Transfer functions for burst firing probability in a model neocortical pyramidal cell

Bruce P. Graham[1*], Jim W. Kay[2], William A. Phillips[3],

1 Computing Science and Mathematics, Faculty of Natural Sciences, University of Stirling, Stirling, U.K.

2 School of Mathematics and Statistics, University of Glasgow, Glasgow, U.K.

3 Psychology, Faculty of Natural Sciences, University of Stirling, Stirling, U.K.

* bruce.graham@stir.ac.uk

Abstract

Neocortical layer 5 thick-tufted pyramidal cells are prone to exhibiting burst firing on receipt of coincident basal and apical dendritic inputs. These inputs carry different information, with basal inputs coming from feedforward sensory pathways and apical inputs coming from diverse sources that provide context in the cortical hierarchy. We explore the information processing possibilities of this burst firing using computer simulations of a noisy compartmental cell model. Simulated data on stochastic burst firing due to brief, simultaneously injected basal and apical currents allows estimation of burst firing probability for different stimulus current amplitudes.

Information-theory-based partial information decomposition (PID) is used to quantify the contributions of the apical and basal input streams to the information in the cell output bursting probability. Different operating regimes are apparent, depending on the relative strengths of the input streams, with output burst probability carrying more or less information that is uniquely contributed by either the basal or apical input, or shared and synergistic information due to the combined streams. We derive and fit transfer functions for these different regimes that describe burst probability over the different ranges of basal and apical input amplitudes. The operating regimes can be

classified into distinct modes of information processing, depending on the contribution of apical input to output bursting: *apical cooperation*, in which both basal and apical inputs are required to generate a burst; *apical amplification*, in which basal input alone can generate a burst but the burst probability is modulated by apical input; *apical drive*, in which apical input alone can produce a burst; and *apical integration*, in which strong apical or basal inputs alone, as well as their combination, can generate bursting. In particular, PID and the transfer function clarify that the apical amplification mode has the features required for contextually-modulated information processing.

Author summary

Pyramidal cells are the dominant cell type of the neocortex and are fundamental to cortical information processing. They are more complex signal processors than the simple computing units used in artificial neural networks. In particular, each pyramidal cell receives two complementary input streams that jointly determine the cell output and hence the information that the cell transmits. One stream comes from sources that convey current sensory information. Another stream carries information from higher in the cortical hierarchy and from other sensory modalities. This stream provides context for the processing of the sensory input stream. Current experimental data and theories suggest that the effect of this stream can vary with the behavioural state of the animal, ranging from active exploration to sleep. In this theoretical study, we explore the possible interactions of these sensory and contextual input streams in determining information transmission in a computer model of a rodent neocortical pyramidal cell. We demonstrate that the cell can operate in a number of modes that encompass the ability to carry out contextually-modulated information processing. This is central to how we perceive and react to the world on the basis of our past experience and knowledge.

Introduction

One of the key challenges in understanding cortical information processing is in building a comprehensive picture of pyramidal cells as two-point processors receiving dual

information-rich input streams that must be combined to produce an output that is part of a meaningful and coherent pattern of activity across the system or sub-system of which those cells are a part. Pyramidal cells are distinctly not single-point processors, as commonly used in artificial neural networks. Neocortical pyramidal cells have essentially two sites of synaptic integration, targetting the basal dendrites and apical tuft dendrites, respectively [1, 2]. The basal inputs are in the perisomatic region and so arrive close to the final site of cell synaptic integration and action potential initiation. These inputs are from feedforward sources, conveying sensory information via specific regions of the thalamus and cortex and provide the primary drive to the receiving pyramidal cell that determines its receptive field. On the other hand, inputs to the apical tuft come from diverse feedback sources in higher cortex, long-range lateral cortex and thalamus. These inputs are hypothesised to provide contextual information that can modulate the cell's response to its primary drive [2]. This has been termed *apical amplification* [2–5]. Contextually-modulated information processing is widespread in the neocortex and underpins many cognitive functions during conscious processing [4–7].

In addition to apical amplification, two other modes have been designated as *apical drive*, in which apical inputs alone can generate spiking output in the pyramidal cell, and *apical isolation* in which the apical inputs have no effect on cell output, which is purely driven by basal (perisomatic) inputs [5, 8]. These different modes can be activated in particular behavioural states through circuit inhibition and neuromodulation. Apical drive is hypothesised to be fundamental to dreaming, whereas apical isolation might be in effect during dreamless sleep [8]. General anesthesia has been shown to decouple pyramidal neurons by effectively cutting off the affects of apical input on cell output [3, 9].

There is considerable experimental evidence from awake, behaving rodents that the basal and apical input streams both contribute to pyramidal cell (PC) responses to sensory stimulation. In the visual system and barrel cortex, PC receptive field responses can be altered in task-dependent ways, indicating contextually-modulated processing via apical amplification [10–13]. There is also evidence for apical drive: visual activity has been recorded due to locomotion in the dark, when visual sensory information is low [14]; in the barrel cortex, increased excitability in PC dendrites leads to an increased detection of apparent whisker stimulation in the absence of actual physical

stimulation [11].

To help shed light on exactly how basal and apical input streams may combine to modulate and refine cellular receptive field responses, here we characterise the integration of apical and basal (perisomatic) inputs in a computational model of a thick-tufted layer 5 neocortical pyramidal cell [15], and interpret what this means for information processing. Given the electrical remoteness of the apical tuft from the final somatic integration zone, the simple additive contribution of apical inputs to basal drive is likely to be modest at best. In thick-tufted layer 5 pyramidal cells, the apical tuft is particularly electrically isolated from the soma, but it contains its own initiation zone for broad calcium spikes. Such spikes are capable of generating a bursting output of sodium spikes from the cell. Apical calcium spikes are often initiated by a combination of apical synaptic input with a back-propagating sodium spike initiated somatically by basal inputs. This has been termed backpropagation-activated calcium (BAC) spike firing [1] and is a candidate for providing the apical amplification underpinning contextual modulation [2].

We focus on this bursting output as the distinctive signature of the combined effects of coincident basal and apical input. Bursting has long been hypothesised to be a powerful information coding signal in the brain [16–18]. Bursts can very effectively be read out downstream through the filtering of chemical synapses via stochastic transmitter release and short-term plasticity [16, 18, 19]. They usually exist within a stream of single spikes and theoretical work has established that bursts and single spikes can coexist and carry different information that can be read out by downstream neurons [18–22]. Within a firing rate framework bursts can be identified as a brief and significant increase in the firing rate [23, 24].

Computer simulations are used to collect data on the bursting output of the model cell for varying strengths of coincident basal and apical inputs. We use information theoretical analysis and transfer function fitting to characterise the information transmission provided by bursts about the dual input streams. Partial information decomposition (PID) [25] is able to quantify the contributions of the apical and basal input streams to the information in the cell output. Depending on the relative strengths of the input streams, output burst probability carries more or less information that is uniquely contributed by either the basal or apical input, or shared and synergistic

information due to the combined streams. Transfer functions allow us to visualise the functional relationship between these inputs and bursting output.

By analysing different ranges for the strengths of basal and apical inputs we can see explicitly the difference in information transmission between different modes of operation. In addition to apical amplification and apical drive, we identify two further modes: *apical cooperation* in which both basal and apical input is required to generate bursting; and *apical integration* in which strong apical or basal inputs alone, as well as their combination, can generate bursting. In all cases the impact of apical input is compared with the mode of apical isolation, in which there is no effect of apical input.

Results

Computer simulations of a reduced-compartmental model of a thick-tufted layer 5 neocortical pyramidal cell [15] are used to explore the contributions of basal and apical input streams to spiking output generation. Details of the model and simulations are given in Materials and methods. We explicitly investigate how basal and apical inputs combine to produce burst firing and we derive and fit transfer functions that describe the burst probability as a function of the strength (amplitude) of simultaneous, brief basal and apical inputs, over different ranges of input strengths. Information transmission of the two input streams is examined using partial information decomposition (PID) of the bursting data

Firing modes in the model cell

As detailed in Materials and methods, the model [15] that we use is able to produce backpropagation-activated calcium (BAC) spike firing in response to coincident, brief basal (somatic) and apical (tuft) inputs, which in turn generates a burst of 2 to 3 somatic (output) spikes. In our simulations, BAC-firing is probabilistic due to the presence of random background activity impinging independently on the soma and apical tuft. A schematic of the model cell and examples of bursting and non-bursting action potential firing in the model due to different amplitudes and durations of basal and apical input are shown in Fig 1.

The range of basal amplitudes considered (0 to 3 nA) is such that stronger

amplitudes of sufficient duration will generate somatic bursting by themselves. Very short basal pulses (2 ms) can only generate a single somatic spike (Fig 1a), and this back-propagating somatic spike can interact with apical input to generate a dendritic calcium spike which then leads to a second or third somatic sodium spike (Fig 1b). Longer duration basal pulses (5 or 10 ms) of sufficient amplitude alone can generate a rapid 2-spike burst through partial activation of calcium currents (Fig 1c). When combined with apical input this can lead to a full calcium spike and a third somatic spike, creating a full burst (Fig 1d; further examples in Supplementary S1 Fig).

Up to an apical amplitude of about 1 nA, the apical input alone usually does not produce a full dendritic calcium spike and hence does not cause a somatic spiking response (S1 Fig a). At amplitudes above 1 nA, apical input is increasingly able to trigger a dendritic calcium spike that leads to somatic burst firing after a small time delay (S1 Fig b). Very occassionally this form of bursting may occur at a lower apical amplitude coincident with a basal input that does not by itself initiate a somatic spike. In the absence of inhibition, apical inputs that are strong enough to elicit output spikes by themselves rarely produce isolated single spikes due to the width of the calcium spike. This is in contrast to basal inputs (Fig 1a).

These ranges of stimulus current amplitudes correspond to the equivalent of at most a few tens of co-occurring synaptic inputs at basal and apical sites.

Our initial analysis concentrates on the situation in which apical input alone does not usually lead to burst firing. This results in subtle nonlinear interactions between the basal and apical inputs in producing bursting output when the combined inputs are of sufficient strength. The rich burst probability data in this case enables the derivation of a transfer function that subsumes the situation in which basal input alone cannot produce a burst and is readily extended to cover the situation in which apical input alone can produce a burst. These remaining situations will be covered in detail later.

Burst probability

The initial data set of simulated spiking output covers an apical input range from 0 to 1 nA, and basal ranges dependent on the duration of the basal stimulus, being 0 to 3 nA for durations of 2 or 5 ms, and 0 to 1 nA for 10 ms duration. Apical input alone very

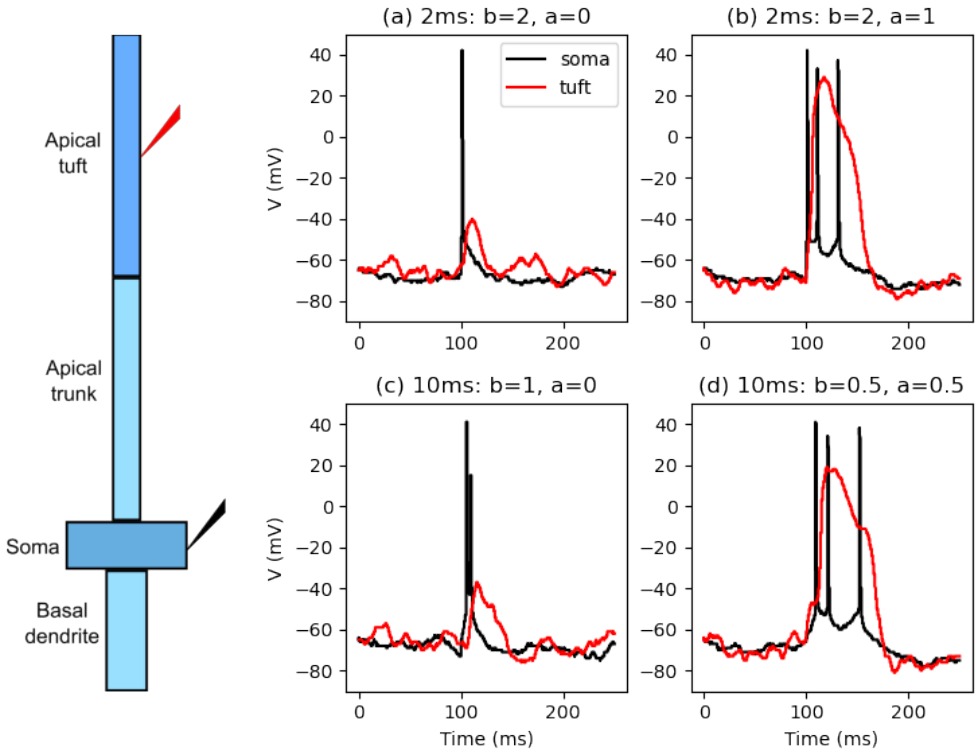


Fig 1. Examples of spiking responses to different stimulus strengths and durations. Cell schematic shows model structure and somatic (black) and apical tuft (red) stimulation and recording sites. For basal amplitude b nA and apical amplitude a nA and basal duration 2 ms: (a) b=2, a=0; (b) b=2, a=1; basal duration 10 ms: (c) b=1, a=0; (d) b=0.5, a=0.5.

