Role of spinal sensorimotor circuits in triphasic command: a simulation approach using Goal Exploration Process

Daniel Cattaert[1], Matthieu Guemann[1], Florent Paclet[1], Bryce Chung[2], Pierre-Yves Oudeyer[3], Aymar de Rugy[1]

[1] Institut de Neurosciences Cognitives et Intégratives d'Aquitaine,

UMR 5287 CNRS & Univ. Bordeaux, Bordeaux, France

[2] Neuroscience Institute, Georgia State University, Atlanta, GA 30302

[3] Inria Bordeaux Sud-Ouest, Talence, France ;

Corresponding author: Daniel Cattaert

INCIA UMR 5287 – 146 rue Léo Saignat, 33076 BORDEAUX CEDEX, France

Manuscript type: Research article

Word count: 12 379

Key points Summary

- Spinal sensorimotor circuits, despite extensive physiological analyses, remain poorly understood in the context of voluntary movement.
- This study models the capabilities of these spinal networks, alongside musculoskeletal dynamics, to achieve single-joint arm flexion following minimum jerk principles.
- Preparatory (SET) and launch (GO) step commands modulate spinal network states, including neuron activity and synaptic strength.
- Notably, arm movement results from sensorimotor interactions within spinal circuits, even with basic step commands that have no dynamics.
- A Goal Exploration Process (GEP), inspired by developmental baby babbling through trial and error, was used to refine step commands.
- The spinal networks, when in appropriate states, were shown to spontaneously generate triphasic commands across a broad range of valid movement amplitudes and speeds.
- These findings highlight the intrinsic capacity of spinal sensorimotor circuits to manage complex movement dynamics, offering novel testable insights as to their possible functional role in voluntary motor control.

Abstract

During rapid voluntary limb movement about a single joint, a stereotyped triphasic pattern is typically observed in the electromyograms (EMGs) of antagonistic muscles acting at this joint. To explain the origin of such triphasic commands, two types of theories have been proposed. Peripheral theories consider that triphasic commands result from sensorimotor spinal networks, either through a combination of reflexes or through a spinal central pattern generator. Central theories consider that the triphasic command is elaborated in the brain. Although both theories were partially supported by physiological data, there is still no consensus about how exactly triphasic commands are elaborated. Moreover, capacities of simple spinal sensorimotor circuits to elaborate triphasic commands on their own have not been tested yet. In order to test this, we modelled arm musculoskeletal system, muscle activation dynamics, proprioceptive spindle and Golgi afferent activities and spinal sensorimotor circuits. Step commands were designed to modify the activity of spinal neurons and the strength of their synapses, either to prepare (SET) the network before movement onset, or to launch the movement (GO). Since these step commands do not contain any dynamics, changes in muscle activities responsible for arm movement rest entirely upon interactions between the spinal network and the musculo-skeletal system. Critically, we selected step commands using a Goal Exploration Process inspired from baby babbling during development. In this task, the Goal Exploration Process proved very efficient at discovering step commands that enabled spinal circuits to handle a broad spectrum of functional behaviors, displayed in a behavioral space characterized by movement amplitude and maximal speed. All over the behavioral space, specific SET and GO commands elicited natural triphasic commands, thereby substantiating the inherent capacity of the spinal network in generating them.

Introduction

During rapid voluntary limb movement about a single joint, a stereotyped triphasic pattern is observed in the electromyograms (EMGs) of antagonistic muscles commanding this joint (Hallett *et al.*, 1975). Prior to limb displacement, the agonist muscle is briefly active (AG1). After the limb begins to move, a burst of muscle activity occurs in the antagonist muscle (ANT) while the agonist is silenced, and then a second burst is observed in the agonist (AG2). The precise organization of this triphasic pattern ensures the smoothness of the movement.

What is the origin of this triphasic command? Two possibilities have been proposed; either a **spinal origin** via sensorimotor circuits (Barnett & Harding, 1955; Terzuolo *et al.*, 1974; Ghez & Martin, 1982) or a **central origin** elaborated in brain networks and transmitted to the spinal cord via descending commands. The **central hypothesis** was supported by observations on humans with severe, though not complete, sensory neuropathy, who could still produce triphasic patterns (Hallett *et al.*, 1975; Rothwell *et al.*, 1982). Although these studies highlight the capacities of the motor cortex and other higher centers that can learn how to control limb movements after partial or total deafferentation, the deep reorganization of motor commands following such deafferentation prevents firm conclusion about how triphasic commands are produced in healthy subjects. Other studies used reversible transient deafferentation using ischemic process of the muscles controlling the joint and their afferent nerves Ia and Ib, the diameter of which are more sensitive to ischemia than axons of alpha motoneurons (Sanes & Jennings, 1984). During such reversible ischemia, the triphasic pattern persists, although passive movements fail to elicit stretch reflexes.

Other experiments indicate that triphasic motor command is also present in the primary motor cortex (M1), and that neural activities in M1 are correlated with output kinematics and kinetics during isometric-force and arm-reaching tasks (Sergio *et al.*, 2005). Physiological model describing the transformation from the neural activity in M1, through the muscle control signal, into joint torques and down to endpoint forces and movements (Trainin *et al.*, 2007), show that optimized solutions for M1 activities display triphasic changes in temporal pattern and instantaneous directionality similar to that found in the EMG (Trainin *et al.*, 2007).

In opposition with the arguments in favor of a central origin of triphasic commands, it was shown that procedures that modify discharge of muscle spindle afferents during movement alter the duration and amplitude of these bursts (Hallett & Marsden, 1979; Brown & Cooke, 1981). In addition, simulation studies concerning the central origin of triphasic commands, although aiming toward physiological models, often lack spinal sensorimotor loops (Trainin *et al.*, 2007). Therefore, if we cannot doubt that M1 neurons and populations of neurons present activities correlated with kinetic, kinematic and dynamic cues of the performed movements, there is still questions about the possible role of spinal sensorimotor circuits in the elaboration of triphasic commands. Notably, since both efference copy from spinal interneurons and afferent feedback from proprioceptors project onto sensorimotor cortical areas, they constitute a potential origin for triphasic patterns observed in cortical areas. Consequently, if the first EMG activity (AG1) can be attributed to a central command from the brain (at least for its onset), there is still debate about ANT and AG2.

The aim of this paper was to explore the capacities of the spinal sensorimotor circuits in shaping triphasic commands, using a modeling approach based on the neuromechanical environment AnimatLab (https://www.animatlab.com/). We modelled the movement of the forearm around the elbow joint, involving two muscles (biceps and triceps). Movements consisted in elbow flexion starting from a fully extended position. Given our aim, we could not use complex descending commands because if descending commands already contain triphasic patterns it would have been impossible to analyze spinal network circuits capacity to produce a triphasic pattern by their own. Therefore, rather than using descending commands from the brain, we aimed at searching states in which sensorimotor circuits could produce by themselves complex dynamic muscle activities observed during rapid arm flexion in human. Our approach is inspired from spinal like regulator (SLR) work (Raphael *et al.*, 2010). In short, spinal interneurons and synapses of sensorimotor circuits were controlled by simple SET and GO step commands. The SET commands shape the sensorimotor circuits (synaptic gains, level of activity of neurons) without producing any movement, and the GO commands trigger and elicit the movement. Since these signals lacked any dynamic content, the spinal sensorimotor circuit

was solely responsible for shaping the dynamics of the commands, particularly in producing the triphasic commands.

We also wanted the neuromechanical system to self-organize its commands as a function of the produced movements, as would do a child exploring the commands of its own arm. Recent advances in developmental machine learning (Baranes & Oudeyer, 2013; Forestier *et al.*, 2017) have shown that a novel form of exploration algorithms, called *Goal Exploration Processes* (**GEP**), can be used to automatically define target behaviors and efficiently discover a map of reachable behaviors, thereby solving problems of unreachable behaviors and local minima encountered by alternative, state-of-the-art optimization methods (Oudeyer & Kaplan, 2007; Forestier *et al.*, 2017; Colas *et al.*, 2018) ; (for related methods using diversity search to solve problems with local minima or sparse rewards, see also Lehman & Stanley, 2011; Mouret & Doncieux, 2012; Pathak *et al.*, 2017; Eysenbach *et al.*, 2019). Moreover, self-conducted exploration of high-dimensional, continuous parameter systems have the advantage to be sample efficient, i.e., they enable the discovery of diverse behavioral domains with a limited number of experiments. So far, such algorithms have been used to discover diverse behaviors in robotics **(**Forestier *et al.*, 2017; Péré *et al.*, 2018) as well as in complex systems such as cellular automata, morphogenetic models (Reinke *et al.*, 2019; Etcheverry *et al.*, 2020) and chemical systems (Grizou *et al.*, 2020), implementing a form of discovery assistant that can leverage interactively guidance of users (Etcheverry *et al.*, 2020). Furthermore, GEP is not only an exceptional tool for exploring behavioral domains, but this type of exploration, where an agent defines its own goals, has recently been recognized as a curiosity-driven learning mechanism. This approach is used in artificial intelligence and machine learning, similar to how it occurs in animals and humans (Gottlieb & Oudeyer, 2018).

Finding new sets of parameters and behaviors from previously achieved behaviors has already been successively used to generate intermediate movement using interpolation (Tsianos *et al.*, 2014). However, due to the large redundancy of the neural circuits, interpolation fails if previous behaviors relate to different strategies (Tsianos *et al.*, 2014). To get a systematic exploration of the capacities of the neuromechanical model, we used a simple GEP protocol inspired from Forestier et al. (2017).

These algorithms use two work spaces: (i) one for the behaviors: here we used a two dimensions space (movement amplitude and movement speed), and (ii) one for the parameters that determine the motor commands from the brain: here we used a series of step commands that modify synaptic strength and neuron activity in spinal cord circuits either in the preparatory phase (SET) or to launch the movement). The GEP finds new valid behaviors by repeating 4 stages: (1) Choose a goal (a goal is a targeted behavior); (2) Find the parameters that produced a behavior close to the aim in the data base of preceding successes and modify slightly these parameters in a random way; (3) Run the model with these parameters; (4) Observe the result and if acceptable, store behavior and parameters in the database.

Several modelling approaches have proposed that the CNS employs a minimum jerk strategy when planning any given movement (Flash & Hogan, 1985). This proposal was refined later (Nakano *et al.*, 1999) showing that minimum angle jerk predicts the actual arm trajectory curvature better than the minimum jerk model. Therefore, we used it here. Once valid behaviors were found, the GEP searched new ones by modifying preexisting valid behaviors. Because the policy was random, we called this process rGEP. Here, rGEP proved efficient at discovering step commands that enabled spinal circuits to handle a broad spectrum of functional behaviors while eliciting natural triphasic commands over the whole range of the behavioral domain.

Methods

A. Software's used in simulations.

During this work, we have used several models of sensorimotor systems controlling the same musculo-skeletal apparatus. All models were developed with the **AnimatLab software** (https://animatlab.com) (Cofer *et al.*, 2010), a neuromechanical environment based on user-friendly graphic interface that can be downloaded from https://www.animatlab.com. Execution of AnimatLab models can be controlled by **Python scripts**, which were used to command execution of model movements by **step command** targeting spinal neurons and synapses (see below). All scripts used in the present work are written in Python3.8 language. For a quick overview of GEP principles, see pseudo script in Supplementary materials. Python scripts for GEP and all simulations included in this study can be downloaded from GitHub, with an installation procedure for all software package involved:

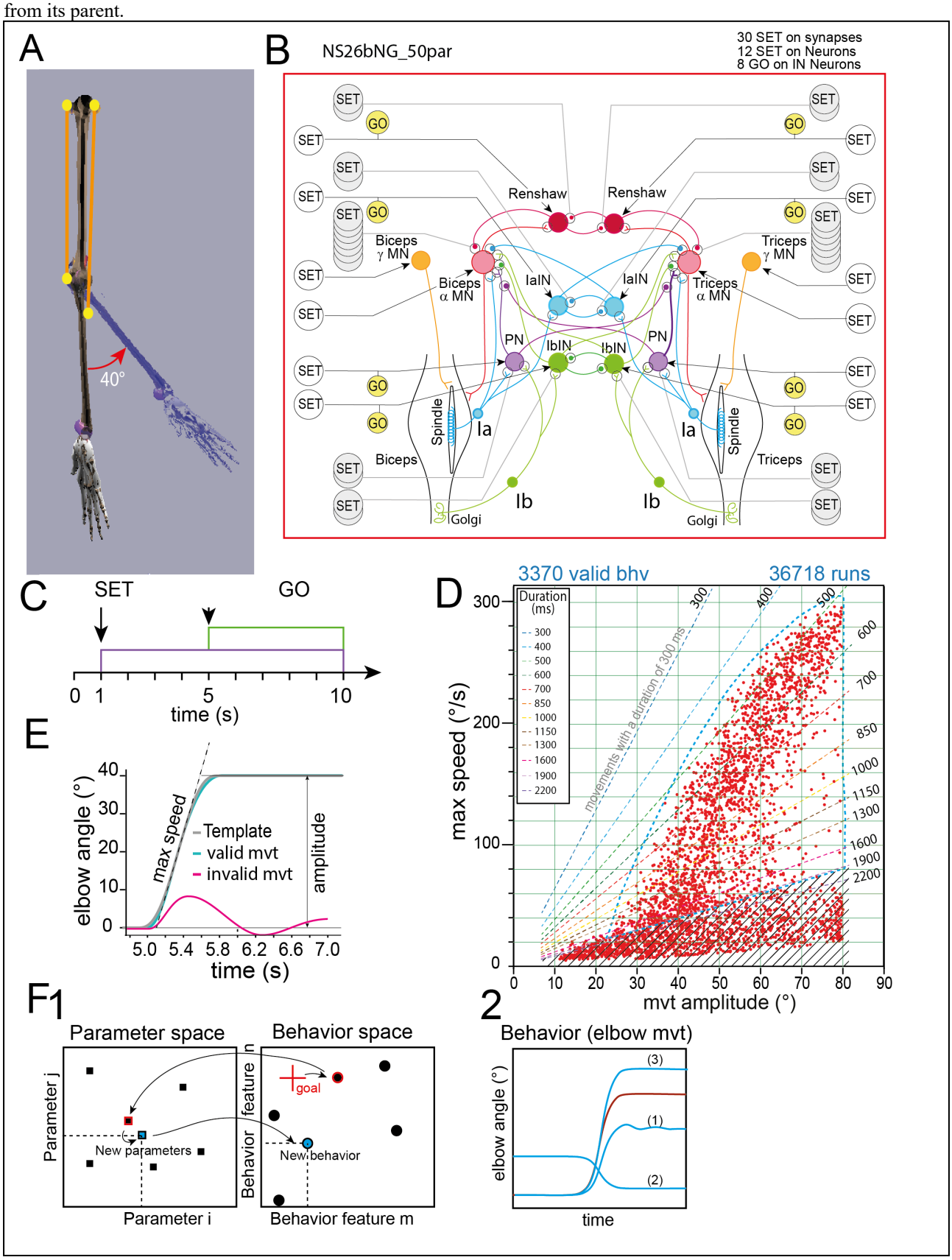
(https://github.com/Cattaert/rGEP/tree/main).

B. Musculoskeletal system

The musculoskeletal model consisted of an upper limb involving an elbow actuated by two antagonistic muscles (biceps and triceps) (Figure 1A). This single musculoskeletal model was used in all simulations, with various spinal sensorimotor circuits and step commands controlling their state (see below). Information about this musculoskeletal model is given in table 1.

Figure 1: Neuromechanical model and GEP principles.
A, Musculoskeletal system.
B, Sensorimotor system comprising αMNs, γMNs, propriospinal interneurons and Ia afferents from muscle spindles, with SET ad GO commands. In this study, simplified sensorimotor networks have been associated to the same musculo-skeletal system (see Figs 4-8).
C, Timing of SET and GO commands used to control the sensorimotor network. The SET commands adjust all neuron activity levels at 1 sec to bring the network in a state ready to trigger the movement (without producing the movement by itself). The GO commands launch the movement. SET and GO commands are the parameters of **parameter space**.
D, Example of Behavior space obtained with rGEP algorithm after 39160 iterations. Valid behaviors are indicated by red dots, and corresponding movement durations are indicated by doted lines.
E, In this work the **behavior space** was defined by the amplitude of elbow movement (abscissa) and its maximal speed (ordinate). Only, movement (green curve) fitting to a minimum jerk profile (gray curve) were accepted (i.e., mean square error (MSE) < 1 - see Methods). An example of rejected movement is presented (red curve).
F, Schematic diagram of the GEP exploration algorithm. **F1**: Process starts from an existing **parameter space** (black filled squares in *Parameter space* – for simplicity reduced to two parameters), which is associated with a **behavior space** (black filled circles in *Behavior space*), a new *goal* is given in the periphery of the known behavioral space (red cross). The algorithm searches for the closest behavior (closest neighbor), it will retrieve the parameters associated with this behavior and replay its parameters including additional noise ("*New parameters"*). This new set of parameters generates a "*New*

behavior" in the space of behaviors that is kept or not, depending on its *Cost* value). **F2**: The movement that was closest to the goal (red circle in behavior space, and red movement trace in F2) is a valid movement. However, when its parameters are changed randomly different issues are possible: i) the obtained new movement is invalid, either because there are oscillations (1) or because it is performed in the wrong direction (2); ii) the new movement is valid and may be close or far from its parent.



Segments	Length (mm)	density	mass
Humerus	337.0911395	8000 Kg/m3	1.164 Kg
Forearm	283.1808226	8000 Kg/m3	0.802 Kg
Hand	193.21	8000 Kg/m3	0.966 Kg

Joints (coordinates world positions)	Z_coord (deepness)	Y_coord (up-down)	X_coord (left-right)
Elbow	-4.586	-144.65	-5.372
wrist	14.883	-427.07	1.789

Muscle insertions (world positions)			
Biceps origin	0	150.69	26.509
Biceps insertion	-0.578	-188.28	24.146
Triceps origin	0	151.13	-21.51
Triceps insertion	0.934	-103.64	-16.507

Biceps insertion/Forearm elbow axis	52.82947367		
Triceps insertion/forearm elbow axis	42.85182289		

Table 1: musculo-skeletal model data. Coordinates and dimensions are given in mm. The origin of coordinates (0,0,0) is the center of the Humerus, that is in vertical position at rest (see Fig. 1A).

