A spiking model of basal ganglia dynamics in stopping behavior supported by arkypallidal neurons

Lorenz Goenner | Oliver Maith | Iliana Koulouri | Javier Baladron | Fred H. Hamker

Department of Computer Science, Chemnitz University of Technology, Chemnitz, Germany

Correspondence
Fred H. Hamker, Department of Computer Science, Chemnitz University of Technology, Chemnitz, Germany.

Funding information
Federal Ministry of Education and Research, Germany, Grant/Award Number: BMBF 01GQ1707; Volkswagen Foundation

Abstract
The common view that stopping action plans by the basal ganglia is achieved mainly by the subthalamic nucleus alone due to its direct excitatory projection onto the output nuclei of the basal ganglia has been challenged by recent findings. The proposed “pause-then-cancel” model suggests that the subthalamic nucleus provides a rapid stimulus-unspecific “pause” signal, followed by a stop-cue-specific “cancel” signal from striatum-projecting arkypallidal neurons. To determine more precisely the relative contribution of the different basal ganglia nuclei in stopping, we simulated a stop-signal task with a spiking neuron model of the basal ganglia, considering recently discovered connections from the arkypallidal neurons, and cortex-projecting GPe neurons. For the arkypallidal and prototypical GPe neurons, we obtained neuron model parameters by fitting their neuronal responses to published experimental data. Our model replicates findings of stop-signal tasks at neuronal and behavioral levels. We provide evidence for the existence of a stop-related cortical input to the arkypallidal and cortex-projecting GPe neurons such that the stop responses of the subthalamic nucleus, the arkypallidal neurons, and the cortex-projecting GPe neurons complement each other to achieve functional stopping behavior. Particularly, the cortex-projecting GPe neurons may complement the stopping within the basal ganglia caused by the arkypallidal and STN neurons by diminishing cortical go-related processes. Furthermore, we predict effects of lesions on stopping performance and propose that arkypallidal neurons mainly participate in stopping by inhibiting striatal neurons of the indirect rather than the direct pathway.

KEYWORDS
basal ganglia, computational modeling, neuronal networks, stop-signal task

Abbreviations: BG, basal ganglia; GPe, globus pallidus pars externa; GPe-Arky, arkypallidal neurons; GPe-Cp, cortex-projecting GPe neurons; GPe-Proto, prototypical GPe neurons; GPi, globus pallidus pars interna; SNr, substantia nigra pars reticulata; SSD, stop-signal delay; STN, subthalamic nucleus; Str, striatum; StrD1, striatal projection neurons expressing the D1-type dopamine receptor; StrD2, striatal projection neurons expressing the D2-type dopamine receptor; StrFSI, striatal fast-spiking interneurons.

Lorenz Goenner and Oliver Maith contributed equally to this work.

Edited by: Yoland Smith

© 2020 Federation of European Neuroscience Societies and John Wiley & Sons Ltd
INTRODUCTION

The basal ganglia (BG) are a collection of subcortical nuclei which are associated with action selection (Haber, 2003; Mink, 1996); however, studies have also shown that the BG support the cancellation of actions (Aron & Poldrack, 2006; Eagle & Robbins, 2003).

Theories and models of the BG have focused primarily on three main pathways: the direct, indirect, and hyperdirect pathways (for review, see Schroll & Hamker, 2013). The direct pathway refers to the inhibitory projections of striatal neurons expressing the D1-type dopamine receptor (StrD1) to the output nuclei: the globus pallidus interna and the substantia nigra pars reticula (GPi/SNr; Albin et al., 1989; DeLong, 1990; Mink, 1996). The hyperdirect pathway is characterized by the direct projections from the subthalamic nucleus (STN), receiving cortical input, to the SNr or GPe (Kita et al., 2006; Nambu et al., 2002). The indirect pathway includes the projections from striatal neurons expressing the D2-type dopamine receptor (StrD2) to the globus pallidus externa (GPe), as well as the direct projections from the GPe to the output nuclei and the indirect projections from the GPe over the STN to the output nuclei. All three pathways converge to the GPi/SNr, which, in turn, inhibit the thalamus and therefore can regulate action initiation (DeLong, 1990; Deniau & Chevalier, 1985; Mink, 1996).

An established paradigm to study the role of the BG in response inhibition is the stop-signal task. In this task, a motor response must be initiated after the occurrence of a Go cue. However, sometimes the Go cue is followed by a Stop cue, signaling that the imminent response must be rapidly suppressed (e.g., Logan et al., 1984). A common hypothesis is that STN activity can function as a global Stop signal, as STN activation correlates with stopping premature responses during conflicting decision processes (Aron & Poldrack, 2006; Baunez et al., 2001; Frank, 2006; Schmidt & Berke, 2017; Schmidt et al., 2013). This function is anatomically supported by its direct excitatory synaptic projection to the GPi/SNr, by which STN activity can contribute to inhibiting the thalamus, thus suppressing action release (Nambu et al., 2000; Robledo & Féger, 1990; Wiecki & Frank, 2013).

There is evidence showing that STN lesions led to deficits in response preparation processes, which include the selection and preparation of an action and the timing of performing an action. These include slower reaction times, impaired choice accuracy, and an increased number of incorrect premature responses to a stimulus (Baunez et al., 2001). Other evidence shows that STN lesions do not slow the Stop signal reaction time, but affected the ability to correctly initiate a Stop process (Eagle et al., 2008) so that the role of the STN in the stop-signal reaction time paradigm is not completely understood.

However, this notion of a key role of STN in stopping has recently been challenged based on two findings: First, Schmidt et al. (2013) observed that STN activation remained highly similar across trials, regardless of whether stopping was successful or not. By contrast, SNr neurons responded preferably to correct rather than failed stopping (Schmidt et al., 2013). Second, STN responded to both cues, the presentation of Go and Stop cues, to a similar degree (Mallet et al., 2016). These findings led Mallet et al. (2016) and Schmidt and Berke (2017) to propose a two-stage model of stopping, in which STN activity serves the role of a Pause signal which is followed by a separate “cancellation” signal.

This “pause-then-cancel” model of stopping was further motivated by the recent characterization of the arky pallidal subpopulation of the GPe (GPe-Arky; Mallet et al., 2012). In a stop-signal task, GPe-Arky neurons responded with high temporal specificity to the onset of a Stop cue, with increased activity in successful Stop trials (Mallet et al., 2016). Furthermore, making direct GABAergic projections to the striatum (Glajch et al., 2016; Hernández et al., 2015; Mallet et al., 2012), these neurons may inhibit Go-related striatal activity which triggers the release of actions by removing thalamic inhibition in the GPi/SNr. On this basis, Mallet et al. (2016) concluded that GPe-Arky neurons are well suited to support the cancellation of ongoing actions by inhibiting striatal Go-related activity.

However, several questions remain unanswered. First, are GPe-Arky Stop responses both necessary and sufficient for stopping imminent actions? More generally, as the precise connectivity patterns of GPe-Arky neurons remain to be determined, what computational constraints for stopping may arise from the dynamics of the BG network? Furthermore, more recently, a new subpopulation of GPe cells which form a cortico-pallido-cortical loop has been characterized (Abecassis et al., 2020; Chen et al., 2015; Saunders et al., 2015). The function of this closed-loop system, which acts in parallel to the cortico-BG-thalamo-cortical loop, is not yet clear. We hypothesize that it also has a role in stopping and complements the short-term inhibition of striatal activity produced through the GPe-Arky cells. In detail, we propose that the GPe-Arky activation produces only a short-term stop while cortical inhibition through the direct pallido-cortical projection produces a long-term cancellation of cortical input to the striatum, which is required due to the brief activation of the GPe-Arky cells.

To better understand the interactions between the STN, the GPe-Arky striatal projections and the cortico-pallido-cortical loop during stopping, we extended our previous model of the BG (Baladron et al., 2019) by (a) GPe-Arky neurons, (b) a population of fast-spiking striatal interneurons (StrFSI), and (c) a subpopulation of GPe neurons which are cortex-projecting (GPe-Cp), thus forming a loop with the cortex. Based on this model, we reproduce key behavioral
results as well as neural BG response dynamics reported by Mallet et al. (2016) in the stop-signal task. By identifying potential failure modes, we derive constraints for successful stopping in terms of network interactions. Furthermore, our model allows to predict the effect of experimental manipulation of the stopping-related nuclei on the behavior. The simulations show that the projection between the GPe-Arky cells and the striatum, which causes a short-term stop, is not sufficient to produce a fully functional stopping behavior but that further, the Pause effect of the STN and the long-term cancellation effect of the cortico-pallido-cortical loop are required. We therefore extend the pause-then-cancel theory from Mallet et al. (2016) and Schmidt and Berke (2017) to include a long-term cancellation component caused by the newly observed pallidal connectivity.

2 | MATERIALS AND METHODS

2.1 | Model overview

We established a neuro-computational model of the BG circuitry based on our previous work (Baladron et al., 2019). As its predecessor, the model includes the direct, indirect and hyperdirect pathways. However, we have further included StrFSIs and two new subtypes of the GPe neurons: GPe-Arky and GPe-Cp, and the corresponding connections.

We model two types of striatal spiny projection neurons, one for each of the two main types of dopamine receptor. The direct pathway is initiated in the StrD1 neurons, which project directly to the GPi/SNr, whereas the indirect pathway is initiated in the StrD2 neurons, which reach the GPi/SNr only through the prototypical GPe neurons (GPe-Proto) neurons. We consider collateral connections within and between the StrD1 and StrD2 nuclei (Burke et al., 2017; Taverna et al., 2008). The StrFSIs inhibit both StrD1 and StrD2 cells. All striatal neurons receive excitatory cortical input (see Figure 1). These and the following mentioned cortical inputs in the model are described in more detail in Section 2.3.

The STN initiates the hyperdirect pathway. It receives excitatory cortical input and provides excitatory input to the GPi/SNr through a direct projection. Additionally, based on our previous work (Baladron et al., 2019), bidirectional connections between the GPe-Proto and the STN are considered (Kita, 2007; Nevada-Holgado et al., 2014).

Consistent with experimental findings of uncorrelated spontaneous activity in the GPi/SNr (Bergman et al., 1998; Boraud et al., 1996; Wichmann et al., 1999), we induced baseline activity in our model GPi/SNr nucleus from Poisson inputs. The thalamus in our model is thus continuously inhibited and only generates a few action potentials, unless the GPi/SNr activity is reduced. Only when the thalamus becomes sufficiently active, the model performs an action (more details in the Section Model details).

**FIGURE 1** Model architecture. (a) Model overview with direct, indirect, and hyperdirect pathways, and crosstalk between these pathways, such as bidirectional connections between the STN and GPe and collateral projections in the striatum. Furthermore, three distinct cortical populations serve as input for the striatum, thalamus, STN, and GPe. The GPe is presented here as a single nucleus, but distinct GPe populations have different projections. (b) The whole GPe consists of three populations, the GPe-Proto, GPe-Arky, and GPe-Cp. Their individual glutamatergic and GABAergic inputs and GABAergic outputs are listed at the arrow ends. Each population consists of 100 neurons. The arrow heads identify the projections as GABAergic or glutamatergic. Not shown are the base line input populations of each BG population and the thalamic feedback projections from thalamus to StrD1, StrFSI and StrD2.
Furthermore, the model includes direct cortico-thalamic connections which bypass the BG pathways. This pathway has been shown to be essential for modeling the effect of GPi lesions, such as in Parkinson's Disease patients, which improve performance in well-learned, everyday movements (Baladron & Hamker, 2015; Baladron et al., 2019; Schroll et al., 2014).

Compared to the previous model (Baladron et al., 2019), we have made significant changes to the structure of the GPe. The GPe cells which are part of the classical indirect pathway are the GPe-Proto cells. However, as they do not only project to the GPi but also to StrFSI (Bevan et al., 1998; Corbit et al., 2016; Glajch et al., 2016; Hernández et al., 2015; Mastro et al., 2014; Saunders et al., 2016), we have implemented this projection in addition to the new StrFSI population.

