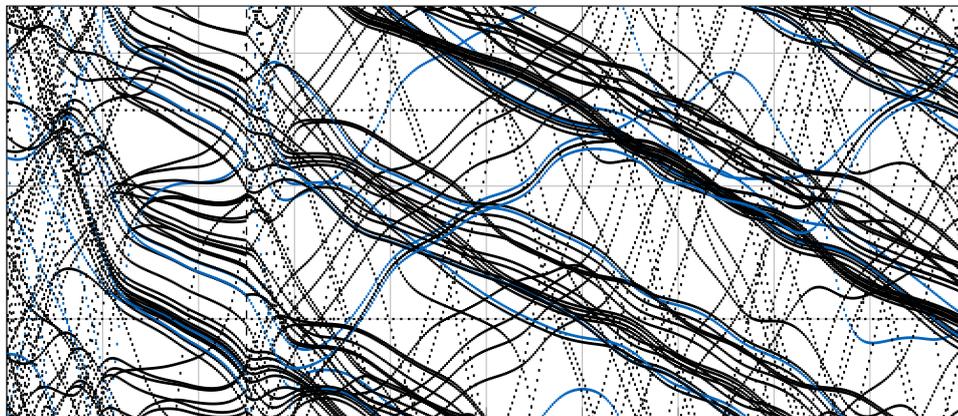Technische Universität München

Department Physik

E19a – Chemische Physik fern vom Gleichgewicht

A Neural Network of Weakly Coupled Nonlinear Oscillators with a Global, Time-Dependent Coupling - Theory and Experiment
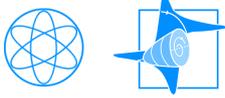

Doktorarbeit von Robert Hölzel

Dezember 2012

Technische Universität München

Department Physik

E19a – Chemische Physik fern vom Gleichgewicht

A Neural Network of Weakly Coupled Nonlinear Oscillators with a Global, Time-Dependent Coupling - Theory and Experiment

Robert Hölzel

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. F. Simmel

Prüfer der Dissertation:

1.   Univ.-Prof. Dr. K. Krischer

2.   Univ.-Prof. Dr. A. Pikovsky, Universität Potsdam

Die Dissertation wurde am 18. 12. 2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 18. 04. 2013 angenommen.

Danksagung

Meiner Doktormutter, Prof. Dr. Katharina Krischer, danke ich für die hervorragende Betreuung meiner Arbeit, bei der ich freie Hand hatte, ein faszinierendes Thema zu erforschen, und mich zugleich jederzeit auf ihre Unterstützung, gepaart mit ihrer großen Begeisterung, verlassen konnte.

Dr. Joachim Wiechers danke ich für die viele Zeit, die er sich genommen hat, um meine grundlegenden Fragen zum Entwurf von Schaltkreisen zu beantworten, sowie für zahlreiche hilfreiche Hinweise zur Optimierung des experimentellen Aufbaus.

Prof. Dr. Rolf Schuster bin ich ebenfalls für hilfreiche Diskussionen und für die Beantwortung diverser Fragen über die verwendeten elektronischen Bauteile dankbar.

Bei meinen Kollegen vom Fachbereich E19a möchte ich mich für die ausgesprochen freundschaftliche und anregende Arbeitsatmosphäre bedanken, deretwegen ich immer gern ins Labor beziehungsweise Büro gekommen bin, selbst in Zeiten großen Stresses.

Meinen Eltern danke ich für ihre beständige, liebevolle Zuwendung und für den bedingungslosen Rückhalt, den ich bei ihnen immer gefunden habe und immer finden werde.

Contents

1. Introduction

For about ten years now, commercially available consumer computers (desktops and laptops) have been experiencing a stall in their processing speeds. Most of the recent improvements in computer performance have been in the fields of energy efficiency and increased parallelization (for an accessible discussion of this development and its reasons, see Sutter (2005)). While these gains are still considerable, due to the mature, well-understood technology as well as the manpower and other resources dedicated to further improvement, the classical von-Neumann computer architecture (von Neumann 1945), which is the basis of modern computers, in principle does not lend itself particularly well to parallelization (Backus 1978).

This thesis is motivated by the continuing quest for an inherently parallel computing device - a device employing a network of many small, interconnected units acting both as data processors and data storage for efficient parallel computation. Such devices are called neural networks, because their parts mimic the function of neurons in human brains. They exhibit a level of parallelization that goes much deeper than the multicore processors and GPUs which today seem to have become almost synonymous with parallel computing. The decisive architectural difference is the real time exchange of information between the network elements through synaptic coupling, which is not a part of conventional parallel computing devices.

Neural networks are based on the notion that the human brain, while governed by dynamic processes many orders of magnitude slower than modern computers, still surpasses them in many areas of information processing, such as pattern matching, speech recognition, finding structure in unstructured data and, in general, the handling of complex, erroneous and sometimes contradictory external inputs. The fact that nowadays a supercomputer with 3000 processor cores and online access to terabytes of data, containing all knowledge of humanity, does only moderately better than humans in a quiz show, rather than proving the superiority of the computer, underlines the incredible efficiency of the human brain (referring to the IBM project "Watson" (Ferrucci, et al. 2010)).

Ideas to construct artificial neural networks are roughly as old as the conventional computer (McCulloch & Pitts 1943). After initial enthusiasm, interest in neural networks somewhat faded due to the success of conventional computers. A major reason for this is that the number of synapses of highly interconnected neural networks increases quadratically with the size of the system. Also, the connections between neurons are not entirely trivial. In general, synaptic coupling strengths must be adaptable to enable some form of Hebbian learning (Hebb 1949), which means that neurons that often act synchronously will have an increased synaptic coupling strength. When the technology was available to build neural networks in sufficient sizes, conventional computers had improved to the point that it was

(and still is) often more efficient (or at least more cost effective) to simulate a neural network on a conventional computer than building the network in reality.

Interest in neural networks was renewed in the 1980s, when a seminal paper by Hopfield (1982) introduced a new type of neural network, working as a content-addressable or associative memory. The intriguing property of Hopfield's approach is its applicability to simple physical networks (in particular spin glasses), in which the computational task of associative pattern recognition is accomplished by the tendency of the network to minimize the energy function

$$E = -\frac{1}{2N} \sum_{i}^{N} \sum_{j \neq i}^{N} w_{ij} V_i V_j, \tag{1.1}$$

where w_{ij} are the synaptic coupling coefficients (or weights) and $V_i \in \{0, 1\}$ is the state of the i-th neuron. For suitably chosen (i.e. according to the Hebbian rule) synaptic weights, equation (1.1) has a minimum for each memorized binary pattern of neuron states.

In spite of the renaissance of neural network research due to the Hopfield network, commercial applications have been few and far between (e.g. the Intel 80170NX electronically trainable neural network chip (Holler, et al. 1989)) and never particularly successful. However, in light of the developments described earlier and in conjunction with the advent of the memristor (Chua 1971, Strukov, et al. 2008), which makes the physical implementation of synapses a lot easier, neuromorphic computers today are once again in the focus of science and computer industry alike (see for example Lohr (2011)), some researchers even speak of a second renaissance of neural networks (Schmidhuber, et al. 2011).

Although these prospects are better than they have been for quite some time now, many problems remain to be tackled on the way to applicable neural network hardware. One of them, as mentioned before, is that all neuromorphic hardware must somehow realize the synaptic coupling. In the case of Hopfield type networks, this means that N^2 connections have to be provided for N neurons, which becomes increasingly difficult if each connection is implemented as a physical object like for example a wire or a resistance in an electric circuit. The neural network presented in this thesis follows a different approach, by using a global coupling which acts on all neurons alike, but affects each of them differently. Once built, a network employing such a coupling technique allows for the addition of further neurons without the need for adding any more neuron-neuron connections. This is made possible by using oscillatory neurons that resonate only with specific frequencies present in an external signal comprising many frequency components. The mathematical basis of this idea was developed by Hoppensteadt & Izhikevich (1999). For this thesis, a working associative network of 8 electrical oscillators using and proving their concept was built. To the best knowledge of the author, this is the first hardware implementation. The experimental realization of the coupling is based on research on globally coupled

electrochemical oscillators, investigated in theory by Mazouz, et al. (1997), Krischer, et al. (2003) and Hölzel (2007) and experimentally by Kiss, et al. (1999), Kiss, et al. (2002) and Mikhailov, et al. (2004). Despite these foundations, one of the challenges of this thesis was the transfer of the abstract idea of an oscillatory neural network to an experiment. Therefore, before the chapter on experimental results, there is a chapter on the design of the experiments, which describes this process.

Besides the experimental part, there is also a theoretical part to this thesis, which has roughly the same extent. Before the motivation of the theoretical analysis is discussed, it is helpful to look a bit closer into the mathematics underlying the neural network at hand: Hoppensteadt & Izhikevich suggested a network of globally coupled sinusoidal phase oscillators (or Kuramoto oscillators, (Kuramoto 1984)) with different frequencies and a complex, time-dependent coupling function. They showed that the dynamics of the network can be effectively written as

$$\dot{\varphi}_i = \frac{1}{N} \sum_j^N w_{ij} \sin(\varphi_j - \varphi_i), \tag{1.2}$$

which describes the dynamics of a network of Kuramoto oscillators with phases ϑ_i defined by $\vartheta_i(t) = \Omega t + \varphi_i(t)$, common frequency Ω and individual coupling strengths w_{ij} between pairs of oscillators. By choosing the phase shifts φ_i as variables, the common rotation was eliminated. At first sight, this seems a bit remote from the topic of neural networks. However, there is a strong relation between (1.2) and the Hopfield network, which will become apparent a little further below.

The dynamical system (1.2) is not only studied in the context of neural networks, but also with respect to synchronization phenomena of coupled oscillatory systems. It is a mathematically relatively tractable toy model for more complicated oscillatory networks and therefore appears, in varied form (i.e. with a frequency distribution instead of equal frequencies, some added noise and/or different sets of coefficients w_{ij}), in a considerable amount of publications (see for example section 3. in the review by Arenas, et al. (2008) and the references there). Most of those, however, do not deal with an all-to-all coupling obeying the Hebbian rule.

Due to the involvement of many researchers from different scientific communities including neurobiologists, mathematicians and physicists, it is surprisingly hard to determine who came up with the idea to build a Hebbian neural network of coupled Kuramoto oscillators first. For this idea to take form, the two concepts of neural networks and synchronized oscillators had to be merged, a process that started in the early 80s and took roughly ten years (this is briefly covered in a review by Acebrón, et al. (2005), which is focused on synchronization phenomena).

In the synchronization community the dynamical system (1.2) was investigated as a model

for coupled oscillators (escpecially chemical oscillators), where the w_{ij} reflected the spatial displacement between oscillators (Neu 1980), a nearest-neighbor coupling (Sakaguchi, et al. 1987, Kowalski, et al. 1988), random interactions (Daido 1987, Daido 1992), or a combination of all those (Niebur, et al. 1991), albeit with a frequency distribution instead of equal oscillator frequencies.

In the context of neural networks it appears that (1.2) was introduced by Cohen, et al. (1982), once again with a frequency distribution. They, and also Strogatz & Mirollo (1988) and Kopell & Ermentrout (1988), were concerned with chains of oscillators modeling undulatory locomotion in certain lower animals. The first usage of (1.2) with some kind of Hebbian learning can be found in Baldi & Meir (1990) and Sompolinsky, et al. (1991), who used several coupled arrays of oscillators, as well as (Abbott 1990), who mentioned the open problem of a neural network using the dynamics (1.2) with the coupling coefficients

$$w_{ij} = \sum_{k=1}^{M} \xi_i^k \xi_j^k, \tag{1.3}$$

which is an application of the Hebbian rule for a set of M memorized binary pattern vectors $\boldsymbol{\xi}^k$ with $\xi_i^k = \pm 1$. The coupling matrix given by (1.3) is the coupling matrix used throughout this thesis.

Meanwhile, coming from a more physical side, Cook (1989) extended the Hopfield model, by considering a Q-state neural network with arbitrarily large Q (instead of the 2-state spins used by Hopfield). In the limit of $Q \to \infty$, i.e. if the spins take on continuous values (and if, other than in Cook's paper, only binary memorized patterns are considered), the Hamiltonian of the system is given by

$$E = -\frac{1}{2N} \sum_{i}^{N} \sum_{j \neq i}^{N} w_{ij} \cos(\varphi_j - \varphi_i), \tag{1.4}$$

where the φ_i are now interpreted as the orientation of the respective spins (Nobre & Sherrington 1986). The energy function (1.4) is a generalization of (1.1). Incidentally, it is also a potential function of (1.2), which means the dynamics are equivalent.

Building on this preliminary work, Arenas & Pérez-Vicente (1994), were first to unequivocally prove the applicability of a Hebbian network of (almost) identical Kuramoto oscillators for pattern recognition. They, like Cook, allowed for real-valued pattern vectors instead of binary patterns only, (1.3) being a special case.

In the years that followed, the understanding of Hebbian networks of Kuramoto oscillators was further improved (Park & Choi 1995, Pérez-Vicente, et al. 1996, Aoyagi & Kitano 1997, Aoyagi & Kitano 1998, Aonishi 1998, Aonishi, et al. 1999, Yamana, et al. 1999, Yoshioka & Shiino 2000, Hong, et al. 2001, Shim, et al. 2002). The most relevant findings for the

experimental realization were that, other than in the Hopfield network, the memorized patterns are not attractors of the network dynamics. Instead, the system settles for states close to the memorized patterns. Also, it has been found that the overlap of the resulting pattern with the memorized pattern is the largest immediately after the onset of pattern recognition.

None of these publications are written from an experimentalists point of view (especially an experimentalist dealing with a small number of oscillatory neurons). Most of them use the thermodynamic limit of $N \to \infty$ in their calculations and all of them use some kind of mean field approach to assess the quality of pattern recognition. For reasons discussed in the theory chapter, this approach has some disadvantages. Instead of concentrating on the mean field, the analysis presented in this thesis deals with the dynamics of individual oscillators and with the structure of the 8-dimensional (or N-dimensional) phase space of the network dynamics. Using this approach, a theoretical understanding of some previously undocumented long-term behavior of the network observed both in simulations and in the experiments was reached. In addition, the effects of several different types of experimental inaccuracies on the quality of pattern recognition were assessed and compared, using numerical simulations.

2. Theoretical background

2.1 Associative networks

An associative network is a mathematical model or physical system that performs a mapping between two vector spaces A and B

$$f : A \to B, \boldsymbol{x} \mapsto \boldsymbol{y}$$

with the following characteristic: For a set of predefined input vectors $\boldsymbol{x}^k \in A, \; k = 1, \dots, M$, there is a set of output vectors $\boldsymbol{y}^k \in B, \; k = 1, \dots, M$ such that any input in a neighborhood of $\boldsymbol{x}^k$ is mapped to $\boldsymbol{y}^k$. This is the defining property of an associative network (Rojas 1996). If $\boldsymbol{x}^k = \boldsymbol{y}^k$, the network is called autoassociative. Due to its capability to associate a fixed output vector with all inputs that are close to a predefined input vector, an associative network can be used to recognize previously stored patterns (given by $\boldsymbol{y}^k$) from noisy input. The network discussed in this thesis is an autoassociative network modeled by a continuous dynamical system in which the memorized patterns are represented by different preferred dynamical states of the system (albeit those states are not attractors). The input pattern is supplied by the choice of the initial conditions. The network is able to store and retrieve (albeit not perfectly) binary patterns $\boldsymbol{\xi}^k$ that only have entries of either 1 or -1:

$$\boldsymbol{\xi}^k = (\xi_1^k, \dots, \xi_N^k)^{\mathrm{T}}, \; \xi_i^k = \pm 1, \; i = 1 \dots N, \; k = 1 \dots M$$

The most important characteristic of an associative network is its loading rate

$$\alpha = \frac{M}{N}.$$

The maximal loading rate α_{max} indicates how many patterns can be reliably memorized and retrieved by a network of a given size. The quality of retrieval depends not only on the architecture of the network but also on the separation of the patterns (i.e. the Hamming distance, the number of components that are different between two vectors) in configuration space. To make different network architectures comparable, storage capacities are generally computed for large N and randomly selected patterns $\boldsymbol{\xi}^k$, which have Hamming distances of roughly $N/2$. The optimal value for the Hamming distance for the neural network under investigation in this thesis is exactly $N/2$, which is the case for orthogonal patterns, i.e. if $\boldsymbol{\xi}^k \cdot \boldsymbol{\xi}^l = 0$.

2.2   Pattern recognition in a network of individually coupled ideal oscillators with equal frequency and constant coupling

Building on previous, more general research (Arenas & Pérez-Vicente 1994, Aoyagi & Kitano 1997), Aonishi (1998) investigated and proved the ability of an associative network of coupled Kuramoto oscillators (Kuramoto 1984) with individually adjustable coupling strengths to recognize binary pattern vectors from a given set of memorized patterns. The oscillator phases $\vartheta_i(t)$ in this network are governed by the following system of differential equations:

$$\dot{\vartheta}_i = \Omega + \frac{1}{N}\sum_{j=1}^{N} w_{ij}\sin(\vartheta_j - \vartheta_i) \tag{2.1}$$

Here N is the number of oscillators, Ω is the common frequency of all oscillators and $w_{ij} = w_{ji}$ is the individual coupling strength between oscillators i and j. A positive value of w_{ij} stabilizes a difference of $\vartheta_i - \vartheta_j = 0$ while a negative value stabilizes $\vartheta_i - \vartheta_j = \pi$.

Since the phase differences are the quantities we are interested in, we monitor the phase shift rather than the phase in our numerical simulations and experiments, where each phase $\vartheta_i(t)$ is written as the sum of a uniform oscillation and a phase shift $\varphi_i(t)$:

$$\vartheta_i(t) = \Omega t + \varphi_i(t)$$

Therefore, $\vartheta_i - \vartheta_j = \varphi_i - \varphi_j$, and (2.1) can be rewritten as

$$\dot{\varphi}_i = \frac{1}{N}\sum_{j=1}^{N} w_{ij}\sin(\varphi_j - \varphi_i). \tag{2.2}$$

The pattern recognition process consists of two steps, namely the initialization and the recognition step.

2.2.1   Initialization of the network

By choosing appropriate values for w_{ij}, a binary pattern $\boldsymbol{\xi}$,

$$\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_n), \;\; \xi_i = \pm 1, \;\; i = 1...N,$$

can be stored in the network. From here on, this process will be called the *initialization* of the network. We consider the network as successfully initialized to $\boldsymbol{\xi}$, if

$$\xi_i\xi_j = 1 \Leftrightarrow \varphi_i = \varphi_j + 2k\pi$$

$$\xi_i\xi_j = -1 \Leftrightarrow \varphi_i = \varphi_j + (2k+1)\pi,$$

where $k \in \mathbb{Z}$. This means there are two groups of synchronized oscillators representing all '1' and all '-1' entries of the pattern vector respectively that differ in phase by π. This is achieved by setting the coupling strengths to

$$w_{ij} = \xi_i \xi_j.$$

The resulting dynamical system is given by

$$\dot{\varphi}_i = \frac{1}{N} \sum_{j=1}^{N} \xi_i \xi_j \sin(\varphi_j - \varphi_i). \tag{2.3}$$

By choosing new variables

$$\varphi_i \rightarrow \varphi'_i = \varphi_i + \pi \text{ for } \xi_i = -1$$
$$\varphi_i \rightarrow \varphi'_i = \varphi_i \text{ for } \xi_i = 1$$

the system (2.3) can be transformed to a system of uniformly coupled Kuramoto oscillators:

$$\dot{\varphi}'_i = \frac{1}{N} \sum_{j=1}^{N} \sin(\varphi'_j - \varphi'_i)$$

In this network, all oscillators synchronize (Kuramoto 1975). Therefore, $\boldsymbol{\xi}$ is the only attractor of (2.3).

The left hand sides of Fig. 1to Fig. 4 show numerical simulation results for the initialization of a pattern over time (up to the dashed vertical line), starting at random initial phases.

Note that after initialization there are two distinct branches of oscillators with phase shifts at 0 and π, respectively. This is not necessarily the case, because (2.3) is invariant under a global shift

$$\varphi_i \rightarrow \varphi'_i = \varphi + u, \quad u \in \mathbb{R}.$$

In all plots in this chapter, the phase shift curves were globally shifted in such a way that after initialization for $\xi_i = -1$ the phase shift is $\varphi_i = 0$ and for $\xi_i = 1$ the phase shift is $\varphi_i = \pi$.

2.2.2 Recognition of a defective pattern

In the recognition step a network that has been initialized to $\boldsymbol{\xi}$ is used to recognize $\boldsymbol{\xi}$ as one of a given set of M memorized patterns $\boldsymbol{\xi}^k$, $k = 1 \ldots M$. This is achieved by adjusting the coupling strengths w_{ij} according to a Hebbian (Hebb 1949) learning rule:

$$w_{ij} = \sum_{k=1}^{M} \xi_i^k \xi_j^k$$
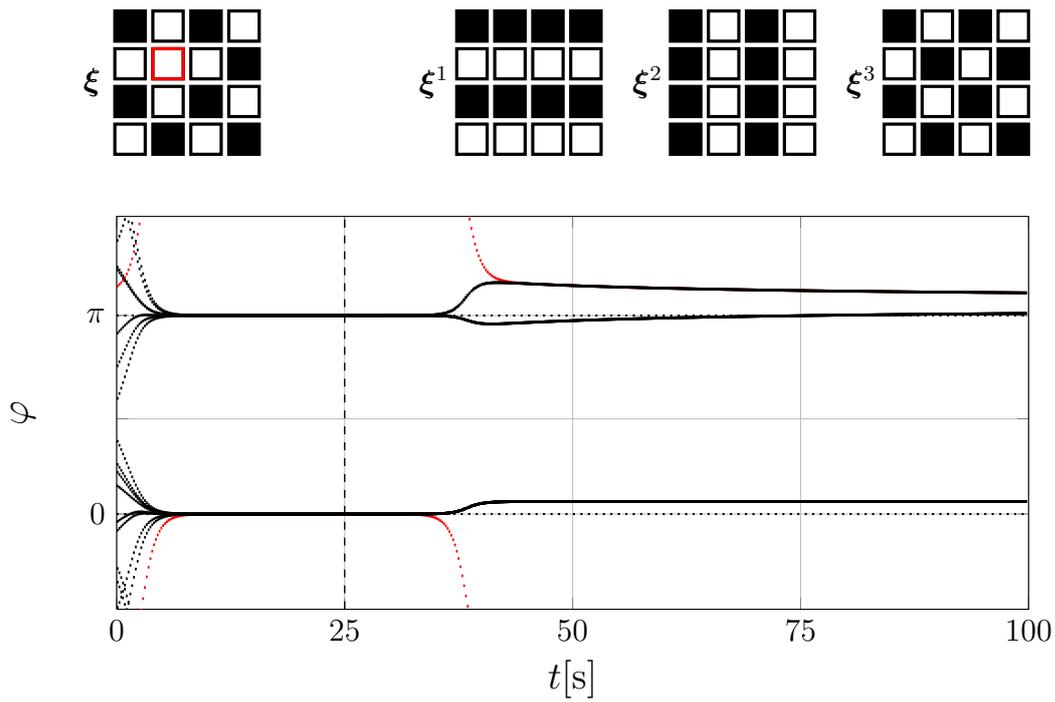
Fig. 1: Numerical time integration of the phase shifts in (2.3) ($t \leq 25\,\mathrm{s}$) and (2.4) ($t > 25\,\mathrm{s}$) with random initial conditions, $N = 16$ oscillators, $M = 3$ memorized patterns and one erroneous bit in the initial pattern. The initial pattern $\boldsymbol{\xi}$ and the memorized patterns $\boldsymbol{\xi}^k$ are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. Both the squares and the phase shift curves belonging to erroneous bits in the initial pattern are marked red. Note that the three memorized patterns are mutually orthogonal.
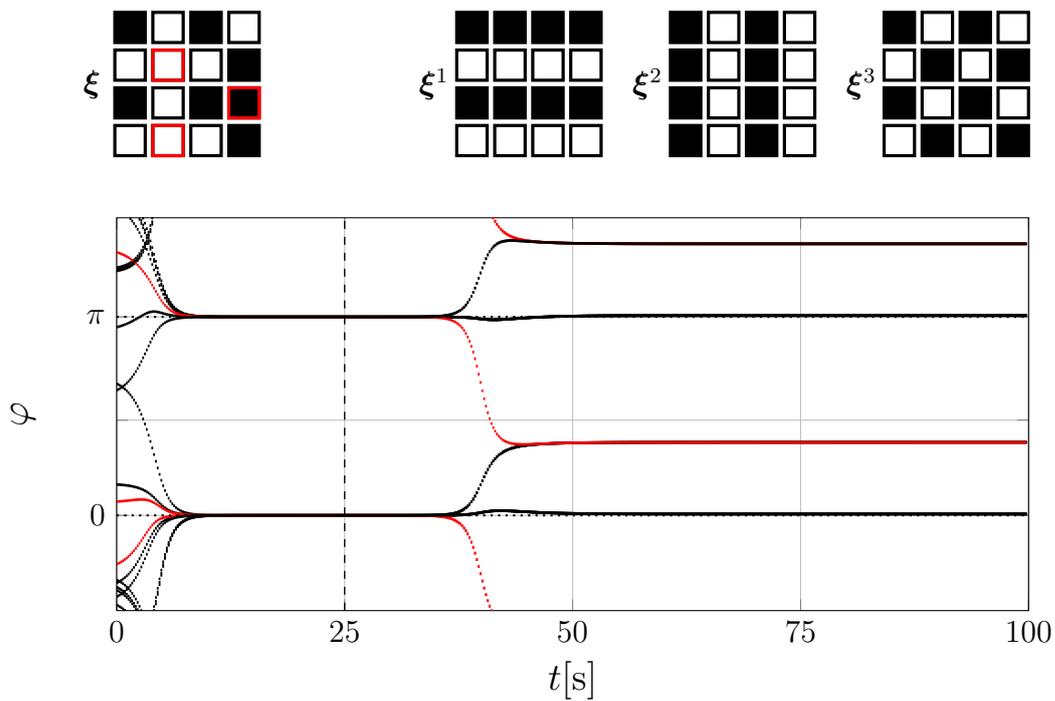
Fig. 2: Simulation results analog to Fig. 1, with three erroneous bits in the initial pattern.
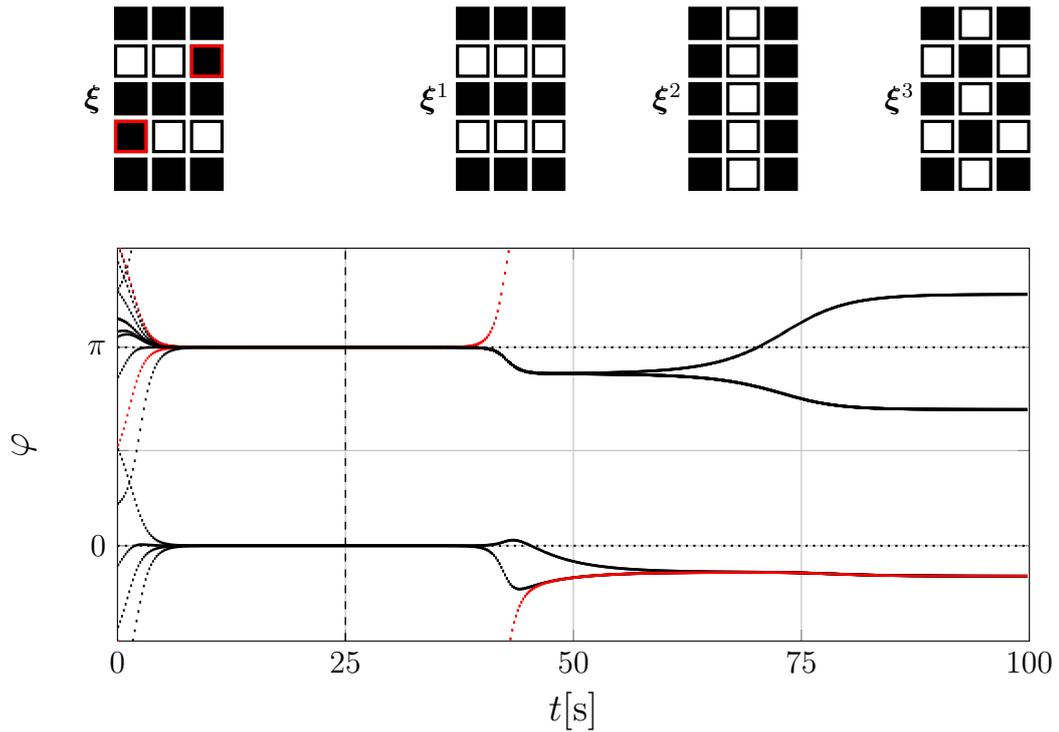
Fig. 3: Simulation results analog to Fig. 1 with $N = 15$ and two erroneous bits in the initial pattern. Note that the three memorized patterns are not orthogonal (which is impossible for an odd number of pixels). Still, the patterns are separate enough to enable a successful pattern recognition.
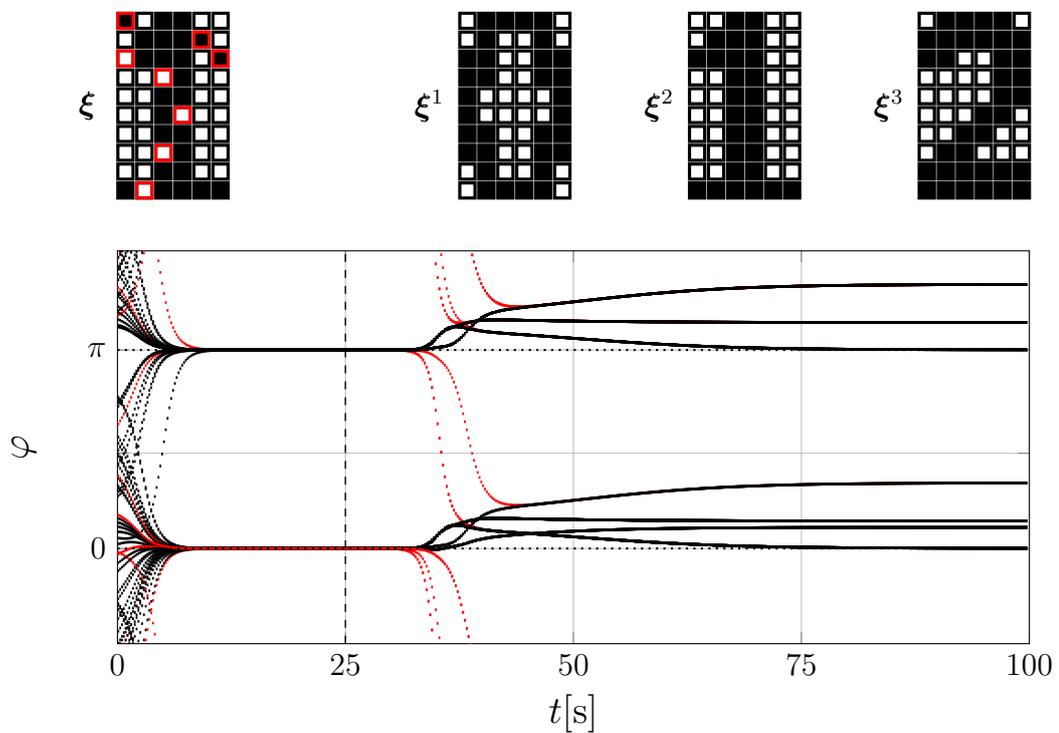
Fig. 4: Simulation results analog to Fig. 1 with $N = 60$ and eight erroneous bits in the initial pattern. Note that the three memorized patterns are not orthogonal. Still, the patterns are separate enough to enable a successful pattern recognition.

$$\dot{\varphi}_i = \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{M} \xi_i^k \xi_j^k \sin(\varphi_j - \varphi_i) \qquad (2.4)$$

The idea behind this choice of coupling is as follows: If the initial pattern ξ is close to ξ^l, the ξ^l-component in the coupling will favor a switching of the few erroneous bits of ξ to their correct values in ξ^l because the overwhelming majority of oscillators already have the correct phase for this pattern. On the other hand, the $\xi^{k \neq l}$-components of the coupling will neither strongly favor switching nor keeping the current phase shift of any erroneous oscillator, if the Hamming distances of the memorized patterns are large enough (ideally they are orthogonal, making the Hamming distances $N/2$ each).

Another motivation for the introduction of the Hebbian learning rule is given by the fact that (2.3) is a gradient system with a potential function

$$E_{\xi} = -\frac{1}{2N} \sum_{i=1}^{N} \sum_{j=1}^{N} \xi_i \xi_j \cos(\varphi_j - \varphi_i),$$

where $\dot{\varphi}_i = -\partial E / \partial \varphi_i$. The globally stable minimum of this potential corresponds to ξ.

The potential function of equation (2.4) is the sum of the potential functions belonging to the patterns ξ^k:

$$E_{\text{Hebb}} = \sum_{k=1}^{M} E_{\xi^k} = -\frac{1}{2N} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{M} \xi_i^k \xi_j^k \cos(\varphi_j - \varphi_i) \qquad (2.5)$$

It is reasonable to expect this combined potential landscape to have a locally stable minimum in the vicinity of each pattern ξ^k. If that is the case, initial patterns that are close to ξ^k are transformed into the pattern corresponding to this attractor.

The question remains whether these attractors actually exist and what their basins of attraction are, which in turn determine the storage capacity $\alpha_{\max}$ of the network. As of now, no analytical way to locate the attractors of (2.4) has been found for the general case. However, Aonishi (1998) showed that in general, for large N and *randomly* assigned patterns ξ^k, the dynamics of the network settles indeed in the vicinity of a memorized pattern if $M/N \leq \alpha_{\max} = 0.042$ (for comparison: the standard Hopfield network (Hopfield 1982) has a loading rate of $\alpha_{\max} \approx 0.14$ (Amit, et al. 1985)) and if the initial pattern is close enough to a particular memorized pattern. These steady states reached in the long run are different from the patterns themselves, which in general are unstable steady states of (2.4). If the patterns are near orthogonal instead of random, the ratio M/N may be larger than $\alpha_{\max}$ and still allow for successful pattern recognition, a fact that was used for the experiments with 8 oscillators and 3 memorized patterns described in chapter 6. If all patterns are mutually orthogonal, it appears that the attractors form a degenerate network

with a common basin of attraction rather than isolated minima of the potential. Section 4.1 presents some arguments that corroborate this conjecture.

