Sparse bursts optimize information transmission in a multiplexed neural code

Richard Naud[a,b,1] and Henning Sprekeler[c,d]

[a]University of Ottawa Brain and Mind Research Institute, Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON K1H 8M5, Canada; [b]Department of Physics, University of Ottawa, Ottawa, ON K1N 6N5, Canada; [c]Bernstein Center for Computational Neuroscience Berlin, 10115 Berlin, Germany; and [d]Modelling of Cognitive Processes, Institute of Software Engineering and Theoretical Computer Science, Technische Universität Berlin, 10587 Berlin, Germany

Edited by Terrence J. Sejnowski, Salk Institute for Biological Studies, La Jolla, CA, and approved May 21, 2018 (received for review December 4, 2017)

Many cortical neurons combine the information ascending and descending the cortical hierarchy. In the classical view, this information is combined nonlinearly to give rise to a single firing-rate output, which collapses all input streams into one. We analyze the extent to which neurons can simultaneously represent multiple input streams by using a code that distinguishes spike timing patterns at the level of a neural ensemble. Using computational simulations constrained by experimental data, we show that cortical neurons are well suited to generate such multiplexing. Interestingly, this neural code maximizes information for short and sparse bursts, a regime consistent with in vivo recordings. Neurons can also demultiplex this information, using specific connectivity patterns. The anatomy of the adult mammalian cortex suggests that these connectivity patterns are used by the nervous system to maintain sparse bursting and optimal multiplexing. Contrary to firing-rate coding, our findings indicate that the physiology and anatomy of the cortex may be interpreted as optimizing the transmission of multiple independent signals to different targets.

neural coding | cerebral cortex | multiplexing | dendritic computation | short-term plasticity

Visual, auditory, and motor processing in the mammalian brain is organized in a hierarchy (1–5). At the bottom of this hierarchy, ensembles of neurons code a dense array of simple features such as local visual contrast or simple movement components. At the top of the hierarchy neurons code more complex features such as complex images and movement sequences. Given that information travels both up and down the hierarchy with the power to drive or modulate responses (6–9), we are compelled to an important question: How do populations that receive both bottom–up and top–down information process these two different types of messages?

Experimental observations argue for several opposing views. In one view, descending inputs modulate the bottom–up responses (7). In a second view, top–down inputs can create responses de novo (8). A third view arises from conceptual requirements. In the theories of unsupervised learning, the same units must simultaneously communicate feature recognition to higher-order units and a feature prediction to lower-order units (10, 11). In supervised learning, the higher-order success signal must percolate down the hierarchy, requiring units to communicate both the credit residual from top to bottom and an activation from bottom to top (12, 13). Also, in the binding problem, neurons are required to simultaneously signal the presence of a lower-order feature and its binding to a high-order one, across modalities (14–16). Hence the third view is that of multiplexing: The same population needs to communicate different functions of ascending and descending information, simultaneously and to possibly different target neurons.

Present neural mechanisms for multiplexing can be separated into three different categories. First, spike-phase multiplexing (15, 17) posits that a population represents bottom–up information by its firing rate and top–down information by the timing of its spikes with respect to distinct frequency bands of a local field potential. This type of frequency-division multiplexing (18) is supported by multiple experimental studies in different systems (15, 17), but the cellular mechanisms for encoding and decoding with the local field potential remain to be fully articulated. A second mechanism exploits the difference between time-averaged firing rate in a single neuron and ensemble-averaged firing rate. Distinct features of the sensory input can be represented in output features as shown in the auditory brainstem (19), in the olfactory bulb (20), and in the visual system of flies (21). Here, however, it is distinct features of the same stimulus that are represented in distinct features of the neural response, so the mechanism is closer to the concept of information synergy (22) than to multiplexing. A third possibility is to allow the neurons to alternate between different modes: one devoted to the transmission of ascending information and another for the propagation of descending information. Time-division multiplexing of this kind is common in artificial neural networks (10, 12). In a similar fashion, time-division multiplexing is a useful mechanism in computational models of synaptic plasticity (13, 23, 24), where the population alternates between sensing and learning phases. Yet it is not clear how time-division multiplexing can be mapped on the ongoing activity of cortical networks (9, 23).

In this article we propose a type of multiplexing based on the separation of bursts and single spikes at the level of an ensemble. Burst coding, we suggest, acts on the level of an ensemble to represent multiple information streams simultaneously and

Significance

Understanding the neural code is to attribute proper meaning to temporal sequences of action potentials. We report a simple neural code based on distinguishing single spikes from spikes in close succession, commonly called "bursts." By separating these two types of responses, we show that ensembles of neurons can communicate rapidly changing and graded information from two sources simultaneously and with minimal cross-talk. Second, we show that this multiplexing can optimize the information transferred per action potential when bursts are relatively rare. Finally, we show that neurons can demultiplex these two streams of information. We propose that this multiplexing may be particularly important in hierarchical communication where bottom–up and top–down information must be distinguished.

Author contributions: R.N. and H.S. designed research; R.N. performed research; R.N. analyzed data; and R.N. and H.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: rnaud@uottawa.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1720995115/-/DCSupplemental.

Published online June 22, 2018.

without ambiguity. We study this idea in the thick-tufted pyramidal neurons (TPNs) as a paradigmatic cell type that receives both bottom–up and top–down signals. Using computational simulations, we show that TPNs can encode two independent streams of information with high temporal precision. The two streams can be decoded by postsynaptic populations using short-term plasticity and disynaptic inhibition. A theoretical analysis demonstrates that information representation is optimal for short and sparse bursts, a regime consistent with bursting in vivo. We further show that this optimal regime can be preserved by a network architecture that shares interesting parallels with the anatomy of dendritic feedback inhibition in the cortex. Finally, we discuss how the proposed burst ensemble multiplexing (BEM) code could allow the central nervous system to distinguish ascending information from descending information and hence suggests a different approach to their analysis.

Results

We consider a neural code where spike-timing patterns—single spikes and bursts—are separated at the level of individual spike trains before being averaged across a neural ensemble (Fig. 1*A*). From classical studies on the firing rate (21, 26, 27), we expect that the resulting time-varying rates of single spikes and bursts can be related to the time-averaged rates, but for time-varying stimuli they are generally not equal. How could rates of distinct spike-timing patterns represent different input streams or features? The simpler variant would be that single spikes and bursts are generated by two independent cellular mechanisms that each depend on one input stream alone. In this case, the ensemble singlet rate and ensemble burst rate would encode these streams independently. This possibility has been explored in the context of single-cell firing of the thalamus (14, 28), hippocampus (29), cortex (30, 31), and electrosensory lateral lobe (ELL) (32–34).

Alternatively, bursts could be generated by a synergy of the two input streams, namely conjunctive bursting. Cellular and molecular mechanisms for burst firing in the thalamus (14, 35), the superficial ELL (34), L2–3 pyramidals (36), CA1 pyramidals (37, 38), and TPNs (30) can be said to burst in response to a conjunction of distinct streams of information. Since in this case both singlet rate and burst rate represent a mixture of the two input streams, contrasting singlet and burst rates is not likely to reveal independent information.

In TPNs, dendritic spikes convert a somatically induced singlet into a burst via the activation of a calcium spike in the dendrites (30). Therefore, we reason that, in TPNs, a dendritic input stream is represented by the probability that a somatically induced spike is converted into a burst (33). On the ensemble level, this burst probability, F, is reflected by the fraction of active cells that emit a burst (Fig. 1*A*). Then, a somatic input stream should be reflected in the rate of either singlet or burst events (22). We termed this quantity the event rate (Fig. 1*A*), E, and it is calculated by taking the sum of the singlet rate and the burst rate. This event rate equals the firing rate only in the absence of bursting and is otherwise smaller. In this nomenclature, the burst rate, B, is by definition a conjunction $B = EF$. Each of these quantities— B, E, and F—represents a distinct signal, which may need to be communicated to a specific target.

Although burst coding was the focus of many theoretical (31, 39–42) and experimental (32–34, 38, 43, 44) studies and although ensemble burst coding may have been implied in some experimental studies (38, 44), its potential as a neural code for multiplexing has not been explored previously. In the following, we use computational modeling and theoretical analyses to show that the anatomy and the known physiology of the neocortical networks are consistent with this neural code for TPNs.

