

Dynamic Hebbian learning in Rowat-Selverston CPG for synchronization in human robot handshaking

Lancelot Caron^{1,2}, Melanie Jouaiti¹, Patrick Hénaff^{1,2}

¹*LORIA UMR 7503, University of Lorraine-INRIA-CNRS, F-54506 Nancy, France*

²*Ecole des Mines de Nancy, Information and Systems department, Nancy, France*

Correspondence*:

Patrick Hénaff

LORIA UMR 7503, University of Lorraine-INRIA-CNRS, F-54506 Nancy, France,
patrick.henaff@loria.fr

2 ABSTRACT

3

4 It is important for a robot to be able to interact properly with humans, ie. in a socially acceptable
5 way. Rhythmic movements are indeed essential in physical interactions because they give way to
6 a synchronisation of the interaction, hence an involvement in the communication, notably with the
7 handshaking gesture. One possible way to achieve this consists in using robot controllers which
8 are intrinsically rhythmic, such as CPGs. The hypothesis of CPGs for rhythmic movements of
9 upper limbs is put forward. This article offers to validate a Hebbian plasticity mechanism model
10 for the Rowat-Selveston CPG model.

11 In the first part, we recall the Righetti approach, the rhythmic neuron model of Rowat-Selveston
12 and the structure of Rybak's CPG. In the second part, we rewrite the Rowat-Selveston model to
13 a Van der Pol form in order to apply the rule introduced by Righetti. In the third part, we apply
14 this model to the command of a robotic arm physically interacting with a human and a simulated
15 robotic arm, by using Righetti's structure. We show that the controller learns to synchronise
16 with the human-imposed rhythm in a given frequency range matching the usual frequencies
17 of handshaking. In the fourth part, we implement this learning model in the complete CPG
18 reproducing Rybak's model and we validate on the same robot and in simulation by showing that
19 the latter learns to synchronise with the human-imposed rhythm.

20 **Keywords:** PhRI Hebbian Learning CPG plasticity

1 INTRODUCTION

21 When two humans interact physically or socially, they adapt to each other and learn from each interaction. If
22 the interaction is rhythmic (handshaking, walking, waving...), synchronisation will emerge. Synchronisation
23 creates links between people and is thus an essential component of interaction. So, when a robot interacts
24 with a human, it needs to learn and adapt to the human for the interaction to be socially acceptable.

- 25 • introduction l'interaction rythmique : la ncessit d'apprendre et de s'adapter l'humain. le role du
 26 rythme et des phnomnes de synchronie dans l'interaction physique et sociale
 27 • PhRI : les etude sur le handshaking en robotique, l'absence d'apprentissage et de mise en avant du
 28 phnomne de synchronie (clamping)
 29 • la plasticit hebbienne
 30 • le modele de righetti
 31 • les CPGs et le modele de RS
 32 • plan de l'article

2 MATERIAL & METHODS

33 2.1 Righetti's Hebbian plasticity model

34 Studied since the 19th Century, the relaxation oscillator has nowadays many versions and extensions.
 35 Here, we will be using the Rowat-Selveston model ?, itself adapted from the Van der Pol relaxation
 36 oscillator widely used in CPG models ?, ?.

37 2.1.1 The three forms of the Van der Pol oscillator

38 First form

39 The one-dimension equation for a Van der Pol oscillator is the following:

$$\boxed{\ddot{x} - \mu (1 - x^2) \dot{x} + x = 0} \quad (1)$$

40 Second form

41 With the following transformation:

$$y = x - \frac{x^3}{3} - \frac{\dot{x}}{\mu} \quad (2)$$

42 we have an equation in \dot{x} :

$$\dot{x} = \mu \left(x - \frac{x^3}{3} - y \right) \quad (3)$$

43 By differentiating equation 2, we have \dot{y} :

$$\dot{y} = \dot{x} - x^2 \dot{x} - \frac{\ddot{x}}{\mu}$$

44 By replacing \ddot{x} with its expression in equation 1

$$\dot{y} = \frac{1}{\mu} (\mu (1 - x^2) \dot{x} - \mu (1 - x^2) \dot{x} + x)$$

45 With equation 3, we now have the second form of a Van der Pol oscillator:

$$\boxed{\begin{aligned}\dot{x} &= \mu \left(x - \frac{x^3}{3} - y \right) \\ \dot{y} &= \frac{x}{\mu}\end{aligned}} \quad (4)$$

46 Third form

47 With the following transformation applied on 1:

$$y = \dot{x} \quad (5)$$

48 we immediately have the third form of a Van der Pol oscillator:

$$\boxed{\begin{aligned}\dot{x} &= y \\ \dot{y} &= \mu (1 - x^2) y - x\end{aligned}} \quad (6)$$

49 2.1.2 Righetti learning for a Van der Pol oscillator

50 In ? Righetti proposed a model for frequency learning of a Van der Pol oscillator. The equations are the
51 following:

$$\begin{aligned}\dot{x} &= y + \epsilon F \\ \dot{y} &= -\alpha (x^2 - 1) y - \omega^2 x \\ \dot{\omega} &= \epsilon F \frac{y}{\sqrt{x^2 + y^2}}\end{aligned} \quad (7)$$

52 It is a forced Van der Pol oscillator. The free form is the following:

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -\alpha (x^2 - 1) y - \omega^2 x\end{aligned} \quad (8)$$

53 This form is really close to the third form (equation 6) of the Van der Pol oscillator.

54 2.2 The CPG model based on the Rowat-Selveston neuron

55 The Rybak model was developed by McCrea and Rybak for mammal locomotion. The CPG is divided
56 into two parts representing the extensor and flexor muscles and has three layers: Rhythm Generator, Pattern
57 Formation and Motoneurons. It also takes sensory feedback into account. While this model is widely used

58 for locomotion (? , ?), very few works apply it to arm movements: ? used it to study the reaching movement
 59 and so far I haven't found anyone else. In this work, we use the Rybak architecture with Rowat-Selveston
 60 neurons as described by ? .

61 To this day, the Rowat-Selveston model is underused. Only few studies employ it ?, ?, preferring the
 62 Matsuoka model. In our case, as we wish to be as biologically close as possible and need an online
 63 adaptation of the parameters, this model is better suited to our purpose. t

64 The unforced form of the Rowat-Selverston model is described by these equations:

$$\begin{aligned}\tau_m \ddot{V} + V - A_f \tanh\left(\frac{\sigma_f}{A_f} V\right) + q &= 0 \\ \tau_s \dot{q} &= -q + \sigma_s V\end{aligned}\tag{9}$$

65 This is closer to the second form of the Van der Pol oscillator. The goal is to write it as the third form to
 66 look for a frequency term, like ω in the Righetti equations.

67 2.3 Mathematical modelling

68 2.3.1 Bringing back the RS model to the first form

69 The first equation of 9 can be differentiated:

$$\tau_m \ddot{V} + \dot{V} - \sigma_f \left(1 - \tanh^2\left(\frac{\sigma_f}{A_f} V\right)\right) \dot{V} + \dot{q} = 0$$

70 \dot{q} is given by the second equation of 9:

$$\tau_m \ddot{V} + \dot{V} - \sigma_f \left(1 - \tanh^2\left(\frac{\sigma_f}{A_f} V\right)\right) \dot{V} + \frac{1}{\tau_s} (\sigma_s V - q) = 0$$

71 We replace q by its expression given by the first equation of 9:

$$\tau_m \ddot{V} + \dot{V} - \sigma_f \left(1 - \tanh^2\left(\frac{\sigma_f}{A_f} V\right)\right) \dot{V} + \frac{\sigma_s}{\tau_s} V + \frac{1}{\tau_s} V + \frac{\tau_m}{\tau_s} \dot{V} - \frac{A_f}{\tau_s} \tanh\left(\frac{\sigma_f}{A_f} V\right) = 0$$

72 By grouping the terms,

$$\tau_m \ddot{V} + \left(\frac{\tau_m}{\tau_s} + 1 - \sigma_f + \sigma_f \tanh^2\left(\frac{\sigma_f}{A_f} V\right)\right) \dot{V} + \frac{1 + \sigma_s}{\tau_s} V - \frac{A_f}{\tau_s} \tanh\left(\frac{\sigma_f}{A_f} V\right) = 0$$

73 2.3.2 Third form of the RS model

74 We apply the same transformation 5 to put the model into the third form:

$$\dot{V} = y$$

$$\dot{y} = \frac{1}{\tau_m} \left(\sigma_f - \frac{\tau_m}{\tau_s} - 1 - \sigma_f \tanh^2 \left(\frac{\sigma_f}{A_f} V \right) \right) y - \frac{1 + \sigma_s}{\tau_s \tau_m} V + \frac{A_f}{\tau_s \tau_m} \tanh \left(\frac{\sigma_f}{A_f} V \right)$$

75 2.3.3 Frequency expression

76 By comparing the last equation with the Righetti model 8, ω could be seen as:

$$\omega = \sqrt{\frac{1 + \sigma_s}{\tau_s \tau_m}}$$

77 The issue with this expression is that we have so many additional terms compared to the Righetti model
 78 (the term in tanh for instance) that this approximation might be very wrong in many cases. Moreover,
 79 in the Righetti model (which is simple compared to Rowat-Silverston), ω is not the only parameter that
 80 influences frequency: α does too (and so does the term in front of y in the last model which contains much
 81 more parameters).

