DTU Library

Bifurcation of learning and structure formation in neuronal maps

Marschler, Christian; Faust-Ellsässer, Carmen; Starke, Jens; van Hemmen, J. Leo

[Link back to DTU Orbit](#)

Bifurcation of learning and structure formation in neuronal maps

Bifurcation of learning and structure formation in neuronal maps

Christian Marschler[1], Carmen Faust-Ellsässer[2], Jens Starke[1] and J. Leo van Hemmen[3]

[1] *Department of Applied Mathematics and Computer Science, Technical University of Denmark*
2800 Kongens Lyngby, Denmark
[2] *Interdisciplinary Center for Scientific Computing, University of Heidelberg - 69120 Heidelberg, Germany*
[3] *Physik Department & BCCN – Munich, Technische Universität München*
85747 Garching bei München, Germany

Abstract – Most learning processes in neuronal networks happen on a much longer time scale than that of the underlying neuronal dynamics. It is therefore useful to analyze slowly varying macroscopic order parameters to explore a network's learning ability. We study the synaptic learning process giving rise to map formation in the laminar nucleus of the barn owl's auditory system. Using equation-free methods, we perform a bifurcation analysis of spatio-temporal structure formation in the associated synaptic-weight matrix. This enables us to analyze learning as a bifurcation process and follow the unstable states as well. A simple time translation of the learning window function shifts the bifurcation point of structure formation and goes along with traveling waves in the map, without changing the animal's sound localization performance.

Introduction. – Combined with its gigantic dimension reduction, the notion of order parameter has turned out to be extremely useful. One of the oldest examples is the "magnetization" m of an Ising ferromagnet, $m \equiv N^{-1} \sum_{i=1}^{N} S_i$ with $S_i = \pm 1$ corresponding to spin up or down. It describes whether the system is in an up ($m > 0$) or down ($m < 0$) state below the critical temperature [1]. An order parameter such as m is a macroscopic characterization of a microscopic system state, and varies slowly in time as compared to its (very many) microscopic constituents. Here, we focus on an important additional advantage of a well-conceived order parameter in that we can now pursue dynamically *un*stable macroscopic states by employing recently developed multiscale methods (equation-free analysis; see, *e.g.*, [2]). One can therefore perform a complete bifurcation analysis.

Biological physics has meanwhile provided a plethora of systems that possess at least two different time scales, a fast one and a slow one, but the accompanying order-parameter choice has hardly attracted any attention and its bifurcation analysis even less. As a generic example, we analyze structure formation in the auditory system of the barn owl [3,4] and the way in which a map of azimuthal sound localization arises in the laminar nucleus. This is the first station where phase-locked signals from

the left and right ear come together to form a map as a consequence of synaptic learning [3–7]. For a review of early results on the relation between the neuronal code and sound perception in mammals instead of birds, a far less understood issue, the reader is referred to Eggermont [8].

In the present context, a map is a neuronal representation of the outside sensory world. What is actually mapped here is the interaural time difference (ITD) corresponding to the horizontal direction. Until now, all maps were static. As we will see, under the high-dimensional microscopic neuronal and synaptic dynamics a sound localization map arises that can be reduced to a bifurcation of a corresponding low-dimensional order parameter. The bifurcation analysis of the macroscopic dynamics is made possible by equation-free methods [2,9–11]. The surprise is both in the type of bifurcation and in the ensuing map that behaves as a slowly traveling wave.

In the auditory system of the barn owl, feeder axons come from the left and right ear (through the cochlear and magnocellular nucleus) and pass along to the laminar nucleus (NL), making a sharp turn when they enter it [12]. There they all have a practically uniform signal velocity of about 4 m/s (see, *e.g.*, [13]) and touch the laminar neurons in an intertwining manner. For our purpose, the latter can be imagined as a row of neurons; cf. fig. 1.
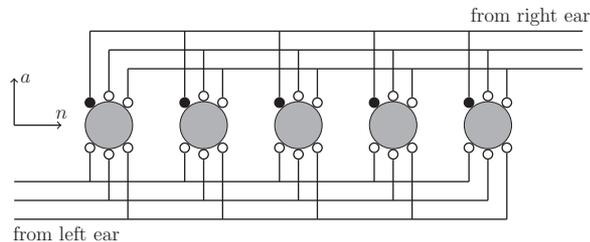
(a) NL model
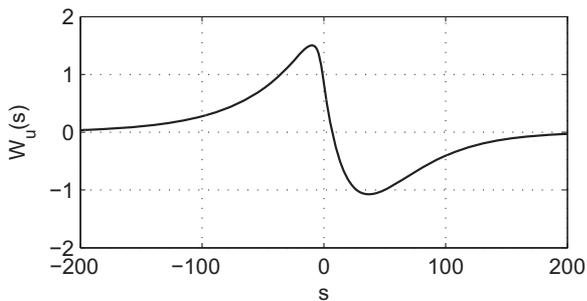
(b) Learning window

