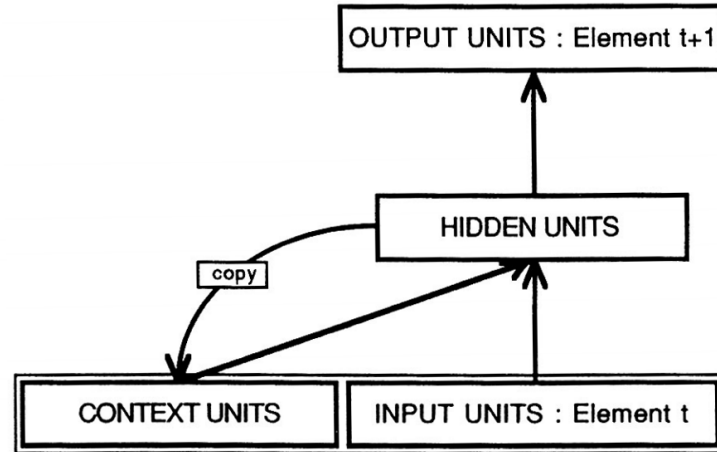


1 Introduction

The *LeabraTI* framework is a mechanistic description and general model of how prediction and temporal integration might work in the brain. The general idea is closely related to the Simple Recurrent Network (SRN) (Elman90Servan-SchreiberCleeremansMcClelland91) an artificial neural network architecture that explicitly represents temporally lagged information in discrete “context” units whose activity gets integrated with more current information to predict what happens in the next time step (Figure 1a). This method of copying a contextual representation from an intermediate representation at discrete intervals was originally shown to be a robust way to leverage error-driven learning to represent latent temporal structure in auditory streams and artificial grammars. More generally, the SRN’s explicit representation of temporal context can capture the latent structure of any stimulus that varies systematically over time, making it a good basis for a generic prediction and temporal integration mechanism.

However, *LeabraTI* differs in several key ways from the traditional SRN architecture, primarily in the way temporally lagged context is represented and used in neural processing. These differences are due to biological constraints imposed by the circuitry of the neocortex, and thus form a number of testable predictions that can be used to evaluate the validity of the *LeabraTI* framework. The central prediction of *LeabraTI* is that temporal context is updated in a periodic manner, approximately every 100 ms, subserved by deep (Layer 5 and 6) neurons. This periodic updating is suggested to contribute to the brain’s alpha rhythm, which has been studied extensively using scalp electroencephalography. This chapter **TODO – also need something here about error driven learning**

a)



b)

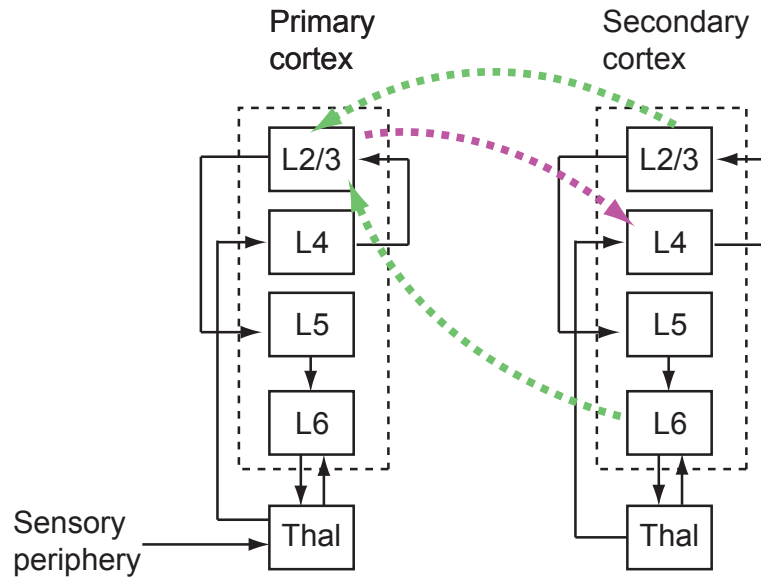


Figure 1: The Simple Recurrent Network (SRN) and microcircuitry of the neocortex. **a)** The SRN represents temporal information explicitly using discrete context units that are updated once per time step. Context is integrated with more current inputs to predict information at the subsequent time step. Reproduced from Servan-SchreiberCleeremansMcClelland91). **b)** The neocortex is laminated with canonical circuitry between neurons across layers and between areas. Intraareal connections are shown in black with interareal feedforward connections in purple and feedback connections in green.

