HillTau: A fast, compact abstraction for model reduction in biochemical signaling networks.

Short title: HillTau: A fast, compact abstraction for modeling signaling pathways

Upinder S. Bhalla[1*]

[1.] NCBS-TIFR, Bangalore, Karnataka, India

* Corresponding Author

E-mail: bhalla@ncbs.res.in

Abstract

Signaling networks mediate many aspects of cellular function. The conventional, mechanistically motivated approach to modeling such networks is through mass-action chemistry, which maps directly to biological entities and facilitates experimental tests and predictions. However such models are complex, need many parameters, and are computationally costly. Here we introduce the HillTau form for signaling models. HillTau retains the direct mapping to biological observables, but it uses far fewer parameters, and is 100 to over 1000 times faster than ODE-based methods. In the HillTau formalism, the steady-state concentration of signaling molecules is approximated by the Hill equation, and the dynamics by a time-course *tau*. We demonstrate its use in implementing several biochemical motifs, including association, inhibition, feedforward and feedback inhibition, bistability, oscillations, and a synaptic switch obeying the BCM rule. The major use-cases for HillTau are system abstraction, model reduction, scaffolds for data-driven optimization, and fast approximations to complex cellular signaling.

2

Author summary

Chemical signals mediate many computations in cells, from housekeeping functions in all cells to memory and pattern selectivity in neurons. These signals form complex networks of interactions. Computer models are a powerful way to study how such networks behave, but it is hard to get all the chemical details for typical models, and it is slow to run them with standard numerical approaches to chemical kinetics. We introduce HillTau as a simplified way to model complex chemical networks. HillTau models condense multiple reaction steps into single steps defined by a small number of parameters for activation and settling time. As a result the models are simple, easy to find values for, and they run quickly. Remarkably, they fit the full chemical formulations rather well. We illustrate the utility of HillTau for modeling several signaling network functions, and for fitting complicated signaling networks.

Keywords: Simulation, systems biology, mass-action, synaptic plasticity

3

Introduction

John von Neumann's elephant haunts mechanistically detailed models. von Neumann was reported to have claimed that he could fit an elephant with 4 parameters, with the implication that models with many parameters are under-constrained and over-fitted (Mayer, Khairy, and Howard 2010). There are two major arguments to exorcise this elephant: that mechanistic detail is needed to address certain kinds of questions; and that in the era of big data it is both easier and less biased to simply build up detailed models with all the available pieces. Here we describe a model formalism, the HillTau form, to help navigate between biological mechanisms and big data on the one hand, and the desirability of condensed model representations that expose the key principles of system function.

Cellular, and particularly synaptic signaling, is notoriously complex. There are an estimated 1400 protein species localized to the postsynaptic density alone (Bayés et al. 2011). These support a range of functions including synaptic transmission, maintenance, plasticity, activity-driven protein synthesis, metabolic control, and traffic (Upinder S. Bhalla 2014).

Mass-action chemistry is a common denominator for mechanistically inspired modeling of these phenomena. This has the key virtue of defining specific biological entities (molecules) and processes (reactions) that map directly to experimental observables. Many studies are based at this level (U S Bhalla and Iyengar 1999; Shouval, Bear, and Cooper 2002). Further levels of mechanistic detail include reaction-diffusion, stochastic chemistry mesoscopic stochastic methods with trapezoidal or cubic meshes (Wils and De Schutter 2009; Oliveira et al. 2010) and even single-particle reaction-diffusion calculations (Stiles and Bartol 2001; Andrews et al. 2010). Note that the additional mechanistic detail comes at a considerable computational cost.

A few studies have found ways to lessen the level of detail, typically by focusing on interactions

without dynamics (e.g, (Sorokina, Sorokin, and Armstrong 2011)) or on dynamics with highly reduced interactions (e.g.,(Barak and Tsodyks 2006)). Model detail may also be abstracted out through model reduction, which starts from a detailed (usually mass-action or Michaelis-Menten ODE form) model and strips it down to core interactions needed to account for model behavior. Another reduction approach is to identify 'fast' reactions in the system, which settle much faster than the overall system, and can be replaced with algebraic relations(Hoops et al. 2006; Deuflhard and Heroth 1996). There is a substantial body of work on model reduction techniques (reviewed by Snowden (Snowden, van der Graaf, and Tindall 2017)). Most of these methods retain the chemical kinetics formalism using ordinary differential equations (ODEs) to represent mass-action.

Biochemical signaling models frequently suffer from incomplete parameterization. Thus 'detailed' models of signaling pathways, which are of course essential for many kinds of mechanistic analyses and design of experiments, are often under-constrained. In this context, a reduced model is preferable as it requires fewer parameters. While some of the formal approaches to this yield very compact models, the mapping to experimental observables may be quite indirect (Snowden, van der Graaf, and Tindall 2017). Hence it is useful to have a compact chemically-inspired formulation to serve as the core for the model reduction while remaining easy to parameterize and predict using the same quantities that are measured in experiments (Maurya et al. 2005; Danø et al. 2006; Taylor, Doyle, and Petzold 2008). Indeed, a compact model with few parameters is arguably a better starting point to understand complex signaling with insufficient data, than is a mechanistically detailed model.

Savageau and colleagues have developed the Design Space Toolbox to facilitate a systematic approach to developing reduced signaling and transcriptional network models with specified properties such as multistability (Lomnitz and Savageau 2016). They cast mechanistic models into a Generalized Mass Action form, and this is then analyzed to realize the required phenotypic repertoire. While this is an effective way of obtaining models with desired multi-state properties, it

differs in objectives from our goal of having a reduced, very efficient representation of dynamic responses of complex reaction networks such as synaptic signaling.

Efficiency is a specific constraint in developing models of synaptic signaling. On the one hand, many neural functions depend on the nuances of signaling. For example, network properties are quite sensitive to different plasticity rules (Dan and Poo 2004), neuromodulators (Roelfsema and Holtmaat 2018), and mutations (Südhof 2017). Network models also are expanding to include diffusible messengers controlling cellular activity and blood flow (Dormanns, Brown, and David 2016). At the single-cell level, explorations of receptor insertion and clustering (Hudmon et al. 2005; Wilson et al. 2016), sequence recognition (Upinder S Bhalla 2017) and synaptic tagging (Frey and Morris 1997; Smolen, Baxter, and Byrne 2012) all require some level of reference to the chemical signaling. The crux of the problem arises when these studies need to scale beyond one synapse to whole-neuron (up to 10^4 synapses, (Upinder S Bhalla 2017)) or even network scales (e.g., 10^9 synapses (Markram et al. 2015)). Clearly, efficiency in memory and computations is important for such models.

The HillTau form addresses several key concerns with modeling of complex signaling networks. It utilizes only those observable states specified by the user to map directly to the chemistry, thus supporting sparse models that are easier to constrain with limited data. This requires very few parameters, yet behaves similarly to chemical cascades involving multiple intervening steps. Since the user specifies their chosen observables, each can be related directly to observations of concentration over time. The models are small and calculations are highly efficient, being closed-form and event-driven.

6

Results

We first provide an overview of the HillTau algorithm. Then we illustrate its use to approximate increasingly complex reaction networks. We then show how one can reduce a mass-action model to its HillTau equivalent, with a tradeoff of greater complexity for better accuracy. Finally we carry out some benchmarks of several reduced HillTau models against the original ODE-chemical kinetic models run on two simulators, MOOSE (Ray and Bhalla 2008) and COPASI (Hoops et al. 2006), and show that HillTau is orders of magnitude faster.

Overview of HillTau algorithm

In brief, a ‘reaction’ in HillTau uses a form derived from the Hill equation to estimate steady-state values of the product(s) of one or several chemical reactions, and then computes how the system approaches this steady-state through an exponential function of time. The motivation for this formalism is that the steady-state value of a cascade of binding reactions, or of enzyme reactions with a fixed rate back-reaction, can be approximated by a Hill function (Methods). Further, the time-course of approach to steady-state is typically governed by the slowest reaction, and this can be approximated as an exponential settling function (Methods). The name HillTau refers to this conjunction of the Hill form and *tau*, the time-course for settling. Since both these calculations are analytic, one can do this calculation in an event-driven manner. HillTau achieves sparseness and simplicity by approximating many steps with a single ‘reaction’, considering only those intermediates that are needed for readouts or for improved precision. It achieves speed because the models are smaller, and by using event-driven calculations rather than numerical integration.

Most reaction networks cascade through many layers of reactions. HillTau evaluates each upstream layer before downstream ones. It first builds a dependency graph of all reactions. This is done by identifying input molecules as layer 0, and successively ranking all reactions that depend only on layer 0 as layer 1, reactions that depend on layers 0 to 1 as layer 2 and so on.

7

HillTau identifies feedback loops by reactions which do not resolve into the above layers. Based on ordering of reactions in the model definition, it picks a reaction to 'break' the loop, and assigns it to layer N+1, where N was the previously deepest layer. It then repeats the process of layer assignment, including further loop-breaking if needed.

During evaluation of a single step in HillTau, all the steady-state and time-course calculations are completed for layer 1, then layer 2 is calculated, and so on. Thus each layer receives the inputs appropriate to the current time before doing its evaluation. This approach yields outputs that are substantially insensitive to sampling and event intervals provided there is no feedback. If there is feedback, then one has to use event intervals shorter than the time-course around the feedback loop. A factor of 10 usually gives good convergence.

In summary, HillTau uses analytic evaluation of reaction outputs based on a Hill-like form and exponential settling, and propagates the evaluation through successive layers of the reaction network for each event time. Events can be stimuli or points in a time-series for sampling system time-evolution.

The HillTau form can model a range of chemical signaling motifs

We implemented a range of elementary chemical signaling functions to illustrate the use of the HillTau form (Methods, Figure 1). The HillTau versions of most of these reactions have an exact fit to their mass-action counterparts (Supplementary Figure S1). We further implemented key signaling motifs, including feedback inhibition, oscillation, and bistables (Figure 2). To do this, we constructed minimal HillTau schemes that incorporated the essential elements of each of these motifs. We developed an optimizer program *mash.py* (MASH: Model Abstraction from SBML to HillTau, see Methods) to tune parameters of the HillTau models to match the outputs to the original mass-action or ODE versions. We used normalized RMS difference between the traces as a measure of goodness of fit. In Fig. 2A-C we compare feedback inhibition implemented in mass-action (5 reactions, 7 species, 2A), HillTau (2 reactions, 3 species, 2B), and run for a square pulse input (2C).

The feedback inhibition model is well approximated by the HillTau version to within 4% normalized RMS deviation.

Next, we implemented a HillTau version of a mitogen-activated protein kinase (MAPK) feedback oscillation model having 11 reactions and 15 species, Fig 2E (Kholodenko 2000). The reduced HillTau version (3 reactions, 4 species, fig 2F) had similar period and amplitude (Fig 2 D), and it fit the waveform to within ~7.6% normalized RMS deviation.

Finally, we made a HillTau version of a chemical bistable switch using just 2 reactions (Fig 2 G). We demonstrated that the HillTau form works with the standard dose-response (null-cline) approach to estimating steady states (Figure 2 H) and showed that the resulting switch exhibits high and low states that are triggered by transient inputs (Figure 2 I).

Thus the HillTau form can efficiently represent a range of important signaling motifs and their dynamics, including feedback inhibition. oscillations, and bistability.