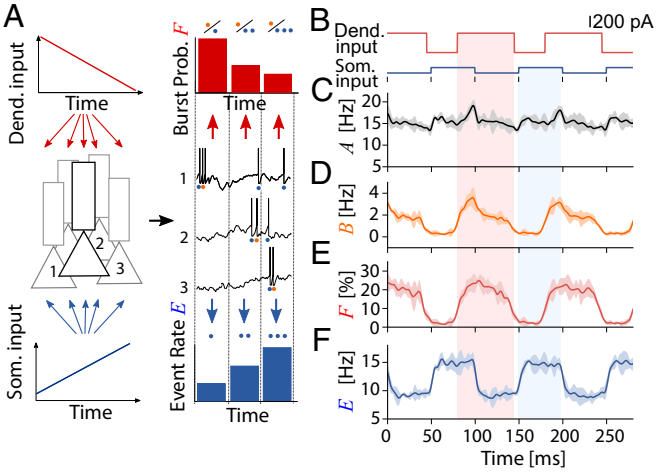

Fig. 1. BEM for representing simultaneously two signals. (*A*) Schema of the suggested neural code: One signal is delivered to the somata and another to the apical dendrites of a neural ensemble. Each neuron responds with a series of action potentials (black traces), which can be classified as isolated spike events or burst events on the basis of the interspike interval. The total number of events (blue dots) can be computed in each time bin to form an event rate (blue bars). The burst probability (red bars) is calculated by taking the ratio of the number of bursts (orange dots) and the total number of events for each time step. (*B–F*) Computational simulation to illustrate the neural code. (*B*) Alternating somatic and dendritic input currents were used as inputs to a population of 4,000 two-compartment neuron models (25). Phase lag and amplitudes were chosen such that (*C*) the output firing rate remains largely constant, to illustrate the ambiguity of the firing rate code. (*D*) The ensemble burst rate reflects a conjunction of somatic and dendritic inputs. (*E*) The burst probability reflects the alternating dendritic input. (*F*) The ensemble event rate reflects the alternating somatic input. Shaded regions show two standard deviations, calculated over five trials.

Encoding: Dendritic Spikes for Multiplexing. To illustrate the BEM code in cortical ensembles, we first consider the firing statistics of model TPNs as they respond to alternating dendritic and somatic input shared among neurons (Fig. 1*B*). Individual TPNs are simulated using a two-compartment model that has been constrained by electrophysiological recordings to capture dendrite-dependent bursting [*SI Appendix*, Fig. S1 *A–D* (30)], a critical frequency for an after-spike depolarization [*SI Appendix*, Fig. S1 *E–H* (45)], and the spiking response of TPNs to complex stimuli in vitro (46). In addition to the shared alternating signals, each cell in the population receives independent background noise to reproduce the high variability of recurrent excitatory networks balanced by inhibition, as well as low burst fraction and the typical membrane potential SD observed in vivo (47–49) (*Materials and Methods*). As a result, simulated spike trains display singlets interweaved with short bursts of action potentials. Both types of events appear irregularly in time and are weakly correlated across the population (*SI Appendix*, Fig. S2, mean pairwise correlation coefficient <0.005). In the example illustrated in Fig. 1, the dendritic and somatic inputs were chosen to yield an approximately constant firing rate (Fig. 1*C*). This illustrates the ambiguity of firing-rate responses: The same response could have arisen from a constant somatic input. The burst rate (Fig. 1*D*) is also ambiguous as it signals the conjunction of somatic and dendritic inputs. However, this ambiguity can be resolved since a strong dendritic input is more likely to convert a single spike into a burst. Indeed, the event rate and burst probability qualitatively recover the switching pattern injected into the dendritic and somatic compartments, respectively (Fig. 1 *E* and *F*). Thus, it emerges that TPNs can simultaneously communicate many different functions of the somatic and dendritic

inputs, depending on the spike-timing patterns considered in the ensemble average.

To determine the dynamic range of this multiplexing, we characterize the input–output (I-O) function of the ensemble by simulating the population response to short 20-ms current pulses of varying amplitude, delivered simultaneously to all compartments (Fig. 2 *A* and *B*). Consistent with earlier computational work and in vitro recordings of the time-averaged firing rate (31, 42, 50, 51), the ensemble firing rate grows nonlinearly with the somatic input, with a gain that is modulated by concomitant dendritic input (Fig. 2*C*). Also consistent with the single-cell notion that bursts signal a conjunction of dendritic and somatic inputs (30), the ensemble burst rate in our simulations strongly depends on both somatic and dendritic inputs (Fig. 2*D*). The event rate increases nonlinearly with the somatic input (Fig. 2*E*), but is less modulated by the dendritic input than the firing rate, consistent with an encoding of the somatic input stream. The dynamic range of the event rate is limited to small and moderate input strengths since the event rate saturates when somatic input is sufficiently strong to produce a burst in all of the cells. Similarly, burst probability is driven by the dendritic input strength. There is a weak modulation by the concomitant somatic input (Fig. 2*F*), but it is overall consistent with our hypothesis that burst probability encodes the dendritic information stream. The dynamic range of burst probability is limited by small to moderate dendritic inputs since strong dendritic inputs produce bursts across the entire population, saturating the I-O function.

We summarize the above results by a phenomenological model of the different responses. The ensemble firing rate, or activity $A(I_s, I_d)$, depends on the somatic input, I_s, but is also modulated by the dendritic input, I_d (Fig. 2*C*). This bivariate function can be separated into a sum of two univariate functions, an event rate $E(I_s)$ and a modulation which corresponds to the number of additional spikes n incurred by bursts controlled by univariate probability $F(I_d)$:

$$A = E(I_s)\,[1 + nF(I_d)]. \qquad [1]$$

Since the burst rate $B = EF$ and the single spike rate $S = E - B$ are both bivariate, the decomposition of the response into event rate and burst probability can recover two independent channels. Without advocating which of these spike-timing patterns, A, E, B, S, and F, are used by the nervous system, we can best assess the information content by focusing on univariate quantities $E(I_s)$ and $F(I_d)$.

To quantify the quality of this multiplexing, we compute a signal-to-noise ratio (SNR_I), which is high for a response that is strongly modulated by the input in one compartment but invariant to input in the other (*SI Appendix*, *Data Analysis Methods*). We find that both the burst probability and the event rate reached larger SNR_I than either the burst rate or the firing rate (maximum $\mathrm{SNR}_I > 250$ for burst probability and >1,000 for the event rate vs. $\mathrm{SNR}_I < 10$ for the firing rate and <5 for the burst rate; *SI Appendix*, Fig. S3). Also, the range of input amplitudes with an $\mathrm{SNR}_I > 1$ is broader for burst probability than for burst rate (gray regions in Fig. 2 *E* and *F*). For very high inputs, the clear invariance of the somatic and the dendritic input in event rate and burst probability breaks down (*SI Appendix*, Fig. S4), because bursts can be triggered by somatic or dendritic input alone and are no longer a conjunctive signal. Therefore, multiplexing of dendritic and somatic streams is possible, unless either somatic or dendritic inputs are very strong. The low firing rates and sparse occurrence of bursts typically observed in vivo (47–49) are in line with this regime.

Information-Limiting Factors in Multiplexing. Previous studies that were not based on ensemble coding have shown that bursts encode the slowly varying part of sensory inputs only (34, 52). Given the need for fast cortical communication (53), we ask whether BEM is limited in terms of how fast it can encode two input streams. To this end, we simulated the response of an ensemble of TPNs receiving two independent input signals, one injected into the dendrites and the other injected into the somata. Both inputs are time dependent and fluctuate with equal power in fast and slow frequencies over the 1- to 100-Hz range (*SI Appendix*, *Computational Methods*). As a first step, we consider the case of a very large ensemble (80,000 cells) to minimize finite-size effects. Since the I-O functions obtained from pulse inputs (Fig. 2) are approximately exponential in the moderate input regime, we use the logarithm of the burst probability as an estimate of the dendritic input and the logarithm of the event rate as an estimate of the somatic input. Although crude compared with decoding methods taking into account pairwise correlation and adaptation (54–56), this simple approach recovers accurately both the somatic and dendritic inputs (Fig. 3 *A* and *B* and *SI Appendix*, Fig. S5), with deficits primarily for rapid dendritic input fluctuations. To quantify the encoding quality at different timescales, we calculate the frequency-resolved coherence between the inputs and their estimates. The coherence between dendritic input and its estimate based on the burst probability (Fig. 3*B*) is close to one for slow input fluctuations, but decreases to zero for rapid input fluctuations. Concurrently, the event rate can decode the somatic input with high accuracy for input frequencies up to 100 Hz (Fig. 3*B*). When somatic input is much stronger than dendritic input, it is still possible to recover slowly changing dendritic information by considering the burst probability (Fig. 3*C* and *SI Appendix*, Fig. S5). In all cases the coherence is at least as high, but typically surpasses the coherence obtained from the classical firing-rate code, indicating that burst multiplexing matches but typically surpasses the information contained in the firing rate. To rule out that bursts would reflect merely the overall strength of the two inputs

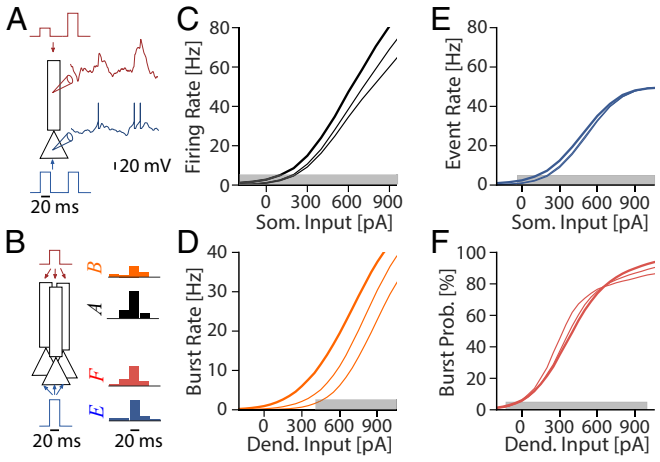

Fig. 2. I-O functions for distinct spike-timing patterns. (*A*) The two-compartment TPN model combines dendrite-dependent burst firing in the presence of high somatic and dendritic inputs with background noise. (*B*) I-O functions are computed by simulating the response of 4,000 TPNs to short current pulses and averaging across the ensemble. (*C*) The ensemble firing rate as a function of the somatic input amplitude is shown in the presence of a concomitant dendritic input (0 pA, 200 pA, 400 pA; thicker line corresponds to larger dendritic input). (*D*) The ensemble burst rate as a function of the dendritic input amplitude is shown in the presence of concomitant somatic input pulses (0 pA, 200 pA, 400 pA; thicker line corresponds to greater somatic input). (*E*) Same as *C* but for the ensemble event rate. (*F*) Same as *D* but for the burst probability, computed by dividing the ensemble burst rate by the event rate. Gray bars indicate the input regimes associated with $\mathrm{SNR}_I > 1$ (*SI Appendix*, *Data Analysis Methods*).