2 LeabraTI biological details

2.1 Laminar structure and microcircuitry of the neocortex

A salient feature of the brain, and potential clue in realizing how an SRN-like computation might be carried out in biological neural circuits, is the laminar structure prevalent across the neocortex (Figure 1b). Incoming information from the sensory periphery is transmitted through the thalamus and targets Layer 4 neurons in the primary sensory cortices (e.g., V1). From there, Layer 4 neurons propagate spikes to superficial neurons (Layers 2 and 3) which in turn target Layer 4 neurons of higher-level cortices, forming the prominent corticocortical feedforward pathways that subserve visual and auditory recognition (FellemanVanEssen91). Corticocortical feedback originates in superficial layers or Layer 6 of the higher-level cortex and generally terminates on superficial neurons of the lower-level cortex (RocklandPandya79). In addition to these interareal pathways, there exists a canonical microcircuit of the form Layer 4 \rightarrow Layer 2/3 \rightarrow Layer 5 \rightarrow Layer 6 that routes spike propagation through the local neuronal structure (DouglasMartin04ThomsonLamy07). This microcircuit forms the core computational unit of LeabraTI, as will be described in this and the following sections.

The importance of the local microcircuit was first suggested by Vernon Mountcastle in his proposal regarding the gross columnar organization of the neocortex (see Mountcastle97, for a comprehensive review). Mountcastle’s proposal states that microcolumns composed of around 80-100 neurons extending vertically through all six lamina with canonical circuitry form the core repeating structure of the neocortex. Neurons within a single microcolumnar circuit possess nearly identical receptive field tunings across lamina while neurons in neighboring microcolumns (radial separation greater than 600 μm) possess very different receptive field tunings but contribute to the higher-order macrocolumn (i.e., hypercolumn) structure (HubelWiesel77Jones00). Microcolumns have been identified in a variety of neural systems with this electrophysiological mapping and are also prominently visible under Nissl staining. Despite this evidence for their structural existence, any function of the microcolumn aside from an organizing principle remains debated (Buxhoeve-

denCasanova02HortonAdams05).

LeabraTI provides a computational role for the microcolumn, by mapping an SRN-like computation onto their Layer 4 \rightarrow Layer 2/3 \rightarrow Layer 5 \rightarrow Layer 6 circuit. In this mapping, superficial neurons continuously integrate feedforward and feedback interareal synapses to process current information. Layer 2/3 \rightarrow Layer 5 \rightarrow Layer 6 provides an intraareal pathway for explicitly representing temporal context deep layers, which are relatively isolated from nonlocal inputs. There is also appropriate circuitry for recirculating this context through the local microcolumn via Layer 4 to drive the learning of temporal associations. This basic idea provides a concise explanation for the strong degree of isotuning throughout a single microcolumn, as deep neurons need to represent the same information as superficial neurons except at a delayed interval.

More generally, LeabraTI's dichotomy of continuous integration in superficial layers and periodic updating of deep layers receives strong support by the literature. Recent studies that have employed depth electrodes to simultaneously record from multiple layers within a patch of cortex have indicated that superficial layers exhibit spectral power at much higher frequencies than deep layers. BuffaloFriesLandmanEtAl11) recorded responses from ventral visual sites V1, V2, and V4 in awake, behaving monkeys during a simple directed attention task, finding a dissociation in spike coherence frequency in superficial (gamma spectrum, peak ~ 50 Hz) and deep layers (alpha spectrum, peak ~ 10 Hz). A similar experimental paradigm expands on these findings by demonstrating cross-frequency coupling between gamma and alpha spectra localized to superficial and deep layers, respectively (SpaakBonnefondMaierEtAl12). The cross-frequency coupling is characterized by a clear nesting of gamma activity within alpha cycles, suggesting that deep neurons' alpha activity might subserve a general pacemaker mechanism. In the context of LeabraTI, this pacemaker property is important to ensure the regular updating of context through deep layers and temporally predictable reintegration with more current information.