82 It remains to be seen if this expression of ω is valid in a certain range of values.

83 2.3.4 Learning with the RS model

84 Considering the last equation in ω is true, we can apply Righetti learning by applying a force in the \dot{V}
 85 equation:

$$\dot{V} = y + \epsilon F$$

86 Of course, F is **not** the injected current that appears in the Rowat-Silverston equation: if we were to take
 87 the injected current into consideration since the beginning, the third form of the RS model would involve
 88 its derivative, which we do not know. Though, it is a force that we want to learn the frequency from.

89 The learning equation would be:

$$\dot{\sigma}_s = 2\epsilon\sqrt{\tau_m \tau_s} \sqrt{1 + \sigma_s} F \frac{y}{\sqrt{V^2 + y^2}}$$

3 RESULTS

90 The experiments have been conducted with the Kinova Mico robotic arm and its simulated version. This
 91 arm has seven degrees of freedom. In the current setup, we work only with two joints. The setup is the
 92 same for simulated or real-world experiments. At the beginning, the robot isn't subjected to any external
 93 force (other than gravity, obviously). At a given time, a sinusoidal perturbation in the vertical plane is
 94 applied to the arm. Finally, the interaction stops. Sensory feedback is taken into account during the whole
 95 process since the input of the CPG is the force exerted on the joint.

96 **3.1 Method**

97 The McCrea and Rybak model previously described is used.

98 The complete CPG for one joint is represented below:

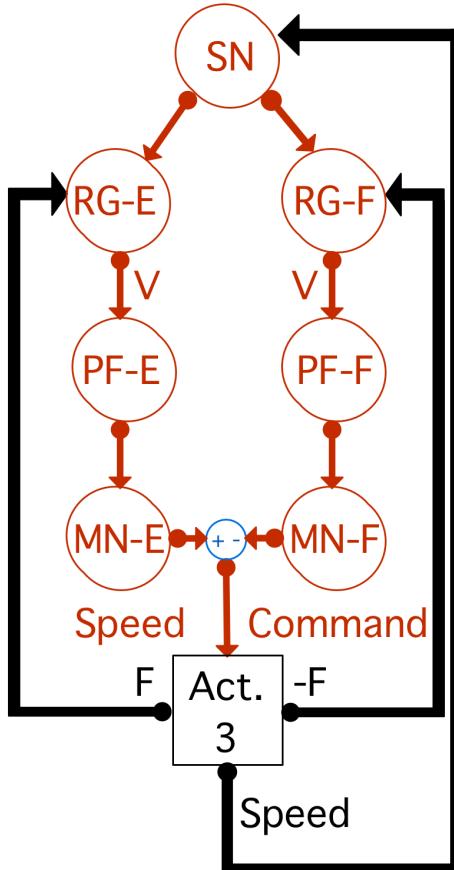


Figure 1. Complete CPG for one joint.

99 The equations for the RS cells are the following:

$$\begin{aligned}\dot{V}_1 &= y_1 - W \frac{y_1}{1 + \exp(-4y_{1other})} + \epsilon F_1 \\ \dot{V}_2 &= y_2 - W \frac{y_2}{1 + \exp(-4y_{2other})} + \epsilon F_2 + \tau_R \sin(\theta(V, y) - \phi) \\ \dot{y}_i &= \frac{1}{\tau_m} \left(\sigma_f - \frac{\tau_m}{\tau_s} - 1 - \sigma_f \tanh^2 \left(\frac{\sigma_f}{A_f} V_i \right) \right) y_i - \frac{1 + \sigma_s}{\tau_s \tau_m} V_i + \frac{A_f}{\tau_s \tau_m} \tanh \left(\frac{\sigma_f}{A_f} V_i \right) \\ \dot{\sigma}_{s_i} &= 2\epsilon F_i \sqrt{\tau_m \tau_s (1 + \sigma_s)} \frac{y_i}{\sqrt{V_i^2 + y_i^2}}\end{aligned}$$

100 With y_{other} being y_E or y_F if the equation concerns the flexor or extensor respectively and $\theta(V, y) =$
 101 $sign(V) \text{acos}(\frac{-y}{\sqrt{V^2+y^2}})$

102 Pattern Formation neuron PF, Sensory neuron S and Motoneurons MN are defined as follows:

$$PF_i = \frac{1}{1 + \exp\left(\frac{-V_i}{2}\right)}$$

$$S_i = \frac{1}{1 + \exp(-0.061342v_{imes})}$$

$$MN_i = \frac{1}{\exp(3(PF_i - S_i))}$$

103 These coefficients have been calculated to match the parameters of the robot. For instance, the sigmoid
 104 slope of the sensory neuron is determined by the range of values of the speed.

105 The choice of the parameters is a crucial step, with inappropriate parameters, the system may not behave
 106 as expected.

107 3.2 Parameter Influence

108 To have an oscillating system, ? determine that $\sigma_F > 1 + \frac{\tau_M}{\tau_S}$ and $\tau_M : \tau_S > 1 : 10$ must be observed.

109 ϵ is the learning step (local oscillation) and sets how much σ_S will oscillate before reaching stability. A
 110 small ϵ will allow for a more robust and stable learning but will require more time to reach stability, if the
 111 interaction doesn't last long enough, it may never be reached. On the other hand, a big ϵ will lead to a more
 112 unstable learning but the final σ_S may be reached.

113 A_F is the amplitude. When A_F is high, σ_S will oscillate globally before reaching stability.

114 For uncoupled CPGs, the natural frequency of the oscillator is determined by τ_M , τ_S and σ_{S0} . Here,
 115 the natural frequency also depends on W. The higher W, the lower the frequency, hence the higher σ_S to
 116 compensate (See Figure 12). For $W \geq 1$, the system isn't able to oscillate.

117 τ_R influences the evolution of σ_{S3} (σ_{S2} is not affected). The higher τ_R , the smaller the final value of σ_{S3}
 118 (See Figure 13). Note that above a given value of τ_R , the system isn't able to oscillate.

119 The initial value of σ_S doesn't change the final value reached. For very high or very low values, the final
 120 σ_S may never be reached if the interaction does not last long enough.