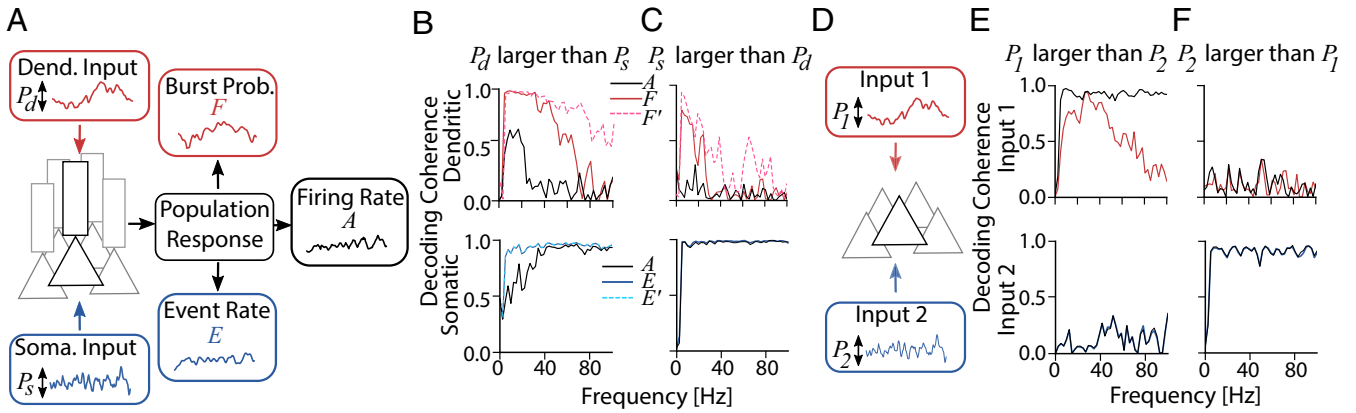

Fig. 3. Decoding multiplexed time-dependent information. (*A*) Two uncorrelated time-dependent inputs are injected into distinct subcellular compartments. Multiplexing suggests that more information can be decoded by considering burst probability and event rate as estimates of the dendritic and somatic inputs, respectively. Classical firing rate coding suggests that information can be decoded by considering the firing rate as an estimate of summed dendritic and somatic inputs. (*B*) For strong dendritic inputs, the coherence of the dendritic input (*Top*) and of the somatic input (*Bottom*) is shown with respect to burst probability (red), event rate (blue), and firing rate (black). Increasing the sensitivity of dendritic excitability increases the bandwidth of burst probability (pink dashed line) without altering the event rate (cyan dashed line). (*C*) Same as *B* but for dendritic input of small amplitude and somatic input of large amplitude. (*D*) For comparison, we simulate the injection of the same inputs as in *A* into the somatic compartment. (*E* and *F*) The coherence for (*E*) the case of stronger input 1 or for (*F*) the case of stronger input 2.

combined, we considered a scenario where the same two inputs are injected into the soma (Fig. 3 *D–F* and *SI Appendix*, Fig. S6). In this case the firing-rate code is most informative and reports only the strongest input. These observations support the idea that active dendrites enable ensembles to increase information representation via multiplexing.

We now ask, What limits the coherence bandwidth of the burst probability (Fig. 3*C*)? If it were limited by the dendritic membrane potential dynamics, coding could in principle be improved by changing membrane properties. If it were limited by the finite duration of bursts, which effectively introduces a long refractory period before the next burst can occur, this could introduce a fundamental speed limit for BEM. To investigate the latter, we performed an information-theoretic analysis of the BEM code, which indicates that the refractory period does not affect the bandwidth of BEM for sufficiently large ensembles, consistent with previous theoretical work (57). Alternatively, the membrane dynamics in the dendrite could limit the bandwidth. Slow passive dynamics in the apical dendrites are not likely to be a limiting factor, because the high density of the hyperpolarization-activated ion channels (58) contributes to a particularly low dendritic membrane time constant (46). The other possibility is that the slow onset of dendritic spikes limits the bandwidth (59). This possibility would be compatible with slow calcium spike onsets observed, arising from the kinetics of calcium ion channels (60). Therefore, we simulated the response to the same time-dependent input shown in Fig. 3 but with a threefold increase of the voltage sensitivity for dendritic spikes, to accelerate the onset of dendritic spikes. This single manipulation considerably improved the encoding of high-frequency fluctuations (Fig. 3 *B* and *C* and *SI Appendix*, Fig. S7). Thus, the slow onset of dendritic spikes indicates that TPNs sacrifice bandwidth to represent slowly changing information.

The total amount of information transmitted depends on a variety of extrinsic and intrinsic factors. The main extrinsic factors are the power and the bandwidth of the input signals. High power is manifested in large input changes, which can strongly synchronize cells and can thus increase the ensemble response. Consistently, the coherences obtained in the previous section depend strongly on our choice of input power. For instance, decreasing the relative power in the dendrites decreases the coherence obtained from the burst probability (Fig. 3 *C* and *D*). To arrive at a more objective assessment of multiplexing, we derived mathematical expressions for the event and burst information rates, assuming low pairwise correlations (61). The neurons are receiving a fixed total input power P separated equally in the two compartments and are emitting a fixed total number of spikes per unit time A_0 (*SI Appendix*, *Theoretical Methods*). The information rate from the event rate increases with the number of neurons N and the frequency bandwidth W, but decreases with the number of spikes per bursts n and the stationary burst probability F_0:

$$\mathcal{M}_E = W \log_2 \left(1 + \frac{NA_0P}{2(1+nF_0)\,W}\right). \quad [2]$$

Similarly, the burst probability can communicate additional information at a rate which increases with stationary burst probability:

$$\mathcal{M}_F = W \log_2 \left(1 + \frac{NA_0F_0(1-F_0)P}{2(1+nF_0)\,W}\right). \quad [3]$$

We can compare the multiplexing information obtained by summing, $\mathcal{M}_E+\mathcal{M}_F$, with the information obtained from the classical firing rate with total input power P and emitting the same total amount of spikes per unit time A_0:

$$\mathcal{M}_A = W \log_2 \left(1 + \frac{NA_0P}{W}\right). \quad [4]$$

Since both the input power and the output power are matched in the multiplexing code and the classical firing-rate code, these expressions can be used to determine the conditions under which multiplexing is advantageous.

First, we find that there is an optimal burstiness, i.e., mean burst probability, for which information transmission is maximized (Fig. 4*A*). This optimum arises from the fact that rare bursting sacrifices information from the dendritic stream, whereas frequent bursting must sacrifice information from the somatic stream to meet the constraint of total number of spikes. This optimal burstiness depends on the number of spikes in a burst and the bandwidth of the two channels. It decreases with the number of spikes per burst (Fig. 4*B*), in line with the notion

