Identifying control ensembles for decision-making within the cortico-basal ganglia-thalamic circuit

Catalina Vich[1,2], Matthew Clapp [3,4] Timothy Verstynen[3,4*¤◕], Jonathan E. Rubin[4,5**¤◕],

1 Dept. de Matemàtiques i Informàtica, Universitat de les Illes Balears, Palma, Spain
2 Institute of Applied Computing and Community Code, Palma, Spain
3 Dept. of Psychology & Neuroscience Institute, Carnegie Mellon University, Pittsburgh, USA
4 Center for the Neural Basis of Cognition, Pittsburgh, USA
5 Dept. of Mathematics, University of Pittsburgh, Pittsburgh, USA

¤These authors contributed equally to this work.
◕Corresponding authors
* timothyv@andrew.cmu.edu (TV)
** jonrubin@pitt.edu (JR)

Abstract

During action selection, mammals exhibit a high degree of flexibility in adapting their decisions in response to environmental changes. Although the cortico-basal ganglia thalamic (CBGT) network is implicated in this adaptation, it features a synaptic architecture comprising multiple feed-forward, reciprocal, and feedback pathways, complicating efforts to elucidate the roles of specific CBGT populations in the process of evidence accumulation during decision-making. In this paper we apply a strategic sampling approach, based on Latin hypercube sampling, to explore how CBGT network properties, including subpopulation firing rates and synaptic weights, map to parameters of a normative drift diffusion model (DDM) representing algorithmic aspects of information accumulation during decision-making. Through the application of canonical correlation analysis, we find that this relationship can be characterized in terms of three low-dimensional control ensembles impacting specific qualities of the emergent decision policy: responsiveness (associated with overall activity in corticothalamic and direct pathways), pliancy (associated largely with overall activity in components of the indirect pathway of the basal ganglia), and choice (associated with differences in direct and indirect pathways across action channels). These analyses provide key mechanistic predictions about the roles of specific CBGT network elements in shifting different aspects of decision policies.

Author summary

Mammals are continuously subjected to uncertain situations in which they have to choose among behavioral options. The cortico-basal ganglia-thalamic (CBGT) circuit is a complicated collection of interconnected nuclei believed to strongly influence the ability to adapt to environmental changes. The roles of specific CBGT components in controlling decisions remains uncertain, however. At a more phenomenological level, drift-diffusion models have been shown to be able to reproduce behavioral data (action

selection probabilities and the time needed to make a decision) obtained experimentally from mammals and to provide an abstract representation of a decision policy. In this work, we use simulated decision-making to establish a mapping from neural activity in the CBGT circuit to behavioral outcomes. This mapping illuminates the importance of three core sets of CBGT subnetworks in the action selection process and how they are involved in adapting decision policies across exploitative and exploratory situations.

1 Introduction

Although making a decision can feel instantaneous, decisions in fact arise gradually from the ongoing processing of external (e.g., sensory) and internal (e.g., learned contingencies) information streams. In this process, the inclination toward selecting one action over others is continually updated until sufficient evidence is reached to allow one action, or a set of actions, to proceed [1]. The parameters associated with this process, such as the speed of integration of incoming information and the level of evidence needed to make a decision, define the decision policy [2]. Shortening the window of time available for evidence accumulation often, although not always, leads to faster but more random or exploratory decisions. Lengthening the time for evidence accumulation often leads to slower and more "greedy", or exploitative, decisions. A fundamental challenge for the brain is to manage this speed-accuracy tradeoff, and the associated exploration-exploitation dilemma, via progressive adjustments to the evidence accumulation process, using both current context and prior experience to promote effective outcomes for any given situation [3].

A likely control center for the adaptation of decision policies is the cortico-basal ganglia-thalamic (CBGT) network, or "loop" (see Fig. 1 upper-right). The canonical CBGT circuit includes two structurally and functionally dissociable control streams: the *direct* (facilitation) and *indirect* (suppression) pathways [4]. Central to the canonical model is the assumption that the basal ganglia are organized into multiple action channels [5–9], each containing direct and indirect pathway components. While in reality actions are likely represented in a less thoroughly segregated fashion across CBGT circuits [10], the concept of independent action channels provides conceptual ease when describing the competition between possible actions without changing the key dynamic properties of the underlying computations. The classical view of these pathways [4, 11, 12] is that activation of the direct pathway, via cortical excitation of D1-expressing spiny projection neurons (SPNs) in the striatum, unleashes GABAergic signals that can suppress activity in the CBGT motor output nucleus (internal segment of the globus pallidus, GPi, in primates or substantia nigra pars reticulata, SNr, in rodents) and hence relieves the thalamus from the tonic inhibition that basal ganglia outputs normally provide. This release from inhibition allows the thalamus to facilitate action execution. Conversely, activation of the indirect pathway, via D2-expressing SPNs in the striatum, can alter firing in the external segment of the globus pallidus (GPe) and the subthalamic nucleus (STN) in a way that strengthens basal ganglia inhibitory output. This result suppresses activity of motor pathways and reduces the likelihood of action selection. Finally, the *hyperdirect* pathway (not shown) is activated by direct cortical input to the STN and acts as a fast, reactive "brake" across action channels [13, 14]. More recent experiments have revealed nuances of pathway interactions that go beyond these principles, however; for example, recordings show that indirect and direct pathway SPNs can increase their activity together in the lead-up to a decision [15]. Overall, the topological encoding of actions in the striatum [10], along with the convergence of projections to the GPi/SNr [16], suggests that the direct and indirect pathways may compete for control over the output of the basal ganglia, encoding the "evidence" favoring any behavioral decision as the relative activation of

the two pathways *within* the corresponding action channel [17–19]. Overall, the subtleties of this interaction across pathways and the details of how decision policies emerge from CBGT circuits remain to be determined.

In addition to playing important roles in decision-making [20, 21] (for review see [17, 22]) and proactive decision control [23, 24], the direct and indirect pathways are known to be modulated by dopaminergic feedback from the substantia nigra pars compacta (SNc) to implement a form of reinforcement learning (RL) [25, 26]. Due to the opposing effects of dopamine (DA) on the direct and indirect pathways [27, 28], these feedback signals serve to reinforce rewarded actions while suppressing actions leading to disappointing or undesirable outcomes. Using a spike-timing dependent plasticity (STDP) model of dopaminergic learning, we have previously shown how these plasticity dynamics in striatal SPNs sculpt the balance of direct and indirect pathway influence such that channels associated with more valuable actions exhibit a greater dominance of direct pathway activity over indirect pathway activity, while the opposite occurs for actions with lower values [29]. These STDP-induced changes in direct-vs-indirect pathway influence have been proposed, based on computational evidence, to lead to changes in two specific parameters of the accumulation of evidence process during decision making [30]: the rate of information accumulation (linked with the difference in direct pathway activity between channels) and the threshold of evidence needed to initiate an action (tuned by overall indirect pathway activity, see also [31, 32]). This prior work suggests that the specific synaptic configuration of CBGT pathways, which shapes the firing dynamics of neurons in CBGT subpopulations, sets the qualities of the decision policy by controlling the way that information is used during future decision-making.

The prior work makes clear that within the high-dimensional space of CBGT circuits, there exists a likely mapping between the activity of specific subnetworks and specific properties of the evidence accumulation process. Here, we refer to these subnetworks as *control ensembles* and the corresponding decision properties that they influence as *factors*. Thus, synaptic plasticity, along with other top-down control signals that result in changes in the activity of control ensembles, will adjust factors and hence affect the behavioral outcomes that arise in decision-making tasks. In this work, as a step towards the big picture goal of mapping and understanding the possible control ensembles within CBGT circuits that drive specific decision factors, we first sought to determine what possible tunings of synaptic weights within CBGT pathways would result in firing rates compatible with experimental observations, and hence are biologically reasonable to consider. Second, we identified how the modulation of these individual weights within these ranges would impact decision-related parameters, and hence could serve as building blocks of the control ensembles and overall network tunings that implement various decision policies. For the latter step, we ran simulations in which basal ganglia output firing rates in two competing action channels were used to determine decision outcomes and response times. We fit the distributions of these simulated behavioral variables with the drift diffusion model (DDM) [2, 33], a canonical formalism for the process of evidence accumulation during decision-making. We subsequently used canonical correlation analysis (CCA) to compute a low-dimensional mapping between CBGT synaptic weights and DDM decision parameters. CCA yielded three collections of weights, each of which provides the strongest impact on a corresponding component of the DDM parameter vector and hence acts as the control ensemble for that factor of the decision-making process. This analysis provides new insights about the possible control mechanisms that could be tuned during ongoing behavior, through reward feedback and associated synaptic plasticity, to select decision policies.

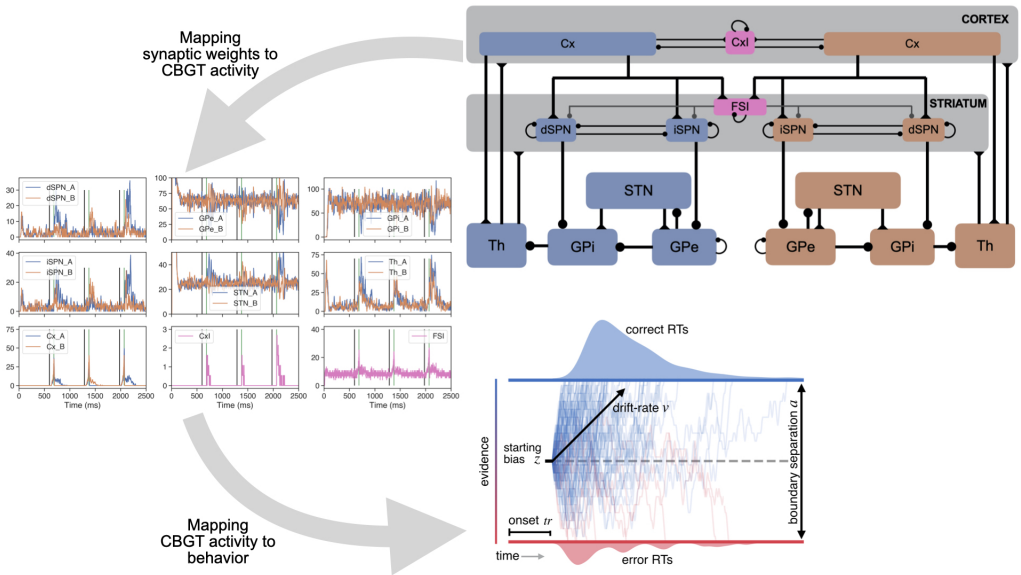

Fig 1. Mapping cortico-basal ganglia-thalamic (CBGT) circuitry to behavioral data. The CBGT network includes neurons from a variety of interconnected populations (see Section 4.1). Blue nodes and orange nodes represent two distinct action channels; pink nodes are those populations that interact with both channels. Connection weights, W, in the CBGT network (upper-right panel) modulate network firing rates, R, (central-left panel) that in turn map to behavior that can be fit by a DDM (lower-right panel) after tuning of its parameters, P. In the central-left panel, black vertical traces indicate the stimulus starting times for three actions while the green vertical lines mark the decision onset times.

