Controlled pattern retrieval in a designed energy landscape

Pablo Kaluza[1,2] and Hildegard Meyer-Ortmanns[2]
[1)] *Abteilung Physikalische Chemie, Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany.*
[2)] *School of Engineering and Science, Jacobs University, P.O.Box 750561, 28725 Bremen, Germany.*

(Dated: 13 July 2011)

We construct a network of N globally coupled phase oscillators with a designed energy landscape. The oscillator dynamics is derived as gradient dynamics of this energy function. The network can store N-component binary patterns in cluster partitions of oscillators that synchronize to phase-locked motion. The phase relations of the cluster partitions correspond to the minima of the energy landscape. We can modulate the energy landscape via suitably chosen external fields in a way that the selected pattern for retrieval becomes the only minimum of the energy function. The phase oscillator dynamics then evolves the system to this minimum and retrieves the required pattern independently of the initial condition, one pattern at a time. When, in addition, our network of phase oscillators is driven by a suitably coupled pacemaker, the whole system performs in analogy to a central pattern generator of neural networks, provided that the time dependence of the external fields is accordingly tuned.

PACS numbers: 05.45.Xt, 89.75.-k, 89.75.Kd
Keywords: phase oscillators, pattern retrieval, cluster partition

Pattern retrieval is an important issue in the context of artificial neural networks. In particular patterns can be encoded in phase differences of oscillators. Usually, models of associative memory are static in the sense that a whole set of possible patterns can in principle be retrieved at a given time; which pattern is actually selected depends on the choice of initial conditions. So the patterns are encoded in static couplings. Also in our approach we encode the patterns in phase differences, but differently from the usual approach, we retrieve only a single pattern at a time, but from an arbitrary initial condition. Retrieving a new pattern requires a new choice of couplings. So our couplings will be time dependent in general. Given a static input that contains information about the required order and the labels for the required patterns, we shall show how a pacemaker can convert this static input into a temporal sequence of binary patterns, retrieved via Kuramoto-like dynamics from a storage device. The storage device is given by a designed energy landscape, on which the selected patterns of phase differences are imposed as local minima. The choice of couplings makes the selected pattern to be the only minimum of this energy landscape. This way our combined system of pacemaker and retrieval dynamics performs in an analogous way to a central pattern generator.

I. INTRODUCTION

Phase oscillators are studied in a variety of applications in physics, biology and information science. A paradigm of coupled phase oscillators is the Kuramoto model[1] (for a review see Acebron[2]). Among many applications to synchronization phenomena in biological systems, phase oscillators were studied in particular to represent the behavior of systems of neurons. The reason is that information in neural networks is not only encoded in the average activities of the neurons, but also in temporal features such as the relative phases of spikes[3]; therefore phase oscillators were supposed to individually represent the behavior of whole groups of neurons. These models for associative memory, based on temporal coding of information, consists of coupled oscillators, mutually interacting with a Hebbian learning rule; patterns are then stored as phase locked motion. Such models are generalizations of the Hopfield[4] model towards phase oscillators rather than spins as basic dynamical variables. They finally amount to modifications of the original Kuramoto model with a specific coupling matrix and with synaptic delay. Such an example for retrieving information about the phases in a network of neural oscillators is given by Aoyagi[5]. Here the author uses synaptic prescriptions that recover the Hebbian rule as a special case. In these generalized Kuramoto models, solutions, which correspond to error-free retrieval, are usually unstable so that the memory capacity is even smaller than the capacity of the Hopfield model. An improved model as considered by Nishikawa *et al.*[6] uses the second-order Fourier mode in the interaction between the oscillators in order to have a similar capacity as in the original Hopfield model.

On the one hand, such models serve to unfold possible mechanisms for memory and pattern retrieval in natural neural networks like the brain. On the other hand, nowadays the engineering aspect of networks plays also a prominent

role. For example, complex dynamical structures in populations of N phase oscillators were engineered by means of nonlinear time-delayed feedback that is implemented in the interactions of the oscillators[7–9]. In a similar spirit we engineer the energy landscape in which a system of phase oscillators moves in a controlled way due to an appropriate choice of the couplings between the oscillators (below we shall call the couplings external fields). Therefore, differently from previous models, in our case the mode of encoding patterns in the couplings is primarily realized as external signal which points to the state that is needed at a time, so that for a whole temporal sequence of states the couplings will become time dependent.

It is well known that coupled phase oscillators may converge to a synchronized state which is characterized by all oscillators sharing the same frequency, but differing by their individual phases, so that the state can be characterized by a phase-locked motion. If we gather all labeled oscillators sharing the same phase difference with respect to some reference oscillator into one cluster, the synchronized state can be described by a cluster partition of the oscillators. In this paper we shall choose the system of phase oscillators such that it assumes only two possible values for the phase difference, 0 or π, so that the maximal number of possible cluster partitions is 2^{N-1} in the state of phase-locked motion of a globally coupled system. A cluster partition ξ of N phase oscillators is then characterized by an $N-1$-component vector of phase differences, which take the values $\xi_i = 0$ or π. Such a vector will be called "pattern" or configuration in the following. (Obviously these patterns are immediately mapped to binary sequences with $N-1$ components.) Now our goal is to retrieve sequences of these 2^{N-1} patterns in a controlled way starting from an arbitrary initial condition. In view of that, we design an energy-landscape over configuration space such that the minima in this landscape correspond to the 2^{N-1} patterns, all of which are of the same height initially. What we call "energy" plays the role of a potential in a gradient dynamics that drives our phase oscillators to a local minimum of this potential (if there were several minima, the minimum approached by the dynamics would be the one with a basin of attraction that includes the given initial condition).

Now, in order to retrieve a selected pattern out of the 2^{N-1} possible ones that are in principle available, we apply external fields which modulate the energy landscape in a way that the selected pattern becomes the only minimum. For the selection of another pattern, the external fields would be appropriately changed. This way it becomes possible to retrieve a whole temporal sequence of patterns. We shall illustrate how we can retrieve a cyclic sequence of patterns and point on parallels to central pattern generators. Central pattern generators usually are considered in the context of neuroscience as explanation of how nervous systems produce movement. They are autonomous neural networks that can endogenously produce rhythmically patterned output like breathing, walking, or heartbeat[10]. For a review on how central pattern generators for locomotion can be driven by pacemaker cells we refer to Brocard *et al.*[11].

