Biased competition in the absence of input bias: predictions from corticostriatal computation

Salva Ardid[1*], Jason S. Sherfey[1,2], Michelle M. McCarthy[1], Joachim Hass[1,3], Benjamin R. Pittman-Polletta[1], Nancy Kopell[1]

[1]Department of Mathematics and Statistics, Boston University, Boston, MA 02215, USA.
[2]Department of Psychological and Brain Sciences, Center for Systems Neuroscience, Center for Memory and Brain, Boston University, Boston, MA 02215, USA.
[3]Department of Theoretical Neuroscience, Bernstein Center for Computational Neuroscience, Central Institute of Mental Health, Heidelberg University, J 5 68159 Mannheim, Germany

*Corresponding author Email: sardid@bu.edu

Classical accounts of biased competition (BC) require an input bias to resolve the competition between neuronal ensembles driving downstream processing. However, flexible and reliable selection of behaviorally-relevant ensembles can occur with unbiased stimulation: striatal D1 and D2 medium spiny neurons (MSNs) receive balanced cortical input, yet their activity determines the choice between GO and NO-GO pathways in the basal ganglia. We present a corticostriatal model identifying three candidate mechanisms that rely on physiological asymmetries to effect rate- and time-coded BC in the presence of balanced inputs. First, tonic input strength determines which MSN phenotype exhibit higher mean firing rate (FR). Second, low strength oscillatory inputs induce higher FR in D2 MSNs but higher coherence between D1 MSNs. Third, high strength inputs oscillating at distinct frequencies preferentially activate D1 or D2 MSN populations. Of these candidate mechanisms, only the latter accommodates observed rhythmic activity supporting rule-based decision making in prefrontal cortex.

Biasing the competition between neuronal ensembles is essential for preferential processing of relevant visual information (*1*). Two computational principles underlie biased competition, as currently understood. First, stimulus-driven neuronal ensembles having distinct stimulus selectivity suppress each other's activity via mutual inhibition. Second, an external input preferentially targets one of the competing ensembles, breaking the symmetry of the system. Computational models of biased competition implementing these principles can differ in considering either an asynchronous (*2, 3, 4, 5*) or a rhythmic (*6*) input bias, as well as in the impact of the bias on neural circuit dynamics, which may increase firing rate (FR) (*2, 3*), coherence (*4, 5*), or both (*6*).

In this work, we introduce an entirely different set of computational principles for biased competition. In the absence of externally imposed (i.e. input) biases, we will show that biased competition is possible between neuronal ensembles endowed with distinct physiological properties. We use corticostriatal processing as a model system for biased competition in the absence of an input bias, because striatal input-output processing is mediated by competition between two distinct GABAergic populations of medium spiny neurons (MSNs), preferentially expressing either D1 or D2 dopamine receptors (*7, 8, 9*), that receive balanced cortical stimulation (*10*).

The manifold differences between D1 and D2 MSNs span anatomical (*11*), network (*12*) and intrinsic properties (*13*), and the two inhibitory populations interact in complex and asymmetrical ways. Thus, it is difficult to predict which physiological asymmetries enable biased competition, and under which conditions. Consequently, we addressed this question in a neural circuit model of corticostriatal processing (Fig. 1A). In this model, D1 and D2 MSNs (which we think of and refer to as ensembles, respectively representing to execute or hold an action) receive balanced cortical stimulation (*10*), and exhibit the three main experimentally observed physiological differences: (i) an asymmetric connectivity profile (Fig. 1A), in which there are

about five times more connections from D2 to D1 MSNs, than vice-versa (*11*); (ii) distinct GABAergic dynamics (Fig. 1B), with efferent synapses from D1 MSNs having higher maximum conductance but more rapid depression than those from D2 MSNs (*12*); and (iii) intrinsic properties (Fig. 1C), such that outward calcium-dependent potassium currents are activated earlier and more strongly in D2 MSNs (*13*).

Functionally, D1 and D2 MSNs represent the first relay of the direct (GO) and indirect (NO-GO) pathways of the basal ganglia (Fig. 1A). GO and NO-GO pathways compete with each other to either trigger or hold an action (*14*). Coactivation of D1 and D2 MSNs during action initiation (*15*) imposes a limitation on winner-take-all competition in the striatum (*16*). Recent modeling work proves, however, that even weak activity biases strongly influence downstream attractor dynamics subserving routing and decision making (*17*). Our corticostriatal model is consistent with this view. While the time course of a selected action may depend on complex interactions between basal ganglia nuclei (*16, 18*), a bias in the activity of D1 and D2 MSNs may be sufficient to determine that selection.

