Bulletin of Mathematical Biology (2023) 85:91
https://doi.org/10.1007/s11538-023-01176-x

Society for Mathematical Biology

ORIGINAL ARTICLE

Predicting the Behavior of Sparsely-Sampled Systems Across Neurobiology and Epidemiology

Eve Armstrong[1,2]

Received: 16 December 2022 / Accepted: 30 May 2023 / Published online: 31 August 2023

Abstract

Inference is a term that encompasses many techniques including statistical data assimilation (SDA). Unlike machine learning, which is designed to harness predictive power from extremely large data sets, SDA is designed for sparsely-sampled systems. This is the realm of study of nonlinear dynamical systems in nature. Formulated as an optimization procedure, SDA can be considered a path-integral approach to state and parameter estimation. Within this formulation, we can use the physical principle of least action to identify optimal solutions: solutions that are consistent with both measurements and a dynamical model assumed to give rise to those measurements. I review examples from neurobiology and an epidemiological model tailored to the coronavirus SARS-CoV-2, to demonstrate the versatility of SDA across the sciences, and how these distinct applications possess commonalities that can inform one another.

Keywords Inference · Neurobiology

1 Introduction

Dynamical models of physical systems in nature are fiercely nonlinear beasts. They may possess multiple timescales and many degrees of freedom, usually a small fraction of which are physically measurable quantities. These characteristics render the task of state and parameter estimation a formidable challenge. Fortunately, while dynamical systems across disciplines are extraordinarily diverse in terms of their context and scale, they possess one commonality: they all obey the principle of least action.

This principle, attributed to de Maupertuis (1750), Euler (1744), and Kabitz (1913) and Gerhardt (1898) (early 1700's), is a variational principle in physics that can be used to derive the Newtonian, Lagrangian, and Hamiltonian equations of motion of a

---

✉ Eve Armstrong
earmst01@nyit.edu

1 Department of Physics, New York Institute of Technology, New York, NY 10023, USA

2 Department of Astrophysics, American Museum of Natural History, New York, NY 10024, USA

dynamical system. It is both a deep and a profoundly useful statement: that among an infinite number of possible paths that a system can take in a state space, the path that the system takes is that which minimizes one scalar quantity: the "action" along that path. In other words, for any physical problem, there exists a scalar quantity that can be used as a metric for identifying optimal solutions. In physics, this quantity is the action. Meanwhile, within the context of inference, this quantity is a "cost function." In this paper, we shall examine—and exploit—the powerful equivalency between the two terms.

The principle of least action underlies path integral approaches to statistical data assimilation (SDA) (Abarbanel 2013; Quinnand and Abarbanel 2011; Restrepo 2008). SDA is an inference procedure (Tarantola 2005) wherein a dynamical system is assumed to underlie any measured quantities. Invented for numerical weather prediction (Kimura 2002; Kalnay 2003; Evensen 2009; Betts 2010; Whartenby et al. 2013; An et al. 2017), SDA differs from many machine learning (ML) techniques in that it is designed specifically for cases wherein measurements of the system are sparse—the opposite of the "big data" wave at work in the ML community. This makes SDA well suited for solving realistic dynamical models of physical systems.

The SDA methodology asks: *Which measurements must be made in order to estimate unknown model components, so that the completed model can predict events that occur outside locations at which the measurements were made?* SDA identifies not only which measurements are required, but also properties of those measurements such as the sampling frequency and the degree of contamination that can be tolerated. Moreover, SDA reveals how Shannon entropy (Fano 1961) flows in a dynamical system, from measured to unmeasured state variables, so that the complete system can be known. The inherent simplicity and power of path-integral-based methods of SDA can tremendously impact the study of dynamical systems across diverse disciplines.

In neurobiology, the technique has gained traction over the past 15 years (Schiff 2009; Toth et al. 2011; Kostuk et al. 2012; Hamilton et al. 2013; Meliza etal. 2014; Nogaret et al. 2016), it was in that field that I first applied SDA to small biological neuronal network models (Armstrong 2020; Abarbanel et al. 2017; Kadakia et al. 2016; Breen etal. 2016). Then in March 2020, the rise of the novel coronavirus SARS-CoV-2 prompted me to learn that SDA was largely unknown in epidemiology. It was presented roughly 10 years ago as a means for epidemiological forecasting (Bettencourt etal. 2007; Rhodesand and Hollingsworth 2009), and there had been two employments of SDA for examining COVID-19 specifically (Sesterhenn 2020; Nadler etal. 2020), neither of which examined the sensitivity of state and parameter estimates to contamination in measurements, which at the time was a severe problem. To explore this, I formed a collaboration with epidemiologists at Northwestern University to apply SDA to a population model tailored to COVID-19 (Armstrong et al. 2021).

This paper is intended to demonstrate the versatility of SDA across biology-related disciplines, by reviewing my work in both neurobiology and epidemiology. First we shall examine the mechanics of the SDA procedure (Sect. 2), and then apply it to two sets of state variables: neurons (Sect. 3) and human beings (Sect. 4). Table 1 and Fig. 1 offer a summary statement that a path-integral-based approach to SDA distills the essence of inference to a governing principle across disciplines, and in this way, disjoint disciplines in fact possess commonality and can inform one another.

Table 1 Two distinct applications of SDA to be examined in this paper, each using the principle-of-least-action optimization framework

	Biological neuronal network	COVID-19 pandemic model
The state	A three-neuron circuit capable of reliable patterned electrical output	Populations of human beings subject to COVID-19 and restrictions on social behavior
Measured quantities	Membrane voltage of each neuron, sampled at 50 kHz over 1 s	Numbers of mildy and severely symptomatic, recovered, and dead cases, sampled once per day over 101 days
Principle of least action leads to:	Estimates of ion channel maximum conductances and synaptic coupling strengths, and prediction of future membrane voltage activity of each neuron	Estimates of all populations, transmission rate, detection probability, and the fraction of the susceptible population that will require hospitalization
Important considerations for each case	• Multiple time scales • Challenge of distinguishing degeneracy due to inadequate measurements vs degeneracy that signifies real redundancy in the system	• High uncertainty in recorded population measurements • Too early in the pandemic to fully know the model—e.g. whether Recovered population is immune or instead should return to Susceptible

2 The SDA Technique

A path-integral formulation of SDA, beginning with the definition of a model and ultimately writing an action on a path in the state space of that model, can be summarized in four equations. One can then cast the estimation as an optimization, wherein the action is a cost function—a succinct and powerful equivalency.

Let us take a dynamical model $\boldsymbol{F}$ to be written as a set of D ordinary differential equations that evolve in some continuous parameterization t as:

$$\frac{\mathrm{d}x_a(t)}{\mathrm{d}t} = F_a(\boldsymbol{x}(t), \boldsymbol{p}(t), \eta(t)); \quad a = 1, 2, \ldots, D, \tag{1}$$

where the components x_a are the model state variables. The array $\boldsymbol{p}(t)$ contains p unknown parameters to be estimated, and $\eta(t)$ is some stochastic component. A subset L of the D variables is associated with measured quantities. One seeks to estimate the unknown parameters and the evolution of all state variables that is consistent with measurements, to predict model evolution at parameterized locations where measurements are not available.

The path integral is an integral representation of the master equation for the stochastic process that is represented by Eq. 1. We seek the probability of obtaining a path $\boldsymbol{X}$ in the model's state space given observations $\boldsymbol{Y}$, or: $P(\boldsymbol{X}|\boldsymbol{Y})$. If we write:

$$P(\boldsymbol{X}|\boldsymbol{Y}) = e^{-A_0(\boldsymbol{X},\boldsymbol{Y})}, \tag{2}$$

we are stating: *the path* **X** *for which the probability—given* **Y***—is greatest is the path that minimizes the quantity* A_0, which we call our action. A formulation for A_0 will permit us to obtain the expectation value of any function
$G(\boldsymbol{X})$ on a path $\boldsymbol{X}$:

$$G(\boldsymbol{X}) = \langle G(\boldsymbol{X})\rangle = \frac{\int d\boldsymbol{X} G(\boldsymbol{X}) e^{-A_0(\boldsymbol{X},\boldsymbol{Y})}}{\int d\boldsymbol{X} e^{-A_0(\boldsymbol{X},\boldsymbol{Y})}}. \tag{3}$$

Expectation values are the quantities of interest when the problem is statistical in nature. That is, $G(\boldsymbol{X})$ can be expressed as a weighted sum over all possible paths, where the weights are exponentially sensitive to A_0. For many estimation problems, the quantity of interest is the path itself: $G(\boldsymbol{X}) = \boldsymbol{X}$.