The paper is organized as follows. In section II we present the model and the rules for choosing the external fields. Section III gives the results for the energy landscape of four oscillators, and the retrieval within a system of eleven oscillators with 2^{10} patterns. We also show first results on the performance of our device in analogy to a central pattern generator. The conclusions are drawn in section IV.

II. MODEL

At a first place we construct an effective energy landscape. The energy landscape is chosen as the function L ($'L'$ shall remind to its role as Lyapunov function) that enters the dynamics like a potential. It depends on N oscillator phases in the following way:

$$L = -\frac{K}{4N} \sum_{i,j=1,i\neq j}^{N} \Big(\cos(\Phi_j - \Phi_i) - f_{ij} \Big)^2 \tag{1}$$

with $\Phi_i \in [0, 2\pi[, i = 1, ..., N$ the phase variables and K a coupling that finally determines the speed of convergence of the associated dynamics towards the stationary state, see Eq. 2 below. Here we just consider an all-to-all coupled system, but similar considerations apply for other topologies like oscillators with nearest neighbor couplings only. Out of the maximally $N(N-1)/2$ different phase differences, only $N-1$ are independent. We choose the phase of the Nth oscillator as reference, and $\Delta\Phi_{iN} = \Phi_i - \Phi_N \equiv x_i, i = 1, ..., N-1$ as independent variables. With f_{ij} we denote the external fields, $f_{ij} \in \mathbf{R}, i, j \in \{1, ..., N\}$.

For vanishing f_{ij}, L has obviously 2^{N-1} minima given by the vectors $\xi^{(k)}, k = 1, ..., 2^{N-1}$ of equal height with components $\xi_i^{(k)} \in \{0, \pi\}$. So the meaning of $\xi_i^{(k)} = \Delta\Phi_{iN}^{(k)}$ is the value that the phase difference between oscillator i and the reference oscillator N takes in the selected pattern k. In general, the f_{ij} are constant external fields appropriately chosen to modulate the energy function L, as we shall explain below. It is these 2^{N-1} local minima which are our candidates for retrieval. The label k of the state $\xi^{(k)}$ is determined by interpreting the pattern of 0s and πs as binary sequence in decimal representation. Local extrema of L are considered as function of $N-1$

independent phase differences. Candidates are all combinations of $\Delta\Phi_{ij}$ being either zero or π, for an arbitrary pair $i, j \in \{1, ..., N\}$. Other zeros of the first derivative of L occur at values of $\Delta\Phi_{ij} = \arccos f_{ij}$ as long as $|f_{ij}| \leq 1$, or, given a choice of f_{ij}, when the sum in the first derivative of L vanishes due to the cancelation of terms with alternating sign.

The oscillator dynamics is chosen to be the gradient dynamics of L according to

$$\dot{\Phi}_i = -\frac{K}{N} \sum_{j=1, j\neq i}^{N} \sin(\Delta\Phi_{ji}) \Big(\cos(\Delta\Phi_{ji}) - f_{ji} \Big). \tag{2}$$

At this place a remark is in order why we call f_{ij} external fields rather than couplings although they enter the dynamics of (Eq. 2) also in a multiplicative way. The reason is to stress that their choice (like usual external fields) is externally controlled, here via the required pattern for retrieval. First we shall choose these fields independently of time, later their time dependence will be controlled via a pacemaker, by the externally required (so far arbitrarily chosen) sequence of patterns. Thus the dynamics of Eq.2 serves only as a tool for pattern retrieval. It is not supposed to model a concrete natural system of oscillators. For vanishing external fields we get back the Kuramoto model with twice the usual frequency in the interaction of oscillators. Similarly to the Kuramoto dynamics the interaction of the oscillators depends only on phase differences $\Delta\Phi_{ij}$, and due to the choice of trigonometric functions the interaction terms are bounded as in the Kuramoto dynamics. The natural eigenfrequencies of the oscillators in our case are all set to zero. Due to the gradient dynamics the phase differences will evolve to a fixed point which is the minimum of the energy function L that is closest to the initial conditions. The reason is that the time derivative of L is negative. It is given as

$$\dot{L} = \sum_{i=1}^{N} \frac{\partial L}{\partial \Phi_i} \dot{\Phi}_i = -\sum_{i=1}^{N} \left(\frac{\partial L}{\partial \Phi_i} \right)^2 < 0 \, . \tag{3}$$

A similar set to our dynamic equations (Eq. 2) and to the energy function Eq. 1 was used by Nishikawa *et al.*[6], who added a stabilizing term (second-order Fourier mode) to the interaction of Kuramoto phase oscillators in order to overcome the shortage in the capacity for error-free retrieval of the associative memory. This way the authors succeeded in improving the capacity to become as high as that of the Hopfield model. Our model corresponds to theirs in the special case when only one pattern ξ is stored in the network, and our external fields are equal to the Hebbian couplings for that pattern. Under these conditions, we have an associative memory with error-free retrieval of just one pattern ξ, but our generic motivation is a very different one. We want to error-free retrieve only one pattern at a time, but from an arbitrary initial condition, possibly far away from the final minimum, which by choice of the external fields is the only minimum in the system; many patterns shall be retrieved only in a time sequence which is controlled by a pacemaker, therefore our couplings (external fields) are in general time dependent. Differently, in Nishikawa *et al.*[6], $N/\log N$ patterns can be stored with N neurons at a given time, but the associative memory in the network of Nishikawa *et al.* is static. The general implementation of patterns in our case is, however, quite different.

Choice of external fields A sufficient condition for a local minimum reads that the Hessian matrix of L is positive definite, i.e., all eigenvalues being larger than zero, due to an appropriate choice of f_{ij}. Now let us select one pattern $\xi^{(s)}$ with $s \in \{1, ..., 2^{N-1}\}$.

Conjecture Let the external fields modulate the interaction between oscillator pairs according to

$$f_{ij}(\alpha, s) = \alpha \left(\frac{2}{\pi} |\xi_i^{(s)} - \xi_j^{(s)}| - 1 \right) \tag{4}$$

for $i, j \in \{1, ..., N, i \neq j\}$ with $\xi_N^{(s)} \equiv 0$, α any real number with $\alpha > 1$ and s the index of the selected pattern. The conjecture then is that the selected configuration remains the only local minimum when these external fields are applied, while all other former local minima become saddles in at least one direction or local maxima. (For $\alpha < 1$ it can be shown that all local minima remain stable, but the selected one becomes the deepest.) The gradient dynamics will then retrieve the local minimum from any initial condition, not necessarily close to the selected minimum. In the appendix we shall show that for this choice of external fields the Hessian is positive definite for each configuration that is selected from the 2^{N-1} patterns, and not positive definite for all other $2^{N-1} - 1$ configurations, which were formerly also minima for vanishing external fields.