9

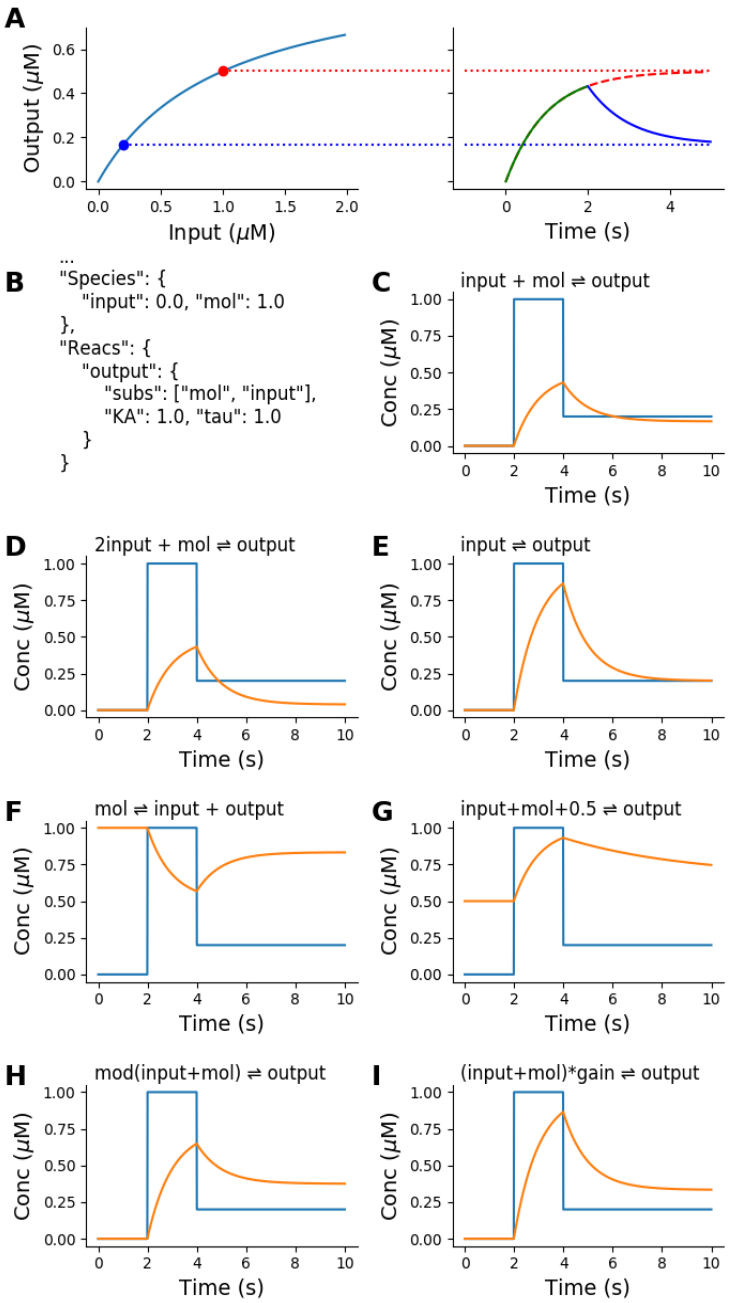

Figure 1. The HillTau formulation and representation of elementary chemical reactions by a single HillTau reaction. A: Principle of HillTau formulation. Left: steady-state output values for different levels of the input molecule, computed by the Hill equation. In all simulations in this figure, two input values are used: first 1 μM (red dot) and later 0.2 μM (blue dot). Right: The simulator starts from the current value of the output, and computes the approach to the steady-state as an exponential time-course. Note that these are algebraic calculations, not numerical integration. In this example the output rises from zero toward the red dotted line for 2 seconds. Then the input is changed to 0.2 μM, and now the simulator approaches the steady-state value for this (blue dotted line) with an exponential time-course. B: Key

section of the JSON code defining this reaction system. C-G: Inputs (blue) and simulated time-course of outputs (orange) for seven different reactions. Each is represented by a similar HillTau reaction but with different parameters (see Supplementary Material). In all cases the HillTau output onset is identical to the output computed using numerical integration of a single reaction expressed as mass-action chemical kinetics. Decay time-course may differ from onset time-course in mass-action. C: Binding. D: 2nd order binding. E: Conversion. F: Inhibition, conceptually equivalent to removal of output molecules by binding of input to the output molecule, and sequestration of the resultant complex. G: Variant of binding reaction, in which there is a fixed baseline of 0.5 μM, and the system has different on (tau = 1s) and off (tau2 = 5s) time courses. H. Same as reaction 1, but with a modifier term that strengthens input affinity. I. Same as reaction 1, but with a gain term that multiplies the output, in this case by a factor of 2.

11

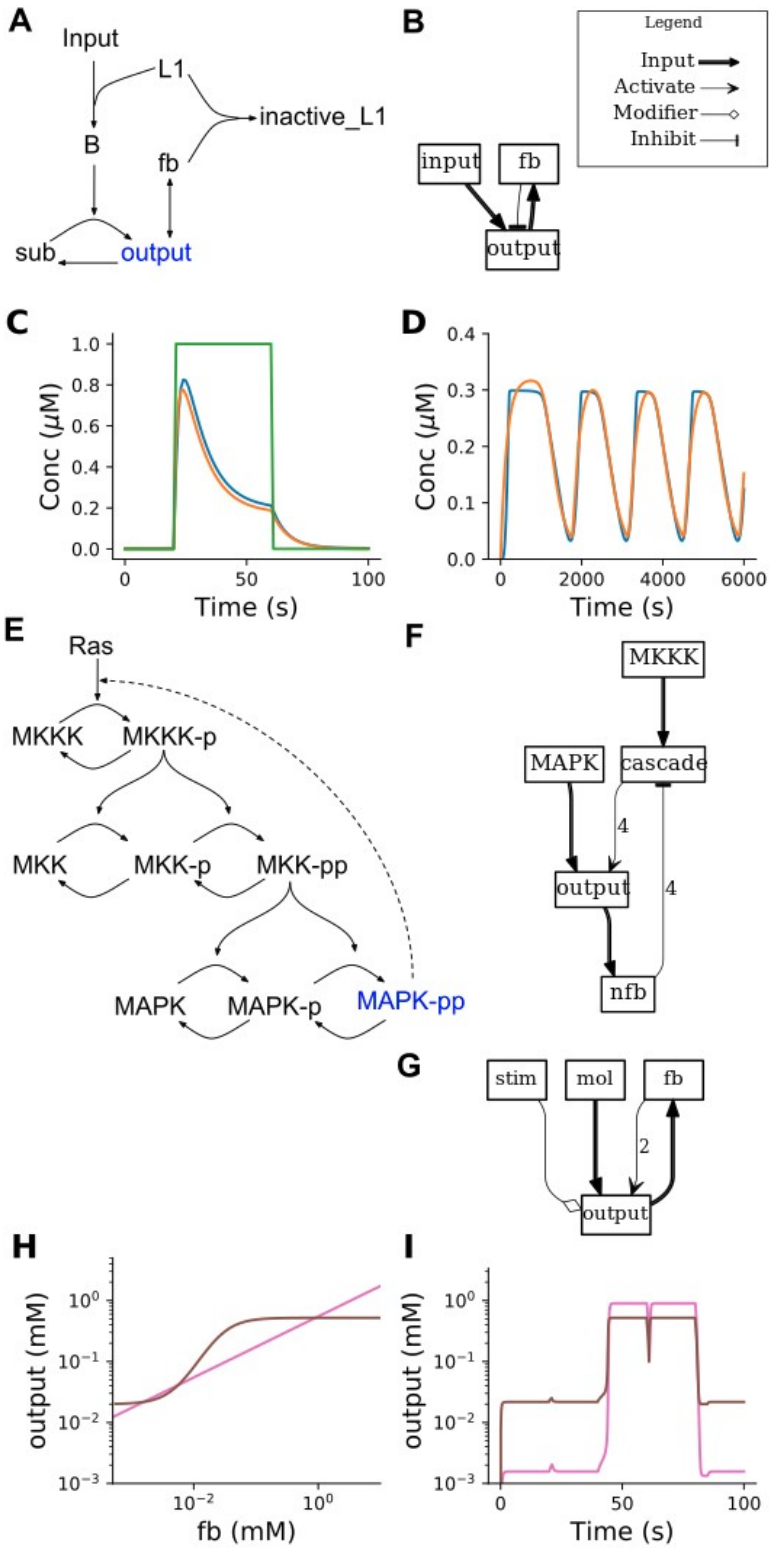

Figure 2. HillTau models of key signaling motifs. A-C: Feedback inhibition. A: Mass-action reaction scheme for feedback inhibition, involving 7 molecules and 5 reactions. B: HillTau version. Each box represents a molecule. If there are input arrows to the box it means there is a reaction whose product is the named molecule. Input arrows can be either

12

inputs (reagents), activators, inhibitors, or modifiers. This reaction consists of 3 molecules (input, fb, and output) and 2 reactions (fb and output). C: Simulations for mass-action (blue) and HillTau (orange) versions of feedback inhibition. The green trace is the input molecule. D-F: Oscillator from ultrasensitive MAPK cascade, taken from (Kholodenko 2000). D: Output of simulation. Blue is ODE output and orange is HillTau. E: ODE model. This uses 15 molecules, and 11 reactions. MAPK-pp is the molecular species used as output of the oscillator. F: HillTau reaction scheme for oscillator, using 5 molecules and 3 reactions. The concentration of the 'output' molecule is plotted. G: HillTau model of bistable system, involving 4 molecules and 2 reactions. H: Phase plot showing stable states of system as the intersection points between the steady-state dose-response curves. This was generated by varying the feedback molecule *fb*, and measuring *output* (brown curve), and then varying the *output* molecule and measuring *fb* (pink curve). I: Time-series illustration of state switching in the bistable. As before, *output* is in brown and *fb* in pink. The Y axes of H and I are the same to show that the steady-state output levels (brown) match. The system starts in the low state. At 20 s a small excitatory input *stim* is given which fails to switch the state. At 40 s a strong input causes switching to the high state. At 60 s a weak inhibitory input fails to turn it off, but at 80 s a strong inhibitory input returns the state to baseline. Excitatory and inhibitory inputs were delivered by transiently setting the level of *stim* to high or low values.

