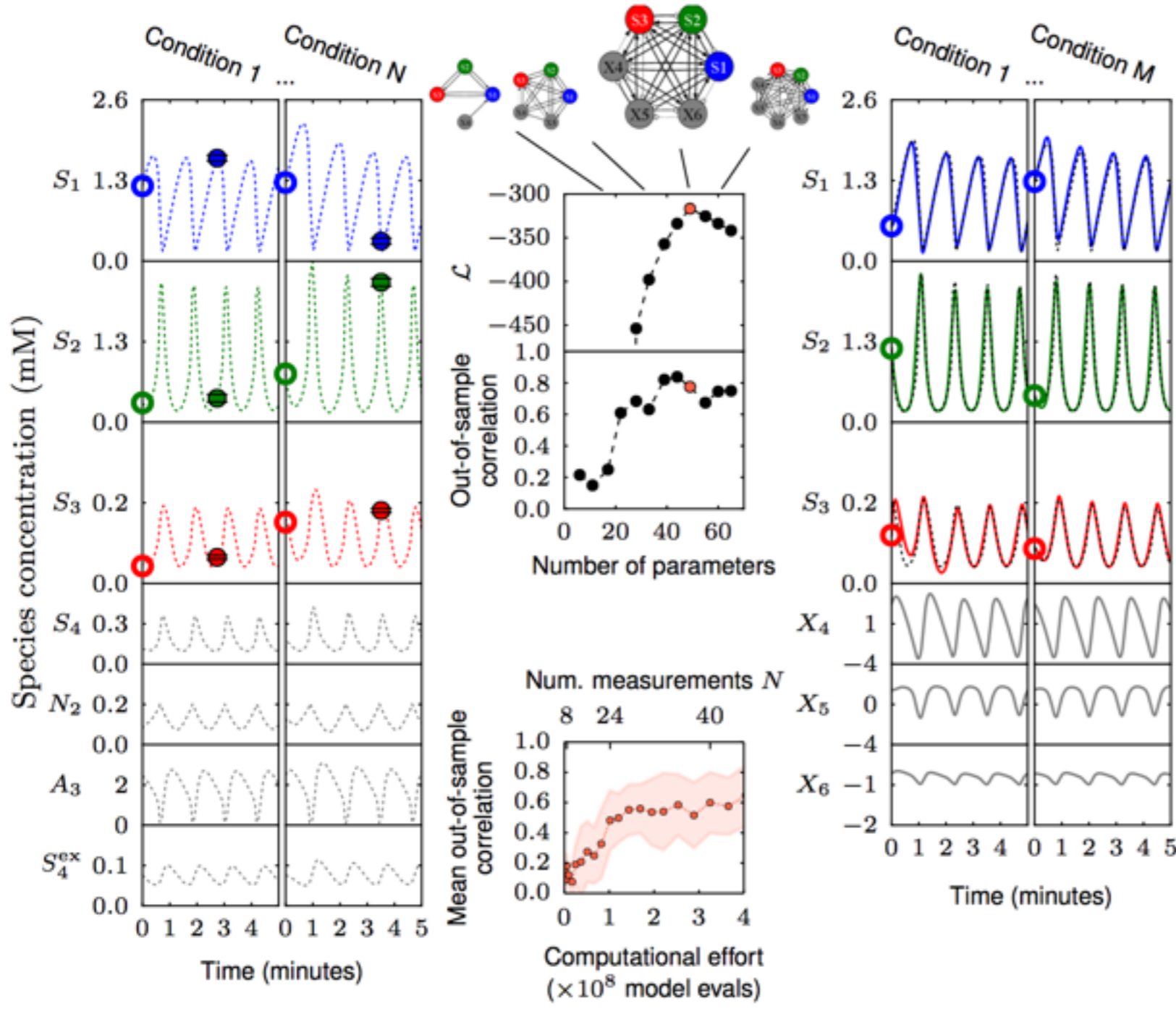
The yeast glycolytic oscillations: Complex dynamics needing complex structure

- Observe only 3/7 of variables; add 10% noise.
- Data: N samples of structure
 - Initial condition of the 3 species;
 - Some random time later;
 - The value of these 3 species at that time.