We cannot analytically exclude that local minima occur for phase configurations at intermediate values between 0 and π, but our numerical calculations do not show any such evidence for systems up to 11 phase oscillators with $2^{10} = 1024$ patterns, for which we have checked the feasibility of an appropriate retrieval.

III. RESULTS

A. Energy landscape

An application is shown in Fig. 1 for three mutually coupled phase oscillators with four patterns $(0,0),(0,\pi),(\pi,0),(\pi,\pi)$, corresponding to local minima of equal height, since the external fields are set to zero. Fig. 2 a-d shows modulated energy landscapes with only one local minimum at $(0,0)$ (a), $(0,\pi)$ (b), $(\pi,0)$ (c), and (π,π) (d), and a choice of f_{ij} according to Eq. 4 with $\alpha = 2$, cf. Table I. Note that cases (b), (c) and (d) in Table I result from a cyclic permutation, so that the dynamics is identical, but the binary patterns to be retrieved are different.

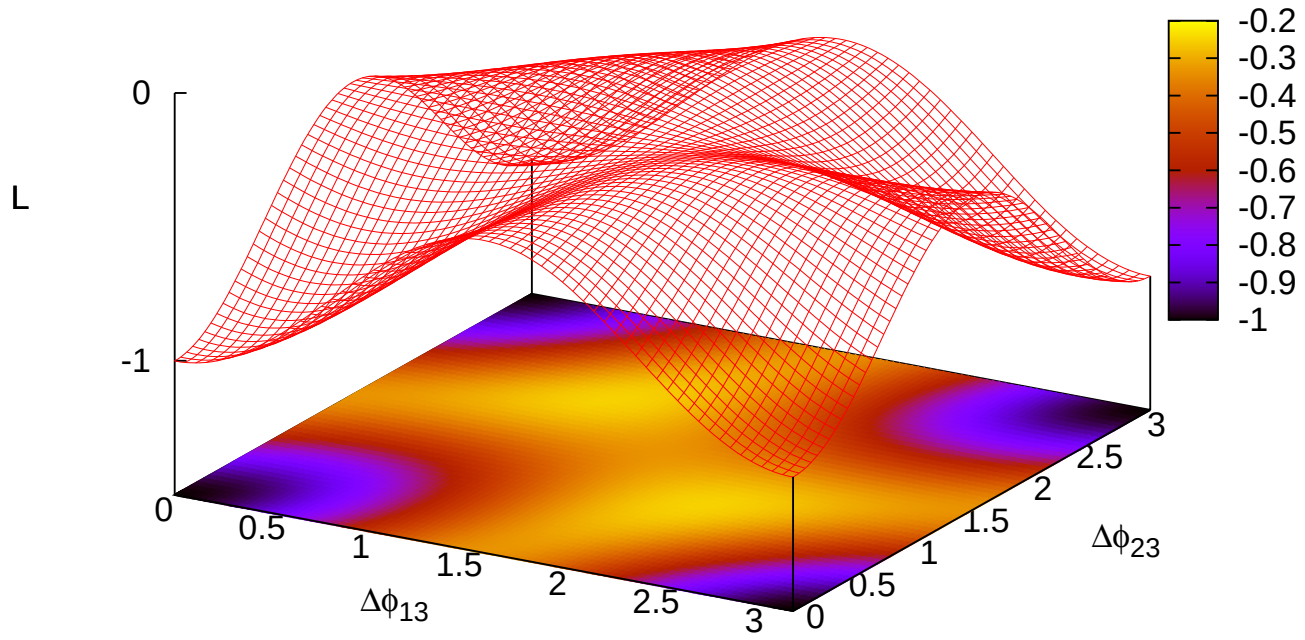

FIG. 1. (Color online) Energy landscape without external fields. Energy function L over configuration space spanned by the phase differences $\Delta\Phi_{13}$ and $\Delta\Phi_{23}$. L has four local minima of equal height at $(0,0)$, $(0,\pi)$, $(\pi,0)$, and (π,π).

TABLE I. **Choice of external fields**

Fig.2	f_{12}	f_{13}	f_{23}
(a)	-2	-2	-2
(b)	+2	-2	+2
(c)	+2	+2	-2
(d)	-2	+2	+2

Choice of external fields for three mutually coupled oscillators to modulate the energy L according to Eq.4 for the different panels of Fig. 2 (a-d) with $\alpha = 2$.