Second, we added a new GPe population, the GPe-Arky. All GPe-Arky neurons target StrD1, StrD2, and StrFSI cells (Glajch et al., 2016; Hernández et al., 2015; Mallet et al., 2012). Based on findings of collateral connections in the GPe (Bevan et al., 1998; Fujiyama et al., 2016; Kita, 2007; Kita & Kitai, 1994; Mallet et al., 2012; Sadek et al., 2007), we have implemented projections within and between all GPe populations. GPe-Arky neurons receive inputs from the striatum (Hernández et al., 2015; Kita, 2007; Kreitzer & Malenka, 2008) and STN (Kita, 2007), like the GPe-Proto neurons. Thus, we considered separate projections from the STN and StrD2 to the GPe-Arky neurons. Mallet et al. (2016) observed that STN neurons respond similarly to both Go and Stop cues. However, the GPe-Arky cells only respond strongly to Stop cues. Thus, their activation does not seem to follow the excitatory input from the STN and therefore, we assume that the response of GPe-Arky cells on Stop cues is caused by an additional excitatory cortical input to the GPe which bypasses the striatum and STN. Various findings show that there is a direct projection from cortex to GPe (Abecassis et al., 2020; Karube et al., 2019; Milardi et al., 2015; Naito & Kita, 1994; Smith & Wichmann, 2015). Therefore, we considered separate cortical inputs to GPe-Arky. These cortical inputs are associated with stopping activity, which is related to movement preparation, as further described in Section 2.3.

### Table 1: Evidence for efferent GPe projections

| Reference          | Key conclusion                                                                 | Supported projection          |
|--------------------|--------------------------------------------------------------------------------|-------------------------------|
| Bevan et al. (1998)| • Main part of the GPe projects “downstream” and partially to striatum (main part and downstream projections are likely GPe-Proto)  |
|                    | • Collaterals within GPe                                                       | GPe-Proto → GPi/SNr           |
|                    | Corbit et al. (2016)                                                           |                               |
|                    | • StrFSIs respond stronger to GPe stimulation than SPNs                        |                               |
|                    | Saunders et al. (2016)                                                         |                               |
|                    | • PV expressing GPe cells project to STN, SNr and StrFSI                       |                               |
|                    | Mastro et al. (2014)                                                          |                               |
|                    | • Lhx6- and PV-expressing GPe cells project to STN and StrFSI                  |                               |
|                    | Dodson et al. (2015)                                                          |                               |
|                    | • GPe-Arky cells express FoxP2                                                 |                               |
|                    | • PV and Lhx6 are widely coexpressed but not with FoxP2                        |                               |
|                    | • → GPe-Proto likely express PV, Lhx6                                          |                               |
|                    | Mallet et al. (2012)                                                          |                               |
|                    | • GPe-Proto project to STN, SNr and partially striatum (likely FSI)            | GPe-Arky → FSI                |
|                    | • GPe-Arky project to SPNs and StrFSIs                                         | GPe-Arky → SPNs               |
|                    | • GPe-Proto cells express PV                                                   | (GPe-Proto → SNr)             |
|                    | • GPe-Arky cells do not express PV                                             | GPe-Proto → STN               |
|                    | Glajch et al. (2016)                                                          |                               |
|                    | • PV expressing cells project to STN and StrFSI, almost not to SPNs            | GPe-Cp → striatum/SPNs        |
|                    | • Npas1-expressing cells project to SPNs, almost not to STN                    | GPe-CP → cortex               |
|                    | Hernández et al. (2015)                                                       |                               |
|                    | • PV and Npas1 are only rarely coexpressed                                     | GPe-Arky → striatum/SPNs      |
|                    | • PV expressing cells project stronger to STN than to striatum                 | GPe-Proto → STN               |
|                    | • Npas1-expressing cells project stronger to striatum than to STN              | GPe-Proto → FSI               |
|                    | Abdi et al. (2015)                                                            |                               |
|                    | • PV and Npas1 are only rarely coexpressed                                     |                               |
|                    | Dodson et al. (2015)                                                          |                               |
|                    | • Npas1-expressing cells can be mainly distinguished into FoxP2- and Nkx2.1-expressing cells |                               |
|                    | Abecassis et al. (2020)                                                       |                               |
|                    | • PV and Npas1 are only rarely coexpressed                                     |                               |
|                    | • Npas1-expressing cells can be mainly distinguished into FoxP2 and Nkx2.1-expressing cells |                               |
|                    | • Npas1-Nkx2.1-expressing cells represent a separate GABAergic cortex-projecting population |                               |
|                    | Saunders et al. (2015)                                                        |                               |
|                    | • There is a direct GABAergic GPe → cortex projection                          |                               |
| Chen et al. (2015) |                                                                                  |                               |
Based on the findings of Mallet et al. (2016), we assume that the GPe-Arky population plays a critical role in stopping go-related striatal activity. Glajch et al. (2016) showed that Npas1-expressing GPe neurons project stronger to StrD2 than to StrD1 cells. Although Npas1 is not only expressed by GPe-Arky neurons (Abdi et al., 2015; Abecassis et al., 2020; Dodson et al., 2015) almost all GPe-Arky neurons express Npas1. Thus, GPe-Arky neurons seem to project stronger to StrD2 than to StrD1. At a first glance, this seems to be a counterintuitive finding, because inhibiting the StrD1 and thus deactivating the direct pathway and disinhibiting the GPi/SNr appears to be the proper way for GPe-Arky to stop movement. We assume that the inhibition of StrD2 is important because the StrD2 population inhibits the whole GPe. Go-related inputs seem to activate both StrD1 and StrD2 (Cui et al., 2013; Tecuapetla et al., 2016). Therefore, GPe-Arky may first inhibit StrD2 to disinhibit the GPe and thus itself, so that it can exert sufficient inhibition on the entire striatum to stop the go-related activity.

Furthermore, we assume that another GPe-population is also associated with stopping. The GPe-Cp population

\[
 f(V, U) = \begin{cases} 
 a(b(V-(-80))-U) & \text{for StrD1 and StrD2} \\
 a(b(V-(-55))^3 - U) & \text{for STN, SNr, GPe-Arky, GPe-Proto, GPe-Cp and thalamus} \\
 -aV & \text{for StrFSI, if } V < -55 \\
 a(b(V-(-55)))^3 - U & \text{for StrFSI, otherwise.} 
\end{cases}
\]

in our model is supported by recent findings of Abecassis et al. (2020), who observed a separate cortex-projecting GPe population besides the GPe-Proto and GPe-Arky. Based on the common gene expression of Npas1 of GPe-Cp and GPe-Arky, we assumed a similar connectivity to that of GPe-Arky; therefore, GPe-Cp gets input from the StrD2, STN, other GPe neurons and further from StrD1 (Saunders et al., 2015). According to our hypothesis, the GPe-Cp complements the GPe-Arky-related short-term inhibition of the striatum during stopping by stopping cortical go-related processes. Therefore, we assume that they also receive the cortical input associated with stopping. Evidence that motivated the implementation of the different GPe populations and the GPe connectivity is summarized in Tables 1 and 2.

### 2.2 Model details

We used a simplified spiking neuron model (Izhikevich, 2004) and implemented our network model in the ANNarchy simulator, version 4.6 (Vitay et al., 2015), using Euler integration with a timestep of 0.1 ms. Simulation code is available at https://github.com/Olima ol/stopsignaltask_BG.

The striatum is composed of three populations (StrD1, StrD2, and StrFSI), so is the GPe (GPe-Proto, GPe-Arky, and GPe-Cp). The GPi/SNr, the thalamus, and the STN are modeled as single populations. All of these populations consist of 100 Izhikevich spiking neurons. The membrane potential of the neurons is given by:

\[
 \frac{dV}{dt} = \begin{cases} 
 n_2 V^2 + n_1 V + n_0 - U C - g_{\text{AMPA}} (V-E_{\text{AMPA}}) - g_{\text{GABA}} (V-E_{\text{GABA}}) & \text{for StrFSI, if } V < -55 \\
 -g_{\text{AMPA}} (V-E_{\text{AMPA}}) & \text{for StrD1 and StrD2} \\
 f(V, U) & \text{for STN, SNr, GPe-Arky, GPe-Proto, GPe-Cp and thalamus} \\
 -aV & \text{for StrFSI, if } V < -55 \\
 a(b(V-(-55)))^3 - U & \text{for StrFSI, otherwise.} 
\end{cases}
\]

Each time the membrane potential reaches a value of \( V_{\text{peak}} \), a spike is considered to be emitted and the value of \( V \) is reset to a value of \( c \), and \( U \) is increased by a fixed amount, \( d \). For the specific values used in our simulations, see Table 3. At all synaptic projections, we used a value of \( \tau_{\text{AMPA}} = 10 \text{ ms}, \tau_{\text{GABA}} = 20 \text{ ms}, E_{\text{AMPA}} = 0 \text{ mV}, \text{and } E_{\text{GABA}} = -90 \text{ mV} \).

The parameters for all striatal neurons were taken from the model of Humphries et al. (2009). For the GPi and STN neurons, the parameters were taken from Thibeault and Srinivas (2013). Both parameter sets were also used in Baladron et al. (2019). For the thalamic neurons, the parameters of the tonic spiking model proposed by Izhikevich (2004) were used. As we are not aware of published parameters approximating the dynamics of GPe-Arky neurons based on a simplified spiking neuron model, we obtained estimates for neuron model parameters for GPe-Arky and GPe-Proto neurons by numerically fitting the responses to published experimental data (Abdi et al., 2015; Bogacz et al., 2016). Specifically, we computed the mean firing rate of our model neurons in response to step current injections, and used the scipy.optimize function to find those model parameters which minimized the root mean square error between the simulated firing rates and published data (Abdi et al., 2015; Bogacz et al., 2016; Figure 2). To obtain a concave shape of the \( f-I \) curve as
observed experimentally, we added an absolute refractory period of $t_{	ext{refr}} = 5$ ms for both model neurons. For all other neuron populations, no absolute refractory period was added. For the GPe-Cp neurons, we used the same parameters as for the GPe-Proto neurons.

Synaptic input to each neuron is modeled by considering the effect of both AMPA and GABA receptors (see Equation 1). The AMPA conductance is increased by a fixed amount (weight) when a presynaptic neuron with an excitatory synapse emits a spike, while the GABA conductance is increased if the synapse is inhibitory. When no presynaptic spike is emitted, both conductances reduce their value to 0 exponentially, with time constants $\tau_{\text{AMPA}}$ and $\tau_{\text{GABA}}$.

The synaptic contacts of a projection between two populations were defined stochastically. For each projection within the BG, each neuron of the postsynaptic population was connected to 10 randomly selected neurons of the presynaptic population. A synaptic contact between two neurons was defined as excitatory if the presynaptic neuron was a part of the STN, and as inhibitory otherwise.

### TABLE 2 Evidence for afferent GPe projections

| Reference | Key conclusion | Supported projection |
|-----------|---------------|----------------------|
| Hernández et al. (2015) | Striatum projects to PV, Npas1- and Lhx6-expressing cells | StrD2 → GPe-Proto |
| Kreitzer and Malenka (2008) | StrD2 is main input of GPe | StrD2 → GPe-Arky |
| Kita (2007) | Striatum and STN project to GPe | STRN → GPe-Proto |
| | STN project to PV$^+$ and PV$^-$ cells | STRN → GPe-Arky |
| | | STRN → GPe-Cp |
| Bevan et al. (1998) | There are collaterals within the GPe | GPe collaterals |
| Kita (2007) | | |
| Mallet et al. (2012) | | |
| Kita and Kitai (1994) | | |
| Sadek et al. (2007) | | |
| Fujiyama et al. (2016) | | |
| Chen et al. (2015) | Striatum projects to GPe-Cp | StrD1 → GPe-Cp |
| Saunders et al. (2015) | StrD1 and StrD2 project to GPe-Cp | (StrD2 → GPeCp) |
| Abecassis et al. (2020) | Various projections from cortex to GPe | Cortex Stop → GPe-Arky |
| Milardi et al. (2015) | | Cortex Stop → GPe-Cp |
| Naito and Kita (1994) | | |
| Smith and Wichmann (2015) | | |
| Karube et al. (2019) | | |
| Dodson et al. (2015) | Movement onset leads to increased GPe-Arky activity and heterogeneous response of GPe-Proto cells | |
| | Our assumption: response after movement onset is caused by excitatory cortical input related to movement | |

### TABLE 3 Neuron model parameters

| Population | $a$ | $b$ | $c$ | $d$ | $n_0$ | $n_1$ | $n_2$ | $C$ | $V_{\text{peak}}$ |
|------------|-----|-----|-----|-----|-------|-------|-------|-----|-----------------|
| StrD1, StrD2 | 0.05 | −20 | −55 | 377 | 61.651 | 2.595 | 0.0228 | 50 | 40 |
| StrFSI | 0.2 | 0.025 | −60 | 0 | 43.75 | 1.5 | 0.0125 | 80 | 25 |
| STN | 0.005 | 0.265 | −65 | 2 | 140 | 5 | 0.04 | 1 | 30 |
| GPi/SNr | 0.005 | 0.585 | −65 | 4 | 140 | 5 | 0.04 | 1 | 30 |
| GPe-Proto, GPe-Cp | 0.0058 | 0.56 | −65 | 3.8 | 117 | 4.86 | 0.043 | 1 | 30 |
| GPe-Arky | 0.0054 | 0.34 | −71 | 9.81 | 113 | 4.47 | 0.04 | 1 | 30 |
| Thalamus | 0.02 | 0.2 | −65 | 6 | 140 | 5 | 0.04 | 1 | 30 |
population (Ten-to-One pattern). All synaptic contacts of a projection were created with the same fixed weight value, which depends on the corresponding projection. All projections within the BG and their weight values are presented in Table 4.

