Subcortical modulation of cortical dynamics for motor planning: a computational framework

Jorge Jaramillo[1,2,] Ulises Pereira[1], Karel Svoboda[3], Xiao-Jing Wang[1]
1 New York University. New York, USA
2 European Neuroscience Institute. Göttingen, Germany.
3 Allen Institute for Neural Dynamics. Seattle, USA

Planning, a prospective form of short-term memory, is a cognitive function that has been predominantly attributed to the cortex. Recent experiments, however, have concluded that the thalamus and other subcortical structures participate in this function. A comprehensive computational framework to link neural dynamics and cognition in the context of large-scale subcortical-cortical circuits is lacking.

In this computational study, we elucidated the dynamical mechanisms by which the cortex, thalamus and other subcortical structures jointly contribute to planning. Recurrent circuitry in the cortex *generates* stimulus-selective activity patterns, which are *maintained* by reciprocal corticothalamic projections across a memory epoch. Subcortical signals are routed through the thalamus to selectively modify these patterns, for example enabling execution after planning. We refer to this dynamical process as *subcortex control of dynamical modes*, as the cortical activity patterns ('dynamical modes') are low-dimensional in comparison to the number of neurons that are modulated by the task.

We evaluated the implications of *subcortex control* by simulating networks of interconnected thalamic and cortical 'rate' units in the context of a motor planning task. In tight link with electrophysiological data from mice, we identified subcortical excitatory and inhibitory contributions to the planning computations during the memory epoch. Our model predicts that the *distinct* computational roles of the pars reticulata (SNr) and thalamic reticular nucleus (TRN) during planning (Wang et al., 2021) are a result of their *specific selectivity-dependent* connectivity patterns with the thalamus. Moreover, the 'switch' from movement planning to execution (Inagaki et al., 2021) is instantiated by a midbrain-mediated thalamic burst, which uncovers a latent motor instruction that is *stored in deep cortical layers* during the memory epoch. Overall, we propose a *novel framework* to analyze planning computations in terms of cortical dynamical modes, which are shaped by subcortical structures via the thalamus based on task demands.

Computational framework: We introduce a firing-rate based framework to model two alternative forced choice tasks (2AFC), in which an action must be made in response to sensory stimuli. Importantly, we characterize choice-related computations in 2AFC via *behaviorally-relevant* and *controllable* population activity patterns in the cortex, referred to as 'dynamical modes' (**Figure 1, see details in Methods**).

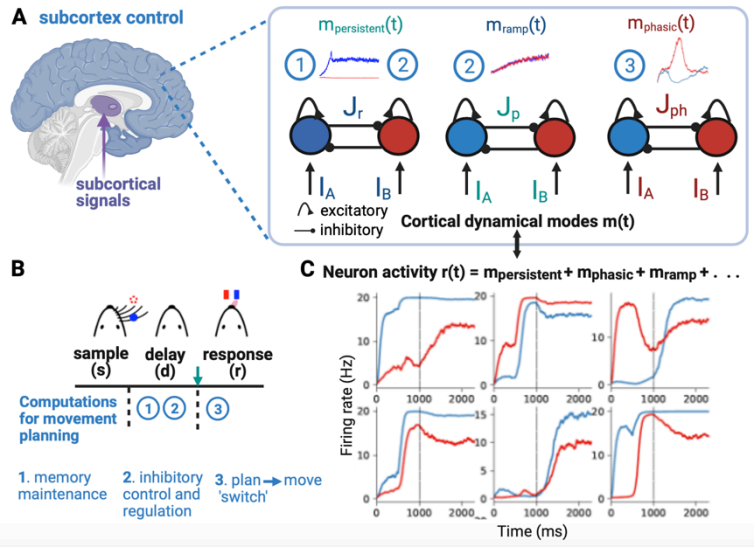

Figure 1. Computational framework: Subcortex control of dynamical modes. A. Dynamical modes are behaviorally-relevant neural activity patterns in the cortex, which are controlled by subcortical inputs onto the thalamus. Modes are generated from excitatory and inhibitory interactions between two selective neural pools in the cortex and modulated via thalamocortical inputs. Synaptic interactions and background currents I_A and I_B determine the dynamics of each individual mode pair. Three dynamical modes are shown, which subserve the computations depicted in B. **B**. Three relevant behavioral computations characterize motor planning across delay (i.e., memory) and response epochs. The green arrow denotes a sensory GO cue, which marks the end of the delay period and beginning of the response period. **C.** Simulated neural activity from one hemisphere of the frontal cortex (blue, contra trial w.r.t. hemisphere; red, ipsi trial).