But, how can balanced input enable a flexible biasing of neuronal ensembles—i.e., one that allows the reliable selection of either ensemble through variation in the properties of their common input—, in the first place? We anticipated two potential candidate mechanisms each exploiting one of two specific neural codes, based on either mean firing rate or precise spike timing (coherence). First, due to a trade-off between network and intrinsic dynamics, tonic input strength is able to induce a flexible bias, in which each neuronal ensemble is preferentially activated by inputs within a characteristic range of intensities (Fig. 2C). This mechanism applies even though the two ensembles have the same baseline activity (Fig. 2A and B). The fact that the two ensembles are differentially activated by high and low intensity inputs results from a trade-off between inhibition and activity-dependent hyperpolarization: higher GABAergic inhibition targeting D1 MSNs predominates at low input strengths, leading to higher FR in D2 MSNS,

whereas higher outward calcium-dependent potassium currents in D2 MSNs reverse this bias at high input strengths (Fig. 1C). This turning point in relative excitability is not apparent for single cells under single-cell stimulation because the network contribution mediated by inhibition is not affected (Fig. 2C, inset).

Second, an oscillatory input can induce a coherence bias by preferentially activating the resonant properties of a specific neuronal ensemble. By varying the frequency of a rhythmic cortical input (Fig. 2D) to the striatal circuit in each of the two excitability regimes (i.e., oscillatory inputs of high and low input strengths), our model revealed two ways in which the resonances of the two MSN ensembles diverge. At low input strength, D1 and D2 MSN populations both resonate at the same (low beta) frequency, but D1 MSNs are much more strongly synchronized by rhythmic input (Fig. 2E), despite their lower FR. This divergence between rate and coherence relies on synaptic inhibition. Higher inhibition decreases the overall FR, but enhances spiking coherence, since cells are pushed closer to baseline by inhibition and thus exhibit a more uniform state when inhibition wears off (*19*). At high input strengths, the resonant frequencies of D1 and D2 MSN populations both increase, and diverge from each other (Fig. 2F). The increases in resonant frequency occur because the external input drives MSNs faster than their network frequency in the low beta band (Fig. 2E). As a result, the resonant frequencies of D1 and D2 MSNs shift beyond low beta, respectively to high and middle beta frequencies, following their mean FR (Fig. 2C).

Thus, Figure 2 predicts three types of inputs supporting flexible biased competition under balanced stimulation, confirmed in Figure 3 (top and middle panels): (i) a rate bias regulated by high vs. low input strength (Fig. 3A and B); (ii) a coherence bias regulated by high strength inputs oscillating at distinct beta bands (Fig. 3D and E); and (iii) coexisting rate and coherence biases in the activity of D2 and D1 MSNs, respectively, resulting from low strength oscillatory inputs at low beta frequency (Fig. 3C).

How reliable is each bias at driving downstream action selection? To address this question we ran the model output through two read-out decoders of striatal activity. The first decoder was a spiking activity accumulator, with a slow integration timescale (τ = 100 ms). The second decoder was a coincidence detector, with a fast integration timescale (τ = 5 ms). Our results show that the nature of the striatal bias must fit the timescale of the decoder to guarantee reliable downstream selection (Fig. 3). Thus, only the activity accumulator reliably selects either the GO or the NO-GO pathway from the FR bias between D1 and D2 MSNs (Fig. 3A and B, bottom panels), while only the coincidence detector flexibly selects either the GO or the NO-GO pathway from the coherence bias between D1 and D2 MSNs (Fig. 3D and E, bottom panels). When rate and coherence biases coexist, the selection between GO and NO-GO pathways depends on the integration timescale of the decoder. Thus, a coincidence detector reads out the coherence bias of D1 MSNs, whereas an activity accumulator reads out the rate bias of D2 MSNs. Flexible action selection in this case requires adjusting the decoder integration timescale, so it behaves as a coincidence detector or as an activity accumulator. One way to accomplish this may be adjusting the amount of balanced inhibition targeting the decoder, which has been shown to regulate temporal precision (*20*).

