



Automated, predictive, and interpretable inference of *Caenorhabditis elegans* escape dynamics

Bryan C. Daniels^a, William S. Ryu^{b,c}, and Ilya Nemenman^{d,e,f,1}

^aArizona State University–Santa Fe Institute Center for Biosocial Complex Systems, Arizona State University, Tempe, AZ 85281; ^bDepartment of Physics, University of Toronto, Toronto, ON M5S 1A7, Canada; ^cThe Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada; ^dDepartment of Physics, Emory University, Atlanta, GA 30322; ^eDepartment of Biology, Emory University, Atlanta, GA 30322; and ^fInitiative in Theory and Modeling of Living Systems, Emory University, Atlanta, GA 30322

Edited by David A. Weitz, Harvard University, Cambridge, MA, and approved February 22, 2019 (received for review September 25, 2018)

The roundworm *Caenorhabditis elegans* exhibits robust escape behavior in response to rapidly rising temperature. The behavior lasts for a few seconds, shows history dependence, involves both sensory and motor systems, and is too complicated to model mechanistically using currently available knowledge. Instead we model the process phenomenologically, and we use the *Sir Isaac* dynamical inference platform to infer the model in a fully automated fashion directly from experimental data. The inferred model requires incorporation of an unobserved dynamical variable and is biologically interpretable. The model makes accurate predictions about the dynamics of the worm behavior, and it can be used to characterize the functional logic of the dynamical system underlying the escape response. This work illustrates the power of modern artificial intelligence to aid in discovery of accurate and interpretable models of complex natural systems.

dynamical systems | nociception | machine learning | phenomenological models

The quantitative biology revolution of recent decades has resulted in an unprecedented ability to measure dynamics of complex biological systems in response to perturbations with the accuracy previously reserved for inanimate, physical systems. For example, the entire escape behavior of a roundworm *Caenorhabditis elegans* in response to a noxious temperature stimulus can be measured for many seconds in hundreds of worms (1, 2). At the same time, theoretical understanding of such living dynamical systems has lagged behind, largely because, in the absence of symmetries, averaging, and small parameters to guide our intuition, building mathematical models of such complex biological processes has remained a very delicate art. Recent years have shown the emergence of automated modeling approaches, which use modern machine-learning methods to automatically infer the dynamical laws underlying a studied experimental system and predict its future dynamics (3–14). However, arguably, these methods have not yet been applied to any real experimental data with dynamics of a priori unknown structure to produce interpretable dynamical representations of the system. Thus, their ability to build not just statistical but physical models of data (15) which are interpretable by humans, answer interesting scientific questions, and guide future discovery remains unclear.

Here we apply the *Sir Isaac* platform for automated inference of dynamical equations underlying time series data to infer a model of the *C. elegans* escape response, averaged over a population of worms. We show that *Sir Isaac* is able not only to fit the observed data, but also to make predictions about the worm dynamics that extend beyond the data used for training. The inferred optimal model has an easily interpretable form, with the identified interactions and the inferred latent dynamical variable connecting naturally to known mechanisms of *C. elegans* sensorimotor control. And by analyzing the dynamical structure of the model—number of dynamic variables, number of attractors (distinct behaviors), etc.—we can generalize these results across many biophysical systems.

Results

Automated Dynamical Inference. *Sir Isaac* (7, 8) is one of the new generation of machine-learning algorithms able to infer a dynamical model of time series data, with the model expressed in terms of a system of differential equations. Compared with other approaches, *Sir Isaac* is able to infer dynamics (at least for synthetic test systems) that are (i) relatively low dimensional, (ii) have unobserved (hidden or latent) variables, (iii) have arbitrary nonlinearities, (iv) rely only on noisy measurements of the system's state variables and not of the rate of change of these variables, and (v) are expressed in terms of an interpretable system of coupled differential equations. Briefly, the algorithm sets up a complete and nested hierarchy of nonlinear dynamical models. Nestedness means that each next model in the hierarchy is more complex (in the sense of having a larger explanatory power) (16–18) than the previous one and includes it as a special case. Completeness means that any sufficiently general dynamics can be approximated arbitrarily well by some model within the hierarchy. That is, the only restriction on the dynamics is that they are continuous and do not have infinite rates of change. Two such hierarchies have been developed, one based on S systems (19) and the other on sigmoidal networks (20). Both progressively add hidden dynamical variables to the model and then couple them to the previously introduced variables using nonlinear interactions of specific forms. *Sir Isaac* then uses a semianalytical formulation

Significance

The cost of an empirical bit in biophysics has fallen dramatically, and high-precision data are now abundant. However, biological systems are notoriously complex, multiscale, and inhomogeneous, so that we often lack intuition for transforming such measurements into theoretical frameworks. Modern machine learning can be used as an aid. Here we apply our *Sir Isaac* platform for automatic inference of a model of the escape response behavior in a roundworm directly from time series data. The automatically constructed model is more accurate than that curated manually, is biophysically interpretable, and makes nontrivial predictions about the system.