Dynamical modes represent latent dynamics in the cortex that subserve behavioral computations, e.g., persistent activity for working memory, ramping activity for evidence accumulation, phasic activity for transition from memory to movement, etc (**Figure 1A**). Thus, the behavioral computations underlying a given task (e.g. motor planning) correspond to a sequence of dynamical modes (**Figure 1B**). The function of the thalamus in this framework is twofold: first, it contributes to the maintenance of selective activity in the cortex and second, it routes subcortical signals through specific thalamocortical pathways to modulate dynamical modes as a function of task demands. Subcortical modulation of dynamical modes can take different forms, including: amplification, suppression, switching from one mode to another, and de novo generation. In our framework, distinct modes are generated by distinct pairs of neural pools. Single neurons, however, may form part of more than one pool and contribute to more than one mode, which provides a

basis for heterogeneous neural activity across time and neurons (**Figure 1C**). Based on this framework, we studied the involvement of thalamic nuclei and other subcortical structures during motor planning.

Inhibitory control and regulation of thalamocortical dynamics. Parts of the thalamus are necessary for the maintenance of information across a delay period, as demonstrated by optogenetic inhibition **(Schmitt, et al., 2017; Guo et al., 2017; Bolkan et al., 2017)**. The thalamus receives inhibitory input from the pars reticulata (SNr), involved in action selection and the thalamic reticular nucleus (TRN). It is not clear what the role of these inhibitory structures are in the context of motor planning, and whether they modulate thalamocortical dynamics in the same manner. Wang et al., 2021 activated and inactivated SNr and TRN projections onto the thalamus in one hemisphere and studied the effects on cortical dynamics during motor planning. The SNr manipulation differentially affected contra and ipsi trials w.r.t. the affected hemisphere. In contrast, the TRN manipulation was more uniform across contra and ipsi trials. We modelled thalamocortical dynamics under the assumption that SNr and TRN inputs targeted selective and non-selective modes, respectively (**Figure 2**). *The model predicts that thalamic neurons receive uniform input from TRN, selective input from SNr, and that they form closed thalamocortical loops*. We suggest that the SNr suppresses incorrect choices depending on context (e.g., suppress a 'lick-left' action in a 'lick-right' trial), while the TRN regulates delay activity to prevent runaway excitation from excitatory thalamocortical loops. Thus, the SNr and TRN manipulate information throughout a stimulus-absent delay period: *cognitive* roles that argue against a generic subcortical inhibition.

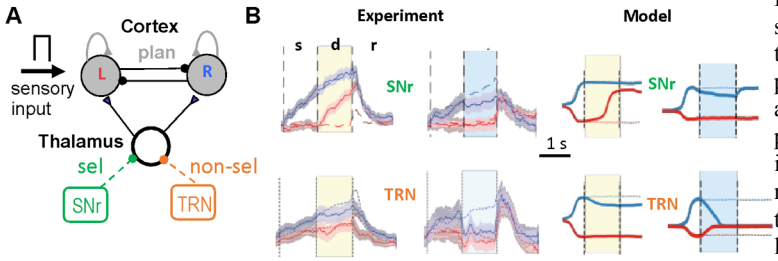

Figure 2. Inhibitory control and regulation. A. Model schematic: two inhibitory structures, SNr and TRN, project to the thalamus and modulate thalamocortical dynamics during a motor planning task. **B**. *Left*, the traces depict 'modes' -population activity patterns in the cortex- for control (dashed) and perturbation (solid) trials during the delay period (yellow shade = inactivation; blue= activation .Wang et al., 2021). *Right*, the model reproduces qualitative features where connectivity to the thalamus (selective vs non-selective in A) is crucial. (blue, right-lick = contra w.r.t. the perturbed hemisphere; red, left-lick = ipsi trials)