2 Results

2.1 CBGT network dynamics and behavior

Our main goal in this work was to establish a mapping between CBGT properties and decision parameters, and thereby to identify control ensembles within CBGT networks. To this end, we simulated a spiking model for the CBGT network that included two action channels (Fig. 1, upper right) and declared an action to be selected when the instantaneous firing rate of the thalamic population for a channel first reached a pre-specified decision threshold (taken for concreteness as 30 Hz [30]). Before we can progress to discussing control ensembles and decision factors, however, we first need to demonstrate that our simulations a) produced realistic dynamics that qualitatively map to known biological observations and b) exhibited a dynamic range of behavior when we performed directed sampling of the underlying synaptic weights. In regards to the dynamics, the integration of the CBGT network with the control conductances g given in Table 2 (2nd column) and with an external input presented simultaneously to the cortical populations Cx for both the A and B channels (see Section 4.1) yielded average firing rates across neurons in each population that agree with those observed experimentally [34–42]. The mean firing rates of all populations as a function of time over a sequence of three actions (B, B, A) are depicted in Fig. 1 (central-left panel).

A clearer picture of how firing rates evolve over the course of deliberation is obtained by averaging over many trials (Fig. 2). We see that the application of a boxcar stimulus to both cortical sensory populations yielded a nonlinear increase in cortical firing rates (upper left) as well as firing rates across striatal populations (dSPN, iSPN, FSI). The

inhibition of SPNs by the higher-firing FSIs allowed FSI activity to rise earlier than SPN firing, but eventually the SPNs did follow the cortical surge. The firing rates of the other populations also remained steady throughout much of the period leading up to an action. In particular, the flat GPi firing rates indicate that its inhibitory inputs from dSPNs and GPe and its excitatory inputs from STN remained balanced, but eventually with this set of conductances and cortical drive, the balance broke, GPi firing decreased, and thalamic firing quickly rose to the decision threshold (cf. [18, 30]). Note that because we had no difference in inputs to the populations Cx_A, Cx_B, both actions had an equal chance of being selected and we expect the firing rates in the A and B subpopulations to be similar for each CBGT region; small differences resulted from randomness in the selection of initial conditions.

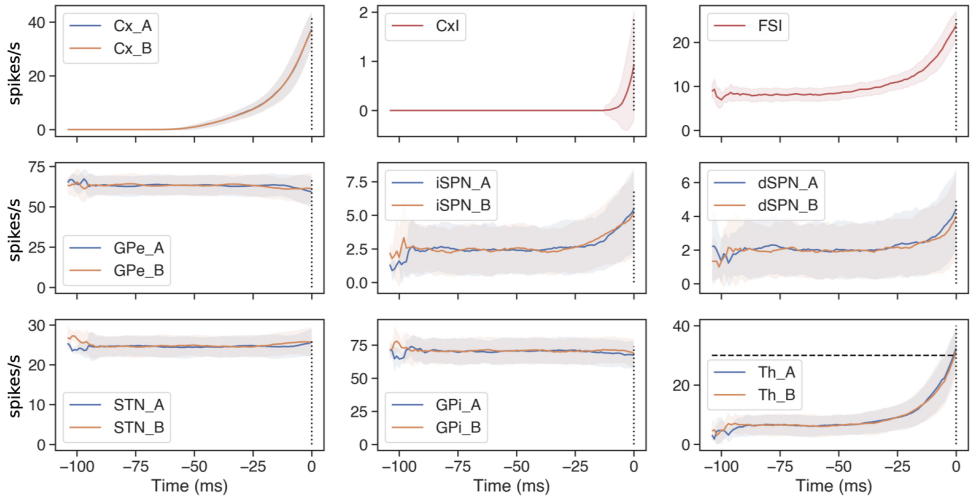

Fig 2. Mean firing rates of the populations in the CBGT network. The firing rate in spikes/s is averaged across 300 different trials with the control set of synaptic conductances (second column in Table 2), from the stimulus onset time at Cx to the decision time (set to be 0, vertical dotted line in each panel). Blue traces correspond to the sub-populations affecting the A action, orange traces represent the sub-populations impacting the B action, and red traces are used for the populations (CxI and FSI) common to both action channels. Shadow areas represent the standard deviation for each population. Since each trial had a different reaction time, we performed the averaging by aligning trials on their decision times.

In our simulations of the baseline network parameters, the synaptic conductances were identical in the A and B channels. Consequently, we expect that the network would have chosen actions A and B with similar likelihood on each trial and with similar frequencies across many trials. Consistent with this assumption, using the baseline parameters, over the course of 300 trials action A was selected 45.3% of the time, with an average reaction time of 86.45 ms across all trials (86.56 ms for A actions, 86.36 ms for B actions).

By independently varying the synaptic conductances of specific connections within the network, but maintaining the same value for each conductance across the two channels, we determined a range of values (see Table 2, 4th and 5th columns) over which neuronal firing rates remained within our pre-specified acceptable ranges (Table 1). This gave us a sampling window for each synapse that would still produce relatively stable and realistic firing rates. We note that these do not represent the largest ranges of values that would preserve these characteristics. Indeed, after we established these maximal ranges, we used them to perform Latin hypercube sampling (LHS; see

Section 4.2) and thus to generate 300 different configurations of network weights. For each permuted network, we simulated 300 choices (i.e., trials). Initially, we found that simultaneously varying parameters within the LHS procedure could push some firing rates out of bounds. Thus, we completed an iterative process of successively shrinking our acceptable ranges of synaptic conductance values and performing LHS until the average pre-selection firing rates were preserved within our prespecified ranges (see Fig. S1).

We next examined the impact of each synapse on the firing rate of the corresponding postsynaptic population, based on the full set of results obtained from varying synaptic conductances via the LHS procedure. Note that the relationship between synaptic weight and postsynaptic rate is not as simple as it sounds, because of the multiplicity of pathways impacting the firing of many populations within the CBGT (Fig. 1, upper right). Fig. 3 shows the degree to which changes in synaptic weights influenced the firing of the post-synaptic neurons, averaged across trials over the full period from trial onset to decision time. This association is represented as a simple linear regression coefficient. The first observation that arises from this calculation is that many post-synaptic populations remained rather insensitive to the variation of synaptic weights. Indeed, while our LHS procedure effectively changed local firing rates of post-synaptic cells, the magnitude of this influence varied substantially across pathways. Second, the directions of influence matched the expected changes given the nature of the synaptic connections involved: changes in inhibitory synapses largely yielded decreases in postsynaptic firing rates, while changes in excitatory synapses induced increases.

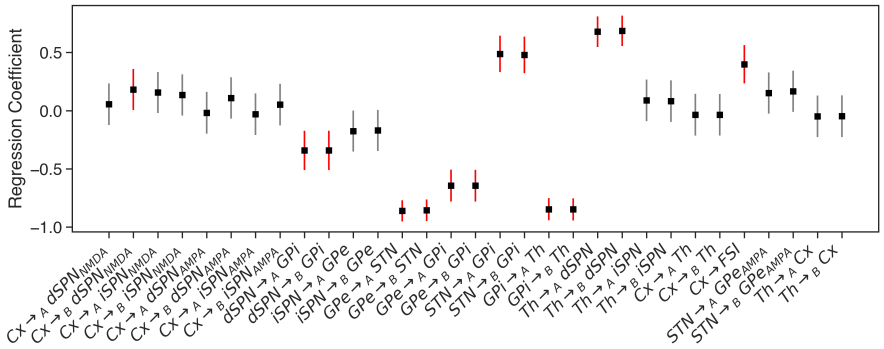

Fig 3. Influence of synaptic weight variation on corresponding postsynaptic firing rates. Black dots represent the estimated regression coefficient for a linear regression model with sampled synaptic weight as an input and post-synaptic unit firing rate as the output. Vertical lines correspond to the 95% confidence interval computed using the standard error of the estimate (gradient) under the assumption of residual normality. Grey error bars correspond to those cases that include the 0 regression coefficient for the slope while red error bars are the others. Both weights and firing rates were z-scored before the analysis.

To verify that our perturbation of synaptic weights also influenced the overt behavior of the network, we looked at how the model behavior varied across the permutations of the synaptic weights. The LHS procedure yielded an approximately normal distribution of average reaction times spread over a relatively broad range, with greatest concentration between 50 and 150 ms (Fig. 4A). This variation is largely consistent with the range of variability seen in humans during speeded RT tasks. That is, while the mean reaction times in our simulations were shorter than what is typically observed in both humans and non-human animals, it is important to note that we were

not simulating either sensory processing time or motor planning and execution processes, which would add approximately $100 - 200\,ms$ to the response time in a full model system (for example, see [43]). As far as the actions selected, across LHS samples, the percentage of A choices remained close to 50% (Fig. 4B), reflecting a relatively unbiased sampling procedure.

The modulation of both choice and reaction time across parameter variations (i.e., unique tunings) of the network, despite *identical* sensory signals across cases, suggests that the LHS sampling effectively impacts how information is used in the deliberation process. To understand how parameter variations modulate the components of this process, we fit each network's choice and reaction time distributions to a hierarchical drift-diffusion model (HDDM, [33]). This procedure returned four separate information processing parameters – boundary height a, drift rate ν, onset time t, and starting bias z – for each version of the network. Histograms indicating the distributions of the estimated DDM parameters (a, ν, t and z) across all permuted networks are shown in Fig. 4C-F.

The first pattern that pops out in these distributions is the presence of bimodality across the a and t parameters, while in contrast the distributions for the drift rate ν and the starting bias z are normally distributed. In particular, the distribution for ν is appropriately centered at 0, which corresponds to an absence of drift towards either option as is suitable for the simulated scenario that lacks the differences in reward or input that would normally impart a direction to the drift rate. The bimodal distributions observed for a and t are related and involve compensation, as can be seen from Fig. 4G. Specifically, the concentration of relatively high a values, which would promote slower decisions, corresponds to relatively low t values, which would accelerate decisions, and vice versa. Moreover, we find a second compensatory relationship in that the largest starting biases, z, arise in combination with strong negative drift rates, within the particle cloud with relatively large t. In contrast, lower z values are positively associated with large values of ν. This suggests the possibility that the specific tunings generated by our LHS procedure give rise to two distinct types of general approaches to decision-making: early onsets of evolution of the process with elevated boundary heights and later onsets with lower boundary heights. Since the directions of the changes between these two parameters have countermanding influences on response speed, the net speed remains approximately conserved across these two clusters of decision policies.

