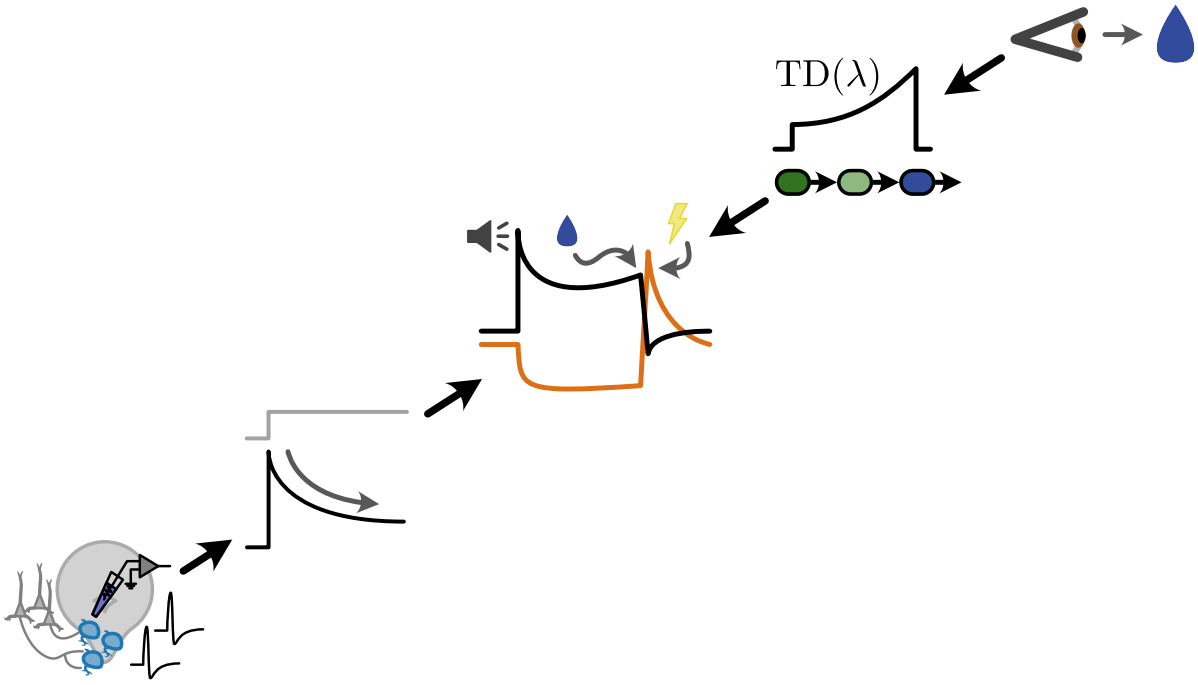
Making Sense of Serotonin Through Spike Frequency Adaptation

Emerson Harkin

Thesis submitted in partial fullfillment of the requirements for the degree of Doctorate in Philosophy in Neuroscience

Department of Cellular & Molecular Medicine
Faculty of Medicine
University of Ottawa


TD(λ)

Abstract

What does serotonin do? Just as the diffuse axonal arbours of midbrain serotonin neurons touch nearly every corner of the forebrain, so too is this ancient neuromodulator involved in nearly every aspect of learning and behaviour. The role of serotonin in reward processing has received increasing attention in recent years, but there is little agreement about how the perplexing responses of serotonin neurons to emotionally salient stimuli should be interpreted, and essentially nothing is known about how they arise. Here I approach these two aspects of serotonergic function in reverse order. In the first part of this thesis, I construct an experimentally-constrained spiking neural network model of the dorsal raphe nucleus (DRN), the main source of forebrain serotonergic input, and characterize its signal processing features. I show that potent spike-frequency adaptation deeply shapes DRN output while other aspects of its physiology are relatively less important. Overall, this part of my work suggests that *in vivo* serotonergic activity patterns arise from a temporal-derivative-like computation. But the temporal derivative of what? In the second part, I consider the possibility that the DRN is driven by an input that represents cumulative future reward, a quantity called state value in reinforcement learning theory. The resulting model reproduces established tuning features of serotonin neurons, including phasic activation by reward predicting cues and punishments, reward-specific surprise tuning, and tonic modulation by reward and punishment context. Because these features are the basis of many and varied existing serotonergic theories, these results show that my theory, which I call value prediction, provides a unifying perspective on serotonergic function. Finally, in an empirical test of the theory, I re-analyze data from an *in vivo* trace conditioning experiment and find that value prediction accounts for the firing rates of serotonin neurons to a precision $\ll 0.1\,\mathrm{Hz}$, outperforming previous models by a large margin. Here I establish serotonin as a new neural substrate of prediction and reward, a significant step towards understanding the role of serotonin signalling in the brain.

Acknowledgements

First and foremost, I would like to thank my supervisor, Richard Naud, for taking a leap of faith five years ago and offering a spot in a computational neuroscience lab to someone without much formal training in mathematics and none at all in programming. Thank you for your support, guidance, and mentorship over the years.

I would also like to thank Jean-Claude Béïque for always finding time for stimulating discussions and to provide feedback on the work of a lab alumnus. I am also grateful for your support and encouragement in the aftermath of the trucker occupation on my doorstep, without which this thesis might have been shorter and less interesting.

Thank you to my thesis advisory committee members Jean-Claude Béïque, Simon Chen, André Longtin, and Leonard Maler for many helpful discussions and feedback on my progress.

Trying to help people to answer their burning scientific questions has been one of the most enjoyable parts of my studies. Thank you to all of my formal and informal collaborators, especially Simon Chen, Candice Lee, Michael Lynn, Saskia de Vries, George Stuyt, and Kai Lun Teh, for sharing interesting problems and broadening my scientific horizons.

Reading and assessing a Ph.D. thesis is not a small committment. Thank you to Tuan Bui, Katalin Tóth, Jérémie Lefebvre, and Zachary Mainen for taking the time to evaluate my work.

Thank you to Michael Lynn, Sébastien Maillé, Éloïse Valérie Giraud, Candice Lee, and all members of the Naud and Béïque labs for many helpful discussions and making RGN a fun place to be.

Although he will probably never read this, I would like to thank James Acaster for inspiring words on the Great British Baking Show.

Last but not least, thank you Rachel. Here's to our next adventure together, whatever that might be.

Contents

List of figures

List of tables

List of abbreviations

4-AP	**4-aminopyridine.** Potassium channel blocker.
5-HT	**5-hydroxytryptamine, serotonin.** Neuromodulator derived from tryptophan with important but poorly-understood roles in regulating learning, behaviour, and mood.
AHP	**After-hyperpolarization potential.** A decrease in membrane voltage following an action potential.
CI	**Confidence interval.** A range of values expected to include the true value of some unknown quantity a specified proportion of the time (typically 95 %).
CS	**Conditioned stimulus.** A stimulus that an animal can be trained to associate with an impending reward or punishment.
DRN	**Dorsal raphe nucleus.** Midbrain nucleus that is the main source of serotonergic input to the forebrain.
vmDRN	**Ventromedial DRN.** Subregion of the DRN enriched in cortical-projecting neurons that co-release glutamate.
FFI	**Feed-forward inhibition.** Inhibitory input driven by an external source.
GABA	**γ-amino-butyric acid.** Inhibitory neurotransmitter.
GIF	**Generalized integrate-and-fire (model).** Extended LIF neuron model with stochastic spiking and adaptation.
aGIF	**Augmented generalized integrate-and-fire (model).** GIF neuron model selectively augmented with Hodgkin–Huxley currents.
IQR	**Interquartile range.** Range from the 25th to 75th percentile of a set of measurements.
ITI	**Inter-trial interval.** The period between two trials in a behavioural experiment.
LIF	**Leaky integrate-and-fire (model).** Simple neuron model consisting of an RC circuit that resets once a voltage threshold is reached.
mPFC	**Medial prefrontal cortex.** Cortical region involved in executive function and mood.
RL	**Reinforcement learning.** Sub-field of machine learning concerned with maximizing reward through trial and error.

RPE **Reward prediction error.** The difference between observed and expected reward in a temporal difference reinforcement learning algorithm.

SD **Standard deviation.** The square root of the mean squared error of a measurement, or the square root of the variance of a random variable.

SOM **Somatostatin.** Neuropeptide marker of a class of GABA neuron abundant in the DRN.

TD **Temporal difference (learning).** Type of reinforcement learning algorithm centred around correcting estimated value based on feedback after a (usually short) time interval.

TEA **Tetraethylammonium.** Potassium channel blocker.

TTX **Tetrodotoxin.** Pufferfish toxin that blocks voltage-gated sodium channels.

US **Unconditioned stimulus.** A stimulus that an animal recognizes as intrinsically rewarding or aversive.

VTA **Ventral tegmental area.** Dopaminergic midbrain nucleus known for reward prediction error-like activity patterns.

Chapter 1

General introduction

> Serotonin is an enigma. It is at once implicated in virtually everything, but responsible for nothing.
>
> Jacobs & Fornal, 1995

1.1 Serotonin

Serotonin has aged well. The evolutionary origin of the synthesis of 5-hydroxytryptamine, better known as serotonin, from the amino acid tryptophan predates the advent of animal life itself [1, 2]. And yet, somehow, this ancient neurotransmitter plays a role in shaping conscious experience that is as mysterious as it is powerful [3].

The clinical pharmacology of the serotonin system exemplifies this duality. Serotonin receptor agonists induce profound subjective changes in thinking and perception in humans, from long-lasting improvements in mood to visual hallucinations [4], while serotonin receptor antagonists are commonly used in the management of psychosis [5, 6]. Drugs that enhance or prolong the action of endogenous serotonin by slowing its breakdown or causing it to linger in the extracellular space have been used for decades to treat psychiatric conditions from major depression to anxiety and obsessive-compulsive disorder [7–10]. At the same time, whether or not serotonergic pathologies are actually the root cause of the despair and low energy characteristic of depression, for example, is still hotly debated [11, 12]. In spite of the prominent place serotonin occupies in medicine (and popular culture [13]), it is difficult to say what exactly this ubiquitous neurotransmitter does in the brain.

To be sure, the serotonin system is very well understood in some important respects. Anatomically, the vertebrate central serotonin system consists mainly of one midbrain region, the dorsal raphe nucleus (DRN), the neurons of which make reciprocal connections with nearly every part of the forebrain [14–19]. The post-synaptic effects of serotonin are mediated by a zoo of serotonin receptors which can be divided into seven types and at least fourteen subtypes [1]. Most of these are G-protein coupled receptors with inhibitory effects on target cells, with the excitatory G_q-coupled 5-HT_2 and ionotropic 5-HT_3 receptors being important exceptions. Serotonin neurons exhibit slow, pacemaker-like firing punctuated by activity that is tightly time-locked to external stimuli *in vivo* [20–25]. *In vitro*, serotonin neurons are characterized by unusually high membrane resistance and large and long-lasting after-hyperpolarization

potentials (AHPs) [26, 27]. The activity levels of serotonin neurons are regulated not only by potent AHPs, but also by feed-forward inhibition (FFI) from somatostatin-expressing GABAergic interneurons [19, 28], which are abundant in the lateral wings of the DRN [29, 30], as well as by 5-HT_{1A} receptor-mediated recurrent inhibition [31, 32]. Anatomical evidence suggests that these network-level sources of inhibition are roughly balanced [17].

Beyond these broad anatomical and electrophysiological features, the serotonin system is difficult to characterize. There is a degree of variability in nearly every aspect of the biology of serotonin neurons [33], including their anatomical connectivity [34–39], electrophysiological features [27, 37, 40], developmental origin [41], and even co-release of other neurotransmitters [42] (but this last point may be species-dependent [30]). Determining to what extent such differences are functionally important is an exciting and active area of research (*e.g.*, [38, 39]).

The behavioural picture is also muddy. Some of the earliest theories in this area proposed that serotonin might be involved in behavioural inhibition [43] or, conversely, facilitation of repetitive movements [3]. While there is evidence of serotonin's involvement in regulating the gain of motor commands in fish [44], and stress-induced changes in serotonin signalling affect locomotion in *Drosophila* [45], activation of serotonin neurons does not affect rodent movement in a consistent way [46]. Indeed, the effects of serotonergic manipulations on locomotion in mice can be inhibitory or facilitatory depending on the duration of the manipulation [47], similar to the duration-dependent effects of serotonin on the crayfish escape reflex [48]. Serotonergic drugs with antidepressant effects in humans have a clear behavioural signature of promoting active escape in rodents [49, 50], and one of the first applications of optogenetics was to show that direct activation of DRN neurons has the same effect [51]. Confusingly, however, there is a possibility that the same medications may decrease movement and increase fatigue in humans [52, 53]. These messy but intriguing conceptual connections across different branches of the tree of life hint that whatever serotonin does, it may be related to the most primitive aspects of behaviour and cognition.

One of the best studied yet most puzzling aspects of serotonergic function is its involve-

ment in reward processing. Optogenetic activation of serotonin neurons seems to promote maintenance of reward-seeking behaviour in rodents [54–56], for example, while chemogenetic inhibition reduces the rate at which depleted sources of reward are abandoned [25, 57]. The apparently paradoxical nature of this effect — activation promotes persistence, inactivation promotes perseveration — is maddeningly typical: serotonin neurons are activated by rewarding stimuli [58] but also by punishments [24, 39, 57], and they are more strongly activated by rewards that are surprising [24, 57] but have no such preference for punishments [57]. These seemingly-contradictory observations have spawned a wide range of theories that are equally difficult to reconcile. Over just the past decade or so, serotonin has been proposed to encode current and expected future reward [59–62]; or current reward and punishment [24, 39]; or surprise [57]; or reward prediction and reward prediction error [63]; or uncertainty, its opposite, and its rate of change [25]. "Puzzling" might be an understatement.

Existing ideas about serotonergic responses to rewards and punishments might seem contradictory, but they are generally not. The main achievement of this work is to show that previous observations and theories reflect different aspects of an adapting code for cumulative future reward. I call this unifying theory value prediction, after the definition of cumulative future reward as value in reinforcement learning [64] and after the role of adaptation in predictive coding [65]. The ability of value prediction to explain such a wide range of observations and ideas turns around the distinction between what the activity patterns of serotonin neurons *resemble* — a menagerie of phasic and tonic responses to rewards and punishments — and what they *encode* — value.

In the remainder of this introduction, I present the meaning of value in more depth (Section 1.2), briefly argue why the serotonin system might encode value through activity patterns that at first glance appear quite different from a value signal (Section 1.3), and introduce the basics of the technical approach I will use to deduce the coding features of the serotonin system (Section 1.4). Finally, I outline the main results that lead to my overall conclusion: serotonin predictively encodes value through adaptation (Section 1.5).

1.2 Reinforcement learning

An increasingly popular way to understand the function of the serotonin system is through the lens of reinforcement learning (RL) (*e.g.*, [25]), a theoretical framework with roots in psychological theories of animal learning that is now widely used in engineering applications to solve problems that involve multi-step decision making and online feedback. This section gives a very brief overview of the intertwined history of RL and neuroscience followed by a tutorial on the elements of RL theory relevant to this work. The main focus of this section is to develop an intuitive understanding of the critical importance of value functions, and in particular the state value function $v_\pi(s)$, in reward-based learning.

1.2.1 History

Of the three main divisions of machine learning, reinforcement learning (RL), which formalizes the process of reward-based learning from experience, has particularly close ties to psychology and neuroscience. (For a short history, see [66, 67].) The main ideas behind RL theory can be traced back to some of the first psychological theories of animal learning, introduced in the late XIX$^{\text{th}}$ and early XX$^{\text{th}}$ centuries. Thorndike's Law of Effect [68], for example, proposes that animal actions that are followed by reward are more likely to be repeated, and Pavlov described a stimulus that drives conditioning as a "reinforcer."[1] The first major intersection of RL and neuroscience occurred towards the end of the XX$^{\text{th}}$ century, after George Tesauro popularized temporal difference (TD) learning by using the eligibility trace-based TD(λ) algorithm to train an artificial neural network with fine-tuned inputs to play backgammon at a professional level [69]. Within a few years, Schultz, Dayan, and Montague [70] brought attention to a population of cells with activity patterns and reinforcing effects strikingly similar to the reward prediction errors of TD learning: dopamine neurons. This seminal result has deeply shaped the dopamine field, and today the RPE hypothesis is still the gold standard

[1] Or at least his English translator did.

against which new theories are measured (see for example [71, 72]). Attempts to pin down the function of other neuromodulators, notably including serotonin, in terms of RL theory [73, 74] have not stood the test of time to the same extent [75, 76], at least so far.

Compared with supervised and unsupervised learning [77], the other two main divisions of machine learning, RL is defined by its focus on learning through trial and error[2]. The apparent simplicity and broad scope of this approach does not limit its power. Over the past decade, RL has once again been ushered into the spotlight by impressive successes in training artificial neural networks to play games [78–81]. However, unlike Tesauro's TD-Gammon, the best-performing artificial agents are now trained *without* hand-crafted, game-specific features [82], or even knowledge of the rules [83]. Apparently, trial and error is enough to achieve human or even superhuman performance in many tasks. (But not all. Games with a very large set of possible actions such as StarCraft and Angry Birds remain out of reach for learning strategies based only on trial and error [84, 85].) The pattern of models built on specialized prior knowledge eventually being superseded by general models built on simple learning rules occurs frequently enough that it has been called the bitter lesson of AI [86].

RL is an attractive place to look for a theory of serotonin because of its close conceptual connection with reward-based learning, in which serotonin seems to play a role [25, 57, 87]; the intertwined history of RL, psychology, and neuroscience; and the impressive ability of extremely simple reinforcement learning rules to find solutions to complex tasks.

[2]The distinction between RL and supervised learning can seem subtle. In RL, a reward signal is available to guide or "supervise" learning, and modern supervised techniques often involve learning through stochastic gradient descent, essentially a type of trial and error. In my view, the critical difference between the two is the type of quantity available to guide learning. In supervised learning, the guiding quantity is the target output, and learning involves minimizing the difference between the target and the actual output of some model. In RL, the guiding quantity is the reward signal, which is not the target. The actual targets are the optimal policy and corresponding value function, neither of which are known (at least until learning is complete). In short, whereas supervised learning *can* occur via trial and error, reinforcement learning *must* use trial and error because the target is not available.

1.2.2 Tutorial

If serotonin participates in RL, what might its role be? To begin answering this question, we must first establish some basic concepts from RL theory and why they are important.

Goal of reinforcement learning

The scope of RL theory is extremely broad, but its core elements are very simple [64]. RL theory focuses on the interactions between an entity called an agent and its environment, where the agent endowed with the ability to choose actions A that influence the state S of the environment, and the environment provides feedback to the agent in the form of rewards $R_t \sim \Pr[R_t \mid S_t]$ and new states $S_{t+1} \sim \Pr[S_{t+1} \mid A_t, S_t]$. The goals of RL algorithms are to find a rule for choosing actions, referred to as a behavioural policy and formally defined as a probability distribution over the actions available in a given state $\pi(a \mid s) = \Pr[A_t = a \mid S_t = s]$, that maximizes the total future reward $R_{t+1} + R_{t+2} + R_{t+3} + \cdots$. Since rare catastrophic events might make it impossible to collect rewards far in the future, it is common to discount distant rewards at a rate that implicitly reflects the assumed level of risk in the environment [88, 89]. Subject to this preference for immediate rewards, the RL objective can be written as

$$\pi^\star = \arg\max_\pi \; \mathbb{E}_{\pi, \Pr[S_{t+1} \mid S_t, A_t]} \left[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1} \;\middle|\; S_t = s \right], \tag{1.1}$$

where $\pi^\star$ is the reward-maximizing policy, $\arg\max_x f(x)$ denotes the value of x that maximizes $f(x)$, $\mathbb{E}_{\Pr[X \mid Y]}[X \mid Y]$ denotes conditional expectation, $0 \leq \gamma \leq 1$ is a discounting factor that controls the relative weighting of immediate and distant rewards, and s is the current state.

To understand the meaning of the abstract elements of RL theory, consider the case of a person playing a boardgame. In this example, the agent is the player, the environment is the game itself, the actions A are the ways that the player can move the gamepieces, the states S are the possible positions of pieces on the board, and the reward R might be a positive

number if the player wins the game on the current turn, negative if the player loses, and zero otherwise. The objective of the player is to win the game, ideally as quickly as possible. In RL, this translates to the idea of maximizing reward by finding a policy for moving pieces based on the current state of the board that wins the game in only a few turns.

Why maximizing reward is not trivial

The main challenge that RL theory aims to solve is that future rewards are very difficult to work out in advance. In the boardgame example, it could be that the opposing player's moves are unpredictable. As a result, the board could look very different a few turns in the future no matter what action the agent takes. In the face of this unpredictability, how can the agent know whether a particular move in the current turn is better than another, let alone which move is best? This example illustrates the general difficulty of finding a policy that maximizes the long-run reward when future states and rewards are uncertain.

State-action value functions enable reward maximization

Fortunately, a critical theoretical result in RL, called the policy improvement theorem [90, 91], states that it is possible to determine whether a given action in the current state is better than another, even without *direct* knowledge of future rewards. This result can be shown in terms of a hypothetical table that relates states and actions to total future reward called a q value function. The q value function, defined as

$$q_\pi(s,a) = \mathbb{E}_{\pi,\Pr[S_{t+1}|S_t,A_t]}\left[\sum_{i=0}^{\infty}\gamma^i R_{t+i+1} \,\middle|\, S_t = s, A_t = a\right], \tag{1.2}$$

is exactly the same as the objective to be maximized from Eq. (1.1) except for its dependence on the current action $A_t = a$. (The subscript π in q_π is used to indicate that actions beyond A_t are chosen according to the behavioural policy.) Assuming we have access to such a table, it is possible to determine whether a given action a taken in the current state s is better

than a different action b by simply looking up the rows corresponding to (s, a) and (s, b) and comparing the values. By extension, it is possible to find the long-term reward-maximizing policy $\pi^\star$ by always taking the action with the highest q value.

The q value function is a useful theoretical concept for comparing actions and finding better policies, but it is not obvious how it addresses the problem of uncertain long-term rewards. After all, the definition of $q_\pi(s, a)$ given in Eq. (1.2) includes an explicit sum over rewards infinitely far in the future $R_{t+1}, R_{t+2}, R_{t+3}, \ldots$ The main insight of the policy improvement theorem comes from the fact that the q value can be broken down into two parts: the immediate reward R_{t+1} and all subsequent rewards $R_{t+2}, R_{t+3}, \ldots$ It turns out that the second part is actually an average over the possible q values of the different actions that can be taken in the very next state S_{t+1}, so the q value function can be rewritten recursively as

$$q_\pi(s, a) = \mathbb{E}_{\pi, \Pr[S_{t+1}|S_t, A_t]} \left[R_{t+1} + \gamma q_\pi(S_{t+1}, A_{t+1}) \mid S_t = s, A_t = a \right]. \tag{1.3}$$

Taking the interpretation of the q value function as a table, the expression above means that the value of taking a particular action in a given state is the reward delivered immediately after taking the action plus an average of only a few other rows in the table. Counter-intuitively, the fact that the q value can be written in terms of itself means that it is possible to select actions that maximize the total future reward without directly adding up rewards far in the future[3].

A simpler value function for neuroscience experiments

From a normative perspective, the q value function introduced above plays a critical role in finding the reward-maximizing policy $\pi^\star$. However, many neuroscience experiments are conducted in highly trained animals which have already settled on a policy[4] (*e.g.*, [25]), or

[3] In this section, I have focused on the theoretical properties of the q value function rather than how it can be estimated from observed rewards in practice because the latter is not critical to my work. For an introduction, refer to [64, 92].

[4] Hopefully a policy close to the optimal policy $\pi^\star$, but not necessarily.

else focus on tasks that do not involve selecting actions, such as classical conditioning (*e.g.*, [58]). In these situations, it is possible to define a new type of value function that can be interpreted as the average q values of the actions taken under a fixed policy (as in overtrained animals), or simply the total future reward independent of the action taken (as in classical conditioning). Because this new value function only depends on the current state, it is called the state value function

$$\begin{aligned} v_\pi(s) &= \sum_a q_\pi(s,a)\ \pi(A_t = a \mid S_t = s) \\ &= \mathbb{E}_{\pi,\Pr[S_{t+1}|S_t,A_t]}\left[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1} \;\middle|\; S_t = s\right]. \end{aligned}$$

Like the q value function, the state value function v depends on the behavioural policy π indicated by the subscript.

Importance of state value

The v value function is not just a particularly simple type of value function induced by the design of neuroscience experiments. Because it represents the average q value under a given policy, the state value $v_\pi(S_{t+1})$ can be used to replace $q_\pi(S_{t+1}, A_{t+1})$ in the recursive definition of the q value given in Eq. (1.3). This means that the future state value $v_\pi(S_{t+1})$ plays an important role in dictating which action is best in a given state; in fact, when rewards are sparse and/or discounting is low, the future state value can be the *main* determining factor. The fact that the future state value $v_\pi(S_{t+1})$ appears directly in the definition of the q value gives it an important normative role in action selection.

The role of the value of the current state $v_\pi(S_t = s)$ is less clear-cut from a normative perspective, but it often appears in useful strategies for learning the optimal policy $\pi^\star$. For example, policy gradient methods, which follow in the footsteps of Pavlov and Thorndike by reinforcing the selected action in proportion to its q value [93], often use the value of the current state $v_\pi(s)$ as an optional baseline to speed learning [94]. In principle, it is also

possible to reinforce the past action A_{t-1} in proportion to the current reward R_t and state value $v_\pi(s)$, since this is roughly equivalent to the q value used by policy gradient methods with a lag of one timestep. Finally, a mixture of the state value and its rate of change appear in the RPEs of TD learning in various ways [70, 95, 96].

Depending on your perspective, the state value v represents either an average over the q values of different immediately-available actions, or an expectation of total future reward. In spite of its simple structure, state value is of enormous normative and practical importance in RL. A neural system that encodes state value could play a central role in learning and behaviour.

1.2.3 Terminology: value function vs. value signal

In principle, the state value function v can be used to evaluate any state under any policy. However, in the section above we saw that the value of roughly the current state (strictly, either the future state S_{t+1} or the current state $S_t = s$, but these are almost the same in real time) under the current policy has special importance. To avoid confusion with the meaning of value function as a time-independent abstract table or mapping, here I use the term value *signal* to refer to a time-dependent quantity that represents the value of the current state under the current policy over time $v_t = v_\pi(s_t)$.

1.2.4 Summary

Value functions are fundamental to RL theory [64, 97]. By extension, they have a strong normative connection to anything that involves learning to behave in a way that maximizes some type of benefit. At the same time, the specific nature of the link between value and behaviour is not obvious: while the policy improvement theorem is the foundation on which many important RL methods are built, there are multiple strategies for finding value-increasing policy adjustments that use value in different ways. In a sense, Jacobs & Fornal's well-known comment that serotonin is involved in everything but sufficient for nothing [3] could also

apply to value. In the coming chapters, we will see that this similarity is not superficial.

1.3 Neural coding

In this work, I argue that the *in vivo* activity patterns of serotonin neurons encode a value signal. However, whether or not serotonergic activity patterns represent such a signal and whether or not they visually resemble one are not the same question [98]. Indeed, dynamic, non-linear representations of even simple underlying quantities are the rule in neuroscience rather than the exception: touch receptors respond to the onset and offset of sustained tactile stimuli [99, 100]; retinal ganglion cells are activated by unexpected patterns of light [101], as are neurons in visual cortex [102]; and high-level cortical representations of choice and reward change over time [103–105], to name just a few examples. This section introduces some basic principles of neural coding to develop an intuitive understanding of why the activity patterns of serotonin neurons might be different from the quantity they represent.

1.3.1 Rate coding

What is an activity pattern in the first place? The basic unit of activity in the central nervous system is the action potential, and because the shape of an action potential carries no information, neurons are said to communicate exclusively via sequences of action potentials called spiketrains [106]. The term "pattern" implies that there is some aspect or function of a spiketrain, in other words a statistic, that is reliable. To understand how serotonin neurons encode a value signal, then, is to understand how these cells convert a value into spiking statistics.

Rate coding presents a particularly intuitive way of thinking about the relationship between time-varying input and spiking statistics. According to this perspective, the intensity of an input is encoded in the number of spikes emitted by a group of neurons over a period of time, a non-negative quantity called the firing rate (Fig. 1.1A; [107]). This provides a simple

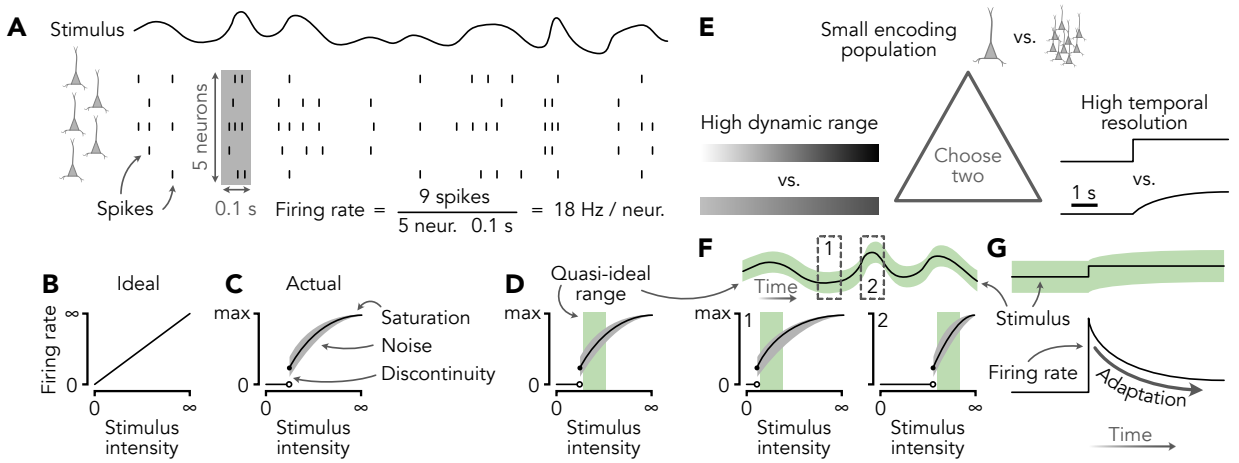


Figure 1.1: Overview of rate coding. **A** A time-varying input is encoded in the number of spikes emitted by a group of neurons over a short period of time, a quantity called the firing rate. **B–D** Relationship between input strength and firing rate for ideal (B) and real (C) rate codes. Real rate codes typically only resemble ideal rate codes within a narrow range of input intensities (D). **E** Rate coding tradeoffs. **F**, **G** Gain modulation and adaptation resolve rate-coding tradeoffs by extending dynamic range.

explanation of why, for example, stronger sensory stimuli evoke more rapid firing in and/or recruit a greater number of peripheral nerve fibres [108, 109] and why higher population firing rates in motor cortex produce faster movements [110].

Do serotonin neurons communicate a value signal using such a rate code? Probably not. In the oversimplified picture of rate coding that I have presented, there is a fixed, linear correspondence between the strength of an input and the number of spikes per unit time (Fig. 1.1B). There are normative reasons to believe that such a simple and inefficient coding scheme might not be used [111–113] as well as practical ones: Spiking is noisy for slowly-varying inputs [114, 115], the fact that neurons cannot fire arbitrarily fast causes spike counts to saturate for strong inputs [108, 116], and even for weak inputs the relationship between input intensity and firing is highly nonlinear in many cell types (Fig. 1.1C; called type II firing [117]). Static rate coding is therefore a conceptually-useful approximation that generally only holds within a narrow range of spike counts, and, therefore, input intensities (Fig. 1.1D). Because the upper limit of the spike count is set by the number of encoding neurons and integration time, static rate coding suggests important tradeoffs between population size,

temporal resolution, and dynamic range of the signal that can be encoded (Fig. 1.1E; *e.g.* [118, 119]). Importantly, this tradeoff hints that for neural systems with particularly small populations of slow-firing neurons, such as neuromodulatory systems, static rate coding might only apply to a very narrow range of inputs indeed.

1.3.2 Dynamic encoding

One way to address the compromises involved in the static rate coding picture is to remove the requirement that the relationship between input intensity and firing rate be fixed. By varying the sensitivity of a neural population to its input over time, it is possible to maintain spike counts within the range where the rate approximation holds (Fig. 1.1F).

The general idea of adjusting input sensitivity to optimize neural coding has many variations which go by different names. Gain modulation implies a purely multiplicative change in the sensitivity of a neural population [120], while normalization generally refers to the level of gain modulation that maintains population-level spike counts within an appropriate operating range [121]. Predictive coding involves removing the part of the input that is expected[5], leaving only a surprise-like signal with a lower dynamic range to drive spiking [65, 123]. Spike frequency adaptation (Fig. 1.1G; [124]) removes the part of the input that is correlated with past input, and can be seen as a specific type of predictive coding in which the surprise-like signal is related to the rate of change of the input [125, 126].

1.3.3 Implications

Neural coding is a thorny issue. The rules according to which neural systems translate inputs into spiking statistics are often messy and sometimes cryptic. Previous work on the serotonin system has side-stepped this problem by implicitly invoking the static rate coding assumption,

[5]Predictive coding is a confusing term because it is used to refer to both coding schemes in which the predicted part of the signal is removed as well as schemes in which a prediction is transmitted [122]. In the former sense, predictive coding can be seen as a type of compression scheme, while the latter sense is more connected to the prior or posterior in a Bayesian view of neural computation. To avoid confusion, here I use predictive coding in only the surprise/compression/redundancy reduction sense.

but a single quantity that directly matches the activity patterns of serotonin neurons has not yet been found. Understanding in what respects the coding features of the serotonin system do and do not agree with the idealized rate coding picture could shed considerable light on its function.

1.4 Single neuron models

Bottom-up modelling is a useful technique for deducing the coding features of neural systems. Given a set of models of the neurons in a particular system of interest, it is possible to determine to what extent individual neurons and neural populations exhibit the static, linear stimulus responses required by a rate code, or, conversely, whether features such as strong spike-frequency adaptation or firing rate discontinuities that suggest other coding schemes are present. For the bottom-up approach to succeed, it is imperative that the models of individual neurons that form the bedrock of this approach are accurate. But accurate in what sense? This section surveys three broad approaches to constructing single neuron models that prioritize accuracy in different respects. Finally, I introduce a new technique that I began developping during my M.Sc. that combines the strengths of previous methods.

1.4.1 Three types of accurate model

The first approach is to prioritize biophysical realism, leading to Hodgkin–Huxley-style models that explicitly describe the interacting effects of multiple ionic currents on membrane voltage. These models have produced deep insights into the mechanisms underlying action potential generation [127, 128] and the contributions of dendritic processing to neural computation [129–131]. At the same time, they are difficult to tune to match the spiking statistics of any individual neuron [132–135], and their parameters can be difficult to interpret in terms of functionally-meaningful phenomena such as firing rate discontinuities and adaptation.

A totally different approach is to construct a statistical model of the input-to-spiketrain

mapping, abstracting away the biology in between. Instead of prioritizing biophysical realism, this approach focuses on explicitly delineating which aspects of neural data can be explained in terms of the input and which aspects can be considered noise [136, 137]. Statistical models are typically much simpler to constrain than Hodgkin–Huxley style models (but see [138] for a statistical approach to fitting HH models) and produce a clearer picture of the relationship between input and spiking statistics, especially at the population level [139–142]. At the same time, they do not provide any direct insight into the biophysical mechanisms that mediate this relationship. They are also limited by design in terms of the types of input-output relationships they can describe; models without some sort of dependence on the past cannot capture adaptation, for example. Finally, because this approach directly models the relationship between input and spikes, skipping the membrane potential as an intermediate step, statistical models can be challenging to fit to cells that fire at very low rates, serotonin neurons included.

A third approach is to select a phenomenological neuron model that can capture a wide range of spiking statistics by adjusting a small number of parameters [143]. This class includes models that are essentially simplifications of Hodgkin and Huxley's biophysical model as well as integrate-and-fire models that abstract away the dynamics of membrane voltage during the action potential [144–149]. These models are not necessarily simple to tune to quantitatively match the spiking statistics of individual neurons, and they also lack the biophysical detail of Hodgkin–Huxley-style models. In spite of these limitations, phenomenological models have been the workhorses of computational neuroscience for many years because they lend themselves well to mathematical analysis and simulation [134, 143]. Compared with Hodgkin–Huxley-style and statistical models, phenomenological models are particularly well-suited to connecting voltage dynamics to coding features.

1.4.2 Augmented generalized integrate-and-fire models: the best of all worlds

An ideal modelling framework would combine the mathematical simplicity and insightfulness of phenomenological models, the ease of fitting of statistical models, and, where needed, the detail of Hodgkin–Huxley-style models. Augmented generalized integrate-and-fire (aGIF) models provide just such an approach (Appendix A).

The three useful features of modern aGIF models were introduced in stages by two different research groups. The generalized linear integrate-and-fire (GLIF) model introduced by Mihalaş and Niebur [150] consists of a leaky integrate-and-fire neuron enhanced with a spike-triggered adaptation current and voltage-dependent threshold. Like many phenomenological models, it is able to produce a wide range of firing statistics in spite of its mathematical simplicity. A few years later, Mensi, Naud, *et al.* [151] introduced the very similar generalized integrate-and-fire (GIF) model. Unlike the Mihalaş–Niebur GLIF model, the Mensi–Naud GIF model includes a statistical model of spiking, making it simple to tune to match the firing statistics of individual neurons with millisecond precision. The fitted parameters of GIF models are simple to interpret in terms of meaningful processing features such as integration timescale, spike frequency adaptation, and spiking variability, leading them to be used to shed light on the functional properties of a wide range of cortical cell types [152]. Finally, for cell types in which the subthreshold voltage dynamics are dominated by just one or two voltage-dependent currents, I introduced GIF models *augmented* with Hodgkin–Huxley conductances (see Appendix A and [153]). aGIF models largely preserve the simplicity and quantitative accuracy of GLIF and GIF models while providing insight into the ionic mechanisms that shape the spiking statistics of different cell types.

1.4.3 Models of cells in the serotonin system

A handful of serotonin neuron models have been reported in the literature (see [154] for review). However, because these models were generally created with a focus on understanding the sub-cellular ionic mechanisms underlying tonic firing and action potential generation [149, 155] or, separately, network-level phenomena such as oscillations [156, 157], only the aGIF model I introduced during my M.Sc. [153] has been validated in terms of its ability to precisely reproduce the spiketrains of serotonin neurons. Somatostatin-expressing GABAergic interneurons make up a large proportion of the cells in the serotonin system, but even fewer models of these cells have been reported. Unfortunately, the electrophysiological properties of these cells have not been characterized *in vitro*, so it is unclear to what extent existing models reproduce even their qualitative features. Overall, almost nothing is currently known about what serotonin system-level coding features are induced by the properties of the cells it contains.

1.5 Outline of the thesis

The overarching goal of this work is to provide a computational theory of serotonin that is precise, simple, and consistent. An ideal serotonergic theory should explain (1) how emotionally salient stimuli such as reward-predicting cues, rewards, and punishments produce the activity patterns of serotonin neurons observed *in vivo* (Fig. 1.2A left), and (2) how these activity patterns are translated into changes in behaviour and learning (Fig. 1.2A right). The fact that a computational theory of serotonin would have to reach almost from the sensory periphery to motor centres in order to be called "complete" is part of what makes this system fascinating. At the same time, this incredible breadth places a complete theory firmly beyond the scope of this thesis. Here, I focus on the first point to develop an explanation of how the reward and punishment tuning features of serotonin neurons arise and how they should be interpreted.

In Chapter 2, I study the signal processing features of the DRN with the goal of reaching a

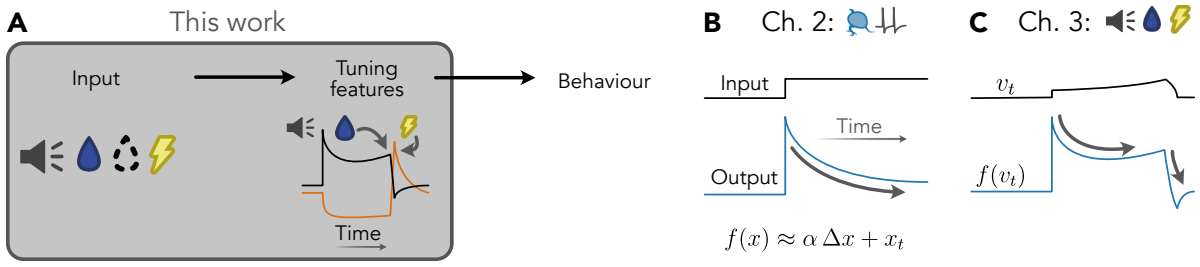


Figure 1.2: Visual table of contents. **A** Scope of the thesis compared with a complete theory of serotonin. I focus on explaining how the responses of serotonin neurons to rewards and punishments are generated, leaving the question of how these responses drive behaviour for future work. **B**,**C** Main ideas of the manuscripts that make up the thesis. In Chapter 2 (B), I build up a simplified picture of how the DRN processes arbitrary time-varying inputs, finding that the defining signal processing feature of the DRN is adaptation. In Chapter 3 (C), I consider how the signal processing features of the DRN might transform an input that represents cumulative future reward (*i.e.*, a value signal), finding that the resulting output is a very good qualitative and quantitative match for the *in vivo* activity patterns of serotonin neurons.

general understanding of how its input is transformed into serotonergic output without making assumptions about what exactly the input might be (Fig. 1.2B). I begin by presenting an experimentally-validated serotonin neuron model that I had created previously [153]. Through a combination of experiments and model-based analysis, I show that this model captures the main biophysical features of serotonin neurons. Using simulations, I explore how these neuron-level properties shape the signal processing features of the DRN as a whole. While some of the effects I uncover are nuanced and context-dependent, I argue that the dominant processing feature of the DRN is to signal the rate of change of its input, a computation that underlies a type of predictive coding.

After establishing the relationship between DRN input and output in a very general sense, I turn my attention to the question of what sort of input could produce the activity patterns of serotonin neurons observed *in vivo* (Fig. 1.2C). In Chapter 3, I use a value signal (see Section 1.2.3) to drive a simplified model of the DRN derived from the results of Chapter 2. By simulating a variety of reward learning experiments from the serotonin literature, I show that serotonergic activity patterns that have so far eluded simple theories are consistent with a predictive code for value. In an empirical test of the theory, I re-analyze a particularly rich

dataset of *in vivo* responses of serotonin neurons to stochastic rewards and show that my value prediction theory provides a quantitatively more accurate description of the activity patterns of serotonin neurons than competing ideas, almost always by a very large margin. Finally, I include a general discussion of the biological basis of value prediction and its potential implications for the role of serotonin in learning, behaviour, and beyond.

Chapter 4 summarizes the main results and briefly discusses my general approach in the broader context of computational neuroscience.

Significance: The history of the serotonin system is not yet written. The value prediction theory may turn out to be wrong in subtle or important ways, and in any case makes no attempt to directly explain the role of serotonin in regulating learning and behaviour. At the same time, this theory represents an important step forward because it provides a consistent interpretation of a wide range of experimental results, including some that were previously believed to be in conflict; makes a bottom-up connection between the biophysical features of serotonin neurons and the function of the serotonin system as a whole; and establishes a top-down link between the *in vivo* activity patterns of serotonin neurons and a specific, fundamental quantity of reinforcement learning theory.

1.6 Summary of research contributions

The main body of this thesis argues for a simple computational theory of serotonin. However, my work on the serotonin system is not restricted to this doctoral thesis, nor has my doctoral work been restricted to serotonin. What follows is a brief summary of my relevant work not included in the main body of this thesis.

- During my M.Sc. (2018), I created the simplified serotonin neuron model that forms the foundation of Chapter 2 as well as the computational component of my collaborator Michael Lynn's study of recurrent inhibition in the DRN [32].

- At the beginning of my Ph.D. (2020), I wrote a textbook chapter describing the conceptual approach I took to designing the serotonin neuron model, reproduced here in Appendix A. The augmented generalized integrate-and-fire (aGIF) approach I used to model serotonin neurons extends leaky integrate-and-fire (LIF) models by accounting for the effects of ionic currents.

- In the manuscript reproduced in Appendix B (2022), my co-authors and I present another possible extension, this time to account for the effects of dendritic morphology on firing patterns. (Reminiscent of [129].)

- Separately, around this time I concluded a collaboration with Candice Lee and Simon Chen (also at the University of Ottawa) to characterize the activity patterns of different motor cortical cell types during reward learning, published in 2022 [105]. My main contribution to this work was to validate, refine, and develop aspects of the analysis of calcium imaging data.

- Finally, in 2019 I joined an ongoing collaboration with the de Vries (Allen Institute for Brain Sciences), Kremkow (Charité University Berlin), and Palmer (University of Melbourne) labs to examine the role of dendritic processing in visual perception as

part of an AIBS OpenScope project. My main contributions to this project were to implement a data processing pipeline to extract the activity patterns of dendrites from calcium imaging data (remotely validated by George Stuyt, a Ph.D. student at the University of Melbourne) and oversee data analysis (remotely conducted by Rémi Renoud-Grappin, a M.Eng. student at the University of Lyon).

In total, over the course of my doctoral studies I have led or co-authored five research manuscripts [32, 105, 158–160]; written one textbook chapter [161]; given talks at Neuromatch [162] and the Canadian Computational Neuroscience Spotlight [163]; and presented posters at COSYNE [164], Bernstein Conference [165], Neuroscience [166], and the Canadian Association for Neuroscience annual meeting[6] [167]. Most recently, an abstract based on the work reproduced in Chapter 3 was accepted for a nanosymposium talk at Neuroscience 2023.

[6]Abstract accepted but spring 2020 conference cancelled due to COVID.

Chapter 2

The dorsal raphe nucleus computes rate of change through adaptation

You can observe a lot by just watching.

Yogi Berra

This chapter is reprinted from a previously-published work.

Citation

E.F. Harkin, M.B. Lynn, A. Payeur, J.-F. Boucher, L. Caya-Bissonnette, D. Cyr, C. Stewart, R. Naud*, and J.-C. Béïque*. Temporal derivative computation in the dorsal raphe network revealed by an experimentally-driven augmented integrate-and-fire modeling framework. *eLife*, **12**, e72951. (2023)

*: Equal contribution.

Statement of contributions

This work can be broadly divided into two parts: I carried out experiments and created the serotonin neuron model during my M.Sc., and carried out simulations and analysis during my Ph.D. A more detailed breakdown is given below.

I conceptualized the augmented generalized integrate-and-fire (aGIF) modelling approach, carried out *in vitro* electrophysiological experiments to characterize 5-HT and SOM neurons, and fitted and validated the 5-HT neuron model during my M.Sc. with Jean-Claude Béïque and Richard Naud. During my Ph.D. with Richard Naud, I finalized the single neuron models, conceptualized the bootstrap-based approach to population-level simulations, carried out all simulations and analysis, and conceptualized the main computational feature of the serotonin system as being a temporal derivative-like operation. I created all figures and wrote the first draft of the manuscript. Subsequent drafts were co-written by Richard Naud, Jean-Claude Béïque, and I with input from André Longtin. Richard Naud and Jean-Claude Béïque provided supervision and valuable input throughout this work.

Statement of contributions (continued)

Many other people made important contributions to various parts of this project. Jean-Claude Béïque, Richard Naud, André Longtin, Alexandre Payeur, and Michael B. Lynn conceptualized the creation of a simplified, experimentally-validated serotonin neuron model. Richard Naud created the generalized integrate-and-fire (GIF) modeling approach on which our model is based. Early efforts to create an experimentally-constrained serotonin neuron model by Michael B. Lynn, Alexandre Payeur, and Sébastien M. Maillé led me to create the aGIF model during my M.Sc. Michael B. Lynn, Jean-François Boucher, Léa Caya-Bissonnette, Dominic Cyr, and Chloe Stewart collected data under my direction.

Abstract

By means of an expansive innervation, the serotonin (5-HT) neurons of the dorsal raphe nucleus (DRN) are positioned to enact coordinated modulation of circuits distributed across the entire brain in order to adaptively regulate behavior. Yet the network computations that emerge from the excitability and connectivity features of the DRN are still poorly understood. To gain insight into these computations, we began by carrying out a detailed electrophysiological characterization of genetically identified mouse 5-HT and somatostatin (SOM) neurons. We next developed a single-neuron modeling framework that combines the realism of Hodgkin-Huxley models with the simplicity and predictive power of generalized integrate-and-fire models. We found that feedforward inhibition of 5-HT neurons by heterogeneous SOM neurons implemented divisive inhibition, while endocannabinoid-mediated modulation of excitatory drive to the DRN increased the gain of 5-HT output. Our most striking finding was that the output of the DRN encodes a mixture of the intensity and temporal derivative of its input, and that the temporal derivative component dominates this mixture precisely when the input is increasing rapidly. This network computation primarily emerged from prominent adaptation mechanisms found in 5-HT neurons, including a previously undescribed dynamic threshold. By applying a bottom-up neural network modeling approach, our results suggest that the DRN is particularly apt to encode input changes over short timescales, reflecting one of the salient emerging computations that dominate its output to regulate behavior.

2.1 Introduction

The forebrain-projecting serotonin (5-HT) neurons of the DRN play a key role in regulating behaviour in dynamic environments, but the precise nature of this role is still not well understood [51, 76, 168, 169]. DRN serotonin neurons have been proposed to modulate a wide range of cognitive processes, such as encouraging patience for future rewards [54, 55], signalling the beneficialness of current actions or states [61], complementing reinforcement signals of dopamine [23, 73, 170–174], and, partially as a corollary, regulating both learning [43, 57, 73, 175, 176] and mood [177–179]. While the remarkable diversity of roles attributed to this single neurotransmitter has historically been perplexing, recent findings are beginning to provide insight (see [33] for review). For example, the unsuspected organization of 5-HT neurons into anatomical sub-modules that differentially regulate behaviour [34–36, 38, 180], or the observation that 5-HT neurons can encode distinct salient features of the environment over different time scales [20, 23, 24, 58, 181] are compelling mechanisms that may contribute to the multiplicity of 5-HT's actions. These anatomical and dynamical perspectives on 5-HT diversity need not be mutually exclusive. A clearer understanding of the biophysical mechanisms that contribute to the coding features of raphe neurons over multiple timescales has the potential to substantially increase our understanding of how 5-HT regulates behavior.

The spiking statistics of 5-HT neurons necessarily shape and constrain their computational role. For instance, the slow firing rate ($\sim$5 Hz) of 5-HT neurons, in large part attributable to a large after-hyperpolarization potential (AHP) [22, 26], may appear to preclude signaling on faster time scales. However, fast signaling despite slow firing can arise naturally in ensemble-rate codes [182, 183]. Consistent with this idea, the in vivo population activity of 5-HT neurons has been observed to track impending rewards over second to sub-second timescales [58], and the trial-averaged ensemble rates of individual 5-HT neurons can track environmental changes over the millisecond time scale [23, 24]. In addition, the fact that 5-HT receptor subtypes can regulate the excitability of target neurons over different time scales, including ionotropic

5-HT3 receptors with millisecond gating kinetics [184–187], at the very least suggests that the 5-HT system is capable of fast information transmission, an observation mirrored by the fast dynamics of neurons which project to the DRN [188, 189]. If fast and slow signaling by the DRN is manifest, it is less clear which cellular mechanisms regulate the interplay between these timescales, nor which input features are represented on which timescales.

Computational modeling is a standard approach to link levels of description and is thus well suited to delineate how network-level function emerges from excitability features identified at the single-cell level. In spite of their conceptual utility, the most detailed single cell models, including those of DRN neurons [154, 155], do not lend themselves with ease to bottom-up modeling efforts because of the substantial technical difficulty of obtaining sufficiently accurate values for a large number of interacting model parameters [190, 191]. Mathematically simpler generalized integrate-and-fire (GIF) models provide a strong foundation for network modeling because their small number of parameters can be estimated with a high degree of precision [151, 152, 192]. This precision comes at a price, however: the process of distilling the effects of many biophysical mechanisms into a small number of model parameters makes it difficult to study a specific mechanism (e.g., a subthreshold ion channel) in isolation. A hybrid approach based on a reductionistic GIF model augmented with a limited set of biophysical mechanisms could leverage the precision of GIFs while allowing the ability to link specific biophysical mechanisms with higher-order network function.

In this study, we developed and validated for DRN neurons a hybrid modeling approach that lies between reductionist GIF and biophysical Hodgkin-Huxley-type models to capture excitability features of individual neurons for accurate simulations of population dynamics and, by extension, network computation inference. To this end, we carried out cellular electrophysiological recordings from genetically-identified DRN 5-HT and SOM neurons to

1. extract and validate, from sets of noisy inputs, parameters for the automatic development of accurate GIF models, and

2. experimentally define complementary biophysical mechanisms to be grafted onto the

GIF models to iteratively improve their prediction accuracy (augmented GIFs).

This approach recapitulated and extended past findings on DRN neurons by showing that the best-performing models of 5-HT neurons featured slow membrane time constants, an A-type potassium current, and strong adaptation mechanisms. Network simulations of optimized GIF models of both 5-HT and GABAergic SOM neurons organized in a feed-forward inhibitory circuit revealed that 5-HT neuron populations context-dependently encode a mixture of the intensity and temporal derivative of their inputs. Our overall approach further allowed us to trace back specific features of these population responses (e.g., gain) to defined excitability features of DRN neurons.

2.2 Results

2.2.1 Salient electrophysiological features of DRN neurons

Our main goal was to develop an experimentally-grounded model of the DRN to better understand its computational properties. As a first step towards this goal, we carried out experiments to constrain a set of single-neuron models of the two main cell types found in the DRN: 5-HT and SOM GABA neurons. We performed whole-cell electrophysiological recordings from genetically-identified 5-HT (Fig. 2.1A1; SERT-Cre::Rosa-TdTomato mice) and SOM (Fig. 2.1A2; Table 1; SOM-Cre::Rosa-TdTomato mice) neurons in slices. In keeping with previous descriptions (e.g., [26, 27]), in the majority of the 5-HT neurons recorded in our dataset, current steps induced strongly adapting action potential firing accompanied by large afterhyperpolarization potentials (AHP), and a characteristic kink in the voltage trace leading up to the first spike (Fig. 2.1B). Qualitatively distinct firing patterns of 5-HT neurons were, however, occasionally observed (Fig. 2.8). Recordings from SOM neurons revealed spiking patterns that were more heterogeneous (Fig. 2.9). Comparing the relationship between the injected currents and firing frequencies between these populations, we found that SOM neurons were generally more sensitive to changes in input current (gain) than 5-HT neurons

Table 2.1: Membrane parameters of DRN neurons. Parameters obtained from recordings from PFC L5 pyramidal neurons used to fit GIF models as a point of comparison are also shown. Data are presented as mean ± SD. Distributions are shown in Fig. 2.10.

Cell type	Resistance R	Capacitance C	Time constant τ	N
5-HT	$1.16 \pm 0.55\,\mathrm{G\Omega}$	$67.0 \pm 17.1\,\mathrm{pF}$	$75.2 \pm 33.8\,\mathrm{ms}$	96
SOM	$1.07 \pm 0.58\,\mathrm{G\Omega}$	$43.5 \pm 15.5\,\mathrm{pF}$	$42.2 \pm 19.8\,\mathrm{ms}$	28
mPFC	$0.188 \pm 0.130\,\mathrm{G\Omega}$	$160.6 \pm 48.2\,\mathrm{pF}$	$27.4 \pm 13.2\,\mathrm{ms}$	25

and responded to weaker inputs (Fig. 2.1B). The gain showed greater variability in SOM neurons than in 5-HT neurons (Brown-Forsythe equality of variance test $p = 0.001$ on $N = 17$ 5-HT and $N = 7$ SOM neurons). In line with this observation, SOM neurons also consistently exhibited a wider range of firing frequencies for a given input (e.g., for a 50 pA input 5-HT neurons fired at 2.81 ± 2.22 Hz vs 8.16 ± 5.70 Hz; Brown-Forsythe test $p = 0.005$ in $N = 17$ 5-HT neurons and $N = 14$ SOM cells). Together, these observations outlined three salient cellular-level features of DRN neurons, namely the strong AHP and voltage kink of 5-HT neurons as well as noticeable heterogeneous excitability of SOM neurons.

The characteristic kink in the voltage leading up to the first spike in 5-HT neurons in principle may be caused by near-threshold activation of voltage-gated potassium channels (VGKCs; [193–195]). We therefore examined whole-cell currents evoked by voltage steps (from -90 mV to -20 mV) in both 5-HT and SOM neurons to look for evidence of such a VGKC. In 5-HT cells, these experiments revealed a large (peak amplitude 928 ± 249 pA, leak-subtracted), partly-inactivating (steady-state amplitude 142 ± 45 pA, leak-subtracted) outward current (Fig. 2.1C1) that was sensitive to Kv4-selective potassium channel blockers (Fig. 2.11). This current activated rapidly (peak latency 7.46 ± 0.21 ms) and inactivated over tens of milliseconds (inactivation time constant $\tau_h = 42.9 \pm 9.4$ ms; kinetics are similar at near-physiological temperature, see Fig. 2.12). The gating and kinetic profile (Table 2, Figs. 2.13 and 2.14) of the inactivating component of this conductance in 5-HT neurons are broadly similar to those expected of the A-type potassium currents (I_A) characterized in great detail in several other cell types (e.g., [196]). Because these parameters are sufficient to construct a model of this conductance (see below), we have not attempted to determine its

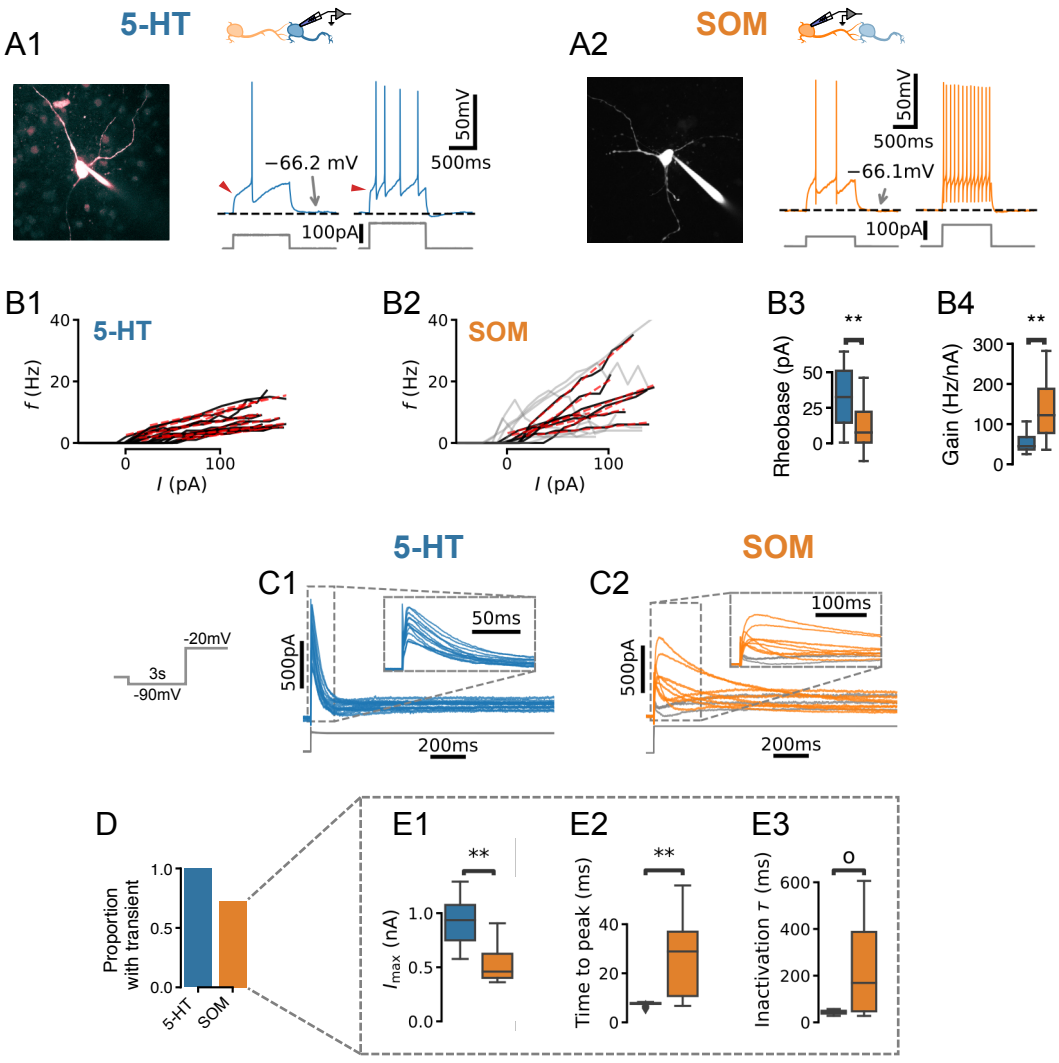


Figure 2.1: Physiology of DRN neurons. **A**, **B** Morphology, current steps (A), and spike frequency vs. input (f/I) curves (B) of genetically-identified DRN neurons. Non-monotonically-increasing f/I curves (gray) and linear fits to monotonically-increasing curves (red dashed) for $N = 17$ 5-HT (B1) and $N = 14$ SOM (B2) neurons. (B3) Rheobase of 33.8 ± 21.0 pA in 5-HT neurons vs. 11.3 ± 16.0 pA in SOM neurons. (B4) Gain of 52.2 ± 22.2 Hz nA^{-1} in 5-HT neurons vs. 87.2 ± 33.0 Hz nA^{-1} in $N = 7$ SOM neurons with monotonically-increasing f/I curves. **C** Leak-subtracted whole-cell currents evoked by a depolarizing step. Each trace is one cell; $N = 13$ 5-HT and $N = 11$ SOM cells. Traces without a transient outward current are shown in gray. **D** Proportion of neurons with a transient outward current by cell type. **E** Quantification of transient outward currents in each cell type. $N = 3$ SOM cells without a transient outward current were excluded from analysis, leaving $N = 13$ 5-HT and $N = 8$ SOM neurons. Annotations reflect Mann-Whitney U-tests. Non-parametric Brown-Forsythe equality of variance tests indicated significantly more variable time to peak ($p = 1.11 \times 10^{-4}$; E2) and inactivation time constant ($p = 1.97 \times 10^{-4}$; E3) in SOM cells.

Table 2.2: 5-HT I_A current gating parameters. Gating curves shown in Fig. 2.13B were fitted with the scaled Boltzmann function $g_\infty(V)/g_\infty(V_{\text{reference}}) = x_\infty = A_x/\left(1+\exp\left[-k_x(V-V_x^*)\right]\right)$. Values are based on experiments from N=13 cells.

Gate	$V_{\text{reference}}$	A_x	k_x	V_x^*
m_∞	$-20\,\text{mV}$	1.61	$0.0985\,\text{mV}^{-1}$	$-23.7\,\text{mV}$
h_∞	$-80\,\text{mV}$	1.03	$-0.165\,\text{mV}^{-1}$	$-59.2\,\text{mV}$
n_∞	$-20\,\text{mV}$	1.55	$0.216\,\text{mV}^{-1}$	$-24.3\,\text{mV}$

molecular identity further. For the sake of simplicity, we refer to the inactivating component herein as I_A (in keeping with the previous literature; see [155, 197]) and the steady-state component as I_K. Thus, an I_A-like inactivating VGKC is a consistent feature of DRN 5-HT neurons.

The same voltage-clamp protocol applied to SOM neurons, in contrast, triggered a mixture of outward and inward currents that varied widely from cell to cell (Fig. 2.1D2). A significant proportion of SOM neurons did not express a transient outward current at all (27.3 %, Fig. 2.1E), while the remaining cells had currents that were significantly smaller ($p = 0.003$), activated more slowly ($p = 0.003$), and exhibited much more heterogeneous kinetic profiles than those found in 5-HT neurons (Fig. 2.1E3). Together, these results show that the expression of this subthreshold voltage-gated current is substantially more variable in SOM neurons than in 5-HT neurons, in line with the distinctive heterogeneity of excitability features observed in this DRN cell type (Fig. 2.1C2 and Fig. 2.9).

2.2.2 I_A regulates initial firing rate via control of spike time jitter

To develop an intuition for how I_A impacts the firing patterns of 5-HT populations, we first created a toy model that captured the effect of this conductance on single cell voltage dynamics (see Methods). In keeping with previous studies, I_A introduced a kink in the subthreshold voltage leading up to spike threshold ([198–200], Fig. 2.2A) and increased the latency to the first spike evoked by a square step stimulus, particularly when starting from a hyperpolarized voltage at which I_A is free from inactivation (Fig. 2.2B, C). The effect of I_A on spike latency

depends at least to some extent on its effective magnitude and inactivation kinetics (defined as the ratio of maximal A-type conductance to inverse membrane resistance and the ratio of the inactivation time constant of I_A to the membrane time constant; see Methods). When we set the corresponding parameters in our toy model to experimentally-determined values from 5-HT neurons, we observed the same qualitative relationship between spike latency and initial voltage (Fig. 2.15), further pointing towards a functional effect of I_A in this cell type. The predicted relationship between initial-voltage and latency was experimentally recapitulated in whole-cell recordings from identified 5-HT neurons (Fig. 2.2D-F). In particular, the onset of spiking was delayed by hyperpolarization (Fig. 2.2D), and the magnitude of this effect was significantly reduced by the partial pharmacological block of I_A with 4-AP ($p = 3.2 \times 10^{-6}$ for initial voltages of $-90\,\text{mV}$ to $-70\,\text{mV}$; Fig. 2.2E). Finally, we also observed an inflection point predicted by the model in the normalized initial-voltage/latency relationship (Fig. 2.2F, compare with model prediction in Fig. 2.2C). In summary, our toy model captured the expected effects of I_A in single 5-HT cells.

Next, we used our experimentally-validated toy model to understand how I_A impacts the spiking responses of whole neuronal populations. To do this, we simulated the effect of a shared step input to a population of 600 toy neurons each receiving independent background noise (corresponding to naturalistic fluctuations in synaptic inputs). Whereas subthreshold fluctuations yielded time-locked spikes without I_A (Fig. 2.2G), they induced spiking with larger jitter across the simulated population when I_A was present (Fig. 2.2H). This desynchronizing effect of I_A also decreased the peak population rate at the time corresponding to the mean latency (Fig. 2.2I) since the peak rate corresponds to the coincidence rate from an ensemble of cells with similar properties. (The same effects were also observed in the toy models with parameters constrained to experimentally-determined values; Fig. 2.15.) Taken together, results from these toy-models revealed a role of I_A in regulating the degree of synchronization of a population following sudden inputs, suggesting that I_A may regulate the gain of the DRN network to time-varying inputs. This intuition gleaned from this toy model is examined

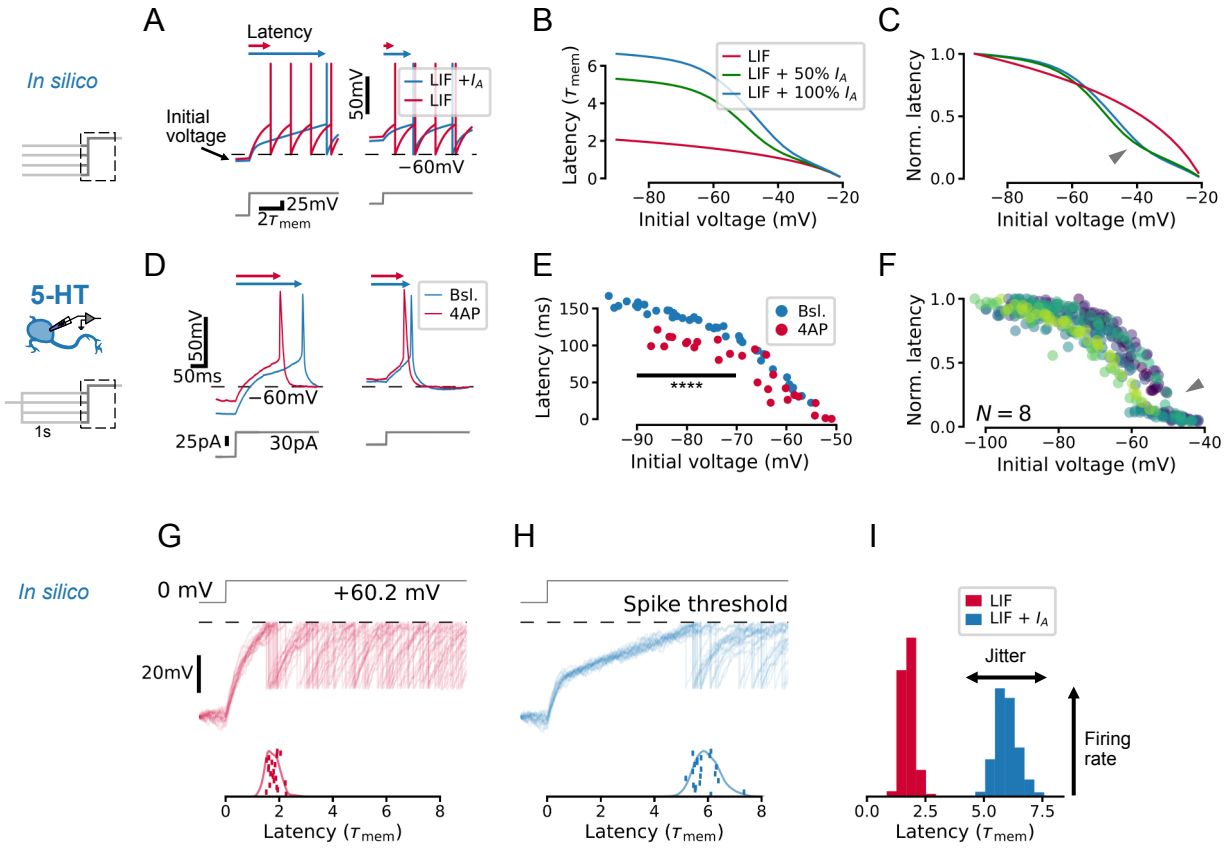


Figure 2.2: I_A qualitatively alters the relationship between initial voltage and spike timing. **A**–**C** A toy model neuron with I_A predicts a non-linear effect of voltage history on spike timing in a simple experiment. **D**–**F** Experiments in 5-HT neurons fulfill predictions of the toy model. (F) Latency curves for $N = 8$ 5-HT neurons, normalized to the maximum latency for each cell. Each colour is one cell. (G–I) I_A causes an increase in spike latency and jitter in the presence of noise. Models and input are the same as in A–C. Spike latency histograms for populations of 600 toy neurons are shown in I. The width of the histogram reflects jitter in the timing of the first spike, while the height of the histogram approximates the peak instantaneous firing rate. Note that as jitter increases, the height of the histogram decreases. The toy model with 100 % I_A has an effective I_A conductance $\bar{g}'_A/g_l = 10$ and effective inactivation time constant $\tau_h/\tau_{\text{mem}} = 1.2$.

in more detail with optimized GIF models (see below).