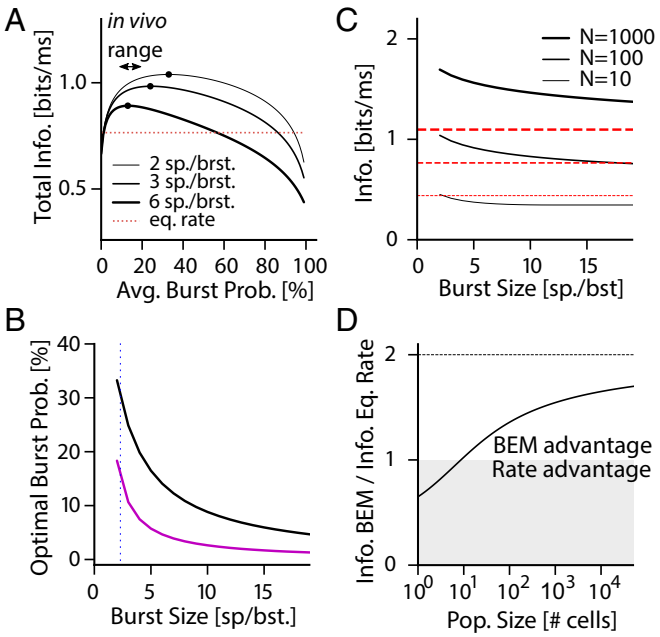

Fig. 4. Short and sparse bursts are optimal for multiplexing. (*A*) Theoretical estimates of total multiplexed information constrained to a fixed total number of spikes (black curves, Eqs. **2** and **4**). The information varies as a function of the stationary burst probability and has a maximum for low burst probability (black circles), consistent with bursting statistics in vivo (47). The three black curves correspond to three burst sizes (2, 3, and 6 spikes per burst), illustrating that the smaller burst size communicates the greater amount of information. The information rate for a firing-rate code with matched input amplitude and stationary firing rate is shown for comparison (red dashed line, Eq. **4**). (*B*) The optimal mean burst probability decreases as a function of burst size. It reaches 31% for typical burst size [corresponding to an average of 2.3 spikes per burst (47), dotted line] for parameters as in *A*. Considering slower dendritic dynamics $W_d = 0.2W_s$ reduces the optimal burst probability (magenta curve). (*C*) The maximum information rate (solid black curves, $N = 10$, 10^2, and 10^3) decreases as a function of burst size. It surpasses the information of the firing-rate code with matched input amplitude and stationary firing rate (red lines, for corresponding N) for sufficiently large ensembles and for small burst sizes. (*D*) For matched input amplitude and output rate, the total multiplexed information gain relative to the information rate of the firing rate (Eq. **2** plus Eq. **3** over Eq. **4**) asymptotes to two (dashed line). The area where the classical firing rate offers an advantageous coding strategy is shaded gray. When not specified, parameter values were $p = 0.5$, $N = 100$, $P = 20$ (49, 62), $F_0 = 0.5$, $n = 1$, $W_s = W_d = 100$ Hz, and $A_0 = 10$ Hz (*SI Appendix, Theoretical Methods*).

that long bursts convey little information per spike and should hence be used more sparsely. The optimal burstiness further decreases with decreasing dendritic bandwidth (Fig. 4*B*), that is, for neurons with slower dendritic dynamics. The information transmitted decreases with the number of spikes per burst (Fig. 4*C*), in line with the intuition that the second spike in a burst marks the event as a burst, whereas additional spikes contain no further information. Hence, for neurons with slow dendritic dynamics, BEM performs best when bursts are short and occur rarely, in line with experimental observations (47). Finally, the preference for short bursts is independent of the number of neurons in the ensemble (Fig. 4*C*), but a minimal number of neurons is required to transmit more information than a rate code with the same firing rate and number of neurons (Fig. 4 *C* and *D*). If the somatic and dendritic compartments have the same bandwidth, the total information transmitted by BEM approaches twice the information of a classical rate code, in the limit of very large ensembles (Fig. 4*D*). In summary, the theoretical analysis suggests that short and sparse bursts in a large ensemble maximize information transmission in burst multiplexing.

Decoding: Cortical Microcircuits for Demultiplexing. For the brain to make use of a multiplexed code, the different streams have to be decodable by biophysical mechanisms. Previous experimental (63) and theoretical (52, 64) studies have argued that short-term plasticity (STP) can play a role in facilitating or depressing the postsynaptic response of bursts. We have argued that ensembles of TPNs may represent distinct information not in the rate of single spikes and bursts, but in the event rate and the burst probability. To test if STP can be used to recover these ensemble features, we simulated cell populations receiving excitatory input from TPNs and studied how the response of these postsynaptic cells depends on the input to TPNs (Fig. 5). As a model for STP, we used the extended Tsodyks–Markram model (63) with parameters constrained by the reported properties of neocortical connections (65) (*SI Appendix, Computational Methods*). By decreasing the postsynaptic effect of additional spikes in a burst, short-term depression (STD) could introduce a selectivity to events, particularly for short bursts and strong depression. Indeed, we find that in a population of cells receiving excitatory TPN inputs, responses correlated slightly more with event rate when STD was present than without STP (*SI Appendix*, Fig. S8 *A* and *B*). The presence of STD affects the range of SNR_I above one only weakly, but increases the maximum reached by SNR_I considerably (SNR_I reaches 120 with STD and is below 30 without STD; *SI Appendix*, Fig. S9 *A* and *B*). STD can hence be interpreted as an event rate decoder. In turn, since the TPN event rate encodes the somatic stream (Fig. 2), it is not surprising that we find STD to further suppress the weak dependence on dendritic inputs, while maintaining the selectivity to somatic input (Fig. 5*B*). Hence STD improves the selective decoding of the somatic stream.

We then ask whether postsynaptic neurons can decode the conjunction of somatic and dendritic inputs. By increasing the postsynaptic effect of later spikes in a burst (63, 64, 66), short-term facilitation (STF) boosts the sensitivity to bursts (*SI*

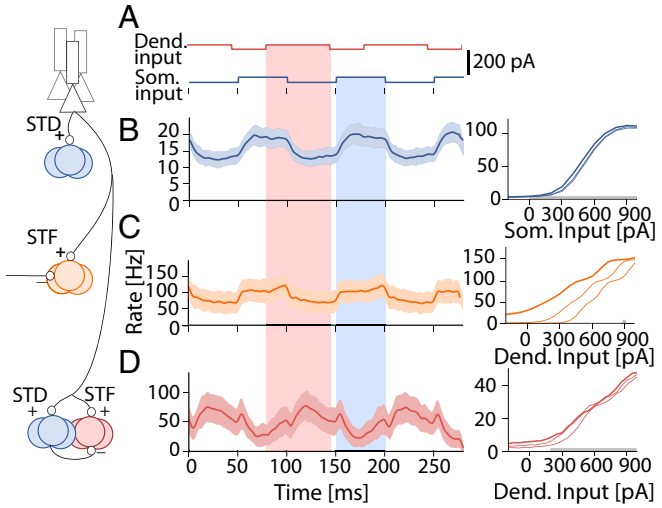

Fig. 5. The role of STP and disynaptic inhibition for separating distinct information streams. (*A*) Alternating somatic and dendritic inputs are injected in 4,000 TPNs. (*B–D*, *Left*) Response of population postsynaptic to the TPNs under alternating inputs (solid line; shaded area shows 2 SDs around the mean). (*Right*) Responses of postsynaptic population for short pulse inputs to TPNs shown as a function of the amplitude of somatic and dendritic pulses. The different lines show the response in the presence of a 200-, 300-, and 400-pA (thicker) concomitant input (somatic input when abscissa is dendritic and vice versa). Gray shading highlights range with $\mathrm{SNR}_I > 1$. (*B–D*) The firing-rate response of neurons postsynaptic to TPNs is shown for synapses (*B*) with STD, (*C*) with STF along *x*-axis and constant inhibition, and (*D*) with STF and disynaptic inhibition.

Appendix, Fig. S8*C*) and hence to the dendritic stream (Fig. 5*C* and *SI Appendix*, Fig. S9*C*). To decode the dendritic stream, neurons should compute a quantity similar to the burst probability (Fig. 2*D*). We reason that neural computation of burst probability could be achieved by combining burst rate sensitivity with divisive disynaptic inhibition from an event rate decoder. Thus, we consider a population receiving facilitating excitatory input from TPNs, combined with disynaptic inhibition from an STD-based event rate decoder. We manually adjusted the weights of these connections to increase the postsynaptic effect of dendritic inputs, while decreasing the postsynaptic effect of somatic input. This was achieved with potent excitation and inhibition, a regime associated with divisive inhibition (39, 67). The output rate of this microcircuit displayed a higher correlation with burst probability than that of a microcircuit without disynaptic inhibition (*SI Appendix*, Fig. S8*D*). This microcircuit can selectively decode dendritic input (Fig. 5*D*) with an SNR_I above 1 over a large range of dendritic input amplitudes (*SI Appendix*, Fig. S9*D*). Is the presence of STD essential to this operation? In line with the weak dependence of TPN firing rate on dendritic input, we find that the presence of STD in this microcircuit is not essential since decoding of the dendritic stream can also result from STF combined with disynaptic inhibition without STD (*SI Appendix*, Fig. S9 *E* and *F*). We conclude that a microcircuit with STP and disynaptic inhibition in a divisive regime can selectively extract different input streams from a multiplexed neural code.

Gain Control of Multiplexed Signals. To transmit significant information, the burst code relies on a graded increase of the burst rate as a function of dendritic and somatic inputs, which is at odds with the all-or-none nature of calcium spikes in single cells (30) and ensembles (31). Three mechanisms can linearize the I-O function and transform an all-or-none response into a graded one. These mechanisms are background noise, spike-frequency adaptation, and feedback inhibition. For fast and reliable encoding, feedback inhibition is the most efficient since linearization is faster than with adaptation and the SNR better than with background noise. Feedback dendritic inhibition (FDI) is mediated in the neocortex by somatostatin-positive neurons (SOMs), which receive input from TPNs (66) and project back to the apical dendrites of the same ensemble. FDI is known to linearize dendritic activity (68, 69) and may therefore linearize the burst probability, while feedback somatic inhibition linearizes the event rate. But since SOMs are activated by the TPNs, both somatic and dendritic inputs may reduce the burst probability, breaking the segregation of dendritic and somatic streams. Therefore, we ask whether FDI inhibition can linearize the burst response without introducing a coupling between the two input streams.