Fig. 1: (a) Sketch of the structural setup of the barn owl's laminar nucleus (NL), an essential part of the auditory system. The $1 \leq n \leq N$ neurons in a row (5 gray circles, in reality and in computations $N = 30$) are connected to $1 \leq a \leq A$ axonal fibers (3 from right and 3 from left). Axons from the left and right ear are connected to laminar neurons through synapses (small circles). Map formation needs a tuning of both the synapses on each neuron through the firing of the neuron they are on [5] and of the synapses on different neurons, which should "know" what the others do, so that a topographically ordered map can emerge [6]. The latter tuning is performed by means of axon-mediated spike-based learning (AMSL), here indicated by solid circles connected to *the same* axon; cf. eq. (6) below. (b) Learning window for $\hat{u} = -1$; cf. eq. (5).

The approximately uniform velocity is essential to map formation, the precise speed in the feeder axons is not.

Model of the barn owl's auditory system. – We adapt the NL model of Kempter *et al.* [6] with $N = 30$ neurons in a row connected to $A = 280$ axons, so that the neurons receive input from 140 axons from each ear. For each synapse there are two indices, the neurons $1 \leq n \leq N$ and the axons $1 \leq a \leq A$ it is connected to. The distance between neurons is given by $d = 27\,\mu$m and the propagation speed of the signals is taken [12] to be $c = 4\,$m/s. According to the finite propagation speed c of the signals, the time delay between neighboring neurons connected to the same axon is $\tilde{\Delta} = d/c = 6.75\,\mu$s. For all numerical simulations, time has been discretized in steps of $\Delta t = 5\,\mu$s, with time delays given in units of Δt. Due to a pre-processing in the brain [12], input signals along different axons arrive at the NL border at different times. Accordingly, a uniform distribution of delays $\Delta_a \in [0, 2T]$ has been chosen with

$$\Delta_a = \begin{cases} \left(\frac{2T}{A/2}\right)a, & 1 \leq a \leq \frac{A}{2} \quad \text{(left ear)}, \\ \left(\frac{2T}{A/2}\right)(a - \frac{A}{2}), & \frac{A}{2} + 1 \leq a \leq A \text{ (right ear)}, \end{cases} \quad (1)$$

and $T = A/4$ is the period of the input cochlear signal, a cochlear best frequency [12], covering two periods each ear. Combining both delay types, eq. (1) gives for synapse (a, n) connected to the left and right ear the total delay

$$\Delta_{an} = \begin{cases} \Delta_a + n\tilde{\Delta} + t_{\text{ITD}}, & 1 \leq a \leq A/2, \\ \Delta_a + (N + 1 - n)\tilde{\Delta}, & A/2 + 1 \leq a \leq A \end{cases} \quad (2)$$

with t_{ITD} as the interaural time difference in units of Δt.

The process of learning is modeled as spike-timing–dependent synaptic plasticity (STDP) [5,7] and described by the change of synaptic weights with time. Hence, in order to study the learning behavior, we assign to each synapse (a, n) a time-dependent synaptic weight $J_{an}(t) \in [J_{\min}, J_{\max}] = [0, 2]$. The local synaptic weight J_{an}^{loc} changes on a slow time scale $1/\xi$ with $\xi = 10^{-3} \ll 1$ in units of Δt in dependence upon the arrival times t_a of presynaptic spikes and the firing times t_n of the postsynaptic neuron n [5,6,14], while $\eta \in [0, 1]$ is the important learning parameter accompanying the learning window $\mathcal{W}$,

$$\frac{\mathrm{d}}{\mathrm{d}t} J_{an}^{\text{loc}}(t) = \xi \Big[w^{\text{in}}\delta(t - t_a) + w^{\text{out}}\delta(t - t_n) + w^{\text{unlearn}}$$
$$+ p(t)\left(J^{\max} - J_{an}\right) + \eta\,\mathcal{W}(t)\Big]. \quad (3)$$