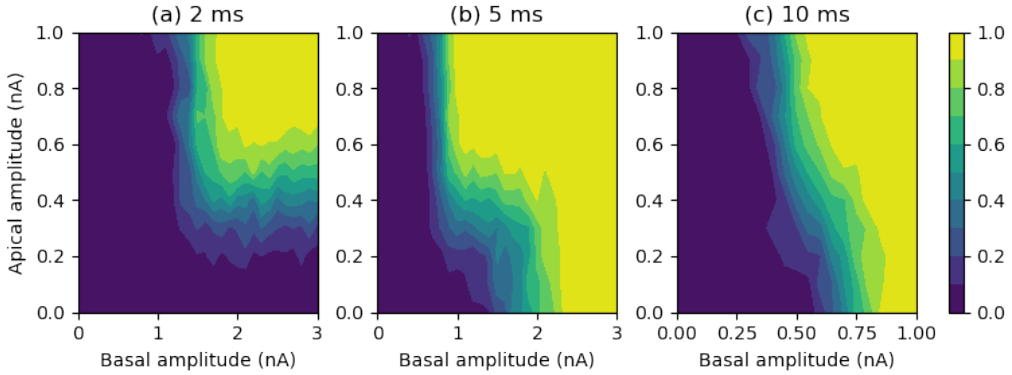


Fig 2. Contour maps of burst probability for different durations of basal input. Left: basal duration 2 ms, basal maximum amplitude 3 nA; Middle: 5 ms, max amp 3 nA; Right: 10 ms, max amp 1 nA; apical maximum amplitude is 1 nA in all cases.

rarely produces spiking output in this range. The basal input alone can produce spiking and for the 5 and 10 ms durations can also produce bursting at the higher end of their ranges. Burst probabilities were calculated from the frequency of output bursts (2 or more spikes with interspike intervals less than 25 ms) over 100 simulations for each combination of amplitudes of basal and apical input for discrete increments in their respective amplitudes within their defined ranges (see Materials and methods). Contour maps of the burst probability across these different combinations of basal and apical input amplitudes and selected durations of basal input are shown in Fig 2.

From the maps it can be seen that for longer duration basal inputs (5 or 10 ms) burst probability is an increasing function of basal amplitude, with high burst probability being achieved at lower basal amplitudes as the amplitude of the apical input is increased. For brief 2 ms basal inputs, once the basal input is sufficiently strong so as to produce a somatic spike then burst probability becomes an increasing function of the apical amplitude. Apical input alone produces a very low (essentially zero) burst probability over the range of amplitudes used.

Information conveyed by bursting

Before deriving a transfer function for burst probability, we first consider the information that a burst output may carry about the apical and basal input streams. We calculate both classical mutual information measures and partial information decompositions (see Materials and methods), using the data displayed in Fig 2.

Classical information measures

The classical information measures for the mutual information (see Materials and methods for details) between the three random variables B (basal input), A (apical input) and Y (bursting output) are provided in Table 1.

Table 1. Classical information measures between B, A and Y for three durations of basal input: 2 ms (B2), 5 ms (B5), 10 ms (B10). $I(Y;B)$ is the mutual information shared between Y and B; $I(Y;A)$ is the information shared between Y and A; $I(Y;B|A)$ is the information shared between Y and B but not with A; $I(Y;A|B)$ is the information shared between Y and A but not with B; $I(Y;B,A)$ information shared between Y and the pair (B, A); $II(Y;B;A)$ is the interaction information; $H(Y)$ is the total output entropy and $H(Y)_{res}$ is the residual output entropy.

Duration	$I(Y;B)$	$I(Y;A)$	$I(Y;B\|A)$	$I(Y;A\|B)$	$I(Y;B,A)$	$II(Y;B;A)$	$H(Y)$	$H(Y)_{res}$
B2	0.260	0.185	0.371	0.256	0.556	0.111	0.883	0.327
B5	0.521	0.050	0.626	0.156	0.677	0.105	0.967	0.291
B10	0.597	0.021	0.658	0.082	0.679	0.061	0.992	0.313

For a basal duration of 2 ms, the transmitted joint mutual information (0.556 bit) is 63% of the output entropy (0.883 bit). The marginal mutual information between basal input and bursting output (0.260 bit) is quite low but is a little higher than the corresponding mutual information between apical input and bursting output (0.185 bit). Thus the unique information asymmetry is fairly low but positive, at 0.075 bit. The value of the interaction information (0.111) is also low, but gives a significant lower bound for the synergy in the system.

For basal durations of 5 and 10 ms, there are higher values for both the joint mutual information and the output entropy, with the transmitted information comprising a larger proportion of the output entropy (around 70% in both cases). The marginal mutual information between the basal input and bursting output is now larger (0.521 and 0.597 bit, respectively) and is much larger than that between the apical input and bursting output (0.05 and 0.021 bit). This means the unique information asymmetry is now substantial in magnitude and positive, at 0.471 and 0.576 bit, respectively. The value of the interaction information is lower than at 2 ms. These effects will be present in the results obtained by *any* method of partial information decomposition (see below).

The residual output entropy is fairly substantial in all cases.

Partial information decomposition

Partial information decomposition provides measures of a number of different quantities relating the two input variables B and A with the output Y: UnqB is the unique information B conveys about Y; UnqA is the unique information A conveys about Y; ShdS is the source shared (or redundant) information that both B and A have about Y, due to correlation between A and B; ShdM is the mechanistic shared information due to the probabilistic mechanism; Syn is the synergy or the information that the joint variable (B, A) has about Y that cannot be obtained by observing B and A separately (see Materials and methods for further details).

Four different numerical decompositions, namely Imin [25], Iproj [26], Ibroja [27, 28] and Idep [29], were applied to the data. In order to facilitate the comparison of the decompositions across the different basal durations, the normalised values of the PID components (see Materials and methods for details) are provided as stacked bar charts in Fig 3. Plots of the absolute values of the PID components are provided in Supplementary S2 Fig. Since there is no way of knowing the true values of these components it is advisable to place more emphasis on the decompositions which provide similar results. By design, the basal and apical input variables are independent and so the source shared information is equal to zero in all PIDs.

When the duration of the basal input is 2 ms, the Idep decomposition differs from those obtained with the other three methods, which are all very similar, and we focus on their results. Thus, for this basal duration, the transmitted information is mainly due to synergy as well as a large amount of shared mechanistic information. For durations of 5 ms or 10 ms, the four methods produce very similar decompositions in which the transmitted information is dominated by unique basal information, and there are much smaller levels of synergy and mechanistic shared information. As the basal duration is increased we see that synergy and mechanistic shared information decrease, and correspondingly the unique basal information increases, as does the level of unique information asymmetry.

For basal input of duration 5 ms, decomposition of the information transmitted by bursting shows that the basal and apical inputs combine to transfer 23% of the joint mutual information as synergy, and they contribute to transferring 7% as mechanistic

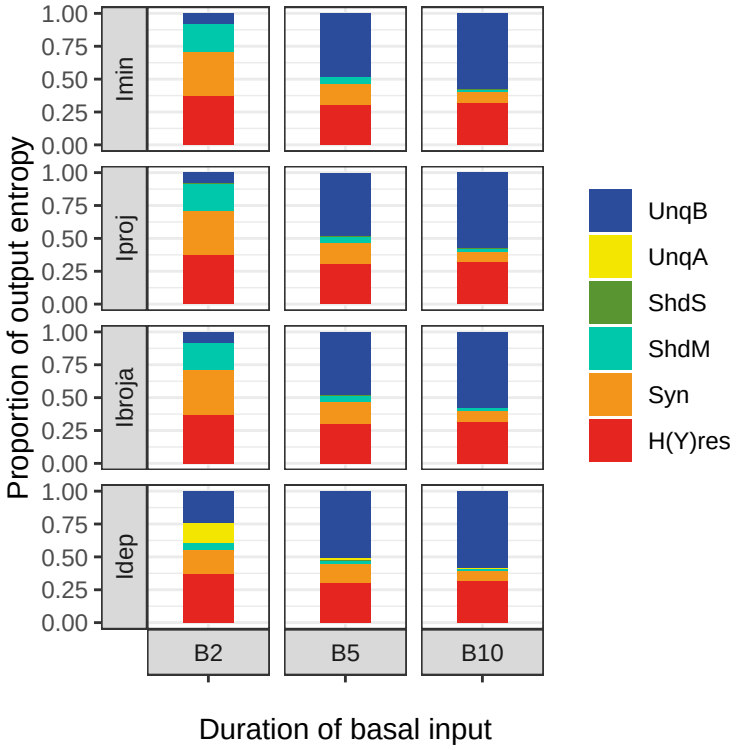


Fig 3. Normalised partial information decompositions obtained by using four different methods (Imin, Iproj, Ibroja, Idep) on burst probability data for apical inputs up to 1 nA and basal inputs up to 3 nA, for each of three durations of basal input: 2 ms (B2), 5 ms (B5), 10 ms (B10).

shared information. There is a marked asymmetry in the estimates of the unique basal and apical components. The unique information transmitted about the basal input is large, at 70% of the joint mutual information, but the unique information about the apical input is essentially equal to zero. As the information transmitted about the apical input in this case is close to zero this indicates that the apical input can have a large effect on output via the synergistic component while conveying little or no information about itself, thus indicating the presence of apical amplification. These observations also hold for a basal duration of 10 ms, with a slightly higher proportion of unique basal input information being transmitted.

Transfer function for burst firing

A transfer function for the probability of burst firing, $P_2(b, a)$, taking the strength (amplitude) of basal (somatic), b, and apical, a, current injections as arguments, was arrived at via the following reasoning and inspection of the burst probability data for

different durations of basal input, shown in Fig 2. This burst probability data and the least-squares transfer function fits are illustrated in Fig 4a-c. The corresponding transfer function parameter values are given in Table 2.

Firstly, the range of apical input used cannot cause spiking output on its own, so the probability of burst firing factors as the product of the probability of a first spike, which is a function, $P_{1b}(b)$, of basal input only, multiplied by the probability of a second or more spikes given that a first spike has occurred, $P_{21}(b, a)$, which is a function of both basal and apical input. So we have:

$$P_2(b, a) = P_{1b}(b)P_{21}(b, a). \tag{1}$$

First spike probability is reasonably a sigmoidal function of basal amplitude as this basal input sums with the noisy (random) membrane current of mean zero, giving:

$$P_{1b}(b) = \frac{1}{1 + \exp(-g_{1b}b + k_{1b})}. \tag{2}$$

Examining the burst probability curves when basal duration is short (Fig 4a), shows that in this case, once a first spike has occurred then the probability of a burst is constant with increasing basal amplitude, but with this constant increasing with apical amplitude. This gives:

$$P_{21}(b, a) = P_{2a}(a)[1 - P_{2b}(b)] + P_{2b}(b), \tag{3}$$

where $P_{2b}(b)$ is the burst probability achieved when apical input is zero; and $P_{2a}(a)$ is the probability of a burst due to apical input following a first somatic spike. The burst probabilities reached with increasing a are well matched by $P_{2a}(a)$ being a sigmoidal function of a:

$$P_{2a}(a) = \frac{1}{1 + \exp(-g_{2a}a + k_{2a})}. \tag{4}$$

Finally, when basal duration is longer than 2ms, basal input alone is able to generate bursts with increasing probability with increasing basal amplitude. This increasing burst probability clearly has a sigmoidal relationship with basal amplitude (Fig 4b,c

blue lines), so we define:

$$P_{2b}(b) = h_{2b}\frac{1}{1+\exp(-g_{2b}b + k_{2b})}, \tag{5}$$

where scaling factor $h_{2b} \ll 1$ when basal duration is 2ms, but $h_{2b} \approx 1$ for longer basal durations.

The full transfer function is thus given by the following set of equations:

$$\begin{aligned}
P_{1b}(b) &= \frac{1}{1+\exp(-g_{1b}b + k_{1b})} && (6)\\
P_{2b}(b) &= h_{2b}\frac{1}{1+\exp(-g_{2b}b + k_{2b})} && (7)\\
P_{2a}(a) &= \frac{1}{1+\exp(-g_{2a}a + k_{2a})} && (8)\\
P_2(b,a) &= P_{1b}(b)[P_{2a}(a)[1-P_{2b}(b)] + P_{2b}(b)] && (9)\\
&\equiv P_{1b}(b)P_{21}(b,a). && (10)
\end{aligned}$$

To support the veracity of this transfer function, the correspondence between the different components and aspects of first spike and bursting probability data is shown in Fig 4d-f. Here, P_{2b} is a direct fit to the underlying data, whereas P_{1b} and P_{2a} emerge implicitly from the optimisation and are not explicit fits, but nonetheless show a strong correspondence to the equivalent data. Note that the data shown for the first spike probability, P_{1b}, has been extracted from the simulation data, but is not explicitly used in the optimisation. The contribution of the apical amplitude to bursting, P_{2a}, is not explicitly available in the data, but good estimates are available using slices through the data for basal amplitude values that give a very low burst probability when the apical amplitude is zero, but have a burst probability near one when combined with strong apical input; these are the data examples shown in Fig 4d-f.

This transfer function (Eq 9) can expanded and rewritten to give the following form:

$$\begin{aligned}
P_2(b,a) &= P_{1b}(b)[P_{2a}(a) + P_{2b}(b) - P_{2a}(a)P_{2b}(b)] && (11)\\
&= P_{1b}(b)P_{2b}(b) + P_{1b}(b)P_{2a}(a)[1-P_{2b}(b)] && (12)\\
&\equiv f(b) + g(b,a). && (13)
\end{aligned}$$

This makes explicit that it contains a component that is only a function $f(b)$ of the basal amplitude, that summates with a combined function of the basal and apical amplitudes, $g(b, a)$. The function $g(b, a)$ is zero if $b = 0$ and is an increasing function of a for $b > 0$, thus it serves to amplify the basal response as a function of the strength of apical input, while not admitting a contribution of the apical input by itself. This is a clear demonstration of the neurobiological plausibility of apical amplification.

For the short 2 ms basal stimulation, the burst probability due to basal input alone is approximately zero. In this case the transfer collapses to approximately $P_2^s(b, a) \approx P_{1b}(b)P_{2a}(a) \equiv g^s(b, a)$. That is, burst probability is a function only of the combined inputs, in the form of the product of the probability that basal input produces a first spike with the ability of the apical input to transform this into BAC firing.