2.2.3 Extensions to GIF models are required to capture the excitability of DRN neurons

We next sought to develop a model able to capture the essential biophysical features of DRN neurons and accurately predict their responses to naturalistic inputs. GIF models offer a flexible modeling framework well-suited to this purpose because they can be trained to accurately reproduce the firing patterns of individual neurons using less than five minutes of electrophysiological data per neuron [151, 152, 201–203]. In this framework, individual neurons are described in terms of three core components:

1. a passive membrane filter, κ, which transforms input currents into a subthreshold membrane potential,

2. a stochastic spiking process, which transforms the subthreshold membrane potential into action potentials, and

3. two adaptation mechanisms, namely a spike-triggered current mediating the commonly observed AHP, η, and change in firing threshold, γ (Fig. 2.3A1, see Methods).

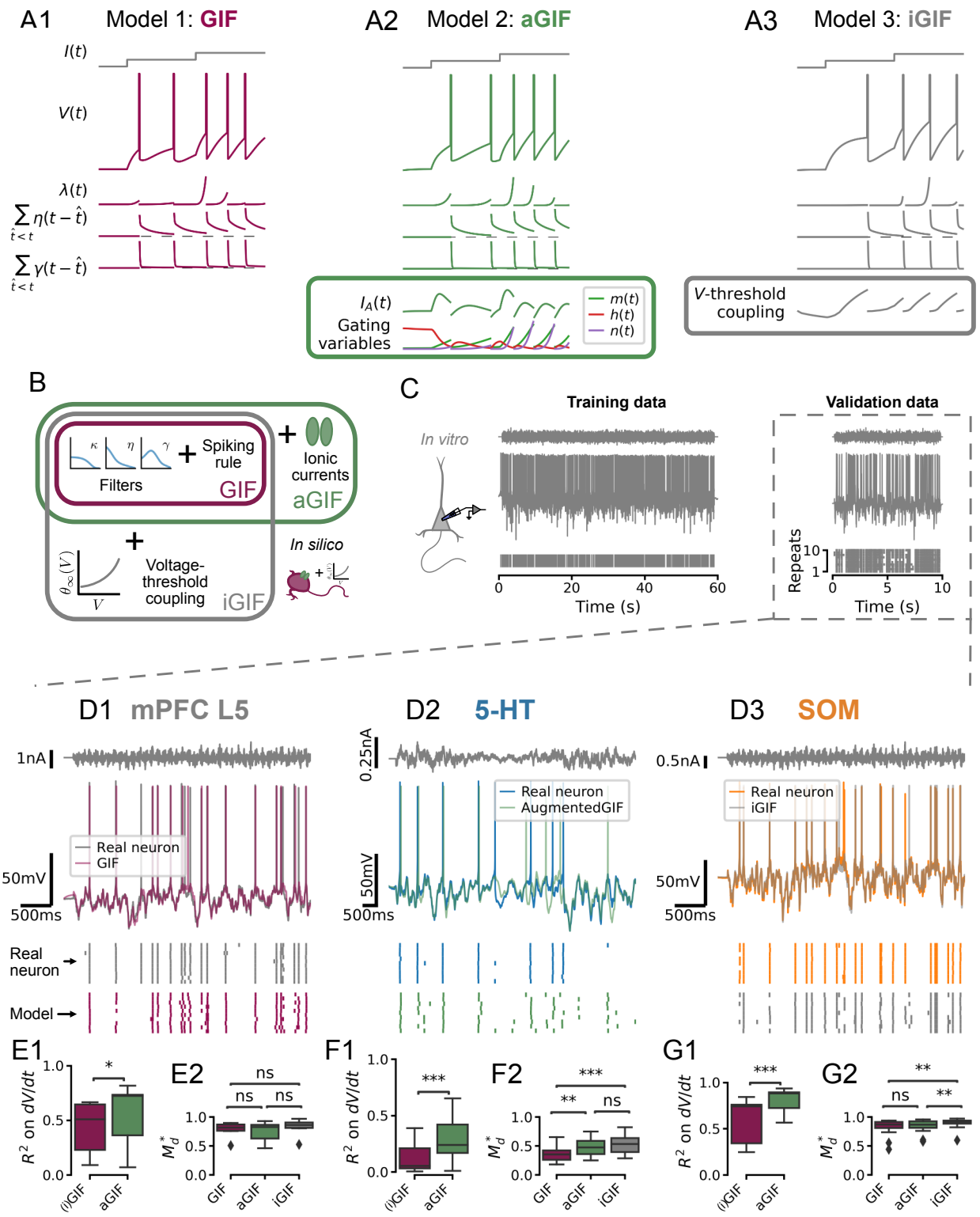
These components are described by parameters the values of which are inferred from the electrophysiological data using a combination of least-squares multi-linear regression and gradient ascent of a likelihood function. The flexibility and data-efficiency of this framework lends itself well to capturing the functional properties of single neurons and, by extension, heterogeneous neural populations.

Our results outlined in Fig. 2.2 show that I_A regulates spike timing in 5-HT neurons because of its nonlinear subthreshold effects. Foreseeing that the presence of this prominent current may limit the accuracy of canonical GIF models — which are not designed to capture nonlinear subthreshold effects — we first *augmented* the canonical GIF model (aGIF; Fig. 2.3A2) with

a simplified Hodgkin-Huxley-type model of the subthreshold voltage-dependent currents we recorded in 5-HT neurons (see Methods). To assess whether incorporating additional biophysical details into the aGIF model might further improve its predictive performance, we turned to the previously-described sodium channel-inactivation GIF model (iGIF; Fig. 2.3A3), which extends the GIF model of Mensi, Naud *et al.* [151] by adding a non-parametric voltage coupling function to the dynamic spike threshold ([204], see Methods). Although this GIF model extension was initially conceived specifically to capture the influence of subthreshold sodium channel inactivation on firing threshold (hence its name), the non-parametric definition of the threshold coupling function gives it the capacity to account for a wide range of other subthreshold biophysical mechanisms which regulate spiking, notably including, but not limited to, I_A. Comparing the performance of the more parsimonious aGIF model to that of the iGIF model enabled us to assess whether accounting for additional mechanisms that regulate spiking beyond I_A might further improve our DRN neuron models.

To establish comparative GIF model benchmarks across cell types, we carried out whole-cell electrophysiological recordings not only from DRN 5-HT and SOM cells, but also from canonical deep-layer pyramidal neurons of the medial prefrontal cortex (mPFC). For each recording, we applied two distinct instantiations of noisy in vivo-like inputs (see Supplemental Methods, Fig. 2.16), one of which was used to determine the model parameters while the other was reserved for post hoc evaluation of the models' accuracy (i.e., "training" data and "validation" data, respectively; see Fig. 2.3C). Accuracy was assessed by comparing models with recorded data across cell types in terms of (Fig. 2.3D): 1) subthreshold voltage changes on training data, R^2; and, 2) spike timing on validation data, M_d^* (where $M_d^* = 1$ is the best possible performance and $M_d^* = 0$ is chance-level; see Methods).

The canonical GIF model predicted both the subthreshold dynamics and spike timing of mPFC pyramidal neurons with high accuracy ($R^2 = 0.431 \pm 0.249$; $M_d^* = 0.783 \pm 0.134$; Fig. 2.3E), consistent with previous reports on cortical pyramidal neurons [151, 152, 203, 204]. While our aGIF model slightly better predicted the voltage of mPFC neurons ($R^2 = 0.544 \pm$


A1
Model 1: GIF
I(t)
V(t)
λ(t)
Σ η(t − t̂)
Σ γ(t − t̂)
A2
Model 2: aGIF
I_A(t)
Gating variables
m(t)
h(t)
n(t)
A3
Model 3: iGIF
V-threshold coupling
B
Filters
Spiking rule
GIF
Ionic currents
aGIF
Voltage-threshold coupling
iGIF
In silico
C
In vitro
Training data
Validation data
Repeats
Time (s)
D1 mPFC L5
D2 5-HT
D3 SOM
1nA
0.25nA
0.5nA
50mV
500ms
Real neuron
GIF
AugmentedGIF
iGIF
Real neuron
Model
E1
E2
F1
F2
G1
G2
R² on dV/dt
M*d
(i)GIF
aGIF
GIF
iGIF
ns
*
**
***

0.280, $p = 0.028$, Fig. 2.3E1), this did not translate into more accurate spike predictions ($M_d^* = 0.743 \pm 0.180$, $p = 0.710$, Fig. 2.3E2), consistent with the observation that I_A is not a significant conductance recorded from the cell body of mPFC pyramidal neurons (Fig. 2.17 and see [205, 206]). On the basis of spike timing prediction, the canonical GIF model thus offered the most parsimonious account of the behaviour of mPFC neurons.

With this point of comparison established, we next quantified the performance of each of our candidate GIF models (GIF, aGIF, and iGIF) in 5-HT neurons. As previously intuited, the canonical GIF model performed rather poorly in 5-HT neurons (Fig. 2.3F), predicting <15 % of the variance of the subthreshold voltage ($R^2 = 0.128 \pm 0.135$) and achieving M_d^* scores less than half of those observed in mPFC neurons ($M_d^* = 0.352 \pm 0.118$). This indicates that the passive membrane filter and adaptation mechanisms included in the canonical GIF model were insufficient to capture the behaviour of 5-HT neurons. By augmenting the GIF model with our experimentally-constrained model of I_A, the aGIF model not only better predicted the voltage ($R^2 = 0.301 \pm 0.200$, $p = 1.96 \times 10^{-4}$; Fig. 2.3F1) but also the spike timing ($M_d^* = 0.481 \pm 0.148$, $p = 0.001$; Fig. 2.3F2) of 5-HT neurons. While the more general iGIF model exhibited a similar improvement in spike timing predictions over the GIF model ($M_d^* = 0.536 \pm 0.154$, $p = 5.89 \times 10^{-4}$), it did not significantly outperform the aGIF model ($p = 0.644$; Fig. 2.3F2), suggesting that accounting for additional biophysical mechanisms

Figure 2.3: (Previous page.) Single neuron models accurately predict the subthreshold voltage and spike times of DRN neurons. **A** Components of candidate single neuron models: λ intensity of stochastic spike-generating process; η spike-triggered current (positive values indicate a hyperpolarizing current); γ spike-triggered threshold movement. **B** GIF model extensions. **C** Representative experiment used to train and validate neuron models. Training set consists of repetitions of 60 s of frozen Ornstein-Uhlenbeck (OU) noise and the validation set consists of repetitions of a different frozen OU noise stimulus lasting 10 s (only one repetition is shown). **D** Representative validation data and model predictions for each cell type. OU noise input current (top), recorded and predicted voltage traces (middle), and recorded and predicted spike times across all repetitions of the validation stimulus. Stimulus parameters were adjusted for each cell type, see Fig. 2.16. **E**–**G** Quantification of model performance in terms of R^2 on the training subthreshold voltage derivative and on the validation spike-train similarity metric M_d^*. GIF and iGIF models have the same subthreshold performance because the subthreshold components of these models are identical (see Methods). Benchmarks are for models fitted to $N = 18$ 5-HT, $N = 14$ SOM, and $N = 7$ mPFC neurons.

that regulate spiking beyond those included in the aGIF model would be unlikely to further improve performance. Repeating this process using data collected closer to physiological temperature yielded the same result (Fig. 2.18). Thus, among the models considered, adding I_A to the subthreshold and spiking mechanisms of the GIF model best accounts for the biophysical mechanisms responsible for shaping the responses of 5-HT neurons to in vivo-like inputs.

Turning to the other main cell type of the DRN, we next analyzed the performance of each model in SOM cells (Fig. 2.3G). In these cells, the canonical GIF model produced highly accurate predictions ($R^2 = 0.600 \pm 0.238$ and $M_d^* = 0.818 \pm 0.149$), consistent with its high performance previously reported for cortical GABAergic neurons [151, 152]. Nonetheless, the iGIF achieved small but significant performance gains ($M_d^* = 0.892 \pm 0.094$, $p = 0.004$ vs. aGIF and $p = 0.003$ vs. GIF; Fig. 2.3G2), leading us to select it as our model of SOM neurons.

2.2.4 Multiple adaptation mechanisms in 5-HT neurons

Our model selection approach identified the most salient components required to capture the input-output functions of individual neurons and allowed us to identify functional differences across cell types. 5-HT neurons were distinguishable from SOM and mPFC cells by their long membrane time constants (Fig. 2.4A and Fig. 2.19A) and by the presence of conspicuously potent and protracted adaptation mechanisms (Fig. 2.4B-D). Indeed, in addition to evoking a characteristically large and prolonged adaptation current (Fig. 2.4C), action potential firing in 5-HT neurons produced a substantial and long-lasting increase in firing threshold (Fig. 2.4D; but note that this effect is somewhat attenuated near physiological temperature, Fig. 2.20). In contrast, SOM neurons most often displayed either negligible or even depolarizing spike-triggered currents (Fig. 2.4B,C) that may underlie the burst firing patterns often observed in this cell type (Fig. 2.9). These observations derived from the parameters of GIF models are not only consistent with our experimental characterization (Figs. 2.1 to 2.3), but significantly

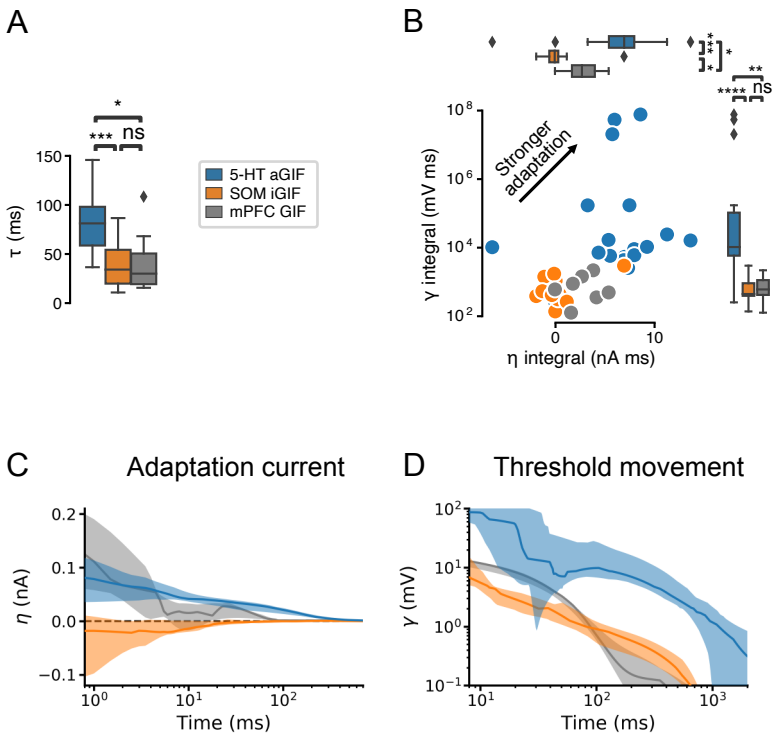


Figure 2.4: 5-HT neurons are distinguished by slow membrane time constants and potent adaptation. **A** Using the features from the best-performing GIF model variant for each cell type (legend), passive membrane time constant. **B** Spike adaptation features: potency of AHP-mediated (spike-triggered current η integral) and AHP-independent (spike-triggered threshold movement γ integral) adaptation. **C**, **D** Comparison of model filters. Presented as median (lines) and interquartile range (bands). Note the long-lasting adaptation currents (C; positive values indicate hyperpolarizing current) and threshold movements (D) of 5-HT neurons. Parameters are from models fitted to $N = 18$ 5-HT, $N = 14$ SOM, and $N = 7$ mPFC neurons.

expand it. Thus, 5-HT neurons are characterized by slow membrane dynamics, I_A, and particularly prominent adaptation mechanisms.

2.2.5 Preferential sensitivity of 5-HT neuron populations to the onset of sudden inputs

The development and validation of accurate single-cell models allowed us to identify the population-level computations operating in the DRN. We took advantage of the one-to-one correspondence between our GIF models and real neurons to construct synthetic populations with realistic neuron-to-neuron heterogeneity by sampling from banks of single-cell models (Fig. 2.5A). In response to step increases of synaptic-like inputs delivered to the entire population (Fig. 2.5B left), the population firing rates (in Hz/neuron; Fig. 2.5B right) of 5-HT, SOM, and mPFC neurons (Fig. 2.5C) transiently increased before relaxing to a significantly lower stationary level. Strong inputs did not produce oscillations in the population firing rates, likely because of population heterogeneity (Fig. 2.23; [124, 207, 208]). The transient and stationary parts of the population input-output functions were approximately rectified linear functions (Fig. 2.23) which we summarized and plotted as the time-varying slope (i.e., gain; Fig. 2.5D). While the gain of the transient response was greater than that of the stationary response in all three cell types, the ratio of transient to stationary gain was substantially higher in 5-HT neurons (Fig. 2.5E; ratio of 3.42 ± 0.07, vs. 1.89 ± 0.04 in SOM and 1.50 ± 0.03 in mPFC; $p < 0.001$ in each case; but note that the gain ratio in 5-HT neurons falls to 2.13 ± 0.03 near physiological temperature [Fig. 2.25], consistent with a smaller spike-triggered threshold movement [Fig. 2.20]). This marked response of 5-HT cells occurred quickly, in the first 100 ms after the onset of the step. Thus, despite 5-HT neurons being characterized by slow membrane time constants, their population activity provided a remarkably strong encoding of the onset of step synaptic inputs.

We next considered the underlying mechanisms giving rise to the distinctive time-dependent gain of 5-HT neurons. We found that the characteristically strong spike-triggered adaptation

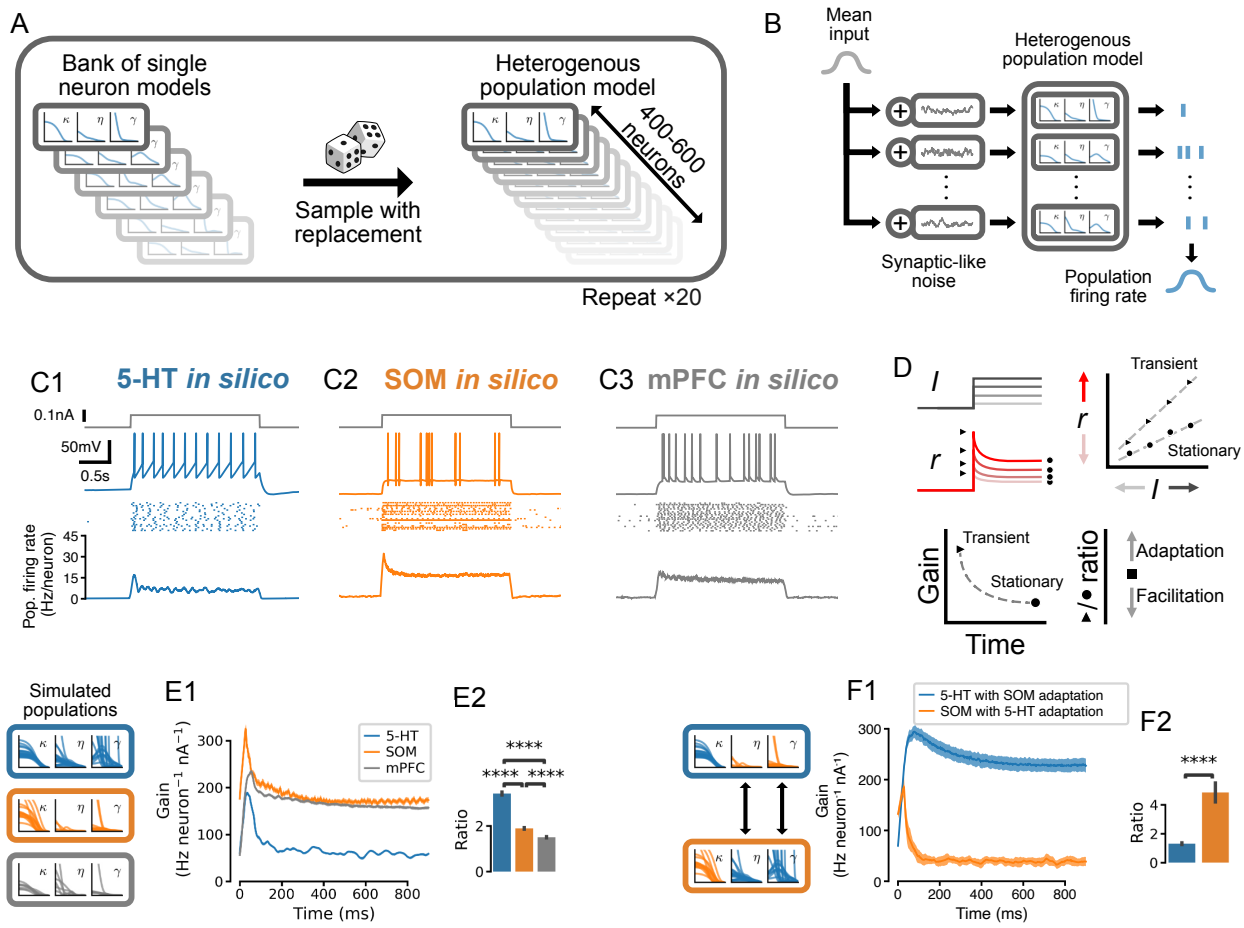


Figure 2.5: Adaptation mechanisms cause a higher gain of the transient vs stationary population response. **A** Generation of heterogeneous population models from experimentally-constrained single neuron models. **B** Schematic of population simulations. Spikes from individual neuron models in the simulated population are added together to produce a population firing rate. **C** Population responses to input step. From top in each column: stimulus (gray); sample voltage trace; spike raster of first 20 neurons; mean population firing rate across 20 independent simulations. **D** Schematic for quantifying the time-varying population input-output function for both the transient and the stationary components of the response. An input-output function is calculated for the population response at each time point after the input step. The slope of each input-output function (gain) is then plotted as a function of time since the step onset. The ratio of the maximum gain to the minimum gain is a measure of the relative amount of population adaptation. **E** Time-resolved gain of step input responses across cell types following the approach shown in D. **F** Time-resolved gain of 5-HT populations with the adaptation parameters of SOM neurons (blue) and of SOM populations with adaptation parameters of 5-HT neurons (orange). Data are presented as mean ± SD in E1 and F1.

of 5-HT neurons (spike triggered hyperpolarizing adaptation current and threshold movement shown in Fig. 2.4) contributed to the observed relaxation of the population response to a lower stationary level: grafting the weak adaptation from SOM neuron models onto 5-HT models dramatically reduced the ratio of transient to stationary gain, and vice-versa (Fig. 2.5F). These findings are consistent with previous models in other cell types showing that spike-triggered adaptation reduces the sensitivity of neural populations to input changes over long timescales ([124, 125, 209]). Therefore the preferential sensitivity of 5-HT neuron populations to sudden changes in synaptic inputs is a natural consequence of strong adaptation at the single neuron level.

2.2.6 Feedforward inhibition and I_A control 5-HT output gain of the DRN

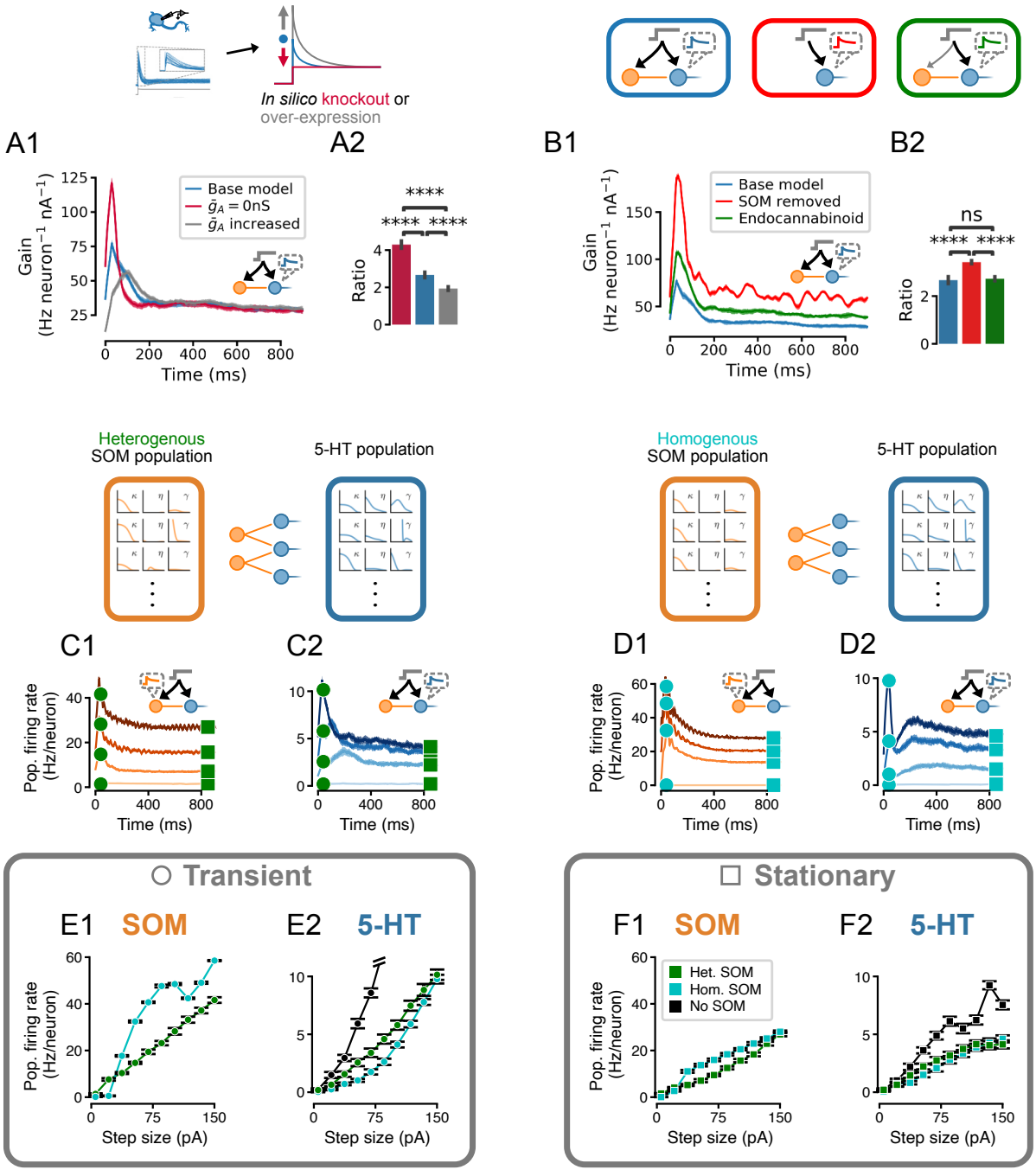
Apart from the strong adaptation mechanisms of 5-HT neurons, two other mechanisms have the potential to dynamically modulate the 5-HT output from the DRN: I_A in 5-HT neurons and the feed-forward inhibition (FFI) enacted by local DRN interneurons [19, 28]. To examine the contributions of these two mechanisms, we first connected our existing SOM population models to 5-HT population models using experimentally-constrained $GABA_A$ receptor-mediated synaptic conductances (see Methods and Fig. 2.24).

To dissect the contribution of I_A in shaping population responses in this connected DRN network, we applied the same inputs to both 5-HT and SOM neuron populations and examined 5-HT neuron population dynamics (as in Fig. 2.5) while varying the maximal conductance of I_A (in 5-HT neurons). The gain of the transient component of the 5-HT response increased markedly when the conductance of I_A was set to zero (Fig. 2.6A), while increasing the potency of I_A substantially dampened and broadened the population response to fast inputs, reminiscent of I_A's modulation of spike timing jitter observed in our toy model (Fig. 2.2I–K). These simulations thus show that I_A substantially regulates the gain of the transient component of DRN 5-HT output evoked by sustained inputs, with negligible effects

on the gain of the slower stationary component.

Previous work has shown that glutamatergic excitatory inputs from the PFC make strong mono-synaptic contacts onto both DRN 5-HT and GABAergic neurons, triggering a classic FFI. Intriguingly, the PFC axonal inputs onto these two cellular elements of the DRN are functionally distinct inasmuch as the PFC synapses onto GABAergic neurons are far more sensitive to endocannabinoid neuromodulation than those onto 5-HT neurons [28]. The computational role of this differential sensitivity to neuromodulation is currently unknown. We began by determining the role of the DRN FFI *per se* by comparing the responses of 5-HT neuron population dynamics with or without SOM cells (Fig. 2.6B). Including FFI onto 5-HT neurons substantially dampened the overall response of the 5-HT population to synaptic inputs, while still sustaining the preferential encoding of the early phase of sudden inputs (Fig. 2.6B2). While introducing FFI did decrease the gain ratio, this decrease was quantitatively smaller than the differences between 5-HT neurons and other cell types shown in Fig. 2.5E and the effect of changing I_A shown in Fig. 2.6A, Fig. 2.6B2. We next directly simulated the effects of endocannabinoid modulation of excitatory input to the DRN observed experimentally [28] by weakening the strength of the input to SOM neuron populations by 30 % while leaving that to 5-HT neurons intact. By favoring the direct monosynaptic excitation of 5-HT neurons by preferentially diminishing the glutamatergic drive of SOM neurons, this neuromodulation led to an increase in the overall gain of the DRN that was unexpectedly apparent across the entire duration of the response to step inputs (i.e., no change in the gain ratio, Fig. 2.6B2). Thus, the target-specific endocannabinoid-mediated modulation of PFC excitatory drive in DRN exerts a normalizing role by increasing the overall gain of 5-HT output evoked by synaptic inputs without altering its preferential encoding of changes in input, which is emerging as a cardinal feature of DRN network dynamics.

Our electrophysiological recordings showed that excitability heterogeneity is a salient feature of the SOM DRN neuron population. Our modeling approach allows us to specifically examine the role of this cellular heterogeneity in shaping the output of the DRN by comparing


In silico knockout or
over-expression
A1
A2
B1
B2
Base model
$\tilde{g}_A = 0$nS
$\tilde{g}_A$ increased
Gain
(Hz neuron^{-1} nA^{-1})
Time (ms)
Ratio
****
ns
Base model
SOM removed
Endocannabinoid
Heterogenous
SOM population
5-HT population
Homogenous
SOM population
5-HT population
C1
C2
D1
D2
Pop. firing rate
(Hz/neuron)
Transient
Stationary
E1 SOM
E2 5-HT
F1 SOM
F2 5-HT
Het. SOM
Hom. SOM
No SOM
Step size (pA)

our DRN model (Fig. 2.6C) to an alternative homogenized version in which the parameters of SOM neurons were set to fixed values (Fig. 2.6D). Thus, while FFI with an experimentally-determined degree of heterogeneity mainly imposed a reduction of the slope of the input-output function (i.e., divisive inhibition), homogeneous FFI mainly shifted the input-output function of the transient component of the population response to the right (i.e., subtractive inhibition; Fig. 2.6E). This subtractive feature can be traced back to a strong non-linearity in the input-output functions of homogenized SOM neuron populations (compare Fig. 2.6E1 and Fig. 2.6F1). In the case of the stationary component, both heterogeneous and homogenized DRN models implemented divisive inhibition (Fig. 2.6F). Therefore, we conclude that heterogeneity among GABAergic neurons implements divisive inhibition.

2.2.7 5-HT neurons linearly encode the temporal derivative of inputs to the DRN

Adaptation plays a critical role in implementing temporal derivative encoding in sensory systems [126, 192], but has not been ascribed a similar role in neuromodulatory systems such as the DRN. To determine whether the DRN also supports this computation, we

Figure 2.6: (Previous page.) Effect of I_A density, feedforward inhibition, and heterogeneity of SOM neurons on the 5-HT neuron population response. Network input is the same set of step stimuli as in Fig. 2.5D-F. **A** Increasing I_A reduces adaptation by selectively suppressing the early part of the response to sudden inputs, and vice-versa. **B** Gain curves with normal feedforward inhibition (blue), with reduced input strength onto the inhibitory population (green), or without inhibition (red). Reduced input strength onto the inhibitory population (green) simulates the effect of endocannabinoid input (Geddes et al., 2016). (A2,B2) Ratio of peak to steady-state gain. **C** Population firing rates of SOM and 5-HT neurons in a network in which both populations are heterogenous. **D** Population firing rates of SOM and 5-HT neurons in a network in which all SOM neurons are identical. Effects of homogeneous (cyan) or heterogeneous (green) SOM populations on the population input-output functions for the transient **E** and stationary **F** components of the response (see square and circle markers in C and D). Note that the input-output function of the heterogenous SOM population is approximately linear, whereas that of the homogenous population is not (E1, F1). Relative to the input-output functions of a 5-HT population receiving no feed-forward inhibition, the effect of the heterogenous SOM population is divisive, but the effect of the homogenous SOM population on the transient part of the 5-HT population input-output function includes a strong subtractive component (E2).

parameterized the rate of change of DRN inputs by applying ramp stimuli with variable slopes (i.e., derivatives; Fig. 2.7A and B). Remarkably, the peak 5-HT neuron population firing rate linearly reported the slope of the ramps, an effect which was enhanced by FFI (Fig. 2.7C). We further found that this linearity was conditional on the presence of slightly depolarizing background input ($\geq$20 pA, Fig. 2.7D). Simulations using aGIF models fitted to data collected near physiological temperature yielded similar results; Fig. 2.26. The potent adaptation mechanisms of 5-HT neurons play a key role in mediating this linear encoding of input derivative, since reducing the strength of adaptation reduces linearity across a wide range of input baselines (Fig. 2.7E and F). Together, these observations suggest that the DRN signals to its brain-wide targets a mixture of the intensity and temporal derivative of its excitatory inputs, and that the derivative-encoding component dominates when the input is increasing rapidly (Fig. 2.27).

The extent to which the output of the DRN signals the temporal derivative of its input is likely to be limited by several factors, notably: the long membrane time constants of 5-HT neurons (Table 1, Fig. 2.19A), which cause rapidly fluctuating to be filtered out; the fact that firing rates cannot be less than zero, limiting the dynamic range available to encode negative input derivatives; the presence of I_A, which filters out inputs with a high temporal derivative (Fig. 2.2); and the level of background input (Fig. 2.7D). Because I_A can be partly inactivated by depolarizing background input, the effects of background input and I_A on the derivative-encoding properties of the DRN are expected to interact. Consistent with this idea, removing I_A from 5-HT neurons in our DRN network models extended the range of background input where the peak 5-HT neuron population firing rate is an approximately linear function of the slope of a ramp stimulus (Fig. 2.7E and G). In summary, we found that the presence of strong spike-frequency adaptation in 5-HT neurons causes the DRN to signal the rate of change of its input to its brainwide targets, but that this core computation is progressively suppressed when a state of hyperpolarization engages I_A.

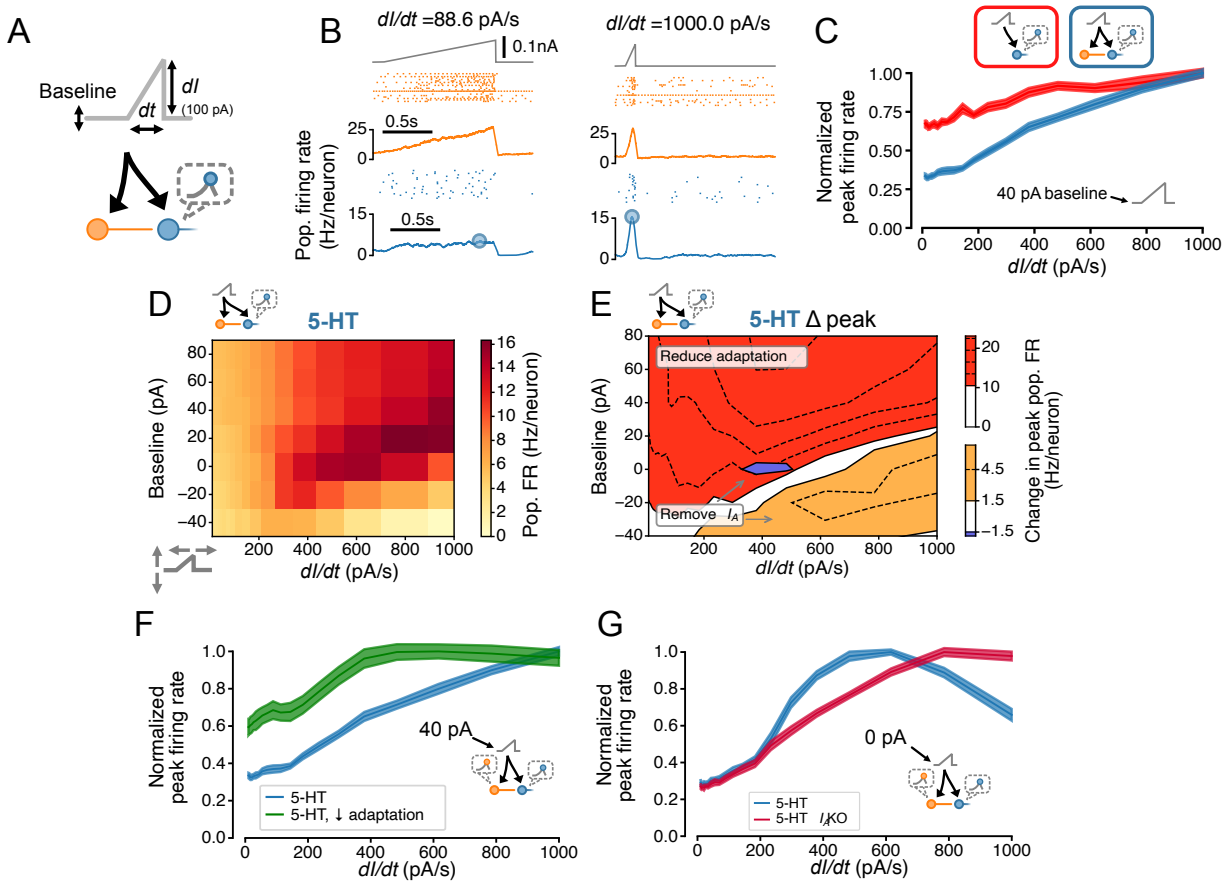


Figure 2.7: DRN 5-HT neuron population output conditionally encodes the temporal derivative of its input. **A** Design of simulations. A ramp stimulus with an adjustable baseline and slope (derivative) is applied to the same network models as in Figs. 2.5 and 2.6 and the peak firing rate of the 5-HT neuron population is extracted. **B** Representative simulated input (top), SOM neuron population activity (middle), and 5-HT neuron population output (bottom). **C** With a baseline input of 40 pA, peak 5-HT neuron population output is approximately linearly related to the derivative of the ramp input, and feed-forward inhibition by SOM neurons enhances this feature. **D** Peak firing rates (FRs) of 5-HT neuron populations depend on interacting effects of input baseline and slope. Panel C shows normalized data from the 40 pA row in blue. **E** 5-HT neuron adaptation and I_A dominate the DRN input-output function under different input regimes. Effect of reducing 5-HT neuron adaptation (following the approach from Fig. 2.5F) is most pronounced for higher levels of background input and more slowly-changing inputs (red), while the effect of removing I_A (following the approach of Fig. 2.6A) is most pronounced for low background input and fast-changing inputs (orange, blue). **F** Effect of reducing adaptation in 5-HT neuron models visualized at a 40 pA baseline. Note that 5-HT output no longer linearly encodes dI/dt when adaptation is reduced. **G** Effect of removing I_A from 5-HT neuron models visualized at a 0 pA baseline. Note that 5-HT output approximately linearly encodes dI/dt when I_A is removed.

2.3 Discussion

Here, we sought to characterize the computational properties of the DRN using a bottom-up approach grounded in experimentally-constrained models of the two most abundant cell types in this region: 5-HT and SOM GABA neurons. Consistent with, and extending, previous work, we found that 5-HT neurons were relatively homogeneous and characterized by potent spike-frequency adaptation (Fig. 2.1) and by the presence of a strong A-type potassium current (Fig. 2.2), while SOM neurons displayed a considerably more heterogeneous excitability profile (Fig. 2.1 and Fig. 2.9). Extensions to classical generalized integrate-and-fire (GIF) models [151, 203] to capture the non-linear subthreshold effects of I_A observed in 5-HT neurons were required to adequately capture the spiking response of 5-HT neurons to naturalistic stimuli (Fig. 2.3). This work introduces a new approach to capturing such non-linear subthreshold effects in the form of the aGIF model, which augments the GIF model of Mensi, Naud *et al.* [151] with experimentally-constrained Hodgkin-Huxley style currents, improving model interpretability without compromising predictive performance. Inspecting the parameters of the best-performing GIF models revealed that the substantial spike-frequency adaptation observed in 5-HT neurons is not fully explained by their distinctively large AHPs, and is partly mediated by a previously undescribed dynamic spike threshold (Fig. 2.4). This model-based approach allowed us to probe causal relationships between specific excitability features and population computations. Thus, we found that the prominent adaptation mechanisms in 5-HT neurons regulated DRN population responses to synaptic inputs (Fig. 2.5), that I_A suppressed the response to sudden inputs, and that heterogenous FFI had a divisive rather than subtractive effect on DRN output (Fig. 2.6). By further exploring DRN population dynamics, our simulations demonstrated that 5-HT neurons linearly reported a mixture of the intensity and temporal-derivative of their synaptic inputs (Fig. 2.7), and that the temporal-derivative dominates DRN output when the input is increasing rapidly (Fig. 2.27). In summary, this work points to a new computational role for the DRN in encoding the

derivative of its inputs, and identifies specific cellular and network mechanisms that give rise to this computation and modulate its expression. These results raise important questions about how the selective responses of the DRN to changing synaptic inputs might support its role in guiding animal's behavior in dynamic environments.

2.3.1 Need for a hybrid biophysical-simplified methodology

The computational and statistical modeling methodology presented here was designed to bridge the gap between specific biophysical mechanisms and network level computation. Closing this gap has also been the target of complex biophysical simulations, motivated by the hope to create tools for testing disease-related treatments and for untangling the computations performed by large neural networks [134, 210]. Preserving the accuracy and identifiability of simpler approaches [151, 152, 191, 192], the "augmented GIF" model developed here explicitly incorporates the most important biophysical features of 5-HT neurons, allowing us to probe their contributions to network-level computation by altering or removing the corresponding model components during network simulations. While the aGIF framework was developed here to capture the effects of inactivating subthreshold potassium currents in 5-HT neurons, it lends itself equally well to capturing the effects of other subthreshold voltage-gated currents. We note that, as in other methods based on linear regression of nonlinear ion channel dynamics [138], adequate experimental estimates of the voltage-dependent gating features of the conductance at play must be available to be inserted in the aGIF model. Altogether, this expanded modeling framework adds to a toolset of computational approaches for interrogating the role of particular microcircuit motifs (e.g., FFI) or excitability features (e.g., spike-triggered adaptation) in shaping network computations, while lending itself to more elaborate inference methodologies [135].

Could the dynamical features identified here have been captured by a simpler modeling framework? Two closely-related approaches that we have not considered here are linear-nonlinear (LNL) and generalized linear models (GLMs), which are trained using only the

spike output and external input to each cell and do not consider the subthreshold voltage [141, 211]. Despite the fact that the GLM approach was not possible here given the very low firing rates of 5-HT neurons and the large number of action potentials required for accurate characterization in the absence of information about the subthreshold voltage, it is worth asking whether GLMs could in principle capture the network-level properties of 5-HT signaling. For instance, the role of spike-triggered adaptation in conveying preferential sensitivity to suddenly changing inputs arises in GLMs [124], but the state-dependence of the input derivative sensitivity identified in 5-HT neurons (Fig. 2.7) could not have been captured by a GLM implementation. In summary, the GIF framework provides a more solid foundation for network modeling than LNL- or GLM-based approaches for cell types with very low firing rates or highly state-dependent output.

Does the aGIF modeling approach represent an unnecessary complication of the GIF model framework or, conversely, an oversimplification of detailed Hodgkin-Huxley models? GIF models that do not explicitly account for the effects of specific ionic conductances produce highly accurate spiketrain predictions in many cell types [151, 152, 191, 192]; indeed, even in 5-HT neurons, the iGIF model predicts the timing of spikes with an accuracy equal to that of our aGIF model. For questions where the biophysical mechanisms that regulate spiking are not of primary interest and for systems where simpler LNL or GLM models are not able to predict the timing of spikes accurately (e.g., due to low firing rates as discussed above), non-augmented GIF models remain suitable tools. In our case, it would not have been possible to probe the effect of I_A on the network-level processing features of the DRN without the aGIF model.

2.3.2 Network level role of I_A current

Previous modeling work has implicated I_A in controlling the sensitivity of the stationary response to sustained inputs [155, 193–195]. These studies contrast with our findings which implicate this current in the control of the transient component but show almost no effect

on the stationary component of the response. This discrepancy can be explained by noting that the AHPs of 5-HT neurons (and thus of our computational model) do not reach the hyperpolarized potentials required to free I_A from inactivation (Fig. 2.1 and Fig. 2.8), in contrast to the model of Connor and Stevens [194]. As a result, I_A remains mostly inactivated during sustained inputs, and the stationary response is mostly regulated by the interplay between spike-triggered adaptation and the strength of the input. Other factors such as a shift in the activation and/or inactivation curves (e.g., by neuromodulators) are expected to influence how I_A controls the transient and stationary components of the response. Finally, it is interesting to note that I_A is also highly expressed in the dendrites of cortical neurons, where it may have an analogous function [212–215]. Our results hint at a possible general role of I_A in suppressing transient responses to sustained inputs in the midbrain, cortex, and other systems.

2.3.3 5-HT neuron heterogeneity

5-HT neurons are not all alike in every respect: recent experimental work has uncovered molecular, electrophysiological [27], developmental, and anatomical [36, 38] differences among 5-HT neurons across raphe nuclei and within the DRN (reviewed in [33]). Most relevant to our work are previously reported quantitative differences in the excitability of serotonin neurons located in the dorsomedial DRN, ventromedial DRN, and median raphe nucleus [27]. These observations suggest that the predictions made by our model, which was fitted primarily to serotonin neurons from the ventromedial DRN, may agree qualitatively but not quantitatively with the behaviour of 5-HT neuron ensembles in these areas. While there is not yet any evidence that serotonin neurons in different parts of the serotonin system perform qualitatively different computational operations, this remains an intriguing possibility for future work.

2.3.4 Heterogeneous properties of SOM neurons ensure divisive inhibition

How the heterogeneity of excitability influences the response properties of neuronal populations depends on a number of factors. Specifically, we and others [216] have argued that heterogeneity of feedback inhibition (and of principal cells) implements a divisive effect on the stationary part of the population input-output function. For feedforward inhibition, a divisive effect on the gain of stationary input-output functions is expected in naturalistic conditions [217]. The findings outlined here further support these theoretical results by showing that the heterogeneous feedforward inhibition remains divisive on the transient part of the response. Divisive inhibition has been proposed to be essential to counteract strong excitation so as to maintain activity within an adequate dynamic range [120, 218], and it is expected that brain circuits will harness cellular and circuit-level mechanisms to tune their sensitivity to relevant inputs while maintaining overall stability. This point is germane to 5-HT neurons given their position at the confluence of many excitatory input streams [17–19, 28, 38, 219]. Thus, while the exact behavioral function of the 5-HT system is still unclear, uncovering important components of its gain control mechanisms might provide useful hints about how it integrates its multifold inputs.

2.3.5 Neuromodulation of neuromodulation

Neuromodulators can dynamically reconfigure information processing in neural circuits that are otherwise anatomically fixed [220, 221]. While 5-HT is considered to be a neuromodulator, the DRN network is itself under neuromodulatory influence, both from distal (e.g., locus coeruleus or ventral tegmental area) or local (e.g., endocannabinoids, 5-HT itself) sources [17, 28, 32, 222, 223]. Whereas previous work has outlined defined cellular metrics that are modulated by specific receptor subtypes (e.g., changes in release probability or direct membrane depolarization/hyperpolarization), the consequences of these neuromodulatory

influences on higher-order network computation are only superficially understood. Here, we showcase two broad neuromodulatory mechanisms that enacts different effects on population coding. Through simulations, we show that reducing the magnitude of I_A (which could be caused for instance in vivo by noradrenergic input the DRN [197]) enhances the sensitivity of the raphe response to the onset of step inputs while leaving the stationary firing rate unchanged. In contrast, the cannabinoid-mediated preferential reduction of FFI onto 5-HT neurons (caused by the tonic activation of DRN endocannabinoid receptors, as expected to occur for instance during marijuana recreational or therapeutic use [28]) rather causes a general reduction in the output gain of the DRN. Together with our simulations probing the temporal derivative-encoding properties of this region, these observations point to a conceptual model in which the output of the DRN represents a mixture of the intensity and temporal derivative of its input where I_A controls the relative balance of the two components and FFI regulates the overall intensity of the output, and where these functions can be rapidly and independently tuned by neuromodulatory control.

Our heuristic model of the DRN helps to illustrate the unexpectedly multifaceted nature of the computations performed by this evolutionarily ancient region, but, like most heuristics, it remains an oversimplification. Some of the qualitative features of DRN processing emerging from our simulations are not explained by our "input intensity plus temporal derivative" heuristic (for example, the ability of FFI to modulate the temporal derivative-encoding properties of the DRN, or the attenuation of these same coding properties by hyperpolarization [Fig. 2.7]), presenting further opportunities to better understand the influence of neuromodulation on network computation in this region.

2.3.6 Role of derivative encoding in reinforcement learning

The role of 5-HT signaling in modulating behavior is increasingly conceptualized through the lens of reinforcement learning (RL) theory. Indeed, 5-HT output has been proposed to loosely encode or modulate every component of classical RL [64, 176], including a reward

signal [174], state value [24, 61], bias in state-action value [62], temporal discounting factor [74, 224], prediction error ([73], but see [75]), and learning rate [25, 57], with varying degrees of experimental support. Might the derivative-like computation described here have a place in an RL-based conception of DRN function? For now, it is only possible to speculate. Existing RL models of DRN function bin time in increments of tens of seconds, obscuring the faster adaptation dynamics that are the subject of our work. How and whether the sub-second fluctuations in DRN 5-HT neuron activity that are consistently observed in reward learning experiments [23–25, 58, 174] should be incorporated into RL models remains unclear. Our results suggest that RL operations that can be seen as computing a temporal derivative are candidates for an RL-based account of DRN function.

If the electrophysiological features of individual 5-HT neurons directly participate in shaping the computations enacted by the DRN, the same is likely true for other neuromodulatory systems and this work may offer overall guiding principles. For instance, dopamine neurons, well-known for their reward prediction error-like coding properties [70], bear some electrophysiological features in common with DRN 5-HT neurons, with both cell types exhibiting strong adaptation and a prominent A-type potassium current [225, 226]. Dopamine neurons have been proposed to encode reward prediction errors partly by approximating a mixture of a value signal and its temporal derivative [227], hinting at a possible role for adaptation in implementing one of the central computations of RL.

If the derivative-like operation identified here does not directly contribute to computing one of the key components of RL, what might its role in the DRN be? One possibility is that strong spike-triggered adaptation may optimize the efficiency of neural coding by filtering out temporally redundant information, a phenomenon referred to as predictive coding and that is ubiquitous in sensory systems [111, 228–230]. As the search for a unified interpretation of DRN 5-HT activity continues, our results provide a new perspective on the fast component of 5-HT neuron dynamics: fluctuations in 5-HT neuron's activity do not solely encode the intensity of their input, but rather how quickly their inputs are changing over time.

2.4 Materials and methods

2.4.1 Experimental methods

Animals

Experiments were performed on male and female C57/Bl6 mice aged 4-8 weeks. Slc6a4-cre::Rosa-TdTomato (SERT-Cre) and Sst-cre::Rosa-TdTomato transgenic lines were used to fluorescently label DRN 5-HT and somatostatin (SOM) GABA neurons, respectively. Animals were group-housed and kept on a 12:12-h light/dark cycle with access to food and water *ad libitum*. All experiments were carried out in accordance with procedures approved by the University of Ottawa Animal Care and Veterinary Services (protocol numbers CMM-164, CMM-176, CMM-1711, CMM-1743, and CMM-2737).

Slice preparation

Animals were deeply anaesthetized using isofluorane (Baxter Corporation) before being euthanized by decapitation. The brain was quickly removed from the skull and submerged into ice-cold dissection buffer containing the following: 119.0 mM choline chloride, 2.5 mM KCl, 4.3 mM $MgSO_4$, 1.0 mM $CaCl_2$, 1.0 mM NaH_2PO_4, 1.3 mM sodium ascorbate, 11.0 mM glucose, 26.2 mM $NaHCO_3$; saturated with 95 % O_2/5 % CO_2. A Leica VT1000S vibratome was used to cut 300 µm coronal sections of midbrain containing the DRN or of the cortex containing the mPFC in the same ice-cold choline dissection buffer. After cutting, slices were placed in a recovery chamber filled with artificial cerebrospinal fluid (ACSF) containing the following: 119.0 mM NaCl, 2.5 mM KCl, 1.3 mM $MgSO_4$, 2.5 mM $CaCl_2$, 1.0 mM NaH_2PO_4, 11.0 mM glucose, 26.2 mM $NaHCO_3$; 298 mOsm, maintained at 37 °C, and continuously bubbled with 95 % O_2/5 % CO_2. The recovery chamber was allowed to equilibrate to room temperature for 1 h before beginning experiments.

In vitro whole-cell electrophysiological recording

Neurons were visualized using an upright microscope (Olympus BX51WI) equipped with differential interference contrast and a 40×, 0.8 NA water-immersion objective. Whole-cell recordings were obtained from fluorescently-labelled DRN 5-HT and SOM neurons and unlabeled mPFC L5 pyramidal neurons using glass electrodes (Sutter Instruments; tip resistance 4 MΩ to 6 MΩ). For most experiments, the following potassium gluconate-based internal solution was used: 135 mM potassium gluconate, 6.98 mM KCl, 10 mM HEPES, 4 mM Mg ATP, 0.40 mM GTP, 10 mM Na phosphocreatine; adjusted to pH 7.25 with KOH, 280 mOsm to 290 mOsm. A subset of experiments (GABA synaptic physiology) were carried out using a cesium-based internal solution (120 mM $CsMeSO_3$, 10 mM EGTA, 5 mM TEA Cl, 1 mM $CaCl_2$, 10 mM Na HEPES, 4 mM Mg ATP, 2 mM GTP, 2 mM QX-314, and 10 mM Na phosphocreatine; adjusted to pH 7.25 with CsOH, 280 mOsm to 290 mOsm) and in the presence of bath-applied 100 µM (2R)-amino-5-phosphonovaleric acid (APV) and 5 µM 2,3-Dioxo-6-nitro-1,2,3,4-tetrahydrobenzo[f]quinoxaline-7-sulfonamide (NBQX). For voltage clamp experiments, whole-cell capacitance compensation was applied manually following break-in, and leak current subtraction was performed post hoc using membrane leak conductance estimated based on a −5 mV pulse at the start of each sweep. Experiments were carried out at room temperature except where noted. For current clamp experiments used to fit GIF models, access resistance was compensated using an active electrode compensation method [203]. For voltage clamp experiments used to characterize I_A in 5-HT neurons at room temperature, recordings had $R_a = 14.7 \pm 6.2$ MΩ (mean ± SD; half of recordings between 12.8 MΩ and 21.6 MΩ) after applying an access resistance cutoff of 30 MΩ (a more stringent cutoff of 20 MΩ yielded statistically indistinguishable estimates of I_A maximal conductance and kinetic parameters; compare Fig. 2.1D and Fig. 2.12). For voltage clamp experiments used to characterize whole-cell currents in SOM neurons, recordings had $R_a = 14.3 \pm 7.0$ MΩ (mean ± SD; half of recordings between 9.8 MΩ and 15.5 MΩ) after applying a similar cutoff of 30 MΩ. For synaptic electrophysiology experiments, recordings had $R_a = 5.7 \pm 0.5$ MΩ

(mean ± SD; range 5.0 MΩ to 6.1 MΩ) after applying a cutoff of 10 MΩ. Recordings were collected with an Axon MultiClamp 700B amplifier and the analog signals were filtered at 2 kHz and digitized at 10 kHz using an Axon Digidata 1550 digitizer.

2.4.2 Models

GIF and related models

The generalized integrate-and-fire (GIF) and Na-inactivation GIF (iGIF) models have been described previously in detail [151, 203, 204]. Briefly, the GIF and iGIF are composed of a subthreshold component which integrates input currents into voltage and a stochastic spiking rule which transforms subthreshold voltage into a series of spikes. The subthreshold dynamics of the GIF and iGIF are given by

$$C\frac{dV}{dt} = -g_l\left(V(t) - E_l\right) - \sum_{\hat{t}_i<t} \eta\left(t - \hat{t}_i\right) + I_{\text{inj}}(t) \tag{2.1}$$

where $\{\hat{t}_i\}$ is the set of spike times and $\eta(t) = \sum_j \beta_j^{(\eta)} e^{\frac{-t}{\tau_j^{(\eta)}}}$ is the spike-triggered adaptation current. Here the $\beta_j^{(\eta)}$ are coefficients estimated from the data and the $\tau_j^{(\eta)}$ are fixed hyper-parameters; see appendix for details. The GIF emits spikes according to an inhomogeneous Poisson process with intensity $\lambda(t)$, given by

$$\lambda(t) = \lambda_0 \exp\left[\frac{V(t) - V_T^* - \sum_{\hat{t}_i<t} \gamma(t - \hat{t}_i)}{\Delta V}\right] \tag{2.2}$$

where V_T^* is the stationary threshold, $\gamma(t) = \sum_j \beta_j^{(\gamma)} e^{-t/\tau_j^{(\gamma)}}$ is the spike-triggered threshold movement (where the $\beta_j^{(\gamma)}$ are coefficients estimated from the data and the $\tau_j^{(\gamma)}$ are fixed; see appendix), ΔV is the threshold sharpness (mV; larger values increase the stochasticity of spiking), and $\lambda_0 = 1$ Hz is a constant such that $\lambda(t)$ is in units of Hz. In the iGIF, an additional variable $\theta(t)$ is added to the numerator of the exponentiated term in Eq. (2.2) to

account for voltage-dependent changes in threshold

$$\lambda(t) = \lambda_0 \exp\left[\frac{V(t) - V_T^* - \sum_{\hat{t}_i<t}\gamma(t-\hat{t}_i) + \theta(t)}{\Delta V}\right]$$
$$\frac{d\theta}{dt} = \frac{\theta_\infty(V) - \theta}{\tau^{(\theta)}}.$$

The equilibrium voltage-dependent change in spike threshold $\theta_\infty(V) = \sum_{j=1}^{N_{\text{step}}} \beta_j^{(\theta)} \text{rect}\left[V; A_j, A_{j+1}\right]$ is a piecewise constant function of voltage where each $\beta_j^{(\theta)}$ defines the value of $\theta_\infty(V)$ over the voltage range $[A_j, A_{j+1})$ and $N_{\text{step}} = 5$. The locations of the steps in the piecewise constant function A_j are selected based on the data. (See [204] for details on the iGIF model.) Our aGIF model is identical to the GIF model except that two Hodgkin-Huxley currents which together capture the voltage-gated potassium currents found in 5-HT neurons (see "Potassium current," below) are added to the subthreshold dynamics given in Eq. (2.1), yielding

$$C\frac{dV}{dt} = -g_l\left(V(t) - E_l\right) - I_A(t) - I_K(t) - \sum_{\hat{t}_i<t}\eta\left(t-\hat{t}_i\right) + I_{\text{inj}}(t) \tag{2.3}$$

as the definition of the subthreshold dynamics of the aGIF model.

The procedures for fitting the GIF and iGIF models to electrophysiological data have also been described previously in detail [151, 203, 204]. Briefly, parameter estimation for both models occurs in two stages: first, the subthreshold parameters are estimated by regression, and second, the threshold parameters are estimated by maximizing the likelihood of the observed spiketrain as a function of the threshold parameters. The fitting procedure for the aGIF is very similar to that of the GIF, with adjustments to the subthreshold fitting procedure to accommodate the extra terms in Eq. (2.3) (see appendix for details). Neurons with non-stationary firing statistics (Pearson correlation between number of spikes and validation sweep number $r > 0.9$) or highly variable spike timing (intrinsic reliability <0.1) were automatically excluded from our analysis.

LIF neuron with an inactivating potassium current

Our toy model of a neuron with an inactivating potassium current is based on a LIF augmented with $I_A(t)$ (see "Potassium current" below):

$$C\frac{dV}{dt} = -g_l(V(t) - E_l) - I_A(t) + I_{\text{inj}}(t),$$

where g_l and E_l are the leak conductance and reversal, respectively, and $I_{\text{inj}}(t)$ is the external input to the model. To reduce the number of free parameters, the model we used is non-dimensionalized with respect to the membrane time constant $\tau_{\text{mem}} = R_{\text{mem}}C$ and leak conductance g_l, yielding

$$\frac{dV}{dt} = E_l - V(t) - \bar{g}'_A m_\infty h\,(V(t) - E_K) + V_{\text{inj}}(t)$$

where t is in units of the membrane time constant, $\bar{g}'_A = \bar{g}_A/g_l$ is the effective maximum conductance associated with I_A, and $V_{\text{inj}}(t) = I_{\text{inj}}(t)/g_l$ is the effective external input. The gating variables m_∞ and h are described below in "Potassium current".

Potassium current

The voltage-gated potassium currents in 5-HT neurons were modeled in terms of an inactivating current and a non-inactivating current we refer to as I_A and I_K, respectively. These were defined as follows

$$\begin{aligned} I_A &= \bar{g}_A\, m_\infty(V)\, h(t)\ (V(t) - E_K) \\ I_K &= \bar{g}_K\, n_\infty(V)\ (V(t) - E_K) \end{aligned} \tag{2.4}$$

where $\bar{g}_i$ is the maximal conductance; m and h are the activation and inactivation gates of I_A, respectively; n is the activation gate of I_K; and $E_K = -101\,\text{mV}$ is the reversal potential of potassium in our recording conditions. Note that although this value is not physiological, the effect of varying this parameter is very similar to the effect of varying $\bar{g}_A$, as we have done

in our results. For simulations involving models fitted to data collected at 29 °C to 30 °C, $E_K = -89.1\,\text{mV}$ was used. The equilibrium state of each gate $x \in \{m, h, n\}$ is a sigmoid function of voltage

$$x_\infty(V) = \frac{A_x}{1 + e^{-k_x(V - V_x^*)}},$$

where V_x^* is the half-activation voltage (mV), k_x is the slope (mV^{-1}), and A_x is a scaling factor.

To keep the number of parameters in our current model to a minimum, we assumed that the m and n gates have instantaneous kinetics (allowing their corresponding equilibrium gating functions m_∞ and n_∞ to be used directly in Eq. (2.4)), and that the h gate activates and deactivates with a single time constant τ_h (ms) that does not depend on voltage. The time dynamics of the h gate are therefore given by

$$\frac{dh}{dt} = \frac{h_\infty - h}{\tau_h}.$$

Quantification of single neuron model performance

R^2 was calculated based on the training set $\frac{dV}{dt}$ predicted by the subthreshold component of a given GIF model (Eqs. (2.1) and (2.3), where the spike times $\hat{t}$ were constrained to match the data), excluding a small window around each spike (from 1.5 ms before to 6.5 ms after in 5-HT neurons, and from 1.5 ms before to 4.0 ms after in SOM and mPFC neurons). M_d^* was calculated based on validation set data as previously described by [231]. This metric is defined as

$$M_d^* = \frac{2n_{dm}}{n_{dd}^* + n_{mm}},$$

where n_{dm} is the number of model-predicted spikes that occur within 8 ms of a spike in the validation data, and n_{dd}^* and n_{mm} are the corresponding numbers of coincident spikes across sweeps in the validation data and model predictions (where n_{dd}^* is corrected for small sample bias). M_d^* can be interpreted as the fraction of model-predicted spikes that occur within 8 ms

of a spike emitted by a real neuron (the spike timing precision is set to 8 ms by inspecting the relationship between precision and intrinsic reliability [232]), corrected such that the chance level is 0 and perfect agreement between predicted and observed spikes is 1.

Population models

DRN network models were constructed by connecting a population of 400 SOM neuron models to a population of 600 5-HT neuron models in a feed-forward arrangement. Population models were bootstrapped by sampling with replacement from a bank of experimentally-constrained GIF models. SOM neuron models were randomly connected to 5-HT neuron models with a connection probability of 2 %, such that the expected number of GABAergic synapses on each 5-HT neuron model was 8. We used a conductance-based model of GABAergic synapses with a fixed reversal potential of $-76.7\,\mathrm{mV}$, conductance of $0.3\,\mathrm{nS}$, and biexponential kinetics with $\tau_{\mathrm{rise}} = 1.44\,\mathrm{ms}$, $\tau_{\mathrm{decay}} = 26.0\,\mathrm{ms}$, and a propagation delay of 2.0 ms.

Simulated 5-HT populations with decreased or increased I_A were generated by setting $\bar{g}_A$ in all single neuron models to 0 nS or 10 nS, respectively. DRN network models with homogenized SOM neuron populations were created by setting all SOM neuron model parameters to their respective median values from the bank of experimentally-constrained single neuron models. Population models in which the adaptation mechanisms of 5-HT and SOM neuron models were swapped were generated by randomly sampling a GIF model of the opposite cell type and substituting in its adaptation filter coefficients $\beta_j^{(\gamma)}$ and $\beta_j^{(\eta)}$. This procedure is summarized in the following pseudocode:

for 5-HT_model in 5-HT_population **do**

SOM_model = random_choice(SOM_models)

5-HT_model.eta.coefficients = SOM_model.eta.coefficients

5-HT_model.gamma.coefficients = SOM_model.gamma.coefficients

end for

2.4.3 Numerical methods

Simulations were implemented in Python and C++ using custom-written extensions of the GIF Fitting Toolbox ([203]; original code archived at `https://github.com/pozzorin/GIFFittingToolbox`). Numerical integration was performed using the Euler method with a time step of 0.1 ms for the GIF model and related models (to match the sampling rate of electrophysiological recordings) and $0.001\tau_{\text{mem}}$ for the toy model of a neuron with I_A.

2.4.4 Statistics

Statistical analysis was carried out using the SciPy and statannot (`https://github.com/webermarcolivier/statannot`) Python packages. Non-parametric tests were used for all two-sample comparisons (Mann-Whitney U test for unpaired samples and Wilcoxon signed-rank test for paired samples). Non-parametric tests were chosen because we often had reason to believe that our data did not come from a Normal distribution, either due to intrinsic qualities of the data, such as being bounded between 0 and 1, or due to skewness apparent in our samples. Whenever multiple tests were performed in the same figure panel, p-values were adjusted for multiple comparisons using the Bonferroni correction. '*', '**', '***', and '****' are used in figures to denote statistical significance at the $p \leq 0.05, 0.01, 0.001$, and 0.0001 levels, respectively, and 'o' is used to indicate a trend towards significance (defined as $0.05 < p \leq 0.1$). Exact p-values are reported in the main text, and summary statistics are presented as mean ± standard-deviation. Sample sizes always refer to biological replicates.