All movements were produced **in the absence of gravity** in order to avoid natural slowdown of movement speed as the forearm is flexed. Muscles used in AnimatLab are based on the Hill muscle model (Scovil & Ronsky, 2006). The parameters of the Hill model for biceps and triceps are given in Table 2 with length/tension and stimulus/tension parameters.

		Biceps	Triceps
B		1 Ns/m	1 Ns/m
Kpe		100 N/m	100 N/m
Kse		1000 N/m	1000 N/m
Length		33.898 cm	27.25 cm
MaxTension		400 N	300 N
Length-Tension	Lwidth	20 cm	11 cm
	Resting Length	40 cm	38 cm
	Range of length	27 – 34 cm (0-100°)	27.5– 32.5 cm (100-0°)
	Range of Tension%	58% - 92%	2% - 70.5%
Stimulus-Tension	Amplitude	400 N	300 N
	Steepness	100	100
	X offset	-20 mV	-30 mV
	Y offset	-4.5 N	-4 N

Table 2: Hill muscle parameters used in models for Biceps and Triceps.

C. Presentation of the sensorimotor networks and their control by step commands

Non-spiking (NS) neurons were used to simulate neurons in spinal sensorimotor networks. NS neurons allow a continuous synaptic activation that depends only on presynaptic threshold potential, presynaptic saturation potential, equilibrium potential for post-synaptic effect and synaptic gain (maximal synaptic conductance). The change in postsynaptic conductance increases from 0 (at presynaptic threshold potential) to the maximal synaptic conductance (for presynaptic saturating potential and over). Between these presynaptic threshold and saturation potentials, the postsynaptic conductance increases linearly. Such neurons present the advantage to avoid discontinuity of spikes and spiking synaptic transmission. This is specially critic in small neuron networks, when the level of activity is low. Using NS neurons can also be considered as representing population neurons instead of single neurons.

In this study, we have used several sensorimotor spinal networks, all derived from the "complete" model (Figure 1B). The complete spinal network (50 parameters model, Figure 1B) received proprioceptive feedback from muscle spindles (Ia) and Golgi Tendon Organ (GTO). The associated sensorimotor circuits were inspired from Raphael et al (2010), and contained 12 neurons (2 alpha motoneurons (MNs), 2 gamma MNs, 2 propriospinal interneurons (INs), 2 Ia INs, 2 Ib INs, and 2 Renshaw INs), and 30 synapses.

The specific aim here was to generate minimum jerk movements (i.e., fitting to this movement shape was in the cost function – see below). We used the same principles for step commands as Raphael *et al.* (2010). Step commands were represented as step functions starting at t=1s (SET command) and at t=5s (GO command) (Figure 1C). The SET commands were responsible for the state of readiness of the network. This first stimulation should activate the network without releasing any movement (Prut & Fetz, 1999; Eliasmith & Anderson, 2003).

The SET command targeted all the elements of the network (neurons and synapses) and their control was maintained constant during all the time of the simulation (up to t=10s). The GO commands were responsible for movement initiation and were maintained constant until the end of the simulation. In this example GO commands modulated the level of activity of propriospinal interneurons (PN), 1a

228 interneurons (IaINs), 1b interneurons (IbINs) and Renshaw INs (Renshaw) (see Figure 1B). Note that
229 at t = 5 seconds, the transient changes induced by the SET command had stabilized before the GO
230 command was issued. In summary, GO commands targeted the 8 INs, SET commands targeted the 12
231 neurons and adjusted the 30 synapses' gains.

232 ### D. Brief overview of rGEP

233 To assess behavior capabilities associated with a given neuromechanical model (i.e., the variety
234 of movements that can be produced (Figure 1D) and meeting some criteria like minimum jerk - see
235 below for the **cost function**), we have developed a rGEP algorithm that uses two spaces (the parameter
236 space and the related behavior space). The movements contained in the behavior space are defined by
237 their amplitude and their maximal speed (Figure 1E). At each iteration, goals (i.e., movements of given
238 amplitude and speed) are self-produced by the algorithm. Goals are disposed either on the periphery
239 of (extend strategy) or inside (fill strategy) the known behavior space. In the illustration of Figure 1F1,
240 the example goal is on the periphery of the behavior domain (see red cross). The closest behavior is
241 then searched and its associated parameters are slightly modified (**random function**) to produce a new
242 parameter set (*New parameters*). We choose a random policy in order not to introduce bias in the
243 setting of new values for parameters. Because the policy used by GEP for parameter changes was a
244 random function, this GEP was named **rGEP.** After a simulation is run (Figure 1F2), if the produced
245 movement fulfils the **acceptance criteria**, it is stored in the behavior domain (see trace (3) in Figure
246 1F2), otherwise, it is rejected (traces (1) and (2)). Acceptance criteria were as follows: (i) the produced
247 movement should fit a **minimum jerk angular kinetics** (Friedman & Flash, 2009); (ii) there should
248 be no oscillations in preparatory and post movement phases; (iii) the movement amplitude should be
249 in the range (10-110°); (iv) no co-contraction should be present during the preparatory phase
250 (controlled by SET commands, see below) and during the stabilized phase after movement (launched
251 by GO commands). A **cost function** ensured valid behaviors fulfill acceptance criteria (see methods)

252 Because rGEP needs valid behaviors to progress, the first step, was to find initial valid
253 behaviors. In order not to interfere with the design of the commands, these initial behaviors (termed
254 **seeds**) were found via a state of the art optimization method (covariance matrix association evolution

255 strategy (CMAes) (Hansen *et al.*, 2003). Briefly, a desired behavior (a movement template) was given
256 to the optimization method that worked on the parameters of the SET and GO commands to find the
257 result that best fitted the desired movement template (see next paragraph for more explanations).
258 During this process of optimization, any valid movement produced was stored and could be used as a
259 seed.

260 ### E. Optimization method used to get first behaviors (seeds).

261 Covariance Matrix Adaptation evolution strategy (CMAes) is an evolution algorithm developed
262 by Hansen (Hansen *et al.*, 2003). It works with a gradient descent with the aim to minimize a cost
263 function. This evolution algorithm is based on the principle of biological evolution. This means that
264 for each generation (iteration), new individuals are generated from variation, in a stochastic way, of
265 the current generation. To initiate the next generation, individuals with the best scores (minimal value
266 of the cost function) are selected to become the parents of the next generation. Following this rule,
267 individuals of the successive generations get closer to the best results.

268 In the evolution algorithm, the solutions that represent the new generations are issued from
269 multivariate-normal distribution based on a variance value (*sigma*) and the covariance matrix of the
270 preceding best solutions. This allows the stochastic changes made to the parameters to be oriented in
271 the direction of the solution. Moreover, from one generation to the next, the *sigma* value is adapted:
272 increased when two generations are in the same direction, and decreased if not. This allows
273 accelerating the progress of optimization.

274 In the CMAes optimization procedure, some parameters have to be given before the simulation
275 was launched including the *limits* of exploration and the *sigma* value. If the *limits* of exploration were
276 too large, new individuals could be generated too far from each other and the optimal solution might
277 not be reached. A small *sigma* will create a new generation not so different from the previous one and
278 a large one will bring more changes. All optimization series were initiated with a small *sigma* (0.005)
279 and *limits* of exploration covering 100% (i.e., between 0 and 1 for all normalized parameter values).
280 *Sigma* and *limits* were given to CMAes at initiation of optimization process.

F. Cost function used to select valid behaviors.

For each movement produced, a **cost function** estimated the proximity of movements produced by the simulation to human movement. As normal human movements follow a template based on the minimum jerk model (Flash & Hogan, 1985), the **cost function** calculates the distance (mean square error) between the produced movement and this template (Figure 1E), with the addition of the penalty for the co-contraction (*Equation 1*). This term was added after a first exploration of the method, which led to systematic co-contraction during the preparation period (result of the SET command). The result was considered valid when the *Loss* value was less than 1.

$$Cost = \sum_{t=2}^{t=10}(T_{desired}(t) - T_{actual}(t))^2 + Penalty_Prep + Penalty_Main$$ **Equation 1**,

with t representing the different sampled times of the simulation (dt=10ms).

In order to avoid solutions in which co-contraction occurred during the preparatory phase (or after movement stabilization), co-contraction penalty was added in each of these phases. *Penalty_Prep* was calculated during the preparatory phase (start=2s, end=5s), and *Penalty_Main* during the maintained position phase (start=7s, end=10s). The movement phase was not included because complex interactions may occur due to the dynamics of sensorimotor interactions. The penalty for co-contraction was calculated as indicated in *Equation 2*.

$$Penalty_i = \frac{1}{end-start} * \left(\int_{t=start}^{t=end}(VMN_{flex} - Vthr) * (VMN_{ext} - Vthr) * dt\right) * \ coeff_i$$ **Equation 2**,

in which VMN_{flex} and VMN_{ext} are the membrane potential of the flexor and extensor MNs, respectively, and V_{thr}=-60mV, is a threshold voltage under which penalty = 0.

The coefficient (*coeff_i*) was set to 100 during the preparatory phase to make sure coactivation would be rejected in this phase, and *coeff_i* was set to 0 during the maintained position after movement to allow coactivation in this last phase.

G. Time course of the rGEP algorithm

An example of rGEP time course is presented in Figure 2, on the same model presented in Figure 1 (see also inset network diagram in Figure 2). Before rGEP starts, **valid behaviors (seeds)** must be given to the algorithm (see below). In the example presented in figure 2, four seeds were given (see blue crosses in Figure 2A1-4).

Four pictures of the behavior domain extension are presented starting from 4 seeds (Figure 2A1), and extending to 100, 500 and 2000 valid behaviors (Figure 2A2-4), corresponding to 394, 1924 and 6926 runs of simulation, respectively. As the rGEP progresses, the behavior domain extends (Figure 2B, violet trace, *behavior domain area*). The behavior domain area was quantified as the number of

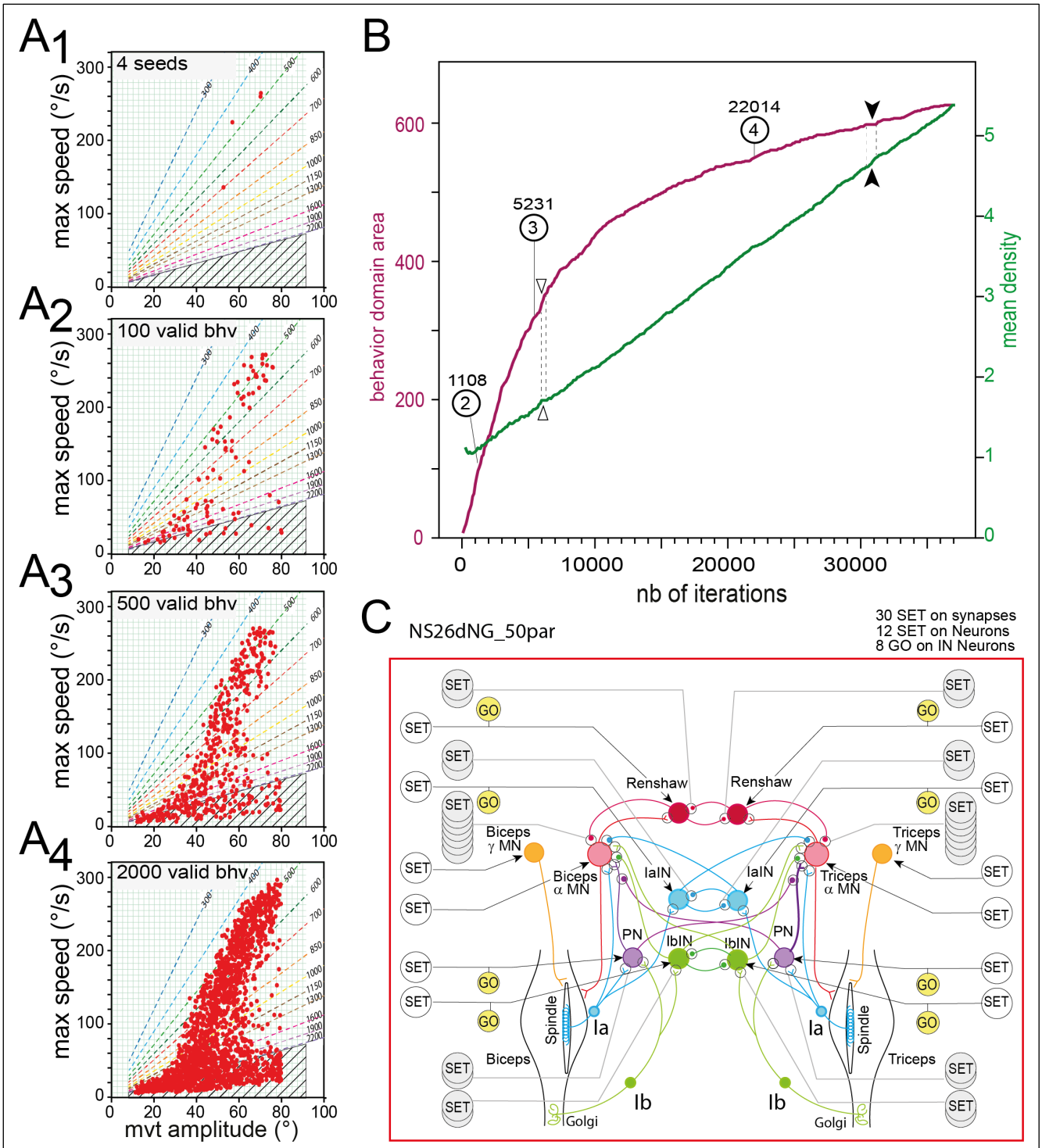


Figure 2: Temporal evolution of rGEP. **A**, Progression of the behavior domain uncovered by rGEP after starting with 4 seeds (A1), and extending to 100 (A2), 500 (A3) and 2000 (A4) valid behaviors. **B**, Time course of behavior domain area expansion (i.e., number of squares containing at least one valid behavior in the green grids presented in A1-4) presented in violet trace. The mean density of behaviors in the behavior domain area is presented in green trace. The numbers 2, 3 and 4 correspond to the behavior domains representations shown in (A2-3). **C**, Diagram of the 50-parameter model.

312 cases of a grid, defined in the behavior work space (see Methods), which contain at least one valid
313 behavior (see green grids in Figure 2A1-4).

314 While the borders of the behavior domain get more and more precise, there were still some
315 empty central zones when the process was stopped after 300 "extend" rounds and 50 "fill" rounds
316 (totalizing 36718 runs). Note that there is no direct relationship between the number of extend rounds
317 and the number of runs, because the number of runs in a given round depends on the size of the
318 behavioral domain at the time the round is launched. Note also that when the rGEP was stopped, the
319 process was almost complete as indicated by the very slow increase of behavioral domain area (violet
320 trace) during the last 15000 runs (Figure2B). The behavior domain at the end of the 300 extend and 50
321 fill rounds (36718 runs) was presented in Figure 1B. We note that some empty zones still persist after
322 the 3000 extend rounds. This domain was then completed by adding three new rGEP with 300 expands
323 and 50 fills (starting from new seeds each time) totalizing 112240 runs (see below "*Behavior domain*
324 *of the complete (50-parameters) model*").

325 We can also see that the progression of rGEP is not linear. Indeed, several transient plateaus were
326 observed (see filled arrow heads in Figure 2B) on the behavior domain area curve (violet trace). During
327 such blockage of exploration, valid results were found in the already discovered behavior domain,
328 resulting in a transient increase of the mean density (see methods) of valid behaviors (green curve in
329 Figure 2B). The situation unblocks when an original valid behavior is found that extends the behavior
330 domain, and allows new behaviors to be found from its parameter set, resulting sometimes in an abrupt
331 increase of the behavior domain area (see open arrow heads in Figure 2B) while the progression of
332 mean density slows down (see green trace, in Figure 2B). As the behavior domain extends (Figure 2A-
333 B), the parameter domain also evolves.