To quantify the clusters that we see visually in Fig. 4G, we applied K-means clustering to the DDM parameters, across network permutations, and computed the silhouette coefficient s (see Section 4.4.2). We found that the best silhouette coefficient is reached for $K = 3$ with value $s = 0.68$, meaning that the DDM parameter space can be clustered into three different groups of parameter values: two subsets that separate high and low values of the onset time and the boundary height, and a third subset that only contains a small set of data points with very high boundary heights that we treat as outliers. As shown in Fig. S2, the aforementioned separation splits the bimodal distributions observed for t and a in Fig. 4E,F, with each cluster effectively representing a different strategy for producing a similar set of RTs and accuracy in the DDM. This analysis largely confirms the clustering patterns we identified visually in the parameter distributions.

2.2 Low-dimensional control ensembles

We used canonical correlation analysis (CCA, see Section 4.4 and [44]) to uncover the low-dimensional relationships between network features and decision parameters that arise as we move up effective levels of analysis in the model. We ran two CCA evaluations, one exploratory and one to address our primary hypothesis of interest; these were a first instance to map synaptic weight schemes W to CBGT firing rates R

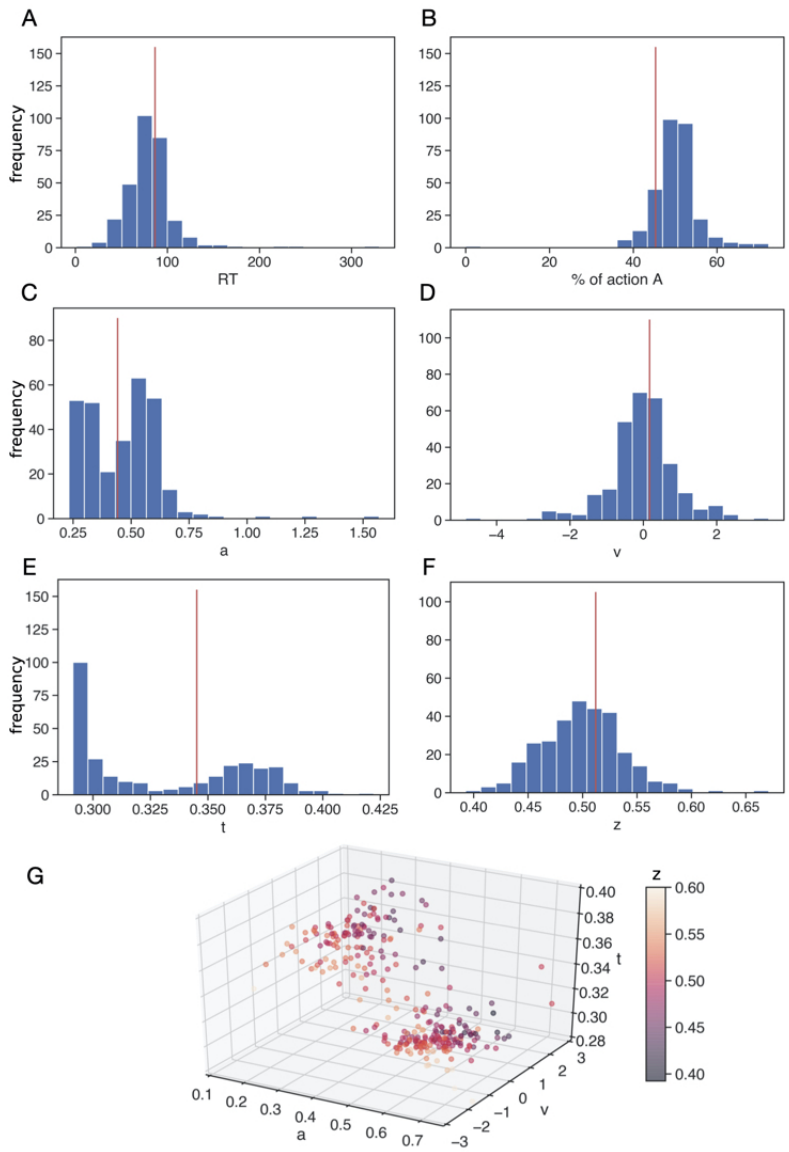

Fig 4. Histograms of simulated behavioral data and best fit DDM parameter values. Panels A and B display the histograms of the reaction times (RT, averaged over 300 trials) and the percentage of choice A (also averaged over 300 trials) across 300 networks generated via the LHS procedure. For each network, the distribution of RTs and accuracy on each trial were used to estimate the DDM parameters (a, ν, t, z), which are represented in histograms in panels C-F. The vertical red lines in panels A-F correspond to the specific value obtained for the permuted network based on the control conductances. Panel G is a 3D representation of the (a, ν, t) parameter space with parameter z represented by the colorbar, which is capped at 0.6 (despite one outlier at $z > 0.65$ and another at $z \in (0.6, 0.65)$; see panel F). In this panel, each dot represents the (a, ν, t, z) vector obtained using the HDDM algorithm for a specific sampled network.

($\mathcal{G} : W \to R$; exploratory evaluation) and a second instance to map R to the DDM parameters P ($\mathcal{H} : R \to P$; hypothesis evaluation). For each evaluation, a 4-fold

cross-validation analysis (see Section 4.4) on each of the data sets $W \to R$ and $R \to P$ showed that a 3-component model best explained variance for $W \to R$ and another 3-component model best explained variance for $R \to P$ (see Fig. S3). Specifically, the best model for the $W \to R$ test explained 20.8% of the variance in the hold out test data, with a standard deviation of 5%. The best model for the $R \to P$ test explained 26.9% of hold out test variance, with a standard deviation of 7%.

Weights to firing rates. The exploratory CCA for $W \to R$ did not reveal any surprising effects and was dominated by a few specific pathways (see Fig. S4 for details). Based on the leading component, the strongest relationship observed represents the straightforward finding that strengthening the inhibitory synapse from GPi to thalamus associates with decreasing thalamic firing rates. The second component resolves a point of minor ambiguity: since GPe and STN form a reciprocal loop, the effects of changing weights of one connection pathway between the two areas are hard to predict, but here we find that increases in the inhibition from GPe to STN associate with decreases in activity in STN and, presumably through the resulting loss of excitatory input, in GPe as well. Interestingly, the third component represents a higher order form of the same relationship found in the first component: it shows that synaptic changes favoring inhibitory over excitatory inputs to the GPi associate with decreasing GPi firing rates and increasing thalamic firing rates. Put together, this exploratory CCA model illustrates that variation in the weights for the primary output nodes of the basal ganglia contribute the most to variability in firing rates. Given, however, the recurrent architecture of the network and its nonlinear properties, more subtle contributions of other synaptic connections to firing rates may simply be washed out by the dominance of GPe, STN, and GPi influences on thalamic nuclei.

Firing rates to decision parameters. The CCA model evaluating our primary hypothesis, that of a relationship between $R \to P$, proved to be more illuminating. For this analysis firing rates were grouped in two ways: 1) overall activity of populations in a CBGT region averaged across channels A and B, and 2) the difference in firing rates in the region between its channel A and channel B populations. Figure 5 shows color-coded representations of the entries, or loadings, of the matrix **U** composed of the firing rate components and of the matrix **V** comprising the DDM parameter components from the second CCA model (see Section 4.4).

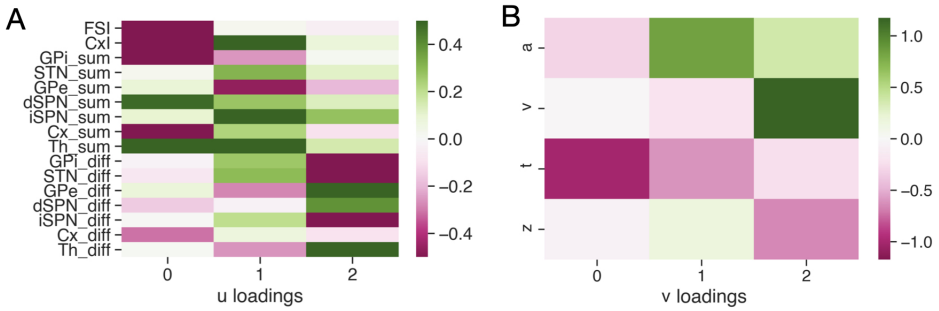

Fig 5. Canonical loading matrices obtained by applying the CCA to the pair $R \to P$. A) The loadings corresponding to the firing rates R. B) The loadings corresponding to the DDM parameters P. In both panels, the values of matrix entries are represented by the color bars.

Right away, we see a clear pattern emerge with three unique components, or factors, in the decision policy (i.e., DDM) space. The first factor has a strong, negative loading on onset time t and a weaker negative loading on boundary height a. This pattern is consistent with a *responsiveness* influence on the decision policy, modulating how

quickly evidence evaluation begins and the overall response speed, with no definitive bias towards one action or another. The second factor, by contrast, manifests as a strong positive loading on boundary height a and a slightly weaker negative loading on onset time t. This pattern is consistent with a *pliancy* component, modulating the degree of evidence necessary to make a decision (i.e., length of evidence accumulation vector) by primarily impacting the domain size for the diffusion process, but not its direction. The final factor is expressed as a positive loading on drift rate ν, with a weaker negative loading on the bias term z. This pattern is consistent with an action *choice* component, modulating the likelihood of selecting a specific option by impacting the direction of the diffusion process. We will examine each of these components in turn.

The *responsiveness* factor was largely associated with the overall activity of corticothalamic systems and the direct pathway projection (Fig. 6A). As expected from the logic of the circuit, greater activity in the thalamic units and the dSPNs, which disinhibit the thalamic units, leads to earlier onset times and, to some extent, lower thresholds (i.e., faster responses). The opposite effect was linked with the overall activity in GPi, FSIs, and cortical units, where greater activity leads to an overall slowing of responses by increasing the onset time and, to some extent, the boundary height. These opposing relationships are consistent with the inhibitory nature of projections from FSIs to dSPNs and from GPi to thalamus, while the cortical impact is more subtle. In contrast, the differential activity between the two channels contributed relatively little to the responsiveness factor. It is worth noting that we display the inhibitory cortical neurons and FSIs in Figure 6 on both the sum and difference diagrams, but there is just one shared population in each of these regions, so the color-coding is the same in both cases.

