The role of adaptation in neural coding
Alison I Weber\(^1,3\) and Adrienne L Fairhall\(^2,3\)

The concept of ‘neural coding’ supposes that neural firing patterns in some sense represent some external correlate, whether sensory, motor, or structural knowledge about the world. While the implied existence of a one-to-one mapping between external referents and neural firing has been useful, the prevalence of adaptation challenges this. Adaptation provides neural responses with dynamics on timescales that range from milliseconds up to many seconds. These timescales are highly relevant for sensory experience in the natural world, in which local statistical properties of inputs change continuously, and are additionally altered by active sensing. Adaptation has a number of consequences for coding: it creates short-term history dependence; it engenders complex feature selectivity that is time-varying; and it can serve to enhance information representation in dynamic environments. Considering how to best incorporate adaptation into neural models exposes a fundamental dichotomy in approaches to the description of neural systems: ones that take an explicitly ‘coding’ perspective versus ones that describe the system’s dynamics. Here we discuss the pros and cons of different approaches to the modeling of adaptive dynamics.

Addresses
\(^1\) Graduate Program in Neuroscience, University of Washington, United States
\(^2\) Department of Physiology and Biophysics and Computational Neuroscience Center, University of Washington, United States
\(^3\) UW Institute for Neuroengineering, United States

Corresponding author: Fairhall, Adrienne L (fairhall@uw.edu)

Neurons and neural circuits are dynamic on a continuum of timescales, undergoing continuous changes in cellular properties such as excitability due to the dynamics of ion channels and synaptic properties such as vesicle availability with timescales of history dependence ranging from milliseconds to seconds [5,6]. Response properties evaluated at the single neuron level are further influenced by network level dynamics. Typically, experiments exploring adaptation have probed the system in two or more different and distinct states. After abruptly changing the value of a stimulus, or the statistical ensemble from which stimulus samples are drawn, one can track the response of the system as the system relaxes into its new state in order to characterize how firing rates change; or fit models for the coding properties of the system in its new ‘adapted’ state, Figure 1. Despite this typical experimental design, however, natural environments are continuously changing, and the underlying neural dynamics responsible for both temporal relaxation and steady-state characterizations of adaptation are always in play as the nervous system is continually driven by ongoing changes in the environment.

In light of these ongoing processes, what does it mean for a neuron to adapt? How might one think about adaptation in the absence of a natural separation of timescales between ‘response’ and ‘adaptation’? How does our understanding of adaptation depend on the models we employ? As the field moves beyond single neuron recordings and isolated stimulus presentations to situations, in which population recordings are made during natural stimulation and behavior, it is worth revisiting adaptation in view of emerging model characterizations.
This can serve to allow a relatively larger response at the onset of a novel stimulus. Adaptation also refers to a change in the steady-state responses properties of a neuron or network following a change in the stimulus ensemble. For example, the spike-triggering features or tuning curve of a neuron might change after a change in stimulus statistics or context. In response to a contrast change, early sensory areas have been shown to adjust both the temporal feature that drives responses and the associated nonlinear input-output function (Figure 1). This adjustment often occurs such that the dynamic range of the responses is matched to the range of inputs, a phenomenon known as gain scaling [3], regarded as a signature of efficient coding. This canonical computation of the nervous system occurs throughout early [4*,7,8] and higher order [9] sensory areas, as well as in circuits that perform multisensory integration [10], motor control [11], and economic decision making [12].

The multiplicity of cellular dynamical timescales leads to multiple timescales of firing rate adaptation [13,14]. A signature of these multiple timescales is that adaptation to a step change in stimulus often does not occur with a fixed time constant, but with a time constant that scales with the frequency of change in the stimulus. When driven with continuously varying stimuli such as sine waves, these responses have been found to be compactly described by a rate response to the stimulus that has power law properties [14,15], or acts like fractional differentiation applied to the input [14,16]. In the case of cortical neurons, this transformation serves to decorrelate output, which may improve coding efficiency of the system [15].