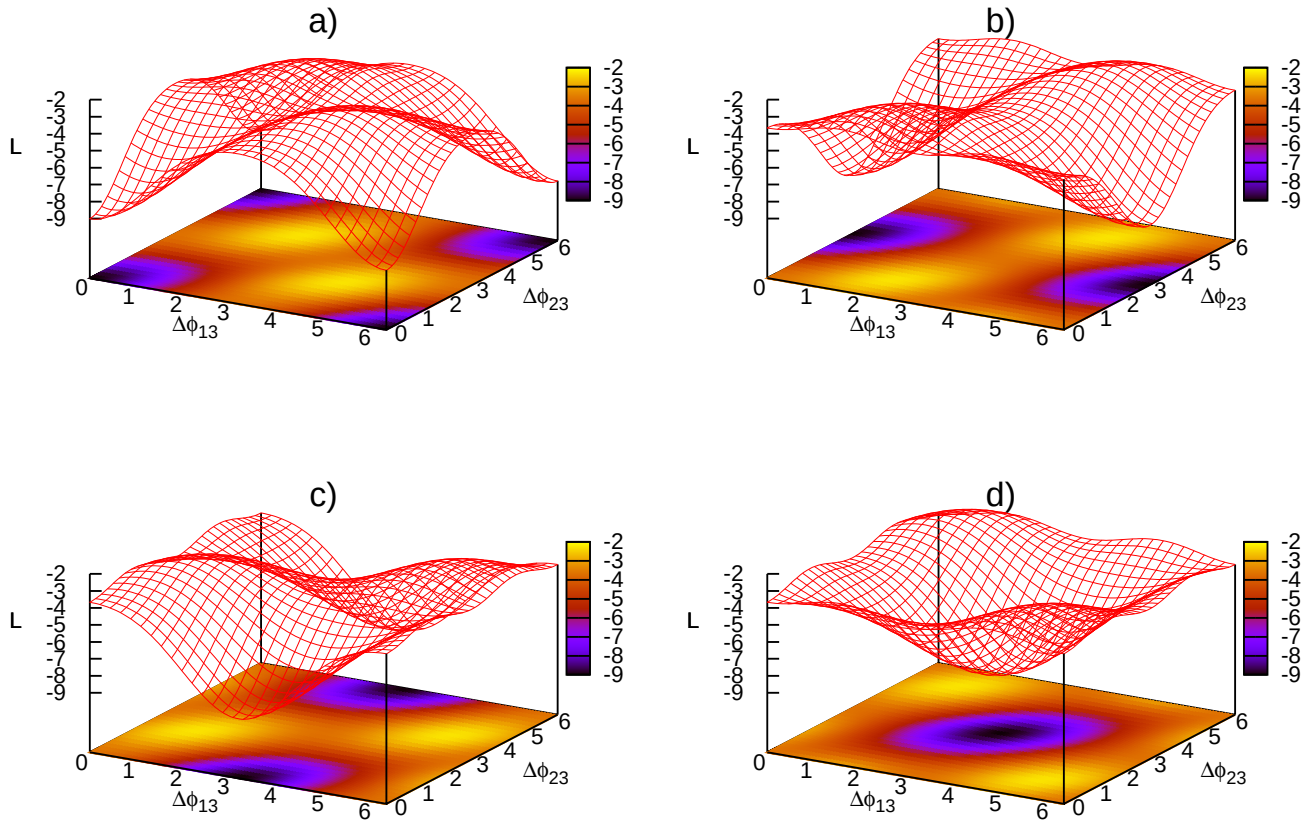

FIG. 2. (Color online) Modulated energy landscape. Modulated energy function L over configuration space after application of external fields such that the only local minimum occurs at $\xi^{(0)} = (0, 0)$ (a), $\xi^{(1)} = (0, \pi)$ (b), $\xi^{(2)} = (\pi, 0)$ (c), and $\xi^{(3)} = (\pi, \pi)$ (d). For the choice of external fields see Table I.

B. Retrieval of a pattern out of a large set

We studied larger systems of 11 phase oscillators with $2^{10} = 1024$ patterns, which are local minima as long as the external fields are kept zero. The external fields were chosen according to our rule Eq.4, $\alpha = 2$ was kept throughout all simulations, and each of the randomly selected patterns was found within a few time units, (with one time unit composed of 100 integration steps, since the integration step size was chosen as 0.01). As a typical example, let us retrieve the pattern $\xi^{(682)} = (\pi, 0, \pi, 0, \pi, 0, \pi, 0, \pi, 0)$ using the dynamics of Eq. 2. We integrate this dynamics using a second order stochastic Runge-Kutta method[12] with white noise. As initial conditions we choose the phases of all other corners of the hypercube in configuration space, corresponding to the 1023 patterns which are not selected, and a set of 10^5 randomly chosen initial phases $\Phi_i \in [0, 2\pi[$ in between. Throughout the simulations we keep the white noise at a low level with an amplitude of $T = 0.001$. This intensity is sufficiently small to let the system evolve to the selected minimum; on the other hand it is needed to give the system a kick towards one of the directions which is now unstable for the initial condition that corresponds to one of the non-selected corners. We find that in all these cases the dynamics retrieves the selected pattern $\xi^{(682)}$. This is the numerical evidence that there is only one local minimum in the landscape. For the retrieval of other patterns we proceeded in the same way and never found numerical evidence for another local minimum than the selected pattern.

C. Time dependent external fields and pattern generators

Obviously we can extend our dynamics to sequential pattern retrieval via time dependent external fields. We choose the external fields constant over a time interval that exceeds the typical transient time for retrieval of a state which starts from an arbitrary initial condition. Let us first label the 2^{N-1} states by decimal labels s as before (that correspond to the dual representation of s which is obtained by interpreting the pattern of 0s and πs of s as binary sequence). Out of these states we want to retrieve a subset of P patterns ${}^r\xi^{s(r)}$, $r = 1, ..., P$, $s(r)$ is then the associated decimal label of the state that is selected as the r^{th} state in the sequence. For cyclic sequences we choose now a pacemaker with constant frequency ω_R and phase ψ, whose time derivative is given by

$$\dot{\psi} = \omega_R. \tag{5}$$

It is the phase ψ of this pacemaker that points to the selected pattern $r, r \in \{1, ..., P\}$ over some period τ_r, chosen as $\tau_r = \frac{2\pi}{\omega_R}\frac{1}{P}(r-1)$. For simplicity we choose the time intervals between different patterns to be the same, namely $\Delta\tau_r = \tau_r - \tau_{r-1}$ for all $r \in \{1, ..., P\}$. When the instantaneous phase ψ reaches a value $\psi_r = \tau_r$, external fields are switched on that guarantee the retrieval of pattern r according to the following dynamical equations

$$\dot{\Phi}_i \;=\; -\frac{K}{N}\sum_{j=1, j\neq i}^{N} \sin(\Delta\Phi_{ji})\Big(\cos(\Delta\Phi_{ji}) - f_{ji}(\psi)\Big) \tag{6}$$

with

$$f_{ij}(\psi) = \sum_{r=1}^{P} \alpha\left(\frac{2}{\pi}|{}^r\xi_i^{s(r)} - {}^r\xi_j^{s(r)}| \;-\; 1\right) g_r(\psi) \tag{7}$$

and

$$g_r(\psi) = \Theta(\psi - \psi_r) - \Theta(\psi - \psi_r - B). \tag{8}$$

This means that the function $g_r(\psi)$ controls the external signal to be given as $\alpha\left(\frac{2}{\pi}|{}^r\xi_i^{s(r)} - {}^r\xi_j^{s(r)}| - 1\right)$ over the phase interval $[\psi_r, \psi_r + B]$, with Θ the Heavyside function and B the duration of the application of f_{ij} for the given choice. In general we should have $\tau_r \geq B > \tau_{trans}$, that is, both the time interval between two pattern retrievals and the time of application of the constant external fields to reach the new pattern should be larger than the transient time τ_{trans} that the system needs to go from one pattern to the next.