The *pliancy* factor, on the other hand, was heavily dependent on the overall activity of indirect pathway regions across the two channels (Fig. 6B). More activity in iSPNs, STN neurons (to some extent), thalamic neurons, and cortical inhibitory interneurons, as well as less activity in pallidal units, all associated with an increase of the boundary height and, to a slightly lesser extent, a decrease of the onset time. The former of these changes effectively increases the amount of information necessary to trigger a decision, while the change in onset time moderates the impact on reaction time somewhat. We also found a weaker positive association with overall firing rates in the direct pathway, specifically the dSPNs, along with the cortical units. As with the responsiveness factor, differential activity between channels had a more modest association with changes to the pliancy factor in the DDM process. The minor effects present again loaded more heavily onto indirect pathway nodes.

For the third factor reflecting action *choice*, the situation reverses (Fig. 6C) from what we see in the responsiveness and pliancy factors. Overall activity across channels had very little impact on the parameters of the DDM process related to the direction of information accumulation towards one action or the other. Between-channel differences in firing rates, however, had a robust association. Greater activity of dSPNs, GPe units, and thalamic neurons in channel A, compared to channel B, led to more positive drift rates and slightly lower onset biases. These effects result in a steeper angle of evidence accumulation towards the boundary representing action A, somewhat compensated by a starting point closer to action B. The opposite held for the indirect pathway units, with greater activity of iSPN, STN, and GPi units in channel A leading to lower drift rates and higher onset biases (i.e., angling the drift direction towards the boundary for action B while starting closer to the boundary for A). Put succinctly, the direction of the evidence accumulation process depends on the relative difference in activity across the action channels in all basal ganglia populations.

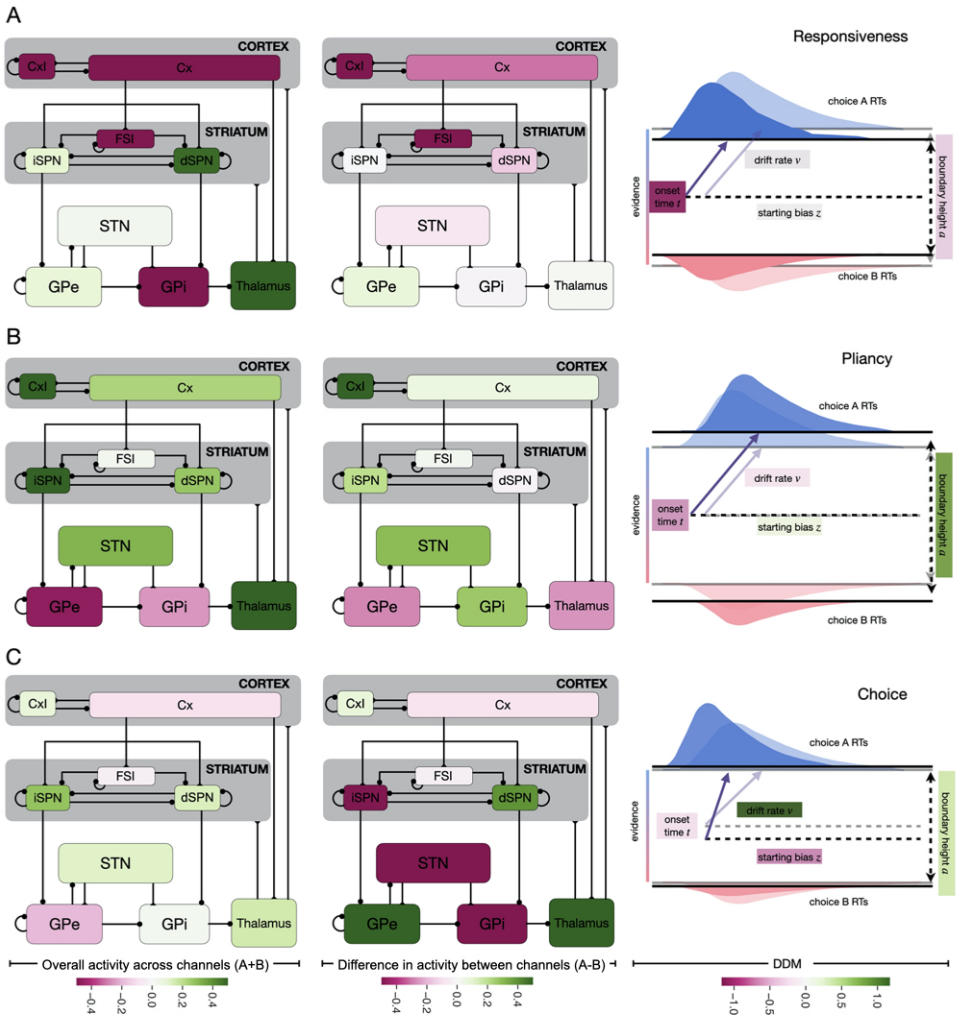

Fig 6. Canonical loadings obtained by applying the CCA to the pair R, P. Panels A, B, and C represent the relation between the firing rate of each population in the CBGT network and the parameters of the DDM obtained in the first, second and third components of the CCA, respectively. 1st column subpanels represent the overall activity summed across both channels while the 2nd column represents the differences in activity between the populations in the A and B channels. 3rd column subpanels depict the associated increases (green) and decreases (magenta) in the DDM parameters.

3 Discussion and Conclusions

The CBGT circuit provides a neural substrate that can control or tune the process of choosing among a set of available actions [5, 45]. After calibrating a computational model of the CBGT network to produce similar firing rates to those observed experimentally, we fit network decision-making behavior with a DDM model and performed CCA to reveal relationships among network synaptic weights, firing rates and the DDM parameters. This analysis demonstrates that different CBGT populations act as control ensembles capable of tuning specific factors that contribute to the action selection process: (1) a responsiveness factor, which highlights complementary changes

in the onset time of evidence accumulation and, to a lesser extent, the boundary separation between actions controlled largely by overall corticothalamic and direct pathway activity; (2) a pliancy factor, which represents antagonistic changes in boundary separation and in onset time impacted primarily by overall indirect pathway activity; and (3) a choice factor, which captures antagonistic changes in the drift rate towards one decision boundary and a shift in the starting position in the DDM domain towards the other decision boundary linked mostly with inter-channel differences in activity between populations.

These factors fit into a broader upwards-mapping conceptual framework [46] that underlies the approach taken in this study (see also [30, 47]). Specifically, by sampling the collection of synaptic weights (W) within the CBGT network (i.e., the *tuning* of the network), we varied the basis set by which *control ensembles* can drive characteristics of the *decision policy* observed in the network's overt behavior. In our framework, the parameters of the DDM (P) provide a description of the information processing dynamics of a given policy, but it was not a priori clear that the CBGT network can adjust its outputs (R) in a way that corresponds to variation of each individual DDM parameter. In practice, we observe that small sets of DDM parameters appear together within the components of the CCA that links firing rates to DDM parameters. We refer to these as the *factors* and they effectively parameterize the space of decision policies. The specific manifestations of factors from the dynamics of underlying control ensembles lead to decision policies with a specific degree of 'greediness' and speed-accuracy tradeoff. Interestingly, these factors were obscured in a CCA applied directly to $W \rightarrow P$, but by performing CCA separately on $W \rightarrow R$ and $R \rightarrow P$, that is the exploratory and hypothesis-focused evaluations respectively, we were able to draw the conclusions on the relationship between network properties and factors of decision policies.

Results obtained from the pliancy and the choice factors, relating to overall indirect pathway activity and inter-channel differences in dSPN activity, respectively, agree with previous computational observations made with corticostriatal synaptic plasticity, which showed that the boundary separation varies with overall iSPN activity while the drift rate varies with the inter-channel difference in dSPN activity [30] (see also reviews [17, 19, 48]). Interestingly, however, our new findings extend this analysis to encompass all of the major basal ganglia populations. In particular our findings indicate that differences across channels in all non-shared CBGT populations correlate with the DDM drift rate (and onset bias), while responsiveness and pliancy factors represent directions in DDM parameter space that are orthogonal to the action choice factor and are driven primarily by changes in onset times and boundary height. The between-channel differences are not independent across all nuclei, of course, as differences in firing across A and B subpopulations in one nucleus will impact the relative firing of A and B subpopulations in all post-synaptic nuclei. This interdependency is reflected in the multiple nuclei with significant loadings in each of the control ensembles identified by the CCA.

While we opted for biological realism in our network design, some abstractions needed to be made for computational efficiency. For example, our CBGT network simulates the GPi output as always suppressing or blocking thalamic activity. Yet this relationship between the GPi and thalamus is more complicated. Evidence suggests that in non-human primates, GPi neurons and their thalamic targets tend to both increase their activity leading up to arm movement in a reaching task [49]. However, this counterintuitive observation may not represent the relationship between GPi and thalamic activity during the learning or decision-making process, since the experiments were made for highly well-trained movements without a choice component. Another biological complication is that direct pathway signals and GPe outputs may not always inhibit GPi/SNr [50, 51], which implies that more complex operational principles may

be at play than those represented in our model.

Of course, the pathways that we included also represent only a subset of the known connections within the broader CBGT circuit. Other populations and pathways, such as arkypallidal neurons [52] or the cortico-STN (i.e., hyper-direct) pathway [53, 54], were not included in our simulations. These may play additional roles as parts of the control ensembles identified here, as well as possible new control ensembles, for example one contributing to a global inhibition factor capable of preventing an action from being performed. We also opted for a primate variant of the CBGT pathways, with GPi used as the output nucleus of the basal ganglia, as opposed to the SNr, which is the primary output nucleus in rodents. It has been shown that SNr targets 42 different brainstem and midbrain regions, which may implement the direct motor impacts of basal ganglia activity, whereas SNr projections to thalamus may represent an efferent copy of the descending signal [55]. Even though our model does not consider this complexity of the output signals from the basal ganglia, in the future it could be modified to include a general brainstem motor target of the output nucleus, where a decision threshold would be imposed, along with one or more separate thalamic targets. If these were tuned similarly and the cortical inputs to the thalamic regions were sufficiently weak, then the model behavior would not be expected to change, despite the physical separation of the decision threshold from the feedback provided from thalamus to striatum and cortex. The weak influence of the cortico-thalamic synaptic weight on thalamic firing rate shown in Fig. 3 supports the likelihood of this robustness, although this would need to be tested to make sure. Increasing the biological realism of the model, however, should be a progressive goal across studies, particularly since our understanding of the anatomy and physiology of these circuits is still exponentially increasing with the advent of new experimental tools that give greater clarity on the underlying circuitry of these pathways.

