CrossMark

Decision-making neural circuits mediating social behaviors

An attractor network model

Julián Hurtado-López[1] (ID) · **David F. Ramirez-Moreno**[2] · **Terrence J. Sejnowski**[3]

Abstract We propose a mathematical model of a continuous attractor network that controls social behaviors. The model is examined with bifurcation analysis and computer simulations. The results show that the model exhibits stable steady states and thresholds for steady state transitions corresponding to some experimentally observed behaviors, such as aggression control. The performance of the model and the relation with experimental evidence are discussed.

Keywords Nonlinear dynamics · Bifurcation analysis · Decision-making mechanisms · Dynamics of social behavior · Mounting and attack circuits

✉ Julián Hurtado-López
  jhurtado@uao.edu.co

  David F. Ramirez-Moreno
  dramirez@uao.edu.co

  Terrence J. Sejnowski
  terry@salk.edu

[1] Department of Mathematics, Universidad Autónoma de Occidente, Cll 25 No. 115-85 Km 2 vía Cali-Jamundí, 760030, Cali, Colombia

[2] Department of Physics, Universidad Autónoma de Occidente, Cali, Colombia

[3] Howard Hughes Medical Institute, the Salk Institute for Biological Studies, La Jolla, California 92037, USA

1 Introduction

Even though neural mechanisms that support aggression and reproductive behaviors in mice have been studied (Nelson and Chiavegatto 2000; Nelson and Trainor 2007; Veenema and Neumann 2007; Choi et al. 2005; Crews 2005; Dulac and Kimchi 2007), these mechanisms are not well understood yet. In addition, how aggression, mating, and the associated synaptic circuits relate to each other remains not clearly understood.

In recent works, optogenetic manipulations in certain brain nuclei in mice have shown the relationship between aggression circuits and those mediating opponent social behaviors, such as mating (Lin et al. 2011; Lee et al. 2014).

Lin et al. (2011) identified cells localized within a small subdivision of ventromedial hypothalamus of males that are activated during male-female mating, and which are mostly different from those activated during fighting. Moreover, optogenetic stimulation of ventrolateral subdivision of the ventromedial hypothalamus (VMHvl) neurons showed that these neurons are necessary and sufficient for offensive aggression. Additionally, authors suggested that a neural correlate of competitive interactions between fighting and mating is likely. This conclusion is supported by the observation that many male-activated units were actively inhibited by females presence, and a higher intensity of photostimulation was required to evoke attack towards a female during mating encounters.

Lee et al. (2014) identified a subset of VMHvl neurons marked by estrogen receptor 1 (Esr1) which at lower photostimulation intensities promote close investigation and mounting rather than attack towards both males an females. Their data suggest that these neurons in VMHvl control the progression of a social encounter in a scalable manner.

Hashikawa et al. (2016) identified the main olfactory pathway as the way for carrying information for social approach and the accessory olfactory system as the pathway activated by social investigation. Both systems have projections into VMHvl and MPOA (medial pre optic area) as well.

Thus, these results should shed some light on how control of opponent innate behaviors by the VMH is performed at a circuit level.

Despite the new findings, relatively little is known about the connectivity of the neurons that control social behaviors. In order to give some insight and provide an explanation for the mechanisms and processes that might control social behaviors, we propose a biologically inspired neural network model based on recent findings.

This model aims to explain social behaviors from a dynamical system perspective that has not been worked in the literature cited.

First, we propose interactions between mounting and aggression circuits in hypothalamus based on optogenetics' experimental data. The hypothalamus plays an essential role in aggression and sexual behaviors (Nelson and Trainor 2007; Swanson 2000; Choi et al. 2005; Sokolowski and Corbin 2012; Hashikawa et al. 2016).

Second, we develop a mathematical model based on continuous attractor neural networks with dynamical stability (Wang 2009), which gives us a framework for understanding the rich and complex neural activity patterns present in recurrent networks. Attractor network models have been increasingly used by neuroscientists (Tegnér et al. 2002; Boucheny et al. 2005; Milton et al. 2010; Trotta et al. 2012; Wang 2008, 2009, 2012; Wang et al. 2013; Hurtado-López and Ramirez-Moreno 1350).

The paper is organized as follows: the methods section describes the possible circuitry supporting the interactions between mounting and attack when mice are artificially stimulated. This section also describes the dynamics of our mathematical model based on the Wilson-Cowan model and a simplified mean-field approach. The results section follows and presents the relevant findings. The performance of the network was validated through simulations. We use numerical bifurcation analysis and find the stationary states and parameters' values where state transitions take place. Section 4 provides conclusions and a discussion of the apparent potential usefulness of the proposed neural network.

2 Methods

2.1 Proposed neural network model

The model proposed in this work describes a possible synaptic circuit underlying the interactions among close investigation (CI), mounting, and attack in the male mouse when artificially stimulated, see Fig. 1. The model assumes that
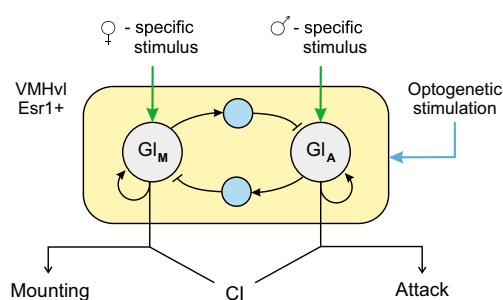
Fig. 1 Possible circuitry underlying the interactions among CI (close investigation), mounting, and attack in VMHvl. Gl_M represents the mount-specific sub-population and Gl_A represents the attack-specific sub-population. Blue circles are inhibitory interneurons populations. *Black lines* ending in a bar are inhibitory synapses and those ending in arrows are excitatory synapses. *Green lines* are inputs and represent specific stimuli coming from females and males. *Blue line* is an input representing the intensity of photostimulation

mounting and attack decisions come from interactions of a mounting circuit, Gl_M, and an attack circuit, Gl_A, (Gl stands for glutamatergic neurons) which are inhibiting each other through interneurons. Gl_M and Gl_A are proposed as pools of excitatory Esr1$^+$ VMHvl neurons making excitatory connections onto themselves. These pools of neurons represent different sex-specific sub-populations of Esr1$^+$ neurons where the Gl_M circuit is activated by females and promotes mounting while the attack-promoting circuit, Gl_A, is activated by males. Additionally, Gl_M and Gl_A circuits receive the same input (a train of light) due to photostimulation.

2.2 Mathematical model

In order to gain understanding of dynamical processes during social encounters we propose a reduced model which has two units (dynamical variables) that follows a mean-field approach. Local interneurons are not taking into account in the mathematical model, because the simplified model yields essentially the same dynamics and behavior.

In the computational neurobiology literature, one of the most influential mean field firing-rate model was developed by Wilson and Cowan (Wilson and Cowan 1972). Based on the Wilson-Cowan model, the dynamics for our network is modeled by the dynamical system given by Eqs. (1) and (2).

$$\tau_1 \frac{dx_1}{dt} = -x_1 + f_1(S_1 + w_{11}x_1 - w_{12}x_2 + L) \tag{1}$$

$$\tau_2 \frac{dx_2}{dt} = -x_2 + f_2(S_2 + w_{22}x_2 - w_{21}x_1 + L) \tag{2}$$

Variables x_i correspond to mean firing rates describing the states of the neural sub-population i. Indeed, x_1 and x_2 represent, respectively, the level of activity of Gl_M and Gl_A circuits. The time constant governing the mean-firing rate of each type of sub-population's response is given by

τ_i. The term w_{ij} represents the magnitude of the connection from unit j to unit i. S_1 and S_2 are female-specific and male-specific signal values that activate, respectively, Gl_M and Gl_A circuits. L represents the photostimulation intensity. f_i represents the transfer function (gain function) of the neuronal sub-population i and is defined in Eq. (3).

$$f_i(u) = \begin{cases} \dfrac{m_i u^{N_i}}{\sigma_i^{N_i} + u^{N_i}} & \text{for } u \geq 0 \\ 0 & \text{for } u < 0 \end{cases} \quad (3)$$

m_i represents the maximum mean firing rate. The constant σ_i is the semi-saturation value and is the point at which f_i reaches half of its maximum. Finally, N_i determines how sharp the transition is between threshold and saturation.

Unless otherwise mentioned, the general set of parameters for this model is shown in Table 1.

In our model, the values proposed for parameters m_1, m_2, τ_i, and photostimulation L are close to the experimental values found in the cited literature. Values estimated for synaptic weights w and semisaturation values σ were heuristically proposed in order to drive the model to elicit specific behaviors.

Parameter values define that Gl_M and Gl_A neurons have different activation thresholds, $\sigma_1 < \sigma_2$. Furthermore, $m_1 < m_2$ means that $Esr1^+$ VMHvl neurons should present a greater degree of activation during attack than during mounting. It is also worth pointing out that an asymmetric and reciprocal inhibition is made between Gl_M and Gl_A neuronal sub-populations as shown by $w_{21} < w_{12}$. These observations were suggested by Kennedy et al. (2015) and then adjusted in our model.

In addition, VMHvl $Esr1^+$ neurons in both sub-populations were optogenetically activated under the same

infrared light. Certainly, it does not assure that both sub-populations have to respond equally. The optogenetic activation can affect differently both sub-populations, fact that is taken into account in our model as different parameters in the activation function are associated to each sub-population, respectively. It leads to possible different responses and behaviors.

2.3 Simulations

Since we have a reduced model it does not require a high temporal resolution. Therefore, the numerical integration of the dynamical equations was executed by a fourth order Runge-Kutta with an integration time step of 0.02. We added noise as a background signal in inputs S_1, S_2, and L using a normal distribution. We applied the bifurcation analysis of the model by the MATCONT package using numerical continuation of equilibria (Dhooge et al. 2003).

3 Results

Numerical simulations of the proposed network were performed at different intensities for S_1, S_2, and L. On the one hand, stimuli S_1 and S_2 run over the interval [0,1] where, in a normal male-female (male-male) encounter, $S_1 \rightarrow 0^+$ ($S_2 \rightarrow 0^+$) describes conditions of low state of arousal or motivation and $S_1 \rightarrow 1^-$ ($S_2 \rightarrow 1^-$) describes extreme state of arousal or motivation. On the other hand, photostimulation intensity L runs over the interval [0,5].