2.4.5 Data and code availability

Raw data is available on Dryad at `https://doi.org/10.5061/dryad.66t1g1k2w`. Code to fit models, run simulations, and reproduce figures is available at `https://github.com/nauralcodinglab/raphegif`.

Acknowledgements

This work was carried out on the unceded and unsurrendered land of the Algonquin Anishinaabe people. We thank Dr. Simon Chen for providing SOM-Cre::Rosa-TdTomato mice and Liwen Cai and David Lemelin for technical assistance with mouse lines. We would also like to thank Sébastien Maillé for contributions to pilot-testing of the GIF model in 5-HT neurons and all other members of the J.-C.B. and R.N. labs for helpful discussions. E.H. is thankful to have received graduate scholarships from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada. This work was supported by grants from the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada, the Canada Foundation for Innovation, Brain Canada (Canadian Neurophotonic Platform), and the Krembil Foundation.

Supplementary figures

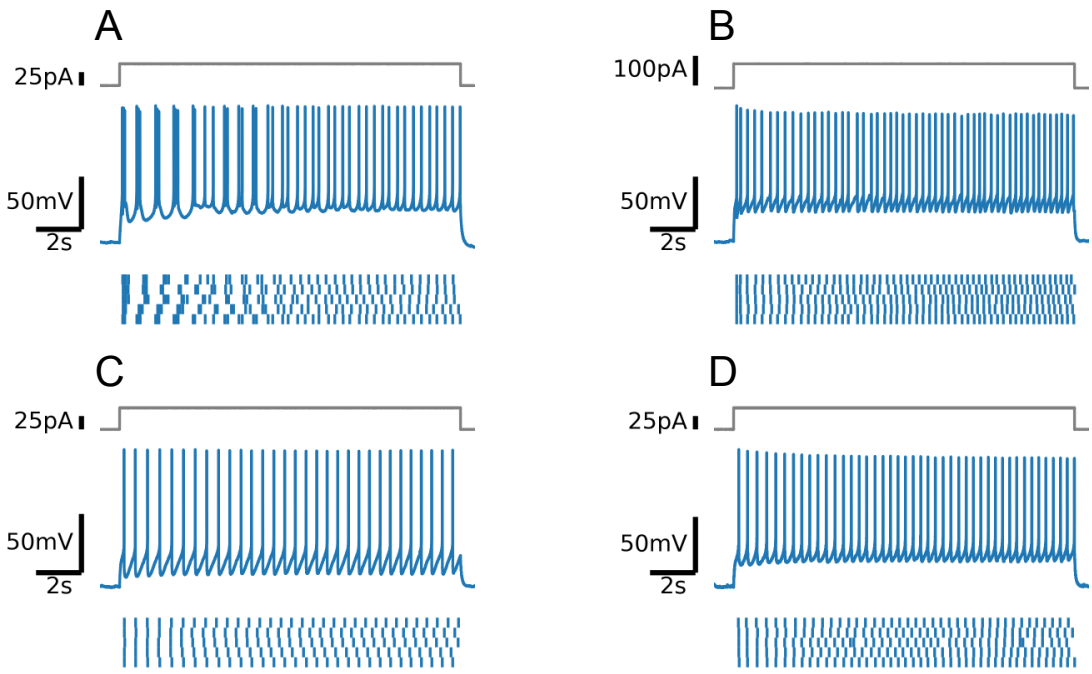


Figure 2.8: Firing patterns of four positively-identified 5-HT neurons. Note intermittent bursting in the first cell.

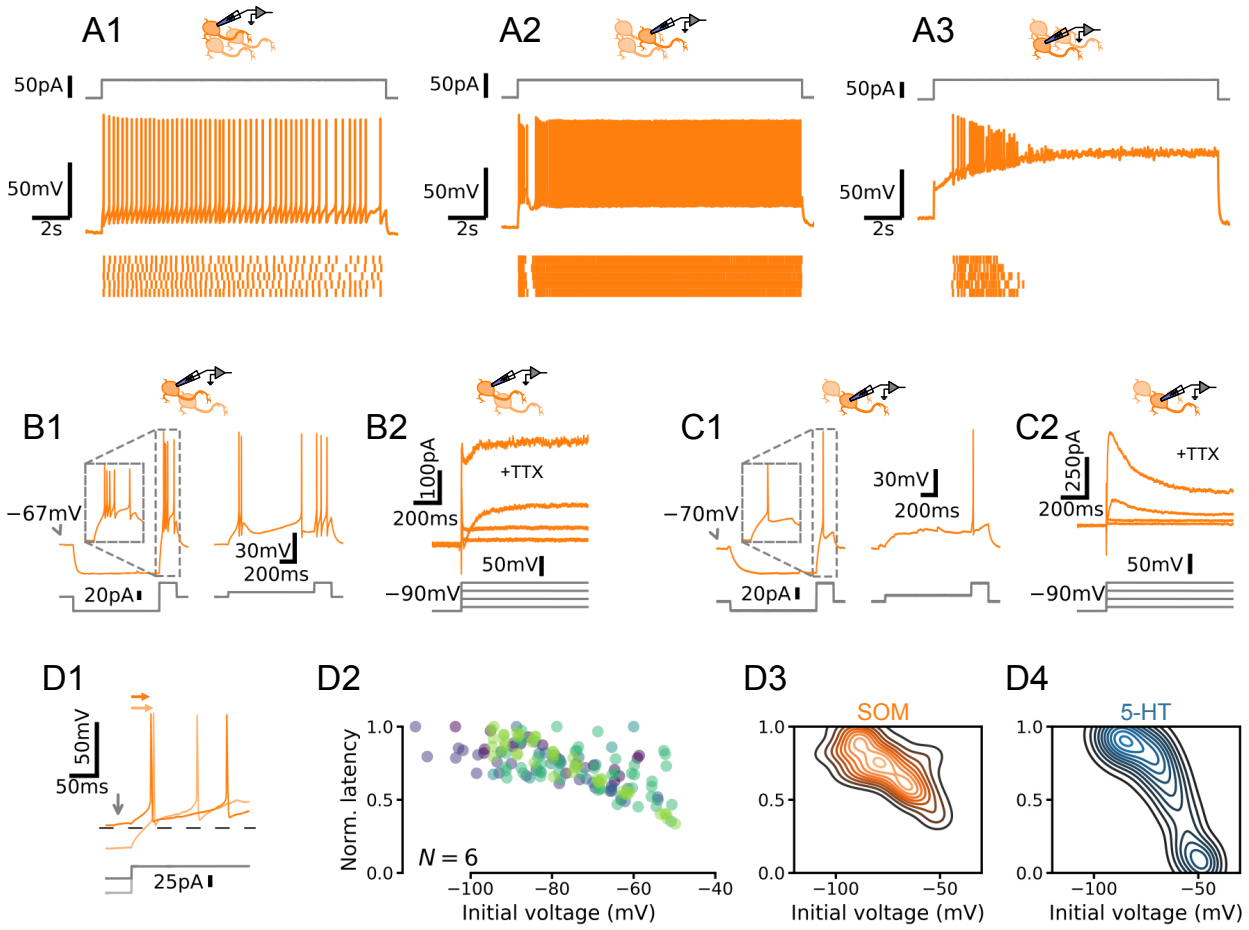


Figure 2.9: DRN SOM neurons are not homogenous. **A** Qualitatively different firing patterns in three SOM neurons. **B** A burst-firing SOM neuron with a transient depolarizing current. **C** A SOM neuron that does not fire bursts with a pronounced hyperpolarizing current. **D** Heterogenous relationships between voltage history and spike timing across cells. Experiment is the same as in Fig. 2.2. (D1) Representative experiment. Arrows indicate spike latency and initial voltage. (D2) Normalized spike latency for $N = 6$ SOM neurons. (D3) Gaussian kernel density estimate of data shown in D2. (D4) Gaussian kernel density estimate of data shown in Fig. 2.2F for comparison.

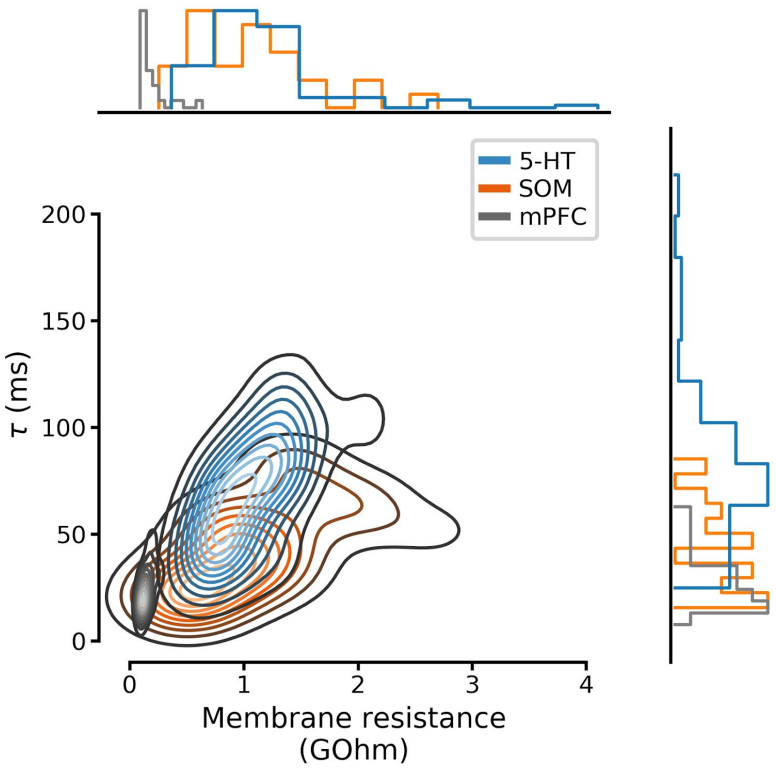


Figure 2.10: Full distributions of membrane parameters listed in Table 1. Curves on the main set of axes are Gaussian kernel density estimates.

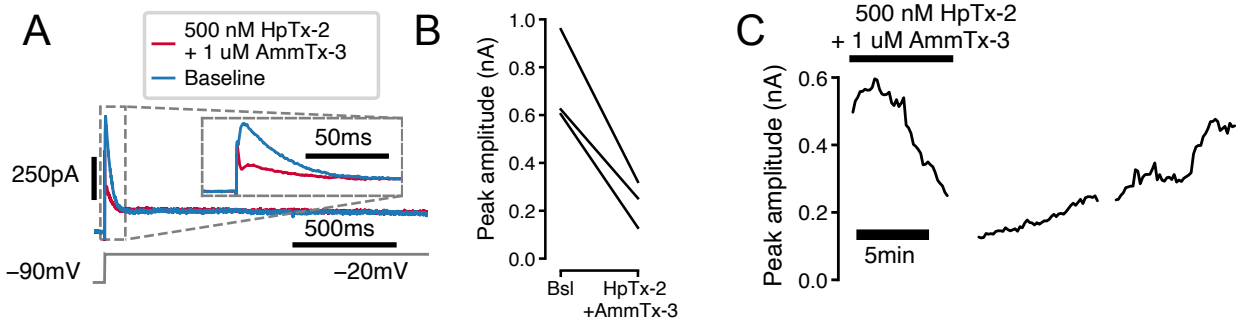


Figure 2.11: Transient outward current found in 5-HT cells (I_A) is sensitive to potassium channel blockers. **A** Representative voltage-clamp trace showing attenuation of transient outward current in a 5-HT neuron by the selective K_v4-channel blockers Heteropodatoxin-2 and AmmTx-3. Currents are leak-subtracted. **B** Attenuation of transient outward current in $N = 3$ 5-HT neurons. **C** Time-course of attenuation of transient outward current in a representative 5-HT neuron.

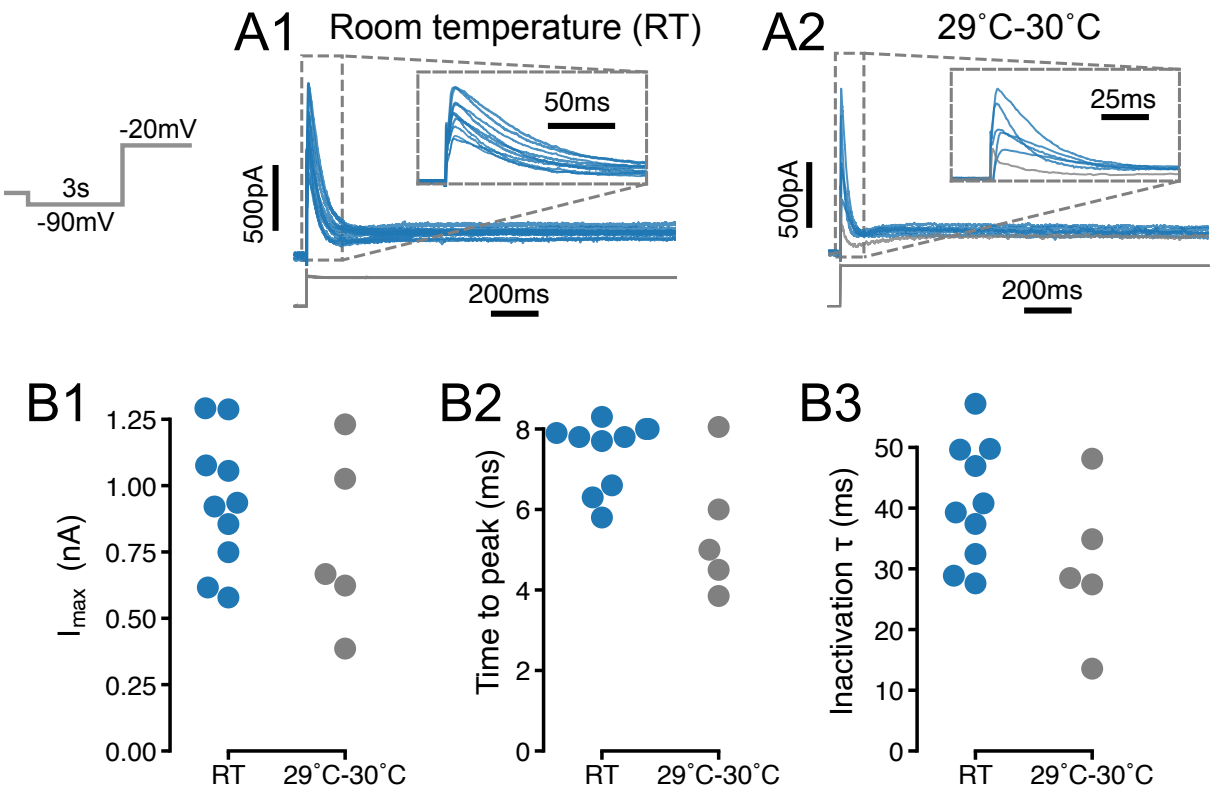


Figure 2.12: Temperature-dependence of I_A amplitude and kinetics in 5-HT neurons. **A** Transient currents evoked by depolarizing steps in 5-HT neurons recorded at room temperature (RT; $N = 10$) and 29 °C to 30 °C ($N = 5$). Traces in A1 are also presented in Fig. 2.1C1. Gray trace in A2 was excluded from further analysis due to lack of a clearly-defined transient current (based on same criteria as in Fig. 2.1D). **B** Comparison of amplitude and kinetic parameters extracted from traces in A. Each point is one neuron. Room temperature (blue) data are also presented in Fig. 2.1D. Recordings with access resistance >20 MΩ were excluded (included room temperature recordings had access resistance of 12.1 ± 4.0 MΩ, 29 °C to 30 °C recordings had access resistance of 13.2 ± 5.2 MΩ; mean ± SD).

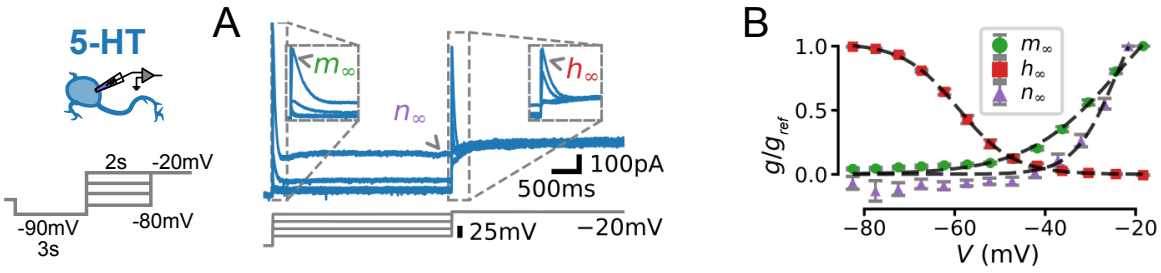


Figure 2.13: Characterization of voltage-dependence of I_A in 5-HT neurons. **A** Experimental protocol. Voltage command designed to reveal voltage-dependent activation and inactivation of I_A (left) and representative leak-subtracted and capacitance-corrected current trace (blue, top right) aligned to voltage command (gray, bottom right). Parts of the voltage trace used to fit the activation (m_∞), inactivation (h_∞), and non-inactivating (n_∞) gates (see Methods) are indicated. **B** Gating curves extracted from $N = 13$ neurons. g_{ref} is the maximal conductance observed for each cell (i.e., the conductance at $V = -20\,\text{mV}$ for m_∞ and n_∞, and at $-80\,\text{mV}$ for h_∞). Dashed lines indicate boltzmann functions with parameters given in Table 2.

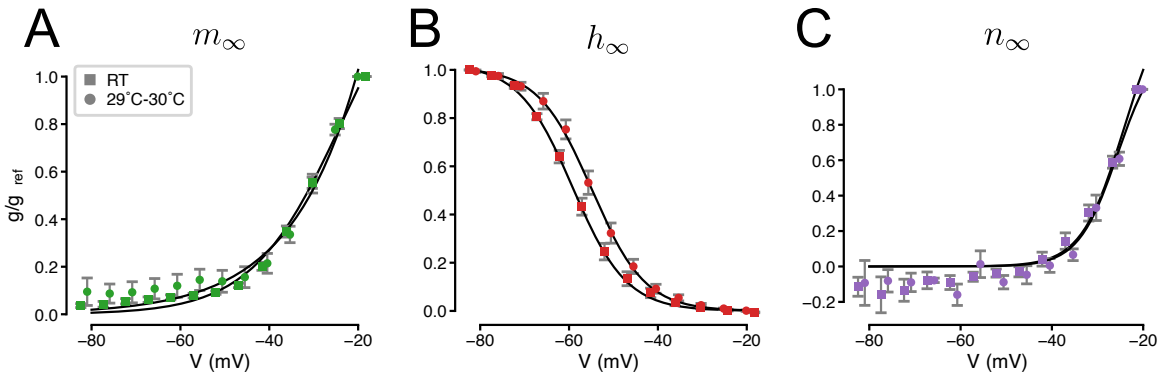


Figure 2.14: Temperature-dependence of the gating of I_A. Squares indicate data collected at room temperature (RT) reproduced from Fig. 2.15. Circles indicate recordings at 29 °C to 30 °C ($N = 13$ 5-HT neurons). Solid lines indicate boltzmann functions. Activation **A**, inactivation **B**, and non-inactivating **C** gates (see Methods and Fig. 2.13).

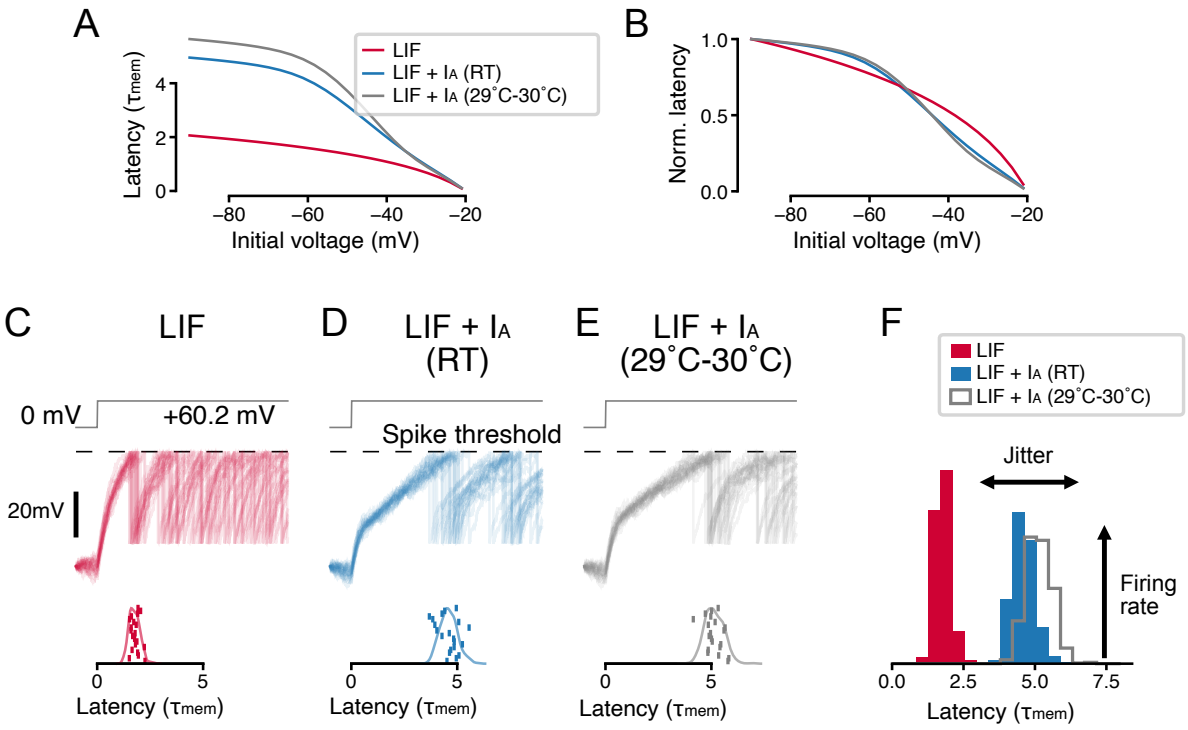


Figure 2.15: Parameter and temperature-dependence of the effect of I_A on spike-timing. **A**, **B** A toy model neuron with I_A predicts a non-linear relationship between voltage-history and latency to the first spike evoked by a step input. Setup is the same as in Fig. 2.2A-C. **C**–**F** I_A causes an increase in spike latency and jitter in the presence of noise. Models are the same as in A and B, setup is the same as in Fig. 2.2G-I. Effective I_A conductance and inactivation time constants were chosen to match experimentally-determined values for 5-HT neurons: the toy model with I_A at room temperature (RT) has an effective I_A conductance $\bar{g}'_A/g_l = 11.4$ and effective inactivation time constant $\tau_h/\tau_{\text{mem}} = 0.545$, the toy model with I_A at 29 °C to 30 °C has an effective I_A conductance $\bar{g}'_A/g_l = 4.54$ and effective inactivation time constant $\tau_h/\tau_{\text{mem}} = 0.941$. Gating parameters are taken from Fig. 2.14.

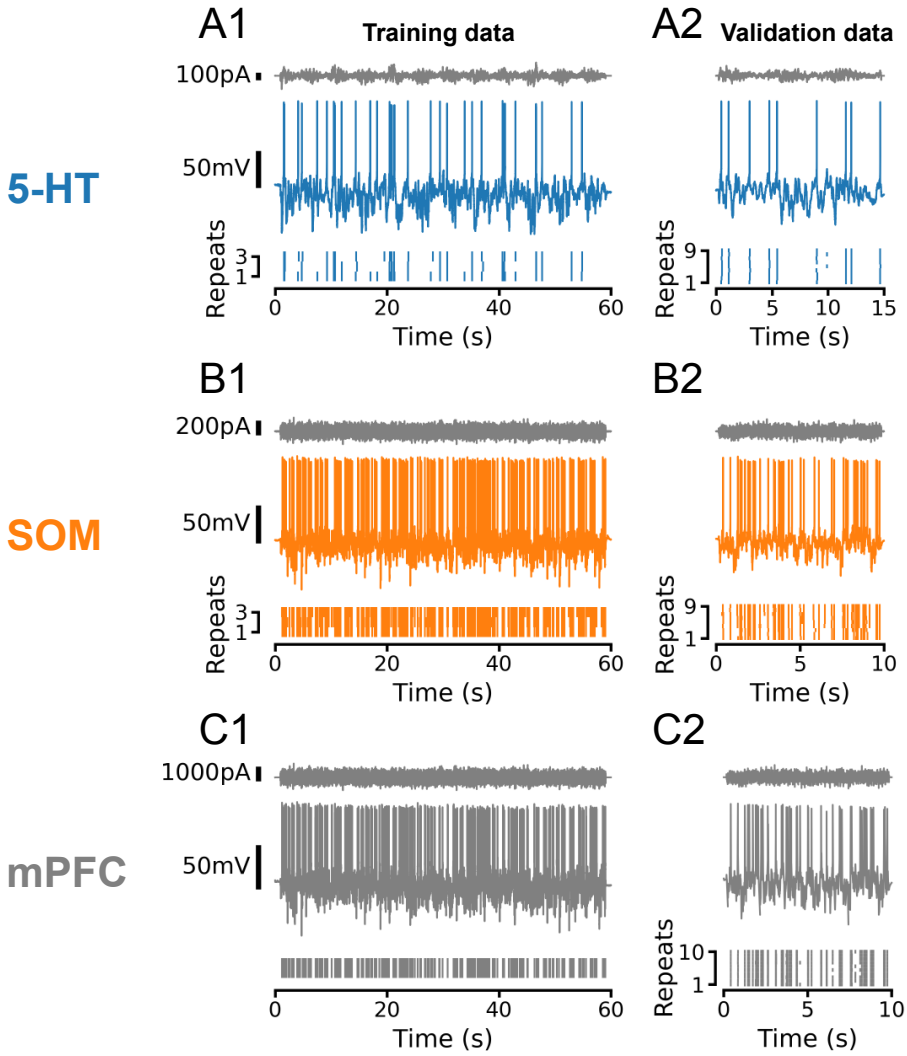


Figure 2.16: Representative training and validation sets for all cell types. In all cases, training and validation sets consisted of frozen Ornstein-Uhlenbeck (OU) noise with a sinusoidally-modulated variance. Noise was rescaled to evoke subthreshold fluctuations of approximately equal amplitude across individual neurons and cell types. **A** Representative dataset from a 5-HT neuron using OU noise with a time constant $\tau = 50\,\text{ms}$ and variance modulations with a period of 5 s. **B**, **C** Representative datasets from SOM and mPFC neurons using OU noise with a time constant of $\tau = 3\,\text{ms}$ and variance modulations with a period of 100 ms. Note that the number and duration of sweeps in the training and test sets varied between cell types, as shown. In 5-HT and SOM neurons, training and validation sweeps were interleaved such that each training sweep would be followed by three of the nine validation sweeps. In all cases, active electrode compensation (see) [203]) was applied based on a single 10 s sweep of frozen OU noise with $\tau = 3\,\text{ms}$ before fitting and validating GIF models.

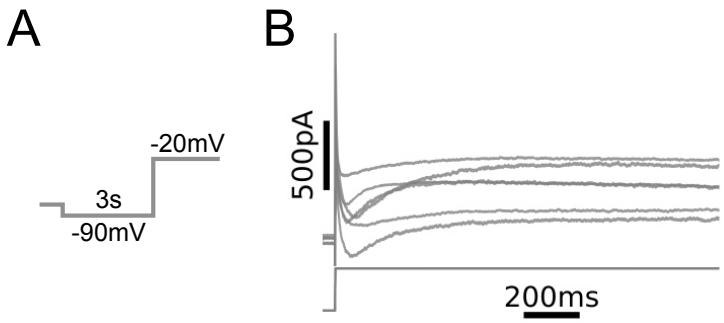


Figure 2.17: Whole cell currents observed in mPFC neurons. **A** Voltage clamp protocol. Same as in Fig. 2.1C. **B** Leak-subtracted whole-cell currents from $N = 6$ L5 mPFC pyramidal neurons (top) and voltage command (bottom). Recordings were carried out in the presence of bath-applied tetrodotoxin (TTX).

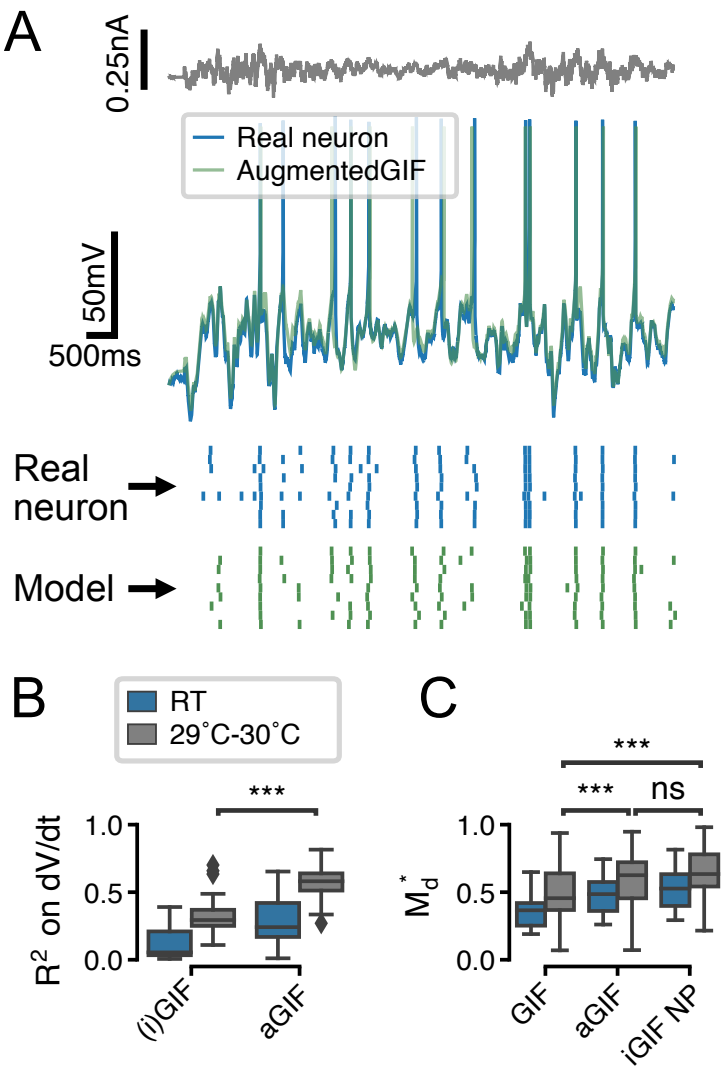


Figure 2.18: The aGIF model accurately predicts the subthreshold voltage and firing patterns of 5-HT neurons recorded at room temperature (RT) and 29 °C to 30 °C. **A** Representative validation data and aGIF model predictions for a 5-HT neuron recorded at 29 °C to 30 °C. Compare with Fig. 2.3D2. **B**, **C** Quantification of model performance in terms of R^2 on the training subthreshold voltage derivative and on the validation spike-train similarity metric M_d^*. Room temperature data from Fig. 2.3F are presented for comparison, but note that differences across experimental conditions (i.e., room temperature vs. 29 °C to 30 °C) in B may be due to confounds such as recording noise. Room temperature $N = 18$, 29 °C to 30 °C $N = 18$.

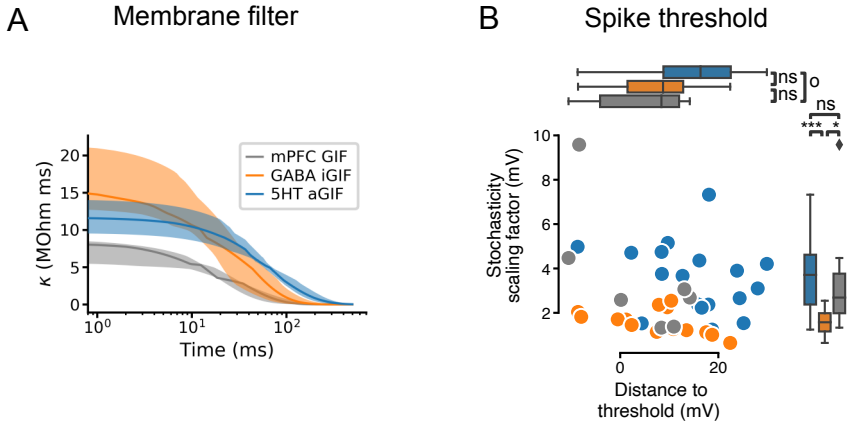


Figure 2.19: Additional features extracted from single neuron models. **A** Passive membrane filter. Presented as median and interquartile range, as in Fig. 2.4C, D. **B** Spike threshold features: stochasticity scaling factor (ΔV, also known as the threshold width; larger values indicate greater stochasticity) and distance to threshold ($V_T^* - V_{\text{rest}}$).

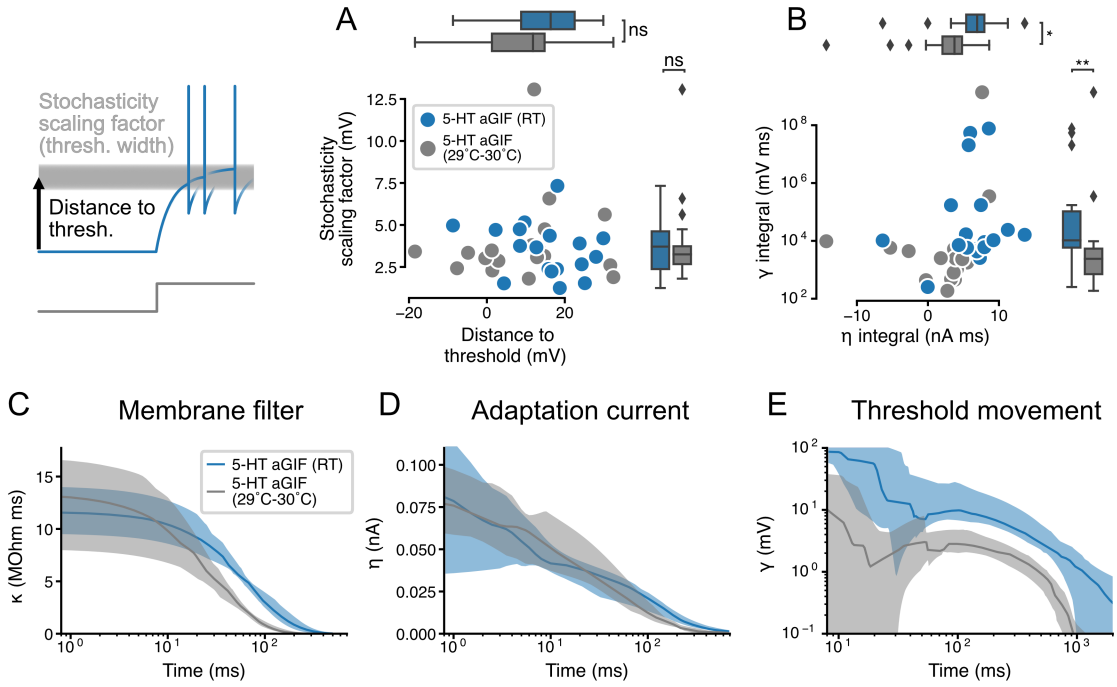


Figure 2.20: Temperature-dependence of features extracted from aGIF models fitted to 5-HT neurons. **A** Spike threshold features. **B** Adaptation strength. Two aGIFs fitted to data collected at 29 °C to 30 °C with γ integral ≤ 0 are not shown for clarity, but are used in statistical tests. (C-E) Passive membrane **C**, spike-triggered current **D**, and spike-triggered threshold movement **E** model filters. Data are presented as median and interquartile range, as in Fig. 2.4 and Fig. 2.19. Data from Fig. 2.4 and Fig. 2.19 are shown in blue for comparison. $N = 18$ models fitted to data collected at room temperature (RT), $N = 18$ models fitted to data collected at 29 °C to 30 °C.

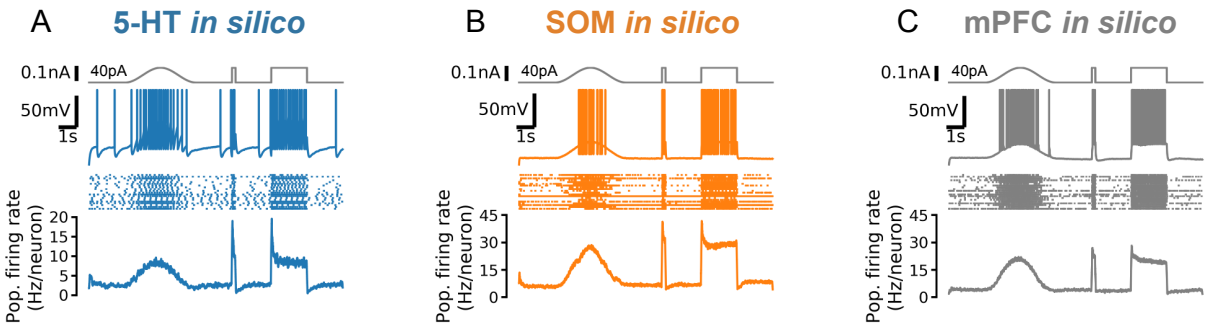


Figure 2.21: Responses to fast and slow inputs are robust to increases in input baseline. Square step is the same as Fig. 2.5C but with 40 pA background input. Note strongly-adapting population firing rate in 5-HT neurons.

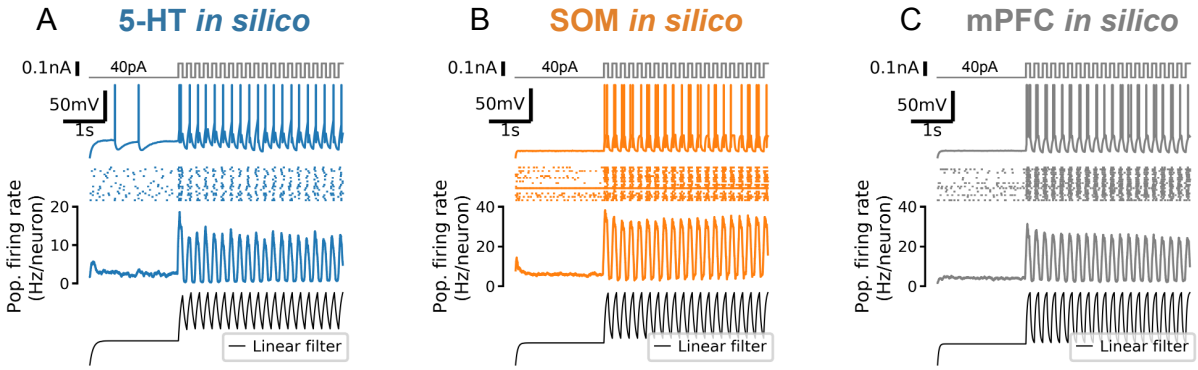


Figure 2.22: Strong responses of neuron populations to sudden inputs are due to a non-linear filtering effect. A 5 Hz square wave is applied to each simulated population on top of a constant 40 pA input. Population firing rate shows the mean of $N = 20$ independent simulations. Bottom row shows the population firing rate (in arbitrary units) predicted by a linear population coding model (a monoexponential filter with a time constant equal to the mean membrane time constant of the corresponding cell type applied to the input signal). Note that the start of the square wave component of the stimulus is amplified by the simulated neuron populations (second row from bottom) but not by the linear model (bottom row).

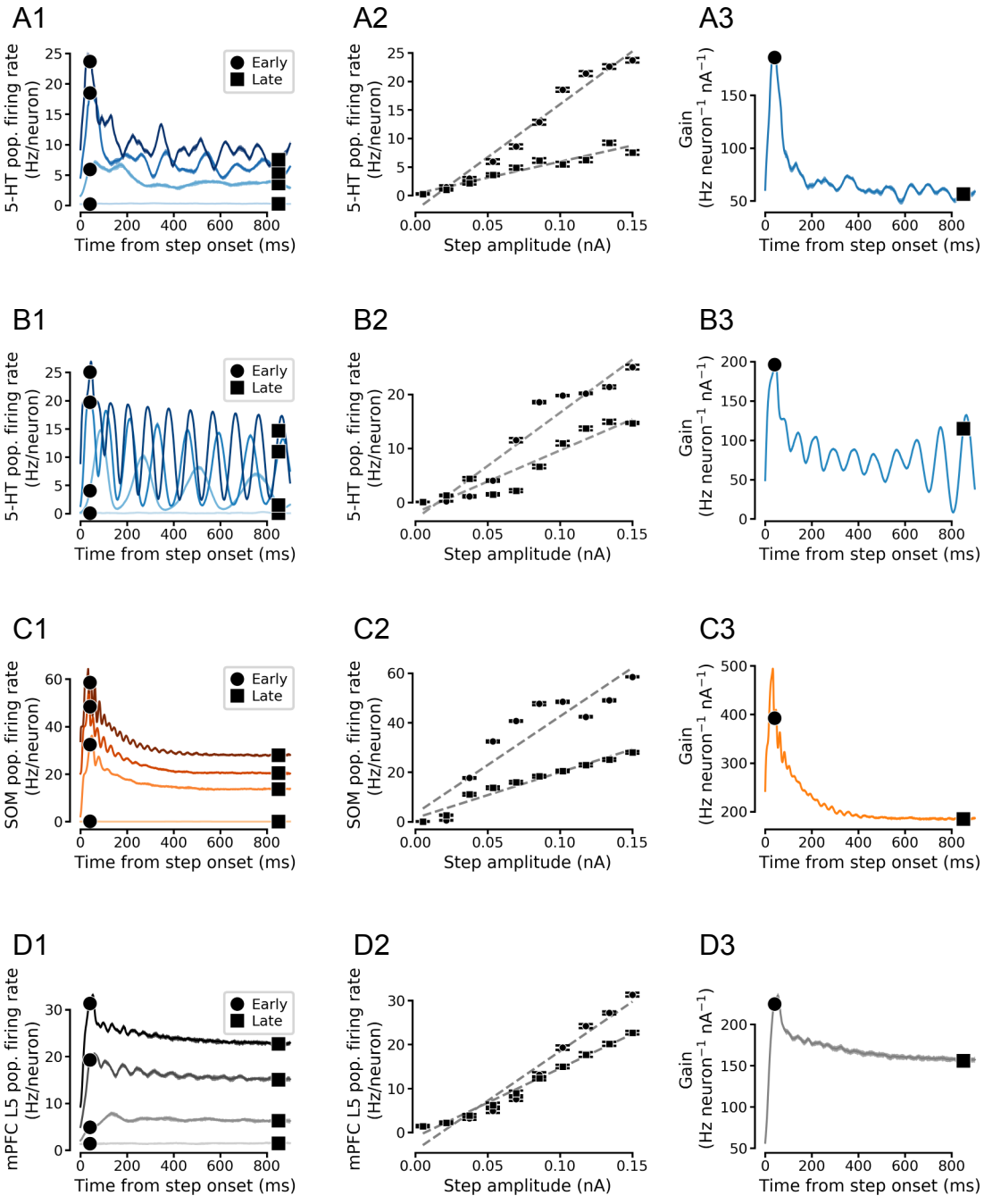


Figure 2.23: Simulated population input-output functions across cell types. **A** Simulated 5-HT neuron population. **B** Simulated homogenous population of 5-HT neurons. 5-HT neuron models were homogenized by setting each parameter to its median value across all aGIF models fitted to 5-HT neurons. Note oscillations. **C** Simulated SOM neuron population. **D** Simulated mPFC neuron population. Note that the f/I curves of (heterogenous) 5-HT and SOM neuron populations are approximately linear in the simulated input range (A and C). See Fig. 2.5D for a schematized version of this figure.

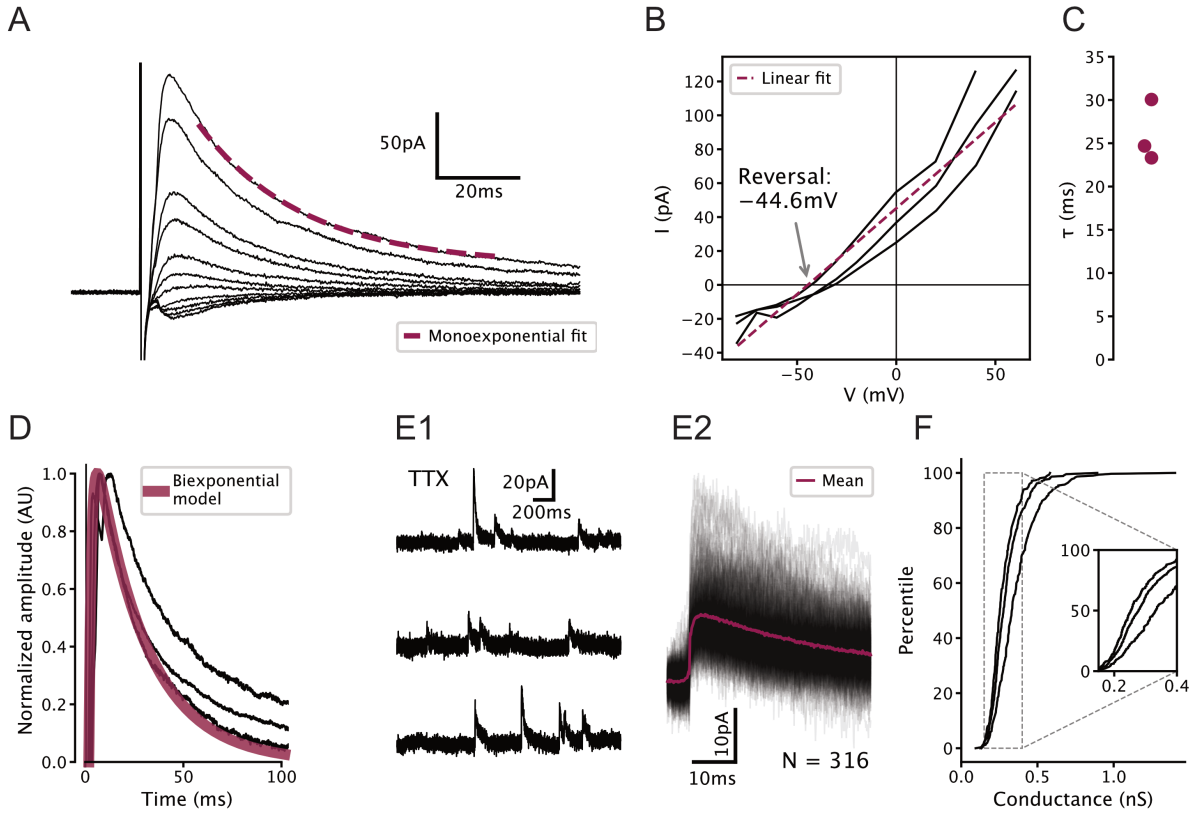


Figure 2.24: Characterization of GABAergic synapses on 5-HT neurons used to constrain network model. All recordings were carried out in 100 µM APV and 5 µM NBQX to block glutamatergic synapses. **A** Representative electrically evoked inhibitory post-synaptic currents (eIPSCs) recorded from a 5-HT neuron. Stimulus artefact can be seen just before the start of the synaptic current. **B** Voltage dependence of eIPSC amplitude in $N = 3$ 5-HT neurons. Observed reversal potential is $-44.6\,\text{mV}$, consistent with a GABA_A receptor-mediated chloride current under our recording conditions. **C** Fitted decay time constants of eIPSCs from $N = 3$ 5-HT neurons held at 60 mV. **D** Biexponential GABA synapse model used in DRN network models overlaid on eIPSCs from $N = 3$ 5-HT neurons. Model uses $\tau_\text{rise} = 1.44\,\text{ms}$, $\tau_\text{decay} = 26.0\,\text{ms}$, and conductance 0.3 nS (see F). **E** Representative unitary IPSCs recorded in a 5-HT neuron held at 20 mV. **F** Distribution of unitary GABA synaptic conductance in 5-HT neurons extracted from unitary IPSCs. Each curve represents $N = 308$ to 448 unitary IPSCs from a single neuron. IPSC amplitude was converted to units of conductance using $g = \frac{A}{V - V_\text{rev}}$, where g is the synaptic conductance, A is the IPSC amplitude, $V = 20\,\text{mV}$ is the holding potential, and $V_\text{rev} = -44.6\,\text{mV}$ is the reversal potential of the synaptic current (see B).

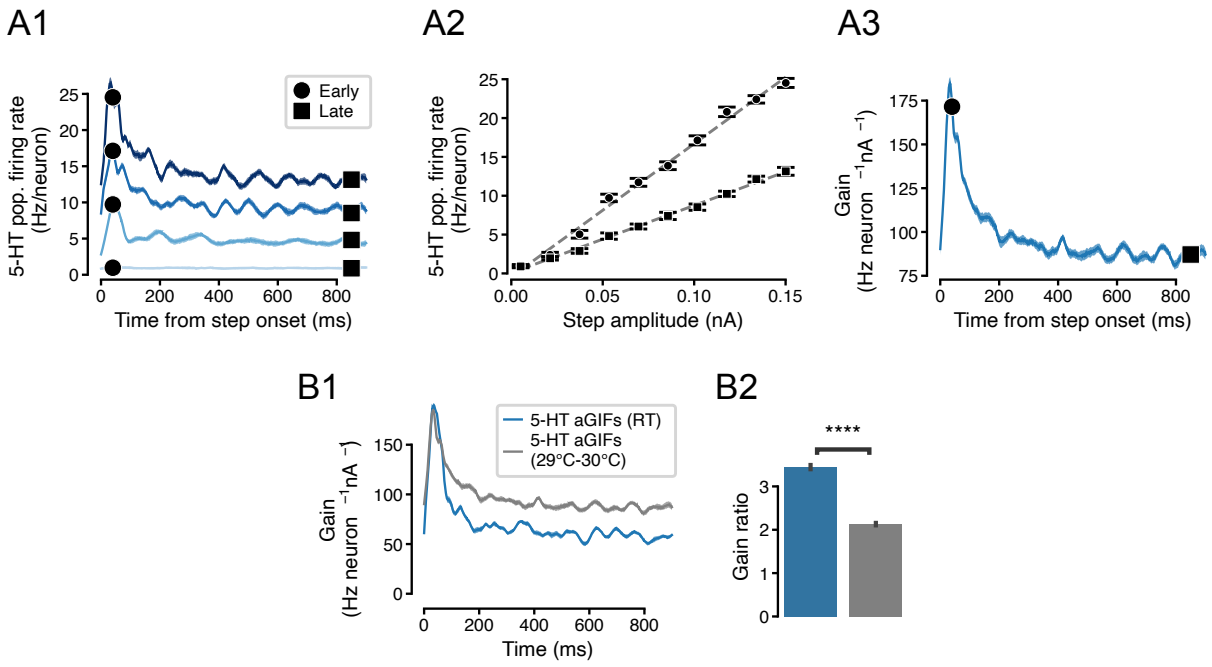


Figure 2.25: Temperature-dependence of simulated 5-HT neuron population gain. **A** Input-output function of a simulated population of 5-HT neurons composed of aGIF models fitted to data collected at 29 °C to 30 °C. Note that the population f/I curves are approximately linear (A2) and that adaptation is apparent in the time-dependent gain curve (A3). Compare with Fig. 2.23A. **B** Temperature-dependence of time-dependent gain. Data from Fig. 2.5E is shown in blue for comparison. $N = 20$ independent simulations.

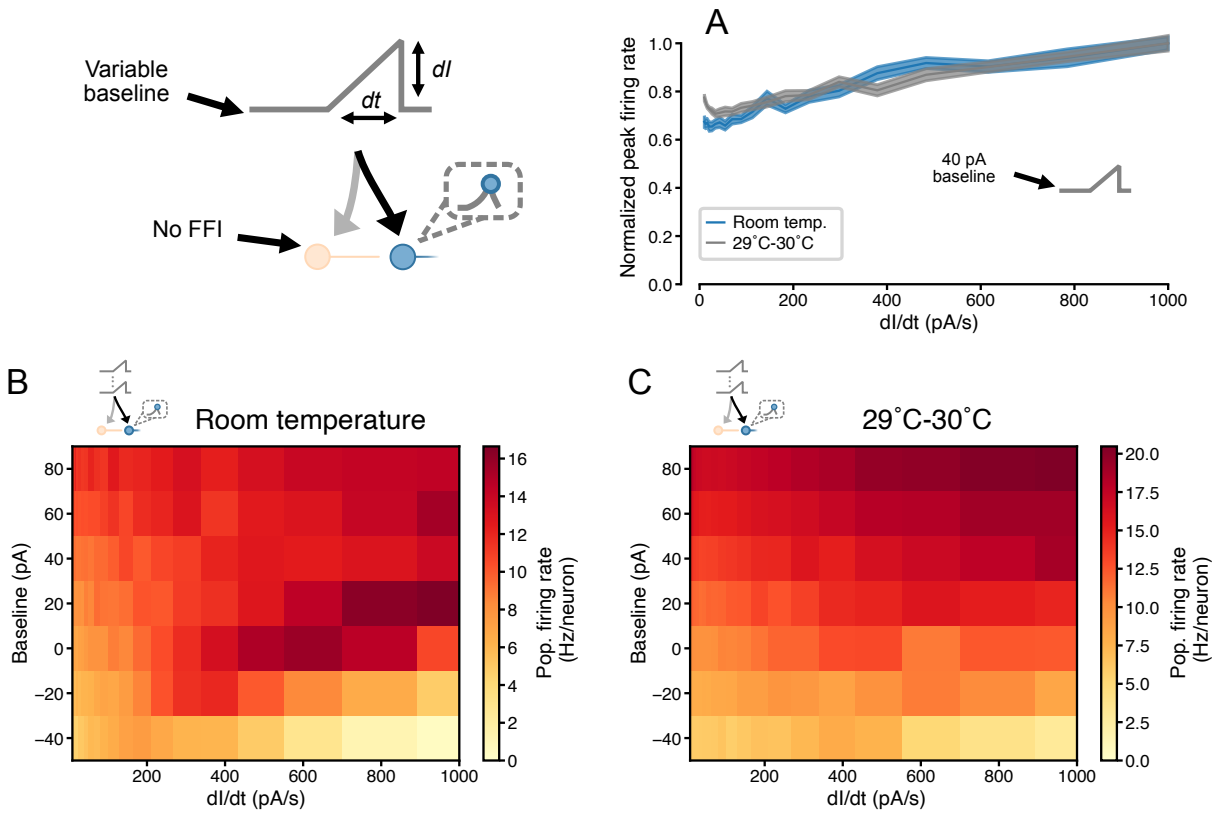


Figure 2.26: Temperature-dependence of temporal derivative encoding by simulated 5-HT neuron populations. **A** When a depolarizing background input is applied, the peak firing rates of simulated 5-HT neuron populations composed of aGIFs fitted to data collected at room temperature and 29 °C to 30 °C linearly encode the slope of a ramp input. Simulation setup is the same as in Fig. 2.7. Data from Fig. 2.7C is presented in blue for comparison. **B**, **C** Relationship between peak population firing rate and ramp stimulus slope across a range of input baselines. Data from the 0 pA and 40 pA rows are presented in A1 and A2, respectively. Darkening colour from left to right across a row indicates a positive relationship between the temporal derivative of the input and population firing rate (a phenomenon we refer to as derivative encoding).

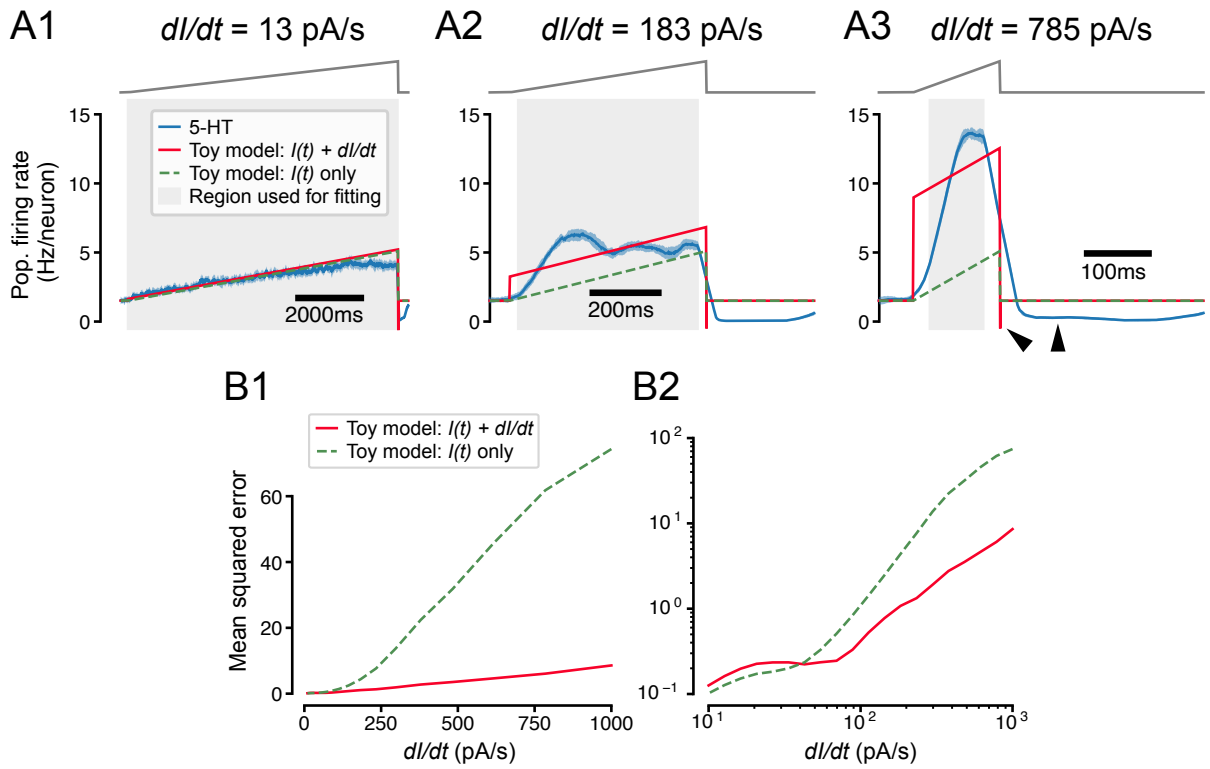


Figure 2.27: Toy model of temporal derivative encoding by the DRN. **A** Toy models that predict the 5-HT population firing rate $r(t)$ (blue) as a function of both the intensity of the applied stimulus and its temporal derivative (red) or the intensity of the stimulus alone (dashed green) make similar predictions for slowly-changing inputs (A1) but very different predictions when the input is increasing quickly (A2, A3). Note that when the input decreases quickly at the end of the ramp ($dI/dt \ll 0$), the population firing rate drops to approximately the $0\,\mathrm{Hz\,neuron^{-1}}$ boundary while the toy model that accounts for the derivative of the input predicts a very negative firing rate and the toy model that ignores the derivative predicts a return to the baseline firing rate (A3, black wedges). **B** Toy models that do and do not account for the temporal derivative of the input predict the 5-HT population firing rate with similar accuracy for slowly-increasing inputs ($0\,\mathrm{pA\,s^{-1}} \leq dI/dt \leq 50\,\mathrm{pA\,s^{-1}}$), but the model that accounts for the temporal derivative makes more accurate predictions for inputs with a larger positive temporal derivative. Note that the performance of both models degrades at larger values of dI/dt, possibly reflecting sensitivity to higher moments of the stimulus or passive membrane filtering not accounted for by our toy models. The toy models are defined as $r(t) = \beta_0 + \beta_1\, I(t) + \beta_2\, dI/dt$, where $\beta_0 = 0.073\,\mathrm{Hz\,neuron^{-1}}$ and $\beta_1 = 0.036\,\mathrm{Hz\,neuron^{-1}\,pA^{-1}}$. β_2 is set to $0.0095\,\mathrm{Hz\,neuron^{-1}\,pA^{-1}}$ in the toy model that accounts for the temporal derivative of the input (red) and $0\,\mathrm{Hz\,neuron^{-1}\,pA^{-1}}$ in the model that does not (dashed green). Coefficients were estimated via a weighted least-squares fit to data from simulations shown in Fig. 2.7 using a stimulus baseline of 40 pA and DRN network models that included FFI. Only data from the ramp portion of the stimulus (excluding 25 ms at the start and end of the ramp to account for smoothing artifacts) was used for fitting and performance evaluation, but the traces in A show that the toy models also accurately predict the baseline firing rate. Weights were set such that all ramps contributed equally to coefficient estimates.

Supplementary methods

Details of the procedure used to fit the parameters in the subthreshold component of the aGIF model described in Harkin *et al.* (2021).

2.5 Definition of aGIF model subthreshold dynamics

The subthreshold dynamics of the aGIF model are given by

$$C\frac{dV}{dt} = -g_l(V(t) - E_l) - \bar{g}_A m_\infty h D_K - \bar{g}_K n_\infty D_K - \sum_{\hat{t}_i<t} \eta(t - \hat{t}_i) + I_{\text{inj}}(t), \tag{2.5}$$

where $V(t)$ is the membrane voltage, $I_{\text{inj}}(t)$ is an externally applied current, $D_K = V(t) - E_K$ is the driving force on potassium, and $\sum_{\hat{t}_i<t} \eta(t - \hat{t}_i)$ is the adaptation current η summated over all past spikes $\{\hat{t}_i \in \mathcal{S} : \hat{t}_i < t\}$ ($\mathcal{S}$ is the set of all spike times).

The adaptation current produced by a single spike $\eta(t - \hat{t}_i)$ is implemented as a sum of k exponentials given by

$$\eta(t - \hat{t}_i) = \begin{cases} \sum_{j=1}^{k} w_j e^{-(t-\hat{t}_i)/\tau_j} & \text{if } t > \hat{t}_i \\ 0 & \text{otherwise,} \end{cases} \tag{2.6}$$

where the timescales $\tau_1, \tau_2, ..., \tau_k$ are treated as hyperparameters. If we substituted this implementation of η back into Eq. (2.5), the term associated with the spike-triggered current $\sum_{\hat{t}_i<t} \eta(t - \hat{t}_i)$ would become $\sum_{\hat{t}_i<t} \sum_{j=1}^{k} w_j e^{-(t-\hat{t}_i)/\tau_j}$. This double sum can be written more

concisely as

$$\sum_{\hat{t}_i<t} \eta(t-\hat{t}_i) = \sum_{j=1}^{k} w_j \hat{\eta}_j(t), \tag{2.7}$$

where $\hat{\eta}_j(t) = \sum_{\hat{t}_i<t} e^{-(t-\hat{t}_i)/\tau_j}$ is a basis for the adaptation current over the timescale τ_j.

Substituting the definition of the adaptation current from Eq. (2.7) into Eq. (2.5), we obtain a detailed definition of the aGIF model subthreshold dynamics as follows:

$$\frac{dV}{dt} = \frac{1}{C}\Big(- g_l(V(t)-E_l) - \bar{g}_A m_\infty h D_K - \bar{g}_K n_\infty D_K - w_1\hat{\eta}_1(t) - \cdots \\ - w_k\hat{\eta}_k(t) + I_{\text{inj}}(t)\Big). \tag{2.8}$$

2.6 Estimating aGIF model subthreshold parameters

Given a training dataset $\mathcal{D} = \{(V(t), I_{\text{inj}}(t)) : 1 \leq t \leq T\}$, knowledge of the equilibrium gating functions $m_\infty, h_\infty,$ and n_∞, and appropriate choices of $\tau_1, \tau_2, ..., \tau_k$ in η, our goal is to estimate the remaining parameters in Eq. (2.8); namely, $g_l, C, E_l, \bar{g}_A, \bar{g}_K, w_1, ..., w_k,$ and τ_h, where τ_h is the time constant of the inactivation gate h. Fortunately, all of these except for τ_h can be estimated easily using linear regression.

We begin by rewriting Eq. (2.8) as the product of a row vector of predictors $\mathbf{x}$ and a column vector of coefficients β as follows

$$\frac{dV}{dt} = \mathbf{x} \cdot \beta \tag{2.9}$$

$$= \begin{bmatrix} V(t) \\ 1 \\ m_\infty h D_K \\ n_\infty D_K \\ \hat{\eta}_1(t) \\ \vdots \\ \hat{\eta}_k(t) \\ I_{\text{inj}}(t) \end{bmatrix}^\top \cdot \begin{bmatrix} -g_l/C \\ g_l E_l/C \\ -\bar{g}_A/C \\ -\bar{g}_K/C \\ -w_1/C \\ \vdots \\ -w_k/C \\ 1/C \end{bmatrix}. \tag{2.10}$$

We solve this subject to $g_l, C, \bar{g}_A, \bar{g}_K \geq 0$ using `scipy.optimize.lsq_linear`.

Next we turn to the question of calculating all of the components of $\mathbf{x}$. Because Eq. (2.8) only reflects the subthreshold dynamics of the aGIF model, we begin by removing all time points in $\mathcal{D}$ within a small window around each spike (from 1.5 ms before each spike until the end of the absolute refractory period). Given the voltages in the cleaned dataset and the set of spike times, it is simple to calculate $m_\infty, h_\infty, n_\infty, D_K$ and $\hat{\eta}_i(t)$. To calculate h for each time point, we order the values of h_∞ according to time and integrate h numerically

using a fixed time step, and the initial condition $h = h_\infty$. This has the effect of assuming that the dynamics of the h gate are paused just before each spike and resumed at the end of the refractory period.

The variance explained by the subthreshold model is a non-convex function of τ_h. We therefore conducted a line search over plausible values of τ_h and chose the value associated with the highest variance explained. This is equivalent to solving

$$\arg\min_{\theta} \left\| \widehat{\frac{dV}{dt}} - \frac{dV}{dt} \right\| \tag{2.11}$$

where $\theta = (\beta, \tau_h)$.

Table 2.3: Single neuron model hyperparameters. For more details on the iGIF model hyperparameter τ_θ, see Mensi *et al.* (2016).

Model	Parameter	Symbol	Cell type	Value (ms)
All	η timescales	$\tau_1, \tau_2, ..., \tau_k$	All	3, 10, 30, 100, 300, 1000, 3000
All	γ timescales	None	All	3, 30, 300, 3000
All	Refractory period	None	5-HT SOM & mPFC	6.5 4.0
iGIF	Candidate threshold-coupling timescales	τ_θ	All	1, 2, 5, 10, 22, 46, 100
aGIF	Candidate inactivation timescales	τ_h	All	10, 13, 18, 25, 33, 45, 61, 82, 111, 150

Chapter 3

Serotonin predictively encodes value

> What I cannot create, I do not understand.
>
> Richard Feynman

This chapter is reprinted from a previously-published work.

Citation

E.F. Harkin, C.D. Grossman, J.Y. Cohen, J.-C. Béïque, and R. Naud. Serotonin predictively encodes value. *bioRxiv*. (2023)

Copyright authorization

Statement of contributions

I conceptualized the project, created the model, performed all simulations and data analysis, performed mathematical analysis, created all figures, and wrote the first and final drafts of the manuscript as well as all drafts of the discussion section.

Richard Naud provided supervision, funding, and extensive input on all aspects of the project as well as performing mathematical analysis and writing the intermediate draft. Jean-Claude Béïque provided extensive input and helpful discussion throughout the project, as well as detailed comments on the manuscript. Cooper Grossman and Jeremiah Cohen provided data and validated the design of our analysis. Cooper Grossman provided helpful discussion and extensive input on comparisons with the uncertainty model, significantly strengthening this aspect of the work. Jeremiah Cohen provided helpful discussion and detailed input on the manuscript.

Abstract

The *in vivo* responses of dorsal raphe nucleus (DRN) serotonin neurons to emotionally-salient stimuli are a puzzle. Existing theories centred on reward, surprise, or uncertainty individually account for some aspects of serotonergic activity but not others. Here we find a unifying perspective in a biologically-constrained predictive code for cumulative future reward, a quantity called state value in reinforcement learning. Through simulations of trace conditioning experiments common in the serotonin literature, we show that our theory, called value prediction, intuitively explains phasic activation by both rewards and punishments, preference for surprising rewards but absence of a corresponding preference for punishments, and contextual modulation of tonic firing—observations that currently form the basis of many and varied serotonergic theories. Next, we re-analyzed data from a recent experiment and found serotonin neurons with activity patterns that are a surprisingly close match: our theory predicts the marginal effect of reward history on population activity with a precision $\ll 0.1\,\mathrm{Hz\,neuron^{-1}}$. Finally, we directly compared against quantitative formulations of existing ideas and found that our theory best explains both within-trial activity dynamics and trial-to-trial modulations, offering performance usually several times better than the closest alternative. Overall, our results show that previous models are not wrong, but incomplete, and that reward, surprise, salience, and uncertainty are simply different faces of a predictively-encoded value signal. By unifying previous theories, our work represents an important step towards understanding the potentially heterogeneous computational roles of serotonin in learning, behaviour, and beyond.