Fig. 1 to Fig. 4 show the pattern recognition process for different network sizes and memorized patterns. Fig. 1 and Fig. 2 depict numerical simulations with 3 orthogonal patterns. Note that the system approaches a steady state φ_{final} that is substantially different from a perfect recognition φ_{perfect}, even more so as more erroneous bits are introduced in the initialized pattern. However, as long as

$$\left|\varphi_{i,\text{final}} - \varphi_{i,\text{perfect}}\right| < \pi/2, \tag{2.6}$$

the two branches corresponding to $\xi_i = 1$ and $\xi_i = -1$ can still be told apart, and recognition is successful. It is possible to modify the coupling such that the ξ^k become attractors (Nishikawa, et al. 2004). However, this only works for strongly coupled oscillators of identical frequency.

Fig. 3 shows a numerical simulation of a pattern recognition with 3 non-orthogonal patterns. The system comes very close to the exact pattern it is supposed to recognize. Since this is an unstable steady state, however, the system evolves away again from the exact representation, but 2.6 remains fulfilled.

Finally, Fig. 4 is an example for a successful pattern recognition in a larger network (pattern taken from Hoppensteadt & Izhikevich (1999)) with non-orthogonal patterns. Here $M/N = 0.05$.

In all of these simulations, a perturbation of the order 10^{-6} was added to the phases when the recognition coupling was switched on, because otherwise the system would have remained at the initial pattern which is a steady state as well, albeit an unstable one.

Please note that, in the network discussed in this thesis, the coupling matrix w_{ij} has to be changed between initialization mode and recognition mode for each pattern recognition process. As a consequence, the memorized patterns are not stored in the network itself but must be kept elsewhere; the network is only a processing unit, other than a classic neural network (which has static synaptic connection strengths w_{ij}, once it is completely trained). On the plus side, this means that the network can switch dynamically between different pattern sets.

2.3   Pattern recognition in a network of globally coupled ideal oscillators with different frequencies and weak time-dependent coupling

The complexity of physically implementing the coupling in (2.1) is of $\mathcal{O}\left(N^2\right)$. If each coupling link is established by a unique physical connection (like a conducting wire between electrical oscillators), adding or removing a new oscillatory neuron becomes increasingly complex with a larger number of oscillators, rendering a modular setup impossible. Depending on the type of oscillators used for the physical realization, it might not be feasible to individually link oscillators to each other at all.

On the other hand, a global coupling can be readily implemented in a number of oscillatory systems (Kiss et al. 1999, Taylor, et al. 2008, Mertens & Weaver 2011, Temirbayev, et al. 2012). Hoppensteadt & Izhikevich (1999) proposed a network of weakly globally coupled oscillators with different frequencies. In their approach, the complexity of $\mathcal{O}(N^2)$ needed to implement individual coupling strengths between pairs of oscillators is transferred away from the physical architecture of the network towards the time-dependence of the global coupling.

If the frequencies are suitably chosen, the dynamics of the system can be reduced to the shape of (2.2), which is explained in section 2.3.1. Section 2.3.2 deals with the difficulties of the search for suitable frequencies.

2.3.1   Equivalence of weak global coupling dynamics and strong individual coupling dynamics

In the network proposed by Hoppensteadt and Izhikevich the time evolution of oscillator phases $\vartheta_i(t)$ is given by:

$$\dot{\vartheta}_i = \Omega_i + \varepsilon a(t)\frac{1}{N}\sum_{j=1}^{N}\sin(\vartheta_j - \vartheta_i) \tag{2.7}$$

Or, in terms of the phase shifts $\varphi_i(t)$:

$$\dot{\varphi}_i = \varepsilon a(t)\frac{1}{N}\sum_{j=1}^{N}\sin((\Omega_j - \Omega_i)t + \varphi_j - \varphi_i) \tag{2.8}$$

These equations differ from the original system of individually coupled oscillators ((2.1), (2.2)) in several points:

- The frequencies Ω_i are different.

- The coupling strength ε is weak compared to the natural frequencies of the oscillators (namely, $\varepsilon N^2 \ll \Omega_{\max} - \Omega_{\min}$, as is explained in section 2.3.2). This means that the time scale of the dynamics of φ_i is of $\mathcal{O}(1/\varepsilon)$ instead of $\mathcal{O}(1)$.

- Instead of a constant individual coupling strength w_{ij} for each pair of oscillators, there is a time-dependent global coupling term $a(t)$. This global coupling must be designed such that an "effective coupling" proportional to w_{ij} arises between oscillators i and j.

What follows now is a multiple time scale analysis (Strogatz 1994) of (2.8) that yields a set of time-averaged equations for the φ_i. This approach is an alternative to the ansatz used by Hoppensteadt and Izhikevich, namely the direct averaging method, that comes to the same result. The approach developed in this work puts more emphasis on the fact that there is a fast time scale depending on the oscillator frequencies Ω_i besides the slow time scale introduced by ε. The quality of the approximation depends on the ratio between fast and slow time scale rather than on the slow time scale alone.

Let δ denote the minimal difference between two frequencies present in the network:

$$\delta = \min_{i \neq j} |\Omega_i - \Omega_j|$$

δ is the smallest frequency that occurs in the oscillatory terms on the right hand side of (2.8). An oscillation with this frequency is still considered as fast. Therefore, it makes sense to define the fast time $\mathcal{T} = \delta \cdot t$ and the slow time $\tau = \varepsilon t$. Now, $\varphi_i(t)$ is written as a Taylor series in ε/δ:

$$\varphi_i(t) = \varphi_i^0(\mathcal{T}, \tau) + \frac{\varepsilon}{\delta} \varphi_i^1(\mathcal{T}, \tau) + \mathcal{O}\left(\frac{\varepsilon^2}{\delta^2}\right)$$

$$\dot{\varphi}_i(t) = \delta \frac{\partial \varphi_i^0(\mathcal{T}, \tau)}{\partial \mathcal{T}} + \varepsilon \left(\frac{\partial \varphi_i^1(\mathcal{T}, \tau)}{\partial \mathcal{T}} + \frac{\partial \varphi_i^0(\mathcal{T}, \tau)}{\partial \tau} \right) + \mathcal{O}\left(\frac{\varepsilon^2}{\delta}\right)$$

Plugging these approximations into (2.8) and separating the terms according to their order in ε/δ yields for zeroth order

$$\frac{\partial \varphi_i^0}{\partial \mathcal{T}} = 0 \Leftrightarrow \varphi_i^0(\mathcal{T}, \tau) = \varphi_i^0(\tau).$$

The first order equation reads

$$\frac{\partial \varphi_i^1}{\partial \mathcal{T}} = a(\mathcal{T}/\delta) \frac{1}{N} \sum_{j=1}^{N} \sin((\Omega_j - \Omega_i)\mathcal{T}/\delta + \varphi_j^0 - \varphi_i^0) - \frac{\partial \varphi_i^0}{\partial \tau}. \tag{2.9}$$

On the right hand side, all fast oscillating terms with a vanishing time average can be neglected. The other, resonant terms including $\frac{\partial \varphi_i^0}{\partial \tau}$ must cancel each other out. Whether there are resonant terms or not depends on the frequencies present in the Fourier spectrum of the global coupling signal

$$a(t) = \int_0^\infty a(\Omega') \cos(\Omega' t + \varphi(\Omega')) d\Omega'.$$

Assume that there is only one frequency present in the coupling ($a(\Omega') = \delta(\Omega' - \Omega_{\mathrm{coup}})$) and $\varphi(\Omega_{\mathrm{coup}}) = 0$. Therefore $a(t) = \cos \Omega_{\mathrm{coup}} t$ and

$$
\begin{aligned}
a(\mathcal{T}/\delta) \sin((\Omega_p - \Omega_q)\mathcal{T}/\delta + \varphi_p^0 - \varphi_q^0) &= \frac{1}{2} \sin((\Omega_p - \Omega_q - \Omega_{\mathrm{coup}})\mathcal{T}/\delta + \varphi_p^0 - \varphi_q^0) \\
&+ \frac{1}{2} \sin((\Omega_p - \Omega_q + \Omega_{\mathrm{coup}})\mathcal{T}/\delta + \varphi_p^0 - \varphi_q^0).
\end{aligned}
$$

This is a fast oscillating term if $|\Omega_{\mathrm{coup}} - |\Omega_p - \Omega_q|| \gg \varepsilon$ and can be neglected in (2.9). For smaller values of $|\Omega_{\mathrm{coup}} - |\Omega_p - \Omega_q||$ the term becomes resonant. For the special case of $|\Omega_{\mathrm{coup}} - |\Omega_p - \Omega_q|| = 0$, the resonant term becomes

$$
a(t) \sin((\Omega_p - \Omega_q)t + \varphi_p^0 - \varphi_q^0) = \frac{1}{2} \sin(\varphi_p^0 - \varphi_q^0).
$$

In this case, (2.9) becomes

$$
\frac{\partial \varphi_i^1}{\partial t} = \delta_p^i \delta_q^j \frac{1}{2N} \sin(\varphi_j^0 - \varphi_i^0) + \delta_q^i \delta_p^j \frac{1}{2N} \sin(\varphi_j^0 - \varphi_i^0) - \frac{\partial \varphi_i^0}{\partial \tau}.
$$

To keep the right hand side nonresonant, $\frac{\partial \varphi_i^0}{\partial \tau}$ must fullfill the following conditions:

$$
\frac{\partial \varphi_{i \neq p,q}^0}{\partial \tau} = 0, \quad \frac{\partial \varphi_p^0}{\partial \tau} = \frac{1}{2N} \sin(\varphi_q^0 - \varphi_p^0), \quad \frac{\partial \varphi_q^0}{\partial \tau} = \frac{1}{2N} \sin(\varphi_p^0 - \varphi_q^0).
$$

Therefore, by introducing a frequency component $\Omega = \Omega_p - \Omega_q$ in the global coupling function $a(t)$, the phase shifts φ_l and φ_q can be selectively manipulated without influencing the other oscillators. This is possible simultaneously for all pairs of oscillators if the constraint

$$
||\Omega_q - \Omega_p| - |\Omega_j - \Omega_i|| \gg \varepsilon \text{ for } \{p, q\} \neq \{i, j\} \text{ and } p \neq q, i \neq j.
$$

is fulfilled. If the minimal difference δ' of two frequency *differences* in the network is defined by

$$
\delta' = \min_{\{p,q\} \neq \{i,j\}, \ p \neq q, i \neq j} ||\Omega_q - \Omega_p| - |\Omega_j - \Omega_i|| \tag{2.10}
$$

this amounts to

$$
\varepsilon \ll \delta'
$$

or, with an appropriate choice of frequencies (see 2.3.2),

$$
\varepsilon \ll \delta. \tag{2.11}
$$

The constraint (2.11) ensures that the frequency differences of two pairs of oscillators cannot coincide to produce spurious resonant terms in (2.9). Under this constraint the
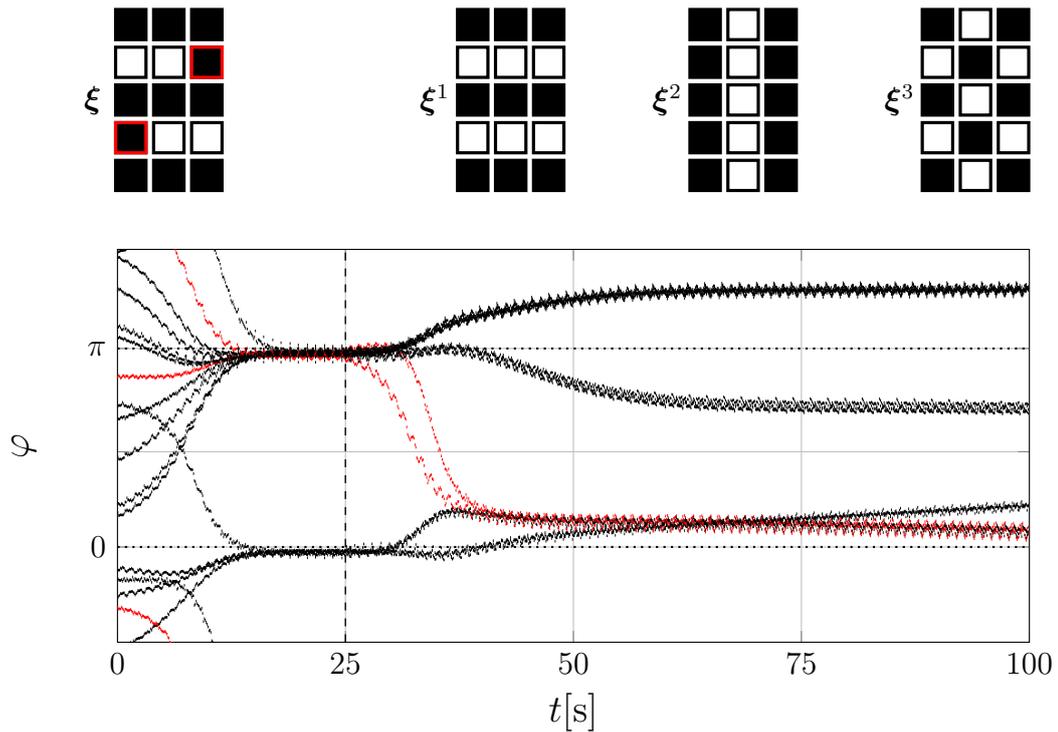
Fig. 5: Numerical time integration of the phase shifts in (2.8) with random initial conditions, $N = 15$ oscillators, $M = 3$ memorized patterns and two erroneous bits in the initial pattern. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. Both the squares and the phase shift curves belonging to erroneous bits in the initial pattern are marked red. Oscillator frequencies were chosen as a Golomb ruler (see next section) between $\Omega_{\min} = 1\,\text{kHz}$ and $\Omega_{\max} = 2\,\text{kHz}$. The coupling strength was set to $\varepsilon = 1$. $a(t)$ was chosen according to (2.12), where w_{ij} was determined by the patterns depicted above as described in section 2.2. The coupling was switched to recognition mode at $t = 25\,\text{s}$.

global coupling function

$$a(t) = \sum_{p=1}^{N} \sum_{q=1}^{N} w_{pq} \cos((\Omega_p - \Omega_q)t) \tag{2.12}$$

inserted in (2.9) yields

$$\frac{\partial \varphi_i^0}{\partial \tau} = \frac{1}{N} \sum_{i=1}^{N} w_{ij} \sin(\varphi_j^0 - \varphi_i^0).$$

Thus, the zero-order approximation for $\dot{\varphi}_i$ in normal time reads:

$$\dot{\varphi}_i = \varepsilon \left(\frac{1}{N} \sum_{j=1}^{N} w_{ij} \sin(\varphi_j - \varphi_i) + \mathcal{O}\left(\frac{\varepsilon}{\delta}\right) \right)$$

Apart from the different time scale, the approximate dynamics of a system of weakly globally coupled oscillators with different frequency (2.8) is identical to the dynamics of a system of coupled oscillators with the same frequency and individual strong coupling (2.2), if the constraint (2.11) is fullfilled and the coupling function $a(t)$ is given by (2.12). Fig. 5 depicts the simulation results of a pattern recognition with the patterns also used in Fig. 3, where $\Omega_{\min} = 1\,\text{kHz}$, $\Omega_{\max} = 2\,\text{kHz}$ and $\varepsilon = 1$. The details of frequency selection are

described in 2.3.2. Note that the average dynamics depicted in Fig. 5 differs qualitatively from the dynamics shown in Fig. 3 in two aspects. First, the unstable steady state is destroyed much quicker compared to Fig. 3, which is expected because of the fluctuations that are present in the weakly coupled case. Second, there is a continuous drift in some of the phase shifts. This drift settles for very long times (not shown in the picture) at a state which does not necessarily correspond to the recognized pattern any more (not even in the sense of (2.6)). The reason is that *all* steady states close to patterns are not particularly robust (see 4.1.2).

2.3.2   Choosing suitable frequencies

A set of integer numbers, where each difference of any two numbers in the set occurs only once, is called a Golomb ruler (Golomb 1997). Finding a set of N frequencies Ω_i that satisfy condition (2.11) is equivalent to finding a Golomb ruler of order N and rescaling the numerical values appropriately to fit the available frequency range.

An example for a Golomb ruler of order $N = 4$ is given by the sequence

$$0, 1, 4, 6.$$

These numbers produce the six differences $1 - 0 = 1$, $6 - 4 = 2$, $4 - 1 = 3$, $4 - 0 = 4$, $6 - 1 = 5$ and $6 - 0 = 6$. They form a *perfect* Golomb ruler because the length L of the ruler (i.e the difference between the highest and the lowest number) is equal to the number of pairs that can be chosen from the set:

$$L_{\mathrm{perfect}} = \frac{N^2 - N}{2}$$

However, it is proven that no perfect Golomb rulers exist for $N > 4$. For higher orders, there exist *optimal* rulers with a minimal length $L_{\mathrm{optimal}} > L_{\mathrm{perfect}}$. There are no known search algorithms for optimal Golomb rulers that only take polynomial time (Drakakis 2009). As a consequence, until now, only Golomb rulers up to $N = 26$ have been found, while the search for $N = 27$ is currently underway (the progress can be seen at `http://www.distributed.net/OGR`). Fig. 6 plots the length of the known optimal rulers versus the order. The length scales worse than L_{perfect} but considerably better than N^2. While there are no known efficient algorithms to find optimal Golomb rulers, *near optimal* Golomb rulers with

$$L_{\mathrm{near\ optimal}} < N^2$$

have been found for $N \leq 65000$ (Apostolos 2002). This gives an estimate of the minimal separation between two frequencies that can be achieved in a Golomb ruler of length N. The ratio of the minimal distance of two marks on the ruler (i.e. 1) to the length of the ruler is
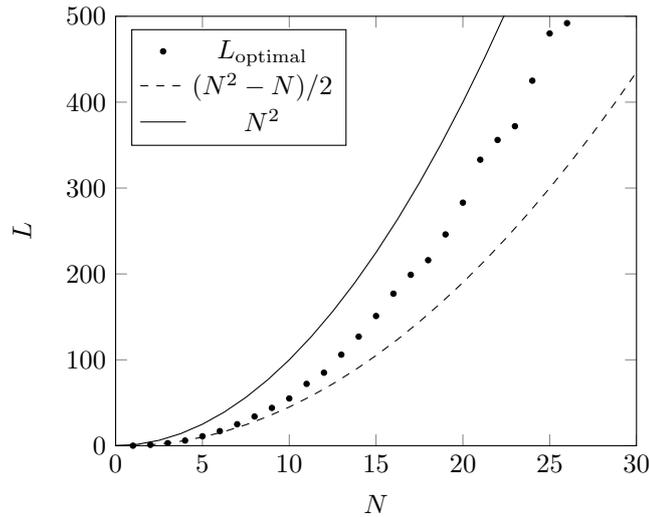
Fig. 6: Length of known optimal Golomb rules vs. the number of marks on the ruler, with differently scaling polynomials for reference.

above or equal to $1/N^2$ if $N \leq 65000$. Rescaling the integer distances on the Golomb ruler to the available range of frequencies yields

$$\min_{i \neq j} |\Omega_i - \Omega_j| \geq \frac{\Omega_{\max} - \Omega_{\min}}{N^2}. \tag{2.13}$$

Therefore, if the frequencies represent a Golomb ruler of length N, the constraint (2.11) is fulfilled if

$$\varepsilon \ll \frac{\Omega_{\max} - \Omega_{\min}}{N^2}. \tag{2.14}$$

This condition is stricter than necessary, because larger values of ε still lead to a successful pattern recognition. A large minimal separation of two frequencies and therefore large ε is desirable because it increases both convergence speed and tolerance of the network for frequency inaccuracies (the reason for that is explained in 4.2.1).

2.4   Phase shift equations for networks of weakly coupled realistic oscillators

The experiments presented in chapter 6. were conducted with electrical oscillators that are described by two differential equations each. The dynamics of the phases and phase shifts is extracted from these equations using a method introduced by Kuramoto, applicable to weakly coupled oscillators. The method is described in 2.4.1. Section 2.4.2 shows that for the specific case of oscillators that exhibit an almost harmonic oscillation, global coupling in one variable leads to phase shift dynamics equivalent to (2.2).

2.4.1   General case

The dynamics of a general nonlinear oscillator is given by a dynamical system

$$\dot{\boldsymbol{X}} = \boldsymbol{f}(\boldsymbol{X}) \tag{2.15}$$

that has a stable limit cycle solution $\boldsymbol{X}_{\mathrm{LC}}(t + T)) = \boldsymbol{X}_{\mathrm{LC}}(t)$, where $T = 2\pi/\Omega$ is the period of the limit cycle. On the limit cycle itself, a uniform phase variable ϑ can be defined with

$$\vartheta(t) = \vartheta(\boldsymbol{X}_{\mathrm{LC}}(t)) = \Omega t + \mathrm{const}, \tag{2.16}$$

$$\dot{\vartheta} = \nabla_{\boldsymbol{X}}\vartheta \cdot \frac{d\boldsymbol{X}}{dt} = \boldsymbol{Z}(\vartheta) \cdot \boldsymbol{f}(\boldsymbol{X}) = \Omega, \tag{2.17}$$

where (2.17) follows from the chain rule. $\boldsymbol{Z}(\vartheta)$ is called the phase-dependent sensitivity of the oscillator. If the current phase of the oscillator is ϑ and the oscillator is subject to a small external perturbation $\mathrm{d}\boldsymbol{X}$ this will result in a jump in phase given by

$$\mathrm{d}\vartheta = \boldsymbol{Z}(\vartheta) \cdot \mathrm{d}\boldsymbol{X} \tag{2.18}$$

The vector $\boldsymbol{Z}(\vartheta) = \nabla_{\boldsymbol{X}}\vartheta$ is only meaningful if there is a phase $\vartheta(\boldsymbol{X})$ defined outside the limit cycle. An asymptotic phase (Kuramoto 1984) $\vartheta(\boldsymbol{P})$ for an arbitrary point $\boldsymbol{P}$ in the basin of attraction of the limit cycle can be defined in the following way (see also Fig. 7):

- If $\boldsymbol{P} \in \boldsymbol{X}_{\mathrm{LC}}$, $\vartheta(\boldsymbol{P})$ is given by (2.16).

- If $\boldsymbol{P} \notin \boldsymbol{X}_{\mathrm{LC}}$, find a point $\boldsymbol{P}' \in \boldsymbol{X}_{\mathrm{LC}}$ such that

$$\lim_{t \to \infty} \|\boldsymbol{P}'(t) - \boldsymbol{P}(t)\| = 0,$$

where $\boldsymbol{P}(t)$ and $\boldsymbol{P}'(t)$ are the trajectories emanating from the initial conditions $\boldsymbol{P}$ and $\boldsymbol{P}'$ respectively. Since $\boldsymbol{P}$ is in the basin of attraction of $\boldsymbol{X}_{\mathrm{LC}}$, it is always possible to
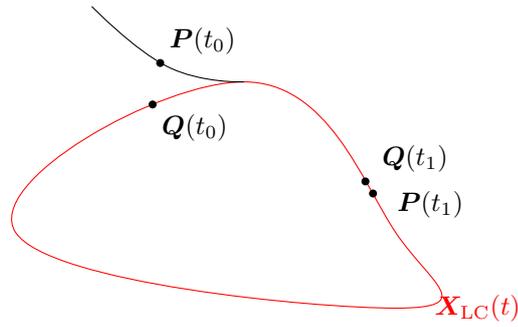
Fig. 7: Attractive limit cycle of (2.15) (red) with a trajectory in its basin of attraction (black). Initial states that do not lie on the limit cycle (like $\boldsymbol{P}(t_0)$) will evolve towards states on the limit cycle ($\boldsymbol{P}(t_1)$). The phase of $\boldsymbol{P}(t_0)$ is equal to the phase of $\boldsymbol{Q}(t_0)$ if $\boldsymbol{P}(t)$ and $\boldsymbol{Q}(t)$ coincide for $t \gg t_0$

find such a point $\boldsymbol{P}'$. The phase of $\boldsymbol{P}'$ is given by (2.16). It is identical to the phase of $\boldsymbol{P}$:

$$\vartheta(\boldsymbol{P}) = \vartheta(\boldsymbol{P}')$$

Now consider a system of N oscillators with frequencies Ω_i subject to a weak coupling

$$\dot{\boldsymbol{X}}_{\boldsymbol{i}} = \boldsymbol{f}_i(\boldsymbol{X}_i) + \varepsilon \boldsymbol{g}_i(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N, t), \quad i = 1 \ldots N, \tag{2.19}$$

where $\varepsilon \ll \Omega_i$. In analogy to (2.17), the phase dynamics in this system is given by

$$\dot{\vartheta}_i = \Omega_i + \varepsilon \boldsymbol{Z}_i(\vartheta_i) \cdot \boldsymbol{g}_i(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N, t). \tag{2.20}$$

This is exact, however it is generally not possible to find an analytic expression for $\boldsymbol{Z}(\vartheta)$. Usually, $\boldsymbol{Z}(\vartheta)$ is determined numerically in a simulation or experimentally (see 3.3.1 in the chapter on methods).

2.4.2 Phase shift equations for a network of weakly coupled near-harmonic oscillators with a global coupling in one variable

This section gives a derivation why both the waveform and the phase-dependent sensitivity of a weakly nonlinear, near-harmonic oscillator take sinusoidal shape. Later on, it is shown how this can be used in conjunction with a suitable global coupling to form a network with properties equivalent to the networks presented in previous sections.

The weakly nonlinear oscillator

$$\begin{aligned} \dot{X} &= -Y + \mu f(X, Y) \\ \dot{Y} &= X + \mu g(X, Y) \end{aligned}$$

with $\mu \ll 1$ has the solution (Strogatz 1994)

$$
\begin{aligned}
X(t) &= r(\tau)\cos(t + \varphi(\tau)) \\
Y(t) &= r(\tau)\sin(t + \varphi(\tau)),
\end{aligned}
\tag{2.21}
$$

where $r(\tau) = 1 + \mathcal{O}(\mu)$, $-\pi \le \varphi(\tau) < \pi$ are varying on the slow time scale $\tau = \mu t$. The instantaneous phase in the sense of (2.16) is given by $\vartheta = t + \varphi(\tau) + \mathcal{O}(\mu) = \tan^{-1}(Y/X) + \mu h(X, Y)$. Therefore, if (2.21) is subjected to an infinitesimal jump $(\mathrm{d}X, \mathrm{d}Y)$ in both variables, there is a change in phase:

$$
\mathrm{d}\vartheta = \tan^{-1}\left(\frac{Y(t) + \mathrm{d}Y}{X(t) + \mathrm{d}X}\right) + \mu h(X + \mathrm{d}X, Y + \mathrm{d}Y) - \tan^{-1}\left(\frac{Y(t)}{X(t)}\right) - h(X, Y)
$$

$$
\mathrm{d}\vartheta = \tan^{-1}\left(\frac{r\sin(t + \varphi) + \mathrm{d}Y}{r\cos(t + \varphi) + \mathrm{d}X}\right) + \mu h(X + \mathrm{d}X, Y + \mathrm{d}Y) - \tan^{-1}\left(\frac{r\sin(t + \varphi)}{r\cos(t + \varphi)}\right) - \mu h(X, Y)
$$

A Taylor expansion to first order in $\mathrm{d}X$ and $\mathrm{d}Y$ leads to

$$
\mathrm{d}\vartheta = -(\sin(t + \varphi) + \mathcal{O}(\mu))\frac{\mathrm{d}X}{r} + (\cos(t + \varphi) + \mathcal{O}(\mu))\frac{\mathrm{d}Y}{r}.
\tag{2.22}
$$

With smaller μ, (2.22) becomes arbitrarily close to

$$
\mathrm{d}\vartheta = -\sin\vartheta\,\mathrm{d}X + \cos\vartheta\,\mathrm{d}Y.
$$

A comparison with (2.18) shows that in the limit of weak nonlinearity, the phase-dependent sensitivity for near-harmonic oscillators is given by

$$
\boldsymbol{Z}(\vartheta) = \begin{pmatrix} -\sin\vartheta \\ \cos\vartheta \end{pmatrix}
$$

if the position in phase space is given by

$$
\boldsymbol{X}(\vartheta) = \begin{pmatrix} \cos\vartheta \\ \sin\vartheta \end{pmatrix}.
$$

The next step is to couple N such oscillators with different frequencies Ω_i, $i = 1 \ldots N$. In order to achieve a dynamics equivalent to (2.2), a suitable weak coupling function (i.e. $\boldsymbol{g}$ in (2.19)) is needed. One possibility to choose the coupling is given by

$$
\boldsymbol{g}(X_1, \ldots, X_N, Y_1, \ldots, Y_N) = a(t)\frac{1}{N}\sum_{j=1}^{N} Y_j(\vartheta_j(t)),
$$

where $a(t)$ is given by (2.12). Using this coupling and the expressions obtained earlier for $\boldsymbol{X}(\vartheta)$ and $\boldsymbol{Z}(\vartheta)$, the general equation (2.20) transforms to

$$\dot{\vartheta}_i = \Omega_i + \varepsilon a(t) \frac{1}{N} \sum_{j=1}^{N} Z_i(t) Y_j(t). \tag{2.23}$$

Here, $Z_j(t) = \sin \vartheta_j(t)$ is the second component of $\boldsymbol{Z}_j$. Since there is only a coupling in the Y variable, $Z_j(\vartheta_j)$ contains all information that is needed about the phase-dependent sensitivity. From here on, the functions $Z(\vartheta_j)$ are referred to as the *phase response curves* of the respective oscillators. Entering $Z_j(t)$ and $Y(t)$ into (2.23) yields

$$\dot{\vartheta}_i = \Omega_i + \varepsilon a(t) \frac{1}{N} \sum_{j=1}^{N} \sin \vartheta_j \cos \vartheta_i. \tag{2.24}$$

This equation can be analyzed in the same manner as (2.7) in section 2.3.1. The final result is (Hölzel & Krischer 2011)

$$\dot{\varphi}_i = \varepsilon_{\mathrm{eff}} \left(\frac{1}{N} \sum_{j=1}^{N} w_{ij} \sin(\varphi_j - \varphi_i) + \mathcal{O}\left(\frac{\varepsilon_{\mathrm{eff}}}{\delta}\right) \right) \tag{2.25}$$

where $\varepsilon_{\mathrm{eff}} = \varepsilon/2$ and, additionally to $\varepsilon_{\mathrm{eff}} \ll \delta$, also

$$||\Omega_q + \Omega_p| - |\Omega_j - \Omega_i|| \gg \varepsilon_{\mathrm{eff}} \text{ for } \{p, q\} \neq \{i, j\} \text{ and } p \neq q, i \neq j \tag{2.26}$$

must hold. A simple way to guarantee this condition is fulfilled is to pick frequencies such that $3\Omega_{\min} - \Omega_{\max} \gg \varepsilon_{\mathrm{eff}}$. Again, apart from the different time scale, (2.25) is equivalent to (2.2). As it turns out, there is a difference in the slower time scale to the ideal weakly coupled network, which results in a qualitatively different dynamics of the phase shifts. In 4.2.5, a multiple time scale analysis of 2.24 with three time scales is done which elaborates on this difference.

3. Methods

3.1   Numerical

3.1.1   Time integration of ODEs

For all numerical integrations of differential equations discussed in the thesis, the equations were implemented in a straightforward C++ program.   All variables and parameters had double precision.   The Livermore ODE solver (Radhakrishnan & Hindmarsh 1993) available from `http://www.netlib.org/odepack` was used for numerical integration.   The flag indicating the integration method was set to `MF=22`, indicating a stiff ODE with a Jacobian internally generated by the solver.   Both, the relative and the absolute error tolerances were set to $1 \cdot 10^{-7}$. For all other settings, standard values were used. The solver was compiled into a library with the GNU gfortran compiler. The library was then linked into the C++ program with the GNU g++ compiler. Appendix A2 gives an example code listing and describes the compiling and linking steps.

3.1.2   Simulations of circuit behavior with SPICE

SPICE (Nagel & Pederson 1973) is a state of the art simulation software for electronic circuits. It realistically models the behavior of circuit elements like operational amplifiers and diodes. The results of SPICE simulations predict the behavior of a circuit more reliably than the numerical integration of differential equations derived from applying Kirchhoff's laws to an ideal circuit.

Before the experiments were conducted, the circuitry was tested with a SPICE model for stability and for plausibility of the observed currents and voltages.   For each non basic circuit element, the SPICE model provided by the vendor was used, with the exception of the tunnel diode, which was modeled as a nonlinear current source with a polynomial characteristic. The characteristic was a fit of a fifth order polynomial to the data depicted in Fig. 9 further below.

There are different implementations of the SPICE engine. For this thesis, LTspiceIV, which is obtainable at `http://www.linear.com/designtools/software`, was chosen.   Appendix A3 shows SPICE netlists for both types of oscillators that were used in the experiment as well as the commands to run the simulations.

3.1.3   Extraction of the period and the phase shift from waveform data

Both numerical and experimental data originally were either available as a time series of the phase shifts (in simulations of networks of phase oscillators) or as time series of an oscillating variable (in the experiment). In the second case, one or more steps of processing

the data were necessary to be able to plot the phase shift curves. The procedure is first described in general, before its application for the data presented in this thesis is discussed.

It is impossible to compute a phase shift from a waveform without deciding what is considered the intrinsic period $T = 2\pi/\Omega$ of an oscillator. The reason is that in principle the instant phase shift $\varphi(t)$ of a signal is computed in the following way:

- Determine the current phase $\vartheta(t)$ of the signal.

- Subtract t from an integral multiple k of the intrinsic period T, such that $-0.5 \cdot T \leq kT - t < 1.5 \cdot T$. This will result in $-0.5 \cdot \pi \leq \varphi(t) < 1.5 \cdot \pi$.