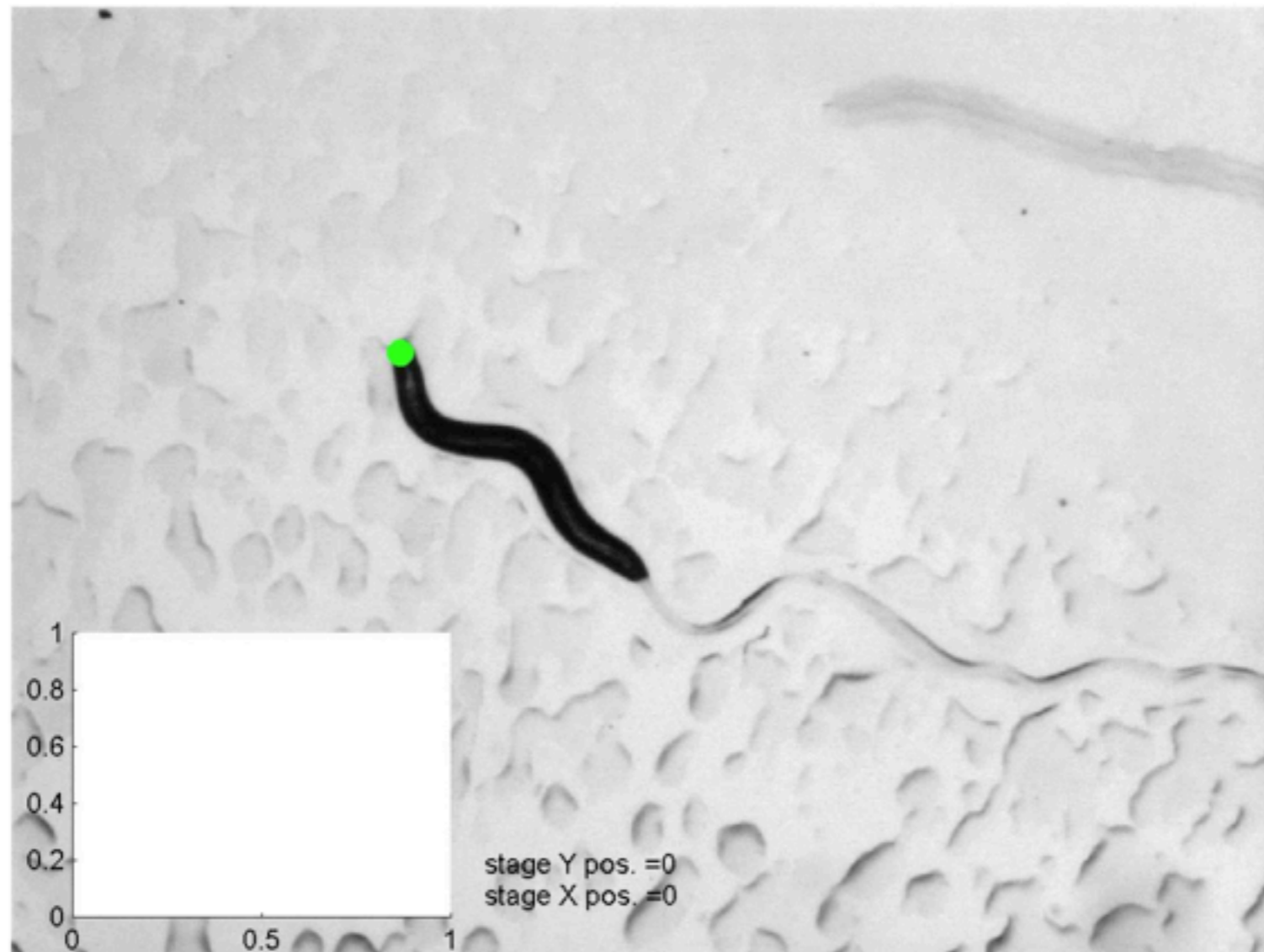
Results

Daniels and Nemenman, Nature Comm 2015; PLoS ONE 2015

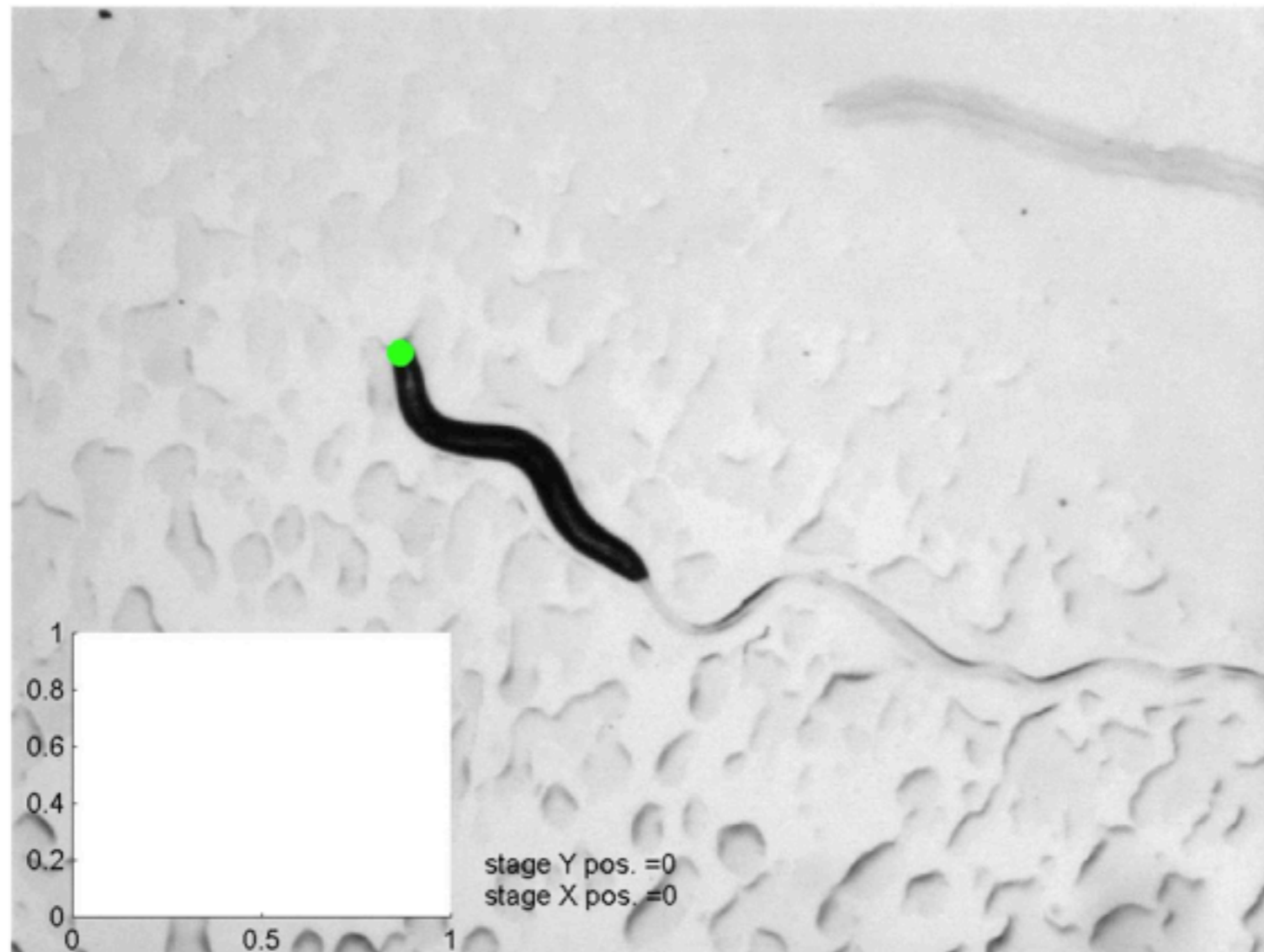


- ~100x fewer evaluations for the same accuracy compared to full search.
- ~1000x fewer data points than full search.
- Better accuracy than curve fitting.
- Linear scaling with the amount of data and with the number of variables.
- Finds conservation laws.