When discretized in times t_n for computational implementation, the action is written in two general terms:

$$\begin{aligned} A_0(\boldsymbol{X}, \boldsymbol{Y}) &= -\sum \log[P(\boldsymbol{x}(n+1)|\boldsymbol{x}(n))] \\ &\quad -\sum \mathrm{CMI}(\boldsymbol{x}(n), \boldsymbol{y}(n)|\boldsymbol{Y}(n-1)). \end{aligned} \tag{4}$$

The first (“model error") term describes Markov-chain transition probabilities governing the model dynamics. Extremizing this term imposes adhesion of the solutions to model dynamics. The second (“measurement error") term is the conditional mutual

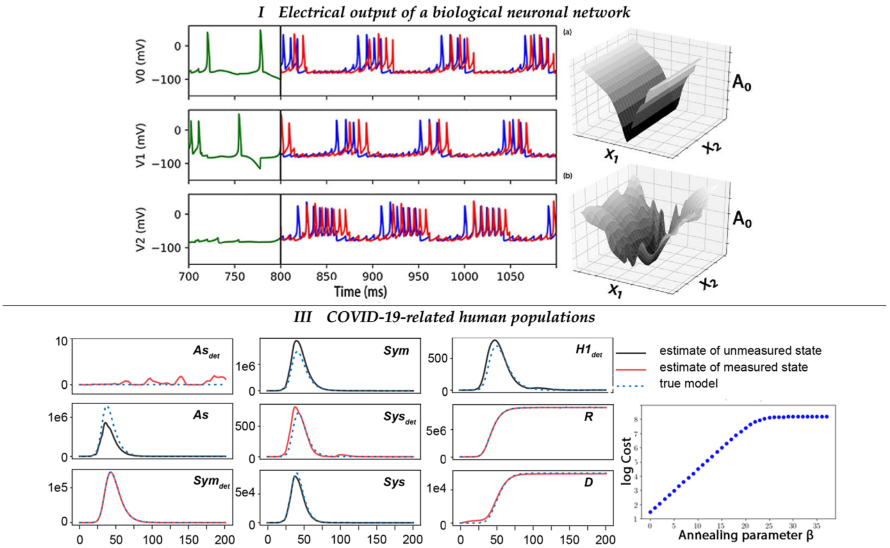

Fig. 1 State estimates and predictions from SDA across two models, with correct solutions corresponding to the path of least action. *Top*: A biological network. *Left*: Estimates (green) yield predictions (red) of timings of voltage bursts, compared to the true model (blue), for three neurons. *Right*: Representative action surfaces with (upper) and without (lower) annealing (reproduced from Armstrong 2020). *Bottom*: Estimate of COVID-19 population model, with low noise in the recovered population R. *Left*: Population estimates are: asymptomatics As versus mild and severe symptomatics Sym and Sys, undetected versus detected (denoted by "det"), one hospitalized population $H1$, Recovered R, and Dead D. *Right*: action (β) showing zero model error at high β, a reflection of the well-matched state (the significance of β is explained at the end of Sect. 2.) Additive 10% noise precludes an estimate of transmission rate (not shown), highlighting the need for accurate record-keeping (reproduced from Armstrong et al. 2021) (Color figure online)

information (Fano 1961) (CMI), which quantifies the amount of information, in bits, transferred from measurements to model. The expression asks, "How much is learned about event $\boldsymbol{x}(n)$ upon observing event $\boldsymbol{y}(n)$, conditioned on having previously observed event(s) $\boldsymbol{Y}(n-1)$?". Extremizing this term imposes the solutions' adhesion to any existing measurements.[1] For a concise derivation of Eq. 4, see Appendix A of Ref. Armstrong etal. (2017); for a complete derivation, see Ref. Abarbanel (2013).

The two terms contained in Eq. 4 are evaluated simultaneously, and it is in this way that information contained in the measurements is transferred to the incomplete model, to estimate unknown parameters. For a sparsely-sampled system, the main question becomes: *whether the available measurements contain sufficient information to succeed in this endeavour.* Simplifications are then made to write a computationally-functional form of A_0, and equality constraints may be added, which

---

[1] This "measurement error" term can be considered a synchronization term, which are often introduced artificially into control problems. Here, however, the measurement term arises naturally through considering the effects of the information those measurements contain. Absent measurements, we live in a state space restricted only by the model's degrees of freedom. Measurements guide us to a subspace in which such measurements are possible.

are model-specific. A complete specific computational implementation of A_0 is given in “Appendix A”.

The SDA problem is then cast as an optimization, where the cost function A_0 of the optimizer is equivalent to the action on paths in the state space that is searched. The $((D+p)\times(N+1))$-dimensional action surface is searched via the variational method (Odenand and Reddy 2012), where N is the number of discretized model locations, taken to be independent dimensions. One seeks the path $\boldsymbol{X}^0 = \boldsymbol{x}(0), \ldots, \boldsymbol{x}(N), \boldsymbol{p}(0), \ldots \boldsymbol{p}(N)$ in state space on which A_0 attains a minimum value.

Now, because these dynamical models are nonlinear, the action surface will be non-convex. To identify a global minimum, a method of simulated annealing is employed (Ye et al. 2015). It is important for the reader to understand this procedure, in order to see how the action formulation can be used as a metric to identify optimal solutions (Armstrong et al. 2020).

The measurement and model error terms have covariance matrices R_m and R_f, respectively. We define them to be diagonal, and render them coefficients. R_m is taken to be 1.0, and R_f is written as: $R_f = R_{f,0}\alpha^{\beta}$, where $R_{f,0}$ is a number much less than 1, α is slightly greater than 1, and β – the "annealing parameter" is initialized at zero. When $\beta = 0$, R_f is small: the action is relatively free from model constraints, and so there exists one minimum of the variational problem that is consistent with the measurements. An estimate of that minimum is obtained. Then R_f is increased slightly, via an integer increment in β, and the action is recalculated. This step is performed recursively, toward the deterministic limit of $R_f \gg R_m$. The aim is to remain sufficiently near to the global minimum so as not to become trapped in a local minimum as the surface becomes resolved. It will be shown, in Sect. 4 on COVID-19, that this method offers a systematic method to identify the lowest minimum, in a specific region of state-and-parameter space, of a non-convex cost function. Specifically: in the deterministic limit (i.e. for high values of β), the optimal solutions are those that correspond to the path of least action.

All “data" used in this paper are simulations generated via forward integration, which serve as consistency checks on the SDA solutions. Having this check is a vital first step, in advance of “flying blind” with real experimental data. Further, in simulations one may use whichever measurements one desires, regardless of whether such measurements are currently possible to take in a laboratory. It is in this way that SDA lends itself readily to informing experimental design. In References Toth et al. (2011), Kostuk et al. (2012) and Meliza etal. (2014), for example, a simulated SDA procedure identified which forms of experimentally-injected electrical currents would yield parameter estimations of a desired precision, and the real laboratory design was amended accordingly.