Table 2. Least squares fits of $P_2(b, a)$ transfer function parameter values for burst firing with different durations of basal input (2, 5 and 10 ms). Values are plus/minus their estimated standard error (apart from the $2b$ parameters at 2 ms duration, which are not well constrained by the data, but where the scaling factor h_{2b} is very small).

Dur	h_{2b}	g_{2b}	k_{2b}	g_{1b}	k_{1b}	g_{2a}	k_{2a}
2	0.0019	0.0012	40.4	7.3 ± 0.41	10.18 ± 0.57	10.45 ± 0.28	4.36 ± 0.12
5	1.02 ± 0.01	4.15 ± 0.16	7.7 ± 0.29	12.66 ± 0.69	10.09 ± 0.54	11.06 ± 0.37	4.13 ± 0.15
10	1.0 ± 0.014	15.43 ± 0.82	10.94 ± 0.57	19.81 ± 1.14	9.09 ± 0.5	8.8 ± 0.49	3.46 ± 0.19

Examples of contextual information processing

Orientation selectivity

To assess the functional implications of the derived transfer function, we consider the simple orientation selectivity example studied by [23]. Here, orientation-specific basal and apical input stimuli are generated from von Mises circular distributions. These stimuli are fed through the transfer function to give the distribution of output bursting probability. Fig 5 illustrates the orientation selectivity of the transfer function derived for a basal input duration of 10 ms.

A maximum basal amplitude of 0.6 is chosen for a preferred orientation of 0 radians. Without apical input, this basal amplitude gives a low but significant probability of burst firing (Fig 4c). This bursting probability is also highly sensitive to apical input. A maximum apical amplitude of 0.7 is chosen for this input's preferred orientation, which is able to significantly amplify the bursting probability when combined with a

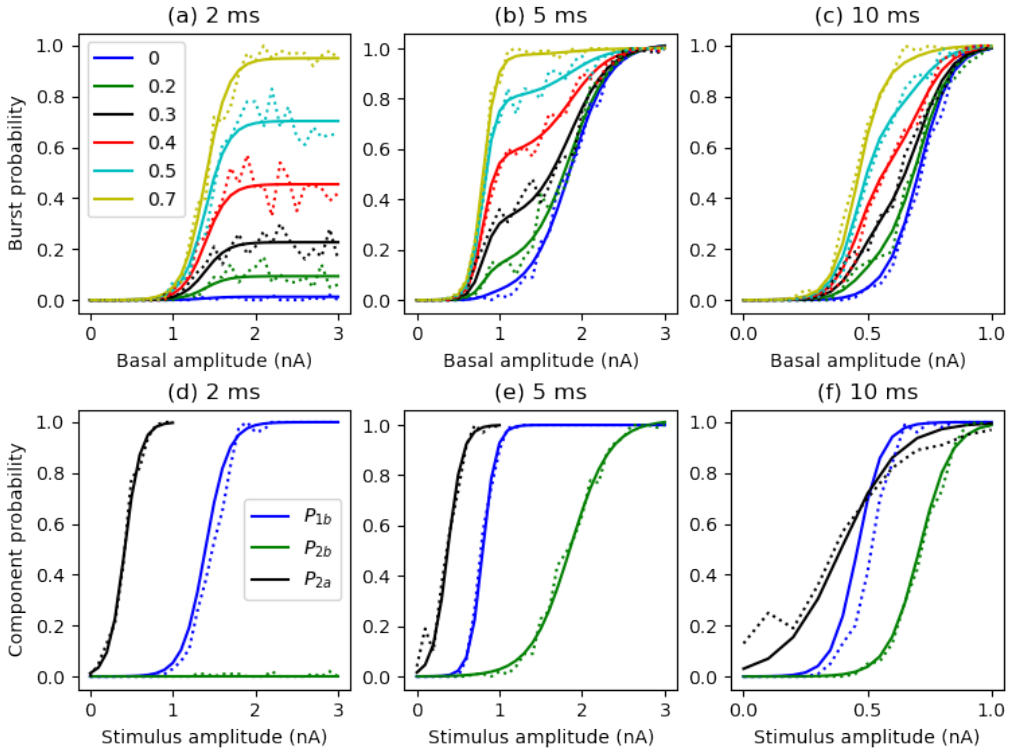


Fig 4. (a-c) Least-squares fit of transfer functions (solid lines), shown as burst probability versus basal amplitude, with plots for selected values of the apical amplitude (legend); dashed lines are the frequency data from simulations. (d-f) Comparison of transfer function components with estimates from the simulation data; first spike probability (P_{1b}) and burst probability when apical amplitude is zero (P_{2b}) are extracted exactly from the data; burst probability due to apical input alone (P_{2a}) is estimated by data slices for particular basal amplitudes (b=2 nA for 2 ms duration, 1.2 nA for 5 ms and 0.6 nA for 10 ms). Left: basal duration 2ms; Middle: 5ms; Right: 10ms.

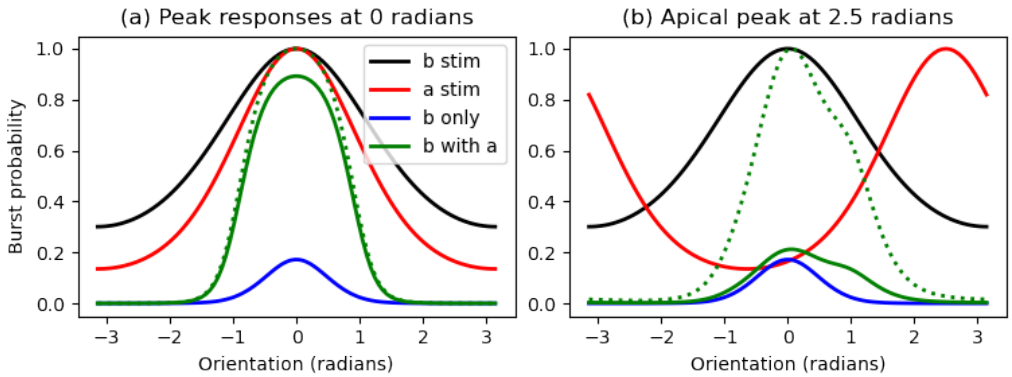


Fig 5. Orientation selectivity example.(a) both basal and apical inputs are tuned to an orientation of 0 radians. (b) apical input now tuned to an orientation of 2.5 radians. The dotted line in both plots indicates the basal with apical response with the output amplitude scaled to 1 for ease of comparison with the phase of the inputs. Basal amplitude is 0.6 nA at its preferred orientation; apical amplitude is 0.7 nA at its preferred orientation; basal duration is 10 ms.

sufficiently strong basal input.

With no apical input, the output response shows a small maximum burst probability at 0 radians (Fig 5 blue lines). Inclusion of the apical stimulus greatly amplifies the output burst probability when the basal and apical input selectivities are both centred at 0 radians (Fig 5a), but the apical stimulus has only a modest effect on the output amplitude and phase when its preferred orientation is shifted to 2.5 radians (Fig 5b).

Lateral amplification

To give a simple example of apical amplification working in a network of spiking cells, we simulated two Bahl model cells receiving basal stimulation via injected currents (as in the simulations for generating bursting data), but with the outputs of each cell feeding as an excitatory synaptic input (same time course as for apical input in previous simulations) to the apical tuft of the other cell. Short (2 ms) but strong basal stimulation to these cells when they are not mutually connected leads to a single somatic spike in each cell (Fig 6a). However, when each cell output is sent to the apical tuft of the other cell, then the single spike is often converted into a burst due to initiation of a dendritic calcium spike (Fig 6b).

This demonstrates that laterally connected cells can act to amplify their neighbours' responses when they respond to the same basal stimulus with an initial somatic spike. Note that such amplification only occurs when the lateral connections are via the apical

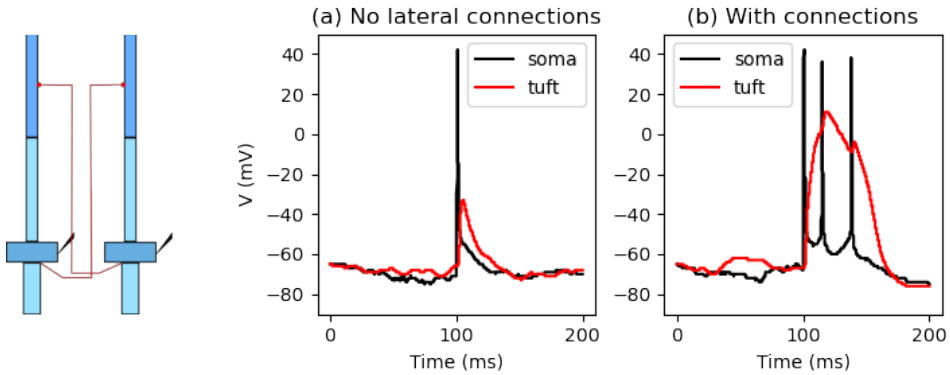


Fig 6. Lateral amplification example of voltage response in one cell of a two cell laterally-connected network (schematic). (a) Short, strong basal input only (basal duration 2 ms; amplitude 3 nA) without lateral connections. (b) Basal input with strong apical input provided via monosynaptic input from other cell (red connection lines in schematic).

tuft dendrites and not via the soma. Somatic connections do not add to the spiking response, at least for the short synaptic delays expected with nearby monosynaptic connections, as the input from a neighbour arrives within the refractory period following the spike originated by the basal stimulus.

Extended analysis of bursting regimes

The use of information theory above reveals two distinct operating regimes for burst generation: (1) a regime in which neither basal nor apical input alone can generate a burst, but together they can, as revealed with the use of short 2 ms duration basal inputs (Fig 3 B2); and (2) a regime in which strong basal inputs alone can produce a burst, as revealed with longer duration (5 ms and 10 ms) basal inputs (Fig 3 B5 & B10). The information transmission characteristics of these two regimes are quite distinct, with mostly synergistic and shared mechanistic information being transmitted in the first regime, and largely unique basal information in the second (Fig 3).

It is possible to examine different operating regimes within data from a single basal duration. So to further explore information transmission in different regimes, we have analysed an extended data set for basal inputs of 10 ms duration, in which larger apical inputs (up to 1.7 nA) are included. The contour map of burst probability across this data set is shown in Fig 7. Note that large apical inputs are able to produce bursts by themselves as they are strong enough to induce an apical calcium spike without the

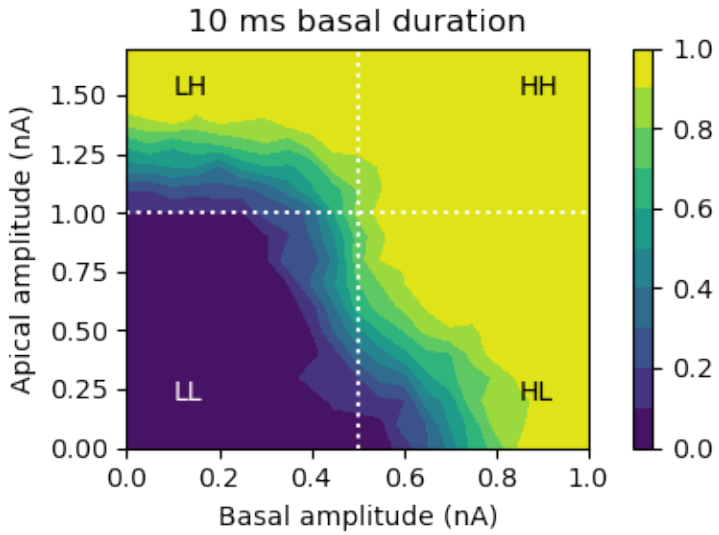


Fig 7. Contour map of burst probability for a basal duration of 10 ms, with basal (b) amplitude up to 1 nA and apical (a) amplitude up to 1.7 nA. Four operating regimes are indicated: (1) LL: *b* up to 0.5 nA, *a* up to 0.5 nA; (2) HL: *b* up to 1 nA, *a* up to 1 nA; (3) LH: *b* up to 0.5 nA, *a* up to 1.7 nA; (4) HH: *b* up to 1.0 nA, *a* up to 1.7 nA.

need for an initial somatic spike.

Taking as boundaries the basal and apical amplitudes at which their ability to produce a burst by themselves starts to increase from 0, this data can be divided into four different regimes. Defining *low* ranges as being basal inputs up to 0.5 nA and apical inputs up to 1.0 nA, and *high* ranges as being basal inputs up to 1 nA and apical inputs up to 1.7 nA, gives the four regimes: (1) LL: *b=low, a=low*; (2) HL: *b=high, a=low*; (3) LH: *b=low, a=high*; (4) HH: *b=high, a=high*. These regimes are indicated in Fig 7; note that the high ranges subsume the low ranges. In the LL regime, only the highest values of basal and apical inputs in these ranges combined result in a significant probability of bursting, with neither input able to produce a burst by itself. The HL regime corresponds to the contextually-modulated information processing regime we have analysed above, in which the strongest basal inputs alone are able to generate a burst, but apical inputs in this range alone cannot produce a burst. The LH regime reverses this so that strong apical inputs alone can produce a burst, but basal inputs cannot. The HH regime covers ranges of apical and basal inputs such that the strongest inputs of either pathway alone can produce a burst.

Classical information measures

We apply information theory to the burst probability distributions corresponding to each regime. Table 3 gives the values of of the classical information measures for these four regimes.

Table 3. Classical information measures (bit) for four regimes of basal and apical inputs, with 'L' denoting 'low' and 'H' denoting 'high'.

Regime	$I(Y;B)$	$I(Y;A)$	$I(Y;B\|A)$	$I(Y;A\|B)$	$I(Y;B,A)$	$II(Y;B;A)$	$H(Y)$	$H(Y)$res
LL	0.157	0.052	0.187	0.082	0.239	0.030	0.520	0.281
HL	0.599	0.028	0.674	0.103	0.702	0.075	0.995	0.293
LH	0.046	0.499	0.130	0.583	0.629	0.084	0.963	0.334
HH	0.301	0.220	0.462	0.381	0.682	0.161	0.952	0.270

Regime LL: *b=low, a=low* The joint mutual information is low, at 0.239 bit, as is the value of the output entropy, at 0.520, so the transmitted information is only 46% of the output entropy. The marginal mutual information between basal input and bursting output is stronger by a factor of about 3 than the corresponding mutual information between apical input and bursting output. The unique information asymmetry is fairly low at 0.105 bit. The value of the interaction information is low.