These mechanisms impose predictions that can be tested experimentally. According to the striatal rate bias, action release must functionally correlate with higher FR of D1 over D2 MSNs (Fig. 3B), whereas according to the striatal coherence bias, action release must functionally correlate with a spike-field coherence peak at either high (Fig. 3D) or low (Fig. 3C) beta frequencies. Note, however, that both rate and coherence biases may correlate with action release, even when only one bias is functionally relevant (e.g., Fig. 3D). Contrasting experimentally short vs. long response time trials (or correct vs. error trials) may help in discriminating the ultimate mechanism supporting BC in the striatum. Thus, our model predicts that such a contrast should be highest for the relevant bias, e.g., a larger rate difference between D1 and D2 MSNs for the

rate bias to be relevant.

We have focused so far on the case in which D1 and D2 MSN ensembles compete for the power to trigger or hold isolated actions, but most frequently goal-directed behaviors require selecting the proper action from multiple available sensory-motor associations, such as in rule-based decision tasks. Rule, category and stimulus selective neural activities have been reported in prefrontal cortex (PFC) (*21, 22, 23, 24*) and striatum (*25, 26*), with coactivation of competing ensembles in PFC, coactivating, in turn, competing pathways in the basal ganglia (*27*). Modeling studies have proposed connectionist and rate coding mechanisms to describe routing of sensory-motor responses according to rule biases (*28, 29, 17*); however, recent experimental evidence highlights the central role of temporal dynamics. In particular, rhythmic activity at high beta frequencies is observed in the interaction between PFC and striatum during category learning (*30*), as well as within PFC while performing a rule-based decision task (*31*), where beta phase-locking was higher for the neuronal ensemble encoding relevant information than for its irrelevant competitors. In the same rule-based decision task, alpha-band prefrontal activity was suggested to mediate suppression of ensembles processing the dominant sensory-motor responses during non-dominant trials, i.e., when these representations were irrelevant (*31*).

Our model sheds light on how high beta and alpha rhythms might affect downstream processing in the basal ganglia during this task, suggesting a coherence bias as a mechanistic explanation for rule-based action selection based on stronger high beta synchronization of relevant ensembles in PFC. We hypothesize that while D1 and D2 MSN ensembles representing the same categorical action receive balanced input, MSNs representing relevant categorical actions receive more synchronized input at high beta frequency than MSNs representing irrelevant categorical actions (Fig. 4A top panel). Higher input synchrony produces more coherent striatal firing (Fig. 4A middle panel), a bias that can be reliably read-out by a coincidence detector, but not by an activity accumulator because the mean FR is the same for relevant and irrelevant

MSN ensembles (Fig. 4A middle and bottom panel). Thus, higher beta coherence in PFC is able to bias relevant over irrelevant GO pathways of the basal ganglia. Importantly, neither of the other two biased competition mechanisms present in our model favored the relevant GO pathway (fig. S1 and supplementary text).

In the basal ganglia, inhibitory control is mediated by the indirect (NO-GO) pathway. For an alpha rhythm in PFC to play a role in downstream inhibitory control, it would have to bias the activity of D2 over D1 MSNs. This is the case for the coherence bias mechanism: a balanced cortical input oscillating at alpha frequencies (Fig. 4B top panel) leads to more coherent firing in D2 MSNs (Fig. 2F and Fig. 4B middle panel), which can be reliably read-out by a coincidence detector (Fig. 4B bottom panel). An activity accumulator, on the contrary, does not support an alpha oscillatory input as an inhibitory control mechanism, since it reads out the higher mean FR of D1 MSNs (Fig. 4B bottom panel). Thus, our model suggests a manner by which cortical inputs oscillating at alpha frequencies synchronize the activity of D2 MSNs more strongly than that of D1 MSNs, hence favoring the selection of the NO-GO pathway. Neither of the other two biased competition mechanisms favored the NO-GO pathway (fig. S2 and supplementary text).