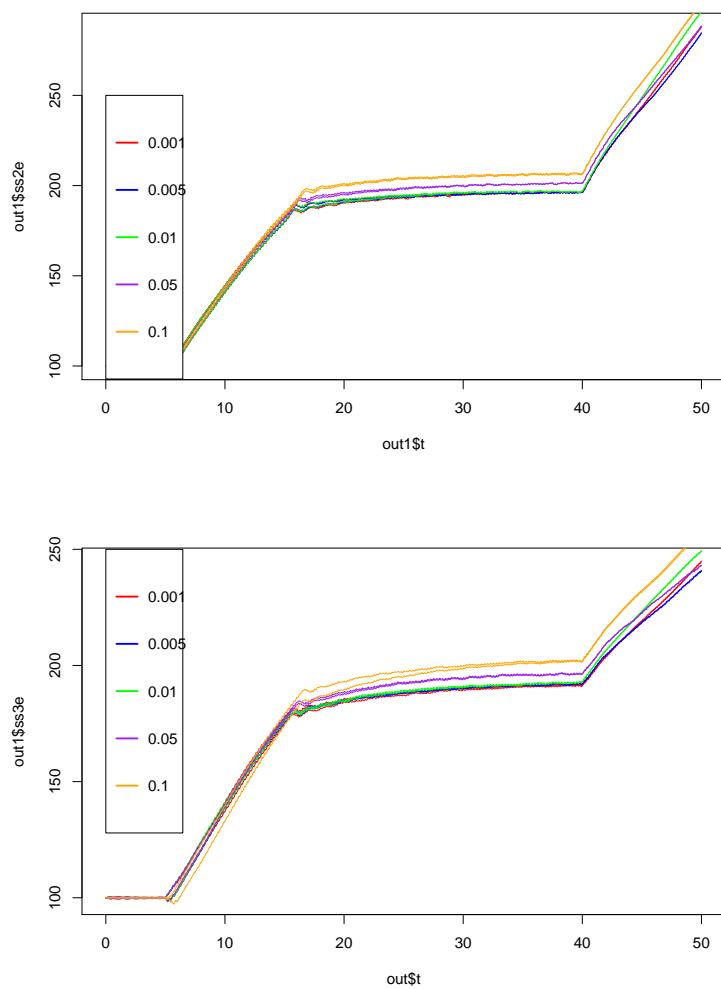


Figure 2. Evolution of σ_{S2E} , σ_{S2F} , σ_{S3E} and σ_{S3F} for various values of W . The initial value is 100 for each σ_S . We can observe that the value of W influences the final value of the σ_S . Below $W = 0.05$, the result is roughly the same. We already observe a demarcation for $W = 0.05$ which is slightly above the others and it's very clear to see for $W = 0.1$

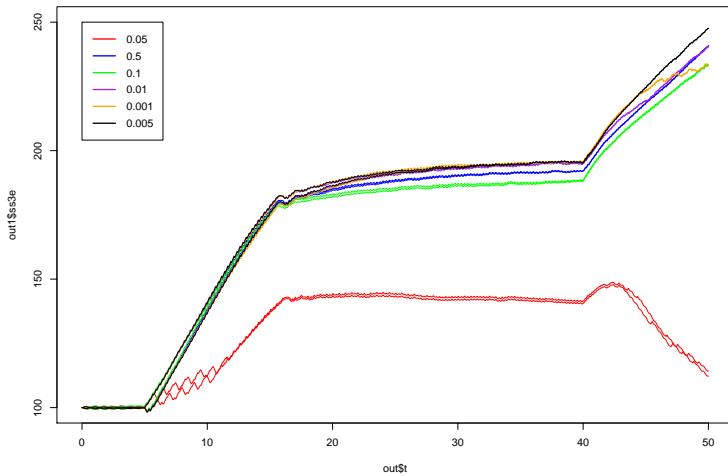


Figure 3. Evolution of σ_{S3E} and σ_{S3F} for various values of τ_R . τ_R does influence the final value of σ_S . For bigger values of τ_R , 0.1 and especially 0.5, the σ_S reached is smaller.

121 3.3 Experimental validation of the plastic CPG

122 The parameters of the algorithm are the following: $\epsilon = 0.1$, $\tau_m = 0.5$, $\tau_s = 10.0$, $\tau_r = 0.005$, $W = 0.1$,
 123 $\sigma_f = 10.0$, $A_f = 1.0$. These parameters have been determined by running simulations beforehand to
 124 adapt well to the robot. When it is not mentioned, the PID constants for the robot are $P = 2.0$, $I = 0.0$,
 125 $D = 0.09$.

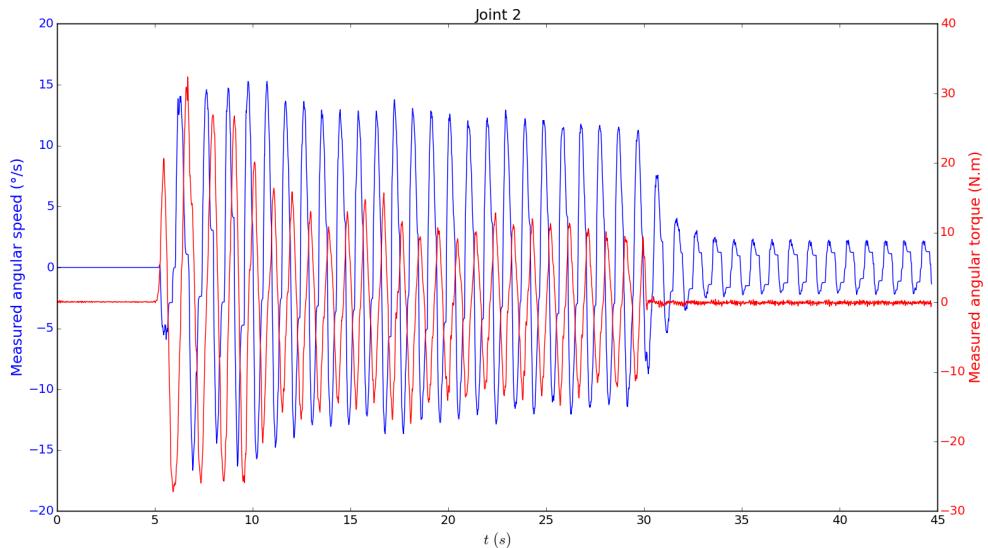


Figure 4. Measure of torque and sent velocity for the second articulation during the experiment with the complete CPG. We can see that, during the first phase of the experiment, the robot arm is not moving. It has not received any perturbation yet, so the motoneurons send the same signal, thus the difference between the two of them is zero. So the command sent to the robot is a zero speed. Right after the robot receives a perturbation, it starts oscillating and adapts to the frequency the human decides (it is seen better on the following figures, with the output of the cell and the torque). Once the arm is released, it continues oscillating at the same frequency. This figure shows that the physical coupling between the human and the robot is achieved.

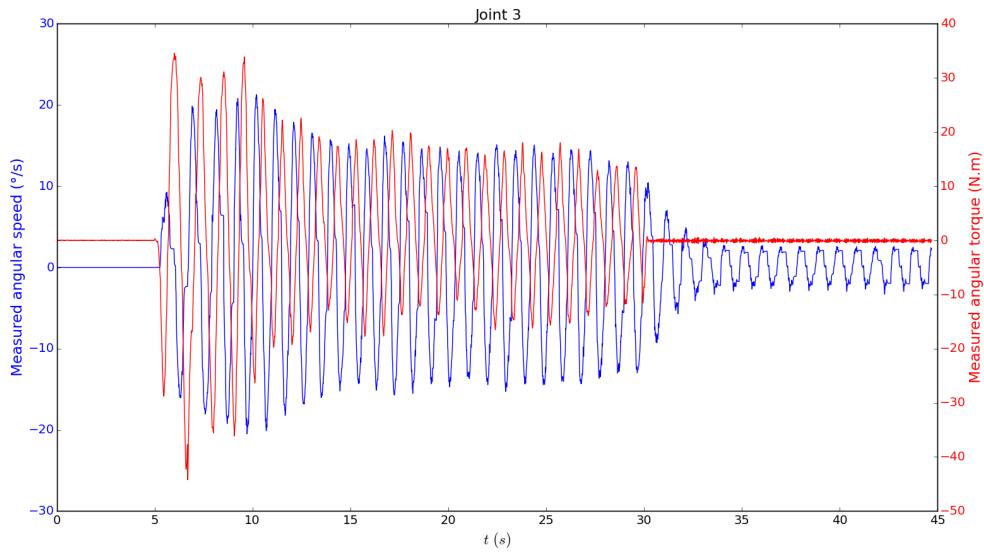


Figure 5. Measures of speed and torque for the third articulation during the experiment with the complete CPG. The same observations can be done, as with the second articulation. Also, the different steps appear roughly at the same time.