Regime HL: *b=high, a=low* Adding higher levels of basal input leads to higher values for both the joint mutual information and the output entropy, with the transmitted information comprising a larger proportion of the output entropy (71%). The marginal mutual information between the basal input and bursting output now has a large value (0.599), much larger than that between the apical input and bursting output (0.028). This means the unique information asymmetry is now substantial in magnitude and positive, at 0.571 bit. The value of the interaction information is larger, giving a larger lower bound for the synergy in the system.

Regime LH: *b=low, a=high* Adding higher levels of apical input also results in higher values for both the joint mutual information and the output entropy, with the transmitted information again comprising a large proportion of the output entropy (65%). However, the situation is reversed here: it is now the marginal dependence between basal input and bursting output that is low, being lower than in the first regime, and much lower than the large value pf 0.499 for the corresponding dependence

between apical input and bursting output. The unique information asymmetry is again substantial in magnitude but negative, at -0.453 bit. The value of the interaction information is again larger, giving a larger lower bound for the synergy in the system.

Regime HH: *b=high, a=high* In the previous comparisons it is found that introducing higher levels of apical or basal input separately to the first regime leads to a large value of unique information asymmetry, but what happens when higher levels of apical and basal input are added in combination? We find that the unique information asymmetry is much reduced in magnitude and that the marginal mutual informations between basal input and bursting output, and between apical output and bursting output, are similar in magnitude.

Partial information decomposition

Since it is of interest to compare the information decompositions across the four regimes we present normalised decompositions in Fig 8; plots of the absolute PID components are provided in Supplementary S3 Fig. The PIDs given by Imin, Iproj and Ibroja are very similar. The Idep approach also gives similar PIDs but only for two of the combinations. Hence we focus on the results given by the first three methods.

Regime LL: *b=low, a=low* Small amounts of information are transmitted that are unique to the basal input and due to mechanistic shared information and synergy, but information that is unique to the apical input is negligible.

Regime HL: *b=high, a=low* The introduction of higher levels of basal input results in the transmission of more information, which is dominated by information unique to the basal input, with some mechanistic shared information and a slightly larger amount of synergy - and negligible information unique to the apical input. Since the apical input combines with the basal input in the transmission of mechanistic shared information and synergy the presence of apical amplification of basal input is indicated.

Regime LH: *b=low, a=high* The introduction of higher levels of apical input results in the transmission of more information, which is dominated by information unique to the apical input, with some mechanistic shared information and a slightly

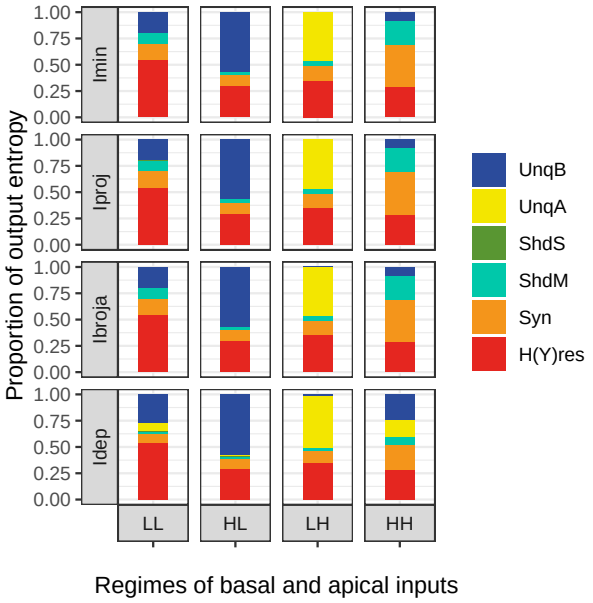


Fig 8. Normalised partial information decompositions obtained by using four different methods (Imin, Iproj, Ibroja, Idep) for each of the four regimes of basal and apical inputs. 'L' is an abbreviation of 'low', and 'H' is an abbreviation of 'high', so for example, 'HL' denotes the regime in which basal input is high and apical input is low.

larger amount of synergy - and negligible information unique to the basal input. Since the basal input combines with the apical input in the transmission of mechanistic shared information and synergy the presence of basal amplification of apical input is indicated.

Regime HH: *b=high, a=high* Adding higher levels of both basal and apical input provides a dramatic change to the information decomposition. The unique components become very small, with the output being largely composed of synergy, and to a lesser extent mechanistic shared information, with a much smaller component unique to the basal input. This regime could therefore be described as being predominantly a generative mode of information processing in which its output depends upon both diverse internal sources and more specific feedforward sources.

Extended transfer function

In what follows, the transfer functions for different operating regimes will be designated with a two letter superscript specifying the range of basal and apical inputs, respectively. For example, the transfer function derived above covers the high basal and low apical ranges and will be designated with the HL superscript, so that $P_2^{\mathrm{HL}}(b,a) \equiv P_2(b,a)$.

The transfer function above (Eq 9), that covers the first two regimes, can be extended to cover the other two regimes in which strong apical inputs alone can generate output bursting by including a factor that accounts for such bursting. The resultant transfer function is the original function fractionally summed with the contribution to bursting by apical input alone:

$$P_2^{\mathrm{HH}}(b,a) = P_2^{\mathrm{HL}}(b,a)[1 - P_{2a}^{\mathrm{H}}(a)] + P_{2a}^{\mathrm{H}}(a), \tag{14}$$

where $P_{2a}^{\mathrm{H}}(a)$ is a sigmoidal function of the apical amplitude (with superscript H designating the high apical range only):

$$P_{2a}^{\mathrm{H}}(a) = \frac{1}{1 + \exp(-g_{2a}^{\mathrm{H}} a + k_{2a}^{\mathrm{H}})}. \tag{15}$$

This can be rearranged to yield a form that is a sum of terms that depend on the basal and apical inputs alone plus a contribution due to the apical and basal inputs combined:

$$P_2^{\mathrm{HH}}(b,a) = f_a(a) + f_b(b) + g^{\mathrm{HH}}(b,a), \tag{16}$$

where $f_a(a) \equiv P_{2a}^{\mathrm{H}}(a)$ and $g^{\mathrm{HH}}(b,a) \equiv g^{\mathrm{HL}}(b,a)[1 - P_{2a}^{\mathrm{H}}(a)] - f_b(b)P_{2a}^{\mathrm{H}}(a)$. Now we are using $f_b(b) \equiv f(b)$ and $g^{\mathrm{HL}}(b,a) \equiv g(b,a)$ from Eq 13.

A least squares fit of $P_{2a}^{\mathrm{H}}(a)$ to the collected burst probabilities when the basal input is 0 yields $g_{2a}^{\mathrm{H}} = 10.35$ and $k_{2a}^{\mathrm{H}} = 12.66$. Using these parameter values, the extended transfer function fits the data well across the entire range of basal and apical input amplitudes (regime HH), as illustrated in Fig 9. This reveals that the influence of basal on apical input and vice versa is now qualitatively symmetric, with both being able to modulate the effect of the other (see also Fig 10d). However, the quantitative modulation is different, as indicated by the shape of the curves in Fig 9a compared to Fig 9b. In neither case is the modulation simply a change in gain or bursting threshold (shift), rather the 'modulating' input increases the likelihood of bursting over a range of the 'driving' input.

In the regime of low basal but high apical input, in which basal input alone does not produce a burst ($f_b(b) \approx 0$), this extended transfer function collapses to $P_2^{\mathrm{LH}}(b,a) = f_a(a) + g^{\mathrm{LH}}(b,a)$, where

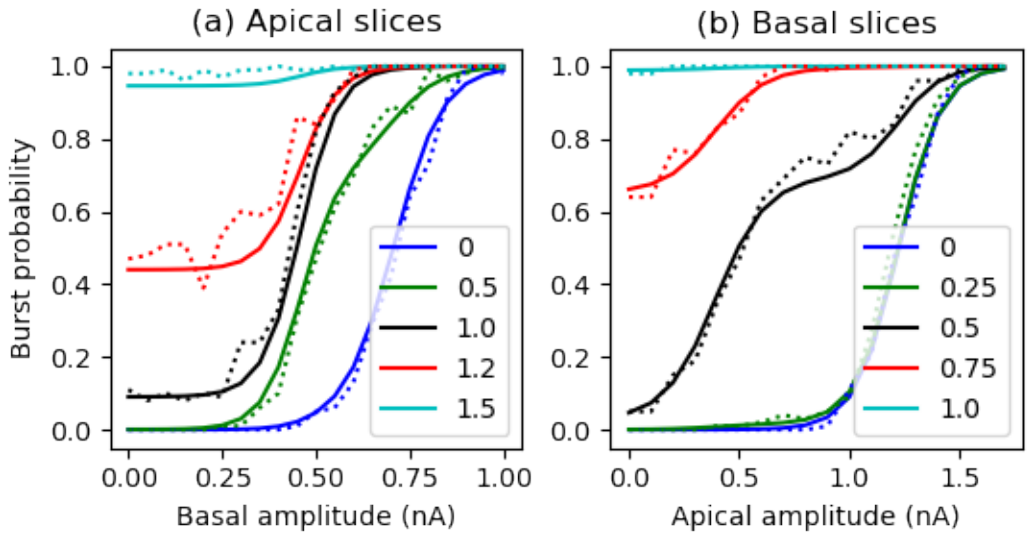


Fig 9. Fit of extended transfer function to full data set for a basal duration of 10 ms. (a) Transfer function for selected values of the apical amplitude (solid lines) plotted against the simulated burst probabilities (dotted lines) across the range of basal amplitudes. (b) Transfer function for selected values of the basal amplitude (solid lines) plotted against the simulated burst probabilities (dotted lines) across the range of apical amplitudes.

$g^{\mathrm{LH}}(b,a) \equiv g^{\mathrm{HL}}(b,a)[1 - P_{2a}^{\mathrm{H}}(a)] \equiv g^{\mathrm{LL}}(b,a)[1 - P_{2a}^{\mathrm{H}}(a)]$ where $g^{\mathrm{HL}}(b,a)$ reduces to $g^{\mathrm{LL}}(b,a) \equiv P_{1b}(b)P_{2a}(a)$ since $P_{2b}(b) \approx 0$ in this regime.

Summary of transfer functions

We have derived dual-input-stream transfer functions for the spike bursting probability of a neocortical pyramidal cell receiving short, concommitant apical and basal inputs of particular amplitudes on top of ongoing background synaptic activity. Four related transfer functions covering different operating regimes have been obtained as follows:

$$P_2^{\mathrm{LL}}(b,a) = P_{1b}(b)P_{2a}(a) \equiv g^{\mathrm{LL}}(b,a) \quad (17)$$

$$P_2^{\mathrm{HL}}(b,a) = P_{1b}(b)P_{2b}(b) + P_2^{\mathrm{LL}}(b,a)[1 - P_{2b}(b)] \equiv f_b(b) + g^{\mathrm{HL}}(b,a) \quad (18)$$

$$P_2^{\mathrm{LH}}(b,a) = P_{2a}^{\mathrm{H}}(a) + P_2^{\mathrm{LL}}(b,a)[1 - P_{2a}^{\mathrm{H}}(a)] \equiv f_a(a) + g^{\mathrm{LH}}(b,a) \quad (19)$$

$$P_2^{\mathrm{HH}}(b,a) = P_{2a}^{\mathrm{H}}(a) + P_2^{\mathrm{HL}}(b,a)[1 - P_{2a}^{\mathrm{H}}(a)] \equiv f_a(a) + f_b(b) + g^{\mathrm{HH}}(b,a) \quad (20)$$

The functional breakdown in Eq 20 is obtained by expanding the second term to give $f_b(b)$ plus $g^{\mathrm{HH}}(b,a) \equiv P_{1b}(b)P_{2a}(a)[1 - P_{2b}(b)][1 - P_{2a}^{\mathrm{H}}(a)] - P_{1b}(b)P_{2b}(b)P_{2a}^{\mathrm{H}}(a)$.

These transfer functions are the combination of a number of distinct components

(Eq 9) which can be interpreted as follows:

$P_{1b}(b) \equiv P(Z_1 = 1|b)$ is the probability of an initial somatic spike (binary random variable $Z_1 = 1$ when spike occurs) due to basal input alone.

$P_{2b}(b) \equiv P(Z_2 = 1|b)$ is the burst probability (binary random variable $Z_2 = 1$ when second spike (burst) occurs) due to basal input alone.

$P_{2a}(a) \equiv P(Z_2 = 1|Z_1 = 1, a)$ is the contribution of apical input to a full (or partial) calcium spike which can lead to a second (or more) somatic spike, following an initial somatic spike.

$P_{2a}^{\mathbf{H}}(a) \equiv P(Z_2 = 1|a)$ is the probability that apical input produces a calcium spike on its own.

$P_{1b}(b)$, $P_{2a}(a)$, $P_{2b}(b)$ and $P_{2a}^{\mathrm{H}}(a)$ are sigmoidal functions of either basal stimulus amplitude, b, or apical stimulus amplitude, a.

Modes of operation

To further appreciate the distinct characteristics of these transfer functions, contour plots of their output across the relevant ranges of basal and apical input amplitude are shown in Fig 10.

For basal and apical inputs both in their low ranges, bursting probability is clearly a function of both the basal and apical amplitudes combined (Fig 10a), as made explicit in Eq 17. This is not a simple summing, as bursting typically results from BAC-firing in which basal input triggers a single somatic spike which in turn combines with apical input to trigger a dendritic calcium spike and subsequent burst output. We might term this a mode of *apical cooperation*.