- The phase shift is now given by $\varphi(t) = \vartheta(t) + 2\pi(kT - t)/T$.

Note that it is not always evident which is the intrinsic period. For example, in a system of coupled oscillators, it is entirely possible that the frequencies change due to the coupling. The data then could be interpreted as both a drift in the phase shift at the old frequency or a constant phase shift at the new frequency. Arguments can be made for both interpretations; the first one clearly separates coupling effects from the intrinsic dynamic of the oscillator, while the second one extracts the relevant information more clearly under some circumstances (this is illustrated with Fig. 24 in section 4.2.1).

The problem of finding the instant phase of the signal remains in both cases. Theoretically, it is possible to obtain a continuous time series of the phase by using a Hilbert transform on the data (a detailed discussion of this approach can be found in Pikovsky (2001)). However, this high time resolution is not needed for the experiments and comes at a considerable computational cost. Therefore, the following path was taken in this thesis:

Let t and U be two vectors representing data of an experimental voltage time series in computer memory, where U_i represents the value of the voltage at the time t_i. From this data, the times of zero crossings t_{cross_i} with positive slope were linearly interpolated (compare Fig. 8):

$$t_{\text{cross}_i} = t_{n_i} - \frac{t_{n_i+1} - t_{n_i}}{U_{n_i+1} - U_{n_i}} U_{n_i}, \quad \text{for} \ \ (n_i | \ t_{n_i} < 0 \wedge t_{n_i+1} \geq 0)$$

The intrinsic period of the oscillation was determined by averaging the time between two zero crossings in a certain time window:

$$T = \frac{t_{\text{cross}_{n+N}} - t_{\text{cross}_n}}{N}$$

The time window was chosen such that the oscillators were either uncoupled, or had settled on the steady state during initialization. Otherwise, the transient changes in the phase shifts due to the coupling could have lead to a biased average value.
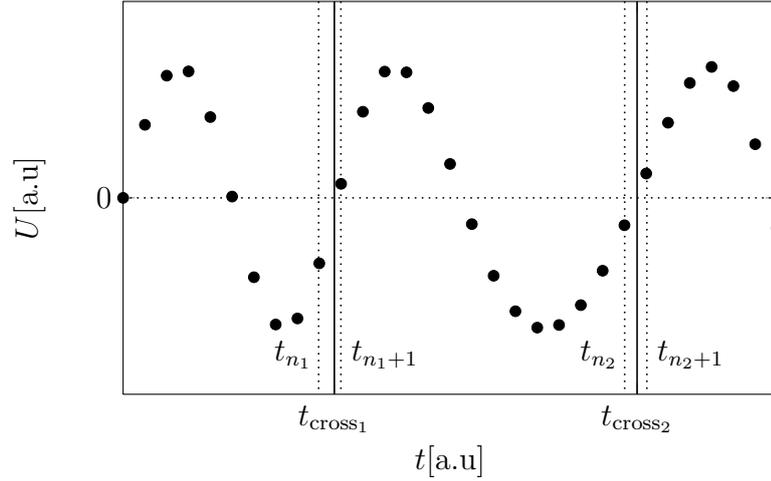
Fig. 8: Extraction of the momentary phase $\vartheta(t)$ from a measured voltage signal (black dots) at discrete points in time. The times of zero crossings with positive slope (t_{cross_i}) are interpolated linearly from neighboring data points (see text). $\vartheta(t_{\mathrm{cross}_1}) = \vartheta(t_{\mathrm{cross}_2}) = 0$.

Even without actually performing the Hilbert transform, the advance in phase between two zero crossings of the signal is known to be π, if the signal is sufficiently close to sinusoidal shape and has no spurious zero crossings. In this thesis, zero crossings with positive slope were assigned the phase value $\vartheta = 0$, fixing the phase of zero crossings with negative slope at $\vartheta = \pi$. With this assignment, the phase of the signal is known at the time of all zero crossings. Only zero crossings with positive slope were taken into account. The phase shifts $\varphi(t_{\mathrm{cross}_i})$ were determined from $\vartheta(t_{\mathrm{cross}_i}) = 0$ as described above. This whole procedure is only necessary if one is interested in a time series of the phase shifts. For example, for pattern recognition, it is sufficient to determine one zero crossing before and one zero crossing sufficiently long after the coupling function was switched, to determine whether the phase shift has changed by π or not.

3.1.4   Extraction of the total harmonic distortion from waveform data

One method to characterize the deviation of an oscillating signal from a sine wave often used in electronics is the total harmonic distortion (THD) (see for example Shmilovitz (2005) for a thorough discussion). The THD describes the ratio of the amplitude of the higher harmonics of a signal $x(t)$ to the amplitude of the ground frequency:

$$x(t) = \sum_{n=1}^{\infty} a_n \sin(n\Omega t + \varphi_n), \quad \mathrm{THD}(x(t)) = \frac{\sqrt{\sum_{n=2}^{\infty} a_n^2}}{a_1}$$

For an ideal sine wave, the THD is zero.

To obtain the THD from a time series the Fourier coefficients a_n were computed for

$n = 1 \ldots 10$ by numerically evaluating the expression

$$a_n = 2\sqrt{\left(\int \sin(n\Omega t)x(t)\mathrm{d}t\right)^2 + \left(\int \cos(n\Omega t)x(t)\mathrm{d}t\right)^2}$$

where the integrals were taken over one period of the signal. These first ten harmonics were then used to compute the approximate THD.

3.1.5   Curve fitting

For the fits in shown in Fig. 9, Fig. 10, Fig. 15, Fig. 31, Fig. 64 and Fig. 65, the `fit` command of the open source software Gnuplot (`http://www.gnuplot.info`) was used with standard settings. This command works with the Levenberg-Marquardt algorithm (Marquardt 1963).

3.2 Electronic circuitry

3.2.1 Circuit elements

Tunnel diode

A tunnel diode (Esaki 1958) is a special type of diode, which exhibits a region of negative differential resistance in its current voltage characteristic; Fig. 9 shows the measured $I - U$-characteristic of the type of tunnel diode (part number 40571, available at `http://www.americanmicrosemiconductor.com`) used in this thesis. The tunnel diodes were used to enable sustained oscillations in oscillatory circuits. The schematic symbol of the tunnel diode is depicted in Fig. 11. A serious drawback of tunnel diodes is that they are somewhat hard to get by, since they are not produced any more. As a consequence, no vendor model for SPICE was available for the tunnel diode; instead, a voltage dependent current source with the polynomial characteristic also shown in Fig. 9 was used.
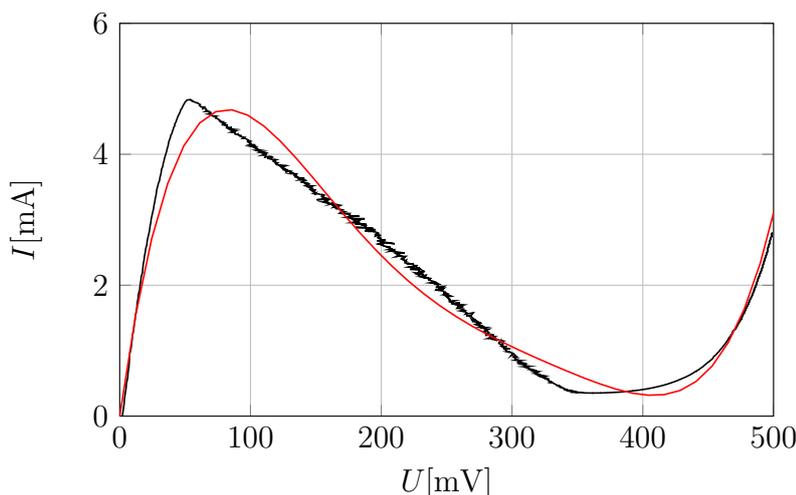
Fig. 9: Current-voltage characteristic of the 40571 tunnel diode used for the experiments (black curve). The curve was measured with a Jaissle PGU-10V-1A-IMP-S potentiostat. Note the wrinkles on the branch of negative slope, testament to the instability of the steady state, which was constantly readjusted by the potentiostat during the voltage sweep. The red curve shows a polynomial fit of fifth order, given by $I(U) = \sum_{n=1}^{5} a_i U^n$ to the characteristic. The coefficients were $a_1 = 1.434\,61 \cdot 10^{-4}\,\mathrm{A\,V^{-1}}$, $a_2 = -1.482 \cdot 10^{-3}\,\mathrm{A/V^2}$, $a_3 = 6.139\,39 \cdot 10^{-4}\,\mathrm{A/V^3}$, $a_4 = 1.177\,77 \cdot 10^{-2}\,\mathrm{A/V^4}$ and $a_5 = 8.657\,29 \cdot 10^{-3}\,\mathrm{A/V^5}$.

Voltage controlled resistor

For this thesis, an n-channel junction field effect transistor (JFET, for a discussion see (Tietze 2008)) of type VCR2N was used as a voltage controlled resistor (VCR). This particular transistor is specifically designed to be used in the linear regime of the current voltage characteristic, where the device behaves as a resistor between source and drain (see Fig. 10a). The resistance R_{SD} is governed by the gate-source voltage U_{GS}; the dependence is shown in Fig. 10b. The schematic symbol of the JFET is depicted in Fig. 11.
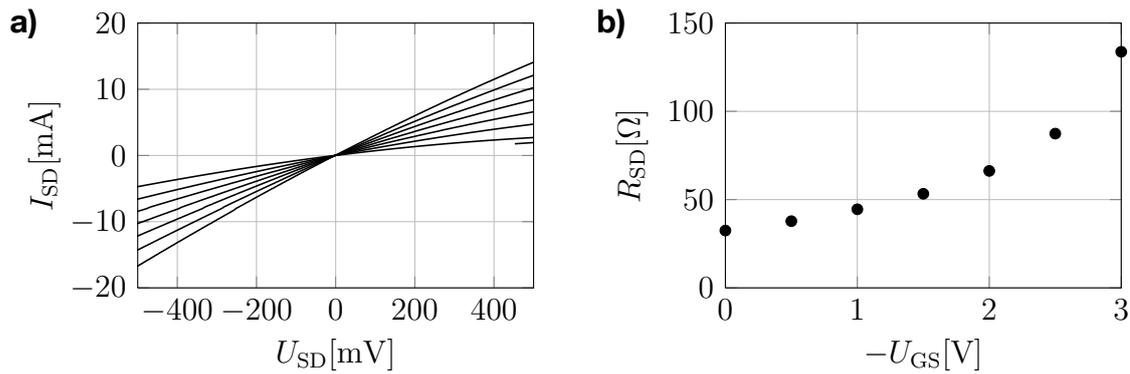
Fig. 10: a) Source-drain current vs. source-drain voltage of a VCR2N JFET for different values of the gate-source voltage U_{GS}. U_{GS} was decreased from $0\,V$ (curve with the largest slope) to $-3\,V$ (curve with the smallest slope) in steps of $-0.5\,V$. The curves were measured with a Jaissle PGU-10V-1A-IMP-S potentiostat. b) Inverse slope of linear fits to the curves in a), approximately equal to the source-drain resistance of the JFET, vs. the gate-source voltage.

Like the tunnel diode, the VCR2N is a discontinued device, but it is replacable by current JFETs.

Operational Amplifier

An operational amplifier, or op-amp, is a device with two high impedance inputs (called the non-inverting and the inverting input) and a low impedance output. The potential at the output is given by $U_{out} = A \cdot (U^+ - U^-)$, where U^+ and U^- are the potentials at the non-inverting input and the inverting input, respectively, and A is a huge gain factor (usually $A > 10^4$ - A depends on the frequency of the input signals, however). If the output voltage is fed back to the inputs in a suitable manner, any circuit containing an op-amp can be treated by applying the two "golden rules" (Horowitz & Hill 1989):

- The inputs draw no current.

- The output voltage will be such that both inputs are at the same potential.

For this thesis, op-amps were used as voltage follower, differential amplifier, nonlinear impedance converter and gyrator, all of which are presented briefly below. The schematic symbol of the op-amp is depicted in Fig. 11.
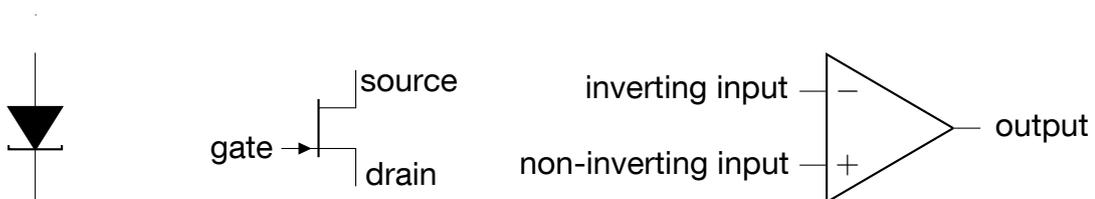
Fig. 11: Circuit symbols for tunnel diode, n-channel JFET and op-amp (from left to right).

3.2.2   Subcircuits

The following five subcircuits that make use of op-amps were used at least once in the experimental network of oscillators. They can be understood by applying the two rules mentioned above. A more detailed description for all of them can be found, for example, in Tietze et al. (2008).

Voltage follower

This device mirrors the input voltage U_{in} at its output: $U_{\mathrm{out}} = U_{\mathrm{in}}$ (see Fig. 12a). No current is drawn from the circuit providing U_{in}. In the experiments, a voltage follower was used to drive the VCR without influencing the dynamics of the coupled oscillators.

Differential amplifier with gain 1

The output of this device is the difference of its inputs: $U_{\mathrm{out}} = U_2 - U_1$ (see Fig. 12b). In the experiments, a differential amplifier was used to compute the effective gate-source voltage of the VCR.

Negative impedance converter (NIC)

This device behaves as if it were a negative resistance to ground: $I = -U/R_{\mathrm{neg}}$ (see Fig. 12c). In the experiments, NICs where used both as part of the nonlinear element of oscillatory circuits and to enhance the range of the coupling, including negative values of $\varepsilon_{\mathrm{eff}}$.

Gyrator

This device acts as an inductance, with $L = C_L \cdot 1\,\mathrm{k\Omega} \cdot R_L$ (see Fig. 12d). In the experiments, gyrators where used instead of regular inductances (i.e. coils), because they are easily tunable.

Inverting adder

The output of this device is the negative sum of its inputs: $U_{\mathrm{out}} = -\sum_{i=1}^{N} U_i$ (see Fig. 12e). In the experiments, an adder was used to sum over all oscillator voltages, which is necessary for one particular implementation of the weak coupling mechanism.

3.2.3   Van der Pol oscillator

Van der Pol oscillators are LC circuits enhanced by an active nonlinear device that drives the oscillation. A detailed description of the oscillation mechanism can be found for example in Kanamaru (2007). The two variants that were used in this thesis are
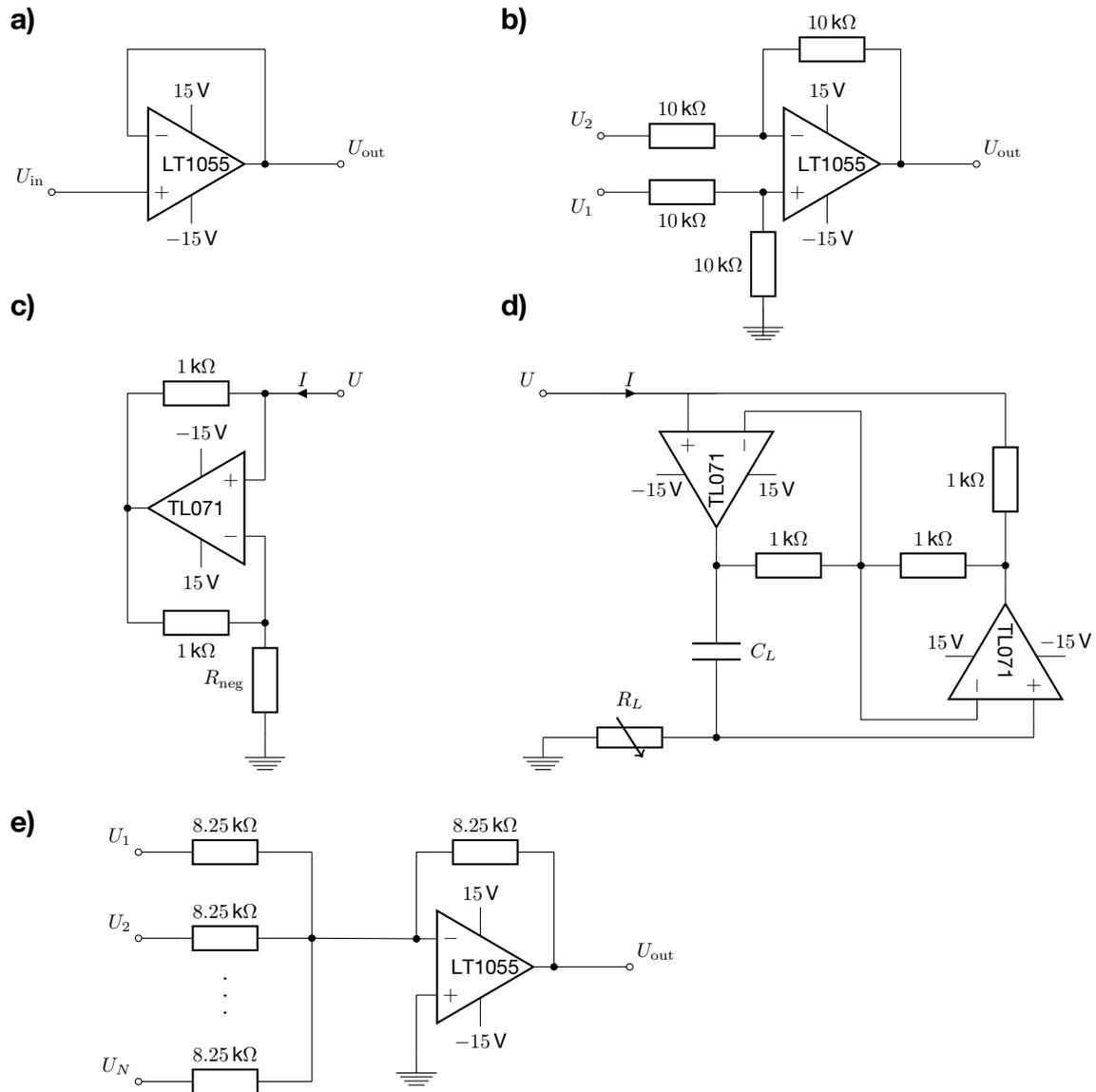
Fig. 12: Diagrams of some subcircuits used in the experiments: **a)** Voltage follower. **b)** Difference amplifier. **c)** Negative impedance converter. **d)** Gyrator. **e)** Inverting voltage adder.

- a van der Pol oscillator with a tunnel diode in series with a voltage source as nonlinear element (see Fig. 13) and

- a van der Pol oscillator with a NIC in parallel connection with two diodes as nonlinear element and a gyrator as tunable inductance (see Fig. 14).

The first design will be referred to as *TD-type* oscillator, the second as *NIC-type* oscillator from here on. If the current-voltage characteristic of the nonlinear element is given by $I_{\mathrm{NL}}(U)$, the dynamical equations of the oscillator are

$$\dot{U} = \frac{1}{C}\left(-I_{\mathrm{NL}}(U) - I\right)$$
$$\dot{I} = \frac{U}{L} \tag{3.1}$$
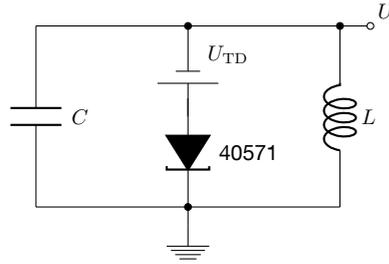
Fig. 13: Circuit diagram of a TD-type oscillator. U is a nearly sinusoidal signal if the specifications if the values for C, L and U_{TD} are chosen appropriately (see 5.1).

Here, U is the potential at the respective circuit nodes indicated in Fig. 13 and Fig. 14, I is the current through the inductance. For sustained oscillations, $I_{\mathrm{NL}}(U)$ must have a region of negative differential resistance around $U = 0$ (i.e. $\mathrm{d}I_{\mathrm{NL}}/\mathrm{d}U|_{U=0} = -1/R_{\mathrm{neg}}$ with $R_{\mathrm{neg}} > 0$). If this is true, the damping is negative and the steady state $(U, I) = (0, I_{\mathrm{NL}}(0))$ of (3.1) is unstable and the system exhibits stable long term oscillations (i.e. the attractor of the system is a stable limit cycle). The frequency of the oscillations is close to the frequency of the ideal LC-circuit, namely $f_0 = 1/2\pi\sqrt{LC}$. The shape of the oscillations as well as the deviation of the actual frequency f from f_0 depend on the ratio of time scales, given by $\gamma = C/L$ (note that usually this parameter is denoted by ε, which, however, is used in this thesis to denote the coupling strength between oscillators) and on $1/R_{\mathrm{neg}}$. If $\gamma \ll 0$, oscillations tend to be relaxational, far from a sinusoidal shape. This is not the case if $1/R_{\mathrm{neg}}$ is sufficiently close to zero (i.e. if the system is close to the Hopf bifurcation). A comprehensive numerical analysis of the parameter regions with sinusoidal oscillations can be found in section 5.1. The analysis presented there also takes into account effects of non-ideal circuit behavior that are not included in (3.1).

Fig. 15a shows the current-voltage characteristic of the tunnel diode. The slope of the branch of negative differential resistance corresponds to $R_{\mathrm{neg}} \approx 60\,\Omega$. To move the region
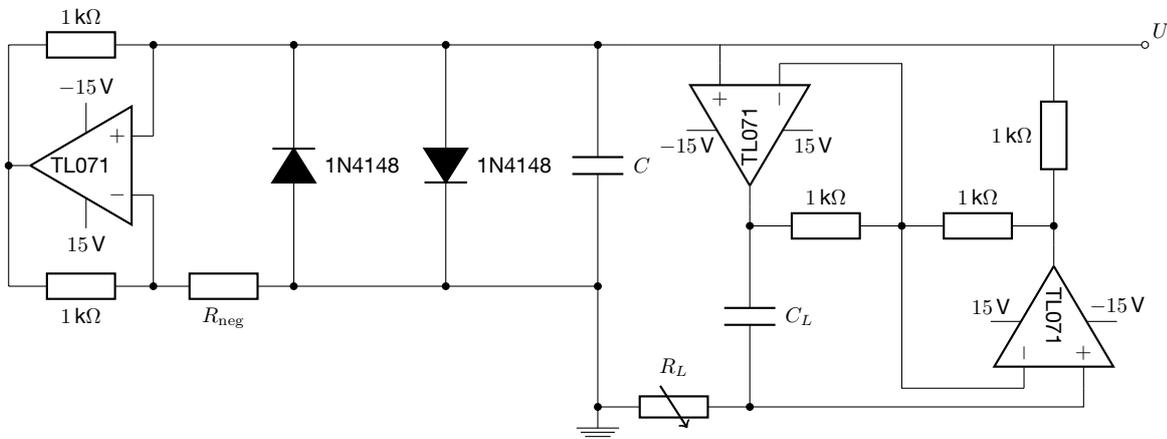
Fig. 14: Circuit diagram of a NIC-type oscillator. U is a nearly sinusoidal signal if the specifications of the TL071 amplifier are met, and if the values for C, C_L and R_L are chosen in compliance with the specifications of the TL071.
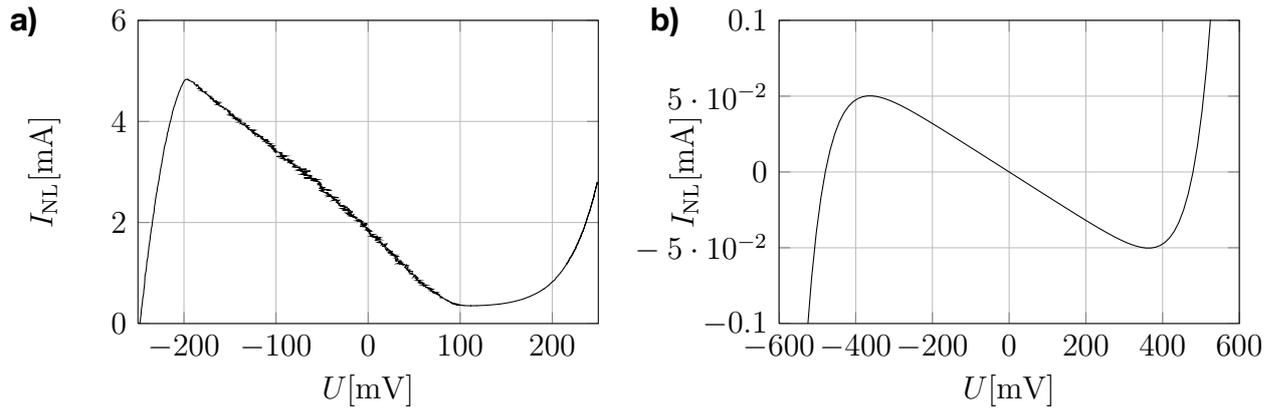
Fig. 15: a) Measured current-voltage characteristic of a 40571 tunnel diode shifted by $U_{\mathrm{TD}} = 250\,\mathrm{mV}$. This is the characteristic of the nonlinear element in Fig. 13. **b)** Current-voltage characteristic of the nonlinear element in the circuit shown in Fig. 14 with $R_{\mathrm{neg}} = 6.19\,\mathrm{k\Omega}$. The characteristic for the 1N4148 diode was measured with a Jaissle PGU-10V-1A-IMP-S potentiostat and then fitted to the Shockley equation (Shockley 1949): $I_{\mathrm{1N4148}}(U) = I_{\mathrm{S}}(\exp(U/U_{\mathrm{T}}) - 1)$ with U_{T} and I_{S} as parameters. The results of the fit were $U_{\mathrm{T}} = 52.39\,\mathrm{mV}$ and $I_{\mathrm{S}} = 8.235\,\mathrm{nA}$. The plot shows $I_{\mathrm{NL}} = I_{\mathrm{1N4148}}(U) - I_{\mathrm{1N4148}}(-U) - U/6.19\,\mathrm{k\Omega}$.

of negative differential resistance to $U = 0$, a voltage source must be connected in series to the tunnel diode. For this voltage source, values in the range $U_{\mathrm{TD}} \in [200\,\mathrm{mV}, 300\,\mathrm{mV}]$ were used throughout the experiments.

Fig. 15b shows the current-voltage characteristic of a nonlinear element consisting of an NIC with $R_{\mathrm{neg}} = 6.19\,\mathrm{k\Omega}$, in parallel with two diodes connected in opposite direction. The effect of this arrangement is that the negative slope provided by the NIC is capped at high voltages by the exponential shape of the characteristic of the diodes.

3.3   Experimental

3.3.1   Phase response curve measurements

To obtain experimental phase response curves, a van der Pol oscillator was connected to an external voltage source U_{ext} via a resistor R (Fig. 16 shows the circuit for a TD-type oscillator). The external voltage was provided by a Spectrum M2i.6021 digital/analog converter PCI card with a sample rate of 50 MS/s. The differential equation for the evolution of U in the extended circuit is given by

$$\dot{U} = \frac{1}{C}\left(-I_{\mathrm{NL}}(U) - I - \frac{U - U_{\mathrm{ext}}}{R}\right).$$

The resistance R is also part of the circuit employed for pattern recognition experiments, namely $R = 500\,\Omega$ for the TD-type oscillator and $R = 8.25\,\mathrm{k\Omega}$ for the NIC-type oscillator (see chapter 5. on the design of the experimental circuits). The same values were chosen for the determination of the phase response curve, because then, the dynamics can be written as

$$\dot{U} = \dot{U}_{\mathrm{uncoupled}} + \frac{U_{\mathrm{ext}}}{CR},$$

which is a perturbation of the dynamics of an uncoupled oscillator. When a sharp external voltage pulse is applied (ideally a delta peak), the voltage of the oscillator makes an instant jump by a defined value ΔU, regardless of the phase of the oscillator. To obtain the phase response curves in Fig. 31, a pulse with a Gaussian profile was applied to the oscillator 100 times at random points in time. For the TD-type oscillator (with $C = 100\,\mathrm{nF}$), the pulse had a FWHM of $0.3\,\mu$s and an amplitude of $3000\,\mathrm{mV}$, while for the NIC-type oscillator (with $C = 1\,\mathrm{nF}$), the pulse had a FWHM of $0.1\,\mu$s and an amplitude of $1500\,\mathrm{mV}$. In both cases, the expected voltage jump is

$$\Delta U = \int \frac{U_{\mathrm{ext}}(t)}{RC}\,\mathrm{d}t \approx 0.02\,\mathrm{V}.$$

This jump in the voltage signal is accompanied by a phase jump $\Delta\vartheta$ which does depend on the phase of the oscillator at the time when the external pulse was applied. This jump in phase can be calculated by comparing the time T' between zero crossings of the voltage signal $U(t)$ (recorded at a rate of 50 MS/s with a Spectrum M2i.4032 A/D converter PCI card) with positive slope directly before and directly after the pulse with the original period T:

$$\Delta\vartheta = 2\pi\left(1 - \frac{T'}{T}\right)$$

Now, the phase response curve can be obtained as $Z(\vartheta) = \Delta\vartheta/\Delta U$, where ϑ is the phase of the signal when the external pulse was applied to the oscillator. The procedure described above was adapted from a procedure used by Kiss, et al. (2005). Note that for the TD-type
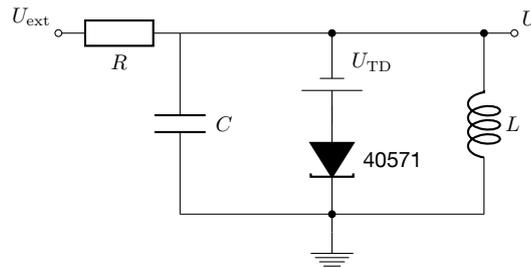
Fig. 16: Schematic of the extended circuit of a TD-type oscillator used to determine the phase response curve of the oscillator.

oscillators, which tend to exhibit parasitic oscillations of very high frequency, a running average of the sampled voltage signal with a window of 21 samples was used instead of the raw signal to suppress spurious zero crossings. Still, spurious zero crossings due to noise happened for both types of oscillators. If the phase at the time of the pulse was close to zero with a spurious crossing occuring immediately afterwards, this resulted in a return time T' close to zero instead of T. These values were ignored in the preparation of Fig. 31b and Fig. 31d.

3.3.2 Pattern recognition experiments

For the pattern recognition experiments, the components comprised in the setup were

- a Dell Precision T7400 workstation (Intel Xeon dual core processor, 2x2Ghz, 4GB RAM) running 64 bit Windows Vista,

- a Spectrum M2i.4032 analog/digital converter PCI card, installed in the workstation,

- a Spectrum M2i.6021 digital/analog converter PCI card, also installed in the workstation,

- a LeCroy WaveRunner 44MXi oscilloscope,

- the circuit board with the oscillatory network, as described in the design chapter (5.3),

- a power supply for the circuit board,

- a metal shield, consisting of a bottom plate and a lid, shielding the circuit board from above, below and two sides

- and a cardboard box, containing holes for the BNC connections of the circuit board, in which the circuit board along with the metal shielding was placed. The shielding improves the performance of the network.

Fig. 17: Schematic setup of the pattern recognition experiments. White on black numbers show the order of the experimental steps described in the text.

The schematic setup is depicted in Fig. 17. The whole setup was controlled by a program written in C++, using the Qt framework (http://qt-project.org) as well as the C++ APIs provided by Spectrum and LeCroy. The communication between the oscilloscope and the workstation was done via a LAN connection. All other external connections between devices were realized with BNC cables.

For a pattern recognition experiment, the following steps were performed in the order: 1, 2a through 7a, 2b through 7b, 8 (compare Fig. 17):

1. The defective pattern and the three recognizable patterns were passed to the controlling software, along with the desired duration T_{record} of the recorded time series and the desired parameters $U_{\text{coup}}^{\text{amp}}$ and $U_{\text{coup}}^{\text{off}}$ of the voltage signal governing the coupling (see also step 2b).

2. (a) The coupling voltage (i.e the designated output of the D/A converter) was set to $U_{\text{DA}}(t) = 0$ and written to a memory buffer.

   (b) Using the frequencies determined in step 7a, the coupling voltage was set to $U_{\text{DA}}(t) = U_{\text{coup}}^{\text{off}} + U_{\text{coup}}^{\text{amp}}/N \cdot a'(t)$. The coupling function $a'(t)$ (slightly varied from (2.12)) is given by

   $$a'(t) = \sum_{p=1}^{N} \sum_{q \neq p} w_{pq} \cos((\Omega_p - \Omega_q)t).$$

   This coupling function contains the same frequency components as $a(t)$ without the homogeneous contribution, which is instead governed by the value of $U_{\text{coup}}^{\text{off}}$ in the coupling signal. $U_{\text{coup}}^{\text{amp}}/N$ is the amplitude of a single sinusoidal component in the coupling signal. There was a low cutoff voltage ($U_{\text{DA}}^{\text{min}}$) as well as a high cutoff

voltage ($U_{\mathrm{DA}}^{\max}$). These values differed between measurements. They were due to the limited range of the M2i.6021 DA converter at a given voltage resolution, the maximum voltage range being $[-3\,\mathrm{V}, 3\,\mathrm{V}]$.

In both cases, the size of the buffer depended on the duration of the measurement: Since the sample rate was fixed at 1 MS/s, the buffer had $T_{\mathrm{record}} \cdot 1\,\mathrm{MHz}$ samples.

3. The D/A output buffer was transferred to the D/A converter.

4. The trigger signal was given for the A/D and D/A converter cards, which was passed on to the oscilloscope with a delay of a few microseconds. This delay was due to some undocumented behavior of the Spectrum cards and therefore not entirely predictable.

5. Simultaneously (i.e. within the small time difference between the oscilloscope and the A/D converter mentioned above), the coupling voltage was applied to the oscillatory network and the voltage signal $U_i(t)$ of all 8 oscillators was collected into recording buffers on the oscilloscope and the A/D converter, respectively, during a time interval of length T_{record}. The sample rate of data acquisition was 1 MS/s.