A 'behavioral switch': from movement planning to execution. In many cognitive tasks, a sensory cue (i.e., a GO cue) signals the subject to make a choice after a preparatory period. The mechanisms underlying the relationship of the cue to motor planning and execution are not well understood. Inagaki et al., 2021 found that the GO cue is processed subcortically and is subsequently relayed to the thalamus and frontal cortex to generate cue-related and movement dynamics. We investigated the possible circuit-level mechanism behind the planning- to-movement transition (**Figure 3**). We hypothesized that the *movement* mode is negligible until amplified and deployed by the GO cue through the thalamus. To explicitly show this, we considered two pairs of neural pools, responsible for generating persistent and movement modes, respectively. Neurons with persistent activity project to the *movement* neural pool during the delay period, but their activity remains low because they are less excitable (e.g. low baseline currents). Our model suggests that the GO input generates selective activity in the cortex during the response epoch, thus enabling movement. Moreover, the GO input shuts down delay period activity by virtue of intracortical inhibitory projections. *The model predicts the existence of neurons in deep layers with low firing during the delay that switch to high firing at the GO cue.* Thus, the thalamus implements a 'behavioral switch' by amplifying latent movement dynamics in the cortex. To conclude, we propose that the thalamus controls behaviorally-relevant cortical dynamics, based on differential contributions of excitatory and inhibitory subcortical projections. This computational study is constrained by electrophysiological data from rodents during motor planning and encourages new experiments based on *anatomical* and *physiological* model predictions. *Importantly, the framework we propose is sufficiently general to address other cognitive computations and the role of other subcortical structures in dynamical interplay with the cortex*.

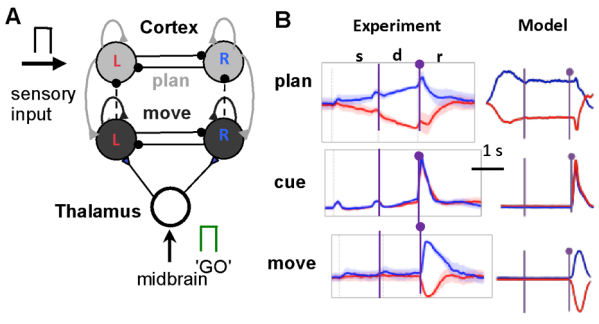

Figure 3. A behavioral switch implemented by the thalamus. A. Model schematic: two pairs of neural pools in the frontal cortex denoted by *plan* and *move* generate planning and movement modes, respectively, for motor planning. The GO input arises from the midbrain, targets the thalamus, and is relayed to the *plan* and *move* pools. **B.** *Left,* neural activity in the frontal cortex is characterized by three modes (Inagaki et al, 2021): a *plan* mode, that represents choice activity during the delay period, a *cue* mode that generates non-selective phasic activity in the cortex, while uncovering a selective *move* mode that is orthogonal to the plan mode. *Right*, the model reproduces qualitative features from the experiment, where intracortical and 'open-loop' thalamocortical connections in A are essential (predicted) circuit parameters.

Methods

Overview

The dynamical modes framework has two major components. For the first component, we construct a dynamical system that represents the activity of neurons in the cortex with biophysically-realistic synaptic dynamics as well as specified low-dimensional dynamics. That is, given a priori low-dimensional patterns of activity (or 'modes'), I obtain the parameters of the dynamical model that gives rise to such modes. The procedure effectively recovers a proxy synaptic connectivity matrix in the cortex, which could represent both intracortical and interareal connections. The second component of the framework is the thalamus, which sends and receives projections from the cortex. The role of the thalamus in this framework is to route control signals from subcortical structures to modulate cortical modes. Recall that such cortical modes were functional, i.e., related to a cognitive process. Thus, subcortical structures can selectively modify these functional modes via thalamo-cortical projections. For example, the thalamus can selectively amplify a mode, or it could modify the stability of a particular mode.