3.1 Introduction

What do the activity patterns of serotonin neurons encode? Over a quarter-century ago, Schultz, Dayan, and Montague [70] persuasively argued that the phasic activity of dopamine neurons might encode the reward prediction errors (RPEs) of reinforcement learning (RL) theory [64]. Given the deep connections between the dopamine and serotonin systems, both of which are neuromodulatory systems with important and well-studied roles in regulating mood, learning, and behaviour [233], it is surprising that no single account of the responses of serotonin neurons enjoys a similar level of support.

There are several possible reasons for this lack of consensus. One possibility is that the serotonin system is not a monolith, but rather a heterogeneous collection of partially-overlapping sub-systems with diverse coding features [33, 38, 39]. Another possibility, in no way mutually-exclusive, arises from the fact that experimental and theoretical work in the serotonin field, including our own, has been deeply shaped by the potentially incorrect assumption that the activity patterns of serotonin neurons can be divided into phasic and tonic components that reflect essentially unrelated quantities [73, 154, 158, 233, 234]. This separation of timescales is reflected in the currently fragmented picture of the dominant tuning features of serotonin neurons. Rejecting this assumption could lead to more clarity about serotonergic function within and across raphe sub-systems.

Over the past decade, detailed experimental characterizations of the diverse *in vivo* responses of genetically-identified serotonin neurons to emotionally-salient stimuli have revealed some common themes in their tuning features, even if these patterns remain difficult to interpret. In trace conditioning experiments, the activity patterns of serotonin neurons are dominated by phasic bumps, plateaus, or ramps preceding expected rewards that emerge over the course of learning and diminish during reversals (Fig. 3.1A; 24, 39, 57, 58), modulation of tonic activity by reward or punishment context (Fig. 3.1B; 24), a phasic preference for unpredicted over predicted rewards (Fig. 3.1C; 24, 57), and phasic activation by punishments

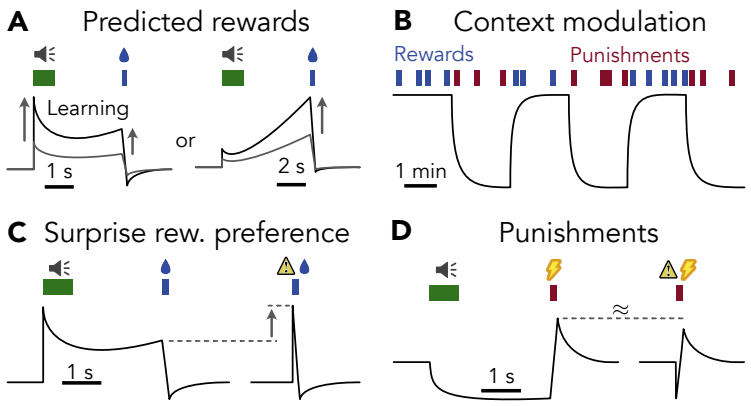


Figure 3.1: Summary of qualitative tuning features of serotonin neurons captured by predictive value coding model. Curves indicate the activity of serotonin neurons over time, measured either as firing rate [24] or calcium fluorescence [39, 57, 58]. **A** Phasic activation by predicted rewards over short timescales [24] emerges gradually during learning [39, 57, 58]. Depending on the experiment, activity takes the form of a phasic cue-associated peak followed by a plateau (left; 24, 57), or a ramp leading up to reward (right; [58]). **B** Tonic activity modulated by reward or punishment context over long timescales [24]. **C** Stronger phasic activation by unpredicted (right) compared with predicted (left) rewards [24, 57]. **D** Phasic activation by punishments whether predicted (left) or not (right) [24, 39, 57].

whether predicted or not (Fig. 3.1D; 24, 39, 57). To explain aspects of these observations, serotonin neurons have been proposed to encode current or future reward [Fig. 3.1A [61] or, separately, B [24], but not D], surprise [Fig. 3.1C but not D [57]], or salience [Figure 3.1D but not C [39]]. The reward [61], surprise [57], uncertainty [25], and salience [39] theories do not offer detailed predictions about the dynamics of serotonin neuron activity, nor can any of them individually account for all of their dominant tuning features (Table 3.1). Other serotonergic theories related to persistence [56], confidence [62], learning rate [87], and discounting [74, 224] focus on explaining the effects of serotonergic manipulations on behaviour and do not connect directly to the naturalistic tuning features of these cells (but see 25). Even the best established tuning features of serotonin neurons therefore lack a consistent interpretation.

Here we argue that existing qualitative serotonergic theories are incomplete, not incorrect, and that reward, surprise, salience, and uncertainty are simply different aspects of a single quantity encoded in the activity patterns of serotonin neurons. To formulate a consistent interpretation of the dominant reward and punishment tuning features of serotonin neurons

outlined above, we combine top-down ideas from theories of RL [64] and predictive coding [65, 122] with recent bottom-up insights into the computational features of the DRN [158]. We hypothesized that serotonin neurons predictively encode a weighted average of future reward, a quantity referred to as state value in RL, via the dominant biophysical feature of this cell type: exceptionally strong and long-lasting spike frequency adaptation. We formalize this hypothesis in a quantitative model that we refer to as as the value prediction theory of serotonin.

To test our value prediction theory, we simulate trace conditioning experiments common in the serotonin literature [24, 25, 57, 58] and show that our model provides a consistent account of the main established tuning features of these cells. We also interpolate and extend previous results, providing intuitive mechanistic connections between seemingly unrelated observations, resolving apparent conflicts in the literature, and making experimentally testable predictions. Next, we re-analyze a recently-published dataset from a trace conditioning experiment [235], finding activity patterns consistent with our hypothesis. Finally, to counter our own confirmation bias, we explicitly compare against quantitative formulations of previous theories and find that value prediction best explains the data by a large margin. Our theoretical and empirical results reveal a surprisingly precise quantitative code for value in the serotonin system.

3.2 Results

3.2.1 Predictive encoding of value signals

Reinforcement learning describes the process by which an agent learns a policy for controlling the state of its environment S, s in order to maximize reward R, r (Figure 3.2A; 64). For example, a mouse learning which lever to press to obtain a food pellet in an operant conditioning experiment. RL conceptualizes the reward estimate as a mapping from states to future rewards referred to as a value function $v(s)$ (Figure 3.2B). For simplicity, here we focus

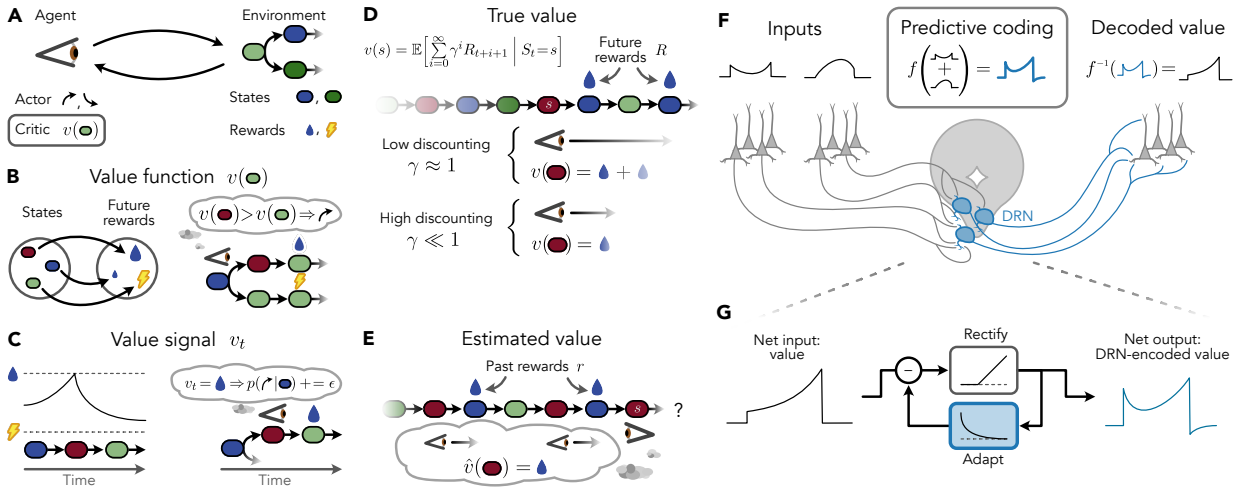


Figure 3.2: Computing future reward. **A** High-level overview of reinforcement learning (RL; 64). **B** A value function $v(s)$ is a mapping from states to future rewards. The value function can be used to drive decision-making (right): by comparing the values of future states, the agent can make choices that lead to rewards [92, 97]. **C** A value signal v_t is the result of evaluating the value function $v(s)$ at the current state s_t over time. The value signal can be used to drive learning (right): by increasing the probability of taking an action in proportion to the value signal just after the action is taken, the agent can learn to take actions that lead to rewards [93]. **D** Normative definition of a value function as the expected sum of discounted future rewards, referred to as the true value $v(s)$. **E** The estimated value function $\hat{v}(s)$ approximates the true value on the basis of past rewards. **F** Predictive value coding model. The dorsal raphe nucleus (DRN) receives a distributed value signal as input, summates its components, and predictively-encodes the result. (Note that although a distributed value code is illustrated, similar to Fig. 2 in ref. 70, it is also possible that the value signal originates in a single upstream region.) The predictively-encoded signal consists of a mixture of the original value signal and its time derivative $f(v(t)) \approx \alpha \frac{dv}{dt} + v(t)$, a transformation implemented by strong spike-frequency adaptation in serotonin neurons [158]. Predictive coding can easily be reversed via leaky integration in downstream regions to recover the original value signal (Section 3.13). **G** Adaptation-based predictive coding model. See Methods.

on the state value $v(s)$ which is equivalent to the average $q(s, a)$ value of actions a available in state s (Section 3.6).

Value functions have been central to RL since the very beginning [97]. One particularly well-known use of value functions is to compare the estimated rewards associated with different hypothetical courses of action (for example, pressing different levers) in order to select the action most likely to lead to the greatest reward (Figure 3.2B right; 236, 237). A lesser-known application is the evaluation of the current state s_t over time, yielding what we call a value *signal* $v_t = v(s_t)$ (Figure 3.2C). Such a value signal is time dependent because the state is continually changing, a feature exploited by temporal difference learning to gradually refine the estimated values of past states [238, 239]. Apart from value learning, value signals can be used to directly reinforce recent actions [93, 94] or promote persistence [236, 240]. Here we present evidence of a close match between the activity patterns of serotonin neurons and value, leaving the question of how value might be used to drive learning and behaviour for future work.

Specifically, we focus on a value signal defined as the expected total discounted future reward in the present state

$$v_t \equiv \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1} \,\middle|\, S_t = s_t\right], \tag{3.1}$$

where R_t is the random reward obtained at time t, $0 \leq \gamma \leq 1$ is a discrete time discounting factor that controls the relative weighting of imminent and distant rewards (imminent rewards are weighted more heavily when γ is closer to zero), and $\mathbb{E}[X|Y = y]$ denotes conditional expectation. Intuitively, it represents a weighted average of future rewards, with closer rewards being weighted more heavily depending on the degree of discounting (Figure 3.2D). The size of the window within which future rewards are summed to calculate the value can be quantified with the discounting timescale τ, defined as $\tau = -dt/\ln\gamma$, where dt is the duration of a discrete time step.

The above definition of a value signal in terms of future reward is precise, extremely general (Section 3.6), and conceptually simple, but unrealistic: animals do not generally have perfect knowledge of future rewards. We therefore distinguish between signals calculated on the basis of future rewards, which we refer to as *true* value (Figure 3.2D), and more realistic ones learned from past experience, which we refer to as *estimated* value (Figure 3.2E).

We propose that the firing rates of serotonin neurons present a predictive code for an estimated value signal (Figure 3.2F). Recently, we showed that potent spike-frequency adapation dominates the signal processing features of the DRN [158]. This removes the part of the signal that is similar to past output, a redundancy-reduction scheme sometimes called predictive coding (65, 122, see 123 for review). In a simplification of this previous work, here we model the firing rate output of the DRN as

$$\rho_t = \mathrm{ReLU}\left[(1 + A)\, v_t - A\, u_t\right],$$

where $\mathrm{ReLU}[x]$ is the rectified linear function used to ensure the firing rate is non-negative, A is a parameter controling the strength of adaptation, and u_t is the adaptation variable, which has a subtractive effect on the output firing rate. We model adaptation as an exponential moving average of past activity (Figure 3.2G and methods). To build an intuition for this model, consider that, from a computational perspective, adaptation can be seen as implementing temporal differentiation [126, 158]. As a consequence, the firing rates of serotonin neurons reflect a mixture of value and its rate of change.

Our main result is that this predictive coding process induces a qualitative change in the value signal. Adaptation has the effect of exaggerating sharp transients, often leading the encoded signal to over- or under-shoot its apparent target (illustrated schematically in Figure 3.2G, simulations in the next section), and hiding the connection between serotonergic activity patterns and value.

3.2.2 Value prediction during trace conditioning

To examine the temporal evolution of this signal in an experimental setup common in the serotonin literature, we simulated our model under trace conditioning. Trace conditioning experiments consist of a series of trials that begin with a sensory cue (e.g., an auditory tone or an odour) and end in a reward (typically a drop of water) with a short delay ($\sim$2 s) separating the two [24, 57, 58]. In this experimental paradigm, true value signals take four distinct phases (see Methods): 1) jumping to a higher value upon receiving the cue since the cue indicates a reward is coming, 2) ramping upward between the time of the sensory cue and the reward delivery due to the effects of time discounting, 3) falling during the reward epoch as the *future* reward left to collect disappears, and 4) staying at a constant non-zero value during the inter-trial interval as the animal waits for the next randomly-timed trial to begin (black line in Figure 3.3A).

These four phases are altered by predictive coding, especially phase 2) where the ramping upward is preceded by the adaptation from the cue-triggered jump and phase 3) where the return to baseline is accompanied by an undershoot (blue line in Figure 3.3A). Multiple research groups have shown that serotonin neurons are transiently activated by reward-predicting cues *in vivo* (24, 57, 58; *e.g.*, Figure 3.3B inset; schematized in Figure 3.1A). Previous value and reward-based serotonergic theories cannot explain this phasic activity (black line in Figure 3.3A), leading to proposals that serotonin might encode some other quantity (*e.g.* surprise 57). In value prediction, phasic cue-associated firing emerges naturally as a result of adaptation (blue line in Figure 3.3A). Unlike previous models, our theory further predicts a subtle, counter-intuitive drop in activity during the reward epoch to complement phasic activation by the cue. Interestingly, this phenomenon is visible in raw experimental data presented in the literature [24, 57, 58], but is generally not quantified. In short, value prediction through adaptation explains why serotonin neurons are phasically activated by reward-predicting cues (Figure 3.1A) and predicts that serotonin neurons should exhibit

decreasing/below baseline activity during reward consumption.

3.2.3 Value prediction captures response to punishments

A significant problem for value and reward-based serotonergic theories is that serotonin neurons are often activated by both rewards and punishments [24, 39, 57]. Since punishments can be seen as negative rewards $R_t < 0$, and value represents an estimate of future reward (Figure 3.2D), then serotonin neurons should be inhibited by punishments — not activated — if they encode a simple value signal (black line in Figure 3.3B).

In contrast, activation by punishments is expected under the value prediction theory. This is because predictive coding through adaptation (Figure 3.2G) creates an overshoot in the level of activity as the punishment ends (blue line in Figure 3.3B). To understand why this happens, recall that the effect of predictive coding through adaptation is to exaggerate positive (and negative) transients in the underlying value signal. The value signal during a punishment trial is the mirror image of the value during a reward trial (black lines in Figure 3.3A and B), increasing as the punishment ends just as the reward trial value signal decreases when the reward is consumed (3 s to 4 s post-cue in Figure 3.3A and B). Through predictive coding, the fast increase in value during the punishment epoch is enhanced, causing the encoded signal to briefly overshoot its baseline (∼4 s post-cue in Figure 3.3B).

The idea that predictively encoding a value signal creates a punishment withdrawal-induced overshoot explains 1) why serotonin neurons are activated by punishments as well as rewards [24, 39, 57], 2) why these seemingly opposite response features are positively correlated across cells [24], and 3) why this activation occurs at the end of a punishment rather than the beginning [38].

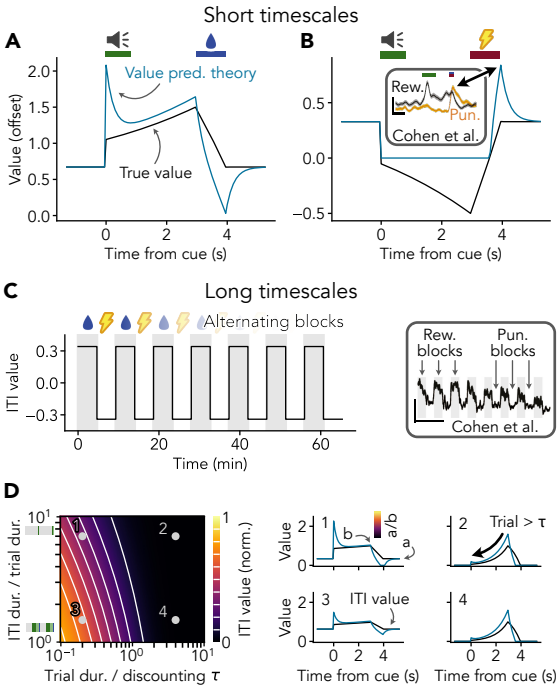


Figure 3.3: Value prediction signals reward and punishment over multiple timescales. **A**,**B** True value signals (black) and their predictively-encoded counterparts (blue) for trace conditioning trials terminating in either a reward (A) or punishment (B). Signals are shifted up by 0.5 AU to capture background firing. Note resemblance between value prediction theory and firing rate of a genetically-identified serotonin neuron from Cohen *et al.* [24] (B inset; modified from ref. Fig. 3A2; scale bar 5 Hz, 1 s). **C** ITI value reflects reward or punishment context. Simulated block length of 10 trials, normalized reward and punishment sizes of 1 and −1, respectively, and all other parameters as in D1 below. Note resemblance with the tonic firing rate of a serotonin neuron (right; modified from ref. 24 Fig. 3B1; scale bar 3 Hz, 10 min). **D** Analytically-derived true value of the inter-trial interval (ITI) is proportional to peak within-trial value. Heatmap shows ITI value as a function of experimental design parameters (mean ITI duration / trial duration; vertical axis; ribbons are to scale, gray represents ITI duration and colours represent trial epochs) and agent parameters (trial duration / discounting τ; horizontal axis). ITI value is presented as a fraction of the peak value during the trial (*i.e.*, value just before reward). Numbered panels at right illustrate the within-trial dynamics of the true (black) and predictively-encoded (blue) value signals for various combinations of experimental and agent parameters indicated on the heatmap. Note that predictive coding has no effect on ITI value because the time derivative of the value signal during the ITI is zero. Since value signals are normalized, different reward sizes can be accommodated by scaling traces. Trial structure same as in A. See Figure 3.8 for an extended range of trial durations and discounting timescales.

3.2.4 Tonic firing during inter-trial intervals reflects reward in future trials

The phasic responses of serotonin neurons to rewards and punishments (Figure 3.1A and D) have historically been difficult to reconcile with tonic activity that tracks reward and punishment context (Figure 3.1B; 24, 76), spawning proposals that serotonin neurons may multiplex unrelated quantities over short intra-trial and long inter-trial timescales [24, 25]. Value prediction explains these phasic responses (see above) while also predicting that tonic activity should track reward and punishment context (Figure 3.3C, Section 3.10), unifying the responses of serotonin neurons to rewards and punishments over short and long timescales.

More interestingly, our theory predicts that trial duration should have pronounced effects on both inter-trial value coding and within-trial activity dynamics of serotonin neurons. Analysis of our model shows that the proportionality between inter-trial and within-trial value depends on two factors: 1) the mean duration of the ITI relative to the trial duration (vertical axis in Figure 3.3D) and 2) the duration of the trial (defined as the time between cue onset and reward delivery) relative to the discounting timescale of the animal τ (horizontal axis in Figure 3.3D). However, while the effect of ITI duration is surprisingly weak in the typical experimental range (*i.e.*, ITIs two to five times the trial duration; 24, 57, 58), the effect of trial duration is quantitatively large and visually obvious. Specifically, when the trial duration is shorter than the discounting timescale, we expect to see both phasic cue-associated activity (Figure 3.3D1 and 3) and inter-trial value coding, while when the trial duration is longer than the discounting timescale, we expect ramping within-trial activity (Figure 3.3D2 and 4) and little to no inter-trial value coding. The transition between these two regimes is sharp and occurs when the trial duration is roughly equal to the discounting timescale. Thus, the ratio between the trial duration and discounting timescale controls both inter-trial value coding and within-trial "peak and plateau" vs. ramping activity dynamics.

The effect of trial duration on within-trial activity dynamics and inter-trial value coding

predicted by our model explains 1) why "peak and plateau" dynamics and tonic value coding co-occur [24], 2) why experiments using longer trials sometimes produce ramping rather than "peak and plateau" dynamics (compare 24 and 58; schematized in Figure 3.1A; but see 57), and 3) why firing during the pre-reward epoch may decrease slightly as the trace duration increases [241].

3.2.5 Value prediction explains reward-specific surprise

While little is known about the effect of trial duration on the pre-reward activity dynamics of serotonin neurons, the effect of trial duration on the amplitude of the reward (or punishment) response itself has received more attention. Compared with rewards delivered at the end of a trace conditioning trial, serotonin neurons are more strongly activated by rewards delivered spontaneously [24] or immediately following a cue [57]. This has been interpreted as evidence for surprise coding [57], defined as activity that reflects an unsigned reward prediction error $|\delta_t| = |R_t - v_t|$, which is believed to be important for learning [25, 57, 242]. However, because the surprise/absolute RPE-like reward responses of serotonin neurons do not evolve on the same timescales as dopaminergic RPEs δ_t [57], are smaller than corresponding dopaminergic responses [24], and seem to be specific to rewards (Figure 3.1C and D; 24, 57), a different explanation is needed.

To understand how surprise tuning for rewards might emerge, we simulated value prediction under progressively shorter trace conditioning trials. As the trial duration shortened, the adaptation-induced cue-associated peak began to overlap with the response to the reward itself (compare Figure 3.4A1 and 2). This phenomenon becomes increasingly pronounced as the trial duration falls below the effective timescale of adaptation (on the order of hundreds of milliseconds; scan from left to right along the top of Figure 3.4B corresponding to the green line in Figure 3.4C), and is maximally strong when the trial duration reaches the zero lower bound, corresponding to an uncued reward (Figure 3.4A3).

It is difficult to differentiate absolute RPE from value prediction on the basis of reward

responses alone because both theories predict stronger responses for surprising rewards (schematized in Figure 3.1C). To rule out absolute RPE, we turn our attention to serotonergic responses to punishments. Whereas the absolute RPE theory predicts that serotonin neurons should respond most strongly to surprising punishments [57], just as they do for rewards, serotonin neurons should have a very slight preference for predicted punishments under value prediction (Figure 3.4A4–6, left to right along the bottom of Figure 3.4B, red line in Figure 3.4C), consistent with experimental observations (57; schematized in Figure 3.1D).

To understand why value prediction implies reward-specific surprise tuning, recall that punishment-associated activity is caused by punishment *withdrawal* under our theory (Figure 3.3B). Shortening the trial duration has no effect on the rate of punishment withdrawal, and even causes the pre-punishment inhibition to slightly overlap the withdrawal-associated peak if the trial duration is sufficiently short, leading to a small decrease in the punishment response (Figure 3.4A4–6). The transition from surprise tuning to lack thereof occurs sharply when the size of the reward passes below zero, but this is obscured by the relatively small responses to near-zero rewards in our model (Figure 3.4B). As with rewards, the transition to a slight preference for unsurprising punishments occurs when the trial duration drops below the effective timescale of adaptation, such that reward and punishment responses are expected to diverge markedly for trial durations on the order of hundreds of milliseconds or less (Figure 3.4C).

These simulations show that value prediction explains surprise tuning for rewards that reverses for punishments, thus providing a more complete account of serotonergic surprise tuning than the existing absolute RPE theory [57].

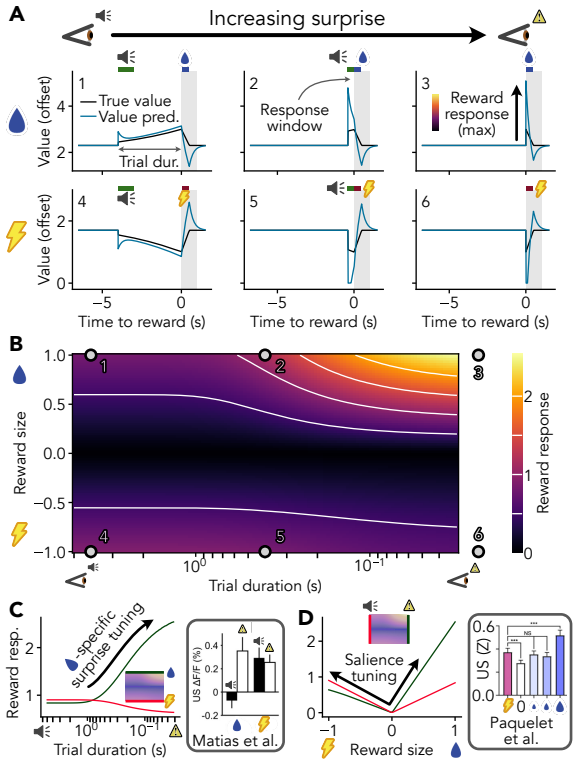


Figure 3.4: Predictively-encoded value resembles surprise and salience. **A**,**B** Reward responses depend on reward size and cue timing. Reward response is defined as the baseline-subtracted maximum DRN-encoded value signal within 1 s of cue onset (gray window). Trial duration is defined as the time between the onset of cue and reward (gray arrow in 1). Sample traces 3 and 6 represent uncued rewards and punishments, respectively. Signals are shifted up by 2 AU to capture decreased activity during punishment trials. **C** Predictively-encoded value yields larger reward responses for uncued vs. cued rewards but similar responses to uncued and cued punishments. Note resemblance to reward-specific surprise coding from Matias *et al.* [57] (inset modified from ref. Fig. 7B; punishment type: air puff). **D** Apparent salience coding is distinct from surprise. Whether cued or uncued, reward responses extracted from predictively-encoded value signals are smaller for neutral outcomes than both rewards and punishments. Note resemblance to salience coding from Paquelet *et al.* [39] (inset modified from ref. Fig. 2D with kind permission from Bradley Miller; punishment type: bitter quinine solution, reward type: sucrose solutions). Green and red lines in B and C are slices of data from A as indicated on mini-heatmaps.

Table 3.1: Short timescale reward tuning features of genetically-identified mouse DRN serotonin neurons qualitatively explained by various theories. ✓ and ✗ indicate empirical observations that are clearly consistent or inconsistent with each theory; ambiguous cases are left blank. Ambiguity is due to a lack of quantitative models to accompany the current and future reward, salience, and surprise theories, as well as variation in experimental design to a lesser extent. Note that refs. [58, 61, 174, 243] focus on a signal that qualitatively reflects both current and future reward. This signal is referred to as reward or beneficialness by the authors, but is most similar to a value signal in our terminology.

Theory	**Tonic firing tracks rew. rate**	**Rew. cue activation** [24, 25, 57, 58, 174, 243]	**Rew. delivery activation** [24, 38, 39, 57, 58, 174, 243]	**Pun. activation** [24, 38, 39, 57]	**Correlated rew. cue and pun. activation** [24]	**Surprise rew. preference** [24, 57]	**Lack of surprise pun. preference** [24, 57]
Value prediction (ours)	✓(Fig. 3.3C)	✓(Fig. 3.3A)	✓(Fig. 3.3A)	✓(Fig. 3.3B)	✓(Fig. 3.4C)	✓(Fig. 3.4B)	✓(Fig. 3.4B)
Current and future reward [58, 61, 174, 243]	✓	✓	✓	✗	✗	✗	✗
Salience [39]	✗	✓	✓	✓		✗	✓
Surprise [57]	✗	✓	✓	✓		✓	✗
Dopamine opponent [73]	✓	✗	✗	✓	✗	✗	✗

3.2.6 Value prediction explains salience tuning for both surprising and unsurprising stimuli

If surprise is defined in the serotonin literature as absolute RPE $|\delta_t|$, then salience is defined as the absolute size of the reward itself $|R_t|$. We have already shown that value prediction explains serotonergic responses to cued rewards and punishments (Figure 3.3A and B), a type of salience tuning, but recent experimental work has focused on this phenomenon in the context of *uncued* rewards and punishments [39].

To show that value prediction explains salience tuning for both uncued and cued rewards, we simulated trace conditioning experiments using a wide range of trial durations and reward sizes. We find stronger responses to both rewards and punishments compared with neutral outcomes across all trial durations, consistent with salience tuning (Figure 3.4B). Interestingly, the salience tuning effect is fairly balanced for cued rewards and punishments ("V"-shaped red line in Figure 3.4D), whereas a clear preference for rewards emerges in very short trials ("✓"-shaped green line in Figure 3.4D). This reward preference can be explained by the interaction of reward-specific surprise tuning (Figure 3.4C) with salience. Thus, evidence for existing salience and surprise theories also supports value prediction.

3.2.7 Slow online learning

So far we have focused on the resemblance between predictively-encoded true value signals and the *in vivo* activity patterns of serotonin neurons. However, it is unrealistic to think that serotonin neurons signal true value, since this would require perfect knowledge of future rewards (Figure 3.2D). Instead, serotonin neurons likely encode a value signal that is estimated on the basis of past rewards (Figure 3.2E).

Does our focus on true rather than estimated value pose a problem for the results presented above? To find out, we applied an online value estimation algorithm (van Seijen's TD(λ) [96], see Methods) to a trace conditioning experiment consisting of hundreds of trials. The

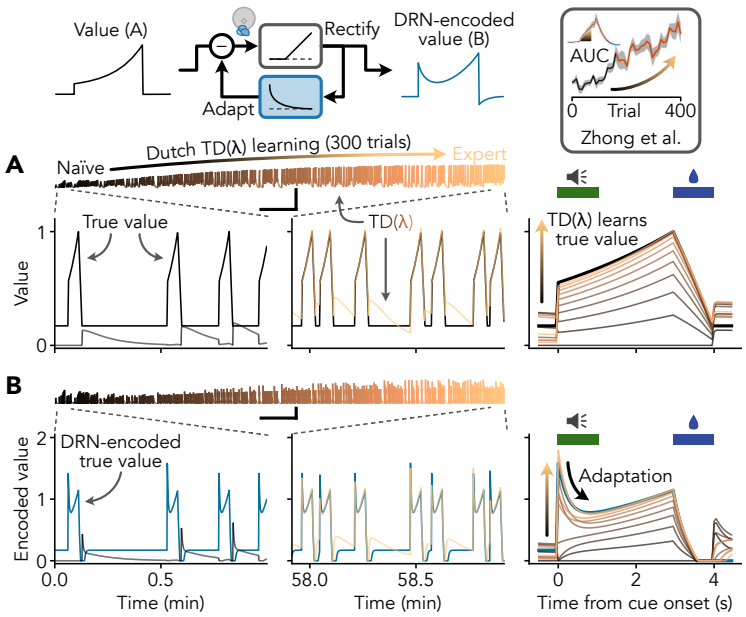


Figure 3.5: Online value estimation. **A** Estimation of value during 300 trials of trace conditioning. Comparison between true value v_t (black) and value estimated using TD(λ) with Dutch eligibility traces (shades of copper). Note that the estimated value signal converges to a close approximation of the true value (right). **B** Value signals predictively encoded by the DRN. Scale bars: 5 min, 1 arbitrary value unit. Vignette above A modified from ref. Fig. 1G of Zhong *et al.* [58].

estimated value signal exhibited a ramp that gradually increased in amplitude, gradually converging to a close approximation of the true value (Figure 3.5A), mirroring observed activity of serotonin neurons in mice [57, 58]. The same was true of the predictively-encoded estimated value (Figure 3.5B). Overall, estimated value signals resembled a true value template rescaled by reward history. These simulations illustrate that the details of how the value signal is calculated (*i.e.* on the basis of future rewards, as in true value Figure 3.2D, or on the basis of past rewards, as in estimated value Figure 3.2E) play only a minor role in shaping the activity patterns of serotonin neurons during trace conditioning, whereas predictive coding and reward history are critical.

If the activity levels of serotonin neurons are scaled by reward history, how large should this effect be? The rate at which the TD(λ)-estimated value signal is scaled up and down by reward history depends on the learning rate α, which can also be expressed as the estimation timescale $\tau_{\text{est}} = -dt/\ln(1-\alpha)$. Previous work suggests that the timescale over which serotonin

neurons integrate rewards is on the order of hundreds of trials [57, 58]. Since the estimation timescale is so long, it is possible to write a first-order approximation of the effect of an uninterrupted string of rewards (or reward omissions) on the firing rates of serotonin neurons (Section 3.12. Using an estimation timescale of $\tau_{\text{est}} = 200$ trials [57], background firing rate of 2 Hz [24, 25], and winning streak of five trials, we expect to see firing rate modulations of only $\sim$0.05 Hz. While longer runs of rewards (or reward omissions) would produce larger effects (*e.g.*, 0.4 Hz for a run of 50 trials), such winning streaks (or losing streaks) are rare in experiments with probabilistic rewards that are best suited to studying the effects of reward history on serotonergic activity [25]. In short, if serotonin neurons encode a value signal that is estimated over a long period of time, as evidence suggests [57, 58], we expect to see only small effects of reward history on serotonin neuron firing ($\ll$1 Hz) in typical experiments.

3.2.8 Serotonin neurons quantitatively encode reward history

Value prediction unifies a wide range of experimental observations that do not have a consistent interpretation under existing serotonergic theories (Figure 3.1, Table 3.1). To assess whether value prediction generalizes beyond the main qualitative results of Cohen *et al.* [24], Zhong *et al.* [58], and Matias *et al.* [57], we re-analyzed a dataset of serotonergic responses to dynamically-varying *in vivo* rewards from Grossman *et al.* [25]. To begin, we sought to determine whether the activity levels of serotonin neurons in this dataset are weakly modulated by reward history, and, in particular, whether this is true of neurons with activity dynamics resembling a predictively-encoded value signal.

A short description of the experiment and our analysis approach follows (see Methods for details). The dataset consists of tetrode recordings of $N = 37$ identified serotonin neurons in mice receiving dynamically-varying probabilistic rewards in a trace conditioning experiment [25, 235]. We extracted the spikes of each neuron in a short window around each trial along with the proportion of the past five trials that were rewarded (Figure 3.6A). Since true value and TD value estimates are both very closely tied to the mean reward (Section 3.7), we use

the five-trial mean reward as an interpretable proxy for value. Because the effect of reward history on serotonin neuron activity is expected to be small (see previous section) and many common statistical tests for value coding are prone to high rates of false positives [244], we used circular trial permutation (Figure 3.6B). This test breaks the putative association between reward history and serotonin neuron activity while preserving all other structure in the data (for example, slow fluctuations related to arousal). If the association between serotonin neuron activity and reward history is stronger when the trials are correctly aligned than when this alignment is broken, we can conclude that the correlation between serotonin neuron activity and reward history cannot easily be explained by random fluctuations in the data.

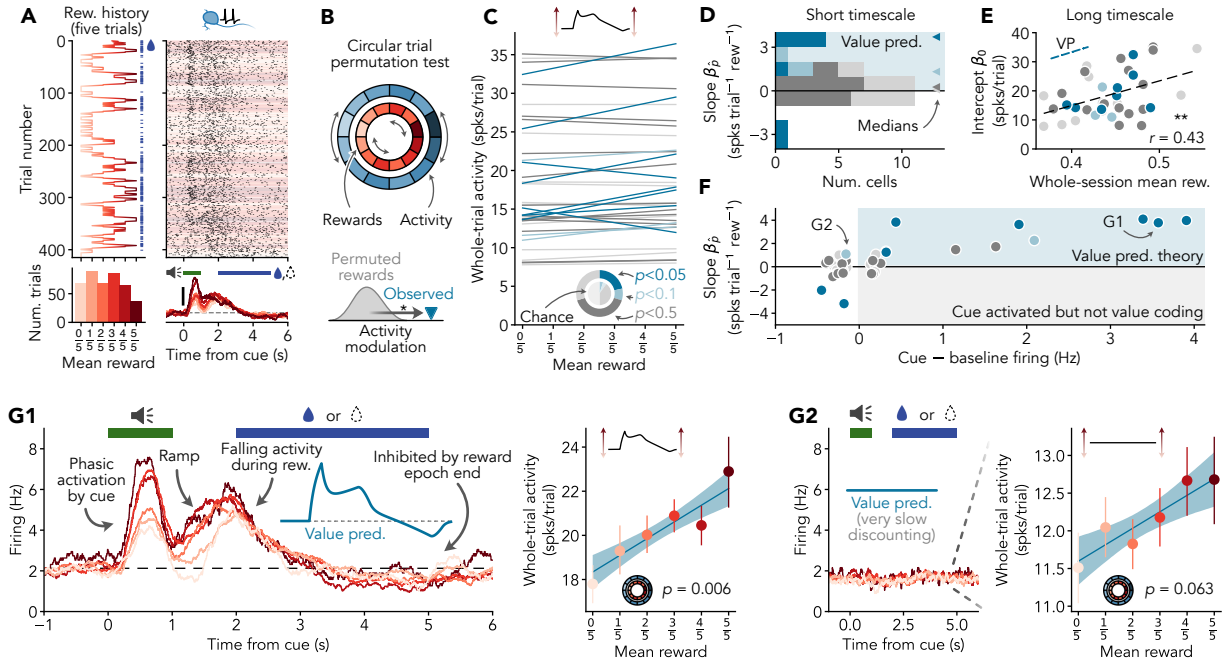

A
Rew. history (five trials)
Trial number
Num. trials
Mean reward
Time from cue (s)
B
Circular trial permutation test
Rewards
Activity
Permuted rewards
Observed
Activity modulation
C
Whole-trial activity (spks/trial)
p<0.05
p<0.1
p<0.5
Chance
Mean reward
D
Short timescale
Value pred.
Medians
Slope $\beta_{\hat{p}}$ (spks trial^{-1} rew^{-1})
Num. cells
E
Long timescale
VP
Intercept β_0 (spks/trial)
r = 0.43
Whole-session mean rew.
F
G2
G1
Value pred. theory
Cue activated but not value coding
Cue − baseline firing (Hz)
G1
or
Phasic activation by cue
Ramp
Falling activity during rew.
Inhibited by reward epoch end
Value pred.
Firing (Hz)
Time from cue (s)
p = 0.006
G2
Value pred. (very slow discounting)
p = 0.063

Figure 3.6: (Previous page.) Individual serotonin neurons exhibit reward coding features consistent with value prediction. **A** Example trace conditioning experiment of Grossman *et al.* [25] (see Methods). Counterclockwise from top left: Bernoulli water rewards (blue lines) and time varying reward probability estimated using a five-trial moving average (red line). Distribution of number of trials at each level of estimated reward probability; note the wide range of reward probabilities in this dataset. Firing rate of the serotonin neuron recorded in this session calculated using a 500 ms PSTH, lines coloured according to estimated reward probability as in histogram at left, scale bar 2 Hz. Spike raster used to calculate PSTH, background is shaded according to the estimated reward probability as in the other plots. **B** Circular trial permutation test used to assess statistical significance of correlations between estimated reward probability and serotonin neuron activity. Reward history is randomly shifted with respect to spiking activity to build up a null distribution against which the observed correlation can be compared. **C** Serotonin neuron whole-trial activity reflects reward history. Each line represents the following regression model fitted to a single neuron: $\hat{y} = \beta_{\hat{p}}\hat{p} + \beta_0$, where $\hat{y}$ is the predicted whole-trial activity (defined as the number of spikes within a 7.5 s period beginning 1.5 s before the start of the cue and ending 1 s after the end of the reward epoch), $\hat{p}$ is the reward probability estimated as in A, and $\beta_{\hat{p}}$ and β_0 are the slope and intercept. Slope $\beta_{\hat{p}}$ represents the effect of recent reward history $\hat{p}$ on activity and intercept β_0 represents the baseline activity level following a short string of unrewarded trials ($\hat{p} = 0/5$). Donut plot shows the distribution of circular trial permutation test p-values against $H_0 : \beta_{\hat{p}} = 0$; $p < 0.05$ occurs significantly more frequently than 5 % chance rate (inner pie chart). Lines are colour-coded according to statistical significance of the slope $\beta_{\hat{p}}$ as in the donut plot. $N = 37$ neurons. **D** Distribution of regression slopes $\beta_{\hat{p}}$. Note tendency towards positive slopes consistent with value prediction. Colour-coded as in C. **E** Distribution of regression intercepts β_0. Note positive correlation between baseline activity and reward rate over very long timescales consistent with value prediction. Colour-coded as in C. **F** Relationship between reward history modulation (vertical axis) and phasic cue-associated firing (horizontal axis). Note that neurons with clear cue-associated firing (G1 and A, for example) tend to be positively modulated by reward history, consistent with value prediction. **G** Firing dynamics and whole-trial activity modulation in representative serotonin neurons. G1 shows a neuron with clear trial-associated activity dynamics and positive activity modulation by reward history, representative of neurons in the upper right of F. Note the striking correspondence between activity dynamics of the value prediction model (blue inset) and the example neuron. G2 shows a neuron with no clear trial-associated activity dynamics and numerically positive (but not statistically significant) activity modulation by reward history, representative of most neurons in the left part of F. Regression plots (G1 right and G2 right) illustrate the analysis used for C–F. Error bars/bands represent 95 % bootstrap confidence intervals (with Monte-Carlo bias correction in the case of error bars) provided for illustration purposes only.

With data and statistical approach in hand, we turned our attention to quantifying the effect of reward history on the activity levels of serotonin neurons. Regressing the number of spikes per trial (which we refer to as whole-trial activity) onto the five trial mean reward for each neuron (Figure 3.6C) revealed statistically-significant modulation by reward history in 8/37 cells (circular permutation test $p < 0.05$ against regression slope $\beta_{\hat{p}} = 0$; significantly above the 5 % chance rate, binomial proportion test $p < 0.001$; see donut plot in Figure 3.6C) somewhat variable levels of background activity (regression intercepts $\beta_0 = 18.6 \pm 8.0\,\text{spikes}\,\text{trial}^{-1}$, mean $\pm$ SD, equivalent to 2.5 ± 1.1 Hz, coefficient of variation 0.43). It is possible that 22 % represents a lower bound on the proportion of cells in our sample that encode value. Consistent with this possibility, the relationship between activity and reward history was generally positive across the population (regression slopes $\beta_{\hat{p}} = 0.64 \pm 1.60\,\text{spikes}\,\text{trial}^{-1}\,\text{reward}^{-1}$, Wilcoxon signed rank test $p = 0.039$; equivalent to 0.09 ± 0.21 Hz, roughly consistent with the effect size calculation in the previous section; $\beta_{\hat{p}} > 0$ in 65 % of cells, one-sided sign test $p = 0.049$), including in many cells that did not cross the $p = 0.05$ significance threshold in the circular trial permutation test (Figure 3.6D, note significance-stratified medians). We observed two neurons with a statistically significant negative effect of reward history on whole trial activity (Figure 3.6C and D). This is consistent with the expected false positive rate ($2/37 = 5.4\,\%$), but it is also possible that value prediction is not universal in the DRN. Finally, we also observed a significant correlation between background activity and the proportion of all trials rewarded in the corresponding session (Pearson $r = 0.43$ between regression intercept β_0 and whole-session mean reward, $p = 0.008$, Figure 3.6E), consistent with value coding over timescales beyond the five trial horizon. This correlation is surprisingly strong considering the many factors that impact background firing rate (differences in the biophysical features or inputs of individual serotonin neurons, differences in thirst or arousal between recording sessions and mice, etc.) and the coarseness of the whole-session reward metric. Due to a high rate of statistically null results which neither confirm nor rule out value coding, additional analyses are needed to determine to what extent value prediction is typical in the DRN (see

"Value prediction dominates population activity" below). For now, we conclude that at least some serotonin neurons exhibit positive reward history modulation consistent with value prediction.

Our theory predicts that the phasic reward cue-associated activity observed in some serotonin neurons is due to value coding, while other theories propose that phasic activity could be unrelated to reward history. To examine the potential connection between this phasic activity and value prediction, we stratified the slopes obtained from our regression analysis according to the amplitude of the cue-associated extremum in the firing rate (Figure 3.6F). Of the small number of cells with a clear cue-associated peak ($N = 7$ cells with >1 Hz increase in firing above baseline), all exhibited numerically positive reward history modulation (one-sided sign test $p = 0.008$, $N = 7$) which was individually statistically significant in just over half of these cells (circular trial permutation test $p < 0.05$ in $4/7$ neurons; exact p values for each neuron are 0.006, 0.007, 0.009, 0.013, 0.071, 0.126, 0.237 in order of decreasing significance, values lower than $p = 0.002$ are not possible). The connection between phasic activity and value coding was surprisingly consistent: out of 37 neurons, we did not observe any with both clear phasic activity and numerically negative reward history modulation (bottom right quadrant in Figure 3.6F). We conclude that there is a strong association between phasic cue-associated activity and positive modulation by reward history, consistent with the idea that value prediction underlies both phenomena.

The main ideas of this analysis are summed up by the two example neurons presented in Figure 3.6G. In Figure 3.6G1, we see a neuron with activity dynamics strikingly similar to the value prediction model, including cue-associated phasic firing, a pre-reward ramp, and falling activity during the reward period. In this neuron, we also observe statistically significant (circular trial permutation test $p = 0.006$) and almost perfectly linear positive scaling of whole-trial activity by reward history, consistent with value prediction. The neuron in Figure 3.6G1 is clearly well-described by our theory, but the same cannot be said of the neuron in Figure 3.6G2. Examining its peri-stimulus time histogram (PSTH) reveals no

clear activity dynamics, and there is no statistically-significant effect of reward history on firing (circular trial permutation test $p = 0.063$). This neuron may not predictively encode value, but it is also possible that the timescale of this experiment is simply too fast. An absence of discernible within-trial activity dynamics is consistent with value prediction if the discounting timescale is much longer than the trial duration[1] (Figure 3.8), and, while not statistically significant, the magnitude of the reward history modulation is actually closer to our predictions than the unusually large effect shown in Figure 3.6G1 (for an increase in reward probability from 1/2 to 1, the calculation in the previous section shows we expect an increase in firing of $\sim$0.05 Hz, compared with 0.25 Hz and 0.07 Hz in example neurons 1 and 2, respectively). The analyses presented in this section prioritize clarity over statistical power. As a result, effects that pass the statistical significance threshold are likely to be unusually strong and not representative of the broader population of value coding serotonin neurons [245]. Our results are best interpreted as a lower bound on the proportion of serotonin neurons that encode value.

In this section, we have shown that value prediction provides a very good description of the activity patterns of at least some serotonin neurons. Could a different model provide an equally good or even better description of serotonin neuron activity? To what extent are the tuning features explained by value prediction dominant at the population level? We turn to these questions next.

3.2.9 Comparison with expected uncertainty

A significant body of literature argues that serotonin neurons encode a quantity related to RPE, typically its absolute value, in order to signal surprising events [25, 57, 73]. Perhaps surprisingly, some of the strongest evidence for this perspective is consistent with our results.

An influential model of the role of serotonin in learning connects trial-to-trial modulations

[1]This does not seem implausible. Trace conditioning trials are typically only a few seconds long—if humans had a discounting timescale on the order of a minute or less, no-one would ever read beyond the first few sentences of this manuscript.

of serotonin neuron activity to a moving average of absolute RPEs called expected uncertainty [25]. To compare expected uncertainty against our model (which predicts an essentially linear relationship between mean reward and serotonergic activity, *e.g.* Figure 3.6G1 right), we analytically derived the relationship between reward probability and mean absolute RPE in models of animal learning (Section 3.11). While the mean absolute RPE is precisely twice the variance of a binary reward in the simplest of RL models (and therefore has an inverted U-shaped relationship with reward probability), sophisticated models of animal learning often include features that profoundly alter this relationship (*e.g.*, 25, 87). As a result, in theory expected uncertainty is usually negatively related to the mean reward in addition to being positively related to variance (Figure 3.11). Re-analyzing the computed expected uncertainty values from the dynamic Pavlovian task in Grossman *et al.* [25] shows that this is also true in practice: expected uncertainty is more strongly correlated with reward probability than variance in 26/28 sessions in this dataset (Figures 3.9 and 3.10; median marginal r^2 between expected uncertainty and five-trial mean reward 0.815, IQR 0.565 to 0.903, compared with median 0.083, IQR 0.016 to 0.216 for variance). The fact that the previously-reported correlation between expected uncertainty and serotonergic activity is negative more often than not also suggests a connection between serotonergic activity and mean reward rather than variance. We conclude that evidence for serotonin neurons signalling expected uncertainty is consistent with value coding.

3.2.10 Comparison with reward variance

To address the possibility that serotonergic activity might encode reward variability in a way that is not captured by expected uncertainty, we repeated the regression analyses described above (Figure 3.6) using reward variance in place of the mean reward (Figure 3.11). We found that the whole-trial activity levels of serotonin neurons are better described by mean reward than variance (Wilcoxon signed-rank test on weighted sum of squared errors from regression fits $p = 0.007$, $N = 37$), and the effects of mean reward are larger (absolute change in activity

of 1.19 ± 1.25 spikes trial^{-1} from a reward of zero to a reward of one and 0.59 ± 0.69 spikes trial^{-1} from a variance of zero to the maximum variance of 0.25, both quantified using the absolute regression slope, Wilcoxon signed-rank $p = 0.001$), more consistent (regression slopes are typically positive for mean reward, Wilcoxon signed-rank $p = 0.039$, but symmetric around zero for variance, Wilcoxon signed-rank $p = 0.281$), and statistically significant twice as often (21.6 % and 10.8 % of cells with circular trial permutation test $p < 0.05$ against regression slope equal to zero for mean and variance, respectively; for comparison, we expect significant p values in up to 12 % of cells even if none actually encode variance, approx. 95 % CI on a proportion of 5 % with $N = 37$).

Because the mean and variance of binary rewards are directly related, we wondered whether the apparent correlations between reward variance and serotonin neuron activity might actually be due to value coding. Consistent with this idea, statistical confounding between mean reward and variance was unusually strong in the subset of cells with a significant correlation between variance and activity (Pearson r between variance and mean reward is higher in cells with circular trial permutation test $p < 0.05$ for the variance slope $\beta = 0$, permutation t test $p = 0.039$).

Our simple analysis does not reveal clear evidence that serotonin neurons encode an unbiased estimate of reward variance. To mitigate the possibility that these findings are sensitive to technical details of our approach, we repeated the above analyses using reward standard deviation and entropy as alternative definitions of variability, using an iterative/TD method to estimate reward statistics rather than a five-trial moving average, and quantifying activity using a pre-trial baseline rather than whole-trial activity (Figures 3.11 and 3.12). None of these variations affected our results.

We are unable to conclude that serotonin neurons encode reward variability in this task. If serotonin neurons do encode reward variability, our results are consistent with a variability code that is significantly weaker and less consistent than the code for value.

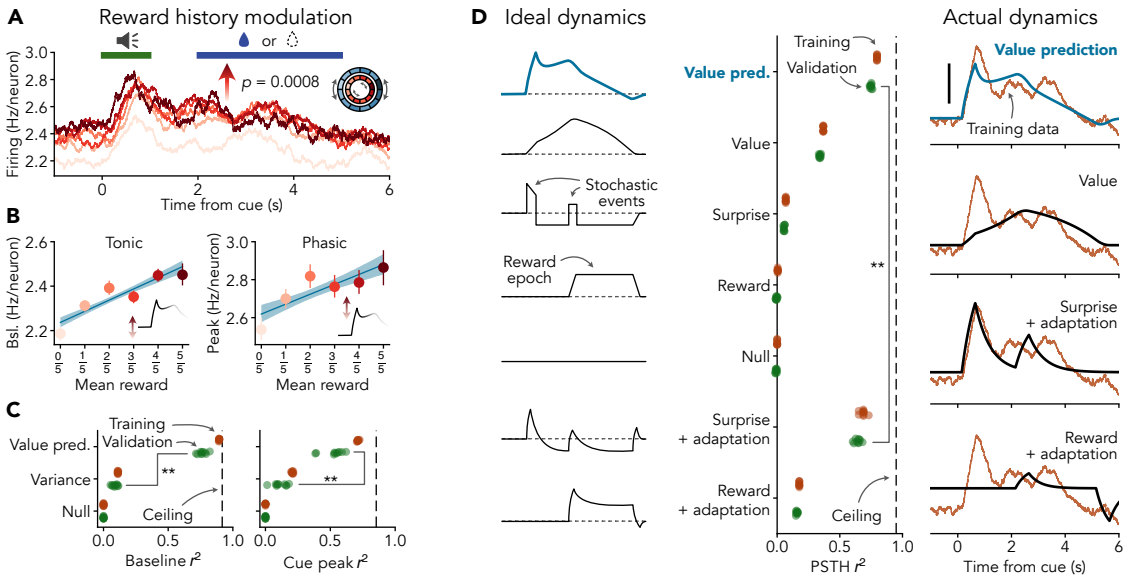


Figure 3.7: Value prediction better explains serotonin neuron population activity than competing theories. **A** Whole-trial population activity encodes reward history. Whole-trial activity and circular permutation test as in Fig. 3.6. **B** Baseline population firing rate quantitatively encodes reward history. Baseline activity defined as mean of PSTH 1 s before start of cue. Peak cue activity is defined as the maximum of the PSTH during the 1 s cue period. Error bars represent 95 % confidence intervals obtained via bootstrap with Monte-Carlo bias correction. Error bands around regression lines represent 95 % confidence interval obtained via bootstrap. Regression slopes are significantly different from zero; bootstrap 99 % CI test. **C** Value prediction better accounts for reward history modulation of population firing rate than variance. x-axis represents the proportion of variance explained (weighted r^2) by each model fitted to data as in B. Performance is presented as the mean five-fold cross-validated accuracy, each point represents one cross-validation repeat. Ceiling line represents the accuracy obtained using the training data to predict the validation data directly (maximum across all repeats). **D** Value prediction better accounts for population firing rate dynamics than competing theories. Schematics at left illustrate predictions of each theory; note that surprise-like signals should decrease during trials, but this does not happen in our model fits. Performance is assessed using repeated five-fold cross-validation as in C. Scale bar 0.2 Hz neuron^{-1}. Trial structure as in A.

3.2.11 Untested models

To address the possibility that a different, untested model might provide a better account of the relationship between reward statistics and serotonergic activity, we inspected the residuals of our regression fits (Figure 3.11). Whereas the regression against reward variance systematically overestimates serotonin neuron activity when the mean reward is low and overestimates activity when the mean reward is high, we did not observe any obvious structure

in the errors of the regression against the mean reward. The marginal relationship between serotonergic activity and mean reward is surprisingly well-described by a straight line, offering no hint as to what a better model might be.

3.2.12 Value prediction dominates population activity

Value prediction provides a good description of the activity dynamics and reward history modulation of at least some serotonin neurons, but are the features captured by our model typical of serotonin neurons in general? To address this question, we constructed a synthetic serotonin neuron population using the $N = 37$ cells in our dataset and tested how well value prediction explains the synthetic population-level activity patterns in comparison with other models. If these cells do not generally predictively encode value, the features predicted by our model could be washed out by noise or masked by other coding features that are subtle at the single neuron level.

We began by considering whether serotonin neuron population activity is positively modulated by reward history. Using a population-level version of the circular trial permutation test, we found that whole-trial population activity is positively modulated by reward history ($p = 0.0008$, Figure 3.7A) at a level consistent with our predictions (expected increase in firing from a mean reward of 1/2 to 1 of $\sim$0.05 Hz neuron^{-1}, see "Serotonin slowly learns value" above, compared with 0.11 Hz neuron^{-1} in our synthetic population based on fitted $\beta_{\hat{p}} = 1.64$ spikes neuron^{-1} trial^{-1}). Encoding of reward history was nearly linear during both the pre-trial baseline and cue epochs (Figure 3.7B, weighted $r^2 = 0.802$ during baseline and $r^2 = 0.656$ during cue), suggesting that tonic and phasic firing participate equally in value prediction. To determine whether a different model could better explain population coding of reward history, we used repeated five-fold cross validation to assess how well mean reward, reward variance, and null models could predict serotonin neuron population activity during the baseline and cue epochs of held-out trials (Figure 3.7C). Consistent with value prediction, the mean reward model exhibited five- to ten-fold better performance than

variability-based alternatives (Table 3.2). Thus, value prediction describes the effects of reward history on serotonin neuron population activity with a precision considerably better than $0.1\,\mathrm{Hz\,neuron^{-1}}$.

The population activity dynamics shown in Figure 3.7A exhibit a cue-associated peak, elevated firing during the trace period, and falling activity during the reward, qualitatively consistent with value prediction. Value (without adaptation), surprise, and reward signals are each missing at least one of these properties (Figure 3.7D left). As a result, quantitative models based on these ideas (see Methods) poorly explain population activity in comparison with value prediction (Figure 3.7D middle and right). The only tested model that offered performance competitive with value prediction was a surprise signal with added adaptation (mean validation weighted $r^2 = 0.749 \pm 0.007$ for value prediction and $r^2 = 0.642 \pm 0.023$ for surprise with adaptation; mean $\pm$ SD of ten cross-validation repeats). However, since the fitted surprise signal with adaptation does not exhibit the expected decrease in activity during the trial (see Methods), the justification for adding adaptation to a model that is already an idealized form of adaptation is dubious, and surprise does not readily explain the effects of reward history on activity (Figure 3.6 and Figure 3.7A–C), nor serotonergic responses to punishments (Figure 3.4) in addition to offering lower performance than value prediction (Figure 3.7D), we believe it can be rejected. In sum, the fast activity dynamics of serotonin neurons are better explained by value prediction than surprise, reward, or a raw value signal.

Overall, we conclude that value prediction provides a remarkably precise and complete account of the population activity patterns of serotonin neurons during trace conditioning, as foreshadowed by our results at the level of individual neurons.

3.3 Discussion

The *in vivo* activity patterns of serotonin neurons are notoriously difficult to explain. In this work, we show that a time-dependent estimate of cumulative future reward predictively encoded through spike-frequency adaptation, which we call value prediction (Figure 3.2), unifies a surprisingly wide range of puzzling observations and conflicting theories from the serotonin literature (Figure 3.1 and Table 3.1). In particular, phasic activation by reward-predicting cues and primary rewards is explained by a rapid increase in proximity to reward, activation by punishments is explained by a rapid increase in value as the end of the punishment approaches, and tonic firing that reflects reward and punishment context is consistent with a code for the value of upcoming trials (Figure 3.3). By simulating trace conditioning experiments with different reward sizes and trial durations, we have shown that the appearance of surprise [57] and salience [39] tuning emerges naturally in our model (Figure 3.4), providing an intuitive link between tuning features that previously seemed conceptually unrelated. To add weight to these qualitative results, we re-analyzed a dataset of serotonin neuron responses to *in vivo* rewards [235]. We observed small modulations in serotonin neuron firing by recent reward history, quantitatively consistent with value estimation over hundreds of trials (Figure 3.6; see also ref. 57). Finally, we directly compared value prediction against competing theories. We found that our theory provides a remarkably precise description of both trial-to-trial changes in activity associated with reward history and within-trial activity dynamics, usually exhibiting predictive performance several times better than alternative models (Figure 3.7). It has been said that "serotonin's many meanings elude simple theories" [76]. Our work shows that several of these meanings — reward, punishment, surprise, salience, and uncertainty — can be merged into one: value prediction.

3.3.1 Hiding in plain sight

Why was a serotonergic code for value not established earlier? The idea that serotonin neurons encode value or something very similar is not new [60, 61], but this perspective has fallen increasingly out of favour in recent years because of evidence that seems to directly contradict a value or reward-based code, first and foremost the fact that many serotonin neurons are activated by punishments [24, 57, 246] as well as model-based analysis that suggests that serotonin neurons do not track reward history on the same timescale as changes in foraging behaviour [25]. Here we have shown that not only do punishment responses not rule out a value code, they are actually expected if reward-predicting cues evoke transient firing, consistent with experimental results (ref. Fig. 5C in 24). Our work also shows that value coding by serotonin neurons is more precise and temporally extended than behaviour would suggest, again consistent with previous results [57] and adding to an emerging pattern of neural systems having population codes that are more precise than behaviour and perception [247, 248].

Turning to the literature and finding results that seemed puzzling at the time but are predicted by our theory (the apparent reward-specificity of surprise tuning being a notable example; 57) became a recurring theme during this project, leading us to feel that evidence for value coding by serotonin, much like dopamine [249], has long been hiding in plain sight. We hope that future work will uncover more such examples, and, conversely, temper our confirmation bias by highlighting tuning features that are clearly incompatible with value prediction.

3.3.2 The meaning of predictively-encoded value

An adapting value signal is not the same as value itself, raising interesting questions about how DRN output might be interpreted by downstream regions (and scientists!).

One possibility, although perhaps not a very exciting one, is that value prediction is simply

reversed to recover the original value signal. The exceptionally large axonal arborizations of serotonin neurons [16] likely place significant metabolic constraints on the activity of serotonin neurons, and predictive coding through adaptation could act as a sort of compression scheme [65, 111, 122, 123, 250] allowing the serotonin system to broadcast a signal widely using a minimum number of spikes. If adaptation compresses the value signal, how is later decompressed? In theory, the exact answer is simple: leaky integration (Section 3.13). In light of the close relationship to predictive coding seen as a compression scheme, it is interesting to note that many of the biological processes involved in decoding serotonin neuron activity implement leaky integration (*e.g.*, accumulation of serotonin in the extracellular space, slow kinetics of G-protein coupled receptors, and the membrane voltage dynamics of downstream cells). Adapting value might be more similar to value than it first appears.

A more intriguing possibility arises if adaptation already visible in the spiketrains of serotonin neurons is further enhanced downstream, for example via depressing synapses. In that case, our work implies that DRN output would be decoded as the rate of change of value [158]. This quantity is required to calculate real-time RPEs (as is value itself; 95, see also 73), and one of the main predictions of the dopaminergic RPE hypothesis of Schultz *et al.* [70] was that the dopamine system should receive input from some region (or collection of regions) that encodes the rate of change of value. Since the dopamine system is one of the main targets of the DRN [251], might serotonin play an important part in computing RPE?

3.3.3 Behaviour

If serotonin predictively encodes value, how is this signal used to drive learning and behaviour? The ways in which value is used in RL are varied and the effects of serotonergic manipulations are perplexing, but here we offer some speculation.

One of the better established roles of serotonin in regulating behaviour is that fast optogenetic activation of serotonin neurons promotes maintenance of behaviours directed at obtaining imminent rewards, an effect commonly called patience or persistence [54–56, 62]. If

serotonin encodes an estimate of relatively immediate reward (due to heavy discounting) that is compared against a longer-term average reward rate, then control policies based on optimal giving up [252] or option interruption [253] could produce similar behaviour. In principle, persistence could also be explained by immediate reinforcement of an ongoing reward-seeking action through policy gradient learning [93] or boosting RPEs (see previous section), but these ideas are difficult to reconcile with evidence that stimulation of serotonin neurons is not generally reinforcing (47, 54, 55, 87, but see 254).

Selfridge's run and twiddle model [240] offers a more elegant potential explanation of persistence in terms of value. According to this simple control model, the current action is maintained as long as value is increasing ("run"), offering an interesting connection to the predictive coding component of our theory, whereas a decrease in value triggers a random action ("twiddle"). Given the evolutionarily-ancient origins of the serotonin system [2] and its involvement in regulating very coarse aspects of behaviour such as the tradeoff between exploration (taking random or sub-optimal actions) and exploitation (taking actions that are expected to lead to reward) [255], we are tempted to speculate that a relatively primitive control policy might provide the best account of the role of serotonin in behaviour. From that perspective, run and twiddle, which was originally conceived to explain the foraging behaviour of bacteria, might be a good place to start.

3.3.4 Learning

Confusingly, activation and inhibition of serotonin neurons both promote maintenance of reward-seeking behaviour. However, whereas optogenetic activation of serotonin neurons produces behaviour directed at obtaining immediate rewards that is usually framed in a positive light, chemogenetic inhibition reduces the rate of abandonment of depleted sources of reward [25, 57], which is framed as perseveration. This effect has been explained in terms of a selective decrease in the rate of learning from reward omissions [25, 57]. There is a normative reason for the learning rate to decrease when rewards are intrinsically variable (Section 3.8)

and increase when the environment is non-stationary, prompting the development of models that separately track variance and volatility to enhance learning [256, 257]. Serotonin has been proposed to modulate the rate of learning from reward omissions via surprise or uncertainty [25, 57], but, as we have shown, these results are more consistent with value. Finally, a model-based analysis showed that optogenetic manipulations of serotonin neuron activity affected behaviour in a way that was consistent with an enhancement of learning rate [87], but this effect was specific to long timescales and also explained relatively well by RPE boosting (ref. Fig. S14; see previous section). Since all components of RL models affect the rate of change of behaviour, it is plausible that many of the apparent qualitative effects of serotonin on learning could be explained through the lens of value.

3.3.5 Mechanistic basis of value prediction

The question of where value prediction comes from can be broken into two parts: where does the value signal input originate, and where does predictive coding through adaptation occur?

Since the DRN receives input from nearly the entire forebrain [17], it is unlikely that a single upstream region is completely responsible for computing value, simply because this input would be drowned out by unrelated information streaming into the DRN. Instead, we believe it is likely that the net value input required by our theory is assembled from multiple sources, for example temporally-extended action values from the mPFC [258], and various aspects of reward, including inverted RPE, from the lateral habenula and lateral hypothalamus [189, 247, 259]. These possibilities are merely speculation. Our theory does not depend on whether the net value input originates in one particular region or is distributed across many others.

As for the predictive coding aspect of our theory, the exceptionally strong spike frequency adaptation of serotonin neurons is not only sufficient [158], it was the main motivation for the present work. This adaptation comes from a combination of apamin-sensitive potassium currents and spike-triggered changes in spike threshold in individual serotonin neurons

[32, 158] as well as network-level recurrent inhibition via 5-HT_{1A} receptors [31, 32]. Feed-forward inhibition [17, 28] does not seem to be functionally involved [158]. Setting aside this physiological evidence, value prediction on its own technically does not exclude the possibility that at least some of the adaptation visible in the spiketrains of serotonin neurons originates upstream of the DRN.

Selectively inhibiting certain DRN inputs or pharmacologically reducing adaptation *in vivo* could shed light on the mechanisms of value prediction in the serotonin system.

3.3.6 Phasic and tonic firing

The idea that serotonin neurons encode essentially unrelated signals in phasic and tonic components of their firing rates is popular in the serotonin literature [73, 154, 233, 234]. Here we have shown that the responses of serotonin neurons to rewards and punishments are well-described by a simple model that does not distinguish between different types of firing. Instead, our model shows that firing patterns that have traditionally been called phasic can be interpreted as increases in firing rate that are short-lived due to adaptation. While we cannot rule out the possibility that serotonin neurons multiplex different quantities in their firing rates in other tasks, it is important to note that the phasic/tonic separation has historically been partly rooted in speculation [73, 234] and the difficulty of formulating a consistent interpretation of serotonergic responses to rewards and punishments [24]. Functionally distinct phasic and tonic firing is an exciting hypothesis, but perhaps no longer a good default modelling assumption.

3.3.7 Heterogeneity

Serotonin neurons are biochemically, developmentally, anatomically, and to some extent electrophysiologically heterogeneous [33]. They are probably computationally heterogeneous as well, but in what sense? In principle, serotonin neurons might be quantitatively computationally heterogeneous, meaning that differences in their activity patterns can be captured by adjusting

the parameters of our value prediction model, or qualitatively computationally heterogeneous, meaning that value prediction simply does not apply to all serotonin neurons. Our results cannot differentiate between these two possibilities (which are not mutually-exclusive in any case). While value prediction dominates the population activity patterns of serotonin neurons in the data we re-analyzed, many individual neurons seemed essentially unresponsive to the task. If serotonin neurons encode value subject to a wide range of discounting timescales (quantitative heterogeneity), echoing distributional coding in the dopamine system [249, 260–262], the cells that appear unresponsive might simply exhibit discounting and/or learning rates that are very slow relative to the structure of the experiment (see Figure 3.8). Consistent with the idea that this experiment is too fast to engage value prediction in most serotonin neurons, the neurons with the clearest responses to reward-predicting cues exhibited reward history modulations much larger than expected based on previous work (see effect size calculation in results). At the same time, the null responses we observed could just as easily be explained by the idea that the neurons in question do not predictively encode value (qualitative heterogeneity). Value prediction subsumes several previous ideas about serotonergic function and is sufficient to explain a surprisingly wide range of results reported in literature, but this does not imply that it is universal.