Here, δ is a Dirac δ-function, $w^{\text{in}} = 0.02$ accounts for a synaptic change on arrival of a spike at t_a and so does $w^{\text{out}} = -0.25$ for t_n, when the postsynaptic neuron fires while $w^{\text{unlearn}} = -3 \cdot 10^{-4}$ describes a slow "forgetting". The $p(t)$ term incorporates intrinsic noise uniformly distributed in $[0, 10^{-4}]$ to model fluctuations on the level of the synaptic strength. Finally, the learning window

$$\mathcal{W}(t) = \sum_{t_n} W_{\hat{u}}(t - t_n)\delta(t - t_a) + \sum_{t_a} W_{\hat{u}}(t_a - t)\delta(t - t_n) \quad (4)$$

describes STDP with the learning window $W_{\hat{u}}$ shifted by $\hat{u} = -1$. Later on, $\hat{u}$ is used as an additional parameter in the study of traveling waves; see fig. 5. With hindsight [5], the underlying philosophy is simple. If we have an excitatory synapse and $t_a - t_n < 0$ so that the arriving spike instructs the neuron to fire and shortly thereafter it does so, the synapse performs a good job. Accordingly, it ought to be strengthened and hence $W_{\hat{u}}(t_a - t_n) > 0$. If, on the other hand, $t_a - t_n > 0$, the spike comes too late and the synapse should be weakened: $W_{\hat{u}}(t_a - t_n) < 0$. The function $W_{\hat{u}}$ has already been specified in [6], as

$$W_{\hat{u}}(s) = \begin{cases} 2\exp\left(\frac{s - \hat{u}}{\tau_2}\right) - \exp\left(\frac{s - \hat{u}}{\tau_0}\right), & s < \hat{u}, \\ \exp\left(-\frac{s - \hat{u}}{\tau_1}\right)\Big[2\left(1 + (s - \hat{u})\frac{\tau_1 + \tau_2}{\tau_1\tau_2}\right) \\ \quad - \left(1 + (s - \hat{u})\frac{\tau_0 + \tau_1}{\tau_0\tau_1}\right)\Big], & s \geq \hat{u}. \end{cases} \quad (5)$$

The time constants are $\tau_0 = 5$, $\tau_1 = 30$ and $\tau_2 = 50$ (in units of Δt). Before proceeding, we need to give map formation the finishing touch through axon-mediated spike-based learning (AMSL); cf. Kempter *et al.* [6]. Let

$\rho = 0.1$ be a small parameter accounting for contributions from neighboring neurons in (6) and define for axon a at neuron n the quantity $S_{an} = \sum_k \dot{J}_{ak}^{\text{loc}}$. The sum over k can be over nearest neighbors, or next nearest neighbors, or here over all synapses connected to the *same* axon. Then, with (3), the full synaptic dynamics is given by the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} J_{an} \equiv \dot{J}_{an} = \dot{J}_{an}^{\text{loc}}(t) + \rho S_{an}. \tag{6}$$

Hebbian learning is in the style of "practice makes perfect" and accordingly η is the important parameter. Once the barn owl's head is full-grown, it takes two more weeks before the animal can perform azimuthal sound localization or, equivalently, before the laminar map has been completed. Hence, we can choose a time much longer than the neuronal time scale (ms) and much shorter than the time of map completion, say 10 minutes, so as to segregate the two time scales of neuronal activity and map formation.

Equations (4) and (6) describe learning, *i.e.*, map formation processes in terms of developing synaptic strengths and, hence, contain the neuronal dynamics explicitly through spike times t_a and t_n, which are generated by a neuronal dynamics, orders of magnitude faster [4] than the synaptic one. The neuronal dynamics may well be taken to be simple, *viz.*, the integrate-and-fire model [15,16], governed by a membrane potential of the form

$$h_n(t) = \sum_{a=1}^{A} \sum_{j:t_n^f < t_{an,j} \leq t} J_{an}(t)\epsilon(t - t_{an,j}), \tag{7}$$

where t_n^f is the most recent firing time of neuron n, $t_{an,j}$ is the j-th arrival spike at synapse (a,n), and

$$\epsilon(t) = \begin{cases} (t/\tau^2)\exp(-t/\tau), & t \geq 0, \\ 0, & t < 0, \end{cases} \tag{8}$$

describes the excitatory postsynaptic response at the soma with time constant $\tau = 20\Delta t$. Despite its simplicity, the model is nonlinear as h_n is reset to 0, once h_n passes a threshold $\theta = 4.2$ from below. The synaptic weights J_{an} in (7) steer the dynamics of the neurons and, conversely, in the long run the latter shape the way in which the former develop.