6. The oscilloscope and A/D buffers were transferred to the memory of the workstation computer.

7. (a) The second half of the buffer data was taken to determine the frequencies of the oscillators as described in 3.1.3. In order to avoid transients from switching on the system or from any previous measurements, the first half was not used.

   (b) The second quarter of the buffer data was taken to determine the frequencies of the oscillators after initialization. This is a time interval in which the oscillators are expected to have settled for their initial configuration, but the coupling according to the Hebbian rule has not yet started. Also, the phase shifts were computed for the whole time interval $[0, T_{\mathrm{record}}]$ using these frequencies as described in 3.1.3. Through this procedure, common drifts in the phase shift of all oscillators during initialization, like they are present in the simulations shown in Fig. 25 or Fig. 29 in the next chapter, do not appear in visualizations of experimental data. This is a more convenient way to display the data and does not affect the assessment of pattern recognition quality.

   Due to small delays in the triggering process, the point of reference (namely $t = 0$) was not the same for all oscillators, but differed by a few microseconds. To eliminate the varying offsets in the phase shift curves caused by this discrepancy, the offsets of all phase shift curves were chosen artificially such that after the time interval $T_{\mathrm{record}}/2$, the pattern represented by the phase shifts was exactly

the desired initial pattern. This means, it was assumed that initialization always worked perfectly. As a consequence, in visualizations of experimental phase shift data, the branches of $\varphi = 0$ and $\varphi = \pi$ are never spread out as for example in Fig. 25 during initialization. These plots therefore cannot be taken to judge the quality of the initialization process, because they are artificially perfect in that respect. However, they *can* be used to judge the quality of the recognition process by monitoring the change in the phase shift between initialization and recognition for each oscillator. After all, these changes do not depend on constant individual offsets of the phase shift curves.

8. The phase shift and frequency data (and optionally also the voltage time series) was stored on the hard drive.

Steps 2 to 7 were repeated, because first the frequencies of the uncoupled oscillators had to be measured exactly (in the sequence 2a, 3, 4, 5, 6, 7a) before the recognition could be performed (in the sequence 2b, 3, 4, 5, 6, 7b). However, once the frequencies are known, the system can be used for a series of pattern recognition measurements (i.e. going through steps 1, 2b, 3, 4, 5, 6, 7b, 8) repeatedly without going through the steps 2a, 3, 4, 5, 6, 7a in between). After a rather short time (usually around half a minute) however, the frequencies must be determined anew, because they have drifted too far from their original values for a successful pattern recognition. The frequencies were determined immediately before the pattern recognition experiment for all experiments presented in chapter 6.

4. Theoretical analysis

This chapter shines light on the restrictions that both the ideal and the realistic network have with respect to the robustness of pattern recognition and scalability of the network.

First, both numerical and analytical evidence is presented suggesting that any stationary state the ideal network settles for in recognition mode is marginally attracting at best (even when the zero eigenvalue of the Jacobian due to the global invariance with respect to common rotations of the phase shifts is not considered), other than the steady state in initialization mode (which is, apart from the aforementioned eigenvalue, an asymptotically stable fixed point, even in the presence of distributed oscillator frequencies (van Hemmen & Wreszinski 1993)). This fact has important consequences for the robustness the pattern recognition process.

Second, the effects of different deviations from ideal behavior of the realistic network are investigated systematically. This includes noise, distortions of the oscillation wave form, phase response curve and coupling function as well as a small separation of time scales. The main result is that noise (especially frequency inaccuracies due to noise) is qualitatively different from other types of non-ideal behavior: its effect worsens with a larger number of oscillators, thus putting a limit on network size.

Third, an estimate is given for the maximum size N of the network given the available frequency accuracy $\Delta\Omega/\Omega$.

4.1   Properties of the pattern states in recognition mode

It has already been discussed that the network described by (2.4) does not deliver recognized patterns perfectly, because the perfect representations of the memorized patterns are unstable steady states of the system (Aonishi 1998). While Aonishi established that up to a loading rate of $\alpha \approx 0.042$, a network that is initialized close enough to a memorized pattern will retrieve the memorized pattern, little is known about the exact location or about the eigenvalue spectra of the steady states corresponding to those retrieval states. Especially the eigenvalues are interesting from an experimentalist's point of view, because even a theoretically stable solution will not be realized in the long term if the level of noise or other deviations from ideal behavior of the network are large enough to drive the system away from it.

The first part of this section shows that for the special case of mutually orthogonal memorized patterns ξ^k, each point on the straight lines connecting any two of the states representing the patterns ξ^k in phase space is a stationary state. This family of degenerate states is either an attracting limit set, or it is part of an even larger set of degenerate states. Whether the larger set is an attractor itself remains an open question. However, it seems

plausible that it is the only attractor of the system, as will be discussed below.

Mutually orthogonal memorized patterns are a highly symmetrical special case. In the second part of this section, an argument is given why any meaningful attractor (i.e. close to a memorized pattern) existing in the network during the recognition phase should have M leading eigenvalues close to zero, where M is the number of memorized patterns, which means that states corresponding to recognized patterns are considerably less robust than the single global attractor in the case of initialization. Also, some numerical evidence is presented suggesting that, even in the general case, there is a family of degenerate steady states interconnecting all steady states corresponding to recognized patterns in phase space, making these states indifferent equilibria, same as in the case of mutually orthogonal patterns.

4.1.1   Orthogonal memorized patterns

General case

Consider again the dynamical equations for the recognition step in the ideal network:

$$\dot{\varphi}_i(\varphi_1, \ldots, \varphi_N) = \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{M} \xi_i^k \xi_j^k \sin(\varphi_j - \varphi_i) \tag{4.1}$$

To analyze this equation it is helpful to abstract from the global phase shift symmetry. Therefore, from here on, a fixed point is referred to as attractive, if all but one eigenvalues of the Jacobian are smaller than zero (the last one being always equal to zero). Also, a stationary state will be called non-isolated or degenerate only if the fixed point in question is not just trivially non-isolated (i.e. some neighboring stationary states must not lie along the $(1, 1, \ldots, 1)^{\mathrm{T}}$-direction for a non-isolated fixed point).

For each pattern ξ^k, there is a corresponding 1-dimensional manifold of fixed points of (4.1). Let φ^{*k} be any of those fixed points. In the case of just one memorized pattern (i.e. for an initialization to that pattern), φ^{*k} is a global attractor. For more than one memorized pattern, the φ^{*k} are generally unstable. However, in the special case of mutually orthogonal patterns (i.e. $\xi^l \cdot \xi^m = N\delta_{lm}$), the φ^{*k} are part of a set of non-isolated fixed points (once again, apart from the fact that they are trivially non-isolated because of the global phase shift invariance). Minimally, this set contains all straight lines connecting two such states in phase space:

$$\dot{\varphi}(\varphi^{*l} + u(\varphi^{*m} - \varphi^{*l})) = 0 \tag{4.2}$$

with the parameter $u \in \mathbb{R}$ and $l, m \in (1, \ldots, M)$.

Equation (4.2) can be verified by explicitly computing the time derivative. Entering the parametrization of the phase shifts along the connection of φ^{*l} and φ^{*m} into a single term

on the right hand side of (4.1) yields

$$\xi_i^k \xi_j^k \sin(\varphi_j(u) - \varphi_i(u)) = \xi_i^k \xi_j^k \sin(\varphi_j^{*l} - \varphi_i^{*l} + u(\varphi_j^{*m} - \varphi_j^{*l}) - u(\varphi_i^{*m} - \varphi_i^{*l})).$$

Since the invariance of a state in phase space does not depend on the coordinates, a new set of coordinates φ' can be chosen such that $\xi_i^{l'} = 1$, $i = 1, \ldots, N$, without loss of generality. In these new coordinates $\varphi_j^{*l'} = \varphi_i^{*l'}$ holds for all i and j, leading to:

$$\begin{aligned} \xi_i^{k'} \xi_j^{k'} \sin(\varphi_j{'}(u) - \varphi_i{'}(u)) &= \xi_i^{k'} \xi_j^{k'} \sin(u(\varphi_j^{*m'} - \varphi_i^{*m'})) \\ &= \xi_i^{k'} \xi_j^{k'} (\xi_j^{m'} - \xi_i^{m'})\beta. \end{aligned}$$

where $\beta = \sin(u\pi)/2$. Therefore, the time derivative of any state on the line defined by variation of the parameter u evaluates to

$$\begin{aligned} \dot\varphi_i{'} &= \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{M} \xi_i^{k'} \xi_j^{k'} (\xi_j^{m'} - \xi_i^{m'})\beta \\ &= \sum_{k=1}^{M} \xi_i^{k'} (\delta_{km} - \delta_{lk}\xi_i^{m'})\beta \\ &= (\xi_i^{m'} - \xi_i^{m'})\beta \\ &= 0. \end{aligned}$$

This result proves that the patterns ξ^k correspond to non-isolated fixed points of (4.1) that are mutually connected by straight lines of non-isolated fixed points in phase space. The potential along these lines is equal to $-N/2$:

$$E = -\frac{1}{2N} \sum_{i,j,k} \xi_i^k \xi_j^k \xi_i^l \xi_j^l = -\frac{1}{2} \sum_{j,k} \xi_j^k \xi_j^l \delta_{lk} = -\frac{N}{2} \sum_{k} \delta_{lk}^2 = -\frac{N}{2}.$$

In the next step, it is proven that $\lambda \leq 0$ for all eigenvalues λ of the Jacobian of any of these states. For the stationary states φ^{*k}, this was already done in section 5.1 of (Hölzel 2007). The following proof is an extension of the proof presented there.

The Jacobian $\boldsymbol{J}$ of (4.1) is given by

$$J_{ij} = \frac{\partial \dot\varphi_i}{\partial \varphi_j} = \frac{1}{N} \left(\sum_{k=1}^{M} \xi_i^k \xi_j^k \cos(\varphi_j - \varphi_i) - \delta_{ij} \sum_{p=1}^{N} \sum_{k=1}^{M} \xi_i^k \xi_p^k \cos(\varphi_p - \varphi_i) \right) \tag{4.3}$$

Along the vector joining φ^{*l} and φ^{*m} in phase space, given by $\varphi(u) = u(\varphi^{*m} - \varphi^{*l})$, the individual summands constituting the entries of the Jacobian evaluate to

$$
\begin{aligned}
\xi_i^k \xi_j^k \cos(\varphi_j(u) - \varphi_i(u)) &= \xi_i^k \xi_j^k \cos(\varphi_j^{*l} - \varphi_i^{*l} + u(\varphi_j^{*m} - \varphi_j^{*l}) - u(\varphi_i^{*m} - \varphi_i^{*l})) \\
&= \xi_i^k \xi_j^k \cos\left((\xi_i^l \xi_j^l - 1)\frac{\pi}{2} + u(\xi_i^m \xi_j^m - \xi_i^l \xi_j^l)\frac{\pi}{2} \right) \\
&= \xi_i^k \xi_j^k \xi_i^l \xi_j^l \cos\left(u(\xi_i^m \xi_j^m - \xi_i^l \xi_j^l)\frac{\pi}{2} \right) \\
&= \xi_i^k \xi_j^k \left(\xi_i^l \xi_j^l (1 - \gamma) + \xi_i^m \xi_j^m \gamma \right)
\end{aligned}
$$

where $\gamma = (1 - \cos(u\pi))/2$. $\boldsymbol{J}$ can now be written as a sum of matrices whose eigenvectors can be computed more easily than those of $\boldsymbol{J}$ itself:

$$
\boldsymbol{J} = (1 - \gamma) \sum_{k=1}^{M} \boldsymbol{A}^{kl} + \gamma \sum_{k=1}^{M} \boldsymbol{A}^{km} - \mathbb{1}. \tag{4.4}
$$

Here, $A_{ij}^{kx} = \xi_i^k \xi_j^k \xi_i^x \xi_j^x / N$.

$\boldsymbol{A}^{kx}$ has M mutually orthogonal eigenvectors given by

$$
\boldsymbol{v}^{kx,q} = \boldsymbol{\xi}^x \circ \boldsymbol{\xi}^q, \quad q = 1 \ldots M,
$$

where ' $\circ$ ' denotes component-wise multiplication. The corresponding eigenvalues are given by

$$
\lambda^{kx,q} = \delta_{kq}, \quad q = 1 \ldots M,
$$

as can be seen by calculating $\boldsymbol{A}^{kx} \cdot \boldsymbol{w}$ with an arbitrary vector $\boldsymbol{w}$:

$$
\left(\boldsymbol{A}^{kx} \cdot \boldsymbol{w} \right)_i = \frac{1}{N} \sum_{j=1}^{N} \xi_i^k \xi_j^k \xi_i^x \xi_j^x w_j = \frac{1}{N} \xi_i^k \xi_i^x \sum_{j=1}^{N} \xi_j^k \xi_j^x w_j = \frac{1}{N} v_i^{kx,k} \boldsymbol{v}^{kx,k} \cdot \boldsymbol{w}
$$

This result proves that all vectors orthogonal to $\boldsymbol{v}^{kx,k}$ are eigenvectors of $\boldsymbol{A}^{kx}$ with eigenvalue 0. Therefore the eigenvalue spectra of the three terms on the right hand side of (4.4) are

- $M \times (1 - \gamma)$ and $(N - M) \times 0$ for the first term,

- $M \times \gamma$ and $(N - M) \times 0$ for the second term and

- $N \times -1$ for the third term.

This means that $\lambda \leq 0$ for all eigenvalues λ of $\boldsymbol{J}$, since the largest eigenvalue of a sum of matrices cannot surpass the sum of the largest eigenvalues of the summands (see for example Knutson & Tao (2001)). $\boldsymbol{J}$ has at least two eigenvalues equal to 0, one belonging to the eigenvector $(1, 1, \ldots, 1)$ because of the global phase shift invariance, the other belonging to the eigenvector $\boldsymbol{\xi}^l \circ \boldsymbol{\xi}^m$, which is the vector pointing from φ^{*l} to φ^{*m}.

If the rest of the eigenvalues is negative, the degenerate set of states as a whole is attracting. This is the case in the example presented below in this section and also in the experiments presented in section 6.1.

If there are additional zero eigenvalues, the line of degenerate fixed points is embedded in a stationary manifold of higher dimension. The reasoning for this is as follows:

Movement along an eigenvector $\chi = (\chi_1, \chi_2, \ldots, \chi_N)$ starting at any steady state φ^* in the phase space of (4.1) can be written as

$$\dot{\varphi}_i(u) = \frac{1}{N} \sum_{j=1}^{N} w_{ij} \sin(\varphi_j^* - \varphi_i^* + (\chi_j - \chi_i)u),$$

where $\dot{\varphi}_i(0) = 0$. If the respective eigenvalue is zero, also $\partial_u \dot{\varphi}_i(0) = 0$ holds. The higher order derivatives with respect to u are given by

$$\partial_u^{(2n)} \dot{\varphi}_i(0) = \frac{1}{N} \sum_{j=1}^{N} w_{ij} (-1)^n (\chi_j - \chi_i)^{2n} \sin(\varphi_j^* - \varphi_i^*) = 0$$

and

$$\partial_u^{(2n+1)} \dot{\varphi}_i(0) = \frac{1}{N} \sum_{j=1}^{N} w_{ij} (-1)^n (\chi_j - \chi_i)^{2n+1} \cos(\varphi_j^* - \varphi_i^*).$$

where $n \in \mathbb{N}$. If all entries of χ have the same absolute value χ (which is the case for the degenerate state considered here, namely $\chi_i = \pm 1$), the last equation reduces to

$$\partial_u^{(2n+1)} \dot{\varphi}_i(0) = (-1)^n (2\chi)^{2n} \partial_u \dot{\varphi}_i(0) = 0.$$

This means, all derivatives with respect to u vanish. Since $\dot{\varphi}_i(u)$ is analytical, it must be identical to zero for all u. Therefore, all points along the eigenvector are stationary states as well; the line of non-isolated fixed points and the eigenvector span a whole plane of fixed points.

Unfortunately, the attempt to find the general structure and stability of the degenerate state for arbitrary sets of mutually orthogonal memorized patterns was not successful and further investigations are necessary. Also, it remains unclear, whether other attractors exist far from the memorized patterns even though they were never observed in the experiment or in simulations.

Example: 8 oscillators, 3 memorized orthogonal patterns

For a better impression of the location of the degenerate state in phase space, consider an example with 8 oscillators and 3 memorized patterns $\xi^1 = (1,1,1,1,1,1,1,1)^\mathrm{T}$, $\xi^2 = (1,1,1,1,-1,-1,-1,-1)^\mathrm{T}$ and $\xi^3 = (1,1,-1,-1,1,1,-1,-1)^\mathrm{T}$, which are mutually orthogonal.

a)

b)
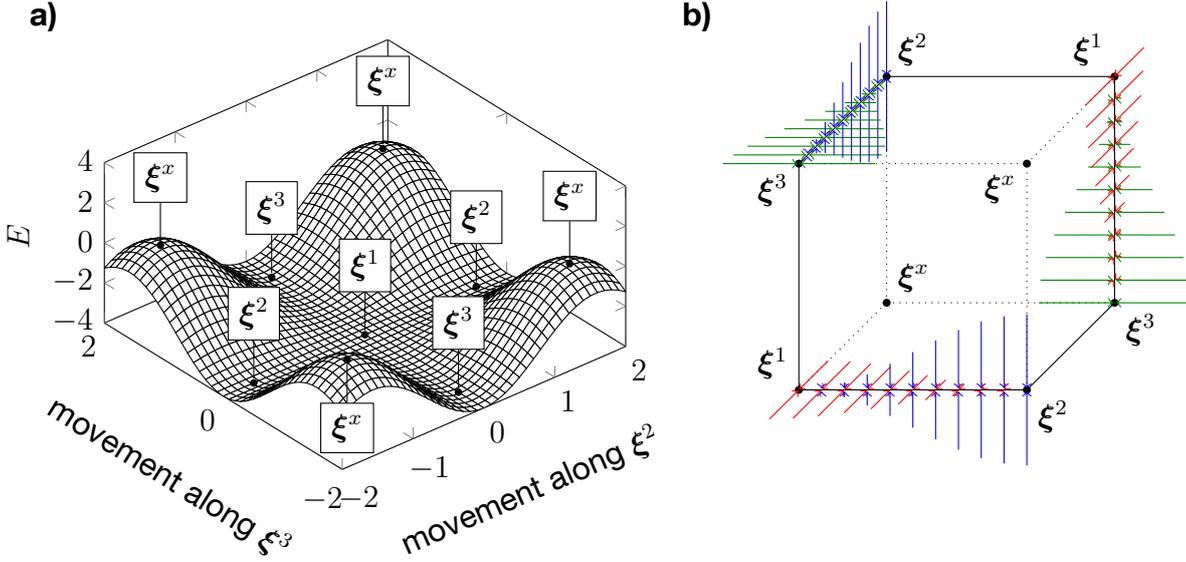
Fig. 18: a) Surface plot of the potential E of the network dynamics in recognition mode (given by (4.5)) on a two-dimensional plane in φ-space spanned by $\boldsymbol{\xi}^2 = \boldsymbol{\xi}^2 \circ \boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^3 = \boldsymbol{\xi}^3 \circ \boldsymbol{\xi}^1$. $\boldsymbol{\xi}^1$ is the origin. The location of special patterns in the plane is marked by labels. **b)** Three dimensional section of φ-space spanned by $\boldsymbol{\xi}^2$, $\boldsymbol{\xi}^3$ and $\boldsymbol{\xi}^x$. $\boldsymbol{\xi}^1$ is the origin. Note that the global invariant direction is orthogonal to the section. Solid lines mark stationary states. The direction of the eigenvectors with negative eigenvalues for these states is indicated by the colored arrows (red: $\boldsymbol{\xi}^x$, green: $\boldsymbol{\xi}^2$, blue: $\boldsymbol{\xi}^3$). The absolute value of the eigenvalue is indicated by the length of the line, ranging from 0 to 1. Note that all eigenvalues along directions other than $\boldsymbol{\xi}^1$, $\boldsymbol{\xi}^2$, $\boldsymbol{\xi}^3$ and $\boldsymbol{\xi}^x$ are equal to -1 for these states.

Fig. 18a visualizes the potential function

$$E = \sum_{k=1}^{M} E_{\boldsymbol{\xi}^k} = -\frac{1}{2N} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{M} \xi_i^k \xi_j^k \cos(\varphi_j - \varphi_i) \qquad (4.5)$$

in the vicinity of the origin on a two-dimensional cross section of phase space. Incidentally, because of the global phase shift invariance of (4.1), the origin $\varphi^1 = 0$ represents pattern $\boldsymbol{\xi}^1$. The cross section is the plane spanned by the vectors $\boldsymbol{\xi}^2$ and $\boldsymbol{\xi}^3$, which are, due to the choice of coordinates, also eigenvectors with zero eigenvalue of the steady state representing $\boldsymbol{\xi}^1$. By moving from the origin a distance of $\pi/2$ along either of the two axes, one arrives at patterns $\boldsymbol{\xi}^2$ and $\boldsymbol{\xi}^3$, respectively: $\varphi^2 = (\pi/2, \pi/2, \pi/2, \pi/2, -\pi/2, -\pi/2, -\pi/2, -\pi/2)^{\mathrm{T}}$ is a representation of $\boldsymbol{\xi}^2$ and the same goes for $\varphi^3 = (\pi/2, \pi/2, -\pi/2, -\pi/2, \pi/2, \pi/2, -\pi/2, -\pi/2)^{\mathrm{T}}$ and $\boldsymbol{\xi}^3$. By moving a distance of $\pi/2$ along both $\boldsymbol{\xi}^2$ and $\boldsymbol{\xi}^3$ from the origin, one arrives at another pattern orthogonal to the other three: $\boldsymbol{\xi}^x = (1, 1, -1, -1, -1, -1, 1, 1)^{\mathrm{T}}$. Since this pattern does not enter the coupling function, the potential function behaves differently here, showing a maximum.

If the dynamics are confined to the $\boldsymbol{\xi}^2$-$\boldsymbol{\xi}^3$-plane, the system can be expected to settle for any state in the potential valleys between $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ or $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^3$, depending on the initial conditions (if there are no perturbations). For pattern recognition, the final state will be closer to one of the patterns, because the initial pattern is closer to this pattern as well.

However, this is not the whole picture. While it is known that at ξ^1 all other directions in phase space are strongly attracting (except for the global invariance of the $(1,1,1,1,1,1,1,1)^{\mathrm{T}}$ direction), this is not true for ξ^2 and ξ^3. For these patterns, the potential landscape looks analog to Fig. 18a in the ξ^2-ξ^x-plane and in the ξ^3-ξ^x-plane, respectively.

Therefore, the subspace containing the possible final states is spanned by all three pattern vectors (once again neglecting the global phase shift invariance). Fig. 18b shows the degenerate steady state in this subspace (where the origin still represents ξ^1). The four patterns lie on the corners of a cube in phase space. Points on the edges that are not adjacent to ξ^x are fixed points with eigenvalues $\lambda \leq 0$. The eigenvectors belonging to negative eigenvalues are embedded in the faces of the cube (the absolute value of λ is visualized by the colored arrows along the edges). As a consequence, the degenerate state as a whole is a stable attractor.

Since each face of the cube is a symmetry plane of the flow, the attractor extends infinitely over phase space, which allows for permanently rotating solutions in the presence of arbitrarily small perturbations. An example for this kind of behavior is presented in Fig. 19, which shows the result of a simulated pattern recognition in the system (4.1) with randomly detuned frequencies. The detuning introduces small additional constant terms in the equations for $\dot{\varphi}_i$:

$$\dot{\varphi}_i = \omega_i + \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{M} \xi_i^k \xi_j^k \sin(\varphi_j - \varphi_i) \tag{4.6}$$

This equation and its dynamics are discussed in detail in section 4.2.1. For the simulation shown in Fig. 19, N values ω_i' were chosen randomly from the interval $[0\,\mathrm{Hz}, \Delta\Omega]$ with a uniform probability distribution. The width of the distribution was $\Delta\Omega = 0.1\,\mathrm{Hz}$. Then, their average value was subtracted from each ω_i' to obtain ω_i. The latter was done to avoid a global change in frequency leading to rotating solutions.

Adding the small terms ω_i to the dynamics is equivalent to introducing a small random tilt in the potential landscape. As a consequence, in the example at hand, after the initial fast recognition (i.e. ξ^1), the system begins to "roll" towards ξ^2. At this crossroad, the system moves towards ξ^3 which appears to be the favored direction, since from then on, the system keeps alternating between ξ^2 and ξ^3. Another way to imagine the motion in phase space is to look at the cube in Fig. 18b. The system starts out close to the ξ^1-corner and gets attracted to the ξ^1-ξ^2-edge, moving slowly towards ξ^2 due to the effect of the perturbation terms. Arriving at ξ^2, the system continues along the ξ^2-ξ^3-edge and periodically cycles through both patterns. Depending on the distribution of the ω_i, different movement patterns occur.

Since perturbations of this kind arise in any experimental system, no truly stationary pattern recognition can be achieved for mutually orthogonal memorized patterns.
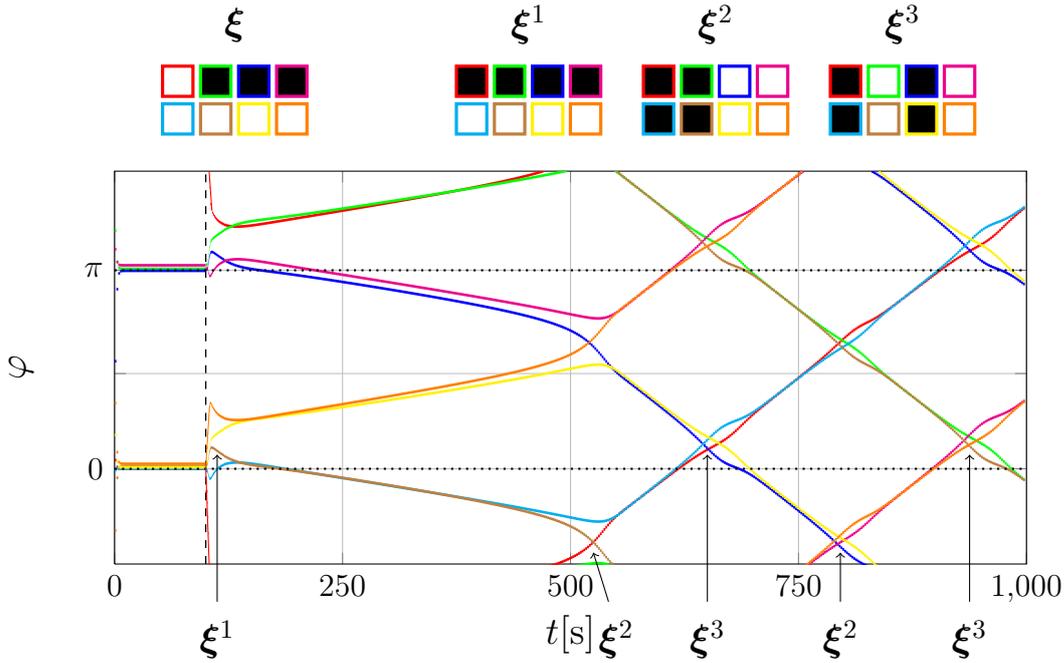
Fig. 19: Numerical time integration of the phase shifts in (4.6) with random initial conditions, $N = 8$ oscillators, $M = 3$ memorized patterns and one erroneous bit in the initial pattern. The initial pattern $\boldsymbol{\xi}$ and the memorized patterns $\boldsymbol{\xi}^k$ are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot. The coupling was switched to recognition mode at $t = 100$ s. The choice of the frequency deviations ω_i in (4.6) is described in the text.

4.1.2 Non-orthogonal memorized patterns

General case

In the following, an argument is made that for any set of randomly selected of memorized pattern vectors $\boldsymbol{\xi}^k$, $k = 1 \ldots M$, the M leading eigenvalues of the Jacobian for an attractor $\boldsymbol{\varphi}^*$ of (4.1) should be

$$\lambda_k \approx 0, \ k = 1, \ldots, M, \tag{4.7}$$

if the attractor is close to a memorized pattern, much like it is the case in for a set of orthogonal memorized patterns. Since the states in φ-space that form the attractors of 4.1 do not in general correspond to pattern vectors $\boldsymbol{\xi}$ with entries 1 or -1, it is useful to define the complex pattern vectors $|\xi\rangle (\boldsymbol{\varphi})$, that are defined for arbitrary states in φ-space:

$$|\xi\rangle_i (\boldsymbol{\varphi}) = e^{i\varphi_i}.$$

With this definition, each vector of phase shifts $\boldsymbol{\varphi}$ is linked to a pattern state $|\xi\rangle$ in an N-dimensional complex Hilbert space. For the proof of (4.7), consider the operator $\mathcal{J}$:

$$\mathcal{J}_{ij} = \frac{1}{N} \left(\sum_{k=1}^{M} \xi_i^k \xi_j^k e^{i(\varphi_j - \varphi_i)} - \delta_{ij} \sum_{p=1}^{N} \sum_{k=1}^{M} \xi_i^k \xi_p^k e^{i(\varphi_p - \varphi_i)} \right)$$

Note that J as defined in (4.3) forms the real part of $\mathcal{J}$. At any state $|\xi\rangle$ in phase space, $\mathcal{J}$ can be written as

$$\mathcal{J} = \frac{1}{N} \sum_{k=1}^{M} \left(\left| \xi^k \circ \xi \right\rangle \left\langle \xi^k \circ \xi \right| - \left\langle \xi^k \middle| \xi \right\rangle \sum_{i=1}^{N} \left\langle \xi^k \circ \xi \middle| i \right\rangle \left| i \right\rangle \left\langle i \right| \right),$$

where $|i\rangle$ is the unit vector along the i-th coordinate direction.

In the limit of large N, it can be shown that $\left\langle \xi^l \middle| \mathcal{J} \middle| \xi^m \right\rangle = 0$ for all combinations of l and m if the state represented by $|\xi\rangle$ corresponds exactly to a memorized pattern. Note that this implies that $\left\langle \xi^l \middle| J \middle| \xi^m \right\rangle = 0$, because $|\xi^l\rangle$ and $|\xi^m\rangle$ are real-valued vectors.

The value of $\left\langle \xi^l \middle| \mathcal{J} \middle| \xi^m \right\rangle$ is determined as follows:

$$
\begin{aligned}
\left\langle \xi^l \middle| \mathcal{J} \middle| \xi^m \right\rangle &= \frac{1}{N} \sum_{k=1}^{M} \left(\left\langle \xi^l \middle| \xi^k \circ \xi \right\rangle \left\langle \xi^k \circ \xi \middle| \xi^m \right\rangle - \left\langle \xi^k \middle| \xi \right\rangle \sum_{i=1}^{N} \left\langle \xi^k \circ \xi \middle| i \right\rangle \left\langle \xi^l \middle| i \right\rangle \left\langle i \middle| \xi^m \right\rangle \right) \\
&= \frac{1}{N} \sum_{k=1}^{M} \left(\left\langle \xi^l \middle| \xi^k \circ \xi \right\rangle \left\langle \xi^k \circ \xi \middle| \xi^m \right\rangle - \left\langle \xi^k \middle| \xi \right\rangle \sum_{i=1}^{N} \left\langle \xi^k \circ \xi \circ \xi^l \circ \xi^m \middle| i \right\rangle \right) \\
&= \frac{1}{N} \sum_{k=1}^{M} \left(\left\langle \xi^l \middle| \xi^k \circ \xi \right\rangle \left\langle \xi^k \circ \xi \middle| \xi^m \right\rangle - \left\langle \xi^k \middle| \xi \right\rangle \left\langle \xi^k \circ \xi \circ \xi^l \circ \xi^m \middle| 1 \dots 1 \right\rangle \right),
\end{aligned}
$$

where $|1 \dots 1\rangle$ is the pattern vector defined by $\xi_i = 1$, $i = 1 \dots N$. The latter expression can be rewritten as

$$\left\langle \xi^l \middle| \mathcal{J} \middle| \xi^m \right\rangle = \frac{1}{N} \sum_{k=1}^{M} \left(\left\langle \xi^k \circ \xi^l \middle| \xi \right\rangle \left\langle \xi \middle| \xi^k \circ \xi^m \right\rangle - \left\langle \xi^k \middle| \xi \right\rangle \left\langle \xi \middle| \xi^k \circ \xi^l \circ \xi^m \right\rangle \right),$$

$|\xi\rangle$ is now assumed to correspond to a memorized pattern. Without loss of generality, this memorized pattern is chosen as $|1 \dots 1\rangle$. Therefore,

$$\left\langle \xi^l \middle| \mathcal{J} \middle| \xi^m \right\rangle = \frac{1}{N} \sum_{k=1}^{M} \left(\left\langle \xi^k \middle| \xi^l \right\rangle \left\langle \xi^k \middle| \xi^m \right\rangle - \left\langle \xi^k \middle| 1 \dots 1 \right\rangle \left\langle \xi^k \middle| \xi^l \circ \xi^m \right\rangle \right).$$

The largest contribution in the left hand term occurs at $k = l$, while the largest contribution in the right hand term occurs for $k = 1$. Both terms cancel each other out. All other individual contributions have the same expected value of $\mathcal{O}(\sqrt{N})$, which means that they should cancel out as well if N is large enough, and therefore $\left\langle \xi^l \middle| \mathcal{J} \middle| \xi^m \right\rangle = 0$. For the rest of the proof, only the symmetric, real-valued operator J and real-valued vectors are considered. Let $|\psi\rangle$ be any such vector in the M-dimensional span of the $|\xi^k\rangle$, given by

$$|\psi\rangle = \sum_{k=1}^{M} a_k \left| \xi^k \right\rangle,$$

with arbitrary coefficients $a_k \in \mathbb{R}$. Then, also

$$\langle \psi \,|\, J \,|\, \psi \rangle = \sum_{k=1}^{M} \sum_{l=1}^{M} a_k a_l \langle \xi^k \,|\, J \,|\, \xi^l \rangle = 0. \tag{4.8}$$

This result allows to prove that J has at least M zero eigenvalues. The proof is done by contradiction:

Suppose that J had only $x < M$ eigenvalues $\lambda_i = 0, \; i = 1 \ldots x$ and $N - x$ eigenvalues $\lambda_i < 0, \; i = x + 1 \ldots N$. As J is a symmetric matrix, all eigenvalues are real. Also, there are no positive eigenvalues, because the state under consideration is assumed to be attractive. Let $|\chi^i\rangle$ denote the corresponding normalized eigenvectors, which form an orthonormal basis of the N-dimensional vector space. Now, choose a vector $|\psi\rangle$ of finite length in the M-dimensional subspace spanned by the $|\xi^k\rangle$, for which $\langle \psi \,|\, \chi^i \rangle = 0$ for $i \leq x$. Such a vector must exist because the dimension of the subspace spanned by the first x eigenvectors is strictly smaller than M, according to the premise. For this particular vector, there is at least one $i > x$ with $\langle \psi \,|\, \chi^i \rangle \neq 0$, and therefore

$$\langle \psi \,|\, J \,|\, \psi \rangle = \sum_{i=1}^{N} \lambda_i \left| \langle \psi \,|\, \chi^i \rangle \right|^2 < 0$$

This is a contradiction to (4.8) and therefore, if an attractive steady state is exactly equal to a memorized pattern, its Jacobian must have M eigenvalues equal to zero. If the steady state is sufficiently close to a memorized pattern, this is still approximately true, as stated in (4.7). The question remains, however, what "sufficiently close" means in this respect.