With basal input in its high range, but apical input low, bursting probability is strongly a function of basal amplitude, but with the transition from low to high bursting probability occurring at lower basal amplitudes with increasing apical amplitude (Fig 10b). This is captured by Eq 18 and has been termed *apical amplification* [2, 5]. This corresponds with contextually-modulated information processing.

With low range basal input but high range apical input (Fig 10c), bursting probability is almost purely a function of apical amplitude, with increased sensitivity in

the upper reaches of the basal low range where the basal input starts to trigger single somatic spikes, which then starts to amplify the apical response, as indicated by Eq 19. This has been termed *apical drive* [5, 8].

For both inputs in their high range (Fig 10d), bursting probability is a function of both the individual amplitudes and their combination, which is not a simple summation of the inputs, as indicated above for the other regimes, which this range subsumes. This is captured by Eq 20 and may be termed a mode of *apical integration*. Close examination of Fig 10d shows how it combines the three other modes of operation. Purely apical drive is evident when the basal amplitude is below around 0.3. For low values of the apical amplitude (below around 0.1) burst probability is purely a function of basal amplitude, corresponding with what has been termed *apical isolation* [5, 8]. The interaction of basal and apical amplitudes is clear between these two extremes, encompassing both apical cooperation and apical amplification.

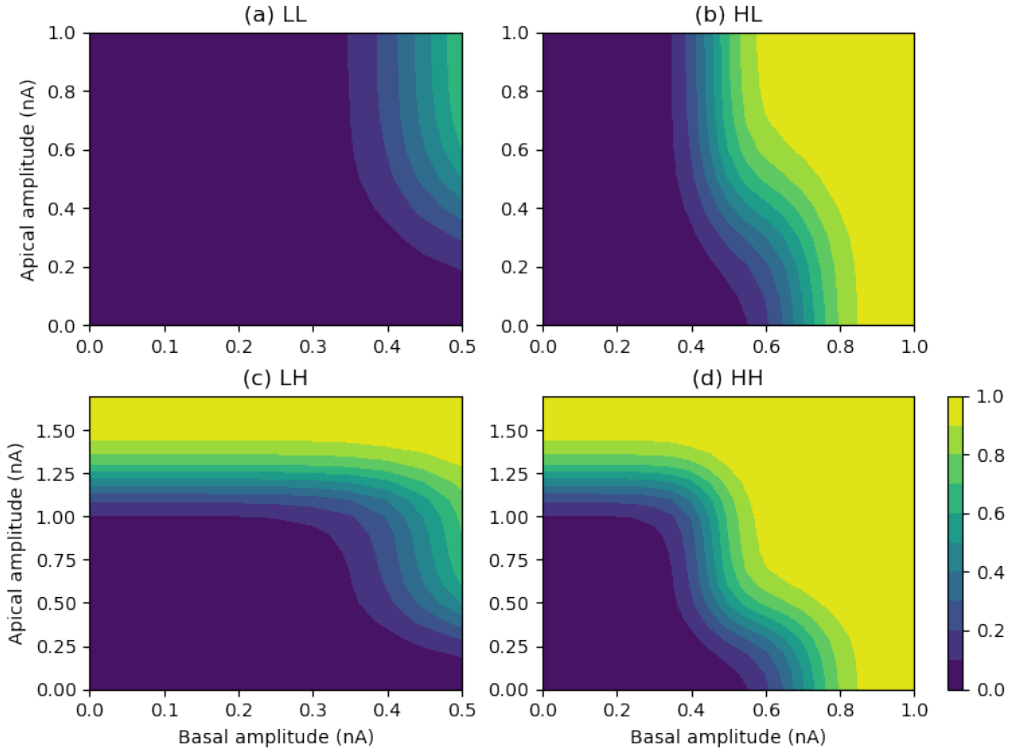


Fig 10. Contour plots of burst probability from the four transfer functions for different operating regimes for burst firing. (a) LL: basal=low, apical=low; (b) HL: basal=high, apical=low; (c) LH: basal=low, apical=high; (d) HH: basal=high, apical=high.

Discussion

We have examined input stream integration and subsequent information processing in one class of neocortical neurons, namely thick-tufted layer 5 pyramidal cells. These cells broadly have two anatomically-separated basal and apical input streams. These streams interact via the generation of a dendritic calcium spike to cause or amplify output burst firing. Though single spiking is also a unit of output currency, bursting is a particular indication of the interaction of the two input streams and we have derived transfer functions to capture this interaction. We now discuss this work in the light of other work on two-input-stream transfer functions, the information processing possibilities of such functions and of bursting outputs, and experimental evidence for the existence of the modes of operation we have identified.

Transfer functions with two input streams

We have concentrated particularly on the range of input amplitudes where strong (large amplitude) basal input is capable of generating a bursting output, but apical input is not. In this range the cell exhibits BAC firing [1] leading to apical amplification (AA) of the output response. The resultant transfer function can be written so that it is explicity the sum of a function of the basal amplitude alone summed with an amplifying component that depends on both basal and apical input amplitudes, but is zero when the basal input is zero (Eq 13). This is the form of transfer function that has been proposed as the basis for contextually-modulated information processing [30, 31] in which the activation of a computational unit (neuron) due to receptive field (driving) inputs is modulated by a separate stream of contextual inputs.

A basic transfer function with these characteristics is of the form [32]:

$$T_{rc} = s_r(k_1 + (1 - k_1)\exp(k_2 s_r s_c)) = s_r + (1 - k_1)s_r(\exp(k_2 s_r s_c) - 1), \tag{21}$$

where s_r and s_c are the accumulated (weighted and summed) receptive field and contextual inputs, respectively. This is of the form $f(x) + g(x, y)$ where $f(x) = x$ and $g(x, y) = x + (1 - k_1)x(\exp(k_2 xy) - 1)$. This is then be passed through a logistic

function to give an output limited to between 0 and 1:

$$O_{rc} = \frac{1}{1 + \exp(-k_3 T_{rc} + k_4)}. \tag{22}$$

The use of computational units (neurons) with this transfer function in artificial neural networks has been demonstrated in a wide range of applications where the networks were trained to achieve contextually-modulated information processing using the *Coherent Infomax (CI)* learning rule derived using information theory [32–35].

Using arguments based on consideration of the log odds for the generation of a second somatic spike, this basic transfer function has been extended in [36] to cover the situation of BAC firing in a pyramidal cell, as with our transfer function. The resultant function is:

$$T_{ba} = \beta_1 + \beta_2 b(1 + \exp(\beta_3 ba(1 + \exp(\beta_4 a)))), \tag{23}$$

where b and a are the strengths of the basal and apical inputs, respectively. The main difference from the basic function (Eq 21) is that the contextual term includes the transformation of apical input into BAC firing via:

$$c = \frac{1}{2} a(1 + \exp(\beta_4 a)), \tag{24}$$

plus the addition of a constant β_1 accounting for the prior odds of a burst.

Least-squares fitting to our simulated bursting data of this function (Eq 23), passed through a logistic (Eq 22) to give bursting probability, reveals that it has the same general characteristics of our more complex function and gives a tolerable fit to the 10 ms basal data (Fig 11a), though not as good as our more complex function (Fig 11b). The T_{ba} function cannot achieve a good fit to the more complex curves for the 2 ms and 5 ms basal data.

The effect of apical input in the T_{ba} function is to increase the gain of the bursting response to basal input. In our P_2 function, the response is left-shifted with increasing apical input, but by different amounts over the basal range, with low bursting probabilities being more quickly enhanced by apical input than higher probabilities. For P_2, the response curves for no apical input and for strong apical input are essentially identical in shape and hence gain, but over different basal ranges. In between, the gain

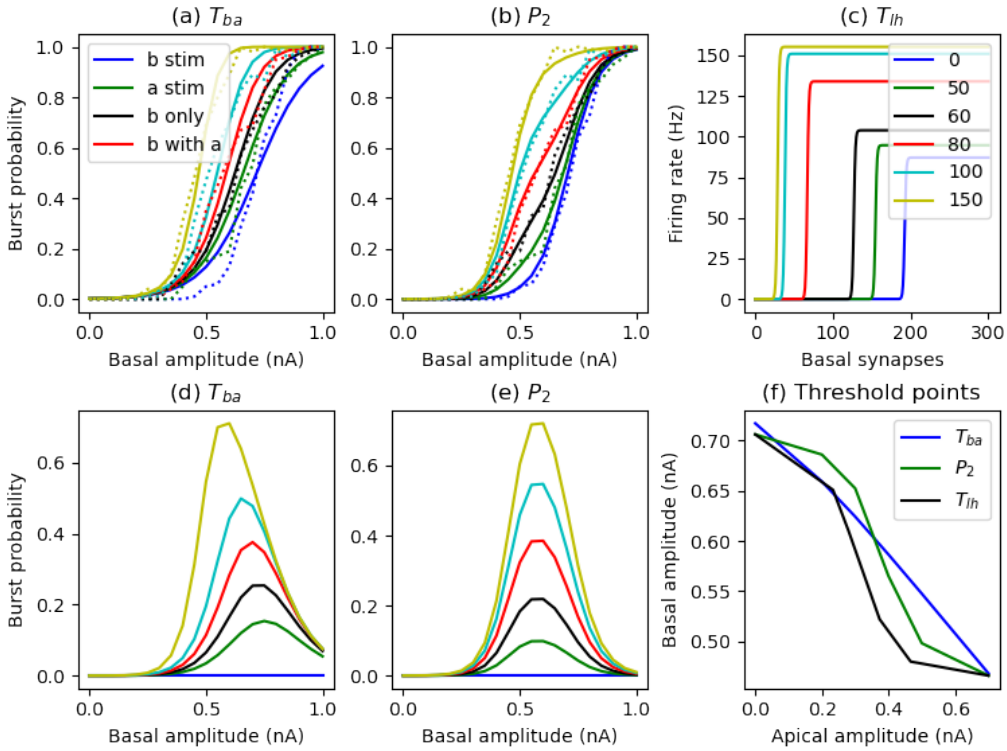


Fig 11. (a) T_{ba} transfer function through logistic (solid lines) with parameter values from least-squares fit to our 10 ms basal duration data, illustrated as burst probability versus basal amplitude, with plots for selected values of the apical amplitude (legend); (b) as for (a) but with our P_2 transfer function (note that dashed lines in (a) and (b) are the frequency data from simulations for a basal duration of 10 ms); (c) Shai et al. [23] transfer function T_{lh} for firing frequency as a function of the number of basal and apical synaptic inputs; (d,e) difference between burst probability curves with apical input and burst probability with basal input alone, for the T_{ba} and P_2 transfer functions, respectively; (f) combinations of apical and basal amplitudes required to achieve a threshold of 0.5 burst probability for the three transfer functions (note that the numbers of synaptic inputs for T_{lh} have been scaled to match the input currents of P_2 for comparison purposes).

is effectively lowered, resulting in a greater dynamic range in the response to basal inputs (see particularly the response when apical input is 0.4 nA in Fig 11b). By considering the difference between response curves with non-zero apical input and the response with only basal input, it can be seen that apical amplification is particularly effective over a limited basal range for both the T_{ba} and P_2 functions (Fig 11d,e). The point of greatest sensitivity to apical input shifts to lower basal values for T_{ba}, but for P_2 (and hence the pyramidal cell) it is constant at around a basal amplitude of 0.6 nA.

In a closely related study, Shai et al. [23] used simulations of a detailed compartmental model of a layer 5 pyramidal cell to examine its firing rate response to

different numbers of concomitant basal and apical synaptic inputs. Though not classifying the output as bursting or not, they demonstrated that the cell exhibited a rapid transition between low and high firing rates, as calculated over a 100 ms period. The high firing rates had a strong dependence on calcium levels underpinning BAC firing. They collected data across a range of numbers of basal and apical synaptic inputs, but with only a single simulation per pair of numbers, resulting in hard thresholds between low and high firing. They fitted a transfer function to this data in which output firing is driven by the basal input, but where the maximum firing rate M and the threshold T between low and high firing, were both functions of the number of apical inputs:

$$T_{lh} = \frac{M(n_a)}{1 + \exp(-(n_b - T(n_a)))}, \tag{25}$$

where n_a and n_b are the number of apical and basal synaptic inputs. The maximum rate $M(n_a)$ is an increasing sigmoidal function of n_a, whilst the threshold $T(n_a)$ is a decreasing sigmoidal function of n_a. This function, as fit by [23] to their simulated data, is shown in Fig 11c.

To compare our transfer function with that of Shai et al. [23], we can define a threshold of 0.5 for the transition from a low to a high bursting probability. The paired apical and basal input amplitudes that achieve this threshold for the T_{ba} and P_2 transfer functions are plotted in Fig 11f. In this same figure, equivalent threshold points (black line) for the T_{lh} function have been found by scaling the firing rate responses shown in Fig 11c to the maximum achieved in each case and then scaling the numbers of apical and basal synaptic inputs to the ranges of input currents of P_2. In this way it can be seen that the data and transfer functions of our model and that of [23] are related, with P_2 and T_{lh} showing a similar non-linear decrease in basal amplitude required to reach the threshold, with increasing apical amplitude. This decline is quite linear for T_{ba}.

Information processing with two input streams

The nature of the information processing carried out by neocortical pyramidal cells that receive anatomically-distinct basal (perisomatic) and apical dendritic inputs from different sources is the subject of many investigations and hypotheses [37]. Our study demonstrates the ability of pyramidal cells that exhibit BAC firing to carry out

contextually-modulated processing of the sensory-derived feedforward basal inputs, where the distinct apical inputs provide the relevant context. Output bursts are the information carriers. A simple orientation-selectivity example illustrates how the appropriate context can greatly amplify the response to the sensory (feedforward) input. The out-of-context response is reduced but still has the same receptive field peak i.e. the conflicting context does not shift this peak.