Author contributions: I.N. coordinated the project; B.C.D., W.S.R., and I.N. designed research; B.C.D. wrote the software; B.C.D., W.S.R., and I.N. performed research; W.S.R. collected experimental data; B.C.D., W.S.R., and I.N. analyzed data; and B.C.D., W.S.R., and I.N. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: Data related to this work have been deposited on figshare (https://figshare.com/articles/Data_and_Code_Archive_for_Automated_predictive_and_interpretable_inference_of_C.elegans_escape_dynamics/7806602). The developed software is available on GitHub (<https://github.com/EmoryUniversityTheoreticalBiophysics/SirIsaac>).

¹To whom correspondence should be addressed. Email: ilya.nemenman@emory.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1816531116/-DCSupplemental.

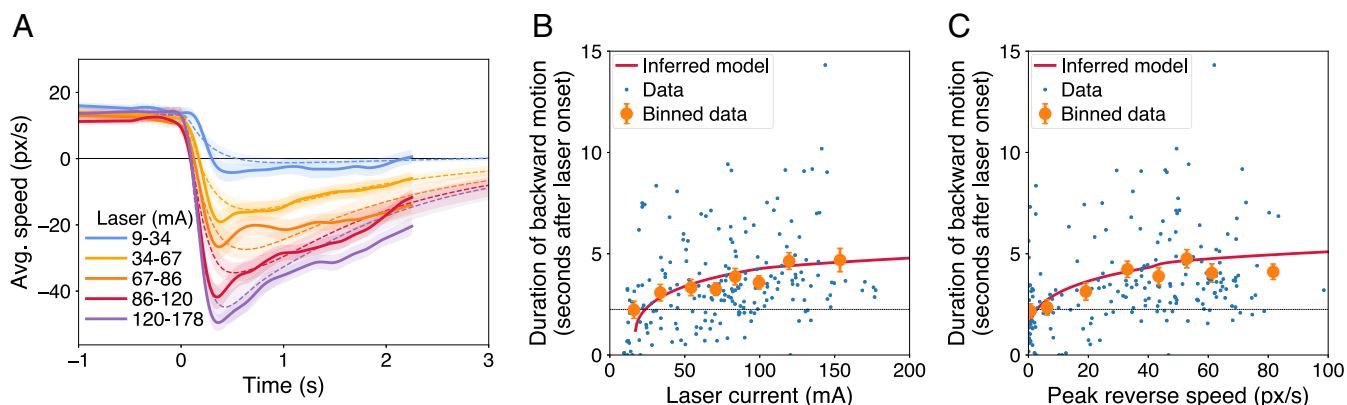


Fig. 1. The escape response behavior is fitted and predicted well by the inferred model. (A) Colored lines and shaded bands represent the empirical mean and the SD of the mean, respectively, of the escape response velocity for five groups of worms stimulated with laser currents in different ranges (38, 42, 39, 41, and 41 subjects in each group). Dashed lines and bands of the corresponding color show means and SDs of the mean (*Materials and Methods*) of fits to these empirical data by the chosen model. Only the velocity in the range of time $[-1, 2.25]$ s relative to the time of the onset of the laser stimulus was used for fitting. (B and C) While most worms are still moving backward during this range of time, the inferred model predicts, without any additional free parameters, the time at which a worm's speed again becomes positive both as a function of (B) applied laser current and (C) peak observed worm reverse speed. These predictions (red curves) agree with the binned averages (orange circles) of individual worms' behavior (blue circles). Bins correspond to eight quantiles each with 25 or 26 trials, and error bars are the SEM.

of Bayesian model selection (8, 18, 21, 22) to choose the model in the hierarchy that best balances the quality of fit vs. overfitting and is, therefore, expected to produce the best generalization. The sigmoidal network hierarchy is especially well suited to modeling biological systems, where rates of change of variables usually saturate over some scale, and it is the sole focus of our study.

Experimental Model System. Nociception evokes a rapid escape behavior designed to protect the animal from potential harm (23, 24). *C. elegans*, a small nematode with a simple nervous system, is a classic model organism used in the study of nociception. A variety of studies have used *C. elegans* to elucidate genes and neurons mediating nociception to a variety of aversive stimuli including high osmolarity and mechanical, chemical, and thermal stimuli (25–28). However, a complete dynamic understanding of the escape response at the neuronal, as well as the molecular, level is not fully known. Recent studies have quantified the behavioral escape response of the worm when thermally stimu-

lated with laser heating (1, 2), and these data are the focus of our study. The response is dynamic: When the stimulus is applied to the animal's head, it quickly withdraws, briefly accelerating backward, and eventually returns to forward motion, usually in a different direction. Various features of this response change with the level of laser heating, such as the length of time moving in the reverse direction and the maximum speed attained.