In view of the marked heterogeneity of the serotonin system in nearly every aspect of its biology, it seems likely that computational heterogeneity in this system is at least partly qualitative. Indeed, there are differences in the tuning features of serotonin neurons across DRN subregions that are difficult for our model to explain [38], as are reported activation by punishment-predicting cues and preference for smaller rewards [24]. Perhaps these observations are the result of competing ON and OFF value prediction pathways in the DRN [32] or something else entirely. By testing value prediction against a battery of alternative models, our work provides a template for assessing computational heterogeneity in the serotonin system.

3.3.8 Top-down meets bottom-up

Here we build on a theme of some of our previous work by adding biological details to a simple model in order to improve its interpretability and performance [158, 159, 161]. Here, the success of our model hinges on combining the normative idea of a value signal with spike-frequency adaptation from a bottom-up model of the DRN [158]. Neither of these are new to the serotonin field [26, 27, 59–61, 63, 155, 156, 171, 246], but, to the best of our knowledge, they have not previously been combined. The fact that serotonergic responses to rewards and punishments only become interpretable after accounting for the effects of adaptation illustrates the usefulness of elements of biological detail even in normatively-focused branches of computational neuroscience.

3.3.9 Conclusion

Here we present a simple theory for serotonin's many meanings: reward, surprise, salience, and uncertainty are different faces of a predictive code for value. Our work links the biology of the serotonin system to a normative account of serotonergic responses to rewards and punishments through value, a quantity that is central to RL theory. On an intuitive level, our definition of value as the expectation of future reward is akin to optimism, providing a conceptual link to the use of serotonergic medications in treating mood disorders. With value prediction, we establish serotonin as "a neural substrate of prediction and reward" [70].

3.4 Methods

3.4.1 True value signal

We define the true value of a state s to be the expected total future reward to be collected by an agent that begins in that state at time t and transitions to future states $S_{t+1}, S_{t+2}, \ldots$ according to the dynamics of the environment. Future rewards are discounted by a factor $0 \leq \gamma \leq 1$ such that rewards after $t+1$ are ignored if $\gamma = 0$ and all rewards are equally valuable if $\gamma = 1$. This textbook definition of value can be written as an explicit sum over future rewards

$$v(s) = \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1} \,\middle|\, S_t = s\right], \tag{3.2}$$

or as an equivalent Bellman recursion

$$v(s) = \mathbb{E}\left[R_{t+1} + \gamma v(S_{t+1}) \mid S_t = s\right].$$

For a trace conditioning experiment with exponentially-distributed inter-trial interval (ITI) durations of mean L_{ITI}, fixed cue, delay, and reward durations $L_{\text{cue}}, L_{\text{delay}}, L_{\text{reward}}$, and fixed reward size, the normalized continuous-time true value signal is given by

$$\frac{v(t)}{v(t_{\text{rew}})} = \begin{cases} \frac{\tau}{L_{\text{ITI}}+\tau} \exp\left[\frac{-(L_{\text{cue}}+L_{\text{delay}})}{\tau}\right] & t \text{ in ITI,} \\ \exp\left[\frac{t-t_{\text{rew}}}{\tau}\right] & t \text{ in cue or delay,} \\ A\,(1 - e^{\frac{t-t_{\text{end}}}{\tau}}) + B & t \text{ in reward,} \end{cases} \tag{3.3}$$

where $\tau = -dt \ln \gamma$ is the discounting timescale, $t_{\text{rew}} = t_0 + L_{\text{cue}} + L_{\text{delay}}$ is the start of the reward epoch (given the start of the trial t_0), $t_{\text{end}} = t_0 + L_{\text{cue}} + L_{\text{delay}} + L_{\text{rew}}$ is the end of the reward epoch, and A, B are scaling and offset factors to ensure continuity. (See Section 3.9

for derivation.) By construction, this normalized definition of the value signal is proportional to the true value signal for any reward size (including negative rewards corresponding to punishments) and can therefore be multiplied by a constant to accommodate different mean rewards.

Note that this model has only one free parameter: the discounting timescale.

3.4.2 Estimated value signal

To simulate the estimated value signal $\hat{v}(t)$ in a trace conditiong experiment, we used the true online TD(λ) algorithm of van Seijen *et al.* [96, 263], which is designed to agree more closely with the forward view of TD learning (Figure 3.2D) than other TD(λ) algorithms. van Seijen's algorithm applies to causal linear value function approximation $\hat{v}(s_t) = \mathbf{w}_{t-1}\mathbf{x}_t$, where $\mathbf{x}_t$ is a vector of state features and $\mathbf{w}_{t-1}$ are weights from the previous time step (since $\mathbf{w}_t$ depends on quantities in the future).

Weights are learned online according to

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \delta_t \mathbf{e}_t + \alpha \left(\overbrace{\mathbf{w}_{t-1}\mathbf{x}_t}^{\hat{v}(s_t)} - \mathbf{w}_t \mathbf{x}_t \right) \mathbf{x}_t$$

$$\mathbf{e}_{t+1} = \gamma\lambda\mathbf{e}_t + \alpha\mathbf{x}_{t+1} - \left(\alpha\gamma\lambda\mathbf{e}_t\mathbf{x}_{t+1}\right)\mathbf{x}_{t+1},$$

where $\mathbf{e}$ is the eligibility trace vector and $\delta_t = R_{t+1} + \gamma\hat{v}(s_{t+1}) - \hat{v}(s_t)$ is the reward prediction error (RPE). Compared with traditional TD(λ) with eligibility traces, this algorithm adds a correction to the weight update and uses Dutch eligibility traces, which are intermediate to accumulating ($\mathbf{e}_{t+1} = \gamma\lambda\mathbf{e}_t + \alpha\mathbf{x}_t$) and replacing traces ($\mathbf{e}_{t+1} = \gamma\lambda\mathbf{e}_t \odot (1 - \mathbf{x}_t) + \alpha\mathbf{x}_t$, where x_t is an indicator vector). Following the notation of van Seijen et al. [263], we place the learning rate α inside the eligibility trace update rather than in front of δ_t in the weight update.

Simulation was performed using tabular features $\mathbf{x}_t = \mathbf{1}_s$, eligibility trace $\lambda = 0.995$, discounting factor $\gamma = 0.99$, learning rate $\alpha = 0.01$, and time step $dt = 50\,\text{ms}$.

3.4.3 Value prediction model

The DRN rate model-based value prediction model is defined as

$$\begin{aligned} \rho(t) &= \mathrm{ReLU}[(1+A)\ v(t) - A\ u(t)] \\ \frac{du}{dt} &= \frac{\rho(t) - u(t)}{\tau_{\mathrm{ad}}}, \end{aligned} \tag{3.4}$$

where $\rho(t)$ is the firing rate of DRN serotonin neurons; $v(t)$ is the time-dependent net input, assumed to be a value signal (either true or estimated, as indicated in main text); $u(t)$ is adaptation; A and τ_{ad} are the strength and timescale of adaptation, respectively; and $\mathrm{ReLU}[x] = \max(x, 0)$ is the rectified linear function. The input is rescaled by a factor $1 + A$ so that $\rho(t) = C$ for any constant input $x(t) = C$ independent of the strength of adaptation A.

We used $A = 3$ and $\tau_{\mathrm{ad}} = 1\,\mathrm{s}$ to achieve effective adaptation amplitude and kinetics similar to those observed in our previous experimentally-constrained semi-biophysical model [158]. Note that the effective adaptation kinetics are faster than τ_{ad} due to feedback between $u(t)$ and $\rho(t)$. Adaptation dynamics were numerically integrated using the second-order Runge-Kutta method. Surprise/salience tuning simulations were carried out using a time step $dt = 1\,\mathrm{ms}$, all others used $dt = 50\,\mathrm{ms}$.

For a true value signal, the value prediction model has only three free parameters: the discounting timescale, the strength of adaptation, and the adaptation timescale.

3.4.4 *In vivo* experiment

The dynamic trace conditioning experiment analyzed here has been reported previously by Grossman *et al.* [25]. A brief summary is as follows:

Tetrode recordings of optogenetically-tagged serotonin (SERT-expressing) neurons were collected from head-fixed and water-restricted C57BL/6J mice presented with odour-cued water rewards. Each trial consisted of a 1 s odour cue followed by a 1 s delay and a 3 s window

during which a fixed-size water reward (approx. 2 µL to 4 µL) could be collected from a lick spout. After the 3 s reward interval, the lick spout was retracted and any remaining water removed via vacuum for 1 s. Inter-trial interval durations were exponentially-distributed with a mean parameter of $3.\overline{3}$ s (actual: 3.31 ± 3.46 s, mean $\pm$ SD, range 0 s to 54.28 s).

Rewards were delivered probabilistically according to a hierarchical Bernoulli point process

$$R_t \sim \begin{cases} \text{Bernoulli}[p_t] & \text{with } 95\,\% \text{ proba. (odour A)} \\ 0 & \text{catch trial; } 5\,\% \text{ proba. (odour B),} \end{cases}$$

where the reward probability p_t varied according to a block structure. Block lengths were uniformly distributed between 20 and 70 trials. Depending on the recording session, reward probabilities were set to $p_t \in \{0.2, 0.5, 0.8\}$ or $p_t \in \{0.2, 0.8\}$. Catch trials and trials in which a reward was available but not collected were deemed unrewarded for the purposes of our analysis.

Mice collected 0.45 ± 0.04 rewards per trial (mean $\pm$ SD; range 0.37 to 0.54) and completed 346 ± 88 trials per session (range 180 to 565). Mice failed to collect available rewards 4.1 ± 4.3 % of the time (median 3.4 %, range 0.0 % to 19.0 %). Sessions lasted approximately 1 h (time from start of first trial to start of last trial: 53.1 ± 12.9 min, range 29.0 min to 88.1 min). Sessions with more trials were not significantly associated with higher or lower reward rates (Pearson $r = 0.206$, $p = 0.293$, $N = 28$ sessions). Data were gathered from five mice (four male, one female) across 28 sessions with 1.32 ± 0.54 neurons recorded per session (range 1 to 3).

Surgical and experimental procedures were approved by the Johns Hopkins University Animal Care and Use Committee and performed in compliance with the *National Institutes of Health Guide for the Care and Use of Laboratory Animals*.

3.4.5 Data analysis

Reward history

Reward history was operationally defined as the mean reward across the past five trials

$$\hat{p}_t = \frac{1}{5}\sum_{i=1}^{5} r_{t-i},$$

which is an unbiased estimate of the true time-varying Bernoulli reward probability. To mitigate boundary effects, we set $r_t = 0.45$ for $t < 0$.

Note that the mean reward can be used as a crude proxy for value because the true value is proportional to the true reward probability. The connection between mean reward and true value is also exploited by temporal difference learning methods, which implicitly define value as an exponential moving average of past rewards (Section 3.7). We choose to use the five-trial mean reward instead of the true reward probability or a TD estimate so that the connection between the activity of serotonin neurons and the animal's estimate of a reward statistic is clear.

Reward variability

We quantified reward variability using three statistics that can be derived from the estimated Bernoulli reward probability $\hat{p}_t$: reward variance

$$\widehat{\mathrm{Var}[R_t]} = \hat{p}_t(1-\hat{p}_t),$$

standard deviation

$$\widehat{\mathrm{SD}[R_t]} = \sqrt{\hat{p}_t(1-\hat{p}_t)},$$

and entropy

$$\widehat{H[R_t]} = \hat{p}_t \log_2 \hat{p}_t + (1-\hat{p}_t)\log_2(1-\hat{p}_t).$$

Variance is the focus of our analysis because it is proportional to the absolute RPE used as a measure of surprise or uncertainty in some reward learning models (Section 3.11). The other statistics are nearly proportional to variance and are included only to illustrate that our results to not depend on technical details of the variability measure.

Quantification of activity

We quantified neural activity using either the peri-stimulus time histogram (PSTH) with 500 ms window width or, as a more coarse-grained metric, the mean number of spikes in a 7.5 s window around each trial (1.5 s pre-trial baseline, 5 s trial, 1 s post-trial baseline), which we refer to as "whole-trial activity". Baseline activity was defined as the mean of the PSTH 1 s before the start of each trial or, similarly, the number of spikes in a 1 s period just before the start of the trial. Cue-associated activity was defined as the extremum (usually maximum) of the PSTH during the 1 s cue period. PSTHs were not smoothed.

All activity metrics were precision-weighted across neurons and reward history conditions as applicable. For example, the population PSTH for the reward probability $\hat{p} = 1/5$ condition was calculated by summing spiketrains from all trials of all neurons where $\hat{p}_t = 1/5$ and dividing by the total number of trials being summed. The resulting population PSTH can be seen as a weighted average of the PSTHs of individual neurons, where each neuron is weighted according to the precision (inverse variance) of its PSTH N/σ^2.

Dynamical model definitions

Value The value prediction and value models of serotonin neuron activity are defined by Eq. 3.3 and 3.4. The discounting timescale, adaptation strength, and adaptation timescale are estimated from the data along with scale and offset parameters.

Surprise Surprise is defined in reward learning as the absolute reward prediction error $|\delta_t|$ and in information theory as $\log_2 p$, both of which are zero for deterministic events and

greater than zero for stochastic events. Therefore, each instant in the inter-trial interval, the beginning of the trial, and the beginning of reward delivery all have non-zero surprise, while all other moments during the trial have zero surprise.

Our surprise-like model is defined as a piecewise constant function

$$\text{SurprishModel}(t) = \begin{cases} A & t \text{ is in ITI,} \\ B & t \text{ is start of trial,} \\ C & t \text{ is the start of reward delivery,} \\ D & \text{otherwise,} \end{cases}$$

subject to the restrictions

$$B \geq A \geq D$$

and

$$C \geq D,$$

where the coefficients A, B, C, D are estimated from the data.

Reward The reward model is a piecewise constant function

$$\text{RewardModel}(t) = \begin{cases} A & t \text{ is in reward epoch,} \\ B & \text{otherwise,} \end{cases}$$

where the coefficients A, B are estimated from the data.

Null The null model is a constant function with an offset parameter estimated from the data.

Surprise with adaptation The surprise-like model with adaptation is modified from the surprise-like model defined above:

$$\begin{aligned}\text{AdSurprishModel}(t) = & A \\ & + B\,\Theta(t - t_{\text{start}})\, e^{\frac{-(t-t_{\text{start}})}{\tau}} \\ & + C\,\Theta(t - t_{\text{rew}})\, e^{\frac{-(t-t_{\text{rew}})}{\tau}} \\ & + D\,\Theta(t - t_{\text{start}})\,\Theta(-(t - t_{\text{end}})) \\ & + E\,\Theta(t - t_{\text{end}})\, e^{\frac{-(t-t_{\text{end}})}{\tau}},\end{aligned}$$

subject to the restrictions

$$\begin{aligned}A &\geq D \\ B &\geq 0 \\ C &\geq 0 \\ E &\geq 0 \\ \tau &\geq 0,\end{aligned}$$

where $\Theta(\cdot)$ is the Heaviside step function and the coefficients A, B, C, D, E, τ are estimated from the data. As before, A, B, C, D represent the activity during the ITI, trial start, reward start, and the remainder of the trial, respectively. E represents the amplitude of the overshoot at the end of the trial and τ is the timescale of adaptation.

While adaptation can be used to compute surprise, we are doubtful that adaptation should be added to a model of surprise itself: the fact that adding adaptation actually slows down the kinetics of the surprise signal is a major conceptual difficulty. We include this model in our analysis only for completeness.

Reward with adaptation The reward model with adaptation is modified from the reward model defined above:

$$\begin{aligned}\text{AdRewardModel}(t) =& A\,\Theta(t - t_{\text{rew}})\,\Theta(-(t - t_{\text{end}})) \\ &+ B \\ &+ C\,\Theta(t - t_{\text{rew}})\,e^{\frac{-(t-t_{\text{rew}})}{\tau}} \\ &+ D\,\Theta(t - t_{\text{end}})\,e^{\frac{-(t-t_{\text{end}})}{\tau}},\end{aligned}$$

subject to the restrictions

$$\begin{aligned}A &\geq B \\ C &\geq 0 \\ D &\leq 0 \\ \tau &\geq 0,\end{aligned}$$

where the coefficients A, B, C, D, τ are estimated from the data. As before, A, B represent the activity during the reward period and at all other times, respectively. C parameterizes the amplitude of the phasic activity associated with reward onset, D parameterizes the amplitude of the undershoot associated with reward offset, and τ is the time constant of adaptation.

Dynamical model fitting

Before fitting, all models were smoothed with a 500 ms boxcar filter to simulate a PSTH and lagged by 150 ms to account for perceptual delays.

All models were fitted by minimizing the mean squared error on the population PSTH (see Quantification of activity). For the reward model (without adaptation) this was accomplished using linear regression. For all other models, this was accomplished using bounded/constrained gradient-based optimization methods provided by `scipy.optimize.minimize` (L-BFGS-B or sequential least-squares quadratic programming).

Performance was assessed using repeated five-fold cross validation. Data was stratified

by neuron identity (which partially reflects reward history beyond the five-trial horizon, Figure 3.6E) and reward history level prior to assigning folds in order to minimize class imbalances between training and validation sets.

We use cross validation rather than AIC or BIC because cross validation does not rely on distributional assumptions that would be needed to formulate a likelihood function for each model.

Analysis of reward modulation

Reward modulation was defined as the slope $\beta_{\hat{p}}$ of a regression line between an activity metric (see Quantification of activity) and the estimated reward probability

$$\hat{y}_i = \beta_{\hat{p}}\hat{p}_i + \beta_0 + \epsilon_i,$$

where y_i is the measured activity, $\hat{p}_i \in \{0/5, 1/5, \ldots, 5/5\}$ is the estimated reward probability (see Reward history), β_0 is the intercept of the regression line, and ϵ_i is a residual. The regression model was fitted using weighted least-squares (observations were precision-weighted; see Quantification of activity).

Statistical significance of the slope $\beta_{\hat{p}}$ was assessed using circular trial permutation or bootstrapping.

Circular trial permutation tests are not sensitive to autocorrelations in timeseries data that can significantly increase the false positive rates of classical and shuffling-based statistical tests [244]. These tests were performed by shifting the per-trial estimated reward probabilities $\hat{p}_t$ to break the alignment between activity and reward history while controlling for other correlations in the data. For example, the reward probabilities for a T trial experiment $\hat{p}_{t=1}, \hat{p}_{t=2}, \ldots, \hat{p}_{t=T}$ can be shifted D places to generate un-aligned probabilities

$$\tilde{p}_t = \hat{p}_{t+D \pmod T},$$

and the analysis described above can be repeated using the shifted probabilities $\tilde{p}$. The value of the slope $\beta_{\tilde{p}}$ obtained using this procedure represents the apparent reward modulation under the null hypothesis that activity and reward history are not actually related (because the trials were shifted). In the case of whole-trial reward modulation of population activity, the permutation procedure was repeated 2500 times to generate a distribution for $\beta_{\tilde{p}}$ used to obtain an approximate p-value for $\beta_{\hat{p}}$. In the case of whole-trial reward modulation of individual neurons, the procedure was repeated exhaustively to generate exact p-values. In both cases, we restricted $10 \leq D \leq T - 10$ due to very high experimental design-related autocorrelations in $\hat{p}$ (specifically, block structure and the fact that $\hat{p}_t$ and $\hat{p}_{t-4}, \ldots, \hat{p}_{t+4}$ are calculated on overlapping sets of trials). Removing this restriction *post hoc* did not meaningfully affect our results.

Bootstrap distributions for activity metrics $\hat{y}_i$, regression predictions y_i, and regression slopes $\beta_{\hat{p}}$ were generated by sampling trials 1000 times with replacement within reward probability levels $0/5, 1/5, \ldots, 5/5$ and neurons as applicable. Activity metric distributions were corrected for Monte-Carlo bias [264]. Statistical significance of the regression slope $\beta_{\hat{p}}$ was assessed using the 99 % confidence interval method.

For consistency with dynamical model comparison, reward modulation models based on reward history $\hat{p}$ and related statistics $\mathrm{Var}[R], \mathrm{SD}[R], H[R]$ (see corresponding section above) were compared using repeated stratified five-fold cross validation. Performance is presented as the mean variance explained across folds for each repeat.

3.4.6 Statistical analysis

Statistical tests are specified in the main text and methods above. Non-parametric tests were used as much as possible; where samples sizes were so small that the loss of power associated with non-parametric tests became prohibitive, robust tests were used. All tests are two-sided unless otherwise stated. Sign tests were one-sided because our theory implies that the relevant effects should have a specific sign. Exact p values are reported in the main text. Results were

considered statistically significant at $p \leq 0.05$ and $p \leq 0.1$ was considered a trend. Sample sizes were not predetermined because only previously published data was used. $N = 37$ neurons in nearly all cases. Because neurons were usually recorded individually (see "*In vivo* experiment" above), we considered them to be independent biological replicates for the purposes of statistical analysis. Error bars represent 95 % confidence intervals. Uncertainties are presented as standard deviation in the main text unless otherwise specified.

3.4.7 Data and code availability

Previously-published data is available on the Dryad repository [235]. Code will be made available on GitHub.

3.4.8 Copyright permissions

Elements of Figure 3.3B and C, Figure 3.4C, and Figure 3.5 have been reproduced from Cohen *et al.* [24], Matias *et al.* [57], and Zhong *et al.* [58], respectively, under the Creative Commons Attribution license (CC-BY 4.0). A bar chart from Paquelet *et al.* [39], which is covered by the Creative Commons Attribution Non-commercial No Derivatives license (CC-BY-NC-ND 4.0), has been included in Figure 3.4D with kind permission from Bradley Miller.

Artwork has been modified by changing colours, replacing axis annotations with icons or larger text, and/or adding schematics to improve clarity. Due to space constraints, small reward and neutral stimulus groups were removed from the vignette in Figure 3.4C, as were statistical annotations. In all cases, our aesthetic modifications do not change the interpretation of the underlying data. References to the specific figures from the original publications are included in our captions.

3.5 Acknowledgements

We wish to acknowledge that this work was carried out on the unceded and unsurrendered land of the Algonquin Anishinaabe people. The data we re-analyzed was collected on the unceded lands of the Piscataway and Susquehannock people.

Emerson Harkin is grateful for PGS-D and Queen Elizabeth II Scholarships in Science and Technology awards from the Natural Sciences and Engineering Research Council and Government of Ontario, respectively. This work was funded by grants to Richard Naud.

We thank Bradley Miller and Paul Albert for providing feedback on an initial draft of this paper. We also thank Michael Lynn and Sébastien Maillé for many helpful brainstorming sessions and input on figure design; John Beninger for assisting with troubleshooting population-level analysis; and all members of the Béïque and Naud labs for helpful discussions.

Supplementary figures

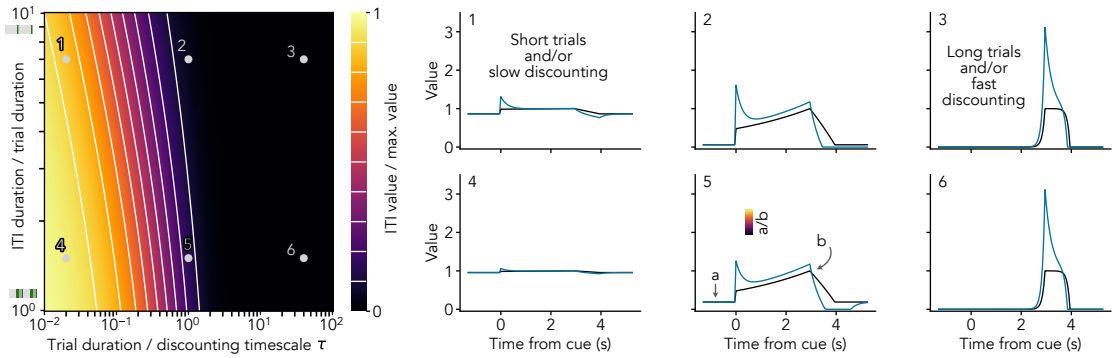


Figure 3.8: Effects of discounting over extreme timescales. This figure extends the horizontal axis of main text Figure 3.3D. Heatmap shows the ITI value relative to the maximal value during the trial (*i.e.*, the value just before the start of reward delivery) as a function of experimental parameters (vertical axis) and the duration of the trial relative to the discounting timescale of the agent (horizontal axis). Ribbons next to the vertical axis are to scale, gray represents the mean ITI duration and colours represent trial epochs. Traces show the true value dynamics (black) and predictively-encoded value (blue) for a trial consisting of a 3 s combined cue and delay epoch and a 1 s reward epoch. True value is normalized to the maximum just before reward as in the heatmap; different reward sizes (and learning) can be accommodated by scaling the traces. Note that the value signal is nearly flat if the discounting timescale is much longer than the trial duration (traces 1 and 4), similar to the dynamics of some serotonin neurons (*e.g.*, Figure 3.6G2).

Table 3.2: Performance statistics for models of population activity modulation by reward history. Related to Figure 3.7C.

	Training r^2		**Validation r^2**	
Model	**Baseline**	**Cue**	**Baseline**	**Cue**
Value pred. (ours)	0.894 ± 0.002	0.713 ± 0.006	0.763 ± 0.030	0.526 ± 0.074
Variance	0.111 ± 0.001	0.208 ± 0.005	0.088 ± 0.019	0.111 ± 0.048
Std. dev.	0.109 ± 0.001	0.168 ± 0.004	0.086 ± 0.018	0.085 ± 0.040
Entropy	0.111 ± 0.001	0.188 ± 0.004	0.088 ± 0.019	0.098 ± 0.044
Null	0.000	0.000	0.000	0.000

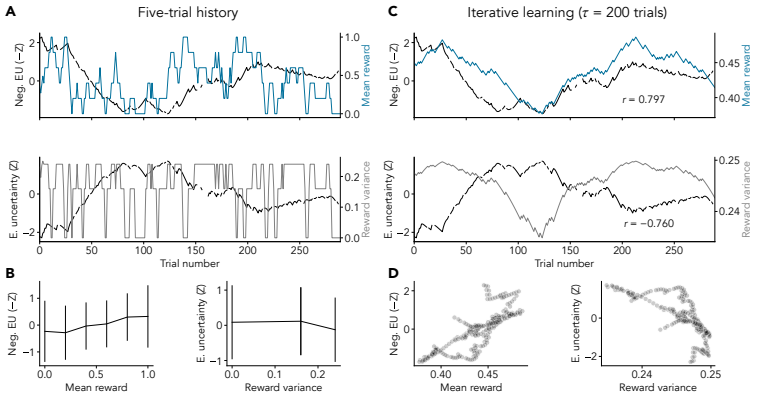


Figure 3.9: Expected uncertainty [25] reflects mean reward, not reward variance, in an example session. **A** Comparison between the time-course of expected uncertainty (black) and mean reward (blue, top) or reward variance (gray, bottom) calculated using a five-trial history. Note that expected uncertainty is theoretically *negatively* related to the mean reward and *positively* related to variance (Section 3.11). Consistent with this, negative expected uncertainty closely resembles five-trial mean reward (top). Gaps in the expected uncertainty timeseries are due to removal of catch trials. **B** Marginal relationship between expected uncertainty and mean reward (left) and reward variance (right) calculated using a five-trial history, as in our main analysis (*e.g.*, Figure 3.7B). Data are presented as mean ± SD. Note connection between negative expected uncertainty and the mean reward (left). **C** Comparison between the time-course of expected uncertainty (black) and mean reward (blue, top) or reward variance (gray, bottom) calculated using an exponential moving average with an estimation time constant of $\tau_{\text{est}} = 200$ trials. The mean reward calculated using an exponential moving average is equivalent to state value under TD learning (Section 3.7). **D** Marginal relationship between expected uncertainty and mean reward (left) and reward variance (right) calculated using an exponential moving average as in C. Note connection between expected uncertainty and mean reward (left). Z-scored expected uncertainty values for recording session `mBB036d20160623` [235] used in [25] kindly provided by Cooper Grossman.

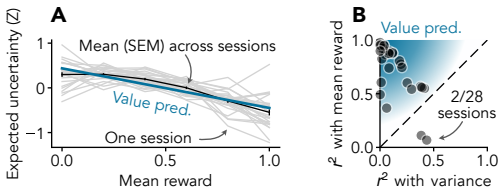


Figure 3.10: Expected uncertainty [25] reflects mean reward, not reward variance, across sessions. **A** Marginal relationship between expected uncertainty and mean reward. Marginals were calculated by averaging expected uncertainty at each level of mean reward calculated using a five-trial history, as in our main analysis (*e.g.*, Figure 3.6A). Black line represents mean ± SEM across $N = 28$ recording sessions. Note that expected uncertainty is negatively related to the mean reward, consistent with our simplified model (Section 3.11). The relationship between expected uncertainty and five-trial mean reward is nearly linear, possibly due to the long timescale over which expected uncertainty is estimated [25]. **B** Expected uncertainty is more closely related to the mean reward than reward variance in nearly all sessions. Unweighted r^2 values represent the fraction of variance in the gray lines from A that is explained by a straight line (mean reward, vertical axis) or a parabola (variance, horizontal axis). Z-scored expected uncertainty values used in [25] kindly provided by Cooper Grossman.

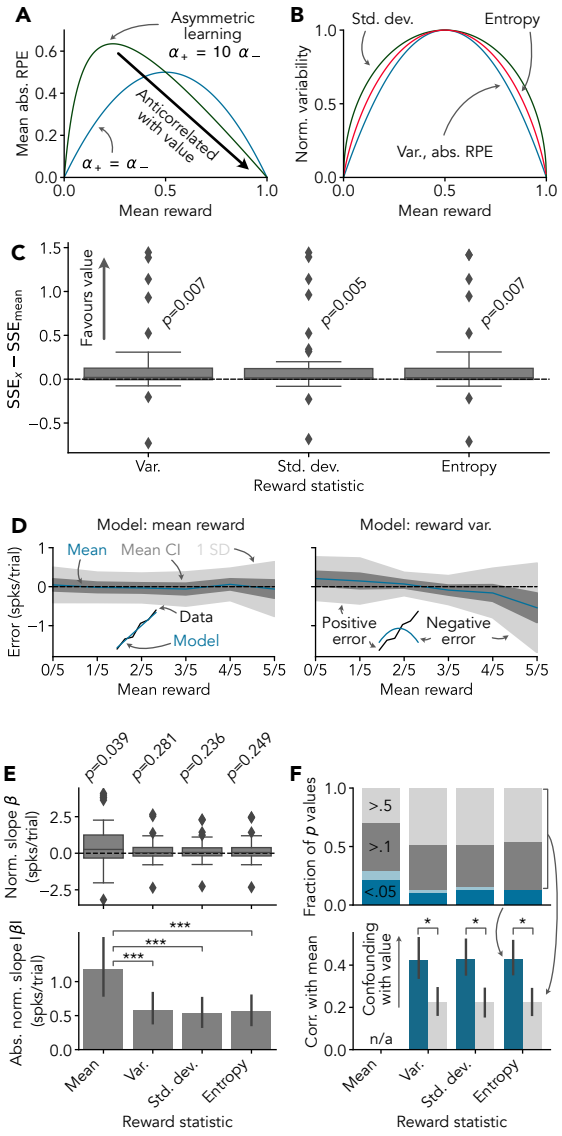

A
Asymmetric learning
α+ = 10 α−
Anticorrelated with value
α+ = α−
Mean abs. RPE
Mean reward
B
Std. dev.
Entropy
Norm. variability
Var., abs. RPE
Mean reward
C
Favours value
p=0.007
p=0.005
p=0.007
SSEx − SSEmean
Var.
Std. dev.
Entropy
Reward statistic
D
Model: mean reward
Model: reward var.
Mean
Mean CI
1 SD
Data
Model
Positive error
Negative error
Error (spks/trial)
Mean reward
E
p=0.039
p=0.281
p=0.236
p=0.249
Norm. slope β (spks/trial)
Abs. norm. slope |β| (spks/trial)
Mean
Var.
Std. dev.
Entropy
Reward statistic
F
Fraction of p values
>.5
>.1
<.05
Corr. with mean
Confounding with value
n/a
Reward statistic

Figure 3.11: Serotonin neuron activity reflects mean reward and is spuriously correlated with reward variability. **A**,**B** Reward variability is inextricably linked to mean reward for binary (Bernoulli) rewards. **A** Expected uncertainty [25], defined as the mean RPE in a model of biased animal learning, has an inverted U-shaped relationship with the mean reward. If learning rates for positive and negative RPEs are highly asymmetric, as in previous work, expected uncertainty is mostly negatively related to the mean reward. **B** Unbiased statistics of reward variability, such as reward variance, also have inverted U-shaped relationships with the mean reward. **C** Serotonin neuron activity is better described by mean reward than reward variability. $\mathrm{SSE}_x - \mathrm{SSE}_{\mathrm{mean}}$ represents the difference in weighted sum of squared errors (SSE) between a linear fit of serotonin neuron whole-trial activity to the reward variance, standard deviation, or entropy versus a linear fit of whole-trial activity to the mean reward (estimated using a five trial history, as in main text). Positive values indicate a better fit (lower error) with the mean reward model. p values are from Wilcoxon signed-rank tests. **D** Serotonin neuron activity does not systematically deviate from the mean reward model (left). Errors are derived from linear fits used in C. Dark gray bands indicate 95 % CIs on the mean (*i.e.*, mean ± 1.96 SEM). No systematic errors are expected if the model is essentially correct (illustrated in inset). Note that the variance model exhibits systematic errors (right), as expected if serotonin neuron activity encodes the mean reward (illustrated in inset). **E** Serotonin neuron activity is relatively strongly positively related to mean reward. Slopes are derived from fits used in C and normalized to the dynamic range of the corresponding reward statistic as in B (0–1 for mean reward, 0–1/4 for variance, 0–1/2 for standard deviation, and 0–1 for entropy). p values are from Wilcoxon signed-rank tests. **F** The activity of individual serotonin neurons is more often correlated with mean reward than reward variability, and significant correlations with variability are likely due to confounding. A significant proportion of serotonin neurons exhibit activity that is significantly correlated with each reward statistic (top; cyclic permutation test, see main text; dark blue, light blue, dark gray, and light gray indicate $p \leq 0.05$, $0.05 < p \leq 0.1$, $0.1 < p \leq 0.5$, and $p > 0.5$, respectively, as in main text). However, in the subset of neurons in which activity is significantly correlated with reward variability, reward variability is unusually strongly confounded with the mean reward (bottom). Stars indicate $0.01 < p \leq 0.05$ (actual range: $p = 0.024$ to $p = 0.037$) in permutation-based two sample t tests (non-parametric tests could not be used because too few neurons had statistically-significant correlations with reward variability statistics). $N = 37$ neurons in all cases.

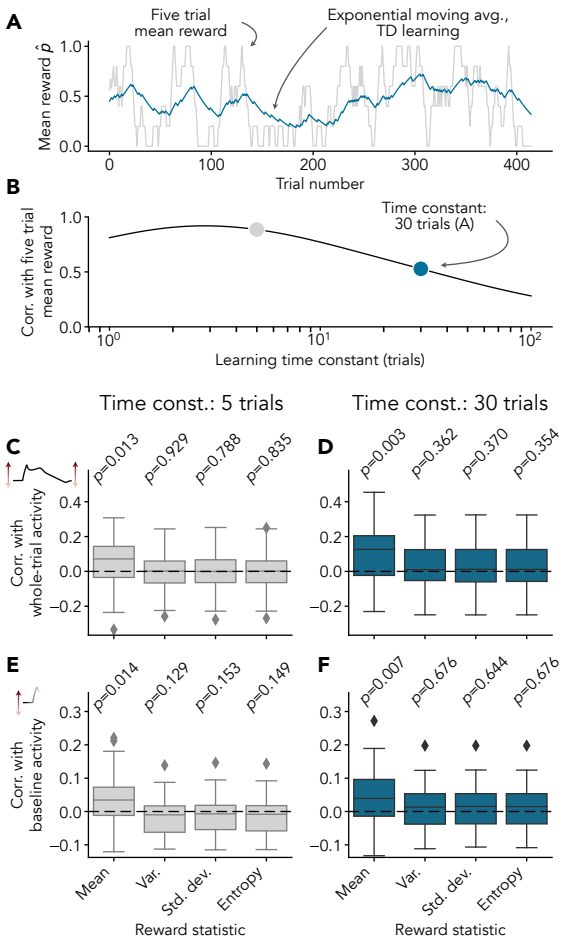


Figure 3.12: Correlation between serotonergic activity and mean reward does not depend on technical details of analysis. **A**,**B** Conclusions are unlikely to depend on the decision to use a five trial reward history as a proxy for value rather than an incrementally-learned estimate because the two are very similar. **A** Comparison of five trial mean reward used in main text with an exponential moving average of past reward for an example session. The blue line represents an exponential moving average with a time constant of 30 trials, which is equivalent to the value estimated by temporal difference (TD) learning using a learning rate of $\alpha = 1/30$ (and no discounting). **B** Pearson correlation between five trial mean reward and moving reward average as a function of the time constant of the moving average (learning time constant) for the example session shown in A. Note that the two are moderately or strongly correlated across a wide range of learning time constants. **C**,**D** Whole trial activity is correlated with mean reward but not reward variability when reward statistics are calculated using a moving average of past rewards rather than a five trial history. Note that this result does not depend on the choice of learning time constant. **E**, **F** Choice of activity metric does not affect conclusions: baseline activity is also correlated with mean reward but not reward variability. In C–F, reward variance, standard deviation, and entropy were calculated from the learned $\hat{p}_t$ rather than moving averages (of squared errors, for example) because this approach is less statistically biased and/or better defined. $N = 37$ cells in all cases.

Supplementary notes

3.6 Connection between state value $v(s)$ and state-action value $q(s, a)$

It is well known that the state-action value depends on the state value:

$$q(s, a) = \mathbb{E}\left[R_{t+1} + \gamma v(S_{t+1}) \mid S_t = s, A_t = a\right]$$

[92]. As we show below, the state value can also be expressed in terms of state-action values.

Following the notational conventions of RL theory, let S, A be random variables denoting the state and action and let s, a be a specific state and action. Let $v(s)$ and $q(s, a)$ be the state and state-action value functions, respectively, under the action-generating policy $\pi(a \mid s) = \Pr[A_t = a \mid S_t = s]$. The state value function $v(s)$ is equivalent to the expected state action value $q(s, a)$ under the policy π:

$$\begin{aligned} v(s) &= \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1} \,\middle|\, S_t = s\right] \\ &= \mathbb{E}_{A_t \sim \pi(a|s)}\left[q(s, A_t) \mid S_t = s\right] \end{aligned}$$

Proof: The state-action value function is the expected cumulative discounted future

reward to be obtained after taking action a in state s, written as

$$q(s,a) = \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1} \;\middle|\; S_t = s, A_t = a\right].$$

Note that this is exactly the same as the state value $v(s)$ except for the dependence on the chosen action $A_t = a$.

To prove that $v(s)$ is the expected q value, we need to incorporate A_t into $v(s)$. To simplify notation, let G_t be the random variable $\sum_{i=0}^{\infty} \gamma^i R_{t+i+1}$. Using G_t and expanding expectation, we can rewrite the state value as follows

$$v(s) = \int_{-\infty}^{\infty} G_t \Pr[G_t \mid S_t = s] dG_t.$$

Incorporating the chosen action A_t by conditioning, we obtain

$$\begin{aligned} v(s) &= \sum_a \overbrace{\int_{-\infty}^{\infty} G_t \Pr[G_t \mid s, a]\, dG_t}^{q(s,a)} \overbrace{\Pr[A_t = a \mid s]}^{\pi(a|s)} \\ &= \mathbb{E}_{A_t \sim \pi(a|s)} \left[q(s, A_t) \mid S_t = s\right], \end{aligned}$$

completing the proof.

3.7 Temporal difference learning averages past rewards

Consider an experiment with a single state (*e.g.*, a trace conditioning experiment where each trial is considered to be a time step) and a stochastic reward $R \sim \text{Bernoulli}[p]$. Learning the state value $\hat{v}_t$ using the RPE $\delta_t = R_t - \hat{v}_t$ causes the state value to be an exponential moving average of past rewards that is an unbiased estimate of the reward probability p

$$\mathbb{E}[\hat{v}] = \mathbb{E}[R] = p$$

Proof: The current state value $\hat{v}_t$ depends on the past state value $\hat{v}_{t-1}$ and RPE δ_{t-1}

$$\hat{v}_t = \hat{v}_{t-1} + \alpha\delta_{t-1},$$

where $0 \leq \alpha \leq 1$ is the learning rate. Expanding the above, we obtain

$$\begin{aligned}
\hat{v}_t &= \hat{v}_{t-1} + \alpha(R_{t-1} - \hat{v}_{t-1}) \\
&= \alpha R_{t-1} + (1-\alpha)\hat{v}_{t-1} \\
&= \alpha R_{t-1} + (1-\alpha)\alpha R_{t-2} + (1-\alpha)^2\hat{v}_{t-2} \\
&= (1-\alpha)^{H+1}v_{t-H-1} + \alpha\sum_{i=0}^{H}(1-\alpha)^i R_{t-i-1},
\end{aligned}$$

where $H \in \mathbb{N}$ is the time horizon. Taking an infinite horizon, the term $(1-\alpha)^{H+1}\hat{v}_{t-H-1}$ vanishes, and we are left with value as a scaled exponential moving average of past rewards

$$\hat{v}_t = \alpha\sum_{i=0}^{\infty}(1-\alpha)^i R_{t-i-1}$$

with an estimation timescale $\tau_{\text{est}} = -dt/\ln(1-\alpha)$.

Taking the expectation shows that the above is an unbiased estimate of the reward probability

$$\begin{aligned}
\mathbb{E}[\hat{v}_t] &= \mathbb{E}\left[\alpha\sum_{i=0}^{\infty}(1-\alpha)^i R_{t-i-1}\right] \\
&= \alpha\sum_{i=0}^{\infty}(1-\alpha)^i\ \mathbb{E}[R_{t-i-1}] \\
&= p
\end{aligned}$$

by linearity of expectation $\mathbb{E}[\alpha X] = \alpha\mathbb{E}[X]$, simplification of the geometric series $\sum_{i=0}^{\infty}(1-\alpha)^i = 1/\alpha$, and expectation of the Bernoulli reward $\mathbb{E}[R_t] = p$.

The reader may verify that the above implies that $\hat{v}$ is an *unbiased* estimator by adding a

bias term b and showing that it is zero: $\mathbb{E}[\hat{v} + b] = p \implies b = 0$. This completes the proof.

3.8 Learning rate controls value variance

Section 3.7 shows that the state value $\hat{v}$ is an unbiased estimate of the mean reward $\mathbb{E}[R] = p$. The variance of this estimate is set by the TD learning rate α according to

$$\mathrm{Var}[\hat{v}] = \frac{\alpha}{2 - \alpha} \mathrm{Var}[R].$$

Proof: Recall that the value estimate can be written as

$$\hat{v}_t = \alpha \sum_{i=0}^{\infty} (1 - \alpha)^i R_{t-i-1}.$$

Taking the variance and moving the constant terms out, we obtain

$$\mathrm{Var}[\hat{v}_t] = \alpha^2 \sum_{i=0}^{\infty} (1 - \alpha)^{2i} \mathrm{Var}[R_{t-i-1}],$$

remembering that $\mathrm{Var}[\alpha \sum_i X_i] = \alpha^2 \sum_i \mathrm{Var}[X_i]$ for constant α and independent X_i. Removing time dependence by assuming $\mathrm{Var}[R_{t-i-1}] = \mathrm{Var}[R]\ \forall t, i$ and simplifying the constant term using convergence of the geometric series completes the proof.

3.9 Derivation of true value in trace conditioning experiments

The normalized true value signal $v(t)/v(t_{\text{rew}})$ in a trace conditioning experiment is given by the piecewise function

$$\frac{v(t)}{v(t_{\text{rew}})} = \begin{cases} \frac{\tau}{L_{\text{ITI}}+\tau} e^{\frac{-(L_{\text{cue}}+L_{\text{delay}})}{\tau}} & t \text{ in ITI,} \\ e^{\frac{t-t_{\text{rew}}}{\tau}} & t \text{ in cue or delay,} \\ A\,(1-e^{\frac{t-t_{\text{end}}}{\tau}})+B & t \text{ in reward,} \end{cases}$$

presented in Equation (3.3). (See Methods for the meaning of each variable.)

To derive this equation, we model a trace conditioning experiment as a Markov reward process (MRP). The states in the MRP are labelled $s_0, s_1, s_2, \ldots, s_{M+N}$, where s_0 is the ITI, $s_1, s_2, \ldots, s_M$ are the trial states representing the M time steps between the start of the cue in s_1 and the end of the delay period in s_M, and $s_{M+1}, s_{M+2}, \ldots, s_{M+N}$ are the reward states representing the N time steps during the reward period. The transition probabilities are

$$\begin{aligned} p(s_1 \mid s_0) &= p_{\text{start}} \\ p(s_0 \mid s_0) &= 1 - p_{\text{start}} \\ p(s_0 \mid s_{M+N}) &= 1 \\ p(s_{i+1} \mid s_i) &= 1 \qquad\qquad 1 \le i < M+N \end{aligned}$$

where $p(s_j \mid s_i)$ represents the probability of transitioning to s_j at the next time step given that the current state is s_i. These transition probabilities were chosen so that the dwell time in the ITI state s_0 follows a geometric distribution with mean $L_{\text{ITI}} = \frac{dt}{p_{\text{start}}}$, reflecting exponentially distributed ITI durations commonly used in experiments, and so that the fixed durations of the cue, delay, and reward epochs are given by $L_{\text{trial}} = L_{\text{cue}} + L_{\text{delay}} = M\ dt$ and

$L_{\text{rew}} = N\ dt$. A reward of size r/N is delivered in each of the N reward states, such that each trial ends in a total reward of size r. As noted in the main text, we define the value of each state in terms of the expected discounted future reward

$$v(s) \equiv \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1} \;\middle|\; S_t = s\right],$$

which can also be written in Bellman form as

$$v(s) = \mathbb{E}\left[R_{t+1} + \gamma v(S_{t+1}) \mid S_t = s\right].$$

In the three sections below, we show how each part of the continuous time normalized true value signal $v(t)/v(t_{\text{rew}})$ can be derived from the MRP and value function $v(s)$ given above in the limit of $dt \to 0$.

3.9.1 True value during the ITI

Writing the value of the ITI state $v(s_0)$ using the Bellman form and expanding the expectation shows that it is dependent on itself and the value of the first trial state

$$v(s_0) = \gamma v(s_1)\ p_{\text{start}} + \gamma v(s_0)\ (1 - p_{\text{start}}),$$

which can be solved in terms of $v(s_0)$, yielding

$$v(s_0) = \frac{\gamma v(s_1)\ p_{\text{start}}}{1 - \gamma(1 - p_{\text{start}})}.$$

The first trial state $v(s_1)$ can be written in terms of the final trial state $v(s_M)$

$$v(s_1) = \gamma^{M-1} v(s_M)$$

since the state transitions are deterministic and no rewards are delivered during the cue and

delay epochs.

Substituting the value of the first trial state $v(s_1)$ into the value of the ITI state $v(s_0)$ and normalizing by the peak value during the trial $v(s_M)$ yields

$$\frac{v(s_0)}{v(s_M)} = \frac{\gamma^M p_{\text{start}}}{1-\gamma(1-p_{\text{start}})}.$$

Using $M = (L_{\text{cue}} + L_{\text{delay}})/dt$ and $p_{\text{start}} = dt/L_{\text{ITI}}$ from the problem definition and letting $\gamma = e^{\frac{-dt}{\tau}}$, the relative ITI value can be rewritten

$$\frac{v(s_0)}{v(s_M)} = \frac{dt/L_{\text{ITI}}}{1-e^{\frac{-dt}{\tau}}(1-dt/L_{\text{ITI}})} e^{\frac{-(L_{\text{cue}}+L_{\text{delay}})}{\tau}}.$$

Taking the limit as $dt \to 0$ completes the derivation.

3.9.2 True value during trace and delay epochs

Using the Bellman form of the value function and the fact that no rewards are delivered during the cue and trace periods, the value leading up to reward can be written

$$v(s_i) = \gamma^{M-i} v(s_M) \qquad \text{for } 1 \le i \le M.$$

Normalizing by $v(s_M)$ and converting to continuous time using $\gamma = e^{\frac{-dt}{\tau}}$ and $M = L_{\text{trial}}/dt$ completes the derivation.

3.9.3 True value during the reward epoch

The true value starting in state s_M and continuing through the reward epoch is

$$v(s_{M+i}) = \gamma^{N-i+1} v(s_0) + \sum_{j=0}^{N-i-1} \gamma^j r/N \qquad \text{for } 0 \le i \le N,$$

abusing notation by allowing $\sum_{j=0}^{-1} x = 0$. Removing the contribution of $v(s_0)$ and normalizing out the reward, we are left with

$$v(s_{M+i}) - \gamma^{N-i+1} v(s_0) \propto \sum_{j=0}^{N-i-1} \gamma^j \qquad \text{for } 0 \leq i \leq N.$$

The sum of discounting factors on the right hand side can be rewritten in continuous time as the integral of an exponential discounting kernel

$$\int_0^{L_{\text{rew}}-t} e^{\frac{-x}{\tau}} dx = \tau \left(1 - e^{\frac{-(L_{\text{rew}}-t)}{\tau}}\right) \qquad \text{for } 0 \leq t \leq L_{\text{rew}}.$$

Incorporating the scaling and offset terms A, B to ensure continuity with the normalized value function at $v(t_{\text{rew}})$ and $v(t_{\text{end}})$ completes the derivation.

3.10 Inter-trial interval value reflects value at the start of the next trial

Let the state a be that the animal is in the ITI, let the state b be that the animal is at the very start of a trial, and let T be a random variable that represents the number of time steps remaining in the ITI before the start of the next trial. The true value of the ITI state $v(a)$ is proportional to the value at the start of the next trial $v(b)$ as follows

$$v(a) = v(b)\ \mathbb{E}\left[\gamma^{T+1} \mid S_t = a\right],$$

where γ is the discounting factor.

Proof: Recall that the true value of a state S, s is defined as the expected sum of exponentially-discounted future rewards

$$v(s) = \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1} \,\middle|\, S_t = s\right].$$

Applying this definition to the states a, b, we obtain

$$v(a) = \mathbb{E}\left[\sum_{i=0}^{T} \gamma^i R_{t+i+1} + \sum_{i=T+1}^{\infty} \gamma^i R_{t+i+1} \;\middle|\; S_t = a\right]$$

and

$$v(b) = \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1} \;\middle|\; S_t = b\right],$$

where the sum in $v(a)$ is split between the value of the rest of the ITI and the discounted value at the start of the next trial. Since the value at the start of the next trial is $v(b)$, we can rewrite the ITI value as follows

$$v(a) = \mathbb{E}\left[\sum_{i=0}^{T} \gamma^i R_{t+i+1} + \gamma^{T+1} v(b) \;\middle|\; S_t = a\right].$$

Observing that the cumulative reward during the ITI is zero by definition $\sum_{i=0}^{T} \gamma^i R_{t+i+1} = 0$, and also observing that $v(b)$ is a constant that can be factored out of the expectation, we find that the ITI value is

$$v(a) = v(b)\ \mathbb{E}\left[\gamma^{T+1} \;\middle|\; S_t = a\right],$$

completing the proof.

Notes:

- This proof can easily be extended to the value of any state in the upcoming trial (not just the first state) by redefining T and b.
- If the ITI durations are drawn from a geometric (or exponential) distribution, then the scaling factor $\mathbb{E}\left[\gamma^{T+1} \;\middle|\; S_t = a\right]$ does not depend on the amount of time spent in the ITI so far. The value during the ITI $v(a)$ is therefore constant.

3.11 Relationship between uncertainty and reward variance

Absolute RPEs $|\delta|$ are sometimes used in RL as a measure of variability or uncertainty that can be used to regulate learning [25], but the precise statistical interpretation of $|\delta|$ is unclear. Here we show that for vanilla TD learning from Bernoulli rewards, the average absolute RPE is proportional to the reward variance

$$\mathbb{E}[|\delta|] = 2\,\mathrm{Var}[R],$$

but this relationship can be distorted by asymmetric learning rates and forgetting.

Proof for vanilla TD: For simplicity, let $R \sim \text{Bernoulli}[p]$, let the RPE be $\delta = R - \hat{v}$, and set value to its fixed point $\hat{v} = p$ (since $\mathbb{E}[\Delta\hat{v}] = 0 \iff \mathbb{E}[\delta] = 0 \iff \hat{v} = p$ under TD learning). Substituting p into the RPE and taking the expectation of the absolute value, we get

$$\mathbb{E}\left[|\delta|\right] = \mathbb{E}\left[|R - p|\right].$$

Since the reward is binary $R \in \{0, 1\}$, we can easily rewrite the expectation above and simplify

$$\begin{aligned}\mathbb{E}\left[|\delta|\right] &= |1 - p|p + |0 - p|(1 - p) \\ &= 2p(1 - p) \\ &= 2\,\mathrm{Var}[R],\end{aligned}$$

completing the proof.

The relationship between absolute RPE $|\delta|$ and reward variance shown analytically above holds well in practice even if the value estimate $\hat{v}$ is not generally exactly equal to its fixed point p (Figure 3.13).

Effect of asymmetric learning: In vanilla TD learning, the value estimate is updated

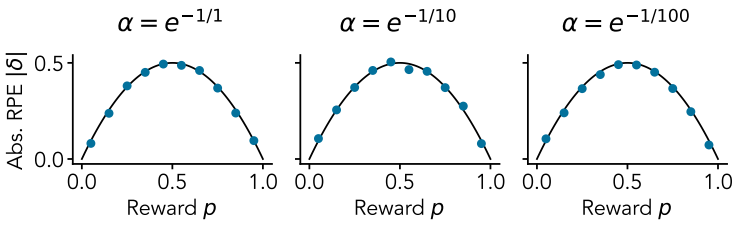


Figure 3.13: Numerical verification of relationship between absolute RPE $|\delta|$ and reward variance $\mathrm{Var}[R]$ under TD learning. Panels show $|\delta|$ mean $\pm$ SEM (blue points; error bars are too small to be visible) calculated from simulated TD learning trials against $2\,\mathrm{Var}[R]$ (black lines). Simulations were burned in for 500 trials and RPE statistics were calculated on 2000 trials. Learning rate α used in the TD model is indicated at top of each panel.

proportional to the RPE $\Delta\hat{v} = \alpha\delta$, but a common extension is to use different learning rates for positive and negative RPEs

$$\Delta\hat{v} = \begin{cases} \alpha_+\delta & \text{if } \delta \geq 0 \\ \alpha_-\delta & \text{if } \delta < 0, \end{cases}$$

called asymmetric learning. Under asymmetric learning, value has a fixed point that is different from the reward probability $\mathbb{E}[\Delta\hat{v}] = 0 \not\Longleftrightarrow \hat{v} = p$. We can find the fixed point by solving

$$\begin{aligned} 0 &= \mathbb{E}[\Delta\hat{v}] \\ 0 &= \alpha_+(1-\hat{v})p - \alpha_-\hat{v}(1-p) \\ \hat{v} &= \frac{p}{p + \frac{\alpha_-}{\alpha_+}(1-p)} \end{aligned}$$

Substituting the above into the expectation of the absolute RPE $\mathbb{E}[|\delta|] = \mathbb{E}[|R - \hat{v}|]$, we find that there is no longer a clear connection to the variance of the Bernoulli reward $\mathrm{Var}[R]$ (Figure 3.14).

Effect of forgetting: Another common modification of TD learning is to include a forgetting rate $0 \leq \zeta \leq 1$ in the RPE $\delta = R - \hat{v} - \zeta\hat{v}$. Solving $\mathbb{E}[\Delta\hat{v}] = 0$ as above, we find

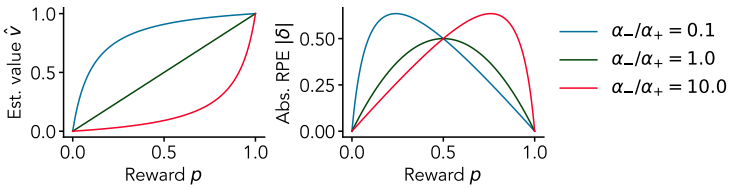


Figure 3.14: Effect of asymmetric learning rates on value $\hat{v}$ and absolute RPE $|\delta|$ in TD learning.

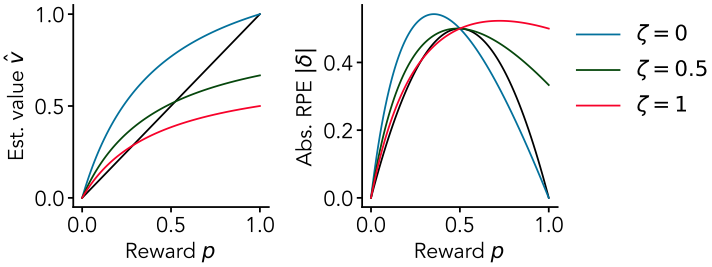


Figure 3.15: Effect of forgetting on value $\hat{v}$ and absolute RPE $|\delta|$ under asymmetric TD learning with $\frac{\alpha_-}{\alpha_+} = 0.3$.

the fixed point

$$\hat{v} = \frac{p}{\left[p + \frac{\alpha_-}{\alpha_+}(1-p)\right](1+\zeta)}.$$

The main effect of forgetting is to decrease the estimated value when the reward probability is high, increasing the absolute RPE in this range (Figure 3.15).

3.12 Perturbation analysis

Setup: According to the value prediction theory, the trial-aligned activity patterns of serotonin neurons in reward learning experiments should resemble the normalized true value signal derived in Section 3.9 scaled by the reward probability. Therefore, if an animal's estimate of the reward probability is dynamically changing, then the activity of serotonin neurons should scale up and down accordingly. Assuming that the animal estimates the reward probability on the basis of the proportion of recent trials that were rewarded, the value prediction theory makes a simple testable prediction: in a given experiment, the firing rates of serotonin neurons should be positively correlated to the proportion of recent trials that were rewarded. We do no

know, however, the precise timescale over which the recent rewards affect the value function. To circumvent this problem, we use perturbation theory to derive a simpler expression relating the value with the fraction of recent trials that were rewarded.

Our value prediction theory states the firing rate of serotonergic neurons is proportional to the predictively coded value signal. Over the slow timescale of the whole-trial, the predictively coded value signal becomes the value signal. Our goal is thus to relate the value signal

$$v_t = \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1} \;\middle|\; S_t = s\right],$$

with fluctuations in the recent reward history $\hat{p}_t = \frac{1}{L}\sum_{k=0}^{L-1} r_k$, for some small number of recent trials L and where r_k refers to the kth reward in the past with respect to time t.

We begin by proving that $v_t \propto \tilde{v}_t \mathbb{E}[R]$, where the normalized value signal $\tilde{v}_t$ from Section 3.9 captures within-trial value dynamics and the expected reward $\mathbb{E}[R]$ is responsible for trial-to-trial fluctuations. By definition,

$$v_t = \tilde{v}_t \; v(s_M),$$

where $v(s_M)$ is the true value just before reward delivery (see Section 3.9), so it remains only to be shown that $v(s_M) \propto \mathbb{E}[R]$. Assuming the reward lasts only one time step, we can write

$$v(s_M) = \mathbb{E}[R] + \gamma v(s_0),$$

where $v(s_0)$ is the true value during the ITI. Using the fact that $v(s_0) \propto v(s_M)$ from Section 3.10, we can introduce a temporary proportionality constant C and simplify

$$v(s_M) = \mathbb{E}[R] + \gamma C v(s_M)$$
$$v(s_M) = \mathbb{E}[R]/(1 - \gamma C)$$
$$v(s_M) \propto \mathbb{E}[R].$$

Thus we have that

$$v(s_M) \propto \mathbb{E}[R] \implies v_t \propto \tilde{v}_t \mathbb{E}[R].$$

The proof can be extended to rewards that last more than one timestep without much difficulty.

There are multiple ways of calculating the expected reward $\mathbb{E}[R]$, but in TD learning methods, the expected reward is an exponential moving average of past rewards (Section 3.7). Therefore, the value signal fluctuates according to an unknown value estimation timescale τ_e

$$\hat{v}_t = \tilde{v}_t(1 - e^{-T/\tau_e}) \sum_{k=0}^{\infty} e^{-Tk/\tau_e} r_k$$

Using the fact that the whole-trial duration $T = M + N + L_{\text{ITI}}$ times the trial horizon L is much smaller than the estimation timescale τ_e, by first order Taylor expansion of the exponential terms we get

$$\hat{v}_t = \tilde{v}_t \left[\frac{T}{\tau_e} L \hat{p}_t + c \right]$$

where $c = \frac{T}{\tau_e} \sum_{k=L}^{\infty} e^{-Tk/\tau_e} r_k \approx \bar{r}(1 + TL/\tau_e)$ where $\bar{r}$ is the average reward on a long time horizon. The relative change in value is then

$$\frac{\hat{v}_t - \bar{v}}{\bar{v}} = (\bar{r}(1 + \tau_e/TL))^{-1} \approx 2.5\,\%$$

using an estimation time scale of $\tau_e = 200$ trials [57, 58], an $L = 5$ trial horizon, and an average reward rate of $\bar{r} = 0.5$. This implies that if the average firing rate of a serotonin neuron is 2 Hz, we expect to observe activity modulations of roughly 0.05 Hz in either direction around the mean based on a five-trial reward history.

Note: As a rough guide to the sensitivity of this calculation, consider that if the estimation time scale were an order of magnitude faster than what has previously been reported [57, 58] ($\tau_e = 20$ trials), then we would expect activity modulations of roughly 0.8 Hz (40 % of a 2 Hz baseline). Therefore, changes in serotonergic activity associated with a five-trial reward history

should be much smaller than the multi-Hz within-trial fluctuations in activity explained by $\tilde{v}$ and observed experimentally [24], even if our estimate of the value estimation timescale τ_e is badly wrong.

3.13 Predictive coding exactly cancels leaky decoding

Let $I(t)$ be the net input to the DRN, let $f(\cdot)$ be the encoding function of the DRN, and let $g(\cdot)$ be the decoding function of some downstream region. The output of the DRN is $(f \circ I)(t)$ and the decoded signal is $(g \circ f \circ I)(t)$. Assuming that g performs leaky integration on its input, the decoded signal is given as follows

$$(g \circ f \circ I)(t) = -\int_0^{-\infty} (f \circ I)(t+z)\ \frac{1}{\tau} e^{z/\tau}\ dz. \tag{3.5}$$

Following our previous work [158], assuming f predictively encodes its input such that

$$(f \circ I)(t) = \tau \frac{dI}{dt} + I(t), \tag{3.6}$$

then

$$(g \circ f \circ I)(t) = I(t).$$

Proof: Substituting Eq. 3.6 in Eq. 3.5 and expanding the resulting integral yields

$$\begin{aligned}(g \circ f \circ I)(t) = &-\int_0^{-\infty} \tau \frac{dI}{dt}\Big|_{t+z}\ \frac{1}{\tau} e^{z/\tau}\ dz \\ &-\int_0^{-\infty} I(t+z)\ \frac{1}{\tau} e^{z/\tau} dz.\end{aligned} \tag{3.7}$$

The time derivative of the input $\frac{dI}{dt}$ can be removed from the first term using integration by parts, yielding

$$-\left[\tau I(t+z)\ \frac{1}{\tau} e^{z/\tau}\right]_{z=0}^{-\infty} + \int_0^{-\infty} I(t+z)\ \frac{1}{\tau} e^{z/\tau} dz.$$

Substituting the above into Eq. 3.7 causes the remaining integrals to cancel and changes the sign on the non-integral term, giving

$$(g \circ f \circ I)(t) = \left[\tau I(t+z)\ \frac{1}{\tau} e^{z/\tau}\right]_{z=0}^{-\infty},$$

which simplifies to

$$(g \circ f \circ I)(t) = I(t),$$

completing the proof.

Chapter 4

General discussion

> The more things change, the more they stay the same.
>
> Jean-Baptiste Alphonse Karr

In this chapter, I summarize the main results of my work and briefly discuss some general implications of my approach.

4.1 Summary

What does serotonin do in the brain? In this thesis, I have made the case that a partial answer can only be found by first posing another question: How? In Chapter 1, I introduced the idea that value could play an important and multifaceted role in regulating learning and behaviour, just as serotonin does. More importantly, I argued that value might not be represented in the brain in a straightforward way. In Chapter 2 (Manuscript I), I studied the representational features of the serotonin system from a signal processing perspective. By constructing an experimentally-validated model of the DRN and simulating its responses to simple inputs such as steps and ramps, I arrived at the conclusion that whatever quantity the serotonin system might receive as an input, that quantity is transformed into a mixture of its intensity and rate of change before being broadcasted throughout the forebrain. This transformation, which can be seen as a type of predictive coding in the redundancy-reduction sense, has its origins in one of the best known electrophysiological features of serotonin neurons: unusually strong and long-lasting spike-frequency adaptation. Finally, in Chapter 3 (Manuscript II), I made a conceptual leap of faith by assuming that the DRN receives a value signal as its net input. Through simulations of trace conditioning experiments from the serotonin literature and re-analysis of previously published data, I showed that the *in vivo* activity patterns of serotonin neurons are surprisingly well-explained by this model, leading to what I call the value prediction theory of serotonin. Overall, my results suggest that the puzzling responses of serotonin neurons to reward-predicting cues, rewards, and punishments are best interpreted as an adapting code for future reward.

4.2 The Goldilocks zone of biological detail

If there is a central theme to this work beyond the computational features of the serotonin system, it is that models with an intermediate level of biological detail can be very helpful. There is a certain amount of tension in modern computational neuroscience between proponents of highly-detailed biophysical models and advocates of extremely simple models, often borrowed from the machine learning field, that abstract away even the most basic elements of neural systems such as electrical potentials and synapses [265–270]. There are striking successes and legitimate criticisms to be made on both sides of this debate, some of which I surveyed in Chapter 1. My work follows in a modelling tradition that attempts to blur the lines between simple and complex [271], capitalizing on the practical strengths of both extremes and providing intuitive links between levels of description [154, 272].

Could I have used simpler or more complex models in this work? Perhaps. A simpler model than the aGIF could certainly have captured the strong spike frequency adaptation of serotonin neurons (Chapter 2). At the same time, such a model might not reproduce the activity patterns of serotonin neurons with the same level of accuracy (Fig. 2.3), could not have elucidated the role (or lack thereof) of I_A in DRN signal processing, and, almost paradoxically, the parameters in the simpler model might even have been more difficult to interpret (Fig. A.5; [273]). Had we chosen to use a detailed biophysical model (and overcome the major technical challenges associated with this approach [134, 191]), the strength and duration of adaptation (Fig. 2.4) could not have been measured in an input-independent manner, and empirically verifying their contribution to DRN signal processing by replacing the potent adaptation of serotonin neurons with the weaker adaptation found in SOM neurons would not have been possible (Fig. 2.5). At least when it comes to interrogating the signal processing features of the DRN, the aGIF modelling framework exists in a relatively-interpretable Goldilocks zone of biological detail.

What about the rate-based DRN model used in Chapter 3? There is very little room to

simplify this model without arriving at a static rate code, a model that at face value seems unlikely to be accurate (Section 1.3) and that is decisively ruled out by this work. A slightly simpler model in which the DRN directly encodes a mixture of the intensity and temporal derivative of its input (rather than simply approximating this mixture through adaptation [126]) would have produced the same results[1] at the expense of a less clear connection to spike frequency adaptation. In principle, the much more detailed spiking model of the DRN developed in Chapter 2 could have been used to simulate the *in vivo* responses of serotonin neuron populations to rewards and punishments. This approach might have shed light on the relevance of the contextual dependence of the derivative-like computation shown in Fig. 2.7, but at considerable computational cost [134] and at the risk of obscuring the striking implications of adaptation for the interpretation of serotonergic activity patterns. Overall, there would have been less to lose but also less to gain from using a simpler or more complex model in the second part of this thesis.

To summarize, this work illustrates that the dichotomy between highly simplified and highly detailed models is a false one. Selectively augmenting simple models with biological detail can be a fruitful approach to uncovering how neural systems process information.