The model cortex has three distinct neuron populations (cortex-Go, cortex-Stop, and cortex-Pause), which consist of 100 neurons each. The cortical neurons randomly generate action potentials according to Poisson processes with a specified firing rate. The three distinct cortex populations are activated dependent on the behavioral task the model performs, which is described in detail in Section 2.3.

The synaptic contacts of the cortical projections were not defined stochastically. For each cortico-BG projection, each neuron of the postsynaptic population was connected to a single neuron of the corresponding presynaptic cortex population (One-to-One pattern). All cortical projections and their weight values are presented in Table 5.

Similar to the cortical inputs (One-to-One pattern), we generated baseline activity in all BG nuclei based on a continuously active Poisson process for each BG population, with a strength depending on the corresponding BG population (Table 6). Within each cortical input population, firing rates per neuron were drawn from a normal distribution with common mean and standard deviation. All projections had synaptic delays drawn from a uniform distribution from 0 to 10 ms.

2.3 Behavioral task

We applied the model on the stop-signal task used by Mallet et al. (2016) to understand the dynamical mechanisms underlying the putative role of GPe neurons in stopping. In this task, rats had to maintain their nose in a central port until a tone (Go signal) instructed them to perform a movement into a neighboring hole. The frequency of the
signal indicated whether the animal should move left or right. On 30% of the trials, the initial tone was followed by a white noise burst (Stop signal), which informed the rat that the action should be canceled and that it should stay in the central port. Specifying the stop-signal delay (SSD) between the onset of the Go and the Stop signal determined whether rats successfully stopped or failed to stop an action. The reaction time was recorded between the onset of the Go signal and the withdrawal of its nose from the port after exposure of the Go signal.

Neural measurements showed that all GPe cells increased their activity following the Stop signal (Mallet et al., 2016). However, the response to the Stop signal was stronger, faster and more consistent in the GPe-Arky than in the GPe-Proto cells. While these experimental results suggest that GPe-Arky responses have the appropriate timing to cancel ongoing actions by inhibiting striatal Go-related activity, we aimed to obtain a more detailed mechanistic understanding of the processes in the BG caused by the Stop signal.

To simulate the task of Mallet et al. (2016), we included three cortical input populations of Poisson
neurons. These cortex populations were activated at different times and for different durations according to the experimental timings from Mallet et al. (2016), particularly with respect to the onset times of the Go signal (Go cue) and Stop signal (Stop cue). The firing rate of each cortex population was regulated by setting a specific target firing rate for a given period, which was achieved as described by Equation 3:

\[
\frac{dr}{dt} = \begin{cases} 
\Delta - r & \text{if } \Delta \geq 0 \\
\Delta / \tau_{\text{UP}} & \text{if } \Delta < 0 
\end{cases}
\]  

(3)

Once a target rate \( r_{\text{target}} \) was defined, the firing rates \( r \) increased with the time constant \( \tau_{\text{UP}} \), and once the target firing rate was reset, it decreased with the time constant \( \tau_{\text{DOWN}} \) depending on the cortex population. An overview of the time course of the simulated cortical inputs of a Go trial (without Stop cue) and a Stop trial (with Stop cue) is shown in Figure 3.

To determine whether the model had executed an action, we introduced a temporal integration process (Integrator-Go) which increased its value by 0.00095 each time a spike was emitted in the GPe-Cp, and decreased its value exponentially to 0 with a time constant of \( \tau = 200 \) ms otherwise. A movement was assumed to be initiated only if the Integrator value reached a threshold of 0.13 before the end of the simulated trial. The same procedure was used in Baladron et al. (2019) and Baladron and Hamker (2015) to extract the selected action of the model. For those trials in which an action was released, the time difference between the Go cue onset and threshold crossing of the Integrator-Go was recorded as the reaction time.

We assume that a Go cue activates a sustained go-related cortical input to the BG which leads to movement initiation. This go-related input could correspond to working memory processes or motor preparation processes in the cortex which are continuously active before a behavioral response (Allen et al., 2017; Chen et al., 2017). We hypothesize that these processes are stopped by the inhibitory GPe-Cp neurons. In the model, these neurons are activated due to cortical stop-related input from sensory sources (Stop cue) and motor-related sources (movement initiation). Dodson et al. (2015) showed that besides all GPe-Arky neurons (which are likely associated with stopping), a part of the remaining GPe neurons (which Dodson et al. (2015) called GPe-Proto, but probably also included GPe-Cp cells) were activated due to movement initiation. This supports our assumption of a stop-related input to the GPe not only after Stop cues (Mallet et al., 2016), but also after movement initiation.

To simulate the effect of the GPe-Cp onto the go-related cortex population, we implemented a second integration process (Integrator-Stop) which increased its value by 0.00095 each time a spike was emitted in the GPe-Cp, and decreased its value exponentially to 0 with a time constant of \( \tau = 30 \) ms otherwise. The go-related cortex population was deactivated if the Integrator value reached a threshold of 0.13. Thus, high GPe-Cp activity was able to stop go-related cortical input.

The target firing rate of the go-related cortical input (cortex-Go) was set to 400 Hz, 75 ms after the Go cue. This caused a delayed, slowly increasing (with \( \tau_{\text{UP}} = 200 \) ms) go-related input after each Go cue. The cortex-Go population projected onto all striatal populations and the thalamus (Table 5). Since the direct pathway was activated more strongly by the cortex-Go inputs than the indirect pathway, the increasing cortex-Go input activated the striatum-SNr-thalamus-striatum loop, which caused the model to perform an action after a certain time due to increased thalamic activity. The cortex-Go target firing rate was not manually reset to zero, but by the model dynamics itself, if GPe-Cp activity was sufficiently high. The cortex-Go activity decreased with \( \tau_{\text{DOWN}} = 10 \) ms.

A second cortex population, cortex-Stop, encoded the presence of a stop-related cortical input. The activation of cortex-Stop was caused by two events, the Stop cue and the movement initiation (Integrator-Go threshold crossing). 50 ms after both events, the target firing rate of cortex-Stop was set to 400 Hz, for a duration of 5 ms after the

| Presynaptic population | Postsynaptic population | Weight | Presynaptic firing rate \((\mu, \sigma) \text{ [Hz]}\) |
|------------------------|-------------------------|--------|---------------------------------|
| GPe-Proto baseline input | GPe-Proto | 0.1 | (400, 80) |
| GPe-Arky baseline input | GPe-Arky | 0.09 | (400, 80) |
| GPe-Cp baseline input | GPe-Cp | 0.1 | (400, 80) |
| SNr baseline input | SNr | 0.1 | (400, 80) |
| STN baseline input | STN | 0.02 | (400, 80) |
| StrD1 baseline input | StrD1 | 0.7 | (10, 2) |
| StrD2 baseline input | StrD2 | 0.8 | (10, 2) |
| StrFSI baseline input | StrFSI | 0.8 | (10, 2) |
| Thalamus baseline input | Thalamus | 0.035 | (150, 30) |

TABLE 6 Weight values and firing rates for Poisson populations used to induce baseline activity in BG populations, connected via AMPA synapses.
Stop-cue and a duration of 200 ms after the movement initiation, with $\tau_{UP} = 1$ ms and $\tau_{DOWN} = 70$ ms. Thus, after each Stop cue, cortex-Stop activity rapidly increased after a delay of 50 ms, and after a 5 ms duration slowly decreased to zero again. After movement initiation, cortex-Stop also increased rapidly after a delay of 50 ms and stayed active for a longer duration (200 ms), after which it decreased to zero again. The cortex-Stop population projected onto GPe-Arky and GPe-Cp (Table 5), thus causing responses in these nuclei both after a Stop cue and movement initiation. The different activation durations after a Stop cue and movement onset were based on the relative brief responses of GPe-Arky after Stop cues (Mallet et al., 2016) and longer responses of GPe-Arky after movement initiation (Dodson et al., 2015).

A third cortical population, cortex-Pause, encoded input to the STN (Table 5). The STN showed fast, unspecific responses after Go cues and Stop cues, regardless of whether stopping was successful or not, as observed by Schmidt et al. (2013) and elaborated by Mallet et al. (2016). Therefore, the cortex-Pause activation was set to 500 Hz directly after Go and Stop Cues, without a delay, for a duration of 5 ms, with $\tau_{UP} = 1$ ms and $\tau_{DOWN} = 150$ ms. Thus, after each Go and Stop cue, cortex-Pause activity rapidly increased, and after 5 ms slowly decreased to 0 again. The pause-then-cancel model (Mallet et al., 2016; Schmidt & Berke, 2017) suggests that the STN-SNr projection pauses the execution of a movement and "buys time" until the movement can be completely canceled.

For all group comparisons in this study, where we tested for significant differences, we applied the Kruskal–Wallis $H$-test (scipy module: stats.kruskal). Multiple testing was either corrected by applying the FDR $p$-value correction (Benjamini & Yekutieli, 2001) or the Bonferroni $p$-value correction. Which correction was used is mentioned at the corresponding results.
3 | RESULTS

To assess the functional contribution of GPe-Arky, GPe-Cp, and STN neurons to BG dynamics during stopping, we extended our previous spiking model of the BG (Baladron et al., 2019), integrating the recently characterized GPe-Arky neurons (Mallet et al., 2012) and a cortico-pallidal loop through the GPe-Cp neurons, as well as a population of StrFSIs (Figure 1). To approximate the different dynamics of GPe-Arky and GPe-Proto neurons, we estimated parameters of a simplified spiking neuron model (Figure 2), based on data by Abdi et al. (2015). For the fast-spiking interneuron model, we used previously published parameter values (Humphries et al., 2009).

We used our network model to simulate both BG dynamics and behavior in the stop-signal task as described by Mallet et al. (2016). We simulated the presentation of a Go cue by cortical activity, driving striatal and thalamic activity. In the absence of specific evidence about the origins of GPe-Arky activity evoked by Stop cues, we assumed that stop-related cortical excitatory inputs might drive GPe-Arky neurons and also the GPe-Cp neurons after a Stop cue (Abecassis et al., 2020; Karube et al., 2019; Milardi et al., 2015; Naito & Kita, 1994; Smith & Wichmann, 2015).

Both the presentation of Go and Stop cues trigger a short-latency activation of the STN by a third cortical input, consistent with a function of STN activity as an unspecific Pause signal (Mallet et al., 2016). In the following, we distinguish between Go trials, in which only a Go cue was presented, and Stop trials, in which a Stop cue followed the Go cue (Figure 3). A Go trial was considered correct if an action was initiated within the trial and a Stop trial was correct if no action was initiated.

The functioning of this model architecture in the initiation and stopping of actions relies strongly on the balance between the direct, indirect, and hyperdirect pathways, converging onto the GPi/SNr. We have chosen model parameters such that the direct pathway is slightly stronger activated by the cortex-Go input. Therefore, in Go trials in which no Stop cue occurred, the cortex-Go input makes the model perform an action. Poisson noise throughout the whole model leads to variability in reaction times. We have chosen model parameters such that a Stop cue with an SSD of 250 ms leads in 70%–80% of Stop trials to correct stopping.