Here we introduce formally the concept of *dynamical mode* with regards to neural activity and the (synaptic) network underlying such (firing-rate) activity. We proceed to describe a *generative* procedure to embed a set of cortical modes onto a neural network. In this self-consistent description, the neural network dynamics, when projected onto a predefined set of directions (equal to the number of modes), gives rise precisely to the cortical modes. Finally, we describe the framework by which the thalamus exerts control on the cortical modes for dynamic computations.

Dynamical modes: distinct population activity patterns in the cortex

We consider a population of N excitatory neurons in the cortex. Such neurons could describe a particular cortical area or more generally, a set of interconnected areas. The N neurons are assumed to be coupled via a synaptic connectivity matrix S_{NXN}. Neurons can inhibit each other through feedforward inhibition mediated by interneurons. This implies that S will generally have positive and negative entries, since we don't consider explicit interneurons for the time being. We denote the (time dependent) firing rate of each neuron $i = 1, 2, ...N$ as $x_i(t)$. The firing rate population vector $\mathbf{x(t)} = \{\mathbf{x_i(t)}\} \in \mathbf{R^N}$ evolves as:

$$\tau_S \frac{\mathrm{d}\mathbf{x(t)}}{\mathrm{dt}} = -\mathbf{x(t)} + \mathbf{\Phi}(\mathbf{S}\mathbf{x(t)} + \mathbf{I}) \tag{1}$$

where τ_S is the time constant of the population, Φ is a non-linear function whose argument is the total current which contains recurrent dynamics from within the population ($S \cdot \mathbf{x}$ as well I is a current that contains contributions as from external inputs (e.g., sensory stimulation or thalamic input). The goal of the framework is to calculate the synaptic connectivity matrix S based on the following properties. First, the population activity $\mathbf{x(t)}$ can be decomposed onto M distinct activity patterns that are specified $a\ priori$. Second, the patterns result from the dynamical interaction between two populations of neurons selective to two stimuli, denoted A and B. Finally, many neurons participate in the generation of any given pattern. Here we describe these conditions in detail.

Relationship between modes and the population firing rate

The population activity (also referred to as the full neural network dynamics) **x(t)** can be decomposed onto distinct activity patterns that are specified *a priori*. Formally, each pattern represents the projection of $\mathbf{x(t)}$ onto a vector $\mathbf{p_j} \in \mathbf{R^N}$, with $j = 1, 2, ...M \ll N$. Each vector $\mathbf{p_j}$ is referred to as a direction, while each pattern $m_j(t) = \mathbf{x(t)} \cdot \mathbf{p_j}$ is referred to as a 'dynamical mode'. Each dynamical mode is thus a scalar function of time. Importantly, we assume that each mode m_j, which is specified *a priori*, arises from the interaction between neural populations. Motivated by two-alternative forced choice tasks, we will consider modes that might be relevant for computations associated to these tasks.

Dynamical modes as projections from the full neural network dynamics

The full neural network dynamics is given by **x(t)**, which are generated as solutions to Equation 1. To obtain an expression for the modes m_j in terms of the full neural network dynamics, we apply projections to direction $\mathbf{p_j}$ to both sides of Equation 1:

$$\tau_S \frac{\mathrm{d}\mathbf{x}(\mathbf{t})}{\mathrm{dt}} = -\mathbf{x}(\mathbf{t}) + \mathbf{\Phi}(\mathbf{S}\mathbf{x}(\mathbf{t}) + \mathbf{I}) \tag{2}$$

$$\mathbf{p_j} \cdot \tau_\mathbf{S} \frac{\mathrm{d}\mathbf{x}(\mathbf{t})}{\mathrm{dt}} = \mathbf{p_j} \cdot [-\mathbf{x}(\mathbf{t}) + \mathbf{\Phi}(\mathbf{S}\mathbf{x}(\mathbf{t}) + \mathbf{I})] \tag{3}$$

$$\tau_S \frac{\mathrm{dm_j(t)}}{\mathrm{dt}} = -m_j(t) + \mathbf{p_j} \cdot [\mathbf{\Phi}(\mathbf{S}\mathbf{x}(\mathbf{t}) + \mathbf{I})] \tag{4}$$