Example: 100 oscillators, 3 memorized random patterns

To see the effect of the M small eigenvalues discussed above, a series of simulations was run in a system with $N = 100$ oscillators and $M = 3$ patterns. The reason for the huge increase in oscillator number compared to the example for orthogonal patterns is that for random patterns, a substantially bigger number than eight oscillators are needed in order to accommodate three memorized patterns. At $N = 100$ and $M = 3$, the load rate of the network is $\alpha = 0.03$, which is sufficiently below Aonishi's threshold that small size effects can be ruled out as the reason for any type of special behavior.

Fig. 20 depicts three simulations of slightly perturbed dynamics

$$\dot{\varphi}_i = \omega_i + \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{M} \xi_i^k \xi_j^k \sin(\varphi_j - \varphi_i) \tag{4.9}$$

where, again, N values ω_i' were chosen randomly from the interval $[0\,\text{Hz}, \Delta\Omega = 0.1\,\text{Hz}]$ with a uniform probability distribution and their average value was subtracted from each ω_i' to

obtain ω_i.

The only difference between the simulations shown in Fig. 20a, Fig. 20b and Fig. 20c were the different random patterns and the different random frequency deviations ω_i that were used. The images show three qualitatively different types of behavior:

In Fig. 20a, the network evolves quickly (faster than the time resolution of the figure) to the correct memorized pattern and stays there for all times. Pattern recognition works in the long time limit.

In Fig. 20b, the network evolves quickly to the correct memorized pattern, but after a transient settles for another, incorrect memorized pattern and stays there for all times. Pattern recognition works on a short time scale only.

In Fig. 20c, the network evolves quickly to the correct memorized pattern, but then drifts cyclically between all memorized patterns, similar to Fig. 19. Pattern recognition works on a short time scale only.

Fig. 20c suggests that the potential landscape looks similar to Fig. 18a even for nonorthogonal patterns. However, in contrast to the orthogonal case, there are truly stable states close to the patterns, to which the system settles in Fig. 20a and Fig. 20b. This begs the question, why the state corresponding to the correct memorized pattern is not always realized for long times, although it is apparently attractive and close to the initial pattern. A possible explanation is that, because some of the eigenvalues are close to zero, $\Delta\Omega = 0.1$ Hz is too large a perturbation, randomly destabilizing the weakly stable attractors.

To determine the effect of the size of the deviations on the behavior, a series of numerical simulations with varying $\Delta\Omega$ was conducted. The results are shown in Fig. 21. Each of the bar stacks represents a series of 100 numerical integrations of (4.9). The size of the perturbations was different for each series, ranging from $\Delta\Omega = 1 \cdot 10^{-5}$ Hz to $\Delta\Omega = 0.1$ Hz.

The most remarkable observation is that the time scale of the error terms ω_i does not play a role at all in the long time limit of the dynamics of 4.9. The relative frequency of the different possible outcomes remains the same for the whole investigated range of $\Delta\Omega$. The ratio of correctly recognized patterns to incorrectly recognized patterns is roughly $1 : 2$, suggesting that the correct pattern is not preferred in any way as long term limit of the dynamics, *regardless* of the size of the perturbation terms. Another interesting result is that for all $\Delta\Omega$, drifting solutions comparable to the one in Fig. 20c exist. Finally, apart from the three different outcomes depicted in Fig. 20, it also happened roughly 20% of the time that a stationary solution that did not correspond to any of the memorized patterns was realized. A possible explanation of this behavior is as follows:

In the case of nonorthogonal memorized patterns, there exists a set of (up to symmetries) M special solutions φ^{k*} that are close to the memorized patterns and have M zero eigenvalues. These solutions are connected by a network of degenerate steady state solutions much like the perfect patterns in the orthogonal case. However, the degenerate state is not
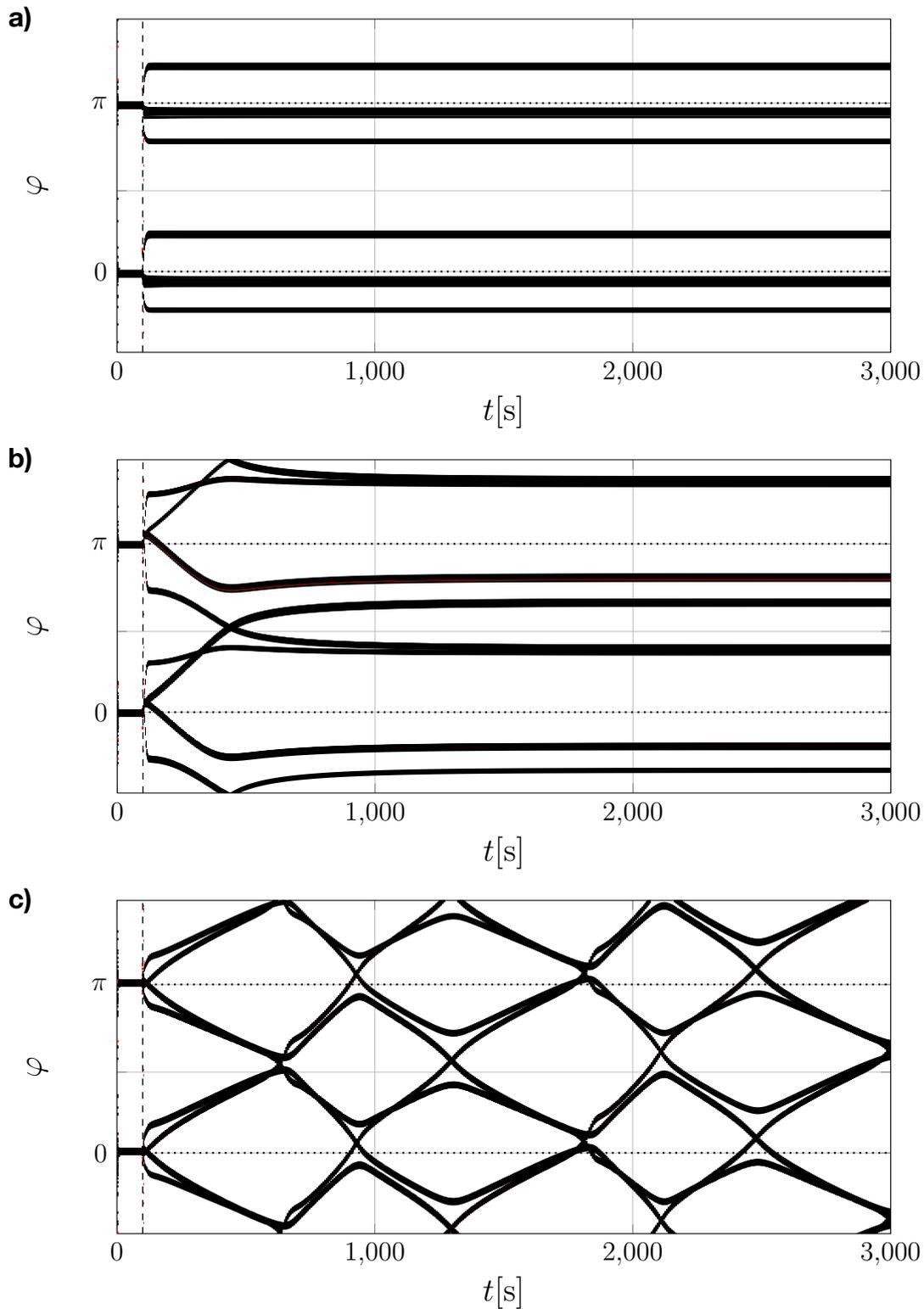
Fig. 20: a) Numerical time integration of the phase shifts in (4.9) with random initial conditions, $N = 100$ oscillators, $M = 3$ randomly selected memorized patterns (with equal probability for both states) and 10 erroneous bits in the initial pattern. The choice of the frequency deviations ω_i in (4.9) is described in the text. b) Same as a), with a different set of random patterns and frequency deviations. c) Same as a), with yet another different set of random patterns and frequency deviations. Note that in all three simulations, the short term pattern recognition was successful in a matter of seconds. Due to the low sampling rate of the data on that time scale, this is unfortunately not recognizable in the images.
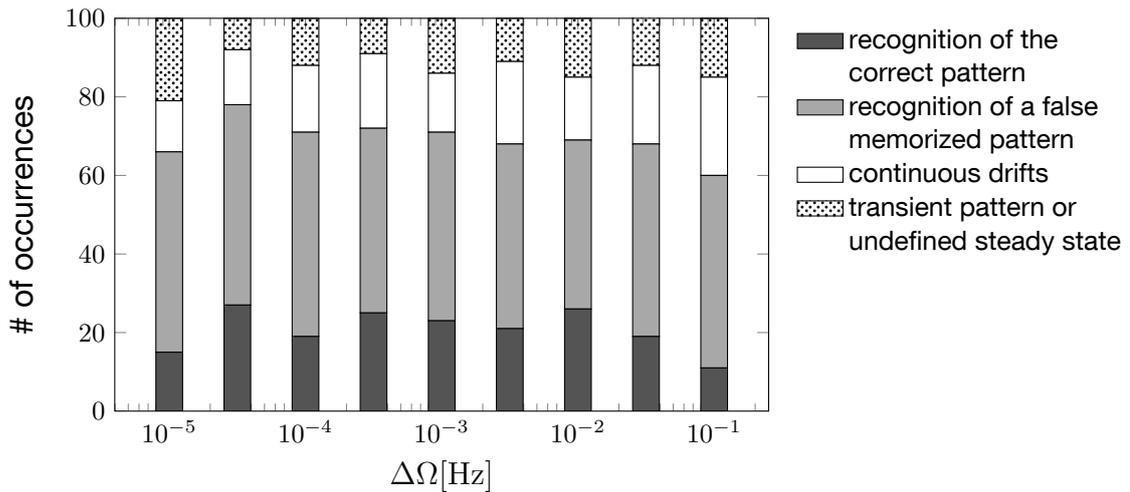
Fig. 21: The plot shows, for different values of the frequency detuning $\Delta\Omega$, how often each possible type of long term behavior occurred in a series of 100 numerical simulations of (4.9) with $N = 100$ oscillators, $M = 3$ randomly selected memorized patterns (with equal probability for both states) and 10 erroneous bits in the initial pattern. The relation between $\Delta\Omega$ and the frequency deviations ω_i in (4.9) is explained in the text. The long term behavior of the system was evaluated at $t = 1000/\Delta\Omega$, the coupling was switched to recognition mode at $t = 100$ s. Note that the short term pattern recognition was always succesful in these simulations.

a network of straight lines now, due to the lesser symmetry of the system. Fig. 22 shows a stylized potential landscape with the curved invariant manifold. While in the ideal system no point in phase space is attractive, an arbitrarily small perturbation can stabilize points close to the invariant manifold, if the effect "tilts" the potential landscape in the right way. This also explains why occasionally a pattern that does not represent one of the memorized patterns is realized in the long term. Note that this stabilizing mechanism does not work in the case of orthogonal patterns, where the invariant curves connecting the φ^k in phase space are straight lines. Please also note that these thoughts as well as the potential landscape in Fig. 22 are quite speculative. However, they do explain the numerical results well.
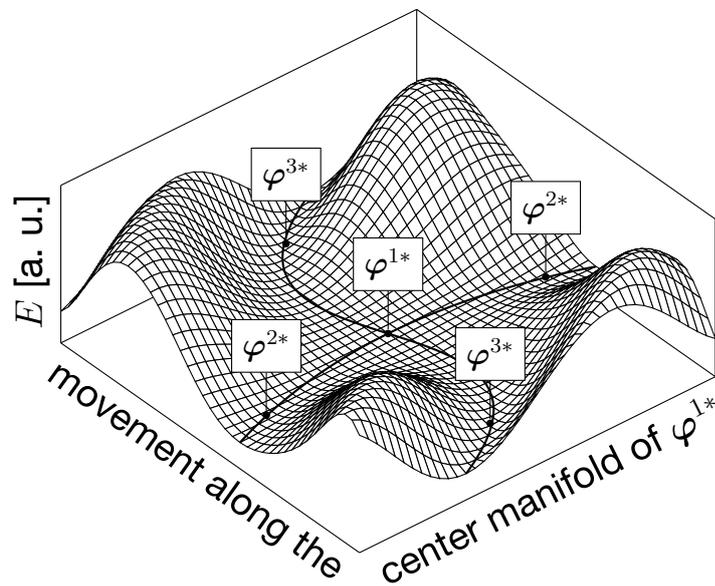
Fig. 22: Schematic plot of the potential E of the network dynamics in recognition mode on a two-dimensional subsection through the three-dimensional center manifold of φ^{1*} in φ-space for a system with $M = 3$ nonorthogonal memorized patterns. The subsection is orthogonal to the $(1, \ldots, 1)$-direction. Fixed points with M zero eigenvalues that represent imperfect memorized patterns are marked with labels. The thick black lines are equipotential curves with minimal potential that connect φ^{1*} to the other memorized patterns.

4.2 Changes of the ideal dynamics under non-ideal conditions present in an experiment

Any experimental realization of a globally coupled network of oscillators will exhibit deviations from (2.25). In the last section, it was shown that reliable long term pattern recognition cannot be achieved in the presence of arbitrarily small perturbations even in the ideal model of the network. However, short term recognition appears to be more robust. Still, it is important to quantify the effect of a non-ideal setup on the quality of short term pattern recognition. This is somewhat difficult, because of the cumulative effect of errors produced by different sources. For example, an offset in oscillator frequencies causes a drift in the solution. A small separation of time scales causes an oscillation of the solution around the slowly drifting average. Both effects together will lead to a faster destruction of the recognized pattern (i.e φ_i leaving the appropriate value range) than the drift alone.

This section investigates different types of deviations from ideal behavior independently, with the caveat that the combined effects may be worse.

Note that the following analysis is different from the one in (Hölzel 2007), as it is given from an experimentalist's point of view, with a special focus on effects that pose severe limitations for a large number of oscillators. It also goes deeper than previous efforts regarding perturbations to the ideal dynamics (2.4) using mean-field models; this is discussed in detail in the concluding section (4.3) of this chapter.

The starting point of the analysis is the general equation for realistic oscillators with a weak global coupling (compare section 2.4.2),

$$
\begin{aligned}
\dot{\vartheta}_i^{real} &= \Omega_i^{real}(t) + \varepsilon a^{real}(t) \frac{1}{N} \sum_{j=1}^{N} Z_i^{real}(t) Y_j^{real}(t)), \\
\dot{\varphi}_i^{real} &= \dot{\vartheta}_i^{real} - \Omega_i.
\end{aligned}
\tag{4.10}
$$

With the definition of the "real" quantities given below, this general equation deals with the fact that each oscillator and the coupling function itself are subject to small deviations in phase and amplitude:

- $\vartheta^{real}(t) = \vartheta(t) + \eta_\vartheta(t)$, $\varphi^{real}(t) = \varphi(t) + \eta_\vartheta(t)$, $\Omega^{real}(t) = \Omega + \dot{\eta}_\vartheta(t)$

  $\eta_\vartheta \neq 0$ accounts for phase noise in the uncoupled oscillations. If $\dot{\eta}_\vartheta(t)$ is a slowly varying function, this error corresponds to a deviation from the original frequency.

- $Z^{real}(t) = \cos(\vartheta^{real}(t)) + \eta_Z(t)$ and $Y^{real}(t) = \sin(\vartheta^{real}(t)) + \eta_Y(t)$

  $\eta_Z \neq 0$ and $\eta_Y \neq 0$ account for a noisy amplitude of the waveforms of both oscillation and phase response as well as for systematic distortions from sinusoidal shape.

- $a^{real}(t) = \sum_{k=1}^{N} \sum_{l=1}^{N} w_{kl} \cos((\Omega_k - \Omega_l)t + \eta_{kl}^{\text{phase}}(t)) + \eta_{kl}^{\text{amp}}(t)$

Each component of the coupling function is treated as a signal with a phase error and an amplitude error, just like each oscillator:

- $\eta_{kl}^{\text{phase}} \neq 0$ accounts for phase noise in a coupling component. Similar to $\dot{\eta}_\vartheta(t)$, a slowly varying $\dot{\eta}_{kl}^{\text{phase}}(t)$ is equivalent to a detuned frequency of the coupling component.

- $\eta_{kl}^{\text{amp}} \neq 0$ accounts for amplitude noise, distortions and small delays in the coupling function.

There is another possible error source, which also must be covered, namely,

- if $\varepsilon/\delta \ll 1$ does not hold i.e if the coupling strength can no longer be considered small compared to the time scale of the slowest oscillations, even ϑ^{ideal} in the system of coupled harmonic oscillators might deviate substantially from ϑ in the system of phase coupled oscillators.

To get an impression of the effects of the different errors, each error type is investigated separately from the others. Since in most physical systems these errors occur simultaneously, the results must be taken with a grain of salt. In the very least, however, they provide a set of minimum requirements for experimental accuracy. Moreover, the analysis is helpful to identify the most important error types with respect to experimental optimization.

Please note that from here on, ϑ^{real}, φ^{real}, Z^{real} and Y^{real} will be denoted ϑ, φ, Z and Y for brevity. The respective variables of the unperturbed system are called ϑ^{ideal}, φ^{ideal}, Z^{ideal} and Y^{ideal}, respectively.

To quantify the deviations from ideal behavior, the quality of pattern recognition must be assessed in some way. The commonly used criterion is the overlap m^l of the final state φ_{final} of the system with the recognized pattern ξ^l, given by

$$m^l = \frac{1}{N} \left| \langle \xi_{\text{final}} | \xi^l \rangle \right| = \frac{1}{N} \sqrt{\left(\sum_{i=1}^{N} \xi_i^l \sin \varphi_{i,\text{final}} \right)^2 + \left(\sum_{i=1}^{N} \xi_i^l \cos \varphi_{i,\text{final}} \right)^2},$$

which is an order parameter arising from a mean-field treatment of the network (Aoyagi & Kitano 1997, Aoyagi & Kitano 1998, Aonishi 1998, Aonishi et al. 1999, Yamana et al. 1999). For the overlap, $0 \leq m^l \leq 1$ is always fulfilled, with $m^l = 1$ in case of a perfect recognition of ξ^l. Note that "final state" here denotes the state of the network at the point in time when the recognized pattern is read out, even if this is not a steady state. The parameter m^l maps all solutions that differ in a common offset in all φ_i to one value, which is convenient for dealing with the global symmetry of the network. If there is only one memorized pattern, the parameter m is equivalent to the order parameter r in the classic Kuramoto model. However, using m as a criterion for pattern recognition has also some disadvantages: Other

than for synchronized oscillators that are all in phase, the order parameter cannot be simply determined by reading out the amplitude of a global signal. Computing the overlap of the final state with all memorized patterns from the individual values φ_i is not feasible, because this could have been done with the original pattern $\boldsymbol{\xi}$ to obtain the pattern with the greatest overlap in the first place. Also, a value of e.g $m = 0.95$ is no guarantee that all oscillators have correct phase shift values. If the number of pixels in the pattern is very large, single defective bits will not alter the average significantly. Since the intended application of the network is the retrieval of an error-free pattern from an erroneous one, m is not necessarily suitable to judge the capabilities of the network.

An alternative measure, which is introduced here to alleviate the aforementioned shortcomings, is the proximity P^l, defined by

$$P^l = \min_i \xi_i \xi_i^l \cos(\varphi_{i,\text{initial}} - \varphi_{i,\text{final}}),$$

where $\varphi_{i,\text{initial}}$ is the initialized state of the network corresponding to $\boldsymbol{\xi}$. The proximity indicates how close the recognized pattern is to perfect recognition in the sense of the recognition condition (2.6). $-1 \leq P^l \leq 1$ always holds, with $P^l > 0$ if and only if all oscillators corresponding to correct bits in the initial pattern changed their phase shift by less than $\pi/2$ during the recognition phase and all oscillators corresponding to defective bits in the initial pattern changed their phase shift by more than $\pi/2$. Since the output pattern $\boldsymbol{\xi}_{\text{final}}$ of the network after recognition was always determined as

$$\xi_{i,\text{final}} = \text{sgn}\left(\xi_i \cos(\varphi_{i,\text{initial}} - \varphi_{i,\text{final}})\right),$$

(based on the assumption that (2.6) is fullfilled; see also the experimental chapter 6.) this means that $P^l > 0$ if and only if the pattern $\boldsymbol{\xi}^l$ was recognized correctly. Note that P, other than m, is not tolerant of common offsets in the phase shifts, which means that the two measures behave differently in the presence of such offsets. However, in the numerical simulations based on (4.10) presented in the rest of this chapter, a value of P clearly larger than zero was always correlated with a value of m close to one.

4.2.1 Change induced by phase noise

In this section, the effect of phase noise on the dynamics of the network is investigated. It turns out that only low frequency phase noise, which is equivalent to slow drifts in the oscillator frequencies, has an non-vanishing average effect on the ideal dynamics.

General case

Phase noise is taken into account in

$$\dot\vartheta_i = \Omega_i + \dot\eta_{\vartheta,i}(t) + \varepsilon a(t)\frac{1}{N}\sum_{j=1}^{N}\cos\vartheta_i\sin\vartheta_j$$

$$\dot\varphi_i = \dot\eta_{\vartheta,i}(t) + \varepsilon a(t)\frac{1}{N}\sum_{j=1}^{N}\cos\vartheta_i\sin\vartheta_j \tag{4.11}$$

where $\vartheta_i(t) = \vartheta_i^{\text{ideal}}(t) + \eta_{\vartheta,i}(t) = \Omega_i t + \varphi_i^{\text{ideal}}(t) + \eta_{\vartheta,i}(t) = \Omega_i t + \varphi_i(t)$ with the phase noise term $\eta_{\vartheta,i}(t)$ that fulfills $\dot\eta_{\vartheta,i}(t) \ll \Omega_i$. Splitting up the right hand side into the ideal dynamics and a perturbation part yields

$$\dot\vartheta_i = \dot\vartheta_{i,\text{ideal}} + \dot\eta_{\vartheta,i}(t),$$

$$\dot\varphi_i = \dot\varphi_{i,\text{ideal}} + \dot\eta_{\vartheta,i}(t).$$

Let $\dot\eta_{\vartheta,i}(t) = \mu h_i(t)$, where $h_i(t) \leq 1$ and μ describes the intensity of the phase noise. In order to preserve the ideal dynamics, all effects of the error term must occur on a slower time scale than the ideal dynamics. Therefore, if the noise term obeys $\mu \ll \varepsilon$, its effect can be neglected. However, depending on the time scale of the noise, this may be an unnecessarily strict requirement. Assume for example that the noise term is a fast oscillating term: $\dot\eta_{\vartheta,i}(t) = \mu\sin\Omega_\eta t$ with constant $\mu \ll \Omega_\eta$. Then, like for the ideal dynamics, the average effect of the error term can be estimated with the multiple time scale analysis described in section 2.3.1:

$$\dot\vartheta_i = \dot\vartheta_{i,\text{ideal}} + \mu\sin\Omega_\eta t = \dot\vartheta_{i,\text{ideal}} + \mathcal{O}\left(\frac{\mu}{\Omega_\eta}\right).$$

As a consequence, depending on the type of phase noise, there are two different conditions on the size of $\dot\eta_{\vartheta,i}(t)$:

For slow time scale noise, or constant offsets in frequency, the condition is given by
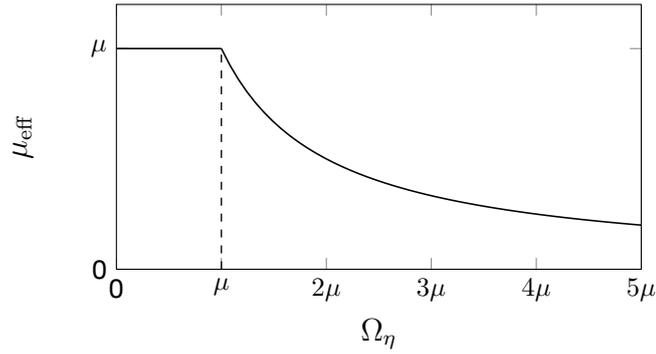
$$\mu \ll \varepsilon_{\text{eff}}.$$

For noise that contains only fast oscillating terms of frequency $\Omega_\eta \gg \mu$ or higher, the condition becomes

$$\frac{\mu}{\Omega_\eta} \ll \varepsilon_{\text{eff}}.$$

Those two conditions can be combined into

$$\mu_{\text{eff}}(\Omega_\eta) \ll \varepsilon_{\text{eff}},$$

Fig. 23: Effective size of the noise term in (4.11) with $\dot{\eta}_{\vartheta,i}(t) = \mu \sin \Omega_\eta t$ vs. the frequency of the noise term Ω_η. The noise term is negligible if $\mu_{\text{eff}} \ll \varepsilon$. Note that this curve is only meant to show the qualitative behavior and the order of magnitude of μ_{eff} in the limit of low and high Ω_η. In particular, the kink at $\Omega_\eta = \mu$ is just an artifact.

where $\mu_{\text{eff}}(\Omega_\eta)$ describes the effective size of the phase noise after averaging (see Fig. 23 for the qualitative behavior of $\mu_{\text{eff}}(\Omega_\eta)$). The largest effect occurs at constant and slowly varying phase noise terms. The special case of constant terms is treated below.

Effect of small frequency inaccuracies of the oscillations

In an experimental network, to provide the coupling function, oscillator frequencies must be measured at some point before the pattern initialization and pattern recognition steps are performed. This can only be done with limited accuracy. Also, frequency drifts over time arise due to changes in the environmental conditions (like temperature and air humidity). Therefore, deviations of the actual frequencies from the frequencies entering the coupling function are to be expected. In this case (4.11) takes the form

$$\dot{\varphi}_i = \Delta\Omega_i + \varepsilon a(t) \frac{1}{N} \sum_{j=1}^{N} \cos \vartheta_i \sin \vartheta_j, \tag{4.12}$$

where $\Delta\Omega_i$ are the frequency deviations. After averaging, these equations become

$$\dot{\varphi}_i = \Delta\Omega_i + \varepsilon_{\text{eff}} \left(\frac{1}{N} \sum_{j=1}^{N} w_{ij} \sin(\varphi_j - \varphi_i) + \mathcal{O}\left(\frac{\varepsilon_{\text{eff}}}{\delta}\right) \right), \tag{4.13}$$

with $\varepsilon_{\text{eff}} = \varepsilon/2$. The averaged equations can only have a steady state if $\overline{\Delta\Omega_i} = 0$ where $\overline{\Delta\Omega_i}$ is the average deviation, because $\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} \sin(\varphi_j - \varphi_i) = 0$. Therefore, it makes sense to analyze the system in a rotating frame of reference. The variable transform $\varphi_i \to \varphi_i' = \varphi_i - \overline{\Delta\Omega_i}t$ leads to

$$\dot{\varphi}_i' = \omega_i + \varepsilon_{\text{eff}} \frac{1}{N} \sum_{j=1}^{N} w_{ij} \sin(\varphi_j' - \varphi_i').$$

Here, $\omega_i = \Delta\Omega_i - \overline{\Delta\Omega_i}$ and therefore $\overline{\omega_i} = 0$. For the initialization case this is equivalent to the Kuramoto model of N coupled oscillators with a frequency distribution:

$$\dot{\varphi}'_i = \omega_i + \varepsilon_{\text{eff}}\frac{1}{N}\sum_{j=1}^{N}\sin(\varphi'_j - \varphi'_i) \tag{4.14}$$

Initialization works if the oscillators form a synchronized cluster for $t \to \infty$. Whether or not oscillators synchronize depends on the distribution of the ω_i and on the coupling strength ε_{eff}. In general, there will be a group of synchronized oscillators and a group of unsynchronized oscillators. Kuramoto introduced the complex order parameter

$$re^{\text{i}\psi} = \frac{1}{N}\sum_{j=1}^{N}e^{\text{i}\varphi'_j}$$

and showed that (4.14) can be rewritten as

$$\dot{\varphi}'_i = \omega_i + \varepsilon_{\text{eff}}r\sin(\psi - \varphi'_i)$$

Here $0 \le r \le 1$ describes the degree of synchronization in the system, where $r = 1$ means full synchronization and ψ is the average phase shift. It is known (Strogatz 2000) that all oscillators with $|\omega_i| \le \varepsilon_{\text{eff}}r$ for $N \to \infty$ form a synchronized group. For a finite number of oscillators, r is not constant. Still, since $r \le 1$ always, $|\omega_i| \le \varepsilon_{\text{eff}}$ is a necessary condition for a stationary value of φ'_i.

The condition on the actual measuring error of the frequencies then becomes

$$|\overline{\Delta\Omega_i} - \Delta\Omega_i| < \varepsilon_{\text{eff}}.$$

This means, if frequency deviations occur symmetrically around zero, no single deviation should be larger than ε_{eff}.