3.1 | Consistency with experimental results

A typical behavioral observation in stop-signal tasks is that failed Stop trials show short reaction times, potentially indicative of separate Go and Stop processes racing for completion (Logan et al., 1984; Mallet et al., 2016). In our simulation results, we observe a similar overlap between failed Stop trials and correct Go trials with short reaction times (Figure 4b). Inspection of the model variable underlying action initiation (Integrator-Go value) confirms faster increases for failed Stop trials than for correct Stop trials already before the Stop cue, consistent with stronger go-related processes leading to early threshold crossing before the Stop process (caused by the Stop cue) becomes effective (Figure 4a).

To further demonstrate the importance of the timing of the Go responses for stopping success, we varied the cortex-Go target firing rate. Increasing the cortex-Go firing rate leads to faster Go responses (Figure 4d) and thus causes a decrease in correct Stop trials (Figure 4c). Decreasing the cortex-Go firing rate leads to slower Go responses (Figure 4d) and thus improves the stopping performance but causes a decrease in correct Go trials (Figure 4c).

Next, we wanted to demonstrate that the neural dynamics in the model qualitatively match the observations of Mallet et al. (2016). Therefore, we recorded the neural responses to Go cues and Stop cues during correct Stop trials and compared our model neural responses to the neural data of Mallet et al. (2016). We simulated 400 Stop trials which led to 292 correct Stop trials. The firing rates of the different populations were averaged over these correct Stop trials and then the mean firing rates were z-transformed for each population separately. As in the data shown by Mallet et al. (2016), STN activity in our model is driven equally strong by Go as by Stop cues (Figure 5). Also, the SNr in our model shows short-latency responses to both Go and Stop cues, similar to the experimental data. As far as the GPe-Arky and GPe-Proto response is concerned, our model replicates the observation of Mallet et al. (2016), that is a slower and more specific response to Stop cues than for STN and SNr (Figures 5 and 6a). Furthermore, the GPe-Arky neurons show stronger and faster responses to Stop cues than the GPe-Proto neurons. The strong response of GPe-Arky after the Stop cue is caused by the cortical input in our model. The later and weaker responses of the GPe-Proto neurons are caused by the inhibition of the StrD2 (Figure 5 lower row). In the data of Mallet et al. (2016), GPe-Arky further showed a weak increase after Go cues, which is missing in our data. This may indicate that the cortical inputs of GPe-Arky also respond weakly to sensory stimulation from the Go cues. This weak response after Go cues could be easily implemented in our model by activating cortex-Stop Go cues as well. However, this would not affect further model behavior.

In addition to the comparison with Mallet et al. (2016), we further studied the responses to Go and Stop cues during correct Stop trials for the other neuron populations of the model (Figure 5 lower row). Except for the STN and the SNr, no other population shows a short-latency response to the Go cue. The GPe-Cp shows similar responses to the GPe-Arky. The strong response of GPe-Cp after the Stop cue is caused by the cortical input like for the GPe-Arky. As the model was tuned to initiate an action around or shortly after
the Stop cue timing, all striatum populations start getting active shortly before the Stop cue. However, after the Stop cue, they all show a rapid decrease caused by the GPe-Arky which exerts inhibition onto the striatum populations in correct Stop trials.

To demonstrate this more clearly, Mallet et al. (2016) further compared the timing of the GPe-Arky responses in correct Stop trials with the timing of the striatum responses in correct Stop trials. To obtain the neural activity during slow Go trials, we ran 400 Go trials, which led to 218 slow Go trials (trials whose reaction times were larger than the 95% quantile of the reaction times of the 108 failed Stop trials). The response of GPe-Arky neurons after the Stop cue directly causes the divergence of StrD1 activity between correct Stop and slow Go trials (Figure 6b). The StrD2 and StrFSI responses show a very similar divergence between correct Stop and slow Go trials (Figure 6c). As the GPe populations are inhibited by StrD2 neurons, they show the opposite response as the striatal populations in slow Go trials. After action initiation in slow Go trials, the pallidal populations are activated by cortex-Stop, which leads to a decrease in the striatal populations.

The thalamus slowly increases its activity before the Stop cue and shows a rapid decrease after the Stop cue (Figure 5). The activity exhibits two separate decreases after the Stop cue. That is caused by the very fast increase in the SNr, due to the STN response, and the slightly delayed increase in the SNr, due to the inhibited striatum (Figures 5 and 6). That supports the pause-then-cancel model according to which the STN first pauses action initiation and then the GPe-Arky neurons cause

---

**FIGURE 4** Dynamics of correct Go, correct Stop, and failed Stop trials. (a) Time course of integrated thalamic activity (Integrator-Go) for correct Go, correct Stop, and failed Stop trials. (b) Distribution of reaction times (time between Go cue and threshold crossing of the integrator-Go). Trial counts for correct Go and failed Stop trials are each normalized to a maximum of one. Failed Stop trials are associated with shorter reaction times, indicating that fast Go processes are more difficult to stop. Integrator time courses and reaction time distributions were obtained from 294 correct Stop, 106 failed Stop, and 400 correct Go trials. (c and d) The correct Stop percentage/stoping performance and mean reaction time, depending on the cortex-Go firing rate. The cortex-Go firing rate was varied from 0.5 (corresponds to 50% of the normal cortex-Go firing rate) to 1.5 (corresponds to 150% of the normal cortex-Go firing rate). The means and standard deviations (indicated by the error bars) were calculated over 20 simulations, each consisting of 200 Go and Stop trials. The cortex-Go firing rate variations further illustrate that faster Go processes are more difficult to stop. Decreasing cortex-Go firing rates delays Go responses and therefore increases stopping performance and vice versa. No reaction times are shown for the failed Stop trials at the lowest cortex-Go rate because no failed Stop trials occurred in this condition. The colors indicate the different trial types.
the cancellation of the action (Mallet et al., 2016; Schmidt & Berke, 2017).

In summary, our model qualitatively reproduces the timing of slow Go and correct Stop trials at the neural level compared to data obtained by Mallet et al. (2016).

### 3.2 Comparison of successful versus failed stopping

Having established that our model behavior is generally consistent with the experimental results reported by Mallet et al. (2016), we next sought to determine the factors controlling the success of stopping in our model network. For this purpose, we compared the mean activity of the BG neural populations between correct ($N = 332$) and failed Stop trials ($N = 68$) in the time window from 300 ms before the Stop cue until 300 ms after the Stop cue. We tested for significant differences between correct and failed Stop trials. We split each trial into bins of 20 ms width, calculated the mean firing rate of these bins, and then compared the bins of all correct Stop and failed Stop trials at the corresponding time steps.

By determining in which population and when the activities differ between correct and failed Stop trials, we identified populations and processes relevant for successful stopping. We specifically highlighted the earliest differences to determine when the neural response between correct and failed Stop trials starts to diverge.

The earliest significant difference ($p < .001$ FDR corrected) occurs in the StrD1 and the SNr 120 ms before the Stop cue (Figure 7). In all striatal populations, the activity raises more quickly in failed Stop trials than in correct Stop trials. Because this already happens before the Stop cue, and the cortical inputs do not differ before the Stop cue, this divergence is caused by the random Poisson fluctuations in the model. In failed Stop trials, the quickly rising striatal activity then causes a drastic inhibition in all GPe populations as well as in the SNr, which, in turn, releases the thalamus from inhibition. Therefore, a significant divergence ($p < .001$ FDR corrected) between failed and correct Stop trials occurs in all GPe populations and the thalamus 100 ms before the Stop cue.

![Comparison of Stop and Go cue responses for the different model populations.](image)
The fact that all responses in the neural populations except for the cortical inputs and the STN differ significantly already before the Stop cue illustrates the importance of the progress of the go-related activities in the BG for stopping.

In correct Stop trials, the striatal neurons are quickly inhibited after the Stop cue. The thalamus and similarly the Integrator-Go display two decreases shortly after the Stop cue, reflecting the STN/SNr Stop responses and the slightly later inhibition of the striatum caused by GPe-Arky (Figure 7). In failed Stop trials, the already advanced go-related activities (active striatum, inhibited GPe and SNr, active thalamus) prevent the occurrence of correct Stop responses. The GPe populations are all stronger inhibited shortly after the Stop cue, making it harder for the cortex-Stop input to activate the GPe-populations, thus delaying their Stop responses. The stop response of the STN does not differ between correct and failed Stop trials, but the effect on the SNr differs significantly. During failed Stop trials, the SNr is already so strongly inhibited by the advanced striatal activity that the excitatory input from the STN has no effect; thus, there is no Stop response of the SNr in failed Stop trials.

The cortical inputs only differ significantly later in the trials, after the Stop cue (Figure 7). In failed Stop trials, the cortex-Go is active for a longer period because it is stopped by the GPe-Cp later in time. The cortex-Stop input also further increases after the Stop cue in failed Stop trials, instead of decreasing like in correct Stop trials. That is because, in failed Stop trials, an action is initiated, and therefore the cortex-Stop input gets active again (see Section 2.3). Finally, of course, the Integrator-Go only reaches its threshold in failed Stop trials.

3.3 | STN, GPe-Arky, and GPe-Cp complement each other to achieve functional stopping behavior

A question relevant to the pause-then-cancel model (Mallet et al., 2016; Schmidt et al., 2013) is the relative contribution of STN pause-related and GPe-Arky stop-related activity to the cancellation of actions. Furthermore, we wanted to investigate the relative contribution of the cortico-pallidal loop. To assess this in our model, we separately regulated the individual Stop components of each of these nuclei. For STN, the Stop component is the cortex-Pause input caused by the Stop Cue. For the GPe-Arky and GPe-Cp, the Stop components are the cortex-Stop inputs caused by the Stop cue. In Figure 8a and b, the stopping performance and reaction time distribution of failed Stop trials are shown for the model with individually activated Stop components. If none of the Stop components is activated, the model does not stop at all, and the failed Stop percentage is 100%. Activating only the GPe-Arky Stop component significantly decreases the failed Stop percentage ($M = 90.93\%$, $SD = 3.43\%$, $p < .05$ Bonferroni corrected). The reaction time distribution, a strong effect of the GPe-Arky Stop component is apparent. An active GPe-Arky Stop component prevents the model from initiating actions for a duration of about 100 ms. This period corresponds to the time during and shortly after GPe-Arky is active. However, many actions are nevertheless initiated later. Activating only the STN Stop component does not affect the stopping performance (100% failed Stop trials), but responses are delayed (Figure 8b). Compared to the GPe-Arky Stop component, the STN Stop component shows
an earlier but weaker effect on the reaction time distribution, which is explained by the time courses of the activities in the model (Figure 6a). Activating only the GPe-Cp Stop component, again significantly reduces the failed Stop percentage ($M = 86.30\%, SD = 5.98\%, p < .05$ Bonferroni corrected). Unlike the other two Stop components, the GPe-Cp Stop component alone has hardly any influence on the reaction time distribution. In summary, individually activated components are not sufficient to enable a functional stopping behavior, suggesting that a combination of them is necessary.

To examine more closely whether all three components are required, we have additionally deactivated the components individually. The corresponding effects on stopping performance and the reaction time distributions are shown in Figure 8 c and d. If none of the Stop components is deactivated, on average only 20.03\% ($SD = 7.87\%$) Stop trials fail. The number of failed Stop trials increases significantly when each Stop component is deactivated individually (all $p < .05$ Bonferroni corrected). This suggests that all three Stop components are needed to provide functional stopping behavior. Furthermore, the Stop component of GPe-Cp seems to be particularly important. By deactivating the GPe-Cp Stop component, the amount of failed Stop trials increases to 81.90\% ($SD = 4.57\%$) because the model lacks the long-term cancellation of the cortex-Go input. That demonstrates the importance of a long-term cancellation process which complements the short-term inhibition in the striatum caused by the GPe-Arky.