where we have used the fact that $m_j(t) = \mathbf{x}(\mathbf{t}) \cdot \mathbf{p_j}$. We now define the vector $\mathbf{m}(\mathbf{t}) = \{\mathbf{m_j}(\mathbf{t})\}$ and the matrix $\mathbf{P}$, where each row of the matrix corresponds to each direction $\mathbf{p_j}$: $\mathbf{P} = [\mathbf{p_j}]$. This implies that $\mathbf{m}(\mathbf{t}) = \mathbf{P}\mathbf{x}(\mathbf{t})$ so that $\mathbf{x}(\mathbf{t}) = \mathbf{P^{-1}}\mathbf{m}(\mathbf{t})$. We then rewrite Equation 4 as

$$\tau_S \frac{\mathrm{d}\mathbf{m}(\mathbf{t})}{\mathrm{dt}} = -\mathbf{m}(\mathbf{t}) + \mathbf{P}\mathbf{\Phi}(\mathbf{S}\mathbf{P^{-1}}\mathbf{m}(\mathbf{t}) + \mathbf{I}) \tag{5}$$

The following step is to specify equations for $\mathbf{m}(\mathbf{t}) = \{\mathbf{m_j}(\mathbf{t})\}$ with the goal of obtaining the full neural network synaptic matrix $\mathbf{S}$ and current $\mathbf{I}$.

Dynamical modes as resulting from the dynamical interaction between two populations

We now specify the dynamics of individual modes $m_j(t)$. We focus on modes that are relevant and/or arise from the study of two alternative forced choice (2AFC) tasks. To model two-alternative forced choice tasks, we consider two pools of excitatory neurons that are selective to two different stimuli,e,g. objects, features, or spatial locations. Within each pool there are recurrent connections and the pools inhibit each other through a common pool of interneurons. respectively. We write the following pair of equations, for *reduced* neural populations j and $j+1$

$$\tau_S \frac{\mathrm{dm_j(t)}}{\mathrm{dt}} = -m_j(t) + \phi(J_{Ej} m_j + J_{Ij} m_{j+1} + I_j)$$

and

$$\tau_S \frac{\mathrm{dm_{j+1}(t)}}{\mathrm{dt}} = -m_{j+1}(t) + \phi(J_{E_{j+1}} m_{j+1} + J_{I_{j+1}} m_j + I_{j+1}) \tag{6}$$

We will typically consider the case where modes are coupled pairwise and independent from other pairwise couplings: a mode m_j is exclusively coupled to mode m_{j+1} (via equations 6) and no other mode. This means that m_1 is coupled (exclusively) to m_2, m_3 is coupled (exclusively) to m_4, etc. (Below we will consider a case where a mode m_j is coupled to both m_{j-1} and m_{j+1}). In the scheme described by Eq.6, a mode $m_j(t)$ will be completely specified by the input current I_j and by the recurrent excitatory and inhibitory weights J_{Ej} and J_{Ij}. We can rewrite equation 6 in vector form as

$$\tau_S \frac{\mathrm{d}\mathbf{m}(\mathbf{t})}{\mathrm{dt}} = -\mathbf{m}(\mathbf{t}) + \mathbf{\Phi}(\mathbf{S_m}\mathbf{m}(\mathbf{t}) + \mathbf{I_m}), \tag{7}$$

where $\mathbf{I_m} = \{I_j\}$, $j = 1, 2, ...N$ and S_m is the *reduced* connectivity matrix and has the following block diagonal form:

$$\mathbf{S_m} = \begin{pmatrix} J_{E1} & J_{I1} & 0 & 0 & \cdots \\ J_{I2} & J_{E2} & 0 & 0 & \cdots \\ 0 & 0 & J_{E3} & J_{I3} & \cdots \\ 0 & 0 & J_{I4} & J_{E4} & \cdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

where each block specifies the pairwise interaction between modes. Moreover, each block is symmetric: $J_{E1} = J_{E2},\ J_{E1} = J_{I2},\ J_{E3} = J_{E4},\ J_{I3} = J_{I4}$, etc. We can now equate Eq. 5 with Eq.7, to obtain the following relation:

$$\mathbf{\Phi}(\mathbf{S_m m(t)} + \mathbf{I_m}) = \mathbf{P}\mathbf{\Phi}(\mathbf{S}\mathbf{P^{-1}}\mathbf{m(t)} + \mathbf{I}). \tag{8}$$

We now assume that the reduced connectivity matrix $\mathbf{S_m}$, the current vector $\mathbf{I_m}$, and matrix of projection directions P are known. The problem is to solve (8) for the *full* recurrent connectivity $\mathbf{S}$ and the *full* network current $\mathbf{I}$. It is important to note that, unless P is specified, the scheme above does not lead to a unique solution. Below we constrain the matrix of projection directions P.