   Fig. 24 illustrates what happens when the system passes through the synchronization bifurcation in a system of two slightly detuned coupled oscillators. In this case, (4.12) may be simplified to

$$\dot{\varphi}_2 - \dot{\varphi}_1 = \Delta\Omega_2 - \Delta\Omega_1 + \varepsilon\cos((\Omega_2 - \Omega_1)t)\sin((\Omega_1 - \Omega_2)t + \varphi_1 - \varphi_2), \tag{4.15}$$

where only terms that are potentially resonant with the coupling function were left in the product on the right hand side. Apart from the drift in the phase shift difference, which is compensated by choosing a rotating frame of reference, there are two effects of the frequency detuning:

The first effect is a shift of the steady states of the phase shift differences. Far below the bifurcation, the stable states are at $\varphi'_2 - \varphi'_1 = 0$ for positive coupling and $\varphi'_2 - \varphi'_1 = \pi$ for
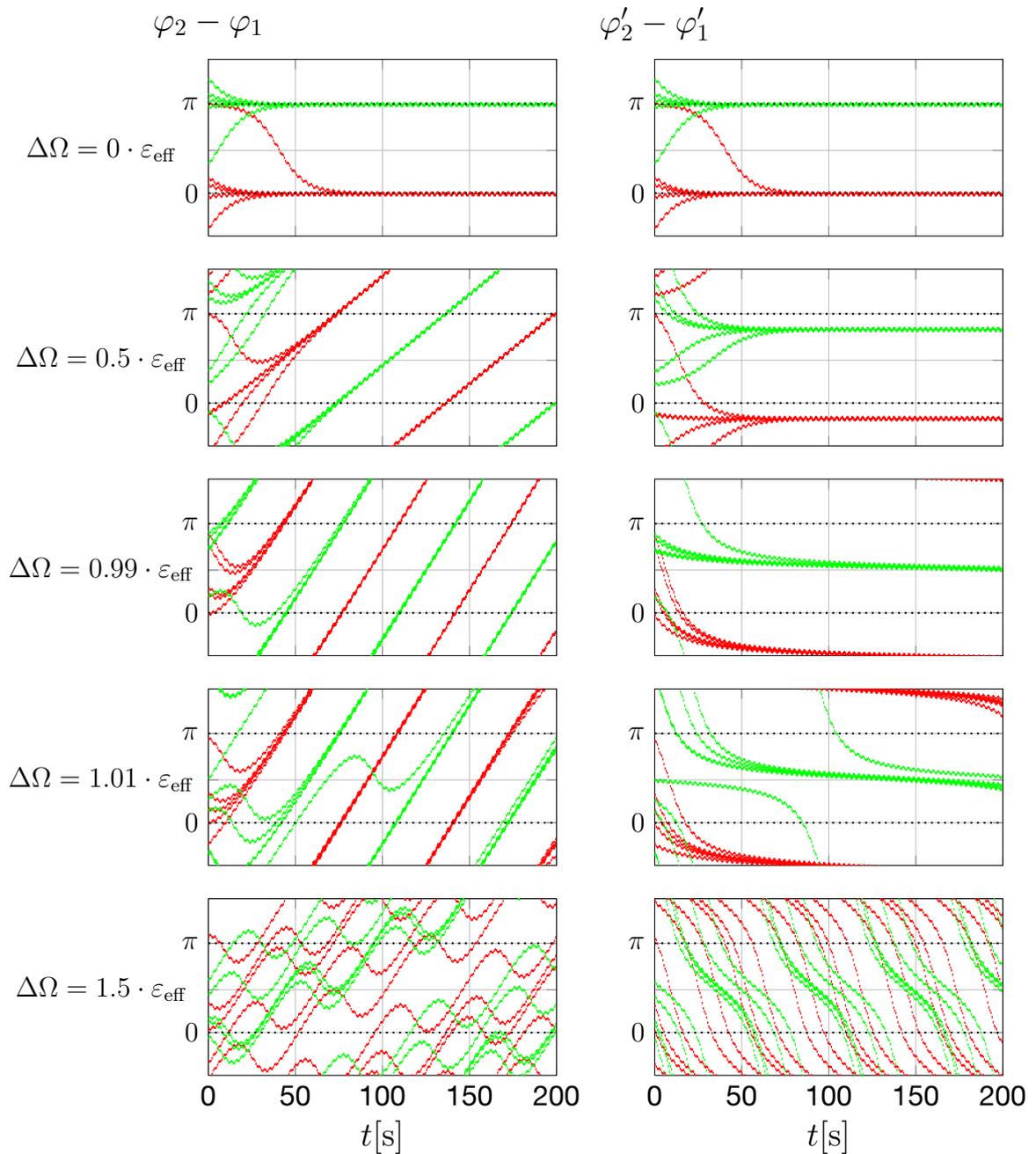
Fig. 24: Left column: Numerical time integration of the phase shift difference of two weakly coupled oscillators with a detuned coupling frequency (ODE given by (4.15)). For each value of the frequency detuning $\Delta\Omega = \Delta\Omega_2 - \Delta\Omega_1$, simulations were repeated with five randomly selected initial conditions for both positive ($w_{12} = 1$, red curves) and negative coupling ($w_{12} = -1$, green curves). The values of the parameters were $\epsilon = 0.2\,\mathrm{s}^{-1}$, $\Omega_1 = 1\,\mathrm{Hz}$, $\Omega_2 = 2\,\mathrm{Hz}$. Right column: Each plot shows the corresponding plot on the left side after a transformation to a rotating coordinate system ($\varphi_2' - \varphi_1' = \varphi_2 - \varphi_1 - \Delta\Omega t$).
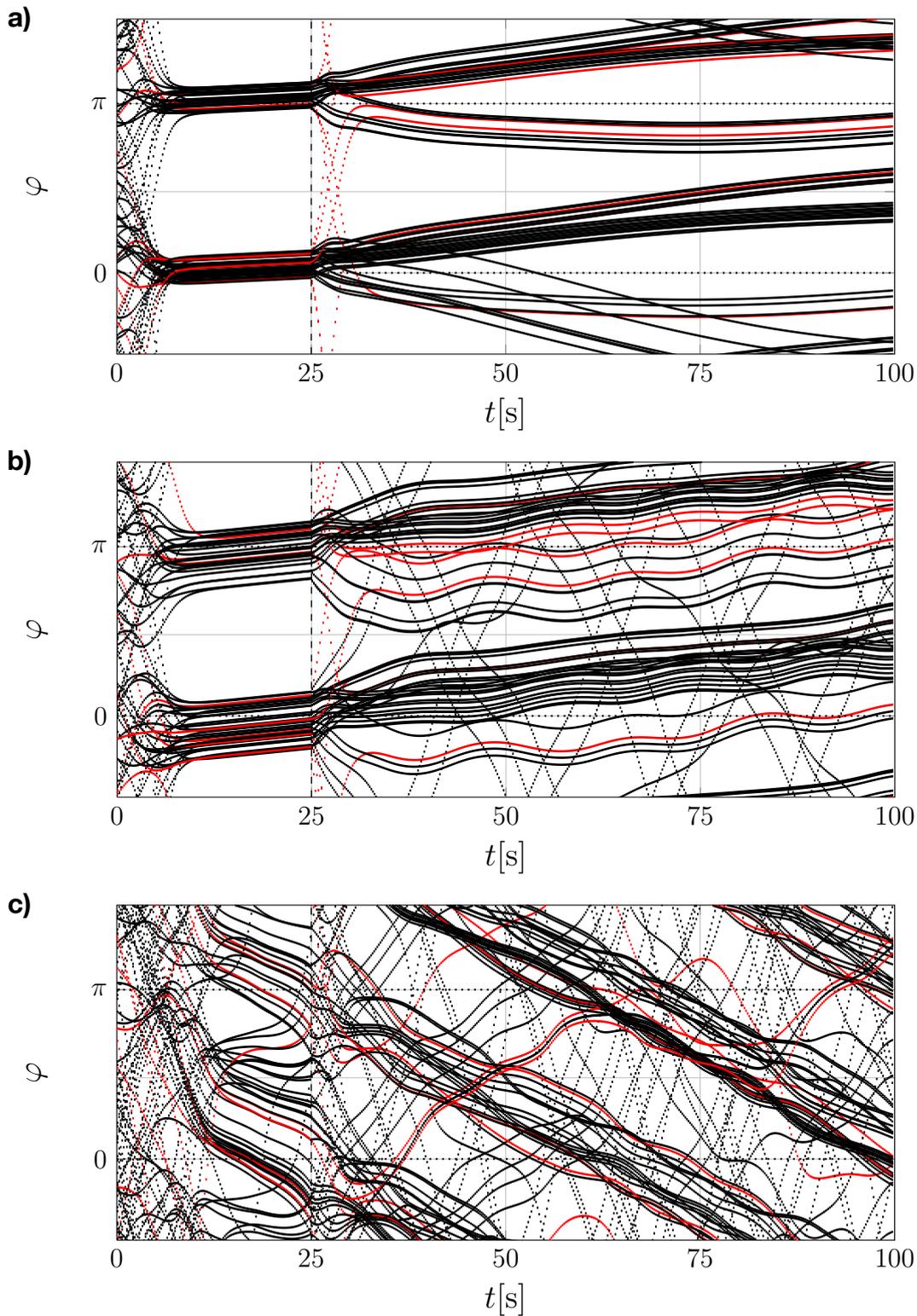
Fig. 25: **a)** Numerical time integration of the phase shifts in (4.13) with random initial conditions, $N = 60$ oscillators, $M = 3$ memorized patterns and 8 erroneous bits in the initial pattern. The initial and memorized patterns are the same as in Fig. 4. Red phase shift curves belong to phase shifts that correspond to erroneous bits in the initial pattern. The coupling strength was set to $\varepsilon_{\text{eff}} = 1$. The coupling was switched to recognition mode at $t = 25\,\text{s}$. The $\Delta\Omega_i$ were chosen randomly from the interval $[-\Delta\Omega/2, \Delta\Omega/2]$ with a uniform probability distribution and $\Delta\Omega = 0.5$. **b)** Same as a) with $\Delta\Omega = 1$. **c)** Same as a) with $\Delta\Omega = 1.5$.

negative coupling. When the detuning approaches the bifurcation (at $\Delta\Omega_2 - \Delta\Omega_1 = \varepsilon/2$), those states move to $\varphi_2' - \varphi_1' = -\pi/2$ and $\varphi_2' - \varphi_1' = \pi/2$ respectively. The second effect comes into play if the system is beyond the bifurcation. In this case, the oscillators still keep their phase shift differences almost constant for a while (this is the "ghost" of the steady state below the bifurcation), but then a jump of 2π occurs.

In a system with many oscillators, there is an average acceleration or deceleration due to frequency inaccuracies as well. The shift in the steady state of the phase shift differences causes a certain spread of the branches at $\varphi = 0$ and $\varphi = \pi$. Both effects can be observed in the numerical results shown Fig. 25, particularly well during the initialization. The plots show three simulations for a network of 60 oscillators governed by 4.13. The range of frequency deviations

$$\Delta\Omega = \max_{i,j} |\Delta\Omega_j - \Delta\Omega_i|$$

was increased from $\Delta\Omega = 0.5 \cdot \varepsilon_{\text{eff}}$ in Fig. 25a to $\Delta\Omega = 1 \cdot \varepsilon_{\text{eff}}$ in Fig. 25b to $\Delta\Omega = 1.5 \cdot \varepsilon_{\text{eff}}$ in Fig. 25c. As one would expect, the pattern recognition does not work any more for $\Delta\Omega = 1 \cdot \varepsilon_{\text{eff}}$, even though initialization still does. At $\Delta\Omega = 1.5 \cdot \varepsilon_{\text{eff}}$, initialization does not work any more as well. The relative robustness of the initialization follows from the robustness of the synchronized steady state in contrast to the steady states during recognition. Another, more intuitive way to explain this robustness goes as follows: During the initialization step, each oscillator is coupled to all other oscillators with the coupling strength $\varepsilon_{\text{eff}}/N$. In recognition mode, this coupling strength is modified by the coupling matrix w_{ij}, which now has entries of different sizes, potentially even including zeros. Therefore, some oscillators will be coupled more strongly than during initialization and others more weakly. It seems logical that a weaker coupling during recognition can drive the system above the bifurcation threshold, even if it is below for initialization.

In a more systematic approach to quantify the effect of the spread of frequency deviations $\Delta\Omega$ on short term pattern recognition, a series of simulations with $N = 100$ oscillators, $M = 3$ randomly selected memorized patterns and ten defective bits in the initial pattern was performed, where $\Delta\Omega/\varepsilon_{\text{eff}}$ was varied from 10^{-2} to 10. For each value of $\Delta\Omega/\varepsilon_{\text{eff}}$, the results of 100 numerical simulations were recorded (see Fig. 26). Fig. 26 plots two different measures of the short term pattern recognition quality over $\Delta\Omega/\varepsilon_{\text{eff}}$. The first is the overlap $m = 1/N |\langle \xi_{\text{final}} | \xi^1 \rangle|$ of the final pattern with the correctly memorized pattern $\boldsymbol{\xi}^1$, where $|\xi_{\text{final}}\rangle_i = e^{\mathrm{i}\varphi_{\text{final},i}}$. The final state was evaluated after a $15\,\mathrm{s}$ ($15/\varepsilon_{\text{eff}}$) recognition period. While m is a good measure for the overall closeness of two patterns, it does not provide a good criterion to decide whether the phase shift of any single oscillator is correct. Therefore, as discussed in the introduction to section 4.2, the proximity

$$P = \min_i \xi_i \xi_i^1 \cos(\varphi_{\text{final},i} - \varphi_{\text{initial},i})$$
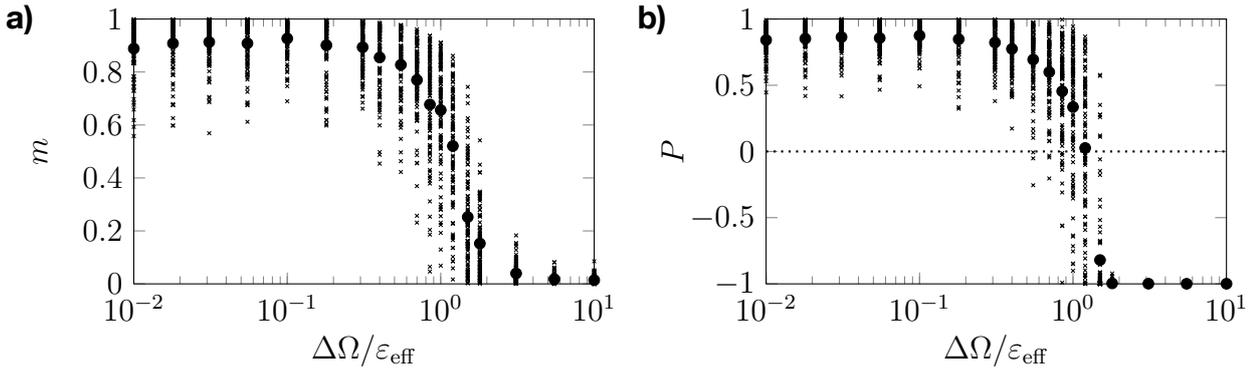
Fig. 26: a) Overlap of the final pattern with the correct memorized pattern vs. the frequency detuning for a series of simulations of (4.13) with $N = 100$ oscillators, $M = 3$ randomly selected memorized patterns (with equal probability for both states) and ten erroneous bits in the initial pattern. The coupling strength was set to $\varepsilon_{\text{eff}} = 1$. The coupling was switched to recognition mode at $t = 25\,$s, the final state was evaluated at $t = 40\,$s. The $\Delta\Omega_i$ were chosen randomly from the interval $[-\Delta\Omega/2, \Delta\Omega/2]$. For each value of $\Delta\Omega$, 100 simulations were run. Small crosses indicate results for individual simulations, the filled circles are the average overlap. **b)** Proximity (see text for definition) of the final pattern to the correct memorized pattern vs. the frequency detuning for the same simulations. Small crosses indicate results for individual simulations, the filled circles are the average proximity.

was introduced as a second criterion. A proximity value of $P = 1$ means that the correct pattern (i.e. ξ^1) has been perfectly recognized, while any value $P > 0$ shows that every oscillator is sufficiently close to its respective branch (i.e. $|\varphi_{\text{final}} - \varphi_{\text{perfect}}| < \pi/2$).

The largest detuning of frequencies for which all pattern recognitions were successful in these simulations was $\Delta\Omega/\varepsilon_{\text{eff}} = 0.4$, although, in at least one instance, recognition was very close to failing. For a larger detuning, recognition starts to fail in more and more cases - at $\Delta\Omega/\varepsilon_{\text{eff}} = 1$ the average value of P is equal to zero, indicating that the number of failures becomes comparable to the number of successes. All in all, the results suggest that

$$\Delta\Omega \leq 0.3 \cdot \varepsilon_{\text{eff}}$$

is a reasonable value for the upper limit of the frequency detuning.

Note that, since the maximal available coupling strength is inversely proportional to N^2 (compare (2.14)), this can be written as

$$\Delta\Omega \ll 0.3 \cdot \frac{\Omega_{\text{max}} - \Omega_{\text{min}}}{N^2}$$

Therefore, the frequency accuracy of the oscillators must increase quadratically with N, which severely limits the scalability of the system. Note that the meaning of "$\ll$" in the last equation is investigated in 4.2.5, when a small separation of timescales is discussed.

4.2.2   Change induced by distorted oscillations with amplitude noise

The phase shift dynamics of a network of weakly coupled near-harmonic oscillators with slightly distorted and/or noisy wave forms and phase response curves is given by

$$\dot{\vartheta}_i = \Omega_i + \varepsilon a(t) \frac{1}{N} \sum_{j=1}^{N} Z_i(t) Y_j(t),$$

where $Z_i(t) = \cos(\vartheta_i(t)) + \eta_{Z,i}(t)$, $Y_j(t) = \sin(\vartheta_j(t)) + \eta_{Y,j}(t)$ and $\eta_Z(t) \ll 1$ and $\eta_Y(t) \ll 1$ are arbitrary, small error terms. The dynamical equation for ϑ_i reads

$$\dot{\vartheta}_i = \Omega_i + \varepsilon a(t) \frac{1}{N} \sum_{j=1}^{N} (\cos \vartheta_i + \eta_{Z,i}(t))(\sin \vartheta_j + \eta_{Y,j}(t))$$

Splitting up the right hand side into the ideal dynamics and a perturbation part yields

$$\dot{\vartheta}_i = \dot{\vartheta}_{i,\text{ideal}} + \varepsilon a(t) \frac{1}{N} \sum_{j=1}^{N} (\eta_{Z,i}(t) \sin \vartheta_i + \eta_{Y,j}(t) \cos \vartheta_j) + \mathcal{O}(\eta_{Z,i}\eta_{Y,j})$$

Since $\eta_{Z,i}(t) \ll 1$ and $\eta_{Y,j}(t) \ll 1$, any deviation from the ideal dynamics occurs on a substantially longer timescale and can be neglected for pattern recognition, regardless of the number of oscillators. Therefore, using noisy oscillators which are not perfectly sinusoidal does not restrict the scalability of the network.

4.2.3   Change induced by phase noise in the coupling function

General Case

If each frequency component $\cos((\Omega_p - \Omega_q)t)$ entering the coupling function is subject to phase noise described by $\eta_{pq}^{\text{phase}}(t)$ with $\dot{\eta}_{pq}^{\text{phase}}(t) \ll \Omega_p - \Omega_q$, the network dynamics is given by

$$
\begin{aligned}
\dot{\vartheta}_i &= \Omega_i + \frac{\varepsilon}{N} \sum_{p=1}^{N} \sum_{q=1}^{N} w_{pq} \cos((\Omega_p - \Omega_q)t + \eta_{pq}^{\text{phase}}(t)) \sum_{j=1}^{N} \cos(\vartheta_i) \sin(\vartheta_j) \\
\dot{\varphi}_i &= \frac{\varepsilon}{N} \sum_{k=1}^{N} \sum_{l=1}^{N} w_{pq} \cos((\Omega_p - \Omega_q)t + \eta_{pq}^{\text{phase}}(t)) \sum_{j=1}^{N} \cos(\vartheta_i) \sin(\vartheta_j) \qquad (4.16)
\end{aligned}
$$

Let $\dot{\eta}_{pq}^{\text{phase}}(t) = \mu h_{pq}(t)$, where $h_{pq}(t) \leq 1$ and μ describes the intensity of the phase noise. Under the assumption, that $\mu \leq \varepsilon$, all noise terms $\eta_{pq}^{\text{phase}}(t)$ can be treated as slowly varying variables just like $\varphi_i(t)$. For now, only this case is considered, and 4.16 can be averaged to

$$\dot{\varphi}_i = \varepsilon_{\text{eff}} \frac{1}{N} \sum_{j=1}^{N} w_{ij} \sin(\varphi_j - \varphi_i - \eta_{ij}^{\text{phase}}(t))) + \mathcal{O}\left(\frac{\varepsilon_{\text{eff}}}{\delta}\right) + \mathcal{O}\left(\frac{\mu_{\text{eff}}}{\delta}\right),$$

where μ_{eff} depends on the frequencies present in the slowest noise term $\eta_{ij}^{\mathrm{phase}}(t)$, analog to Fig. 23. μ_{eff} is indeed the correct effective scale of the phase noise term, because single terms $\eta_{ij}^{\mathrm{phase}}(t)$ can be redeclared as phase noise of either φ_i or φ_j. Therefore, their effective contribution scales in the same way and constant noise terms have the biggest effect, which is investigated below.

Effect of small frequency inaccuracies in the coupling

The averaged equations of a network with frequency inaccuracies in the coupling function are given by

$$\dot{\varphi}_i = \varepsilon_{\mathrm{eff}} \frac{1}{N} \sum_{j=1}^{N} w_{ij} \sin(\varphi_j - \varphi_i - \Delta\Omega_{ij}). \tag{4.17}$$

Fig. 27 shows three simulations for a network of 60 oscillators governed by 4.17. The range of frequency deviations

$$\Delta\Omega = \max_{i,j,p,q} |\Delta\Omega_{ij} - \Delta\Omega_{pq}|$$

was increased from $\Delta\Omega = 0.1 \cdot \varepsilon_{\mathrm{eff}}$ in Fig. 27a to $\Delta\Omega = 0.2 \cdot \varepsilon_{\mathrm{eff}}$ in Fig. 27b to $\Delta\Omega = 0.5 \cdot \varepsilon_{\mathrm{eff}}$ in Fig. 27c.

There are two important differences to the dynamics of a network with detuned oscillator frequencies. First, even during initialization, no synchronized state emerges, but each oscillator drifts at its own pace, because there is no transformation to a rotating frame of reference that maps (4.17) to a system of coupled Kuramoto oscillators. Second, successful pattern recognition breaks down at a substantially lower value of $\Delta\Omega$.

This can be seen especially well in Fig. 28, which shows the quality of short term pattern recognition depending on $\Delta\Omega/\varepsilon_{\mathrm{eff}}$ in a series of numerical simulations of (4.17) with $N = 100$ oscillators, $M = 3$ randomly selected memorized patterns and ten defective bits in the initial pattern. For each value of $\Delta\Omega/\varepsilon_{\mathrm{eff}}$, the results of 100 numerical simulations the overlap and the proximity of the final pattern were recorded. The threshold for the detuning up to which all 100 runs resulted in a successful pattern recognition was at

$$\Delta\Omega \leq 0.1 \cdot \varepsilon_{\mathrm{eff}}.$$

Like the frequency inaccuracies of the oscillators, the frequency inaccuracies of the coupling function must decrease with $1/N^2$. On top of that, it appears that the frequencies of the coupling function must be more accurate than those of the oscillators by a factor of 3.
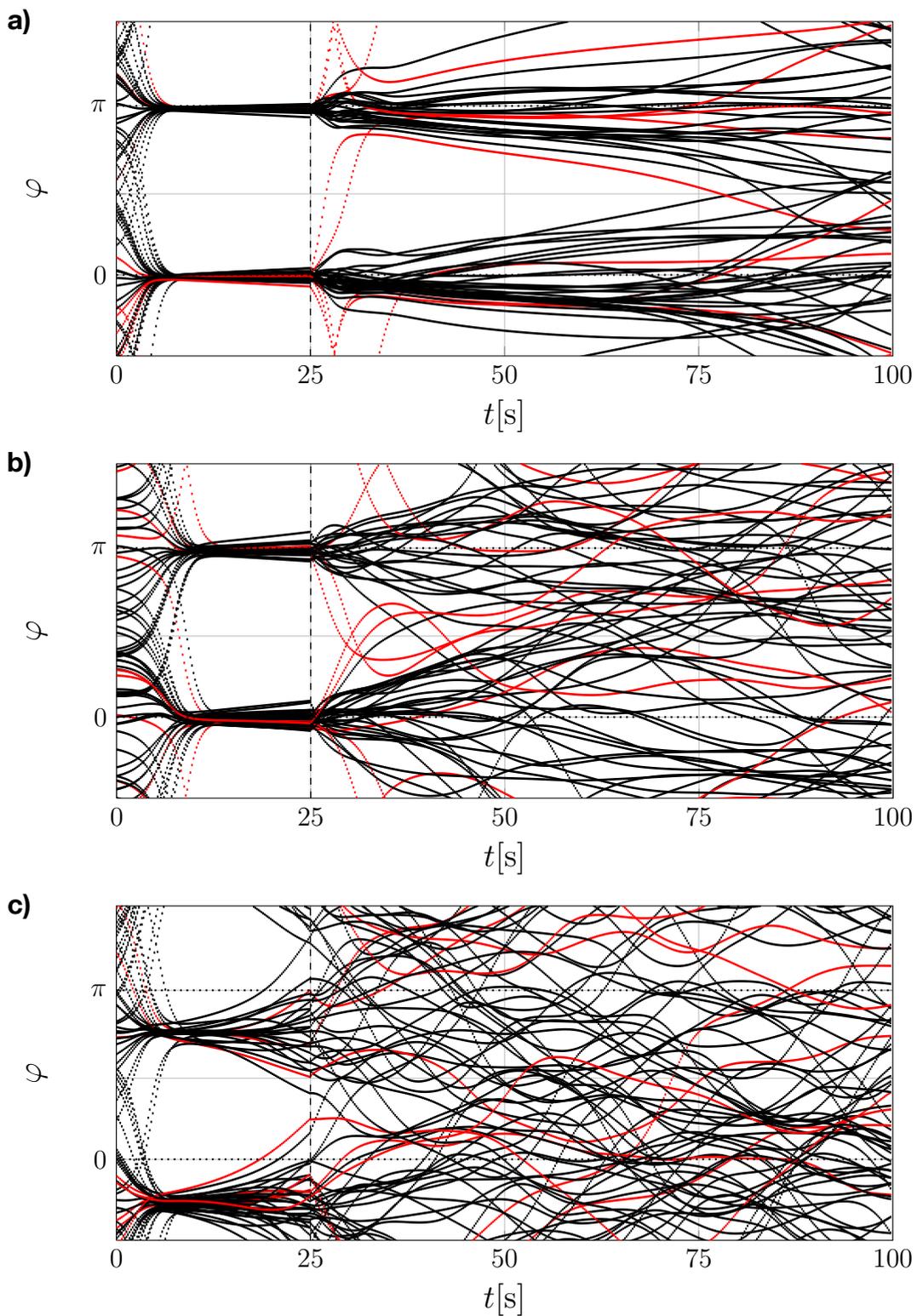
Fig. 27: a) Numerical time integration of the phase shifts in (4.17) with random initial conditions, $N = 60$ oscillators, $M = 3$ memorized patterns and 8 erroneous bits in the initial pattern. The initial and memorized patterns are the same as in Fig. 4. Red phase shift curves belong to phase shifts that correspond to erroneous bits in the initial pattern. The coupling strength was set to $\varepsilon_{\text{eff}} = 1$. The coupling was switched to recognition mode at $t = 25\,\text{s}$. The $\Delta\Omega_{ij}$ were chosen randomly from the interval $[-\Delta\Omega/2, \Delta\Omega/2]$ with a uniform probability distribution and $\Delta\Omega = 0.1$. **b)** Same as a) with $\Delta\Omega = 0.2$. **c)** Same as a) with $\Delta\Omega = 0.5$.

Fig. 28: a) Overlap of the final pattern with the correct memorized pattern vs. the frequency detuning for a series of simulations of (4.17) with $N = 100$ oscillators, $M = 3$ randomly selected memorized patterns (with equal probability for both states) and ten erroneous bits in the initial pattern. The coupling strength was set to $\varepsilon_{\mathrm{eff}} = 1$. The coupling was switched to recognition mode at $t = 25\,\mathrm{s}$, the final state was evaluated at $t = 40\,\mathrm{s}$. The $\Delta\Omega_i$ were chosen randomly from the interval $[-\Delta\Omega/2, \Delta\Omega/2]$. For each value of $\Delta\Omega$, 100 simulations were run. Small crosses indicate results for individual simulations, the filled circles are the average overlap. **b)** Proximity of the final pattern to the correct memorized pattern vs. the frequency detuning for the same simulations. Small crosses indicate results for individual simulations, the filled circles are the average proximity.

4.2.4   Change induced by a distorted coupling function with amplitude noise

If each frequency component $\cos((\Omega_p - \Omega_q)t)$ entering the coupling function is subject to a small additive error $\eta_{pq}^{\mathrm{amp}}(t) \ll 1$, the network dynamics is given by

$$\dot{\vartheta}_i = \Omega_i + \frac{\varepsilon}{N} \sum_{p=1}^{N} \sum_{q=1}^{N} (w_{pq} \cos((\Omega_p - \Omega_q)t) + \eta_{pq}^{\mathrm{amp}}(t)) \sum_{j=1}^{N} \cos\vartheta_i \sin\vartheta_j$$

Splitting up the right hand side into the ideal dynamics and a perturbation part again yields

$$= \dot{\vartheta}_{i,\mathrm{ideal}} + \frac{\varepsilon}{N} \sum_{k=1}^{N} \sum_{l=1}^{N} \eta_{pq}^{\mathrm{amp}}(t) \sum_{j=1}^{N} \cos\vartheta_i \sin\vartheta_j$$

$$= \dot{\vartheta}_{i,\mathrm{ideal}} + \varepsilon N \overline{\eta^{\mathrm{amp}}}(t) \sum_{j=1}^{N} \cos\vartheta_i \sin\vartheta_j$$

where $\overline{\eta^{\mathrm{amp}}}(t) = \sum_{p=1}^{N} \sum_{q=1}^{N} \eta_{pq}^{\mathrm{amp}}(t)/N^2$ is the error average over all entries of the coupling matrix.

The effect of the error term is especially large, if $\overline{\eta^{\mathrm{amp}}(t)}$ can be written as a Fourier sum of one or more frequency differences present in the system. Let for example $\overline{\eta^{\mathrm{amp}}(t)} = \mu \cos((\Omega_p - \Omega_q)t)$, where μ is a small constant. In this case, the perturbed dynamics becomes

$$\dot{\vartheta}_i = \dot{\vartheta}_{i,\mathrm{ideal}} + \varepsilon N \mu \cos((\Omega_p - \Omega_q)t) \sum_{j=1}^{N} \cos\vartheta_i \sin\vartheta_j.$$

The last term can only be neglected in the equations for $i = p, q$, if $N\mu \ll 1$. This means that the signal-to-noise ratio of the coupling function increases linearly with the number of oscillators. This is a rather mild restriction on scalability because, as it has been already shown, frequency accuracy must improve quadratically with the number of oscillators. Besides, it is reasonable to expect that there are also frequencies present in the noise term that do not contribute to the slow phase shift dynamics but cancel out in the time averaging. $N\mu \ll 1$ is a constraint derived from a worst case scenario.

4.2.5 Change induced by a small separation of time scales

In the experiment, it is desirable to keep the ratio of the time scales given by $\varepsilon_{\text{eff}}/\delta$ as large as possible. The largest usable value of $\varepsilon_{\text{eff}}/\delta$ determines how many oscillators that fulfill condition (2.11) can be built within a given window of available frequencies. On the other hand, if the number of oscillators is fixed, a larger value of ε_{eff} means faster convergence. In the following, a multiple time scale analysis is conducted on the dynamics

$$\dot{\vartheta}_i = \Omega_i + \varepsilon a(t) \frac{1}{N} \sum_{j=1}^{N} \cos \vartheta_i \sin \vartheta_j \tag{4.18}$$

of an ideal network of harmonic oscillators. The fast time is given by $\mathcal{T} = \delta \cdot t$ and the slow time is given by $\tau = \varepsilon t$. A multiple time scale analysis of (4.18), carried out with these two time variables analog to the one in section 2.3 will yield an average solution $\varphi_i(t)$ which is guaranteed to fulfill

$$|\varphi_i(t) - \varphi_{\text{real},i}(t)| = \mathcal{O}\left(\frac{\varepsilon}{\delta}\right)$$

for times up to $t \sim 1/\varepsilon$. Around a hyperbolic fixed point of the ideal dynamics, this approximation is true for all times (Guckenheimer & Holmes 1983). Unfortunately, due to the invariance of the averaged dynamics with respect to a global rotation of phase shifts, no fixed point is hyperbolic, not even the final state during an initialization. Thus, a statement for the proximity of the averaged solution to the true solution can only be made for $t \sim 1/\varepsilon$. In contrast, the time it takes until a pattern is initialized or recognized is at the very least ten times longer (compare any of the simulations shown so far - for the strongly coupled network, $\varepsilon = 1\,\text{s}^{-1}$). Therefore, another, very slow time variable given by $\tau' = \varepsilon^2/\delta \cdot t$ is introduced.

Writing the first three terms of the Taylor series in ε/δ for $\varphi_i(t)$ gives

$$
\begin{aligned}
\varphi_i(t) &= \varphi_i^0(\mathcal{T},\tau,\tau') + \frac{\varepsilon}{\delta}\varphi_i^1(\mathcal{T},\tau,\tau') + \frac{\varepsilon^2}{\delta^2}\varphi_i^2(\mathcal{T},\tau,\tau') \\
\dot{\varphi}_i(t) &= \delta\frac{\partial\varphi_i^0(\mathcal{T},\tau,\tau)}{\partial\mathcal{T}} \\
&\quad + \varepsilon\left(\frac{\partial\varphi_i^1(\mathcal{T},\tau,\tau')}{\partial\mathcal{T}} + \frac{\partial\varphi_i^0(\mathcal{T},\tau,\tau')}{\partial\tau}\right) \\
&\quad + \frac{\varepsilon^2}{\delta}\left(\frac{\partial\varphi_i^2(\mathcal{T},\tau,\tau')}{\partial\mathcal{T}} + \frac{\partial\varphi_i^1(\mathcal{T},\tau,\tau')}{\partial\tau} + \frac{\partial\varphi_i^0(\mathcal{T},\tau,\tau')}{\partial\tau'}\right).
\end{aligned}
$$

The explicit form of $\dot{\varphi}_i$, if $a(t) = \sum_p\sum_q w_{pq}\cos((\Omega_p - \Omega_q)t)$, is given by

$$
\begin{aligned}
\dot{\varphi}_i(t) &= \delta\cdot 0 \\
&\quad + \varepsilon\left(\frac{1}{4N}\sum_{j,p,q} w_{pq}(\sin(\Omega_{iq}^{jp}t + \varphi_j^0 - \varphi_i^0) + \sin(\Omega_{ip}^{jq}t + \varphi_j^0 - \varphi_i^0)\right. \\
&\qquad\left. + \sin(\Omega_q^{ijp}t + \varphi_i^0 + \varphi_j^0) + \sin(\Omega_p^{ijq}t + \varphi_i^0 + \varphi_j^0)\right) \\
&\quad + \frac{\varepsilon^2}{\delta}\left(\frac{1}{4N}\sum_{j,p,q} w_{pq}(\cos(\Omega_{iq}^{jp}t + \varphi_j^0 - \varphi_i^0)(\varphi_j^1 - \varphi_i^1) + \cos(\Omega_{ip}^{jq}t + \varphi_j^0 - \varphi_i^0)(\varphi_j^1 - \varphi_i^1)\right. \\
&\qquad\left. + \cos(\Omega_q^{ijp}t + \varphi_i^0 + \varphi_j^0)(\varphi_i^1 + \varphi_j^1) + \cos(\Omega_p^{ijq}t + \varphi_i^0 + \varphi_j^0)(\varphi_i^1 + \varphi_j^1))\right)
\end{aligned}
$$

with the shorthand notation $\Omega_{iq}^{jp} = \Omega_j + \Omega_p - \Omega_i - \Omega_q$, $\Omega_q^{ijp} = \Omega_i + \Omega_j + \Omega_p - \Omega_q$ and so on. The zeroth order immediately yields

$$
\frac{\partial\varphi_i^0}{\partial\mathcal{T}} = 0 \Leftrightarrow \varphi_i^0(\mathcal{T},\tau,\tau') = c_i^0(\tau,\tau') = \varphi_i^0(\tau,\tau'),
$$

which is the already known fact that the zero-order approximation of $\varphi_i(t)$ is constant at the fast time scale. The first order is given by

$$
\begin{aligned}
\frac{\partial\varphi_i^1(\mathcal{T},\tau,\tau')}{\partial\mathcal{T}} + \frac{\partial\varphi_i^0(\mathcal{T},\tau,\tau')}{\partial\tau} &= \frac{1}{4N}\sum_{j,p,q} w_{pq}(\sin(\Omega_{iq}^{jp}\mathcal{T}/\delta + \varphi_j^0 - \varphi_i^0) + \sin(\Omega_{ip}^{jq}\mathcal{T}/\delta + \varphi_j^0 - \varphi_i^0) \\
&\quad + \sin(\Omega_q^{ijp}\mathcal{T}/\delta + \varphi_i^0 + \varphi_j^0) + \sin(\Omega_p^{ijq}\mathcal{T}/\delta + \varphi_i^0 + \varphi_j^0))
\end{aligned}
$$

On the right hand side, all terms containing an Ω with three upper indices are fast oscillating and therefore non-resonant. The same goes for all terms containing Ω_{iq}^{jp} and Ω_{ip}^{jq} as long as the upper and lower indices are not identical. Considering only the potentially resonant

terms in the first order equation leads to

$$\frac{\partial \varphi_i^1(\mathcal{T}, \tau, \tau')}{\partial \mathcal{T}} + \frac{\partial \varphi_i^0(\mathcal{T}, \tau, \tau')}{\partial \tau} = \frac{1}{2N} \sum_j w_{ij} \sin(\varphi_j^0 - \varphi_i^0)$$

Resonant terms in the fast time derivative of φ_i^1 are avoided if

$$\frac{\partial \varphi_i^0}{\partial \tau} = \frac{1}{2N} \sum_{j=1}^N w_{ij} \sin(\varphi_j^0 - \varphi_i^0), \tag{4.19}$$

which is the result already given in 2.4.2. Integrating the non-resonant terms of the first order equation yields the fast part of φ_i^1:

$$
\begin{aligned}
\varphi_i^1(\mathcal{T}, \tau, \tau') &= -\frac{1}{4N} \sum_{j,p,q} w_{pq} \left(\frac{\delta}{\Omega_{iq}^{jp}\big|_{jp \neq iq}} \cos(\Omega_{iq}^{jp} \mathcal{T}/\delta + \varphi_j^0 - \varphi_i^0) + \frac{\delta}{\Omega_{ip}^{jq}\big|_{jq \neq ip}} \cos(\Omega_{ip}^{jq} \mathcal{T}/\delta + \varphi_j^0 - \varphi_i^0) \right. \\
&\quad \left. + \frac{\delta}{\Omega_q^{ijp}} \cos(\Omega_q^{ijp} \mathcal{T}/\delta + \varphi_i^0 + \varphi_j^0) + \frac{\delta}{\Omega_p^{ijq}} \cos(\Omega_p^{ijq} \mathcal{T}/\delta + \varphi_i^0 + \varphi_j^0) \right) + c_i^1(\tau, \tau') \\
&= -\frac{1}{4N} \sum_{j,p,q} w_{pq} \left(\frac{\delta}{\Omega_{iq}^{jp}\big|_{jp \neq iq}} A_{iq}^{jp}(\mathcal{T}) + \frac{\delta}{\Omega_{ip}^{jq}\big|_{jq \neq ip}} A_{ip}^{jq}(\mathcal{T}) \right. \\
&\quad \left. + \frac{\delta}{\Omega_q^{ijp}} B_q^{ijp}(\mathcal{T}) + \frac{\delta}{\Omega_p^{ijq}} B_p^{ijq}(\mathcal{T}) \right) + c_i^1(\tau, \tau') \tag{4.20}
\end{aligned}
$$

This is a fast oscillating term that is responsible for the small oscillations on top of the slow dynamics observed in the weakly coupled network.