Evidence that lesions of the STN impair stopping performance in stop-signal tasks (Eagle et al., 2008), motivated us to investigate how actual lesions of the stop-related nuclei affect the stopping performance of the model. Therefore, we...
ran simulations consisting of 200 Stop trials where we varied the SSD (from 10 to 500 ms in 5 ms steps) in different lesion conditions (None, STN, GPe-Arky, or GPe-Cp). To simulate a lesion, we deactivated the complete output of the corresponding population. The percentage of the correct Stop trials depending on the SSD and the lesion conditions are shown in Figure 9. First of all, in the normal model without any lesions, the stopping performance decreases with increasing SSD, which is a fundamental finding in stop-signal tasks (Eagle et al., 2008; Logan et al., 1984). Lesioning the STN decreases the stopping performance of the model, but only during late SSDs. Thus, the STN has only a short-term effect on the BG dynamics, which disturbs stopping only when the go-related processes in the BG are already advanced. Thus, the model STN matches well with the Pause function. Unlike the STN, lesioning the GPe-Arky disturbs stopping performance at lower SSDs. This underlines the importance of the GPe-Arky in stopping. At very low SSDs, the Stop response of GPe-Arky takes place before the go-related processes in the BG. Nevertheless, the GPe-Arky lesion has an influence on the stopping performance even at low SSDs. This suggests that the tonic activity of GPe-Arky, which acts independent
of the actual Stop response and exerts a constant inhibition on the striatum, has a relevant effect on stopping. Finally, models with lesioned GPe-Cp are no longer able to stop actions at all. This illustrates the extreme importance of a long-term cancelation mechanism, which not only affects the striatum but also cancels the cortical go-related inputs. Compared with the results of Figure 8c, lesioning the GPe-Cp has even a stronger effect than only deactivating its stop-related input. This is because, unlike the GPe-Cp lesion, deactivating only the stop-related input of GPe-Cp does not prevent it totally from becoming active and causing the Integrator-Stop to reach its threshold.

Overall, we showed that all three nuclei: the STN, the GPe-Arky, and the GPe-Cp, are important for a functional stopping behavior. Deactivating the individual Stop responses and lesioning the whole nuclei has slightly different effects, but overall, they significantly impair stopping performance.

**3.4 | GPe-Arky contributes in stopping by inhibiting StrD2**

Besides investigating the relative contribution of the different stop-related populations in stopping, we aimed at complementing the findings of Mallet et al. (2016) by better understanding the mechanisms by which GPe-Arky contributes to stopping. GPe-Arky inhibits StrD1, the origin of the direct pathway, StrD2, the origin of the indirect pathway, and StrFSI. We investigated the role of each of these projections and determined which is critical for stopping. If the weight from the GPe-Arky to one of its targets (e.g., StrD1) is changed, the overall balance between the direct and indirect pathway is disturbed (e.g., StrD1 is constantly less inhibited, thus action initiations are faster). Therefore, to investigate the actual stop-related effect, we changed the effects on the output nuclei only during stopping without changing the tonic inhibition to the output nuclei. This was done by implementing a copy of the GPe-Arky population in the model. This GPe-Arky copy receives the same inputs except for the cortex-Stop input. Thus, we could change the weight from GPe-Arky to one of its targets and compensate for the loss of tonic input. Therefore, the output of GPe-Arky differs only when the cortex-Stop input is active.

Figure 10 shows the mean performance and reaction times of Go and Stop trials dependent on the GPe-Arky output weights. To investigate the relative contribution of the individual GPe-Arky output projections, we additionally varied the weight from cortex-Stop to GPe-Arky, which affects the whole Stop component of GPe-Arky (like in the previous investigations). As also shown in Figure 8c, decreasing the cortex-Stop input onto GPe-Arky decreases the stopping performance significantly \((M(1) = 76.52\%, SD(1) = 9.52\%, M(0) = 56.38\%, SD(0) = 9.76\%, p < .05\) Bonferroni corrected). Furthermore, it significantly increases the mean reaction time of failed Stop trials \((M(1) = 286.16\ ms, SD(1) = 4.02\ ms, M(0) = 342.43\ ms, SD(0) = 14.83\ ms, p < .05\) Bonferroni corrected). Importantly, changing the strength of this GPe-Arky Stop component does not affect the mean reaction time of correct Go trials \((M(1) = 347.99\ ms, SD(1) = 12.58\ ms, M(0) = 347.84\ ms, SD(0) = 14.95\ ms, p < .05\) Bonferroni corrected). Therefore, the change in stopping performance is not simply caused by slower Go responses like in Figure 4c,d, but actually by the change of the stop-related activity. The mean reaction times of correct Go trials also do not change significantly for the variations of the GPe-Arky output weights. This demonstrates that the compensation by the GPe-Arky copy works and only the Stop component is affected. Only varying the weight from GPe-Arky to StrD2 significantly decreases the stopping performance \((M(1) = 80.50\%, SD(1) = 6.35\%, M(0) = 47.78\%, SD(0) = 5.64\%, p < .05\) Bonferroni corrected) and increases...
the reaction times of failed Stop trials \( M(1) = 287.55 \text{ ms}, \ SD(1) = 3.65 \text{ ms}, \ M(0) = 398.98 \text{ ms}, \ SD(0) = 19.91 \text{ ms}, p < .05 \text{ Bonferroni corrected} \). This shows that the GPe-Arky Stop component participates in stopping by inhibiting StrD2 (indirect pathway) and not, as one might have expected, by inhibiting StrD1 (direct pathway). Reducing the weights from GPe-Arky to StrD2 does not cause a significant change in performance.

**FIGURE 10** The stopping performance and reaction times depending on the cortex-Stop—GPe-Arky weight and the individual GPe-Arky output weights. The corresponding weights were varied from 1 (corresponds to the standard weight in the normal model) to 0. The error bars indicate the standard deviations (SD). Decreasing the weight from cortex-Stop to GPe-Arky, that is, the whole GPe-Arky Stop component, decreases the stopping performance and increases the mean reaction time of failed Stop trials. This effect is mainly caused by the inhibitory projection from GPe-Arky to StrD2, as only decreasing the weight from GPe-Arky to StrD2 causes a significant decrease in stopping performance and an increase in mean reaction time of failed Stop trials. The variation of the different weights has neither an effect on the Go trial performance nor on the mean reaction time of Go-trials. Therefore, decreases in stopping performance are not caused by faster Go processes. The mean and standard deviation values are calculated over 20 simulations, each consisting of 200 Go and Stop trials. GPe: external globus pallidus, GPe-Arky: arky pallidal neurons, StrD1: striatal projection neurons expressing the D1-type dopamine receptor, StrD2: striatal projection neurons expressing the D2-type dopamine receptor, StrFSI: striatal fast-spiking interneurons.

the reaction times of failed Stop trials \( M(1) = 287.55 \text{ ms}, \ SD(1) = 3.65 \text{ ms}, \ M(0) = 398.98 \text{ ms}, \ SD(0) = 19.91 \text{ ms}, p < .05 \text{ Bonferroni corrected} \). This shows that the GPe-Arky Stop component participates in stopping by inhibiting StrD2 (indirect pathway) and not, as one might have expected, by inhibiting StrD1 (direct pathway). Reducing the weights from GPe-Arky to the StrFSI and StrD1 does not cause a significant change in performance.

### 3.5 | Stopping mechanisms after action initiation

Besides the neuronal activities related to stopping processes, we also investigated processes that occur around action initiation. From 400 Go and Stop trials, we obtained responses of failed Stop \( N = 62 \) and fast Go trials \( N = 98 \). In both types of trials, there occur action initiations with similar reaction times (Figure 4b). They only differ in the occurrence of the Stop cue in failed Stop trials. Thus, differences in neuronal responses are mainly related to the Stop cue. Since stop-related processes are not only triggered by the Stop cue but also by action initiation (see Section 2.3), the neuronal activities are qualitatively very similar (Figure 11). The earlier cortex-Stop activity, caused by the Stop cue (in failed Stop trials), which on average occurs before the action initiations, leads to small but significant differences between the two types of trials \( p < .01 \text{ FDR corrected, calculated like in the Section Comparison of successful vs. failed stopping} \). The responses in failed Stop trials are somewhat earlier, but qualitatively they are the same as in fast Go trials and thus also related to the action initiation (despite in STN).

All striatum populations are driven by the cortex-Go input. They start to increase slightly before the action initiation and decrease again afterward due to the stop mechanisms (active GPe-Cp, GPe-Arky). The STN gets active after the Go and Stop cue (in failed Stop trials) but there is no response during or after the action initiation period in fast Go trials.
Thus, the STN does not show responses related to action initiation but only to the Stop cue. All GPe populations show biphasic responses with an initial decrease and a subsequent increase. The decreases are caused by the inhibition from the active StrD2, and the increases by the excitation from the cortex-Stop input and the inhibition of StrD2. For GPe-Proto, the decrease is much stronger than the increase, in contrast to GPe-Arky, where the increase is much stronger than the decrease. This is mainly caused by the missing cortex-Stop input in GPe-Proto and by a stronger lateral inhibition from GPe-Proto to the other two GPe populations than vice versa.

4 | DISCUSSION

We have proposed and analyzed a spiking network model of rapid stopping behavior in the BG, integrating the recently described arkypallidal neurons (Abdi et al., 2015; Mallet et al., 2012) and a cortico-pallido-cortical loop (Abecassis et al., 2020; Chen et al., 2015; Saunders et al., 2015). Conceptually, our model follows and extends the pause-then-cancel model (Mallet et al., 2016; Schmidt & Berke, 2017), which assumes a two-stage process of behavioral inhibition: First, activity in STN is rapidly triggered by unspecific cues, serving to pause current go-related processes. Second, activity in arkypallidal neurons is triggered only by specific cues and acts as a cancellation signal (Mallet et al., 2016; Schmidt & Berke, 2017). This conceptual framework rests on the observation that arkypallidal responses evoked by Stop cues are temporally well suited to stop go-related striatal activity and that arkypallidal responses were found to be higher for correct Stop than for failed Stop trials, supporting a functional role in stopping (Mallet et al., 2016). Our model can be used as a tool to discover more details on how the arkypallidal...
Stop responses may participate in stopping. Our simulation results show that an additional cortex-projecting subpopulation of the GPe, which also gets excited after Stop cues, complements the short-term cancellation caused by the arkypallidal neurons by a long-term cancellation of go-related cortical processes.

To validate the model, we first reproduced several key observations of Mallet et al. (2016): The model reaction time distribution of failed Stop trials primarily overlaps with reaction times of fast Go trials, which is a common finding in stop-signal tasks (Eagle et al., 2008; Logan et al., 1984; Mallet et al., 2016), suggesting a race between Go and Stop processes, that is, stopping is more likely to fail for fast Go processes. At the level of response timing of BG populations, our network model qualitatively reproduces the relative timing of Stop responses in STN, SNr, arkypallidal, and prototypical GPe neurons.

4.1 The subthalamic nucleus, arkypallidal, and cortex-projecting GPe neurons in stopping

According to the pause-then-cancel theory, the STN and the arkypallidal neurons together achieve the stopping of movement initiations (Mallet et al., 2016; Schmidt & Berke, 2017). Our neuro-computational model results support this idea. In our model, the STN shows a delaying effect on action initiation due to its excitatory projection on the SNr. In the same way, the response of arkypallidal neurons specific to Stop cues prevented the initiation of actions for about 100 ms. However, our simulation results additionally suggest that the Stop responses of STN and arkypallidal neurons alone are not sufficient to enable a functioning stopping behavior. The lack of cancellation is due to the fact that the cortical go-related processes are not stopped by the arkypallidal or STN Stop responses and continue to activate the striatum after the Stop responses resulting in late responses. Apart from the poor stopping performance, findings on stop-signal-tasks speak strongly against failed Stop trials with reaction times greater than the reaction times of Go trials (Eagle et al., 2008; Logan et al., 1984; Mallet et al., 2016). Thus, we propose that cortex-projecting GPe neurons (Abecassis et al., 2020; Chen et al., 2015; Saunders et al., 2015), which also receive stop-related cortical input like arkypallidal neurons, cancel the go-related cortical processes in the model, thus preventing late failed Stop trials. Only if all three components complement each other during stopping, a fully functional stopping behavior can be achieved.

In failed Stop trials, all GPe populations and the SNr are less active compared to correct Stop trials. This is caused by stronger inhibition from the striatum during failed Stop trials. The comparison between correct and failed Stop trials, first of all, shows that the most critical issue for successful stopping is the progress of the go-related activities during the Stop cue. If the go-related activities are too advanced, stopping fails. A related finding is that the STN Stop responses of a constant magnitude across trials (regardless if correct or failed Stop) vary considerably in their effectiveness in pausing ongoing actions, owing to considerable variations of activity levels in the SNr. If the SNr is already too inhibited by the striatum, the STN does not excite the SNr anymore, which is the typical Stop response in correct Stop trials. Also, the GPe Stop responses are prevented due to strong striatal inhibition during failed Stop trials. Exactly this behavior of the STN and SNr during correct and failed Stop trials has already been confirmed by single-unit recordings in rats (Schmidt et al., 2013). This highlights the need for simultaneous recordings across BG nuclei to study their dynamical interactions on a trial-by-trial basis.