Unless otherwise mentioned, system's evolution is considered from the resting state $(0, 0)$ taken as the initial state. Furthermore, stable steady state levels of activity in the controlling neural circuits will correspond to behavior modes in either male-female or male-male encounters.

3.1 Network performance in normal male social behavior

We first executed simulations of the neural circuits when they are activated by specific stimuli coming from females or males in absence of photostimulation. Therefore, circuits might exhibit normal male-female and male-male encounters.

Figure 2 shows the states of the VMHvl neuronal activity. Indeed, in a normal male-female encounter, our model shows, as expected, that only Gl_M is activated and just two different states may be presented, close investigation or mounting. These states are represented by stable equilibrium points shown in Fig. 2a. Black line represents social or close investigation states whereas the blue line represents mounting behavior. Therefore, regardless of the initial level of activity of Gl_M only mounting is promoted when $S_1 > 0.184$.

Table 1 Default parameters for the model

Network parameters	Symbol	Value
Recurrent strength of Gl_M	w_{11}	0.5
Recurrent strength of Gl_A	w_{22}	0.7
Synaptic weight from Gl_M to Gl_A	w_{21}	0.6
Synaptic weight from Gl_A to Gl_M	w_{12}	2
Synaptic time constants	τ_i	0.4 sec
Maximum mean firing rate of Gl_M	m_1	5
Maximum mean firing rate of Gl_A	m_2	10
Semi-saturation value of Gl_M	σ_1	1.3
Semi-saturation value of Gl_A	σ_2	3.6
Maximum slope of f	N_1, N_2	2

Fig. 2 a,c Stimulus-response curves. *Solid and green dashed lines* correspond to the stable and unstable equilibrium points, respectively. *Red asterisks* show the points where the system switches from a system with one stable state to a system with three equilibrium points and then to one with one stable state again. *Black, blue and red line* represent, respectively, close investigation, mounting and attack states. **b,d** Six time-mean firing rate sample trials are shown for a given value of S_1 and S_2 (*dashed line* in (**a**) and (**c**), respectively). The switch from close investigation occurs when the level of activity of VMHvl is greater than a particular threshold, saddle point (*green dot*)
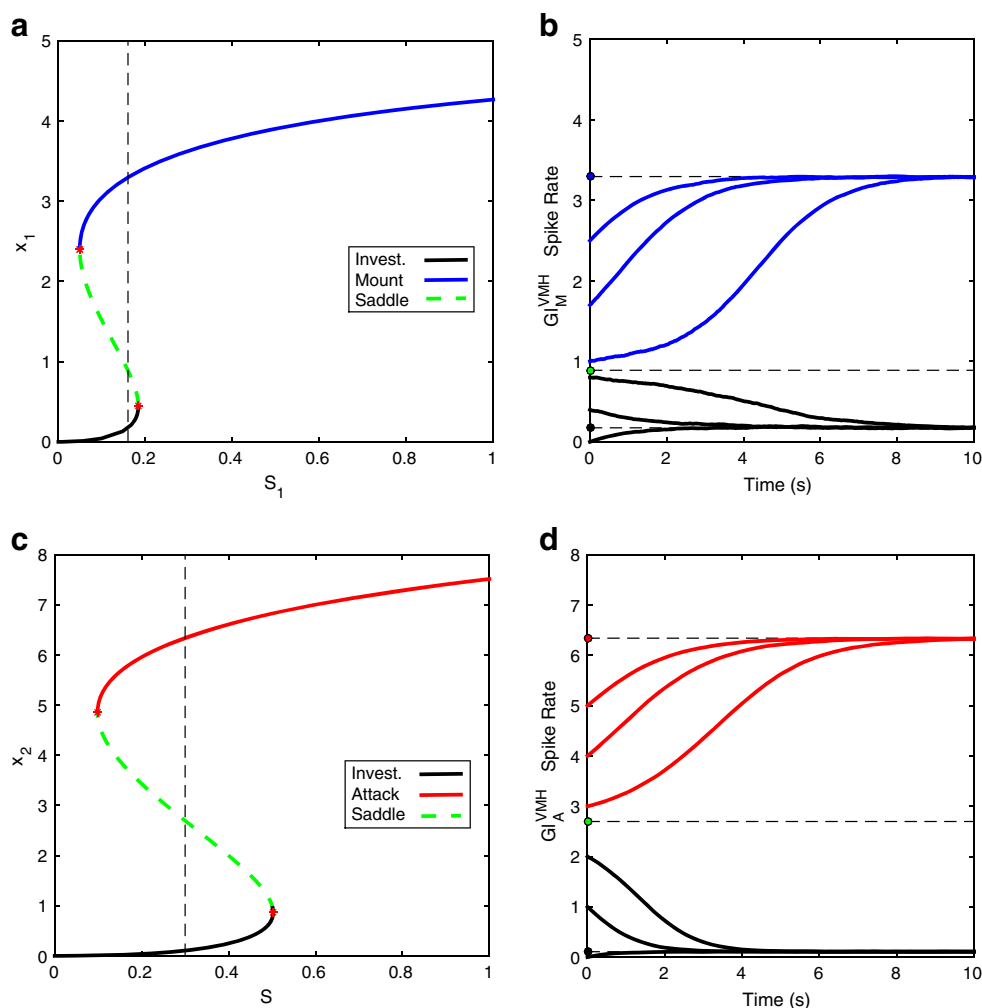
Figure 2a also shows that the system is bistable when $S_1 \in [0.049, 0.184]$. It means that close investigation and mounting are elicited depending on the level of activity of Gl_M. The vertical dashed line exhibits a transition from close investigation to mounting when the level of activity of Gl_M is greater than 0.889, which is the saddle point as shown in Fig. 2b, and for an stimulus value given by $S_1 = 0.16$.

On the other hand, in a normal male-male encounter the model shows that only Gl_A is activated and just two different states may be presented, close investigation or attack. In the same way, close investigation and attack are represented by stable equilibrium points shown in Fig. 2c. Black line represents social or close investigation states whereas the red line represents attack behavior. Hence, only attack is promoted when $S_2 > 0.502$, regardless of the level of activity of Gl_A.

The system is bistable when $S_2 \in (0.099, 0.502)$ and consequently, close investigation and attack are elicited depending on the level of activity of Gl_A. The vertical dashed line exhibits a transition from close investigation to attack when the level of activity of Gl_A is greater than 2.696

as shown in Fig. 2d, and for an stimulus value given by $S_2 = 0.3$.

From this bifurcation analysis, we conclude that our model promotes transitions from close investigation to mounting in male-female encounters and transitions from close investigation to attack in male-male encounters. Additionally, a greater state of arousal or motivation is needed for promoting attack in a male-male encounter than for mounting in a male-female encounter. Moreover, depending on the value of the state of arousal or motivation, a greater level of activity in VMHvl is required to shift from close investigation to attack in a male-male encounter than to shift from close investigation to mounting in a male-female encounter. For inputs close to the threshold, trajectories converge slowly, causing time-delay.

In order to test whether Gl_A and Gl_M neurons are able to inhibit mounting and attack, respectively, we performed two simulations. These simulations were run for 20 s and shown in Fig. 3.

On the one hand, Fig. 3a shows, at the top, that mounting does not present changes when the stimulus $S_1 = 0.5$
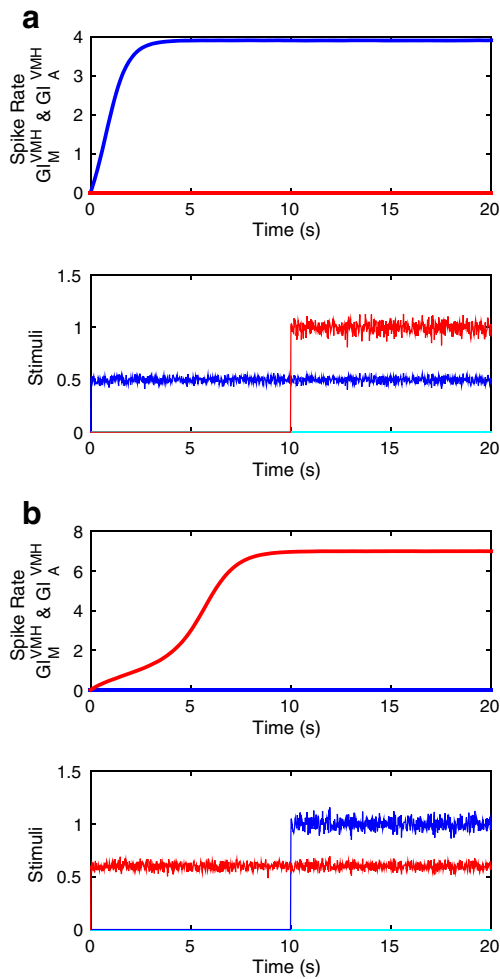
Fig. 3 Time-Mean Firing Rate diagrams. At the *top blue and red lines* represent, respectively, Gl_M and Gl_A neuronal activities. At the *bottom blue and red signals* represent, respectively, S_1 and S_2 stimuli. **a** With $S_1 = 0.5$ Gl_M neuronal activity increases to settle on mounting and does not present changes once $S_2 = 1$ is applied. **b** With $S_2 = 0.6$ Gl_A neuronal activity increases to settle on attack and does not present changes once $S_1 = 1$ is applied

is applied at zero time while $S_2 = 1$ is applied 10 s later (bottom of the figure).

On the other hand, Fig. 3b shows that attack does not present changes when the stimulus $S_2 = 0.6$ is applied at zero time while $S_1 = 1$ is applied 10 s later.

These results show that the model is able to simulate either CI/mounting or attack behaviors but does not present opponent control between mounting and attack when $S_1, S_2 \in [0, 1]$. The opponent control was just possible when the difference between both values ($|S_1 - S_2|$) is great enough. For instance, with $S_1 = 0.6$, once the system has entered a mounting mode, it began to be reversed by an input $S_2 = 2.6$ and then attack was evoked. While at $S_2 = 0.6$ the attack mode was just reversed by a much stronger input, $S_1 = 14$.

3.2 Network performance in male social behavior when artificially stimulated

The following numerical results from the model provide a different role for VMHvl in male mouse when optogenetically stimulated.