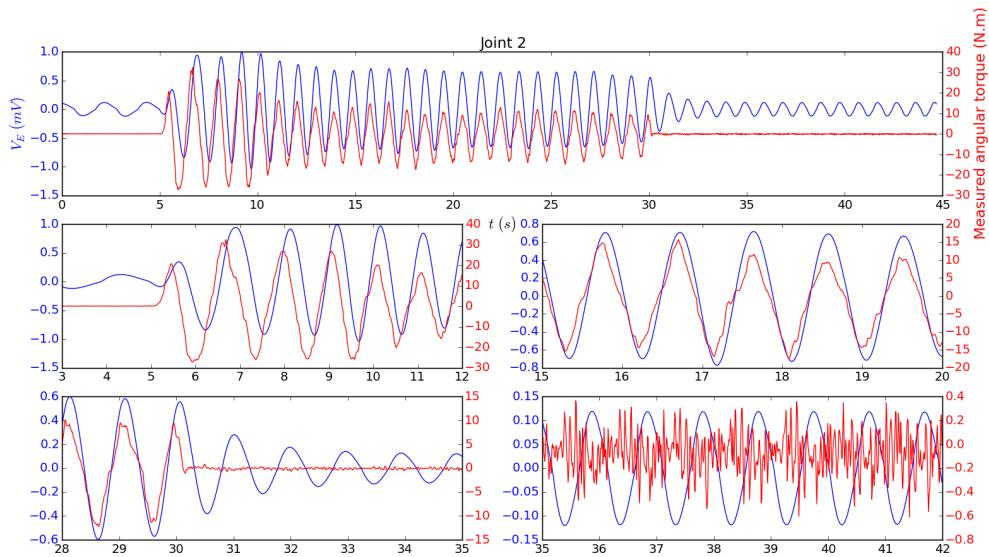


Figure 6. Measures of V_E and torque for the second articulation during the experiment with the complete CPG. Here we can see the detailed phases of the experiment. We can see first the extensor cell oscillating at its intrinsic frequency. When the interaction begins, it adapts to the frequency of the torque. It is synchronised with the torque during the interaction, showing the robot has adapted. Finally, the frequency remains the same when the arm is released. It does not take into account the residual noise after the interaction. It shows that this algorithm makes the robot learn the right frequency.

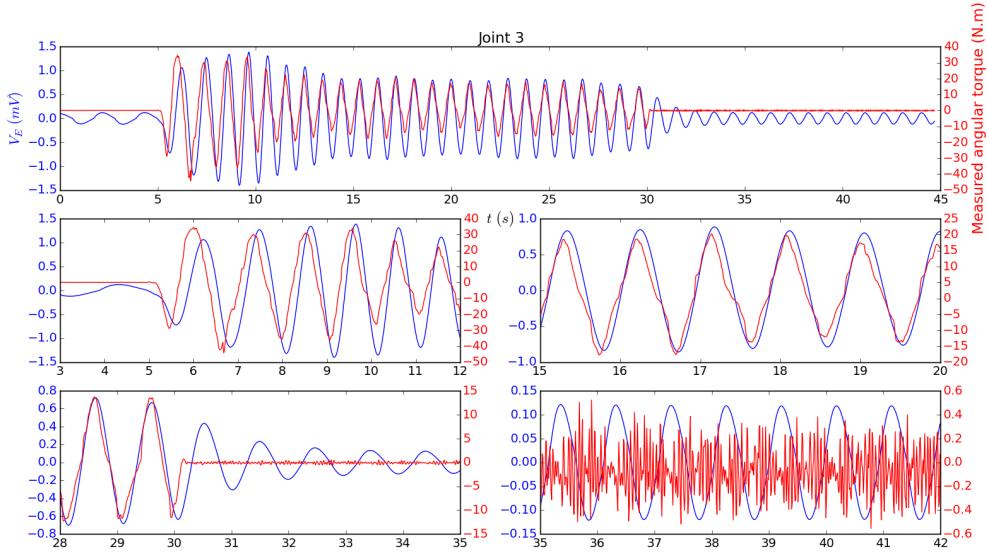


Figure 7. Measures of V_E and torque for the third articulation during the experiment with the complete CPG. Once again, the same observations can be done with the different phases of the experiment. We can also see that the different phases happen roughly at the same time as with the second joint.

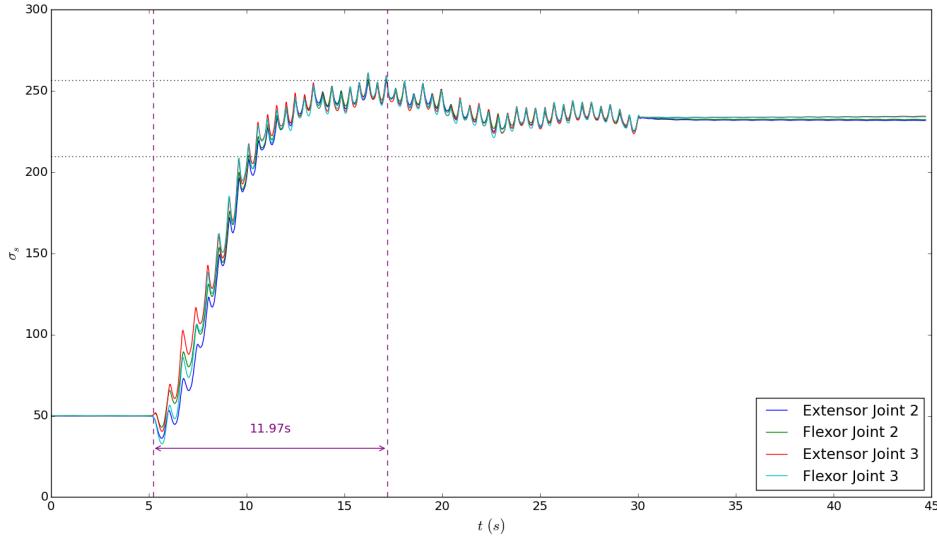


Figure 8. Evolution of σ_s with time during the same experiment. This figure takes into account the four values of σ_s that are present in the system: second joint extensor, second joint flexor, third joint extensor, third joint flexor. The four of them follow roughly the same direction, though some are ahead of the others at the start (based on this image and other experiments, we can't predict which one is going to be ahead of the others). The diagram shows the response time at 10%, which is 11.97s. It is an acceptable response time for this interaction. We can even consider that the robot learned the frequency at around 11s (when σ_s arrives to a plateau), which reduces the response time to 6s. The oscillations before stabilisation can also be affected by the human slightly changing the frequency (because we do not keep exactly the same frequency during a handshaking) although during the experiments, caution has been taken to reproduce the same movement. When the interaction stops, all of the σ_s values remain constant: the robot has learned the right frequency.

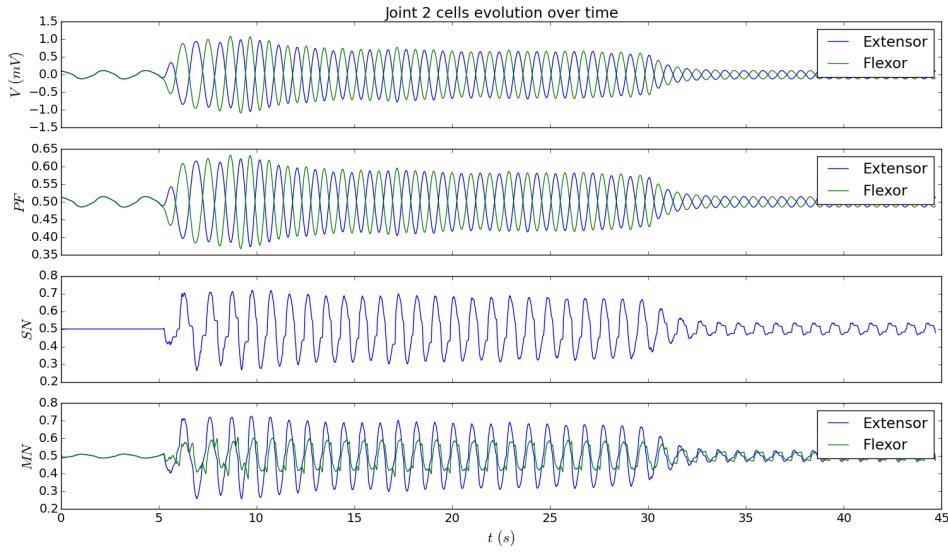


Figure 9. Evolution of the different cells with time during the same experiment, for the second joint. We can see that the RG cells (V on the figure) are in phase during the first phase of the experiment. This is because the robot has not interacted with the human yet. Then, the frequency adapts to the torque frequency and the two cells are in opposition of phase. Finally, when the arm is released, the two cells continue on being in opposition of phase. The pattern formation neurons are here to normalise the rhythm generator cells, and we can see that they follow the same pattern. The sensory neuron has to normalise the speed measured by the robot. Once again, the phases of the experiment appear clearly on this diagram. Finally, the last diagram shows the motoneurons output. We can see the two motoneurons do not have the same frequency during the interaction. This difference can be explained by the fact that we take two signals in opposition of phase from PF and we compare both of them against one signal coming from SN. The difference of phase between the two signals coming from PF explains the difference in amplitude for the motoneurons.