The results reported in this work reveal novel computational principles underlying biased competition in support of goal-directed behaviors, such as action selection in the striatum. These mechanisms extend previous approaches that considered unbalanced inputs as the source of the bias between competing neuronal ensembles. In the context of corticostriatal processing, such an approach (*32*) is challenged by the evidence of balanced cortical input to MSNs (*10*). In contrast, our model predicts, to our knowledge for the first time, that flexibly biasing basal ganglia dynamics toward activation of either the direct or the indirect pathway can be accomplished by tuning either the strength or the spectral properties of a balanced cortical input. Of the candidate mechanisms in our model, only the coherence bias mechanism is consistent with observed rhythmic activity in PFC in the context of rule-based decisions (*31*). In fact, our

model of corticostriatal processing suggests a mechanistic explanation for how alpha and high beta rhythms in PFC support, respectively, inhibitory control and rule-based action selection in the basal ganglia.

The validity of these computational principles may extend beyond corticostriatal processing. Thus, a rate bias may arise wherever a diference in relative excitability exists between competing neuronal ensembles (*33*), and a coherence bias may be induced whenever competing neuronal ensembles resonate at distinct frequencies (*34*). For both biases to exist simultaneously, there must be a trade-off between FR and coherence. In our model, this trade-off relies on competing neuronal ensembles receiving different amounts of inhibition, internally generated within the striatal microcircuit, despite balanced cortical input. We suspect that the FR-coherence trade-off may also be present when competing ensembles have different AMPA to NMDA conductance ratios, resulting in different synaptic decay timescales: while more AMPA excitation may enhance coherent dynamics (*6*), less NMDA excitation reduces the overall excitability and, hence, decreases FR.

We analyzed biased competition between distinct neuronal ensembles receiving the same inputs, the inverse condition of "classical" biased competition, which occurs between identical ensembles receiving unbalanced input. In general, however, biased competition may occur between competing neuronal ensembles that differ both physiologically and in their input. While this situation is more complex, it may also be more ubiquitous in the brain and, hence, important to consider systematically. Our work provides a foundation from which to start addressing this challenge.