An example is shown in Fig. 3. Out of the 1024 pattern we randomly select the sequence ${}^1\xi^{672}$, ${}^2\xi^{0}$, ${}^3\xi^{942}$, ${}^4\xi^{477}$, ${}^5\xi^{1023}$ in the indicated order. The external fields are changed at time $\psi_r = 0,\ 2\pi/5,\ 4\pi/5,\ 6\pi/5$ and $8\pi/5$. At these instances we see in Fig. 3a that the energy function L jumps from its minimal value at $L \approx -500$ to some larger value around $L \approx -300$, where it remains as long as the system of oscillators searches the new minimum, corresponding to the new choice of external fields. When the new minimum is found, L drops to the minimal value again. During such an interval of duration τ_r, the Euclidean distance $D(t)$ in configuration space between the actual state and the closest pattern has a peak at an intermediate time interval where the system is moving from one to the next selected state. We see these peaks in Fig. 3b. For about half of the period τ_r this distance is zero, indicating that the system is in the required pattern. In Fig. 3c we plot the states which are closest to the instantaneous states of the system as function of time. Obviously the closest states are just the selected ones, but this does not mean that the actual states (evolving with time) are identical with the selected ones over the whole duration of the plateau; as mentioned before, the distance to the selected states vanishes only for roughly half of the period as it is seen from Fig. 3b. The width of the peaks in the distance from the closest states can be tuned by the coupling parameter K, large K accelerates the convergence to a new pattern, once the external fields are changed; also the strength of the external fields, parameterized by α, determines the speed of convergence.

In Fig. 4 we show three permutations of the same subset of five out of 2^{10} patterns. Indicated are the five plateaus in time where a certain pattern remains the closest to the current state and where this pattern agrees with the system's state over roughly half of the period (the analogous figures to Fig. 3b are not displayed here). It should be noticed that a change in the pattern sequence from Fig. 4a to 4b and 4c only amounts to a permutation of $\{{}^r\xi^{s(r)}\}_{r\in\{1,...,5\}}$ to $\{{}^{\pi(r)}\xi^{s(\pi(r))}\}_{\pi(r)\in\{1,...,5\}}$, where $\pi(r)$ denotes a permutation of the original order.

If a certain sequence of patterns would code a certain sequence of locomotion, a switch from one to another mode in locomotion amounts to a mere permutation of the order in which the pointer of the pacemaker clock addresses the pattern of external fields, but the "hardware" of patterns, stored in the energy function L along with the retrieval

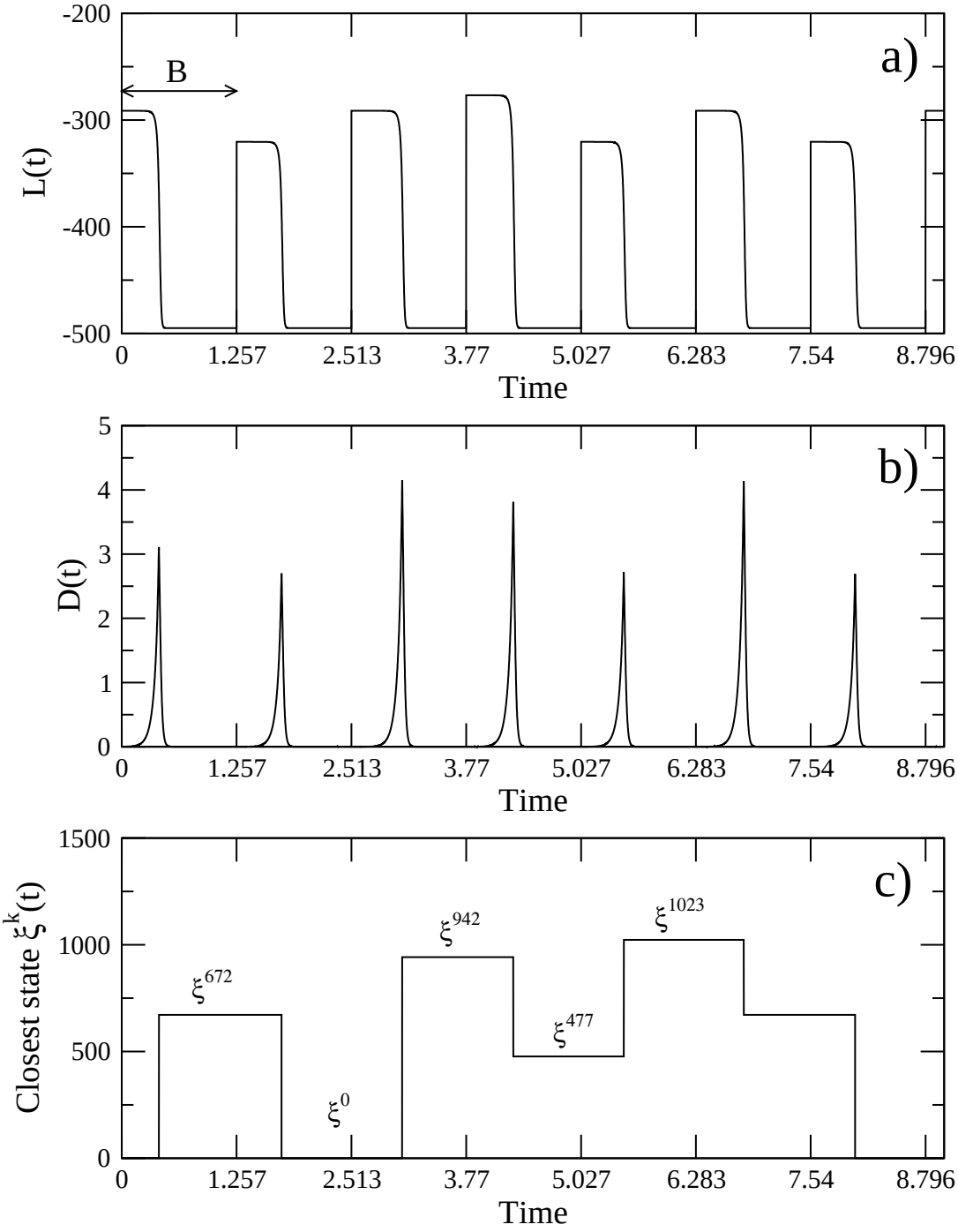

FIG. 3. Sequential pattern retrieval of a sequence $s_1 = \{\xi^{672}, \xi^0, \xi^{942}, \xi^{477}, \xi^{1023}\}$. (a) Energy $L(t)$ as function of time. (b) Euclidean distance of the actual state of the system to the closest state (corresponding to patterns on the corners of the hypercube). It vanishes for roughly half of the period B, so that the system then has retrieved the desired state. (c) Closest states to the current evolving state as function of time. As seen from the figure, the closest states themselves vary with time. The set of closest states agrees with the set of selected states. The states carry their decimal labels. The external fields are time-dependent as explained in the text. The parameters are chosen as $K = 10$, $N = 11$, $P = 5$, $B = 2\pi/P$, $\alpha = 2$ and $\omega_R = 1$. The dynamical equations of the pattern generator are integrated with the second order stochastic Runge-Kutta method with white noise with $\Delta t = 0.01$ and noise intensity $T = 0.001$.

rules, stored in f_{ij}, remains the same. So the energy landscape would play the role of a "multi-item working memory buffer", while the pacemaker would select the patterns from the memory at the required time instants. In our model

the very choice of external fields f_{ij} is externally designed in view of the pattern that we want to retrieve. In an extended dynamics the choice may be internally determined.