Fits and Predictions. We use the worm center-of-mass speed, v , as the variable whose dynamics need to be explained in response to the laser heating pulse. We define $v > 0$ as the worm crawling forward and $v < 0$ as the worm retreating backward. The input to the model is the underlying temperature, $h(t)$, which can be approximated as $h(t) = I h_0(t)$, where I is the experimentally controlled laser current, and h_0 is the temperature template, described in *Materials and Methods*. Based on trajectories of 201 worms in response to laser currents ranging between 9.6 mA and 177.4 mA, we let *Sir Isaac* determine the most likely dynamical system explaining these data within the sigmoidal networks

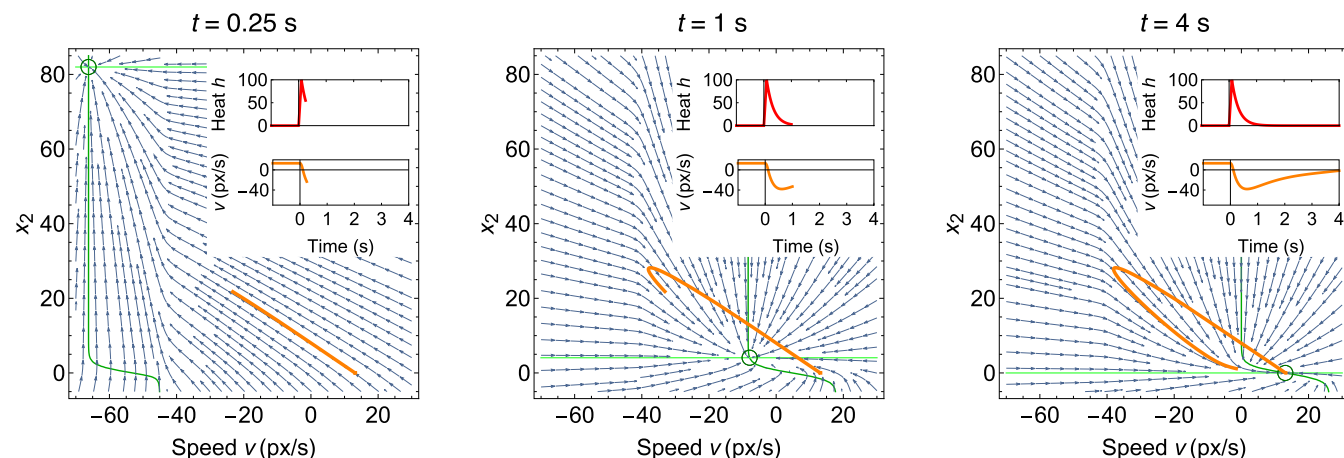
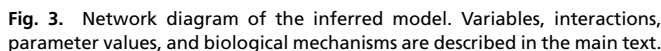


Fig. 2. Phase-space structure of the inferred model. With one hidden variable x_2 , the model dynamics can be visualized in the 2D phase space. As the instantaneous heat input h returns to zero after a brief pulse (red curve in *Insets*), a single fixed point in the 2D (v, x_2) dynamics moves from negative speed (escape) to positive (forward motion). As the velocity of the worm trails the fixed point, this produces first a fast escape and then a slow return to forward motion (speed trajectory in orange in *Insets* and in the phase portrait plots). Blue arrows indicate flow lines, circles indicate stable fixed points, and green lines indicate nullclines (dark green where $dv/dt = 0$ and light green where $dx_2/dt = 0$).



PNAS Latest Articles | 3 of 6

Table 1. Parameters used for input data, model inference, and dynamics of sensory input

Parameter	Value
Input data	
Equilibration time	10 s
Inference	
Complexity step size	1
Maxiter	100
Avegtol	10^{-2}
StopfittingN	5
Type order	"Last"
Tested ensemble members	10
Ensemble temperature	100
Ensemble generation steps	1,000
Connection order	"Node"
Prior σ	100
Sensory input	
τ_{on}	0.1 s
τ_{decay}	0.25 s