References

1. R. Desimone, *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **353**, 1245 (1998).

2. G. Deco, E. T. Rolls, *J. Neurophysiol.* **94**, 295 (2005).

3. S. Ardid, X. J. Wang, A. Compte, *J. Neurosci.* **27**, 8486 (2007).

4. C. Borgers, S. Epstein, N. J. Kopell, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18023 (2008).

5. C. I. Buia, P. H. Tiesinga, *J. Neurophysiol.* **99**, 2158 (2008).

6. S. Ardid, X. J. Wang, D. Gomez-Cabrero, A. Compte, *J. Neurosci.* **30**, 2856 (2010).

7. R. L. Albin, A. B. Young, J. B. Penney, *Trends Neurosci.* **12**, 366 (1989).

8. G. E. Alexander, M. D. Crutcher, *Trends Neurosci* **13**, 266 (1990).

9. C. R. Gerfen, *et al.*, *Science* **250**, 1429 (1990).

10. B. Ballion, N. Mallet, E. Bezard, J. L. Lanciego, F. Gonon, *Eur J Neurosci* **27**, 2313 (2008).

11. S. Taverna, E. Ilijic, D. J. Surmeier, *J Neurosci* **28**, 5504 (2008).

12. F. Tecuapetla, L. Carrillo-Reid, J. Bargas, E. Galarraga, *Proc Natl Acad Sci USA* **104**, 10258 (2007).

13. M. A. Arias-Garcia, *et al.*, *Front Syst Neurosci* **7**, 63 (2013).

14. C. R. Gerfen, D. J. Surmeier, *Annu Rev Neurosci* **34**, 441 (2011).

15. G. Cui, *et al.*, *Nature* **494**, 238 (2013).

16. I. A. Oldenburg, B. L. Sabatini, *Neuron* **86**, 1174 (2015).

17. S. Ardid, X. J. Wang, *J Neurosci* **33**, 19504 (2013).

18. R. Bogacz, E. Martin Moraud, A. Abdi, P. J. Magill, J. Baufreton, *PLoS Comput. Biol.* **12**, e1005004 (2016).

19. C. Borgers, N. Kopell, *Neural Comput* **15**, 509 (2003).

20. M. Wehr, A. M. Zador, *Nature* **426**, 442 (2003).

21. W. F. Asaad, G. Rainer, E. K. Miller, *Neuron* **21**, 1399 (1998).

22. I. M. White, S. P. Wise, *Exp Brain Res* **126**, 315 (1999).

23. J. D. Wallis, K. C. Anderson, E. K. Miller, *Nature* **411**, 953 (2001).

24. D. J. Freedman, M. Riesenhuber, T. Poggio, E. K. Miller, *Science* **291**, 312 (2001).

25. S. P. Wise, E. A. Murray, C. R. Gerfen, *Crit Rev Neurobiol* **10**, 317 (1996).

26. E. G. Antzoulatos, E. K. Miller, *Neuron* **71**, 243 (2011).

27. A. F. Marquand, K. V. Haak, C. F. Beckmann, *Nat Hum Behav* **1**, 0146 (2017).

28. J. D. Cohen, K. Dunbar, J. L. McClelland, *Psychol Rev* **97**, 332 (1990).

29. N. P. Rougier, D. C. Noelle, T. S. Braver, J. D. Cohen, R. C. O'Reilly, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7338 (2005).

30. E. G. Antzoulatos, E. K. Miller, *Neuron* **83**, 216 (2014).

31. T. J. Buschman, E. L. Denovellis, C. Diogo, D. Bullock, E. K. Miller, *Neuron* **76**, 838 (2012).

32. J. Bahuguna, A. Aertsen, A. Kumar, *PLoS Comput. Biol.* **11**, e1004233 (2015).

33. T. P. Vogels, L. F. Abbott, *Nat. Neurosci.* **12**, 483 (2009).

34. T. Akam, D. M. Kullmann, *Neuron* **67**, 308 (2010).

Acknowledgments

This research was supported by ARO Grant W911NF-12-R-0012-02; N.K. and S.A. were also supported by NSF Grant DMS-1042134. M.M.M. was supported by CRCNS NIH Grant CRCNS 1R01NS081716. S.A. designed research; N.K. supervised research; All authors contributed to guide research in regular discussions; S.A. implemented the model, with contributions from M.M.M. and J.S.S.; S.A. ran the simulations, analyzed data, prepared the figures and wrote the manuscript; All authors contributed to edit and revise the text. We thank T. Womelsdorf for helpful suggestions on the manuscript.

Supplementary materials

Materials and Methods

Supplementary Text

Figs. S1 to S4

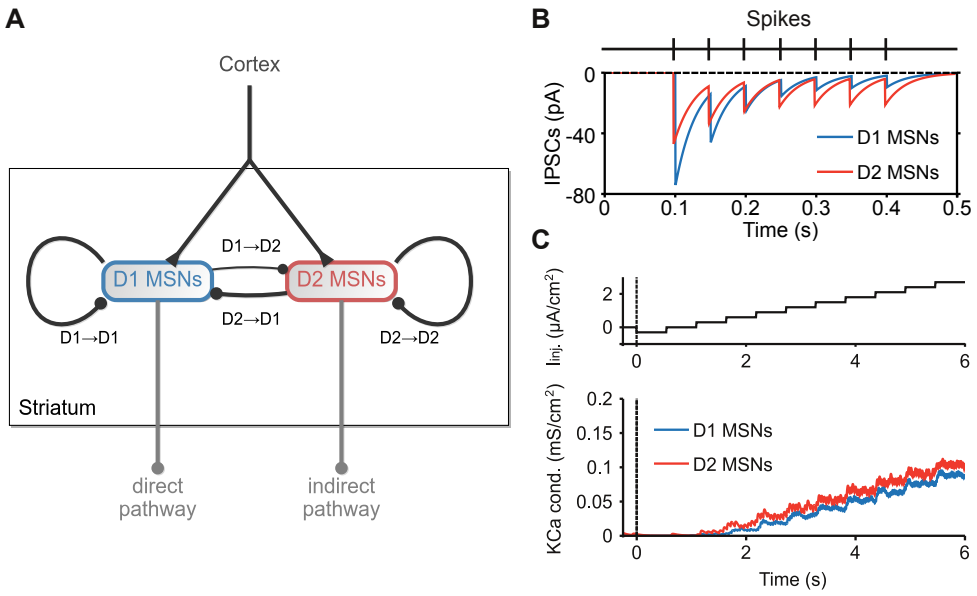

Fig. 1. Corticostriatal circuit model. (**A**) The model of the striatum is composed of D1 and D2 medium spiny neurons (MSNs) according to expressed dopamine receptor. Both phenotypes receive external, balanced input from cortex. D1 and D2 MSNs respectively represent the first stage of the direct (GO) and the indirect (NO-GO) pathways of the basal ganglia. There is an asymmetric connectivity between D1 and D2 MSNs: about 5% of D1 MSNs target D2 MSNs, whereas other synaptic connections in the microcircuit vary within the 20-35% range. (**B**) Distinct GABAergic dynamics between D1 and D2 MSNs. Synapses emerging from D1 MSNs have higher GABAergic conductance (see the difference in amplitude of the first IPSC), but they get depressed more rapidly (see the evolution of IPSC amplitudes). (**C**) Higher activation of outward calcium-dependent potassium currents in D2 MSNs. Top panel shows the protocol of injected current that is applied to D1 and D2 MSNs in the model. Bottom panel shows earlier and stronger activation of the channel for D2 MSNs.