334 The final parameter domain is presented in Figure 3. In this figure, all valid values of the 50
335 parameters are presented in 25 graphs, each displaying a pair of parameters: 4 GO stim pairs: [(*Triceps*

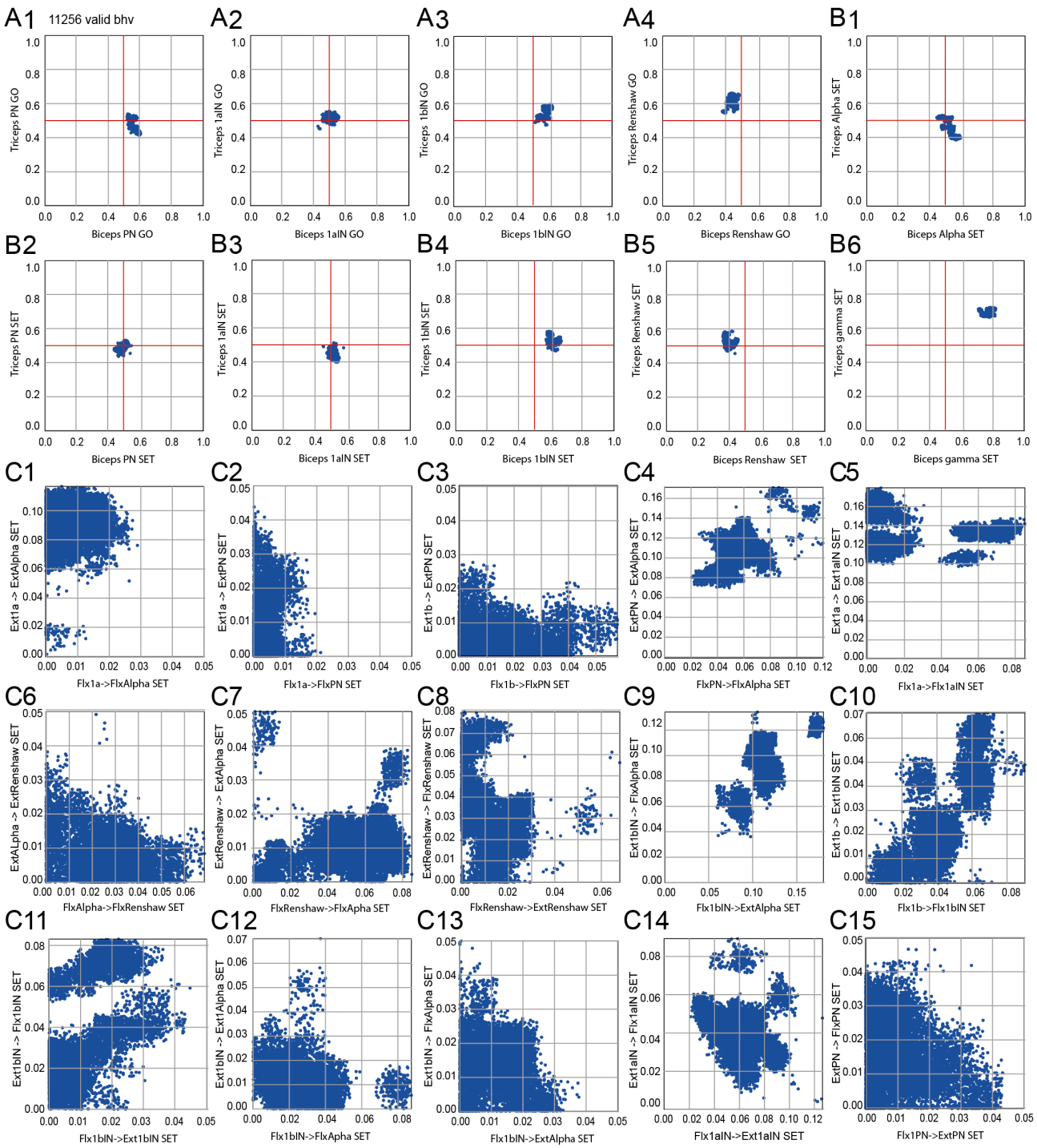


Figure 3: Parameter domains of the complete model. The 50 parameters are represented in 25 graphs displaying 2 parameters each: 4 stimulus GO graphs (A1-4), 6 stimulus SET graphs (B1-6) and 15 synapse strength graphs (C1-15) obtained after 39160 runs. Axis names: stimulus graphs represent intensity of stimulus sent to a neuron in the GO (A1-4) and SET (B1-6) commands; synapse strength graphs represent strength setting of a synapse between a presynaptic and a post synaptic neuron separated by "->" (for example Flx1a->FlxAlpha SET indicate the SET command for synaptic strength between 1a spindle afferent and alpha MN).

PN GO, Biceps PN GO), (Triceps 1aIN GO, Biceps 1aIN GO), (Triceps 1bIN GO, Biceps 1bIN GO), (Triceps Renshaw GO, Biceps Renshaw GO)] (Figure 3A1-4); 6 SET stim pairs: [(*Triceps Alpha SET, Biceps Alpha SET*), (*Triceps PN SET, Biceps PN SET*), (*Triceps 1aIN SET, Biceps 1aIN SET*), (*Triceps*

339 *1bIN SET, Biceps 1bIN SET), (Triceps Renshaw SET, Biceps Renshaw SET), Triceps gamma SET,*
340 *Biceps gamma SET*] (Figure 3B1-6); 15 SET syn pairs: [*Ext 1a-> ExtAlpha*, *Flx 1a->FlxAlpha*]
341 (Figure 3C1), [*Ext 1a->ExtPN, Flx 1a>FlxPN*] (Figure 3C2), [*Ext 1b- >PN, Flx 1b- >PN*] (Figure
342 3C3), [*Ext PN- >ExtAlpha, Flx PN- >FlxAlpha*] (Figure 3C4), [*Ext 1a- >Ext1aIN, Flx 1a- >Flx1aIN*]
343 (Figure 3C5), [*ExtAlpha- >ExtRenshaw, FlxAlpha- >FlxRenshaw*] (Figure 3C6),
344 [*ExtRenshaw- >ExtAlpha, FlxRenshaw- >FlxAlpha*] (Figure 3C7), [*ExtRenshaw- > FlxRenshaw,*
345 *FlxRenshaw- > ExtRenshaw*] (Figure 3C8), [*Ext1bIN- > FlxAlpha, Flx1bIN- > ExtAlpha*] (Figure
346 3C9), [*Ext1b- >Ext1bIN, Flx1b- > Flx1bIN*] (Figure 3C10), [*Ext1bIN- >Flx1bIN, Flx1bIN- >*
347 *Ext1bIN*] (Figure 3C11), [*Ext1bIN- >ExtAlpha, Flx1bIN- > FlxAlpha*] (Figure 3C12),
348 [*Ext1bIN- >FlxAlpha, Flx1bIN- > ExtAlpha*] (Figure 3C13), [*Ext1aIN- >Flx1aIN, Flx1aIN- >*
349 *Ext1aIN*] (Figure 3C14), [*ExtPN- >FlxPN, FlxPN- > ExtPN*] (Figure 3C15), in which *stim* = stimulus
350 sent to a neuron, and *syn* = strength setting of a synapse between a presynaptic and a post synaptic
351 neuron separated by "->" (see Figure 3 legend for more explanations).

352 ### H. Modelling coding properties of muscle spindle

353 Muscle spindles were added to each muscle. In AnimatLab (see Methods), muscle spindles have
354 the same "attachments" as the muscles. Their tension-length ratio, maximum force capacity, and
355 viscoelasticity are based on the Hill muscle model (Scovil & Ronsky, 2006). To assess the quality of
356 this neuro-musculoskeletal model, imposed movements of flexion and extension were simulated, while
357 muscle spindle activity was recorded (Figure 4A). Our first work was to adjust by hand the muscle
358 spindle parameters that controlled their coding of movements so that their response were comparable
359 to electrophysiological data from the literature (Boyd *et al.*, 1977) (Figure 4B), in the absence and in
360 the presence of activation of a Gamma MN.

361 Note that the AnimatLab platform proposes a single type of gamma command that is a mixture of
362 gamma static and gamma dynamic properties. Therefore, in all simulations we used this single type of
363 gamma MN (see Discussion F for more details).

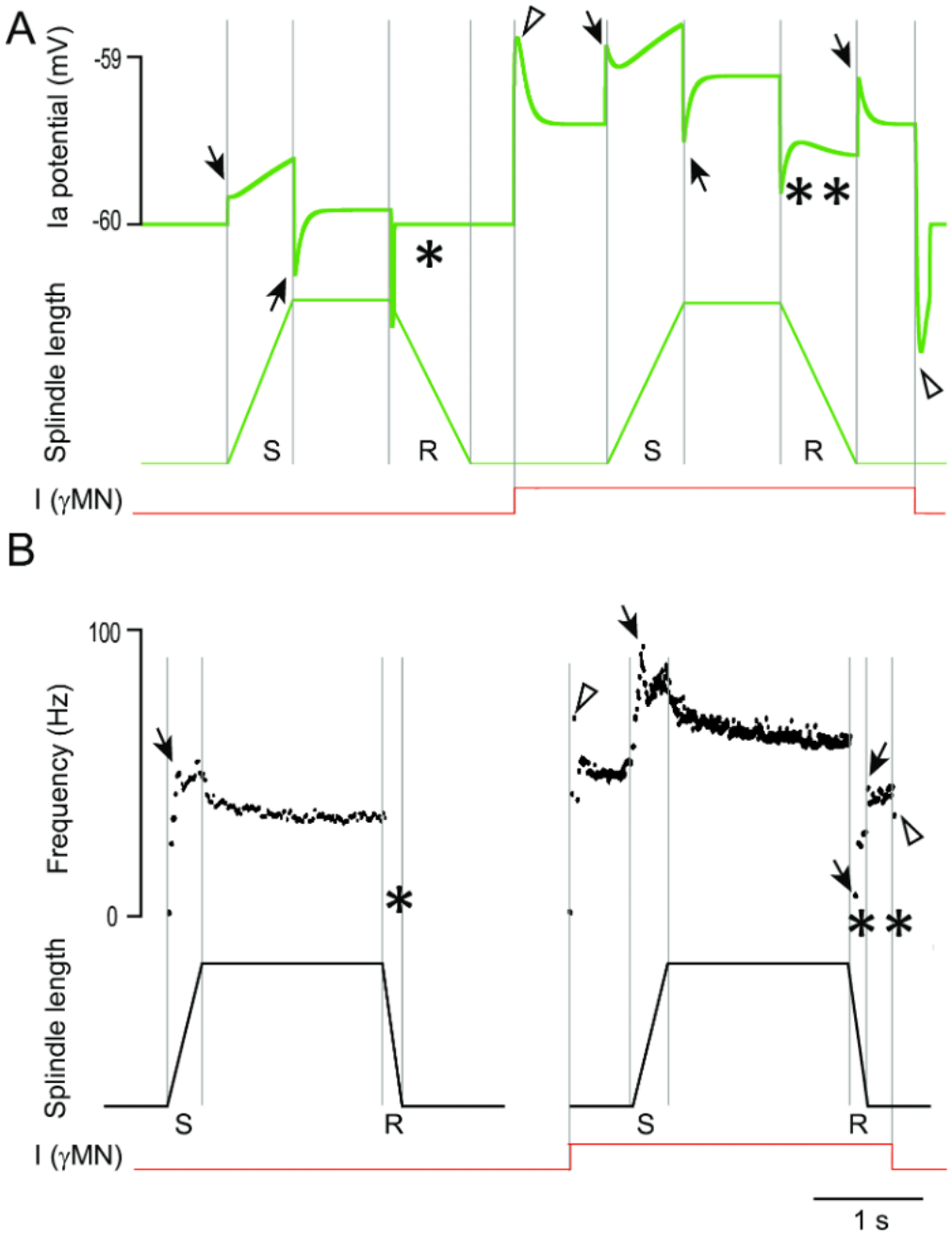


Figure 4. Muscle spindle coding properties

A. Modeled extensor muscle spindle activity during stretch (S) and release (R) ramps, without (left), and with (right) activation of the gamma motoneuron (I γMN). **B.** Electrophysiological recording of the Ia fiber during the same imposed movements as in **A** modified from (Boyd *et al.*, 1977)

Note that, in both simulation and physiological recording, the first release of the spindle does not give any response (see *) due to lack of tension of the spindle. However, the activation of the γMN allows the spindle to give an answer during the second release (see **). Note also that in **A**, the "acceleration" component produced by the serial spring velocity (in the Hill model) is responsible for the transient activation at each movement change (arrows), while the "tonic" component produced by the tension of the spindle (in the Hill model) produces an exponential decay shape. Transients are also produced at onset and offset of the γMN stimulation. These features are present in **B** except for transients at the end of stretch movements.

Results

A. Goal exploration process explores the capacity of spinal sensorimotor circuits and their step commands to produce behaviors.

In this section, we provide a brief overview of the neuromechanical models and the process (rGEP) used to explore their behavioral capabilities.

1) Behavior domain of the complete (50-parameters) model with both Ia and Ib afferents and associated network.

The behavior domain found for the 50 parameters model by the rGEP is presented in Figure 5B. It was obtained by cumulating 4 rGEP rounds, each composed of 300 "extend" rounds and 100 "fill" rounds, and totalizing 112240 runs, in which 37087 behaviors were valid. The behavior domain is defined by two parameters of the produced movements: maximal speed (ordinate) and movement amplitude (abscissa). A third characteristic of movements (movement duration) is presented on the graphs by oblique lines labelled with the represented duration (in ms). Only physiologically plausible movements have been retained (slow movements the duration of which was > 2s have been rejected – see striped area). Only movements with an amplitude in the range [10°, 110°] have been retained, to avoid too small movements and too large movements that could reach the flexion limits. With both Ia and Ib sensorimotor circuits complete, this model could perform large movement (up to 80°) at a rapid speed (>300°/s) with movement durations < 500ms. Note for each amplitude, a great variety of speeds was found.

A representative example of a large movement (amplitude 75.95°) is given in Figure 5C1. Note the profile of the movement that fitted a minimum jerk profile with small MSE value (0.096). Several other variables are also presented for this example: Force, Membrane potentials of MNs, and sensory neurons. At the initiation of the movement, the force recorded in the Biceps was large (reaching >120 N within 110 ms) and decayed to 0 in 200 ms (Figure 5C2). The force recorded in the triceps increased at the end of the rapid decrease of biceps force, and reached 75 N at t=545ms, and then decreased in the next 150 ms, at the end of the flexion movement. Note the small second peak of flexion force while movement stops.

The activity of alpha MNs present transient activities during the movement phase: the biceps MN (*FlxαMN*) presents two peaks of depolarization, the second being smaller, and an absence of activity between the two peaks, during which the triceps MN (*ExtαMN*) presents a peak of activity (Figure 5C3). Once the flexed position is reached, the activity of both alpha MNs totally vanishes. During the movement phase, the proprioceptive feedback from Ib afferents was much larger than that

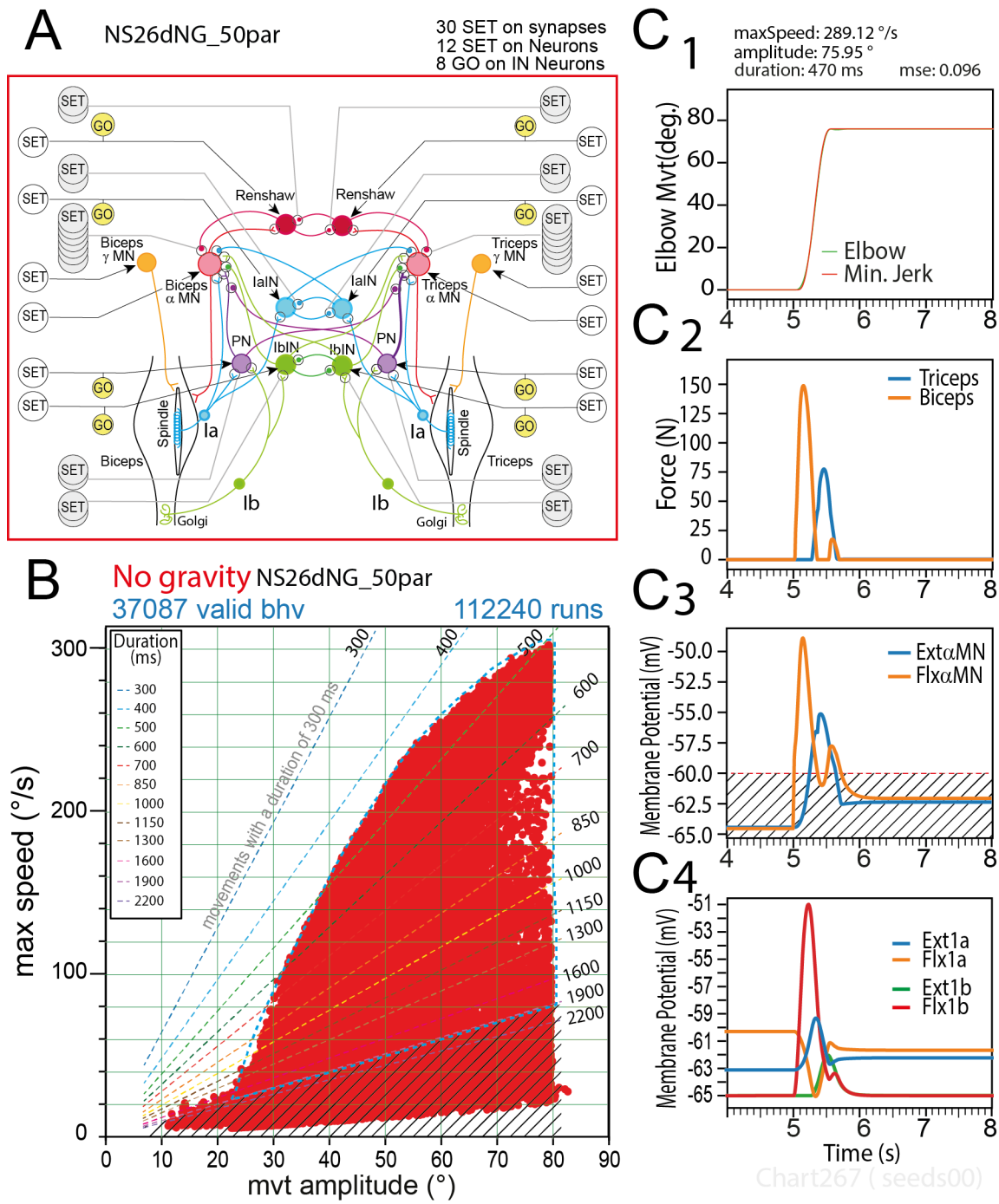


Figure 5: Movements produced by rGEP with the 50-parameters model. **A**, Sensorimotor spinal network of the 50-parameters model. **B**, Behavior domain uncovered by rGEP on this model. The perimeter of the domain is outlined with a blue dashed line. **C**, Representative example of one valid movement in the behavior domain (C1), with force time course of Biceps and Triceps muscles (C2), Flexor and Extensor MN membrane potential (see FlxαMN and ExtαMN in C3), and 1a & 1b afferent neurons potential (see Flx1a, Ext1a, Flx1b, Ext1b in C4). Striped area in C3 indicates MN under-threshold potentials, under which no muscle response was observed.

399 of the Ia afferents (Figure 5C4). The shape of Ib activations was also different from that of Ia afferents:
400 during the rapid flexion movement, the Ib activity of the flexor muscle (*Flx1b*) present a large peak
401 during the acceleration of the flexion movement, in accordance with its role in sustaining the flexion
402 movement via the Flexor PN excitatory circuit (see Figure 5A); this initial *Flx1b* peak is followed by
403 a smaller peak of the extensor 1a activity (*Ext1a*) that contributes to slow down the flexion movement
404 due to its excitatory effect on the extensor αMN (see Figure 5A); this slowing effect is prolonged by
405 the delayed increase of *Ext1b* activity that contributes to the activation of the extensor αMN via the
406 PN interneuron of the extensor circuit (see Figure 5A). Note that during the flexion movement, the
407 activity of the Flx1a presents a trough. Note also the small peak of Flx1a just before stabilization at -62
408 mV.