13

The HillTau form compactly represents bidirectional synaptic plasticity

Synaptic plasticity is one of the most-modeled neuronal signaling processes (U S Bhalla and Iyengar 1999; Lisman 1989). The key features that have been represented include stimulus strength-dependence, timing dependence, and long-term state storage (Upinder S Bhalla 2014). A few studies have come up with rather detailed models to implement each of these processes (Smolen, Baxter, and Byrne 2012; Hayer and Bhalla 2005; Kim et al. 2013). As an illustration of all these properties in the HillTau system, we implemented bidirectional synaptic plasticity including long-term synaptic state changes (Figure 3). One of the interesting aspects of synaptic plasticity is that in many systems, the same input modality (typically read out as Ca^{2+} concentration) can give rise to both synaptic depression and potentiation. This has significant theoretical implications and an abstract rule for this bidirectional plasticity was proposed by Bienenstock et al (the BCM rule, (Bienenstock, Cooper, and Munro 1982)). We first devised a simplified mass-action version of the BCM rule using 9 molecules and 6 reactions (Figure 3A). The species p_AMPAR is the phosphorylated form of the receptor, assumed to be inserted into the synapse. Here, resting Ca^{2+} does not alter the state of the model; low Ca^{2+} causes depotentiation (that is, reduction of receptor levels), and high Ca^{2+} causes potentiation (Figure 3C, D, E). We then implemented a BCM model in just 3 reactions in HillTau (Figure 3B). We used the program *mash.py* to fit the HillTau model to the reference mass-action version using a set of generic time-series and dose-response stimuli (Methods). We obtained a normalized RMS fit of ~2.3%. When we used the fitted model for Figure 3, we obtained fits of ~5.2%, 8.9% and 2.3% for panels C, D and E respectively even though the model had not been tuned to these stimuli. As a further elaboration, we introduced a bistable switch for long-term retention of synapse state, which was driven bidirectionally by the BCM rule (Figure 3F). The bistable switch, derived from Calcium-calmodulin Type II kinase (CaMKII) signaling, controls receptor insertion. Using this model we delivered a typical potentiating stimulus (strong but

14

brief Ca^{2+} input), leading to sustained synaptic AMPAR elevation. We followed this with a typical long-term depression stimulus (modest but sustained Ca^{2+} input), which turned the switch off again and led to reduction in AMPAR (Figure 3G). This composite model required 4 reactions and one summation function in the HillTau form. Several mass-action models of synaptic state switches include these elements (e.g., (U S Bhalla and Iyengar 1999; Lisman 1989; Hayer and Bhalla 2005; Singh and Bhalla 2018)) and they typically involve far more molecules and reactions (e.g., the Hayer and Bhalla 2005 model used 133 molecules and 215 reactions (Hayer and Bhalla 2005)).

Overall, these examples illustrate how compact HillTau models can represent both the bidirectional induction of plasticity, and also long-term maintenance of synaptic state.

15

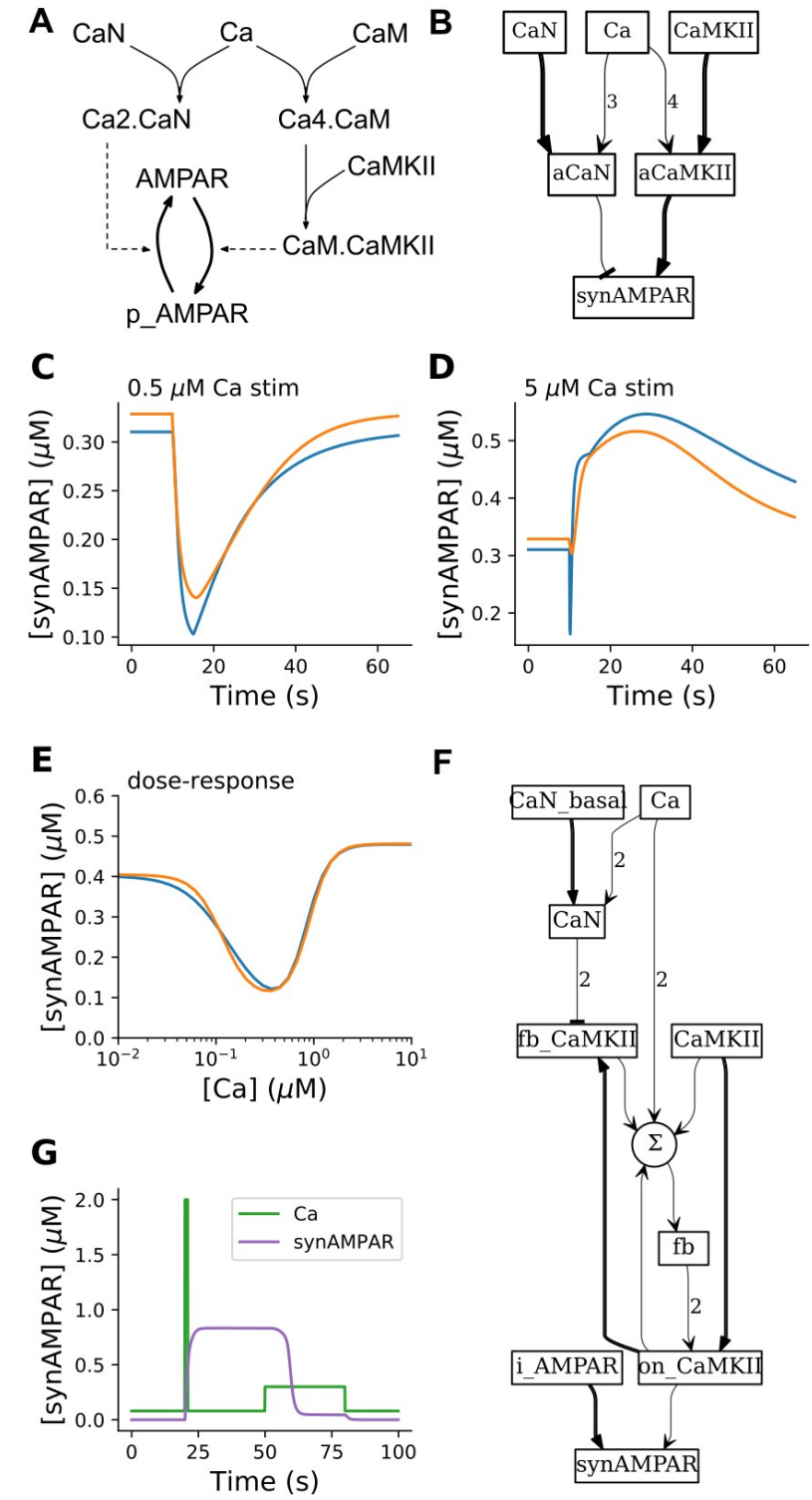

Figure 3: HillTau version of synaptic plasticity rules. A. Mass-action model for generating Beinenstock-Cooper-Munro (BCM) rule for synaptic plasticity. p_AMPAR is the phospho-receptor, and is the output of the model. It is assumed to localize to the synapse and is thus also referred to as synAMPAR. Calcium triggers both an inhibitor (Calcineurin, CaN)

16

and a stimulus (CaMKII) for receptor phosphorylation and insertion into the synapse. CaN activates at lower [Ca^{2+}], so there is initially a reduction in p_AMPAR. CaMKII is present at very high levels, so at higher Ca^{2+} it out-competes CaN to give an increase in p_AMPAR. B. BCM rule implemented in HillTau. Here the species synAMPAR is the output fo the model. C-E: Comparison of mass-action model p_AMPAR with HillTau model synAMPAR. Orange is HillTau, blue is mass action. C: 1s stimulus at 0.5 μM Ca^{2+} gives a reduction in synaptically localized AMPAR (synAMPAR). D: 1s stimulus at 5 μM Ca^{2+} gives an increase in synAMPAR. E: Dose-response curve of steady-state synAMPAR as a function of [Ca^{2+}] for mass-action (blue) and HillTau (orange) models. In both cases settling time for each point was 1000s. F: Schematic for BCM rule model feeding into bistable model, implemented in HillTau. The circular node labeled Σ represents weighted summation of multiple inputs. G: time-course of simulation of bidirectional plasticity using different Ca^{2+} stimuli. At t =20, a 1s stimulus of 2μM Ca^{2+} (green trace) causes a transition to the active state, using synaptic AMPAR (maroon trace) as a readout. At t = 50s, a 30s stimulus of 0.3 μM Ca^{2+} pulls the system back to resting state.

HillTau models can be optimized to fit biochemical measurements

The above examples illustrate how HillTau can represent biological signaling motifs, and build them up into networks with interesting computational properties. We next approached a complementary problem in signaling, namely, to take a complex signaling system, and fit simple HillTau models to it. This provides a way to perform model reduction and to infer computational properties. The basic flowchart is illustrated in Figure 4. This approach works in the same way for model construction from experimental response curves, and for model reduction from response curves taken from detailed models.

17

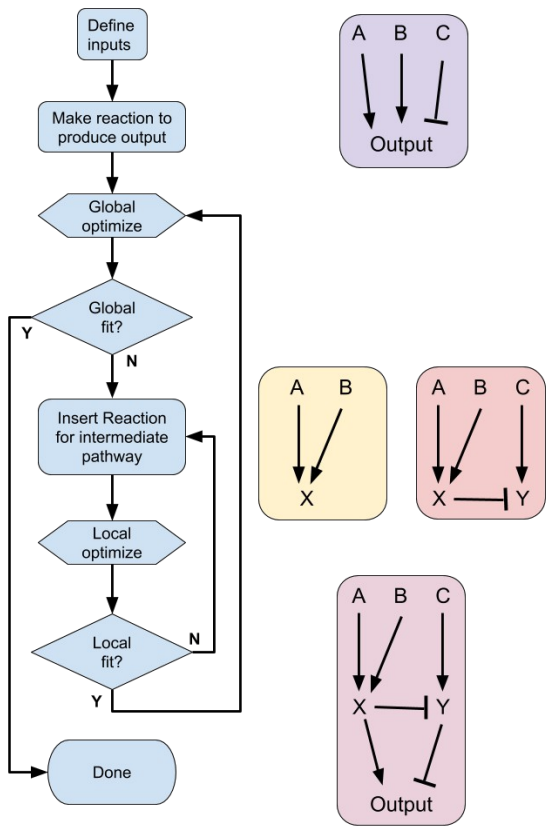

Figure 4: Flowchart for model building using HillTau. Left: flowchart. Right top: initial reaction with inputs A, B, and C. Right middle: successive local increments to model, introducing reactions X and Y respectively. Right lower: Final HillTau scheme with good fit at all stages.

As an example of this flowchart and the use of HillTau fitting to match an existing, detailed chemical ODE model, we derived a HillTau model of synaptic activity-triggered protein synthesis. Our reference data was obtained by running a series of ‘experiments’ on a published model implemented in mass-action kinetics (Jain and Bhalla 2009). The original model was based on numerous experiments, and included 123 molecules and 120 reactions (Fig 5A). The input pathways were Ca^{2+} and brain-derived neurotrophic factor (BDNF), and the final output was protein synthesis rate.

We started with the most reduced form, a single reaction to replace the entire synaptic protein

synthesis network. We specified BDNF as an activator, and Ca as a modifier (Fig 5B) to obtain our starting reduced model. We used MASH to carry out the optimization (Methods). In a model with a single reaction, MASH obtained a fit of about 11% normalized RMS. This is remarkable for such a simple model. It does well with the dose-response experiments (Scores of 7% and 5%, Supplementary Fig 5.1). However, it does not do a good job of replicating the dynamics of the experimental data, achieving scores in the range of 10% to 42%. (Supplementary Fig 5.1). We therefore increased the HillTau model detail. To do this, we introduced two additional key intermediates into the HillTau model: S6K, and CaMKIII. We first made a HillTau model involving inputs to S6K alone. The S6K responses to complex inputs from MASH fit to within 3.3% (Individual panel fits were between 3% and 22%, see Supplementary Fig 5.2). We then held S6K parameters fixed while we fit CaMKIII. MASH gave a CaMKIII fit of 1.9% (Individual panels 2% to 20% but most of the poor scores were small differences at baseline; Supplementary Fig 5.3). Collectively, the optimizations for S6K and CaMKIII correspond to the inner loop of Figure 4. Finally, we added a final protein synthesis reaction that took the already fitted S6K and CaMKIII activity as activator and modifier. We held the earlier reactions (S6K and CaMKIII) fixed, and optimized only the protein synthesis reaction. After this, the composite model fit the optimization waveform in MASH to within 3.2% (normalized RMS), and the figure panels fit within a mean of 9.3% (Fig 5D-K). While we could have optimized to the same stimuli as in the figure panels, we instead chose to give more general input time-series and dose-response stimuli to the MASH optimizer to see how well the model would generalize. Overall, the abstract HillTau model captures many of the key properties of the mass-action system. These include steady-state and time-series responses of two inputs (BDNF and Ca), and three readouts: S6K, CaMKIII, and the end-product protein (Supplementary Fig 5.2, Supplementary Fig 5.3 and Fig 5 respectively). This is a highly effective dimension reduction, from over 360 to 31 parameters.