Based on unexpected experimental results following optogenetic stimulation of $Esr1^+$ neurons, we proposed the following changes in the model when the stimulus value L is provided: once L is delivered, S_1 and S_2 values go to zero and σ_2 goes to 5.

The annihilation of S_1 and S_2 may be justified by the fact that mounting was evoked toward normal and castrated males as well as females depending on photostimulation
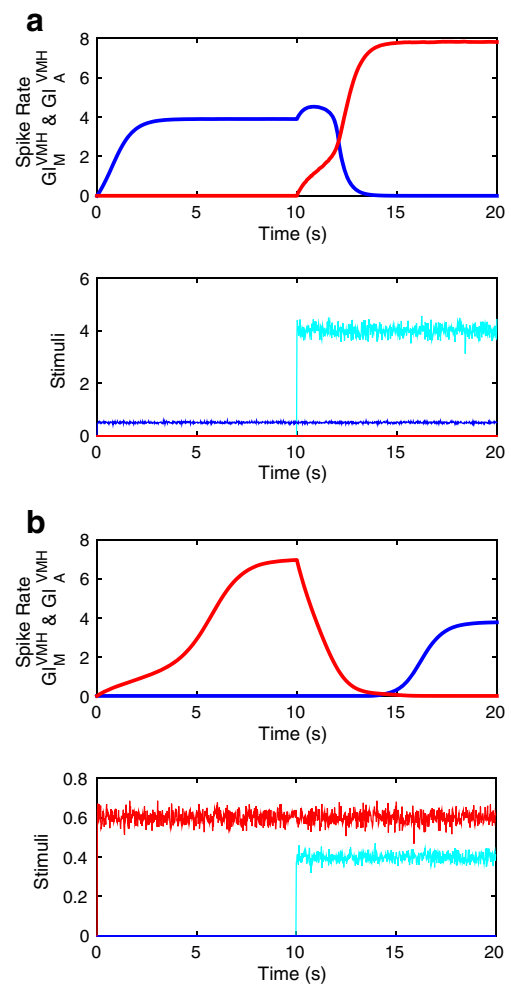
Fig. 4 Time-Mean Firing Rate diagrams. At the *top* of each figure *blue and red lines* represent, respectively, Gl_M and Gl_A neuronal activities. At the *bottom blue*, *red* and *cyan signals* represent, respectively, S_1, S_2, and L stimuli. **a** With $S_1 = 0.5$ Gl_M neuronal activity increases to settle on mounting but once $L = 4$ is applied, Gl_A neuronal activity starts increasing until Gl_M neuronal activity disappears and attack is elicited. **b** With $S_2 = 0.6$ Gl_A neuronal activity increases to settle on attack but when $L = 0.4$ is applied, it starts decreasing while Gl_M neuronal activity increases and mounting is elicited

intensity. The light stimulation acts as if the natural stimuli were erased or switched off. The molecular mechanism involved is still unclear. Moreover, optogenetic activation produced offensive attacks directed toward males, females, and inanimate objects. On the other hand, an increase of the threshold (σ_2) is needed since photostimulation activates both Gl_M and Gl_A at the same time. It might be explained by the fact that optogenetic activation is applied to targeted regions, ruling out the possibility of activation of undesired regions.

Figures 4 and 5 display time-mean firing rate diagrams where the photostimulation effect on model's performance is observed. Stimuli S_1 and S_2 are introduce at zero time while L is applied 10s later.
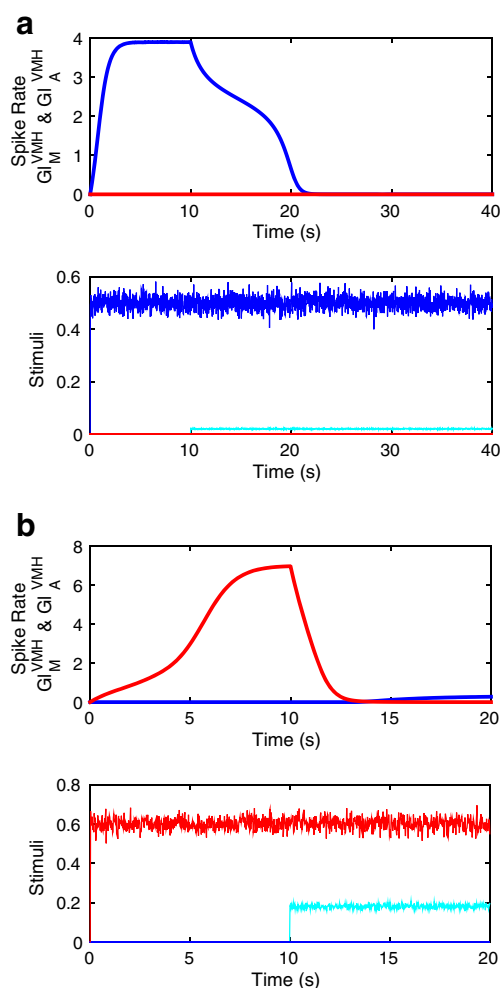
Fig. 5 Time-Mean Firing Rate diagrams. At the top of each figure *blue* and *red lines* represent, respectively, Gl_M and Gl_A neuronal activities. At the *bottom blue*, *red* and *cyan signals* represent, respectively, S_1, S_2, and L stimuli. **a** With $S_1 = 0.5$ Gl_M neuronal activity increases to settle on mounting but once $L = 0.02$ is applied, Gl_M neuronal activity starts to disappear until close investigation is elicited. **b** With $S_2 = 0.6$ Gl_A neuronal activity increases to settle on attack but when $L = 0.18$ is applied, it starts decreasing and close investigation is evoked

In Fig. 4a, for a set of inputs given by $S_1 = 0.5$, $S_2 = 0$, and $L = 4$, Gl_M neuronal activity increases to settle on mounting but once photostimulation is applied, Gl_A neuronal activity starts increasing up to its threshold and then exert a strong inhibition to Gl_M until its neuronal activity disappears and attack is elicited. This result could simulate an offensive attack directed toward females at high levels of optogenetic stimulation.
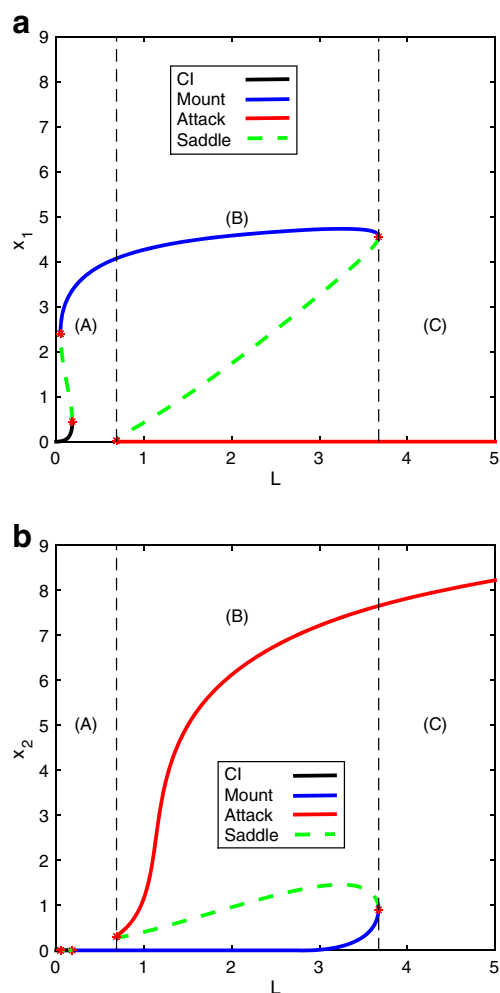
Fig. 6 Bifurcation diagrams, (**a**) x_1 vs L and (**b**) x_2 vs L, of the model showing stimulus-response curves when L runs over the interval [0,5]. *Solid* and *green dashed lines* correspond to the stable and unstable equilibrium points, respectively. *Red asterisks* show the bifurcation points where the system switches from a system with one stable state to a system with three equilibrium points and then to one with one stable state again. Bifurcation points divide the bifurcation interval in three main zones labeled as (**A**): CI/mounting, (**B**): mixed behavior, mounting and attack, and (**C**): attack when L runs, respectively, in the intervals (0, 0.687), (0.687, 3.675), and (3.675, 5]. *Black*, *blue*, and *red lines* represent equilibrium points when the system settles on close investigation, mounting, and attack respectively. *Green dashed lines* represent saddle points which separate the coexisting stable equilibria within the bistable zone

Now, considering $S_1 = 0$, $S_2 = 0.6$, and $L = 0.4$, Fig. 4b illustrates that Gl_A neuronal activity increases to settle on attack but once photostimulation is delivered it starts decreasing while Gl_M neuronal activity is ramping up to promote mounting. This result could simulate an attempted mounting toward males at low levels of optogenetic stimulation.

In Fig. 4a the switch from mounting (x_1 dominance) to attack (x_2 dominance) occurs at a high light stimulation ($L = 4$). The postsynaptic potential or net input reaches and surpasses σ_2 in Gl_A cells since w_{12} is greater than w_{21} and attack is elicited. In Fig. 4b, the switch from attack (x_2 dominance) to mounting (x_1 dominance) occurs at a low

light stimulation ($L = 0.4$) since activity in Gl_M is ramped up as σ_1 is smaller than σ_2 and Gl_A neurons do not reach activation.

When $S_1 = 0.5$, $S_2 = 0$, and $L = 0.02$, Fig. 5a shows that mounting is evoked but when L is applied, Gl_M neuronal activity decreases until it almost disappears and settles on close investigation.