409 The time course of 1a and 1b afferent activities (larger Flx1b peak followed by a smaller Ext1a
410 peak, and by a delayed peak of the Ext1b activity, concomitant with the recovery of Flx1a trough)
411 were regularly observed in the 50-parameter model, with some small variations all along its behavior
412 domain. This first study will serve to compare the capacities of other models issued from this one, but
413 lacking some neural elements. To facilitate comparison, the perimeter of the behavior domain of the
414 50-parameter model (see dashed blue line in Figure 5B) will be superimposed on their behavior
415 domains.

416 *2) Behavioral domain produced in the absence of Ib feedback (with central cross connections).*

417 In order to study what proprioceptive feedback was necessary to produce valid movements, in
418 a first step, we suppressed the Ib input, the Ib INs and their synapses (Figure 6A). When rGEP was
419 run on this model, the resulting behavior domain was close to the one obtained with the complete
420 spinal network (compare Figure 5B and Figure 6B), although maximal speed for large movements was
421 slightly reduced, and all movement durations were longer than 450 ms.

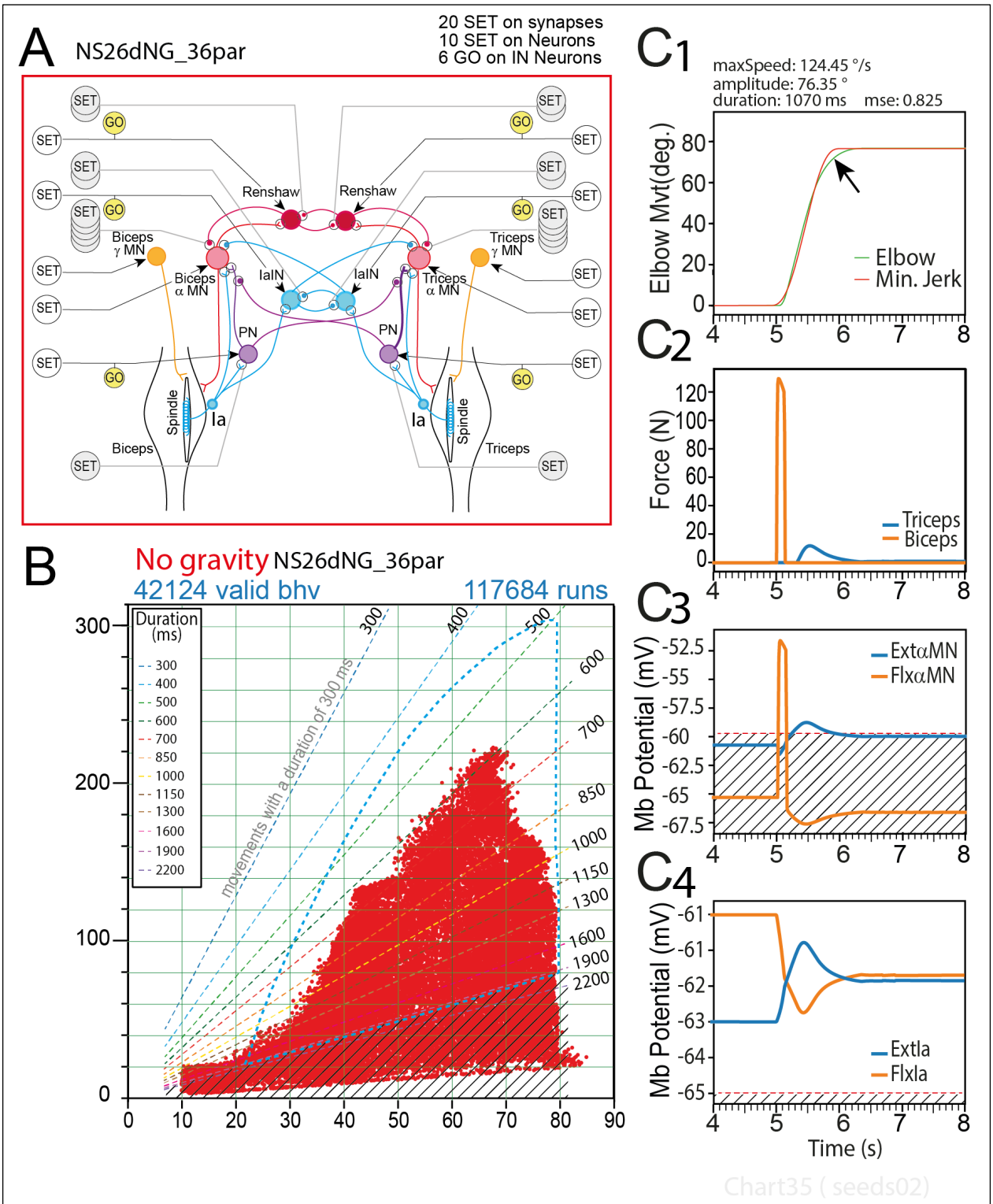


Figure 6: Movements produced by rGEP with the 36 parameters model. **A**, Sensorimotor spinal network of the 36 parameters model. **B**, Behavior domain uncovered by rGEP on this model. For comparison, the perimeter of the domain of the 50-parameter domain is represented with blue dashed line. **C**, Representative example of one valid movement in the behavior domain (C1), with force time course of Biceps and Triceps muscles (C2), Flexor and Extensor MN membrane potential (see FlxαMN and ExtαMN in C3), and 1a afferent neurons potential (see Flx1a, Ext1a in C4). Striped area in C3 indicates the potentials under which no muscle response mas observed. Striped area in C4 indicates under threshold activity.

Forces, MN activities and muscle spindles feedback (Figure 6C) were very similar to that of the complete model (Figure 5C) for comparable movements. A representative example of a large movement (amplitude 76.35°) is given in Figure 6C1, with a maximal speed of 124.45°/s. However,

425 the fitting to a minimum jerk template (MSE=0.825) was not as good as in the complete model. This
426 was likely due to the slowdown at the end of movement (see arrow in Figure 6C1) being much more
427 progressive than in the complete model. The amount of force developed by the Biceps muscle (Figure
428 6C2) was slightly smaller (#130 N) than that observed for the complete model. Strikingly, in this model
429 deprived of Ib afferent circuits, the second peak of Flx muscle force and Flx Alpha MN activity was
430 never observed (Figure 6C2-3). This second peak was also absent from the Flx1a activities (Figure
431 6C4).

432 *3) Behavioral domain produced in the absence of Ib feedback (without central cross*
433 *connections).*

434 To go further in the analysis of the involvement of Ia proprioceptive feedback in the production
435 of minimum jerk movements, we drastically reduced the sensorimotor network, by suppressing all
436 relationships between extensor and flexor parts of the network, as well as all interneurons except
437 propriospinal INs. The resulting spinal network is the 14-parameters simplified network presented in
438 Figure 7A with the same design as other models presented in Result section. Interestingly, this
439 extremely simplified spinal network, reduced to Ia proprioceptive inputs connecting their
440 homonymous MN either directly or via their propriospinal IN (PN), was capable of producing some
441 valid behaviors (Figure 7B). The resulting behavior domain was, however much smaller than the one
442 of the 36-parameters mode (no movement shorter than 1150 ms were produced and maximal speed
443 was under #100°/s). A representative behavioral example (amplitude 84.29°) is given in Figure 7C1.
444 Note that in most of the behaviors produced by this model, the fitting to a minimum jerk template was
445 not as good as in the complete model (here MSE=0.725). Moreover, the second peak in Flx force, Flx
446 alpha MN and Flx1a activities was never observed (Figure 7C2-3).

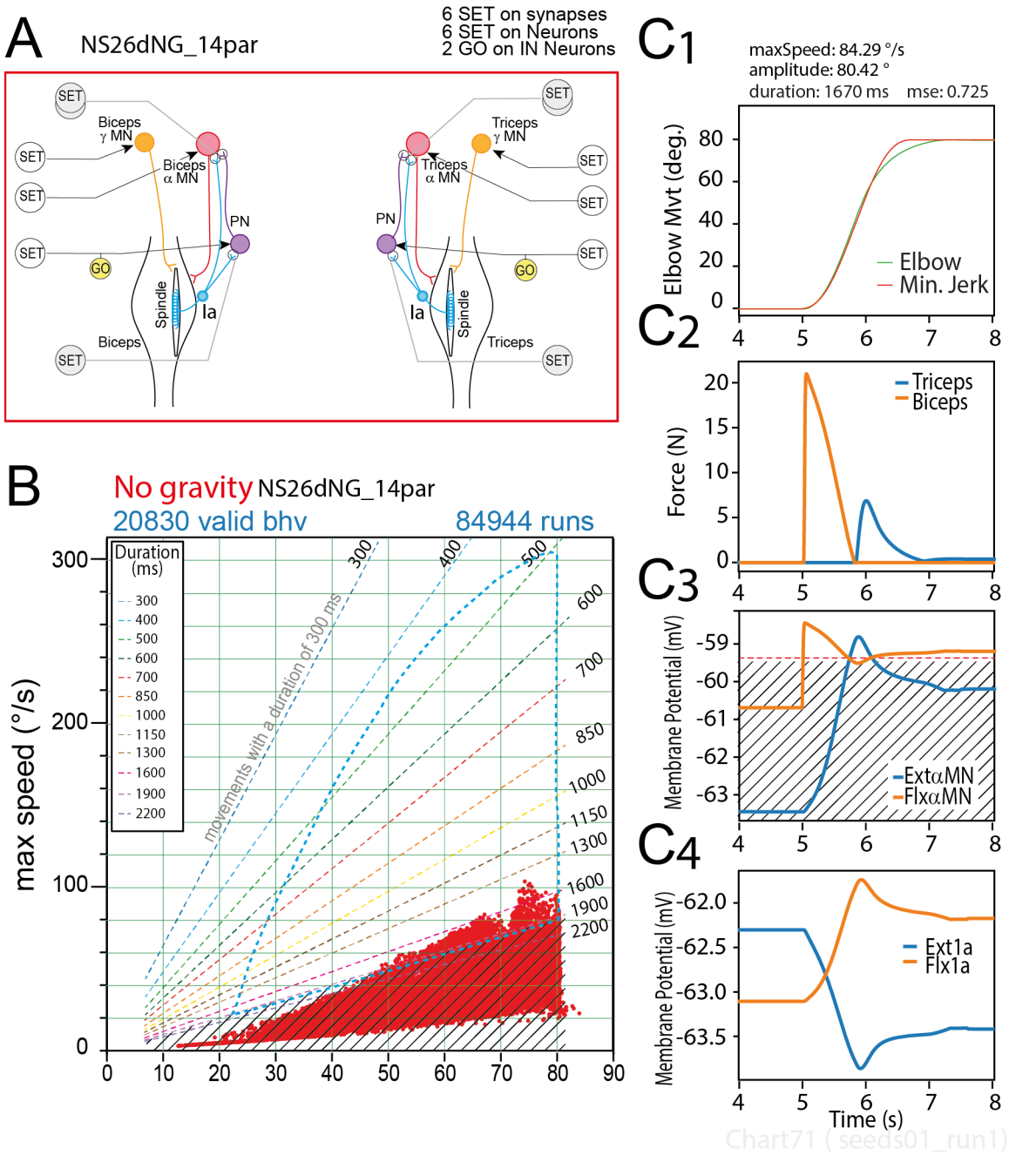


Figure 7: Movements produced by rGEP with the 14-parameters model. **A**, Sensorimotor spinal network of the 14 parameters model. **B**, Behavior domain uncovered by rGEP on this model. For comparison, the perimeter of the domain of the 50-parameter domain is represented with blue dashed line. **C**, Representative example of one valid movement in the behavior domain. Striped area in C3 indicates the potentials under which no muscle response mas observed.

447 *4) Behavioral domain produced in the absence of Ia feedback with central cross-connections.*

448 We then applied the same method to test Ib afferents involvement in the production of valid
449 movements via rGEP. Starting from the complete (50-parameters model), we suppressed Ia afferents,
450 Ia INs and their synapses (Figure 8A). This Ib network could not produce any valid behaviors (Figure
451 8B).

5) *Behavioral domain produced in the absence of any feedback but with all INs.*

Because reciprocal connections seemed important to perform valid behaviors, we tested if complete spinal network with all reciprocal connections but deprived of any proprioceptive feedback (Figure 9A) could produce valid behaviors. However, no valid behaviors were found with this model (Figure 9B).

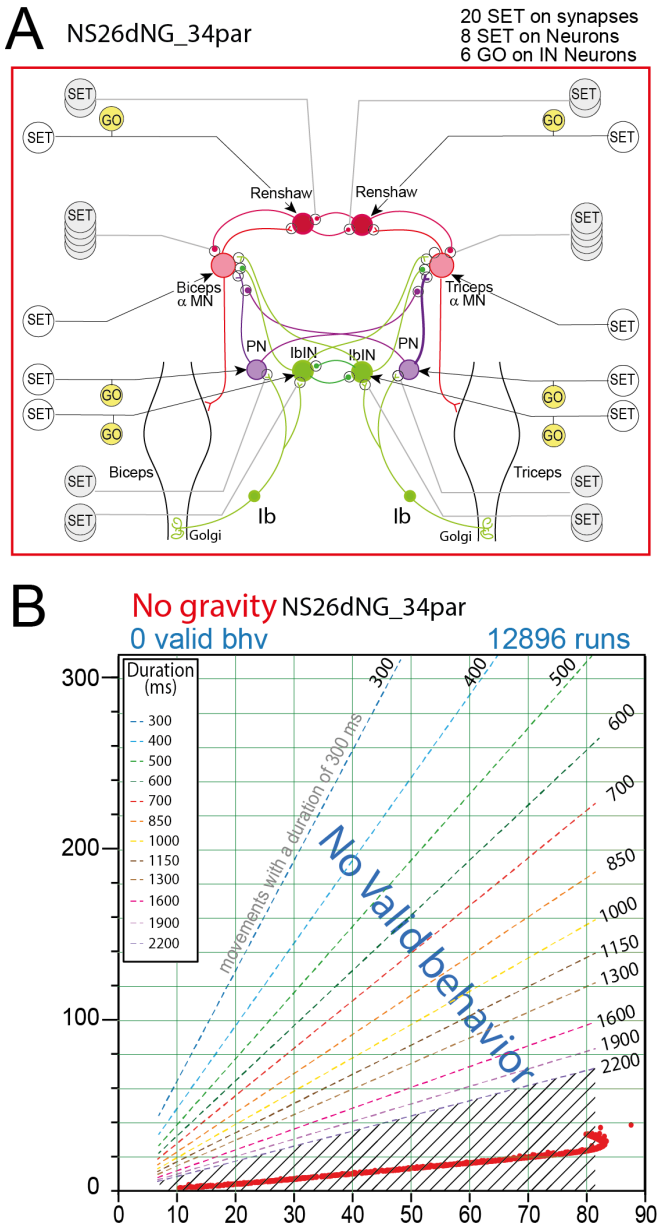


Figure 8: Movements produced by rGEP with the 34-parameters model. **A**, Sensorimotor spinal network of the 34 parameters model. **B**, Behavior domain uncovered by rGEP on this model. The very slow movements produced by this model did not meet criteria for valid movements.

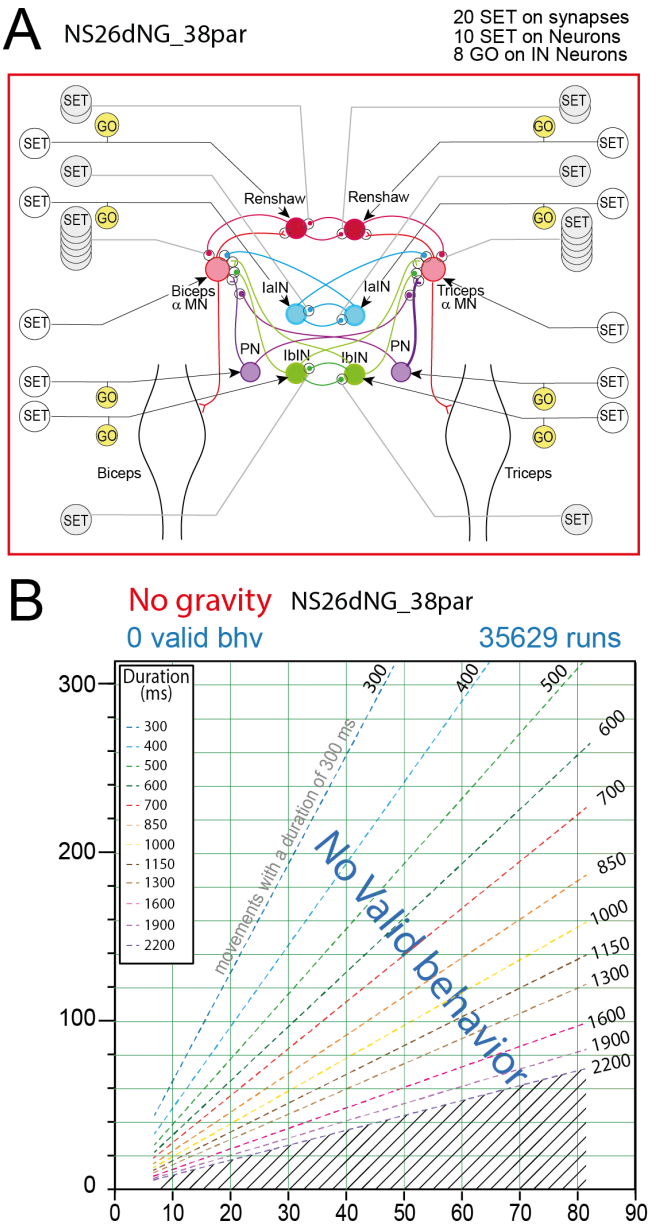


Figure 9: Absence of Valid Movements produced by rGEP with a complete model deprived of proprioceptive feedback. **A**, Sensorimotor spinal network of the 38-parameters model. **B**, Absence of any valid behavior in the behavior domain uncovered by rGEP on this model

457 ## B. Analysis of triphasic patterns in the various models.