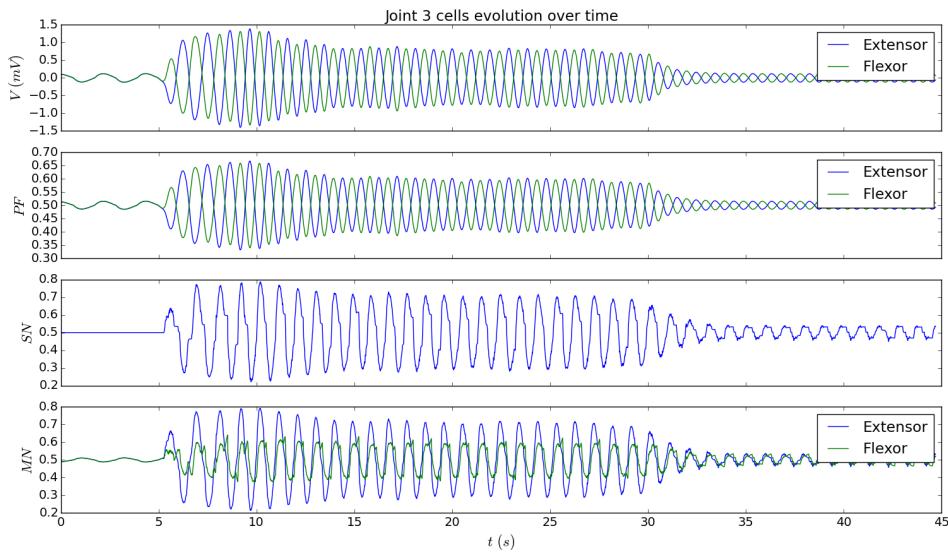


Figure 10. Evolution of the different cells with time during the same experiment, for the third joint. We can observe exactly the same patterns as with the second joint.

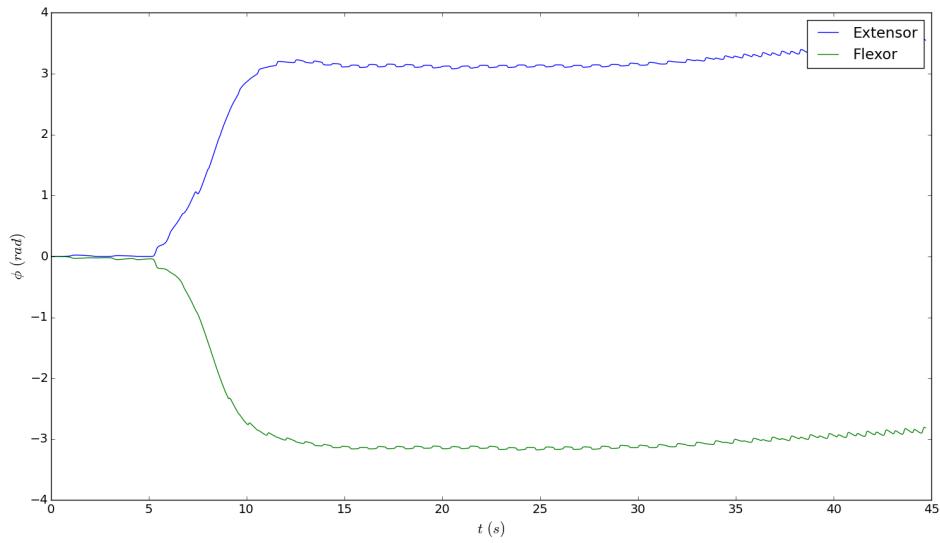


Figure 11. Evolution of ϕ (phase difference between the two joints) for the extensor and the flexor parts of the system. The two articulations are coupled in order to match the phase difference between the two corresponding joints on the human. It is an adaptation system based on the Righetti model, not a learning model (meaning that it adapts as long as the interaction is going, but when the interaction ends, the difference of phase diverges). The robot has adapted its phase difference roughly at the same time it has learned the frequency (with σ_s). We can see here that joint 2 is almost in opposition of phase with joint 3. Also, the phase difference in the flexor part is the same as the phase difference in the extensor part (modulo 2π), which is to be expected.

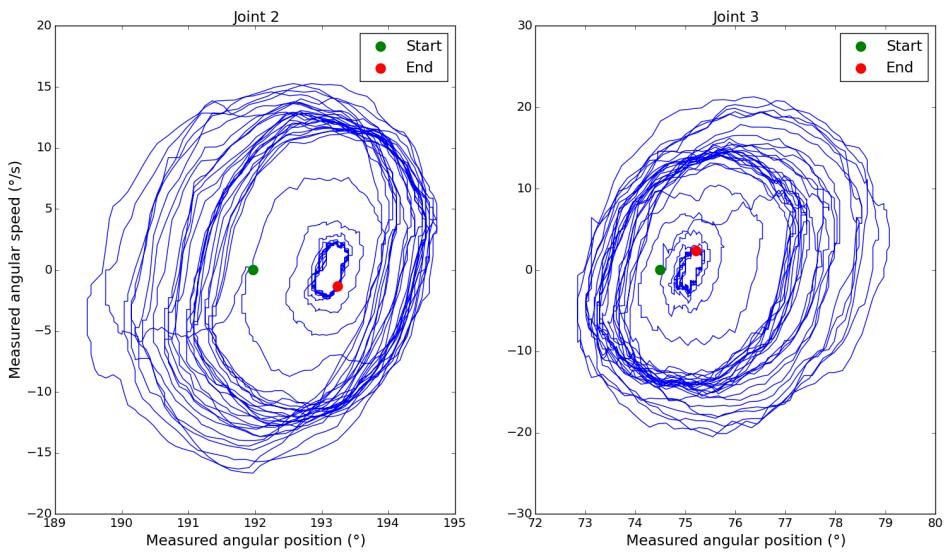


Figure 12. Phase portrait of the arm movement for the second and third articulations, during the same experiment. Here we can only see two cycles. In fact, the first five seconds of the experiment are located at the starting point, since the robot does not move at all. The outer cycle is the interaction part, while the inner cycle is the "arm released" part. Once again, the cycle does not change in shape, but changes in size (due to the amplitude modification) when the arm is released.

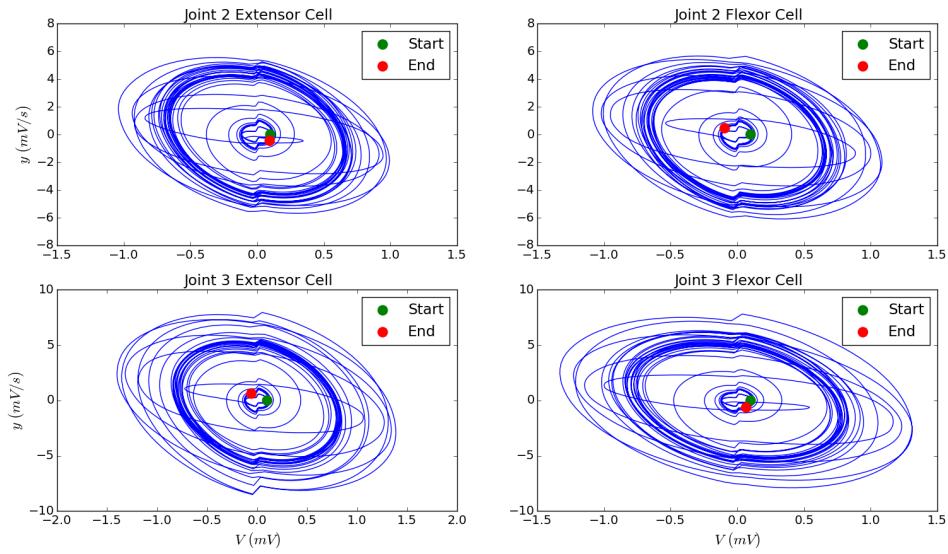


Figure 13. Phase portrait in the (V, y) plan for the second and third articulations, during the same experiment. Three different cycles can be seen here. First, the starting cycle (most inner circle on the four diagrams), when the rhythmic cells oscillate at their intrinsic frequency. Then, the interaction cycle (most outer circle) when the human and the robot are interacting. Finally, the middle circle is the end cycle: it has the same shape as the interaction cycle, but it changes in size. The fact that cycles appear show that the system is stable, meaning the frequency is learned.