In our former example the static input would consist of the sequence $\{{}^r\xi^{s(r)}\}_{r\in\{1,...,5\}}$, by which the time ordering is specified via the order of the sequence and the patterns to be retrieved are coded as decimal numbers. The pacemaker clock then transforms this input into a temporally ordered retrieval of patterns, labeled by $s(r)$. Our device may therefore be regarded as a stylized version of the central pattern generator of the leeches' heartbeat. According to Hooper[10], the central pattern generator of the leech heartbeat can be divided into two sets, the rhythm generator (corresponding to our pacemaker), and the pattern generator (corresponding to our phase-oscillator dynamics). The pattern generator there generates the actual motor pattern in response to driving input from the rhythm generator (in our case in response to the driving input of external fields from the pacemaker). A detailed description of the leech heartbeat central pattern generator can be found in Hooper[10].

IV. CONCLUSIONS

We store binary patterns or, equivalently, sequences of zeros and πs in the phase differences of synchronized phase oscillators. The 2^{N-1} patterns are vectors of length $N-1$. Differently from other retrieval mechanisms, our procedure can start from an arbitrary configuration of oscillator phases. Due to the gradient dynamics the system will converge to the next local minimum of the energy landscape, and due to our choice of external fields there is only one such local minimum. Our external fields play also the role of (in general) time dependent couplings, and in common with other approaches of implementing memory in phase oscillators or neural networks, the information about the required patterns is stored in these couplings. Retrieval of a different pattern at another instant of time would amount to another adapted choice of external fields. The external fields should be kept constant over a time interval that is larger than the transient dynamics of the phase oscillators to approach the required pattern. Otherwise we may choose them as function of time to retrieve a whole temporal sequence of patterns. We have illustrated sequential pattern retrieval for a cyclic sequence via a pacemaker whose pointer controls the appropriate choice of external fields. So we consider our device as a possible building block in a larger (artificial) network. The fields f_{ij} would then no longer be external, and their time evolution would be controlled by the inherent pacemaker dynamics in such a larger unit. This way our retrieval device may then function in an analogous way to a central pattern generator.

V. ACKNOWLEDGMENTS

We would like to thank Alexander S. Mikhailov for valuable discussions.

VI. APPENDIX

Consider the energy function for an all-to-all coupled system of N phase oscillators:

$$L = -\frac{K}{4N}\sum_{ij,i\neq j}^{N}(\cos(\Phi_j-\Phi_i)-f_{ij})^2. \tag{9}$$

From now on we set the coupling $K=1$. Let the external fields be chosen according to

$$f_{ij}(\alpha,s)=\alpha\left(\frac{2}{\pi}|\xi_i^s-\xi_j^s|-1\right)\text{ for } i,j\in\{1,...,N,i\neq j\} \tag{10}$$

with $\Delta\Phi_{ij}=x_i-x_j$, $\xi_N^s=x_N\equiv 0$, α any real number with $\alpha>1$ and s the index of the selected pattern. We now prove a sufficient condition that the Hessian matrix with respect to the $N-1$ independent phase differences is positive definite for the selected pattern and not positive definite for all other $2^{N-1}-1$ patterns, provided we choose the external fields f_{ij} according to Eq.4. From the first derivative $\partial L/\partial x_i$ we immediately see that candidates for extrema are $x_i\in\{0,\pi\}$, where x_i was defined as $\Delta\Phi_{iN}=\Phi_i-\Phi_N$, while for $f_{iN}>1$ the individual cos-dependent terms are different from zero. For a particular given choice of f_{ij}, f_{iN} with possibly alternating signs the first derivatives can vanish also at intermediate values of x_i which we project on $[0,2\pi[$ that can lead to further extrema. This part we treat numerically in order to exclude that these extrema compete with the selected minimum that shall be retrieved.

Next let us consider the Hessian of L as function of the phase differences. Apart from the normalization factor, its diagonal elements are given as:

$$\begin{aligned}\frac{\partial^2 L}{\partial x_k^2} &= -\sin^2(x_k) + \cos(x_k)\big(\cos(x_k) - f_{kN}\big) \\ &- \sum_{j=1, j\neq k}^{N-1} \Big(\sin^2(x_k - x_j) - \cos(x_k - x_j) \\ &\cdot \big(\cos(x_k - x_j) - f_{jk}\big)\Big) ,\end{aligned} \tag{11}$$

its off-diagonal elements are:

$$\frac{\partial^2 L}{\partial x_k \partial x_l} = \sin^2(x_k - x_l) - \cos(x_k - x_l)\big(\cos(x_k - x_l) - f_{lk}\big). \tag{12}$$

For a choice of the external fields f_{ij} according to the rule (4), the matrix simplifies to

$$\frac{\partial^2 L}{\partial x_k^2} = -\Big\{ \mp 1\big(\pm 1 - f_{kN}\big) + \sum_{j=1, j\neq k}^{N-1} \Big(\mp 1\big(\pm 1 - f_{jk}\big)\Big)\Big\} \tag{13}$$

for the diagonal elements and $k = 1, ..., N-1$, and to

$$\frac{\partial^2 L}{\partial x_k \partial x_l} = \mp 1\big(\pm 1 - f_{lk}\big) \tag{14}$$

for the off-diagonal elements, with $k, l \in \{1, ..., N-1\}, k \neq l$. The upper (lower) sign in front of the the bracket with f_{lk} stands for the case that the difference of components $|x_k^s - x_l^s|$, read off from the selected configuration $\vec{\xi}^s$, is zero (π), respectively.
Next we study the positive definiteness of this matrix for all 2^{N-1} configurations which may be selected for retrieval. Here it is convenient to classify the configurations in terms of their Hamming distance from the selected pattern, i.e. the number of mismatches of components between $\vec{\xi}$ and $\vec{\xi}^s$, which varies between zero and $N-1$. By a suitable permutation of the coordinate axis in configuration space we can always achieve that the k mismatches occur in the first k coordinates of ξ so that the corresponding Hessian H is chosen as representative for all patterns with k mismatches.