Constraints on projection directions $\mathbf{p_j}$: space multiplexing through dynamical modes

Here we consider the computational paradigm of space multiplexing: the possibility that any given neuron participates in the generation of multiple modes. An alternative, is that each neuron is associated to only cortical mode such that the ensemble of units associated to the modes are segregated anatomically. Space multiplexing across neurons sets an important constraint on the directions $\mathbf{p_j}$. Recall that each mode m_j lies along a projection $\mathbf{p_j}$ as $m_j(t) = \mathbf{x(t)} \cdot \mathbf{p_j}$ or equivalently, $\mathbf{m(t)} = \mathbf{P x(t)}$. This last expression can be rewritten in terms of the columns P_i of $\mathbf{P}$ as

$$m_i = \sum_i x_i P_i. \tag{9}$$

From this expression, it becomes clear that for a given unit x_i to participate in more than one mode, then the entries of P_i should not be too sparse. For example, for x_3 to participate in modes m_1 and m_3, then P_2 should be of the form $P_2 = [\alpha, \beta, \gamma, ...]$ and α (entry 1) and β (entry 3) are different from zero. In conclusion, a constraint is that there will be multiplexing if the columns P_i are dense. The degree of density will dictate how many neurons participate in a given mode, and the individual entries P_{ij} will dictate the participation a neuron with activity x_i in mode m_j.

Constraints on projection directions $\mathbf{p_j}$: firing rate distribution

The matrix of projection directions $\mathbf{P}$, the modes vector $\mathbf{m(t)}$ and the reduced connectivity matrix S_m determine the firing rate $\mathbf{x(t)}$ and hence its distribution. An approximate intuition to the actual form and magnitude of the distribution comes from analyzing the expressions for the 'full network': the current $\mathbf{I}$ and the actual rate $x(t)$:

$$\mathbf{I} = \mathbf{P^{-1}}\mathbf{I_m} \tag{10}$$

$$\mathbf{x(t)} = \mathbf{P^{-1}}\mathbf{m(t)}. \tag{11}$$

If $\mathbf{P}$ is an orthogonal matrix, then $\mathbf{P^{-1}} = \mathbf{P^T}$. Similarly as for Eq. 9, we can write the components of the firing rate vector $\mathbf{x(t)}$ in terms of the columns of $P^{-1} = P^T$. Notice that the columns of P_T are precisely the rows of P, which are the projection directions $\mathbf{p_i}$:

$$x_i = \sum_i m_i \mathbf{p_i}. \tag{12}$$

Control of cortical modes through thalamo-cortical projections

Given the generative procedure above to embed cortical modes onto a neural network, we now consider the issue of control. By control, we mean the mechanism by which cortical modes are modified based on some desired computational principle. For example, for a given computation it might be desirable to amplify a particular mode or change its stability. The cortex is reciprocally connected to the thalamus so the thalamus is a source of such control input. We first define and calculate thalamo-cortical inputs that can selectively target a mode. From that analysis, we proceed to define a thalamus model such that an extra-thalamic input from subcortical structures (e.g., superior colliculus, basal ganglia, cerebellum, etc.) results in a modification of the cortical modes.