Our model allows us to investigate how certain network modifications of the BG affect the stopping performance. In this respect, we performed lesions of stop-related nuclei in the model. The individual lesions of STN, arkypallidal and cortex-projecting GPe neurons all have in common that they reduce the stopping performance. This has already been demonstrated in rats with lesions of the STN (Eagle et al., 2008). For lesions of the arkypallidal and cortex-projecting GPe neurons, our model predicts even greater deficits in stopping performance. Unlike lesions in the STN, lesions in the GPe also worsen the stopping performance even at very low SSDs. This is because the STN’s main participation in stopping is to cause a short delay by exciting the SNr. However, this delay does not occur if go-related activities have not yet been developed already (at low SSDs). In contrast, the lesion of the arkypallidal neurons not only causes the strong short-term inhibition of the striatum but also causes continuous disinhibition of the striatum. This makes it more difficult for the cortex-projecting GPe neurons to become active enough to stop the go-related cortical processes in the model. Consequently, a lesion of the cortex-projecting neurons completely prevents the model from stopping action initiation.

For the arkypallidal neurons, besides their involvement in stopping together with the STN and the cortex-projecting GPe neurons, we investigated the relative importance of their different output projections for stopping. Surprisingly, successful stopping requires mainly the inhibition of the striatal projection neurons of the indirect pathway. This fits well with findings of striatum-projecting GPe neurons, which project stronger on spiny projection neurons of the indirect pathway than of the direct pathway (Glajch et al., 2016). This is necessary because, in our model, both the direct and indirect pathways become active, which has been shown to be important for action initiation (Cui et al., 2013; Tecuapetla et al., 2016). Due to the active striatal neurons of the direct...
and indirect pathways, not only the SNr is inhibited but also the GPe. Thus, to achieve a sufficiently strong inhibition in the striatum during stopping, the arkypallidal neurons must first inhibit the striatal neurons of the indirect pathway. This allows the arkypallidal neurons to get sufficiently active to inhibit the complete striatum and additionally allows the cortex-projecting GPe neurons to become active enough to stop the Go-related cortical processes.

The short inhibitory input from the arkypallidal neurons to the striatal neurons of the direct pathway and the fast-spiking striatal interneurons during stopping does not seem to be important for stopping performance. Of course, this does not mean that the activity of these two striatal populations is not important for the stopping dynamics at all. We did not show this specifically because it is quite self-evident that when the striatal neurons of the direct pathway are more active, the go-related processes develop faster, and action initiation is harder to stop in the model. Furthermore, the fast-spiking interneurons directly regulate the activity of the striatal projection neurons.

The performance in stop-signal tasks is moderately but consistently reduced in Tourette syndrome (Ganos et al., 2014, 2018; Morand-Beaulieu et al., 2017), which is associated with an altered distribution of interneurons in the BG (Kalanithi et al., 2005). Among others, a reduction in striatal inhibition has been proposed as a mechanism underlying Tourette syndrome (Jahanshahi & Rothwell, 2017) and has been shown to produce motor tics in a rodent model of Tourette syndrome (Vinner et al., 2017). A reduction in striatal inhibition due to a lack of interneurons would also cause faster Go processes (SNr stronger inhibited), which would be harder to stop (GPe stronger inhibited) in our model.

### 4.2 Relation to previous computational models

Inspired by experiments in which a lesion to the STN caused premature responding in rats (Baunez et al., 2002, 2007) and others in which the STN showed stopping-related activation (Aron & Poldrack, 2006), several neuro-computational models have focused on the role of the STN in the inhibition of actions (Baladron & Hamker, 2015; Frank, 2006; Frank et al., 2007). In these models, the STN can initially activate the GPi through its excitatory connection, increasing the afferent inhibition at the thalamus and therefore increasing the threshold to initiate an action. Numerical experiments have shown that this function is useful in conflicting situations. Furthermore, Frank et al. (2007) have related the increased impulsivity in Parkinson's disease patients with STN deep brain stimulation to potential dysfunctions in this process.

Wiecki and Frank (2013) have extended their previous computational models to simulate a stop-signal task. In this approach, the right inferior frontal cortex is included, with direct connections to the STN, adding a global stopping mechanism. The presence of the stop signal activates the frontal cortex, which transiently activates the STN. Next, STN excites the SNr to cancel inhibition produced by the direct pathway. In our model, in addition to the STN, the arkypallidal and cortex-projecting GPe neurons are required for successful stopping. Certainly, it would also be conceivable in our model that the STN alone would be able to prevent actions for a long time. For this, however, the STN would have to show a significantly longer Stop response to continuously excite the SNr, which would not be consistent with the recent findings of Mallet et al. (2016).

Another spiking model of the BG, additionally with a pallido-striatal projection, has been presented by Wei and Wang (2016). They also simulated a stop-signal task with their model, in which a single Stop signal drives STN activity and is further relayed to the SNr and the striatum via the pallido-striatal projection. They neither divided the striatum nor the GPe into different cell types. Therefore, they could not investigate the relative contributions of these cell types and the resulting projections to stopping, as we did. Furthermore, these simplifications led to the fact that several neuronal responses they observed in their model differ from experimental findings (Mallet et al., 2016; Schmidt et al., 2013), which we could mainly replicate with our model. However, their main focus was on the model's behavioral performance rather than reproducing the time course of neural responses observed experimentally.

Suryanarayana et al. (2019) have proposed a computational model of the BG, which focuses on the topology of GPe cells, dividing the neurons into inner, outer, arkypallidal, and prototypical. Although the authors studied the role of each type of cell on action selection rather than stopping, their numerical experiments indicate that the arkypallidal cells can stop striatal activity and thus prevent action selection. They suggest that the prototypical cells are the main regulator of the arkypallidal neurons and that they can switch them off and on. Also, in our model, we considered relatively high inhibitory projections from the prototypical cells to the arkypallidal cells. However, regarding the observations of Mallet et al. (2016), which show that also the prototypical neurons get active during stopping, we consider it unlikely that during stopping, the prototypical neurons activate the arkypallidal neurons. We rather assume that in this situation, a cortical input (Abecassis et al., 2020; Karube et al., 2019; Milardi et al., 2015; Naito & Kita, 1994; Smith & Wichmann, 2015) is the main contributor to the activity of the arkypallidal neurons.

Nevado-Holgado et al. (2014) investigated in their model of the STN-GPe network the role of arkypallidal and prototypical GPe neurons for oscillations occurring in Parkinson's disease. By fitting their model to
physiological data, they proposed a connectivity regime of the STN-GPe network showing some similarities to our model. First, they proposed a stronger connection from prototypical GPe neurons to arkypallidal neurons than the other way round, which is also implemented in our model. Furthermore, their fitting process led to very low weights from arkypallidal neurons to the STN, which is consistent with findings showing that the arkypallidal neurons mainly target the striatum instead of downstream nuclei such as the STN (Glajch et al., 2016; Hernández et al., 2015; Mallet et al., 2012). Based on these findings, we did not implement the projection from arkypallidal neurons to STN at all. However, their model also made some suggestions that are not compatible with our model. First, their fitted model suggests strong connections from the thalamus to STN and arkypallidal neurons. These connections are missing in our model. In the model of Nevado-Holgado et al. (2014), these connections were likely necessary because STN and arkypallidal neurons should fire in phase to fit their physiological measurements. This is, of course, favored by a common excitatory input. In our model, unlike Nevado-Holgado et al. (2014), both the arkypallidal and STN neurons and not only the STN neurons receive direct cortical input. During phases of general cortical activation and cortical slow-wave activity, these direct connections could also enable both the STN and arkypallidal neurons to fire in phase with the cortex. Nevertheless, a direct thalamic projection to arkypallidal neurons could be a potential extension to our model and may be related to activation of the arkypallidal neurons by motor-related activity during movement onset (Dodson et al., 2015).

Second, their adapted model suggests that the striatum projects more strongly on prototypical neurons than on arkypallidal neurons. This contradicts the connectivity of our model since the inhibitory projections of striatal neurons of the indirect pathway on arkypallidal neurons are an important functional component of our model. It would certainly be possible to implement that arkypallidal neurons stop go-related striatal activity by inhibiting the striatal neurons of the direct pathway instead of the indirect pathway when no strong inhibition from striatal neurons of the indirect pathway on the arkypallidal neurons exists. However, this would eliminate the need for the stronger projection of arkypallidal neurons on the striatal neurons of the indirect pathway (Glajch et al., 2016). Furthermore, Nevado-Holgado et al. (2014) propose a stronger connection from STN to arkypallidal neurons than to prototypical neurons, which is conflicting with the finding of Mallet et al. (2016) that arkypallidal neurons do not reliably follow the activation of the STN. These controversies underline the importance of revealing more accurate information about the inputs of arkypallidal neurons in the future.

4.3 | GPe responses related to action initiation

Although our focus is on the stopping mechanisms, especially in the GPe, we further investigated the behavior of our model around the action initiation period. To combine the seemingly counterintuitive findings that arkypallidal neurons get active during both stopping (Mallet et al., 2016) and movement onset (Dodson et al., 2015), we hypothesize that in both cases, the movement preparation processes, which may be represented by sustained cortical activities, must be stopped. This could be realized or at least supported by the stop-related arkypallidal and cortex-projecting GPe populations. Mallet et al. (2016) and Dodson et al. (2015) together showed that arkypallidal neurons reacted more consistent and stronger than the prototypical GPe neurons during stopping and movement onset, which suggests a similar underlying activation process, like stop-related cortical inputs which preferably target arkypallidal neurons. In our model, the arkypallidal neurons respond with a stronger increase during stopping and action initiation than the prototypical GPe neurons. Nevertheless, if we compare the responses that take place around action initiation exactly with those of Dodson et al. (2015) or Mallet et al. (2016), there are some differences. The data on failed Stop and fast Go trials from Mallet et al. (2016) do not show a clear temporal profile. During slow Go trials, they show a decrease in both arkypallidal and prototypical GPe neurons, but an increase is missing. However, they only show the activities around action initiation for a very small fraction of the GPe neurons, and they further show that there are other cells with a significant response during failed Stop and fast/slow Go trials around action initiation, as predicted by our model. Dodson et al. (2015) focused on the period around action initiation, and they show that prototypical GPe neurons both increase and decrease around action initiation but do not show biphasic responses as in our model and arkypallidal neurons show only increases. Thus, our model responses only partially fit their data without any particular parameter tuning.

Although they did not distinguish between prototypical and arkypallidal neurons, several other studies have investigated the responses of GPe during action initiation. These showed that the responses are modulated by many factors, such as direction and velocity of movement (Georgopoulos et al., 1983), duration of movement (Anderson & Turner, 1991), or reward prediction (Arkadir et al., 2004). These are all factors that probably differ between Mallet et al. (2016) and Dodson et al. (2015) and thus may cause different responses during the action initiation period. The findings of Turner and Anderson (2005) are particularly interesting in this respect. They showed that the response of GPe neurons during action initiation is dependent on the context of the task. Decreases are more likely to be associated with
sensory-driven movements (e.g., due to stimuli as in Mallet et al., 2016), and increases are more likely to be associated with self-driven movements (e.g., spontaneous movements as in Dodson et al., 2015). Our model does not take these various factors which affect the GPe response into account. Furthermore, action initiation in our model is modeled very simplified by a thalamus integrator neuron, which reaches a threshold and actual movement-related processes are not modeled, except for a cortical input, which may represent a kind of motor feedback. Thus, our model does not map the exact responses of Dodson et al. (2015) and Mallet et al. (2016), but it does provide possible mechanisms for the biphasic responses around action initiation that have been observed in other studies (Anderson & Turner, 1991; Turner & Anderson, 2005). We hypothesize that movements that are more likely to lead to decreased firing in GPe (sensory-driven) should be associated with increased StrD2 activity, and movements that are more likely to lead to increased firing in GPe (self-driven) should be associated with increased motor feedback from the cortex or thalamus toward GPe. In future research, stopping (successful and failed) and action initiation should be examined together using physiological recordings of multiple BG nuclei, taking into account the various factors that influence the GPe response during action initiation.