H with no mismatches According to our choice of f_{ij}, their signs are opposite to those of $\cos(\Delta\Phi_{lk}^s)$, that is $\Delta\Phi_{lk}^s = 0$ or (π), so that $\cos\Delta\Phi_{lk}^s = 1$ or (-1) and $f_{lk} = -\alpha$ (or $+\alpha$), $\alpha > 1$, respectively. The diagonal elements then simplify to $\partial^2 L/\partial x_k^2 = (N-1)(1+\alpha)$, the off-diagonal elements to $\partial^2 L/\partial x_k \partial x_l = -(1+\alpha)$. The Hessian therefore takes the form of an $(N-1)\times(N-1)$ dimensional circulant matrix, whose eigenvalues turn out to be $\lambda_1 = 1+\alpha$ with multiplicity 1 and $\lambda_2 = N(1+\alpha)$ with multiplicity (N-2). (Here we have used the following: Eigenvalues of an $n \times n$ circulant matrix, specified by the vector $(c_0, c_1, ..., c_{n-1})$, are known to be given as $c_j' = \sum_{k=0}^{n-1} e^{2\pi i j k/n} c_k$ with $j = 0, -1, -2, ..., -(n-1)$. In our case the Hessian has a particularly simple form, for which one element in each row is $(N-1)(1+\alpha)$, while all other $N-2$ elements are $-(1+\alpha)$. Using these values and the fact that the sum over all n roots of the unit circle adds up to zero leads to our results for the eigenvalues.) Now, since for $\alpha > 1$ all eigenvalues are positive, the selected configuration corresponds to a local minimum in configuration space, whatever pattern has been chosen for retrieval. (In order to have only a local minimum at the selected configuration, obviously $\alpha > -1$ would be sufficient, but at the same time, the other patterns should become saddles or local maxima, and in view of that we shall need $\alpha > 1$, see below.)

H with one mismatch Next we evaluate the Hessian for a configuration that differs from the selected pattern in a single phase difference. Without loss of generality we assume the mismatch to happen in the first coordinate, affecting the Hessian in the first column and the first row according to $H_{11} = (N-1)(1-\alpha)$, $H_{1j} = (\alpha - 1) = H_{j1}$ for $j = 2, ..., N-1$, while the remaining $(N-2)(N-2)$ submatrix remains circulant. The Sylvester criterion, applied to the positive definiteness of the overall $(N-1)\times(N-1)$ matrix, is now violated due to the first element $H_{11} = (N-1)(1-\alpha) < 0$ for $\alpha > 1$, so that the configuration with one mismatch is no longer a local minimum of the energy function L. (As necessary and sufficient condition for a Hermitian matrix to be positive definite, the Sylvester criterion requires that all leading principal minors of the matrix are positive.)

H with $k > 1$ mismatches Now the configuration has k mismatches with the selected configuration which we arrange to occur in the first k coordinates. Here it should be noticed that f_{iN} will have the "wrong" sign with respect to $\Delta\Phi_{iN}$, $i = 1, .., k$, but f_{il} will have the "right" sign with respect to $\Delta\Phi_{il}$ for $i, l \in \{1, ..., k\}$, since two mismatches compensate in the relative phase differences ("wrong" (or "right") refer to the feature which prevents (or ensures) the property of becoming a local minimum, respectively.) This explains why the components of the $k \times k$ submatrix $S_k(H)$ in the upper left corner of the Hessian are given by

$$S_k(ii) = (N-1) - \alpha(N-2k+1), \quad i = 1, ..., k \tag{15}$$

for the diagonal elements and

$$S_k(ij) = -(1+\alpha) \;\; i, j = 1, ..., k, \;\; , i \neq j \tag{16}$$

for the off-diagonal elements. The submatrix $S_k(H)$ is again circulant and has eigenvalues $\lambda_1 = (N-k)(1-\alpha)$ with multiplicity 1 and $\lambda_2 = N - \alpha(N-2k)$ with multiplicity $k-1$, so that the determinant of this submatrix reads $|S_k(H)| = \lambda_1 \lambda_2^{k-1}$. Now we have to distinguish the following cases:
1. k odd. For k odd, λ_2^{k-1} is always positive while $\lambda_1 < 0$ for $\alpha > 1$, so that $|S_k(H)| < 0$ for odd k and $\alpha > 1$ and the Sylvester criterion for H being positive definite is violated as it should be for any positive number of mismatches.
2. k even. For k even, both eigenvalues may be negative so that $|S_k(H)| > 0$. In order to see that the Sylvester criterion is still violated, we have to distinguish the following cases:
(i) For $\alpha > 1$ and $k > N/2$ we have $\lambda_1 < 0$ and $\lambda_2 > 0$, so that the Sylvester criterion is violated.
(ii) For $\alpha > 1$ and $k < N/2$, $\lambda_2 < 0$ for $\alpha > N/(N-2k)$, so that $|S_k(H)| < 0$ only for $1 < \alpha < N/(N-2k)$.
(iii) To finally see what happens for $\alpha > 1$ and $\alpha > N/(N-2k)$ let us consider the determinant of the submatrix of size $l = k-1$ in the upper left corner of H. This matrix has eigenvalues $\sigma_1 = (N-k+1) - \alpha(N-k-1)$ and $\sigma_2 = N - \alpha(N-2k)$ with even algebraic multiplicity $(k-2)$, so that again the sign of λ_1 determines the sign of this subdeterminant. Now $\sigma_1 < 0$ for $1 < \frac{N-k+1}{N-k-1} < \alpha$, but this is certainly satisfied, since in the considered case $k \geq 2$ and α was even larger than $N/(N-2k)$ by assumption. So this $(k-1) \times (k-1)$-dimensional subdeterminant violates the Sylvester criterion for H to be positive definite.

In particular, for the maximal number of mismatches $k = N-1$, $\lambda_1 = 1-\alpha < 0$ for $\alpha > 1$ and $\lambda_2 = N + \alpha(N-2) > 0$ for $N > 2$, and for $N = 2$, $\lambda_2^{k-1} = \lambda_2^{N-2} = 1 > 0$, so that the corresponding pattern again ceases to be a local minimum of the energy function.