4.3 Conclusion: Neural coding far from the sensorimotor periphery

The insights of this work into the function of the serotonin system turn around the distinction between the meaning of serotonin neuron activity and how this meaning is encoded in spiking statistics. This distinction was not at the top of my mind when I first became interested in serotonin, but is a central fixture of some areas of neuroscience (*e.g.*, [274]).

A great deal of effort has been dedicated to understanding how interactions with the external world are represented in the vertebrate brain. In some lucky cases, elementary

[1] In fact, did produce the same results in a preliminary version of this work.

aspects of these representations are fairly simple and can be intuitively grasped without the aid of mathematical models. In the case of primary visual cortex, Adrian's analogy of cortex as a screen onto which sensory stimuli are projected [275] does not entirely miss the mark [276, 277], and the vectorized representation of reaching direction in motor cortex is straightforward to understand (at least with the aid of a figure) [278]. Other cases strain intuition. Receptive fields in higher visual areas are contorted into psychedelic patterns [279], and neural representations of high-level decisions and movement preparation exhibit drifting or rotational dynamics that are challenging to interpret [103, 104, 280]. It often seems as though the closer one looks, the more abstract and difficult to understand neural representations become—especially far from the sensorimotor periphery.

Together with these anecdotes, my work contains an important lesson for the study of brain structures involved in higher-order aspects of learning and behaviour: when the *what* of neural activity under the microscope seems obtuse or inconsistent [76], it might only be possible to make progress by first asking *how*?

Appendices

Appendix A

A user's guide to generalized integrate-and-fire models

This chapter is reprinted from a previously-published work.

Citation

E.F. Harkin, J.-C. Béïque, and R. Naud. A user's guide to generalized integrate-and-fire models. In *Computational Modelling of the Brain: Modelling Approaches to Cells, Circuits and Networks*, pp. 69–86. (Springer, 2021).

Statement of contributions

I conceptualized the general approach of selectively augmenting the generalized integrate-and-fire model of Richard Naud with Hodgkin-Huxley conductances. I wrote this textbook chapter in 2020 with input and feedback from Richard Naud and Jean-Claue Béïque.

I had previously designed a GIF model augmented with a Hodgkin-Huxley model of an A-type potassium current to capture the activity patterns of serotonin neurons during my M.Sc. in neuroscience [153]. However, the concept of Hodgkin-Huxley-augmented GIF models as a general technique for modeling the activity patterns of different cell types falls within the scope of my doctoral work.

Abstract

The generalized integrate-and-fire (GIF) neuron model accounts for some of the most fundamental behaviours of neurons within a compact and extensible mathematical framework. Here, we introduce the main concepts behind the design of the GIF model in terms that will be familiar to electrophysiologists, and show why its simple design makes this model particularly well-suited to mimicking behaviours observed in experimental data. Along the way, we will build an intuition for how specific neuronal behaviours, such as spike-frequency adaptation, or electrical properties, such as ionic currents, can be formulated mathematically and used to extend integrate-and-fire models to overcome their limitations. This chapter will provide readers with no previous exposure to modelling a clear understanding of the strengths and limitations of GIF models, along with the mathematical intuitions required to digest more detailed and technical treatments of this topic.

A.1 Introduction to leaky integrate-and-fire models

From an electrical perspective, a neuron is like a rubber balloon in the process of being inflated. The cell membrane of the neuron separates electrically charged ions inside and outside the cell just as the balloon separates molecules of air, and there is a difference in the distribution of charges (i.e., voltage) across the membrane of a cell just at there is a difference in pressure across the membrane of the balloon. Synaptic inputs to the neuron alter the voltage across the membrane just as adding or removing air from the balloon alters the pressure difference. The amount of air needed to appreciably change the pressure inside the balloon depends on its size, just as the number of charges needed to change the membrane voltage depends on the surface area of the membrane. If the balloon is not tied shut, the air inside will slowly leak out, just as charges leak across the cell membrane. Finally, if the difference in pressure across the rubber membrane of the balloon becomes too great, the balloon will pop, just as a neuron will fire an action potential, also called a spike, if the voltage exceeds a threshold. Once the balloon pops, the process restarts with a new, uninflated balloon, just as a neuron will reset to a lower voltage after firing an action potential.

This rubber balloon neuron model captures several of the most important electrical features of neurons, but some important details related to spiking behaviour are still missing. We will see how these can be addressed in Section A.2. For now, however, we will focus on developing a mathematical description of our rubber balloon model which will serve as a foundation for the rest of the chapter.

The leaky integrate-and-fire (LIF) neuron model first introduced by [147], rooted in the work of [144] from over a century ago, captures the same properties of neurons as our rubber balloon analogy. The LIF models the membrane voltage $V(t)$ in terms of its rate of change

dV/dt and behaviour at spike threshold V_T as follows

$$\frac{dV}{dt} = \frac{1}{C}\left(-g_l(V(t) - E_l) + I_{\text{ext}}(t)\right), \qquad V(t) < V_T \tag{A.1}$$

$$\begin{cases} V(t + t_{\text{ref}}) & \leftarrow V_{\text{reset}} \\ t & \leftarrow t + t_{\text{ref}} \end{cases}, \qquad V(t) \geq V_T \tag{A.2}$$

where g_l represents the leakiness of the membrane, which is equivalent to $1/R_m$, where R_m is the membrane resistance; E_l represents its equilibrium voltage, also called the resting membrane potential; $I_{\text{ext}}(t)$ represents external inputs to the neuron; C is the membrane capacitance, which reflects the membrane surface area; and t_{ref} is the duration of the absolute refractory period after a spike. According to Eq. (A.1), the membrane voltage relaxes towards its equilibrium E_l in the absence of any input $I_{\text{ext}}(t)$. This happens more quickly if the membrane is very leaky (dV/dt increases with g_l) or if the voltage is far from equilibrium (dV/dt increases with $V(t) - E_l$), and more slowly if the membrane surface area is very large (dV/dt decreases with increasing C), consistent with our balloon analogy. Whenever the voltage reaches V_T a spike occurs instantaneously, like the popping of a balloon, and the dynamics of Eq. (A.1) no longer apply. Instead, Eq. (A.2) specifies that the voltage should be reset to a lower value V_{reset} after a short refractory period t_{ref}. Unlike in many other neuron models, the membrane voltage *during* the spike is not defined—for a LIF model, a "spike" is not spike-shaped at all, it has no shape because the voltage is not defined. This simplification takes advantage of the fact that the shape of the action potential does not carry any meaningful information.

Comparison of LIF and Hodgkin-Huxley models

The decision not to model the dynamics of the membrane voltage during the action potential is one of the most important conceptual differences between LIF models, which have their roots in the early 20th-century work of [144], and the biophysically-realistic models first introduced by [127] half a century later.
Hodgkin and Huxley's account of the biological mechanisms of action potential generation in squid axon included a model of membrane voltage dynamics as a sum of voltage-dependent ionic currents

$$C\frac{dV}{dt} = I_{\mathrm{Na}}(t) + I_{\mathrm{K}}(t) + I_{\mathrm{ext}}(t),$$

where $I_{\mathrm{Na}}(t)$ and $I_{\mathrm{K}}(t)$ are voltage-dependent sodium and potassium currents, respectively. Unlike in Lapicque's LIF model, the membrane voltage in Hodgkin and Huxley's model is always defined, including during the action potential. This detailed and realistic approach to modelling the voltage dynamics of neurons has been enormously influential, to the point that neuron models that are based on detailed descriptions of ionic currents are commonly referred to as Hodgkin-Huxley-style models.
With the advent of more realistic and detailed Hodgkin-Huxley models, are simplified LIF models still of any use? When comparing these two types of models, it is important to keep in mind that they were created for different purposes. The original model of Hodgkin and Huxley was designed to explain how the interactions of two specific ionic currents give rise to voltage spikes with a particular shape, which we call action potentials. The original model of Lapicque, on the other hand, was created to describe how different electrical stimuli impacted the rate of action potential discharge in frog nervous tissue. For Lapicque's work, a description of the shape of the action potential (and the biophysical mechanisms responsible for it) was not necessary, and LIF models continue to be used in cases when the relationship between electrical input and the *timing* or *rate* of output spikes is of primary interest.
The negative consequences of omitting necessary biological components from a model are obvious, but there are also more subtle drawbacks of including unnecessary detail. More complicated models are usually more difficult to design and constrain to mimic specific neurons of interest, and can be liable to produce inaccurate predictions as a result (for reasons we will touch on in Section A.3.1). Highly detailed models also take more time to simulate simply because more calculations are required per time step.
To summarize, LIF and Hodgkin-Huxley models each describe neurons at different levels of detail because they were created for different purposes, and each style of model comes with its own compromises.

The highly simplified LIF model provides an intuitive account of some of the most basic electrical features of neurons. In particular, the one-to-one correspondence between

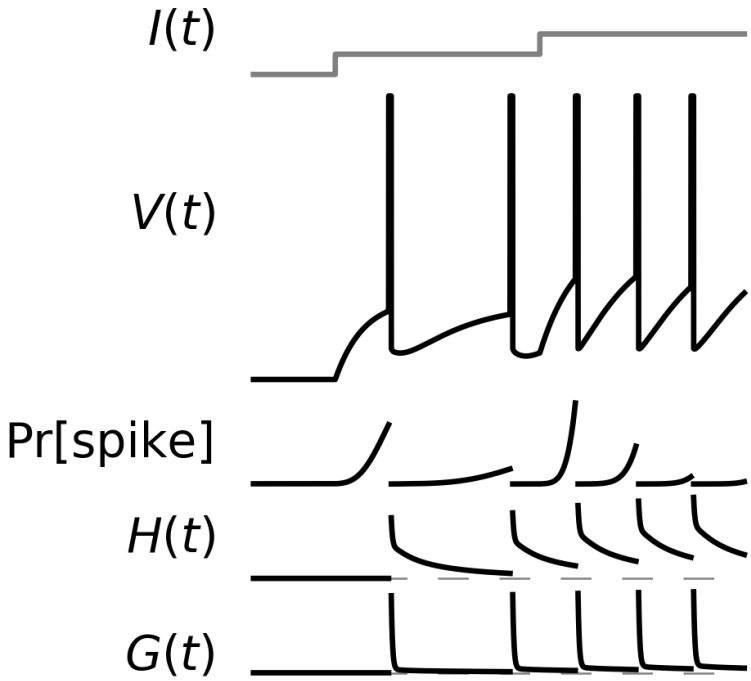


Figure A.1: The generalized integrate-and-fire model. A current input $I(t)$ produces a subthreshold voltage $V(t)$ which is translated into a probability of emitting a spike Pr[spike]. Spikes are emitted stochastically and cause the voltage to reset to a lower value after a short delay, as in the leaky integrate-and-fire model. Spike frequency adaptation is caused by the spike-triggered adaptation current $H(t)$ and threshold movement $G(t)$. Notice that spikes introduce discontinuities into the spike probability and adaptation mechanisms because the dynamics of the GIF model during a spike are not defined. There are also discontinuities in the membrane voltage $V(t)$ for the same reason; however, to improve clarity, here we have set the voltage to an arbitrary high value during each spike. The dashed gray lines in $G(t)$ and $H(t)$ indicate zero.

model parameters and fundamental properties such as resting membrane potential and spike threshold makes the LIF model straightforward to interpret. In the coming sections, we will first discuss how the LIF model can be extended to account for spike-frequency adaptation and stochastic firing in a similarly intuitive way, then show how these intuitive definitions lend themselves particularly well to being fitted to data.

A.2 Generalizing the leaky integegrate-and-fire model

A.2.1 Spike-triggered adaptation

In response to a step stimulus, many neurons initially fire action potentials at a high rate which then decreases gradually to a lower rate. This phenomenon, known as spike-frequency adaptation (SFA), plays many important roles in neural systems such as enhancing the detection of weak stimuli and computing the rate of change of an input [126]. SFA arises from the effects of spike-triggered adaptation currents, which tend to push membrane voltage away from spike threshold, as well as the inactivation of the biophysical mechanisms that cause spiking, resulting in an effective change in the spike threshold. In this section, we will show how to extend the LIF model to incorporate these ideas in an explicit and intuitive way.

On a conceptual level, SFA mechanisms can be divided into two categories: mechanisms which move the membrane voltage away from threshold, such as adaptation currents; and mechanisms which move the threshold away from the membrane voltage, such as inactivation of spiking mechanisms. We can capture these two categories mathematically in terms of a pair of functions $\eta(x)$ (adaptation current) and $\gamma(x)$ (threshold movement) for $x > 0$ where x is the time since a spike. In general, adaptation currents and threshold changes are large just after a spike, but fade away over time. In order to give $\eta(x)$ and $\gamma(x)$ these properties, we can define them mathematically using decaying exponentials $e^{-x/\tau}$, which fade away to zero as x increases at a rate dictated by the timescale τ. Depending on the cell type and even the individual neuron, these adaptation mechanisms may fade away quickly, slowly, or even over multiple timescales. We can accommodate this variability by implementing $\eta(x)$ and $\gamma(x)$ as *weighted sums* of exponentials with different timescales τ

$$f(x) \equiv \sum_i w_i^{(f)} e^{-x/\tau_i^{(f)}}, \qquad \tau_i^{(f)} > 0, x > 0 \tag{A.3}$$

where the $w_i^{(f)}$ are the weights and f corresponds to η or γ as appropriate. If a particular

timescale $\tau_i^{(f)}$ is not found in a given neuron, its associated weight $w_i^{(f)}$ can simply be set to zero to remove its contribution to the overall adaptation function $f(x)$. To allow the effects of adaptation to build up over multiple spikes, we can define $H(t)$ and $G(t)$ as the sums of η and γ over the set of all past spikes $\{s \in \mathcal{S}; s < t\}$

$$H(t) \equiv \sum_{\{s \in \mathcal{S}; s < t\}} \eta\,(t - s) \tag{A.4}$$

$$G(t) \equiv \sum_{\{s \in \mathcal{S}; s < t\}} \gamma\,(t - s) \tag{A.5}$$

where $\mathcal{S}$ is the set of all spike times, s is the time of a specific spike, and t is the current time. This means that if several spikes occur in quick succession, the adaptation mechanisms $H(t)$ and $G(t)$ will be engaged more strongly than if fewer spikes had occurred. Additionally, since both η and γ go to zero as the time since a spike $t - s$ increases, past spikes contribute less and less to the adaptation functions $H(t)$ and $G(t)$ as time goes on.

The LIF neuron model presented in the previous section does not account for spike-frequency adaptation, but this is easily remedied by incorporating an adaptation current and spike-triggered threshold movement into the model via $H(t)$ and $G(t)$. In order to do that, we must first redefine the fixed spike threshold V_T to be a function $V_T(t)$ that returns the spike threshold at a specific time t, taking the effects of previous spikes into account via the threshold movement $G(t)$

$$V_T(t) \equiv V_T^* + G(t),$$

where V_T^* is the spike threshold after all adaptation has faded away. This equation shows why we call $G(t)$ the threshold movement: it gives the amount that the spike threshold has moved as a result of adaptation.

Next, we must incorporate the adaptation current $H(t)$ into the subthreshold dynamics given in Eq. (A.1) so that it can influence the subthreshold voltage. Simply subtracting $H(t)$ from the other currents in Eq. (A.1) completes the definition of our LIF model with

adaptation

$$\frac{dV}{dt} = \frac{1}{C}\left(-g_l(V(t) - E_l) - H(t) + I_{\text{ext}}(t)\right), \qquad V(t) < V_T(t) \tag{A.6}$$

$$\begin{cases} V(t + t_{\text{ref}}) & \leftarrow V_{\text{reset}} \\ t & \leftarrow t + t_{\text{ref}}, \end{cases} \qquad V(t) \geq V_T(t). \tag{A.7}$$

Notice that Eq. (A.6) tells us that when the adaptation current $H(t)$ increases after a spike, it is subtracted away from the input current $I_{\text{ext}}(t)$. This means that the adaptation current can effectively reduce the strength of an input stimulus. Since the LIF fires less rapidly in response to weaker inputs, the adaptation current $H(t)$ causes the firing rate evoked by constant stimulus to drop off after the first few spikes, consistent with the SFA effect we seek.

A.2.2 Stochasticity

The LIF model presented here fires a spike instantaneously when the voltage exceeds a threshold $V_T(t)$, but in real neurons this process is less precise [114]. In general, neurons are more *likely* to spike when their membrane voltage exceeds a threshold, but it is also possible for spikes to occur when the voltage is below threshold or vice-versa. It can therefore be helpful to think of the relationship between voltage and spiking in terms of a *probability* that depends on the current voltage and threshold, rather than as a hard cutoff. We can express this idea mathematically in terms of a spike probability function

$$\rho\left(\frac{V(t) - V_T(t)}{\sigma}\right) = \Pr[\text{spike between time } t \text{ and } t + \Delta t \;\; | \, V(t) - V_T(t)], \tag{A.8}$$

where $\Pr[x \mid y]$ denotes the probability of x given y, σ is a scaling factor that sets the degree of stochasticity, and Δt is a small time interval. To capture our intuition that spikes are more likely when $V(t)$ is above $V_T(t)$, $\rho\left(\frac{V(t)-V_T(t)}{\sigma}\right)$ is usually a sigmoidal function that is close to 1 when $V(t) - V_T(t) > 0$ and close to 0 when $V(t) - V_T(t) < 0$, such as the one shown

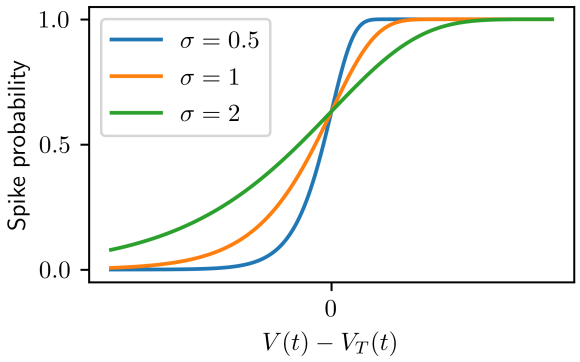


Figure A.2: Typical shape of the spike probability function given in Eq. (A.8). The vertical axis shows the probability of spiking within a small time window $[t, t + \Delta t)$ and the horizontal axis shows the distance between the membrane voltage $V(t)$ and spike threshold $V_T(t)$. Notice that increasing the stochasticity scaling factor σ increases the probability of spiking below threshold. The spike probability function shown here is $\rho\left(\frac{V(t)-V_T(t)}{\sigma}\right) = \exp\left[-e^{\frac{V(t)-V_T(t)}{\sigma}}\Delta t\right]$, which is used by the GIF model as previously described [151, 158, 203, 204, 281].

in Fig. A.2. The scaling factor σ allows us to control how quickly the probability of spiking increases when the voltage is above threshold. In more intuitive terms, σ sets the threshold *sharpness*.

Incorporating this stochastic spiking behaviour into the LIF model with adaptation will complete our definition of the generalized integrate-and-fire model (GIF). To accomplish this, we simply need to invoke the spiking rule given in Eq. (A.7) *probabilistically* according to the spike probability function given in Eq. (A.8) rather than deterministically whenever the voltage exceeds a given threshold. The pseudocode in Algorithm 1 sketches how this can be done, and an example of a model simulated this way is shown in Fig. A.1.

A.2.3 Simplifications, generalizations, and limitations

The generalized integrate-and-fire model we have presented here accounts for more features of neuronal excitability than the leaky integrate-and-fire model from Section A.1, but it is still a highly simplified model. Compared with the LIF model, the GIF model accounts for two important phenomena related to spiking: spike-triggered adaptation and a stochastic threshold. However, the GIF model does not account for subthreshold adaptation that is

Algorithm 1 Simulation procedure for the GIF model.

Require: $I_{\text{ext}}(t)$ ▷ Input current.
Require: $C, g_l, E_l, \eta, \gamma, V_T^*, \sigma, \rho$ ▷ Model components.
Require: $V_0, \Delta t, T$ ▷ Initial voltage, simulation timestep, and duration of simulation.
Require: $\xi(t)$ ▷ Random number between 0 and 1 sampled at time t.

▷ Set initial condition.
$t \leftarrow 0$
$V(t) \leftarrow V_0$
$\mathcal{S} \leftarrow \{\emptyset\}$
while $t < T$ **do**
  $G(t) \leftarrow \sum_{\{s \in \mathcal{S}; s<t\}} \gamma(t-s)$ ▷ Compute threshold movement from Eq. (A.5).
  $V_T(t) \leftarrow V_T^* + G(t)$ ▷ Compute spike threshold with adaptation.
  if $\xi(t) \leq \rho\left(\frac{V(t)-V_T(t)}{\sigma}\right)$ **then**
    $\mathcal{S} \leftarrow \mathcal{S} \cup t$ ▷ Add t to the set of spike times.
    ▷ Emit a spike according to Eq. (A.7).
    $V(t+t_{\text{ref}}) \leftarrow V_{\text{reset}}$
    $t \leftarrow t + t_{\text{ref}}$
  else
    $H(t) \leftarrow \sum_{\{s \in \mathcal{S}; s<t\}} \eta(t-s)$ ▷ Compute adaptation current from Eq. (A.4).
    ▷ Integrate membrane dynamics from Eq. (A.6)
    $\Delta V \leftarrow \frac{\Delta t}{C}\left(-g_l(V(t)-E_l) - H(t) + I_{\text{ext}}(t)\right)$
    $V(t+\Delta t) \leftarrow V(t) + \Delta V$
    $t \leftarrow t + \Delta t$ ▷ Increment time.
  end if
end while
return $V(t), \mathcal{S}$ for $0 \leq t < T$

not related to spiking, nor does it account for the effects of the mechanisms that give rise to stochastic spiking on the subthreshold voltage.

Subthreshold adaptation can be produced by voltage-activated currents that oppose the very changes in voltage that cause them to activate, creating the appearance of a sag in the voltage response to step inputs. Prominent examples include the hyperpolarization-activated current I_h and subthreshold voltage-activated potassium currents. From a functional perspective, the main effect of these currents is to filter out inputs that produce slow changes in voltage, while letting through inputs that cause the voltage to change quickly. In real neurons, this frequency filtering effect is specific to the range of subthreshold voltages in which the current activates — if an input current produces a change in voltage that does not activate the current causing subthreshold adaptation, no adaptation is produced. In GIF models, subthreshold adaptation is most easily accounted for by adding a slowly activating leak current. This approach is mathematically simple to implement but disregards potential voltage-specificity of subthreshold adaptation. An alternative approach is to augment the GIF model with explicit Hodgkin-Huxley models of the currents that give rise to adaptation. We will discuss how to augment GIF models with Hodgkin-Huxley components in Section A.3.4.

Stochastic firing in real neurons arises partly from the fast but not instantaneous dynamics of the sodium channels that mediate spiking. In real neurons, sodium channels begin to open when the membrane potential approaches threshold. This pushes the membrane potential even closer to threshold, creating a positive feedback loop between membrane potential and sodium channel activation. Eventually, this process passes a threshold or point of no return and a spike is produced. In the GIF model, this soft positive feedback process is replaced by a probability of spiking. When the positive feedback process is strong enough to very rapidly pass the point of no return (i.e., when the threshold is very sharp), the probabilistic threshold of the GIF model can be a good approximation. However, if this positive feedback process is initially very weak, the activity of sodium channels can impact the subthreshold voltage dynamics without leading to spiking. In this case, the probabilistic threshold of the GIF

model is a worse approximation because it does not capture the effects of spiking mechanisms on the subthreshold dynamics.

The GIF model differs from real neurons in one additional important respect: real neurons exhibit a complex morphological structure, but the GIF model represents a voltage at a single point. The GIF model is most accurate when only inputs and outputs at the level of the cell body are considered. Fortunately, an electrode located at the cell body is often both the main source of input and instrument for measuring output during *in vitro* electrophysiological experiments. Unfortunately, under physiological conditions, most of the input to a neuron arrives via synapses located on potentially electrically-distant dendrites, and the point-neuron simplification of the GIF model may be less appropriate.

In sum, while the GIF model offers a simple and intuitive description of many of the electrical properties of neurons, it does not capture *all* of the properties of neurons equally well. When the subthreshold effects of ionic currents, spiking mechanisms, or neuronal morphology are of primary interest, it is important to consider whether the GIF model can be adapted to account for these mechanisms to a satisfactory extent, or whether an entirely different modelling framework should be chosen.

A.3 Fitting the generalized integrate-and-fire model

A.3.1 Finding parameter values: experiments vs. optimization

So far we have presented the GIF model as an intuitive single neuron model formulated in terms of interpretable input and output *variables*, which include input current, voltage, and spike times, and *parameters*, such as membrane leakiness and spike threshold. Next, we turn to the question of how to choose appropriate values for the parameters. In practice, there are two main approaches: carry out a series of detailed electrophysiological experiments to measure each parameter in the model, for example, by applying current steps of various amplitudes to measure the action potential threshold, or use mathematical optimization techniques to

find the parameter values that cause the output of the model to mimic that of a real neuron. Choosing between these two approaches involves a trade-off between the interpretation of model parameters and the accuracy of model predictions. Following the experimental approach, parameter values will have a precise and familiar physiological interpretation, but the output of the model might not mimic that of any particular neuron or cell type very closely. Following the optimization-based approach, the parameter values will represent effective quantities with a potentially less precise physiological interpretation, but the output predicted by the model will match that of a particular neuron as closely as possible. For readers with a strong background in experimental neuroscience, it may come as a surprise that models constructed using experimentally-measured parameter values can produce poor predictions in common cases. In this section, we discuss why this happens and present an alternative optimization-based approach.

One situation in which models constructed using experimentally-measured parameter values produce poor predictions arises when both of the following conditions are met:

1. the true values of the parameters in the neuron population of interest are correlated or are otherwise not statistically independent, and

2. it is not possible to experimentally determine the values of all of the parameters in a single neuron, or this data is not available even if it is possible to collect.

This situation arises very often in practice. This is partly because the first condition is surprisingly easy to meet since it is sufficient for only two of the model parameters to be related (resting membrane potential and action potential threshold, for instance). The second condition is also usually met because it is often impractical to measure each of the model parameters in every neuron, even if it is technically possible. If both conditions are met, the distribution of the experimentally-determined model parameters will not match the true distribution in the population. Since the behaviour of the GIF model is controlled by its parameters, this can easily lead to models that exhibit unexpected behaviours that are not

found in the population of neurons they are intended to mimic.

To see why, consider a hypothetical population of neurons with a similar degree of spike frequency adaptation. Suppose that this adaptation is caused by a variable mixture of an adaptation current and moving threshold, but that the sizes of the adaptation current and moving threshold are anticorrelated such that the overall degree of spike frequency adaptation is roughly constant. If it is not possible to measure both the adaptation current and moving threshold in the same neurons, a researcher might observe that both the adaptation and moving threshold range from small to large, but not realize that they are anticorrelated. This might lead them to create a set of models that includes neurons with both a large moving threshold and adaptation current, even though no such neurons exist in the population. As a result, the models might exhibit more variable spike frequency adaptation than the neuron population [133]. Even if the researcher decides to create a single model of an "average" neuron by setting both the adaptation current and moving threshold to a moderate amplitude, the degree of spike frequency adaptation in the model might be very different from any of the neurons in the population. This is because the non-linear interactions between the adaptation current and moving threshold cause the total amount of spike frequency adaptation to be different from the sum of its parts. Similarly, the average spike frequency adaptation in a population can be different from the averages of its parts. This phenomenon is sometimes called the failure of averaging [132].

In an alternative approach to parameter estimation, an experimenter records the voltage response of a neuron to an input current delivered via an intracellular electrode and the output of the model in response to the same input is forced to match that of the neuron as closely as possible [203]. This approach involves defining a mathematical measure of similarity between the output of the model and the observed behaviour of a neuron and finding a set of parameter values that maximize this measure. In some cases, it is possible to find the similarity-maximizing parameter values directly by taking advantage of mathematical properties of the similarity measure. However, it is often necessary to simply take an initial

guess at the values of the model parameters and then adjust them repeatedly in the direction of increasing similarity until a maximum level of similarity is reached, a process known as gradient ascent. Whether they are obtained via gradient ascent or found directly, the similarity-maximizing values are referred to as the *optimal* values for the corresponding parameters, and the process of finding them is called *optimization* or fitting.

The optimization process can be understood more intuitively by visualizing it in terms of a similarity landscape. If we focus on only two of the model parameters at a time, we can imagine that the two parameters give the latitude and longitude of a point on this landscape and that the similarity defines the altitude of that point. Following this analogy, points of maximal similarity correspond to the tops of hills in this landscape. We can imagine gradient ascent as the process of starting somewhere on the landscape and proceeding uphill. In the machine-learning literature, it is more common to define a mathematical measure of dissimilarity and use gradient descent to find optimal points lying at the bottoms of valleys; for the sake of simplicity, we will focus on similarity/gradient ascent for now. Depending on the model and similarity measure, there can be more than one hill in the similarity landscape[1]. When this is the case, finding the point at the top of the tallest hill becomes quite difficult because we cannot reach it by proceeding uphill from any starting point. (Fortunately, in the case of the GIF model, there is only one hill and finding the parameter values that produce the highest possible similarity is not difficult[2].) This process works the same way when there are more than two parameters, but visualizing a landscape with additional dimensions stretches the imagination.

The main advantages of this optimization-based approach over experimentally determining parameter values are that optimization is less labour intensive and yields models that produce

[1]In this case, the points at the tops of hills are higher than all points within a small neighbourhood, but not necessarily all points in the landscape. After all, there might be taller hills elsewhere. Points that are only optimal within a small neighbourhood are called locally optimal and the point at the top of the tallest hill is called *globally* optimal.

[2]Although the measures of similarity and dissimilarity used by the GIF model will be presented briefly in Section A.3.3, the reasons that these measures are associated with landscapes that have a particular structure are beyond the scope of this chapter. For a thorough introduction, see [136, 281].

more accurate predictions. On the other hand, a potential drawback of this approach is that the parameter values obtained via optimization do not necessarily correspond exactly to experimental measurements. To see why, consider a neuron that is exactly like the GIF model except that it has an additional voltage-dependent conductance

$$\frac{dV}{dt} = \frac{1}{C}\left(-g_l(V(t) - E_l) - g_v(V)(V(t) - E_l) - H(t) + I(t)\right),$$

where $g_v(V)$ is the voltage-dependent conductance. If a researcher were to fit the GIF model to this neuron, the effect of the voltage-dependent conductance would be mixed into the leak conductance of the GIF model

$$-g_l(V(t) - E_l) - g_v(V)(V(t) - E_l) = -(g_l + g_v(V))(V(t) - E_l) \implies \hat{g}_l = g_l + g_v(V),$$

where $\hat{g}_l$ is the value obtained by fitting the GIF model. Notice that not only is there a mismatch between the fitted leak conductance and the true leak conductance of the neuron, $\hat{g}_l \neq g_l$, but, since $g_v(V)$ depends on voltage and $\hat{g}_l = g_l + g_v(V)$, the value obtained for $\hat{g}_l$ depends on the voltage of the neuron! This illustrates that unless the neuron to which the GIF model is fitted *is* a GIF model, the optimal parameter values should not be interpreted as exact substitutes for experimentally-determined values. Instead, they should be interpreted as *effective* values, in the sense that they specify how much a particular variable affects the output of the GIF model within the context of a specific input and set of model components.

The mismatch between optimal model parameter values and experimental measurements is hardly unique to the GIF model. In fact, it arises to some extent in all models that are not exact copies of the systems they are intended to mimic. (Such a model would, of course, be of very little use!) Model simplification involves combining multiple components of a more complex model into a smaller number of components in a simpler model, often introducing approximations in the process. This might seem to suggest that more accurate parameter values could be obtained by fitting models with fewer simplifications, but in practice this is

not often the case. Complex neuron models with detailed representations of the elaborate morphology or diverse ionic composition of real neurons are very difficult to fit to data because these complexities introduce many hills and valleys into the similarity landscape, making the best parameter values hard to find.

In this chapter, we will focus on an optimization-based approach for determining parameter values mainly because GIF models constructed in this way yield more accurate predictions of neural output. In addition, the fact that all of the parameters in the GIF model can be fitted simultaneously to small amounts of data means that the optimized values of those parameters provide a detailed window into the effective properties of *individual cells*. However, the reader should be aware that GIF model parameter estimates obtained via optimization come with important caveats: value estimates may depend on the conditions under which they were obtained (for example, the voltage range, as discussed above) and the extent to which the GIF model resembles the neuron being fitted. In the coming sections, we will show how each of these caveats can be addressed.

A.3.2 Choosing an input

Optimization-based methods for choosing GIF model parameter values involve matching the output of the GIF model to that of a real neuron for a given input. How, then, to choose the input? More importantly, why does the choice of input matter? In Section A.3.1, we saw that the parameter values found via optimization can depend on the range of voltages experienced by the neuron to which the GIF model is fitted. Since the voltage of a neuron depends on its input, the parameter values found via optimization depend on the input as well. This implies that the choice of input is important for two reasons: first, the effective parameter values found by fitting the GIF model are specific to the input used during fitting; second, since the predictions made by the GIF model depend on the values of its parameters, the neuronal outputs (i.e., voltage, spike times) predicted by the GIF model are most accurate for the input used during fitting.

Whether the GIF model is to be used to predict neuronal outputs or gain insight into the effective properties of individual cells, both goals are more easily accomplished when the input used for fitting the GIF model is chosen appropriately. Which input is most appropriate depends on the research question at hand; each researcher must ask themselves whether specific types of input (e.g., synaptic vs. artificial), frequency bands (e.g., θ oscillations), or voltage ranges (e.g., close to action potential threshold) are most relevant to their question. In practice, however, noise is often used as input because the fluctuations present in noise cover a wide range of simpler inputs (slow rise, fast rise, rise-then-fall, etc.). Ornstein-Uhlenbeck noise is a particularly popular choice because it approximates the random fluctuations produced by the synaptic bombardment neurons receive *in vivo* [203].

Carefully selecting the input used to fit the GIF model, or simply using noise, mitigates one of the important caveats associated with GIF models; namely, that parameter estimates — and, by association, model predictions — are somewhat specific to a given input. The degree of specificity can be quantified by simply comparing the parameter estimates or predictions on different subsets of data collected from a given neuron [203]. In the machine learning literature, it is common to divide a dataset of independent samples into a portion used for fitting a model and a separate portion for testing the accuracy of model predictions using new inputs, referred to as the training and test datasets, respectively. In the case of data used to fit the GIF model, completely independent samples are difficult to obtain due to the long-lasting effects of adaptation mechanisms. In practice, nearly independent samples of neuronal output are obtained by applying at least two distinct inputs to each neuron to be fitted, each separated by an equilibration period of several seconds. The data associated with some of the inputs will be used for fitting while the rest is reserved for testing, effectively dividing the experiment into separate training and test phases. Since our goal is to evaluate the accuracy of GIF model predictions on the test input, the experimenter should attempt to ensure that the training and test phases of the experiment are as similar as possible apart from the choice of input. For a detailed discussion of the best practices for evaluating the

accuracy of GIF models, see [203]. For a more general treatment of model selection and accuracy estimation, see [282, 283].

A.3.3 Optimization

Quantifying model accuracy

All optimization techniques rely on a precise definition of the similarity or dissimilarity between the actual and desired outputs of a function. In the case of the GIF model, this means we must define how we will measure the similarity between the output of the GIF model and that of a real neuron. Because the GIF model is divided into two parts, each with their own outputs and parameters, we will in fact need to define two separate measures: one for the subthreshold component of the model, and one for the stochastic spiking rule.

The subthreshold component of the GIF model given in Eq. (A.6) defines the relationship between an external input $I_{\text{ext}}(t)$ and the derivative of the membrane voltage $\widehat{\frac{dV}{dt}}$ given the current voltage of the neuron $V(t)$ and the timing of spikes $\mathcal{S}$. Notice that we have added a hat $\widehat{\ }$ to the voltage derivative to indicate that this quantity is predicted by the model rather than measured experimentally. We take the subthreshold voltage $V(t)$ and the set of spike times $\mathcal{S}$ to be given because they are easily measured experimentally, and because doing so allows us to isolate the effect of the *parameters* of the subthreshold model $\hat{g}_l$, $\widehat{E}_l$, $\widehat{C}$, $\widehat{w}_i^{(\eta)}$, and $\widehat{\tau}_i^{(\eta)}$ on its output $\widehat{\frac{dV}{dt}}$. Because the voltage derivative is a continuous signal, we can quantify the level of dissimilarity between the predicted derivative $\widehat{\frac{dV}{dt}}$ and the experimentally measured derivative $\frac{dV}{dt}$ using the sum of squared errors

$$J(\hat{g}_l, \widehat{E}_l, \widehat{C}, \widehat{w}_1^{(\eta)}, \ldots, \widehat{w}_k^{(\eta)}, \widehat{\tau}_1^{(\eta)}, \ldots, \widehat{\tau}_k^{(\eta)}; I(t), \mathcal{S}, V(t)) = \sum_{t \notin \mathcal{S}'} \left(\frac{dV}{dt} - \widehat{\frac{dV}{dt}} \right)^2, \tag{A.9}$$

where $t \notin \mathcal{S}'$ is time excluding a small window around each spike. Notice that if we choose values of the parameters $\hat{g}_l$, $\widehat{E}_l$, etc. that cause the predicted voltage derivative $\widehat{\frac{dV}{dt}}$ to be far above or below the measured derivative $\frac{dV}{dt}$, this dissimilarity function will have a large

value. On the other hand, if we choose parameter values that cause the output of the model to match the measured derivative exactly, the dissimilarity will be zero. Our objective is therefore to find the parameter values that minimize $J(\hat{g}_l, \ldots; I(t), \mathcal{S}, V(t))$, which can be expressed mathematically as

$$\underset{\hat{g}_l, \widehat{E}_l, \widehat{C}, \widehat{w}_1^{(\eta)}, \ldots, \widehat{w}_k^{(\eta)}, \widehat{\tau}_1^{(\eta)}, \ldots, \widehat{\tau}_k^{(\eta)}}{\arg\min} J\left(\hat{g}_l, \widehat{E}_l, \widehat{C}, \widehat{w}_1^{(\eta)}, \ldots, \widehat{w}_k^{(\eta)}, \widehat{\tau}_1^{(\eta)}, \ldots, \widehat{\tau}_k^{(\eta)}; I(t), \mathcal{S}, V(t)\right)$$

where $\arg\min_x f(x)$ finds the value of x that minimizes $f(x)$. This notation is common in the machine learning literature, in which $J(x)$ is referred to as an objective function or loss function.

The stochastic spiking rule given in Eq. (A.8) defines the probability of emitting a spike at a particular time t given the subthreshold voltage predicted by the model $\widehat{V}(t)$ and the timing of previous spikes $\{s \in \mathcal{S}; s < t\}$ (indirectly through Eq. (A.5)). Our goal is therefore to find values for the parameters in the spiking rule $\widehat{V}_T^*$, $\widehat{\sigma}$, $\widehat{w}_i^{(\gamma)}$, and $\widehat{\tau}_i^{(\gamma)}$ that maximize the probability that the spikes emitted by the model occur at the same time as spikes are observed experimentally

$$\underset{\widehat{V}_T^*, \widehat{\sigma}, \widehat{w}_1^{(\gamma)}, \ldots, \widehat{w}_k^{(\gamma)}, \widehat{\tau}_1^{(\gamma)}, \ldots, \widehat{\tau}_k^{(\gamma)}}{\arg\max} \Pr\left[\widehat{\mathcal{S}} = \mathcal{S} \mid \widehat{V}(t)\right]$$

where $\widehat{\mathcal{S}}$ and $\mathcal{S}$ are the model predicted and experimentally observed sets of spike times, respectively. More precisely, we would like to maximize the joint probability that the model emits a spike when a spike is observed experimentally and that the model does not emit a spike when a spike is not observed experimentally

$$\Pr\left[\widehat{\mathcal{S}} = \mathcal{S} \mid \widehat{V}(t)\right] = \Pr\left[\bigcap_t \begin{cases} t \in \widehat{\mathcal{S}} \text{ if } t \in \mathcal{S} \\ t \notin \widehat{\mathcal{S}} \text{ if } t \notin \mathcal{S} \end{cases} \middle| \widehat{V}(t) \right].$$

Unfortunately, the probability that the model emits a spike at a particular time t depends on

the timing of previous spikes $\{\hat{s} \in \widehat{\mathcal{S}}; \hat{s} < t\}$ due to spike frequency adaptation induced by the moving threshold $G(t)$ from Eq. (A.5). This means that the probabilities that the model does or does not emit a spike at any particular set of times $t_i, t_j, t_k, \ldots$ are not independent, and consequently we cannot write the previous equation as a product of probabilities. However, if we assume that all of the spikes up to a given time t were emitted at the correct times $\{\hat{s} \in \widehat{\mathcal{S}}; \hat{s} < t\} \leftarrow \{s \in \mathcal{S}; s < t\}$, then we can take advantage of the fact that the spiking probabilities under this assumption are independent to rewrite the previous equation using the product rule

$$\Pr\left[\widehat{\mathcal{S}} = \mathcal{S} \mid \widehat{V}(t)\right] = \prod_t \begin{cases} \rho\left(\frac{\widehat{V}(t)-\widehat{V}_T(t)}{\widehat{\sigma}}\right) & \text{if } t \in \mathcal{S} \\ 1 - \rho\left(\frac{\widehat{V}(t)-\widehat{V}_T(t)}{\widehat{\sigma}}\right) & \text{if } t \notin \mathcal{S} \end{cases} \tag{A.10}$$

where $\rho\left(\frac{\widehat{V}(t)-\widehat{V}_T(t)}{\widehat{\sigma}}\right)$ gives the probability of spiking at time t (see Eq. (A.8)). This gives us a concrete expression for the similarity between the output of the spiking rule of the GIF model and the experimentally observed spike times which we can maximize.

Solving for parameter values

Now that we have expressed the degree of agreement between the outputs of the GIF model and of a real neuron, we can turn to the question of how to adjust the model parameters to minimize the dissimilarity in Eq. (A.9) and maximize the similarity in Eq. (A.10). If we fix the timescales $\tau_1^{(\eta)}, \ldots, \tau_k^{(\eta)}, \tau_1^{(\gamma)}, \ldots, \tau_k^{(\gamma)}$, we can solve for the remaining parameters to minimize this restricted form of Eq. (A.9) directly, and maximize the restricted form of Eq. (A.10) by gradient ascent.

It is possible to solve for the optimal values of the unknown parameters in Eq. (A.9) because the term representing the subthreshold dynamics of the GIF model $\widehat{\frac{dV}{dt}}$ can be rewritten as a linear equation. The parameter values that minimize the sum of squared errors of any linear model can be found using ordinary least-squares regression. Simple linear models

include $\hat{y} = \hat{m}x + \hat{b}$ for a single input variable x, or

$$\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x_1 + \widehat{\beta}_2 x_2 + \ldots + \widehat{\beta}_n x_n$$

for n input variables. The $\widehat{\beta}_i$ are called regression coefficients or model parameters. In a geometric interpretation of this equation, $\widehat{\beta}_0$ is called an intercept and $\widehat{\beta}_1, \ldots, \widehat{\beta}_n$ are called slopes. It might not be immediately obvious that the GIF model is such a linear model, but if the subthreshold dynamics from Eq. (A.6) are expanded and rearranged, we obtain the following

$$\widehat{\frac{dV}{dt}} = \frac{\widehat{g}_l \widehat{E}_l}{\widehat{C}} + \frac{-\widehat{g}_l}{\widehat{C}} V(t) + \frac{1}{\widehat{C}} I_{\text{ext}}(t) + \frac{-\widehat{w}_1^{(\eta)}}{\widehat{C}} \sum_{\{s \in \mathcal{S}; s<t\}} e^{\frac{s-t}{\tau_1^{(\eta)}}} + \cdots + \frac{-\widehat{w}_k^{(\eta)}}{\widehat{C}} \sum_{\{s \in \mathcal{S}; s<t\}} e^{\frac{s-t}{\tau_k^{(\eta)}}}$$

which can be rewritten as

$$\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 V(t) + \widehat{\beta}_2 I_{\text{ext}}(t) + \widehat{\beta}_3 \sum_{\{s \in \mathcal{S}; s<t\}} e^{\frac{s-t}{\tau_1^{(\eta)}}} + \cdots + \widehat{\beta}_{k+2} \sum_{\{s \in \mathcal{S}; s<t\}} e^{\frac{s-t}{\tau_k^{(\eta)}}}, \tag{A.11}$$

where $\widehat{C} = 1/\widehat{\beta}_2$, $\widehat{g}_l = -\widehat{\beta}_1/\widehat{\beta}_2$, and so on. This is a linear model where the output $\hat{y} = \widehat{\frac{dV}{dt}}$ is the voltage derivative; the inputs x_i are the injected current $I_{\text{ext}}(t)$, the subthreshold voltage $V(t)$, and the exponential basis functions of the adaptation current $\eta(t-s)$ from Eq. (A.3) summated over past spikes $\{s \in \mathcal{S}; s < t\}$ (obtained by decomposing $H(t)$ from Eq. (A.4)); and the regression coefficients $\widehat{\beta}_i$ are the unknown parameters. This linear form of the subthreshold dynamics shows why it is necessary to fix the values of $\tau_i^{(\eta)}$: if these values were not fixed, the dynamics could not be written in terms of known input variables multiplied by unknown regression coefficients, and it would not be possible to use ordinary least squares regression to estimate all of the unknown parameters. Importantly, this form also shows that more components can be added to the subthreshold part of the GIF model as long as they can be written in terms of a known variable scaled by an unknown amount (this will be discussed in detail in Section A.3.4).

Unfortunately, there are no similar techniques to find the values of the threshold parameters that maximize Eq. (A.10) directly. Instead, we must begin with a set of initial guesses for the values of these parameters and incrementally improve them using gradient ascent. In practice, the spike probability function used in the GIF model, $\rho(\frac{\widehat{V}(t)-\widehat{V}_T(t)}{\widehat{\sigma}})$ (see Eq. (A.8)), guarantees that gradient ascent will eventually lead us to the best possible values for these parameters [136, 284]. Similarly to the subthreshold optimization process discussed above, this guarantee is subject to certain constraints, and the simplest way to satisfy these constraints is to require that the term inside the spike probability function, $\frac{\widehat{V}(t)-\widehat{V}_T(t)}{\widehat{\sigma}}$ in this case, can be written as a linear function of its parameters. By expanding $\widehat{V}_T(t)$ and rearranging, we can rewrite the term inside the spike probability function as

$$\begin{aligned}\frac{\widehat{V}(t)-\widehat{V}_T(t)}{\widehat{\sigma}} &= \frac{-\widehat{V}_T^*}{\widehat{\sigma}} + \frac{1}{\widehat{\sigma}}\widehat{V}(t) + \frac{-\widehat{w}_1^{(\gamma)}}{\widehat{\sigma}} \sum_{\{s\in\mathcal{S};s<t\}} e^{\frac{s-t}{\tau_1^{(\gamma)}}} + \cdots + \frac{-\widehat{w}_k^{(\gamma)}}{\widehat{\sigma}} \sum_{\{s\in\mathcal{S};s<t\}} e^{\frac{s-t}{\tau_k^{(\gamma)}}} \\ &= \widehat{\beta}_0 + \widehat{\beta}_1\widehat{V}(t) + \widehat{\beta}_2 \sum_{\{s\in\mathcal{S};s<t\}} e^{\frac{s-t}{\tau_1^{(\gamma)}}} + \cdots + \widehat{\beta}_{k+2} \sum_{\{s\in\mathcal{S};s<t\}} e^{\frac{s-t}{\tau_k^{(\gamma)}}},\end{aligned}$$

where $\widehat{V}(t)$ is the voltage predicted by the subthreshold component of the model. This expression for $\frac{\widehat{V}(t)-\widehat{V}_T(t)}{\widehat{\sigma}}$ is of course linear with respect to the unknown parameters. Just as with the subthreshold component of the model, this linear form illustrates that additional components can be added to the spiking rule of the GIF model as long as they can be written as a known variable scaled by an unknown amount.

A.3.4 Extending the subthreshold model

The GIF model, like all other neuron models, is subject to the No Free Lunch theorem: no single model is best for all cases [285]. The simplifications used to construct the GIF model — for example, that the subthreshold dynamics of neurons are not voltage-dependent — might present problems for particular cell types or research questions. In this section, we will use the serotonin neurons of the dorsal raphe nucleus as a case study to illustrate how the GIF

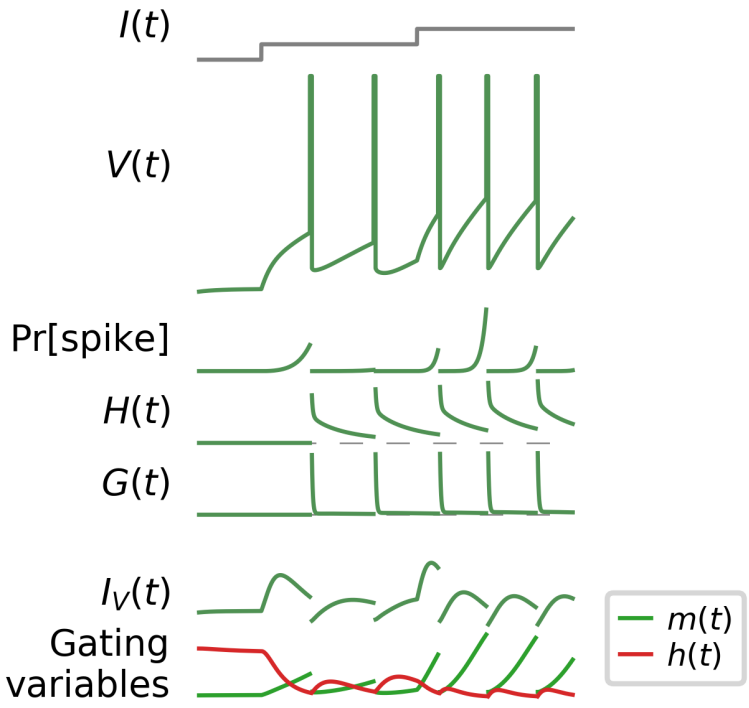


Figure A.3: A GIF model augmented with a voltage-dependent current.

model can be extended to address limitations of the subthreshold model.

The subthreshold electrical properties of serotonin neurons are characterized by an unusually large membrane resistance (equivalent to a very small leak conductance) and a potent voltage-dependent ionic current [158]. These characteristics violate one of the core assumptions of the GIF modelling framework, namely, that the subthreshold electrical properties of neurons are dominated by a voltage-independent leak conductance. Fortunately, it is possible to augment the subthreshold dynamics of the GIF model with a voltage-dependent component to account for the specific characteristics of serotonin neurons. To see how, we will start by adding a voltage-dependent current $I_V(t)$ to the subthreshold dynamics of the GIF model defined in Eq. (A.6)

$$\frac{dV}{dt} = \frac{1}{C}\left(-g_l(V(t) - E_l) + I_V(t) - H(t) + I_{\text{ext}}(t)\right). \tag{A.12}$$

We can model $I_V(t)$ following the usual Hodgkin-Huxley approach

$$I_V(t) = g_V m(t)h(t)(V(t) - E_V) \tag{A.13}$$

where g_V is the maximum conductance; $m(t)$ and $h(t)$ are activation and inactivation gating functions, respectively; and E_V is the reversal potential of the current. The details of the gating functions and reversal potential are not important; suffice it to say that these can usually be determined experimentally[3] or controlled. With this in mind, we can group together the known terms from Eq. (A.13) into a new variable $a(t) = m(t)h(t)(V(t) - E_V)$. Substituting this back into Eq. (A.13), we obtain a definition of the voltage-dependent current in terms of a *known variable* scaled by an *unknown parameter*

$$I_V(t) = g_V a(t). \tag{A.14}$$

Recall from Section A.3.3 that the unknown parameters in the subthreshold component of the GIF model can be found easily as long as the subthreshold dynamics can be written in a linear form; in other words, as a sum of known variables scaled by unknown parameters. Substituting Eq. (A.14) back into Eq. (A.12) and expanding and rearranging the terms, we obtain the following linear form in analogy with Eq. (A.11)

$$\widehat{\frac{dV}{dt}} = \frac{\widehat{g}_l \widehat{E}_l}{\widehat{C}} + \frac{-\widehat{g}_l}{\widehat{C}} V(t) + \frac{\widehat{g}_V}{\widehat{C}} a(t) + \frac{1}{\widehat{C}} I_{\text{ext}}(t) + \frac{-\widehat{w}_1^{(\eta)}}{\widehat{C}} \sum_{\{s \in \mathcal{S}; s<t\}} e^{\frac{s-t}{\tau_1^{(\eta)}}} + \cdots + \frac{-\widehat{w}_k^{(\eta)}}{\widehat{C}} \sum_{\{s \in \mathcal{S}; s<t\}} e^{\frac{s-t}{\tau_k^{(\eta)}}}.$$

Because the subthreshold dynamics can still be written in a linear form, the regression approach to estimating the unknown model parameters presented in Section A.3.3 can still be used, ultimately yielding an augmented GIF model with a voltage-dependent ionic current

[3]The gating functions in Hodgkin-Huxley current models are usually expressed in terms of an equilibrium gating function, which is a sigmoidal function of voltage, and one or more gating time constants, which may themselves depend on voltage. Readers with a background in whole-cell electrophysiology will likely already be familiar with techniques for measuring these quantities. For a comprehensive treatment, see [286].

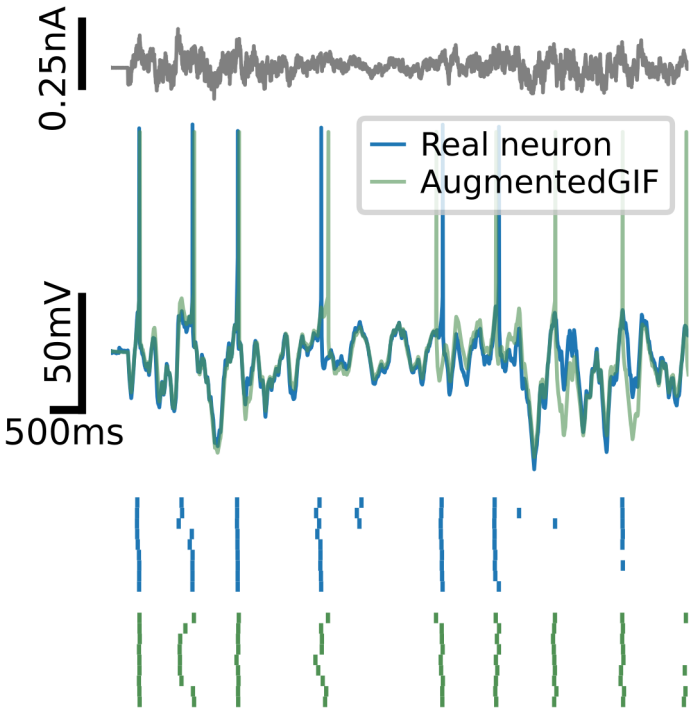


Figure A.4: A GIF model augmented with a voltage-dependent current fitted to a serotonin neuron. The noisy input stimulus at the top was not used during fitting. Notice that the augmented GIF model (teal) accurately predicts the subthreshold voltage (middle) and spike times (raster, bottom) of the real neuron (blue). The spike rasters at bottom are for several repetitions of the test stimulus shown at top. These illustrate that the augmented GIF model captures some of the natural stochasticity of spiking observed in serotonin neurons.

(see Fig. A.3)[4].

Augmenting the GIF model with additional components such as ionic currents can bring the assumptions of the neuron model into closer agreement with the known features of particular neurons. This improves the interpretability of the model by reducing the extent to which multiple electrical features are mixed into a single model term. Evidence of this mixing can be seen in the estimated values of $\widehat{g}_l$ in serotonin neurons using the GIF model and its augmented counterpart presented above. Even though the leak term used to estimate $\widehat{g}_l$ is the same in both models, the values of $\widehat{g}_l$ in the augmented model are closer to the corresponding true values g_l because the effects of the voltage-dependent current are mixed

[4]For an example of a similar approach used to estimate the parameters of ionic currents in a more detailed model, see [138].

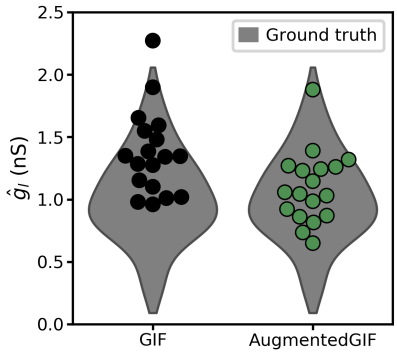


Figure A.5: Leak conductance estimated by GIF models fitted to serotonin neurons (circles). Ground-truth estimates for this parameter based on experimental measurements from a large number of cells are shown in gray; wider areas indicate a relatively greater proportion of cells. Notice that the estimates produced by the augmented GIF model are closer to the ground truth.

into it to a lesser extent (see Fig. A.5). Of course, unless the Hodgkin-Huxley model of the current used in the augmented GIF model is exactly correct, a certain amount of mixing will always occur, which explains why the distribution of $\widehat{g}_l$ in Fig. A.5 does not exactly agree with the ground truth. This reduced mixing can also increase the accuracy of model predictions by decreasing the dependence of model parameter estimates on the input used to fit the model (as discussed in Section A.3.1). In the case of serotonin neurons, the augmented GIF predicts the timing of spikes significantly more accurately than the base GIF model [158], even though the differences between the two models are limited to the subthreshold dynamics. (See Fig. A.4 for an example of an augmented GIF model fitted to a serotonin neuron.) These results illustrate how adjustments to the GIF model can improve accuracy and interpretability.

A.4 Summary

In this chapter, we have seen that generalized integrate-and-fire (GIF) models build on leaky integrate-and-fire models to capture three of the most fundamental features of neurons: leaky subthreshold integration, stochastic spiking, and spike-frequency adaptation. The simplified mathematical structure of the GIF model provides a one-to-one correspondence between

model components and electrophysiological features, making it intuitive to understand. We also saw that the subthreshold and spiking components of the GIF model are rooted in linear models that are easily fitted to data. This allows the GIF model to be constrained to mimic the behaviour of individual neurons based on very little data using a two-step optimization procedure, often with better results than more labour-intensive experimental approaches that require measuring the values of model parameters one at a time. Finally, we showed how the linear components of the GIF model can be extended to account for non-linear ionic currents that are not present in the GIF model as it was initially defined.

A.5 Further reading

The leaky integrate-and-fire model as it is used today was first introduced by [147], and the generalized integrate-and-fire model as it is presented here was introduced by [151]. See [287] for an earlier model similar to the GIF model which also captures a wide range of neural behaviours. A detailed derivation of the GIF model and its two-step fitting procedure can be found in [281]. For a very practical description of how to carry out experiments that can be used to constrain the GIF model, and an overview of how GIF models can be fitted using publicly-available software, see [203]. For examples of how different variations of the GIF framework can be used to capture the behaviours of various types of cortical neurons, see [152, 204]. The definitive textbook on the physiology of ion channels and neuronal membranes is [286].

The GIF modelling framework is conceptually related to other classes of neuron models that are based on linearity assumptions. Like the GIF framework, generalized linear models (GLMs) [141] and linear-nonlinear Poisson models (LNPs) [139] are based on a linear function of a set of inputs which is then passed through a non-linear link function to produce a firing rate (GLMs) or spike probability (LNPs). However, unlike GIF models, these models do not attempt to predict the subthreshold voltage as an intermediate step. This means that GLMs

and LNPs are not well-suited to cases where the subthreshold voltage is of primary interest, or when spiking data is very sparse. On the other hand, these models can be fitted to spiking data even when the subthreshold voltage is not known, as is the case during extracellular recording.

In this chapter, we have introduced optimization methods that rely on either exact knowledge of the relationship between model parameter values and predictive accuracy to solve for the best possible parameter values directly, or that use only local knowledge of this relationship to find the best values within a neighbourhood (gradient-based methods introduced in Section A.3.3). [135] recently introduced an intermediate approach which uses an artificial neural network to approximate the global relationship between parameter values and model outputs, allowing approximately optimal parameter values to be found even for complex models.

Appendix B

Parallel and recurrent cascade models as a unifying force for understanding subcellular computation

This chapter is reprinted from a previously-published work.

Citation

E.F. Harkin*, P.R. Shen*, A. Goel, B.A. Richards†, and R. Naud†. Parallel and recurrent cascade models as a unifying force for understanding subcellular computation. *Neuroscience*, **489**, 200–215. (2022)

*,†: Equal contribution.

Statement of contributions

Richard Naud conceptualized the model, which was finalized in collaboration with Peter R. Shen and I. I implemented the surrogate gradient-based optimization approach used to train spiking network models to perform classification tasks. I perfomed simulations and data analysis for Figs. 6 to 8, with contributions from Anish Goel to Fig. 8. Blake A. Richards and Richard Naud wrote the first draft of the paper, which was finalized with Peter R. Shen, Anish Goel, and I. Blake A. Richards and Richard Naud conceptualized the project.

Abstract

Neurons are very complicated computational devices, incorporating numerous non-linear processes, particularly in their dendrites. Biophysical models capture these processes directly by explicitly modelling physiological variables, such as ion channels, current flow, membrane capacitance, etc. However, another option for capturing the complexities of real neural computation is to use cascade models, which treat individual neurons as a cascade of linear and non-linear operations, akin to a multi-layer artificial neural network. Recent research has shown that cascade models can capture single-cell computation well, but there are still a number of sub-cellular, regenerative dendritic phenomena that they cannot capture, such as the interaction between sodium, calcium, and NMDA spikes in different compartments. Here, we propose that it is possible to capture these additional phenomena using *parallel*, *recurrent* cascade models, wherein an individual neuron is modelled as a cascade of parallel linear and non-linear operations that can be connected recurrently, akin to a multi-layer, recurrent, artificial neural network. Given their tractable mathematical structure, we show that neuron models expressed in terms of parallel recurrent cascades can themselves be integrated into multi-layered artificial neural networks and trained to perform complex tasks. We go on to discuss potential implications and uses of these models for artificial intelligence. Overall, we argue that parallel, recurrent cascade models provide an important, unifying tool for capturing single-cell computation and exploring the algorithmic implications of physiological phenomena.

B.1 Introduction

One of the greatest success stories in modern neuroscience is the development of biophysical models of single-cell computation. Though there are still many mysteries to be explored, and new discoveries are still being made, our core understanding of the biophysics of excitable membranes as described by circuit equivalence models, cable theory, and Hodgkin & Huxley-style models has stood the test of time and can reasonably be considered as an accepted theory in neuroscience [127, 288–292]. This has provided the foundation for countless computational studies on single-cell computation using detailed models that explicitly incorporate physiological variables including membrane capacitance, ion channels, reversal potentials, etc. Such models have proven very useful for understanding a variety of phenomena, particularly dendritic computation [129, 293–305]. Without these models we would understand much less than we do about how dendrites contribute to computation in neural circuits.

However, due to their complexity, biophysical models are very difficult to link to algorithmic models of neural computation. To some extent, this is part of the dilemma we always face in science, i.e., "How detailed should our models be in order to achieve our scientific goals?" [293, 306]. But, one thing that we can say is that it would be beneficial for understanding the functional importance of dendritic computation if we possessed models of intermediate complexity that could capture many of the phenomena of single-neuron computation while still being sufficiently mathematically tractable to use for explaining complex animal behaviour. Moreover, if we could develop such intermediate models we would be better placed to use insights on dendritic computation to inform the development of novel machine learning approaches [306, 307].

To this end, previous work has explored the use of "cascade models" to capture dendritic computation [131, 214, 308, 309] (but see [310]). Cascade models use a cascade of linear and non-linear operations, which are effectively tree-structured, feedforward, multi-layer artificial neural networks (ANNs) [293]. Research has shown that these models can capture

more variance in single-cell activity than standard point neuron models (which consist of a single linear operation and non-linear activation function) [131, 214, 308]. Thus, cascade models treat *individual* neurons as deep feedforward ANNs in order to capture single-cell computation with a mathematically tractable model (Figure B.1A). However, such models still have not been compared to many facets of dendritic computation, including calcium spikes and N-methyl-D-aspartate (NMDA) receptor mediated plateaus.

Here, we show that it is possible to capture these phenomena using *parallel and recurrent cascade models* (PRC models), i.e. models that use a set of parallel cascades of linear and non-linear operations that are recurrently connected to one another. This is equivalent to treating individual neurons as multi-layer, recurrent ANNs (Figure B.1B). We show that these PRC models can successfully reproduce a number of experimentally observed regenerative phenomena in dendrites, all while being mathematically tractable. Moreover, because these models are fully differentiable, we show that they can easily be incorporated into machine learning approaches that utilize gradient descent for model optimization [306]. This opens the door to exploring the possibility that dendrites and dendritic computation may provide important inductive biases for brains that could be mimicked by artificial intelligence (AI) [306, 307]. Thus, we conclude this paper by providing some speculation as to the possible advantages for AI of dendritic PRC models. Altogether, this work helps to lay the ground for better integration between our well-developed understanding of the biophysics of neural computation and our ever increasing understanding of algorithms for complex behaviour.

B.2 The use of recurrent cascade models to capture single-cell computation

Models made of a cascade of linear-nonlinear (LNL) operations have had a long history in systems neuroscience, where such models were conceptually implied by early work on retinal ganglion cells [311] and cortical cells [312]. These models consist of a linear filter

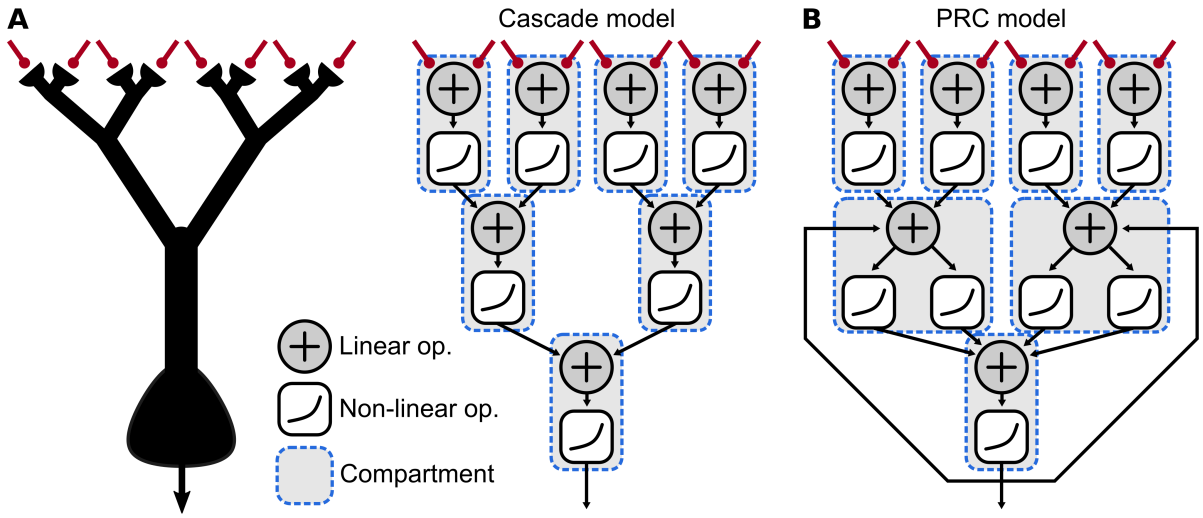


Figure B.1: Illustration of cascade models and PRC models: **A** Dendritic computation can be modelled using a cascade of linear and non-linear operations [131, 214, 308]. **B** Cascade models can be extended with the use of parallel pathways and recurrence in the operations, which can enable the modeling of more complicated regenerative phenomena.

of the stimulus followed by a nonlinear readout to generate predictions of a non-negative observable. Important refinements to improve the accuracy of these models were the inclusion of spike-triggered adaptation [139, 140], the composition of linear-nonlinear (LNL) operations in a cascade [313, 314], and the addition of recurrent interactions between linear-nonlinear subunits [141, 315]. Together, these various extensions are described by an LNL subunit with multiple possible types of interconnection motifs. The LNL subunit is composed of a linear-nonlinear operation with feedback from the output of the nonlinearity to capture spike-triggered adaptation (Fig. B.2A; Methods B.5.1). Wiring between subunits can create multiple types of motifs including a strictly feedforward cascade, a common input to two units having possibly different filters or nonlinearities (parallel feature processing, Fig. B.2B), or recurrent interconnections (Fig. B.2C). In most systems neuroscience applications, the output of the nonlinearity is the Poisson intensity, used to capture the stochastic spike-timing responses of real neurons. Our approach with PRC models here is slightly different, as we will consider the output of the nonlinearity to represent deterministic voltage excursions. Also, in most systems neuroscience applications, the filters and nonlinearities may arise from a

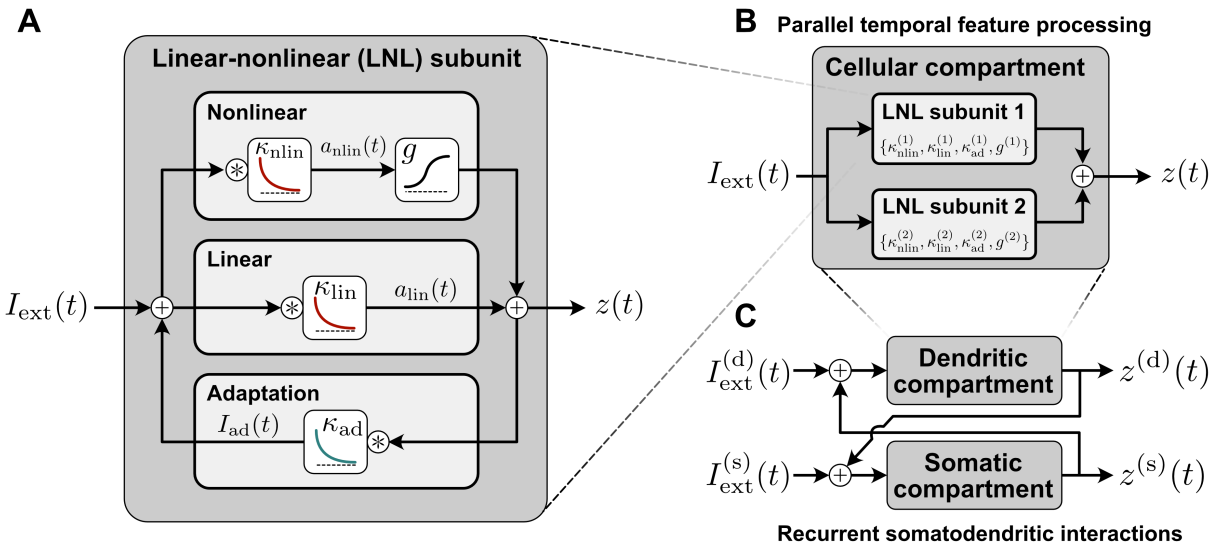


Figure B.2: Linear-nonlinear operations and interconnection motifs. **A** The linear-nonlinear (LNL) subunit. This architecture combines linear filtering of an input $I_{\mathrm{ext}}(t)$ (middle) with a nonlinear readout (top, g) and feedback (bottom, κ_{ad}), which together flexibly capture the passive filtering effects of neuronal membranes, the nonlinear effects of voltage-dependent ionic conductances, and adaptation. The contributions of each of these effects to the output $z(t)$, which may loosely represent neuronal voltage or a spiketrain, can be tuned via the parameters of the filters κ_{nlin}, κ_{lin}, and κ_{ad}, and the choice of nonlinear function $g(\cdot)$. **B** Multiple LNL subunits arranged in parallel can model the effects of multiple ionic conductances in a single cellular compartment. **C** Recurrent connections between model compartments, each composed of one or more LNL subunits, capture interactions between cellular compartments.

large number of possible — yet undefined — mechanisms. These are typically thought of as interactions within and between cells, but may also include dendritic computations [316]. In order to better understand the cellular mechanisms underlying such information processing, we focus on the application of a deterministic LNL framework within a single cell.