**Using models to probe adaptation**

How can one model adaptation? There are two main classes of models in sensory neuroscience. One may try to develop _dynamical models_ of a system, approximate dynamical equations for neural activity $r(t)$ that may be driven by external variables $s(t)$, generically:

$$\frac{dr}{dt} = F(r,s)$$

(1)

The function $F$ represents a general nonlinear, possibly multidimensional, set of dynamical equations, such as the Hodgkin-Huxley equations driven by a current input. Alternatively, _statistical models_ evaluate and express the correlational relationships between the firing of a neuron or population, $r(t)$, at the current time $t$, and external variables $s_r(t)$ in a time window extending over a window of time $\tau$ relative to $t$,

$$r(t) = f(s_r(t))$$

(2)

for some functional form $f$.

While at some level equivalent, the first class of models aims to explicitly capture dynamical mechanisms while the
second is often thought of as defining ‘coding’. Examples of Eq. (2) include simple, phenomenological models such as linear-nonlinear (LN) models [17]. Typically, fitting such models under changing stimulus conditions reveals changes in a neuron’s response properties (Figure 1) [18–21]. A general way to capture such adaptive effects is to extend Eq. (2) to incorporate parameters \( \theta \) describing the stimulus-response mapping, such as the shape of a filter or the gain of a tuning curve, which may themselves be time-varying:

\[
\mathbf{r}(t) = f(s_t(t), \theta_T(t))
\]

(3)

Here, the time variation of \( \theta_T \) captures the adaptation of the system: it is a function of the longer-term history of the stimulus, \( s_T(t) \), where \( T \) is much longer than \( \tau \):

\[
\theta_T(t) = g(s_T(t))
\]

(4)

Ideally a complete system characterization according to Eqs. (3–4) would include an evaluation of how the model parameters, \( \theta \), depend on longer term stimulus properties. Studies often identify slowly varying properties of stimuli such as mean and variance [22,23] or offset and envelope [23] as examples of relevant long timescale properties. For example, in the phenomenon of gain scaling discussed above, the system’s gain curves exhibit a dynamic range that is modulated by the stimulus’ time-varying standard deviation \( \sigma(t) \) [13]. In the fly neuron H1, at the same time, the mean slow-varying firing rate \( R \) behaves as a fractional derivative of \( \sigma(t) \), so in that case the full response model can be written as [3]

\[
\mathbf{r}(t) = R(\sigma(t))f(s_t(t)/\sigma(t))
\]

(5)

This can be considered as a temporally multiplexed code [13,23,24], whereby detailed spike timing conveys information about short timescale stimulus variations, while the slow-varying rate conveys information about \( \sigma(t) \).

**Models and mechanisms**

It is important to note that a system that does not have any explicit adaptive dynamics can demonstrate dependence on stimulus conditions when viewed through the lens of an LN model [25,26]. This effect has been demonstrated in model neurons without adaptive currents, including leaky integrate-and-fire neurons [27] and simple conductance-based models [25], and in computational instantiations of the Rechardt correlator [28]. In these models, fitted linear filters and nonlinearities show changes between stimulus conditions despite no slow temporal dynamics or parameter changes. Further, in some cases one can analytically predict the changes in filter time constants by analysis of the underlying system [29]. Thus the ‘mechanism’ for these changes need not be specific channel or synaptic dynamics, but rather an understanding of how the components of the LN model description arise from the underlying neural dynamics [29,30,31].

A key property of overt adaptation mechanisms is that response properties are modulated by the history of activity of the system. A coding model framework that takes this explicitly into account might therefore be expected to better fit the timecourse of adaptive responses. Recurrent linear-nonlinear models, or generalized linear models (GLMs), incorporate response history dependence into an LN model framework (Figure 1) [17]. In the GLM, the gain of the response to a stimulus feature is multiplicatively modulated by the history of firing via a ‘spike history filter’. This explicit history dependence allows GLMs to straightforwardly capture simple adaptive phenomena, like spike frequency adaptation, by accumulating negative feedback with each subsequent spike [32].

GLMs can also capture more complex adaptive phenomena (Figure 2a). A GLM including selectivity to a single stimulus feature and a single spike history filter can demonstrate gain rescaling, such that all nonlinearities overlay when normalized by stimulus contrast, as observed experimentally [30]. These apparent adjustments in the nonlinearity occur without any corresponding adjustment or explicit adaptation of the nonlinearity in the GLM. Further, by allowing a sufficiently long history filter, GLMs are able to capture experimentally observed multiple-timescale dynamics. A GLM can exhibit signatures of fractional differentiation, such as near-constant phase lead, to the same extent as a Hodgkin-Huxley model with multiple afterhyperpolarization currents with a range of timescales (Figure 2b).