126 3.4 Experimental validation of the CPG in simulation

127 The simulations have been run in V-REP Simulation software with the Kinova Mico robotic arm. A
 128 gripper has been added to the robotic arm. We simulate the handshake with a ball placed inside the gripper.
 129 It moves up and down according to a 2 Hz sinusoidal signal of amplitude 0.16. Since both objects are
 130 collidable, it forces the arm to move.

131 The method is the same as described before. Each simulation lasts 50 s. The interaction starts at $t = 5$
 132 s and stops at $t = 40$ s. The parameters are slightly different: $\epsilon = 0.01$, $\tau_M = 0.35$, $\tau_S = 3.5$, $\tau_R = 0.05$,
 133 $W = 0.005$, $\sigma_F = 10$ and $A_F = 0.05$.

134 In the simulation setup, in the first phase (before the interaction starts), the velocity sent to the joints
 135 is not zero because the ball is already inside the gripper before the interaction. The contact between the
 136 gripper and the ball and the ball counter-acting gravity generate some noise forces. These velocities are
 137 negligible. When the interaction starts, we could observe a massive increase in the force applied to the
 138 joints. The magnitude stays the same roughly 12 s, until σ_S and ϕ stabilise.

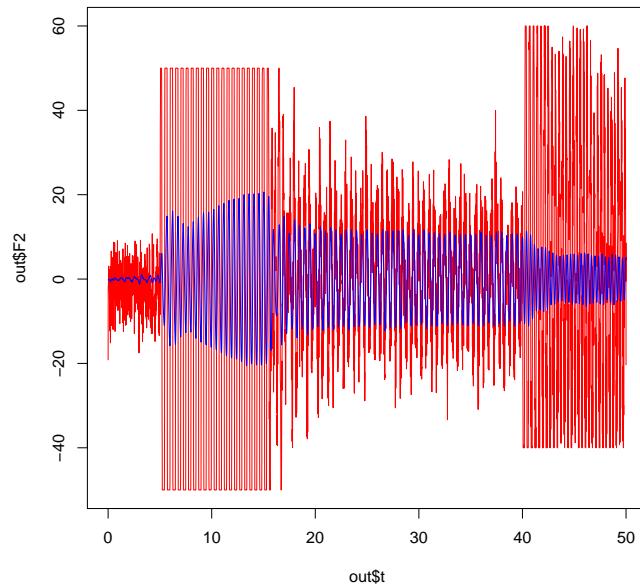


Figure 14. Evolution of s_2 and F_2 . We can observe that s is not zero before the interaction starts, contrary to the real robotic arm. When the simulation starts, the ball is already inside the gripper before the interaction. The contact between the gripper and the ball and the ball counter-acting gravity generate some noise forces. When the synchronisation phase is over, the force exerted on the arm decreases and F_2 and s_2 can be observed to be perfectly synchronised. When the interaction stops, the joint goes on oscillating, though at a lower amplitude.

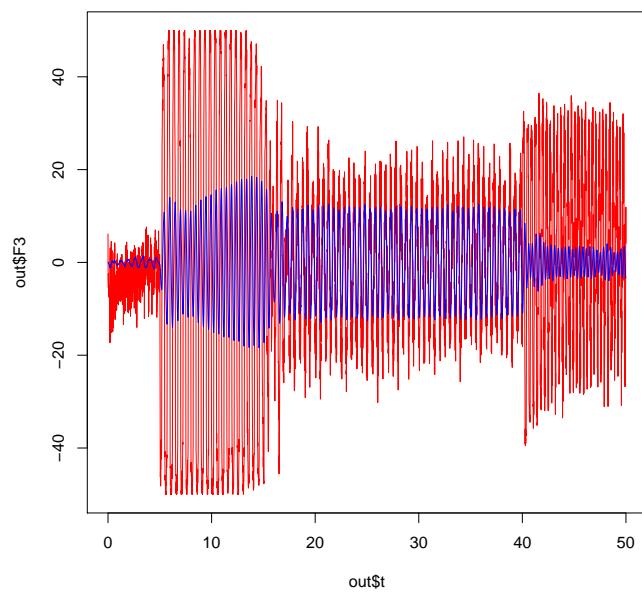


Figure 15. Evolution of s_3 and F_3 (Input of the cell) during the experiment.

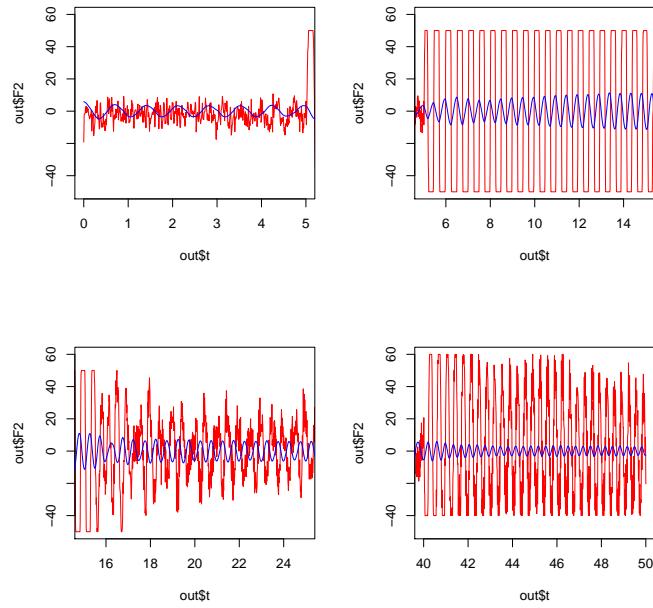


Figure 16. Evolution of V_{2E} and F_2 (Input of the cell) during the experiment. It can be divided into four parts (depicted in the bottom pictures). During the first phase, there is no physical interaction so the arm oscillates at its natural frequency. The second phase corresponds to the synchronisation phase. We can observe the natural frequency of the oscillator changing until it matches the input frequency. In the third phase, the two signals are synchronised. Finally, at 40 s, the interaction stops but the arm keeps oscillating at the frequency learned during the interaction

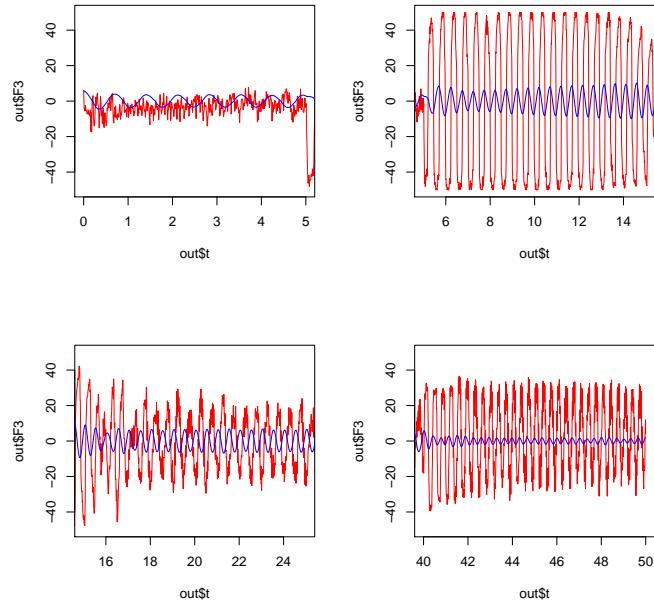


Figure 17. Evolution of V_{3E} and F_3 (Input of the cell) during the experiment. It can be divided into four parts (depicted in the bottom pictures). During the first phase, there is no physical interaction so the arm oscillates at its natural frequency. The second phase corresponds to the synchronisation phase. We can observe the natural frequency of the oscillator changing until it matches the input frequency. In the third phase, the two signals are synchronised. Finally, at 40 s, the interaction stops but the arm keeps oscillating at the frequency learned during the interaction