It is important to note that the efficiency of HillTau made the optimization calculations quite

19

tractable. In the final optimization run for the entire model, the single reference ODE run took ~60 seconds, and the cumulative time for 746 HillTau evaluations was around 18 seconds. The optimization algorithm itself took about 75 seconds, excluding function evaluations.

Can we create a reverse mapping from these simplified HillTau models to ODE forms? A close but not exact mapping is obtained by taking the small-time limit of the HillTau event-driven form (Equations 3.x) and converting to an ODE (rate) form (Equations 4.x). ODE equations are supported by many systems biology simulators. It is not an exact mapping because HillTau may use different values for rising and falling time-courses (tau and tau2), whereas the ODE form can accommodate just a single value, tau. We implemented this conversion in a program, *ht2sbml.py*, which is provided on the GitHub repository for HillTau. Using this mapping we were able to export HillTau to SBML, and tested that SBML-capable simulators such as COPASI could run the reduced, ODE form models, and give approximately matching results (Methods, Supplementary Figure 5.4). Thus we can use the HillTau toolchain to make reasonably reduced ODE models, though these are neither as efficient as their HillTau counterparts, nor do they have the same capabilities to use two time-courses to improve model fitting.

In summary, we developed a systematic procedure for developing reduced HillTau models to fit mass-action simulations, including a model optimization utility MASH. We illustrate this procedure by developing a HillTau model of 10 molecules and 4 reactions to fit a mass-action model having 123 molecules and 120 reactions. The resultant HillTau models generalize well and the fit improves when intermediate reaction steps are added.

20

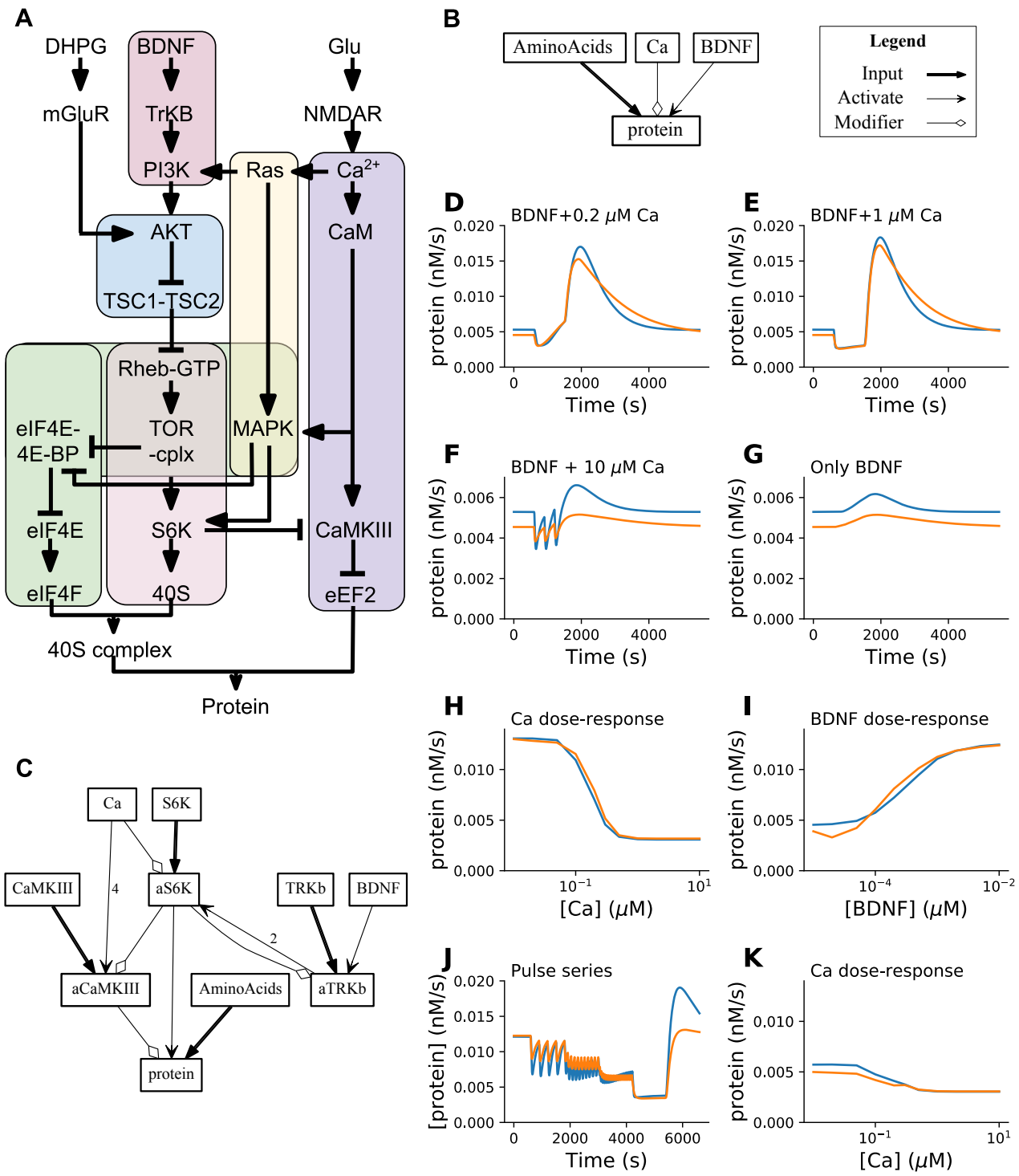

Figure 5: Model fitting and model reduction. A: Block diagram of source model with 123 molecules and 120 reactions, from (Jain and Bhalla, 2009). B: First pass reduced model in HillTau, with 1 reaction and 4 molecules. C: Final reduced

21

HillTau model including activated S6K (aS6K) and activated CaMKIII (aCaMKIII) as intermediate readouts which also were fit to data. This model has 4 reactions and 10 molecules. D-K: Eight 'experiments' on reference and HillTau models, not part of stimulus setused to tune parameters for HillTau version. In all cases protein production rate is readout. Blue plots are reference, orange are HillTau. D: BDNF@3.7 nM + Ca^{2+}@0.2 μM, 900 seconds. E: BDNF@3.7nM, Ca^{2+}@1μM. F: 3 pulses of BDNF@3.7 nM for 5s, coincident with Ca^{2+}@10μM for 1s, pulses separated by 300 s. G: Same as F, but Ca^{2+} held at baseline of 0.08 μM. H: Dose-response of protein vs. Ca^{2+}, holding BDNF fixed at 3.7 nM. I: Dose-response of protein vs BDNF, holding Ca^{2+} fixed at 0.08 μM. J: Protein production rate for fixed BDNF@3.7 nM, where Ca^{2+} was given in 1 second pulses at intervals of 300, 120, 60 and 10 seconds; each pulse sequence lasting for 1200 s. . K: Dose response of protein production rate vs. Ca^{2+}, holding BDNF at basal levels of 0.05 nM. The average normalized RMS difference across the eight 'experiments' was under 10%, and in all cases the qualitative properties such as direction of change, matched well.

HillTau models are compact and efficient

We next took a set of HillTau models of various levels of complexity, and compared various measures of computational cost with the ODE equivalents (Table 1, Figure 6.)

We first compared model complexity, measured as the number of parameters needed to specify the model. The number of parameters scales roughly as

*# of molecular species + 2 * # of reactions*.

This is a slight overestimate, since some of the molecules are state variables and we do not need initial concentration values for them. Each reaction needs two parameters, *Kf* and *Kb* for conversion reactions, and *Km* and *kcat* for enzymes. We sampled from among the mass-action models presented in the above sections, ranging from 3 to over 360 parameters, and included an additional model with almost 750 parameters. (Fig 6A). The HillTau form had a similar scaling with molecules and reactions, except that HillTau also allows for a number of optional terms in the

22

reactions so the average scaling is somewhat larger than *2 * # of reactions*. We found that the HillTau form became increasingly effective at model reduction for larger models. Note that here the optimization goal was to obtain a single end-point response (3 end points in the case of the model in Figure 5). Further reactions would be needed to also represent intermediate pathway readouts.

We then examined run-time efficiency. We took run-times for two ODE simulators, MOOSE and COPASI, whose numerical cores are in C++. We compared these with the standard Pybind11/C++/Python implementation of matching HillTau models (Methods, Figure 6B). For small models, HILLTAU was about 100 times faster than the ODE calculations , but for large models HillTau became over 3 to 4 orders of magnitude faster (Figure 6B). This set of HillTau models fit their ODE counterparts within 3 to 9% accuracy, but note that the approximations were inherent in the HillTau model structure, and did not arise from lack of numerical convergence. Interestingly, the run-time for HillTau models grew almost linearly with the number of parameters (Fig 6C, slope = 0.013 us/s, R^2 = 0.98). This suggests that the HillTau calculation cost depends almost entirely on model complexity as opposed to attributes such as numerical stiffness, in contrast to variable-timestep methods for ODE solutions frequently used for chemical kinetic calculations.

23

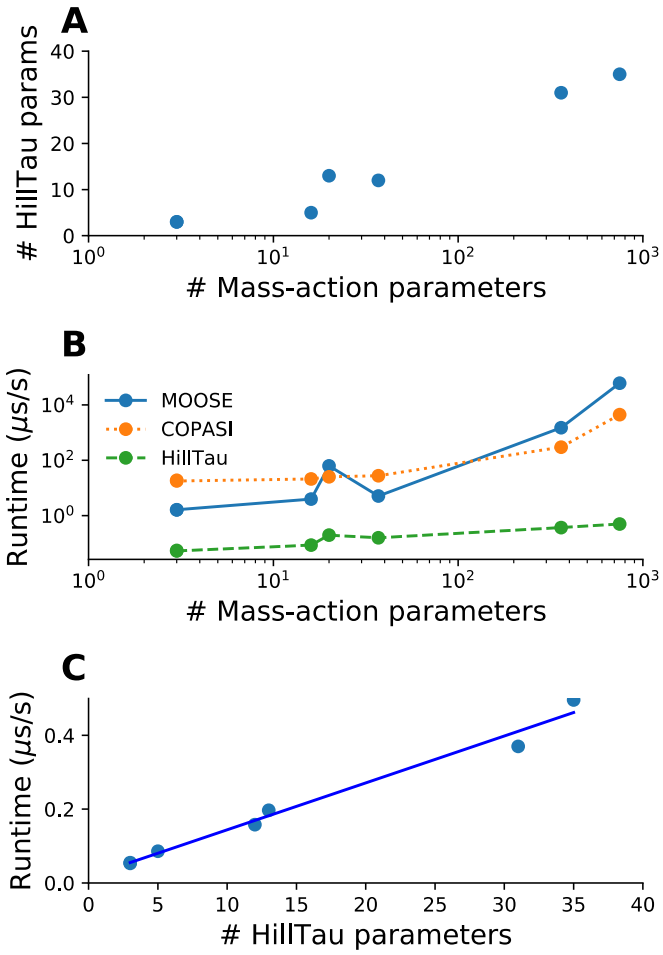

Figure 6. Efficiency of HillTau method. A. Scaling of number of parameters. The HillTau form uses far fewer parameters than mass-action, and becomes relatively more concise with larger models. B, C: run-time scaling. All run-times are expressed in μs of wall-clock time to execute 1 second of simulation time. Calculations were done on an Intel(R) i7-7700HQ processor, running Ubuntu 18.04. B. Comparison of run-times of the HillTau and mass-action forms, where the numerically intensive sections of HillTau were implemented in C++, and C++ was also used for ODE calculations in two simulators: MOOSE and COPASI. Due to the combination of model reduction and efficient calculation, HillTau has a huge efficiency advantage which grows to over 3 orders of magnitude for larger models. The accuracy with which this set of HillTau models fit their ODE counterparts was in the range 3 to 9%. C. HillTau model run-time scales nearly linearly with number of parameters. R^2 = 0.98, slope = 0.013, intercept = 0.017, units in μs/s.