In summary, the laminocolumnar organization of the neocortex provides the dual pathways necessary for continuous information processing and the SRN's explicit temporal context representation. One question that remains, however, concerns the 10 Hz alpha periodicity of deep

neurons. The Layer 4 → Layer 2/3 → Layer 5 → Layer 6 microcircuit only contains four synapses including the thalamus and intracolumnar monosynaptic latencies for regular spiking neurons are on the order of 5 ms or faster (Armstrong-JamesFoxDas-Gupta92LumerEdelmanTononi97). This relatively small amount of tissue, if driven with constant input, would cause deep neurons to spike at a rate much faster than 10 Hz. How such a circuit could produce the strong alpha power observed in recent depth recordings is described next.

2.2 Pacemaker properties of Layer 5 and thalamic bursting neurons

Layer 5 neurons can be roughly divided into 5a and 5b subtypes (ThomsonLamy07). Layer 5a neurons have relatively small cell bodies and exhibit “regular spiking” depolarization responses. They collect input from other Layer 5a neurons and pass it to 5b neurons and thus, likely play a simple information integration role. Layer 5b neurons, in contrast, have larger cell bodies and exhibit “intrinsic bursting” properties at ~ 10 Hz when over threshold (ConnorsGutnickPrince82SilvaAmitaiConnors91; FranceschettiGuatteoPanzicaEtAl95).

Thalamic neurons...

Layer 5b neurons project to Layer 6 whose neurons can also be roughly divided into cortico-cortical (CC) and corticothalamic (CT) subtypes (Thomson10).

Both Layer Layer 6 CC neurons have properties similar to Layer 5a neurons – they collect inputs from other Layer 6a neurons and pass it to Layer 6 CT neurons. Layer 6 CT neurons project specifically to the thalamus and also receive direct thalamic input forming a small circuit. They have

This rhythmic firing has been shown to persist even with constant sensory stimulation *in vivo* (LuczakBarthoHarris13), suggesting that Layer 5 neurons’ alpha rhythmicity could implement a roughly 10 Hz gating function for spikes relayed to Layer 6 neurons.

Thus, Layer 6 specifically becomes the neural substrate of the SRN’s temporally lagged context representation, representing information that is, on average, one alpha cycle (approximately 100 ms) in the past. This contextual storage occurs at an automatic interval due to the intrinsic

pacemaking properties of Layer 5 neurons, and might implement a reference frame that essentially would allow the brain to know *when* to anticipate inputs. As such, intrinsic oscillations have been shown to phase lock to environmental stimulation (WillBerg07LakatosKarmosMehtaEtAl08SchroederLakatos09S ensuring environmental events coincide with key events like Layer 5 bursts in cortex.

Layer 6 sends axons toward the thalamus, completing the microcircuit within the local column and allowing the temporally lagged Layer 6 information to integrate with more current Layer 4 inputs. There also exists a direct connection between Layer 6 and Layer 4, that could be used for this purpose, although it has been noted as being weak compared to other intracolumnar connections (HirschMartinez06b). In either case, temporal associations could be created by simple Hebbian learning mechanisms that track high probability co-occurrences across past and present events.

The Leabra algorithm (OReillyMunakata00OReillyMunakataFrankEtAl12), however, also makes use of powerful error-driven learning (in addition to more standard Hebbian learning). In the context of temporal integration, error-driven learning would allow computation of error signals based on the difference between what is predicted to happen at a given moment (given the previous moments context as an input) and what actually happens. However, this computation requires that both the prediction and the actual sensation are represented subsequently within a single alpha cycle, which is not possible if the sensory periphery is always transmitting incoming inputs. To resolve this issue, the LeabraTI framework posits a mechanism to modulate or even block the transmission of inputs from the sensory periphery. A subset of cells in the thalamus exhibit alpha spectrum bursting properties similar to that of Layer 5 neurons (LopesdaSilva91; HughesLorincz-CopeEtAl04; LorinczCrunelliHughes08; LorinczKekesiJuhaszEtAl09), and thus perhaps perform a similar gating function. Specifically, these neurons appear to shift the balance of inputs to Layer 4 and superficial neurons between exogenous environmental inputs and endogenous inputs local to the microcolumn.

When environmental inputs are downmodulated or blocked, Layer 6 context relayed via the thalamus is the dominant input to the microcolumn, which can be used to predict the incoming sensory event during the latter part of the alpha cycle. Importantly, during both the prediction and

sensation phases, feedforward and feedback projections are constantly transmitting between lower and higher cortical areas. As previously mentioned, these projections originate and terminate predominantly in superficial layers, boosting their spike coherence to higher frequency spectra. This could potentially explain the differentially high gamma power in superficial layers compared to deep layers, and provides a compelling link between gamma oscillations and predicting specific details about the next sensory event.