Modeling *C. elegans* temperature nociception escape response

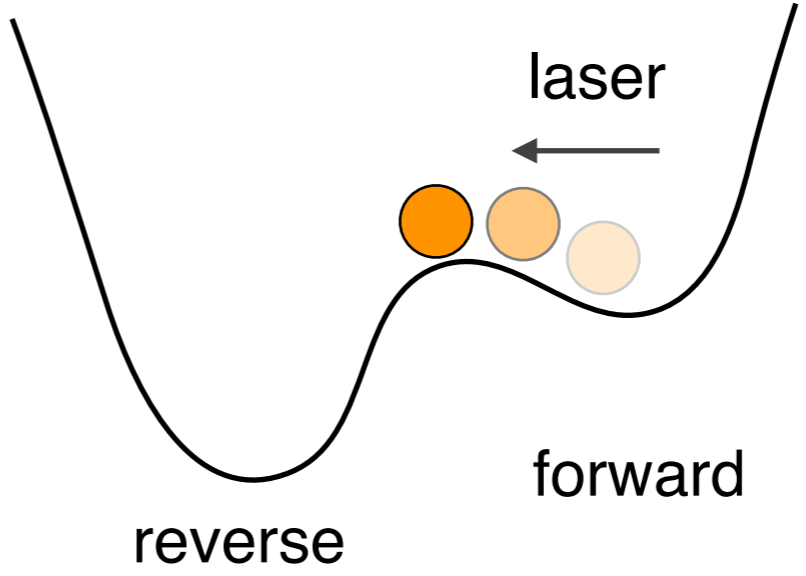


Modeling *C. elegans* temperature nociception escape response

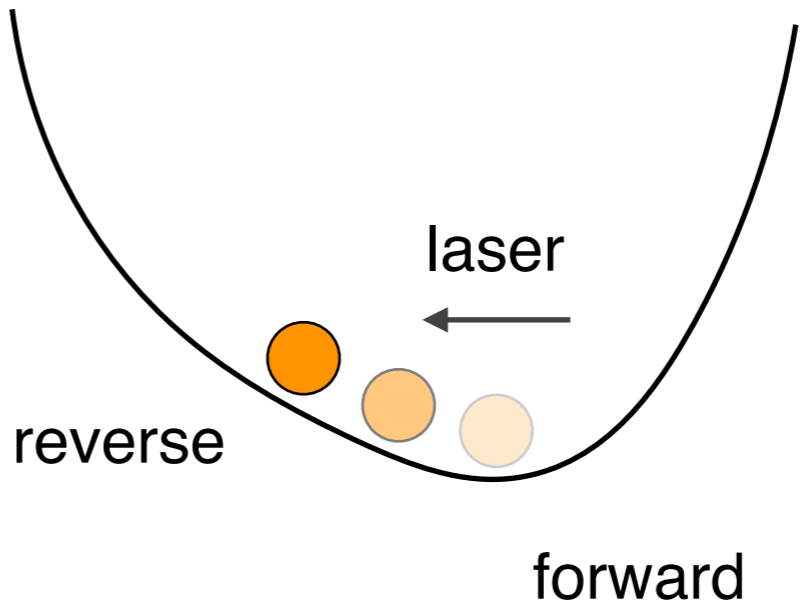


Question: what is the logic of this behavior?