While GLMs capture a wide array of adaptive phenomena, what insights, if any, do we gain into neural systems by fitting GLMs to data? One might hope that this formulation could allow one to separate out stimulus-driven and history-driven effects and determine, for example, a single characteristic spike-triggering stimulus feature that is invariant under different stimulus conditions, while the spike history filter might isolate activity-driven internal neural dynamics. However, generally this is not the case. Upon a change in mean input, for example, while the presence of the spike history filter captures the dynamics of the change in firing rate, in general both the form of the optimal spike history filter and of the stimulus feature depend on the stimulus condition [22,32]. Thus, there is no clear interpretation of these model components in terms of cellular mechanism.

**Coding on multiple timescales**

More complex models can provide a richer basis to fit multiple-timescale response properties in experimental data. For example, recent work has shown that cortical neurons exhibit a surprisingly complex form of temporal
Simple models can capture adaptive phenomena. (a) A GLM fit to responses of a Hodgkin-Huxley neuron that exhibits gain scaling, the adjustment of the dynamic range of the input/output curve to the range of the stimulus, also shows gain scaling. Left: Input-output functions for the GLM, plotted as a function of stimulus value for four different stimulus standard deviations (σ). Right: Input-output functions plotted by normalized stimulus values overlay, indicating gain rescaling. (b) A GLM fit to responses of a Hodgkin-Huxley neuron with three different afterhyperpolarization currents also shows signatures of fractional differentiation, including near-constant phase lead across frequencies (top) and a power law relationship between gain and stimulus period (bottom).

sequence sensitivity: they adapt to, apparently ‘binding’, repeated lengthy sequences of images, and respond strongly when the sequence is interrupted [33]. This behavior can be interpreted as a form of predictive coding, whereby a neural system responds with high amplitude to an input that violates a recently built expectation, here the missing arrival of an element of a sequence. These apparently sophisticated responses may also arise from multiple-timescale adaptive dynamics. Latimer et al. show that similar offset responses are observed in multiple cortical areas [34] (like those observed in the retina [35]), and can be predicted from an extension of an LN model that nonlinearly integrates contributions from temporal ‘subunits’. This model combines 3–4 stimulus filters that span a wide range of timescales—up to many seconds. In contrast to Eqs. (3–4), this model describes the response, in a single step, as a function of the entire longer-term history of the stimulus, $s_T(t)$,

$$r(t) = h(s_T(t))$$

albeit where $h$ is a multidimensional model that can accommodate effects such as multiplicative modulation.

This model describes coding of multiple stimulus timescales on the same footing; long-term stimulus properties are not explicitly parsed out, as in Eq. (5). However, fitting such a model requires fitting to an increasingly high dimensional stimulus as $T$ increases. It may be more efficient to fit models like Eq. (6) to inputs, in which slow-varying statistical components have been extracted and specified at lower temporal resolution.

**Network models**

One can extend the statistical modeling approach to arbitrarily complex models. Recently, McIntosh et al. [36**] used population data from simultaneously recorded retinal ganglion cells to fit a multilayer convolutional neural network (CNN), analogous to cascaded LN operations [35]. After fitting this model using natural movies, the model was able to predict responses to natural scenes and, perhaps surprisingly, to generalize responses to other stimulus categories. As a feed-forward model, the CNN model framework of [36**] lacks dynamics beyond the 400 ms temporal filtering of the initial layer and so does not reproduce the full range of long-timescale adaptive dynamics that the retina exhibits. However, the inclusion of a long short-term memory (LSTM) unit, a standard and powerful method for incorporating long-timescale memory in deep learning networks, allows the network to capture firing rate adaptation following a change in stimulus contrast.

Such increasingly complex models are able to capture complex response properties, but what, if anything, is learned from them about the system? Multilayer network models of the retina can exhibit properties similar to retinal neurons without explicitly training the network to include these properties. For example, a two-layer
subunit model of the retina showed nodes in the first layer to be consistent with properties of bipolar cells, such as receptive field size and temporal structure [37], and a CNN appears to self-organize into multiple cell types resembling those seen in the retina [38]. It is possible that this approach will eventually provide new and testable predictions about network structure. However, models like CNNs are sufficiently large and complex that functional interrogation of the network becomes difficult or impossible without additional analysis. Even when a network model captures a system’s complex response properties, it does not ‘reveal’ these computations: rather, the model must be probed, like the experimental system, with stimuli designed to test for the presence of these properties.