Inference parameters control the *Sir Isaac* algorithm: "Complexity step size" sets which models in the hierarchy are tested, with 1 indicating that no models are skipped as parameters are added; "Ensemble" refers to the parameter ensemble used to avoid local minima during fitting; "Avegtol" and "Maxiter" adjust the local minimization phase of parameter fitting; "StopfittingN" sets the number of higher-complexity models with poorer performance tested before selecting a given model; "Connection order" and "Type order" set the order of models in the hierarchy (7); and "Prior σ " is the SD of Gaussian priors on all parameters.

down the forward drive. This is again consistent with the literature indicating that the worms pause with even reasonably small temperature perturbations (1). An additional effect of the stimulus is to directly inject a negative drive $-V_1 h$ into the dynamics of the velocity. When the stimulus is large, the V_1 term is sufficiently negative to result in the velocity overshooting the pause into the negative, escape range. We identify the V_1 term with the reverse command interneurons AVA, AVD, and AVE, activated by the thermosensory neurons AFD and FLP (32). When $v > -\theta_1 \approx -42$ pixels/s (px/s), the W_{11} term is suppressed. However, during fast escapes this suppression is lifted, activating the positive drive W_{11} , which leads to faster recovery of the forward velocity. We identify the W_{11} term with internal recovery dynamics of the reverse command interneurons whose molecular mechanisms of activity are only partially understood (33). The velocity does not just relax to zero over some characteristic time, but crosses back into the forward crawl once x_2 has decreased sufficiently to reactivate the W_{12} term. Overall, the model suggests a number of neuronal correlates for its features. And where there is no direct match between known worm biology and the model, the model guides us to look for specific predicted features, such as the neural and molecular mechanisms for sensing both the heat stimulus and its recent average.

Another notable feature of the network diagram is that it is similar to other well-known sensory networks, namely chemotaxis (and the related thermotaxis) in *Escherichia coli* (34, 35) and chemotaxis in *Dictyostelium discoideum* (36, 37). In all these cases, the current value of the stimulus is sensed in parallel to the stimulus integrated over the recent time. They are later brought together in a negative feedback loop (*E. coli*) or an incoherent feedforward loop (*D. discoideum*), resulting in various adaptive behaviors. In contrast, the worm's behavior is more complicated: The ambient temperature participates in a negative feedback loop on the speed through W_{11} , which results in adaptation. At the same time, the integrated temperature (through W_{12}) and the current temperature add coherently to cause the escape in response to both the temperature and its rate of change. This

illustrates the difference between sensing, when an organism needs to respond to stimulus changes only, and escape, where more complicated dynamics are needed.

Discussion

In this article, we have used modern machine learning to learn the dynamics underlying the temperature escape behavior in *C. elegans*. The resulting automatically inferred model is more accurate than the model curated by hand. It uses the dynamics within what is normally considered as discrete behavioral states to make extremely precise verifiable (and verified) predictions about the behavior of the worm beyond the range of time used for training. The model is fully interpretable, with many of its features having direct biological, mechanistic equivalents. Where such biological equivalents are unknown, the model makes strong predictions of what they should be and suggests what future experiments need to search for.

One can question whether describing the *C. elegans* nociceptive behavior, which typically is viewed as stochastic (1, 2) and switching between discrete states, with the deterministic dynamics approach of *Sir Isaac* is appropriate. The quality of the fits and predictions is an indication that it is. This is likely because the dynamics of the mean behavior, which are modeled by *Sir Isaac*, are deterministic even for a stochastic system. Furthermore, as the stimulus intensity increases, the entire escape, and not just its mean, becomes more and more deterministic (1). In this regime, the discretized behavior states (forward, backward, pause, ...) have their own pronounced internal dynamics with different time-dependent velocities. These dynamical changes are comparable to differences between the states. That is, the range of forward speeds is similar to the difference between the mean forward and the mean backward speeds. Unlike most probabilistic methods, *Sir Isaac* correctly models these dynamical changes within and across the states.

Crucially, the model discovers that, at least in the context of stimuli that are scalar multiples of the same single-pulse time course, the behavior of an average worm is not a simple one-to-one mapping of the input signal: The instantaneous stimulus and its temporally integrated history (one latent variable) are both important for driving the behavior. The behavior is driven by one fixed point in the velocity-memory phase space, and the worm changes its speed while chasing this fixed point, which in turn changes in response to the stimulus. This is in contrast to other possibilities, such as the worm being able to exhibit both the forward and the backward motion at any stimulus value, and the stimulus and its history merely affect the probability of engaging in either of these two behaviors. In this initial study, we have focused on explaining the first few seconds of the worm's response to a single heat pulse. Future experimental analysis will reveal whether the same dynamical model also explains the response to more complex temporal stimulus profiles, which might include habituation and other longer-term phenomena.