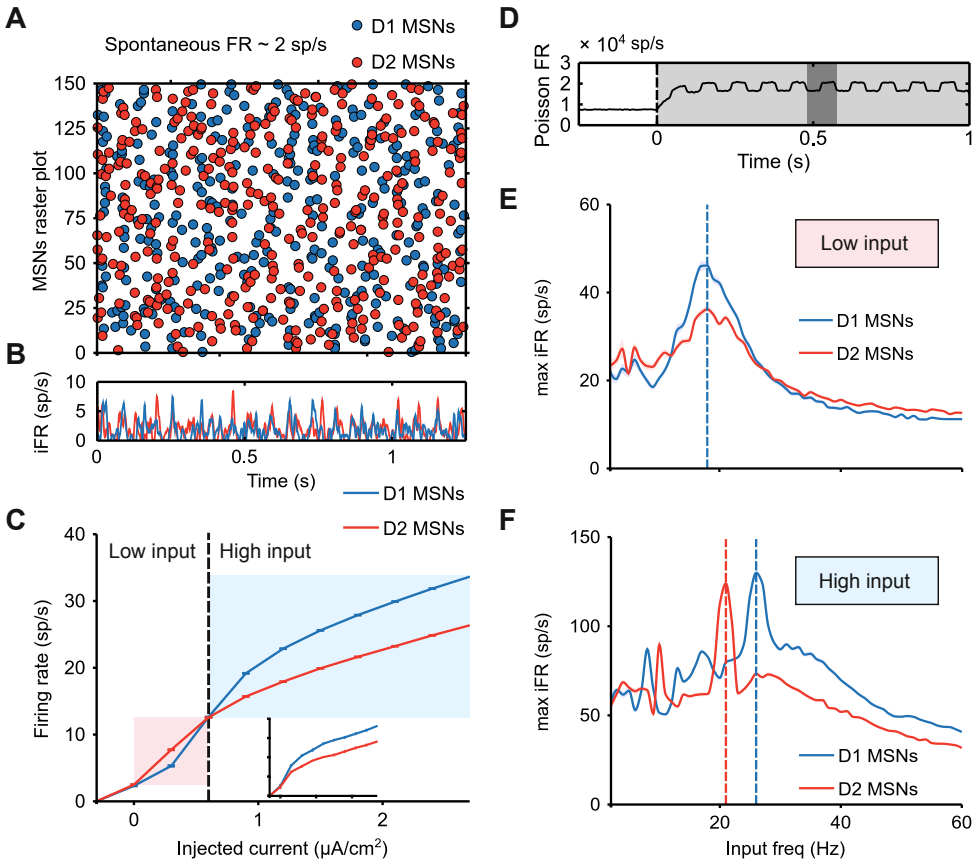

Fig. 2. D1 and D2 MSNs respond differently to balanced input. (**A**) Raster plot of the spontaneous activity of MSNs. (**B**) Instantaneous firing rate (iFR, average firing of the population varying in time) of MSNs. (**C**) Averaged f-I curve of each neuronal ensemble (input-output transfer function between injected current, as in Fig. 1C top panel, and time-averaged population firing rate). The higher excitability of D2 MSNs is shaded in red. The higher excitability of D1 MSNs is shaded in blue. Inset plot: f-I curve when the injected current is applied only to a single cell of each population. (**D**) Poisson rate of the oscillatory input to the striatal circuit: when the stimulus is on (shaded in light gray), the Poisson rate increases and oscillates at a given frequency (inverse of the period shaded in dark gray). (**E**) Resonance of MSNs for low strength input. The resonance is quantified in terms of maximum iFR, a measure of local population synchronization: across input frequencies (x-axis), the average of the peak iFR through all cycles is computed (y-axis). (**F**) Resonance of MSNs for high strength input. Computed as in (E).