Finally, Fig. 5b points out that for a set of inputs given by $S_1 = 0$, $S_2 = 0.6$, and $L = 0.18$, Gl_A neuronal activity increases to settle on attack but when photostimulation is delivered it starts decreasing until it disappears while Gl_M neuronal activity increases slightly to promote close investigation.
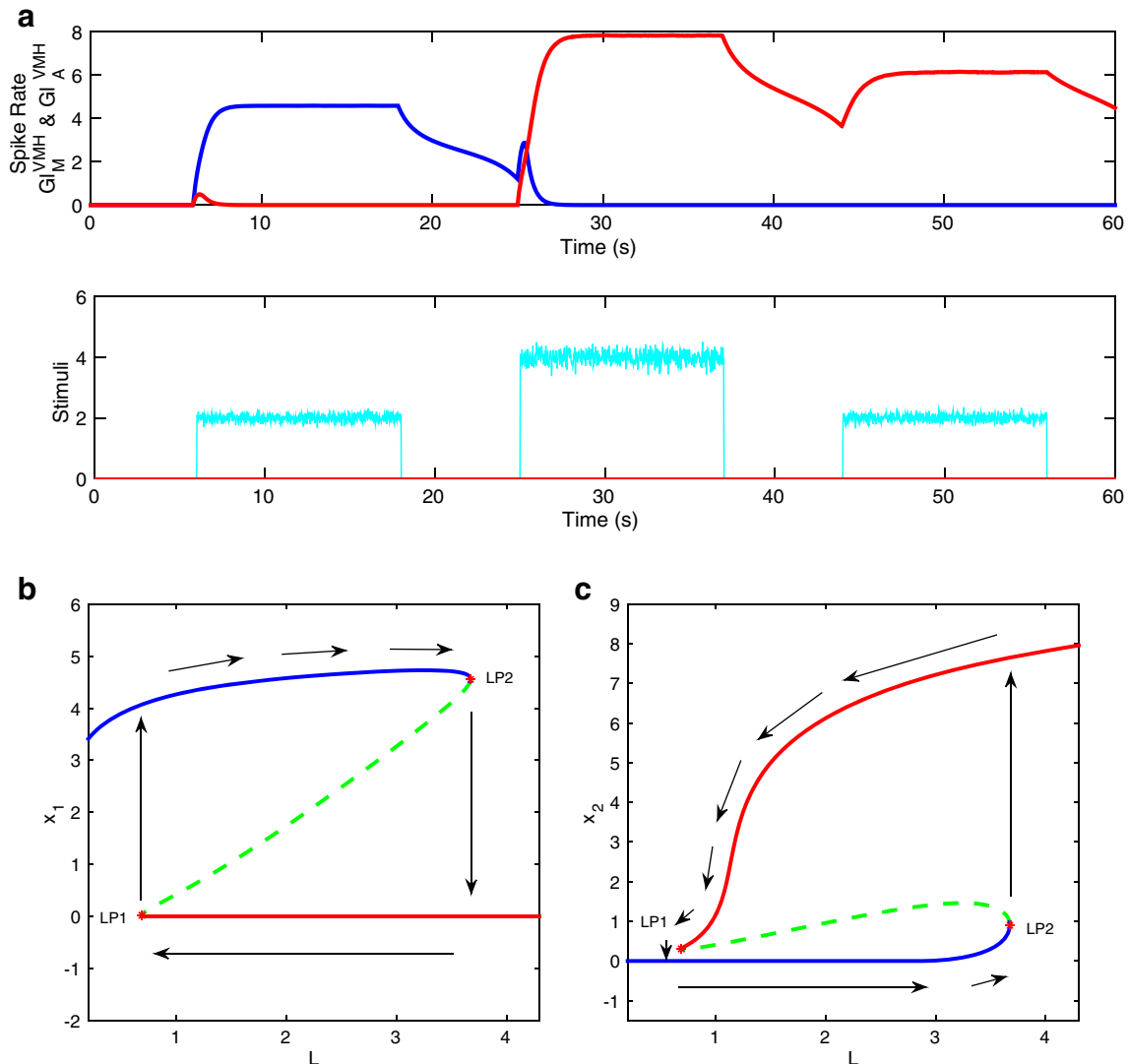
Fig. 7 Time-Mean Firing Rate diagrams. Typical hysteresis phenomenon. **a** At the *top*, Gl_M and Gl_A neuronal activities of the model show a transition from mounting to attack and once attack was elicited, reducing the stimulus intensity no longer evoked mounting, but simply promoted attack. At the *bottom*, three 12-s pulses representing the stimulus intensity with amplitudes 2.0, 4.0, and 2.0 respectively and seven-seconds-inter-pulse periods. **b-c** Hysteresis loop and bifurcations LP1 and LP2 in the presence of stimulus L. Between LP1 and LP2 two steady states are asymptotically stable nodes, while an unstable saddle point is in between. If L is swept back and forth across range LP1−LP2, the system will trace out the hysteresis loop shown by the arrows

So far, the model proposed is able to simulate the opponent control between mounting and attack in a scalable manner depending on the photostimulation intensity. These results are consistent with the optogenetic experiments reported in (Lin et al. 2011; Lee et al. 2014).

3.3 Numerical bifurcation analysis

In order to find the whole dynamics of the model when optogenetically stimulated, we performed a bifurcation analysis running the bifurcation parameter L over the interval [0,5]. Fig. 6 illustrates bifurcation diagrams of the system for variables x_1 and x_2.

In these diagrams, branches of stable points are indicated by solid curves and unstable points by the dashed part of the curve. Red asterisks show cod-1 bifurcation points called limit points (LP) or fold.

LP points indicate where the system switches from a monostable system to a bistable system and then to a monostable system again. Thus, the LP bifurcation, also called saddle-node, explains disappearance or appearance of a new stable state depending on the direction of movement of the bifurcation parameter. In any way, the qualitative behavior of the system changes exactly at the saddle-node.
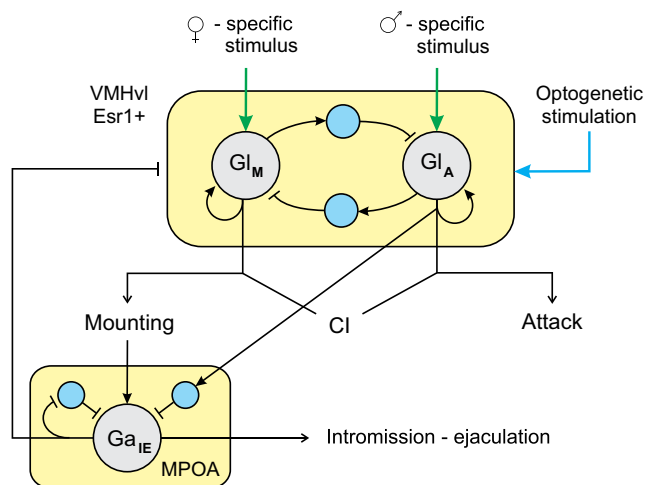
Bifurcation points occur when $L \approx 0.049$, $L \approx 0.184$, $L \approx 0.687$, and $L \approx 3.675$, which divide the bifurcation interval in three main zones of social behaviors.

Therefore, when we consider the photostimulation intensity in the interval $(0, 0.687)$, the system always presents CI/mounting behavior. A photostimulation intensity in the
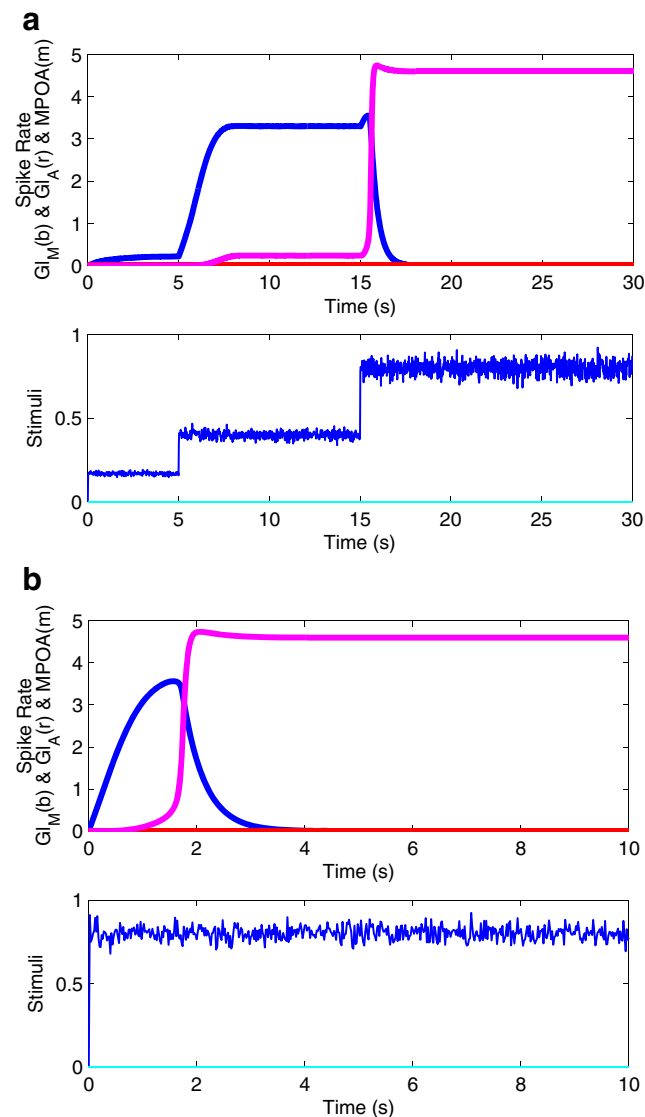
interval $(0.687, 3.675)$ leads the system to a mixed mounting and attack behavior, depending on initial conditions, and only attack is evoked when L is considered in the interval $(3.675, 5]$ regardless of the initial condition.

It is worth noting that a different qualitative change for greater values of L was not observed. It means that only fighting behavior is evoked for $L > 5$.

On the other hand, blue lines in bifurcation diagrams (Fig. 6) represent equilibrium points where the system is settled on mounting. We might consider close investigation

Fig. 9 Time-Mean Firing Rate diagrams. At the *top* of each figure *blue* and *magenta lines* represent, respectively, GI_M and Ga_{IE} neuronal activities. At the *bottom blue signal* represents the stimulus S_1. **a** GI_M and Ga_{IE} neuronal activities exhibit transitions from close investigation to mounting and then to intromission when the stimulus S_1 changes from 0.17 to 0.4 and then to 0.8 **b** For $S_1 = 0.8$ the model promotes intromission behavior showing it as a transition from mounting

Fig. 8 Possible circuitry underlying the interactions among CI (close investigation), mounting, attack, and intromission in VMHvl and MPOA. GI_M represents the mount-specific sub-population, GI_A represents the attack-specific sub-population, and Ga_{IE} represents the intromission-ejaculation-specific population. *Blue circles* are inhibitory interneurons populations. *Lines ending in a bar* are inhibitory synapses and those ending in *arrows* are excitatory synapses

behavior for lower stimulus intensity, e.g when $0 < L < 0.687$. It could be connected with early stages of a social encounters, for instance, when a resident male exhibits close investigation of the intruder with approaching, anogenital and head sniffing or staying together (Zou et al. 2008; Lee et al. 2014). Red lines represent equilibrium points when the system is settled on attack.