Despite these limitations, our results clearly show that within the canonical CBGT circuits, specific subpopulations of cells contribute to specific aspects of the information processing capabilities of the network during decision-making. These low-dimensional relationships lead to very specific predictions for future experimental work. For example, global versus differential (i.e., between action channels) stimulation of direct and indirect pathways should manifest as changes in response speed and choice behavior, respectively. Stimulation of the indirect pathway alone, but globally across action representations, should result in contrasting effects on behavior, largely focused on threshold of evidence (i.e., pliancy), since the responsiveness factor is dominated by the direct pathway. These are just a few of the many experimental predictions generated by our results. In this way, the work described here can be seen as a guide for directed hypotheses in future experimental work.

4 Methods

Overall, the goal of this work is to elucidate the complicated nonlinear mapping, call it $\mathcal{F} : W \to P$, from synaptic weights within the CBGT network to parameters in the DDM that are computed by fitting the distribution of choices and reaction times (RT) produced by various tunings of the CBGT network. The way that weight changes translate into different RT distributions is by impacting firing rates throughout the network. Moreover, firing rates are currently much more experimentally accessible than synaptic weights. Hence, we aim to decompose $\mathcal{F}$ into a composition of two relations: the association between firing rates and changes in DDM parameters, $\mathcal{H} : R \to P$, and the association between synaptic weights and network firing rates, $\mathcal{G} : W \to R$; that is, $\mathcal{F} = \mathcal{H} \circ \mathcal{G}$ (see Fig.1 for an overview). In this section, we describe the CBGT network model, the DDM model, and the data analysis that we use to link the two.

4.1 Neural activity: CBGT network

In this study, we simulate behavioral data using a cortico-basal ganglia-thalamic (CBGT) network model adapted from previous works [18, 30, 47]. The network has been designed such that, when a stimulus is presented to the cortex, a decision between two different choices, A or B, is eventually made based on the firing rates of thalamic neurons, which are in turn impacted by neuronal firing throughout the network. The network consists of 8 different populations: the cortical interneurons (denoted by CxI) and the excitatory cortical neurons (Cx); the striatum, which includes the D1 and D2-expressing spiny projection neurons (dSPNs and iSPNs, respectively) and the fast-spiking interneurons (FSIs); the internal and prototypical-external globus pallidus (GPi and GPe, respectively); the subthalamic nucleus (STN); and the thalamus (Th). The model includes two groups of neurons in each population, one for the A channel and one for B, except for the CxI and FSIs, which are shared between the two.

Each neuron evolves according to the integrate-and-fire-or-burst model [18, 56] given by

$$C\frac{dV}{dt} = -g_L(V(t) - V_L) - g_T h(t) H(V(t) - V_h)(V(t) - V_T) - I_{syn}(t) + I_{ext}(t)$$

$$\frac{dh}{dt} = \begin{cases} -h(t)/\tau_h^- & \text{when } V \geq V_h \\ -(1 - h(t))/\tau_h^+ & \text{when } V < V_h \end{cases}$$

where the equation for the membrane potential $V(t)$ includes a leak current with conductance g_L and reversal potential V_L; a low-threshold Ca^{2+} current with maximal conductance g_T, reversal potential V_T, gating variable $h(t)$, and time constants τ_h^+ and τ_h^- representing the rate of change of the gating variable $h(t)$ before and after the membrane potential reaches a certain constant voltage threshold V_h, respectively; a synaptic current, $I_{syn}(t)$; and an external current $I_{ext}(t)$.

The synaptic current itself includes excitatory $AMPA$ and $NMDA$ and inhibitory $GABA$ components, such that

$$\begin{aligned} I_{syn} &= g_{AMPA} s_{AMPA}(t)(V(t) - V_E) + \frac{g_{NMDA} s_{NMDA}(t)(V(t) - V_E)}{1 + e^{-0.062V(t)/3.57}} \\ &+ g_{GABA} s_{GABA}(t)(V(t) - V_I), \end{aligned}$$

where each g_i denotes the maximal net channel conductance for $i \in \{AMPA,\ NMDA,\ GABA\}$. V_E and V_I are the reversal potentials for excitation and inhibition, respectively, and $s_i(t)$ are open channel fractions, with dynamics given by

$$\begin{aligned} \frac{ds_{AMPA}}{dt} &= \sum_j \delta(t - t_j) - \frac{s_{AMPA}}{\tau_{AMPA}}, \\ \frac{ds_{NMDA}}{dt} &= \alpha(1 - s_{NMDA}) \sum_j \delta(t - t_j) - \frac{s_{NMDA}}{\tau_{NMDA}}, \\ \frac{ds_{GABA}}{dt} &= \sum_j \delta(t - t_j) - \frac{s_{GABA}}{\tau_{GABA}}, \end{aligned}$$

for spike onset times t_j, rate constant α, and decay rates τ_i for each choice of i. Note that because NMDA has the slowest decay rate, we design the s_{NMDA} equation to explicitly prevent this variable from exceeding one.

The external current, which is used to tune the baseline firing rate of each population, is given by

$$I_{ext}(t) = S_{ext,AMPA}(V(t) - V_E) + S_{ext,GABA}(V(t) - V_I)$$

where $S_{ext,X}$ is a mean-reverting random walk depending on the external input frequency, the efficacy of the external connections, and the number of external connections (see Table 3).

The neurons within a population that correspond to the same choice are connected to each other and, in some cases, to those in FSI and CxI, resulting in two different channels, which we call channel A and channel B. To simplify notation, we will label a specific population with the subscript A or B if we want to specify that we consider the A-channel sub-population or the B-channel one, respectively. The established synaptic pathways between regions that we include in the model can be found in Fig. 1 (upper-right). In all simulations, the synaptic conductances for the two channels are identical. Within each channel, the connections from Cx to $dSPN$ and from Cx to $iSPN$ are equal in our control parameter set, but we allow these to differ in subsequent simulations.

While Fig. 1 (upper-right) shows the pathways between populations, the individual neurons within populations are connected with specified probabilities and synaptic conductances g in nS (which we also call weights). The connection probabilities and weight ranges considered, which we calibrated to obtain similar firing rates to those observed experimentally [34–42] during resting states (baseline) and decision processes (Table 1), are presented in Table 2.

Population	**baseline FR range (Hz)**	**full FR range (Hz)**	**References**
dSPN	$[0, 5]$	$[0, 35]$	[34–37]
iSPN	$[0, 5]$	$[0, 35]$	[34–37]
GPe	$[40, 90]$	$[40, 150]$	[38–40]
GPi	$[40, 90]$	$[40, 150]$	[40]
STN	$[10, 35]$	$[10, 55]$	[38–40]
Th	$[5, 20]$	$[5, 85]$	[41]
Cx		$[0, 100]$	[42]
FSI	$[5, 40]$	$[5, 70]$	[42]

Table 1. Firing rate ranges during baseline and during decision tasks for different populations in the brain. These ranges reflect experimental data from both primates and rats.

To represent the presentation of a stimulus to the network, we increase the external input frequency for Cx_A and Cx_B from $2.2\,Hz$ to $2.5\,Hz$. Subsequently, both channels compete to make a decision. If the thalamic firing rate associated with channel A reaches a specific threshold, which we take as $30\,Hz$, before the thalamic firing rate associated with channel B, then we say that decision A is selected, with a similar condition for decision B. We define the time from the stimulus onset to the decision as the *reaction time* for a trial. If the decision thresholds for both channels are not reached in a time window of $800\,ms$, then we say that no decision has been performed and we end the trial.

Unlike [29, 47], in the version of the network implemented in this work, spike-timing-dependent plasticity from the cortex to the spiny projection neurons is not included. The rest of the structure and parameters remain as in [47] except for two changes. First, the leak conductance of thalamic neurons, g_L, is decreased from 25 to $18\,nS$ to fit experimental information about their firing rate under baseline conditions. Second, after a decision is made, the external input presented to cortex as a stimulus is maintained for $300\,ms$, as in the earlier work, but here we only maintain this input at 75% of its original strength.

4.2 Network specification and network behavior

By modifying the values of the maximal conductances or weights g of the connections from one population to another, we obtain different configurations of the network that we call the network *tunings*. Given the large number of weights in the full model, for tractability, we only explore variations across the main feed-forward connections in the CBGT network; in simulations, we also see that these connections have the strongest influence on reaction times. The weights for the rest of the connections in the network are kept constant. Therefore, each tuning is generated by considering different values of the 14 specific connections highlighted in bold in Table 2 (1st column).

For each varied connection type, we identify the interval of weights over which the population firing rate remains in our allowed range, thus establishing upper and lower bounds on the g values to be considered for that connection (see Table 2, 4th and 5th columns).

We performed Latin hypercube sampling (LHS) on values of the 14 varied weights to randomly specify $N = 300$ different weight configurations, and thus 300 unique network tunings. In brief, for each connection, we partitioned the allowed range into 300 bins of the same size. For each of our 300 iterations, we selected one bin uniformly at random for each weight, independently across weights. These selections were made without replacement, so that each bin was used exactly once, resulting in 300 multi-dimensional bins. Within each multi-dimensional selected bin, we randomly specify 14 weight values that together correspond to one configuration of the network. To perform the LHS we use the lhsmdu function implemented in Python.

Finally, for each tuning, we simulated 300 decision trials, each of which ended when one of the thalamic subpopulations reached the decision threshold or when the allowed time for the decision expired, if no decision had been reached by that time; note that the use of 300 for both the number of trials and the number of LHS bins is inconsequential. On each trial on which a decision was made, we considered as the CBGT output the reaction time from the introduction of the stimulus to the decision; the average firing rates of all populations over that time period; and the decision made.

4.3 Drift-diffusion model

For each tuned network, we fit the simulated behavioral data from the CBGT network, specifically the reaction time and the choice made on each trial, with the drift-diffusion model (DDM), a stochastic model that instantiates the statistically optimal performance on two-alternative decision tasks (see [30, 57, 58]). The DDM assumes that decisions are made by an accumulative stochastic process. Given two different boundaries, one for each choice, which are separated by a distance called the decision threshold (a), and given also a specific starting bias (z), the accumulated evidence θ is a function of time τ, defined from $\tau = 0$ up to a stopping time that occurs when θ reaches one of the two thresholds, determined by the following stochastic differential equation and starting condition:

$$\begin{aligned} d\theta &= \nu d\tau + \sigma dW \quad && \text{if } \tau > t \\ \theta &= z/a && \text{if } \tau \leq t \end{aligned}$$

where ν denotes the rate of evidence accumulation, t is the time of onset of evidence accumulation, and σ represents the level of noise in the process, which is given by the standard deviation of a white noise process W. For each trial, the choice and the reaction time for the DDM are determined by which boundary is reached and the time elapsed between the onset time t and the moment when θ reaches one of the boundaries, respectively.