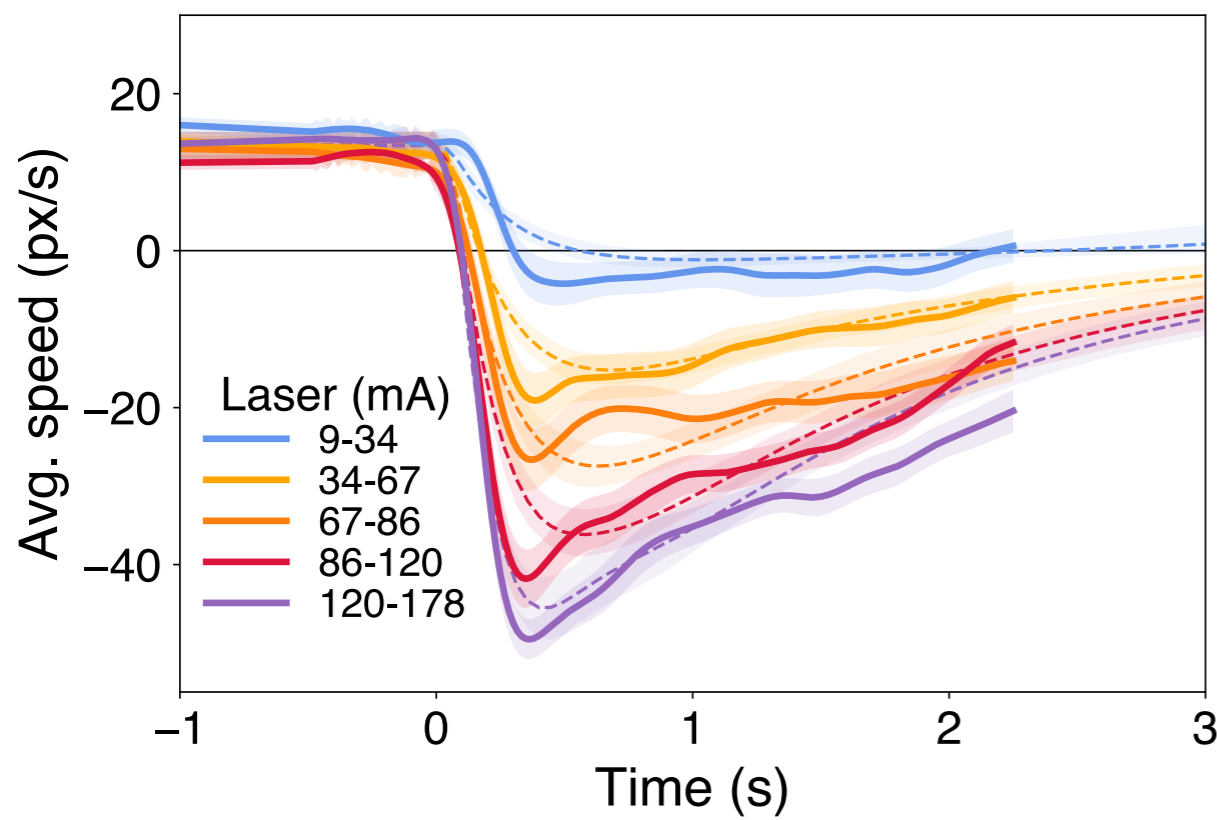
Distinct behavioral attractors



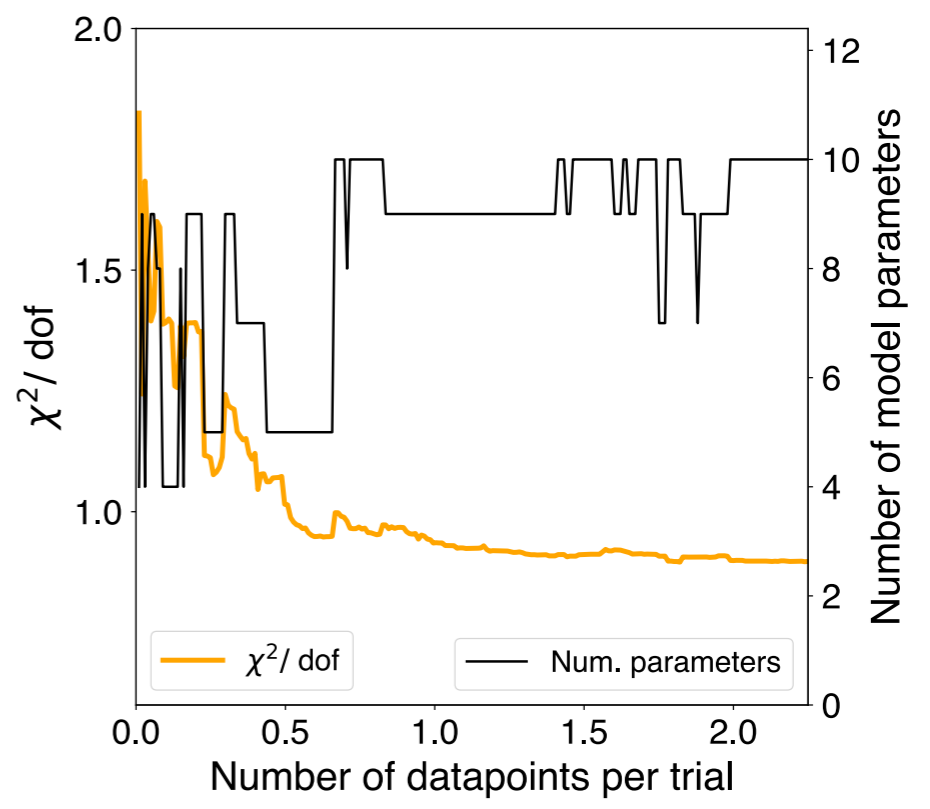
Continuously varying behavior