Additionally, the mixed zone (B), in Fig. 6, displays a "Z" shape, Fig. 6a, and an "S" shape, Fig. 6b, each of them composed by two coexisting stable equilibrium branches and one unstable equilibrium branch, green dashed line, in the middle. The saddle-node bifurcation connects the stable and unstable branches. These shapes are typical in bistable systems where the initial conditions play an important role depending on the basin of attraction of the attractors.

The steady-state shows a mounting behavior when the system's solution starts in the attraction domain of the attractor representing mounting. Similar explanation for attack, it means, the steady state shows an attack behavior when the system's solution starts in the attraction domain of the attractor representing attack.

Based on Fig. 6, we hypothesize that the model is adequate for capturing the dynamics of social encounters in male mice, dynamics observed in (Lin et al. 2011; Lee et al. 2014). For example, at low levels of photostimulation intensity our model exhibits mounting behavior rather than attack, as well as close investigation. When increasing the photostimulation intensity the model promotes a transition from mounting to attack, as well as an increase in the average level of the mean firing rate in VMHvl nucleus.

Figure 6 also shows that in most cases, attack requires activation of a larger number of neurons than those related with mounting. Thereby this model performs a scalable control of mounting and attack by certain neurons in the ventromedial hypothalamus, within a single social encounter as found in (Lee et al. 2014).

An interesting phenomenon was also found when we considered different values of L. In this case, increasing the photostimulation intensity promoted a transition from mounting to attack. Once attack was elicited, reducing the stimulus intensity no longer evoked mounting, but simply promoted attack. This phenomenon is called "hysteresis" and may be observed in Fig. 7.

Figure 7a at the top shows the evolution of the system for one minute using three 12-s pulses as stimuli (at the bottom). The first pulse with amplitude $L = 2.0$ is presented at $t = 6$ s, a second pulse with amplitude $L = 4$ is delivered 7s after the end of pulse 1, and the third pulse with amplitude $L = 2.0$ is presented 7s after the end of pulse 2. Over pulse 1 the network showed mounting behavior, once pulse 2 is delivered the network switches to attack and when the pulse 3 was presented with the same amplitude of the first pulse, the system did not evoke mounting but attack again.

Therefore, hysteresis phenomenon means that the present state of our neural network is determined not only by the present state of stimulation but also by the history of stimulation. This is mathematically explained with the initial conditions. In the previous simulation when the first pulse was delivered the model was in the resting state, but when the third pulse was presented the state at that moment had a different initial condition, not the resting state. It therefore seems that that initial condition, approx. (0,3.648), is within the attraction domain of the attractor representing attack when $L = 2.0$.
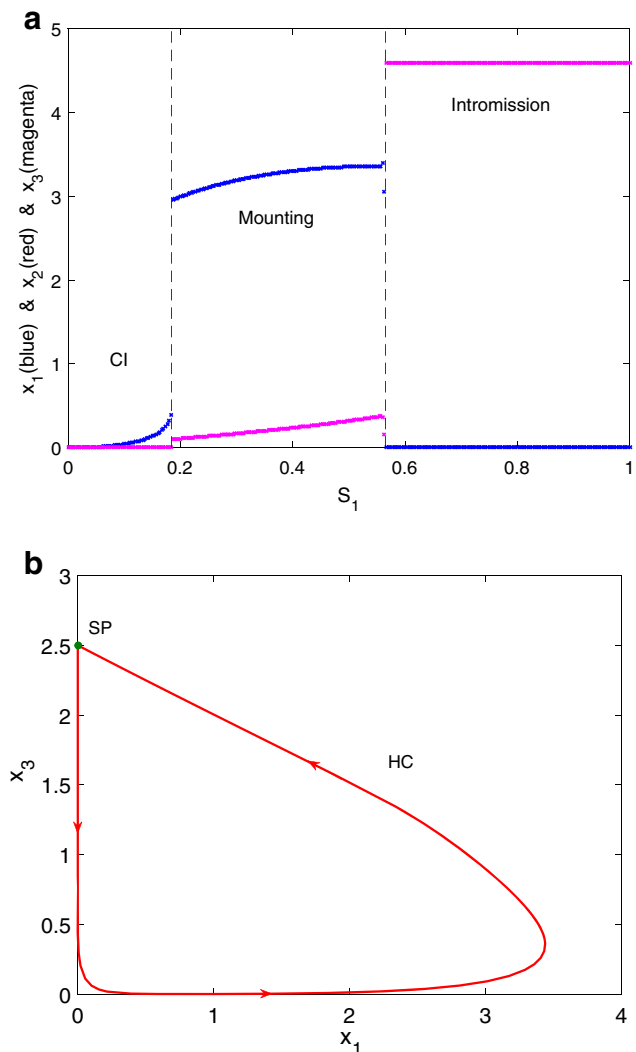

Fig. 10 **a** S_1-response curves. *Vertical dashed lines* point out values where the system switches from one state to another. These values divide the bifurcation interval in three main zones: CI (close investigation), mounting, and intromission when S_1 runs, respectively, in the intervals $(0, 0.184)$, $(0.184, 0.565)$, and $(0.565, 1]$. *Blue and magenta lines* correspond to the stable equilibrium points, starting from the resting state, when the system settles on CI/mounting and intromission respectively. **b** $x_1 - x_3$ phase portrait shows the homoclinic orbit (HC) and the saddle point (SP) that cause the SHC bifurcation (where the transition from mounting to intromission occurs)

Additionally, if L is varied slowly back and forth across the range (0.687, 3.675), the neural response will trace out a loop shown by the arrows in Fig. 7b,c which is known as a hysteresis loop. Therefore, if we start stimulating the system with $L > 3.675$ and decrease the stimulation intensity during the attack phase, x_1 in (a) (x_2 in (b)) slides along the lower (upper) branch to the left. When it reaches the end of the branch, LP1, it jumps (falls downward) to the upper (lower) branch corresponding to mounting, and slides along this branch to the right. When it reaches the right end, LP2, of the upper (lower) branch, it falls (jumps) to the lower (upper) branch, and thereby closes the hysteresis loop.

Fig. 11 Time-Mean Firing Rate diagrams showing the model's performance when artificially stimulated. At the *top* of each figure *blue, red, and magenta lines* represent, respectively, GI_M, GI_A, and Ga_{IE} neuronal activities. At the *bottom blue, red, and cyan signals* represents the stimulus S_1, S_2, and L respectively
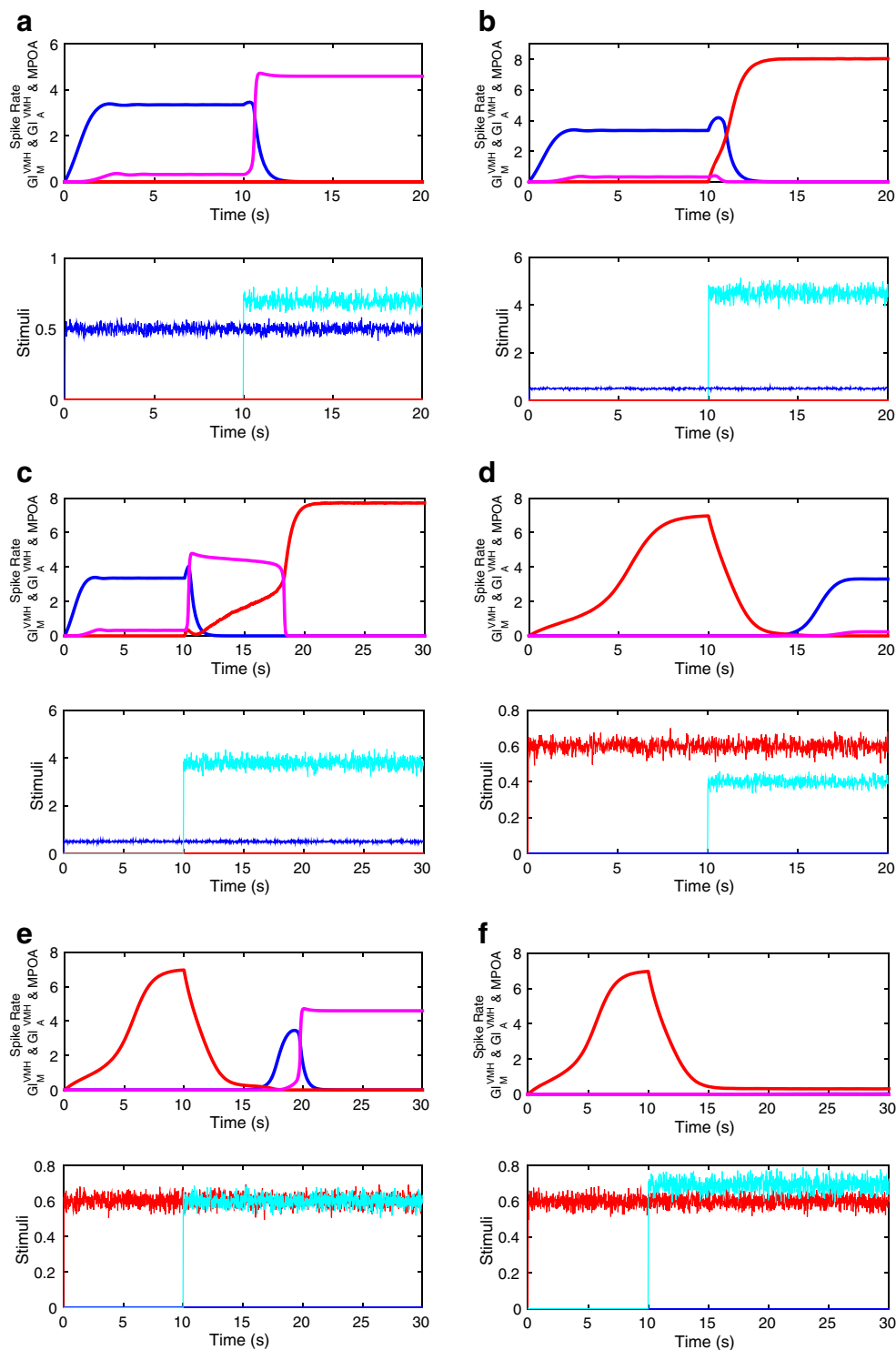
The hysteresis phenomenon was often exhibited during some of the experiments reported in (Lee et al. 2014). Further, they showed a representative raster plot as an example of hysteresis, illustrating a shift from mounting to attack while increasing photostimulation intensity. There, once attack was elicited during a single trial, reducing the photostimulation intensity back to a lower intensity no longer evoked mounting, but simply failed to elicit attack (Extended Data Fig. 9. (Lee et al. 2014)). In Fig. 7, our model portrays the hysteresis phenomenon observed as a behavioral feature as reported in (Lee et al. 2014).