458 By design, valid behaviors found by rGEP followed the minimum jerk rule. However, as we
459 will see, this rule does not imply conformity with a triphasic command. Triphasic command is defined
460 from the EMG recordings of two antagonistic muscles that were involved in a rapid movement around
461 a joint (Wachholder, 1928; Angel, 1974; Wadman, 1979; Brown & Cooke, 1981). Typically, the
462 agonist muscle produced two bursts of activity: one (AG1) at the onset of movement and one (AG2)
463 at end of the movement. The amplitude of the initial agonist burst (AG1) was larger than that of the
464 second (AG2) (Bullock, 1992). In between the two agonist bursts, an antagonist burst (ANT) is present
465 (Sanes & Jennings, 1984), the role of which is to slow down the movement speed so that, ideally, the
466 speed is 0 when target angle is reached. However, in rapid movements, if the activity of the antagonistic
467 muscle is too large, the movement goes back. The asymmetrical rise- and fall-times inherent to muscle
468 activation may also induce a prolonged effect of ANT. These effects are counteracted by AG2. This
469 excess of ANT activity (in amplitude and/or in duration) and the second burst of the agonist (AG2)
470 produce a small oscillation of the joint angle. Such small oscillations at the movement end were
471 accepted as long as the error (MSE) calculated between the movement and its minimum-jerk profile
472 was less than 1. Movements with too large end oscillation (see Figure 1F, right, profile (1)) were
473 rejected.

474 In order to apply the definition of the triphasic command, we have used the electrical activity of
475 muscles. However, since all models used in this study were based on non-spiking neurons (and hence
476 non-spiking muscle), we used the non-spiking muscle activity to analyze triphasic command
477 occurrences. In order to illustrate EMG profiles, for each model, we give examples of EMGs produced
478 during a large movements (#80°), during a medium amplitude movement (#60°), and during a smaller
479 movement (#45°) (see Figure 10C1-3).

480 ### *1) Triphasic commands with the complete (50-parameters) model*

481 The occurrence of triphasic commands with AG1, ANT, AG2 peaks of activity were only
482 observed with the **50-parameters model** (see the full network in Figure 10A). Triphasic patterns with
483 two Flexor peaks, separated by a silent period during which an Ext peak was present, were observed

in the whole domain of the valid behaviors produced by the 50 parameters model (Figure 10B). Three examples of triphasic pattern are given in Figure 10C1-3, for large movement (80°; Figure 10C1), medium movement (60°; Figure 10C2) and small movement (45°; Figure 10C3). It is interesting to note that for the largest movements the Ext EMG peak was narrow, and its width increased for smaller movements (compare orange trace in Figure 10C1-3). It is also interesting to note that all of these movements were very close to a minimum jerk profile (their MSE were 0.069, 0.036 and 0.012, respectively). Finally, in all of the valid movements with a triphasic pattern, the activity between the two AG peaks went back to 0.

Here we represented only the movements associated with a clear triphasic pattern. However, such movements associated with a true triphasic pattern represent only 5% of the valid movements found by rGEP. Indeed, in the great majority of movements found by rGEP, a single Flx EMG peak is

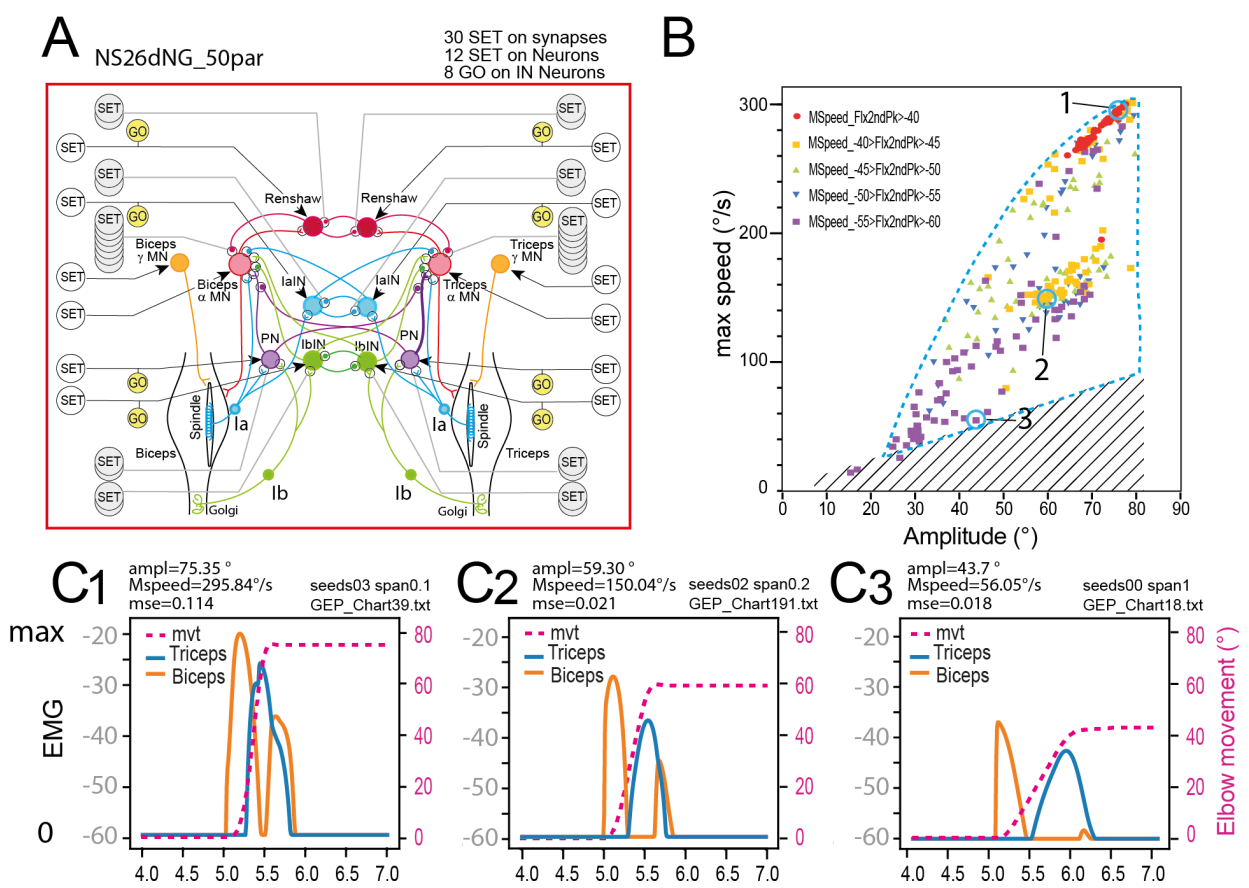


Figure 10: Presence of triphasic pattern in the behavior space with the complete 50 parameter model. **A**, Sensorimotor spinal network of the 50-parameters model. **B**, Presence triphasic pattern in the behavior domain uncovered by rGEP on this model. As can be seen triphasic pattern were found all over the domain (see blue dashed lines for the limits of the 50-parameter behavior domain). **C**, Examples of triphasic pattern for large (80°, C1), medium (60 °, C2) and small (45 °, C3) movements. These 3 examples are identified in B, with numbers 1,2 and 3.

495 followed by a single Ext EMG peak. This is generally enough to produce a perfect minimum jerk
496 movement.

497 *2) Triphasic commands with the other models*

498 In the other models, we could not find any valid movement with a triphasic pattern. Even in
499 the 36-parameters model (complete model deprived of the 1b afferent and Ib Interneurons), there were
500 valid movements very close to minimum jerk (MSE<0.05) but the second Flx EMG peak was always
501 absent. Nevertheless, we could find very rare pseudo triphasic patterns. They were rejected either
502 because the second peak was so small that it did not elicit any force in the Biceps, or because their
503 profile did not meet the minimum jerk criterium (MSE >1). Moreover, all of these pseudo triphasic
504 profiles were found in slow movements (data not shown).

505 *3) Elaboration of Triphasic commands in spinal sensorimotor networks.*

506 An example of triphasic commands with AG1, ANT, AG2 peaks of activity was used to
507 examine the temporal relationships between EMG, Alpha MN activity, PN activity and sensory 1a and
508 1b activity (Figure 11). These neurons represent all the spinal sensorimotor neurons that participate to
509 the dynamical activity leading to the triphasic command (Figure 11A, B). Only gamma MNs are not
510 represented because they only display step activation. Constitutively, FlxEMG and ExtEMG are
511 directly dependent on the activity of FlxAlpha and ExtAlpha, respectively (Figure 11C). In the same
512 way, the activity of FlxAlpha and ExtAlpha should depend on FlxPN and ExtPN, respectively (see
513 Figure 11A), but the time course of the peaks observed in Alpha MNs and their corresponding PN are
514 different (Figure 11D). The difference is small for FlxPN, the peak of which is delayed compared to
515 FlxAlpha peak, but larger for ExtPN, the peak of which is earlier than that of ExtAlpha. These
516 mismatches can be explained because Alpha MNs receive also other input from their corresponding
517 1a afferents (Figure 11A). The role of Flx1a will be more precisely analyzed in the next paragraph.
518 Note simply that Flx1a activity (yellow curve in Figure 11E) decreases during most of the FlxAlpha.
519 By contrast the role of Ext1a is more directly understandable because Ext1a depolarizes and peaks at
520 the same time as ExtAlpha (compare light-blue and green traces in Figure 11E). Interestingly, 1b

afferent present peak activities that correspond to their corresponding Alpha MN activities (Figure 11F). For example, the large early peak of Flx1b contributes to the activation of FlxAlpha via its connection to FlxPN (compare pink and red traces in Figure 11F). Finally, the activity of Ext1b also contributes to the late part of ExtAlpha activation via its connection to ExtPN (see the second bump on ExtAlpha activity that corresponds to the peak of Ext1b and to the peak of ExtPN in Figure 11F).

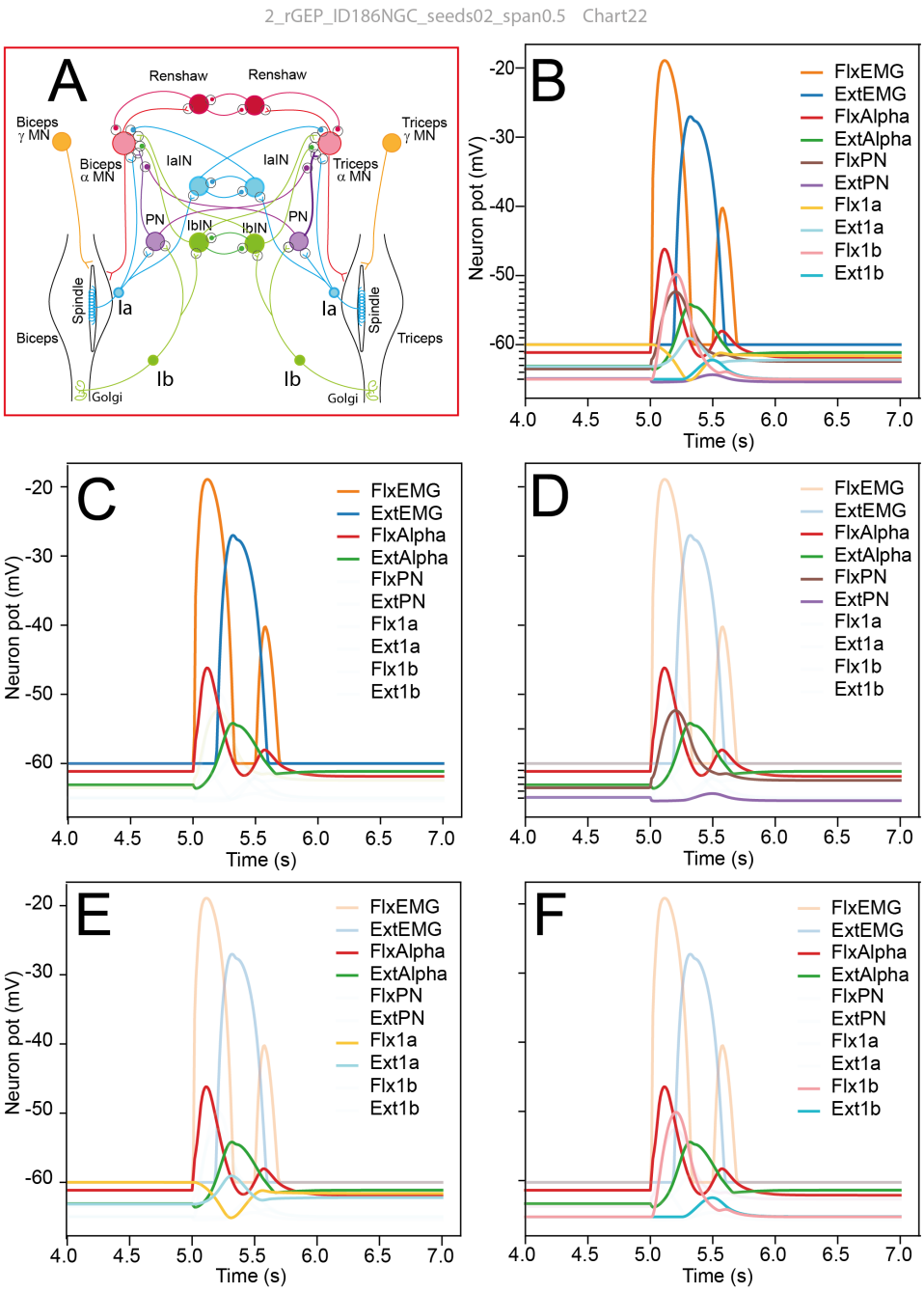


Figure 11: Activity of spinal neurons during triphasic pattern with the complete 50 parameter model. A, Sensorimotor spinal network of the 50-parameters model. **B**, Activity of Alpha MNs, PN interneurons, 1a and 1b sensory neurons during a triphasic command. **C**, Relationship between biceps EMG (FlxEMG), triceps EMG (ExtEMG), biceps alpha MN (FlxAlpha) and triceps alpha MN (ExtAlpha). **D**, Relationship between propriospinal interneurons (FlxPN and Ext PN) and their corresponding MNs (FlxAlpha and ExtAlpha). **E**, Relationship between 1a sensory afferents (Flx1a and Ext1a) and their corresponding MNs (FlxAlpha and ExtAlpha). **F**, Relationship between 1b sensory afferents (Flx1b and Ext1b) and their corresponding MNs (FlxAlpha and ExtAlpha).

To go further in the analysis of how triphasic pattern can be elaborated in spinal sensory-motor circuits, we have examined more precisely the impact of each sensory feedback, by replaying the same example of movement as presented in Figure 11, after suppression one of the four sensory feedback. The results are presented in Figure 12 for the movement launch, and Figure 13 for the movement stop.

(i) Role of Flx1a in movement launch

When Flx1a was inactivated (i.e., its membrane potential maintained artificially constant at -65 mV), we observed that the membrane potential of FlxAlpha was hyperpolarized to -68 mV in the preparing phase (see dotted red trace in Figure 12A). Similarly, in the absence of Flx1a activity, FlxPN was also hyperpolarized to -65 mV during the preparatory phase (see dotted brown trace in Figure 12A). These results indicate that during the preparatory phase, in the normal configuration, Flx1a was continuously activating FlxAlpha (and FlxPN), but FlxAlpha did not fire because of the hyperpolarizing SET command it received. When the GO command is given at t=5s, FlxAlpha is immediately freed from this hyperpolarization and will be activated by Flx1a. Interestingly, this source of activation will be transient because as the forearm get flexed, the Flx1a signal

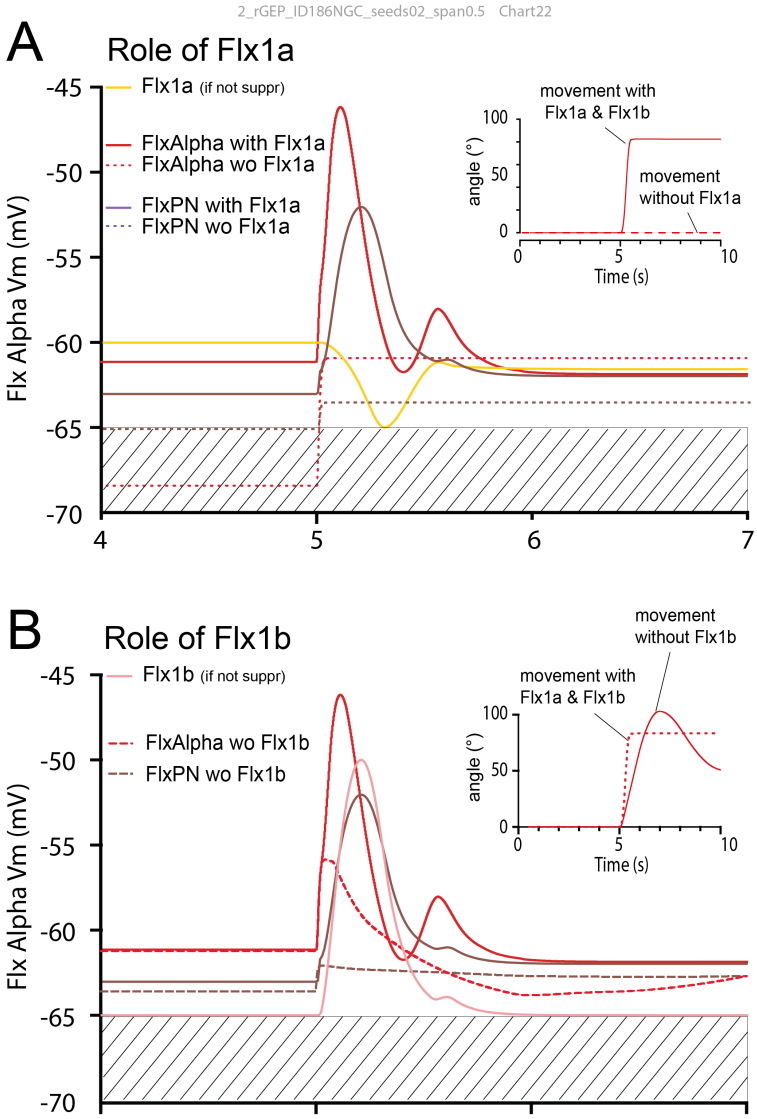


Figure 12: Role of Flx1a and Flx1b in onset of triphasic pattern. **A**, Role of Flx1a. The same movement as presented in Figure 11 was replayed with inactivation of the Flx1a afferent. The control FlxAlpha and FlxPN activities are presented in red and brown traces, respectively. The control Flx1a activity is presented in a yellow trace. After inactivation of the Flx1a afferent, no movement is produced (see dashed line in inset) **B**, Role of Flx1b. Same disposition as in **A**. The control Flx1b activity is presented in a pink trace After Flx1b afferent is inactivated, a slow movement is still produced (see dotted line in inset), but it is unable to stabilize. Flx1Alpha activity is also smaller (see red dashed line). Moreover, the FlxPN peak (full brown line) disappears when Flx1b is inactivated (dashed brow line).