To this end, we simulated TPNs receiving feedback inhibition from a burst-probability decoder (Fig. 6*A* and *SI Appendix*, Fig. S10 *A–D*). We find that the presence of such FDI reduces the average burst length (Fig. 6*B*), consistent with similar experimental manipulations in the hippocampus (70). Also, there is a weaker gain modulation of the firing rate I-O function compared with TPNs without FDI (Fig. 6*C*). Inhibition from the burst probability decoder motif does not abolish bursting in the TPNs but reduces both the overall proportion of bursts and the gain of the burst-probability I-O function (Fig. 6*D*). Importantly, this form of FDI does not change the invariance of the burst probability to somatic input, so the multiplexed code is conserved (Fig. 6*D*). We suggest that this invariance requires that FDI is primarily driven by dendritic input to TPNs. The effect of somatic input to TPNs is counteracted by the disynaptic inhibition. When divisive disynaptic inhibition in the burst-probability decoder is replaced by a constant hyperpolarizing current (Fig. 6*E* and *SI Appendix*, Fig. S10 *E–H*), FDI becomes strongly modulated by somatic input (*SI Appendix*, Fig. S10*H*). Although bursts

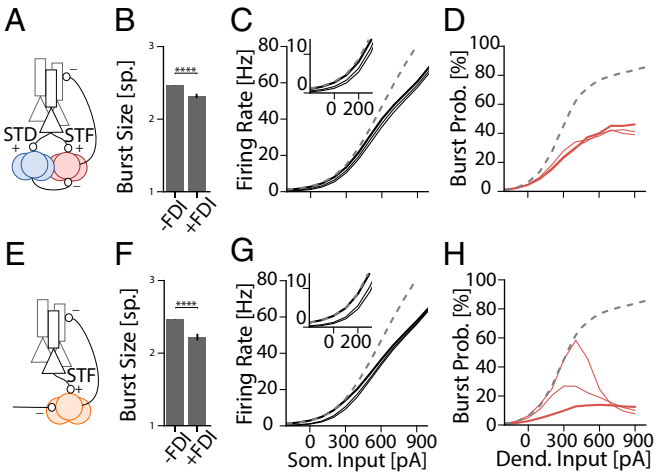

Fig. 6. Dendritic feedback inhibition controls multiplexing gain. (*A*) Schematic representation of the simulated network. (*B*) Burst size averaged across simulated input conditions with and without FDI. (*C*) Dependence of TPN firing rate on somatic input for three different amplitudes of dendritic input pulses (0 pA, 200 pA, and 400 pA). *Inset* shows zoom-in close to firing-rate onset. (*D*) Dependence of burst probability on the dendritic input for three different somatic input amplitudes (0 pA, 200 pA, and 400 pA). The thicker line shows the response with the strongest somatic input. (*E–H*) As in *A–D*, but replacing disynaptic inhibition by a hyperpolarizing current (430 pA). In this case, the burst probability generally decreases with somatic input (*H*). Asterisks in *B* and *F* indicate significant difference (Welch's *t* test, $P < 0.0001$). Dashed lines in *C*, *D*, *G*, and *H* show response in the absence of FDI (corresponding to thick lines in Fig. 2).

remain shorter (Fig. 6*F*) and sparser (Fig. 6*H*), and although the gain of the firing-rate response is reduced in a very similar fashion (Fig. 6*G*), the burst probability loses its invariance with respect to somatic input (Fig. 6*H*). Therefore, FDI from the burst-probability decoder motif seen in anatomical studies (71) is required to control the gain of the dendritic signal while at the same time maintaining short and sparse bursts. Such a mechanism is essential for optimal information transfer in the multiplexed neural code.

Discussion

We have introduced a neural code able to simultaneously communicate two streams of information through a single neural ensemble. This neural implementation of multiplexing is distinct from time-division (23, 72) and frequency-division multiplexing (18) and is specific to communication with spike trains. Contrary to single-cell burst coding (34), we have found that ensemble burst coding can encode quickly changing inputs, although processing speed may be limited by the biophysical properties of active dendrites. This code is optimal for short and sparse bursts, which is consistent with observed bursting in L2–3 and L5B cells (47). Finally, we have illustrated in simulations how ensemble multiplexing suggests specific connectivity motifs to demultiplex burst coded information. We believe that this neural code satisfies the need to communicate different quantities in top–down and bottom–up directions through the same neurons.

Extensions of Multiplexing. Is BEM antagonizing frequency-division multiplexing? Frequency-division multiplexing has been suggested on the basis of experimental observations (15, 17). Burst coding can in theory supplement this type of code since distinct types of events may synchronize to distinct frequency bands. This idea is supported by experimental evidence from two studies in nonhuman primates. In the first study (73), bursts were shown to synchronize to distinct bands of the local field potential compared to single spikes. In the second study (17), mutual

information between sensory stimuli and neural responses is shown to increase when spike patterns are taken into account in addition to the local field potential. Also, in this study it is interesting to note that mutual information increases further when spike patterns are taken into account in addition to the firing rate, in an amount consistent with our derivations. We can therefore not rule out that spike timing patterns can be combined with frequency-division multiplexing.

For clarity of the exposition, we have limited our discussion to a distinction between singlets and bursts. It is has been suggested that the size of bursts may encode additional information (29, 74). One can generalize BEM to capture three distinct inputs controlling three types of events, namely singlets, short bursts, and long bursts. A postsynaptic decoding of such a complex temporal ensemble code would probably require intricate dynamics of STP (75), however, potentially combined with microcircuits.

Potential Roles for Multiplexing. Multiplexing allows one to increase the information conveyed by a fixed number of axons. It could therefore be a solution to the energy and space constraints associated with sending information over long distances. TPNs are known to project to subcortical targets including the colliculus, pons, and medulla (76). The axons of TPNs in the motor cortex axons are notoriously long since these synapse to motor neurons in the spinal cord (77). Neurons in layer 5 also send a smaller number of axons to other cortical targets intrahemispheric (78) or interhemispheric (76). Since in the majority of these projections the distance is considerable, the energetic cost of spikes is high, and multiplexing may have arisen to reduce the amount of energy per bit.

Another line of experimental evidence suggests that the same cellular and synaptic mechanisms enabling multiplexing are essential for learning (79) and perception (80). In multiple areas outside neocortex, bursting in particular has been associated with both learning (81, 82) and perception (14, 44). While further theoretical work is required, this suggests that bursting does more than allowing cells to preserve two types of information unaltered through multiple processing stages (Fig. 7*A*). Instead, a learning signal, in accordance with a powerful learning algorithm (12), can be sequentially altered by the sensory information as it is back propagating through the network (Fig. 7*B*). All of the while, sensory information can climb the hierarchy unaltered by the learning-related signal. Experimental design and analysis in view of this neural code are required to shine a light on these issues.

Limitations of Multiplexing. Theoretical studies on frequency-division multiplexing have shown that this type of multiplexing can encode time-varying information only for slowly changing inputs (83). This theoretical constraint is supported by experimental assessment of the relative increase in information in spike phase compared with spike rate (17). Although theoretically BEM does not suffer from this constraint and can be used to communicate rapidly changing information, the properties of dendritic spike initiation will in practice limit multiplexing to slowly changing information as well. The slower fluctuations observed in higher-order areas (84) suggest that this can be an effective coding strategy given slower top–down information impinging on the dendrites.

The multiplexing mode of burst coding described here is limited to a regime where inputs are small or moderate. In TPNs, strong somatic inputs could trigger bursts of action potentials, which would challenge an association between dendritic inputs and burst probability. Since the regime where multiplexing can arise (Fig. 2) is clearly reflected in the ensemble event rate and burst probability, experiments could assess whether TPNs are maintained in this regime or whether the ensemble can switch between multiplexing and classical rate coding. Whether the inhibitory motifs described here can pick up this switch of modes is a question that lies beyond the scope of this study.