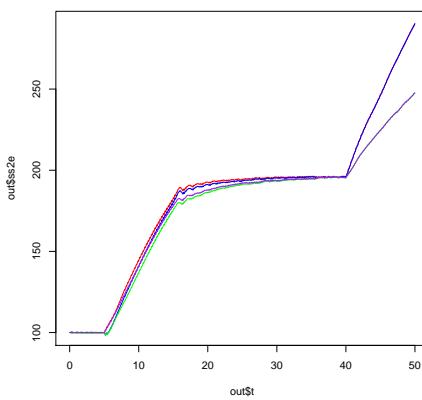
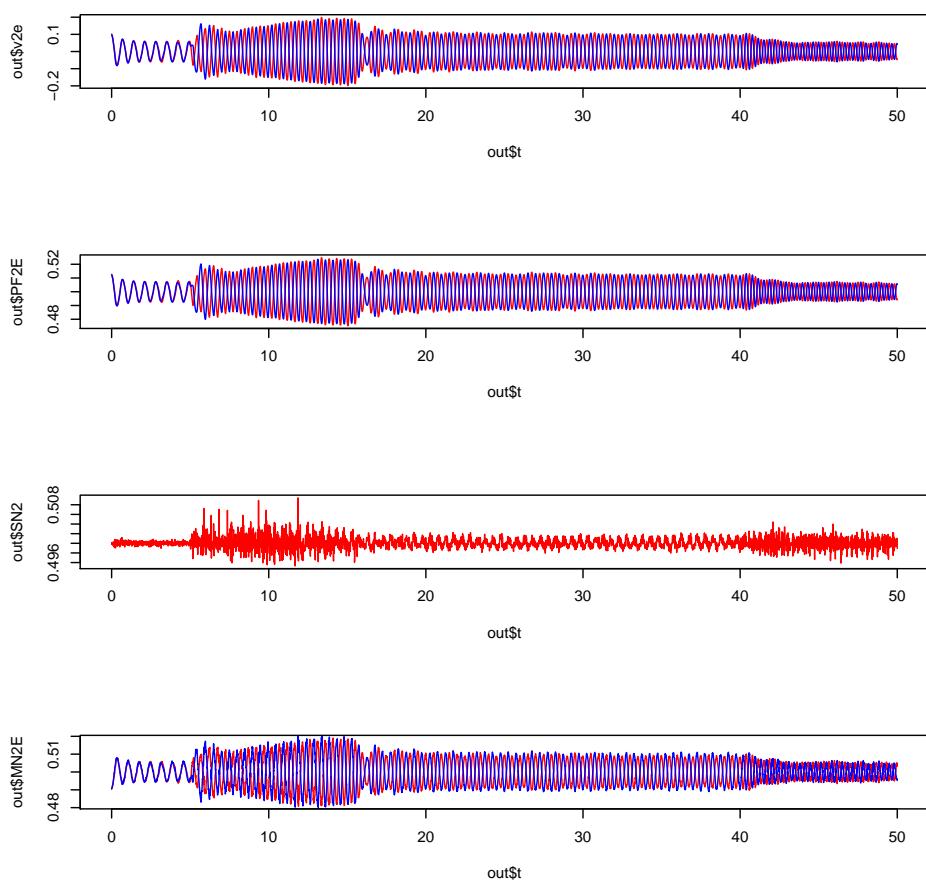
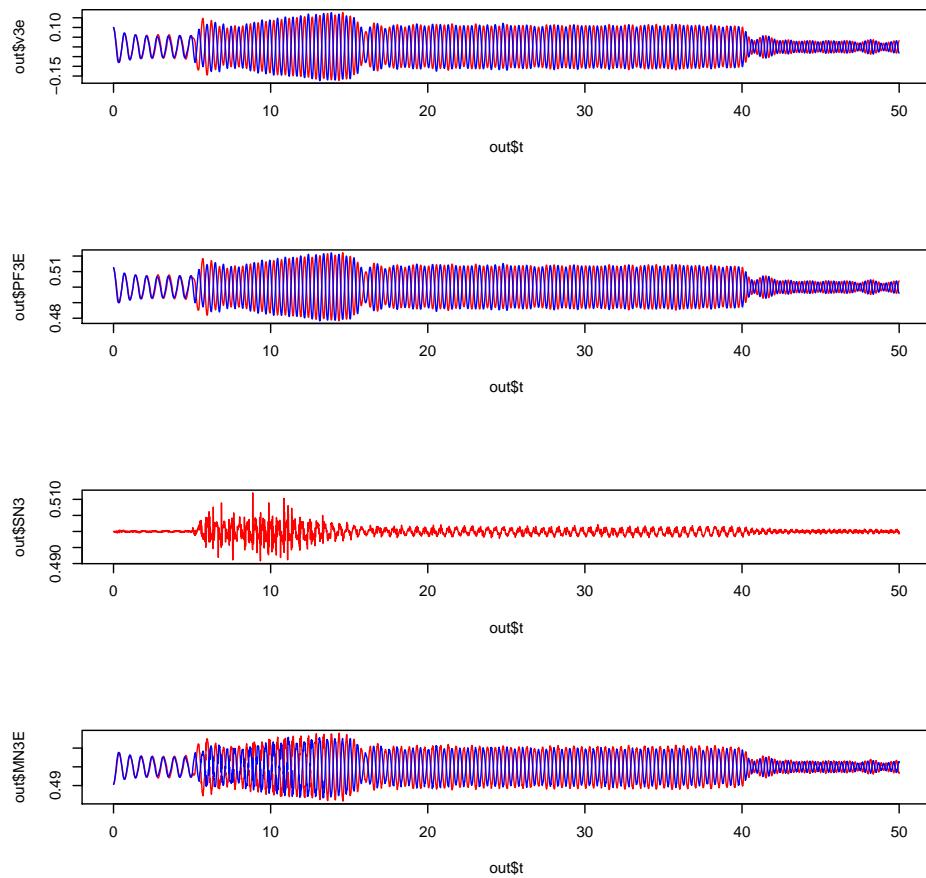


Figure 18. Evolution of σ_{S2E} , σ_{S2F} , σ_{S3E} and σ_{S3F} . The initial value is 100 for each σ_S . They stay mainly stable at 100 until $t = 5$ s when the interaction starts. Then they start increasing, all following the same direction, though some are slightly slower than others they finally catch up around $t = 15$ s. From $t = 19$ s onwards, the σ_S are mostly stable around 190. When the interaction stops at $t = 40$ s, the σ_S remain stable, showing that the new value has indeed been learned. The final values for each joint are slightly different but the difference is negligible.

**Figure 19.**

**Figure 20.**

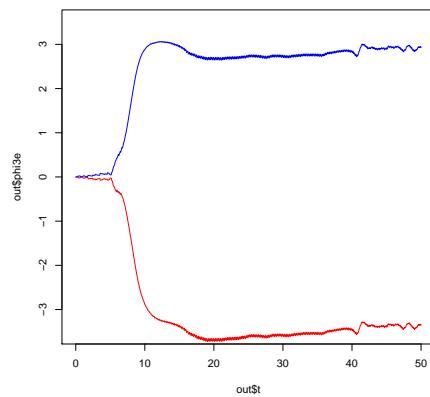


Figure 21. Evolution of ϕ during the experiment for the extensor and flexor of the second joint. Both ϕ start at 0. When the interaction starts, we can observe a strong divergence leading both ϕ to be opposed. They remain stable until the interaction stops and then they start increasing.