Worm nociception: Data and fits



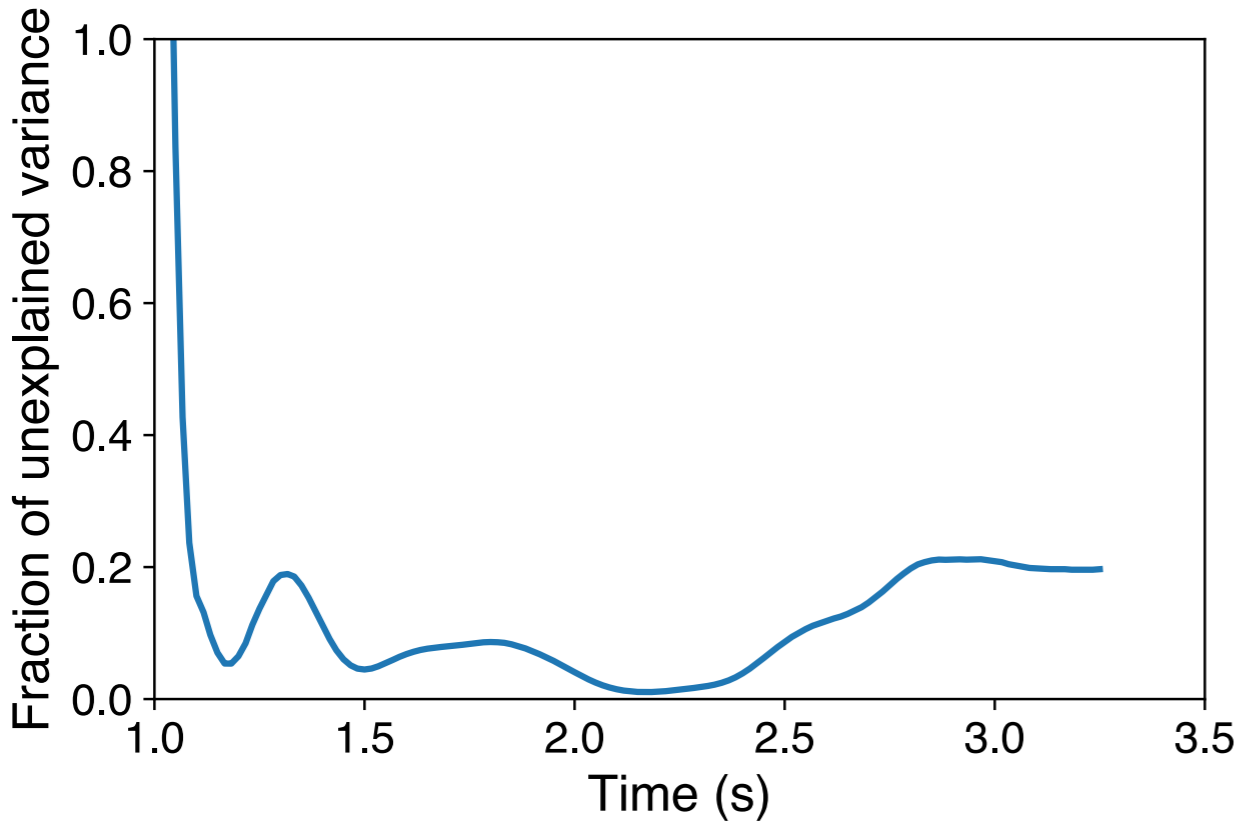
SGD style training



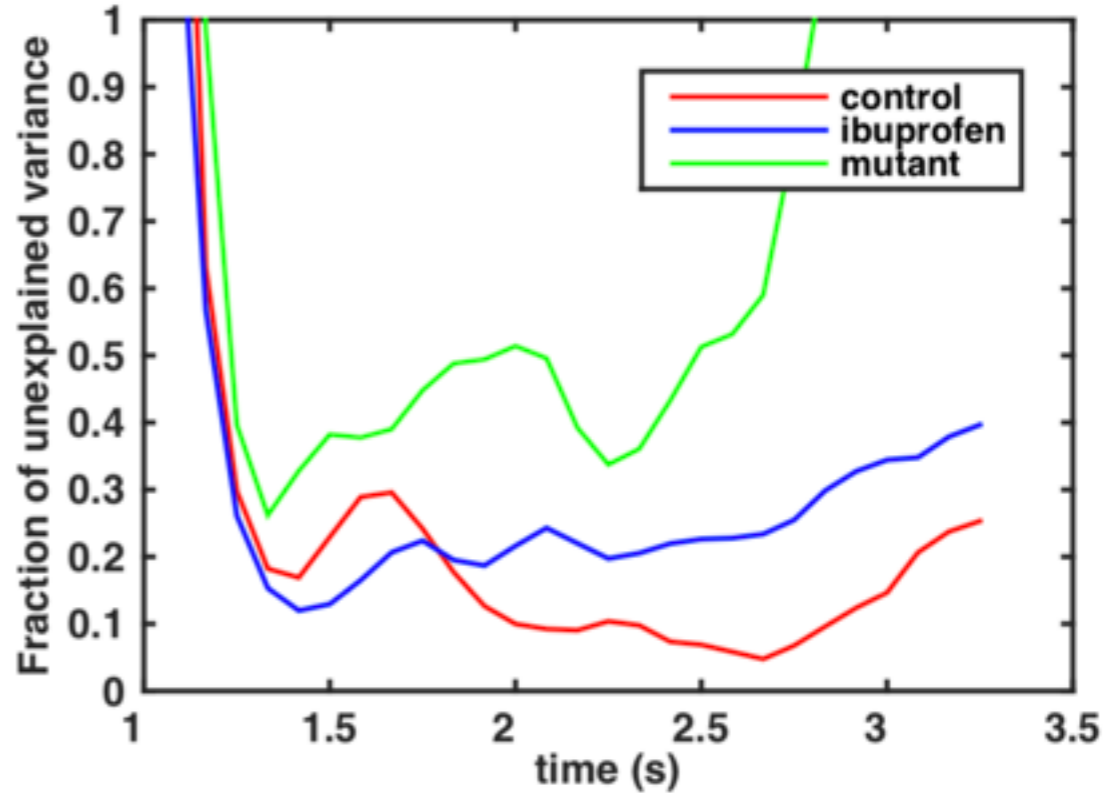
Daniels, Ryu, IN, in prep.