3.4 A more complex neural network

Optogenetic studies (Lee et al. 2014; Lin et al. 2011) have also evidenced that optogenetic inhibition of Esr1$^+$ neurons in the VMHvl did not interrupt ongoing male-female mounting. Additionally, the average activity during a male-female interaction is suppressed during subsequent sexual behaviors, intromission and ejaculation.

Since MPOA may promote the consummatory phase of a sexual encounter (Dominguez and Hull 2005; Greenberg and Trainor 2016) and considering the results obtained in (Lee et al. 2014; Lin et al. 2011), we proposed a population of inhibitory neurons in MPOA, denoted Ga$_{IE}$ (Fig. 8). This speculative scenario was proposed in (Kennedy et al. 2015).

The dynamics for the new network is modeled by the dynamical system given by Eqs. (4)–(6).

$$\tau_1 \frac{dx_1}{dt} = -x_1 + f_1(S_1 + w_{11}x_1 - w_{12}x_2 - w_{13}x_3 + L) \quad (4)$$

$$\tau_2 \frac{dx_2}{dt} = -x_2 + f_2(S_2 + w_{22}x_2 - w_{21}x_1 - w_{23}x_3 + L) \quad (5)$$

$$\tau_3 \frac{dx_3}{dt} = -x_3 + f_3(w_{33}x_3 + w_{31}x_1 - w_{32}x_2) \quad (6)$$

Now, x_3 represents the level of activity of the Ga$_{IE}$ circuit that may promote the consummatory phase of a sexual encounter. The new parameters for this model are given by $\tau_3 = 0.05$, $w_{13} = 1$, $w_{23} = 0.6$, $w_{33} = 2.8$, $w_{31} = 0.8$, $w_{32} = 0.5$, $m_3 = 5$, $\sigma_3 = 7$, and $N_3 = 4$.

3.5 Network performance of the new model in normal male-female social behavior

Similarly to the case without MPOA, we executed simulations of the neural circuits when they are activated by specific stimuli coming from females or males in absence of photostimulation.

Figure 9 displays time-mean firing rate diagrams showing the model's performance for different values of S_1. Figure 9a exhibits transitions from close investigation to mounting and then to intromission when the stimulus S_1

changes from 0.17 to 0.4 and then to 0.8 as pointed out at the figure's bottom.

Figure 9b shows the transition from mounting to intromission at $S_1 = 0.8$. In this case the dominance of the mounting circuit over the intromission circuit begins to be reversed, after 2 seconds, when the inhibition exerted from the Ga$_{IE}$ circuit to the mounting circuit is stronger than the excitation over the Ga$_{IE}$ circuit coming from the mounting circuit.
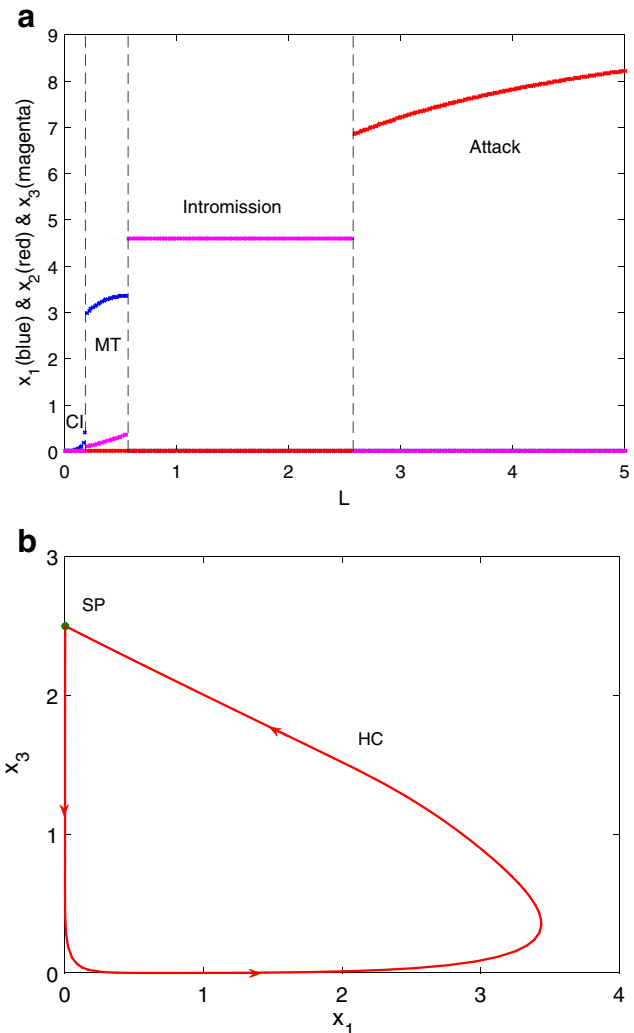

Fig. 12 **a** L-response curves. *Vertical dashed lines* point out points where the system switches from one state to another. These points divide the bifurcation interval in four main zones: CI (close investigation), MT (mounting), intromission, and attack when L runs, respectively, in the intervals (0, 0.184), (0.184, 0.565), (0.565, 2.575), and (2.575, 5]. *Blue, magenta, and red lines* correspond to the stable equilibrium points, starting from the resting state, when the system settles on CI/mounting, intromission, and attack respectively. **b** $x_1 - x_3$ phase portrait shows the homoclinic orbit (HC) and the saddle point (SP) that cause the SHC bifurcation (where the transition from mounting to intromission occurs)

In Fig. 10a, the stable steady states of Gl_M and Ga_{IE} are illustrated. In a normal male-female encounter, this new model shows that three different states may be presented, either close investigation, mounting, or intromission as the parameter S_1 is varied. These states are represented by stable equilibrium points. Blue lines represent the level of activity of the Gl_M circuit deciding between close investigation or mounting whereas the magenta line represents the level of activity of Ga_{IE} circuit that triggers intromission-ejaculation.

Therefore, from the resting state the model may switch as the stimulus increases, from close investigation to mounting and then to intromission as shown from the $x_1 - x_3$ phase portrait in Fig 10a. Thus, close investigation is promoted when $S_1 < 0.184$, mounting is promoted when $0.184 < S_1 < 0.565$, and $0.565 < S_1 \leq 1$ is the interval where intromission is evoked.

Similarly, as in the previous model, the transition from close investigation to mounting occurs at $S_1 = 0.184$ where the system undergoes a LP bifurcation. However,
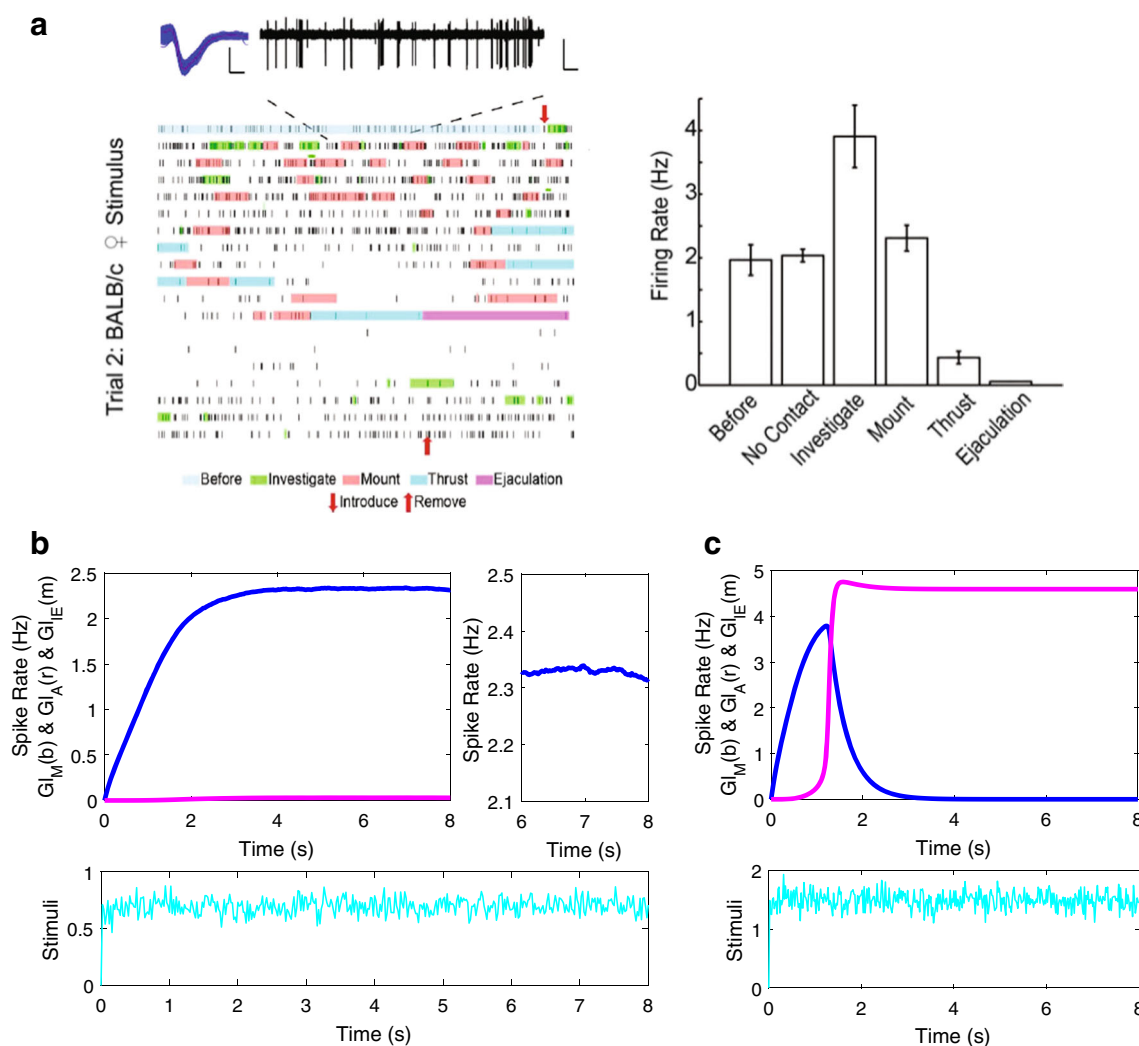
Fig. 13 **a** A sample raster plot of a VMHvl cell response to a female during mating and a histogram showing the average firing rate during mounting and ejaculation in addition to other stages (Adapted by permission from Macmillan Publishers Ltd: Nature. (Lin et al. 2011), Supplementary Fig. S9, copyright 2011). **b-c** At the *top*, Gl_M and Ga_{IE} neuronal activities of the model. At the *bottom cyan signal* represents the L stimulus. **b** With $L = 0.7$ and a maximum mean firing rate of Gl_M given by $m_1 = 3.5$, Gl_M neuronal activity increases to settle on mounting with a mean firing rate ≈ 2.33Hz (the activity signal