552 decreases and rapidly vanishes (see red trace in Figure 12A).

553 *(ii) Role of Flx1b in movement launch*

554 When we replayed the parameters of the same movement presented in Figure 11 and Figure
555 12, after inactivating Flx1b, we still observed a movement, but it was slower (compare red trace with
556 dotted red trace in Figure 12B inset). This
557 result is in accordance with the reinforcing
558 role of 1b afferent on their muscle during the
559 onset of movement observed in physiology
560 (Donelan & Pearson, 2004).

561 *(iii) Role of Ext1a and Ext1b in movement*
562 *stop*

563 We then replayed the same
564 movement parameters after inactivating
565 Ext1a (Figure 13A). In the absence of Ext1a,
566 the preparatory phase was too much
567 disturbed (movement started at t=0s instead
568 of t=5s) to conclude about the role of Ext1a
569 in the movement stop. The role of Ext1a in
570 the activation of ExtAlpha was however
571 presented in Figure 11E and in the
572 corresponding text (see above).

573 The role of Ext1b was easier to
574 decipher, because in its absence, the second
575 FlxAlpha peak totally vanishes (see dashed
576 red curve in Figure 13B). The suppression of
577 the 2nd FlxAlpha peak is also linked with the

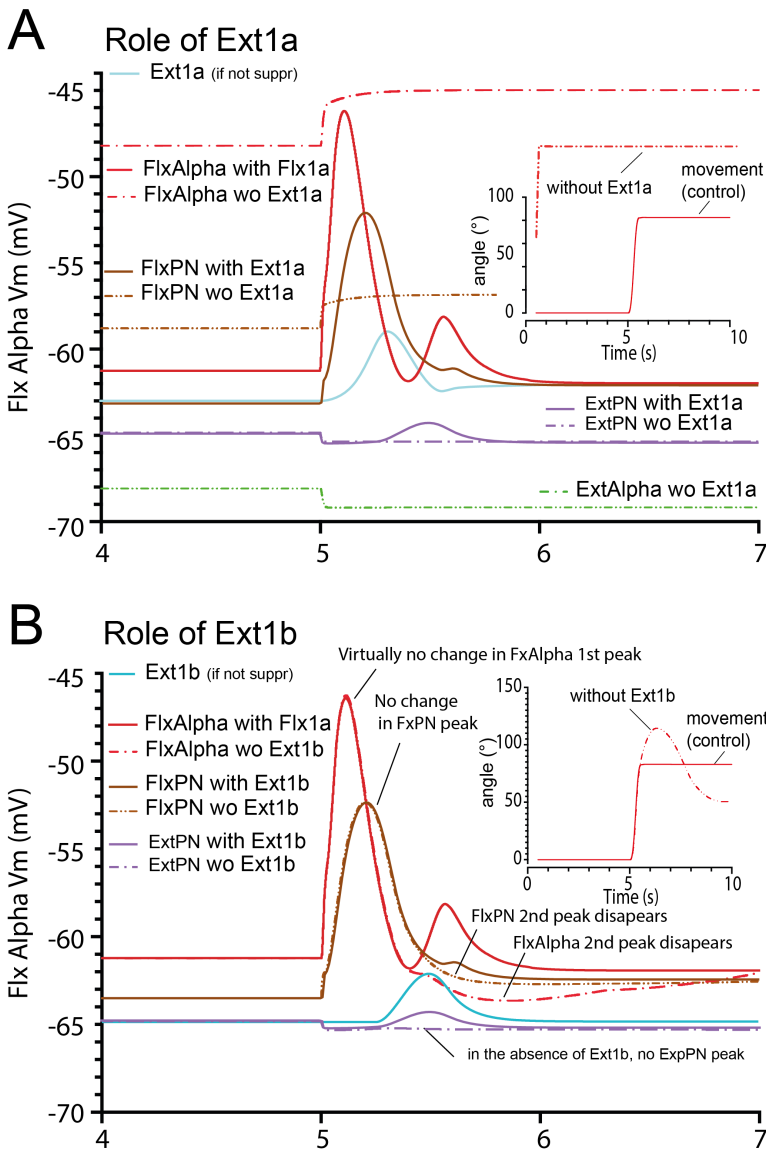


Figure 13: Role of Ext1a and Ext1b in shaping triphasic pattern. **A**, Effect of Ext1a inactivation. As in Figure 12, control FlxAlpha and FlxPN activities are presented in red and brown traces, respectively. The control Ext1a activity is presented in a lightBlue trace. After inactivation of the Ext1a afferent, movement starts at the onset of the preparatory phase (see dashed line in inset), and was therefore rejected. **B**, Role of Flx1b. Same disposition as in **A**. The control Ext1b activity is presented in a blue trace After Ext1b afferent was inactivated, a movement is still produced (see dotted line in inset), but with a huge overshoot. Moreover, the second FlxAlpha peak (full red curve), the second FlxPN peak (brown curve) and the ExtPN peak (violet curve) disappear when Flx1b is inactivated (dashed lines).

disappearance of the 2nd FlxPN peak (see dashed brown curve in Figure 13B). Moreover, the ExtPN peak also disappears (see violet curve in Figure 13B).

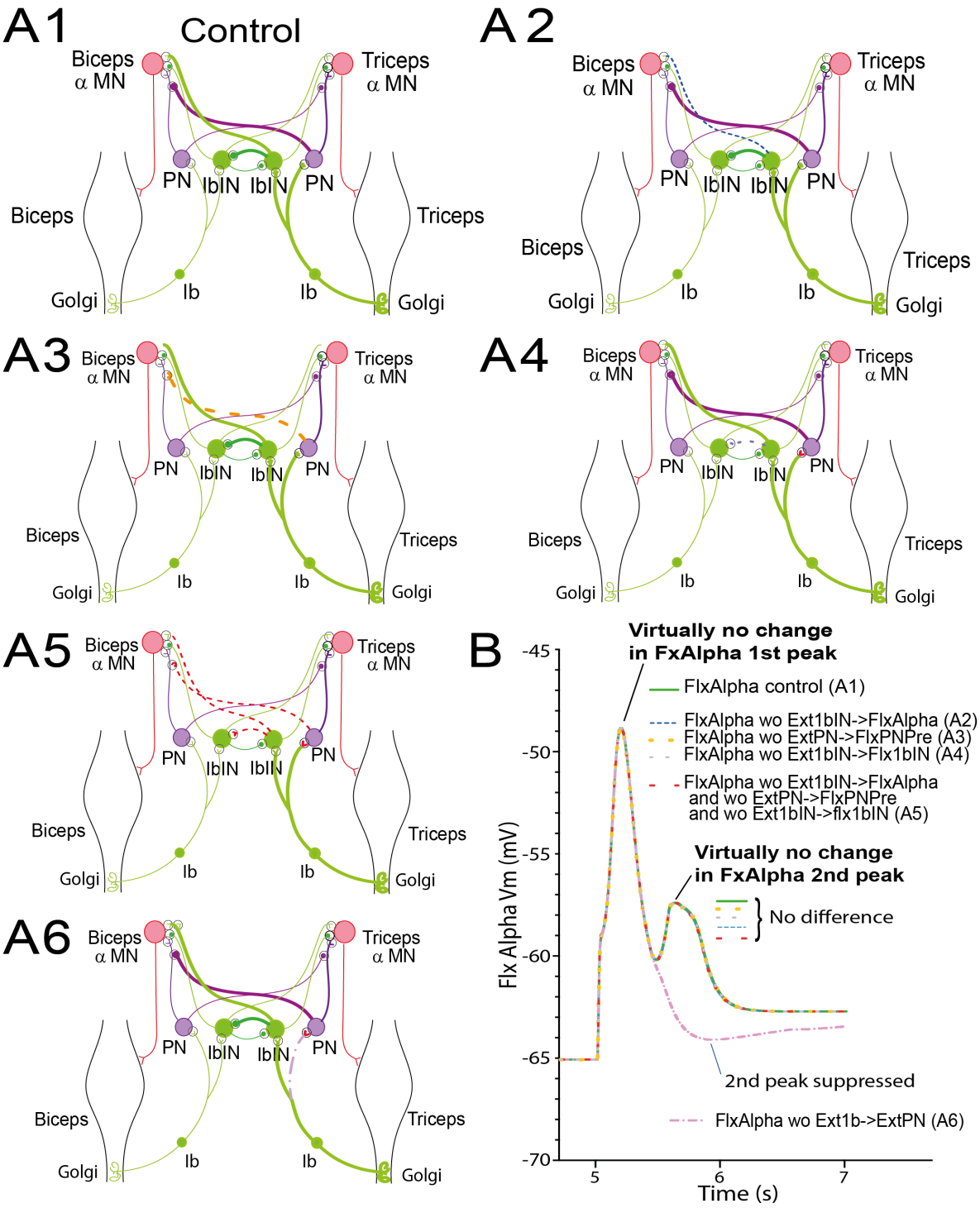


Figure 14: Role of Ext1b pathways in the production of the 2nd FlxAlpha peak of triphasic pattern.
A, Exploration of the various pathways from Ext1b to FlxAlpha (A1). Several replays of the simulation of the same movement as in Figure 12, 12 and 13) have been performed in which one of the involved synapses has been suppressed in each run (A2, A3, A4 and A5). In A6 the three synapses involved in crossing pathways have been suppressed at once. The result of the run of the model with thes diverse ablations are presented in B using the same color as the suppressed synapse. (see text for explanations).

580 *(iv) Which Ext1b pathway is involved in movement stop and 2nd FlxAlpha peak in triphasic pattern?*

581 From the result presented in Figure 13B, we conclude that Ext1b is involved in the FlxAlpha
582 2nd peak of triphasic pattern. Then, we dissected out all pathways between Ext1b and FlxAlpha (Figure
583 14). If we consider all these pathways (Figure 14A), it is easy to replay the simulation of this
584 movement, after selectively suppressing a single synapse to estimate its role in the transmission of
585 Ext1b Feedback. When the Ext1bIN->FlxAlpha synapse was suppressed (blue dashed curve in Figure
586 14A2) the FlxAlpha second peak was only slightly decreased (see blue dashed curve in Figure 14B).
587 When the ExtPN->FlxPNPre was suppressed (orange dashed line in Figure 14A3), the FlxAlpha
588 second peak was unchanged (see orange dashed curve in Figure 14B). The same absence of effect was
589 also observed when the Ext1bIN->Flx1bIN synapse was suppressed (orange dashed line in Figure
590 14A5 and in Figure 14B). A slight decrease in the second peak was also observed wen these three
591 synapses were suppressed at the same time (red dashed curves in Figure 14A6 and Figure13B).

592 By contrast, when the Ext1b->ExtPN synapse was suppressed (dashed violet curve in Figure
593 14A4), the second FlxAlpha peak totally disappeared (dashed violet curve in Figure 14B). These
594 results indicate that the second FlxAlpha peak is not due to a cross connection from Ext1b to FlxAlpha.
595 Therefore, if this 2ndFlxAlpha peak requires Ext1b feedback to be present, its effect is not central but
596 rather is a consequence of strengthening of ExtAlpha by the Ext1b->ExtPN->ExtAlpha pathways.

597 This hypothesis was then tested in Figure 15 in which we assessed the consequence of
598 suppressing the Ext1b->ExtPN synapse, onto FlxAlpha (red curves in Figure 15A), ExtAlpha (green
599 curves in Figure 15A) and Ext1b activity (blue curves in Figure 15A). As can been observed, the
600 suppression of the Ext1b->ExtPN synapse not only suppresses the FlxAlpha 2nd peak, but also alters
601 ExtAlpha activity profile and reduces the Ext1b peak itself. These later effects are produced via the
602 mechanical system, because Ext1b->ExtPN synapse is excitatory as is the ExtPN->ExtAlpha synapse.
603 Therefore, Ext1b activity contributes to reinforcing the ExtAlpha peak. Note that because of the
604 positive feedback of Ext1b onto ExtAlpha, the increase of strength will in turn increase the Ext1b
605 activity (compare Ext1b activity with and without Ext1b->ExtPN synapse in blue and dashed blue
606 curves in Figure 15A). This effect strengthens the braking action he ExtAlpha on the flexion movement

that will end more rapidly. This rapid change in movement speed is then perceived by Flx1a as can be observed in Figure 15B (see yellow curves). In the absence of Ext1b->ExtPN synapse, due to the lack of brake, speed decreases less rapidly and the Flx1a activity is reduced in its last part (after t=5.3s, see yellow dashed curve in Figure 15B). More importantly, the second Flx1a peak is also suppressed.

The rapid change in movement speed is also perceived by Flx1b as can be observed in Figure 15B (see pink curves): this effect can be observed as a second peak on the Flx1b activity (plain pink curve in Figure 15B), that disappears when Ext1b->ExtPN synapse is suppressed (dashed pink curve in Figure 15B). The presence of these two sensory 2nd peaks of Flx1a and Flx1b are responsible for the FlxAlpha 2nd peak via direct connection and via FlxPN, respectively (see FlxPN 2ndPeak brown trace in Figure 13B).

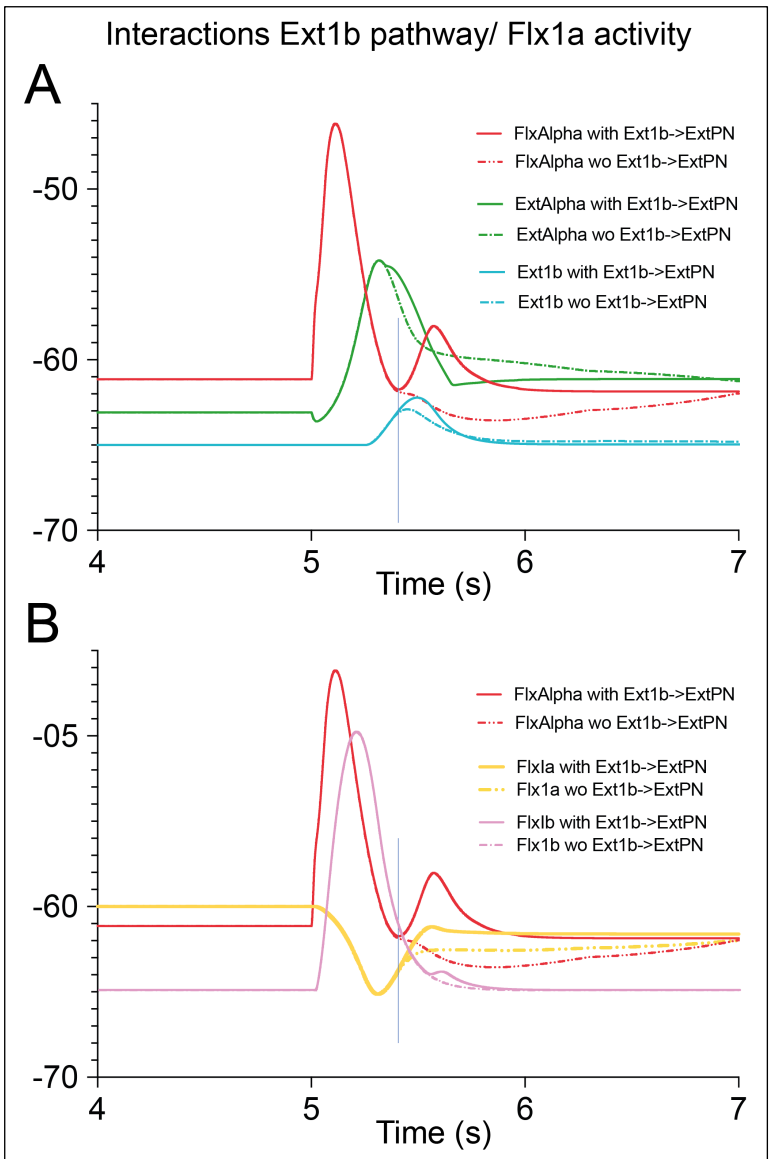


Figure 15: Interaction of Ext1b pathways and Flx1a activity. **A**, Contribution of Ext1b->ExtPN synapse on ExtAlpha activity and Ext1b feedback. **B**, Contribution of Ext1b->ExtPN synapse on Flx1a and Flx1b feedback. The FlxAlpha activity is presented in both graphs (A and B) to indicate the timing of the second peak.

(v) Confirmation of triphasic elaboration process in spinal circuits: role of 1a and 1b feedbacks?

In order to confirm the determinant role of Ext1b feedback in enhancing both ExtPN peak and brake action on flexion movement without role of crossing pathways, we have built a simplified version containing 1a and 1b

feedbacks on PN and MNs, with no crossing pathways, no 1aINs and 1bINs, and no Renshaw INs. This simplified 16 parameters model is presented in Figure 16A.