Comparing to hand-curated model

Sir Isaac model

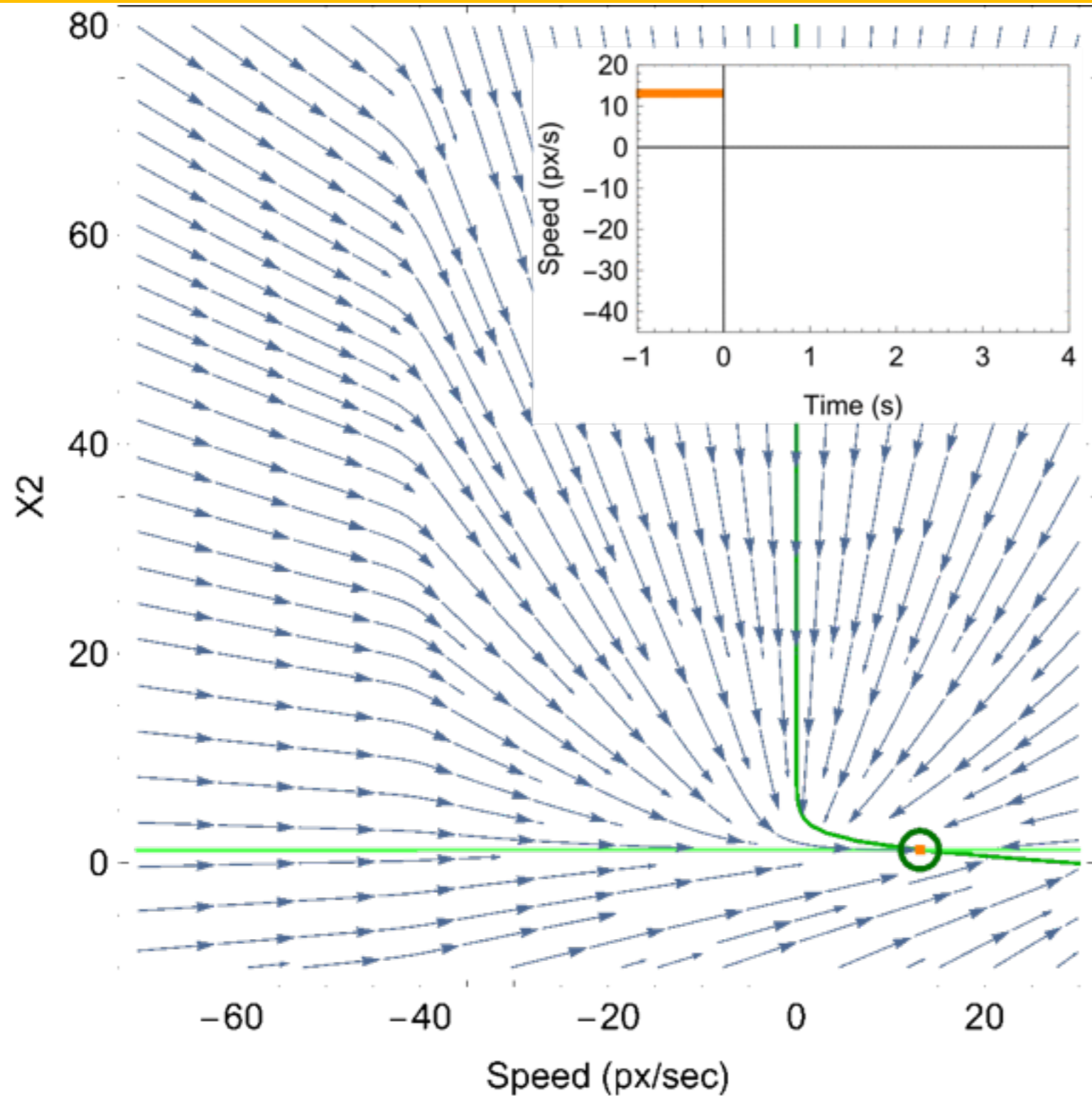


Curated model



What is the logic of the behavioral dynamics?

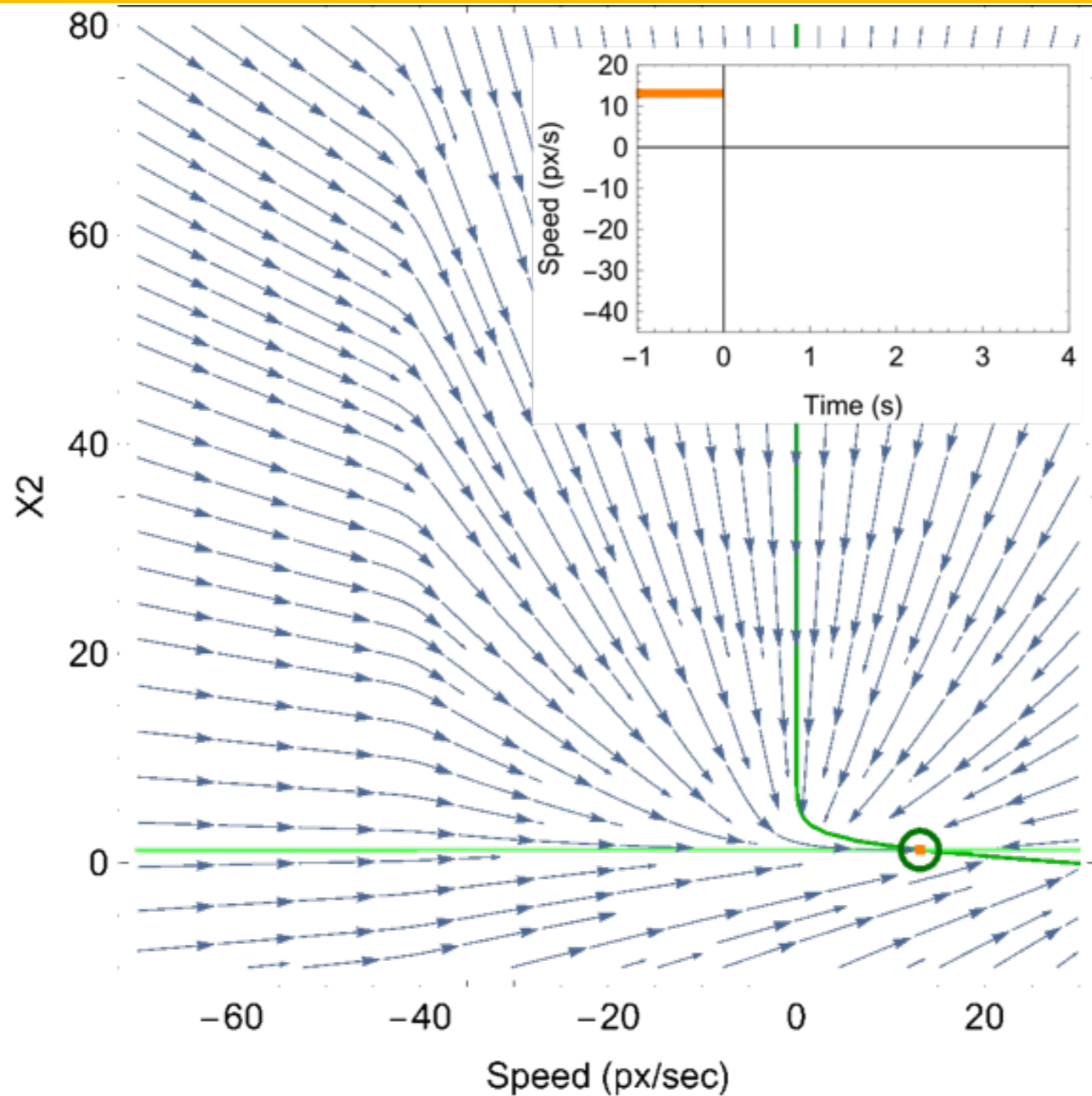
A single attractor



Continuously varying
behavioral attractor

What is the logic of the behavioral dynamics?