3 Neurobiology

Biological circuits can generate patterned electrical outputs that manifest in rhythmic motor behaviors vital for survival, such as respiration and heartbeat. The means by which neurons within such a circuit act in coordination to yield reliable output is a largely open problem. Progress has been made chiefly via the examination of small (fewer than ten-cell) networks that are functional even when isolated from the animal (Marder and Calabrese 1996; Kristan et al. 2005; Marder et al. 2005, 2007; Mulloneyand and Hall 2007; Smarandache et al. 2009; Grashow et al. 2009, 2010; Turrigiano 2011; Marder and Gutierrez 2016; Gunaratne et al. 2017), circuits known as "central pattern generators." Those references just cited have shown that relationships among cellular and synapse parameters are co-dependent. For example, couplings have been identified between two particular circuit parameters, where a change in the value of one is associated with a change in the value of the second, such that circuit output is maintained [e.g. Ref (Grashow et al. 2010)]. One major obstacle to dissecting these relationships is the difficulty of estimating more than just a few parameters *simultaneously*. This problem has been tackled extensively within the context of neurobiology, as the references above indicate.

With this context in mind, the question for SDA is: within a small biological circuit, can we use simulated measurements of cellular properties—that is, simulations of quantities that are currently obtainable in a laboratory—to simultaneously estimate cellular and synaptic properties of the circuit? At the time of publication, inference-based procedures had been applied to estimate electrophysiology of single neurons (e.g. Toth et al. 2011; Kostuk et al. 2012; Meliza etal. 2014; Nogaret et al. 2016; Abarbanel et al. 2017; Kadakia et al. 2016; Breen etal. 2016), but not yet to circuits.

3.1 Model

The model is a small network that yields a specific pattern of electrical output depending on certain values of electrophysiological and synaptic parameters. It was set forth in Armstrongand and Abarbanel (2016), wherein it corresponded to a functional representation of the avian song-related nucleus HVC. For the purposes of state-and-parameter estimation with SDA, the important aspects of the model are twofold. First, it is biologically relevant in that it exhibits distinct modes of circuit output, depending on parameter values. Second, it offers a means to quantify the success of the SDA procedure: in the predictive phase, the parameter estimates must reproduce the pattern of electrical circuit activity that is associated with the known parameter values.

The model is a three-neuron network with all-to-all inhibitory chemical synapses. Figure 2 depicts the two modes of circuit output that emerge when the circuit is exposed to a steady background current. Left and right panels show, respectively, simultaneous firing—for the case of low coupling strengths among neurons, and sequential firing—for higher values of couplings.

The neurons are Hodgkin-Huxley-type (Hodgkin et al. 1952), wherein a neuron is represented as a capacitor. Its membrane voltage as a function of time, $V(t)$, evolves in terms of currents entering and leaving the cell. The specific model is based on the electrophysiological studies of HVC inhibitory interneurons (Daou et al. 2013). The

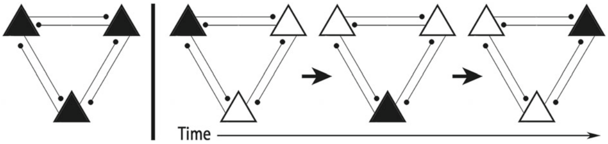

Fig. 2 Two functional modes of circuit activity, which can be expressed by the three-neuron structure when it receives a low-amplitude background current. Each triangle represents one inhibitory neuron, and they are connected all-to-all. Darkened and white shapes correspond to neurons that are currently active and inactive above spike threshold, respectively. *Left*: simultaneous firing of the three nodes, for sufficiently low coupling. *Right*: sequential firing, for a higher range of coupling strengths. (Reproduced from Armstrong 2020)

time course of membrane voltage for each neuron i is written as:

$$C_i \frac{\mathrm{d}V_i(t)}{\mathrm{d}t} = I_{L,i}(t) + I_{Na,i}(t) + I_{K,i}(t) + I_{CaT,i} + \sum_{j \neq i} I_{syn,ij}(t) + I_{inj,i}. \tag{5}$$

Parameter C is the membrane capacitance. On the right, the first four terms are ion channel currents, I_{syn} represent synaptic input currents to Neuron i from Neurons j, and I_{inj} is a current injected by the experimenter. The ion channel currents for the ith neuron are:

$$\begin{aligned} I_{L,i}(t) &= g_{L,i}(E_{L,i} - V_i(t)) \\ I_{Na,i}(t) &= g_{Na,i} m_i(t)^3 h_i(t)(E_{Na,i} - V_i(t)) \\ I_{K,i}(t) &= g_{K,i} n_i(t)^4 (E_{K,i} - V_i(t)) \\ I_{CaT,i}(t) &= g_{CaT,i} a_i(t)^3 b_i(t)^3 GHK(V_i(t), [Ca]_i(t)), \end{aligned} \tag{6}$$

for leak, sodium, potassium, and calcium currents, respectively. See “Appendix B” for details.

The synapse dynamics follow the formalism for chemically-delivered neurotransmitter pulses (Destexhe and Sejnowski 2001; Destexhe et al. 1994):

$$I_{syn,ij} = g_{ij} s_{ij}(t)(E_{syn,i} - V_i(t)), \tag{7}$$

where $I_{syn,ij}$ is the current entering cell i from cell j. The g_{ij} are the synapse maximum conductances – these are the quantities whose values determine the specific mode of firing (Fig. 2). $E_{syn,i}$ is the synaptic reversal potential of Cell i, and s_{ij}(t) is the synaptic gating variable. All state variables and parameters to be estimated in this study are contained in the equations shown above; for the interested reader, “Appendix B” gives further detail.

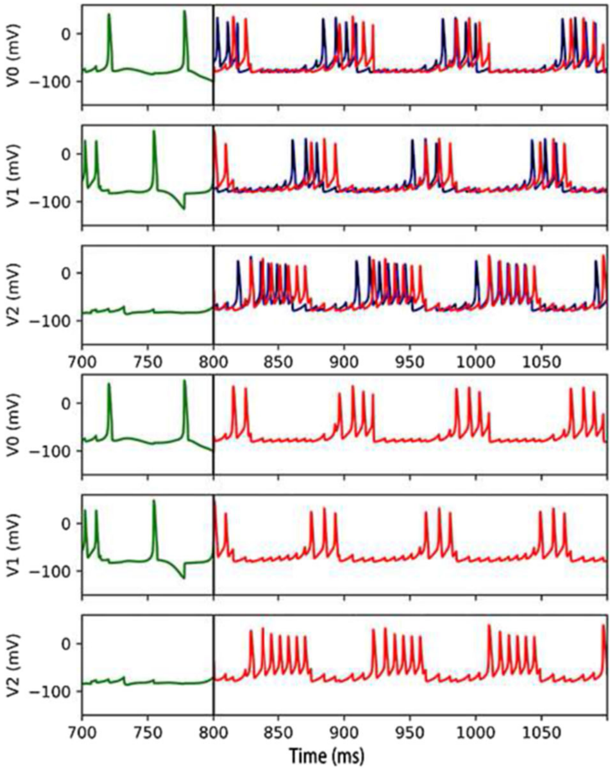

Fig. 3 Estimates and predictions of voltage time series for g_{ij} values corresponding to series activity (Fig. 2, Right), where the injected current during the prediction window is a step current that yields sequential-bursting activity in the known model. Injected currents during prediction, not shown, are: 0.4, 0.5, and 0.3 nA for each neuron, respectively. *Top three rows*: Prediction (red) is shown along with true simulation (blue). *Bottom three rows*: To aid the eye in discerning the series activity, prediction is shown alone. Forty out of forty randomly-initialized paths converged to this solution (reproduced from Armstrong 2020). (The step current was used to represent ambient input current to the circuit. More generally, for parameter estimation, it can be useful to instead use a more complex waveform as input current, which forces the system to explore its full dynamical range (Meliza etal. 2014)) (Color figure online)

3.2 SDA Procedure

The aim is to infer the six inhibitory maximum conductances g_{ij} (Eq. 7) required to reproduce the particular mode of firing, the synaptic reversal potentials $E_{syn,j}$ (Eq. 7), and the maximum conductances of ion channels on the three cells: $g_{L,i}$, $g_{Na,i}$, $g_{K,i}$, $g_{CaT,i}$ (Eq. 6). The question is: which measurements—if any—will allow us to do this? For details of the procedure, see “Appendix B”.