139 **3.5 Influence of Rigidity**

140 Several studies have shown that arm stiffness greatly influences the nature of the movement. The stiffer
 141 the joints of the arm are, the easier it will be to control them. As reported in ?, during cyclic tasks, the
 142 natural frequency of the human arm is adapted to match the first harmonic frequency. In human-robot
 143 systems, if the robot is not stiff enough, the system is highly unstable, leading the human to increase
 144 stiffness in his arm, thus creating even more instability. Hence, appropriate stiffness of the robotic arm is a
 145 paramount parameter for a successful interaction. This was demonstrated by studies showing that a system
 146 able to adapt its stiffness lead to smoother and more successful interactions. Based on that understanding,
 147 we postulate here, that stiffness also affects the ability of the arm to adapt to external stimulations and to
 148 learn new frequencies.

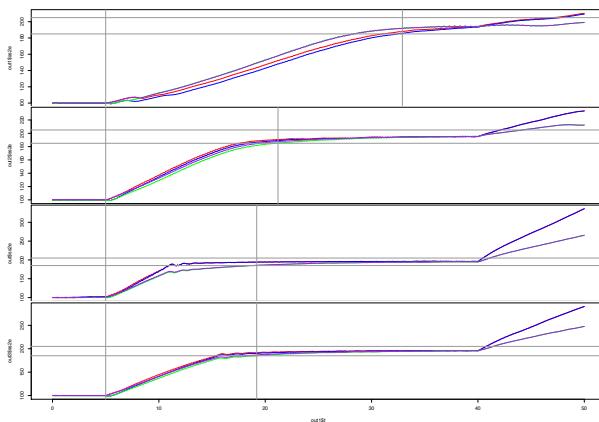


Figure 22. Evolution of σ_{S2E} and σ_{S2F} for various values of P. For $P = 0.001$, response time is 12.4 s, which is close to 12 s for $P = 0.005$. For $P = 0.02$ and $P = 0.05$, response times are similar at 7.7 and 8.3 s respectively. For $P = 1$, response time is 9.4 s but note that for that value of P, the arm was behaving erratically, due to high instability. (this can be observed by looking at the recorded position).

149 The evolution of σ_S was recorded for various values of P and D. Varying P and D allows us to change the
 150 stiffness of the arm. It can be observed that P has indeed a strong influence on the response time of the

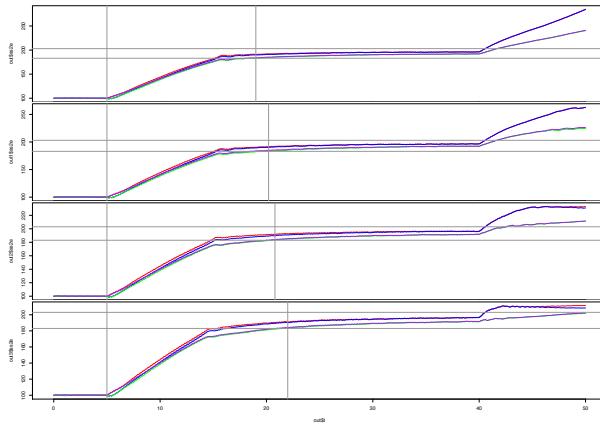


Figure 23. Evolution of σ_{S2E} and σ_{S2F} for various values of D (0, 0.0001, 0.0005 and 0.001). We can observe that the higher D is, the less the greater the system response is but the better it retains the learning of σ_S .

151 system, ie. how fast it adapts to the new frequency. For values of P too low or too high, the response time is
152 higher. We can also note that from $P = 0.5$ onwards, the arm is not stiff enough and oscillates erratically
153 before and during the interaction. The role of D on the learning process is less obvious we can still infer
154 from our experiments that it does influence how well the joint retains the learning.

155 The same observations were made on the real robot.

156 The stability and speed of the algorithm are better when P and D are increased.

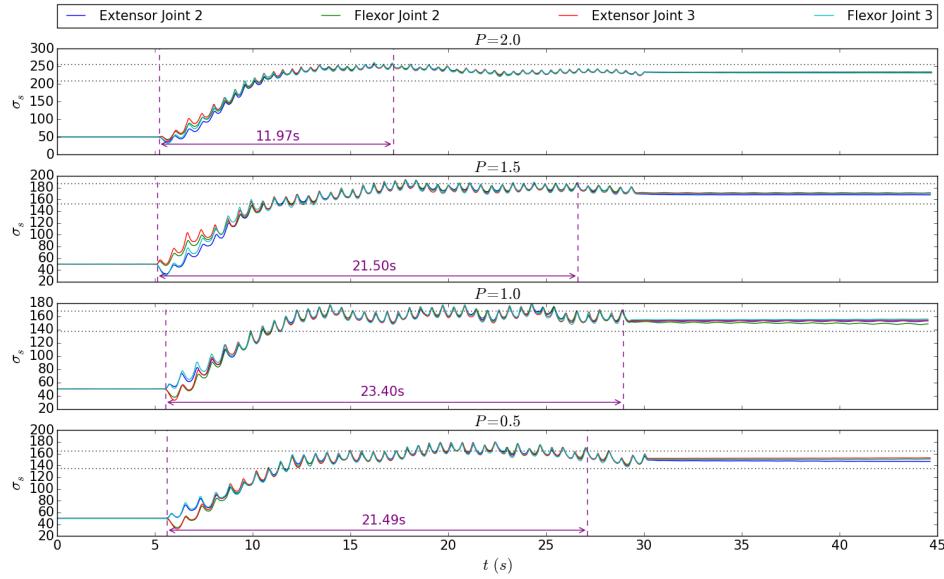


Figure 24. Influence of P in the learning process (the other values, I and D , are the default values, 0 and 0.05). We can clearly see that the greater the P , the faster the system is (its response time is reduced). It is to be expected, of course. Another interesting fact is that the final value of σ_s decreases as P decreases: this means the robot arm cannot achieve high frequency values when P is too small. This is also to be expected, since the system takes more time to get to the position it wants to go to.

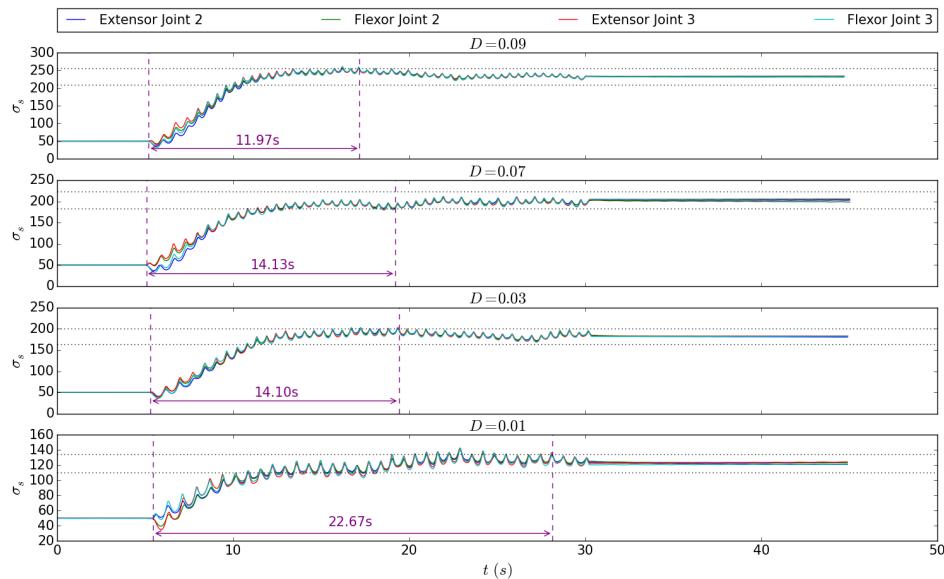


Figure 25. Influence of D in the learning process (the other values, P and D , are the default values, 2 and 0). We can make the same observations as with P , though the response time remains correct for D greater than 0.03.

4 DISCUSSION

157 Discussion text

DISCLOSURE/CONFLICT-OF-INTEREST STATEMENT

158 The authors declare that the research was conducted in the absence of any commercial or financial
159 relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

160 Contributions text

ACKNOWLEDGMENTS

161 Acknowledgments text

162 *Funding:* Funding text

SUPPLEMENTAL DATA

163 Supplemental data text