Connection type	Connection Probability	$g(nS)$ control	Lower g bound	Upper g bound	Receptor(s)
Cx-Cx	0.43	0.0127			AMPA
Cx-Cx	0.43	0.15			NMDA
Cx-CxI	0.2417	0.113			AMPA
Cx-CxI	0.2417	0.525			NMDA
CxI-Cx	1.00	1.75			GABA
CxI-CxI	1.00	3.5833			GABA
Cx-dSPN	1.00	0.027	0.012	0.033	NMDA
Cx-iSPN	1.00	0.027	0.0195	0.0885	NMDA
Cx-dSPN	1.00	0.018	0.008	0.022	AMPA
Cx-iSPN	1.00	0.0986	0.008	0.059	AMPA
Cx-Th	1.00	0.035	0.008	0.0548	NMDA, AMPA
Cx-FSI	1.00	0.198	0.1905	0.63	AMPA
dSPN-dSPN	0.45	0.28			GABA
dSPN-iSPN	0.45	0.28			GABA
dSPN-GPi	1.00	2.09	0.418	2.413	GABA
iSPN-dSPN	0.5	0.28			GABA
iSPN-iSPN	0.45	0.28			GABA
iSPN-GPe	1.00	4.07	2.47	4.46	GABA
GPe-STN	0.067	0.35	0.33	0.39	GABA
GPe-GPe	0.067	1.75			GABA
GPe-GPi	1.00	0.06	0.05733	0.067	GABA
GPi-Th	1.00	0.3315	0.32017	0.357	GABA
FSI-dSPN	1.00	1.7776			GABA
FSI-iSPN	1.00	1.66987			GABA
FSI-FSI	1.00	3.2583			GABA
STN-GPe	0.1617	0.07	0.05	0.1	AMPA
STN-GPe	0.1617	1.51			NMDA
STN-GPi	1.00	0.038	0.036	0.03833	NMDA
Th-Cx	0.83	0.03	0.021	0.035	NMDA
Th-CxI	0.83	0.015			NMDA
Th-dSPN	1.00	0.3825	0.0015	0.3915	AMPA
Th-iSPN	1.00	0.3825	0.3525	0.3975	AMPA
Th-FSI	0.83	0.1			AMPA

Table 2. CBGT connectivity. From left to right, the connections included, specified in terms of the pre- and the post-synaptic populations that they link; the probability that two cells in the pre- and post-synaptic populations will be synaptically linked; a fixed synaptic conductance or weight (measured in nS) used in our control case; the lower and upper bounds on weights based on maintaining realistic neuronal firing rate; and finally the receptor types for each specific connection. Those connections that are not in bold font (i.e. without an upper and lower bound) are fixed to the control value during all simulations. The topology of each connection, which can be either diffuse (they project to left and right action channels) or focal (restricted connections within each channel), has not changed from [30](Table 3).

In our simulations, for each tuned network, we estimate the values of the quadruple (a, ν, t, z) for which the DDM behavior best fits the reaction times and the selected choices obtained with the CBGT network. For this purpose, we use the Hierarchical Drift-Diffusion Model (HDDM) implemented in Python (see [33]) using as inputs both the reaction time (RT) and the specific decision made on each successful trial (i.e.,

distinguishing between A and B), after first removing the trials where no decision was made.

4.4 Data analysis

We consider three different sets of data: the weights selected that define the various CBGT network configurations, W; the averaged firing rates obtained from all populations of the CBGT network, R; and the DDM parameter values obtained using the HDDM, P. For clearer visualization and analysis of the results, we considered the averaged firing rates of each CBGT population summed across both channels, A and B, and the difference in the averaged firing rates of the populations between the two channels. Since there are 7 populations in each channel plus two populations that the channels share, R is a 300×16 data table, while the dimensions of W and P are 300×14 and 300×4, respectively. Once we form these matrices, we perform additional analyses to investigate relations between them.

4.4.1 Canonical Correlation Analysis

We use canonical correlation analysis (CCA, [44]) to infer the relation between the pair of data sets (W, R) and again to infer information relating the pair of data sets (R, P). In general, given two data sets $\mathbf{X} = (x_1, \ldots, x_n)^\top$ and $\mathbf{Y} = (y_1, \ldots, y_m)^\top$, CCA searches for vectors $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^m$ such that the correlation between $\mathbf{u}^\top\mathbf{X}$ and $\mathbf{v}^\top\mathbf{Y}$ is maximal. The vectors $\mathbf{u}, \mathbf{v}$ are the first pair of canonical variables. Subsequently, for each $i = 2, \ldots, N$, we compute the ith pair of canonical variables by maximizing the same correlation subject to the restriction that the new pair is uncorrelated with the previous $i - 1$. We call N the *number of components* and select it to be less than or equal to $\min\{n, m\}$. To compute the CCA, we use the *CCA* function of the *sklearn* package in Python. To discern the minimum number of components necessary to fit the data, we consider the number of components providing the best R^2 score for the predicted data, which is computed for the *CCA* function. This score is defined as $1 - a/b$ for a equal to the sum of $(y_{true} - y_{pred})^2$ and b equal to the sum of $(y_{true} - \bar{y}_{true})^2$ over all data points, where for each point, y_{true} denotes the true data value and y_{pred} the predicted value from the dimension reduction learned during the CCA. This expression quantifies the extent to which the CCA represents the data, relative to chance. The best possible fit would yield an R^2 score equal to 1; notice that, although a 0 score represents a poor fit, the score can in fact be negative, meaning that the model based on CCA is worse than would be expected by chance.

To ensure that we do not obtain spurious results based on our data sample, we apply a 4−fold cross-validation analysis. In this step, we split the 300 tuned networks into two different blocks: the training set, containing 75% of the tuned networks, and the testing set, comprising the remaining 25% of the data. To achieve this splitting, we order the tuned networks from 1 to 300 and then in each i-fold, with $i \in \{1, 2, 3, 4\}$, we consider the tunings in positions $(75(i-1), 75i]$ as the testing set and the rest as the training set. Hence, for different numbers of total components, we compute the CCA using the training set and we test the resulting model using the testing set, from which we compute the CCA score.

Conceptually, the canonical variables expose the variable groupings that most significantly contribute to correlations between specific data sets. However, they are subject to variability across samples and can be highly affected by multicollinearity. Hence, to interpret the CCA results, we consider canonical loadings, which represent the correlations between the original variables and the canonical variables.

4.4.2 K-means clustering

To identify different clusters of data in the HDDM output, we performed K-means clustering using the *KMeans* function of the *sklearn* package in Python. For each tuned network, we considered the point in $\mathbb{R}^4$ given by the (a, ν, t, z) obtained from DDM; together, these comprise a data set of 300 points. We separated these data into K different clusters such that each observation lies in the cluster with the nearest mean, minimizing the variance within clusters.

Before using the clustering technique, we performed a sensitivity-based normalization of the HDDM output (a, ν, z, t). In this procedure, each component $x \in \{a, \nu, z, t\}$ is modified as

$$x \leftarrow (x - \bar{x})\Delta_x$$

where $\bar{x}$ refers to the mean of x over the different tuned networks and Δ_x is the centered difference formula to compute the change in the reaction time relative to the change in x, that is

$$\Delta_x = \frac{RT(x+h) - RT(x-h)}{2h},$$

where $h = 0.1$ and $RT(y)$ is the mean reaction time obtained after running the HDDM 10^5 times with input variable y and fixed values of the components in $\{a, v, z, t\} \setminus y$.

To determine the optimal number of clusters into which to split the data, we subject the data to K-means clustering for each $K \in [1, 10] \subset \mathbb{N}$. In each case, we compute the mean of the silhouette coefficients over samples. The silhouette coefficient is a measure of the proximity of each cluster point to neighboring cluster points. This coefficient is calculated as $(b - a)/\max\{a, b\}$ where a denotes the mean distance between points within a cluster and b is the distance between a data point and the nearest cluster to which it does not belong. This calculation yields a value lying in $[-1, 1]$, where a value close to 1 indicates that the specific sample is not close to the neighboring clusters and hence it has been assigned to the correct cluster, a value close to 0 indicates that the specific sample is very close to or on the boundary defining the different clusters, and negative values close to -1 indicate that the specific sample could be wrongly assigned. Hence, higher values of the mean silhouette coefficient indicate a better clustering performance. We compute the silhouette coefficient using the *Silhouette_score* function in the *sklearn* package in Python.

The rest of the parameters used in the K-means clustering process are standard values of the *KMeans* function in the *sklearn* package in Python. That is, 10 repetitions of the K-means algorithm are run using different centroid seeds, with a maximum of 300 iterations per run. To check for convergence of the method, the Frobenius norm of the difference in the cluster centers of two consecutive iterations is evaluated by the KMeans function itself against a relative tolerance of 10^{-4}. The K-means problem is solved using Elkan's algorithm with precomputed distances, 0 verbosity mode and "None" as the random state.

5 Acknowledgements

CV is supported by the PCI2020-112026 project, funded by MCIN/AEI/10.13039/501100011033 and by the European Union "NextGenerationEU"/PRTR as part of the CRCNS program. TV and JER are partly supported by NIH award R01DA053014 as part of the CRCNS program.