Macroscopic bifurcation analysis. – To investigate the map dynamics, we would like to average out the neuronal dynamics and focus on the bifurcation behavior of map formation in dependence on the learning window's amplitude η and the AMSL coupling strength ρ. In fact, we expect the details of the neuronal dynamics to be immaterial. It is here that equation-free bifurcation analysis proves to be quite helpful [2,9,10]. For the analysis, we need to define a suitable macroscopic order parameter X that should measure for the problem at hand how well the map represents the sound source direction. To get an

(a) left ear, initial state

(b) right ear, initial state

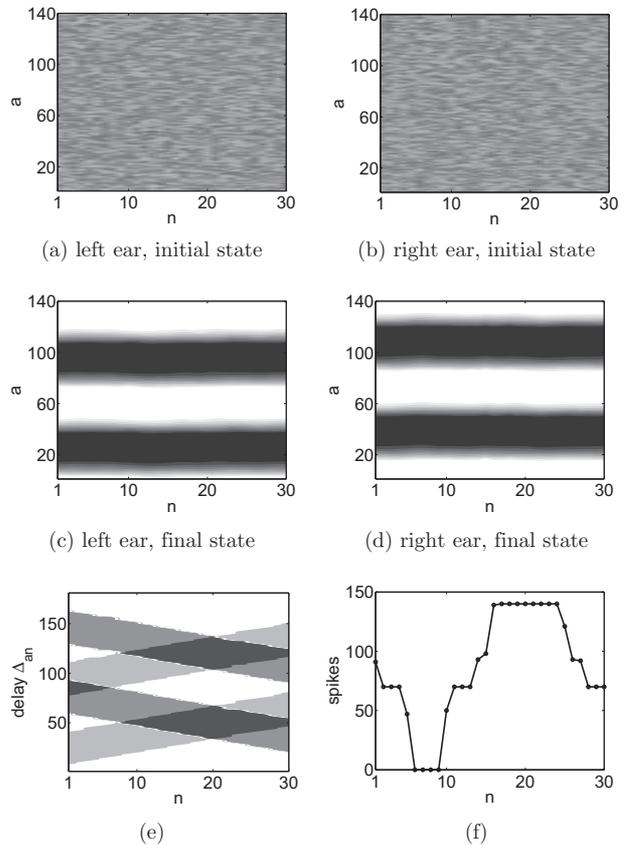(c) left ear, final state

(d) right ear, final state

(e)

(f)

Fig. 2: (a), (b): initial condition $J_{an}(0)$ for synaptic strengths connected to axons from the left and right ear. (c), (d): after a successful learning session employing model eqs. (1)–(7), a striped structure emerges in the synaptic weight matrix J, *i.e.*, a map, enabling the barn owl to azimuthally localize its prey. A gray scale indicates the strength of the synapse from the minimum weight (white) to the maximum weight (black). Synapses are spatially arranged on *parallel* axons connected to the left and right ear (see fig. 1). (e) According to (2) the delay for an incoming signal depends not only on the axon a but also on the neuron n. Plotting the delay Δ_{an} for strong synapses (black stripes in (c), (d)) results in an overlap where signals arrive at the same time (here at $n = 20$). A neuron being a threshold element, the overlap of left and right white stripes in a delay plot practically defines the map width. (f) During a simulation of 10000 time steps with the fully developed connectivity matrix the spiking activity shows a maximum at neuron 20 encoding the angle of prey.

impression of what we can expect, we turn to fig. 2. If a map is random or unstructered, *i.e.*, in a state without stripe structure, we require $X \approx 0$. On the other hand, $X \approx 1$ should signal a clearly organized map representing the relative periodicity (cf. fig. 2) of the synaptic matrix $J = (J_{an})$. In view of these requirements the choice below is quite natural,

$$X = \frac{1}{(A - 2T)N} \sum_n \left[\sum_{a=1}^{A/2-T} \mathcal{J}_{an} + \sum_{a=A/2+1}^{A-T} \mathcal{J}_{an} \right], \tag{9}$$

$$\mathcal{J}_{an} = (J_{an} - J_{a+T/2,n})\,\delta(J_{an}, J_{(a+T)n})\Theta(J_{an} - \tilde{J}),$$

where X depends on the time t, $\tilde{J} = 1.3$ is a significance level for the synaptic weights (see below), and Θ is the Heaviside step function. The Kronecker $\delta(J_{an}, J_{(a+T)n})$ assures a correct period across the axons and the prefactor $1/(A - 2T)N$ normalizes the sum. The Kronecker δ is implemented with tolerance 0.01. If the period is T, a synapse is weighted with the difference to the synapse half a period in front. For a fully developed matrix J, this is usually $J_{a+T/2,n} - J_{an} = J_{\max} - J_{\min} = 2$ with $J_{\min} = 0$.