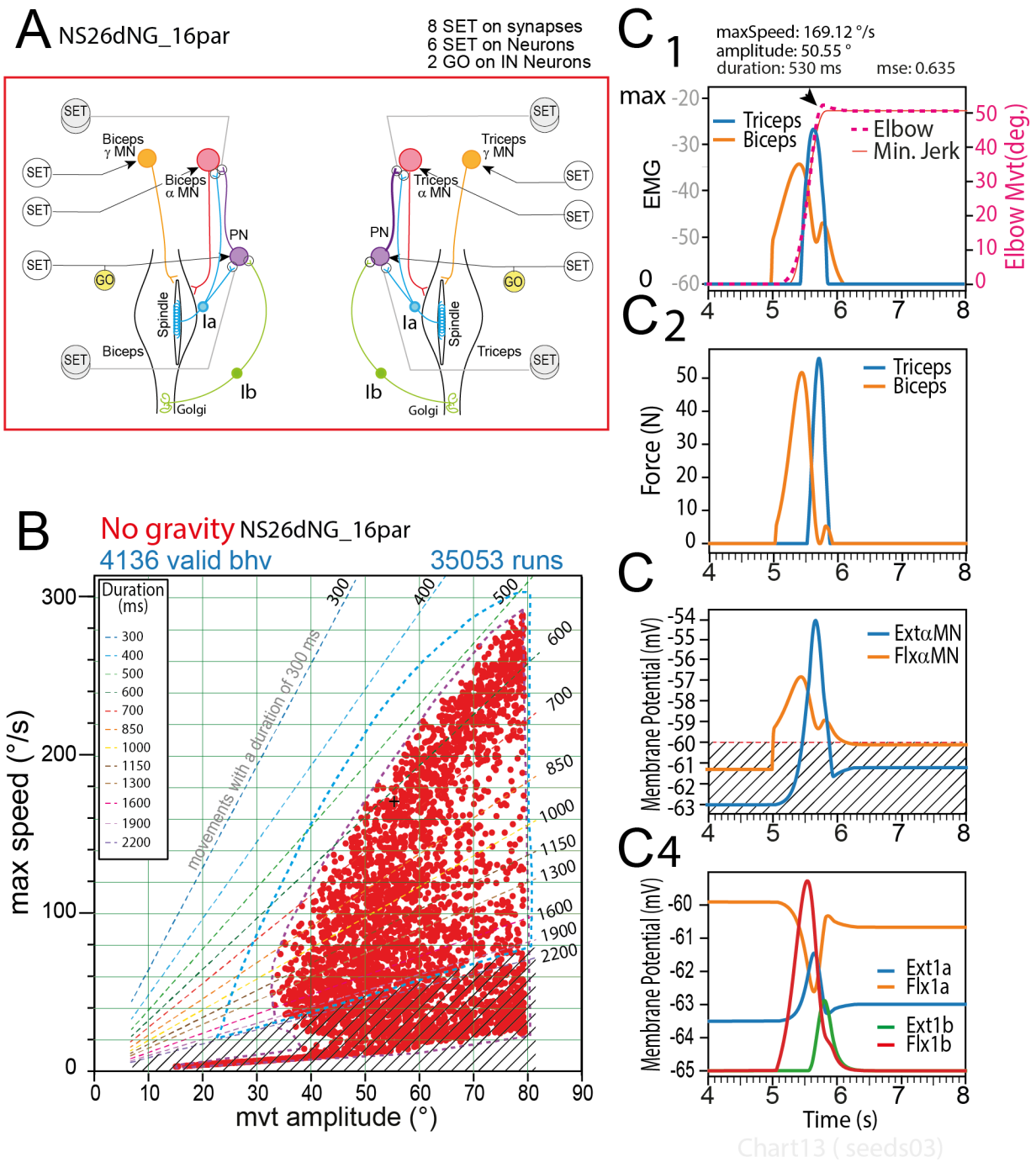


Figure 16: Minimal model with 1a and 1b feedbacks and no cross connections.
A, Presentation of the 16-parameters model. **B**, Behavior space obtained with the 16-parameters model. Note that this behavior domain is almost as large as the one of the complete 50-parameters model (see blue dashed line). **C**, representative example of a valid movement (scale presented on the right side); EMGs from Biceps and Triceps (orange and blue traces, respectively) display a triphasic command; **C1**, Time course of the flexion movement with an amplitude of 50.5° and a maximal speed of 169.12 °/s; **C2**, Time course of muscle forces for the biceps (orange trace) and the triceps (blue trace). Note the two peaks of Biceps force; **C3**, Flexor and extensor MNs activities during the movement (the stripped domain represents subthreshold activity); **C4**, Muscle spindle and Golgi Tendon Organ feedbacks.

As can be seen, the behavior domain of this simplified 16-parameters model is almost as large as that of the complete 50-parameters model (Figure 16B). A representative valid behavior displaying a triphasic command is presented in Figure 16C. Note that on this example, movement presents a small

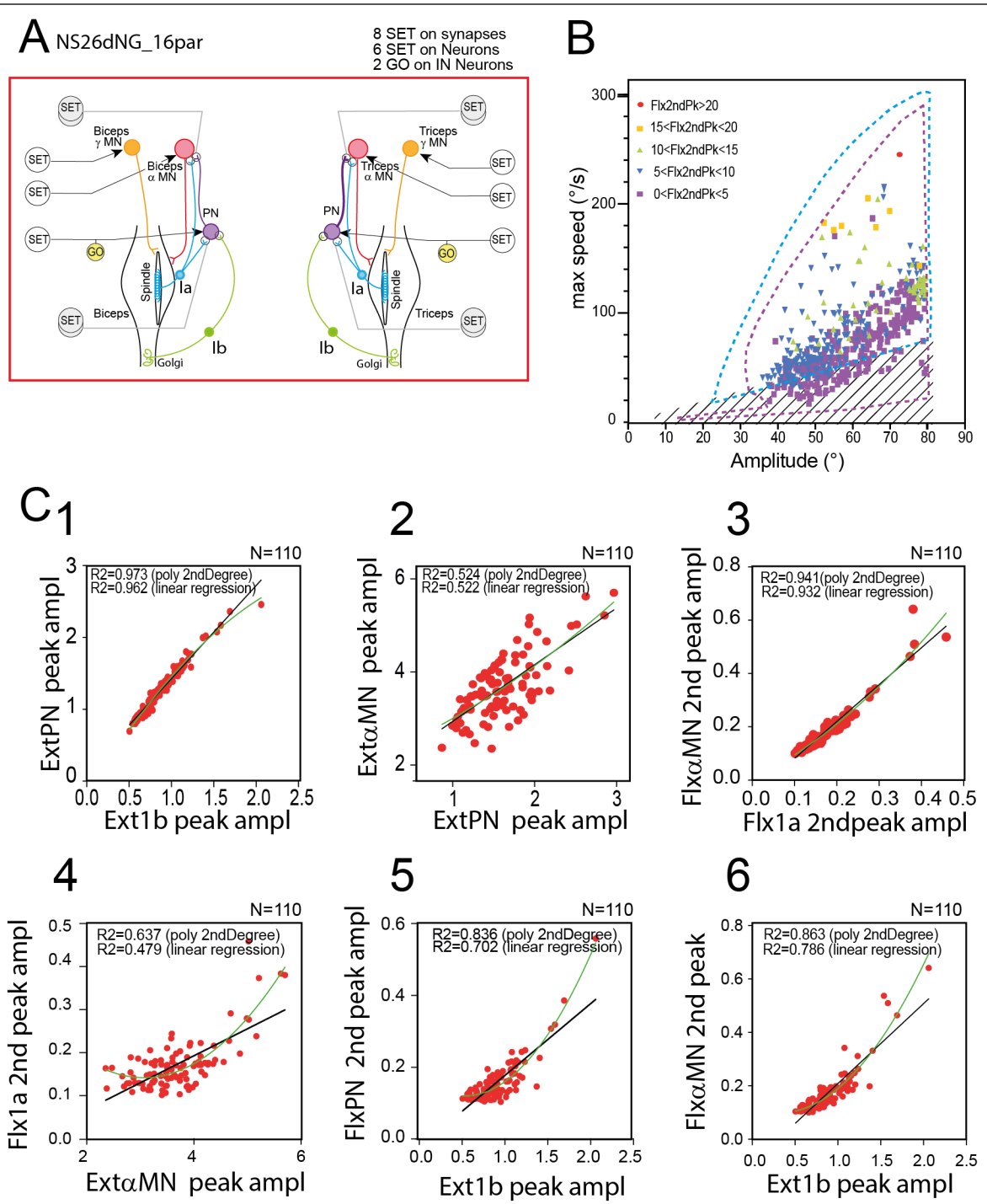


Figure 17: Analysis of triphasic patterns in the 16-parameters model.
A, 16-parameters model circuits **B**, Repartition of triphasic pattern in the behavior space obtained with the 16-parameters model. Each occurrence is presented with a symbol indicating the amplitude of the 2nd FlxMN peak: ranges are [0, 5] (violet squares), [5, 10] (blue triangles), [10, 15] (green triangles), [15, 20] (yellow squares) and >20 mV (red circle). **C**, Correlations between the amplitudes of Ext1b peak and ExtPN Peak (**C1**), ExtPN peak and ExtAlpha MN peak (**C2**), Flx1a 2nd peak and FlxAlpha 2nd peak (**C3**). These correlations indicate the strength of the functional synaptic links between the involved neurons. Correlations also exist between neuron activities involving a mechanical link (via proprioceptive feedback (**C4-C5**). For example, the amplitude of the ExtAlpha MN peak (indicating the brake power) is correlated to the amplitude of the Flx1a 2nd peak (**C4**). Similarly, the amplitude of the Ex1b peak is correlated both to the amplitude of the FlxPN 2nd peak (**C5**) and FlxAlpha 2nd Peak (**C6**).

636 overshoot (arrowhead in Figure 16C1). Note also that in this rapid movement the two flexor EMG
637 peaks partially overlap, as was observed in physiological reports for rapid movements (see Figure1C-
638 D in Brown and Cooke, 1981). Valid behaviors presenting a typical triphasic command, were found
639 in most of the behavior domain of the 16-parameters model (Figure 17B).
640
641 The behavior domain presented in Figure 16B was issued from a single seed. This choice was
642 made to obtain the largest behavior domain based on the same strategy (see discussion). The presence
643 of a single strategy in the whole behavior domain greatly facilitated the unravelling of the links
644 between the various actors presented in Figures 13, 14, 15. These links were established by correlation
645 analysis made on peak amplitudes of the various actors involved in the elaboration of triphasic
646 commands (Figure 17C). All the links were thereby validated by significant correlation coefficients.
647 As expected, Ext1b peak is linked to ExtPN peak (Figure 17C1, R^2=0.962, p<0.00001), which is itself
648 linked to the ExtAlpha peak (Figure 17C2, R^2=0.522, p<0.00001). Similarly, the FlxAlpha 2nd peak is
649 linked to the Flx1a 2nd Peak (Figure 17C3, R^2=0.932, p<0.00001). All these links are supported by
650 synaptic connections, but other correlations are the consequence of a mechanical link, via
651 proprioceptive feedback (Figure 17 C4, C5 and C6): the amplitude of ExtAlpha peak is correlated to
652 the amplitude of Flx1a 2nd peak (Figure 17C4, R^2 = 0.479, p<0.00001); the amplitude of Ext1b peak
653 is correlated both to the amplitudes of FlxPN 2nd peak (Figure 17C5, R^2 = 0.702, p<0.00001) and
654 FlxAlpha 2nd peak (Figure 17C6, R^2 = 0.786, p<0.00001). In addition, from this last graph, it is clear
655 that when Ext1b peak amplitude reaches 0, FlxApha 2nd peak also vanishes, thereby confirming the
656 result of Ext1b "numeric ablation" (Figure 13B, 13B and Figure 15).
657

658 # Discussion

659 In the present modelling work, we explored the capacities of spinal sensorimotor networks to
660 produce triphasic EMG pattern when activated by simple SET and GO step commands. We observed
661 spontaneously occurring triphasic pattern EMGs produced by spinal sensorimotor circuits (50-
662 parameters model) during the goal exploration process used to establish the behavioral capabilities of
663 the system. Critically, the timing of such triphasic patterns were not predetermined in the simple step
664 functions controlling their state. Particularly, the GO command, initiated the movement (t=5s) but did
665 not give any indication for the observed triphasic EMG patterns. It is important to note that such
666 triphasic pattern were observed in the whole range of the behavior domain (Figure 10B). **Therefore,**
667 **triphasic pattern production for any movement seems to be an intrinsic property of sensorimotor**
668 **spinal network in interaction with musculoskeletal system via MN activities and proprioceptive**
669 **feedback**. Since the 50-parameters model involved both Ia spindle and GTO afferent feedback, we
670 have used a dissecting out approach to evaluate the role of Ia and Ib afferents and their respective
671 sensorimotor circuits in the elaboration of the triphasic commands.

672 ## A. Necessity of spindle and GTO feedbacks in triphasic pattern.

673 Removing Ia sensory feedback totally prevented the occurrence of any valid behavior (Figure
674 8B), indicating that the inherent capacity of spinal sensorimotor circuits to elaborate valid behaviors
675 when interacting with a musculoskeletal system, requires spindle afferent signals.

676 By contrast, removing Ib sensory feedback and Ib sensorimotor circuits did not prevent spinal
677 sensorimotor circuits to produce minimum jerk movements (Figure 6B). However, in the absence of
678 GTO feedback, the extend of the valid behavior's domain was moderately reduced (compare Figure
679 6B and Figure 5B). More strikingly, the absence of GTO feedback totally prevented the occurrence of
680 triphasic commands, and indicates that although Ia feedback and Ia sensorimotor spinal circuits can
681 handle the elaboration of minimum jerk movements, Ib afferent signals and Ib sensorimotor circuits
682 are necessary (in addition to 1a sensory-motor circuits) to produce triphasic commands in our model.
683 Note, however, that spinal reciprocal connections in sensorimotor circuits, seem to be important for
684 fast valid movements to be produced (compare Figure 5B with Figure 16B). It seems that although

685 valid movements with triphasic command in the same range of amplitudes can be produced in the
686 absence of reciprocal connections (Figure 16B), for each amplitude, the maximal speed is significantly
687 reduced. Yet, it is important to stress that reciprocal connections via spinal interneurons (Renshaw, Ia,
688 Ib, PN) are not sufficient on their own to produce valid movements, i.e., in the absence of
689 proprioceptive feedback (Figure 9B). Therefore, this study demonstrates that spinal sensorimotor
690 networks fed by spindle afferent signals and GTO afferent signals, have the capacity to handle the
691 production of triphasic command pattern in the whole range of movements of the behavior space.

692 It is then highly likely that spindle and GTO afferents participate in the elaboration of this
693 triphasic command in animals and human. But if this conclusion is true, how to explain the finding
694 that triphasic patterns were still produced in transient deafferentation experiments (Sanes & Jennings,
695 1984)? In fact, this experiment does not totally invalidate the **participation of sensorimotor circuits**
696 in the shaping of the triphasic command, because the ischemic block affected the afferent activity of
697 the muscles below the elbow, leaving intact both the efferent copy of their motoneurons via Renshaw's
698 interneurons and all proprioceptive feedbacks from all proximal muscles. The widespread extent of
699 these feedback circuits (Cavallari *et al.*, 1992)**,** and the strong modulation of proximal muscles to
700 stabilize posture during distal phasic movements, were not fully appreciated until later (Massion,
701 1992).

702 B. Role of spindle feedback in triphasic pattern

703 We have shown that 1a feedback is essential for the spinal sensorimotor circuits to handle valid
704 movements. Indeed, it is interesting to address the question of how simple step commands can elicit
705 dynamic responses in motor activities, i.e. transient activation of flexor at two times ((i) for launching
706 the movement, (ii) just before stabilization at the end of the movement), and transient activation in the
707 extensor to stop the movement.

708 Let's start with the first Flex transient activation. When we look at the result of suppressing
709 Flx1a feedback (Figure 12A), we had noted that without Flx1a feedback the movement did not start,
710 because the strategy uncovered by GEP consisted in a strong enhancement of Flx1a sensitivity to
711 maximize its discharge during the preparatory phase. To avoid this sensory activity to activate

712 FlxAlpha MN, the SET command sent a hyperpolarizing order to FlxAlpha MN. When the GO
713 command was sent to FlxAlpha MN, this hyperpolarising action was suppressed and the Flx1a sensory
714 feedback could strongly activate FlxAlpha MN. But, because movement was launched, the biceps
715 muscle started to shorten and Flx1a signal decreased (see yellow curve in Figure 12A). This is how
716 the step command elicited a transient activation of FlxAlpha MN, via sensory coding of muscle length
717 (i.e. via the movement itself).

718 The transient activation of ExtAlpha MN (Figure 11E, green curve) is also directly related to
719 Ext1a feedback (Figure 11E, blue curve). But its transient nature is related to the **dynamic response**
720 **of Ext1a stretch** (i.e., to the speed of the stretch – see Figure 4A). Note that a possible improvement
721 of our model of spinal sensorimotor circuit would be to introduce both dynamic and static gamma
722 commands. This would improve handling of 1a feedback transients. This was not possible in
723 AnimatLab but future work should address this question.

724 ### C. Role of GTO feedback in triphasic pattern

725 It seems that 1b feedback is necessary for occurrence of triphasic command, but not sufficient
726 because it cannot produce triphasic command without spindle feedback. So, how 1b feedback can help
727 1a afferent feedback for triphasic pattern production in spinal sensorimotor circuits? If we look at the
728 activity of FlxPN (Figure 11D) and Flx1b (Figure 11F) in the 50 parameters model, we note that they
729 display similar activity along time course. This indicates that Flx1b signals play a great role in assisting
730 the Flx activity at movement onset. When Ib feedback is suppressed, maximal speed is reduced (see
731 inset in Figure 12B). More importantly, the transient FlxAlpha peak totally disappears (compare red
732 dashed line and red curve in Figure 12B). This means that Flx1b and Flx1a activities are both necessary
733 to launch the movement. Then while the muscle shortens, Flx1a activity decreases rapidly, and because
734 the force produced by the Biceps muscle decreases due to (i) muscle shortening and (ii) to FlxAlpha
735 MN actvity decrease (consecutive to Flx1a decrease), the activity of Flx1b becomes transient. This
736 involvement in force enhancement was also described in physiological conditions in cat (Donelan &
737 Pearson, 2004). It seems that 1b contribution consists mainly in this enhancement in the flexor phase
738 of the movement (i.e., movement initiation), but 1b feedback seems also involved in reinforcing

739 ExtAlpha MN activity when the extensor is stretched during the flexion movement (compare ExtAlpha
740 activity with and without Ext1b feedback in Figure 13B) and certainly contributes to shortening the
741 brake phase. Therefore, in the absence of 1b feedback, the brake is less powerful and this phase is too
742 slow for a minimum jerk profile (compare end of movements in the complete model – Figure 5C1,
743 with the model derived of 1b feedback – black arrow in Figure 6C1). Ext1bpeak induces a rapid
744 slowdown at the end of the movement, which is responsible for the second peak of Flx1a activity (due
745 to **velocity and acceleration sensitivity of 1a afferents** - see yellow trace in Figure 11E and Figure
746 12A). Consequently, this Flx1a 2nd peak, in turn activates the Flx alpha MN (red trace in Figure 12A).
747 In addition, Flx1b also sense this rapid brake and produces a second peak (see pink trace in Figure
748 12B) that is transmitted to FlxPN (see brown trace in Figure 12 A, B). Finally both Flx1a 2nd peak and
749 FlxPN 2nd peak converge onto FlxAlpha MN (see yellow, brown and red trace in Figure 12A).