Finally, the second order equation reads

$$
\begin{aligned}
\frac{\partial \varphi_i^2(\mathcal{T}, \tau, \tau')}{\partial \mathcal{T}} &+ \frac{\partial \varphi_i^1(\mathcal{T}, \tau, \tau')}{\partial \tau} + \frac{\partial \varphi_i^0(\mathcal{T}, \tau, \tau')}{\partial \tau'} = \\
&\frac{1}{4N} \sum_{j,p,q} w_{pq} (A_{iq}^{jp} + A_{ip}^{jq})(\varphi_j^1 - \varphi_i^1) + (B_q^{ijp} + B_p^{ijq})(\varphi_i^1 + \varphi_j^1)) \tag{4.21}
\end{aligned}
$$

where the expressions defined in (4.20) must be inserted on the right hand side. This is a tedious procedure resulting in a sum over 6 indices where for each set of indices there are 32 oscillating terms. To obtain the dynamics on the very slow time scale, the resonant terms on the right hand side must be determined. At least in the case of initialization, this process can be simplified, because the fixed point of (4.19) is hyperbolic with the exception of the $(1, 1, \ldots, 1)$ eigenvector of the Jacobian with eigenvalue zero. Therefore, the only relevant dynamics on the very slow time scale is the dynamics of the average phase shift $\overline{\varphi} = 1/N \sum_i \varphi_i$. The average is constant even at the slow time scale because

$$\frac{\partial \overline{\varphi^0}}{\partial \tau} = \frac{1}{N} \sum_{i=1}^N \frac{\partial \varphi_i^0}{\partial \tau} = 0.$$

The average of (4.21) during initialization is given by

$$\frac{\partial \overline{\varphi^2}(\mathcal{T}, \tau, \tau')}{\partial \mathcal{T}} + \frac{\partial \overline{\varphi^1}(\mathcal{T}, \tau, \tau')}{\partial \tau} + \frac{\partial \overline{\varphi^0}(\mathcal{T}, \tau, \tau')}{\partial \tau'} = \frac{1}{4N^2} \sum_{i,j,p,q} (B_q^{ijp} + B_p^{ijq})(\varphi_i^1 + \varphi_j^1)), \qquad \text{(4.22)}$$

where still the φ_i^1 given by (4.20) must be inserted. In (4.22), the coordinates were chosen such that $w_{ij} = 1$ for all i and j. To obtain the very slow dynamics of $\overline{\varphi}$, only the resonant terms on the right hand side are considered:

$$\frac{\partial \overline{\varphi^1}(\mathcal{T}, \tau, \tau')}{\partial \tau} + \frac{\partial \overline{\varphi^0}(\mathcal{T}, \tau, \tau')}{\partial \tau'} = \text{res.terms} \left(\frac{1}{4N^2} \sum_{i,j,p,q} (B_q^{ijp} + B_p^{ijq})(\varphi_i^1 + \varphi_j^1)) \right)$$

All potentially resonant terms on the right hand side contain either a factor of the form $A_{cd}^{ab} B_h^{efg}$ or a factor of the form $B_d^{abc} B_h^{efg}$, where each of the letters $a-h$ stands for an arbitrary index. Resonances in the first factor can only occur if the sum over five frequencies present in the system is equal to the sum of three other frequencies. This is possible, though not very likely, but could be avoided entirely by choosing $\Omega_{\min} > 0.6 \cdot \Omega_{\max}$. Therefore, the overwhelming resonant contribution will be by terms of the second form, which will be given by

$$-\frac{\delta}{\Omega_d^{abc}} B_d^{abc} B_h^{efg} = -\frac{\delta}{2\Omega_d^{abc}} \left(1 + \cos(2\Omega_d^{abc} t + \varphi_a^0 + \varphi_e^0) \right)$$

for any combination of indices that fulfills $\Omega_d^{abc} = \pm \Omega_h^{efg}$. The resonant part of these terms is strictly negative, which means that on the very slow time scale, the average of the phase shifts decreases with a constant rate:

$$\frac{\partial \overline{\varphi^1}(\mathcal{T}, \tau, \tau')}{\partial \tau} + \frac{\partial \overline{\varphi^0}(\mathcal{T}, \tau, \tau')}{\partial \tau'} = \frac{\partial \overline{c^0}(\tau, \tau')}{\partial \tau'} + \frac{\partial \overline{c^1}(\tau, \tau')}{\partial \tau} = \text{const} < 0$$

During the recognition phase, the picture is more complicated. At any steady state, there are M eigenvalues equal to zero, therefore the very slow dynamics in an M-dimensional subspace of the system is relevant. As a consequence, much like it is the case for inaccurate oscillator frequencies, the branches of the recognized state will tend to split up along those eigenvectors. Still, there will be a negative tendency in the phase shift average. These drift terms, along with the small oscillations of the slow dynamics with an amplitude of ε/δ and the imperfections that are present even in the ideal dynamics, lead to a breakdown of reliable pattern recognition if ε is not sufficiently smaller than δ.

Fig. 29 illustrates the effect of a small separation of time scales for $\varepsilon/\delta = 0.5$, $\varepsilon/\delta = 1$ and $\varepsilon/\delta = 1.5$. The initialization succeeds in all three cases, although there is a common drift towards smaller values of the phase shift variable for all oscillators. This is the very slow time scale effect discussed above. Like the amplitude of the oscillations around the average dynamics, this effect increases with increasing ε. While in Fig. 29a, with $\varepsilon/\delta = 0.5$,
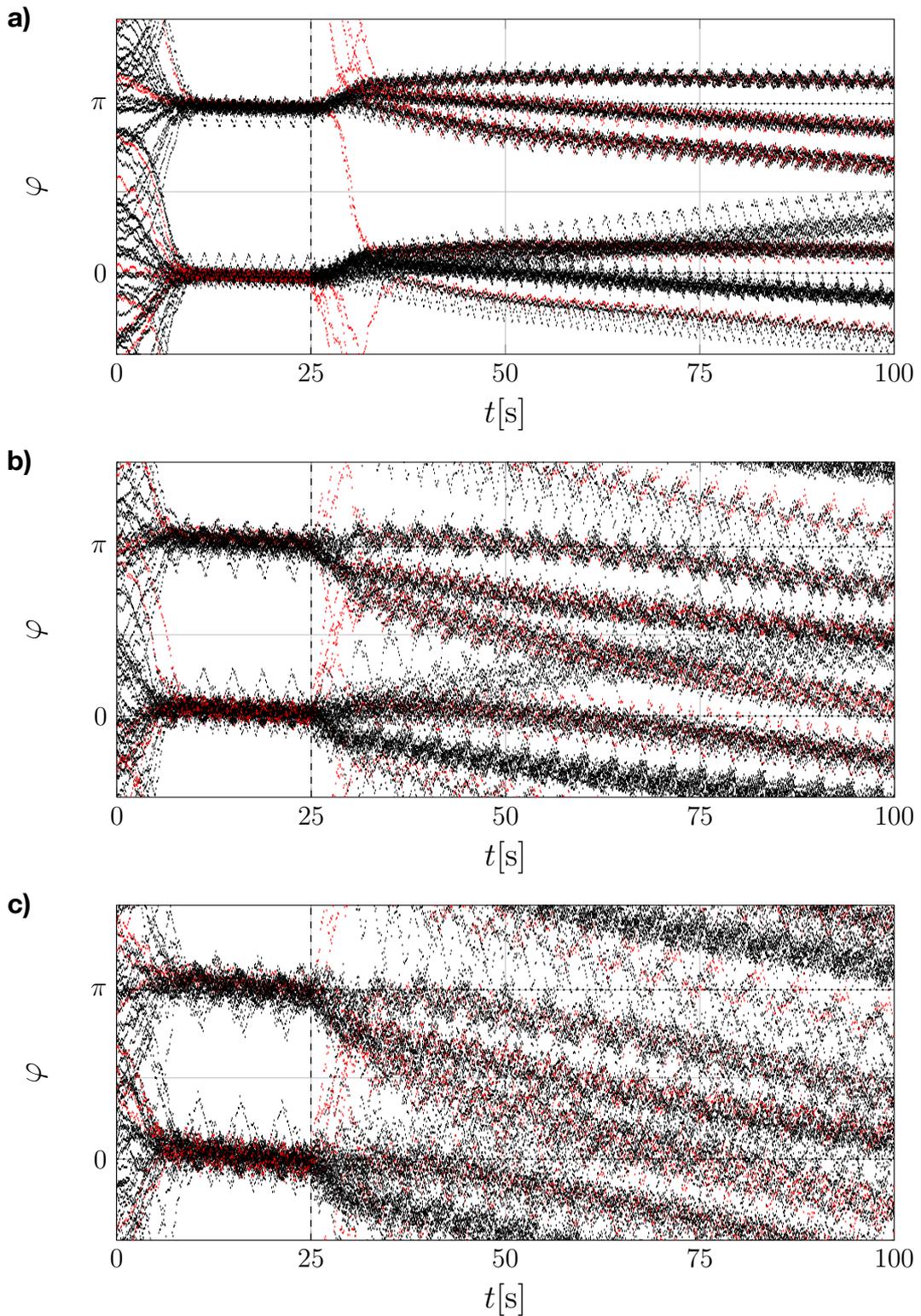
Fig. 29: a) Numerical time integration of the phase shifts in (4.18) with random initial conditions, $N = 60$ oscillators, $M = 3$ memorized patterns and 8 erroneous bits in the initial pattern. The initial and memorized patterns are the same as in Fig. 4. Red phase shift curves belong to phase shifts that correspond to erroneous bits in the initial pattern. Oscillation frequencies were chosen as a Golomb ruler between $10\,\text{kHz}$ and $20\,\text{kHz}$. The coupling strength was set to $\varepsilon = 0.5 \cdot \delta$. The coupling was switched to recognition mode at $t = 25\,\text{s}$. **b)** Same as a), with oscillation frequencies between $5\,\text{kHz}$ and $10\,\text{kHz}$ and $\varepsilon = 1 \cdot \delta$. **c)** Same as a), with oscillation frequencies between $3.5\,\text{kHz}$ and $7\,\text{kHz}$ and $\varepsilon = 1.5 \cdot \delta$.
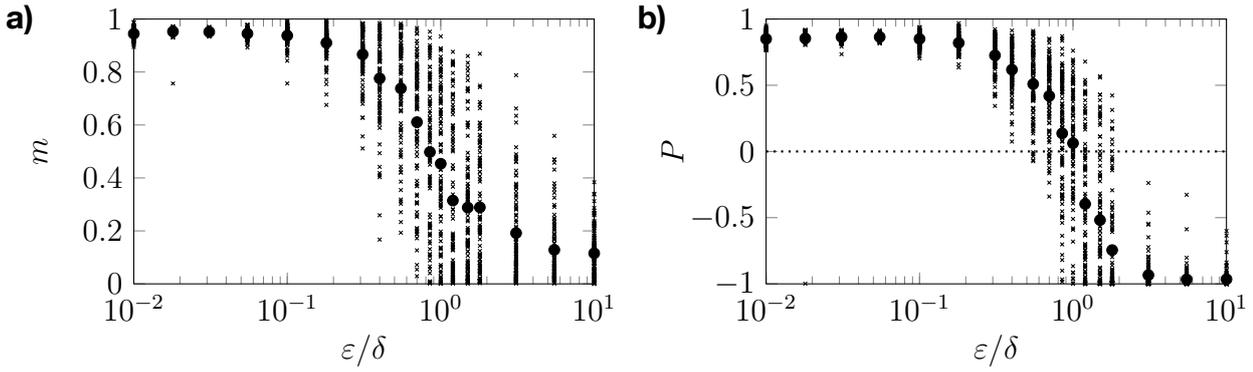
Fig. 30: a) Overlap of the final pattern with the correct memorized pattern vs. the ratio of time scales for a series of simulations of (4.18) with $N = 16$ oscillators, $M = 3$ randomly selected orthogonal patterns and one erroneous bit in the initial pattern. The coupling was switched to recognition mode at $t = 25\,\mathrm{s}$, the final state was evaluated at $t = 50\,\mathrm{s}$. Oscillation frequencies were chosen as a Golomb ruler with $\delta = 1.595\,\mathrm{Hz}$ and $\Omega_{\mathrm{min}} = 0.5 \cdot \Omega_{\mathrm{max}}$. Small crosses indicate results for individual simulations, the filled circles are the average overlap. **b)** Proximity of the final pattern to the correct memorized pattern vs. the ratio of time scales for the same simulations. Small crosses indicate results for individual simulations, the filled circles are the average proximity.

there is a time window of roughly one minute, in which the correctly memorized pattern is represented by the network, the recognized state is barely realized in Fig. 29b, with $\varepsilon/\delta = 1$, and gets destroyed immediately because the phase shift curves are wildly drifting and oscillating. In Fig. 29c, with $\varepsilon/\delta = 1.5$, finally the deviations from ideal behavior are so large that the correct pattern never is realized in the network.

 To get an idea of how large a value of ε/δ still allows for reliable pattern recognition, a series of simulations of (4.18) over a wide range of values for ε/δ was performed for a network with 16 oscillators, 3 randomly chosen mutually orthogonal patterns and one defective bit. The results are presented in Fig. 30. The highest ratio of time scales for which every single simulation was successful is $\varepsilon/\delta = 0.4$, although recognition nearly failed in some of them. Thus, $\varepsilon/\delta \leq 0.3$, corresponding to

$$\varepsilon_{\mathrm{eff}} \leq 0.15 \cdot \delta.$$

should be a safe threshold value under which no problems with pattern recognition are to be expected. Together with the limitations for $\Delta\Omega$ that were investigated earlier, this condition can be used to derive a relation between the required frequency accuracy and the number of oscillators in the network; this is done in the next section.

Note that $\varepsilon_{\mathrm{eff}}$ and ε differ by a factor of two because the oscillators are only coupled in one of the two dynamical variables. It is conceivable to let the coupling function act on both variables, resulting in a phase equation of

$$\dot{\vartheta}_i = \Omega_i + \varepsilon a(t)\frac{1}{N}\left(\sum_{j=1}^{N} \sin\vartheta_j \cos\vartheta_i - \sum_{j=1}^{N} \cos\vartheta_j \sin\vartheta_i\right)$$

Such a coupling would both provide the additional coupling strength to guarantee that $\varepsilon_{\text{eff}} = \varepsilon$ and avoid the global drift during initialization; the resulting equations are exactly

$$\dot{\vartheta}_i = \Omega_i + \varepsilon a(t) \frac{1}{N} \sum_{j=1}^{N} \sin(\vartheta_j - \vartheta_i),$$

as suggested by Hoppensteadt & Izhikevich (1999).

4.3   Consequences for the scalability of the network

Before the consequences of the findings described above for the experimental network are discussed, these findings should be considered in the context of previous research. There have been investigations of two types of Hebbian networks of Kuramoto oscillators with distributed frequencies and white noise before. The first type was investigated by Arenas & Pérez-Vicente (1994), Park & Choi (1995) and Pérez-Vicente et al. (1996), who considered the network

$$\dot{\vartheta}_i = \Omega_i + \gamma_i(t) + \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{M} \cos(\xi_i^k - \xi_j^k) \sin(\varphi_j - \varphi_i)$$

where the ξ_i^k are random phases, the Ω_i are narrowly distributed around a common frequency and $\gamma_i(t)$ is an additive white noise term. The second type, investigated by Aonishi et al. (1999), Yamana et al. (1999) (both without the white noise term) and Yoshioka & Shiino (2000), is given by

$$\dot{\vartheta}_i = \Omega_i + \gamma_i(t) + \frac{1}{N} \sum_{j=1}^{N} w_{ij} \sin(\varphi_j - \varphi_i + \beta_{jj}).$$

This model allows for complex pattern vectors of the form $\left|\xi^k\right\rangle_i = e^{\varphi_i^k}$ with arbitrary φ_i^k. The coupling coefficients are given by $w_{ij} = |C_{ij}|$ and the synaptic delays by $\beta_{ij} = \arg C_{ij}$ with $C_{ij} = \sum_{k=1}^{m} \left\langle \xi^k \,\middle|\, j \right\rangle \left\langle i \,\middle|\, \xi^k \right\rangle$. Note that both types of network contain the ideal dynamics given by (2.4) as a special case.

The important result common to both networks is that in the thermodynamic limit $N \to \infty$, there is a phase transition between a so-called "glassy state" and a "condensed" memorized pattern (Park & Choi 1995, Aonishi et al. 1999). This transition depends on the three macroscopic parameters α, σ and T, where α is the load rate of the network, σ (in general the standard deviation of a normal distribution) governs the width of the frequency distribution and T is a temperature determining the level of the white noise. In the limit of $\alpha \to 0$ and $T \to 0$, this phase transition is equivalent to the transition in the original Kuramoto model.

This macroscopic view on the network establishes that there is a surface of critical values in the three dimensional parameter space that separates the regime in which the system is capable of pattern recognition from the glassy regime. It does not, however, deal with the question whether and how the condensed patterns are separated from each other within the "memorized" state. After all, the states representing different patterns should be thermodynamically equivalent and therefore not separated by a phase transition. This question is addressed in section 4.1, which presents a proof that in the case of mutually orthogonal memorized patterns, all condensed patterns form part of a larger set of steady states; it also presents analytical and numerical results suggesting that the restriction of

mutually orthogonality can be dropped.

As a result of this inner structure of the memorized state, it appears that, in the network investigated in this thesis, there is no critical value for the width σ of the frequency distribution, above which the degradation of condensed memorized patterns sets in; in contrast, an arbitrarily slow degradation will be found for arbitrarily narrow distributions. This degradation is also not influenced by the number of memorized patterns determined by α, because as soon as there is more than one memorized pattern, the system may alternate between patterns in the presence of frequency inaccuracies. This is why, instead of critical values depending on the frequency distribution and the load rate, section 4.2 gives constant threshold values for different types of non-ideal behavior.

Because, other than in the aforementioned studies, the Hebbian Kuramoto model with the network at hand only arises by time-averaging a more complicated dynamics, there is an additional relation between the number of oscillators and the available frequency accuracy of the system (i.e. the width σ of the frequency distribution) in the ideal dynamics, posing severe requirements on the latter.

As a result, with respect to the scalability of an experimental network, two of the error sources discussed in this chapter stand out because their degrading effect scales with N^2, namely frequency inaccuracies in both the oscillators themselves and in the coupling function.

It is reasonable to assume that the coupling mechanism can be implemented with much greater frequency accuracy than the oscillators themselves, simply because the phases of the components of the coupling function are fixed. No sensitivity for any kind of feedback is needed. Therefore, any frequency stabilizing measure, up to using an atomic clock, in principle is available for the implementation of the coupling. With the oscillators, it is a different story. Since their phases need to be adaptable, the frequencies are not entirely fixed. The wiggle room needed for pattern recognition to occur makes the oscillation susceptible to undesired frequency changes due to noise. In conclusion, the single largest factor limiting scalability is the frequency accuracy of the oscillators used in the experiment, for which the condition

$$\Delta\Omega \leq 0.3 \cdot \varepsilon_{\text{eff}}$$

must hold (see section 4.2.1).The coupling strength ε_{eff} must fulfill

$$\varepsilon_{\text{eff}} \leq 0.15 \cdot \delta = 0.15 \cdot \min_{i \neq j} |\Omega_i - \Omega_j|$$

(taken from section 4.2.5). When the two conditions are combined, the result is that the expected error in the frequency should be roughly twenty times smaller than the minimal

difference of two frequencies present in the network:

$$\Delta\Omega_i \leq 0.05 \cdot \min_{i \neq j} |\Omega_i - \Omega_j|$$

There is a more concise formulation for this condition, if some reasonable assumptions are made about the frequency accuracy of the oscillators and the frequency range. First, if only one type of oscillator is used, the relative frequency accuracy $\Delta\Omega_i/\Omega_i$ should have roughly the same value across all oscillators. Let this value be denoted by $\Delta\Omega/\Omega$. Second, the frequencies $\Omega_{\min}$, $\Omega_{\max}$ as well as the range $\Omega_{\max} - \Omega_{\min}$ should all be in the same order of magnitude, denoted by Ω. If not, then either the available frequency space is used poorly (if $\Omega_{\max} - \Omega_{\min} \ll \Omega_{\min}$) or the higher frequency oscillators must have a much higher relative accuracy than the lower oscillators (if $\Omega_{\min} \ll \Omega_{\max}$). With these assumptions and with the scaling of the Golomb ruler taken from (2.13), one can write:

$$\Delta\Omega \leq 0.05 \cdot \frac{\Omega_{\max} - \Omega_{\min}}{N^2} \approx 0.05 \cdot \frac{\Omega}{N^2}.$$

Therefore, to construct a network of N oscillators, a frequency accuracy of

$$\frac{\Delta\Omega}{\Omega} \approx 0.05 \cdot \frac{1}{N^2}$$

or better is needed, although for low numbers of oscillators, where the length of available Golomb rulers is close to $N^2/2 \cdot \delta$, a factor of 0.1 instead of 0.05 will be sufficient.
Note that an increase in the number of oscillators is also accompanied by a decreasing coupling strength which should not surpass

$$\varepsilon_{\text{eff}} \approx 0.15 \cdot \frac{\Omega}{N^2}.$$

This means that the minimal time it takes to recognize a pattern ($t_{\text{run}} \approx 2 \times 10/\varepsilon_{\text{eff}}$ for both initialization and recognition) is in the order of

$$t_{\text{run}} \approx 100 \cdot \frac{N^2}{\Omega},$$

where Ω is a typical frequency of an oscillator. In other words, in this model, an increase in parallelization is not accompanied by an increase in computation speed, because although quadratically more computational steps are executed simultaneously, they also take a quadratically longer time. When a time-dependent coupling is used to implement neural synaptic coupling, this must be taken into account.

5. Design of the experiments

In principle, a set of arbitrary harmonic oscillators could be used to implement a network that behaves like 2.24, including mechanical, chemical and electrical oscillators. Depending on the type of oscillator, there is a wide range of possibilities for the coupling as well. To narrow down the field of possibilities, at the beginning of the thesis the choice was made to implement the network with electrical van der Pol oscillators. For a proof of principle, those excel in terms of simplicity and are easy to tune to a wide range of frequencies. There might be better choices if large numbers of oscillators and/or especially high frequencies are required; however, this kind of performance tuning is beyond the scope of the work at hand.

Even so, there still is a number of possibilities to realize the global coupling mechanism. The coupling variable could be both the voltage across the oscillator or the current through it, or even a combination of both. This depends mainly on the way the individual oscillators are connected (i. e., either all in parallel or all in series, because all of them must be interchangeable for a global coupling) and on the physical quantity governing the coupling strength (i. e., whether the coupling is inductive, capacitive or resistive, or some combiniation).

For this thesis, a global coupling through an external resistance was chosen. This approach has been studied theoretically (Mazouz et al. 1997, Krischer et al. 2003, Hölzel 2007) and experimentally (Kiss et al. 1999, Kiss et al. 2002, Mikhailov et al. 2004) but mainly for strong coupling (in the sense, that the coupling strength was beyond the Kuramoto transition) and never in the weak, time-dependent form that is treated in this thesis.

5.1   Selection of the individual oscillators

The ideal oscillator for the associative network should

- exhibit harmonic oscillations with a harmonic phase response curve,

- have a stable frequency

- have a wide range of available frequencies and

- be simple to build.

The first requirement follows from the analysis presented in 2.4.2. The second is a consequence of the sensitivity to frequency deviations shown by the network, as discussed in 4.2.1. The third condition allows to freely adjust the frequencies to meet the marks of a Golomb ruler, while the fourth is obviously desirable in any realization.

Van der Pol oscillators are almost ideal in terms of waveform and phase response

a)

b)

c)

d)
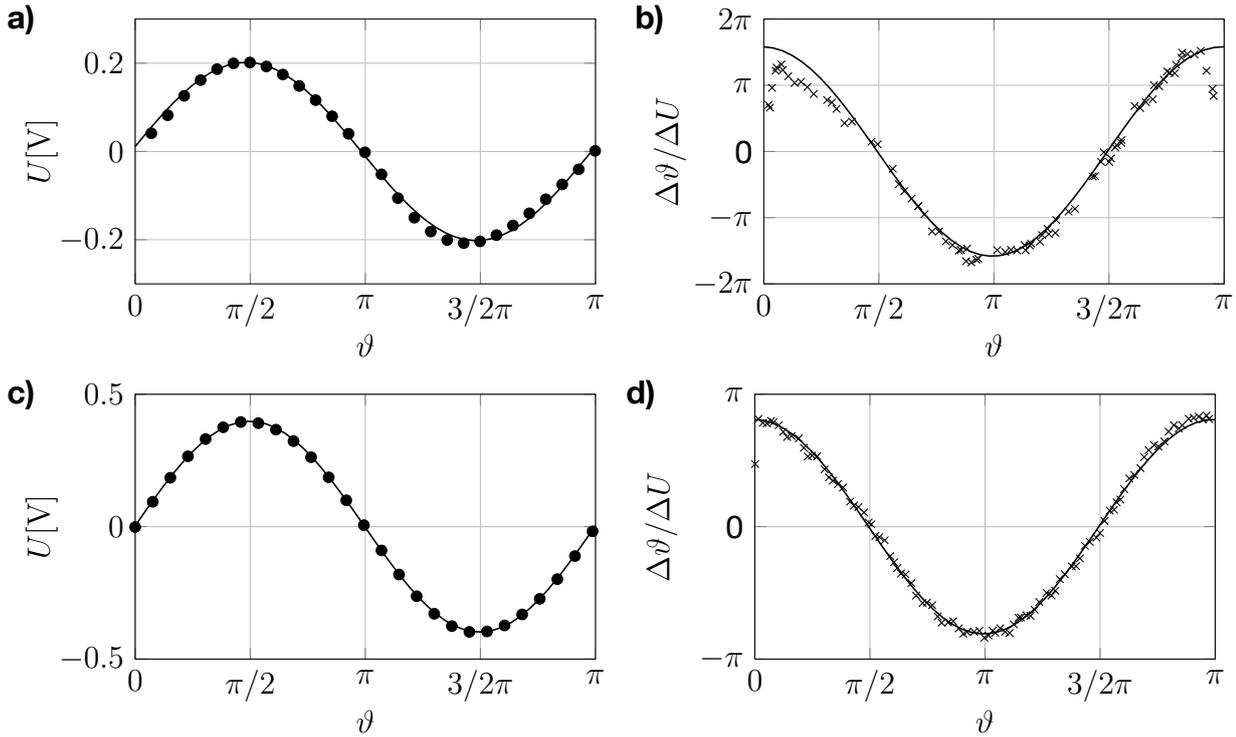
Fig. 31: a) Waveform of a TD-type oscillator with $L = 47\,\mu\mathrm{H}$, and $C = 100\,\mathrm{nF}$ (black dots). **b)** Results of a series of phase response measurements of the same TD-type oscillator (crosses). **c)** Waveform of a NIC-type oscillator with $C = 1\,\mathrm{nF}$ and $f = 67\,\mathrm{kHz}$ (black dots). **d)** Results of a series of phase response measurements of the same NIC-type oscillator (crosses). In panels a), c) and d), the solid line is a fit of the function $f(\vartheta) = A\sin(\vartheta + \phi)$ to the experimental data with A and ϕ as parameters. In panel b), the experimental data used for the fit was restricted to $\vartheta \in [0.5, 2\pi - 0.5]$. This was done because due to the parasitic oscillations of the tunnel diode, which occur especially in the rising flank of the waveform (i.e. around $\vartheta \approx 0$), the phase response is systematically underestimated around $\vartheta \approx 0$.

characteristics (see Fig. 31), if the parameters are chosen appropriately, with a frequency stability of up to $\Delta f/f \approx 10^{-4}$ (Horowitz & Hill 1989).

 In the methods chapter, two possible designs for a van der Pol oscillator were presented (3.2.3). Both of these circuits are basically LC-circuits with a negative damping for small amplitudes of the oscillation, governed by the dynamical equations

$$\dot{U} = \frac{1}{C}\left(-I_{\mathrm{NL}}(U, R_{\mathrm{neg}}) - I\right)$$
$$\dot{I} = \frac{U}{L} \tag{5.1}$$

The TD-type oscillator has the free parameters $f_0 = (2\pi\sqrt{LC})^{-1}$ and $\gamma = C/L$, which determine the frequency and the separation of time scales for U and I. Note that $\sqrt{\gamma}$ is the conductance of both the capacitive and the inductive branch of the LC-circuit at the natural frequency f_0. In case of the NIC-type oscillator there is the additional parameter R_{neg}, which for the TD-type oscillator is fixed at $R_{\mathrm{neg}} \approx 60\,\Omega$. All of these parameters influence both oscillation frequency and shape. There are parameter regions, where oscillations are distorted and/or have a frequency f far away from the frequency $f_0 = (2\pi\sqrt{LC})^{-1}$ of the
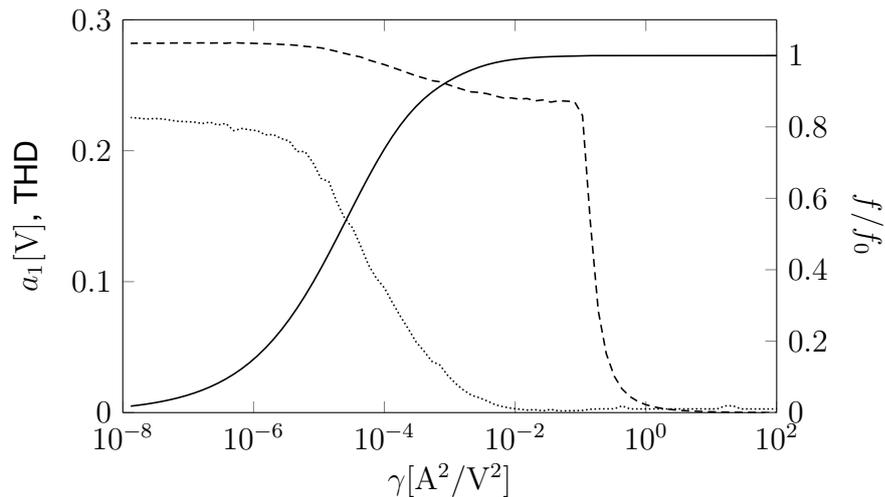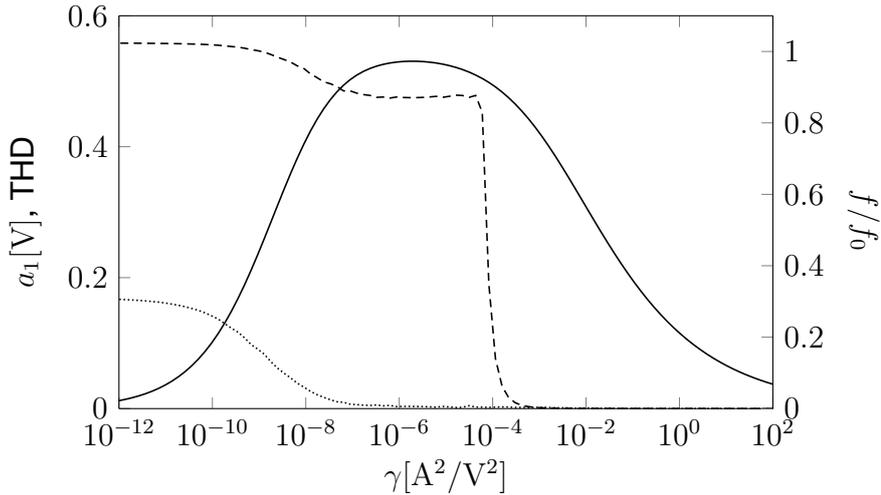
Fig. 32: Ratio of actual frequency f to $f_0 = (2\pi\sqrt{LC})^{-1}$ (solid line), amplitude of the first harmonic a_1 (dashed line) and total harmonic distortion (dotted line) of the oscillations of the oscillator in Fig. 13 vs. $\gamma = C/L$. The plots are based on SPICE simulation results.

ideal LC-circuit. There are also large parameter regions in which (5.1) exhibits sustained oscillations which immediately break down in the presence of small parasitic resistances or some noise in the circuit, because the system is very close to the Hopf bifurcation. All these parameter regions have to be ruled out for the use in the experiments.