B.2.1 Somatic spikes

Since the pioneering work of Richard Stein [147], the leaky integrate-and-fire (LIF) model has become the workhorse of interrogations of information processing of large numbers of interconnected neurons. In itself, the LIF can be seen as a special case of the LNL subunit. When a depolarizing current is injected into an LIF model, it first passes through the membrane filter and produces a voltage (leaky integration; a linear operation) which is

eventually reset to a lower value if it reaches a threshold (firing; a nonlinear operation). This leaky integrate-and-fire behaviour can be captured by an LNL subunit with a filter that corresponds to the membrane filter of an LIF model, together with a Heaviside nonlinearity that is triggered exactly at spike threshold and a Dirac delta-shaped adaptation filter which resets the voltage to a lower value. Such deterministic LNL models can capture both the time-dependent membrane voltage response and the spike timing responses to complex inputs [317]. Adding multiple time-scales to the kernel for spike-triggered adaptation makes these models highly predictive of the responses of a variety of cell types [151, 152, 192, 232, 287]. Furthermore, considering smooth nonlinearities and surrogate gradients can ensure that the LNL unit remains differentiable [318].

In addition it is worth noting that since LNL subunits are a special case of PRC models, and LIF models are a special case of LNL models, LIF models are, in fact, a restricted class of PRC models. As such, they can capture some of the phenomena that the broader class of PRC models are capable of capturing [281, 319]. Part of our contribution here is to illustrate some of the more complicated sub-cellular phenomena that cannot be recapitulated with a pure LIF model, such as dendritic non-linearities and coincidence detection (see Sections B.2.2 to B.2.5, below).

B.2.2 LNL models for dendritic spikes

To circumscribe a systems-level function for dendritic computation, many studies have focused on the role of intrinsic nonlinear dendritic operations—first using models of dendritic trees in a stationary state [131, 320, 321], then using models capturing the dynamics of dendrites and their intrinsic nonlinear properties [214, 309, 322]. These contributions are examples of what we call PRC models because they involve a cascade of nonlinearities, but they leave out the recurrent motifs. Also, most have not considered the parallel processing introduced by [214]. Recurrence was, however, part of other efforts focusing on simplified models of the interaction between somata and dendrites [323–326]. Thus our goal in this section is

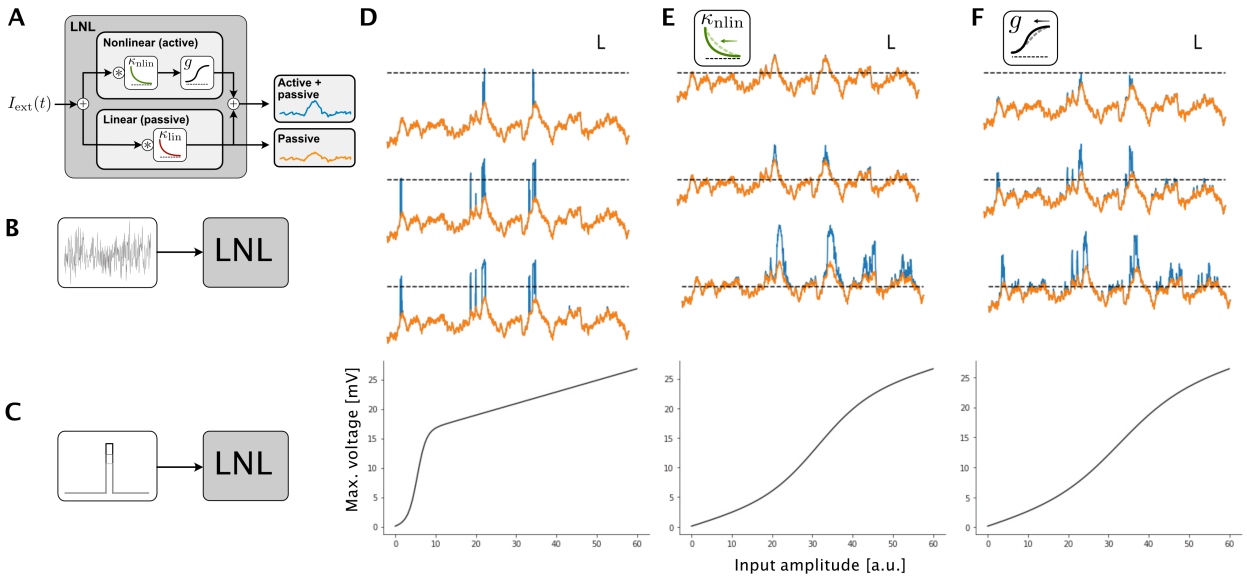


Figure B.3: Effects of changing parameters in the linear-nonlinear model. **A** Schema of the LNL model. An external input undergoes two parallel processing steps: a linear (passive) filter and a linear filter (possibly distinct from passive) followed by a nonlinear readout. The nonlinear processing step (top) produces a nonlinear component (blue) which is added to a passive component (orange). As input to the model, we consider both **B** a noisy time-dependent signal representing a bombardment of asynchronous synaptic currents, and **C** a short pulse. **D, top 3 traces** Response to noisy inputs having three different baselines (lower baseline is topmost trace,dashed line represents the activation threshold of the nonlinearity). The nonlinear (blue) is added to the passive component (orange). **D, bottom** Maximum amplitude of response as a function of pulse amplitude (bottom). **E** Same as D but using a model with an increased timescale of the linear filter. **F** Same as D but with a model having a decreased sensitivity (steepness) of the nonlinearity. Scale bars correspond to 10 ms and 5 mV.

to unify these complementary perspectives by presenting data showing that PRC models can reproduce *qualitative* features of sub-cellular physiological phenomena, such as dendritic spikes, coincidence detection, etc.

In comparison to the action potentials generated in the proximity of the cell body, the spikes observed in dendrites display less stereotypical amplitudes and durations [327–330]. These features pose a problem for the LIF framework, but they are captured naturally by the LNL framework. In Figure B.3, we illustrated the response of a LNL subunit to noisy inputs and to short pulses. For a fast filter preceding a sharp nonlinearity, the model produced short spikes on top of a noisy voltage trace. The short spikes arise when the low-pass filtered input reaches the activation threshold of the nonlinear readout. Sometimes, the fluctuation is only able to activate a fraction of the nonlinearity, which produces spikes of variable amplitudes. Increasing the mean of the input makes those spikes more frequent as there are more chances that the random fluctuations reach the activation threshold of the nonlinearity. Thus in the configuration where a sharp sigmoidal nonlinearity is fed by a linear filter that is considerably faster than the membrane filter, we observe variable amplitude spikes akin to dendritic sodium spikes. We note that a very similar model architecture was able to be reproduce with great precision the response to noisy dendritic inputs in the presence of dendritic sodium spikes [309].

Next, we examined the effects of changing the parameters in the LNL subunit (Fig. B.3E). We began by increasing the timescale of the filter preceding the nonlinearity. This reduced the number of suprathreshold fluctuations, and when a fluctuation in the low-pass filtered input did cross the activation threshold it tended to stay for a longer time period. This produced less frequent but longer spikes, akin to calcium spikes [328, 331–333] or, with an even longer timescale, NMDA spikes [330]. As a consequence of changing the filter, the aspect of the nonlinearity that can be observed when presenting a pulse input is altered, and appears more graded. When, instead of changing the timescale of the linear filter, we only changed the gain of the nonlinearity (Fig. B.3F), then the spikes had a more variable amplitude and duration.

Changing the timescale of the filter and the steepness of the nonlinearity had a similar effect on the maximum voltage observed in response to a pulse input. This can be explained by the fact that in response to a short pulse, the voltage increases linearly with a slope given by the timescale of the filter [281] and thus acts as a multiplicative factor identical to the steepness of the nonlinearity. In these simulations, we have not included an adaptation-like recurrence, although the formalism can include this mechanism. Thus, altogether, by changing the parameters of the LNL model, we can simulate some basic electrophysiological features of various types of dendritic spikes.

B.2.3 Dendritic sodium spikes

To test whether a PRC model can capture other features observed in electrophysiological recordings, we focused on experimental findings reported by [327] pertaining to dendritic sodium spikes. In one of the experiments reported (Fig. B.4), recordings were made simultaneously in a proximal dendrite and in the soma. A variable-amplitude synaptic-like stimulus was injected in the dendrite. The recordings showed that in one dendrite, low input amplitudes initiated a spike in the soma which produced a back-propagating action potential in the dendrite and at high amplitudes initiated a spike in the dendrite *before* triggering an action potential in the soma. In another dendritic recording, a low amplitude stimulus initiated a dendritic spike unaccompanied by an action potential at the soma, and only a large input produced an action potential at the soma. We found that we can reproduce these phenomena by changing the parameters of two LNL subunits wired in a recurrent fashion (Fig. B.4A and C). To capture how different recordings initiated spikes preferentially in the dendrite or the soma, we varied the relative threshold of activation of the dendritic and somatic nonlinearities.

B.2.4 Dendritic NMDA spikes

Next we considered the electrophysiological recordings of NMDA spikes reported in Schiller et al. (2000) [330] (Fig. B.5). We focused on the threshold input-pulse amplitude to trigger

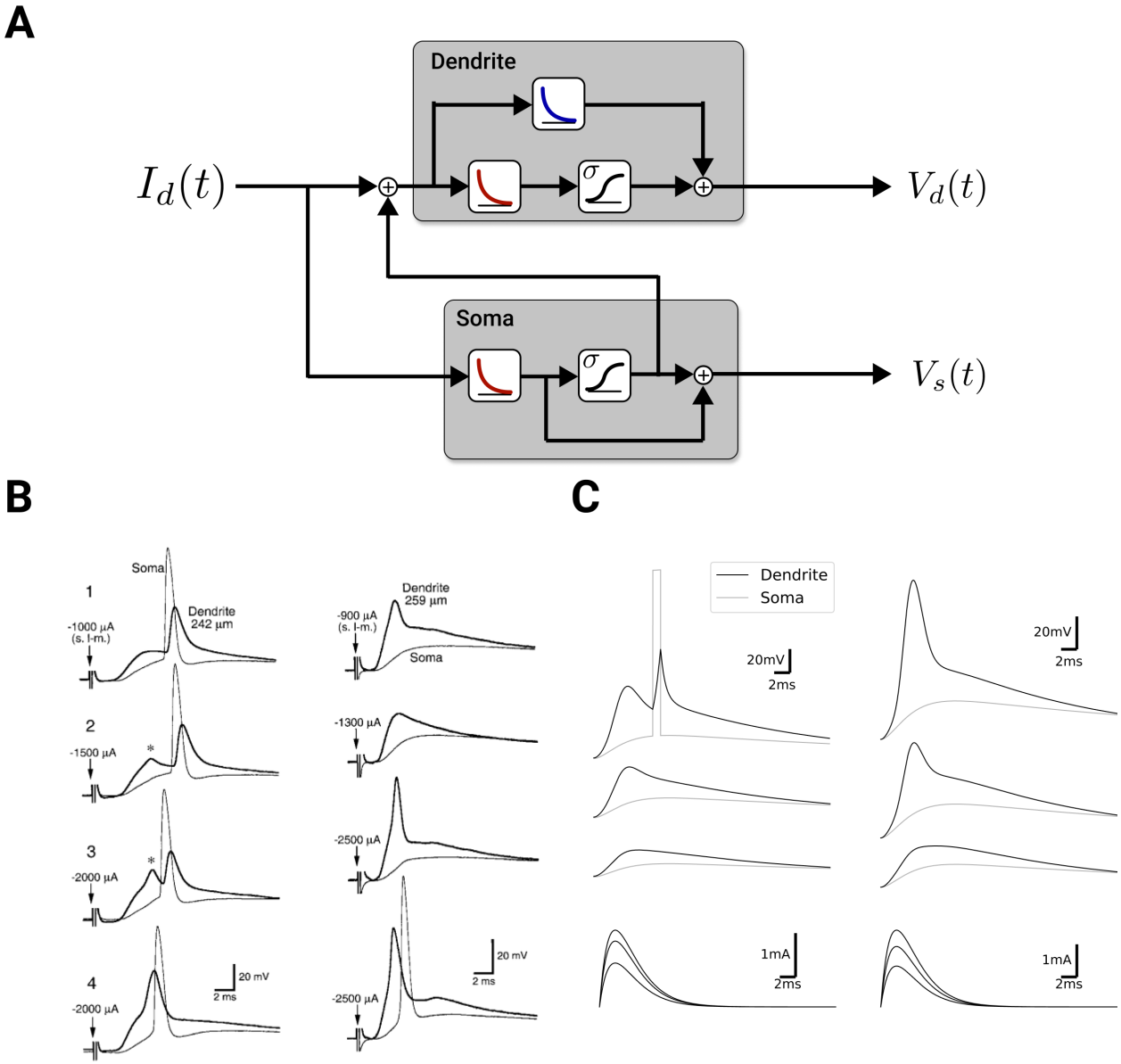


Figure B.4: A recurrent motif of linear-nonlinear models for the dendritic sodium spikes. **A** Schematic of the model: A dendritic current (I_d) impinges on two LNL subunits with recurrent interactions, one corresponding to a dendrite, another corresponding to the soma. When the somatic compartment reaches the threshold of its nonlinearity, a spike in the form of a square pulse is added to the dendritic current. **B** Experimental data showing injection of synaptic-like pulses of varying amplitudes in the dendrite (topmost traces have lowest input), where the voltage in the dendrite (thick trace) and soma (thin trace) are shown. Two exemplars are shown in different columns. Reproduced from [327] Fig. 1. **C** Model responses showing three amplitude levels of a synaptic-like input in the model shown in A. To reproduce the exemplars in B the model in the right column has a lower threshold for the dendritic nonlinearity and a higher threshold for the somatic nonlinearity.

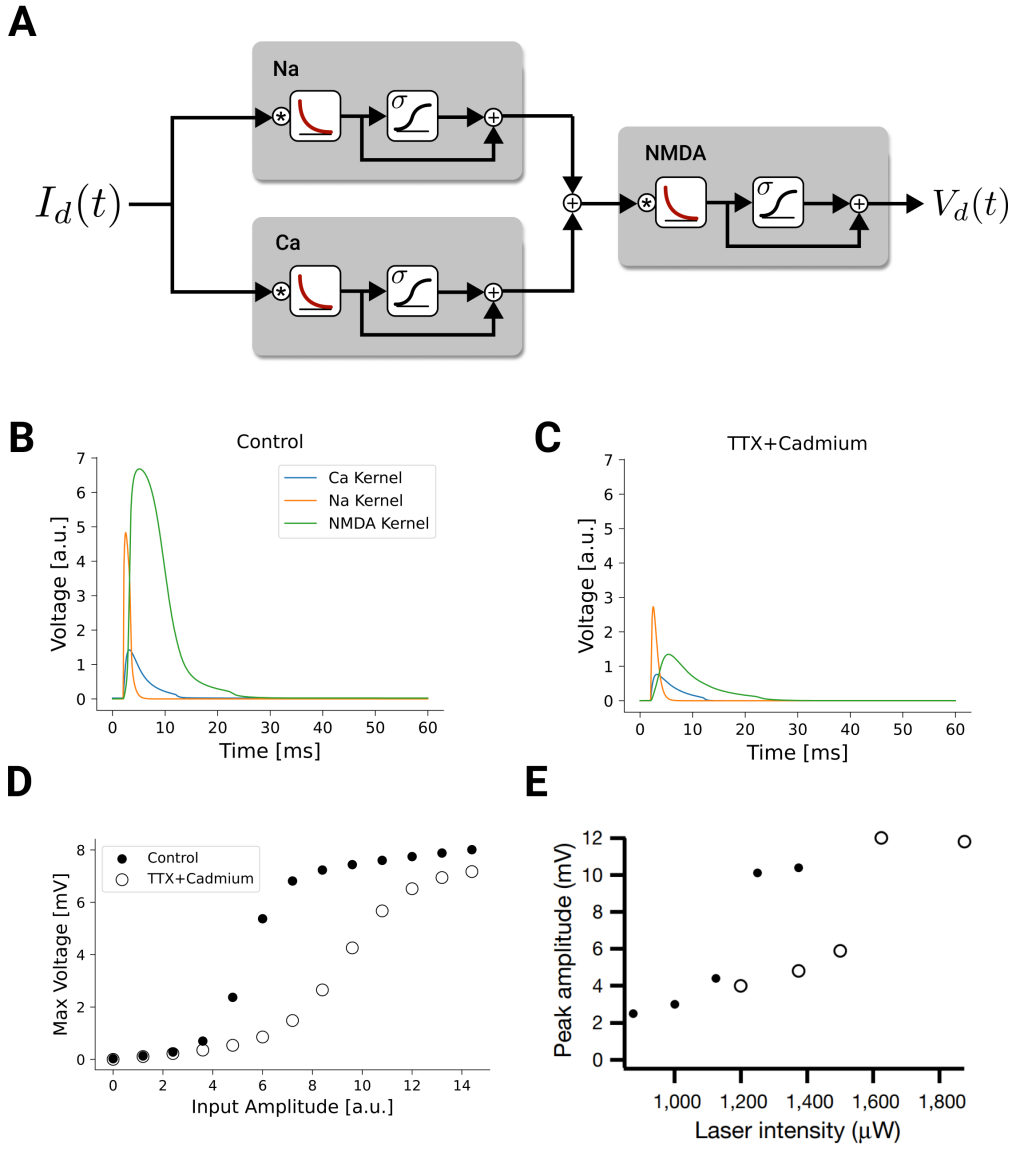


Figure B.5: Linear nonlinear model of an NMDA spike as a combination of cascade and parallel processing. **A** Schematic of the model: current impinging on the dendrite is passed through two LNL operations in parallel before feeding into another LNL operation to produce an NMDA spike. **B** Response of the model to a supra-threshold pulse input showing the calcium (blue) sodium (orange) and NMDA (green). **C** The response to the model with the same pulse input as in B but with the nonlinear component of sodium and calcium set to zero, simulating the application of TTX and Cadmium. **D** Maximum voltage as a function of the amplitude of the input pulse for the model in B (full circles) and the model in C (open circles). **E** Experimental recordings of peak membrane potential as a function of stimulation power in control (full circles) and the presence of calcium and sodium ion channel blockers (TTX and cadmium, open circles). Figure reproduced from [330] Fig. 3C.

an NMDA spike which was lowered by the addition of blockers of sodium and calcium ion channels (TTX and cadmium). This observation suggested that the nonlinear effects of sodium and calcium ion channels participated in the initiation of the NMDA spikes. Since calcium and sodium ion channels are characterized by distinct timescales, we considered a parallel connectivity motif shown in Fig. B.5A. We chose the filter of the sodium LNL to be fast (5 ms), the filter of the calcium filter to be slower (40 ms), and the filter of the NMDA LNL to be even slower (80 ms). Simulating the response of this model to pulse inputs produced a long depolarization that was clearly initiated by contributions from sodium and calcium (Fig. B.5B). To simulate the blockade of these mechanisms by TTX and cadmium, we reduced the amplitudes of their corresponding nonlinearities to zero, which prevented the occurrence of spikes for a range of input amplitudes (Fig. B.5C). The effect on the initiation threshold of removing the nonlinearity in the PRC model mimicked that of pharmacological manpulations (Fig. B.5D-E).

B.2.5 Dendritic calcium spikes

We then considered how bidirectional interactions between somatic spiking and calcium spikes can be captured by a PRC model. The tuft dendrites of cortical pyramidal cells express a high density of voltage-gated calcium channels which produce dendritic plateau potentials when sufficiently strong inputs are injected into the soma and tuft dendrites simultaneously [331, 332]. These dendritic plateau potentials mediate burst firing at the soma, producing coincidence detection and modulating the gain of somatic responses to peri-somatic input. To capture these effects in our PRC framework, we created a model with two recurrently-connected compartments, corresponding to the soma and apical tuft dendrites (Fig. B.6A). Appropriately tuned filtering and nonlinear operations in each compartment (see Methods B.5.4) caused the somatic compartment to emit single spikes when input was delivered to the soma alone and intermittent bursts when input was delivered to both compartments simultaneously (Fig. B.6B). Inputs delivered to either compartment alone evoked small

responses in the dendritic compartment, while simultaneous inputs to both compartments evoked burst-associated plateau potentials in the simulated dendrite. In cortical pyramidal neurons, dendritic inputs produce somatic bursts most potently when they are delivered just before or at the same time as somatic input, creating an asymmetric coincidence-detection effect [331]. Injecting a synaptic-like pulse into the dendritic compartment of our model evoked a dendritic plateau potential and somatic burst only when it preceded or arrived at the same time as a somatic input pulse (Fig. B.6C), recapitulating the coincidence-detection effect observed in pyramidal neurons [331]. Dendritic input to our model also modulates somatic gain by increasing the number of spikes evoked by a given input (Fig. B.6D), consistent with an effect of dendritic input on gain observed experimentally [334]. Together, these simulations add to previous efforts [325] in showing that the PRC framework is able to capture features of dendritic excitability and somatodendritic interactions.

B.3 Potential applications of recurrent cascade models for learning theory

Aside from the additional capabilities to capture biological phenomena in dendrites that we have illustrated here, PRC models may have relevance for machine learning applications. Notably, thanks to the use of differentiable computational graphs (see [335] for an approach to making our somatic units differentiable), a PRC model can be incorporated into any artificial neural network model and trained with gradient descent [214, 336]. As such, PRC models open the door to investigating whether sub-cellular dendritic computations have any potential utility for improving machine learning applications. The answer to this question will depend, in large part, on whether dendritic computations can provide important inductive biases for an artificial neural network [336, 337].

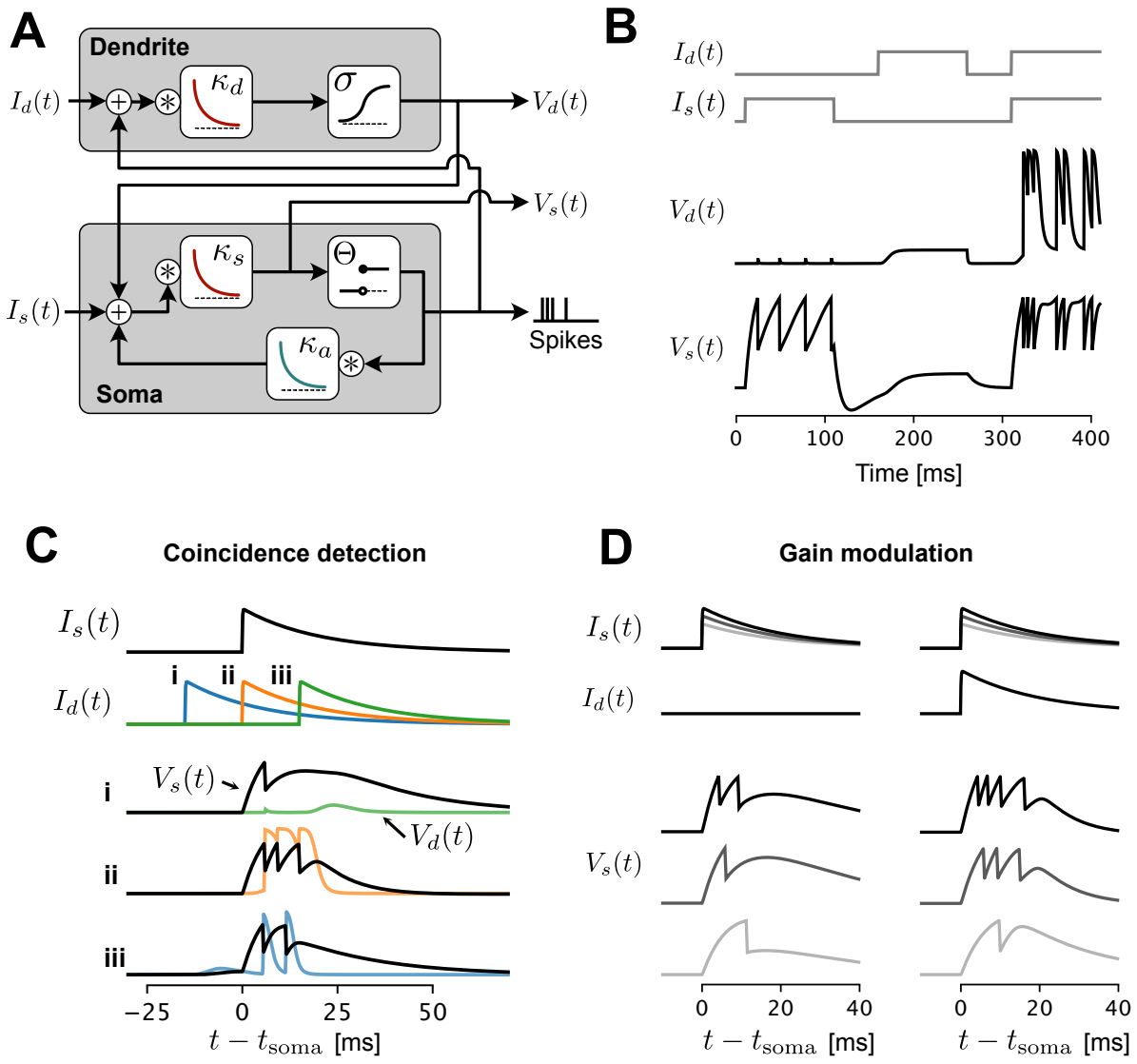
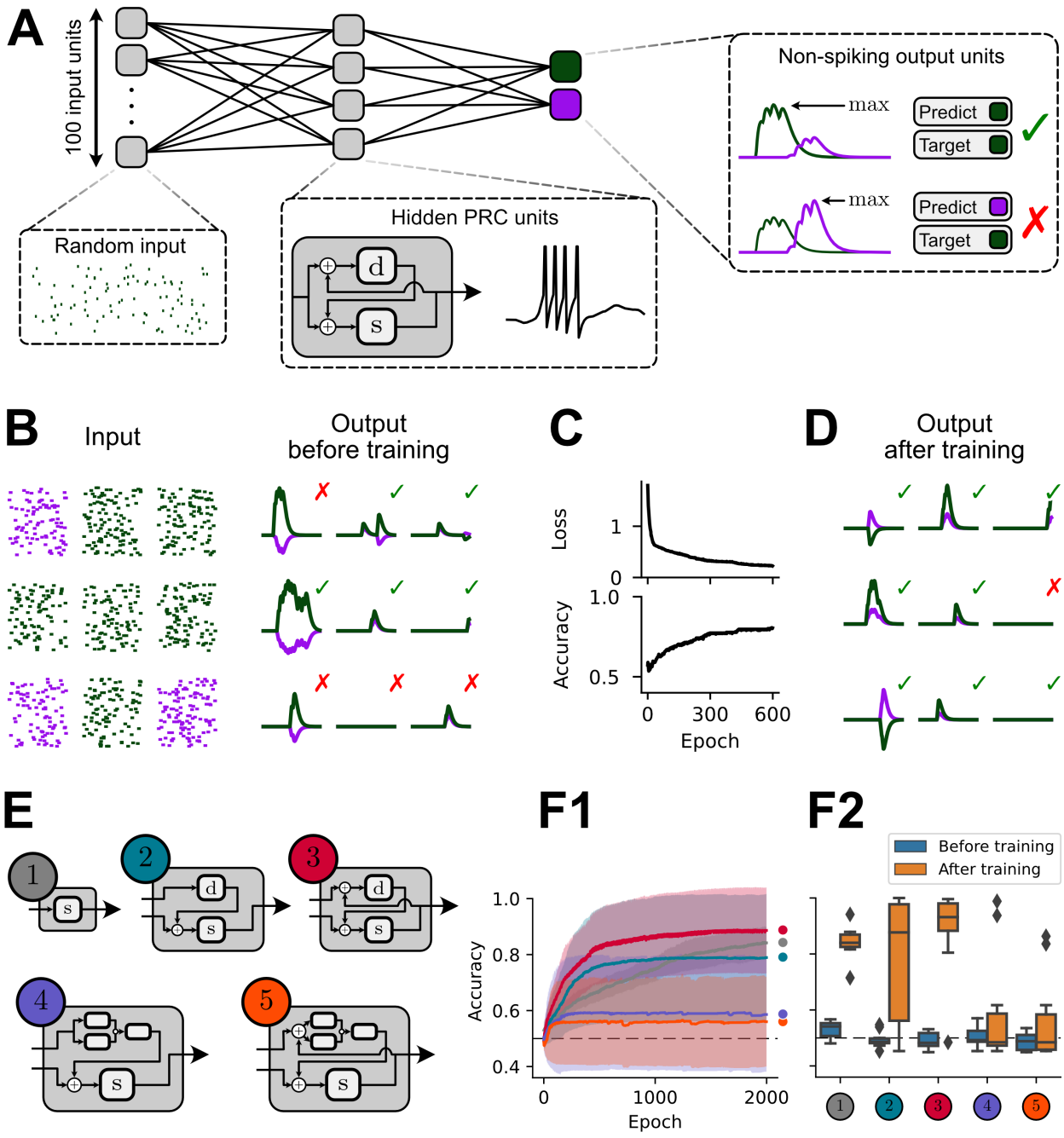


Figure B.6: A recurrent cascade model of the interaction between the back-propagating action potential and calcium spikes. **A** Schematic of the model: external current impinging on the dendrite (I_d) and soma (I_s) each engage a LNL operation. The output of the nonlinearities is combined to the input current of the other unit, forming a recurrent interaction. **B** Step current injections in the soma alone, the dendrite alone and then in both compartments simultaneously produces regular bursting only for the conjunction of inputs. **C** A strong synaptic-like current pulse is injected in the both the soma (I_s) and the dendrite (I_d). Three simulations are shown for three relative timings of the dendritic input (i blue, ii orange, iii green). Responses for the somatic (V_s, black traces) and dendritic (V_d, color traces) compartments are overlaid for each condition. Somatic spikes are denoted by a clear reset but the full depolarization is not shown. A burst of action potentials arise from the near-coincident condition (ii, orange) recapitulating experimental observations in pyramidals of the cortex [331]. **D** Responses to increasing amplitudes of synaptic-like input to soma in the absence of dendritic input (left) and in the presence of a concomitant input in the dendrite (right), a simulation of the gain modulation property of dendritic input reported experimentally [334].

B.3.1 Networks of PRC models can be trained using gradient descent

To illustrate that complex PRC models can be incorporated into multi-layer artificial neural networks and while remaining trainable via gradient descent, we next trained an artificial neural network containing a single hidden layer made of four PRC neurons to solve a simple binary classification problem: memorizing labels associated with random patterns of synaptic input (Fig. B.7A). Before training, our network correctly predicted the label ("green" or "purple") that had been randomly assigned to each randomly-generated pattern of synaptic input ∼50 % of the time, no better than chance. Adjusting the synaptic weights in the direction of decreasing error using backpropagation through time (with a surrogate gradient for the somatic spiking function; see Methods) gradually improved the prediction accuracy to well above the chance level (>75 %, Fig. B.7C, D), indicating that the network had successfully learned a mapping from synaptic input patterns to class labels. We repeated the training procedure with different types of hidden units (Fig. B.7E), ranging from an integrate-and-fire unit without a dendrite (model ①) to the full PRC model (model ⑤), and including recurrently-connected compartments without paralel processing (as in Fig. B.6). For the models with a dendrite (②–⑤ in Fig. B.7), two independent sets of weights were used, such that each input unit connects with one trainable weight to the soma and with another trainable weight to the dendrite. Also, for the more complex models (②–⑤ in Fig. B.7), we fixed as hyper-parameters the connection strength between the somatic and dendritic compartment as well as the time scale of the filters (see Methods). We found that, with the exception of the models with the parallel motif, all models trained well and achieved above 80 % accuracy within 2000 epochs, demonstrating these models are trainable using standard techniques.

Importantly, as with any machine learning model, increasing PRC model complexity increases the number of parameters to be trained and hyperparameters to be tuned. Good


A
100 input units
Random input
Hidden PRC units
d
s
Non-spiking output units
max
Predict
Target
max
Predict
Target
B
Input
Output
before training
C
Loss
1
0
Accuracy
1.0
0.5
0
300
600
Epoch
D
Output
after training
E
1
2
3
4
5
F1
Accuracy
1.0
0.8
0.6
0.4
0
1000
2000
Epoch
F2
Before training
After training
1
2
3
4
5

machine learning practice involves a systematic methodology for tuning the hyper-parameters [338, 339]. Here, however, the particular choice of hyper-parameters was based on our comparison with electrophysiological observations. With these hyper-parameters, good accuracy using the parallel motif was achieved but only in a small portion of the initializations. It is possible that the slow time constant used in the parallel subunits tended to prevent these models from achieving good accuracy. Thus, we have demonstrated that various forms of PRC models can be trained on a memorization task and that combining these models with techniques for hyper-parameter optimization [338, 339] opens the door to comparing model architectures on the basis of training efficacy.

Because few problems solved by machine learning systems or the brain are likely to closely resemble memorization of random patterns of inputs, we next assessed whether our PRC models could be trained to solve a more realistic problem: classification of spoken digits. The Heidelberg digits dataset contains spiketrains produced by a detailed model of the inner ear in response to audio recordings of the numbers zero to nine spoken by multiple people

Figure B.7: (Previous page.) PRC neurons in a network can be trained to memorize patterns of synaptic input. **A** Network architecture with PRC units in the hidden layer. 100 input units send synaptic input to a hidden layer composed of four PRC neurons with recurrently-connected somatic and dendritic subunits (as shown in inset and Fig. B.6A), which in turn send synaptic input to an output layer consisting of two non-spiking single-compartment neurons. Synaptic input patterns are classified as "green" or "purple" based on the maximum voltage attained in the output layer. **B** Randomly-generated patterns of synaptic inputs (left) are incorrectly classified approximately 50 % of the time by a PRC network with randomly-initialized synaptic weights. **C**, **D** Training synaptic weights using backpropagation through time allows the network to memorize labels associated with specific input patterns, resulting in improved predictions. The mean negative log likelihood loss is shown (top). A surrogate gradient was used for the Heaviside nonlinearity in the somatic compartment of hidden PRC neurons (see Methods) in order to train the input weights. **E** Different architectures of the PRC unit used in the hidden layer: ① one-compartment neuron, ② two-compartment neuron with no recurrent connection from the soma to dendrite, ③ two-compartment neuron with recurrent connections between soma and dendrite (as in B.2C and B.6), ④ two-compartment neuron with parallel processing in the dendritic compartment (as in B.5), and ⑤ two-compartment neuron with both parallel processing and recurrent somatodendritic connections. **F1** Mean ± SD accuracy during training for $N = 10$ randomly-initialized networks of each type. Dashed lines indicate chance-level accuracy. **F2** Training set accuracy before and after training is shown for the different models.

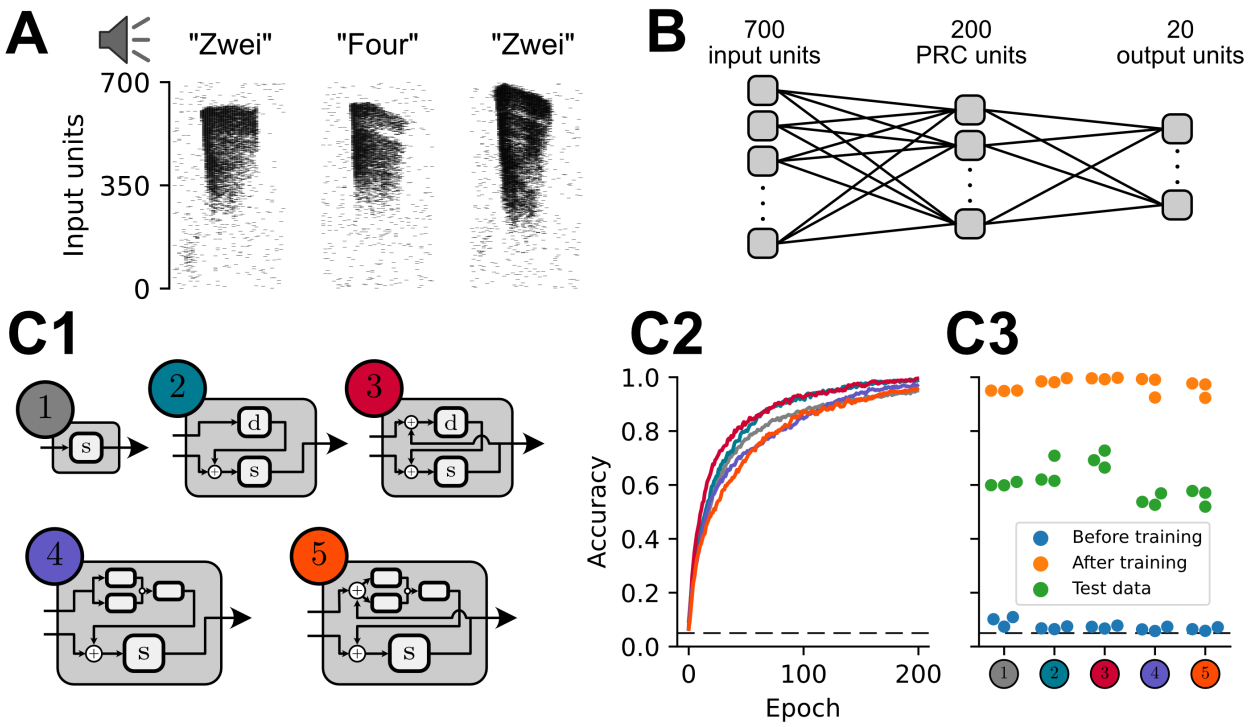


Figure B.8: Multi-layered networks of PRC neurons can be trained to recognize spoken digits, and generalize to sound samples from new speakers. **A** Training examples from the Heidelberg spoken digits dataset [340]. Audio recordings of digits spoken in English and German were transformed into spiketrains using a detailed model of the inner ear. **B** Multi-layered PRC network architecture, an expanded version of the architecture shown in Fig. B.7. **C** Performance of multi-layered PRC networks. Hidden unit architectures (C1) are the same as in Fig. B.7E. C2 shows the mean training accuracy across $N = 3$ randomly-initialized networks. Dashed lines indicate chance-level accuracy. C3 shows test set accuracy (green) as well as initial and end training set accuracy for the different models.

in two different languages (English and German; Fig. B.8A). As a benchmark for assessing the generalization performance of spiking neural network models, the Heidelberg dataset is divided into a standard set of training and test examples, where the test examples include digits spoken by two individuals not included in the training set. To solve this more complex task, we created a set of larger multi-layered PRC network models containing a hidden layer with 200 PRC units in addition to the 700 input units and 20 output units required by the structure of the dataset (Fig. B.8B). Training multi-layered networks containing the same set of five PRC models shown in Fig. B.7E for 200 epochs consistently improved the training set prediction accuracy from the chance level of 5 % to ≥95 % and as high as 99.5 % for the network containing PRC units which capture the effects of backpropagating action potentials (i.e., ③ in Figs. B.7 and B.8; analogous to Fig. B.6; all values are averages across $N = 3$ randomly-initialized networks; Fig. B.7C). The multi-layered PRC networks successfully generalized to new examples, achieving a test set accuracy of 54 % to 70 %, comparable to that previously reported using networks of LIF neurons [340]. The prediction accuracy could be further improved by systematic selection of model hyperparameters, including PRC neuron architecture [338, 339]. Importantly, these preliminary results provide a clear proof-of-concept that multi-layered networks containing PRC neurons can be trained to solve complex tasks.

B.3.2 Parallel-recurrent cascade models as architectural inductive biases

One of the realities that any machine learning system faces is that no learning agent can be efficient at learning all types of problems and/or tasks. Due to the No Free Lunch Theorem for Optimization [285], we know that there is always a trade-off: for a learning system to achieve good performance in one set of problems, it must sacrifice its performance in other sets of problems. When an engineer introduces some design components into an optimization system that help to bias a learning system towards a particular set of problems, we refer to these components as "inductive biases". Inductive biases are key to developing useful machine

learning systems, because without appropriate inductive biases, learning systems will provide mediocre performance on all tasks, as opposed to superior performance on the subset of tasks that we may care about [341, 342]. For example, if an engineer creates a machine learning system that is built with an inductive bias to seek out relationships between discrete objects, then that can help the system learn about spatio-visual object relations such as "there is a blue ball to the left of a red cube" [343, 344].

These insights about the importance of inductive biases actually helped to lay the foundations for the modern deep learning approach in AI [342, 345]. The early proponents of deep learning proposed that the set of problems we most care about in AI are those that humans and/or animals are good at, e.g. image processing, motor control, language comprehension, etc. [345]. Given this, they argued that machine learning researchers should seek inspiration from real brains in order to find appropriate inductive biases for AI [346]. At the time, the principal inductive bias that these researchers were interested in was network *depth* (hence the name "deep" learning). They believed that the macroscopic architecture of the brain, with multiple stages of processing involved in sensorimotor transformations, provided inductive biases to promote hierarchical representations, which was proposed to solve the type of problems where humans and animals excel [345]. In hindsight, it appears that their intuitions were largely correct: deep ANNs have consistently outperformed other types of machine learning approaches in exactly the sort of problem/task domains that humans and animals excel at [346]. Moreover, theoretical analyses have provided some explanations for why ANNs with deep architectures are particularly well suited to such applications [347, 348].

There is a broader, two-fold point within the story of deep learning and inductive biases. First, it is clear now that the inductive biases of an ANN are determined in large part by architecture, i.e. how the linear and non-linear operations are arranged in a computational graph within the ANN. This is because architecture determines both how information flows through the ANN, but also the shape of the loss landscape, which can directly impact the efficiency with which different types of representations can be learned [349, 350]. Second,

the success of using the brain's hierarchical structure to inspire the architecture of ANNs demonstrates that, in principle, it can be beneficial for AI to seek inspiration from the brain when seeking new inductive biases [351]. Importantly, PRC models provide a new way of incorporating biological insights into the design of ANN architectures. If we were to replace the standard units of an ANN with PRC models based on real neurons, this would represent a major change in the architecture of the ANN, one that may provide useful inductive biases.

B.3.3 Potential advantages of PRC-based inductive biases

When considering the potential inductive biases that PRC models introduce, the natural question is, would these inductive biases actually help or hinder AI? Though it is true that inspiration from brains have provided useful inductive biases for machine learning in the past [346, 351], there is no rule that says that neural phenomena always provide such utility. Indeed, some aspects of physiology may be more related to phylogenetic history and biological constraints than they are to improved learning performance. However, there are a few reasons to think that dendrite inspired PRC models may provide useful inductive biases.

As noted above, research over the last decade has confirmed that network depth is an important architectural consideration for AI [347, 348]. However, not all depth is equal. Researchers have found that increasing depth is most useful when additional architectural features are included, such as skip connections [350]. If we were to replace the units of an ANN with PRC models based on real neurons that would increase the depth of the network, but in a very particular way. The central question, then, is would the specific form of increased depth that one would obtain from using PRC model units would be helpful?

One reason that PRC models may provide a useful form of depth is that they would help to promote sparsity, which has also been shown to be useful in neural networks [352]. PRC models would help to promote sparsity because they would provide distinct computational sub-units with non-linear interactions that could handle different components of a task. For example, it would be possible to have individual dendrites that are responsible for processing

distinct types of features, e.g. one set of dendrites for processing facial features, another set for processing body parts, etc. Thus, depending on the input provided, the network could activate only a very sparse set of all the dendrites for processing.

Related to this, there is a growing recognition in the machine learning community that a desirable form of inductive biases for AI would be those that promote the emergence of specialized modules that can be flexibly composed [353, 354]. This may turn out to be critical to overcoming the limitations of current ANN approaches. Specifically, if an ANN was provided with an architecture that promoted the learning of distinct, specialized modules that can be composed, then it should be possible both to learn more intuitive part-whole relationships that capture the underlying structure of objects in the world more accurately [353, 354] and to avoid the catastrophic forgetting that can plague standard ANNs [355]. However, current systems to promote specialized modules are only loosely inspired by real neural circuits. Therefore, an interesting open question is whether PRC models could provide a good mechanism for implementing brain-inspired inductive biases to promote composability. Notably, dendrites have functionally clustered inputs [356], synaptic dynamics can perform LNL operations [357], and these phenomena interact [358]. Given the fact that PRC models would allow an ANN to learn dendrite-like, flexible, non-linear, recurrent interactions between functionally clustered inputs, it is plausible that they could help with composability. Thus, we would argue that future research should investigate the potential for ANNs that use PRC models inspired by sub-cellular dendritic computation to show better specialization and composability, and less catastrophic forgetting.

B.4 Discussion

In this article, we have illustrated the capabilities of PRC models to capture the various experimentally observed features of dendritic computation, and discussed how this modeling framework may be key to understand the role of dendrites in learning and neuronal compu-

tation. In doing so, we have identified modular, composable connectivity motifs (Fig. B.2), with the parallel, recurrent and cascade elements forming the basic building blocks of the framework. We have illustrated how the interaction between sodium spikes in the dendrite and the soma [327] can be captured using a recurrent motif of LNL subunits (Fig. B.4). Furthermore, we have demonstrated that a parallel motif of LNL subunits within a cascade can reproduce the dependence of NMDA spikes on the activation of sodium and calcium channels [330]. Finally, we demonstrated that the interaction between back-propagating action potentials and calcium spikes can be captured by a recurrent motif of LNL subunits with a different set of parameters (Fig. B.3) [331, 334]. In closing, we have argued that these PRC models of dendritic computation could have an important role in shaping inductive biases (Section B.3), and thus contribute to the optimization of learning capabilities of the brain.

Our LNL models, however, bear some important limitations. Firstly, the set of operations that are possible within the LNL framework correspond to a subset of the operations that are achievable by the type of dynamical systems used in detailed simulations of dendritic integration [214, 330, 359, 360]. For instance, modeling the NMDA spike as two LNL units in parallel followed by another nonlinearity ignores the nonlinear impact that sodium ion channels can have on calcium ion channels via rapid increase of the membrane potential. Another element not captured by the phenomenological LNL model is that ion channel time constants almost always depend on the mean depolarization, which implies an adaptive filter instead of a fixed filter assumed in the LNL model. A second important limitation is that we have assumed that the dendrites remain in a fluctuation-driven regime where the net input is low on average but highly variable. If we were to give strong and sustained inputs to the LNL model, these would saturate the nonlinearity and nonlinear transients would disappear. Such a sustained-depolarization regime has been observed in some experiments [334], but it remains to be seen whether these take place *in vivo* or whether homeostatic mechanisms preserve the fluctuation-driven regime [361]. One last limitation of our model is that a very high input variability is capable of making nonlinear operations effectively linear [214]. In this

case, the complex linear-nonlinear structure operates in a way that can actually be captured by an effective model that is entirely linear. Although some *in vivo* manipulation of dendritic inputs [333, 362, 363] argue against this point of view, the full relationship between inputs and outputs of neurons in a naturalistic condition is far from being fully known.

We included in this paper some discussion of the potential machine learning applications of PRC models. As we outlined, there are some reasons to believe that dendritic computations may provide useful inductive biases for machine learning systems. We are hopeful that future research will demonstrate this. However, it also has to be recognized that real dendrites may be solving an implementation problem for neurons, i.e. how can you actually integrate thousands of distinct signals in a physical circuit with space and energy constraints? It is possible that this is the problem which dendrites solve for real neurons, and that dendritic computation is, itself, not important at the algorithmic level. Only by exploring the potential advantages of training AI systems with PRC models inspired by real neurons will we be able to get some initial insight on this mystery.

In summary, our work shows how PRC models can be used to model sub-cellular dendritic computation with a computationally tractable approach. This lays the groundwork for future explorations of the algorithmic implications of dendritic computation, both in the brain, and in machine learning applications. We believe that PRC models will help open the door to exploring the true computational power of dendrites.

B.5 Methods

B.5.1 Linear-nonlinear subunit

The basic component of the modeling framework presented here is the linear-nonlinear subunit, which receives a net time-varying input $I_i(t)$ and produces an activation $z(t)$ as its output

$$z(t;\kappa_{\text{nlin}},\kappa_{\text{lin}},\kappa_{\text{ad}}) = g(a_{\text{nlin}}(t;\kappa_{\text{nlin}},\kappa_{\text{ad}})) + a_{\text{lin}}(t;\kappa_{\text{lin}},\kappa_{\text{ad}}) \tag{B.1}$$

$$a_x(t;\kappa_x,\kappa_{\text{ad}}) = \left[\kappa_x * (I_{\text{ext}} + I_{\text{ad}})\right](t) \tag{B.2}$$

$$I_{\text{ad}}(t;\kappa_{\text{ad}}) = [\kappa_{\text{ad}} * z](t), \tag{B.3}$$

where $g(\cdot)$ is a nonlinear activation function; $a_{\text{nlin}}(t)$ and $a_{\text{lin}}(t)$ are the pre-activations for the nonlinear and linear parts of the output, respectively; $I_{\text{ext}}(t) = \sum_i I_i(t)$ is the total input from all external sources; $I_{\text{ad}}(t)$ is a recurrent adaptation current; and $(\kappa * x)(t) = \int_{-\infty}^{t} \kappa(t-\tau)x(\tau)d\tau$ denotes the causal convolution of a filter κ with a signal x evaluated at time t. The subunit is parameterized in terms of the activation filters κ_{nlin} and κ_{lin}, the adaptation filter κ_{ad}, and the choice of nonlinear activation function $g(\cdot)$. Except where noted, all filters in this work are defined as exponential functions

$$\kappa(t) = \frac{1}{\tau} e^{\frac{-t}{\tau}} \tag{B.4}$$

with time constant τ. For models with multiple cellular compartments, $I_{\text{ext}}(t)$ may include inputs from an external source as well as from other compartments (eg, Methods B.5.4). Depending on the model, $z(t)$, $a_{\text{nlin}}(t)$, or $a_{\text{lin}}(t)$ may correspond loosely to the voltage of a compartment denoted $V_x(t)$ where x is the name of the compartment.

B.5.2 Two-compartment model subject to a single dendritic input

The two-compartment model with a recurrent connection from the somatic to the dendritic compartment is defined as follows

$$z^{(\mathrm{s})}(t) = g(a^{(\mathrm{s})}_{\mathrm{nlin}}(t;\kappa^{(\mathrm{s})}_{\mathrm{nlin}})) + a^{(\mathrm{s})}_{\mathrm{lin}}(t;\kappa^{(\mathrm{s})}_{\mathrm{lin}}) \tag{B.5}$$

$$z^{(\mathrm{d})}(t) = \sigma(a^{(\mathrm{d})}_{\mathrm{nlin}}(t;\kappa^{(\mathrm{d})}_{\mathrm{nlin}})) + a^{(\mathrm{d})}_{\mathrm{lin}}(t;\kappa^{(\mathrm{d})}_{\mathrm{lin}}) \tag{B.6}$$

$$I^{(\mathrm{s})}_{\mathrm{ext}}(t) = I^{(\mathrm{d})}_{\mathrm{ext}}(t) = I_d(t), \tag{B.7}$$

where the dendritic nonlinearity $\sigma(\cdot)$ is the sigmoid function. $g(\cdot)$ is the somatic spiking nonlinearity, a function which emits a 1 ms square pulse of amplitude $A = 2$ AU when $z^{(\mathrm{s})}(t)$ crosses the somatic spike threshold from below

$$g(t;v) = rect(t) * \sum_{t^{(f)}<t} \delta(t - t^{(f)}) \tag{B.8}$$

$$rect(t) = \begin{cases} A & \text{for } 0 < t < 1ms \\ 0 & \text{otherwise} \end{cases}, \tag{B.9}$$

where $t^{(f)}$ denotes the time of a threshold crossing. Exponential functions with the following time constants were used for the pre-activation filters: $\tau^{(\mathrm{s})}_{\mathrm{lin}} = 40$ ms, $\tau^{(\mathrm{d})}_{\mathrm{lin}} = 20$ ms, $\tau^{(\mathrm{s})}_{\mathrm{nlin}} = 40$ ms, and $\tau^{(\mathrm{d})}_{\mathrm{nlin}} = 2$ ms. Adaptation filters were set to zero and the terms associated with them were omitted from the above model for simplicity. The activation output $z^{(x)}(t)$ corresponds loosely to the voltage of each compartment; $V_s(t) \equiv z^{(s)}(t)$ and $V_d(t) \equiv z^{(d)}(t)$ are therefore used to refer to these terms in figures and the main text for ease of interpretation.

B.5.3 Multi-subunit model with parallel processing

The model is composed of three linear-nonlinear subunits which loosely capture the contributions of sodium, calcium, and NMDA voltage-dependent conductances (denoted by the

superscripts (1), (2), and (3), respectively) to nonlinear processing of synaptic inputs in a dendritic compartment. Their dynamics are defined as follows

$$z^{(\mathrm{x})}(t) = \sigma\left(a^{(\mathrm{x})}(t;\kappa^{(\mathrm{x})})\right) + a^{(\mathrm{x})}(t;\kappa^{(\mathrm{x})}) \qquad \text{for } \mathrm{x} \in \{1,2,3\} \tag{B.10}$$

$$\kappa^{(\mathrm{x})} = \frac{1}{\tau^{(\mathrm{x})}} e^{\frac{-t}{\tau^{(\mathrm{x})}}}, \tag{B.11}$$

with inputs

$$I^{(1)}_{\mathrm{ext}}(t) = I^{(2)}_{\mathrm{ext}}(t) = I_d(t) \tag{B.12}$$

$$I^{(3)}_{\mathrm{ext}}(t) = z^{(1)}(t) + z^{(2)}(t). \tag{B.13}$$

In all three subunits, the non-linear and linear pre-activation filters were set to be equal, such that

$$\kappa^{(\mathrm{x})} = \kappa^{(\mathrm{x})}_{\mathrm{lin}} = \kappa^{(\mathrm{x})}_{\mathrm{nlin}} \tag{B.14}$$

where $\tau^{(1)} = 5\,\mathrm{ms}$, $\tau^{(2)} = 40\,\mathrm{ms}$, and $\tau^{(3)} = 80\,\mathrm{ms}$. The adaptation filters were set to zero and the associated terms dropped for simplicity. The dendritic synaptic-like input current is given by the alpha function

$$I_d(t) = Ate^{\frac{-t}{\tau}} \qquad \text{for } t \geq 0 \tag{B.15}$$

with amplitude A and time constant $\tau = 2\,\mathrm{ms}$. The voltage output shown in the figures and main text is analogous to the activation output $V_{\mathrm{x}}(t) \equiv z^{(\mathrm{x})}(t)$ for each respective subunit.

B.5.4 Two-compartment model with bi-directional dendro-somatic interactions

The model with bi-directional dendro-somatic interactions is composed of two reciprocally-connected linear-nonlinear subunits (see Section B.5.1) as follows

$$z^{(\mathrm{s})}(t) = \Theta\left(a^{(\mathrm{s})}_{\mathrm{nlin}}(t; \kappa^{(\mathrm{s})}_{\mathrm{nlin}}, \kappa^{(\mathrm{s})}_{\mathrm{ad}})\right) \tag{B.16}$$

$$I^{(\mathrm{s})}_{\mathrm{ext}}(t) = I_s(t) + z^{(\mathrm{d})}(t) \tag{B.17}$$

$$z^{(\mathrm{d})}(t) = \sigma\left(a^{(\mathrm{d})}_{\mathrm{nlin}}(t; \kappa^{(\mathrm{d})}_{\mathrm{nlin}})\right) \tag{B.18}$$

$$I^{(\mathrm{d})}_{\mathrm{ext}}(t) = I_d(t) + z^{(\mathrm{s})}(t), \tag{B.19}$$

where $\Theta(\cdot)$ is the Heaviside step function and $\sigma(\cdot)$ is the sigmoid function. The nonlinear activation filters $\kappa^{(x)}_{\mathrm{nlin}}$ are defined as exponential functions with $\tau^{(\mathrm{s})}_{\mathrm{nlin}} = 10\,\mathrm{ms}$ and $\tau^{(\mathrm{d})}_{\mathrm{nlin}} = 5\,\mathrm{ms}$. The somatic adaptation filter is defined as

$$\kappa^{(\mathrm{s})}_{\mathrm{ad}}(t) = \begin{cases} 1 & \text{for } t = 0 \\ \frac{1}{\tau_{\mathrm{ad}}} e^{\frac{-t}{\tau_{\mathrm{ad}}}} & \text{for } t > 0 \end{cases} \tag{B.20}$$

with $\tau_{\mathrm{ad}} = 20\,\mathrm{ms}$. In this model, the linear activation filters $\kappa^{(\mathrm{s})}_{\mathrm{lin}}$ and $\kappa_{(\mathrm{d})}$ are set to zero, along with the dendritic adaptation filter $\kappa^{(\mathrm{d})}_{\mathrm{ad}}$. (The terms associated with these filters have been dropped from the above model definition for simplicity.) $I_s(t)$ and $I_d(t)$ correspond loosely to synaptic inputs to the somatic and dendritic compartments, respectively. The somatic pre-activation $a^{(\mathrm{s})}_{\mathrm{nlin}}$ and the dendritic activation $z^{(\mathrm{d})}$ loosely correspond to the voltage in their respective compartments. For clarity, we use $V_s(t) \equiv a^{(\mathrm{s})}_{\mathrm{nlin}}(t)$ and $V_d(t) \equiv z^{(\mathrm{d})}(t)$ to refer to these quantities in the figures and main text.

B.5.5 Multi-layered networks of PRC models

Artificial spiking neural networks with a hidden layer of PRC neurons were constructed following the approach of Refs. [318, 340, 364]. Briefly, spikes in each layer were integrated as exponentially-decaying synaptic currents in each neuron/PRC-compartment in the following layer, which in turn were integrated by the dynamics of the corresponding PRC subunit as $I_{\text{ext}}(t)$. The neurons in the output layer consisted of leaky integrators without threshold or reset. The time-varying voltage of the neurons in the output layer was transformed into a set of class probabilities by applying a softmax operation to the maximum voltage attained by each output unit during each example.

The PRC filters κ used in network models were defined as

$$\kappa_{\text{x}}(t) = e^{\frac{-t}{\tau_{\text{x}}}} \qquad \text{for } t \geq 0 \tag{B.21}$$

where the membrane time constant $\tau_{\text{nlin}} = 10\,\text{ms}$, and the synaptic time constant $\tau_{\text{syn}} = 5\,\text{ms}$. Time constants of sodium, calcium, and NMDA filters took the same values as in Methods B.5.3. The linear forward filters κ_{lin} were dropped for simplicity. The adaptation filter in the somatic compartment was defined as

$$\kappa_{\text{ad}}^{(\text{s})} = \begin{cases} 15 & \text{for } t = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{B.22}$$

In order to train multi-layered networks of PRC neurons, it was necessary to use a surrogate gradient for the Heaviside step function used to generate spikes in the somatic compartment (since its gradient is zero almost everywhere). Following the approach of [335, 364], we used the normalized gradient of a sigmoid function to approximate the gradient of the Heaviside function

$$\frac{d}{dx}\Theta(x) \approx \frac{\alpha}{(s|x|+1)^2},$$

where α is a proportionality constant and $s = 10$ is a scale parameter that sets the slope of the sigmoid.

All training was carried out using the Adam optimization algorithm with a learning rate of 0.002 and the negative log likelihood of making a correct class prediction as the loss function.

B.5.6 Numerical methods

Simulations were implemented in Matlab and Python 3.8 using NumPy 1.18.5, SciPy 1.5.0, and ez-ephys 0.4.2. Figures were prepared in Python using Matplotlib 3.2.2, Jupyter 1.0.0, and ez-ephys. Code is available at .