3.3 Result

When the SDA procedure was given as measurements the time course of membrane voltage $V(t)$ for each neuron (at 50 kHz over 0.8 s), the parameter estimation worked well. Out of a total of forty randomly-initialized trials, those that corresponded to the minimum value of the action reliably yielded state and parameter estimates corresponding to a strong prediction, while trials corresponding to higher values of the action contained various estimate errors. This has important implications for a laboratory setting, as it is possible to simultaneously measure $V(t)$ from all cells in a three- or four-neuron circuit (Sakuraiand and Katz (2016) and the references to the Marder laboratory cited in the beginning of this Section).

Figure 3 shows the resulting predictions of V(t) for each neuron, for the sequentially-bursting mode (Fig. 2, right). It shows two triads of voltage traces. The top triad shows estimation (green), prediction (red), and true model evolution (blue) for all three cells. Importantly, the estimations window is that in which the measurements were provided to the procedure, and the predictions window is that in which no measurements were provided. The metric for an optimal solution is that it not only fit the measurements within the estimations window, but also possess predictive power outside of that window. Here, the prediction was obtained by replacing the parameters in the forward-integration model with the estimated parameter values. To aid the eye in identifying the sequential-firing activity, the bottom triad shows the estimation and prediction alone—with true simulation removed. The neuron order, stability of that order, constancy of the relative phases, spikes per burst for each neuron, and constancy of rotation rate are preserved. The predicted rotation rate is slightly faster, which may be due either to inherent chaos in the system (in which case even an excellent—but not exact—state estimate may yield a divergent outcome), or to specific parameter estimates that are not exact. The associated parameter estimates are in “Appendix B”.

Figure 4 shows the same result (of Fig. 3) with an extended x-axis. The series of activations can no longer be discerned by eye, but the figure is intended to demonstrate that the activity persists at a reliable rate for 2700 ms: roughly ten times longer than the required duration for the associated animal behavior described in Armstrongand and Abarbanel (2016). Result for simultaneous-firing mode is of similar quality (see Armstrong 2020).

3.4 Discussion

It is encouraging that sufficient information is contained in the cells’ voltage traces to yield accurate simultaneous estimates of cellular and synaptic properties. As noted, however, a biological circuit can employ multiple configurations that enact identical circuit output. We have shown (Armstrong et al. 2020) that, within the context of an astrophysical model, an SDA procedure can be designed to identify the existence of degenerate model solutions. For the case of a biological system, however, how might SDA discriminate between degeneracy due to insufficient measurements versus degeneracy that represents real redundancy in the system for the purpose of robustness? This remains a largely open question.

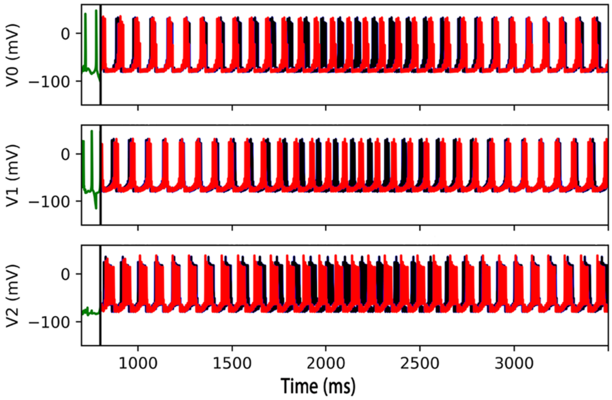

Fig. 4 Estimates and predictions of Fig. 3, with an extended x-axis, showing that the series-firing prediction continues at a stable—and slightly faster—rate of rotation (reproduced from Armstrong 2020)

4 COVID-19

The alarming events that unfolded in March 2020 surrounding COVID-19 compelled me to reach out to epidemiologists, via colleagues at Northwestern University, to ascertain whether SDA could inform predictive models for guiding policy on social behavior. Quickly I learned that it was important to remind the epidemiology community that this framework exists.

At the time, the coronavirus disease was burdening health care systems worldwide. Mathematical modeling had been used to aid policymakers' plans for hospital capacity needs, and to understand the minimum criteria for effective contact tracing (Murray et al. 2020). Insufficient testing capacity, however, especially at the beginning of the epidemic in the United States, and other data reporting issues, meant that surveillance data on COVID-19 was biased and incomplete. Moreover, the creation of a predictive epidemiological model tailored to COVID-19 had been hindered by gross uncertainties in the recordings of the populations affected by the disease (Heggeness 2020; Weinberger et al. 2020; Li et al. 2020).

Meanwhile, inference was not commonly used in epidemiology. It was presented roughly 10 years prior as a means for epidemiological forecasting (Bettencourt etal. 2007; Rhodesand and Hollingsworth 2009), and there had been two employments of SDA for examining COVID-19 specifically (Sesterhenn 2020; Nadler etal. 2020). None of these publications examined the sensitivity of estimates to contamination in measurements.

Within this context, my collaborators and I sought a means to quantify what data must be recorded in order to estimate specific unknown quantities in an epidemiological model of COVID-19 transmission, aiming to show how SDA can inform hospital policy if employed with real data.

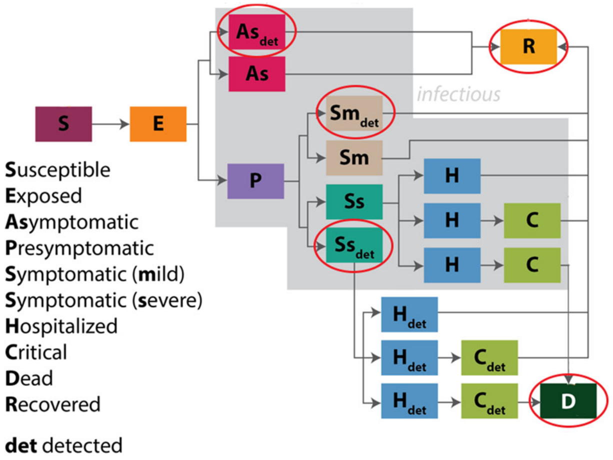

Fig. 5 Schematic of the model. Each rectangle represents a population. Note the distinction of asymptomatic cases, undetected cases, and the two tiers of hospitalized care: H and C. The aim of including this degree of resolution is to inform policy on social behavior so as to minimize strain on hospital capacity. The red ovals indicate the variables that correspond to measured quantities in the inference experiments (reproduced from Armstrong et al. 2021) (Color figure online)

4.1 Model

The standard “compartment" model in epidemiology consists of populations of people. The simplest is the susceptible-infectious-removed (SIR) model, where “removed" includes both recovered and dead populations.

Generally, the model is written as:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta I S}{N},\\ \frac{dI}{dt} &= \frac{\beta I S}{N} - \gamma I,\\ \frac{dR}{dt} &= \gamma I,\end{aligned} \tag{8}$$

where N is the sum of all three populations, β is the average number of contacts per person per time, multiplied by the probability of disease transmission in a contact between a susceptible and an infectious subject, and $1/\gamma$ is the average duration over which an infected person is contagious.

We expanded this model to capture key features of COVID-19, such that the structure is relevant for informing policy on social behavior and contact tracing so as to avoid exceeding hospital capacity. Figure 5 shows a schematic. There are 22 populations, including (i) asymptomatic, presymptomatic, and symptomatic populations, (ii) undetected and detected cases, and (iii) two hospitalized populations: those who do and do not require critical care. Absent additive noise in the measurements, the process is deterministic. The circled populations denote those given as measurements in the

SDA experiments. For details, see “Appendix C”, and for the complete equations of motion, see “Appendix A” of Ref. (Armstrong et al. 2021).

4.2 SDA Procedure

The measured state variables (the red-ovaled populations in Fig. 5) are: the detected asymptomatics, detected mild symptomatics, detected severe symptomatics, recovered, and dead. We simulated the time evolution of these populations based on initial conditions set by the onset of the pandemic in New York City in late winter of 2020 (see Armstrong et al. (2021) for details.) The time series lasted 101 days, with daily sampling.