We emphasize that our automatic phenomenological inference of dynamics of complex biological systems relies very little on help from constraints imposed by knowledge of the underlying biology. Indeed, the choice of the sigmoidal model hierarchy assumes only continuous dynamics without infinite rates of change. The emphasis on interpretable, physical models allows extrapolation well beyond the data used for training, which is difficult for purely statistical methods. The automation allows for a comprehensive search through the model space, so that the automatically inferred model is better than the human-assembled one, especially when, as even for the best-studied biological systems, we do not have the necessary set of measurements to model them from the ground up. Our work illustrates the power of a phenomenological modeling approach, which allows for top-down modeling, adding interpretable constraints to our understanding of the system.

Table 2. Maximum-likelihood parameters for the model inferred by *Sir Isaac*

Parameter	Value
τ_1	1.33 s
V_1	−2.30 px/(s ² ·mA)
W_{11}	81.0 px/s ²
W_{21}	−0.1 s ^{−1}
$v(t=0)$	12.6 px/s
θ_1	42.1 px/s
V_2	1.7 (s mA) ^{−1}
W_{12}	44.1 px/s ²
W_{22}	5.5 s ^{−1}
$x_2(t=0)$	1.6

We do not calculate parameter covariances since posteriors are sloppy (39) and non-Gaussian around the maximum likelihood. Instead we estimate the SD of the fits themselves (Fig. 1A). Parameter values are reported to an accuracy of about one digit beyond the least statistically significant one.

Materials and Methods

Data Collection. A detailed description of the experiments has been previously published (2). We raised wild-type, N2 *C. elegans* using standard methods, incubated at 20 °C with food. Worms were washed to remove traces of food and placed on the surface of an agar plate for 30 min at 20 °C to acclimatize. Worms were then transferred to an agar assay plate seeded with bacteria (food) and left to acclimatize for 30 more minutes. The worms were then stimulated with an infrared laser (2-mm beam diameter) focused (100-mm focal length) to a diffraction-limited spot (~90 μm) directed at the “nose” of the worm. The intensity of the stimulus was randomized by selecting a laser current between 0 mA and 200 mA, with a duration of 0.1 s. The laser pulses caused nearly instantaneous temperature increases up to a maximum of 2 °C for our current range. Each worm was stimulated only once and then discarded. A video of each worm’s response was recorded at 60 Hz and processed offline using custom LabVIEW and MATLAB scripts.

Input Data. Data used for the inference are as described in ref. 2. Speed data for 201 worms were extracted from video frames at 60 Hz and smoothed using a Gaussian kernel of width 74.5 ms (*SI Appendix*). We use data between 0.5 s before and 2.25 s after the start of the laser stimulus. Aligning the data by laser start time, the stimulus happens at the same time in each trial. Naively used, this can produce models that simply encode a short delay followed by an escape, without requiring the stimulus. To ensure that instead the stimulus causes the response in the model, for each trial we add a random delay between 0 s and 1 s to the time data.

Additionally, we are not interested in capturing any dynamics in the prelaser free-crawling state. If we use only the small amount of prelaser data measured in this experiment, the inference procedure is free to include models with complicated transient behavior before the stimulus. For this reason, we include a copy of prelaser data at a fictitious “equilibration time” long before the stimulus time (10 s), artificially forcing the model to develop a prestimulus steady state of forward motion. Finally, we weight the prelaser data such that they appear with equal frequency as postlaser data in fitting, so that the inference algorithm is not biased toward capturing postlaser behavior more accurately than the prelaser one. Data used in the work are available from figshare (https://figshare.com/articles/Data_and_Code_Archive_for_Automated_predictive_and_interpretable_inference_of_C.elegans_escape_dynamics_/7806602).

Estimating Explainable Variance. The observed variance in worm speed can be partitioned into that caused by the input (changes in laser current) and that caused by other factors (individual variability, experimental noise, etc.). As in ref. 2, we focus on our model’s ability to capture the former, “explainable” variance. We treat the latter variance as “unexplainable” by our model, and it is this variance that we use to define uncertainties on data points for use in the inference procedure. We estimate these variances by splitting the data into trials with similar laser current I_μ (five bins), producing variances $\sigma_\mu^2(t)$ that depend on the laser current bin and on the time relative to laser start. For simplicity, we use a constant uncertainty for all data points that is an average over laser current bins and times relative to the stimulus onset, which is equal to $\sigma = 14.2$ px/s. We use this uncertainty when calculating the χ^2 that defines the model’s goodness of fit,

$$\chi^2(\Theta) = \sum_{i=1}^{N_D} \left(\frac{x_i^{\text{model}}(\Theta) - x_i^{\text{data}}}{\sigma} \right)^2, \quad [3]$$

where the index i enumerates data points used for the evaluation, N_D is the number of points, and Θ is the vector of all parameters.