Bibliography

[1] D. E. Nichols and C. D. Nichols. Serotonin receptors. *Chemical Reviews* **108**, no. 5, 1614–1641. (2008)

[2] E. C. Azmitia. Chapter 1: Evolution of serotonin: sunlight to suicide in *Handbook of Behavioral Neuroscience*, pp. 3–22. C. P. Müller and K. A. Cunningham, eds. (Elsevier, 2020)

[3] B. L. Jacobs and C. A. Fornal. Chapter 41: Serotonin and behavior in *Psychopharmacology: The Fourth Generation of Progress*, pp. 461–469. F. E. Bloom and D. J. Kupfer, eds. (Raven Press, 1995)

[4] R. R. Griffiths, M. W. Johnson, M. A. Carducci, A. Umbricht, W. A. Richards, B. D. Richards, M. P. Cosimano, and M. A. Klinedinst. Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomized double-blind trial. *Journal of Psychopharmacology* **30**, no. 12, 1181–1197. (2016)

[5] H. Meltzer and B. Massey. The role of serotonin receptors in the action of atypical antipsychotic drugs. *Current Opinion in Pharmacology* **11**, no. 1, 59–67. (2011)

[6] S. Leucht, A. Cipriani, L. Spineli, *et al.* Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *The Lancet* **382**, no. 9896, 951–962. (2013)

[7] R. S. Goldberg and W. E. Thornton. Combined tricyclic–MAOI therapy for refractory depression: A review, with guidelines for appropriate usage. *The Journal of Clinical Pharmacology* **18**, no. 2–3, 143–147. (1978)

[8] G. M. Soomro, D. G. Altman, S. Rajagopal, and M. O. Browne. Selective serotonin re-uptake inhibitors (SSRIs) versus placebo for obsessive compulsive disorder (OCD). *Cochrane Database of Systematic Reviews*. (2008) doi: `https://doi.org/10.1002/14651858.CD001765.pub3`

[9] A. Cipriani, T. A. Furukawa, G. Salanti, *et al.* Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *The Lancet* **391**, no. 10128, 1357–1366. (2018)

[10] E. Jakubovski, J. A. Johnson, M. Nasir, K. Müller-Vahl, and M. H. Bloch. Systematic rewview and meta-analysis: Dose–response curve of SSRIs and SNRIs in anxiety disorders. *Depression and Anxiety* **36**, no. 3, 198–212. (2019)

[11] J. Moncrieff, R. E. Cooper, T. Stockmann, S. Amendola, M. P. Hengartner, and M. A. Horowitz. The serotonin theory of depression: a systematic umbrella review of the evidence. *Molecular Psychiatry* **28**, 3243–3256. (2022)

[12] S. Jauhar, D. Arnone, D. S. Baldwin, *et al.* A leaky umbrella has little value: evidence clearly indicates the serotonin system is implicated in depression. *Molecular Psychiatry* **28**, 3149–3152. (2023)

[13] C. Cantor. Study shows brain's "pleasure chemical" triggered by pain too. *Columbia University Department of Psychiatry*. Web. (2022) `https://www.columbiapsychiatry.org/news/study-shows-brains-pleasure-chemical-triggered-pain-too`

[14] H. Steinbusch. Distribution of serotonin-immunoreactivity in the central nervous system of the rat—Cell bodies and terminals. *Neuroscience* **6**, no. 4, 557–618. (1981)

[15] R. Vertes. A PHA-L analysis of ascending projections of the dorsal raphe nucleus in the rat. *The Journal of Comparative Neurology* **313**, 643–668. (1991)

[16] D. Gagnon and M. Parent. Distribution of VGLUT3 in highly collateralized axons from the rat dorsal raphe nucleus as revealed by single-neuron reconstructions. *PLoS ONE* **9**, no. 2, e87709. (2014)

[17] B. Weissbourd, J. Ren, K. DeLoach, C. Guenthner, K. Miyamichi, and L. Luo. Presynaptic partners of dorsal raphe serotonergic and GABAergic neurons. *Neuron* **83**, no. 3, 645–662. (2014)

[18] I. Pollak Dorocic, D. Fürth, Y. Xuan, Y. Johansson, L. Pozzi, G. Silberberg, M. Carlén, and K. Meletis. A whole-brain atlas of inputs to serotonergic neurons of the dorsal and median raphe nuclei. *Neuron* **83**, no. 3, 663–678. (2014)

[19] L. Zhou, M.-Z. Liu, Q. Li, J. Deng, D. Mu, and Y.-G. Sun. Organization of functional long-range circuits controlling the activity of serotonergic neurons in the dorsal raphe nucleus. *Cell Reports* **18**, no. 12, 3018–3032. (2017)

[20] M. E. Trulson and B. L. Jacobs. Raphe unit activity in freely moving cats: Correlation with level of behavioral arousal. *Brain Research* **163**, no. 1, 135–150. (1979)

[21] G. Aghajanian and C. Vandermaelen. Intracellular identification of central noradrenergic and serotonergic neurons by a new double labeling procedure. *The Journal of Neuroscience* **2**, no. 12, 1786–1792. (1982)

[22] G. Aghajanian and C. Vandermaelen. Intracellular recordings from serotonergic dorsal raphe neurons: pacemaker potentials and the effects of LSD. *Brain Research* **238**, no. 2, 463–469. (1982)

[23] S. P. Ranade and Z. F. Mainen. Transient firing of dorsal raphe neurons encodes diverse and specific sensory, motor, and reward events. *Journal of Neurophysiology* **102**, no. 5, 3026–3037. (2009)

[24] J. Y. Cohen, M. W. Amoroso, and N. Uchida. Serotonergic neurons signal reward and punishment on multiple timescales. *eLife* **4**, e06346. (2015)

[25] C. D. Grossman, B. A. Bari, and J. Y. Cohen. Serotonin neurons modulate learning rate through uncertainty. *Current Biology* **32**, no. 3, 586–599. (2022)

[26] C. Vandermaelen and G. Aghajanian. Electrophysiological and pharmacological characterization of serotonergic dorsal raphe neurons recorded extracellularly and intracellularly in rat brain slices. *Brain Research* **289**, no. 1-2, 109–119. (1983)

[27] L. H. Calizo, A. Akanwa, X. Ma, Y.-z. Pan, J. C. Lemos, C. Craige, L. A. Heemstra, and S. G. Beck. Raphe serotonin neurons are not homogenous: Electrophysiological, morphological and neurochemical evidence. *Neuropharmacology* **61**, no. 3, 524–543. (2011)

[28] S. D. Geddes, S. Assadzada, D. Lemelin, A. Sokolovski, R. Bergeron, S. Haj-Dahmane, and J.-C. Béïque. Target-specific modulation of the descending prefrontal cortex inputs to the dorsal raphe nucleus by cannabinoids. *Proceedings of the National Academy of Sciences* **113**, no. 19, 5429–5434. (2016)

[29] M. de Léon, R. Coveñas, J. Narváez, G. Tramu, J. Aguirre, and S. González-Barón. Distribution of somatostatin-28 (1–12) in the cat brainstem: an immunocytochemical study. *Neuropeptides* **21**, no. 1, 1–11. (1992)

[30] W. Fu, E. Le Maître, V. Fabre, J.-F. Bernard, Z.-Q. David Xu, and T. Hökfelt. Chemical neuroanatomy of the dorsal raphe nucleus and adjacent structures of the mouse brain. *The Journal of Comparative Neurology* **518**, no. 17, 3464–3494. (2010)

[31] R. Wang and G. Aghajanian. Correlative firing patterns of serotonergic neurons in rat dorsal raphe nucleus. *Journal of Neuroscience* **2**, 11–16. (1982)

[32] M. B. Lynn, S. Geddes, M. Chahrour, S. Maillé, E. Harkin, É. Harvey-Girard, S. Haj-Dahmane, R. Naud, and J.-C. Béïque. A slow 5-HT_{1A}R-mediated recurrent inhibitory network in raphe computes contextual value through synaptic facilitation. *bioRxiv*. (2022) doi: `https://doi.org/10.1101/2022.08.31.506056`

[33] B. W. Okaty, K. G. Commons, and S. M. Dymecki. Embracing diversity in the 5-HT neuronal system. *Nature Reviews Neuroscience* **20**, no. 7, 397–424. (2019)

[34] J. K. Abrams, P. L. Johnson, J. H. Hollis, and C. A. Lowry. Anatomic and functional topography of the dorsal raphe nucleus. *Annals of the New York Academy of Sciences* **1018**, no. 1, 46–57. (2004)

[35] C. A. Lowry, P. L. Johnson, A. Hay-Schmidt, J. Mikkelsen, and A. Shekhar. Modulation of anxiety circuits by serotonergic systems. *Stress* **8**, no. 4, 233–246. (2005)

[36] K. G. Commons. Two major network domains in the dorsal raphe nucleus. *Journal of Comparative Neurology* **523**, no. 10, 1488–1504. (2015)

[37] S. P. Fernandez, B. Cauli, C. Cabezas, A. Muzerelle, J.-C. Poncer, and P. Gaspar. Multiscale single-cell analysis reveals unique phenotypes of raphe 5-HT neurons projecting to the forebrain. *Brain Structure and Function* **221**, no. 8, 4007–4025. (2016)

[38] J. Ren, D. Friedmann, J. Xiong, *et al.* Anatomically defined and functionally distinct dorsal raphe serotonin sub-systems. *Cell* **175**, no. 2, 472–487. (2018)

[39] G. E. Paquelet, K. Carrion, C. O. Lacefield, P. Zhou, R. Hen, and B. R. Miller. Single-cell activity and network properties of dorsal raphe nucleus serotonin neurons during emotionally salient behaviors. *Neuron* **110**, no. 16, 2664–2679. (2022)

[40] J. S. Templin, S. J. Bang, M. Soiza-Reilly, C. B. Berde, and K. G. Commons. Patterned expression of ion channel genes in mouse dorsal raphe nucleus determined with the Allen Mouse Brain Atlas. *Brain Research* **1457**, 1–12. (2012)

[41] A. Alonso, P. Merchán, J. E. Sandoval, L. Sánchez-Arrones, A. Garcia-Cazorla, R. Artuch, J. L. Ferrán, M. Martínez-de-la-Torre, and L. Puelles. Development of the serotonergic cells in murine raphe nuclei and their relations with rhombomeric domains. *Brain Structure and Function* **218**, no. 5, 1229–1277. (2013)

[42] E. W. Prouty, D. J. Chandler, and B. D. Waterhouse. Neurochemical differences between target-specific populations of rat dorsal raphe projection neurons. *Brain Research* **1675**, 28–40. (2017)

[43] P. Soubrié. Reconciling the role of central serotonin neurons in human and animal behavior. *Behavioral and Brain Sciences* **9**, no. 2, 319–335. (1986)

[44] T. Kawashima, M. F. Zwart, C.-T. Yang, B. D. Mensh, and M. B. Ahrens. The serotonergic system tracks the outcomes of actions to mediate short-term motor learning. *Cell* **167**, no. 4, 933–946.e20. (2016)

[45] A.-S. Ries, T. Hermanns, B. Poeck, and R. Strauss. Serotonin modulates a depression-like state in Drosophila responsive to lithium treatment. *Nature Communications* **8**, no. 1, 15738. (2017)

[46] C. Seo, A. Guru, M. Jin, *et al.* Intense threat switches dorsal raphe serotonin neurons to a paradoxical operational mode. *Science* **363**, no. 6426, 538–542. (2019)

[47] P. A. Correia, E. Lottem, D. Banerjee, A. S. Machado, M. R. Carey, and Z. F. Mainen. Transient inhibition and long-term facilitation of locomotion by phasic optogenetic activation of serotonin neurons. *eLife* **6**, e20975. (2017)

[48] T. Teshiba, A. Shamsian, B. Yashar, S.-R. Yeh, D. H. Edwards, and F. B. Krasne. Dual and opposing modulatory effects of serotonin on crayfish lateral giant escape command neurons. *The Journal of Neuroscience* **21**, no. 12, 4523–4529. (2001)

[49] D. A. Slattery and J. F. Cryan. Using the rat forced swim test to assess antidepressant-like activity in rodents. *Nature Protocols* **7**, no. 6, 1009–1014. (2012)

[50] A. Can, D. T. Dao, M. Arad, C. E. Terrillion, S. C. Piantadosi, and T. D. Gould. The mouse forced swim test. *Journal of Visualized Experiments* **59**, e3638. (2012)

[51] M. R. Warden, A. Selimbeyoglu, J. J. Mirzabekov, M. Lo, K. R. Thompson, S.-Y. Kim, A. Adhikari, K. M. Tye, L. M. Frank, and K. Deisseroth. A prefrontal cortex–brainstem neuronal projection that controls response to behavioural challenge. *Nature* **492**, no. 7429, 428–432. (2012)

[52] A. C. Volkers, J. H. Tulen, W. W. Van Den Broek, J. A. Bruijn, J. Passchier, and L. Pepplinkhuizen. 24-Hour motor activity after treatment with imipramine or fluvoxamine in major depressive disorder. *European Neuropsychopharmacology* **12**, no. 4, 273–278. (2002)

[53] M. V. Heinz, G. D. Price, F. Ruan, R. J. Klein, M. Nemesure, A. Lopez, and N. C. Jacobson. Association of selective serotonin reuptake inhibitor use with abnormal physical movement patterns as detected using a piezoelectric accelerometer and deep learning in a nationally representative sample of noninstitutionalized persons in the US. *JAMA Network Open* **5**, no. 4, e225403. (2022)

[54] K. Miyazaki, K. Miyazaki, K. Tanaka, A. Yamanaka, A. Takahashi, S. Tabuchi, and K. Doya. Optogenetic activation of dorsal raphe serotonin neurons enhances patience for future rewards. *Current Biology* **24**, no. 17, 2033–2040. (2014)

[55] M. Fonseca, M. Murakami, and Z. Mainen. Activation of dorsal raphe serotonergic neurons promotes waiting but is not reinforcing. *Current Biology* **25**, no. 3, 306–315. (2015)

[56] E. Lottem, D. Banerjee, P. Vertechi, D. Sarra, M. oude Lohuis, and Z. F. Mainen. Activation of serotonin neurons promotes active persistence in a probabilistic foraging task. *Nature Communications* **9**, no. 1, 1000. (2018)

[57] S. Matias, E. Lottem, G. P. Dugué, and Z. F. Mainen. Activity patterns of serotonin neurons underlying cognitive flexibility. *eLife* **6**, e20552. (2017)

[58] W. Zhong, Y. Li, Q. Feng, and M. Luo. Learning and stress shape the reward response patterns of serotonin neurons. *The Journal of Neuroscience* **37**, no. 37, 8863–8875. (2017)

[59] E. S. Bromberg-Martin, O. Hikosaka, and K. Nakamura. Coding of task reward value in the dorsal raphe nucleus. *The Journal of Neuroscience* **30**, no. 18, 6262–6272. (2010)

[60] K.-i. Okada, K. Nakamura, and Y. Kobayashi. A neural correlate of predicted and actual reward-value information in monkey pedunculopontine tegmental and dorsal raphe nucleus during saccade tasks. *Neural Plasticity* **2011**, 1–21. (2011)

[61] M. Luo, Y. Li, and W. Zhong. Do dorsal raphe 5-HT neurons encode "beneficialness"? *Neurobiology of Learning and Memory* **135**, 40–49. (2016)

[62] K. Miyazaki, K. W. Miyazaki, A. Yamanaka, T. Tokuda, K. F. Tanaka, and K. Doya. Reward probability and timing uncertainty alter the effect of dorsal raphe serotonin neurons on patience. *Nature Communications* **9**, no. 1, 2048. (2018)

[63] Y.-Y. Feng, E. S. Bromberg-Martin, and I. E. Monosov. Dorsal raphe neurons signal integrated value during multi-attribute decision-making. *bioRxiv*. (2023) doi: `https://doi.org/10.1101/2023.08.17.553745`

[64] R. S. Sutton and A. Barto. *Reinforcement Learning: An Introduction*, 2nd ed. (MIT Press, 2018)

[65] M. Srinivasan, S. Laughlin, and A. Dubs. Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society of London* **216**, no. 1205, 427–459. (1982)

[66] R. S. Sutton and A. G. Barto. "Early history of reinforcement learning" in Chap. 1: Introduction of *Reinforcement Learning*, 2nd ed., pp. 13–22. (The MIT Press, 2018)

[67] R. S. Sutton and A. G. Barto. Chapter 14: Psychology in *Reinforcement Learning*, 2nd ed., pp. 341–376. (The MIT Press, 2018)

[68] E. L. Thorndike. The law of effect. *The American Journal of Psychology* **39**, 212–222. (1927)

[69] G. Tesauro. Temporal difference learning and TD-Gammon. *Communications of the Association for Computing Machinery* **38**, no. 3, 58–68. (1995)

[70] W. Schultz, P. Dayan, and P. R. Montague. A neural substrate of prediction and reward. *Science* **275**, no. 5306, 1593–1599. (1997)

[71] L. T. Coddington, S. E. Lindo, and J. T. Dudman. Mesolimbic dopamine adapts the rate of learning from action. *Nature* **614**, no. 7947, 294–302. (2023)

[72] H. Jeong, A. Taylor, J. R. Floeder, M. Lohmann, S. Mihalas, B. Wu, M. Zhou, D. A. Burke, and V. M. K. Namboodiri. Mesolimbic dopamine release conveys causal associations. *Science* **378**, no. 6626. (2022) doi: `https://doi.org/10.1126/science.abq6740`

[73] N. D. Daw, S. Kakade, and P. Dayan. Opponent interactions between serotonin and dopamine. *Neural Networks* **15**, no. 4-6, 603–616. (2002)

[74] K. Doya. Metalearning and neuromodulation. *Neural Networks* **15**, no. 4, 495–506. (2002)

[75] Y.-L. Boureau and P. Dayan. Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology* **36**, no. 1, 74–97. (2011)

[76] P. Dayan and Q. Huys. Serotonin's many meanings elude simple theories. *eLife* **4**, e07390. (2015)

[77] T. Hastie, R. Tibshirani, and J. Friedman. Chapter 1: Introduction in *The Elements of Statistical Learning*, 2nd ed., pp. 1–8. (Springer, 2013)

[78] V. Mnih, K. Kavukcuoglu, D. Silver, *et al.* Human-level control through deep reinforcement learning. *Nature* **518**, no. 7540, 529–533. (2015)

[79] D. Silver, A. Huang, C. J. Maddison, *et al.* Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, no. 7587, 484–489. (2016)

[80] M. Jaderberg, W. M. Czarnecki, I. Dunning, *et al.* Human-level performance in 3D multiplayer games with population-based reinforcement learning. *Science* **364**, no. 6443, 859–865. (2019)

[81] J. Perolat, B. De Vylder, D. Hennes, *et al.* Mastering the game of Stratego with model-free multiagent reinforcement learning. *Science* **378**, no. 6623, 990–996. (2022)

[82] D. Silver, J. Schrittwieser, K. Simonyan, *et al.* Mastering the game of Go without human knowledge. *Nature* **550**, no. 7676, 354–359. (2017)

[83] J. Schrittwieser, I. Antonoglou, T. Hubert, *et al.* Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature* **588**, no. 7839, 604–609. (2020)

[84] O. Vinyals, I. Babuschkin, W. M. Czarnecki, *et al.* Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* **575**, no. 7782, 350–354. (2019)

[85] E. Nikonova and J. Gemrot. Deep Q-Network for Angry Birds. *arXiv.* (2019) doi: `https://doi.org/10.48550/arXiv.1910.01806`

[86] R. S. Sutton. The bitter lesson. *Incomplete Ideas.* Web. (2019) `http://www.incompleteideas.net/IncIdeas/BitterLesson.html`

[87] K. Iigaya, M. S. Fonseca, M. Murakami, Z. F. Mainen, and P. Dayan. An effect of serotonergic stimulation on learning rates for rewards apparent after long intertrial intervals. *Nature Communications* **9**, no. 1, 2477. (2018)

[88] P. Samuelson. Some aspects of the pure theory of capital. *The Quarterly Journal of Economics* **51**, no. 3, 469–496. (1937)

[89] W. Fedus, C. Gelada, Y. Bengio, M. G. Bellemare, and H. Larochelle. Hyperbolic discounting and learning over multiple horizons. *arXiv.* (2019) doi: `https://doi.org/10.48550/arXiv.1902.06865`

[90] R. E. Bellman and S. E. Dreyfus. The policy improvement routine in *Applied Dynamic Programming*, pp. 304. (Princeton University Press, 1962)

[91] R. S. Sutton and A. Barto. "Policy improvement" in Chap. 4: Dynamic programming of *Reinforcement learning: An introduction*, 2nd ed., pp. 76–80. (The MIT Press, 2018)

[92] C. Watkins. *Learning from Delayed Rewards.* Ph.D. thesis. (King's College, 1989)

[93] R. S. Sutton, D. A. McAllester, S. P. Singh, and Y. Mansour. Policy gradient methods for reinforcement learning with function approximation. *Advances in Neural Information Processing Systems* **12**. (1999)

[94] R. J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning* **8**, 229–256. (1992)

[95] K. Doya. Temporal difference learning in continuous time and space. *Neural Information Processing Systems* **8**, 1073–1079. (1996)

[96] H. van Seijen, A. R. Mahmood, P. M. Pilarski, M. C. Machado, and R. S. Sutton. True online temporal-difference learning. *Journal of Machine Learning Research* **17**, 1–40. (2016)

[97] R. E. Bellman and S. E. Dreyfus. *Applied Dynamic Programming.* (Princeton University Press, 1962)

[98] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski. "The problem of neural coding" in Chap. 7: Variability of spike trains and neural codes of *Neuronal Dynamics*, 1st ed., pp. 190–198. (Cambridge University Press, 2014)

[99] U. Lindblom. The relation between stimulus and discharge in a rapidly adapting touch receptor. *Acta Physiologica Scandinavica* **56**, no. 3-4, 349–361. (1962)

[100] A. L. Eastwood, A. Sanzeni, B. C. Petzold, S.-J. Park, M. Vergassola, B. L. Pruitt, and M. B. Goodman. Tissue mechanics govern the rapidly adapting and symmetrical response to touch. *Proceedings of the National Academy of Sciences* **112**, no. 50. (2015) doi: `https://doi.org/10.1073/pnas.1514138112`

[101] G. Schwartz, R. Harrix, D. Shrom, and M. J. Berry. Detection and prediction of periodic patterns by the retina. *Nature Neuroscience* **10**, 552–554. (2007)

[102] M. Garrett, S. Manavi, K. Roll, *et al.* Experience shapes activity dynamics and stimulus coding of VIP inhibitory cells. *eLife* **9**, e50340. (2020)

[103] L. N. Driscoll, N. L. Pettit, M. Minderer, S. N. Chettih, and C. D. Harvey. Dynamic reorganization of neuronal activity patterns in parietal cortex. *Cell* **170**, no. 5, 986–999.e16. (2017)

[104] M. C. Aoi, V. Mante, and J. W. Pillow. Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nature Neuroscience* **23**, no. 11, 1410–1420. (2020)

[105] C. Lee, E. F. Harkin, X. Yin, R. Naud, and S. Chen. Cell-type-specific responses to associative learning in the primary motor cortex. *eLife* **11**, e72549. (2022)

[106] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski. "Spike trains" in Chap. 1: Elements of neuronal systems of *Neuronal Dynamics*, pp. 5. (Cambridge University Press, 2014)

[107] W. Gerstner and J. L. van Hemmen. Universality in neural networks: the importance of the 'mean firing rate'. *Biological Cybernetics* **67**, no. 3, 195–205. (1992)

[108] E. D. Adrian. The impulses produced by sensory nerve endings: Part I. *The Journal of Physiology* **61**, no. 1, 49–72. (1926)

[109] M. C. Liberman. Single-neuron labeling in the cat auditory nerve. *Science* **216**, no. 4551, 1239–1241. (1982)

[110] D. W. Moran and A. B. Schwartz. Motor cortical representation of speed and direction during reaching. *Journal of Neurophysiology* **82**, no. 5, 2676–2692. (1999)

[111] N. Brenner, W. Bialek, and R. De Ruyter Van Steveninck. Adaptive rescaling maximizes information transmission. *Neuron* **26**, no. 3, 695–702. (2000)

[112] S. Panzeri and S. R. Schultz. A unified approach to the study of temporal, correlational, and rate coding. *Neural Computation* **13**, no. 6, 1311–1349. (2001)

[113] M. St-Hilaire and A. Longtin. Comparison of coding capabilities of type I and type II neurons. *Journal of Computational Neuroscience* **16**, 299–313. (2004)

[114] Z. Mainen and T. Sejnowski. Reliability of spike timing in neocortical neurons. *Science* **268**, no. 5216, 1503–1506. (1995)

[115] S. Deneve, P. E. Latham, and A. Pouget. Reading population codes: a neural implementation of ideal observers. *Nature Neuroscience* **2**, no. 8, 740–745. (1999)

[116] R. Tomar, C. E. Smith, and P. Lansky. A simple neuronal model with intrinsic saturation of the firing frequency. *Biosystems* **222**, 104780. (2022)

[117] A. L. Hodgkin. The local electric changes associated with repetitive action in a non-medullated axon. *The Journal of Physiology* **107**, no. 2, 165–181. (1948)

[118] E. Salinas and L. F. Abbott. Vector reconstruction from firing rates. *Journal of Computational Neuroscience* **1**, no. 1-2, 89–107. (1994)

[119] A. P. Nikitin, N. G. Stocks, R. P. Morse, and M. D. McDonnell. Neural population coding is optimized by discrete tuning curves. *Physical Review Letters* **103**, no. 13, 138101. (2009)

[120] K. A. Ferguson and J. A. Cardin. Mechanisms underlying gain modulation in the cortex. *Nature Reviews Neuroscience* **21**, no. 2, 80–92. (2020)

[121] M. Carandini and D. J. Heeger. Normalization as a canonical neural computation. *Nature Reviews Neuroscience* **13**, no. 1, 51–62. (2012)

[122] M. Chalk, O. Marre, and G. Tkačik. Toward a unified theory of efficient, predictive, and sparse coding. *Proceedings of the National Academy of Sciences* **115**, no. 1, 186–191. (2018)

[123] M. Spratling. A review of predictive coding algorithms. *Brain and Cognition* **112**, 92–97. (2017)

[124] R. Naud and W. Gerstner. Coding and decoding with adapting neurons: A population approach to the peri-stimulus time histogram. *PLoS Computational Biology* **8**, no. 10, e1002711. (2012)

[125] J. Benda and A. V. Herz. A universal model for spike-frequency adaptation. *Neural Computation* **15**, no. 11, 2523–2564. (2003)

[126] B. N. Lundstrom, M. H. Higgs, W. J. Spain, and A. L. Fairhall. Fractional differentiation by neocortical pyramidal neurons. *Nature Neuroscience* **11**, no. 11, 1335–1342. (2008)

[127] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology* **117**, no. 4, 500–544. (1952)

[128] Z. F. Mainen, J. Joerges, J. R. Huguenard, and T. J. Sejnowski. A model of spike initiation in neocortical pyramidal neurons. *Neuron* **15**, no. 6, 1427–1439. (1995)

[129] Z. F. Mainen and T. J. Sejnowski. Influence of dendritic structure on firing pattern in model neocortical neurons. *Nature* **382**, no. 6589, 363–366. (1996)

[130] P. Poirazi, T. Brannon, and B. W. Mel. Arithmetic of subthreshold synaptic summation in a model CA1 pyramidal cell. *Neuron* **37**, no. 6, 977–987. (2003)

[131] P. Poirazi, T. Brannon, and B. W. Mel. Pyramidal neuron as two-layer neural network. *Neuron* **37**, no. 6, 989–999. (2003)

[132] J. Golowasch, M. S. Goldman, L. F. Abbott, and E. Marder. Failure of averaging in the construction of a conductance-based neuron model. *Journal of Neurophysiology* **87**, no. 2, 1129–1131. (2002)

[133] A. Balachandar and S. A. Prescott. Origin of heterogeneous spiking patterns from continuously distributed ion channel densities: a computational study in spinal dorsal horn neurons: Heterogeneous spiking patterns in spinal dorsal horn neurons. *The Journal of Physiology* **596**, no. 9, 1681–1697. (2018)

[134] Y. N. Billeh, B. Cai, S. L. Gratiy, *et al.* Systematic integration of structural and functional data into multi-scale models of mouse primary visual cortex. *Neuron* **106**, no. 3, 388–403.e18. (2020)

[135] P. Gonçalves, J.-M. Lueckmann, M. Deistler, *et al.* Training deep neural density estimators to identify mechanistic models of neural dynamics. *eLife* **9**, e56261. (2020)

[136] L. Paninski, E. P. Simoncelli, and J. W. Pillow. Maximum likelihood estimation of a stochastic integrate-and-fire neural model. *Advances in Neural Information Processing Systems* **16**, 1311–1318. (2003)

[137] B. M. Yu, J. P. Cunningham, G. Santhanam, S. I. Ryu, K. V. Shenoy, and M. Sahani. Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *Journal of Neurophysiology* **102**, no. 1, 614–635. (2009)

[138] Q. J. M. Huys, M. B. Ahrens, and L. Paninski. Efficient estimation of detailed single-neuron models. *Journal of Neurophysiology* **96**, no. 2, 872–890. (2006)

[139] W. Truccolo, U. T. Eden, M. R. Fellows, J. P. Donoghue, and E. N. Brown. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *Journal of Neurophysiology* **93**, no. 2, 1074–1089. (2005)

[140] J. W. Pillow, L. Paninski, V. J. Uzzell, E. P. Simoncelli, and E. Chichilnisky. Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *Journal of Neuroscience* **25**, no. 47, 11003–11013. (2005)

[141] J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. J. Chichilnisky, and E. P. Simoncelli. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* **454**, no. 7207, 995–999. (2008)

[142] S. L. Keeley, D. M. Zoltowski, M. C. Aoi, and J. W. Pillow. Modeling statistical dependencies in multi-region spike train data. *Current Opinion in Neurobiology* **65**, 194–202. (2020)

[143] E. Izhikevich. Simple model of spiking neurons. *IEEE Transactions on Neural Networks* **14**, no. 6, 1569–1572. (2003)

[144] L. Lapicque. Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation. *Journal de Physiologie et de Pathologie Générale* **9**, 620–635. (1907)

[145] R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal* **1**, no. 6, 445–466. (1961)

[146] J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE* **50**, no. 10, 2061–2070. (1962)

[147] R. B. Stein. A theoretical analysis of neuronal variability. *Biophysical Journal* **5**, no. 2, 173–194. (1965)

[148] L. F. Abbott and T. B. Kepler. Model neurons: From Hodgkin-Huxley to hopfield. *Statistical Mechanics of Neural Networks* **368**. (1990) doi: `https://doi.org/10.1007/3540532676_37`

[149] H. C. Tuckwell, Y. Zhou, and N. J. Penington. Simplified models of pacemaker spiking in raphe and locus coeruleus neurons. *arXiv*. (2015) doi: `https://doi.org/10.48550/arXiv.1508.05468`

[150] Ş. Mihalaş and E. Niebur. A generalized linear integrate-and-fire neural model produces diverse spiking behaviors. *Neural Computation* **21**, no. 3, 704–718. (2009)

[151] S. Mensi, R. Naud, C. Pozzorini, M. Avermann, C. C. H. Petersen, and W. Gerstner. Parameter extraction and classification of three cortical neuron types reveals two distinct adaptation mechanisms. *Journal of Neurophysiology* **107**, no. 6, 1756–1775. (2012)

[152] C. Teeter, R. Iyer, V. Menon, *et al.* Generalized leaky integrate-and-fire models classify multiple neuron types. *Nature Communications* **9**, no. 1, 709. (2018)

[153] E. F. Harkin. *A simplified serotonin neuron model*. M.Sc. thesis. (University of Ottawa, 2018)

[154] K. Wong-Lin, D.-H. Wang, A. A. Moustafa, J. Y. Cohen, and K. Nakamura. Toward a multiscale modeling framework for understanding serotonergic function. *Journal of Psychopharmacology* **31**, no. 9, 1121–1136. (2017)

[155] H. C. Tuckwell and N. J. Penington. Computational modeling of spike generation in serotonergic neurons of the dorsal raphe nucleus. *Progress in Neurobiology* **118**, 59–101. (2014)

[156] K. Wong-Lin, G. Prasad, and T. M. McGinnity. A spiking neuronal network model of the dorsal raphe nucleus. *The 2011 International Joint Conference on Neural Networks*, 1591–1598. (2011)

[157] J. Jalewa, A. Joshi, T. M. McGinnity, G. Prasad, K. Wong-Lin, and C. Hölscher. Neural circuit interactions between the dorsal raphe nucleus and the lateral hypothalamus: An experimental and computational study. *PLoS ONE*. (2014) doi: `https://doi.org/10.1371/journal.pone.0088003`

[158] E. F. Harkin, M. B. Lynn, A. Payeur, J.-F. Boucher, L. Caya-Bissonnette, D. Cyr, C. Stewart, A. Longtin, R. Naud, and J.-C. Béïque. Temporal derivative computation in the dorsal raphe network revealed by an experimentally-driven augmented integrate-and-fire modeling framework. *eLife* **12**, e72951. (2023)

[159] E. F. Harkin, P. R. Shen, A. Goel, B. A. Richards, and R. Naud. Parallel and recurrent cascade models as a unifying force for understanding subcellular computation. *Neuroscience* **489**, 200–215. (2022)

[160] E. F. Harkin, C. D. Grossman, J. Y. Cohen, J.-C. Béïque, and R. Naud. Serotonin predictively encodes value. *bioRxiv*. (2023)

[161] E. F. Harkin, J.-C. Béïque, and R. Naud. A user's guide to generalized integrate-and-fire models in *Computational Modelling of the Brain: Modelling Approaches to Cells, Circuits and Networks*, pp. 69–86. M. Giugliano, M. Negrello, and D. Linaro, eds. (Springer, 2021)

[162] E. F. Harkin, A. Payeur, M. B. Lynn, J.-F. Boucher, L. Caya-Bissonnette, D. Cyr, C. Stewart, A. Longtin, R. Naud, and J.-C. Béïque. Multiple mechanisms of gain modulation in the serotonin system. Talk presented at Neuromatch. (Online, 2020)

[163] E. F. Harkin, A. Payeur, M. B. Lynn, J.-F. Boucher, L. Caya-Bissonnette, D. Cyr, C. Stewart, A. Longtin, R. Naud, and J.-C. Béïque. Multiple mechanisms of gain modulation in the serotonin system. Talk presented at Canadian Computational Neuroscience Spotlight. (Online, 2020)

[164] E. F. Harkin, A. Payeur, M. B. Lynn, J.-F. Boucher, L. Caya-Bissonnette, D. Cyr, C. Stewart, A. Longtin, R. Naud, and J.-C. Béïque. From single neurons to networks in the dorsal raphe nucleus. Poster presented at COSYNE. (Online, 2021)

[165] E. F. Harkin, A. Payeur, M. B. Lynn, J.-F. Boucher, L. Caya-Bissonnette, D. Cyr, C. Stewart, A. Longtin, J.-C. Béïque, and R. Naud. Temporal derivative computation in the dorsal raphe network revealed by an experimentally-driven augmented integrate-and-fire modeling framework. Poster presented at Bernstein Conference. (Online, 2021)

[166] E. F. Harkin, A. Payeur, M. B. Lynn, J.-F. Boucher, L. Caya-Bissonnette, D. Cyr, C. Stewart, A. Longtin, R. Naud, and J.-C. Béïque. Temporal derivative computation in the dorsal raphe network revealed by an experimentally-driven augmented integrate-and-fire modeling framework. Poster presented at Neuroscience. (San Diego, 2022)

[167] E. F. Harkin, A. Payeur, M. B. Lynn, J.-F. Boucher, A. Longtin, R. Naud, and J.-C. Béïque. Multiple mechanisms of gain modulation in the serotonin system. Poster presented at Canadian Association for Neuroscience annual meeting. (Montréal, 2020) Abstract accepted but conference cancelled due to COVID.

[168] S. N. Young, S. E. Smith, R. O. Pihl, and F. R. Ervin. Tryptophan depletion causes a rapid lowering of mood in normal males. *Psychopharmacology* **87**, 173–177. (1985)

[169] P. L. Delgado, L. H. Price, H. L. Miller, R. M. Salomon, G. K. Aghajanian, G. R. Heninger, and D. S. Charney. Serotonin and the neurobiology of depression: effects of tryptophan depletion in drug-free depressed patients. *Archives of General Psychiatry* **51**, no. 11, 865–874. (1994)

[170] S. F. Maier and L. R. Watkins. Stressor controllability and learned helplessness: the roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. *Neuroscience & Biobehavioral Reviews* **29**, no. 4-5, 829–841. (2005)

[171] K. Nakamura, M. Matsumoto, and O. Hikosaka. Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus. *Journal of Neuroscience* **28**, no. 20, 5331–5343. (2008)

[172] M. Tops, S. Russo, M. A. Boksem, and D. M. Tucker. Serotonin: modulator of a drive to withdraw. *Brain and Cognition* **71**, no. 3, 427–436. (2009)

[173] R. Cools, K. Nakamura, and N. D. Daw. Serotonin and dopamine: unifying affective, activational, and decision functions. *Neuropsychopharmacology* **36**, no. 1, 98–113. (2011)

[174] Y. Li, W. Zhong, D. Wang, *et al.* Serotonin neurons in the dorsal raphe nucleus encode reward signals. *Nature Communications* **7**, no. 1, 10503. (2016)

[175] J. Deakin. The role of serotonin in panic, anxiety and depression. *International Clinical Psychopharmacology* **13**. (1998) doi: `https://doi.org/10.1097/00004850-199804004-00001`

[176] P. Dayan and Q. J. Huys. Serotonin in affective control. *Annual Review of Neuroscience* **32**, 95–126. (2009)

[177] J. Savitz, I. Lucki, and W. C. Drevets. 5-HT_{1A} receptor function in major depressive disorder. *Progress in Neurobiology* **88**, no. 1, 17–31. (2009)

[178] M. Fava and K. S. Kendler. Major depressive disorder. *Neuron* **28**, no. 2, 335–341. (2000)

[179] Z. R. Donaldson, K. M. Nautiyal, S. E. Ahmari, and R. Hen. Genetic approaches for understanding the role of serotonin receptors in mood and behavior. *Current Opinion in Neurobiology* **23**, no. 3, 399–406. (2013)

[180] A. Muzerelle, S. Scotto-Lomassese, J. F. Bernard, M. Soiza-Reilly, and P. Gaspar. Conditional anterograde tracing reveals distinct targeting of individual serotonin cell groups (B5–B9) to the forebrain and brainstem. *Brain Structure and Function* **221**, 535–561. (2016)

[181] J. Schweimer and M. Ungless. Phasic responses in dorsal raphe serotonin neurons to noxious stimuli. *Neuroscience* **171**, no. 4, 1209–1215. (2010)

[182] B. W. Knight. Dynamics of encoding in a population of neurons. *Journal of General Physiology* **59**, no. 6, 734–766. (1972)

[183] W. Gerstner. Population dynamics of spiking neurons: fast transients, asynchronous states, and locking. *Neural Computation* **12**, no. 1, 43–89. (2000)

[184] J.-C. Béïque, B. Campbell, P. Perring, M. W. Hamblin, P. Walker, L. Mladenovic, and R. Andrade. Serotonergic regulation of membrane potential in developing rat prefrontal cortex: coordinated expression of 5-hydroxytryptamine (5-HT) 1A, 5-HT_{2A}, and 5-HT_7 receptors. *Journal of Neuroscience* **24**, no. 20, 4807–4817. (2004)

[185] J.-C. Béïque, M. Imad, L. Mladenovic, J. Gingrich, and R. Andrade. Mechanism of the 5-hydroxytryptamine 2A receptor-mediated facilitation of synaptic activity in prefrontal cortex. *Proceedings of the National Academy of Sciences* **104**, no. 23, 9870–9875. (2007)

[186] R. Andrade. Serotonergic regulation of neuronal excitability in the prefrontal cortex. *Neuropharmacology* **61**, no. 3, 382–386. (2011)

[187] V. Varga, A. Losonczy, B. Zemelman, Z. Borhegyi, G. Nyiri, A. Domonkos, B. Hangya, N. Holderith, J. Magee, and T. Freund. Fast synaptic subcortical control of hippocampal circuits. *Science* **326**, no. 5951, 449–453. (2009)

[188] R. Amo, F. Fredes, M. Kinoshita, *et al.* The habenulo-raphe serotonergic circuit encodes an aversive expectation value essential for adaptive active avoidance of danger. *Neuron* **84**, no. 5, 1034–1048. (2014)

[189] M. Matsumoto and O. Hikosaka. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447**, no. 7148, 1111–1115. (2007)

[190] A. A. Prinz, D. Bucher, and E. Marder. Similar network activity from disparate circuit parameters. *Nature Neuroscience* **7**, no. 12, 1345–1352. (2004)

[191] W. Gerstner and R. Naud. How good are neuron models? *Science* **326**, no. 5951, 379–380. (2009)

[192] C. Pozzorini, R. Naud, S. Mensi, and W. Gerstner. Temporal whitening by power-law adaptation in neocortical neurons. *Nature Neuroscience* **16**, no. 7, 942–948. (2013)

[193] J. A. Connor and C. F. Stevens. Voltage clamp studies of a transient outward membrane current in gastropod neural somata. *The Journal of Physiology* **213**, no. 1, 21–30. (1971)

[194] J. Connor, D. Walter, and R. McKown. Neural repetitive firing: modifications of the Hodgkin–Huxley axon suggested by experimental results from crustacean axons. *Biophysical Journal* **18**, no. 1, 81–102. (1977)

[195] G. Drion, T. O'Leary, and E. Marder. Ion channel degeneracy enables robust and tunable neuronal firing rates. *Proceedings of the National Academy of Sciences* **112**, no. 38. (2015) doi: `https://doi.org/10.1073/pnas.1516400112`

[196] J. F. Storm. An after-hyperpolarization of medium duration in rat hippocampal pyramidal cells. *Journal of Physiology* **409**, 171–190. (1989)

[197] G. Aghajanian. Modulation of a transient outward current in serotonergic neurones by α-1-adrenoceptors. *Nature* **315**, 501–503. (1985)

[198] P. A. Getting. Mechanisms of pattern generation underlying swimming in *Tritonia*. III. Intrinsic and synaptic mechanisms for delayed excitation. *Journal of Neurophysiology* **49**, no. 4, 1036–1050. (1983)

[199] M. Segal. A potent transient outward current regulates excitability of dorsal raphe neurons. *Brain Research* **359**, no. 1-2, 347–350. (1985)

[200] D. A. McCormick. Functional properties of a slowly inactivating potassium current in guinea pig dorsal lateral geniculate relay neurons. *Journal of Neurophysiology* **66**, no. 4, 1176–1189. (1991)

[201] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski. *Neuronal Dynamics*, 1st ed. (Cambridge University Press, 2014)

[202] L. Paninski, J. Pillow, and E. Simoncelli. Comparing integrate-and-fire models estimated using intracellular and extracellular data. *Neurocomputing* **65-66**, 379–385. (2005)

[203] C. Pozzorini, S. Mensi, O. Hagens, R. Naud, C. Koch, and W. Gerstner. Automated high-throughput characterization of single neurons by means of simplified spiking models. *PLoS Computational Biology* **11**, no. 6, e1004275. (2015)

[204] S. Mensi, O. Hagens, W. Gerstner, and C. Pozzorini. Enhanced sensitivity to rapid input fluctuations by nonlinear threshold dynamics in neocortical pyramidal neurons. *PLoS Computational Biology* **12**, no. 2, e1004761. (2016)

[205] Y. Dong and F. J. White. Dopamine D1-class receptors selectively modulate a slowly inactivating potassium current in rat medial prefrontal cortex pyramidal neurons. *The Journal of Neuroscience* **23**, no. 7, 2686–2695. (2003)

[206] Y. Dong, F. Nasif, J. Tsui, W. Ju, D. Cooper, X.-T. Hu, R. Malenka, and F. White. Cocaine-induced plasticity of intrinsic membrane properties in prefrontal cortex pyramidal neurons: adaptations in potassium currents. *Journal of Neuroscience* **25**, no. 4, 936–940. (2005)

[207] J. F. Mejias and A. Longtin. Optimal heterogeneity for coding in spiking neural networks. *Physical Review Letters* **108**, no. 22, 228102. (2012)

[208] S. J. Tripathy, K. Padmanabhan, R. C. Gerkin, and N. N. Urban. Intermediate intrinsic diversity enhances neural population coding. *Proceedings of the National Academy of Sciences* **110**, no. 20, 8248–8253. (2013)

[209] B. Ermentrout. Linearization of F-I curves by adaptation. *Neural Computation* **10**, no. 7, 1721–1729. (1998)

[210] H. Markram. The Blue Brain Project. *Nature Reviews Neuroscience* **7**, no. 2, 153–160. (2006)

[211] J. W. Pillow and E. P. Simoncelli. Dimensionality reduction in neural models: An information-theoretic generalization of spike-triggered average and covariance analysis. *Journal of Vision* **6**, no. 4, 9. (2006)

[212] D. A. Hoffman, J. C. Magee, C. M. Colbert, and D. Johnston. K^+ channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons. *Nature* **387**, no. 6636, 869–875. (1997)

[213] M. Harnett, N.-L. Xu, J. Magee, and S. Williams. Potassium channels control the interaction between active dendritic integration compartments in layer 5 cortical pyramidal neurons. *Neuron* **79**, no. 3, 516–529. (2013)

[214] B. B. Ujfalussy, J. K. Makara, M. Lengyel, and T. Branco. Global and multiplexed dendritic computations under *in vivo*-like conditions. *Neuron* **100**, no. 3, 579–592.e5. (2018)

[215] A. Payeur, J.-C. Béïque, and R. Naud. Classes of dendritic information processing. *Current Opinion in Neurobiology* **58**, 78–85. (2019)

[216] J. F. Mejias and A. Longtin. Differential effects of excitatory and inhibitory heterogeneity on the gain and asynchronous state of sparse cortical networks. *Frontiers in Computational Neuroscience* **8**. (2014)

[217] J. F. Mejias, A. Payeur, E. Selin, L. Maler, and L. A. Subtractive, divisive and non-monotonic gain control in feedforward nets linearized by noise and delays. *Frontiers in Computational Neuroscience* **8**. (2014)

[218] F. S. Chance and L. F. Abbott. Divisive inhibition in recurrent networks. *Network* **11**, 119–129. (2000)

[219] S. Ogawa, J. Cohen, D. Hwang, N. Uchida, and M. Watabe-Uchida. Organization of monosynaptic inputs to the serotonin and dopamine neuromodulatory systems. *Cell Reports* **8**, no. 4, 1105–1118. (2014)

[220] E. Marder. Neuromodulation of neuronal circuits: back to the future. *Neuron* **76**, no. 1, 1–11. (2012)

[221] B. Tsuda, S. C. Pate, K. M. Tye, H. T. Siegelmann, and T. J. Sejnowski. Neuromodulators enable overlapping synaptic memory regimes and nonlinear transition dynamics in recurrent neural networks. *bioRxiv*. (2021)

[222] J. M. Baraban and G. Aghajanian. Noradrenergic innervation of serotonergic neurons in the dorsal raphe: demonstration by electron microscopic autoradiography. *Brain Research* **204**, no. 1, 1–11. (1981)

[223] T. K. Aman, R.-Y. Shen, and S. Haj-Dahmane. D2-like dopamine receptors depolarize dorsal raphe serotonin neurons through the activation of nonselective cationic conductance. *Journal of Pharmacology and Experimental Therapeutics* **320**, no. 1, 376–385. (2007)

[224] N. Schweighofer, M. Bertin, K. Shishida, Y. Okamoto, S. C. Tanaka, S. Yamawaki, and K. Doya. Low-serotonin levels increase delayed reward discounting in humans. *Journal of Neuroscience* **28**, no. 17, 4528–4532. (2008)

[225] A. A. Grace and S.-P. Onn. Morphology and electrophysiological properties of lmmunocytochemically identified rat dopamine neurons recorded *in vitro*. *Journal of Neuroscience* **9**, no. 10, 3463–3481. (1989)

[226] Z. M. Khaliq and B. P. Bean. Dynamic, nonlinear feedback regulation of slow pacemaking by A-type potassium current in ventral tegmental area neurons. *Journal of Neuroscience* **28**, no. 43, 10905–10917. (2008)

[227] H. R. Kim, A. N. Malik, J. G. Mikhael, *et al.* A unified framework for dopamine signals across timescales. *Cell* **183**, no. 6, 1600–1616.e25. (2020)

[228] H. Barlow. The exploitation of regularities in the environment by the brain. *Behavioral and Brain Sciences* **24**, no. 4, 602–607. (2001)

[229] N. Ulanovsky, L. Las, and I. Nelken. Processing of low-probability sounds by cortical neurons. *Nature Neuroscience* **6**, no. 4, 391–398. (2003)

[230] A. Kohn. Visual adaptation: physiology, mechanisms, and functional benefits. *Journal of Neurophysiology* **97**, no. 5, 3155–3164. (2007)

[231] R. Naud, F. Gerhard, S. Mensi, and W. Gerstner. Improved similarity measures for small sets of spike trains. *Neural Computation* **23**, no. 12, 3016–3069. (2011)

[232] R. Jolivet, R. Kobayashi, A. Rauch, R. Naud, S. Shinomoto, and W. Gerstner. A benchmark test for a quantitative assessment of simple neuron models. *Journal of Neuroscience Methods* **169**, no. 2, 417–424. (2008)

[233] C. D. Grossman and J. Y. Cohen. Neuromodulation and neurophysiology on the timescale of learning and decision-making. *Annual Review of Neuroscience* **45**, 317–337. (2022)

[234] D. Asher, A. Craig, A. Zaldivar, A. Brewer, and J. Krichmar. A dynamic, embodied paradigm to investigate the role of serotonin in decision-making. *Frontiers in Integrative Neuroscience* **7**. (2013)

[235] J. Cohen, C. Grossman, and B. Bari. Serotonin neurons modulate learning rate through uncertainty. Dryad dataset. (2021) doi: `https://doi.org/10.5061/dryad.cz8w9gj4s`

[236] C. Watkins. Chapter 5: Modes of control of behaviour in *Learning from Delayed Rewards*, pp. 55–71. Ph.D. thesis. (King's College, 1989)

[237] C. Watkins and P. Dayan. Q-Learning. *Machine Learning* **8**, 279–292. (1992)

[238] R. S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning* **3**, no. 1, 9–44. (1988)

[239] C. Watkins. Chapter 7: Primitive learning in *Learning from Delayed Rewards*, pp. 81–113. Ph.D. thesis. (King's College, 1989)

[240] O. Selfridge. Some themes and primitives in ill-defined systems in *Adaptive Control of Ill-Defined Systems*, pp. 21–26. (1984)

[241] K. Miyazaki, K. W. Miyazaki, and K. Doya. Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. *Journal of Neuroscience* **31**, no. 2, 469–479. (2011)

[242] J. M. Pearce and G. Hall. A model for pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review* **87**, no. 6, 532–552. (1980)

[243] Z. Liu, J. Zhou, Y. Li, *et al.* Dorsal raphe neurons signal reward through 5-HT and glutamate. *Neuron* **81**, no. 6, 1360–1374. (2014)

[244] L. Elber-Dorozko and Y. Loewenstein. Striatal action-value neurons reconsidered. *eLife* **7**, e34248. (2018)

[245] L. V. Hedges. Estimation of effect size under nonrandom sampling: The effects of censoring studies yielding statistically insignificant mean differences. *Journal of Educational Statistics* **9**, no. 1, 61–85. (1984)

[246] K. Hayashi, K. Nakao, and K. Nakamura. Appetitive and aversive information coding in the primate dorsal raphé nucleus. *The Journal of Neuroscience* **35**, no. 15, 6195–6208. (2015)

[247] E. L. Sylwestrak, Y. Jo, S. Vesuna, *et al.* Cell-type-specific population dynamics of diverse reward computations. *Cell* **185**, no. 19, 3568–3587. (2022)

[248] C. Stringer, M. Michaelos, D. Tsyboulski, S. E. Lindo, and M. Pachitariu. High-precision coding in visual cortex. *Cell* **184**, no. 10, 2767–2778.e15. (2021)

[249] W. Dabney, Z. Kurth-Nelson, N. Uchida, C. K. Starkweather, D. Hassabis, R. Munos, and M. Botvinick. A distributional code for value in dopamine-based reinforcement learning. *Nature* **577**, no. 7792, 671–675. (2020)

[250] Y. Dan, J. J. Atick, and R. C. Reid. Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. *The Journal of Neuroscience* **16**, no. 10, 3351–3362. (1996)

[251] K. Beier, E. Steinberg, K. DeLoach, S. Xie, K. Miyamichi, L. Schwarz, X. Gao, E. Kremer, R. Malenka, and L. Luo. Circuit architecture of VTA dopamine neurons revealed by systematic input-output mapping. *Cell* **162**, no. 3, 622–634. (2015)

[252] J. N. McNair. Optimal giving-up times and the marginal value theorem. *The American Naturalist* **119**, no. 4, 511–529. (1982)

[253] R. S. Sutton, D. Precup, and S. Singh. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence* **112**, no. 1-2, 181–211. (1999)

[254] J. Qi, S. Zhang, H.-L. Wang, H. Wang, J. De Jesus Aceves Buendia, A. F. Hoffman, C. R. Lupica, R. P. Seal, and M. Morales. A glutamatergic reward input from the dorsal raphe to ventral tegmental area dopamine neurons. *Nature Communications* **5**, no. 1, 5390. (2014)

[255] J. C. Marques, M. Li, D. Schaak, D. N. Robson, and J. M. Li. Internal state dynamics shape brainwide activity and foraging behaviour. *Nature* **577**, no. 7789, 239–243. (2020)

[256] T. E J. Behrens, M. W. Woolrich, M. E. Walton, and M. F S. Rushworth. Learning the value of information in an uncertain world. *Nature Neuroscience* **10**, no. 9, 1214–1221. (2007)

[257] P. Piray and N. D. Daw. A model for learning based on the joint estimation of stochasticity and volatility. *Nature Communications* **12**, no. 1, 6587. (2021)

[258] B. A. Bari, C. D. Grossman, E. E. Lubin, A. E. Rajagopalan, J. I. Cressy, and J. Y. Cohen. Stable representations of decision variables for flexible behavior. *Neuron* **103**, no. 5, 922–933.e7. (2019)

[259] G. D. Stuber and R. A. Wise. Lateral hypothalamic circuits for feeding and reward. *Nature Neuroscience* **19**, no. 2, 198–205. (2016)

[260] D. Kim, H. H. Schütt, and W. J. Ma. Reward prediction error neurons implement an efficient code for reward. *bioRxiv*. (2022) doi: `https://doi.org/10.1101/2022.11.03.515104`

[261] M. Sousa, P. Bujalski, B. Cruz, K. Louie, D. McNamee, and J. Paton. Dopamine neurons encode a multidimensional probabilistic map of future reward. *bioRxiv*. (2023) doi: https://doi.org/10.1101/2023.11.12.566727

[262] P. Masset, P. Tano, H. R. Kim, A. N. Malik, A. Pouget, and N. Uchida. Multi-timescale reinforcement learning in the brain. *bioRxiv*. (2023) doi: https://doi.org/10.1101/2023.11.12.566754

[263] H. van Seijen and R. Sutton. True online TD(λ). *International Conference on Machine Learning* **32**, no. 1, 692–700. (2014)

[264] T. C. Hesterberg. What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum. *The American Statistician* **69**, no. 4, 371–386. (2015)

[265] D. G. Kelty-Stephen, P. E. Cisek, B. D. Bari, L. H. Favela, F. Hasselman, F. Keijzer, J. B. Wagman, and B. J. Thomas. In search for an alternative to the computer metaphor of the mind and brain. *arXiv*. (2022) doi: https://doi.org/10.48550/arXiv.2206.04603

[266] D. L K. Yamins and J. J. DiCarlo. Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience* **19**, no. 3, 356–365. (2016)

[267] A. M. Zador. A critique of pure learning and what artificial neural networks can learn from animal brains. *Nature Communications* **10**, no. 1, 3770. (2019)

[268] B. A. Richards, T. P. Lillicrap, P. Beaudoin, *et al*. A deep learning framework for neuroscience. *Nature Neuroscience* **22**, no. 11, 1761–1770. (2019)

[269] A. Saxe, S. Nelli, and C. Summerfield. If deep learning is the answer, what is the question? *Nature Reviews Neuroscience* **22**, no. 1, 55–67. (2021)

[270] D. Haufler, S. Ito, C. Koch, and A. Arkhipov. Simulations of cortical networks using spatially extended conductance-based neuronal models. *The Journal of Physiology* **601**, no. 15, 3123–3139. (2023)

[271] A. Destexhe, Z. F. Mainen, and T. J. Sejnowski. Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. *Journal of Computational Neuroscience* **1**, no. 3, 195–230. (1994)

[272] T. Poggio. Afterword: Marr's vision and computational neuroscience in *Vision*, pp. 362–367. (MIT Press, 2010)

[273] J. Sprenger and N. Weinberger. Simpson's paradox. *The Stanford Encyclopedia of Philosophy*. Web. (2021) https://plato.stanford.edu/archives/sum2021/entries/paradox-simpson/

[274] D. Marr. *Vision*. (WH Freeman and Company, 1982)

[275] E. Adrian. *The Physical Background of Perception*. (Clarendon Press, 1947)

[276] M. E. Garrett, I. Nauhaus, J. H. Marshel, and E. M. Callaway. Topography and areal organization of mouse visual cortex. *Journal of Neuroscience* **34**, no. 37, 12587–12600. (2014)

[277] J. Zhuang, L. Ng, D. Williams, M. Valley, Y. Li, M. Garrett, and J. Waters. An extended retinotopic map of mouse cortex. *eLife* **6**, e18372. (2017)

[278] A. P. Georgopoulos, A. B. Schwartz, and R. E. Kettner. Neuronal Population Coding of Movement Direction. *Science* **233**, no. 4771, 1416–1419. (1986)

[279] P. Bashivan, K. Kar, and J. J. DiCarlo. Neural population control via deep image synthesis. *Science* **364**, no. 6439, eaav9436. (2019)

[280] M. M. Churchland, J. P. Cunningham, M. T. Kaufman, J. D. Foster, P. Nuyujukian, S. I. Ryu, and K. V. Shenoy. Neural population dynamics during reaching. *Nature* **487**, no. 7405, 51–56. (2012)

[281] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski. *Neuronal dynamics: From single neurons to networks and models of cognition*. (Cambridge University Press, 2014)

[282] T. Hastie, R. Tibshirani, and J. Friedman. Chapter 7: Model Assessment and Selection in *The Elements of Statistical Learning*, 2nd ed., pp. 219–260. (Springer, 2009)

[283] R. Kohavi. A study of cross-validatoin and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence*. (1995)

[284] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski. Chapter 10: Estimating parameters of probabilistic neuron models in *Neuronal Dynamics*, 1st ed., pp. 243–266. (Cambridge University Press, 2014)

[285] D. Wolpert and W. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation* **1**, no. 1, 67–82. (1997)

[286] B. Hille. *Ion Channels of Excitable Membranes*, 3rd ed. (Sinauer Associates, 2001)

[287] R. Kobayashi, Y. Tsubo, and S. Shinomoto. Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold. *Frontiers in Computational Neuroscience* **3**, 9. (2009)

[288] N. Brunel, V. Hakim, and M. J. Richardson. Single neuron dynamics and computation. *Current Opinion in Neurobiology* **25**, 149–155. (2014)

[289] A. V. Herz, T. Gollisch, C. K. Machens, and D. Jaeger. Modeling single-neuron dynamics and computations: a balance of detail and abstraction. *Science* **314**, no. 5796, 80–85. (2006)

[290] C. Koch. *Biophysics of Computation: Information Processing in Single Neurons*. (Oxford university press, 2004)

[291] T. M. McKenna, J. L. Davis, and S. F. Zornetzer. *Single Neuron Computation*. (Academic Press, 2014)

[292] W. Rall. Membrane potential transients and membrane time constant of motoneurons. *Experimental Neurology* **2**, no. 5, 503–532. (1960)

[293] P. Poirazi and A. Papoutsi. Illuminating dendritic function with computational models. *Nature Reviews Neuroscience* **21**, no. 6, 303–321. (2020)

[294] A. T. Schaefer, M. E. Larkum, B. Sakmann, and A. Roth. Coincidence detection in pyramidal neurons is tuned by their dendritic branching pattern. *Journal of Neurophysiology* **89**, no. 6, 3143–3154. (2003)

[295] P. Vetter, A. Roth, and M. Häusser. Propagation of action potentials in dendrites depends on dendritic morphology. *Journal of Neurophysiology* **85**, no. 2, 926–937. (2001)

[296] A. S. Shai, C. A. Anastassiou, M. E. Larkum, and C. Koch. Physiology of layer 5 pyramidal neurons in mouse primary visual cortex: coincidence detection through bursting. *PLoS Computational Biology* **11**, no. 3, e1004090. (2015)

[297] M. Psarrou, S. S. Stefanou, A. Papoutsi, A. Tzilivaki, V. Cutsuridis, and P. Poirazi. A simulation study on the effects of dendritic morphology on layer V prefrontal pyramidal cell firing behavior. *Frontiers in Cellular Neuroscience* **8**, 287. (2014)

[298] A. Gidon, T. A. Zolnik, P. Fidzinski, F. Bolduan, A. Papoutsi, P. Poirazi, M. Holtkamp, I. Vida, and M. E. Larkum. Dendritic action potentials and computation in human layer 2/3 cortical neurons. *Science* **367**, no. 6473, 83–87. (2020)

[299] J. L. Krichmar, S. J. Nasuto, R. Scorcioni, S. D. Washington, and G. A. Ascoli. Effects of dendritic morphology on CA3 pyramidal cell electrophysiology: a simulation study. *Brain Research* **941**, no. 1-2, 11–28. (2002)

[300] E. P. Cook and D. Johnston. Active dendrites reduce location-dependent variability of synaptic input trains. *Journal of Neurophysiology* **78**, no. 4, 2116–2128. (1997)

[301] S. Gasparini, M. Migliore, and J. C. Magee. On the initiation and propagation of dendritic spikes in CA1 pyramidal neurons. *Journal of Neuroscience* **24**, no. 49, 11046–11056. (2004)

[302] G. Ariav, A. Polsky, and J. Schiller. Submillisecond precision of the input-output transformation function mediated by fast sodium dendritic spikes in basal dendrites of CA1 pyramidal neurons. *Journal of Neuroscience* **23**, no. 21, 7750–7758. (2003)

[303] G. Eyal, H. D. Mansvelder, C. P. de Kock, and I. Segev. Dendrites impact the encoding capabilities of the axon. *Journal of Neuroscience* **34**, no. 24, 8063–8071. (2014)

[304] Y. Deitcher, G. Eyal, L. Kanari, M. B. Verhoog, G. A. Atenekeng Kahou, H. D. Mansvelder, C. P. De Kock, and I. Segev. Comprehensive morpho-electrotonic analysis shows 2 distinct classes of L2 and L3 pyramidal neurons in human temporal cortex. *Cerebral Cortex* **27**, no. 11, 5398–5414. (2017)

[305] A. Papoutsi, G. Kastellakis, and P. Poirazi. Basal tree complexity shapes functional pathways in the prefrontal cortex. *Journal of Neurophysiology* **118**, no. 4, 1970–1983. (2017)

[306] B. A. Richards, T. P. Lillicrap, P. Beaudoin, *et al.* A deep learning framework for neuroscience. *Nature Neuroscience* **22**, no. 11, 1761–1770. (2019)

[307] F. H. Sinz, X. Pitkow, J. Reimer, M. Bethge, and A. S. Tolias. Engineering a less artificial intelligence. *Neuron* **103**, no. 6, 967–979. (2019)

[308] A. Tzilivaki, G. Kastellakis, and P. Poirazi. Challenging the point neuron dogma: FS basket cells as 2-stage nonlinear integrators. *Nature Communications* **10**, no. 1, 1–14. (2019)

[309] B. E. Kalmbach, R. Gray, D. Johnston, and E. P. Cook. Systems-based analysis of dendritic nonlinearities reveals temporal feature extraction in mouse L5 cortical neurons. *Journal of Neurophysiology* **117**, no. 6, 2188–2208. (2017)

[310] V. Francioni and M. T. Harnett. Rethinking single neuron electrical compartmentalization: dendritic contributions to network computation *in vivo*. *Neuroscience* **489**, 185–199. (2021)

[311] S. W. Kuffler. Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology* **16**, no. 1, 37–68. (1953)

[312] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology* **195**, no. 1, 215–243. (1968)

[313] B. Vintch, J. A. Movshon, and E. P. Simoncelli. A convolutional subunit model for neuronal responses in macaque V1. *Journal of Neuroscience* **35**, no. 44, 14829–14841. (2015)

[314] J. M. McFarland, Y. Cui, and D. A. Butts. Inferring nonlinear neuronal computation based on physiologically plausible inputs. *PLoS Comput Biol* **9**, no. 7, e1003143. (2013)

[315] W. Truccolo, J. A. Donoghue, L. R. Hochberg, E. N. Eskandar, J. R. Madsen, W. S. Anderson, E. N. Brown, E. Halgren, and S. S. Cash. Single-neuron dynamics in human focal epilepsy. *Nature Neuroscience* **14**, no. 5, 635. (2011)

[316] W. R. Taylor, S. He, W. R. Levick, and D. I. Vaney. Dendritic computation of direction selectivity by retinal ganglion cells. *Science* **289**, no. 5488, 2347–2350. (2000)

[317] R. Jolivet, A. Rauch, H.-R. Lüscher, and W. Gerstner. Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *Journal of Computational Neuroscience* **21**, no. 1, 35–49. (2006)

[318] E. O. Neftci, H. Mostafa, and F. Zenke. Surrogate gradient learning in spiking neural networks: Bringing the power of gradient-based optimization to spiking neural networks. *IEEE Signal Processing Magazine* **36**, no. 6, 51–63. (2019)

[319] A. I. Weber and J. W. Pillow. Capturing the dynamical repertoire of single neurons with generalized linear models. *Neural Computation* **29**, no. 12, 3260–3289. (2017)

[320] B. W. Mel. Synaptic integration in an excitable dendritic tree. *Journal of Neurophysiology* **70**, no. 3, 1086–1101. (1993)

[321] P. Poirazi and B. W. Mel. Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron* **29**, no. 3, 779–796. (2001)

[322] R. Legenstein and W. Maass. Branch-specific plasticity enables self-organization of nonlinear computation in single neurons. *Journal of Neuroscience* **31**, no. 30, 10787–10802. (2011)

[323] P. F. Pinsky and J. Rinzel. Intrinsic and network rhythmogenesis in a reduced Traub model for CA3 neurons. *Journal of Computational Neuroscience* **1**, no. 1, 39–60. (1994)

[324] R. Urbanczik and W. Senn. Learning by the dendritic prediction of somatic spiking. *Neuron* **81**, no. 3, 521–528. (2014)

[325] R. Naud, B. Bathellier, and W. Gerstner. Spike-timing prediction in cortical neurons with active dendrites. *Frontiers in Computational Neuroscience* **8**, 90. (2014)

[326] R. Naud, A. Payeur, and A. Longtin. Noise gated by dendrosomatic interactions increases information transmission. *Physical Review X* **7**, 031045. (2017)

[327] N. L. Golding and N. Spruston. Dendritic sodium spikes are variable triggers of axonal action potentials in hippocampal CA1 pyramidal neurons. *Neuron* **21**, no. 5, 1189–1200. (1998)

[328] M. E. Larkum and J. J. Zhu. Signaling of layer 1 and whisker-evoked Ca^{2+} and Na^{+} action potentials in distal and terminal dendrites of rat neocortical pyramidal neurons in vitro and in vivo. *Journal of Neuroscience* **22**, no. 16, 6991–7005. (2002)

[329] S. L. Smith, I. T. Smith, T. Branco, and M. Häusser. Dendritic spikes enhance stimulus selectivity in cortical neurons in vivo. *Nature* **503**, no. 7474, 115–120. (2013)

[330] J. Schiller, G. Major, H. J. Koester, and Y. Schiller. NMDA spikes in basal dendrites of cortical pyramidal neurons. *Nature* **404**, no. 6775, 285–289. (2000)

[331] M. E. Larkum, J. J. Zhu, and B. Sakmann. A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature* **398**, no. 6725, 338–341. (1999)

[332] J. C. Magee and M. Carruth. Dendritic voltage-gated ion channels regulate the action potential firing mode of hippocampal CA1 pyramidal neurons. *Journal of Neurophysiology* **82**, no. 4, 1895–1901. (1999)

[333] N.-l. Xu, M. T. Harnett, S. R. Williams, D. Huber, D. H. O'Connor, K. Svoboda, and J. C. Magee. Nonlinear dendritic integration of sensory and motor input during an active sensing task. *Nature* **492**, no. 7428, 247–251. (2012)

[334] M. E. Larkum, W. Senn, and H.-R. Lüscher. Top-down dendritic input increases the gain of layer 5 pyramidal neurons. *Cerebral Cortex* **14**, no. 10, 1059–1070. (2004)

[335] F. Zenke and S. Ganguli. Superspike: Supervised learning in multilayer spiking neural networks. *Neural Computation* **30**, no. 6, 1514–1541. (2018)

[336] I. S. Jones and K. P. Kording. Do biological constraints impair dendritic computation? *arXiv*. (2021) doi: `https://doi.org/10.48550/arXiv.2103.03274`

[337] A. D. Bird and H. Cuntz. Dendritic normalisation improves learning in sparsely connected artificial neural networks. *bioRxiv*. (2020)

[338] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl. Algorithms for hyper-parameter optimization. *Advances in Neural Information Processing Systems* **24**. (2011)

[339] M. Feurer and F. Hutter. Chapter 1: Hyperparameter optimization in *Automated Machine Learning*, pp. 3–33. (Springer, 2019)

[340] B. Cramer, Y. Stradmann, J. Schemmel, and F. Zenke. The Heidelberg spiking data sets for the systematic evaluation of spiking neural networks. *IEEE Transactions on Neural Networks and Learning Systems*. (2020)

[341] M. Hessel, H. van Hasselt, J. Modayil, and D. Silver. On inductive biases in deep reinforcement learning. *arXiv*. (2019) doi: `https://doi.org/10.48550/arXiv.1907.02908`

[342] A. Goyal and Y. Bengio. Inductive biases for deep learning of higher-level cognition. *arXiv*. (2020) doi: `https://doi.org/10.48550/arXiv.2011.15091`

[343] P. W. Battaglia, J. B. Hamrick, V. Bapst, *et al.* Relational inductive biases, deep learning, and graph networks. *arXiv*. (2018) doi: `https://doi.org/10.48550/arXiv.1806.01261`

[344] A. Santoro, D. Raposo, D. G. Barrett, M. Malinowski, R. Pascanu, P. Battaglia, and T. Lillicrap. A simple neural network module for relational reasoning. *arXiv*. (2017) doi: `https://doi.org/10.48550/arXiv.1706.01427`

[345] Y. Bengio and Y. LeCun. Scaling learning algorithms towards AI. *Large-Scale Kernel Machines* **34**, no. 5, 1–41. (2007)

[346] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature* **521**, no. 7553, 436–444. (2015)

[347] J. Lee, Y. Bahri, R. Novak, S. S. Schoenholz, J. Pennington, and J. Sohl-Dickstein. Deep neural networks as gaussian processes. *arXiv*. (2017) doi: `https://doi.org/10.48550/arXiv.1711.00165`

[348] Q. Li and H. Sompolinsky. Statistical mechanics of deep linear neural networks: The back-propagating renormalization group. *arXiv*. (2020) doi: `https://doi.org/10.48550/arXiv.2012.04030`

[349] M. Belkin, D. Hsu, S. Ma, and S. Mandal. Reconciling modern machine-learning practice and the classical bias–variance trade-off. *Proceedings of the National Academy of Sciences* **116**, no. 32, 15849–15854. (2019)

[350] S. Du, J. Lee, H. Li, L. Wang, and X. Zhai. Gradient descent finds global minima of deep neural networks. *International Conference on Machine Learning*, 1675–1685. (2019)

[351] D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick. Neuroscience-inspired artificial intelligence. *Neuron* **95**, no. 2, 245–258. (2017)

[352] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research* **15**, no. 1, 1929–1958. (2014)

[353] A. Goyal, A. Lamb, J. Hoffmann, S. Sodhani, S. Levine, Y. Bengio, and B. Schölkopf. Recurrent independent mechanisms. *arXiv*. (2019) doi: `https://doi.org/10.48550/arXiv.1909.10893`

[354] G. E. Hinton, S. Sabour, and N. Frosst. Matrix capsules with EM routing. *International Conference on Learning Representations*. (2018)

[355] N. Y. Masse, G. D. Grant, and D. J. Freedman. Alleviating catastrophic forgetting using context-dependent gating and synaptic stabilization. *Proceedings of the National Academy of Sciences* **115**, no. 44, E10467–E10475. (2018)

[356] G. Kastellakis, D. J. Cai, S. C. Mednick, A. J. Silva, and P. Poirazi. Synaptic clustering within dendrites: an emerging theory of memory formation. *Progress in Neurobiology* **126**, 19–35. (2015)

[357] J. Rossbroich, D. Trotter, J. Beninger, K. Tóth, and R. Naud. Linear-nonlinear cascades capture synaptic dynamics. *PLoS Computational Biology* **17**, no. 3, 1–27. (2021)

[358] D. E. Wilson, D. E. Whitney, B. Scholl, and D. Fitzpatrick. Orientation selectivity and the functional clustering of synaptic inputs in primary visual cortex. *Nature Neuroscience* **19**, no. 8, 1003–1009. (2016)

[359] E. Hay, S. Hill, F. Schürmann, H. Markram, and I. Segev. Models of neocortical layer 5b pyramidal cells capturing a wide range of dendritic and perisomatic active properties. *PLoS Computational Biology* **7**, no. 7, e1002107. (2011)

[360] L. Beaulieu-Laroche, E. H. Toloza, M.-S. Van der Goes, M. Lafourcade, D. Barnagian, Z. M. Williams, E. N. Eskandar, M. P. Frosch, S. S. Cash, and M. T. Harnett. Enhanced dendritic compartmentalization in human cortical neurons. *Cell* **175**, no. 3, 643–651. (2018)

[361] T. P. Vogels, H. Sprekeler, F. Zenke, C. Clopath, and W. Gerstner. Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science* **334**, no. 6062, 1569–1573. (2011)

[362] F. Gambino, S. Pagès, V. Kehayas, D. Baptista, R. Tatti, A. Carleton, and A. Holtmaat. Sensory-evoked LTP driven by dendritic plateau potentials in vivo. *Nature* **515**, no. 7525, 116–119. (2014)

[363] G. Doron, J. N. Shin, N. Takahashi, *et al*. Perirhinal input to neocortical layer 1 controls learning. *Science* **370**, no. 6523. (2020)

[364] F. Zenke. SpyTorch v0.3. Software. (Zenodo, 2019) doi: `https://doi.org/10.5281/zenodo.3724018`