Information theory has been used to understand what information our transfer functions transmit about the basal and apical input streams. Partial information decomposition reveals that in a regime where sufficiently strong basal input is able to generate a burst, but apical input alone is not, then the presence or not of a burst carries significant information about the strength of the basal input, but little about the apical input (Fig 8 HL). There is some shared and synergistic information. This has been termed the *apical amplification* processing mode [2, 5]. The situation is reversed if the apical input is strong enough to cause a burst by itself, but the basal input is not (Fig 8 LH), which has been termed the *apical drive* mode [5, 8]. In the situation where neither the basal nor apical inputs alone can generate a burst, the output contains significant shared and synergistic information, with some unique information about the basal strength as it contributes to bursting through the generation of an initial sodium spike, even though it cannot generate a burst by itself (Fig 8 LL). We denote this the *apical cooperation* mode. When both inputs are strong enough to generate bursting by themselves (HH) then most information is shared and synergistic, as the probability of bursting is a function of the nonlinear summation of the basal and apical inputs (Fig 8 HH). We denote this the *apical integration* mode.

In the *apical cooperation* mode (low input regime, LL), where the individual inputs do not generate bursts by themselves, [19] demonstrates that single spike and bursting outputs can carry multiplexed information about the basal and apical drives to a pyramidal cell, with different downstream circuits able to readout these different signals. In layered networks, short-term depression in ascending basal connections allows the bottom-up spike rate to be transmitted, irrespective of bursting induced by top-down inputs. Short-term facilitation in descending dendritic connections, on the other hand, allows information on top-down inputs to be transmitted, in the form of burst probability or burst rate. This separation of signals relies on operating in a regime

where basal inputs alone cannot produce bursts.

The *apical integration* mode opens up the possibility of new forms of information processing that have been termed *information modification* [38] or *coding with synergy* [39], in which new information is generated in the output that is not available in either of the input streams individually.

Two-stream signal interaction in cortical pyramidal cells

There is increasing *in vivo* experimental evidence as to the influence of contextual signals on the receptive field responses of cortical pyramidal cells (PCs) . In the visual system, there is considerable evidence that PC receptive field responses can alter in task-dependent ways, indicating contextually-modulated processing [10, 13]. Experiments in primates show that neural responses to visual inputs in their receptive field can be altered by nearby flanking signals that may be relevant or irrelevant to the task at hand, enabling such tasks as contour integration [10]. Our simple network example of lateral amplification provides an indication of a possible mechanism underpinning such effects (Fig 6).

New experimental techniques are now allowing the cellular and network effects and mechanisms of contextual modulation to begin to be unpicked, particularly in the mouse visual system [13]. In rodents, attention and locomotion increase the response amplitude and selectivity of PC receptive fields to visual input [13], such as orientation selectivity [40]. Firing rate increases during locomotion, compared to rest, have been measured in individual PCs, corresponding to both additive and multiplicative changes in receptive fields [12]. Importantly, receptive field orientation preference is not changed, but the increased response amplitude increases orientation discriminability. This corresponds with the effects seen in the example of orientation selectivity with our identified transfer function (Fig 5). In cortex, such effects likely will be mediated by cellular and network-level effects including changes in inhibition and neuromodulation [12, 13], but also through the increased influence of contextual excitatory inputs to PCs on their spiking output. The combination of locomotion and visual input can be a better predictor of PC activity than visual input alone [14]. Further, visual activity has been recorded due to locomotion in the dark, indicating that

contextual (apical) drive can generate output spiking on its own, without corresponding feedforward sensory inputs [14].

Dendritic calcium spiking in layer 5 PCs is correlated with cognition in somatosensory cortex [41]. More specifically, experiments in mice show that calcium spikes in the apical dendrites of layer 5 PCs affect the detection of whisker stimuli in barrel cortex, corresponding with a lowering of stimulus intensity threshold for detection [11]. These experiments showed that blocking of calcium spikes shifts the detection threshold to the right (larger stimulus amplitude), indicating that dendritic spiking modulates cell output to lower the response threshold to sensory input. Further, upregulating excitability in dendrites leads to an increased false detection rate, giving evidence of apical drive. Somatic bursts do correlate with detection though they are less predictive of detection than the overall firing rate [11].

We have identified different operating regimes for layer 5 pyramidal cells that depend on the relative strengths of the excitatory basal and apical input streams. Which regime may be in operation in the behaving animal is determined not purely by the strengths of the excitatory inputs, but also is controlled dynamically by inhibition and neuromodulation. In cortical circuits, the interaction of apical and basal excitatory inputs is strictly controlled through inhibition, mediated by a variety of classes of interneuron that make layer-specific connections onto pyramidal cells and are preferentially driven by top-down, bottom-up or lateral connections [42]. Perisomatic inhibition controls output spiking, with the potential to limit bursting output despite the initiation of dendritic calcium spikes. On the other hand, dendritic calcium spikes are controlled by inhibitory interneurons that target the apical tuft dendrites, affecting calcium spike initiation, or the apical trunk, affecting their transmission to the soma [43, 44]. [45] used detailed computer simulations of a layer 5 pyramidal cell to explore the spatio-temporal effects of inhibition on dendritic spiking and subsequent burst firing.

Excitatory pathways influence their own impact on target pyramidal cells through which classes of inhibitory interneuron they also connect with. Of particular interest is the disinhibition of apical dendrites that promotes the generation of calcium spikes, through interneurons that specifically target other interneurons that inhibit the apical tuft and trunk [42, 46]. This disinhibition is quite focal within neocortical columns [46]

and allows gating of interactions between the basal and apical inputs that can mediate contextually-modulated information transmission [47]. It appears that cross-modal contextual inputs preferentially disinhibit their target pyramidal cells, whereas within the same modality such inputs largely inhibit their target PCs [47, 48].

The influence of apical tuft activity on layer 5 pyramidal cell output strongly depends on the neuromodulatory state of the cell. High levels of acetylcholine and norepinephrine are present in the neocortex during attention and active behaviour in the awake animal. These neuromodulators regulated both pyramidal cell and interneuron activity. In particular, activation of muscarinic acetylcholine receptors in the apical tuft of layer 5 pyramidal cells promotes calcium spiking through upregulation of R-type calcium channels [49]. Adrenergic modulation also increases tuft excitability, putatively through blocking of dendritic HCN channels [50]. General anesthesia acts to decouple the apical tuft from the cell body, putatively through blocking metabotropic glumate and cholinergic receptors in the apical trunk that promote coupling in the awake animal [3, 9].

While neocortical pyramidal cells in general usually have the same basic arrangement of apical and basal inputs, thick-tufted layer 5 pyramidal cells have the most distinct nonlinear interaction between contextual apical and feedforward basal inputs mediated by apical dendritic calcium spikes. However, even within this class of cells, BAC firing and subsequent bursting is limited to the largest cells [51, 52]. Increasing apical trunk length results in greater dendritic compartmentalisation and a greater propensity for generating large calcium spikes that are still able to promote burst firing at the soma. BAC firing is enabled by active, sodium-channel-mediated back propagation of somatic action potentials [52]. Shorter cells only exhibit limited calcium spikes that do not trigger bursts but apical inputs can generate single spike firing in the soma. Shorter cells are more electrically compact which is enhanced by a reduced axial resistance in the apical trunk [51]. These differences correspond with cortical location and its corresponding thickness, with a rostral-caudal gradient from large to small in primary visual cortex [51] and a distinct difference between primary (larger cells) and secondary (smaller cells) visual areas [52] in the rat. At the other extreme, human pyramidal cells are larger and consequently exhibit even greater compartmentalisation of the distal apical dendrites, to the extent that calcium spiking in the dendrites does not lead to

bursting in the soma [53]. Thus two-stream signal interaction in pyramidal cells takes a variety of forms depending on cortical area and animal species.

Issues arising

The use of tightly controlled stimuli here has allowed for PID analysis and transfer function fitting in well defined operating regimes. This opens the door for exploring more naturalistic settings. In particular, here the operating regimes are defined according to the amplitudes of the input stimuli, with no other changes to intrinsic cell properties and with no inhibitory inputs. In the whole animal, the effect of particular strengths of excitatory input will be determined by spatio-temporal patterns of inhibition and changes to cell membrane properties through neuromodulation. These effects can be modelled explicitly to demonstrate if our identified information processing modes, corresponding to different operating regimes, do appear in particular behavioural states, ranging from active wakefulness to dreaming sleep [5, 7, 8].

It has previously been conjectured that the apical amplification mode is primary to cognition in the awake, behaving animal [2], whereas apical drive comes to the fore in dreaming sleep [8]. From our analysis here, we would also conjecture that our mode of apical integration captures a mechanism underpinning the formation of thoughts and imagination in the awake animal [5].

The effect of apical input may be controlled in different ways that may seem similar from the cell output point of view, but have distinct effects within the cell. Inhibition and neuromodulation may limit the summed strength of apical tuft synaptic inputs, or the ability of these inputs to generate a large calcium spike, or the transmission of this spike to the cell body [45]. All of these will limit the contribution of the apical input stream to bursting output. Dendritic inhibition of synaptic input is closest to the simple change in amplitude we have studied here. Limiting transmission of the calcium spike to the cell body, on the other hand, while still allowing a dendritic calcium spike, may be vital for enabling synaptic plasticity in the apical tuft, while preventing its contribution to cell output [45]. We have not considered synaptic plasticity here, but it is of fundamental importance to establishing information storage and transmission in which regenerative activity in the apical dendrites and bursting may play a significant

part [54, 55].

As indicated above, the nature of the interaction between basal and apical inputs is very much pyramidal cell-type specific. The approach we have taken here, using a combination of transfer function fitting and information theory can be used to characterise other classes of pyramidal cell and provide a comprehensive picture of contextually-modulated information processing across the neocortex.

Conclusion

Partial information decomposition and transfer function fitting have been used to characterise the input-output properties of burst firing in a stochastic model of a layer 5 neocortical pyramidal cell that can exhibit BAC firing [15]. It is revealed that the cell can operate in different information processing modes, depending on the amplitude ranges of the basal and apical input streams. Highlighting the contribution of the apical stream, these modes have been termed apical amplification, apical drive [2, 5, 8], apical cooperation and apical integration. Different modes are plausibly obtained *in vivo* through the activation of targeted inhibitory pathways and network neuromodulation. The encompassing theme of these modes is contextually-modulated information processing in which contextual apical inputs from diverse brain regions refine signal transmission of feedforward sensory inputs.

Materials and methods

Model pyramidal cell

Simulations of probabilistic spike firing were run using the [15] reduced 20-compartment model of a layer 5 pyramidal cell. We used model 2 of [15], which can generate backpropagation-activated calcium (BAC) spike firing. Model code for the NEURON simulator (*neuron.yale.edu*) was obtained from ModelDB (*modeldb.science*). The compartmental structure and active biophysics of the model are outlined in Table 4. A cell schematic is shown in Fig 1. Full details of the parameter values are available in [15] and in the code itself.

Table 4. Structure and active biophysics of the Bahl model. Cpts - number of numerical compartments; L - length (μm); Diam - diameter (μm); Nat - transient sodium current; Kfast - fast potassium; Kslow - slow potassium; Nap - persistent sodium; Km - muscarinic potassium; HCN - hyperpolarization-activated cation; Cas - slow calcium; KCa - calcium-activated potassium; CP - calcium pump.

Section	Cpts	L	Diam	Nat	Kfast	Kslow	Nap	Km	HCN	Cas	KCa	CP
soma	1	23.1	23.1	x	x	x	x	x				
basal dendrite	1	257	8.7						x			
apical dendrite	5	500	5.9	x	x	x			x			
apical tuft	2	499	6	x	x	x			x	x	x	x
axon hillock	5	20	3.5-2	x								
initial segment	5	25	2-1.5	x								
axon	1	500	1.5									

Model stimuli

Random background synaptic activity was modelled as a noisy current described by an Ornstein-Uhlenbeck (OU) process of the form used by [56]. Our NEURON *nmodl* code was adapted from the synaptic conductance code of [57]. The current waveform is given by:

$$I(t + dt)) = I(t) + (\mu - I(t))(dt/\tau) + \sigma G_t \sqrt{2dt/\tau}, \tag{26}$$

with mean current $\mu = 0$ nA, standard deviation $\sigma = 0.1$ nA and correlation length $\tau = 3$ ms. G_t is a Gaussian random number with mean 0 and a standard deviation of 1, chosen at each sample time $dt = 25\mu$s. Independent noisy waveforms were injected into the middle of the soma and apical tuft sections.

The noisy currents themselves produced a very low probability of somatic spiking. Defined stimuli that could produce spiking were given simultaneously in the form of a short square-wave current pulse to the soma and an EPSP-like dual exponential current waveform to the apical tuft (rise time 0.5 ms; decay time 5 ms). Combinations of such stimuli of sufficient amplitude could produce BAC-firing, as in the experimental work of Larkum [1, 58, 59] and in pyramidal cell models [15, 60].

Simulations of stochastic spiking

Simulations of a single pyramidal cell with particular basal and apical stimulus amplitudes and time course were repeated 100 times with different noisy currents each time to generate a record of stochastic responses. Simulations were for 250 ms with the defined stimuli applied after 100 ms. Basal amplitudes were varied from 0 to 3 nA in 0.1

nA increments, for step durations of 2 or 5 ms, and from 0 to 1 nA in 0.05 nA increments for a step duration of 10 ms. Apical amplitudes were varied from 0 to 1 nA in increments of 0.1 nA, always with an EPSP rise time of 0.5 ms and fall time of 5 ms. These amplitude ranges were chosen so that strong basal input alone could produce spiking output, whereas apical input alone could not. During each simulation, somatic spikes were counted, based on a somatic voltage threshold of -25 mV, from the onset of the defined stimuli to the end of the simulation. During analysis, groups of spikes recorded with interspike intervals of less than 25 ms were counted as bursts. Burst probability for given defined stimuli amplitudes was calculated as the fraction of the 100 simulations in which a burst occurred.