4.4 | Limitations

A few remarks are in place regarding the limitations of our study. First, we used a fixed stop-signal delay throughout our simulations, whereas this parameter is typically varied between sessions in experimental tasks. For example, Mallet et al. (2016) varied the SSD to maintain a desired balance between successful and failed Stop trials. However, our model considers no learning or anticipation effects as in animals; therefore, this simplification is not critical. Regarding the comparison with the data of Mallet et al. (2016), we have to note that despite a predominantly very good agreement with the data, a few phenomena in our model differ slightly. First, the z-scored firing rates in our Figure 5 have a different pre-stimulus baseline than in the data shown by Mallet et al. (2016). We determined our z-scores over the mean population firing rate averaged over all trials of a condition (e.g., failed Stop trials). Mallet et al. (2016) calculated their z-scores using the session-wide mean and standard deviation of the individual units. The z-scores of the SNr and GPe neurons of Mallet et al. (2016) start at values greater than or similar to one, that is, the session-wide mean used by Mallet et al. (2016) includes activities that are below the shown data. This is not the case in our Figures. For our z-score calculation, we only used the shown activities over the time course of one trial, which usually show only a short increase in activity, so that our z-scores usually start at values below 0. Nevertheless, the changes and timings of the activities overall match very well for the data of Mallet et al. (2016). The only exception is the second increase in SNr activity after Stop cues. This second increase is caused by the inhibition of the striatal neurons of the direct pathway. This indicates that the inhibition of the striatum in our model is a bit too strong.

Although there exists encouraging evidence for sustained go-related cortical input to the striatum (Allen et al., 2017; Chen et al., 2017), direct cortical input to the GPe (Abecassis et al., 2020; Karube et al., 2019; Milardi et al., 2015; Naito & Kita, 1994; Smith & Wichmann, 2015), and GABAergic projections from the GPe to the cortex (Abecassis et al., 2020; Chen et al., 2015; Saunders et al., 2015), we still made several assumptions about the timings and the target specificity of the cortical processes in our model. In the absence of concrete evidence, we assumed that stop-related cortical processes specifically target the stop-related arkypallidal and cortex-projecting GPe cells, which may drive their selective responses to Stop cues. Increasing evidence suggests at least a direct projection from cortical motor-related areas to GPe cells (Karube et al., 2019; Milardi et al., 2015; Naito & Kita, 1994) but also that both striatum-projecting GPe cells and subthalamic-projecting GPe cells receive such cortical input (Karube et al., 2019). Nevertheless, we assume that there are different cortical inputs for the GPe, and that the stop-related inputs mainly target the stop-related arkypallidal and cortex-projecting GPe neurons. This could also be related to the fact that especially arkypallidal neurons increase their firing rate during stopping (Mallet et al., 2016) but also after movement onset (Dodson et al., 2015), to stop processes related to the preparation of movement in both cases.

More generally, any network-level neuronal model is faced with a tradeoff between biological fidelity and computational constraints. As such, while we chose our neuron model parameters in good agreement with physiological data, especially the parameters of arkypallidal and prototypical GPe neurons, which we directly fitted to available physiological data, the weight strengths between model nuclei, the cortical inputs to the model, and the baseline inputs are rather abstract and were determined mainly by functional constraints. Furthermore, we assume that stopping cortical processes is a very interesting possible function of cortex-projecting GPe neurons, but its implementation in our model is still very abstract. The self-sustained go activity in the cortex could probably be produced by a more complex network (e.g., attractor network with two states, Durstewitz et al., 2000) whose active state could be switched to the inactive state by the direct inhibitory projection from the GPe-Cp. Anyway, the output of the cortex-projecting GPe neurons is probably not the only mechanism that can stop the go-related cortical inputs. Therefore, the results regarding the role of the cortex-projecting GPe neurons in stopping are probably overestimated. A lesion of a specific GPe population does not lead to a complete inability in stopping. Nevertheless, we
hypothesize that these GPe neurons make a major contribution to stopping cortical processes. By contrast, more detailed modeling approaches that allow to incorporate published data about synaptic strength values (e.g., Corbit et al., 2016) are difficult to scale to the network or even behavioral level because of prohibitive computational complexity.

5 | CONCLUSION

In summary, our model provides a detailed mechanistic account for the role of arkypallidal neurons in the stopping of ongoing actions. We demonstrate that STN, arkypallidal, and cortex-projecting GPe neurons are all involved in successful stopping and thus extend the pause-then-cancel model, which assumes only the STN and arkypallidal components. We provide testable predictions for both causal interventions and large-scale recordings across basal ganglia nuclei, which may help to advance our understanding of the dynamical circuit mechanisms of action initiation and suppression.

ACKNOWLEDGMENTS

This work has been supported by a CRCNS US-German-Israeli collaboration on computational neuroscience grant “Multi-level neuro-computational models of basal ganglia dysfunction in Tourette syndrome” funded by the Federal Ministry of Education and Research, Germany (BMBF 01GQ1707) and by the SMARTSTART Programme of the Bernstein Network funded by the Volkswagen Foundation. We thank Jonathan E. Rubin and Izhar Bar-Gad for their helpful comments on a previous version of this manuscript. Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHORS’ CONTRIBUTIONS

Designed the research: FH and JB. Performed the research: IK, LG, and OM. Guided the research: FH and JB. Programming: JB, IK, LG, and OM. Data analysis: IK, LG, and OM. Acquired funding: FH. Writing (first draft): LG, JB, and IK. Writing (reviewing and editing): LG, JB, FH, and OM.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15082.

DATA AVAILABILITY STATEMENT

Our study is theoretical (computational modeling). We did not collect any data from human subjects or animals. All simulation code is available on GitHub (https://github.com/Olimaol/stopsignaltask_BG).

REFERENCES

Abdi, A., Mallet, N., Mohamed, F. Y., Sharott, A., Dodson, P. D., Nakamura, K. C., Suri, S., Avery, S. V., Larvin, J. T., Garas, F. N., Garas, S. N., Vinciati, F., Morin, S., Bezd, E., Baufreton, J., & Magill, P. J. (2015). Prototypic and arkypallidal neurons in the dopamine-intact external globus pallidus. Journal of Neuroscience, 35, 6667–6688.

Abecassis, Z. A., Berceau, B. L., Win, P. H., García, D., Xenias, H. S., Cui, Q., Pamukcu, A., Cherian, S., Hernández, V. M., Chon, U., Lim, B. K., Kim, Y., Justice, N. J., Awatramani, R., Hooks, B. M., Gerfen, C. R., Boca, S. M., & Chan, C. S. (2020). Npas1+–Nkx2.1+ neurons are an integral part of the cortico-pallido-cortical loop. Journal of Neuroscience, 40, 743–768.

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. Trends in Neurosciences, 12, 366–375.

Allen, W. E., Kauvar, I. V., Chen, M. Z., Richman, E. B., Yang, S. J., Chan, K., Gradinaru, V., Deverman, B. E., Luo, L., & Deisseroth, K. (2017). Global representations of goal-directed behavior in distinct cell types of mouse neocortex. Neuron, 94, 891.e6–907.e6. https://doi.org/10.1016/j.neuron.2017.04.017

Anderson, M. E., & Turner, R. S. (1991). A quantitative analysis of pallidal discharge during targeted reaching movement in the monkey. Experimental Brain Research, 86, 623–632.

Arkadir, D., Morris, G., Vaadia, E., & Bergman, H. (2004). Independent coding of movement direction and reward prediction by single pallidal neurons. Journal of Neuroscience, 24, 10047–10056.

Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to stop signal response inhibition: Role of the subthalamic nucleus. Journal of Neuroscience, 26, 2424–2433. https://doi.org/10.1523/JNEUROSCI.4682-05.2006

Baladron, J., & Hamker, F. H. (2015). A spiking neural network based on the basal ganglia functional anatomy. Neural Networks, 67, 1–13.

Baladron, J., Nambu, A., & Hamker, F. H. (2019). The subthalamic nucleus-external globus pallidus loop biases exploratory decisions towards known alternatives: A neuro-computational study. European Journal of Neuroscience, 49, 754–767. https://doi.org/10.1111/ejn.13666

Baunez, C., Amalric, M., & Robbins, T. W. (2002). Enhanced food-related motivation after bilateral lesions of the subthalamic nucleus. Journal of Neuroscience, 22, 562–568.

Baunez, C., Christakou, A., Chudasama, Y., Forni, C., & Robbins, T. W. (2007). Bilateral high-frequency stimulation of the subthalamic nucleus on attentional performance: Transient deleterious effects and enhanced motivation in both intact and parkinsonian rats. European Journal of Neuroscience, 25, 1187–1194.

Baunez, C., Humby, T., Eagle, D. M., Ryan, L. J., Dunnett, S. B., & Robbins, T. W. (2001). Effects of STN lesions on simple vs choice reaction time tasks in the rat: Preserved motor readiness, but impaired response selection. European Journal of Neuroscience, 13, 1609–1616.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. Annals of Statistics, 29, 1165–1188.

ORCID

Lorenz Goenner https://orcid.org/0000-0002-2840-8791
Oliver Maith https://orcid.org/0000-0002-6446-9110
Iliana Koulouri https://orcid.org/0000-0001-9203-7799
Fred H. Hamker https://orcid.org/0000-0001-9104-7143
Bergman, H., Feingold, A., Nini, A., Raz, A., Slovin, H., Abeles, M., & Vaadia, E. (1998). Physiological aspects of information processing in the basal ganglia of normal and parkinsonian primates. *Trends in Neurosciences*, 21, 32–38.

Bevan, M. D., Booth, P. A. C., Eaton, S. A., & Bolam, J. P. (1998). Selective innervation of neostriatal interneurons by a subclass of neuron in the globus pallidus of the rat. *Journal of Neuroscience*, 18, 9438–9452.

Bogacz, R., Moraud, E. M., Abdi, A., Magill, P. J., & Baufreton, J. (2016). Properties of neurons in external globus pallidus can support optimal action selection. *PLoS Computational Biology*, 12, e1005004. https://doi.org/10.1371/journal.pcbi.1005004

Boraut, T., Bezard, E., Bioulac, B., & Gross, C. (1996). High frequency stimulation of the internal Globus Pallidus (GPi) simultaneously improves parkinsonian symptoms and reduces the firing frequency of GPe neurons in the MPTP-treated monkey. *Neuroscience Letters*, 215, 17–20. https://doi.org/10.1016/S0304-3940(96)12943-8

Burke, D. A., Rotstein, H. G., & Alvarez, V. A. (2017). Striatal local circuitry: A new framework for lateral inhibition. *Neuron*, 96, 267–284. https://doi.org/10.1016/j.neuron.2017.09.019

Chen, M. C., Ferrari, L., Sacchet, M. D., Foland-Ross, L. C., Qiu, B., Boraud, T., Bezard, E., Bioulac, B., & Gross, C. (1996). Identification of a direct GABAergic pallidocortical pathway in rodents. *European Journal of Neuroscience*, 41, 748–759.

Chen, T.-W., Li, N., Daie, K., & Svoboda, K. (2017). A map of anticipatory activity in mouse motor cortex. *Neuron*, 94, 866–879.e4. https://doi.org/10.1016/j.neuron.2017.05.005

Corbit, V. L., Whalen, T. C., Zitelli, K. T., Crilley, S. Y., Rubin, J. E., & Gittis, A. H. (2016). Palindromic projections promote β oscillations in a dopamine-depleted biophysical network model. *Journal of Neuroscience*, 36, 5556–5571.

Cui, G., Jun, S. B., Jin, X., Pham, M. D., Vogel, S. S., Lovinger, D. M., & Costa, R. M. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature*, 494, 238–242. https://doi.org/10.1038/nature11846

DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in Neurosciences*, 13, 281–285.

Deniau, J. M., & Chevalier, G. (1985). Disinhibition as a basic process in the expression of striatal functions. II. The striatogniral influence on thalamocortical cells of the ventromedial thalamic nucleus. *Brain Research*, 334, 227–233. https://doi.org/10.1016/0006-8993(85)90214-8

Dodson, P. D., Larvin, J. T., Duffell, J. M., Garas, F. N., Doig, N. M., Kessaris, N., Duguid, I. C., Bogacz, R., Butt, S. J. B., & Magill, P. J. (2015). Distinct Developmental origins manifest in the specialized encoding of movement by adult neurons of the external globus pallidus. *Neuron*, 86, 501–513. https://doi.org/10.1016/j.neuron.2015.03.007

Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000). Neurocomputational models of working memory. *Nature Neuroscience*, 3, 1184–1191.