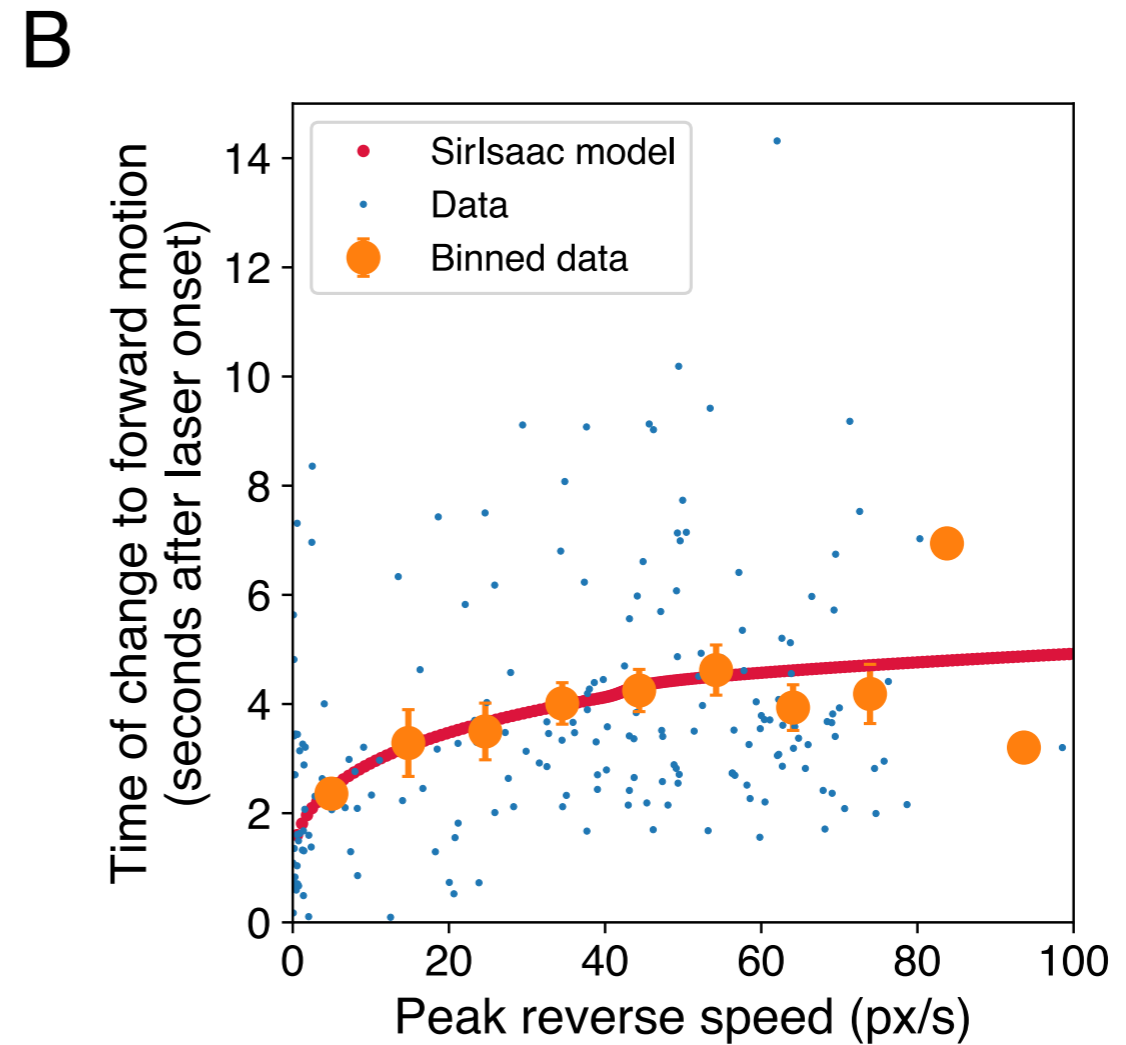
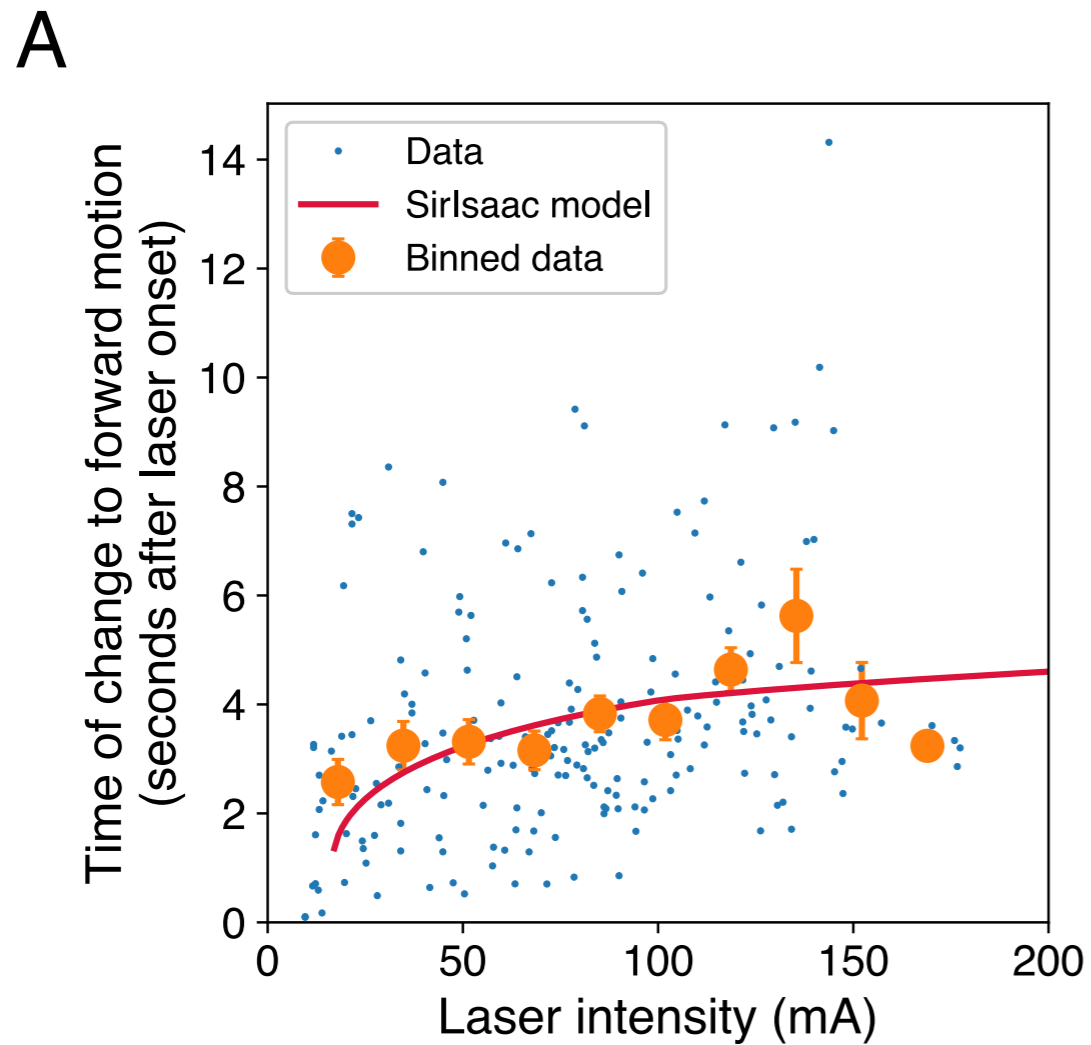
A single attractor