24

Dose-response experiments are particularly efficient to compute using HillTau. An inefficient way to do these for ODE models is to run them out to steady-state for each successive dose. This may take a long while especially if the system is stiff or converges slowly. It is also possible to use linear algebraic root-finding to find the steady-state value in one step, but this is prone to numerical challenges even in modern simulators like COPASI (Clarke, Bruce L. 1981). In HillTau, the form itself incorporates the steady-state value, so in principle one could leap to the steady value in one step. To be more conservative, the HillTau does so in 10 steps to smooth out transients and to allow any feedback signals to propagate through the system. As an illustration, COPASI performed the steady-state calculation for a large model (accession 92, DOQCS, ~100 reactions, (Jain and Bhalla 2009)) in ~1 second after some relaxation of convergence criteria. The HillTau equivalent model (same as final model in Fig 5) took about 4 microseconds.

Overall, HillTau models are compact and highly efficient compared to ODE-solved mass-action models. The efficiency improves for larger models.

Discussion

We have designed HillTau, a compact, computationally efficient abstraction of chemical signaling that is particularly effective in building reduced models of complex signaling networks. It uses an event-driven algebraic representation based on the Hill equation and exponential relaxation to steady state. HillTau is effective in representing a range of chemical signaling motifs and complex synaptic models, using biological observables of molecules, reactions, association constants and time-courses. We show its applicability for model reduction by optimizing the fit of its responses to those of a reference mass-action model. This generates very compact models. A similar optimization approach works to build a HillTau model directly from experimental data. Thus HillTau addresses many of the concerns of model-building with limited data, and serves as a

25

scaffold for eventual development of more detailed models.

HillTau is phenomenological and semi-heuristic, in that it uses the Hill equation to achieve concentration dependencies that fit well, but ignores many intervening chemical steps. This combination gives it the strong points indicated above, namely speed, compactness, and consistent mapping to experimental observables, but it also sets out clear limitations. Foremost among these is that it can only make limited predictions on detailed pathway chemistry, since it is missing many reaction steps. For example, a HillTau model would be limited in its ability to predict drug targets or side-effects because it may have lumped together potential molecular targets into a single reaction step. It is, however, quite effective in representing and predicting emergent signaling properties because it captures dynamics and topology of signaling networks.

The current HillTau formulation is limited in its handling of two important aspects of signaling in neurons: stochasticity and diffusion. These phenomena are out of the scope of our current implementation, which has focused on development, validation, simplicity and speed. Many biochemical signaling processes experience substantial stochasticity, particularly in small-volume systems such as the synapse which is a target of our modeling. One possible way to introduce stochasticity would be through the linear noise approximation of the chemical Langevin equation (Wallace et al. 2012), which if used in an event-driven manner could be quite efficient. We anticipate it will take extensive validation to establish its utility in the HillTau framework. Similarly, there are potential ways to elaborate upon HillTau to use an event-driven approximation to diffusion, but these will require later follow-up.

Based on these attributes, we discuss four major use-cases for HillTau: model reduction, system abstraction, scaffolds for data-driven optimization, and efficient approximations to complex cellular signaling.

26

Model reduction

Several algorithmic approaches have been brought to bear on the *model reduction* problem, including collapsing multiple mass-action steps into one (Radulescu et al. 2008), and power-law generalizations of mass-action signaling (Savageau 2001). Simulators such as COPASI provide utilities for partitioning a reaction system into fast and slow reactions, thus allowing one to approximate the fast steps with their algebraic counterparts (Hoops et al. 2006; Deuflhard and Heroth 1996). In our test model (Accession 92, DOQCS, (Jain and Bhalla 2009)) we were able to partition about 80% of the steps into the fast domain (<100 s) using the ILDM method implemented in COPASI. With HillTau one can use a well-known heuristic/optimization approach to simplifying large networks (Quaiser et al. 2011; Apri et al. 2014). This reduced the same test model down to 4 reactions, about 25-fold (Figures 4, 5). The approach reported in our current study relies on the modeler starting from a minimal input-> output mapping, then iteratively picking relevant major nodes in the network, and optimizing each subset of the model to fit the data (Figure 4).

Thus one can converge on the minimal set of intermediate nodes (illustrated in Figures 4 and 5, and Supplementary Figures S5.1, S52. and S5.3) to achieve the desired accuracy of model fit to data. Like other model reduction approaches, this minimal set of nodes is a compromise between available data and model accuracy (Snowden, van der Graaf, and Tindall 2017). Unlike several other reduction approaches, HillTau retains a direct mapping to observable biological entities. Indeed, the HillTau representation of a signaling node may be closer to the conventional intuition based on pathway schematics, than is a full mass-action reaction representation. Like pathway diagrams, each HillTau reaction receives excitatory, inhibitory and modulatory inputs. A further point of similarity is the HillTau models may condense several intermediate steps into a single node on the reaction network. A more subtle point of similarity is that pathway block diagrams

typically assume implicit back reactions and decay of activity when stimuli are removed. This too is built into how HillTau reactions work. In contrast to these simple mappings from pathway diagrams to the HillTau form, it is often difficult to map between signaling diagrams and the full mass-action reaction schemes (Upinder S. Bhalla 2002; Upinder S Bhalla 2003). While previous model reduction studies have worked on different pathways than the synaptically biased set explored in our study, a comparison with the survey of methods in (Snowden, van der Graaf, and Tindall 2017), suggests that HillTau achieves as good or better model reduction for large models than most other methods.

System abstraction

Next, *system abstraction* and functional modules help to make sense of complex biological signaling. We propose that HillTau forms a useful tool for arriving at functional modules in complex signaling networks. Such modules have long been considered a conceptual basis for understanding complex signaling (U S Bhalla and Iyengar 2001). Typically they have been ascertained by manual inspection and dynamical analysis of components of signaling networks, for example, the nested feedback loops in the cell cycle (Novák and Tyson 2004). A more scalable approach to uncovering such modules is to use graph theory for motif analysis on detailed mass-action models, but this approach loses key aspects of system dynamics (Alon 2007). With the HillTau formalism and our procedures for model reduction, we are able to generate highly reduced reaction graphs that nevertheless support rather accurate dynamics. The formalism encourages models that can be readily mapped to biology. Thus HillTau supports a data-driven approach to arrive at functional modules.

While functional modules are good for analysis, we note that biology does not necessarily partition signaling networks into neat modules (Upinder S Bhalla 2003). Indeed, cross-talk between pathways is common. HillTau supports explicit cross-talk interactions, but does not introduce

implicit interactions. In this regard it differs from mass-action reaction systems, in which mechanisms such as back-reactions or enzyme sequestration may introduce subtle effects, such as implicit feedback. For example, one can achieve bistability through multistage phosphorylation/dephosphorylation cascades (Markevich, Hoek, and Kholodenko 2004). To represent such effects in HillTau one would have to introduce explicit feedback steps between reactions, such as in the bistability example in Figure 2. Similarly, interesting behavior emerging from chemical saturation, such as zero-order ultrasensitivity (Goldbeter and Koshland 1981) would require the use of the explicit math expressions supported by HillTau.

Thus HillTau promotes abstraction through model reduction. The abstracted models expose all interactions explicitly.

Scaffolds for model fitting

Third, HillTau is a useful intermediate step, or *scaffold*, for model fitting of large mass-action models. Direct model-fitting is difficult in at least two ways: there are typically far fewer experiments than parameters, and it is computationally costly to run a large ODE model many times for carrying out an optimization approach to model fitting. We propose that the HillTau form may provide a useful bridge on both these counts. As we have illustrated in Figures 2, 3 and 5, HillTau models lend themselves to fitting to experiments because they have few parameters and they run quickly. Several advantages accrue from an initial pass to make and fit a HillTau model. 1. In building and optimizing a HillTau model, the optimization dataset will be use-tested, and gaps identified. 2. The essential pathway structure of the model will be defined by the HillTau model, and key interactions identified. The mass-action model must, at minimum, incorporate these interactions. 3. The parameters of the HillTau model set bounds for those of the detailed reaction sets. For example, the time-course of any individual mass-action reaction step must be faster than the HillTau reaction in which it is embedded.

29

Efficient approximations to complex signaling

Finally, HillTau models are useful because they are *efficient*. One of the key use-cases envisioned for HillTau was to model complex cellular signaling, with synaptic signaling as an exemplar. Several of the examples in the current paper are in this domain, specifically the BCM curve (Fig 3A-E, the coupled BCM curve with bistables (Fig 3F,G), and a synaptic protein synthesis pathway (Figure 5). While even these large ODE signaling models run somewhat faster than wall-clock time (Figure 6), there are at least two cases where far greater efficiency is desirable. First, as mentioned above, model parameter optimization requires a large number of evaluations of complete simulations (100s to 1000s in our experience). To perform a single evaluation, these synaptic simulations may have to run for many thousands of seconds of simulation time to compare with typical plasticity experiments (Ajay and Bhalla 2004). Further, one typically optimizes a pathway to fit numerous experiments, all of which must be simulated for each evaluation. Together, this is computationally expensive. In our HillTau optimizations using MASH, the total simulation time for large numbers of pathway simulations was typically even smaller than the time spent by the minimizer code itself.

A second use case for highly efficient signaling models is in synaptic signaling. A single neuron may have over 10,000 spines, and there may be many such neurons in a network. If each synapse is to implement a complex biochemical pathway the computational costs may be formidable. Network plasticity models (Higgins, Graupner, and Brunel 2014), and cellular sequence selectivity models (Upinder S Bhalla 2017) are examples of this scale of model. Indeed, Higgins et al. (Higgins, Graupner, and Brunel 2014) have used an efficient event-driven calculation of synaptic weights with a similar exponential decay calculation as in HillTau. HillTau signaling provides a way to implement biologically detailed synaptic dynamics in every synapse, even in large networks.

In summary, the HillTau form and its supporting toolkits for running and optimizing models

30

provide a compact, efficient way to perform data-driven abstraction of complex signaling models.