Dynamical models
An alternative approach to the modeling of adaptation is to construct dynamical models as in Eq. (1). Simplified dynamical models can capture fundamental dynamical mechanisms at the single neuron level [39]. Recurrent network models can produce adaptive dynamics without incorporating explicit neuronal adaptation mechanisms [40]. However, recent work suggests that artificial spiking neural networks can improve performance by incorporating neuronal adaptation, while at the same time reducing the total number of spikes produced, improving network efficiency [41].

Hybrid approaches combine the interpretability of coding-type models with dynamical mechanism by augmenting statistical models with a dynamical component that allows for time dependence. This could model a specific biophysical mechanism like synaptic replenishment, such as the linear-nonlinear-kinetic model of Ozysyal and Baccus [42] or nonlinear threshold dynamics [43]. Such models can provide mechanistic predictions that can be tested against data [44]. One might also incorporate a postulated computational principle such as sensitivity of the system to multiple dimensions [26,28] or divisive suppression [45,46]. Such models can act as a bridge between low-level mechanism and identification of algorithm. As for the statistical models above, one must still probe the system with simplified stimuli in order to probe algorithmic computational motifs.

Outlook
We have discussed here a number of model frameworks that allow one to capture how adaptive dynamics shape responses to naturalistic inputs, including time-dependent coding properties that are modulated by activity and stimulus properties. More generally, these models account for encoding of multiple stimulus timescales. In moving from explicit coding models to more complex statistical or dynamical models, one may gain predictive accuracy but lose interpretability, in particular in terms of theories of efficient or predictive coding. All of the characterizations of adaptation we have discussed have primarily been applied to single neurons, but current recording strategies enable a deeper consideration of the meaning and role of adaptation at the network level. Some studies examine how the cumulative effects of single neuron adaptation affect population level coding [47] or arise from circuit mechanisms [48,49,50**], but there has been little work to date on exploring adaptive population codes, where coding elements are defined at a multineuronal level [51]. Understanding the properties of population codes in these key intermediate timescales, so critical for natural environments, is an interesting avenue for future work.

Conflict of interest statement
Nothing declared.

Acknowledgements
We thank Kenneth Latimer for useful discussions and for contributing results on the ability of GLMs to capture adaptive properties (Figure 2), and Mátr Lengyel for valuable comments on the manuscript. ALF was supported by the Washington Research Foundation UW Institute for Neuroengineering, the Human Frontiers in Science Program and the Simons Foundation’s Collaboration for the Global Brain.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as: ** of special interest and * of outstanding interest


   This provocative and thoughtful article challenges the usefulness of the concept of coding in neural systems.


   This review summarizes and links the phenomenology and various theoretical interpretations of adaptation.


This primer compares different models used to characterize neural spike trains, focusing on variants of the linear-nonlinear cascade framework, including models that incorporate multiple stimulus features (such as through spike-triggered covariance methods) and models that incorporate history dependence (such as generalized linear models).


Neurons in mouse primary visual cortex show adaptation to repeated presentations of the same sequence of visual stimuli and elevated responses to novel or mismatched stimuli in the sequence.


A multi-layer convolutional neural network is shown to capture responses of retinal ganglion cells to natural movies better than other classes of models, including LN models and GLMs. Inclusion of a recurrent layer is necessary to capture contrast adaptation effects.


Upon exposure to a particular stimulus orientation, a recurrent neural network model of orientation tuning shows tuning curve shifts similar to those observed in primary visual cortex. These tuning curve shifts occur without any change in network properties or explicit adaptive mechanisms.


The authors use a novel model formulation to identify different mechanisms in ON and OFF pathways of the retina that contribute to contrast adaptation.


A single model that incorporates a divisive interaction between excitatory and suppressive LN components can account for ganglion cell responses at high and low contrast.


This review discusses mechanisms that give rise to adaptive phenomena, with emphasis on how mechanisms operate at, and potentially interact on, multiple spatial scales.