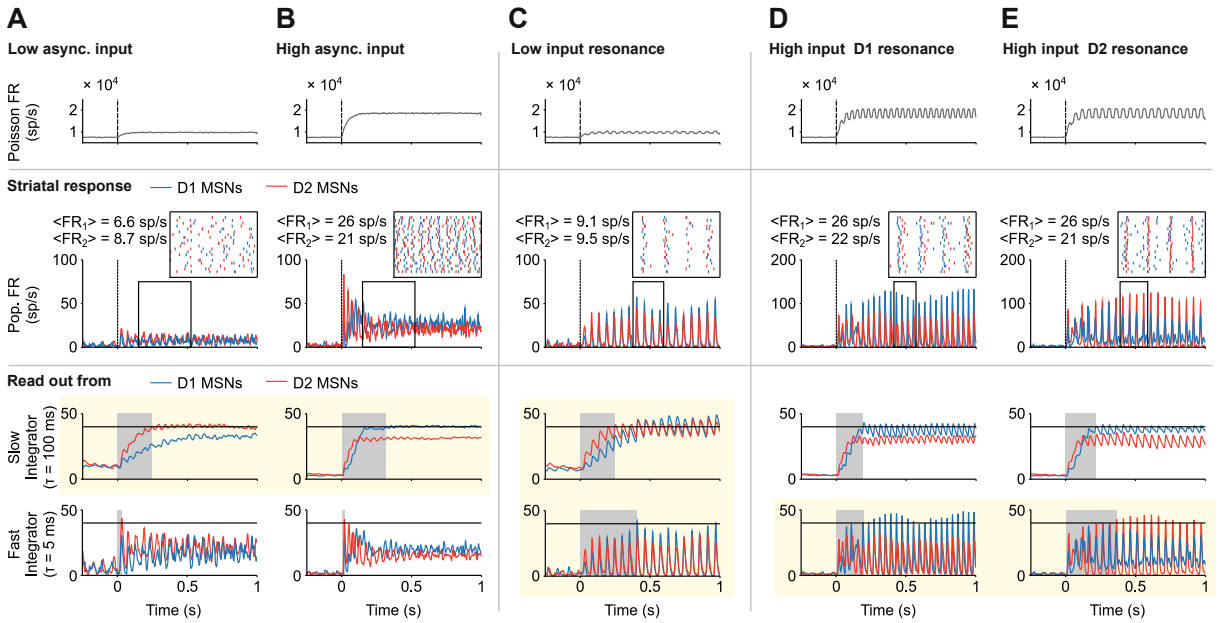

Fig. 3. Three successful mechanisms that flexibly bias the striatal circuit under balanced input. (**A** vs. **B**) Biasing between GO vs NO-GO pathways may depend on the overall strength of the balanced cortical input. Top panels: Poisson rate of the balanced, asynchronous input to the striatal circuit. Middle panels: Population FR varying in time (iFR) Inset plot: raster activity of (n = 20) D1 and D2 MSNs (time window indicated by the black frame in main plot). Bottom panels: Downstream readout of the activity of MSNs using distinct integration timescale. Only slow timescale integration shows flexibility in biasing between D1 and D2 MSNs (highlighted in yellow background). Response time is shaded in light gray. Response threshold at 40 sp/s (solid horizontal line). (**C**) Biasing between GO vs NO-GO pathways under balanced inputs of low strength may depend on resonant properties of MSNs and a dynamic tuning of readout timescale. Top panel: Poisson rate of the balanced, oscillatory input to the striatal circuit that matches the resonant frequency of MSNs (low beta). Middle panel: Population FR varying in time (iFR, which amplitude is a measure of local population synchronization). Inset plot: raster activity of (n = 20) D1 and D2 MSNs (time window indicated by the black frame in main plot). Bottom panel: Downstream readout of the activity of MSNs using distinct integration timescale. On-the-fly tuning of the readout timescale allows flexibility in biasing between D1 and D2 MSNs (highlighted in yellow background). Response time is shaded in light gray. Response threshold at 40 sp/s (solid horizontal line). (**D** vs. **E**) Biasing between GO vs NO-GO pathways under balanced inputs of high strength may depend on resonant properties of MSNs and the spectral content of the balanced cortical input. Top panels: Poisson rate of the balanced, oscillatory input to the striatal circuit that matches the resonant frequency of either MSN type (high and middle beta, respectively). Middle panels: Population FR varying in time (iFR, which amplitude is a measure of local population synchronization; note the higher scale of (D) and (E) compared to (A)-(C)). Inset plot: raster activity of (n = 20) D1 and D2 MSNs (time window indicated by the black frame in main plot). Bottom panels: Downstream readout of the activity of MSNs using distinct integration timescale. Only fast timescale integration shows flexibility in biasing between D1 and D2 MSNs (highlighted in yellow background). Response time is shaded in light gray. Response threshold at 40 sp/s (solid horizontal line).