31

Methods

HillTau formulation

The HillTau formulation is an event-driven variant of a Hill equation with modifiers. It is specified and evaluated in two stages: the steady-state value, and an exponential time-course of approach to steady-state. As detailed below, reactions in a HillTau model are evaluated in successive layers such that the next estimate for steady-state value of a given layer depends only on boundary conditions and on outputs from preceding layers. In the equations below we expand out the equations for a range of use-cases. In equations 1.x we specify the steady-state values for each use case. In equations 2.x we define the approach to steady state. In equations 3.x we combine equations 1 and 2 to summarize the evaluations done in HillTau. In equations 4.x we provide interpretations of the HillTau equations as rate terms which can be evaluated by regular ODE solvers and form the basis for the SBML export of HillTau model systems. However, the definitive form of HillTau is event-driven and the rate-term form should be regarded as a convenient but approximate mapping to conventional mass-action solvers. In equations 5.x we provide a motivation for the form of HillTau as an approximation to complex signaling chemistry.

HillTau steady-state.

The HillTau formulation stipulates that the steady-state level Y_∞ of each signaling step (which may involve multiple chemical steps) is approximated by a Hill function

$$Y_\infty/Y_{input} = \Theta = L^n/(KA^n + L^n) \qquad \text{Eq 1.1}$$

Here Y_{input} is the input concentration to this signaling step, where Y_{input} is either a molecule with a predefined concentration (i.e, a boundary condition) or coming from an upstream reaction.

32

Likewise, the reactant L can either be predefined or come from an upstream reaction. *KA* is the association constant of L with Y.

We elaborate this slightly to accommodate an optional gain term:

$$Y_{\infty} = Y_{gain}.Y_{input}L^n/(KA^n + L^n) \qquad \text{Eq 1.2}$$

We utilize a slightly modified form to permit the Ligand *L* to act in an inhibitory manner:

$$Y_{\infty} = Y_{gain}.Y_{input}(1-L^n/(KA^n + L^n)) \qquad \text{Eq 1.3}$$

In cases where there are modifiers, we include a further term based on the analysis of Hofmeyer and Cornish-Bowden (1997):

$$Y_{\infty} = Y_{gain}.Y_{input}.L^n/(KA^n(1+M/K_{mod})^h)/(1+A_{mod}(M/K_{mod})^h)+ L^n) \qquad \text{Eq 1.4}$$

Here M is the concentration of the modifier, K_{mod} is the half-effect value, A_{mod} is the modifier action, and h is the order of the modifier. The modifier acts in an inhibitory manner when $A_{mod} < 1$, and as an activator when $A_{mod} > 1$. When $A_{mod} == 1$, clearly, the modifier has no effect.

Similarly we define the action of the modifier on an inhibitory reaction:

$$Y_{\infty}= Y_{gain}Y_{input}(1 - L^n/(KA^n(1+M/K_{mod})^h)/(1+A_{mod}(M/K_{mod})^h)+ L^n)) \qquad \text{Eq 1.5}$$

We use a different equation to define steady-state behavior of a system where a single substrate molecule Y_{input} is converted into a product *Y*:

$$Y_{\infty} = Y_{input}^n / KA \qquad \text{Eq 1.6}$$

For the rare cases where a non-chemical formulation is needed to describe the system, we provide

33

an alternative algebraic expression for Y_∞:

$$Y_\infty = f(Y1, Y2\ldots.) \qquad \text{Eq 1.7}$$

where f is an arbitrary algebraic function and Y1, Y2… are concentrations of other molecules. The use of this algebraic form is discouraged as it weakens the mapping of the model to the underlying chemistry. This form does not admit of modifiers.

Note that these equations are entirely feed-forward: the concentrations of molecules L, M, and Y_{input} are not affected by their participating in any downstream reactions.

HillTau time course

Equations 1 define the steady state estimate for molecule Y, given a set of molecule concentrations in the preceding layer. We then assume that the approach of the system to steady-state is governed by a simple exponential with characteristic time τ (Figure 1A):

$$(Y(t+\Delta t) - Y(t))/(Y_\infty - Y(t)) = 1 - \exp(-\Delta t / \tau) \qquad \text{Eq 2}$$

where Y(t) is the value of Y at time t. For a simple binding reaction, the time course τ is an experimental observable, and is approximated by $\tau \sim 1/(kf+kb)$ where *kf* and *kb* are the forward and backward rates for the first-order Hill binding reaction (Figure 1, Supplementary Figure S1.1). As a slight extension to this, we permit an optional separate time course τ_2 when Y is falling:

If $Y_\infty > Y(t)$ $\quad (Y(t+\Delta t) - Y(t))/(Y_\infty - Y(t)) = 1 - \exp(-\Delta t / \tau)$ $\quad$ Eq 2.1

If $Y_\infty < Y(t)$ $\quad (Y(t+\Delta t) - Y(t))/(Y_\infty - Y(t)) = 1 - \exp(-\Delta t / \tau_2)$ $\quad$ Eq 2.2

The occurrence of different time-courses for buildup and decay is quite common. It happens in a

34

simple binding reaction (Supplementary Figure S1.1, panels A, C, E). It also occurs when there are different chemical steps, such as enzymes with different rates, mediating competing processes for buildup and decay.

HillTau composite form

Putting Eq 1 and Eq 2 together, we have the following closed-form expression for the value of Y at time t + Δt:

$$Y(t+\Delta t) = Y(t) + (Y_\infty - (Y(t)-Y_{baseline})) (1-\exp(-\Delta t/ \tau)) \qquad \text{Eq 3}$$

The term $Y_{baseline}$ is an optional (positive) baseline level of molecule Y. Δt is the timestep. Note that this is a closed form: Δt can be as large as the end of the simulation.

In the limit of large Δt, we have,

$$Y_{(t = \infty)} = Y_\infty + Y_{baseline} \qquad \text{Eq 3.1}$$

The initial conditions for molecule Y are either specified in the model definition, or as a fallback we estimate the steady-state concentration as in Eq 3.1.

Rate interpretation of HillTau

Formally, HillTau can be seen as an approximation to the following rate equations:

$$Y'(t) = (Y_\infty - Y(t))/\tau \qquad \text{Eq 4}$$

In cases where we have a separate τ_2 , and $Y_\infty < Y(t)$:

$$Y'(t) = (Y_\infty - Y(t))/\tau_2 \qquad \text{Eq 4b}$$

SBML, and most ODE solvers, will not readily handle this switch between τ and τ_2. For the purposes of the equations below we just use τ.

Expanding out Y_∞, we have five variants of equation 4:

Basic activation reaction:

$$Y'(t) = (Y_{gain}.Y_{input}L^n/(KA^n + L^n) - Y(t))/\tau \qquad \text{Eq 4.1}$$

Basic inhibition reaction:

$$Y'(t) = (Y_{gain}.Y_{input}(1-L^n/(KA^n + L^n)) - Y(t))/\tau \qquad \text{Eq 4.2}$$

Activation reaction with modifier:

$$Y'(t) = (Y_{gain}.Y_{input}.L^n/(KA^n(1+(M/K_{mod})^h)/(1+A_{mod}(M/K_{mod})^h)+ L^n) - Y(t))/\tau \qquad \text{Eq 4.3}$$

Inhibition reaction with modifier:

$$Y'(t) = (Y_{gain}.Y_{input}.(1-L^n/(KA^n(1+(M/K_{mod})^h)/(1+A_{mod}(M/K_{mod})^h)+ L^n)) - Y(t))/\tau \qquad \text{Eq 4.4}$$

Conversion reaction:

$$Y'(t) = ((Y_{input}^n/KA) - Y(t))/\tau \qquad \text{Eq 4.5}$$

Although the HillTau calculations can be done using equations 4.x with a regular ODE solver, the HillTau definition envisions event-driven calculations. Further, the complete HillTau form uses τ_2 extensively, which is not handled by the above equations. In principle, the optimization for the HillTau model could just use τ throughout, in which case the reduced HillTau model could be rendered in mass-action form with a reasonable degree of accuracy (Supplementary Figure 5.4A).

Here we summarize the parameters for HillTau. All but the first apply to Reactions.

Parameter	Units	Meaning	Default	Required?
ConcInit	Concentration (can be any of M, mM, uM, nM)	Initial concentration. Only applies to species definitions	0	No
KA	Concentration	Association	N/A	Yes

36

		constant		
τ	Time (seconds)	Time course for relaxation to steady state	N/A	Yes
τ2	Time (seconds)	Time course of relaxation if output is falling.	τ	No
Gain	None	Scaling factor for reaction output. Used to indicate enzymatic amplification.	1	No
Baseline	Concentration	Baseline value of reaction output	0	No
Kmod	Concentration	Half-saturation concentration for modifier	N/A	Only if modifier molecule is specified.
Amod	None	Activation term for modifier	4	No
Nmod	None	Order of modifier action	1	No

Table 1: Parameters used to define a HillTau model. Concentration units can be any of M, mM, uM, and nM, and are specified in the Json file. The default concentration units are mM.

Simple HillTau models only need species concentrations, and the KA and τ terms for the reactions. The optional terms greatly facilitate the design goals of compactly specifying diverse signaling reactions.

Computing time-evolution and steady-states.

To build complex reaction systems, we permit cascading of reactions so that any molecule can be a substrate or equation term in any other reaction. To reiterate, this is a purely feed-forward formulation, so substrates are not affected by any of their downstream targets. We obtain a

37

dependency tree so that on each timestep the updates are carried out in an order which ensures that inputs ripple in order through the cascade of reactions. However, multi-step systems may include feedback. In such cases the program has to explicitly break the dependency chain, and then it becomes necessary for the modeler to specify a timestep Δt which is smaller than the time-course (τ) of the fastest reaction in the loop. Since one normally performs time-series sampling of reactions at a time finer than the fastest reaction, this restriction usually has little impact on run-time. However, there are cases where the HillTau system is used to compute steady state values (e.g., in a dose-response curve). These could ideally be solved by taking an infinitely long time-step. Given the possible presence of feedback, we instead take a long, user-defined settling time and subdivide it into 10 equal steps so as to allow feedback reactions to also settle. A similar jump to user-defined time-points applies for event-driven calculations.

Motivation for the HillTau formalism

The first motivation for the HillTau form was empirical and observational. A large number of pathway outputs in a database of biochemical experiments have a saturating, Hill-like dependence on input strength (Viswan and Bhalla, unpublished data). Likewise, many multi-step cascades in this and other datasets have time-courses that settle with an exponential-like time course to their final value.

To mathematically support this observation, we considered outcomes of an enzymatic cascade with back reactions. We approximate the rate of production using a Michaelis-Menten- form:

$$dP/dt = ES.kcat/(Km + S) \qquad \text{Eq 5.1}$$

This is balanced by a first-order back reaction:

$$dP/dt = -kr.P \qquad \text{Eq 5.2}$$

Then the steady-state at dP/dt = 0 is obtained by combining these:

$$P_1 = E.S_1.kcat_1/((Km_1 + S_1) * kr_1) \qquad \text{Eq 5.3}$$

In the above case we add a subscript *1* to indicate that it is the first reaction in the cascade. Now we

38

stipulate that P_1 is the catalyst for the next step, resulting in product P_2:

$P_2 = (E.S_1.kcat_1/((Km_1 + S_1) * kr_1)).S_2.kcat_2/((Km_2 + S_2)*kr_2))$ Eq 5.4

And so on for multiple steps. Now, suppose that we only have 2 variable inputs to this pathway: E and one of the substrates. All other substrates are held fixed. This is a reasonable assumption for HillTau, because any further variable inputs should be treated explicitly either as modulators or as separate reaction steps. Then, all the $(Km_n + S_n)$ terms are constant except the one variable substrate S_v. By combining all the constant terms into $K_{cascade}$, we end up with:

$P_n = E.S_v.K_{cascade}/(Km_v + S_v)$ Eq 5.5

This is equivalent to the Hill equation form at the basis of HillTau (See Eq 1.2). We treat inhibition using the same analysis, except resulting in depletion of a substrate (Eq 1.3) and following Hofmeyer and Cornish-Bowden (ref), we can also add modifiers to this as done in Eq. 1.4 and 1.5. For the time-course, we assume that one of the reactions is rate-limiting:

$dP/dt = ES.kcat/(Km + S) - kr.P$ Eq 5.6

This yields an exponential settling curve with a final value Pinf given by Eq 5.3:

$P(t) = Pinf – (Pinf - P0) \exp(-t/tau)$ Eq 5.7

where

tau = 1/kr and P0 is the initial value of P.