A model of the thalamus and thalamo-cortical projections

We consider a population of $N_{\text{thal}} < N$ excitatory neurons in the thalamus. Notably, recurrent excitatory collaterals are absent in the thalamus. We denote the firing rate of each neuron $i = 1, 2, ...N_{\text{thal}}$ as $y_i(t)$. The thalamic firing rate vector $\mathbf{y(t)} \in \mathbf{R^{N_{\text{thal}}}}$ evolves as:

$$\tau_{\text{thal}} \frac{d\mathbf{y}}{dt} = -\mathbf{y} + \mathbf{\Phi}_{\text{thal}}(\mathbf{I}_{\text{thal}}) \tag{13}$$

where τ_{thal} is the time constant of the thalamic population, Φ_{thal} is a non-linear function, and $\mathbf{I}_{\text{thal}}$ is a current that contains contributions from extra-thalamic sources (e.g., cortical and subcortical sources).

$$\mathbf{I}_{\text{thal}} = \mathbf{S}_{\text{cx}\to\text{thal}}\mathbf{x(t)} + \mathbf{I}_{\text{sub}} \tag{14}$$

where $S_{\text{cx}\to\text{thal}}$ is the cortico-thalamic connectivity matrix. Analogously, in Eq. 1, we write the external current $\mathbf{I}$ arriving to the cortex as as a sum of sensory and thalamic contributions:

$$\mathbf{I} = \mathbf{I}_{\text{sen}} + \mathbf{I}_{\text{thal}\to\text{cx}} \tag{15}$$

where $I_{\text{thal}\to\text{cx}} = S_{\text{thal}\to\text{cx}}\mathbf{y(t)}$ We calculate the steady state firing rate of the thalamic population $y_s s$ by setting the left-hand side of Eq. 13 to zero:

$$\mathbf{y}_{\text{ss}} = \mathbf{\Phi}_{\text{thal}}(\mathbf{I}_{\text{thal}}) \tag{16}$$

We further assume that the thalamic population reaches a steady state firing rate more quickly than the cortical population. This is reasonable given that there are no recurrent excitatory collaterals in thalamic cells and further requiring that $\tau_{\text{thal}} << \tau_{\text{cx}}$. This allows us to insert $\mathbf{y_{ss}} \approx \mathbf{y(t)}$ in the equation for the cortical dynamics, Eq. 1 to obtain

$$\begin{aligned}
\tau_S \frac{\mathrm{d}\mathbf{x(t)}}{\mathrm{dt}} &= -\mathbf{x(t)} + \mathbf{\Phi}(\mathbf{S}\mathbf{x(t)} + \mathbf{I}_{\text{sen}} + \mathbf{I}_{\text{thal}\to\text{cx}}) && (17)\\
&= -\mathbf{x(t)} + \mathbf{\Phi}(\mathbf{S}\mathbf{x(t)} + \mathbf{I}_{\text{sen}} + \mathbf{S}_{\text{thal}\to\text{cx}}\mathbf{y(t)}) && (18)\\
&= -\mathbf{x(t)} + \mathbf{\Phi}(\mathbf{S}\mathbf{x(t)} + \mathbf{I}_{\text{sen}} + \mathbf{S}_{\text{thal}\to\text{cx}}\mathbf{\Phi}_{\text{thal}}(\mathbf{I}_{\text{thal}})) && (19)\\
&= -\mathbf{x(t)} + \mathbf{\Phi}(\mathbf{S}\mathbf{x(t)} + \mathbf{I}_{\text{sen}} + \mathbf{S}_{\text{thal}\to\text{cx}}\mathbf{\Phi}_{\text{thal}}(\mathbf{S}_{\text{cx}\to\text{thal}}\mathbf{x(t)} + \mathbf{I}_{\text{sub}})) && (20)
\end{aligned}$$

where we used the expression for $\mathbf{I}_{\text{thal}}$ in Eq. 14.

We now focus on the argument of the cortical transfer function Φ, i.e., the second term of the right hand side of Eq. 20. We assume near-linear dynamics for the function Φ_{thal} so that the argument I_{total} is