An analogous goodness-of-fit measure can be calculated for the predictions in Fig. 1B and C, where we use bins corresponding to the displayed orange circles. For predicting the duration of backward motion as a function of laser current (Fig. 1B), $\chi^2/\text{dof} = 1.12$, and as a function of peak reverse speed (Fig. 1C), $\chi^2/\text{dof} = 1.14$.

***Sir Isaac* Inference Algorithm.** We use the *Sir Isaac* dynamical inference algorithm (8) to find a set of ordinary differential equations (ODEs) that best describes the data without overfitting. Based on previous studies of simulated biological systems, we use the continuous-time sigmoidal network model class, which produces a set of J ODEs of the form

$$\frac{dx_i}{dt} = -x_i/\tau_i + V_i h(t) + \sum_{j=1}^J W_{ij} \xi(x_j + \theta_j), \quad [4]$$

where $\xi(y) = 1/(1 + \exp(y))$ and $h(t)$ is the sensory input defined in Eq. 7. The algorithm infers both the number of parameters (controlled by the total number of dynamical variables J) and the parameter values themselves (the timescales τ_i , interaction strengths V_i and W_{ij} , and biases θ_j). Note that the algorithm has the freedom to make use of linear or nonlinear parts of the sigmoid function. The first dynamical variable x_1 is taken to be the signed speed v of the worm’s center of mass, measured in pixels per second, with negative values corresponding to backward motion. The choice of units for v defines a relevant scale $v_s = 1$ px/s, which does not affect the form of the equations, but which we include where necessary for a correct interpretation of units. Further dynamical variables x_i with $i > 1$ correspond to latent (unmeasured) dynamical variables.

The fitting procedure starts by using one data point (one random time) from each of a few trials and then gradually adds trials and eventually multiple data points per trial, refitting model parameters at each step (see Table 1 for algorithm parameters). The resulting model fits are scored based on their performance in predicting the entire time series (see Fig. 4 for the fit quality). When the performance and model complexity of the winning model saturate, we use the resulting model as our description of the system. In this way, parameters are fitted using only a small subset of the available data—we find that using ~2 randomly chosen time points per trial (of the total 165) is sufficient (Fig. 4). This approach significantly reduces computational effort (which scales linearly in the number of data points used) and minimizes the effects of correlations between data points that are close to one another in time. Finally, it prevents the optimization from getting stuck at local minima and saddles, which change as new data points are added or data are randomized. These reasons are similar to the reasons behind stochastic gradient descent approaches (38). The developed software is available from <https://github.com/EmoryUniversityTheoreticalBiophysics/SirIsaac>.

Inferred models. The inference procedure produces the following differential equations:

$$\frac{dv}{dt} = -\frac{v}{\tau_1} + V_1 h(t) + \frac{W_{11}}{1 + \exp(\theta_1 + v/v_s)} + \frac{W_{12}}{1 + \exp(x_2)}, \quad [5]$$

$$\frac{dx_2}{dt} = -\frac{x_2}{\tau_2} + V_2 h(t) + \frac{W_{21}}{1 + \exp(\theta_1 + v/v_s)} + \frac{W_{22}}{1 + \exp(x_2)}. \quad [6]$$

The model includes one latent dynamical variable x_2 . The maximum-likelihood fit parameters are shown in Table 2. (Note that the selected model

Table 3. Maximum-likelihood parameters for the reduced model

Parameter	Value
τ_1	1.33 s
V_1	−2.30 px/(s ² ·mA)
W_{11}	76.0 px/s ²
$v(t=0)$	12.1 px/s
θ_1	42.1 px/s
V_2	1.5 (s mA) ^{−1}
W_{12}	19.7 px/s ²

did not include variables τ_2 and θ_2 , so they are set to their default values: $\tau_2 = 1$ s and $\theta_2 = 0$.)

We note that some of the inferred parameters are close to zero, and so we check whether the model can be simplified by setting each of them to zero one at a time and in combinations and then measuring the approximate Bayesian model selection posterior-likelihood score (8) for the original *Sir Isaac* inferred model and for each of the reduced models. The original model has log-likelihood -197.2 , and the best model, Eqs. 1 and 2, has the highest Bayesian log-likelihood of -192.9 . This becomes our chosen model, maximum-likelihood parameters for which can be found in Table 3.

Bayesian Model of Uncertainty. To quantify uncertainty in model structure and parameters, we take a Bayesian approach and sample from the posterior distribution over parameters. Assuming independent Gaussian fluctuations in the data used in fits, the posterior is simply $\exp(-\chi^2(\Theta))$, with $\chi^2(\Theta)$ from Eq. 3. We use a standard Metropolis Monte Carlo sampler, as implemented in SloppyCell (40), to take 100 samples (each separated by 1,000 Monte Carlo steps) used to quantify uncertainty in the fit (Fig 1A).