2.3 Summary of LeabraTI computation

The overall computation of LeabraTI is shown in Figure 2 and summarized here. When thalamic cells burst (roughly every 100 ms), information from the sensory periphery is the primary driving force for Layer 4 neurons in primary cortex. This information is relayed downstream to higher-level cortical areas via the strong feedforward Layer 4 \rightarrow Layer 2/3 \rightarrow Layer 4 pathway (FellemanVanEssen91). Within the local microcolumn, Layer 5 neurons integrate this information, until thalamic bursting quiets (generally around 50 ms). At this point, Layer 5 cells burst, sending outputs to Layer 6 and shifting and inputs to the microcolumn endogenously. The information represented by Layer 6 neurons is temporally lagged (from the previous 50 ms) and is relayed to Layer 4 via non-bursting (regular spiking) thalamic neurons or via the direct Layer 6 \rightarrow Layer 4 connection (not pictured in Figure 2), and might be maintained by reciprocal thalamocortical drive back to Layer 6. This information can be used as a prediction as to what will happen next when thalamic bursting resumes and veridical sensory information serves as the input once again. In the context of Leabra’s error-driven learning these two phases correspond to the plus phase (sensation) and minus phase (prediction), which can be used to compute a sensory prediction error signal. This error signal modifies Layer 5 \rightarrow Layer 6 synapses to minimize differences between predictions and sensations over time.

Critically, for the LeabraTI computation to work, thalamic and Layer 5 oscillatory phases need to have an approximately antiphase relationship in order for the error-driven learning scheme described here to work so that Layer 2/3 neurons can represent the current moment’s prediction with

Layer 6 context as their primary input and then subsequently represent the veridical sensory input while Layer 5 neurons are queuing up the next contextual event. Such a relationship has not yet been shown yet, but very few studies have recorded simultaneously from thalamic and cortical neurons in *in vivo* in the awake behaving animal. It is also possible that the brain implements error-driven learning in such a way that does not require representing predictions and sensations temporally interleaved on the same neural substrate or even that the brain accomplishes temporal integration completely without supervision, which in case thalamic gating is not required.

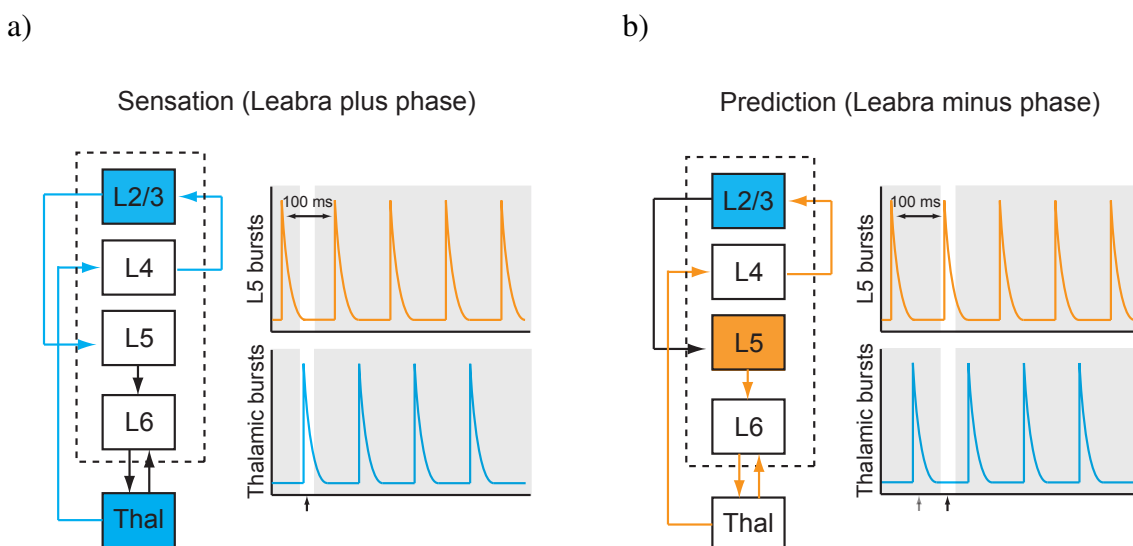


Figure 2: The LeabraTI model computation.

3 LeabraTI predictions

3.1 Relation to existing models

3.2 Testable predictions

TODO