Accordingly, a uniform matrix without stripe structure results in a small positive value of X. Furthermore, the Θ function in (9) takes only synaptic weights $J_{an} > \tilde{J}$ into account. The point is that a synaptic weight matrix with stripe structure but too small $|J_{an}|$ values is not able to let the neurons fire. Accordingly, prey detection would be impossible. We note that, due to the rectangularly shaped spatial patterns across the axons, a Fourier representation to check the periodicity is not useful [17].

We like to gain insight in the way in which the macroscopic order parameter X evolves in time in dependence upon η. Hence, we need a bifurcation analysis on a coarse level and disregarding the details of the underlying microscopic dynamics of the neurons. Exactly here equation-free analysis [2,9] comes in and allows to investigate the macroscopic behavior of the neuronal network as it leads to a map. The number of synapses is 2 to 3 orders of magnitude larger than that of the neurons and the dynamics of the former is adiabatic with respect to the latter, *i.e.*, at least 6 orders of magnitude slower to account for the fact that, after the hatching period, a young barn owl needs two more weeks to develop the neuronal map.

Equation-free methods. – In equation-free computations (see [2] for a review), a microscopic dynamical system

$$\dot{x}(t) = f(x(t), \eta), \quad x \in \mathbb{R}^m, \quad f : \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m \quad (10)$$

is described by a model function f depending on a parameter $\eta \in \mathbb{R}$. Here the microscopic model (10) summarizes the learning of the neuronal network defined in eqs. (1)–(7). By choosing a suitable order parameter X such as the one in (9), the high-, *i.e.*, m-dimensional dynamics in (10) can be reduced to a coarse, say M-dimensional scale,

$$\dot{X}(t) = F(X(t), \eta), \quad X \in \mathbb{R}^M, \quad F : \mathbb{R}^M \times \mathbb{R} \to \mathbb{R}^M, \quad (11)$$

where F describes the not explicitly but only implicitly given macroscopic dynamics and $M \ll m$; here $M = 1$. We can obtain the macroscopic dynamical properties of F by running suitably chosen short simulation bursts, *i.e.*, short simulations, of the microscopic system f and using (see [2,10])

$$X(t + \delta t) = \mathcal{R}[s^k(\mathcal{L}(X(t)), \mathrm{d}t)] \quad (12)$$

where s solves (10) numerically with time step $\mathrm{d}t$ where $\delta t = k\mathrm{d}t, k \in \mathbb{N}$. To be able to perform the continuation
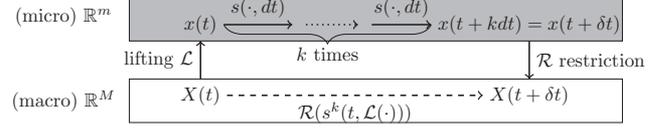

Fig. 3: Sketch of the macroscopic time stepper used in the equation-free method. A macroscopic state $X(t)$ is mapped onto a microscopic state $x(t)$ by using the lifting operator (15). The available microscopic model represented by eqs. (1)–(7) constitutes the microscopic time stepper s that is applied for k steps of size $\mathrm{d}t$ to yield a microscopic state $x(t + k\mathrm{d}t)$ at time $t + k\mathrm{d}t$. Finally, the microscopic state is mapped onto a macroscopic state $X(t+\delta t)$ through the restriction operator (9). This procedure defines the macroscopic time stepper (12).

of the macroscopic quantity depending on the parameter η, *i.e.*, the amplitude of the learning window in (3), the lifting operator $\mathcal{L}: \mathbb{R}^M \to \mathbb{R}^m$ and restriction operator $\mathcal{R}: \mathbb{R}^m \to \mathbb{R}^M$ are used repeatedly to switch between microscopic and macroscopic levels. This procedure is shown in fig. 3. In order to perform a macroscopic time step, a macroscopic state $X(t)$ is lifted with $\mathcal{L}$ (see also (15)) to a microscopic state $x(t)$, which is then evolved in time by the underlying model (eq. (1)–eq. (7)) with the time stepper s. The resulting state $x(t + \delta t)$ is subsequently used to compute the macroscopic state $X(t + \delta t)$ at time $t + \delta t$ by applying the restriction operator $\mathcal{R}$ (defined by (9)).