We used these measurements—with varying noise levels added—to estimate the time course of the unmeasured populations over the same time period, as well as five model parameters taken to be unknown. Three are time-varying: (i) transmission rate $K_i(t)$, and (ii) detection probabilities of asymptomatic and symptomatic cases ($d_{Sym}(t)$ and $d_{Sys}(t)$. Two parameters taken to be intrinsic properties of the disease, and therefore static, are: the fraction of the exposed population that acquires symptoms (f_{sympt}), and the fraction of those symptomatics so severe that they require hospitalization (f_{severe}). We aimed to demonstrate the ability of SDA to quantify the accuracy to which these parameters can be estimated, given certain properties of the data including sampling rate, temporal baseline of sampling, and noise level. Note that this study differs from those of Sect. 3 in that the results are estimates, not predictions.

4.3 Result

With noiseless data, a baseline of 101 days, and measurements of susceptible, detected asymptomatics, detected mild symptomatatics, detected severe symptomatics, recovered and dead: the estimates of all unmeasured states and parameters were good, except for the time-varying transmission rate $K_i(t)$ (Fig. 6). Estimates of the static parameters f_{sympt} and f_{severe} were, respectively, 0.59 (true: 0.6) and 0.07 (true: 0.07). The action plot as a function of annealing parameter β (Fig. 7) levels off around $\beta = 30$, indicating that a solution compatible with both measurements and model has been found.

We next removed the measurement of the recovered population R and repeated the experiment. Estimates of most time-varying quantities are poor (Fig. 8) and the estimates of the static parameters f_{sympt} and f_{severe} failed to converge. The corresponding action plot increases indefinitely (Fig. 9), indicating that a solution compatible with both model and measurements is not found. This has worrisome implications for estimation with real data, as recovered people will not necessarily report the fact.

Further, adding low ($\sim 5\%$) Gaussian additive noise in the recovered population R yields poor estimates of all quantities. Doubling the temporal baseline of state variable measurements in that with-noise version yields good state variable estimates (Fig. 1, bottom panel) but poor parameter estimates (not shown). Finally, applied to real publicly available data (from Italy posted by Johns Hopkins (2020)), the procedure failed to converge.

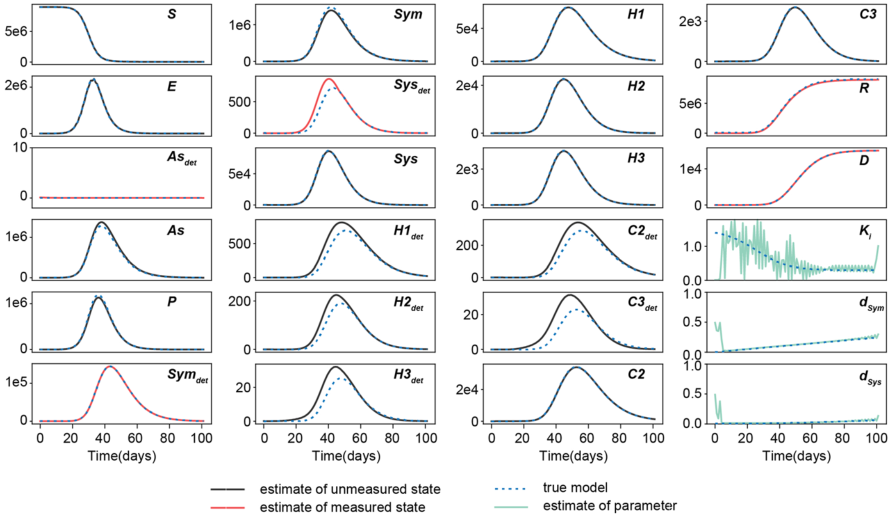

Fig. 6 Estimates of the state—measured and unmeasured—variables, the time-varying transmission rate $K_i(t)$, and detection rates of mild and severe symptomatics $d_{Sym}(t)$ and $d_{Sys(t)}$, respectively, where the y-axes are numbers of cases, and the measured populations are: As_{det}, Sm_{det}, Ss_{det}, R, and D. Excellent estimates are obtained of all states and parameters, except early values of K_i prior to the implementation of social distancing (reproduced from Armstrong et al. 2021)

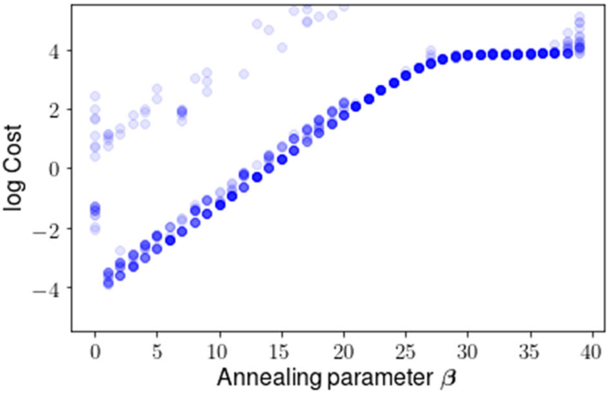

Fig. 7 Cost function plotted at each annealing step $\boldsymbol{\beta}$, corresponding to the state estimates of Fig. 6, for twenty paths in state space, where β scales the rigidity of the imposed model constraint. At low β the procedure endeavours to fit the measured variables to the simulated measurements. As β increases, the weight of the model dynamics is increasingly imposed and thus the cost increases—until it approaches a plateau (around $\beta = 30$), indicating that a solution has been found that is consistent with both measurements and model (reproduced from Armstrong et al. 2021)

4.4 Discussion

The high sensitivity of parameter estimates to contamination in the data indicated that we were in for a long and largely unpredictable journey with this disease, given the difficulties with accurate reporting (although we did not publish that statement in the paper). The failure of convergence for the real Italy data might also be due in-part to the

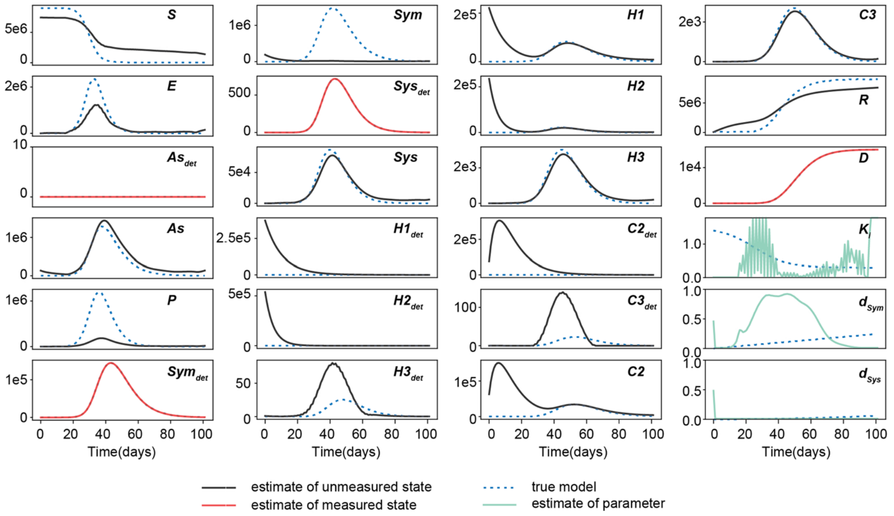

Fig. 8 Estimates akin to those of Fig. 6, without a measurement of Population ***R***. Results are significantly poorer (reproduced from Armstrong et al. 2021)

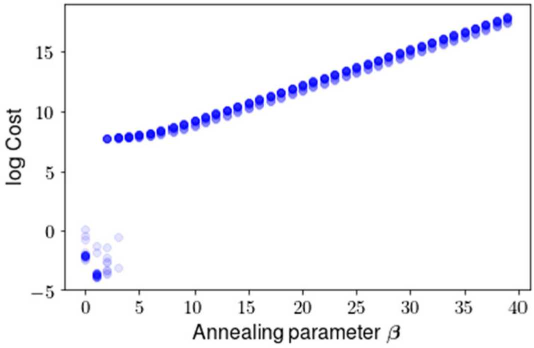

Fig. 9 Cost versus $\boldsymbol{\beta}$ without a measurement of the recovered population, over 20 trials. As β increases, the cost increases indefinitely, indicating that no solution has been found that is consistent with both measurements and model dynamics (reproduced from Armstrong et al. 2021)

fact that the SDA procedure assumes Gaussian-distributed errors in the measurements. That is not necessarily the case for this type of data.