Information theory notation and definitions

We consider a trivariate probabilistic system involving three discrete random variables: an output Y and two inputs B and A. Hence, underlying the discrete data sets we consider is a probability mass function $\Pr(Y = y, B = b, A = a)$, where y, b, a belong the the finite alphabets $\mathcal{A}_y, \mathcal{A}_b, \mathcal{A}_a$, respectively.

The alphabets $\mathcal{A}_b$ and $\mathcal{A}_a$ consist of the sets of discrete basal and apical amplitudes used in the cell simulations. Each basal and apical input amplitude is treated as equally probable. The output spike count was categorised into two categories as 0-1 (no burst) and 1+ (burst) to allow calculation of burst probabilities.

In the first analysis with different durations of basal input, three categorical distributions are considered, of sizes $31 \times 11 \times 2$ for each of the 2 ms and 5 ms durations of basal input, and $21 \times 11 \times 2$ for the 10 ms duration. The probabilities in each case are computed for each combination of basal input, apical input and spike count output as the number of occurrences of this combination divided by the product of the number of basal input amplitudes, the number of apical input amplitudes and 100.

In the analysis of the extended data set for a basal duration of 10 ms, we consider basal and apical amplitudes in steps of 0.1 nA up to their range limits for each of four operating regimes. The probability distributions for the four regimes are of size:
$6 \times 11 \times 2$ when basal and apical amplitudes are both low (range 0 to 1 nA for both);
$6 \times 18 \times 2$ when basal amplitude is low and apical amplitude is high (basal 0-0.5 nA;

apical 0-1.7 nA); $11 \times 11 \times 2$ for high basal amplitude and low apical amplitude (basal 0-1.0 nA; apical 0-1.0 nA); $11 \times 18 \times 2$ when both amplitudes are high (basal 0-1.0 nA; apical 0-1.7 nA).

Classical information theory

We now define the standard information theoretic terms that are required in this work and they are based on results in [61]. We denote by the function H the usual Shannon entropy, and note that any term with zero probabilities makes no contribution to the sums involved. The joint mutual information that is shared by Y and the pair (B, A) is given by,

$$I(Y; B, A) = H(Y) + H(B, A) - H(Y, B, A). \tag{27}$$

The information that is shared between Y and B but not with A is

$$I(Y; B|A) = H(Y, A) + H(B, A) - H(A) - H(Y, B, A), \tag{28}$$

and the information that is shared between Y and A but not with B is

$$I(Y; A|B) = H(Y, B) + H(B, A) - H(B) - H(Y, B, A). \tag{29}$$

The information shared between Y and B is

$$I(Y; B) = H(Y) + H(B) - H(Y, B) \tag{30}$$

and between Y and A is

$$I(Y; A) = H(Y) + H(A) - H(Y, A) \tag{31}$$

The interaction information [62] is a measure of information involving all three variables, Y, B, A and is defined by

$$II(Y; B; A) = I(Y; B, A) - I(Y; B) - I(Y; A) \tag{32}$$

McGill's interaction information has been used as a measure of synergy [63], with a

positive value indicating the presence of synergy and a negative value indicating redundancy. See also [64]. The negative of McGill's measure has been termed *coinformation* [65], and it has been used as an objective function in an artificial neural network possessing two distinct sites of integration [32]. In the case of three variables, the negative of McGill's measure is a special case of the general O-information measure [66]. The O-information measure can be particularly useful with systems which have a large number of interacting variables [67].

Partial Information Decomposition

The output entropy, $H(Y)$, may be written as

$$H(Y) = I(Y; B, A) + H(Y)_{res},$$

where $H(Y)_{res}$ is the residual output entropy.

The decomposition of the joint mutual information can be expressed as [39]:

$$I(Y; B, A) = I_{unq}(Y; B|A) + I_{unq}(Y; A|B) + I_{shdS}(Y; B, A) + I_{shdM}(Y; B, A) + I_{syn}(Y; B, A), \tag{33}$$

where the shared information $I_{shd}(Y; B, A)$ has been split into two separate components of source shared or mechanistic shared information [68] as:

$$I_{shd}(Y; B, A) = I_{shdS}(Y; B, A) + I_{shdM}(Y; B, A).$$

Adapting the notation of [39] we express our joint input mutual information in four terms as follows:

For an excellent tutorial on information theory and partial information decomposition, with illustrations from neuroscience, see [69]. It is possible to make deductions about a PID by using the following four equations which give a link between the components of a PID and certain classical Shannon measures of mutual information.

$\text{Unq}B \equiv I_{unq}(Y;B\|A)$	denotes the unique information that B conveys about Y;
$\text{Unq}A \equiv I_{unq}(Y;A\|B)$	is the unique information that A conveys about Y;
$\text{ShdS} \equiv I_{shdS}(Y;B,A)$	gives the shared (or redundant) information that both B and A have about Y, due to the correlation between A and B. This is termed source shared information.
$\text{ShdM} \equiv I_{shdM}(Y;B,A)$	gives the shared (or redundant) information that both B and A have about Y, due to the probabilistic mechanism. This is termed mechanistic shared information.
$\text{Syn} \equiv I_{syn}(Y;B,A)$	is the synergy or information that the joint variable (B,A) has about Y that cannot be obtained by observing B and A separately.

The following are in [39] (eqs. 4, 5 with amended notation); see also [25].

$$I(Y;B) = \text{Unq}B + \text{ShdS} + \text{ShdM} \tag{34}$$

$$I(Y;A) = \text{Unq}A + \text{ShdS} + \text{ShdM}, \tag{35}$$

$$I(Y;B|A) = \text{Unq}B + \text{Syn}, \tag{36}$$

$$I(Y;A|B) = \text{Unq}A + \text{Syn}. \tag{37}$$

Using (33), (34), (35) we may deduce the following connections between classical information measures and partial information components:

$$II(Y;B;A) = \text{Syn} - \text{ShdS} - \text{ShdM} \tag{38}$$

$$I(Y;B) - I(Y;A) = \text{UnqB} - \text{UnqA} \tag{39}$$

The term in (39) defines the unique information asymmetry. A positive value suggests that the information processing is being driven mainly by input B, whereas a negative value suggests that input A provides more drive.

When the partial information components are known *a priori* to be non-negative, we may deduce the following from (27), (34), (35). When the interaction information in (32) is positive, a lower bound on the synergy of a system is given by the interaction information [62]. Also, the expression in (39) provides a lower bound for UnqB, when

$I(Y;B) > I(Y;A)$. Thus some deductions can be made without considering a PID. While such deductions can be useful in providing information bounds, it is only by computing a PID that the actual values of the partial information components can be obtained.

When making comparisons between different systems it is sometimes necessary to normalise the information decomposition by dividing each term by their sum, the output entropy, $H(Y)$. Such normalisation will be applied in the analyses when comparing decompositions obtained under different conditions of an experimental or grouping factor. For probability distributions in which the inputs B and A are marginally independent the source shared information, ShdS, should be equal to zero, and hence the total shared information is entirely mechanistic shared information - shared information due to the probabilisitic mechanism involved in the information processing.

Software for PID

The Ibroja PID was estimated using compute UI [70]. The discrete information theory library dit [71] was used to estimate the Imin and Iproj PIDs. R [72] code was also used to estimate the Idep PID. Python code was called from RStudio [72] by using the reticulate package [73]. The graphics were produced by using the ggplot2 package [74] in RStudio.

Transfer function fitting

Burst probabilities were extracted from the simulation data in the form of burst frequencies over 100 repetitions at each basal and apical current amplitude. Bursts were counted as groups of spikes separated by interspike intervals of at most 25 ms. The transfer functions were fit to this data by a least-squares fitting procedure using the SciPy (*scipy.org*) `least_squares` function. Optimisation proceded simultaneously against all available basal and apical strengths for a particular duration of basal input (which was either 2, 5 or 10 ms). For each optimisation step, the burst probability as a function of basal amplitude when there was no apical input was calculated as $P_{2b}(b)$ (Eq 7). The resulting values were then used in the calculation of the full transfer function, $P_2(b,a) = P_{1b}(b)[P_{2a}(a)(1 - P_{2b}(b)) + P_{2b}(b)]$ (Eq 9), for all apical amplitudes

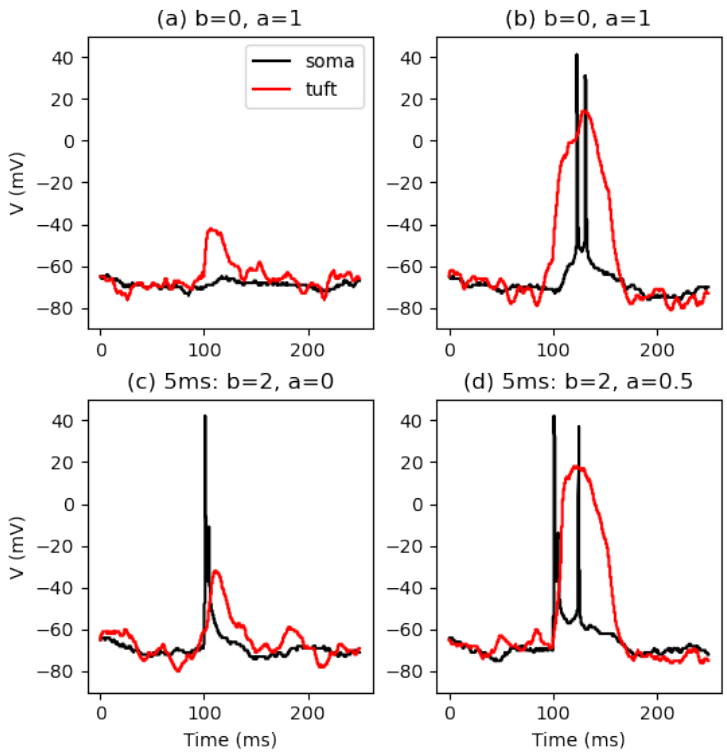
greater than zero but in the low range, leading to the parameterisation of $P_{1b}(b)$ (Eq 6) and $P_{2a}(a)$ (Eq 8) and also allowing the calculation of the transfer function P_2^{LL} (Eq 17).

For apical input in the high range, the burst firing probability $P_{2a}^{*}(a)$ (Eq 15) was obtained by fitting to the full range of apical amplitudes when the basal input is zero. This function was then used to complete the transfer functions P_2^{LH} (Eq 19) and P_2^{HH} (Eq 20).

Standard errors in the parameter values were estimated from the Jacobian matrix and residuals. Fitting quality was also tested by refitting using weighted least squares and binomial nonlinear regression, both of which produced similar results.

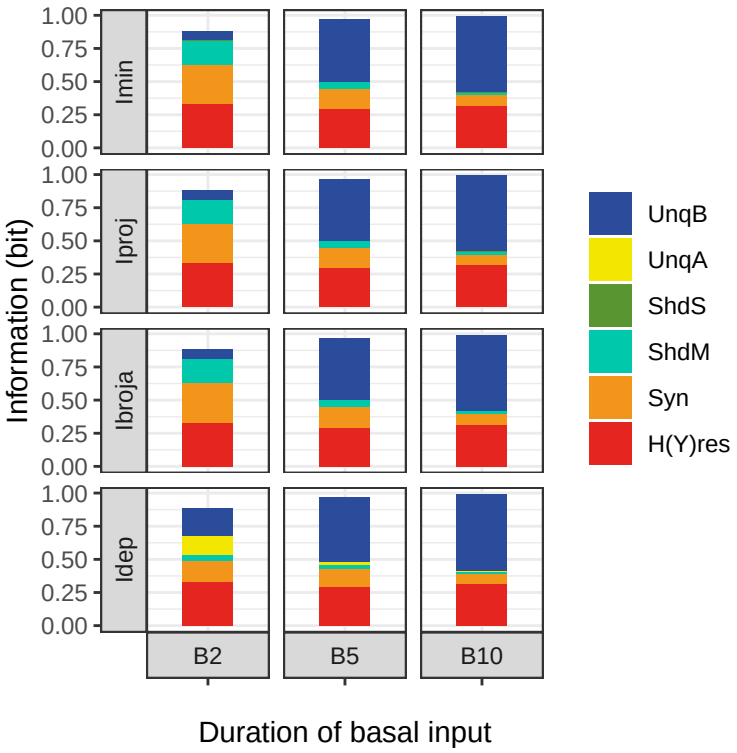
Supporting information

S1 Fig. Further examples of spiking responses. (a) No basal input, apical input a=1 nA: no burst; (b) No basal, a=1 nA: burst occurs after a time delay; (c) Basal duration 5 ms: b=2, a=0; (d) Basal duration 5 ms: b=2, a=0.5.

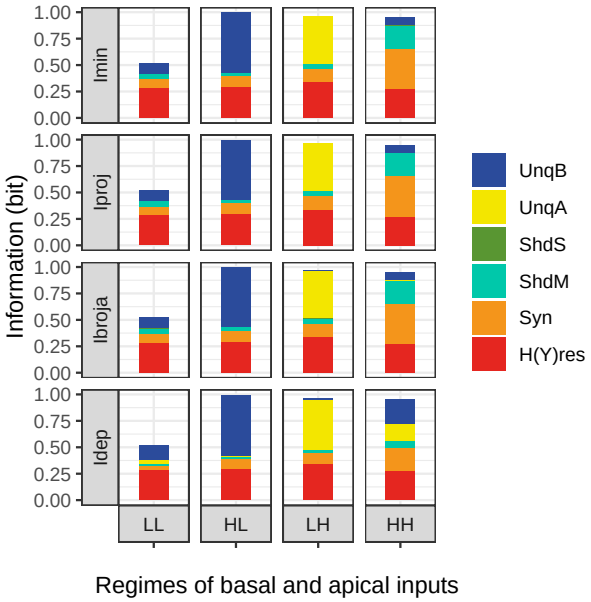


S2 Fig. Partial information decompositions for different basal durations. Partial information decompositions obtained by using four different methods (Imin, Iproj, Ibroja, Idep) on burst probability data for apical inputs up to 1 nA and basal inputs up to 3 nA, for each of three durations of basal input: 2 ms (B2), 5 ms (B5), 10 ms (B10).