750 ### D. Role of descending commands and spinal networks in triphasic pattern.

751 The fact that Ia spindle and Ib feedbacks and their spinal sensorimotor circuits are capable of
752 handling triphasic commands all over the behavior domain, does not exclude, however, that in
753 physiological conditions, descending command also participate in the elaboration of triphasic
754 commands. Recordings from M1 motor cortex indicate that unit and global activities can be correlated
755 either with static and dynamic forces and torques around single joints (Cabel *et al.*, 2001), or with
756 muscle synergies (Holdefer & Miller, 2002). However, the question of the origin of triphasic activities
757 in M1 is still open because correlation is not causation, and triphasic activities in M1 could simply
758 result from efference copy from spinal interneurons and afferent feedback from proprioceptors.
759 Whatever the origin of M1 triphasic activity, it is likely that such precise descending command interact
760 with spinal sensorimotor networks. Future research in this domain will have to unravel how complex
761 descending commands interact with spinal sensorimotor process that can handle by itself triphasic
762 commands.

763 In the present report, we used simple step commands to control spinal sensorimotor circuits.
764 Similarly, some authors have argued that because EMGs are organized in triphasic pattern with silent
765 periods, this indicated that descending commands did not operate in a smooth and continuous manner

766 **(Leib *et al.*, 2020)**. Indeed, the simple model (bang-bang optimal control model) proposed by these
767 authors is more structured in time than the simple step functions we used here. In the bang-bang
768 control, the control signal abruptly switches between a maximum and a minimum value at certain
769 sparse points in time. The control function is made of three constant segments (with two transitions).
770 The bang-bang model explains both abrupt changes in muscles activity (triphasic command) and the
771 smooth reaching trajectory. Interestingly, rGEP was capable of uncovering a set of valid behaviors
772 (minimum jerk movements) relying incidentally on triphasic commands (with abrupt changes in
773 activity), although this triphasic command was not explicitly searched for nor included in the process
774 with simple step commands. Therefore, it is possible that spinal network properties (and the associated
775 biomechanical effector system) possess the capacities of handling such complex dynamic commands
776 just by giving the correct level of stimulation of target spinal neurons and precise synaptic gain settings.
777 In a more general perspective, these spinal sensorimotor circuits could be considered as the
778 physiological support of movement primitives proposed by Degallier and Ijspeert (2010) that can be
779 handled by simple step commands from the brain. In this case, the motor primitives not only encompass
780 movement' start dynamics but also movement' stop dynamics with a unique step command.

781 ### E. Triphasic pattern and minimum jerk.

782 Simulation studies have suggested that minimum-jerk models (Flash & Hogan, 1985) and
783 minimum torque-change models (Uno *et al.*, 1989) were likely not capable to predict muscle actions
784 (Osu *et al.*, 2004) and thus triphasic pattern. Here, we used a different approach that involved a
785 musculoskeletal model associated to spinal sensorimotor networks, and a search method (rGEP) to
786 revisit the capability of spinal sensorimotor circuits to produce triphasic pattern of muscle activation.

787 Indeed, rGEP uncovered a great diversity of parameters associated with a given valid
788 movement. Among these parameters, some will result in triphasic commands. This solution was found
789 all over the behavior domain of the model. But other parameter sets do not fit the definition of triphasic
790 command. For example, although the silence of the agonist between the two agonist peaks (AG1 and
791 AG2) was found in most of the 50-parameters model valid movements (Figure 10B), there were also
792 solutions with no 2nd Flexor peak all over the behavior domain of this complete model.

793 The fact that we imposed a minimum jerk profile for movements had consequences on the type
794 of triphasic command retained by rGEP. Indeed, when we consider fast movements that were retained
795 to define the triphasic pattern (Hallett *et al.*, 1975), a small oscillation at the end of the movement is
796 often present. The origin of this final oscillation comes from ANT peak that was slightly too strong,
797 resulting in a final oscillation that required the second agonist peak (AG2) to be damped (Berardelli *et*
798 *al.*, 1996). The minimum jerk profile virtually imposes the absence of such oscillation at the end of the
799 movement, and thereby reduces dramatically the amplitude of AG2.

800 ## F. Limits and perspectives of the study.

801 **No separate Static and Dynamic gamma MNs**. All models of the present study used a mixed
802 type of gamma MN (one for the biceps and one for the triceps). Ideally, we should have used two types
803 of gamma MNs for each muscle: a static gamma MN and a dynamic gamma MN, as was the case in
804 the model of spinal-like regulator (Raphael *et al.*, 2010), but these distinct gamma control does not
805 exist at present in AnimatLab. However, the activation of the gamma MN of our model increased
806 spindle sensitivity to both position and velocity, and this single type of gamma MN was enough to
807 produce the desired behavior of the model under rGEP. Certainly, the possibility for the step
808 commands to act differently on spindle length coding and speed/acceleration coding feedbacks should
809 benefit to the possibilities of control (see "Role of spindle feedback in triphasic pattern" above), and
810 likely extend the behavior domain of the model. This point will need future work to estimate the
811 changes brought by the two gamma MN types.

812 **Absence of gravity**. We presented here the results of movements conducted in the absence of
813 gravity. The presence of gravity during elbow flexion exerts an increasing torque that opposes the
814 action of the biceps and could potentially obscure the role of the triceps in slowing down the
815 movement, particularly when the movements are not sufficiently fast. Note that in many experiments
816 the effects of gravity are eliminated by performing movements on a frictionless surface, in the
817 horizontal plane. The present results also apply to this experimental condition.

818 It will be interesting in future investigations to add gravity to see how triphasic pattern evolve,
819 depending on movement trajectory, and how spinal network can handle it. Such studies should show

820 smaller ANT burst when gravity is present and tend to slow down the movement as the forearm
821 becomes more horizontal during flexion movements. Inversely, if movements start from a flexed
822 position (say 110° in our model), and stops at 20°, ANT burst should be dramatically increased and
823 thus AG2 burst should increase accordingly.

824 **Simple flexion movement**. In the present study, we used a single joint (elbow) and movement
825 consisted in elbow flexion. This was made because minimum jerk is more adapted to this situation
826 (Nakano *et al.*, 1999)**.** It would be interesting to generalize the results of the present study by using
827 more complex models including two joints, and mono- and bi-articular muscles, and considering
828 endpoint reaching movements as was used in Tsianos et al. (2014).

829 **GEP and exploration of redundancy**. The systematic investigation of the behavioral
830 constraints of reduced spinal circuitry highlights the functionality of a system that might otherwise be
831 considered excessively redundant. For organisms navigating real-world environments, it may be more
832 important to get the largest possible range of performances from their musculoskeletal systems rather
833 than concentrating on the narrow range of tasks typically examined in laboratory settings. In this line
834 of thought, a potential development of the present work would be to evaluate the adaptative role of
835 redundancy by introducing unexpected perturbations or constraints, and see how redundancy helps the
836 system to find appropriate solutions. Notably, such research would highlight the importance of spinal
837 networks in facilitating rapid adaptation under the brain's supervision as is the case in reservoir
838 computing (Tanaka *et al.*, 2019).

839

840 Finally, the present modelling study offers an opportunity to test the physiological roles of Ia
841 and Ib proprioceptive feedback in voluntary movements. Since the model pointed out the specific roles
842 of Ia and Ib at specific timings, it should be possible to test if the involved circuits are responsible for
843 the elaboration of the triphasic commands observed during rapid arm flexion, for example by
844 delivering perturbations at these specific time windows and see if muscle responses are in accordance
845 with similar perturbations delivered in the model.

Author Contributions

Conceptualization: Daniel Cattaert, Aymar de Rugy, Pierre-Yves Oudeyer.

Funding acquisition:

Methodology: Daniel Cattaert, Bryce Chung, Florent Paclet, Pierre-Yves Oudeyer, Aymar de Rugy.

Project administration: Aymar de Rugy.

Investigation: Daniel Cattaert, Matthieu Guemann.

Software: Daniel Cattaert, Bryce Chung, Florent Paclet.

Supervision: Aymar de Rugy.

Validation: Daniel Cattaert, Aymar de Rugy, Pierre-Yves Oudeyer.

Writing – original draft: Daniel Cattaert.

Writing – review & editing: Daniel Cattaert, Matthieu Guemann, Florent Paclet, Bryce Chung, Pierre-Yves Oudeyer, Aymar de Rugy

Acknowledgments

This work was supported by The Centre National de la Recherche Scientifique (CNRS) and the University of Bordeaux.

863 # References

864 Angel, R.G. (1974). Electromyography during voluntary movement: the two-burst pattern.
865 *Electroencephalogr Clin Neurophysiol* **36,** 493-498.

866 Baranes, A. & Oudeyer, P.Y. (2013). Active learning of inverse models with intrinsically motivated goal
867 xploration in robots. *Robotics and Autonomous Systems* **61**.

868 Barnett, C.H. & Harding, D. (1955). The activity of antagonist muscles during voluntary movement.
869 *Ann Phys Med* **2,** 290-293.

870 Berardelli, A., Hallett, M., Rothwell, J.C., Agostino, R., Manfredi, M., Thompson, P.D. & Marsden, C.D.
871 (1996). Single-joint rapid arm movements in normal subjects and in patients with motor
872 disorders. *Brain* **119 (Pt 2),** 661-674.

873 Boyd, I.A., Gladden, M.H., McWilliam, P.N. & Ward, J. (1977). Control of dynamic and static nuclear
874 bag fibres and nuclear chain fibres by gamma and beta axons in isolated cat muscle spindels.
875 *J Physiol* **265,** 133-162.

876 Brown, S.H. & Cooke, J.D. (1981). Amplitude- and instruction-dependent modulation of movement-
877 related electromyogram activity in humans. *J Physiol* **316,** 97-107.

878 Bullock, D.G., S. (1992). Emergence of tri-phasic muscle activation from the nonlinear interactions of
879 central and spinal neural network circuits. *Human Movement Science* **11,** 157-167.

880 Cabel, D.W., Cisek, P. & Scott, S.H. (2001). Neural activity in primary motor cortex related to
881 mechanical loads applied to the shoulder and elbow during a postural task. *J Neurophysiol* **86,**
882 2102-2108.

883 Cavallari, P., Katz, R. & Penicaud, A. (1992). Pattern of projections of group I afferents from elbow
884 muscles to motoneurones supplying wrist muscles in man. *Exp Brain Res* **91,** 311-319.

885 Cofer, D., Cymbalyuk, G., Reid, J., Zhu, Y., Heitler, W.J. & Edwards, D.H. (2010). AnimatLab: a 3D
886 graphics environment for neuromechanical simulations. *J Neurosci Methods* **187,** 280-288.

Colas, C., Sigaud, O. & Oudeyer, P.-Y. (2018). GEP-PG: Decoupling Exploration and Exploita- tion in Deep Reinforcement Learning Algorithms. In *International Conference on Machine Learning (ICML)*. Stockholm, Sweden.

Degallier, S. & Ijspeert, A. (2010). Modeling discrete and rhythmic movements through motor primitives: a review. *Biol Cybern* **103,** 319-338.

Donelan, J.M. & Pearson, K.G. (2004). Contribution of force feedback to ankle extensor activity in decerebrate walking cats. *JNeurophysiol* **92,** 2093-2104.

Eliasmith, C. & Anderson, C.H. (2003). *Neural engineering : Computation, representation, and dynamics in neurobiological systems.* . MIT press.

Etcheverry, M., Moulin-Frier, C. & Oudeyer, P.Y. (2020). Hierarchically Organized Latent Modules for Exploratory Search in Morphogenetic Systems. *Advances in Neural Information Processing Systems*.

Eysenbach, B., Gupta, A., Ibarz, J. & Levine, S. (2019). Diversity is all you need: Learning skills without a reward function. In *7th Internarional Conference on Learning Representations*. New Orleans, LA, USA, .

Flash, T. & Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *J Neurosci* **5,** 1688-1703.

Forestier, S., Mollard, Y. & Oudeyer, P.-Y. (2017). Intrinsically Motivated Goal Exploration Processes with Automatic Curriculum Learning. arXiv preprint

Friedman, J. & Flash, T. (2009). Trajectory of the index finger during grasping. *Exp Brain Res* **196,** 497-509.

Ghez, C. & Martin, J.H. (1982). The control of rapid limb movement in the cat. III. Agonist - antagonist coupling. *Exp Brain Res* **45,** 115-125.

Gottlieb, J. & Oudeyer, P.Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nat Rev Neurosci* **19,** 758-770.

Grizou, J., Points, L.J., Sharma, A. & Cronin, L. (2020). A curious formulation robot enables the discovery of a novel protocell behavior. . *Science advances* **6**.

Hallett, M. & Marsden, C.D. (1979). Ballistic flexion movements of the human thumb. *J Physiol* **294,** 33-50.

Hallett, M., Shahani, B.T. & Young, R.R. (1975). EMG analysis of stereotyped voluntary movements in man. *J Neurol Neurosurg Psychiatry* **38,** 1154-1162.

Hansen, N., Muller, S.D. & Koumoutsakos, P. (2003). Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evol Comput* **11,** 1-18.

Holdefer, R.N. & Miller, L.E. (2002). Primary motor cortical neurons encode functional muscle synergies. *Exp Brain Res* **146,** 233-243.

Lehman, J. & Stanley, K.O. (2011). Abandoning objectives: evolution through the search for novelty alone. *Evol Comput* **19,** 189-223.

Leib, R., Russo, M., d'Avella, A. & Nisky, I. (2020). A bang-bang control model predicts the triphasic muscles activity during hand reaching. *J Neurophysiol* **124,** 295-304.

Massion, J. (1992). Movement, posture and equilibrium: interaction and coordination. *Prog Neurobiol* **38,** 35-56.

Mouret, J.B. & Doncieux, S. (2012). Encouraging behavioral diversity in evolutionary robotics: an empirical study. *Evol Comput* **20,** 91-133.

Nakano, E., Imamizu, H., Osu, R., Uno, Y., Gomi, H., Yoshioka, T. & Kawato, M. (1999). Quantitative examinations of internal representations for arm trajectory planning: minimum commanded torque change model. *J Neurophysiol* **81,** 2140-2155.

Osu, R., Kamimura, N., Iwasaki, H., Nakano, E., Harris, C.M., Wada, Y. & Kawato, M. (2004). Optimal impedance control for task achievement in the presence of signal-dependent noise. *J Neurophysiol* **92,** 1199-1215.

Oudeyer, P.Y. & Kaplan, F. (2007). What is Intrinsic Motivation? A Typology of Computational Approaches. *Front Neurorobot* **1,** 6.

Pathak, D., Agrawal, P., Efros, A.A. & Darrell, T. (2017). Curiosity-driven exploration by self-supervised prediction. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops* pp. 16-17. IEEE, Honolulu, HI, USA.

Péré, A., Forestier, S., Sigaud, O. & Oudeyer, P.Y. (2018). Unsupervised learning of goal spaces for intrinsically motivated goal exploration. In *ICLR 2018*.

Prut, Y. & Fetz, E.E. (1999). Primate spinal interneurons show pre-movement instructed delay activity. *Nature* **401,** 590-594.

Raphael, G., Tsianos, G.A. & Loeb, G.E. (2010). Spinal-like regulator facilitates control of a two-degree-of-freedom wrist. *J Neurosci* **30,** 9431-9444.

Reinke, C., Etcheverry, M. & Oudeyer, P.Y. (2019). Intrinsically Motivated Discovery of Diverse Patterns in Self-Organizing Systems. In *ICLR 2019*.

Rothwell, J.C., Traub, M.M., Day, B.L., Obeso, J.A., Thomas, P.K. & Marsden, C.D. (1982). Manual motor performance in a deafferented man. *Brain* **105 (Pt 3),** 515-542.

Sanes, J.N. & Jennings, V.A. (1984). Centrally programmed patterns of muscle activity in voluntary motor behavior of humans. *Exp Brain Res* **54,** 23-32.

Scovil, C.Y. & Ronsky, J.L. (2006). Sensitivity of a Hill-based muscle model to perturbations in model parameters. *J Biomech* **39,** 2055-2063.

Sergio, L.E., Hamel-Paquet, C. & Kalaska, J.F. (2005). Motor cortex neural correlates of output kinematics and kinetics during isometric-force and arm-reaching tasks. *J Neurophysiol* **94,** 2353-2378.

Tanaka, G., Yamane, T., Héroux, J.B., Nakane, R., Kanazawa, N., Takeda, S., Numata, H., Nakano, D. & Hirose, A. (2019). Recent advances in physical reservoir computing: A review. *Neural Networks* **115,** 100–123.

963 Terzuolo, C.A., Soechting, J.F. & Ranish, N.A. (1974). Studies on the control of some simple motor
964 tasks. V. Changes in motor output following dorsal root section in squirrel monkey. *Brain Res*
965 **70,** 521-526.
966 Trainin, E., Meir, R. & Karniel, A. (2007). Explaining patterns of neural activity in the primary motor
967 cortex using spinal cord and limb biomechanics models. *J Neurophysiol* **97,** 3736-3750.
968 Tsianos, G.A., Goodner, J. & Loeb, G.E. (2014). Useful properties of spinal circuits for learning and
969 performing planar reaches. *J Neural Eng* **11,** 056006.
970 Uno, Y., Kawato, M. & Suzuki, R. (1989). Formation and control of optimal trajectory in human
971 multijoint arm movement. Minimum torque-change model. *Biol Cybern* **61,** 89-101.
972 Wachholder, K. (1928). Willkürliche Haltung und Bewegung. *Ergeb Physiol* **26,** 568-575.
973 Wadman, W.J.D.v.d.G., J..J.; Geuze, R.H.; Mol, C.R. (1979). Control of fast goal-directed arm
974 movements. *Journal of Human Movement Studies* **5,** 3-17.
975