Beyond that, it has been our aim to plant a seed in the epidemiology community: that SDA is a versatile and powerful tool for forecasting, specifically for the case wherein measurements are not only sparse but grossly contaminated. I am elated to see that, indeed, since our publication, the idea has taken stronger hold (Ghostine et al. 2021; Evensen 2020; Marinoschi 2021; Sun et al. 2021a, b).

5 Summary

We have explored applications of SDA over two distinct state variable distributions, measurement and sampling sets, and spatial and temporal scales. Several salient lessons emerge. First, the scale (i.e. choice of discretized step size) makes no difference to the SDA procedure. One must take care, however, to choose it wisely, given the competing interests of computational efficiency and model resolution. That is, there exists a tradeoff between the degree of biological detail (i.e. realism) and computational ease: the greater the number of degrees of freedom comprising the model, the greater the computational cost.

Second, the sensitivity of estimates and predictions to the sparsity of measurements varies widely across models. For the COVID-19 compartment model, with 5% additive noise in the Recovered population, the temporal baseline of measurements had to be doubled to maintain the quality of estimates of the unmeasured state variables. It could be worthwhile to conduct a formal study of how Shannon entropy flows among the state variables and parameters of a given dynamical system.

Third, and intimately related to the second point above: the sensitivity of estimates and predictions to contamination in measurements also varies widely across models. Predictions for the biological circuit weathered 10% additive noise well, while the COVID-19 model suffered severely with 5% noise added to just one of the five measured populations. That result might in-part be due to the SDA procedure's assumption of Gaussian-distributed noise in measurements—a limitation that should be loosened.

It should be noted that, for realistic biological systems, the main challenge for SDA is estimating errors on unknown model parameters. Parameter estimation is notoriously difficult, largely because parameters—unlike model state variables—do not obey a known dynamical law. Thus there exists no straightforward way to correlate errors in parameter estimates with state variable evolution or with errors in measurements (Carrassi and Vannitsem 2011). One must consider this carefully when tackling an optimization problem wherein the model is poorly constrained by measurements.

Finally, I refer the reader back to Table 1 for a summary of the applications of a path-integral inference methodology described in this paper. Examining applications outside one's specific discipline is an opportunity to view one's own through a slightly different filter, and a slightly fresher perspective.

Acknowledgements E. A. acknowledges NSF Grant 2139004.

Appendix A: The SDA Procedure

The specific computational implementation of the cost function used for the neuronal network model of Ref. Armstrong (2020) can be written as:

$$
\begin{aligned}
A_0 &= R_f A_{\text{model}} + R_m A_{\text{meas}} + A_{\text{other constraints}} \\
A_{\text{model}} &= \frac{1}{ND}\sum\nolimits_{n\in\{\text{odd}\}}^{N-2}\sum\nolimits_{a=1}^{D}\left[\left\{x_a(t_{n+2}) - x_a(t_n) - \frac{\delta t}{6}[F_a(\boldsymbol{x}(t_n),\boldsymbol{p}) + 4F_a(\boldsymbol{x}(t_{n+1}),\boldsymbol{p})\right.\right. \\
&\quad \left. + F_a(\boldsymbol{x}(t_{n+2}),\boldsymbol{p})]\right\}^2 \\
&\quad \left. + \left\{x_a(t_{n+1}) - \frac{1}{2}\left(x_a(t_n) + x_a(t_{n+2})\right) - \frac{\delta t}{8}[F_a(\boldsymbol{x}(t_n),\boldsymbol{p}) - F_a(\boldsymbol{x}(t_{n+2}),\boldsymbol{p})]\right\}^2\right] \\
A_{\text{meas}} &= \frac{1}{N_{\text{meas}}}\sum\nolimits_j\sum\nolimits_{l=1}^{L}(y_l(t_j) - x_l(t_j))^2.
\end{aligned}
\tag{9}
$$

Here, the "model error" term A_{model} is a simplification of the first term in Eq. 4 governing Markov-chain transition probabilities. The "measurement error" term A_{meas} is a simplification of the second term of Eq. 4 concerning the conditional mutual information.

A_{model} in Eq. (9) incorporates the model evolution of all D state variables x_a. Here, the outer sum on n is taken over all odd-numbered discretized radial locations of the model equations of motion. The sum on a is taken over all D state variables. A Simpson's rule method of finite differences is used to discretize the state space, so that the first and second terms for A_{model} represent the first and second derivative of the model, respectively.

A_{meas} governs the transfer of information from measurements y_l to model states x_l. Here, the summation on j runs over all discretized radial locations J at which measurements are made, which may be some subset of all integrated locations of the model. The summation on l is taken over all L measured quantities.

$A_{\text{other constraints}}$ are model-dependent.

To generate simulated data, the forward integration is performed by Python's odeINT package, which discretizes via an adaptive step size. The optimization is performed by the open-source Interior-point Optimizer (Ipopt) (Wächter 2009), which employs a Hermite-Simpson method of discretization and a constant step size. The discretization of state space, calculations of the model Jacobean and Hessian matrices, and the annealing procedure are performed via an interface with Ipopt that was written in C and Python; our complete computational implementation is available at the github repository of Ref. Lorenz (1963). Simulations currently run on a computing cluster equipped with 201 GB of RAM and 24 GenuineIntel CPUs (64 bits), each with 12 cores.

Appendix B: Neurobiology

Details of the Model

The $GHK(V_i(t), [Ca]_i(t))$ of Eq. 6 is:

$$GHK(V_i(t), [Ca]_i(t)) = V_i(t)\frac{[Ca]_i(t) - Ca_{ext}e^{-2FV_i(t)/RT}}{e^{-2FV_i(t)/RT} - 1}.$$

[Ca](t) is the intracellular Ca^{2+} concentration as a function of time. Ca_{ext} is the extracellular concentration of Ca^{2+} ions. In the GHK current, F is the Faraday constant, R is the gas constant, and T is temperature, which is taken to be 37 °C. The calcium dynamics evolve as:

$$\frac{\mathrm{d}[Ca_i](t)}{\mathrm{d}t} = \phi_i I_{CaT_i} + \frac{Ca_{0,i} - [Ca_i](t)}{\tau_{Ca,i}}.$$

Ca_0 is the equilibrium concentration of calcium inside the cell, and ϕ is a constant that summarizes the effects of volume and surface area.

The gating variables $U_i(t)$ = m(t), h(t), n(t), a(t), b(t) of Eq. 6 satisfy:

$$\begin{aligned}
\frac{\mathrm{d}U_i(t)}{\mathrm{d}t} &= (U_\infty(V_i(t)) - U_i(t))/\tau_{Ui}(V_i(t)) \\
U_\infty(V_i) &= 0.5[1 + \tanh((V_i - \theta_{U,i})/\sigma_{U,i})] \\
\tau_{Ui}(V_i) &= t_{U0} + t_{U1}[1 - \tanh^2((V_i - \theta_{U,i})/\sigma_{U,i})].
\end{aligned} \tag{10}$$

In Eq. 7 for the synapses, the gating variables s_{ij}(t) evolve as:

$$\begin{aligned}
\frac{\mathrm{d}s_{ij}(t)}{\mathrm{d}t} &= \nu T(V_j(t))[1 - s_{ij}(t)] - \gamma s_{ij}(t) \\
T(V_j(t)) &= \frac{T_{max}}{1 + \exp(-(V_j(t) - V_P)/K_P)}.
\end{aligned} \tag{11}$$

The rate constants ν and γ have units of 1/time. V_P and K_P are parameters governing the shape of the distribution of neurotransmitter rise and fall as it drives gating variables s_{ij}. The neurons and synapses are distinguishable via different values of all electrophysiological and kinetic parameters. For a list of the parameters that were taken to be known and fixed during the SDA procedure, see Ref. Armstrong (2020).