S3 Fig. Partial information decompositions for different regimes. Partial information decompositions obtained by using four different methods (Imin, Iproj, Ibroja, Idep) for each of the four regimes of basal and apical inputs with a basal duration of 10 ms. 'L' is an abbreviation of 'low', and 'H' is an abbreviation of 'high', so for example, 'HL' denotes the regime in which basal input is high and apical input is


Information (bit)
Imin
Iproj
Ibroja
Idep
1.00
0.75
0.50
0.25
0.00
B2
B5
B10
Duration of basal input
UnqB
UnqA
ShdS
ShdM
Syn
H(Y)res

low.


Information (bit)
Imin
Iproj
Ibroja
Idep
1.00
0.75
0.50
0.25
0.00
LL
HL
LH
HH
Regimes of basal and apical inputs
UnqB
UnqA
ShdS
ShdM
Syn
H(Y)res

References

1. Larkum M. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. Trends in neurosciences. 2013;36(3):141–151.

2. Phillips WA. Cognitive functions of intracellular mechanisms for contextual amplification. Brain and Cognition. 2017;112:39–53. doi:https://doi.org/10.1016/j.bandc.2015.09.005.

3. Phillips WA, Bachmann T, Storm JF. Apical function in neocortical pyramidal cells: a common pathway by which general anesthetics can affect mental state. Frontiers in neural circuits. 2018;12:50.

4. Marvan T, Polák M, Bachmann T, Phillips WA. Apical amplification—a cellular mechanism of conscious perception? Neuroscience of consciousness. 2021;2021(2):niab036.

5. Phillips WA. The Cooperative Neuron: Cellular Foundations of Mental Life. Oxford University Press; 2023.

6. Phillips W, Clark A, Silverstein SM. On the functions, mechanisms, and malfunctions of intracortical contextual modulation. Neuroscience & Biobehavioral Reviews. 2015;52:1–20.

7. Aru J, Suzuki M, Larkum ME. Cellular mechanisms of conscious processing. Trends in Cognitive Sciences. 2020;24(10):814–825.

8. Aru J, Siclari F, Phillips WA, Storm JF. Apical drive—A cellular mechanism of dreaming? Neuroscience & Biobehavioral Reviews. 2020;119:440–455.

9. Suzuki M, Larkum ME. General anesthesia decouples cortical pyramidal neurons. Cell. 2020;180(4):666–676.

10. Gilbert CD, Li W. Top-down influences on visual processing. Nature Reviews Neuroscience. 2013;14(5):350–363.

11. Takahashi N, Oertner TG, Hegemann P, Larkum ME. Active cortical dendrites modulate perception. Science. 2016;354(6319):1587–1590.

12. Dadarlat MC, Stryker MP. Locomotion enhances neural encoding of visual stimuli in mouse V1. Journal of Neuroscience. 2017;37(14):3764–3775.

13. Pakan JM, Francioni V, Rochefort NL. Action and learning shape the activity of neuronal circuits in the visual cortex. Current opinion in neurobiology. 2018;52:88–97.

14. Keller GB, Bonhoeffer T, Hübener M. Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. Neuron. 2012;74(5):809–815.

15. Bahl A, Stemmler MB, Herz AVM, Roth A. Automated optimization of a reduced layer 5 pyramidal cell model based on experimental data. Journal of Neuroscience Methods. 2012;210(1):22–34. doi:https://doi.org/10.1016/j.jneumeth.2012.04.006.

16. Lisman JE. Bursts as a unit of neural information: making unreliable synapses reliable. Trends in neurosciences. 1997;20(1):38–43.

17. Krahe R, Gabbiani F. Burst firing in sensory systems. Nature Reviews Neuroscience. 2004;5(1):13–23.

18. Zeldenrust F, Wadman WJ, Englitz B. Neural coding with bursts—current state and future perspectives. Frontiers in computational neuroscience. 2018;12:48.

19. Naud R, Sprekeler H. Sparse bursts optimize information transmission in a multiplexed neural code. Proceedings of the National Academy of Sciences. 2018;115(27):E6329–E6338.

20. Oswald AMM, Chacron MJ, Doiron B, Bastian J, Maler L. Parallel processing of sensory input by bursts and isolated spikes. Journal of Neuroscience. 2004;24(18):4351–4362.

21. Williams E, Payeur A, Gidon A, Naud R. Neural burst codes disguised as rate codes. Scientific Reports. 2021;11(1):15910.

22. Friedenberger Z, Harkin E, Tóth K, Naud R. Silences, spikes and bursts: Three-part knot of the neural code. The Journal of Physiology. 2023;.

23. Shai AS, Anastassiou CA, Larkum ME, Koch C. Physiology of Layer 5 Pyramidal Neurons in Mouse Primary Visual Cortex: Coincidence Detection

through Bursting. PLoS Comput Biol. 2015;11(3):e1004090. doi:https://doi.org/10.1371/journal.pcbi.1004090.

24. Kay JW, Schulz JM, Phillips WA. A comparison of partial information decompositions using data from real and simulated layer 5b pyramidal cells. Entropy. 2022;24(8):1021.

25. Williams PL, Beer RD. Nonnegative decomposition of multivariate information. arXiv preprint arXiv:10042515. 2010;.

26. Harder M, Salge C, Polani D. Bivariate measure of redundant information. Physical Review E. 2013;87(1):012130.

27. Bertschinger N, Rauh J, Olbrich E, Jost J, Ay N. Quantifying unique information. Entropy. 2014;16(4):2161–2183.

28. Griffith V, Koch C. Quantifying synergistic mutual information. In: Guided self-organization: Inception. Springer; 2014. p. 159–190.

29. James RG, Emenheiser J, Crutchfield JP. Unique information via dependency constraints. Journal of Physics A: Mathematical and Theoretical. 2018;52(1):014002.

30. Kay JW, Ince RA, Dering B, Phillips WA. Partial and entropic information decompositions of a neuronal modulatory interaction. Entropy. 2017;19(11):560.

31. Kay JW, Phillips WA. Contextual modulation in mammalian neocortex is asymmetric. Symmetry. 2020;12(5):815.

32. Phillips W, Kay J, Smyth D. The discovery of structure by multi-stream networks of local processors with contextual guidance. Network: Computation in neural systems. 1995;6(2):225.

33. Kay J, Phillips WA. Activation functions, computational goals, and learning rules for local processors with contextual guidance. Neural Computation. 1997;9(4):895–910.

34. Kay J, Floreano D, Phillips WA. Contextually guided unsupervised learning using local multivariate binary processors. Neural Networks. 1998;11(1):117–140.

35. Kay JW, Phillips W. Coherent Infomax as a computational goal for neural systems. Bulletin of mathematical biology. 2011;73:344–372.

36. Kay JW, Phillips WA, Aru J, Graham BP, Larkum ME. Bayesian modeling of BAC firing as a mechanism for apical amplification in neocortical pyramidal neurons. bioRxiv preprint bioRxiv:604066. 2019;.

37. Payeur A, Béïque JC, Naud R. Classes of dendritic information processing. Current opinion in neurobiology. 2019;58:78–85.

38. Wibral M, Finn C, Wollstadt P, Lizier JT, Priesemann V. Quantifying information modification in developing neural networks via partial information decomposition. Entropy. 2017;19(9):494.

39. Wibral M, Priesemann V, Kay JW, Lizier JT, Phillips WA. Partial information decomposition as a unified approach to the specification of neural goal functions. Brain and cognition. 2017;112:25–38.

40. Niell CM, Stryker MP. Modulation of visual responses by behavioral state in mouse visual cortex. Neuron. 2010;65(4):472–479.

41. Xu Nl, Harnett MT, Williams SR, Huber D, O'Connor DH, Svoboda K, et al. Nonlinear dendritic integration of sensory and motor input during an active sensing task. Nature. 2012;492(7428):247–251.

42. Gentet LJ. Functional diversity of supragranular GABAergic neurons in the barrel cortex. Frontiers in neural circuits. 2012;6:52.

43. Murayama M, Pérez-Garci E, Nevian T, Bock T, Senn W, Larkum ME. Dendritic encoding of sensory stimuli controlled by deep cortical interneurons. Nature. 2009;457(7233):1137–1141.

44. Schulz JM, Kay JW, Bischofberger J, Larkum ME. GABA B receptor-mediated regulation of dendro-somatic synergy in layer 5 pyramidal neurons. Frontiers in cellular neuroscience. 2021;15:718413.

45. Leleo EG, Segev I. Burst control: synaptic conditions for burst generation in cortical layer 5 pyramidal neurons. PLoS Computational Biology. 2021;17(11):e1009558.

46. Jiang X, Wang G, Lee AJ, Stornetta RL, Zhu JJ. The organization of two new cortical interneuronal circuits. Nature neuroscience. 2013;16(2):210–218.

47. Wang XJ, Yang GR. A disinhibitory circuit motif and flexible information routing in the brain. Current opinion in neurobiology. 2018;49:75–83.

48. Shen S, Jiang X, Scala F, Fu J, Fahey P, Kobak D, et al. Distinct organization of two cortico-cortical feedback pathways. Nature Communications. 2022;13(1):6389.

49. Williams SR, Fletcher LN. A dendritic substrate for the cholinergic control of neocortical output neurons. Neuron. 2019;101(3):486–499.

50. Labarrera C, Deitcher Y, Dudai A, Weiner B, Amichai AK, Zylbermann N, et al. Adrenergic modulation regulates the dendritic excitability of layer 5 pyramidal neurons in vivo. Cell reports. 2018;23(4):1034–1044.

51. Fletcher LN, Williams SR. Neocortical topology governs the dendritic integrative capacity of layer 5 pyramidal neurons. Neuron. 2019;101(1):76–90.

52. Galloni AR, Laffere A, Rancz E. Apical length governs computational diversity of layer 5 pyramidal neurons. Elife. 2020;9:e55761.

53. Beaulieu-Laroche L, Toloza EH, Van der Goes MS, Lafourcade M, Barnagian D, Williams ZM, et al. Enhanced dendritic compartmentalization in human cortical neurons. Cell. 2018;175(3):643–651.

54. Payeur A, Guerguiev J, Zenke F, Richards BA, Naud R. Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. Nature neuroscience. 2021;24(7):1010–1019.

55. Francioni V, Harnett MT. Rethinking single neuron electrical compartmentalization: dendritic contributions to network computation in vivo. Neuroscience. 2022;489:185–199.

56. Larkum ME, Senn W, Lüscher HR. Top-down dendritic input increases the gain of layer 5 pyramidal neurons. Cerebral cortex. 2004;14(10):1059–1070.

57. Destexhe A, Rudolph M, Fellous JM, Sejnowski TJ. Fluctuating synaptic conductances recreate in vivo-like activity in neocortical neurons. Neuroscience. 2001;107(1):13–24.

58. Larkum ME, Zhu JJ, Sakmann B. A new cellular mechanism for coupling inputs arriving at different cortical layers. Nature. 1999;398(6725):338–341.

59. Larkum ME, Zhu JJ, Sakmann B. Dendritic mechanisms underlying the coupling of the dendritic with the axonal action potential initiation zone of adult rat layer 5 pyramidal neurons. The Journal of physiology. 2001;533(2):447–466.

60. Hay E, Hill S, Schürmann F, Markram H, Segev I. Models of neocortical layer 5b pyramidal cells capturing a wide range of dendritic and perisomatic active properties. PLoS computational biology. 2011;7(7):e1002107.

61. Cover TM, Thomas JA. Elements of information theory. John Wiley and Sons New York; 1991.

62. McGill W. Multivariate information transmission. Transactions of the IRE Professional Group on Information Theory. 1954;4(4):93–111.

63. Schneidman E, Bialek W, Berry MJ. Synergy, redundancy, and independence in population codes. Journal of Neuroscience. 2003;23(37):11539–11553.

64. Gat I, Tishby N. Synergy and redundancy among brain cells of behaving monkeys. Advances in neural information processing systems. 1998;11:111–117.

65. Bell AJ. The co-information lattice. In: Proceedings of the fifth international workshop on independent component analysis and blind signal separation: ICA. vol. 2003; 2003.

66. Rosas FE, Mediano PA, Gastpar M, Jensen HJ. Quantifying high-order interdependencies via multivariate extensions of the mutual information. Physical Review E. 2019;100(3):032305.

67. Varley TF, Pope M, Faskowitz J, Sporns O. Multivariate information theory uncovers synergistic subsystems of the human cerebral cortex. Communications biology. 2023;6(1):451.

68. Pica G, Piasini E, Chicharro D, Panzeri S. Invariant components of synergy, redundancy, and unique information among three variables. Entropy. 2017;19(9):451.

69. Timme NM, Lapish C. A Tutorial for Information Theory in Neuroscience. eNeuro. 2018;5(3). doi:10.1523/ENEURO.0052-18.2018.

70. Banerjee PK, Rauh J, Montúfar G. Computing the unique information. In: 2018 IEEE International Symposium on Information Theory (ISIT). IEEE; 2018. p. 141–145.

71. James RG, Ellison CJ, Crutchfield JP. dit: a Python package for discrete information theory. The Journal of Open Source Software. 2018;3(25):738. doi:https://doi.org/10.21105/joss.00738.

72. R Core Team, et al.. R: a language and environment for statistical computing.; 2021. Available from: `https://www.R-project.org/`.

73. Ushey K, Allaire J, Tang Y. Reticulate: interface to 'Python'. R package version 1.16; 2020.

74. Wickham H. ggplot2: Elegant Graphics for Data analysis. Springer; 2016.