Another limitation lies in the assumption that spikes are probabilistically converted into bursts independently across neurons. Dendrites should hence be in an asynchronous state with weak pairwise correlations. This asynchronous state was suggested to enable rapid (85) and efficient (86) encoding. In our simulations,

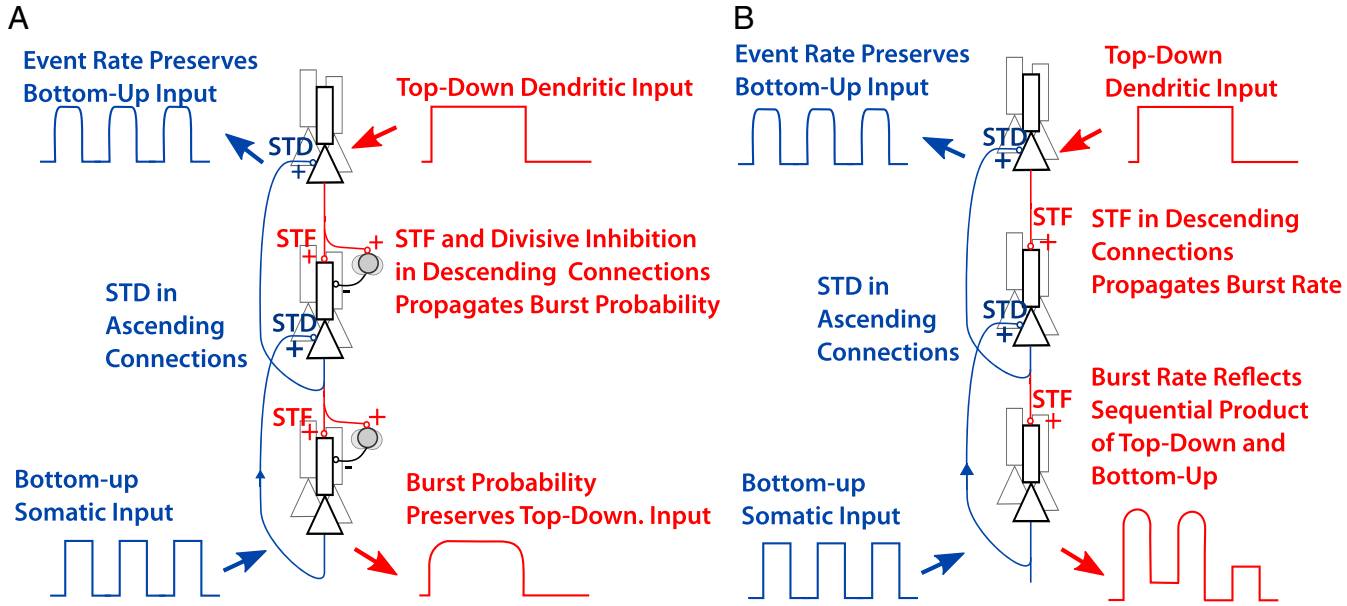

Fig. 7. Potential functional roles of BEM. (*A*) When the descending connections have STF with disynaptic inhibition, top–down information can propagate down unaltered by bottom–up information, via burst probability. Ascending connections with potent STD can communicate bottom–up information even in the presence of potent descending drive. (*B*) When descending connections have STF without potent disynaptic inhibition, a conjunction of ascending and descending information, the burst rate, is communicated down.

the asynchronous state was mimicked by a substantial background noise that effectively desynchronized and linearized the responses of both somatic and dendritic compartments. A physiological basis for this background noise could lie in the stochastic activation of ion channels (87–89) as well as in a state of balanced excitation and inhibition, which is known to favor pairwise decorrelation (48, 90) and could be supported by homeostatic inhibitory synaptic plasticity (91). We have argued that FDI can keep TPNs in a state of short and sparse bursts; it is natural to suggest that FDI can also ensure asynchronous dendritic activity. Burst ensemble coding hence requires a synergistic coordination among single-cell bursting mechanisms, the morphological targets of different input streams on these cells, and neuronal circuit motifs. Whether hallmarks of such a coordination can be found in different neuronal systems remains an open question.

General Relevance. The proposed code may act in other parts of the central nervous system where bursting has been observed. We can now delineate several structural requirements and predictions:

Sparse bursting. As recorded from a single cell in a stationary and awake condition, bursts should be relatively rare and short.

Multiplexed tuning. A difference should be observed between the tuning of trial-averaged burst probability and the tuning of event rate.

Burst code linearization. Feedforward inhibition, feedback inhibition, adaptation, and intrinsic noise are cellular-level mechanisms that can mediate a linearization of an ensemble rate. These mechanisms should be observed in the compartment responsible for modulating the burst probability.

We briefly discuss these in relation to bursting in the ELL, the cerebellum, the thalamus, the hippocampus, and the neocortex.

Sparse Bursting. Across the nervous system, bursts are generally much less numerous than isolated spikes. Specifically, the burst probability associated with the quiet awake state is in the range of 10–30% in sensory cortex (47), prefrontal cortex (73), ventral thalamus (92), and Purkinje cells (93). In all these areas, bursts are also short, with an average number of spikes per burst ranging between 2 and 4. In the CA1 region of the hippocampus, bursts are generally much longer, but also much less frequent [stationary burst probability between 1% and 2% (38)]. These observations are generally in accordance with the maximization of multiplexed information (Fig. 4). We are aware of only one exception: Granule cells of the dentate gyrus fire a great proportion of relatively long bursts (94).

Multiplexed Tuning. Experiments rarely report the tuning of burst responses. When this is done, single spike rate is often contrasted with burst rate instead of burst probability. In the neocortex, burst probability is associated with attention (73), perception (80), and, theoretically, an error signal (13). In the ELL, burst probability signals different dynamic features of the electric stimulation than tonic spikes (44). These observations are loosely consistent with multiplexing, but they warrant targeted investigations. The question of burst-probability tuning was specifically addressed in a recent study in Purkinje cells (93). It was shown that the burst probability is tuned to the direction of saccade errors and that this tuning is opposite to the tuning direction of simple spikes. This forms a more rigorous agreement with the principle of burst multiplexing.

Burst Code Linearization. In the neocortex, linearization is likely implemented through feedback and feedforward inhibition of the apical dendrites. Characteristically, neocortical pyramidal cells send facilitating connection to SOMs and short-term depression to parvalbumin-positive (PV) cells (63, 95). A subtype of SOM cells, the Martinotti cells, inhibits specifically the dendrites and acts as a powerful control of dendritic spikes (30, 68, 69). These cells also receive inhibition from PV cells (71). We therefore find a feedback inhibition motif consistent with Fig. 6*A*. In addition, since both thalamocortical projections and top–down cortical projections are thought to project both onto the apical dendrites and to SOM cells (96, 97), SOM cells can control the ensemble response via feedforward as well as feedback inhibition. Interestingly, thalamocortical projections follow the rule STF onto SOM cells and STD onto PV cells (96). The slow membrane potential dynamics of Martinotti cells also appear ideal for the encoding of slowly changing burst-coded information (Fig. 3). This suggests that Martinotti cells are well poised to encode burst probability and provide targeted feedback inhibition to the dendrites of pyramidal cells.

In the hippocampus, SOM cells share similar connection motifs and intrinsic properties with Martinotti cells. These cells may therefore preserve sparse and short bursts while simultaneously linearizing the ensemble response of pyramidal cells in the CA1 region through feedback and feedforward inhibition of the apical dendrites (70, 98). Similarly, dendrite-targeting inhibition can have the same role of linearizing ensemble burst responses in the cerebellum (99) and the ELL (100). In the thalamus, bursting via potent inhibition (101) would require a linearization and burst dependence of local inhibition.

To conclude, we have exposed how active dendrites need microcircuits and STP to support a multiplexed neural code. Direct experimental assessment of this code may help us to establish whether it generalizes across the central nervous system and what algorithmic function it subserves.

Materials and Methods

Data analysis. Bursts are defined as a set of spikes followed or preceded by an interspike interval smaller than 16 ms. Burst rate is calculated by finding all bursts and summing across the population. Event rate is calculated by finding all isolated spikes and the first spike in a burst before summing across the population. Burst probability is calculated as the ratio of the burst rate over the event rate. A smoothing kernel of 10 ms is used for displaying the time-dependent rates.

Network Simulations. The network consists of four types of units: pyramidal-cell basal bodies, pyramidal-cell distal dendrites and inhibitory cells from population **a** receiving STF and from population **b** receiving STD. We used the lowercase letters **s**, **d**, **a**, and **b** to label the different units, respectively. Somatic dynamics follow generalized integrate-and-fire dynamics described by a membrane potential u and a generic recovery variable w to account for subthreshold-activated ion channels and spike-triggered adaptation (27). For the ith unit in population x we used

$$\frac{d}{dt}u_i^{(x)} = -\frac{u_i^{(x)} - E_L}{\tau_x} + \frac{g_x f(u_i^{(d)}) + c_x K(t - \hat{t}_i^{(s)}) + I_i^{(x)} + w_i^{(x)}}{C_x} \quad [5]$$

$$\frac{d}{dt}w_i^{(x)} = -w_i^{(x)}/\tau_w^{(x)} + a_w^{(x)}\left(u_i^{(x)} - E_L\right)/\tau_w^{(x)} + b_w^{(x)}S_i^{(x)}, \quad [6]$$

where E_L is the reversal potential, C_x the capacitance, $a_w^{(x)}$ the strength of subthreshold adaptation, $b_w^{(x)}$ the strength of spike-triggered adaptation, $\tau_w^{(x)}$ the timescale of the recovery variable, and τ_x the timescale of the membrane potential (see *SI Appendix, Computational Methods* for parameter values). An additional term controlled by g_x (Eq. **5**) models the regenerative activity in the dendrites as described below, but is absent from inhibitory cells ($g_a = g_b = 0$). Also, an additional term controlled by c_x and the kernel K (*SI Appendix, Computational Methods*) models how the last action potential at $t_i^{(s)}$ is back propagating from the soma in the dendrites and is absent from all other units (c_s, c_a, c_b = 0). When units **s**, **a**, and **b** reach a threshold at V_T, the membrane potential is reset to the reversal potential after an absolute refractory period of 3 ms and a spike is added to the spike train $S_i^{(x)}$ in the form of a sum of Dirac δ functions.