Note that Eq 5.7 can be rearranged to give Eq 2.

Overall, this approximation yields a consolidated HillTau ‘reaction’ in which we have one stimulus (E), one reactant (S_v), and optional modifiers to represent a cascade of enzymatic steps with back reactions.

Next, we consider binding reactions in the pathway. These give the same steady-state Hill Equation form, by definition. From Equation 1.1, setting n = 1, and considering the first reaction generating Y1:

$Y1 = L1.R1/(L1+KA1) = R1/(1+KA1/L1)$ Eq 5.8

A cascade of such reactions, where Y1 feeds into reaction 2 and so on, can also be reduced into the same form:

$$Y2 = R2/(1+KA2/Y1) = R2/(1+KA2*(1+KA1/L1)/R1)$$

$$=R1R2/(R1+KA2 +KA1KA2/L1) = Rx/(1+KAx/L1) \qquad \text{Eq 5.9}$$

where $Rx = R1R2/(R1+KA2)$ and $KAx = KA1.KA2/(R1+KA2)$

Note that enzyme/back-reaction steps may also be folded into such a cascade as they have a similar form, from Eq. 5.5.

A standard analysis shows that this too has an exponential time-course if we assume one rate limiting step.

$$dY/dt = Kf.L.R – Kb.Y \qquad \text{Eq 5.10}$$

which yields the same exponential settling time-course as equation 5.7:

$$Y(t) = Yinf – (Yinf - Y0) \exp(-t/tau) \qquad \text{Eq 5.11}$$

where $tau = 1/(L.Kf + Kb)$ and Y0 is the initial value of Y.

It is important to note that this value of tau has a dependence on a variable, L, and hence the use of a constant value of tau is approximate. This is partially mitigated by utilizing different values of tau for rising and falling phases of the signal Y. During the rising phase, L will have a different (typically larger) mean value than during the falling phase. The use of tau for rising and tau2 for falling phases of the response reflects this.

Thus the steady-state terms for cascading binding and enzymatic reactions can be consolidated into single HillTau ‘reaction’ steps of the Hill form, and the time-course can be approximated by an exponential when there is one rate-limiting step.

Model definition and reference implementation

HillTau reaction systems are set up in a simple JSON format (Figure 1C), for which we have provided a schema. We have implemented a small reference driver program in Python (hillTau.py)

40

that loads the model, runs it with optional stimuli, and plots or saves the output. The hillTau.py file also provides a set of library functions for use in larger programs. An equivalent implementation in C++ using PyBind11 for identical Python bindings is also provided. The source files, schema file, examples and documentation are all available on GitHub (https://github.com/BhallaLab/HillTau). Three additional utilities are also provided, as described below: for model illustration, model abstraction, and model conversion to SBML. Python scripts for generating the figures in this paper (except figure 4, which is a schematic only) are also provided in supplementary material and on the GitHub site. Benchmarking was done using the program *fig6.py*, which calls MOOSE and HillTau through their Python interfaces, and calls COPASI through the PyCoTools Python interface (Welsh et al. 2018). The output values of multiple benchmarking runs were averaged and used for generating the graphs in *fig6_plotting.py*. All HillTau and supporting code is licensed under GPL version 3 or later.

Model illustration

We developed a utility *htgraph.py*, which generates a reaction diagram for HillTau models specified in the .json format. This diagram gives a complete specification of the model topology, in that one can rebuild the structure of the HillTau model by inspection of the reaction diagram, though of course the parameters are not provided in the diagram. htgraph.py uses the *dot* module of the open-source package *graphviz* (Gansner and North 2000) to generate the network graphs.

Model reduction and abstraction: MASH

We provide a utility for performing Model Abstraction from SBML to HillTau: MASH. Briefly, MASH runs the original SBML model with a reference stimulus to explore the key dimensions of its input-output mappings. Typical reference stimuli (built into MASH) include dose-response curves and pulsatile time-series stimuli. MASH then uses standard minimization routines (scipy.optimize library, method “L-BFGS-B”, (Zhu et al. 1997)) to tweak the HillTau model parameters to improve its fit to the original model.

41

MASH is implemented as a Python script *mash.py* which uses an ODE model (SBML) as a reference to which it fits a HillTaul model. The user provides an initial HillTau model and specifies a series of stimuli to deliver. As part of this the user also defines which are the input molecules, and which are the readouts, and the ODE and HillTau models. MASH generates a topologically identical HillTau model to the original, with parameters optimized to fit. It also reports initial and final scores, expressed as normalized RMS difference between reference model and HillTau. As per Figure 4, the user may introduce additional intermediate steps in the pathway in order to achieve the target model fit. The user specifies a set of stimuli (typically a combination of dose-response and time-series calculations) that explore the model response space. MASH generates a reference response to these stimuli using an ODE solver (MOOSE). The function evaluation for the minimization is carried out by running the HillTau model with the same stimulus, and comparing the HillTau output point-by-point with the reference. The normalized RMS difference over all points is returned as the score. A score of below 0.05 means that on average the original response differs from the HillTau response by less than 5%. MASH uses the scipy.optimize library to tweak the HillTau model parameters to improve the fit, as measured by this RMS score. MASH documentation and examples are provided on the HillTau website. MASH was used to fit the HillTau models for figures 2, 3 and 5, and the scores are reported. Supplementary figures and data specify how these fits were done.

Model conversion to SBML

The utility *ht2sbml.py* performs a conversion of HillTau models defined in the reference JSON format, into equivalent ODE models defined in SBML. It uses *simplesbml* (https://github.com/sys-bio/simplesbml for generating the SBML. This uses the forms defined in equations 4.x in Methods. The conversion is approximate on two counts, first, HillTau is an event-driven, not continuous method, and second, HillTau may use different time-courses for rising and falling phases as a

42

reaction proceeds, whereas the ODE form uses only a single time-course. If the HillTau model is generated (for example, after model reduction) such that each reaction only utilizes tau and not tau2, then the approximation is very good. In Supplementary Figure S5.4, we performed *ht2sbml.py* conversion of 3 HillTau models to SBML and then compared the HillTau output with COPASI calculation of the converted model, under 5 conditions. We obtained an excellent fit (<1% normalized RMS) for an oscillatory model that did not use tau2. The feedback-inhibition model used in Figure 2, which does use tau2, gave a mediocre fit of 29%. The full protein synthesis model was tested under 3 conditions: protein response to BDNF, S6K response to BDNF, and CaMKIII response to Ca. These gave fits of 30%, 26% and 1.7% respectively, though the qualitative response was similar in all cases. Thus the *ht2sbml.py* conversion works for all HillTau models, but the conversion may be approximate for models which have very different tau and tau2 parameters in their reactions.

43

Acknowledgments

USB received support from NCBS-TIFR and the Department of Atomic Energy, Government of India, under project no. 12-R&D-TFR-5.04-0800. USB is a JC Bose Fellow. Nisha Viswan provided useful feedback from use-testing HillTau. G.V. HarshaRani did much of the coding for htgraph.py.

44

References

Ajay, Sriram M, and Upinder S Bhalla. 2004. “A Role for ERKII in Synaptic Pattern Selectivity on the Time-Scale of Minutes.” *The European Journal of Neuroscience* 20 (10): 2671–80. https://doi.org/10.1111/j.1460-9568.2004.03725.x.

Alon, Uri. 2007. “Network Motifs: Theory and Experimental Approaches.” *Nature Reviews. Genetics* 8 (6): 450–61. https://doi.org/10.1038/nrg2102.

Andrews, Steven S, Nathan J Addy, Roger Brent, and Adam P Arkin. 2010. “Detailed Simulations of Cell Biology with Smoldyn 2.1.” *PLoS Computational Biology* 6 (3): e1000705. https://doi.org/10.1371/journal.pcbi.1000705.

Apri, Mochamad, Maarten de Gee, Simon van Mourik, and Jaap Molenaar. 2014. “Identifying Optimal Models to Represent Biochemical Systems.” *PLOS ONE* 9 (1): e83664. https://doi.org/10.1371/journal.pone.0083664.

Barak, Omri, and Misha Tsodyks. 2006. “Recognition by Variance: Learning Rules for Spatiotemporal Patterns.” *Neural Computation* 18 (10): 2343–58. https://doi.org/10.1162/neco.2006.18.10.2343.

Bayés, Alex, Louie N. van de Lagemaat, Mark O. Collins, Mike D. R. Croning, Ian R. Whittle, Jyoti S. Choudhary, and Seth G. N. Grant. 2011. “Characterization of the Proteome, Diseases and Evolution of the Human Postsynaptic Density.” *Nature Neuroscience* 14 (1): 19–21. https://doi.org/10.1038/nn.2719.

Bhalla, U S, and R Iyengar. 1999. “Emergent Properties of Networks of Biological Signaling Pathways.” *Science (New York, N.Y.)* 283 (5400): 381–87.

———. 2001. “Functional Modules in Biological Signalling Networks.” *Novartis Foundation Symposium* 239: 4–13; discussion 13-15, 45–51.

Bhalla, Upinder S. 2002. “The Chemical Organization of Signaling Interactions.” *Bioinformatics* 18 (6): 855–63.

Bhalla, Upinder S. 2003. “Understanding Complex Signaling Networks through Models and Metaphors.” *Progress in Biophysics and Molecular Biology* 81 (1): 45–65.

Bhalla, Upinder S. 2014. “Multiscale Modeling and Synaptic Plasticity.” *Progress in Molecular Biology and Translational Science* 123: 351–86. https://doi.org/10.1016/B978-0-12-397897-4.00012-7.

Bhalla, Upinder S. 2014. “Molecular Computation in Neurons: A Modeling Perspective.” *Current Opinion in Neurobiology*, Theoretical and computational neuroscience, 25 (April): 31–37. https://doi.org/10.1016/j.conb.2013.11.006.

———. 2017. “Synaptic Input Sequence Discrimination on Behavioral Timescales Mediated by Reaction-Diffusion Chemistry in Dendrites.” *ELife* 6 (April). https://doi.org/10.7554/eLife.25827.

Bienenstock, E L, L N Cooper, and P W Munro. 1982. “Theory for the Development of Neuron Selectivity: Orientation Specificity and Binocular Interaction in Visual Cortex.” *The Journal of Neuroscience* 2 (1): 32–48.

Clarke, Bruce L. 1981. “Complete Set of Steady States for the General Stoichiometric Dynamical System.” *Journal of Chemical Physics* 75 (10): 4970–79.