[1] Kuramoto Y (1984) Chemical Oscillations, Waves, Turbulence, (Springer, New York).
[2] Acebrón J A, Bonilla L L, Pérez Vicente C J, Ritort F and Spigler R (2005) The Kuramoto model: A simple paradigm for synchronization phenomena *Rev Mod Phys* 77:137-185.
[3] O'Keefe, J and Recce, M L (1993) Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, 3, 317-330.
[4] Hopfield J J (1982) Neural networks and physical systems with emergent collective computational abilities, *Proc Nat Acad Sci USA* 79(8):2554-2558.
[5] Aoyagi T (1995) Network of neural oscillators for retrieving phase information *Phys Rev Lett* 74(20):4075-4078.
[6] Nishikawa T, Lai Y-C and Hoppensteadt FC (2004) Capacity of oscillatory associative-memory networks with error-free retrieval. *Phys. Rev. Lett.* 92:108101.
[7] Hansel D, Mato G and Meunier C (1993) Clustering and slow switching in globally coupled phase oscillators. *Phys. Rev.* E 48(5): 3470-3477.
[8] Kiss I Z, Rusin C G, Kori H, and Hudson JL (2007) Engineering complex dynamical structures: sequential patterns and desynchronization *Science* 316:1886-1889.
[9] Manrubia S C, Mikhailov A S and Zanette D H (2004) Emergence of Dynamical Order: Synchronization Phenomena in Complex Systems *World Scientific Lecture Notes in Complex Systems, Vol. 2. World Scientific, Singapore.*
[10] Hooper SL (2000) Central pattern generators *Current Biology* 10(5):R176-R177.
[11] Brocard F, Tazerart S, and Vinay L (2010) Do pacemakers drive the central pattern generator for locomotion in mammals?, *The Neuroscientist* 16(2):139-155.
[12] Honeycutt R L (1992) Stochastic Runge-Kutta algorithms. I. White noise. *Phys. Rev.* A 45(2):600-603.

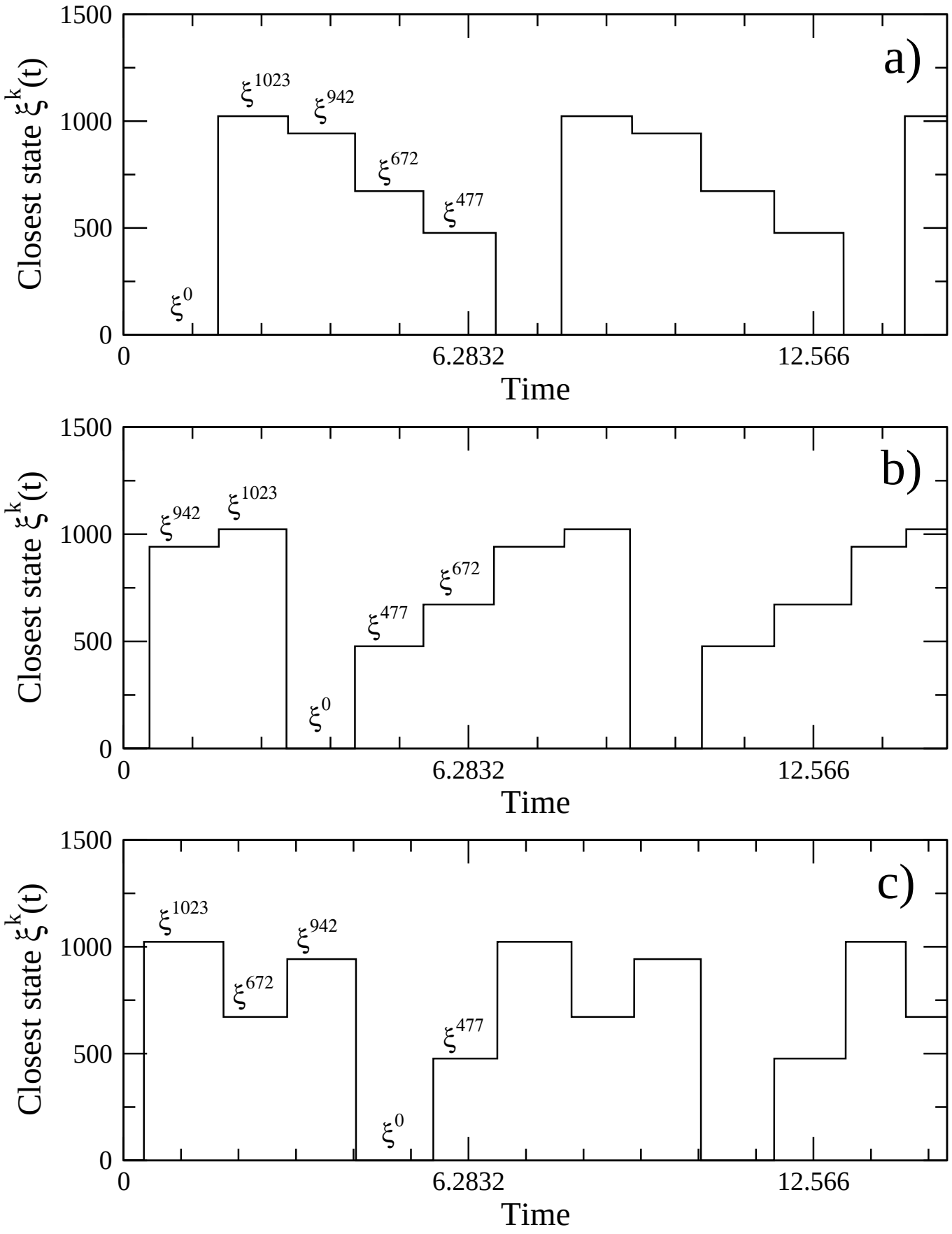

FIG. 4. Sequential pattern retrieval of different sequences with the same states. (a) Time evolution of the sequence $s_2 = \{\xi^0, \xi^{1023}, \xi^{942}, \xi^{672}, \xi^{477}\}$. (b) Time evolution of the sequence $s_3 = \{\xi^{942}, \xi^{1023}, \xi^0, \xi^{477}, \xi^{672}\}$. (c) Time evolution of the sequence $s_4 = \{\xi^{1023}, \xi^{672}, \xi^{942}, \xi^0, \xi^{477}\}$. In all cases we show only the states which are closest to the current evolving state. For roughly half a period the distance between the closest state and the actual state of the system vanishes, which is interpreted as pattern retrieval.