The unknown macroscopic right-hand side F is usually approximated as

$$F(X) = \frac{X(t + \delta t) - X(t)}{\delta t}. \quad (13)$$

This explicit scheme makes it difficult to determine the correct dynamics, since the lifting operator usually initializes the microscopic state away from the slow manifold where the macroscopic dynamics takes place (lifting error). Recently introduced implicit equation-free methods [10] allow to determine the not explicitly given right-hand side of eq. (11) as

$$F(X) = \frac{\mathcal{R}[s^{k+k_{\mathrm{skip}}}(\mathcal{L}(X(t)), \mathrm{d}t)] - \mathcal{R}[s^{k_{\mathrm{skip}}}(\mathcal{L}(X(t)), \mathrm{d}t)]}{\delta t}, \quad (14)$$

where $t_{\mathrm{skip}} = k_{\mathrm{skip}}\mathrm{d}t$ is the "healing" time. The implicit scheme (14) circumvents these lifting errors [10].

To numerically compute bifurcation diagrams, the stationary states of the order parameter X defined by $F(X, \eta) = 0$ are continued with respect to the parameter η using a predictor-corrector method [18,19]. A linear secant prediction and a Newton method as corrector are used. The Jacobian is computed by means of a finite-difference scheme.

The restriction operator $\mathcal{R}$ is determined by the definition of the macroscopic variable (9) and the lifting operator $\mathcal{L}(X) = J$ is chosen as

$$J^{\mathrm{new}} = J^{\mathrm{old}} + \mathrm{rand}(-\epsilon, \epsilon) + \alpha\,\mathrm{bin}(J^{\mathrm{old}} > 1). \quad (15)$$

The above computation adapts the previous state J^{old} of the continuation using noise with amplitude $\epsilon = 0.005$. $\text{rand}(-\epsilon, \epsilon)$ is a $A \times N$ uniformly distributed random matrix in $[-\epsilon, \epsilon]$ and $\text{bin}(J^{\text{old}} > 1)$ is a binarization of the connectivity matrix J. The binarization modifies strong synapses that are relevant to the restriction operator $\mathcal{R}$. The random matrix with small elements is added to obtain a small perturbation to the known state J^{old}. Note that the definition of the lifting operator $\mathcal{L}$ is not unique. The choice to modify known solutions through (15) on a branch is a convenient way to utilize the knowledge from previous solutions in order to construct such a lifting close to the low-dimensional slow manifold (see, *e.g.*, [9] for a description of different lifting operators). To initialize J close to a desired macroscopic value $\hat{X}$, the coefficient α is adjusted so that $\mathcal{R}(J^{\text{new}}) = \hat{X}$.

We start the continuation of the macroscopic stable branch with the fully developed connectivity matrix for $\eta = 1$ (cf. fig. 2) corresponding to a macroscopic fixed point of the system with $X(\eta) = 0.8$; see fig. 4, where the stable branch for $\eta = 1$ is found at the macroscopic value of $X = 0.8$. Continuing this fixed point in η yields a stable branch of solutions. In fig. 4, the branch obtained by implicit equation-free continuation (dots) coincides with the results from direct simulations in an up- (crosses) and down-sweep (circles), respectively. Here the macroscopic solution $X > 0$ encodes the stripe structure, *i.e.*, the correct connectivity matrix for sound localization. By investigating the eigenvalue of the (1×1) Jacobian of the linearized system along the branch, a bifurcation point is found at $\eta^* \approx 0.16$. For $\eta < \eta^*$, the solution $X = 0$ corresponds to a non-structured connectivity matrix, leading to a system that is not able to perform sound localization. Choosing a microscopic state with $X = 0$ as initial condition for an arbitrary $\eta < \eta^*$, it is possible with the equation-free continuation techniques presented here to find an unstable macroscopic solution at $X = 0$ even for $\eta > \eta^*$. The detection of an unstable branch gives information about the underlying dynamics and is never possible by direct simulations. By visual inspection of the bifurcation diagram in fig. 4, the bifurcation is reminiscent of a transcritical one, and not the usual pitchfork.