Model of Sensory Input. Each experimental trial begins with a forward-moving worm, which is stimulated with a laser pulse of duration $\tau_{\text{on}} = 0.1$ s starting at time $t = 0$. The worm's nose experiences a quick local temperature increase $h(t)$, which we model as a linear increase during the pulse, with slope proportional to the laser current I . The timescale of the temperature decay of the heated area to the ambient temperature due to heat diffusion is 0.15 s (41). However, the stimulation area is broad (220 μm , FWHM), and as the worm retreats, its head with the sensory neurons first moves deeper into the heated area, before the temperature eventually decreases. Thus, the dynamics of the sensory input are complex and multiscale. However, since each individual behaves differently, and we do not measure individual head temperatures, we model the average sensory stimulus past the

heating period as a single exponential decay with the longer timescale of $\tau_{\text{decay}} = 0.25$ s:

$$h(t)/\alpha = \begin{cases} 0 & t \leq 0, \\ It/\tau_{\text{on}} & 0 \leq t \leq \tau_{\text{on}}, \\ I \exp[-(t - \tau_{\text{on}})/\tau_{\text{decay}}] & t \geq \tau_{\text{on}}. \end{cases} \quad [7]$$

Here α has units of temperature per unit of laser current. For convenience and without loss of generality, we set $\alpha = 1$ and absorb its definition into the V_i parameters that multiply h in Eqs. 5 and 6, giving V_i units of the time derivative of x_i per unit current.

Comparison to the Model of Ref. 2. A quantitative model of *C. elegans* nociceptive response was constructed by hand in ref. 2. Contrasting with our automated inference, the model was not fitted by optimizing parameters, but rather includes the entire measured escape behavior as a templated, stereotyped response. It was based on careful study of the data, which revealed that the mean response of worms is nearly stereotypical, with the response amplitude depending nonlinearly on the stimulus intensity. Of the variability in the response that is explainable by the stimulus, the model of ref. 2 could not account for $\sim 20\%$ in the range from a few hundred milliseconds to about 2 s poststimulation. Using *Sir Isaac*, with very little human input, our selected model captures more of the data variance. It achieves this using a model of a very different form (a set of coupled ODEs), leaving only $\sim 10\%$ of the explainable variance unexplained over much of the same time range, and it also captures the approximate stereotypy in the response to the same accuracy as is present in the data (SI Appendix, Fig. S1).

ACKNOWLEDGMENTS. We are grateful to the NVIDIA Corporation for donated GPUs and to the Aspen Center for Physics, supported by NSF Grant PHY-1607611, for its "Physics of Behavior" programs. This work was partially supported by NIH Grants EB022872 and NS084844, James S. McDonnell Foundation Grant 220020321, a Natural Sciences and Engineering Research Council of Canada Discovery Grant, and NSF Grant IOS-1456912.