References

1. Gold JI, Shadlen MN. The neural basis of decision making. Annu Rev Neurosci. 2007;30(30):535–561.
2. Ratcliff R. A theory of Memory Retrival. Psychol Rev. 1978;85(2):59–108.
3. Schulz E, Gershman SJ. The algorithmic architecture of exploration in the human brain. Current opinion in neurobiology. 2019;55:7–14.
4. Albin RL, Young AB, Penney JB. The functional anatomy of basal ganglia disorders. Trends in Neurosciences. 1989;12(10):366–375.
5. Mink JW. The basal ganglia: focused selection and inhibition of competing motor programs. Progress in neurobiology. 1996;50(4):381–425.
6. Alexander GE, Crutcher MD. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. Trends Neurosci. 1990;13(7):266–271.
7. Gurney K, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. I. A new functional anatomy. Biological Cybernetics. 2001;84(6):401–410.
8. Humphries MD, Stewart RD, Gurney KN. A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. Journal of Neuroscience. 2006;26(50):12921–12942.
9. Bogacz R, Gurney K. The basal ganglia and cortex implement optimal decision making between alternative actions. Neural Comput. 2007;19(2):442–477.
10. Klaus A, Martins GJ, Paixao VB, Zhou P, Paninski L, Costa RM. The Spatiotemporal Organization of the Striatum Encodes Action Space. Neuron. 2017;95(5):1171 – 1180.e7. doi:https://doi.org/10.1016/j.neuron.2017.08.015.
11. Nambu A, Tokuno H, Takada M. Functional significance of the cortico–subthalamo–pallidal 'hyperdirect'pathway. Neuroscience research. 2002;43(2):111–117.
12. Kravitz AV, Tye LD, Kreitzer AC. Distinct Roles for direct and indirect pathway striatal neurons in reinforcement. Nat Neurosci. 2012;15(6):816–818.
13. Fife KH, Gutierrez-Reed NA, Zell V, Bailly J, Lewis CM, Aron AR, et al. Causal role for the subthalamic nucleus in interrupting behavior. Elife. 2017;6.
14. Aron AR, Poldrack RA. Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. J Neurosci. 2006;26(9):2424–2433.
15. Cui G, Jun SB, Jin X, Pham MD, Vogel SS, Lovinger DM, et al. Concurrent activation of striatal direct and indirect pathways during action initiation. Nature. 2013;494(7436):238–242.
16. Smith Y, Bevan M, Shink E, Bolam JP. Microcircuitry of the direct and indirect pathways of the basal ganglia. Neuroscience. 1998;86(2):353–387.
17. Dunovan K, Lynch B, Molesworth T, Verstynen T. Competing basal ganglia pathways determine the difference between stopping and deciding not to go. Elife. 2015;4:e08723.

18. Wei W, Rubin JE, Wang XJ. Role of the indirect pathway of the basal ganglia in perceptual decision making. J Neurosci. 2015;35(9):4052–4064.

19. Bariselli S, Fobbs W, Creed M, Kravitz A. A competitive model for striatal action selection. Brain Research. 2018;doi:https://doi.org/10.1016/j.brainres.2018.10.009.

20. Ding L, Gold JI. Caudate encodes multiple computations for perceptual decisions. J Neurosci. 2010;30(47):15747–15759.

21. Yartsev MM, Hanks TD, Yoon AM, Brody CD. Causal contribution and dynamical encoding in the striatum during evidence accumulation. Elife. 2018;7:e34929.

22. Ding L, Gold JI. The basal ganglia's contributions to perceptual decision making. Neuron. 2013;79(4):640–649.

23. Dunovan K, Verstynen T. Believer-Skeptic meets Actor-Critic: Rethinking the role of basal ganglia pathways during decision-making and reinforcement learning. Frontiers in neuroscience. 2016;10:106.

24. Majid DSA, Cai W, Corey-Bloom J, Aron AR. Proactive selective response suppression is implemented via the basal ganglia. J Neurosci. 2013;33(33):13259–13269.

25. Hollerman JR, Schultz W. Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci. 1998;1(4):304–309.

26. Perrin E, Venance L. Bridging the gap between striatal plasticity and learning. Current opinion in neurobiology. 2019;54:104–112.

27. Shan Q, Ge M, Christie MJ, Balleine BW. The acquisition of goal-directed actions generates opposing plasticity in direct and indirect pathways in dorsomedial striatum. Journal of Neuroscience. 2014;34(28):9196–9201.

28. Collins AGE, Frank MJ. Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol Rev. 2014;121(3):337–366.

29. Vich C, Dunovan K, Verstynen T, Rubin J. Corticostriatal synaptic weight evolution in a two-alternative forced choice task: a computational study. Communications in Nonlinear Science and Numerical Simulation. 2020;82:105048. doi:https://doi.org/10.1016/j.cnsns.2019.105048.

30. Dunovan K, Vich C, Clapp M, Verstynen T, Rubin J. Reward-driven changes in striatal pathway competition shape evidence evaluation in decision-making. PLoS computational biology. 2019;15(5):e1006998.

31. Ratcliff R, Frank MJ. Reinforcement-Based Decision Making in Corticostriatal Circuits: Mutual Constraints by Neurocomputational and Diffusion Models. Neural Comput. 2012;24:1186–1229.

32. Frank MJ, Gagne C, Nyhus E, Masters S, Wiecki TV, Cavanagh JF, et al. fMRI and EEG predictors of dynamic decision parameters during human reinforcement learning. J Neurosci. 2015;35(2):485–494.

33. Wiecki TV, Sofer I, Frank MJ. HDDM: hierarchical bayesian estimation of the drift-diffusion model in python. Frontiers in neuroinformatics. 2013;7:14.

34. Kimura M. Behavioral modulation of sensory responses of primate putamen neurons. Brain research. 1992;578(1-2):204–214.

35. Aosaki T, Graybiel AM, Kimura M. Effect of the nigrostriatal dopamine system on acquired neural responses in the striatum of behaving monkeys. Science. 1994;265(5170):412–415.

36. Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature. 2005;437(7062):1158–1161.

37. Panigrahi B, Martin KA, Li Y, Graves AR, Vollmer A, Olson L, et al. Dopamine is required for the neural representation and control of movement vigor. Cell. 2015;162(6):1418–1430.

38. Pavlides A, Hogan SJ, Bogacz R. Computational models describing possible mechanisms for generation of excessive beta oscillations in Parkinson's disease. PLoS computational biology. 2015;11(12):e1004609.

39. Tachibana Y, Iwamuro H, Kita H, Takada M, Nambu A. Subthalamo-pallidal interactions underlying parkinsonian neuronal oscillations in the primate basal ganglia. European Journal of Neuroscience. 2011;34(9):1470–1484.

40. Nambu A, Tachibana Y. Mechanism of parkinsonian neuronal oscillations in the primate basal ganglia: some considerations based on our recent work. Frontiers in systems neuroscience. 2014;8:74.

41. Pessiglione M, Guehl D, Rolland AS, François C, Hirsch EC, Féger J, et al. Thalamic neuronal activity in dopamine-depleted primates: evidence for a loss of functional segregation within basal ganglia circuits. Journal of Neuroscience. 2005;25(6):1523–1531.

42. de Lafuente V, Jazayeri M, Shadlen MN. Representation of accumulating evidence for a decision in two parietal areas. Journal of Neuroscience. 2015;35(10):4306–4318.

43. Weinrich M, Wise S, Mauritz KH. A neurophysiological study of the premotor cortex in the rhesus monkey. Brain. 1984;107(2):385–414.

44. Wang HT, Smallwood J, Mourao-Miranda J, Xia CH, Satterthwaite TD, Bassett DS, et al. Finding the needle in a high-dimensional haystack: Canonical correlation analysis for neuroscientists. NeuroImage. 2020;216:116745.

45. Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annual review of neuroscience. 1986;9(1):357–381.

46. Frank MJ. Linking Across Levels of Computation in Model-Based Cognitive Neuroscience. In: An Introduction to Model-Based Cognitive Neuroscience. Springer, New York, NY; 2015. p. 159–177.

47. Rubin JE, Vich C, Clapp M, Noneman K, Verstynen T. The credit assignment problem in cortico-basal ganglia-thalamic networks: A review, a problem and a possible solution. European Journal of Neuroscience. 2021;53(7):2234–2253.

48. Mikhael JG, Bogacz R. Learning Reward Uncertainty in the Basal Ganglia. PLoS Comput Biol. 2016;12(9):e1005062.

49. Schwab BC, Kase D, Zimnik A, Rosenbaum R, Codianni MG, Rubin JE, et al. Neural activity during a simple reaching task in macaques is counter to gating and rebound in basal ganglia–thalamic communication. PLoS biology. 2020;18(10):e3000829.

50. Freeze BS, Kravitz AV, Hammack N, Berke JD, Kreitzer AC. Control of basal ganglia output by direct and indirect pathway projection neurons. J Neurosci. 2013;33(47):18531–18539.

51. Phillips RS, Rosner I, Gittis AH, Rubin JE. The effects of chloride dynamics on substantia nigra pars reticulata responses to pallidal and striatal inputs. Elife. 2020;9:e55592.

52. Mallet N, Micklem BR, Henny P, Brown MT, Williams C, Bolam JP, et al. Dichotomous Organization of the External Globus Pallidus. Neuron. 2012;74(6):1075–1086.

53. Nambu A, Tokuno H, Hamada I, Kita H, Imanishi M, Akazawa T, et al. Excitatory cortical inputs to pallidal neurons via the subthalamic nucleus in the monkey. Journal of neurophysiology. 2000;84(1):289–300.

54. Baladron J, Nambu A, Hamker FH. The subthalamic nucleus-external globus pallidus loop biases exploratory decisions towards known alternatives: a neuro-computational study. European Journal of Neuroscience. 2017; p. 1–14. doi:10.1111/ejn.13666.

55. McElvain LE, Chen Y, Moore JD, Brigidi GS, Bloodgood BL, Lim BK, et al. Specific populations of basal ganglia output neurons target distinct brain stem areas while collateralizing throughout the diencephalon. Neuron. 2021;109(10):1721–1738.

56. Smith GD, Cox CL, Sherman SM, Rinzel J. Fourier analysis of sinusoidally driven thalamocortical relay neurons and a minimal integrate-and-fire-or-burst model. J Neurophysiol. 2000;83(1):588–610.

57. Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychological Review. 2006;113(4):700.

58. Ratcliff R, Smith PL, Brown SD, McKoon G. Diffusion Decision Model: Current Issues and History. Trends Cogn Sci. 2016;20(4):260–281.