The dendritic compartment has nonlinear dynamics dictated by the sigmoidal function f to model the nonlinear activation of calcium channels (46). This current propagates to the somatic unit such that g_s controls the somatic effect of forward calcium spike propagation. In the dendritic compartment, the parameter g_d controls the potency of local regenerative

activity. The dendritic recovery variable controls both the duration of the calcium spike consistent with potassium currents (102, 103) and resonating subthreshold dynamics consistent with the h current (58). Note that the differential equations are similar to those used to model NMDA spikes (104), but the strong and relatively fast recovery variable ensures that the calcium spikes have shorter durations (10–40 ms). Membrane time constants were matched to values found in vivo (95) (*SI Appendix*, Table S1).

Each unit receives a combination of synaptic input, external input, and background noise (*SI Appendix, Computational Methods*). Synapses were modeled as exponentially decaying changes in conductance. Connection probability was chosen to be 0.2 for excitatory connections and 1 for inhibitory connections consistent with experimental observations (95, 105). Background noise was modeled as a time-dependent Ornstein–Uhlenbeck process independently drawn for each unit. We chose the amplitude of background fluctuations to be such that the SD of membrane potential fluctuations is around 6 mV as observed in V1 L2–3 pyramidal neurons (49). To model STP, we used the extended Tsodyks–Markram model (63) with parameters consistent with experimental calibrations in vitro (65) (*SI Appendix, Computational Methods*).

ACKNOWLEDGMENTS. We thank Guillaume Hennequin, Jean-Claude Béïque, and Matthew Larkum for helpful discussions. We thank Loreen Hertäg, Alexandre Payeur, and Stephen E. Clarke for critical reading of the manuscript as well as Greg Knoll for an independent verification of the numerical results. This work was supported by a Bernstein Award (01GQ1201) from the German Federal Ministry for Science and Education and by Natural Sciences and Engineering Research Council Discovery Grant (06872). Part of this work was conducted (R.N. and H.S.) at the Computational and Biological Learning Laboratory, Department of Engineering, University of Cambridge, Cambridge, United Kingdom.

1. Hubel D, Wiesel T (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160:106–154.
2. Felleman D, van Essen D (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47.
3. Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nat Neurosci* 12:718–724.
4. Graziano M (2006) The organization of behavioral repertoire in motor cortex. *Annu Rev Neurosci* 29:105–134.
5. Mineault PJ, Khawaja FA, Butts DA, Pack CC (2012) Hierarchical processing of complex motion along the primate dorsal visual pathway. *Proc Natl Acad Sci USA* 109:E972–E980.
6. De Pasquale R, Sherman SM (2011) Synaptic properties of corticocortical connections between the primary and secondary visual cortical areas in the mouse. *J Neurosci* 31:16494–16506.
7. Niell CM, Stryker MP (2010) Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* 65:472–479.
8. Keller GB, Bonhoeffer T, Hübener M (2012) Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. *Neuron* 74:809–815.
9. Gilbert CD, Li W (2013) Top-down influences on visual processing. *Nat Rev Neurosci* 14:350–363.
10. Dayan P, Hinton GE, Neal RM, Zemel RS (1995) The Helmholtz machine. *Neural Comput* 7:889–904.
11. Bastos AM, et al. (2012) Canonical microcircuits for predictive coding. *Neuron* 76:695–711.
12. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323:533–536.
13. Guergiuev J, Lillicrap TP, Richards BA (2017) Towards deep learning with segregated dendrites. *eLife* 6:e22901.
14. Crick F (1984) Function of the thalamic reticular complex: The searchlight hypothesis. *Proc Natl Acad Sci USA* 81:4586–4590.
15. Engel AK, Singer W (2001) Temporal binding and the neural correlates of sensory awareness. *Trends Cogn Sci* 5:16–25.
16. Larkum M (2013) A cellular mechanism for cortical associations: An organizing principle for the cerebral cortex. *Trends Neurosci* 36:141–151.
17. Kayser C, Montemurro M, Logothetis N, Panzeri S (2009) Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* 61:597–608.
18. Akam T, Kullmann DM (2014) Oscillatory multiplexing of population codes for selective communication in the mammalian brain. *Nat Rev Neurosci* 15:111–122.
19. Konishi M (2003) Coding of auditory space. *Annu Rev Neurosci* 26:31–55.
20. Friedrich RW, Habermann CJ, Laurent G (2004) Multiplexing using synchrony in the zebrafish olfactory bulb. *Nat Neurosci* 7:862–871.
21. Fairhall AL, Lewen G, Bialek W, van Steveninck R (2001) Efficiency and ambiguity in an adaptive neural code. *Nature* 412:787–792.
22. Brenner N, Strong SP, Koberle R, Bialek W, de Ruyter van Steveninck R (2000) Synergy in a neural code. *Neural Comput* 12:1531–1552.
23. O'Reilly RC (1996) Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Comput* 8:895–938.
24. Kaifosh P, Losonczy A (2016) Mnemonic functions for nonlinear dendritic integration in hippocampal pyramidal circuits. *Neuron* 90:622–634.
25. Meyer HS, et al. (2013) Cellular organization of cortical barrel columns is whisker-specific. *Proc Natl Acad Sci USA* 110:19113–19118.
26. Knight BW (1972) Dynamics of encoding in a population of neurons. *J Gen Physiol* 59:734–766.
27. Gerstner W, Kistler W, Naud R, Paninski L (2014) *Neuronal Dynamics* (Cambridge Univ Press, Cambridge, UK).
28. Wang X, et al. (2007) Feedforward excitation and inhibition evoke dual modes of firing in the cat's visual thalamus during naturalistic viewing. *Neuron* 55:465–478.
29. Kepecs A, Wang XJ, Lisman J (2002) Bursting neurons signal input slope. *J Neurosci* 22:9053–9062.
30. Larkum M, Zhu J, Sakmann B (1999) A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature* 398:338–341.
31. Shai AS, Anastassiou CA, Larkum ME, Koch C (2015) Physiology of layer 5 pyramidal neurons in mouse primary visual cortex: Coincidence detection through bursting. *PLoS Comput Biol* 11:e1004090.
32. Gabbiani F, Metzner W, Wessel R, Koch C (1996) From stimulus encoding to feature extraction in weakly electric fish. *Nature* 384:564–567.
33. Bastian J, Nguyenkim J (2001) Dendritic modulation of burst-like firing in sensory neurons. *J Neurophys* 85:10–22.
34. Oswald A, Chacron M, Doiron B, Bastian J, Maler L (2004) Parallel processing of sensory input by bursts and isolated spikes. *J Neurosci* 24:4351–4362.
35. Deschênes M, Paradis M, Roy J, Steriade M (1984) Electrophysiology of neurons of lateral thalamic nuclei in cat: Resting properties and burst discharges. *J Neurophys* 51:1196–1219.
36. Brumberg JC, Nowak LG, McCormick DA (2000) Ionic mechanisms underlying repetitive high-frequency burst firing in supragranular cortical neurons. *J Neurosci* 20:4829–4843.
37. Magee JC, Carruth M (1999) Dendritic voltage-gated ion channels regulate the action potential firing mode of hippocampal CA1 pyramidal neurons. *J Neurophys* 82:1895–1901.
38. Bittner KC, et al. (2015) Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons. *Nat Neurosci* 18:1133–1142.
39. Doiron B, Longtin A, Berman N, Maler L (2001) Subtractive and divisive inhibition: Effect of voltage-dependent inhibitory conductances and noise. *Neural Comput* 13:227–248.
40. Körding KP, König P (2001) Supervised and unsupervised learning with two sites of synaptic integration. *J Comput Neurosci* 11:207–215.
41. Izhikevich EM (2007) *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting* (MIT Press, Cambridge, MA).
42. Giugliano M, La Camera G, Fusi S, Senn W (2008) The response of cortical neurons to in vivo-like input current: Theory and experiment: II. Time-varying and spatially distributed inputs. *Biol Cybern* 99:303–318.
43. Harris KD, Hirase H, Leinekugel X, Henze DA, Buzsáki G (2001) Temporal interaction between single spikes and complex spike bursts in hippocampal pyramidal cells. *Neuron* 32:141–149.
44. Clarke SE, Maler L (2017) Feedback synthesizes neural codes for motion. *Curr Biol* 27:1356–1361.
45. Larkum ME, Kaiser K, Sakmann B (1999) Calcium electrogenesis in distal apical dendrites of layer 5 pyramidal cells at a critical frequency of back-propagating action potentials. *Proc Natl Acad Sci USA* 96:14600–14604.
46. Naud R, Bathellier B, Gerstner W (2014) Spike-timing prediction in cortical neurons with active dendrites. *Front Comput Neurosci* 8:90.
47. De Kock C, Sakmann B (2008) High frequency action potential bursts (> 100 hz) in L2/3 and L5B thick tufted neurons in anaesthetized and awake rat primary somatosensory cortex. *J Physiol* 586:3353–3364.
48. Renart A, et al. (2010) The asynchronous state in cortical circuits. *Science* 327:587–590.
49. Polack PO, Friedman J, Golshani P (2013) Cellular mechanisms of brain state-dependent gain modulation in visual cortex. *Nat Neurosci* 16:1331–1339.
50. Larkum ME, Senn W, Lüscher HR (2004) Top-down dendritic input increases the gain of layer 5 pyramidal neurons. *Cereb Cortex* 14:1059–1070.
51. Hay E, Segev I (2014) Dendritic excitability and gain control in recurrent cortical microcircuits. *Cereb Cortex* 25:3561–3571.
52. Middleton JW, Yu N, Longtin A, Maler L (2011) Routing the flow of sensory signals using plastic responses to bursts and isolated spikes: Experiment and theory. *J Neurosci* 31:2461–2473.
53. Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381:520–522.
54. Bialek W, Rieke F, de Ruyter van Steveninck R, Warland D (1991) Reading a neural code. *Science* 252:1854–1857.
55. Pillow J, et al. (2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454:995–999.
56. Naud R, Gerstner W (2012) Coding and decoding with adapting neurons: A population approach to the peri-stimulus time histogram. *PLoS Comput Biol* 8:e1002711.
57. Deger M, Schwalger T, Naud R, Gerstner W (2014) Fluctuations and information filtering in coupled populations of spiking neurons with adaptation. *Phys Rev E* 90:062704.
58. Kole MHP, Hallermann S, Stuart GJ (2006) Single IH channels in pyramidal neuron dendrites: Properties, distribution, and impact on action potential output. *J Neurosci* 26:1677–1687.