Eagle, D. M., Baunze, C., Hutcheson, D. M., Lehmann, O., Shah, A. P., & Robbins, T. W. (2008). Stop-signal reaction-time task performance: Role of prefrontal cortex and subthalamic nucleus. *Cerebral Cortex*, 18, 178–188. https://doi.org/10.1093/cercor/bhm094

Eagle, D. M., & Robbins, T. W. (2003). Inhibitory control in rats performing a stop-signal reaction-time task: Effects of lesions of the medial striatum and d-amphetamine. *Behavioral Neuroscience*, 117, 1302–1317. https://doi.org/10.1037/0735-7044.117.6.1302

Frank, M. J. (2006). Hold your horses: A dynamic computational role for the subthalamic nucleus in decision making. *Neural Networks*, 19, 1120–1136. https://doi.org/10.1016/j.neunet.2006.03.006

Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007). Hold your horses: Impulsivity, deep brain stimulation, and medication in parkinsonism. *Science*, 318, 1309–1312. https://doi.org/10.1126/science.1146157

Fujiyama, F., Nakano, T., Matsuda, W., Furuta, T., Udagawa, J., & Kaneko, T. (2016). A single-neuron tracing study of arky pallidal and prototypical neurons in healthy rats. *Brain Struct. Funct.*, 221, 4733–4740.

Ganos, C., Kühn, S., Kahl, U., Schonke, O., Feldheim, J., Gerloff, C., Roessner, V., Bäumer, T., Thomalla, G., Haggard, P., & Münchau, A. (2014). Action inhibition in Tourette syndrome. *Movement Disorders*, 29, 1532–1538.

Ganos, C., Rothwell, J., & Haggard, P. (2018). Voluntary inhibitory motor control over involuntary tic movements. *Movement Disorders*, 33, 937–946.

Georgopoulos, A. P., DeLong, M. R., & Crutcher, M. D. (1983). Relations between parameters of step-tracking movements and single cell discharge in the globus pallidus and subthalamic nucleus of the behaving monkey. *Journal of Neuroscience*, 3, 1586–1598. https://doi.org/10.1523/JNEUROSCI.03-08-01586.1983

Glajch, K. E., Kelver, D. A., Hegeman, D. J., Cui, Q., Xenisias, H. S., Augustine, E. C., Hernández, V. M., Verma, N., Huang, T. Y., Luo, M., Justice, N. J., & Savio Chan, C. (2016). Npas1+-pallidal neurons target striatal projection neurons. *Journal of Neuroscience*, 36, 5472–5488.

Haber, S. N. (2003). The primate basal ganglia: Parallel and integrative networks. *Journal of Chemical Neuroanatomy*, 26, 317–330.

Hernández, V. M., Hegeman, D. J., Cui, Q., Kelver, D. A., Fiske, M. P., Glajch, K. E., Pitt, J. E., Huang, T. Y., Justice, N. J., & Chan, C. S. (2015). Parvalbumin+-neurons and Npas1+-neurons are distinct neuron classes in the mouse external globus pallidus. *Journal of Neuroscience*, 35, 11830–11847.

Humphries, M. D., Wood, R., & Gurney, K. (2009). Dopamine-modulated dynamic cell assemblies generated by the GABAergic striatal microcircuit. *Neural Networks.*, 22, 1174–1188.

Izhikevich, E. M. (2004). Which model to use for cortical spiking neurons? *IEEE Transactions on Neural Networks*, 15, 1063–1070.

Jahanshahi, M., & Rothwell, J. C. (2017). Inhibitory dysfunction contributes to some of the motor and non-motor symptoms of movement disorders and psychiatric disorders. *Philosophical Transactions of the Royal Society of London. Series B*, 372, 20160198.

Kalaniithi, P. S. A., Zheng, W., Kataoka, Y., DiFiliga, M., Grantz, H., Saper, C. B., Schwartz, M. L., Leckman, J. F., & Vaccarino, F. M. (2005). Altered parvalbumin-positive neuron distribution in basal ganglia of individuals with Tourette syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 13307–13312.

Kurube, F., Takahashi, S., Kobayashi, K., & Fujiyama, F. (2019). Motor cortex can directly drive the globus pallidus neurons in a projection neuron type-dependent manner in the rat. *eLife*, 8, e49511.

Kita, H. (2007). Globus pallidus external segment. In J. M. Tepper, E. D. Abercrombie, & J. P. Bolam (Eds.), *GABA and the basal ganglia, progress in brain research* (pp. 111–133). Elsevier.

Kita, H., Chiken, S., Tachibana, Y., & Nambu, A. (2006). Origins of GABA and GABAB receptor-mediated responses of globus pallidus induced after stimulation of the putamen in the monkey. *Journal of Neuroscience*, 26, 6554–6562. https://doi.org/10.1523/JNEUROSCI.1543-06.2006

Kita, H., & Kitai, S. T. (1994). The morphology of globus pallidus projection neurons in the rat: An intracellular staining study. *Brain Research*, 636, 308–319. https://doi.org/10.1016/0006-8993(94)91030-8
Kreitzer, A. C., & Malenka, R. C. (2008). Striatal plasticity and basal ganglia circuit function. *Neuron, 60*, 543–554. https://doi.org/10.1016/j.neuron.2008.11.005

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: A model and a method. *Journal of Experimental Psychology, 10*, 276–291.

Mallet, N., Micklem, B. R., Henny, P., Brown, M. T., Williams, C., Bolam, J. P., Nakamura, K. C., & Magill, P. J. (2012). Dichotomous organization of the external globus pallidus. *Neuron, 74*, 1075–1086. https://doi.org/10.1016/j.neuron.2012.04.027

Mallet, N., Schmidt, R., Leventhal, D., Chen, F., Amer, N., Boraud, T., & Berke, J. D. (2016). Arky pallidal cells send a stop signal to striatum. *Neuron, 89*, 308–316. https://doi.org/10.1016/j.neuron.2015.12.017

Mastro, K. J., Bouchard, R. S., Holt, H. A. K., & Gittis, A. H. (2014). Transgenic mouse lines subdivide external segment of the globus pallidus (GPe) neurons and reveal distinct GPe output pathways. *Journal of Neuroscience, 34*, 2087–2099.

Milardi, D., Gaeta, M., Marino, S., Arrigo, A., Vaccarino, G., Mormina, E., Rizzo, G., Milazzo, C., Finocchio, G., Baglieri, A., Anastasi, G., & Quartarone, A. (2015). Basal ganglia network by constrained spherical deconvolution: A possible cortico-pallidal pathway? *Movement Disorders, 30*, 342–349. https://doi.org/10.1002/mds.25995

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology, 50*, 381–425.

Morand-Beaulieu, S., Grot, S., Lavoie, J., Leclerc, J. B., Luck, D., & Lavoie, M. E. (2017). The puzzling question of inhibitory control in Tourette syndrome: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 80*, 240–262.

Naito, A., & Kita, H. (1994). The cortico-pallidal projection in the rat: An anterograde tracing study with biotinylated dextran amine. *Brain Research, 653*, 251–257. https://doi.org/10.1016/0006-8993(94)90397-2

Nambu, A., Tokuno, H., Hamada, I., Kita, H., Imanishi, M., Akazawa, T., Ikeuchi, Y., & Hasegawa, N. (2000). Excitatory cortical inputs to pallidal neurons via the subthalamic nucleus in the monkey. *Journal of Neurophysiology, 84*, 289–300.

Nambu, A., Tokuno, H., & Takada, M. (2002). Functional significance of the cortico-subthalamic-pallidal “hyperdirect” pathway. *Neuroscience Research, 43*, 111–117.

Nevado-Holgado, A. J., Mallet, N., Magill, P. J., & Bogacz, R. (2014). Effective connectivity of the subthalamic nucleus–globus pallidus network during Parkinsonian oscillations. *Journal of Physiology, 592*, 1429–1455.

Robledo, P., & Féger, J. (1990). Excitatory influence of rat subthalamic nucleus to substantia nigra pars reticulata and the pallidal complex: Electrophysiological data. *Brain Research, 518*, 47–54. https://doi.org/10.1016/0006-8993(90)90952-8

Sadek, A. R., Magill, P. J., & Bolam, J. P. (2007). A single-cell analysis of intrinsic connectivity in the rat globus pallidus. *Journal of Neuroscience, 27*, 6352–6362.

Saunders, A., Huang, K. W., & Sabatini, B. L. (2016). Globus pallidus externus neurons expressing parvalbumin interconnect the subthalamic nucleus and striatal interneurons. *PLoS One, 11*, e0149798. https://doi.org/10.1371/journal.pone.0149798

Saunders, A., Oldenburg, I. A., Berezovskii, V. K., Johnson, C. A., Kingery, N. D., Elliott, H. L., Xie, T., Gerfen, C. R., & Sabatini, B. L. (2015). A direct GABAergic output from the basal ganglia to frontal cortex. *Nature, 521*, 85–89. https://doi.org/10.1038/nature14179

Schmidt, R., & Berke, J. D. (2017). A pause-then-cancel model of stopping: Evidence from basal ganglia neurophysiology. *Philosophical Transactions of the Royal Society of London. Series B, 372*, 20160202.

Schmidt, R., Leventhal, D. K., Mallet, N., Chen, F., & Berke, J. D. (2013). Canceling actions involves a race between basal ganglia pathways. *Nature Neuroscience, 16*, 1118–1124.

Schroll, H., & Hamker, F. (2013). Computational models of basal-ganglia pathway functions: Focus on functional neuroanatomy. *Frontiers in Systems Neuroscience, 7*, 1–18.

Schroll, H., Vitay, J., & Hamker, F. H. (2014). Dysfunctional and compensatory synaptic plasticity in Parkinson's disease. *European Journal of Neuroscience, 39*, 688–702.

Smith, Y., & Wichmann, T. (2015). The cortico-pallidal projection: An additional route for cortical regulation of the basal ganglia circuitry. *Movement Disorders, 30*, 293–295. https://doi.org/10.1002/mds.26095

Suryanarayana, S. M., Hellgren Kotaleski, J., Grillner, S., & Gurney, K. N. (2019). Roles for globus pallidus externa revealed in a computational model of action selection in the basal ganglia. *Neural Networks, 109*, 113–136.

Taverna, S., Ilijic, E., & Surmeier, D. J. (2008). Recurrent collateral connections of striatal medium spiny neurons are disrupted in models of parkinson's disease. *Journal of Neuroscience, 28*, 5504–5512.

Tecuapetla, F., Jin, X., Lima, S. Q., & Costa, R. M. (2016). Complementary contributions of striatal projection pathways to action initiation and execution. *Cell, 166*, 703–715. https://doi.org/10.1016/j.cell.2016.06.032

Thibeault, C. M., & Srinivas, N. (2013). Using a hybrid neuron in physiologically inspired models of the basal ganglia. *Frontiers in Computational Neuroscience, 7*, 1–17.

Turner, R. S., & Anderson, M. E. (2005). Context-dependent modulation of movement-related discharge in the primate globus pallidus. *Journal of Neuroscience, 25*, 2965–2976.

Vinner, E., Israelashvili, M., & Bar-Gad, I. (2017). Prolonged striatal disinhibition as a chronic animal model of tic disorders. *Journal of Neuroscience Methods, 292*, 20–29.

Vitay, J., Dinkelbach, H., & Hamker, F. (2015). ANNarchy: A code generation approach to neural simulations on parallel hardware. *Frontiers in Neuroinformaticas, 9*, 1–20.

Wei, W., & Wang, X.-J. (2016). Inhibitory control in the cortico-basal ganglia-thalamocortical loop: Complex regulation and interplay with memory and decision processes. *Neuron, 92*, 1093–1105. https://doi.org/10.1016/j.neuron.2016.10.031

Wichmann, T., Bergman, H., Starr, P. A., Subramanian, T., Watts, R. L., & DeLong, M. R. (1999). Comparison of MPTP-induced changes in spontaneous neuronal discharge in the internal pallidal segment and in the substantia nigra pars reticulata in primates. *Experimental Brain Research, 125*, 397–409. https://doi.org/10.1007/s002210050696

Wiecki, T., & Frank, M. (2013). A computational model of inhibitory control in frontal cortex and basal ganglia. *Psychological Review, 120*, 329–355.

How to cite this article: Goenner L, Maith O, Koulouri I, Baladron J, Hamker FH. A spiking model of basal ganglia dynamics in stopping behavior supported by arky pallidal neurons. *Eur J Neurosci*. 2021;53:2296–2321. https://doi.org/10.1111/ejn.15082