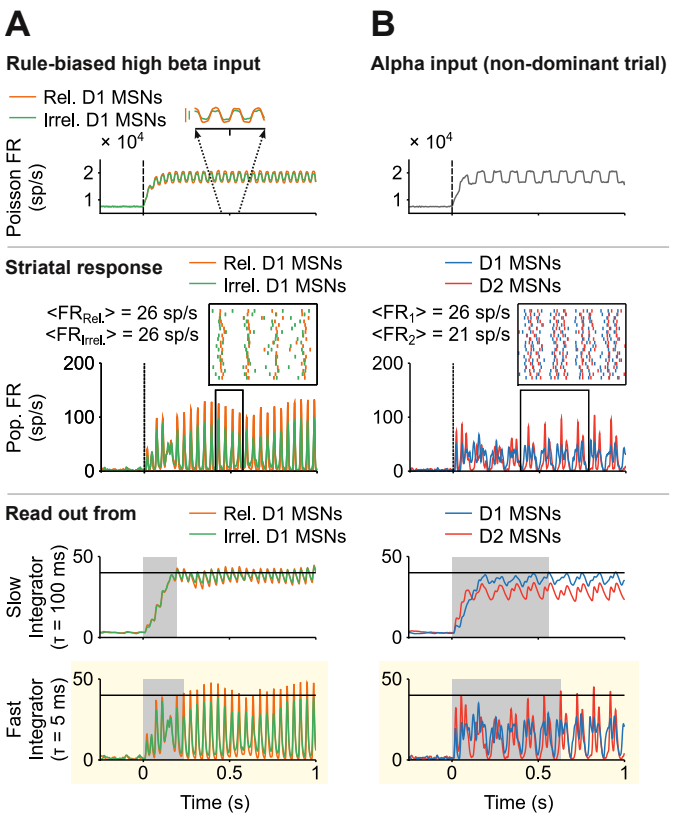

Fig. 4. Striatal processing of rhythmic cortical inputs involved in rule-based decisions. (**A**) Rule-based biased competition between GO pathways: stimulus-driven high beta rhythmic input from PFC biases action selection in the basal ganglia. Top panel: Poisson rate of the high beta oscillatory input to the striatal circuit. The frequency (high beta) and distinct amplitude of the oscillatory input to MSNs (higher for relevant MSNs; see vertical lines of inset plot) are constrained by reported synchronous activity in PFC (see text for details). The input was balanced with respect to D1 and D2 MSNs. Middle panels: Population FR varying in time (iFR, which amplitude is a measure of local population synchronization). Inset plot: raster activity of (n = 20) relevant and irrelevant D1 MSNs (time window indicated by the black frame in main plot). Bottom panels: Downstream readout of the activity of D1 MSNs using distinct integration timescale. Only fast timescale integration shows a reliable bias in favor of the relevant D1 MSN ensemble (highlighted in yellow background). Response time is shaded in light gray. Response threshold at 40 sp/s (solid horizontal line). (**B**) Striatal dynamics toward the NO-GO pathway can be favored by an alpha oscillatory input, present in PFC ensembles in non-dominant trials (see text for details). Top panels: Poisson rate of the balanced, alpha oscillatory input to the striatal circuit. Middle panels: Population FR varying in time (iFR). Inset plot: raster activity of (n = 20) D1 and D2 MSNs (time window indicated by the black frame in main plot). Bottom panels: Downstream readout of the activity of MSNs using distinct integration timescale. For the alpha rhythm to be associated with inhibitory control, D2 MSN bias should prevail over D1 MSN bias. This is only the case through fast integration timescale (highlighted in yellow background). Response time is shaded in light gray. Response threshold at 40 sp/s (solid horizontal line).