59. Wei W, Wolf F (2011) Spike onset dynamics and response speed in neuronal populations. *Phys Rev Lett* 106:088102.
60. Pérez-Garci E, Gassmann M, Bettler B, Larkum M (2006) The GABAB1b isoform mediates long-lasting inhibition of dendritic Ca^{2+} spikes in layer 5 somatosensory pyramidal neurons. *Neuron* 50:603–616.
61. Dettner A, Münzberg S, Tchumatchenko T (2016) Temporal pairwise spike correlations fully capture single-neuron information. *Nat Commun* 7:13805.
62. Pozzorini C, Naud R, Mensi S, Gerstner W (2013) Temporal whitening by power-law adaptation in neocortical neurons. *Nat Neurosci* 16:942–948.
63. Markram H, Wang Y, Tsodyks M (1998) Differential signaling via the same axon of neocortical pyramidal neurons. *Proc Natl Acad Sci USA* 95:5323–5328.
64. Izhikevich EM, Desai NS, Walcott EC, Hoppensteadt FC (2003) Bursts as a unit of neural information: Selective communication via resonance. *Trends Neurosci* 26:161–167.
65. Costa RP, Sjöström PJ, Van Rossum MC (2013) Probabilistic inference of short-term synaptic plasticity in neocortical microcircuits. *Front Comput Neurosci* 7:75.
66. Silberberg G, Markram H (2007) Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. *Neuron* 53:735–746.
67. Murphy BK, Miller KD (2003) Multiplicative gain changes are induced by excitation or inhibition alone. *J Neurosci* 23:10040–10051.
68. Murayama M, et al. (2009) Dendritic encoding of sensory stimuli controlled by deep cortical interneurons. *Nature* 457:1137–1141.
69. Gidon A, Segev I (2012) Principles governing the operation of synaptic inhibition in dendrites. *Neuron* 75:330–341.
70. Royer S, et al. (2012) Control of timing, rate and bursts of hippocampal place cells by dendritic and somatic inhibition. *Nat Neurosci* 15:769–775.
71. Jiang X, et al. (2015) Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* 350:aac9462.
72. Xie X, Seung HS (2003) Equivalence of backpropagation and contrastive Hebbian learning in a layered network. *Neural Comput* 15:441–454.
73. Womelsdorf T, Ardid S, Everling S, Valiante TA (2014) Burst firing synchronizes prefrontal and anterior cingulate cortex during attentional control. *Curr Biol* 24:2613–2621.
74. Mease RA, Kuner T, Fairhall AL, Groh A (2017) Multiplexed spike coding and adaptation in the thalamus. *Cell Rep* 19:1130–1140.
75. Chamberland S, Timofeeva Y, Evstratova A, Volynski K, Toth K (2018) Action potential counting at giant mossy fiber terminals gates information transfer in the hippocampus, 10.1073/pnas.1720659115.
76. O'Leary DD, Koester SE (1993) Development of projection neuron types, axon pathways, and patterned connections of the mammalian cortex. *Neuron* 10:991–1006.
77. Gao WJ, Zheng ZH (2004) Target-specific differences in somatodendritic morphology of layer V pyramidal neurons in rat motor cortex. *J Comp Neurol* 476:174–185.
78. Glickfeld LL, Andermann ML, Bonin V, Reid RC (2013) Cortico-cortical projections in mouse visual cortex are functionally target specific. *Nat Neurosci* 16:219–226.
79. Cichon J, Gan WB (2015) Branch-specific dendritic Ca^{2+} spikes cause persistent synaptic plasticity. *Nature* 520:180–185.
80. Takahashi N, Oertner TG, Hegemann P, Larkum ME (2016) Active cortical dendrites modulate perception. *Science* 354:1587–1590.
81. Yang Y, Lisberger SG (2014) Purkinje-cell plasticity and cerebellar motor learning are graded by complex-spike duration. *Nature* 510:529–532.
82. Bittner KC, Milstein AD, G1rienberger C, Romani S, Magee JC (2017) Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science* 357:1033–1036.
83. Akam T, Kullmann DM (2010) Oscillations and filtering networks support flexible routing of information. *Neuron* 67:308–320.
84. Murray JD, et al. (2014) A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci* 17:1661–1663.
85. Gerstner W (2000) Population dynamics of spiking neurons: Fast transients, asynchronous states, and locking. *Neural Comput* 12:43–89.
86. Sompolinsky H, Yoon H, Kang K, Shamir M (2001) Population coding in neuronal systems with correlated noise. *Phys Rev E* 64:051904.
87. Cannon RC, O'Donnell C, Nolan MF (2010) Stochastic ion channel gating in dendritic neurons: Morphology dependence and probabilistic synaptic activation of dendritic spikes. *PLoS Comput Biol* 6:e1000886.
88. O'Donnell C, van Rossum MC (2014) Systematic analysis of the contributions of stochastic voltage gated channels to neuronal noise. *Front Comp Neurosci* 8: 105.
89. Naud R, Payeur A, Longtin A (2017) Noise gated by dendrosomatic interactions increases information transmission. *Phys Rev X* 7:031045.
90. van Vreeswijk C, Sompolinsky H (1996) Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274:1724–1726.
91. Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W (2011) Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science* 334:1569–1573.
92. Urbain N, et al. (2015) Whisking-related changes in neuronal firing and membrane potential dynamics in the somatosensory thalamus of awake mice. *Cell Rep* 13:647–656.
93. Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R (2015) Encoding of action by the Purkinje cells of the cerebellum. *Nature* 526:439–442.
94. Pernía-Andrade AJ, Jonas P (2014) Theta-gamma-modulated synaptic currents in hippocampal granule cells in vivo define a mechanism for network oscillations. *Neuron* 81:140–152.
95. Pala A, Petersen CC (2015) In vivo measurement of cell-type-specific synaptic connectivity and synaptic transmission in layer 2/3 mouse barrel cortex. *Neuron* 85:68–75.
96. Tan Z, Hu H, Huang ZJ, Agmon A (2008) Robust but delayed thalamocortical activation of dendritic-targeting inhibitory interneurons. *Proc Natl Acad Sci USA* 105:2187–2192.
97. Leinweber M, Ward DR, Sobczak JM, Attinger A, Keller GB (2017) A sensorimotor circuit in mouse cortex for visual flow predictions. *Neuron* 95:1420–1432.
98. Urban-Ciecko J, Barth AL (2016) Somatostatin-expressing neurons in cortical networks. *Nat Rev Neurosci* 16:401–409.
99. Orduz D, Llano I (2007) Recurrent axon collaterals underlie facilitating synapses between cerebellar Purkinje cells. *Proc Natl Acad Sci USA* 104:17831–17836.
100. Mehaffey W, Doiron B, Maler L, Turner R (2005) Deterministic multiplicative gain control with active dendrites. *J Neurosci* 25:9968–9977.
101. Lesica NA, Stanley GB (2004) Encoding of natural scene movies by tonic and burst spikes in the lateral geniculate nucleus. *J Neurosci* 24:10731–10740.
102. Cai X, et al. (2004) Unique roles of SK and Kv4.2 potassium channels in dendritic integration. *Neuron* 44:351–364.
103. Harnett MT, Xu NL, Magee JC, Williams SR (2013) Potassium channels control the interaction between active dendritic integration compartments in layer 5 cortical pyramidal neurons. *Neuron* 79:516–529.
104. Major G, Polsky A, Denk W, Schiller J, Tank DW (2008) Spatiotemporally graded NMDA spike/plateau potentials in basal dendrites of neocortical pyramidal neurons. *J Neurophysiol* 99:2584–2601.
105. Fino E, Yuste R (2011) Dense inhibitory connectivity in neocortex. *Neuron* 69:1188–1203.