A Supporting information

A.1 Supplementary figures

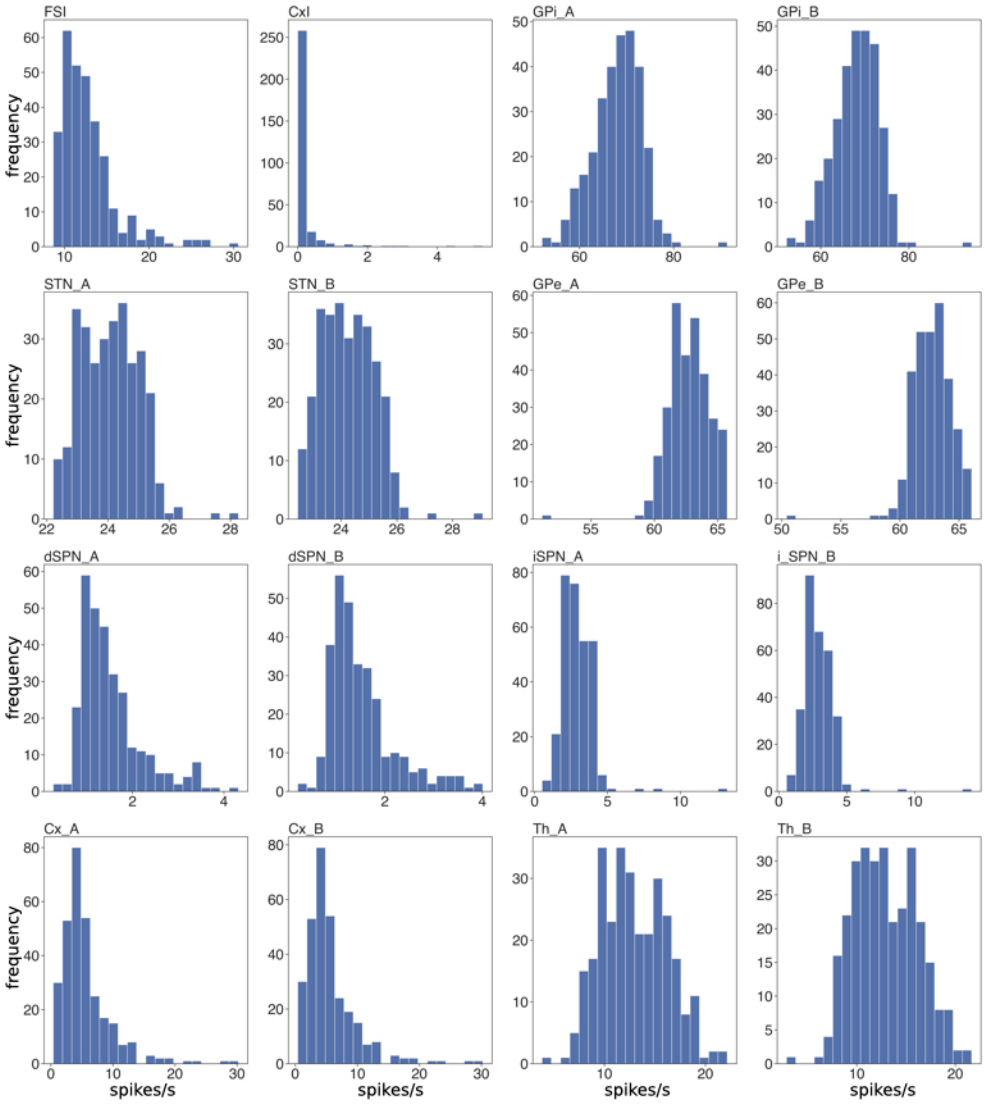

Fig S1. Firing rate histograms across tuned networks. Each panel depicts, for each neuronal population, the histogram of the mean firing rates from stimulus to decision across the 300 different tuned networks.

A.2 Supplementary tables

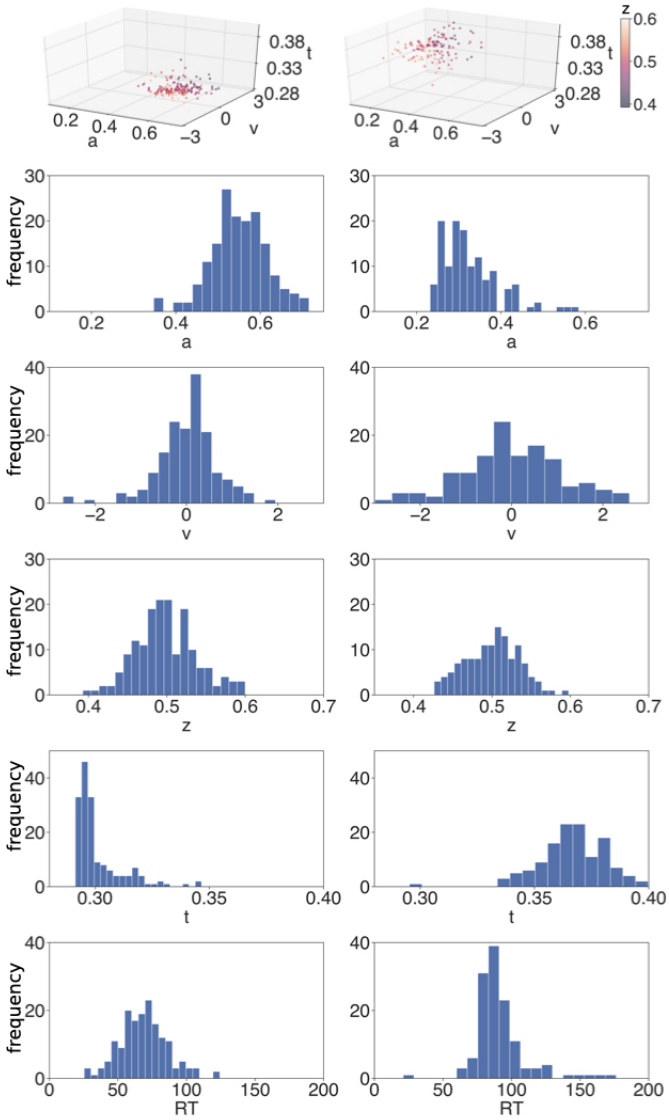

Fig S2. 3-means clustering results. The first row shows the two main clusters of DDM parameter sets obtained after applying K-means clustering with $K = 3$. The third cluster consisting of just a few points with high a values is not shown here. The next four rows present the histograms of the DDM parameters a, ν, z and t, respectively, corresponding to the cluster at the top of each column. The bottom row represents the reaction time histograms for each cluster.

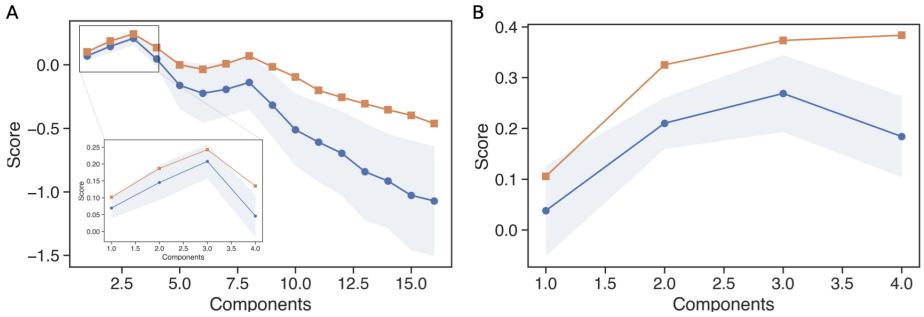

Fig S3. Scores obtained when different numbers of components are considered in the CCA. Panel A represents the scores for the CCA considering the data sets W and R while panel B shows the scores when using R and P. Blue traces depict the mean scores over the 4-fold CVA while the orange traces show the scores for the full data sets (no CVA performed). The shadowed blue zone corresponds to the standard deviation of the scores in the CVA.

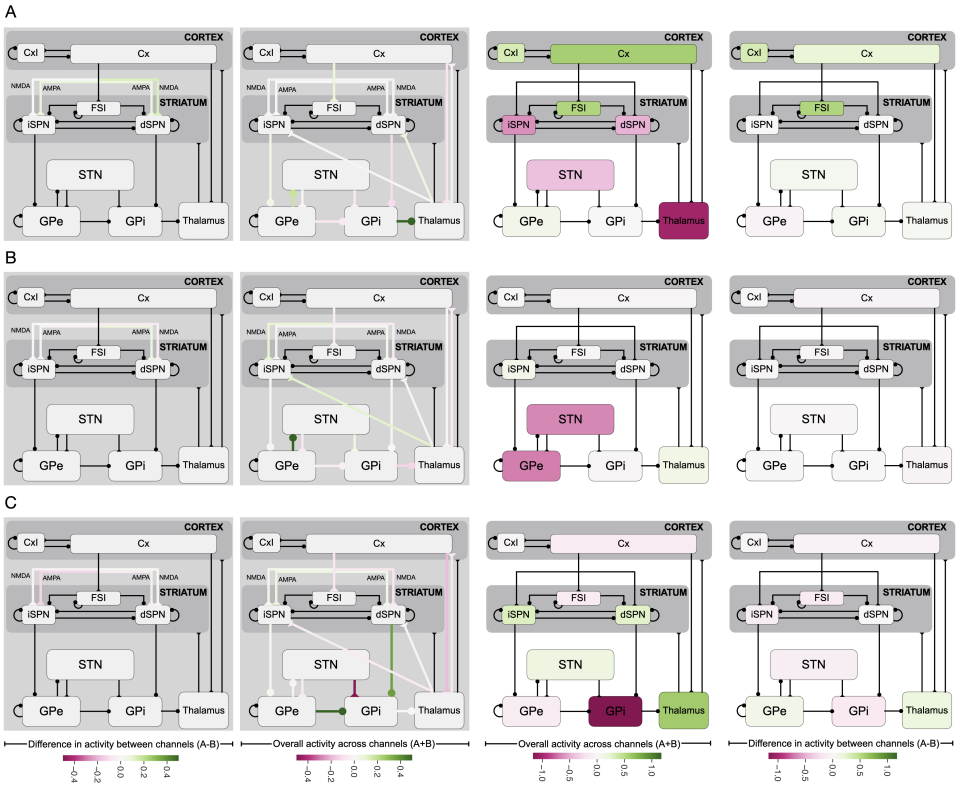

Fig S4. Canonical loadings obtained after applying the CCA to the pair W, R. Relation between the different synaptic weights and the firing rate of each population in the CBGT network. Panels A, B and C stand for the first, second and third components of the CCA. 1st column subpanels correspond to the weight difference between channels. 2nd column subpanels correspond to the overall weights across the two channels. The color-coded loadings of the weights are quite weak overall. The 3rd column subpanels show the channels' overall activity while those in the 4th column depict the difference in activity between the A and B channels. The color-coded loadings of the firing rates are localized in in each component.

Population	Receptor	External frequency	External efficacy	Num. of ext. connections
CxI	AMPA	1.05	1.2	640
Cx	AMPA	2.2	2.0	800
dSPN	AMPA	1.3	4.0	800
iSPN	AMPA	1.3	4.0	800
FSI	AMPA	3.6	1.55	800
GPi	AMPA	0.8	5.9	800
GPe	GABA	2.0	2.0	2000
	AMPA	4.0	2.0	800
STN	AMPA	4.45	1.65	800
Th	AMPA	2.2	2.5	800

Table 3. External current parameters used to simulate the external input arriving at the various CBGT populations. From the first column to the last, we specify the receiving population, the receptor type of the external current, the frequency of the external input, the efficacy of the specific external connection, and, finally, the number of external connections projecting to the population. The time decay constant is $\tau = 2\,ms$ for the AMPA receptor and $\tau = 5\,ms$ for the GABA receptor.