Details of the SDA Procedure for the Network Model

The simulated data were generated from two model versions, each defined by a unique set of synapse strengths g_{ij}. The first set corresponds to a mode that—when the injected current I_{inj} is a steady, low-noise back ground current—expresses sequential activations of the interneurons. These values are on the order of 0.1 μS. The second set of g_{ij} corresponds to a mode that—with the same steady injected current—expresses simultaneous firing of the neurons. These values are an order of magnitude lower: roughly 0.01 μS.

During the estimation window, the cells received three distinguishable input currents $I_{inj,i}$: the x-, y-, and y-phase-offset output of a chaotic Lorenz-63 model (Lorenz 1963), with steps spliced into each current at intermittent locations. The integration time step for the simulated data and the time step of measurement sampling was 0.1 ms; the estimation window was 799 ms.

During the prediction window, the estimated model was exposed to two novel currents: (1) a continuation of the chaotic Lorenz-63 output used in estimation; (2) the low-noise step that is known to produce either synchronous or sequential firing, depending on the synapse maximum conductances g_{ij}. For the first, second, and third neuron, this injected background current was: 0.4, 0.5, and 0.3 nA, respectively (Table 2).

For each of the two functional modes of circuit activity, multiple versions of the experiment were performed, each version employing a distinct set of measurements: (1) membrane voltage of all three neurons; (2) membrane voltage of two out of the three neurons; (3) calcium concentration of all three neurons.

Table 2 Estimates for g_{ij} corresponding to sequential bursting

Parameter	Estimated value	Correct value	Lower bound	Upper bound
E_{01}	− 82.98	− 83.0	− 90.0	10.0
E_{02}	− 83.29	− 83.3		
E_{10}	− 82.67	− 82.7		
E_{12}	− 82.57	− 82.5		
E_{20}	− 83.26	− 83.2		
E_{21}	− 82.88	− 82.9		
g_{01}	0.248	0.25	0.01	10.0
g_{02}	0.403	0.4		
g_{10}	0.283	0.28		
g_{12}	0.177	0.18		
g_{20}	0.211	0.21		
g_{21}	0.314	0.32		
$g_{L,0}$	2.88e−3	3.0e−3	9e−4	9e−2
$g_{L,1}$	3.25e−3	3.3e−3		
$g_{L,2}$	2.87e−3	2.9e−3		
$g_{Na,0}$	1.18	1.2	0.2	1.8
$g_{Na,1}$	0.96	1.0		
$g_{Na,2}$	1.24	1.4		
$g_{K,0}$	0.197	0.2	0.02	0.8
$g_{K,1}$	0.210	0.22		
$g_{K,2}$	0.150	0.17		
$g_{CaT,0}$	1.01e−4	1.0e−4	e−5	e−2
$g_{CaT,1}$	1.16e−4	1.1e−4		
$g_{CaT,2}$	1.16e−4	9.0e−5		

The columns are: *Estimated value*: parameter estimation from the D.A. procedure; *Correct value*: value used to generate the simulated data that was provided to the D.A. procedure; *Lower bound*: User-imposed lower bound on the parameter value, for the D.A. procedure; *Upper bound*: user-imposed upper bound. Note that the bounds used for the reversal potentials E_{ij} permit the possibility that synapses are either excitatory or inhibitory. Units: reversal potentials are in mV; ion channel and synapse maximum conductances are in μS. Notation: g_{ij} denotes the weight of the synapse entering cell i from cell j; $g_{L,i}$ denotes the value of leak current in cell i. Estimates were obtained for annealing parameter values: $R_m = 1$, $R_{f,0} = 0.01$, $\alpha = 1.5$, and $\beta = 27$. Forty paths were searched, all of which converged to this solution (reproduced from Armstrong 2020)

Appendix C: Details of the Epidemiological Model for COVID-19

The resolution of asymptomatic versus symptomatic cases was motivated by an interest in what interventions are necessary to control the epidemic. For example, is it sufficient to focus only on symptomatic individuals, or must we also target and address asymptomatic individuals who may not even realize they are infected?

The detected and undetected populations exist for two reasons. First, we sought to account for underreporting of cases and deaths. Second, we desired a model structure that can simulate the impact of increasing detection rates on disease transmission, including the impact of contact tracing. Thus the model was structured from the beginning so that we might examine the effects of interventions that were imposed later on. The ultimate aim was to inform policy on the requirements for containing the epidemic.

We included both H and C populations (patients in the hospital requiring / not requiring ICU care, respectively) because hospital inpatient and ICU bed capacities were the key health system metrics that we aimed to avoid straining. Any policy considered must include predictions on inpatient and ICU bed needs. Preparing for those needs is a key response if or when the epidemic grows uncontrolled.

References

Abarbanel H (2013) Predicting the future: completing models of observed complex systems. Springer, Berlin

Abarbanel HD, Shirman S, Breen D, Kadakia N, Rey D, Armstrong E, Margoliash D (2017) A unifying view of synchronization for data assimilation in complex nonlinear networks. Chaos Interdiscip J Nonlinear Sci 27:126802

An Z, Rey D, Ye J, Abarbanel HD (2017) Estimating the state of a geophysical system with sparse observations: time delay methods to achieve accurate initial states for prediction. Nonlinear Process Geophys 24:9–22

Armstrong E (2020) Statistical data assimilation for estimating electrophysiology simultaneously with connectivity within a biological neuronal network. Phys Rev E 101:012415

Armstrongand E, Abarbanel HD (2016) Model of the songbird nucleus HVC as a network of central pattern generators. J Neurophysiol 116:2405

Armstrong E, Patwardhan AV, Johns L, Kishimoto CT, Abarbanel HD, Fuller GM (2017) An optimization-based approach to calculating neutrino flavor evolution. Phys Rev D 96:083008

Armstrong E, Patwardhan AV, Rrapaj E, Ardizi SF, Fuller GM (2020) Inference offers a metric to constrain dynamical models of neutrino flavor transformation. Phys Rev D 102:043013

Armstrong E, Runge M, Gerardin J (2021) Identifying the measurements required to estimate rates of COVID-19 transmission, infection, and detection, using variational data assimilation. Infect Dis Model 6:133

Bettencourt LM, Ribeiro RM, Chowell G, Lant T, Castillo-Chavez C (2007) Towards real time epidemiology: data assimilation, modeling and anomaly detection of health surveillance data streams. In: NSF workshop on intelligence and security informatics. Springer, pp 79–90

Betts JT (2010) Practical methods for optimal control and estimation using nonlinear programming, vol 19. SIAM, Philadelphia

Breen D, Shirman S, Armstrong E, Kadakia N, Abarbanel H (2016) HVC interneuron properties from statistical data assimilation. arXiv:1608.04433

Carrassi A, Vannitsem S (2011) State and parameter estimation with the extended Kalman filter: an alternative formulation of the model error dynamics. Q J R Meteorol Soc 137(655):435–451

Daou A, Ross MT, Johnson F, Hyson RL, Bertram R (2013) Electrophysiological characterization and computational models of HVC neurons in the zebra finch. J Neurophysiol 110:1227

de Maupertuis P (1750) Essai de cosmologie (Amsterdam, 1750); Accord de différentes lois de la nature qui avaient jusqu'ici paru incompatibles.(1744) Mém As Sc Paris 417

Destexhe A, Sejnowski TJ (2001) Thalamocortical assemblies: how ion channels, single neurons and large-scale networks organize sleep oscillations. Oxford University Press, Oxford