Dan, Yang, and Mu-Ming Poo. 2004. “Spike Timing-Dependent Plasticity of Neural Circuits.” *Neuron* 44 (1): 23–30.

Danø, Sune, Mads F. Madsen, Henning Schmidt, and Gunnar Cedersund. 2006. “Reduction of a Biochemical Model with Preservation of Its Basic Dynamic Properties.” *The FEBS Journal*

45

273 (21): 4862–77. https://doi.org/10.1111/j.1742-4658.2006.05485.x.
Deuflhard, P., and J. Heroth. 1996. “Dynamic Dimension Reduction in ODE Models.” In *Scientific Computing in Chemical Engineering*, edited by Frerich Keil, Wolfgang Mackens, Heinrich Voß, and Joachim Werther, 29–43. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-80149-5_4.
Dormanns, K., R. G. Brown, and T. David. 2016. “The Role of Nitric Oxide in Neurovascular Coupling.” *Journal of Theoretical Biology* 394 (April): 1–17. https://doi.org/10.1016/j.jtbi.2016.01.009.
Frey, U, and R G Morris. 1997. “Synaptic Tagging and Long-Term Potentiation.” *Nature* 385 (6616): 533–36. https://doi.org/10.1038/385533a0.
Gansner, Emden R., and Stephen C. North. 2000. “An Open Graph Visualization System and Its Applications to Software Engineering.” *Software - Practice and Experience* 30 (11): 1203–33.
Goldbeter, A, and D E Koshland. 1981. “An Amplified Sensitivity Arising from Covalent Modification in Biological Systems.” *Proceedings of the National Academy of Sciences of the United States of America* 78 (11): 6840–44.
Hayer, Arnold, and Upinder S Bhalla. 2005. “Molecular Switches at the Synapse Emerge from Receptor and Kinase Traffic.” *PLoS Computational Biology* 1 (2): 137–54. https://doi.org/10.1371/journal.pcbi.0010020.
Higgins, David, Michael Graupner, and Nicolas Brunel. 2014. “Memory Maintenance in Synapses with Calcium-Based Plasticity in the Presence of Background Activity.” *PLoS Computational Biology* 10 (10): e1003834. https://doi.org/10.1371/journal.pcbi.1003834.
Hoops, Stefan, Sven Sahle, Ralph Gauges, Christine Lee, Jürgen Pahle, Natalia Simus, Mudita Singhal, Liang Xu, Pedro Mendes, and Ursula Kummer. 2006. “COPASI--a COmplex PAthway SImulator.” *Bioinformatics (Oxford, England)* 22 (24): 3067–74. https://doi.org/10.1093/bioinformatics/btl485.
Hudmon, Andy, Eric Lebel, Hugo Roy, Attila Sik, Howard Schulman, M Neal Waxham, and Paul De Koninck. 2005. “A Mechanism for Ca2+/Calmodulin-Dependent Protein Kinase II Clustering at Synaptic and Nonsynaptic Sites Based on Self-Association.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 25 (30): 6971–83. https://doi.org/10.1523/JNEUROSCI.4698-04.2005.
Jain, Pragati, and Upinder S Bhalla. 2009. “Signaling Logic of Activity-Triggered Dendritic Protein Synthesis: An MTOR Gate but Not a Feedback Switch.” *PLoS Computational Biology* 5 (2): e1000287. https://doi.org/10.1371/journal.pcbi.1000287.
Kholodenko, B. N. 2000. “Negative Feedback and Ultrasensitivity Can Bring about Oscillations in the Mitogen-Activated Protein Kinase Cascades.” *European Journal of Biochemistry* 267 (6): 1583–88. https://doi.org/10.1046/j.1432-1327.2000.01197.x.
Kim, BoHung, Sarah L Hawes, Fawad Gillani, Lane J Wallace, and Kim T Blackwell. 2013. “Signaling Pathways Involved in Striatal Synaptic Plasticity Are Sensitive to Temporal Pattern and Exhibit Spatial Specificity.” *PLoS Computational Biology* 9 (3): e1002953. https://doi.org/10.1371/journal.pcbi.1002953.
Lisman, J. 1989. “A Mechanism for the Hebb and the Anti-Hebb Processes Underlying Learning and Memory.” *Proceedings of the National Academy of Sciences of the United States of America* 86 (23): 9574–78.
Lomnitz, Jason G., and Michael A. Savageau. 2016. “Design Space Toolbox V2: Automated Software Enabling a Novel Phenotype-Centric Modeling Strategy for Natural and Synthetic Biological Systems.” *Frontiers in Genetics* 7: 118. https://doi.org/10.3389/fgene.2016.00118.
Markevich, Nick I, Jan B Hoek, and Boris N Kholodenko. 2004. “Signaling Switches and Bistability Arising from Multisite Phosphorylation in Protein Kinase Cascades.” *The*

 46

Journal of Cell Biology 164 (3): 353–59.
Markram, Henry, Eilif Muller, Srikanth Ramaswamy, Michael W. Reimann, Marwan Abdellah, Carlos Aguado Sanchez, Anastasia Ailamaki, et al. 2015. "Reconstruction and Simulation of Neocortical Microcircuitry." *Cell* 163 (2): 456–92. https://doi.org/10.1016/j.cell.2015.09.029.
Maurya, M R, S J Bornheimer, V Venkatasubramanian, and S Subramaniam. 2005. "Reduced-Order Modelling of Biochemical Networks: Application to the GTPase-Cycle Signalling Module." *IEE Proceedings. Systems Biology* 152 (4): 229–42.
Mayer, Jürgen, Khaled Khairy, and Jonathon Howard. 2010. "Drawing an Elephant with Four Complex Parameters." *American Journal of Physics* 78 (6): 648–49. https://doi.org/10.1119/1.3254017.
Novák, Béla, and John J Tyson. 2004. "A Model for Restriction Point Control of the Mammalian Cell Cycle." *Journal of Theoretical Biology* 230 (4): 563–79.
Oliveira, Rodrigo F, Anna Terrin, Giulietta Di Benedetto, Robert C Cannon, Wonryull Koh, MyungSook Kim, Manuela Zaccolo, and Kim T Blackwell. 2010. "The Role of Type 4 Phosphodiesterases in Generating Microdomains of CAMP: Large Scale Stochastic Simulations." *PloS One* 5 (7): e11725. https://doi.org/10.1371/journal.pone.0011725.
Quaiser, Tom, Anna Dittrich, Fred Schaper, and Martin Mönnigmann. 2011. "A Simple Work Flow for Biologically Inspired Model Reduction - Application to Early JAK-STAT Signaling." *BMC Systems Biology* 5 (1): 30. https://doi.org/10.1186/1752-0509-5-30.
Radulescu, Ovidiu, Alexander N. Gorban, Andrei Zinovyev, and Alain Lilienbaum. 2008. "Robust Simplifications of Multiscale Biochemical Networks." *BMC Systems Biology* 2 (October): 86. https://doi.org/10.1186/1752-0509-2-86.
Ray, Subhasis, and Upinder S Bhalla. 2008. "PyMOOSE: Interoperable Scripting in Python for MOOSE." *Frontiers in Neuroinformatics* 2: 6. https://doi.org/10.3389/neuro.11.006.2008.
Roelfsema, Pieter R., and Anthony Holtmaat. 2018. "Control of Synaptic Plasticity in Deep Cortical Networks." *Nature Reviews. Neuroscience* 19 (3): 166–80. https://doi.org/10.1038/nrn.2018.6.
Savageau, Michael A. 2001. "Design Principles for Elementary Gene Circuits: Elements, Methods, and Examples." *Chaos* 11 (1): 142–59. https://doi.org/10.1063/1.1349892.
Shouval, Harel Z, Mark F Bear, and Leon N Cooper. 2002. "A Unified Model of NMDA Receptor-Dependent Bidirectional Synaptic Plasticity." *Proceedings of the National Academy of Sciences of the United States of America* 99 (16): 10831–36. https://doi.org/10.1073/pnas.152343099.
Singh, Dilawar, and Upinder Singh Bhalla. 2018. "Subunit Exchange Enhances Information Retention by CaMKII in Dendritic Spines." *ELife* 7. https://doi.org/10.7554/eLife.41412.
Smolen, Paul, Douglas A. Baxter, and John H. Byrne. 2012. "Molecular Constraints on Synaptic Tagging and Maintenance of Long-Term Potentiation: A Predictive Model." *PLoS Comput Biol* 8 (8): e1002620. https://doi.org/10.1371/journal.pcbi.1002620.
Snowden, Thomas J., Piet H. van der Graaf, and Marcus J. Tindall. 2017. "Methods of Model Reduction for Large-Scale Biological Systems: A Survey of Current Methods and Trends." *Bulletin of Mathematical Biology* 79 (7): 1449–86. https://doi.org/10.1007/s11538-017-0277-2.
Sorokina, Oksana, Anatoly Sorokin, and J. Douglas Armstrong. 2011. "Towards a Quantitative Model of the Post-Synaptic Proteome." *Molecular BioSystems* 7 (10): 2813–23. https://doi.org/10.1039/c1mb05152k.
Stiles, Joel R, and Thomas M Bartol. 2001. "Monte Carlo Methods for Simulating Realistic Synaptic Microphysiology Using MCell." In *Computational Neuroscience: Realistic Modeling for Experimentalists, Editor Erik De Schutter*, 87–127. CRC Press.
Südhof, Thomas C. 2017. "Synaptic Neurexin Complexes: A Molecular Code for the Logic of

 47

Neural Circuits." *Cell* 171 (4): 745–69. https://doi.org/10.1016/j.cell.2017.10.024.
Taylor, Stephanie R., Francis J. Doyle, and Linda R. Petzold. 2008. "Oscillator Model Reduction Preserving the Phase Response: Application to the Circadian Clock." *Biophysical Journal* 95 (4): 1658–73. https://doi.org/10.1529/biophysj.107.128678.
Wallace, E W J, D T Gillespie, K R Sanft, and L R Petzold. 2012. "Linear Noise Approximation Is Valid over Limited Times for Any Chemical System That Is Sufficiently Large." *IET Systems Biology* 6 (4): 102–15. https://doi.org/10.1049/iet-syb.2011.0038.
Welsh, Ciaran M., Nicola Fullard, Carole J. Proctor, Alvaro Martinez-Guimera, Robert J. Isfort, Charles C. Bascom, Ryan Tasseff, Stefan A. Przyborski, and Daryl P. Shanley. 2018. "PyCoTools: A Python Toolbox for COPASI." *Bioinformatics (Oxford, England)* 34 (21): 3702–10. https://doi.org/10.1093/bioinformatics/bty409.
Wils, Stefan, and Erik De Schutter. 2009. "STEPS: Modeling and Simulating Complex Reaction-Diffusion Systems with Python." *Frontiers in Neuroinformatics* 3: 15. https://doi.org/10.3389/neuro.11.015.2009.
Wilson, Daniel E., David E. Whitney, Benjamin Scholl, and David Fitzpatrick. 2016. "Orientation Selectivity and the Functional Clustering of Synaptic Inputs in Primary Visual Cortex." *Nature Neuroscience* 19 (8): 1003–9. https://doi.org/10.1038/nn.4323.
Zhu, Ciyou, Richard H. Byrd, Peihuang Lu, and Jorge Nocedal. 1997. "Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization." *ACM Transactions on Mathematical Software* 23 (4): 550–60. https://doi.org/10.1145/279232.279236.

48