Strict detection of the bifurcation type is, however, complicated by two factors. First, because of the very nature of X, only the non-negative part of the bifurcation ($X \geq 0$) can be observed, thereby obscuring the branch in the strictly negative domain. Second, inherent noise in the macroscopic dynamics because of stochasticity in the underlying microscopic model complicates determining higher-order derivatives and thereby the bifurcation type. Additional analysis has shown that even though a variation of the AMSL coupling strength ρ induces spatial inhomogeneities, the periodicity measure X hardly changes.

Traveling waves in the synaptic connectivity matrix. – Besides the dependence upon η, the position of
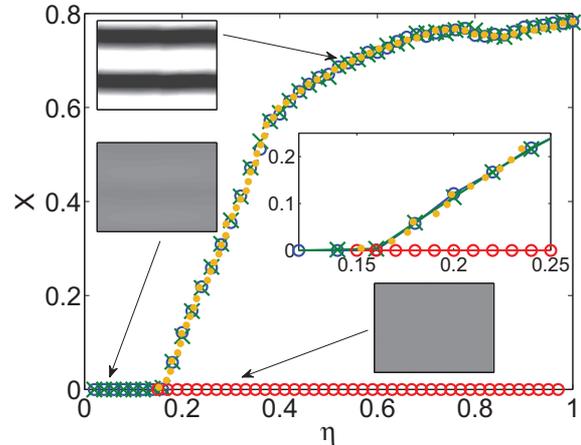

Fig. 4: (Colour on-line) Macroscopic bifurcation diagram of order parameter X as a function of the learning window amplitude η. Using equation-free bifurcation analysis, it is possible to continue branches and to detect bifurcation points of the network learning. A stable branch is found for $X > 0$ using different methods, down-sweep (circles) and up-sweep (crosses) in direct simulations and implicit equation-free continuation (dots). The bifurcation diagram shows a bifurcation reminiscent of a transcritical bifurcation at $\eta^* = 0.16$. The branch at $X = 0$ changes stability and the unstable branch after the bifurcation is continued using one-sided Newton corrections (circles). The big inset shows a magnification of the bifurcation region. The three small insets show the corresponding microscopic states for comparison. See the supplementary file `LearningNL.mp4` for a video of the learning process.

the bifurcation point also depends on the learning window shift $\hat{u}$ (see eq. (4)) and corresponds to *traveling waves* in the map that can be observed for all values $\eta > \eta^*$. The maps found until now in previous work were always stationary. It turns out that the striped structure, although a fixed point of the macroscopic order parameter X, is not a stationary state of the connectivity matrix J. Instead, very slowly traveling waves are observed; see fig. 5. The traveling waves have the same speed on axons from the left and right ear, which is the reason why the function of the network does not change and prey detection remains possible as the maximum activity of the network stays at the same neuron; cf. fig. 2. The speed of the traveling waves depends on the shift $\hat{u}$ in the learning window. By using a certain value (here $\hat{u} = -10$), it is possible to obtain standing waves, corresponding to a stationary state also on a microscopic level. Surprising as they are, the present findings are purely numerical but invite biological interpretation and experimental verification, *e.g.*, through periodically waxing and waning of dendritic spines [20]. The computation time of the traveling wave dynamics can be significantly reduced by using a coarse projective integration method, *i.e.*, an Euler scheme with time step Δt for the macroscopic dynamics of eq. (11). The speed-up depends basically on the required accuracy and is determined by the ratio $\Delta t / \delta t$; cf. also eq. (13).