expanded in the two last seconds is illustrated at the *right upper corner*), a similar firing rate value compared against the recording data (shown in (a)). **c** With $L = 1.5$ Gl_M neuronal activity increases and mounting is briefly presented just before intromission-ejaculation is promoted. The transition from mounting to intromission-ejaculation shows a strong inhibition to Gl_M until its neuronal activity disappears, a similar firing rate value compared against the recording data (shown in (a))

the transition from mounting to intromission is caused by a saddle homoclinic bifurcation (SHC) which occurs at $S_1 \approx 0.56505$. The homoclinic orbit (HC) and the saddle point (SP) that cause the SHC bifurcation are shown in Fig. 10b.

3.6 New model's performance in male social behavior when artificially stimulated

Now, we present the model's performance and a similar bifurcation analysis when the system is optogenetically stimulated.

Figure 11 displays time-mean firing rate diagrams showing the photostimulation effect on model's performance when the Ga_{IE} neuronal activity is considered. Stimuli S_1 and S_2 are introduce at zero time while L is applied 10s later.

For a set of inputs given by $S_1 = 0.5$ and $S_2 = 0$, Gl_M neuronal activity increases to settle on mounting during the first 10 seconds, as shown in Fig. 11a–c. Once photostimulation is applied for different values of L, the system switches to intromission or attack behavior. When $L = 0.7$

it switches to intromission (Fig. 11a), at $L = 4.5$ the system switches to attack (Fig. 11b). On the other hand, at $L = 3.8$ the system initially switches to intromission and after 8 seconds the system switches to attack (Fig. 11c).

Now, at $S_1 = 0$ and $S_2 = 0.6$, Gl_A neuronal activity ramps up to settle on attack during the first 10 seconds, as shown in Fig. 11d–f. When photostimulation is presented, the system changes to either mounting, intromission or close investigation, depending on the stimulus value L. For instance, at $L = 0.4$, Fig. 11d displays a transition to mounting. When $L = 0.6$, Fig. 11e shows that not only intromission is promoted but also mounting is briefly presented just before intromission, which corresponds to natural sexual behavior. Finally, at $L = 0.7$, a transition to close investigation is evoked (Fig. 11f).

Regarding the bifurcation analysis as the parameter L is varied, Fig. 12a shows Gl_M, Gl_A, and Ga_{IE} neuronal activities' stable steady states.

Blue lines represent the level of activity of Gl_M circuit deciding between close investigation or mounting, the magenta line represents the level of activity of Ga_{IE} circuit

Fig. 14 a A sample raster plot of a VMHvl cell response to a male during social behaviors and a histogram showing the average firing rate during attack in addition to other stages (Adapted by permission from Macmillan Publishers Ltd: Nature. Lin et al. (2011), Supplementary Fig. S8, copyright 2011). b At the *top*, Gl_M and Gl_A neuronal activities of the model. At the *bottom cyan signal* represents the L stimulus. With $L = 5$ and a maximum mean firing rate of Gl_A given by $m_2 = 15$, Gl_A neuronal activity increases to settle on attack with a mean firing rate ≈ 13.4Hz (the activity signal expanded in the two last seconds is illustrated at the *right upper corner*), a similar firing rate value compared against the recording data (shown in (a))
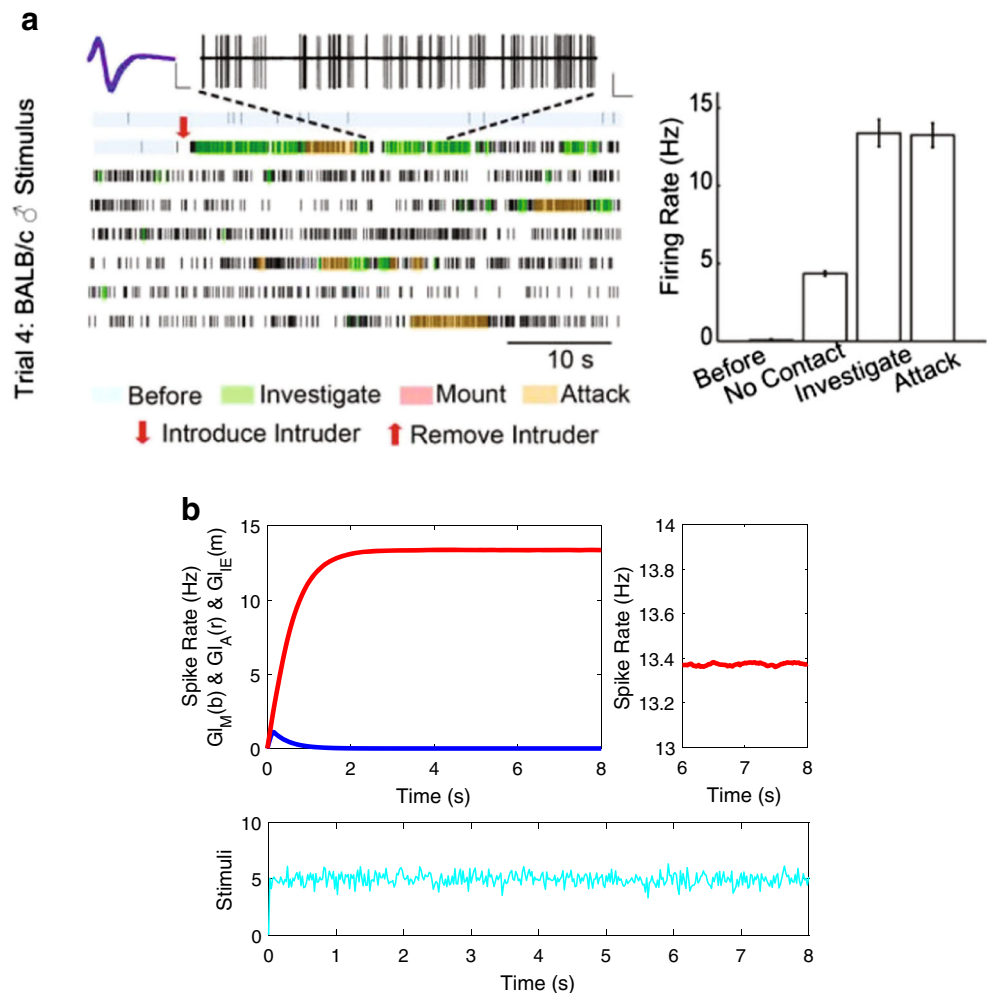
deciding intromission. The level of activity of Gl$_A$ circuit which decides attack is represented by red lines.

Relevant bifurcation points occur at $L \approx 0.184$ (LP), where the system switches from close investigation to mounting, and at $L \approx 0.5655$ (SHC, shown in Fig. 12b), where the system switches from mounting to intromission. The transition from intromission to attack occurs at $L \approx 2.575$. At this value, the stable manifold, corresponding to the saddle point, passes through the origin and then the basin of attraction of intromission behavior no longer contains the resting state, but it is in the basin of attraction of the attack behavior instead. However, if $L > 3.702$ attack may be evoked regardless of the initial condition. At $L \approx 3.702$ the system undergoes a LP bifurcation.

4 Conclusions and discussion

We have reached full agreement with the data collected by David Anderson's Lab Team and reported by Lee et al. (2014). Our model yields not only the behaviors evoked by optogenetic stimulation and natural stimuli but also exhibits the hysteresis phenomenon as an emergent dynamical feature not foreseen from the neural circuitry proposed.

We took into consideration that there exist strong evidence that cells activated during aggression are enriched with Esr1$^+$, as mentioned by Walter et al. (1985).

The photostimulation activation of Esr1$^+$ neurons was confirmed *in vitro* using whole-cell patch-clamp recording in acute hypothalamic slices and *in vivo*, as well as by extracellular recordings (Lee et al. 2014).

Next, the more relevant experimental results replicated or approximated by our model and its most important emergent feature (hysteresis phenomenon):

1. The effect of optogenetic stimulation of VMHvl Esr1$^+$ neurons in resident males −over 87% of ChR2-expressing animals in circa 90% of trails− was shown eliciting intense, time-locked attack towards both castrated male and female intruders under infrared light stimulation (20 Hz, 30 s, 20 ms pulse-width).
2. Attack was initiated within $\sim 5s$ of photostimulation when light pulses were delivered while the resident was facing the intruder and within one mouse body-length.
3. Whereas male-directed mounting was only observed during photostimulation, its latency ($\sim 8 - 12s$) was longer than for attack ($\sim 5s$). Photostimulation-induced mounting towards male intruders was not observed in mCherry-expressing controls.
4. In some animals, as the intensity of stimulation was increased during a single trial, evoked behaviors could

be observed to switch from mounting only (with a photostimulation of 0.5 mW mm^{-2}), to mixed mounting and attack (with a photostimulation of 1.4 mW mm^{-2}) to attack only (with a photostimulation of 2.6 mW mm^{-2}), over 30 s. This effect often exhibited hysteresis and is mentioned below.