$$\begin{aligned}
I_{\text{total}} &= \mathbf{S}\mathbf{x(t)} + \mathbf{S}_{\text{thal}\to\text{cx}}\alpha_{\text{thal}}(\mathbf{S}_{\text{cx}\to\text{thal}}\mathbf{x(t)} + \mathbf{I}_{\text{sub}}) + \mathbf{I}_{\text{sen}} && (21)\\
&= [\mathbf{S} + \alpha_{\text{thal}}\mathbf{S}_{\text{thal}\to\text{cx}}\mathbf{S}_{\text{cx}\to\text{thal}}]\mathbf{x(t)} + \alpha_{\text{thal}}(\mathbf{S}_{\text{thal}\to\text{cx}}\mathbf{I}_{\text{sub}}) + \mathbf{I}_{\text{sen}} && (22)\\
&= \mathbf{S}_{\text{eff}}\mathbf{x(t)} + \mathbf{I}_{\text{eff}} + \mathbf{I}_{\text{sen}} && (23)
\end{aligned}$$

where $\mathbf{S}_{\text{eff}} = \mathbf{S} + \alpha_{\text{thal}}\mathbf{S}_{\text{thal}\to\text{cx}}\mathbf{S}_{\text{cx}\to\text{thal}}$ is the effective cortical connectivity matrix and I_{eff} is the effective (non-thalamic) subcortical current. The overlap between the matrices $\alpha_{\text{thal}}S_{\text{thal}\to\text{cx}}$ and $S_{\text{cx}\to\text{thal}}$ determine the recurrent loops between the thalamus and cortex, the de novo generation of modes through thalmo-cortical connections, and the degree to which thalamic control is open or closed- loop, as well as the ge. Similarly, the degree of overlap between between the current I_{sub} and the matrix $S_{\text{thal}\to\text{cx}}$ dictates the contribution from (non-thalamic) subcortical inputs.

General structure of the thalamo-cortical current

For the thalamocortical current $\mathbf{I}_{\text{thal}\to\text{cx}}$ to have an impact on cortical dynamics along a particular direction $\mathbf{p_j}$, the current should be parallel to this direction. Thus, a current $\mathbf{I}_{\text{thal}\to\text{cx}} = \alpha\mathbf{p_j}$ will have an amplifying (reducing) effect on mode m_j if $\alpha > 0$ ($\alpha < 0$). We now sketch the form of the thalamo-cortical projections

such that an extra-thalamic input results in a modification of a given cortical mode m_j. We focus on a particular mode m_j and we assume that there are thalamo-cortical projections (formally a thalamo-cortical matrix S) such that such mode can be targeted. As before, the direction (vector) associated with the mode m_j is $\mathbf{p_j}$. We can construct a thalamo-cortical matrix of the form $S_{\mathrm{thal}\to\mathrm{cx}} = \mathbf{p_j}\mathbf{h_j^T}$ where $\mathbf{h_j} \in \mathbf{R}^{\mathbf{N}_{\mathrm{thal}}}$ is a vector of weights that couple the thalamic output to a mode m_j. Along these lines, a thalamo-cortical current would have the form $\mathbf{p_j}\mathbf{h_j^T}\mathbf{y(t)}$ where $\mathbf{y(t)}$ is the firing rate of the thalamic population. The expression for the for the matrix given a set of mode directions $\mathbf{p_j}$ is thus:

$$S_{\mathrm{thal}\to\mathrm{cx}} = \sum_j \mathbf{p_j}\mathbf{h_j^T} \tag{24}$$

We note that components of $\mathbf{r}_{\mathrm{thal}}$ parallel to h_j will contribute to the thalamo-cortical current, while components orthogonal to h_j will not. In other words, the thalamocortical matrix is such that it effectively 'filters out' directions orthogonal to h_j. This formulation suggests the design of extra-thalamic currents that can serve as control inputs to control dynamical modes. The extra-thalamic control currents that target a particular mode m_j have the form

$$I_{\mathrm{cont}_j} = \alpha_{\mathrm{thal}} h_j.$$

Assuming that $h_j,\ j = 1, 2, ..., M$ form an orthonormal set, a current $I_{\mathrm{cont}_j} = \alpha_{\mathrm{thal}} h_j$ applied to the thalamus will produce an effective thalamo-cortical contribution $I_{\mathrm{thal}\to\mathrm{cx}}$ proportional to $\alpha\alpha_{\mathrm{thal}} M_j$. In addition to supra-threshold mode-specific control, it is also possible to control the cortical output via the baseline thalamic current I_b such that I_b sets the operation point of the thalamus.