(a) $\hat{u} = -1$
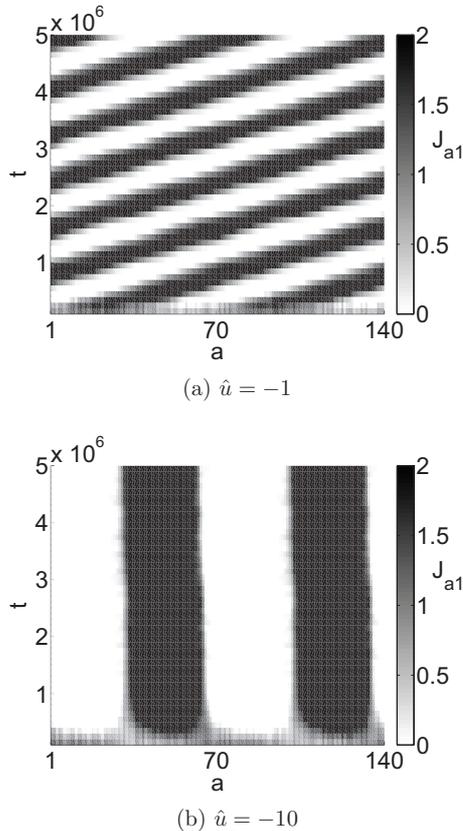

(b) $\hat{u} = -10$

Fig. 5: Traveling waves in the synaptic weight matrix for neuron $n = 1$ at $\eta = 1$ for different values of the shift $\hat{u}$ in the learning window; cf. (5). The other neurons ($n > 1$) show the same traveling-wave behavior. (a) For $\hat{u} = -1$, the stripes travel through the network at a nonzero but slow speed. (b) The traveling waves turn into a standing wave for $\hat{u} = -10$. The temporal multiplication factor of 10^6 clearly indicates that synaptic dynamics leading to map formation is adiabatic with respect to the neuronal one.

Conclusion. – In conclusion, the application of equation-free methods in conjunction with a bifurcation analysis of learning in neuronal networks gives rise to new insights into previously inaccessible phenomena such as: 1) study of unstable solutions; 2) detection and classification of macroscopic bifurcation points. Here we exhibit a transcritical bifurcation replacing the well-known pitchfork and identify the onset of structure formation in, *e.g.*, maps of synaptic connectivities; 3) significant reduction of computing time. The macroscopic time step Δt can be chosen much larger than the microscopic time δt. A striking example is provided by our finding of traveling waves instead of a time-invariant pattern in the sound localization map of the barn owl, as shown in fig. 5.

$* * *$

REFERENCES

[1] Huang K., *Statistical Mechanics*, 2nd edition (Wiley, New York) 1987, Chapts. 14–16.
[2] Kevrekidis I. G. and Samaey G., *Annu. Rev. Phys. Chem.*, **60** (2009) 321.
[3] Sullivan W. E. and Konishi M., *Proc. Natl. Acad. Sci. U.S.A.*, **84** (1986) 8400, fig. 3A.
[4] Konishi M., *Sci. Am.*, **268** (1993) 66.
[5] Gerstner W., Kempter R., van Hemmen J. L. and Wagner H., *Nature*, **383** (1996) 76.
[6] Kempter R., Leibold C., Wagner H. and van Hemmen J. L., *Proc. Natl. Acad. Sci. U.S.A.*, **98** (2001) 4166.
[7] Bi G.-Q. and Poo M.-M., *Annu. Rev. Neurosci.*, **24** (2001) 139.
[8] Eggermont J. J., *Hear. Res.*, **157** (2001) 1.
[9] Gear C. W., Hyman J. M., Kevrekidis P. G., Kevrekidis I. G., Runborg O. and Theodoropoulos C., *Commun. Math. Sci.*, **1** (2003) 715.
[10] Marschler C., Sieber J., Berkemer R., Kawamoto A. and Starke J., *SIAM J. Appl. Dyn. Syst.*, **13** (2014) 1202.
[11] Marschler C., Sieber J., Hjorth P. G. and Starke J., in *Traffic and Granular Flow '13*, edited by Chraibi M., Boltes M., Schadschneider A. and Seyfried A. (Springer) 2015.
[12] Carr C. E. and Konishi M., *Proc. Natl. Acad. Sci. U.S.A.*, **85** (1988) 8311.
[13] Carr C. E. and Konishi M., *J. Neurosci.*, **10** (1990) 3227.
[14] Kempter R., Gerstner W. and van Hemmen J. L., *Phys. Rev. E*, **59** (1999) 4498.
[15] Burkitt A. N., *Biol. Cybern.*, **95** (2006) 1; 97.
[16] Scarpetta S. and de Candia A., *PLoS ONE*, **8** (2013) e64162.
[17] van Hemmen J. L., Longtin A. and Vollmayr A. N., *Chaos*, **21** (2011) 047508.
[18] Doedel E., Keller H. B. and Kernevez J. P., *Int. J. Bifurcat. Chaos*, **1** (1991) 493; 745.
[19] Beyn W.-J., Champneys A., Doedel E., Govaerts W., Kuznetsov Y. A. and Sandstede B., in *Handbook of Dynamical Systems*, Vol. **2**, edited by Fiedler B. (Elsevier) 2002.
[20] Nägerl U. V., Willig K. I., Hein B., Hell S. W. and Bonhoeffer T., *Proc. Natl. Acad. Sci. U.S.A.*, **105** (2008) 18982.