At the beginning, parameters for the model were set up (Table 1) in order to simulate the network performance when it is exhibiting both a normal male-female or male-male social encounter. However, when VMHvl Esr1+ neurons were experimentally photostimulated in our model, attack was elicited towards both male or female, at high levels of stimulation. Since mounting is evoked at low levels of light stimulation and attack is elicited at high levels, a high threshold σ_2 is needed in our model in order to describe an scalable control by light stimulation of mounting and attack as experimentally observed.

The experimental data exhibits that optogenetic activation surpasses natural stimuli in an overwhelming manner and can trigger unexpected behaviors as referred by Kennedy et al. (2015) "optogenetic stimulation experiments should be interpreted with caution, because artificial patterns of activation could produce abnormal behaviors". Therefore, this response is portrayed in our model when S_1 and S_2 go to zero when light is delivered.

We propose a neural circuit supporting the interactions between mounting and attack in order to explain the experimental data collected, and behavioral and neuroanatomical findings as well. This neural network was biologically inspired from recently findings by optogenetic manipulations within the male mouse hypothalamus as mentioned before.

The neural network was modeled by a nonlinear dynamical system based on attractor network models with dynamical stability. Although the network was inspired on experimental results using artificial stimulation, the network's behavior as a response to a natural encounter was also considered.

When artificial stimulation was considered, our results suggested a scenario for controlling social behaviors in a scalable manner, such as mounting and attack. The control of the transition between these social behaviors exhibits a sort of decision-making mechanism by which sub-populations of VMHvl neurons make decisions via steady state transitions. These transitions commit the animal to a given behavioral mode by just changing the amount of stimulation, i.e., the decision is made apparently by altering the level of activity of a single nucleus. The amount of stimulation required could be found when the model was examined with bifurcation analysis and computer simulations

The mathematical model proposed in Eqs. (1)–(3) includes parameters w_{ij}, σ_i and m_i that have different values for each neuron population. This fact, added to the mutual

inhibitory synaptic circuit suggested, allows the deployment of the social behaviors observed experimentally under photostimulation and in natural conditions in mice. The results obtained are in good agreement with the progression of a social interaction, from its appetitive through its consummatory phase, as seen in (Lee et al. 2014) and that a single nucleus is responsible for the entire behavioral progression.

We expanded our first model given in Eqs. (1)–(3) to a more complex model given by Eqs. (4)–(6), which takes into consideration the consummatory phase, namely, the intromission and ejaculation stages. Figures 10–12 show transitions between different behaviors under natural and artificial stimuli that result from our more complex model. These transitions describe the real dynamics observed and recorded in Lin's work (Lin et al. 2011). For certain values in some specific parameters, Figs. 13 and 14 show a close agreement with data collected experimentally in the work cited above.

The neural circuit proposed has followed the observations and discussions pointed out by Kennedy et al. (2015). Particularly, our model has taken the option of cellular heterogeneity instead of an intensity coding for explaining the role of VMHvl during a natural social encounter. The progression from close investigation to mounting or attack suggests a detection of sex-specific pheromonal clues at short range which led to consider sex-specific sub-populations among VMHvl in our model. The sub-populations suggested in VMHvl and their synaptic connections not only encode the intensity of an escalating state of arousal but also the coupling of this state to behavioral decisions via the attractor dynamics that appears accordingly. This mechanism theoretically explains how an escalating state of arousal or motivation produces or induces the transitions between different behaviors as observed under artificial stimulation or natural conditions.

Encoding the escalating state of arousal or motivation requires the addition in time of signals coming from closer and closer investigation of the intruder by the resident individual that results in an increasing detected level of sex-specific pheromones. The proposed VMHvl sub-populations responding to sex-specific clues must be able to trace and encode this addition in time for the signals detected during the close investigation or inspection. Our model provides this capacity through time-circuit constants τ_i that allow the proper time-window necessary for this escalating state encoding.

To our knowledge, there are not other proposed neural circuits and mathematical models addressing the synaptic connectivity and neural substrates details behind the scalable control of the described social behaviors in mice under natural and artificial stimulation. More experimental data have to be sought for the ongoing effort of completing this neural computational model.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

References

Boucheny, C., Brunel, N., & Arleo, A. (2005). A continuous attractor network model without recurrent excitation: maintenance and integration in the head direction cell system. *Journal of Computational Neuroscience*, *18*(2), 205–227. doi:10.1007/s10827-016-0621-9.

Choi, G.B., Dong, Hw., Murphy, A.J., Valenzuela, D.M., Yancopoulos, G.D., Swanson, L.W., & Anderson, D.J. (2005). Lhx6 delineates a pathway mediating innate reproductive behaviors from the amygdala to the hypothalamus. *Neuron*, *46*(4), 647–660.

Crews, D. (2005). Evolution of neuroendocrine mechanisms that regulate sexual behavior. *Trends in Endocrinology & Metabolism*, *16*(8), 354–361.

Dhooge, A., Govaerts, W., & Kuznetsov, Y.A. (2003). Matcont: a matlab package for numerical bifurcation analysis of odes. *ACM Transactions on Mathematical Software*, *29*(2), 141–164.

Dominguez, J.M., & Hull, E.M. (2005). Dopamine, the medial preoptic area, and male sexual behavior. *Physiology & Behavior*, *86*(3), 356–368.

Dulac, C., & Kimchi, T. (2007). Neural mechanisms underlying sex-specific behaviors in vertebrates. *Current Opinion in Neurobiology*, *17*(6), 675–683. motor systems / Neurobiology of behaviour.

Greenberg, G.D., & Trainor, B.C. (2016). Chapter 4 - sex differences in the social behavior network and mesolimbic dopamine system. In Shansky, R. M. (Ed.) *Sex differences in the central nervous system* (pp. 77–106). San Diego: Academic Press.

Hashikawa, K., Hashikawa, Y., Falkner, A., & Lin, D. (2016). The neural circuits of mating and fighting in male mice. *Current Opinion in Neurobiology*, *38*, 27–37.

Hurtado-López, J., & Ramirez-Moreno, D.F. (1350). (2013) Nonlinear dynamics of an oscillatory neural network acting as a motor central pattern generator. *International Journal of Bifurcation and Chaos*, *23*(08), 142.

Kennedy, A., Asahina, K., Hoopfer, E., Inagaki, H., Jung, Y., Lee, H., Remedios, R., & Anderson, D.J. (2015). Internal states and behavioral decision-making: Toward an integration of emotion and cognition. Cold Spring Harbor Laboratory Press, pp. 1–12.

Lee, H., Kim, D.W., Remedios, R., Anthony, T.E., Chang, A., Madisen, L., Zeng, H., & Anderson, D.J. (2014). Scalable control of mounting and attack by esr1+ neurons in the ventromedial hypothalamus. *Nature*, *509*(7502), 627–632.

Lin, D., Boyle, M.P., Dollar, P., Lee, H., Lein, E.S., Perona, P., & Anderson, D.J. (2011). Functional identification of an aggression locus in the mouse hypothalamus. *Nature*, *470*(7333), 221–226.

Milton, J., Naik, P., Chan, C., & Campbell, S.A. (2010). Indecision in neural decision making models. *Mathematical Modelling of Natural Phenomena*, *5*, 125–145.

Nelson, R.J., & Chiavegatto, S. (2000). Aggression in knockout mice. *ILAR Journal*, *41*(3), 153–162.

Nelson, R.J., & Trainor, B.C. (2007). Neural mechanisms of aggression. *Nature Reviews Neuroscience*, *8*(7), 536–546.

Sokolowski, K., & Corbin, J.G. (2012). Wired for behavior: from development to function of innate limbic system circuitry. *Frontiers in Molecular Neuroscience*, 5(55).

Swanson, L.W. (2000). Cerebral hemisphere regulation of motivated behavior. *Brain Research*, *886*(1–2), 113–164. towards 2010, A brain Odyssey, The 3rd Brain Research Interactive.

Tegnér, J., Compte, A., & Wang, X.J. (2002). The dynamical stability of reverberatory neural circuits. *Biological Cybernetics*, *87*(5-6), 471–481.

Trotta, L., Bullinger, E., & Sepulchre, R. (2012). Global analysis of dynamical decision-making models through local computation around the hidden saddle. *PLoS ONE*, 7(3).

Veenema, A.H., & Neumann, I.D. (2007). Neurobiological mechanisms of aggression and stress coping: A comparative study in mouse and rat selection lines. *Brain, Behavior and Evolution*, *70*(4), 274–285.

Walter, P., Green, S., Greene, G., Krust, A., Bornert, J.M., Jeltsch, J.M., Staub, A., Jensen, E., Scrace, G., & Waterfield, M. (1985). Cloning of the human estrogen receptor cdna. *Proceedings of the National Academy of Sciences*, *82*(23), 7889–7893.

Wang, C.T., Lee, C.T., Wang, X.J., & Lo, C.C. (2013). Top-down modulation on perceptual decision with balanced inhibition through feedforward and feedback inhibitory neurons. *PLoS ONE*, 8(4).

Wang, X.J. (2008). Decision making in recurrent neuronal circuits. *Neuron*, *60*(2), 215–234.

Wang, X.J. (2009). Attractor Network Models. In Squire, L.R. (Ed.) *Encyclopedia of Neuroscience* (pp. 667–679). Oxford: Academic Press.

Wang, X.J. (2012). Neural dynamics and circuit mechanisms of decision-making. *Current Opinion in Neurobiology*, *22*(6), 1039–1046. decision making.

Wilson, H.R., & Cowan, J.D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, *12*(1), 1–24.

Zou, H., Zhang, C., Xie, Q., Zhang, M., Shi, J., Jin, M., & Yu, L. (2008). Low dose MK-801 reduces social investigation in mice. *Pharmacology Biochemistry and Behavior*, *90*(4), 753–757.