- Mohammadi A, Byrne Rodgers J, Kotera I, Ryu W (2013) Behavioral response of *Caenorhabditis elegans* to localized thermal stimuli. *BMC Neurosci* 14:66.
- Leung K, Mohammadi A, Ryu W, Nemenman I (2016) Stereotypical escape behavior in *Caenorhabditis elegans* allows quantification of nociceptive stimuli levels. *PLoS Comput Biol* 12:e1005262.
- Schmidt M, Lipson H (2009) Distilling free-form natural laws from experimental data. *Science* 324:81–85.
- Schmidt M, et al. (2011) Automated refinement and inference of analytical models for metabolic networks. *Phys Biol* 8:055011.
- Sussillo D, Abbott L (2009) Generating coherent patterns of activity from chaotic neural networks. *Neuron* 63:544–557.
- Neuert G, et al. (2013) Systematic identification of signal-activated stochastic gene regulation. *Science* 339:584–587.
- Daniels B, Nemenman I (2015) Efficient inference of parsimonious phenomenological models of cellular dynamics using S-systems and alternating regression. *PLoS One* 10:e0119821.
- Daniels B, Nemenman I (2015) Automated adaptive inference of phenomenological dynamical models. *Nat Commun* 6:8133.
- Brunton S, Proctor J, Kutz J (2016) Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc Natl Acad Sci USA* 113:3932–3937.
- Mangan N, Brunton S, Proctor J, Kutz J (2016) Inferring biological networks by sparse identification of nonlinear dynamics. *IEEE Trans Mol Biol Multi Scale Commun* 2:52–63.
- Lu Z, et al. (2017) Reservoir observers: Model-free inference of unmeasured variables in chaotic systems. *Chaos* 27:041102.
- Pathak J, Hunt B, Girvan M, Lu Z, Ott E (2018) Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. *Phys Rev Lett* 120:024102.
- Pandarinath C, et al. (2018) Inferring single-trial neural population dynamics using sequential auto-encoders. *Nat Methods* 15:805–815.
- Henry A, Hemery M, François P (2018) Phi-evo: A program to evolve phenotypic models of biological networks. *PLoS Comput Biol* 14:e1006244.
- Nelson P (2015) *Physical Models of Living Systems* (W. H. Freeman and Co., New York).
- Vapnik V (1998) *Statistical Learning Theory* (Wiley, New York).
- Rissanen J (1989) *Stochastic Complexity in Statistical Inquiry Theory* (World Scientific, River Edge, NJ).
- Bialek W, Nemenman I, Tishby N (2001) Predictability, complexity, and learning. *Neural Comput* 13:2409–2463.
- Savageau M, Voit E (1987) Recasting nonlinear differential-equations as S-systems—A canonical nonlinear form. *Math Biosci* 87:83–115.
- Beer R, Daniels B (2010) Saturation probabilities of continuous-time sigmoidal networks. arXiv: 1010.1714. Preprint, posted October 8, 2010.
- MacKay DJ (1992) Bayesian interpolation. *Neural Comput* 4:415–447.
- Balasubramanian V (1997) Statistical inference, occam's razor, and statistical mechanics on the space of probability distributions. *Neural Comput* 9:349–368.
- Eaton R (2013) *Neural Mechanisms of Startle Behavior* (Springer, New York).
- Pirri J, Alkema M (2012) The neuroethology of *C. elegans* escape. *Curr Opin Neurobiol* 22:187–193.
- Bargmann C, Thomas J, Horvitz H (1990) Chemosensory cell function in the behavior and development of *Caenorhabditis elegans*. *Cold Spring Harbor Symp Quant Biol* 55:529–538.
- Hilliard M, et al. (2005) In vivo imaging of *C. elegans* ash neurons: Cellular response and adaptation to chemical repellents. *EMBO J* 24:63–72.
- Kaplan J, Horvitz H (1993) A dual mechanosensory and chemosensory neuron in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 90:2227–2231.
- Wittenburg N, Baumeister R (1999) Thermal avoidance in *Caenorhabditis elegans*: An approach to the study of nociception. *Proc Natl Acad Sci USA* 96:10477–10482.
- Stephens GJ, Johnson-Kerner B, Bialek W, Ryu WS (2008) Dimensionality and dynamics in the behavior of *C. elegans*. *PLoS Comput Biol* 4:e1000028.
- Goodman MB, Sengupta P (2018) The extraordinary AFD thermosensor of *C. elegans*. *Pflügers Arch* 470:839–849.
- Liu S, Schulze E, Baumeister R (2012) Temperature- and touch-sensitive neurons couple CNG and TRPV channel activities to control heat avoidance in *Caenorhabditis elegans*. *PLoS One* 7:e32360.
- White JG, Southgate E, Thomson JN, Brenner S (1986) The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci* 314:1–340.
- Gao S, et al. (2015) The NCA sodium leak channel is required for persistent motor circuit activity that sustains locomotion. *Nat Commun* 6:6323.
- Sourjik V, Wingreen NS (2012) Responding to chemical gradients: Bacterial chemotaxis. *Curr Opin Cell Biol* 24:262–268.
- Paulick A, et al. (2017) Mechanism of bidirectional thermotaxis in *Escherichia coli*. *eLife* 6:e26607.
- Jilkine A, Edelstein-Keshet L (2011) A comparison of mathematical models for polarization of single eukaryotic cells in response to guided cues. *PLoS Comput Biol* 7:e1001121.
- Levchenko A, Iglesias P (2002) Models of eukaryotic gradient sensing: Application to chemotaxis of amoebae and neutrophils. *Biophys J* 82:50–63.
- Mehta P, et al. (2018) A high-bias, low-variance introduction to machine learning for physicists. arXiv:1803.08823. Preprint, posted March 23, 2018.
- Transtrum M, et al. (2015) Perspective: Slowness and emergent theories in physics, biology, and beyond. *J Chem Phys* 143:010901.
- Myers C, Gutenkunst R, Sethna J (2007) Python unleashed on systems biology. *Comput Sci Eng* 9:34–37.
- Mohammadi A (2013) Quantitative behavioral analysis of thermal nociception in *Caenorhabditis elegans*: Investigation of neural substrates spatially mediating the noxious response, and the effects of pharmacological perturbations. PhD thesis (Univ of Toronto, Toronto).