Continuously varying
behavioral attractor

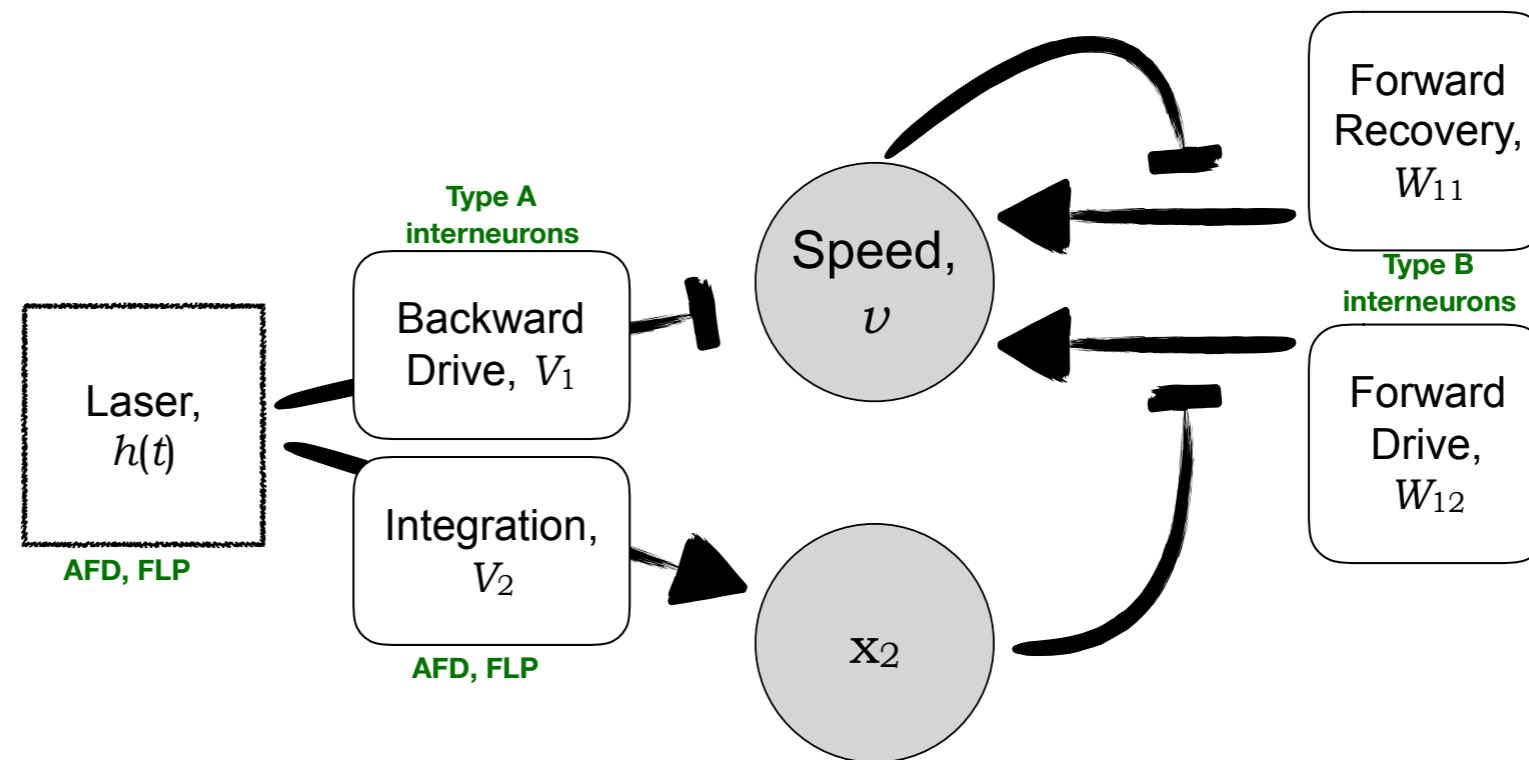
Why should you believe our model?

The predictions



Interpreting the model

$$\frac{dv}{dt} = -\frac{v}{\tau_1} + V_1 h(t) + \frac{W_{11}}{1 + e^{v+\theta_1}} + \frac{W_{12}}{1 + e^{x_2}},$$
$$\frac{dx_2}{dt} = -x_2 + V_2 h(t).$$



Summary II

Summary II

- Deemphasize “truth” of a model, focus on accuracy of prediction and on **refining** (not coarse-graining) **phenomenological** dynamics.
- **Complete, nested** model families of dynamics allow to use Bayesian model selection to adapt model complexity.
- Such phenomenological models make accurate and interpretable predictions, at least for the worms.
- Why do this? Go from **models** to phenomenological **theories**
 - Repeat Hookean approach in biology: build effective models of similar systems and look for patterns (e.g., chemotaxis in *C. elegans* and *E. coli*).