Destexhe A, Mainen ZF, Sejnowski TJ (1994) Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. J Comput Neurosci 1:195

Euler L (1744) Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes. apud Marcum-Michaelem Bousquet

Evensen G (2009) Data assimilation: the ensemble Kalman filter. Springer, Berlin

Evensen G, Amezcua J, Bocquet M, Carrassi A, Farchi A, Fowler A, Houtekamer PL, Jones CK, de Moraes RJ, Pulido M et al (2020) An international assessment of the COVID-19 pandemic using ensemble data assimilation. medRxiv

Fano RM (1961) Transmission of information: a statistical theory of communications. Am J Phys 29:793

Gerhardt CI (1898) Ueber die vier Briefe von Leibniz, die Samuel König in dem Appel au public. Leide MDCCLIII, veröffentlich hat

Ghostine R, Gharamti M, Hassrouny S, Hoteit I (2021) An extended SEIR model with vaccination for forecasting the COVID-19 pandemic in Saudi Arabia using an ensemble Kalman filter. Mathematics 9:636

Github Repository for Optimization Procedures. https://github.com/AA-Ahmetaj/SLURM_minAone. Accessed 11 Oct 2022

Grashow R, Brookings T, Marder E (2009) Reliable neuromodulation from circuits with variable underlying structure. Proc Natl Acad Sci 106:11742

Grashow R, Brookings T, Marder E (2010) Compensation for variable intrinsic neuronal excitability by circuit-synaptic interactions. J Neurosci 30:9145

Gunaratne CA, Sakurai A, Katz PS (2017) Variations on a theme: species differences in synaptic connectivity do not predict central pattern generator activity. J Neurophysiol 118:1123–1132

Hamilton F, Berry T, Peixoto N, Sauer T (2013) Real-time tracking of neuronal network structure using data assimilation. Phys Rev E 88:052715

Heggeness M (2020) The need for data innovation in the time of covid-19. https://www.minneapolisfed.org/article/2020/the-need-for-data-innovation-in-the-time-of-covid-19. Accessed 17 May 2020

Hodgkin AL, Huxley AF, Katz B (1952) Measurement of current-voltage relations in the membrane of the giant axon of Loligo. J Physiol 116:424

Johns Hopkins University Center for Systems Science and Engineering. Novel coronavirus cases. https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series. Accessed 14 May 2020

Kabitz W (1913) Über eine in Gotha aufgefundene Abschrift des von S. König in seinem Streite mit Maupertuis und der Akademie veröffentlichten, seinerzeit für unecht erklärten Leibnizbriefes

Kadakia N, Armstrong E, Breen D, Morone U, Daou A, Margoliash D, Abarbanel HD (2016) Nonlinear statistical data assimilation for HVC_{RA} neurons in the avian song system. Biol Cybern 110:417

Kalnay E (2003) Atmospheric modeling, data assimilation and predictability. Atmospheric modeling, data assimilation and predictability. Cambridge University Press, Cambridge

Kimura R (2002) Numerical weather prediction. J Wind Eng Ind Aerodyn 90:1403

Kostuk M, Toth BA, Meliza CD, Margoliash D, Abarbanel HD (2012) Dynamical estimation of neuron and network properties II: path integral Monte Carlo methods. Biol Cybern 106:155

Kristan WB Jr, Calabrese RL, Friesen WO (2005) Neuronal control of leech behavior. Prog Neurobiol 76:279

Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, Shaman J (2020) Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science 368:489

Lorenz EN (1963) Deterministic nonperiodic flow. J Atmos Sci 20:130

Marder E, Calabrese RL (1996) Principles of rhythmic motor pattern generation. Physiol Rev 76:687

Marder E, Bucher D, Schulz DJ, Taylor AL (2005) Invertebrate central pattern generation moves along. Curr Biol 15:R685

Marder E, Tobin A-E, Grashow R (2007) How tightly tuned are network parameters? Insight from computational and experimental studies in small rhythmic motor networks. Prog Brain Res 165:193

Marder E, Gutierrez G, Nusbaum MP (2016) Complicating connectomes: electrical coupling creates parallel pathways and degenerate circuit mechanisms. Dev Neurobiol 77:597–609

Marinoschi G (2021) Parameter estimation of an epidemic model with state constraints. Appl Math Optim 84:1903

Meliza CD, Kostuk M, Huang H, Nogaret A, Margoliash D, Abarbanel HD (2014) Estimating parameters and predicting membrane voltages with conductance-based neuron models. Biol Cybern 108:495

Mulloneyand B, Hall WM (2007) Local and intersegmental interactions of coordinating neurons and local circuits in the swimmeret system. J Neurophysiol 98:405

Murray CJ, Alamro NMS, Hwang H, Lee U (2020) Digital public health and COVID-19. Lancet Public Health 5(9):e469–e470

Nadler P, Wang S, Arcucci R, Yang X, Guo Y (2020) An epidemiological modelling approach for Covid19 via data assimilation. arXiv:2004.12130

Nogaret A, Meliza CD, Margoliash D, Abarbanel HD (2016) Automatic construction of predictive neuron models through large scale assimilation of electrophysiological data. Sci Rep 6:1

Odenand JT, Reddy JN (2012) Variational methods in theoretical mechanics. Springer, Berlin

Quinnand JC, Abarbanel HD (2011) Data assimilation using a GPU accelerated path integral Monte Carlo approach. J Comput Phys 230:8168

Restrepo JM (2008) A path integral method for data assimilation. Phys D Nonlinear Phenom 237:14

Rhodesand C, Hollingsworth TD (2009) Variational data assimilation with epidemic models. J Theor Biol 258:591

Sakuraiand A, Katz PS (2016) The central pattern generator underlying swimming in Dendronotus iris: a simple half-center network oscillator with a twist. J Neurophysiol 116:1728

Schiff SJ (2009) Kalman meets neuron: the emerging intersection of control theory with neuroscience. In: 2009 annual international conference of the IEEE engineering in medicine and biology society. IEEE, pp 3318–3321

Sesterhenn JL (2020) Adjoint-based data assimilation of an epidemiology model for the Covid-19 pandemic in 2020. arXiv:2003.13071

Smarandache C, Hall WM, Mulloney B (2009) Coordination of rhythmic motor activity by gradients of synaptic strength in a neural circuit that couples modular neural oscillators. J Neurosci 29:9351

Sun Q, Richard S, Miyoshi T (2021a) Analysis of COVID-19 in Japan with Extended SEIR model and ensemble Kalman filter. arXiv:2110.14892

Sun C, Richard S, Miyoshi T (2021b) Agent-based model and data assimilation: analysis of COVID-19 in Tokyo. arXiv:2109.00258

Tarantola A (2005) Inverse problem theory and methods for model parameter estimation. SIAM, Philadelphia

Toth BA, Kostuk M, Meliza CD, Margoliash D, Abarbanel HD (2011) Dynamical estimation of neuron and network properties I: variational methods. Biol Cybern 105:217

Turrigiano G (2011) Too many cooks? Intrinsic and synaptic homeostatic mechanisms in cortical circuit refinement. Annu Rev Neurosci 34:89

Wächter A (2009) Short tutorial: getting started with ipopt in 90 minutes. In: Dagstuhl Seminar Proceedings. Schloss Dagstuhl-Leibniz-Zentrum für Informatik

Weinberger D, Cohen T, Crawford F, Mostashari F, Olson D, Pitzer VE, Reich NG, Russi M, Simonsen L, Watkins A et al (2020) Estimating the early death toll of COVID-19 in the United States. Medrxiv

Whartenby WG, Quinn JC, Abarbanel HD (2013) The number of required observations in data assimilation for a shallow-water flow. Mon Weather Rev 141:2502

Ye J, Rey D, Kadakia N, Eldridge M, Morone UI, Rozdeba P, Abarbanel HD, Quinn JC (2015) Systematic variational method for statistical nonlinear state and parameter estimation. Phys Rev E 92:052901