Fig. 32 and Fig. 33 show, for a TD-type oscillator with $f_0 = 30\,\text{kHz}$ and a NIC-type oscillator with $f_0 = 30\,\text{kHz}$ and $R_{\text{neg}} = 10\,\text{k}\Omega$, respectively, how three criteria relevant for the assessment of the experimental suitability of the oscillator depend on the choice of γ. These criteria are the THD of the waveform, its amplitude and the ratio f/f_0. The values of these quantities were determined from the results of SPICE simulations (see A3 for the source code) of the respective oscillator at the particular set of parameter values. The simulations were started at the unstable equilibrium $U = I = 0$, with the external voltage source starting out at zero voltage (in that way, the system was driven to oscillate). They were run for a time interval of $50/f_0$, where THD, amplitude and f/f_0 were determined for the first complete period after a time interval of $40/f_0$.

The following trends can be observed:

- For both oscillator types, in the limit of very small γ, the oscillations become relaxational and lose their sinusoidal shape, reflected by a substantial nonzero value of the THD. A detailed discussion of this effect of different time scales can be found for example in (Kanamaru 2007). The distortion from the sinusoidal shape is accompanied by a decrease in frequency, because the system now acts like an RL-circuit with a negligible parallel capacitance and a characteristic time scale of L/R_{neg} rather than an LC-circuit with a time scale of $\sqrt{LC}$. The transition occurs in the parameter region where the impedances of the nonlinear branch and the capacitive branch are roughly equal, namely around $R_{\text{neg}} \approx (2\pi f_0 C)^{-1} = \sqrt{1/\gamma}$.

Fig. 33: Ratio of actual frequency f to $f_0 = (2\pi\sqrt{LC})^{-1}$ (solid line), amplitude of the first harmonic a_1 (dashed line) and total harmonic distortion (dotted line) of the oscillations of the oscillator in Fig. 14 for $R_{\mathrm{neg}} = 10\,\mathrm{k\Omega}$ vs. $\gamma = C/L$. The plots are based on SPICE simulation results.

Therefore, a larger value of R_{neg} moves the transition towards smaller values of γ.

- In the limit of large γ on the other hand, the system acts like a weakly perturbed harmonic oscillator, because the currents through the inductive branch become huge compared to the currents through the nonlinear element, due to the fact that $R_{\mathrm{neg}} \gg 2\pi f_0 L = \sqrt{1/\gamma}$. There is a sharp drop to zero amplitude for both oscillators above a certain value of γ (i.e. $\gamma \approx 0.03\,\mathrm{A^2/V^2}$ for the TD-type and $\gamma \approx 1 \cdot 10^{-4}\,\mathrm{A^2/V^2}$ for the NIC-type oscillator). This drop, however, is an artifact of the way the amplitude was computed, namely by giving the oscillator only a certain time (i.e 50 periods) to spiral out of the unstable steady state. By waiting an arbitrarily long time, the threshold could be pushed to arbitrarily large values of γ. Beyond some value of γ, these oscillations will be never observed experimentally; it is, however, not self-evident where the threshold should be placed to correctly reflect the experimental breakdown of oscillations. Therefore, while this absence of sustained oscillations for too large values of γ was an important factor for the choice of the experimental parameters in the course of this thesis, it is hard to use as a rigorous criterion to rule out certain parameter regions. A more rigorous way to establish a threshold on γ is the introduction of small parasitic serial resistance R_L in the inductance L. This will introduce a Hopf bifurcation at $\gamma = 1/(R_{\mathrm{neq}}R_L)$ beyond which there are no stable oscillations. For example, in the case of the TD-type oscillator Fig. 32 deals with, the amplitude cliff is around $\gamma \approx 0.03\,\mathrm{A^2/V^2}$, which, at $f_0 = 30\,\mathrm{kHz}$, corresponds to $L \approx 20\,\mu\mathrm{H}$. For a Hopf bifurcation at $\gamma = 0.03\,\mathrm{A^2/V^2}$, the parasitic resistance of this inductance should be $R_L \approx 50\,\mathrm{m\Omega}$, which is close to the specifications for the coils used to build the TD-type oscillators. In other words, the γ value of the amplitude cutoff depicted in Fig. 32 is realistic, although it is based on an initially arbitrarily
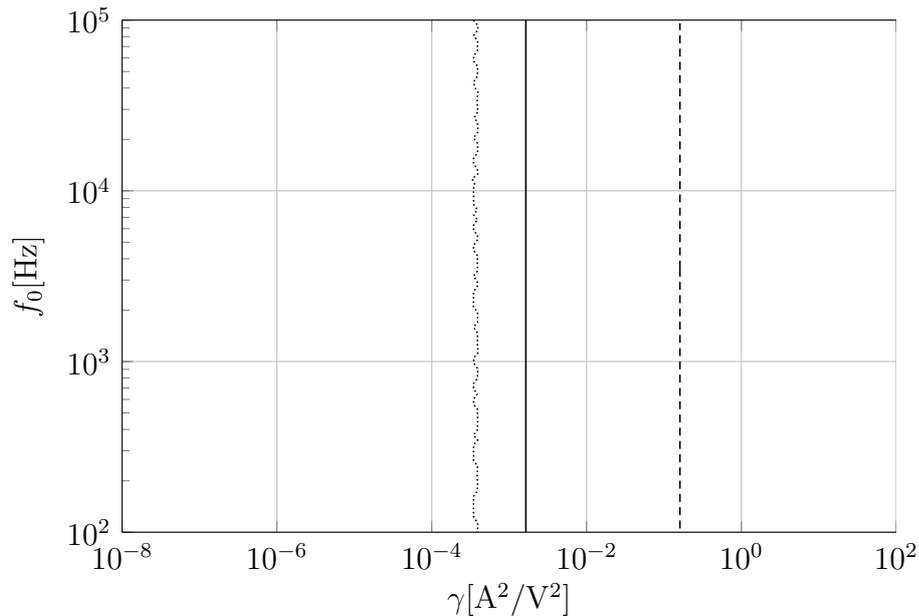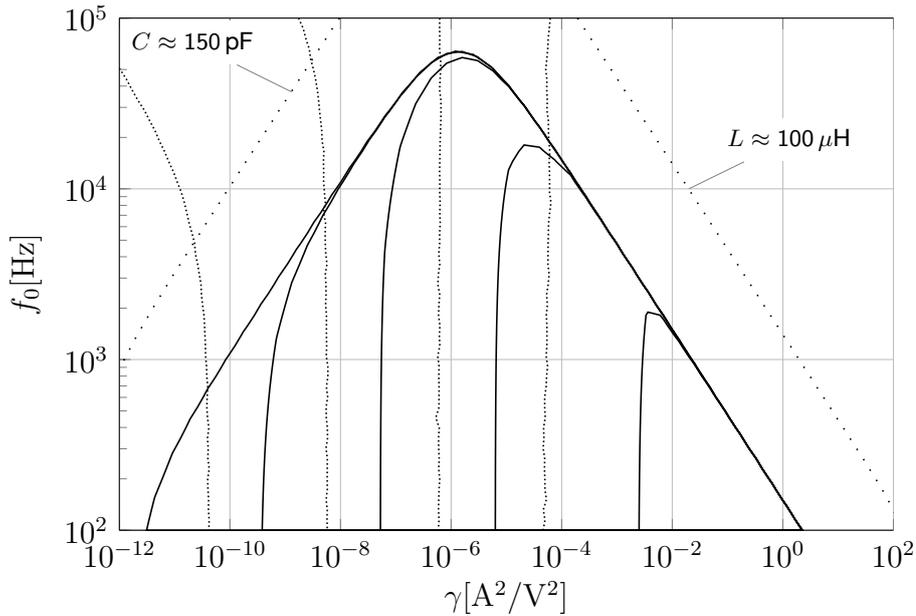
Fig. 34: Parameter space of a TD-type oscillator. Shown are the regions of sustained oscillations (area left of the dashed line), a THD below 0.05 (area right of the dotted line) and a ratio f/f_0 above 0.95 (area right of the solid line). The THD and f/f_0 curves are contour lines determined with a grid of 101×101 SPICE simulations with even logarithmic spacing of the parameter values. The dashed line is the Hopf bifurcation for $R_L = 50\,\text{m}\Omega$, which, according to the datasheet, is the serial resistance R_L of the $47\,\mu\text{H}$ coils that were used in the experiment.

chosen time interval of the simulation of an ideal model system. Unfortunately, this argument cannot be simply extended to the NIC-type oscillator, because the gyrator was, among other reasons, designed to overcome the problems of the parasitic serial resistance of other types of inductors. Still, there are other effects present in NIC-type oscillators as well as in TD-type oscillators, like noise for example, which are also prone to suppress oscillations if γ becomes too large.

- Finally, also in the limit of large γ, but only for the NIC-type oscillator, the frequency ratio f/f_0 decreases, similar to the limit of very small γ. The reasons for this effect will be discussed further below.

In the analysis of the parameter space so far, the frequency f_0 of oscillators was excluded, as well as the negative differential resistance R_{neg} of the nonlinearity. The first two effects discussed above are not expected to depend on f_0, because they are only governed by a relationship between R_{neg} and γ. This is illustrated by Fig. 34, which shows the boundaries of the region in parameter space which is eligible for the use in experiments for TD-type oscillators with a parasitic resistance $R_L = 50\,\text{m}\Omega$.

The f_0-γ-plane in the parameter space of the NIC-type oscillator was investigated for several values of R_{neg}, because its negative differential resistance can be adjusted much more flexible than that of the TD-type oscillator. The results are shown in Fig. 35. The main finding illustrated in this figure is that the area where $f \approx f_0$ is considerably smaller than for

Fig. 35: f_0-γ-plane in the parameter space of a NIC-type oscillator. Shown are the regions with a THD below 0.05 (area right of the densely dotted line) and a ratio f/f_0 above 0.95 (area beneath the solid line) for five different values of R_{neg}. Those values are, from right to left, $100\,\Omega$, $1\,\mathrm{k\Omega}$, $10\,\mathrm{k\Omega}$, $100\,\mathrm{k\Omega}$ and $1\,\mathrm{M\Omega}$. The THD and f/f_0 curves are contour lines determined with a grid of 101×101 SPICE simulations with even logarithmic spacing of the parameter values. The THD contour line for $R_{\mathrm{neg}} = 1\,\mathrm{M\Omega}$ is located at lower values than the depicted range of γ and therefore not shown. The loosely dotted lines indicate the location of all oscillators with either $C = 150\,\mathrm{pF}$ or $L = 100\,\mu\mathrm{H}$, which appear to be the minimal values that can be reached with the setup in Fig. 14.

the TD-type. The boundary of this region (solid lines in Fig. 35) has three distinct segments: At $L \approx 1\,\mathrm{mH}$ and below, constant stray inductances in the system lower the frequency f with respect to f_0 (segment of negative slope). Below $C \approx 1\,\mathrm{nF}$, constant stray capacitances play a similar role (segment of positive slope). More precisely, the location of the boundary indicates that the circuit has a minimal capacitance $C_{\mathrm{min}} \approx 150\,\mathrm{pF}$ and a minimal inductance $L_{\mathrm{min}} \approx 100\,\mu\mathrm{H}$. As a result, the effectively accessible parameter region for the NIC-type oscillator regarded here is bounded by the two loosely dotted lines in Fig. 35. These minimal values depend on the internal properties of the integrated circuits that were used. They appear to be rather high; they result, however, from SPICE calculations using tested vendor models for all parts involved and should be valid nonetheless. Also note that the segment with positive slope in the $f/f_0 = 0.95$ contour line appears only if R_{neg} is large enough that the system still acts as an LC-circuit rather than an RL-circuit. Otherwise, there is a vertical segment of the boundary, as it is the case for the TD-type oscillator. This vertical segment, along with the region of a THD below 0.05, moves towards smaller values of γ with increasing R_{neg}.

In summary, the conclusions from the analysis of presented above are:

- The ratio of time scales γ should be chosen such that $\gamma \geq 1/R_{\mathrm{neg}}^2$, to avoid relaxational oscillations but not be so large that $\gamma \gg 1/R_{\mathrm{neg}}^2$, in order to still have

sustained oscillations under non-ideal conditions. This leaves a preferred value range of roughly two orders of magnitude for γ for any given R_{neg}.

- For the NIC-type oscillators, frequencies much larger than $1\,\mathrm{MHz}$ are unattainable because of the effective reduction of the available parameter space due to stray capacitances and inductances in the circuit.

Apart from these limitations, the parameters can be chosen freely in a manner most suited for the demands of the coupling mechanism (see next section).

5.2   Selection of the coupling mechanism

The first part of this section explains a global coupling with an external resistor, which is a common way to couple oscillators with the electric potential as oscillating variable. In this thesis, a time-dependence is introduced to the coupling resistance - it turns out, however, that the straightforward implementation has some undesirable limitations that restrict its usability for the type of coupling needed for pattern recognition. In particular, it is very hard to implement a coupling function varying sinusoidally around zero in time with a substantial amplitude. The second section presents an improved version of the coupling circuitry that eliminates these problems, but still has a minor issue, namely an unwanted distortion of the coupling function from a sum of purely sinusoidal signals. This problem can be circumvented by using an entirely different approach, in which the correct coupling input to the oscillators is calculated from an external voltage signal with the help of active analog circuit elements. This approach is presented in the third part of the section. Both the second and the third approach were used in the experiments.

5.2.1   Coupling through an external resistor

The straightforward way to implement a coupling via an external resistor is depicted in Fig. 36a. Each oscillator is connected to a common node via an "internal" resistance R_{int}. The common node is connected to ground via the time-dependent external resistance $R_{\mathrm{ext}}(t)$. The presence of R_{int} provides the degree of freedom for each oscillator to run at its own frequency. R_{int} is part of the local dynamics of each individual oscillator (therefore "internal"). The dynamical equations for a single oscillator in the coupled network, following from Kirchhoff's laws, are:

$$
\begin{aligned}
\dot{U}_i &= \frac{1}{C_i}\left(-I_{\mathrm{NL}}(U_i) - I + \frac{U_{\mathrm{ext}} - U_i}{R_{\mathrm{int}}}\right) \\
\dot{I}_i &= \frac{U_i}{L_i}
\end{aligned}
$$

Where the common external potential U_{ext} is given by

$$
U_{\mathrm{ext}} = \frac{R_{\mathrm{ext}}(t)}{R_{\mathrm{int}} + N R_{\mathrm{ext}}(t)} \sum_{j=1}^{N} U_j
$$

The inclusion of R_{int} effectively changes the negative differential resistance of the nonlinear element (compare 3.2.3): The new value is given by

$$
R_{\mathrm{neg,eff}} = \frac{R_{\mathrm{neg}} R_{\mathrm{int}}}{R_{\mathrm{int}} - R_{\mathrm{neg}}},
$$

where

$$R_{\mathrm{neg}} = \left(-\frac{\mathrm{d}I_{\mathrm{NL}}}{\mathrm{d}U} \right)^{-1}_{U=0}.$$

To preserve the stable oscillations, it is important that $R_{\mathrm{int}} > R_{\mathrm{neg}}$, because otherwise, $-1/R_{\mathrm{neg,eff}} \geq 0$, which means there is a positive damping, rendering the steady state stable. There is an intuitive explanation for this fact: The oscillations are stable, because the NIC constantly works to increase the voltage drop across the oscillator. At high voltages (around $0.4\,\mathrm{V}$), this effect is balanced by the diodes, thus limiting the amplitude. The rate with which the NIC can build up a voltage difference is proportional to $1/R_{\mathrm{neg}}$. If only R_{int} were in place instead of the NIC, any voltage drop would decrease at a rate proportional to $1/R_{\mathrm{int}}$. Therefore, if $R_{\mathrm{neg}} > R_{\mathrm{int}}$, no substantial voltage drop develops, because the build-up rate through R_{neg} is smaller than the rate of consumption through R_{int}.

The dynamics can be separated into a term pertinent to the individual oscillator and a coupling term:

$$\dot{U}_i = \dot{U}_{i,\mathrm{individual}} + \frac{R_{\mathrm{ext}}(t)}{C_i R_{\mathrm{int}}(R_{\mathrm{int}} + N R_{\mathrm{ext}}(t))} \sum_{j=1}^{N} U_j$$

$$\dot{I}_i = \dot{I}_{i,\mathrm{individual}}$$

If the coupling term is small, this separation allows to derive an equation for the time evolution of the phases using the phase response approach discussed in section 2.4:

$$\dot{\vartheta}_i = \Omega_i + \frac{R_{\mathrm{ext}}(t)}{C_i R_{\mathrm{int}}(R_{\mathrm{int}} + N R_{\mathrm{ext}}(t))} \sum_{j=1}^{N} Z_i(t) U_j(t)$$

Here, $Z_i(t)$ is the phase response curve of the i-th oscillator and $U_j(t)$ is the waveform of the j-th oscillator. As Fig. 31 shows, the waveforms and phase response curves for the oscillators used in the experiments are close to

$$U_i(t) = U_{\mathrm{max}} \cdot \sin(\vartheta_i(t)), \quad Z_i(t) = Z_{\mathrm{max}} \cdot \cos(\vartheta_i(t)).$$

For the particular TD-type oscillator shown in Fig. 31a,b, the measured values were $U_{\mathrm{max}} = 0.20\,\mathrm{V}$ and $Z_{\mathrm{max}} = 5.0\,\mathrm{V}^{-1}$ For the NIC-type oscillator shown in Fig. 31c,d, the values were $U_{\mathrm{max}} = 0.40\,\mathrm{V}$ and $Z_{\mathrm{max}} = 2.5\,\mathrm{V}^{-1}$. Both oscillators are representative for all oscillators of the respective type used in the experiments.

For an almost harmonic oscillator, the expected value of the product of the amplitudes of waveform and phase response curves is $U_{\mathrm{max}} Z_{\mathrm{max}} = 1$. For both types of oscillator, this is fulfilled (within the error margins of the fitting algorithm). The dynamics of the phase is

therefore approximately given by

$$\dot{\vartheta}_i = \Omega_i + \frac{R_{\text{ext}}(t)}{C_i R_{\text{int}}(R_{\text{int}}/N + R_{\text{ext}}(t))} \frac{1}{N} \sum_{j=1}^{N} \cos \vartheta_i \sin \vartheta_j.$$

A comparison with (2.24) gives a relation between the coupling strength ε and the coupling function $a(t)$ in the phase model on the one hand and the values of $R_{\text{ext}}(t)$, R_{int}, C_i on the other hand:

$$a(t) = \frac{R_{\text{ext}}(t)}{\varepsilon C_i R_{\text{int}}(R_{\text{int}}/N + R_{\text{ext}}(t))}$$

To obtain the function $R_{\text{ext}}(t)$ that provides a certain coupling function $a(t)$, the expression above is solved for $R_{\text{ext}}(t)$

$$R_{\text{ext}}(t) = \frac{R_{\text{int}}}{N} \cdot \frac{a(t)}{(\varepsilon C_i R_{\text{int}})^{-1} - a(t)} \tag{5.2}$$

There are two problems with this relation of $R_{\text{ext}}(t)$ and $a(t)$. First, since $a(t)$ alternates between positive and negative values, $R_{\text{ext}}(t)$ must do the same to deliver the correct coupling function. This is not necessarily a huge problem, as an offset to the external resistance, such that $R_{\text{ext}}(t) > 0$ at all times, is tolerable if it is not too large, because its effect will cancel out on average. Second, and more gravely, if the denominator on the right hand side of (5.2) becomes very small or even changes its sign, $R_{\text{ext}}(t)$ will be hugely distorted compared to $a(t)$. For the Hebbian coupling

$$a(t) = \sum_{p=1}^{N} \sum_{q=1}^{N} w_{pq} \cos((\Omega_p - \Omega_q)t)$$

and any set of values for ε, C_i and R_{ext}, this is bound to happen at some point if the number of oscillators is increased, because the value range of $a(t)$ increases with N. The only way to avoid this is to ensure that $(\varepsilon C_i R_{\text{int}})^{-1} \gg a(t)$ which in turn limits the amplitude of R_{ext}.

5.2.2 Improved coupling through an external resistor

To avoid the problems discussed above, the original coupling was modified by replacing the external resistance $R_{\text{ext}}(t)$ with an external impedance $Z_{\text{ext}}(t)$ (see Fig. 36b). A negative impedance Z_{ser} was included in series to a variable resistor $R_{\text{var}}(t)$ to allow for positive as well as negative values of $Z_{\text{ext}}(t)$. The parallel negative impedance $Z_{\text{par}} = -R_{\text{int}}/N$ was chosen with the goal that the nonlinear dependence of $R_{\text{ext}}(t)$ from $a(t)$ is changed into a simple proportionality. With the new choice of the external impedance, the external
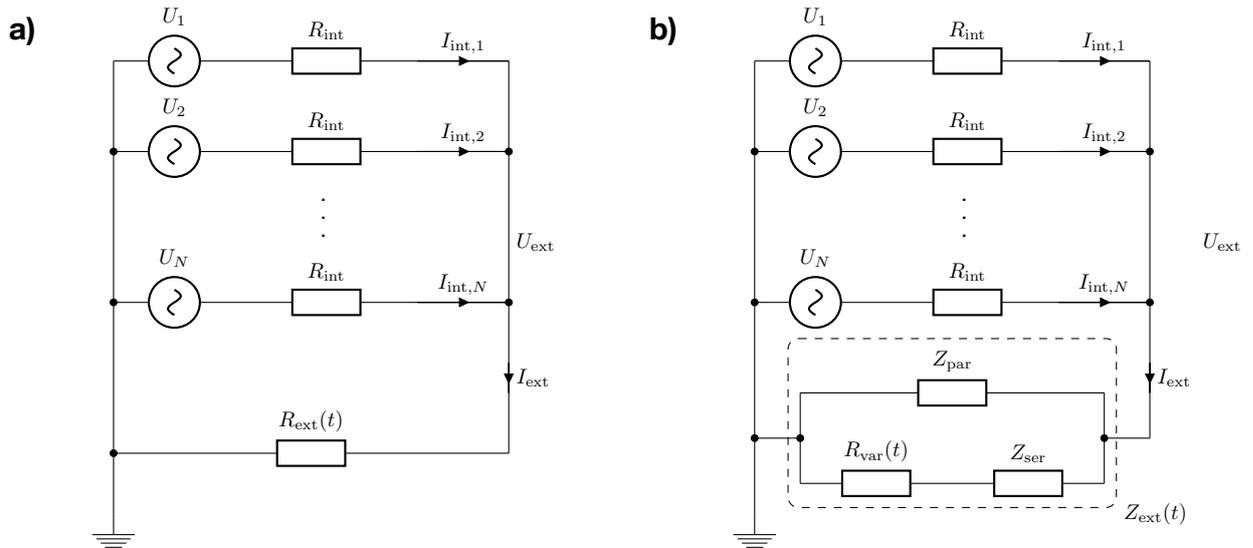
Fig. 36: a) Circuit of coupled oscillators with a straightforward implementation of a time-dependent global coupling provided by $R_{ext}(t)$. The relation between the coupling strength ε and the value of $R_{ext}(t)$ is defined by (5.2). **b)** Circuit of coupled oscillators with a time-dependent coupling proportional to $R_{var}(t) + Z_{ser}$, as defined by (5.4).

potential is given by

$$U_{ext} = \frac{Z_{ser} + R_{var}(t)}{R_{int}} \sum_{j=1}^{N} U_j \qquad (5.3)$$

Again, the single oscillator dynamics can be broken down in an individual and a coupling term:

$$\dot{U}_i = \dot{U}_{i,individual} + \frac{R_{ext}(t)}{C_i R_{int}(R_{int} + N R_{ext}(t))} \sum_{j=1}^{N} U_j$$

$$\dot{I}_i = \dot{I}_{i,individual}$$

The phase equations for weak coupling are now given by

$$\dot{\vartheta}_i = \Omega_i + \frac{N(R_{var}(t) + Z_{ser})}{C_i R_{int}^2} \frac{1}{N} \sum_{j=1}^{N} \cos \vartheta_i \sin \vartheta_j.$$

Once again, a sinusoidal shape of both waveform and phase response curve was assumed. Now, the relation between $R_{var}(t)$ and $a(t)$ is simply

$$R_{var}(t) = -Z_{ser} + \frac{\varepsilon C_i R_{int}^2}{N} \cdot a(t) \qquad (5.4)$$

This relation between $R_{var}(t)$ and $a(t)$ has a number of advantages over the original design (5.2). First of all, by making Z_{ser} sufficiently smaller than zero, the variable resistance can be kept at positive values, which allows for the application of a simple VCR as presented in 3.2.1. Second, there is no distortion any more in the shape of the coupling function
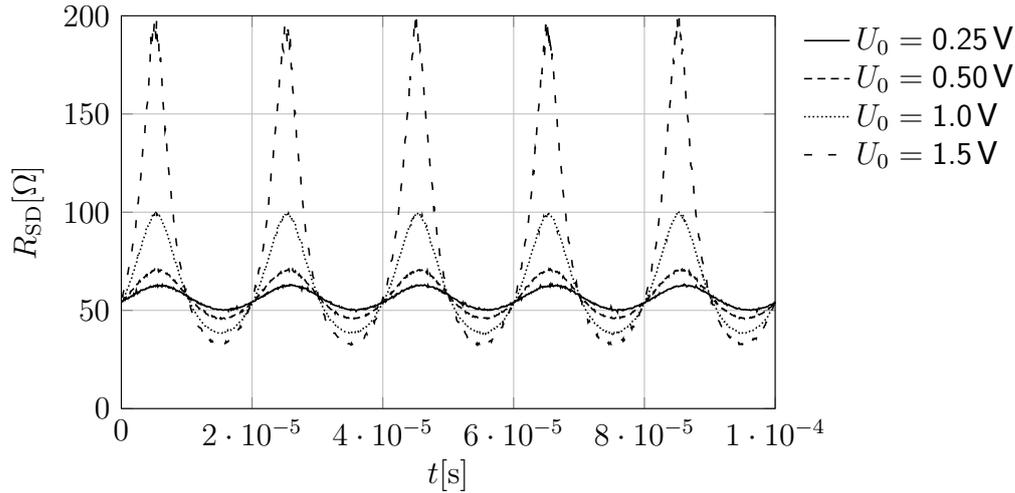
Fig. 37: Source-drain resistance of a VCR2N JFET for a time-dependent sinusoidal gate source voltage $U_{\mathrm{GS}} = -1.5\,\mathrm{V} - U_0 \sin(2\pi f t)$ with $f = 50\,\mathrm{kHz}$. The variable resistance of the JFET was connected in series to a $50\,\Omega$ resistor with 0.1% tolerance. A voltage of $0.5\,\mathrm{V}$ was applied across both resistances. The voltage drops U_{SD} and $U_{50\,\Omega}$ across both resistors where measured with a Waverunner44MXi oscilloscope. From these voltages, the source drain resistance was computed as $R_{\mathrm{SD}} = 50\,\Omega \cdot U_{\mathrm{SD}}/U_{50\,\Omega}$, using the math capabilities of the oscilloscope. The output curves are shown in the plot.

implemented as a variable resistance. Both $a(t)$ and $R_{\mathrm{var}}(t)$ are discrete Fourier sums with the same limited number of frequency components. Such a function is in general easier to provide than the complex nonlinear function required by 5.2. Last but not least, there is now a direct way to relate the variable external resistance to the coupling strength ε of the phase model. Assume that the variable resistance is given by

$$R_{\mathrm{var}}(t) = R_{\mathrm{coup}}^{\mathrm{off}} + \frac{R_{\mathrm{coup}}^{\mathrm{amp}}}{N} \cdot a(t),$$

where $R_{\mathrm{coup}}^{\mathrm{amp}}/N$ is the amplitude of one single frequency component in $R_{\mathrm{var}}(t)$ and $R_{\mathrm{coup}}^{\mathrm{off}}$ is a positive offset. Then, if Z_{ser} is chosen such that $Z_{\mathrm{ser}} = -R_{\mathrm{coup}}^{\mathrm{off}}$, the coupling strength is given by

$$\varepsilon = \frac{R_{\mathrm{coup}}^{\mathrm{amp}}}{C_i R_{\mathrm{int}}^2} \tag{5.5}$$

As a consequence, to ascertain equal coupling coefficients across all oscillators, it is mandatory that all capacitances C_i be the same. If not, oscillators with larger capacitances will feel a weaker influence of the coupling. Another consequence is the quadratic influence of the internal resistances R_{int} on the coupling strength. This means that R_{int} is not an entirely free parameter when it comes to the selection of the individual oscillator, even if $R_{\mathrm{int}} > R_{\mathrm{neg}}$, because its value is tied to the value of the coupling strength.

Although there is now a linear relationship between the coupling strength ε and the coupling resistance R_{coup}, the coupling function $a(t)$ provided by a variable resistance might still not be free from distortions, depending on the implementation. If, like it was done for this

thesis, a VCR2N-type voltage controlled resistor is used with a variable gate source voltage U_{GS}, a sinusoidal shape of the voltage signal does not necessarily translate into a sinusoidal variation of the source drain resistance R_{SD}, which was used as $R_{\mathrm{var}}(t)$ in the experiments. Fig. 37 illustrates this point, showing the behavior of R_{SD} for a sinusoidal gate source voltage of frequency $50\,\mathrm{kHz}$. This nonlinearity could be compensated by appropriately attenuating the voltage signal for larger amplitudes. Alternatively, the VCR2N could be replaced by a more sophisticated VCR with a linear $R - U$ characteristic (see for example Senani, et al. (2009)). Both approaches, while promising to yield better results, arguably take something away from the simplicity and elegance of the coupling mechanism.

5.2.3 Coupling through analog computing devices

A variable resistance is not the only possible source of a time-dependent coupling. Since the time evolution of an oscillator's voltage is given by

$$\dot{U}_i = \dot{U}_{i,\mathrm{individual}} + \frac{U_{\mathrm{ext}}}{C_i R_{\mathrm{int}}},$$

any coupling circuitry producing an external voltage $U_{\mathrm{ext}}(t) = c \cdot a(t)/N \sum_i U_i$ with a sufficiently small proportionality constant c will produce the correct phase shift dynamics, with a coupling strength of

$$\varepsilon = \frac{c}{C_i R_{\mathrm{int}}} \tag{5.6}$$

In particular, it is possible to explicitly sum over the voltages of all oscillators employing a voltage adder, multiply the resulting signal with an external coupling signal containing the correct frequency components and feed back the product to each oscillator.

5.3   Complete Circuit

Two circuit boards with 8 coupled oscillators each were built, one using TD-type oscillators, one using NIC-type oscillators. Fig. 38 and Fig. 39 show photographs of the TD-type and the NIC-type network respectively. For the circuit board with TD-type oscillators, only the improved coupling through a VCR, as described in 5.3.2 was implemented; for the circuit board with NIC-type oscillators, additionally also the coupling mechanism using analog computing devices was built. In the first part of this section, the specific parameter values of the individual oscillators used in the experiments are given. Then, the detailed circuit schematics for both coupling mechanisms are discussed.

5.3.1   Parameters of the individual oscillators

Oscillator ID	TD-type		NIC-type
	C	f	f
1	660 nF	29 376 Hz$\pm$7 Hz	35 048 Hz$\pm$3 Hz
2	517 nF	35 620 Hz$\pm$80 Hz	36 071 Hz$\pm$4 Hz
3	330 nF	40 625 Hz$\pm$20 Hz	39 126 Hz$\pm$5 Hz
4	220 nF	55 615 Hz$\pm$30 Hz	44 093 Hz$\pm$3 Hz
5	147 nF	61 374 Hz$\pm$8 Hz	50 158 Hz$\pm$3 Hz
6	132 nF	66 092 Hz$\pm$8 Hz	57 115 Hz$\pm$2 Hz
7	122 nF	67 309 Hz$\pm$9 Hz	67 022 Hz$\pm$20 Hz
8	100 nF	73 800 Hz$\pm$7 Hz	69 037 Hz$\pm$15 Hz

Table 5.1: Characterization of the 8 oscillators used for both network types. The second column shows the nominal value of the capacitance C_i of each experimental TD-type oscillator. The third column shows the value range over which the frequency of the same, uncoupled oscillator varied during a series of pattern recognition experiments that took roughly 5 min (the results of this series are shown in Fig. 45 in the next chapter). The fourth column shows the value range over which the frequency of each uncoupled NIC-type oscillator varied during a series of pattern recognition experiments that took roughly 15 min (the results of this series are shown in Fig. 54 in the next chapter).

The parameter range of the TD-type oscillators was $\gamma = 2.1 \cdot 10^{-3}$ A^2/V^2 ... $1.4 \cdot 10^{-2}$ A^2/V^2 and $f_0 = 29$ kHz ... 73 kHz; each oscillator had an inductance $L = 47\,\mu$H (with 10% tolerance), while the values of the capacitance C_i, $i = 1 \ldots 8$ differed between oscillators depending on the frequency. Both the values C_i (with 20% tolerance) and the resulting actual oscillation frequencies as they were measured in one particular series of pattern recognition experiments are given in table 5.1. The range of experimental parameters is also visualized in Fig. 40.

In the network of NIC-type oscillators, each oscillator had a negative differential resistance $R_{\mathrm{neg}} = 6.19$ kΩ (the resistors had a tolerance of 1%) and a capacitance $C = 1$ nF (again using capacitors with 20% tolerance); for each oscillator, the frequency was tuned by adjusting the
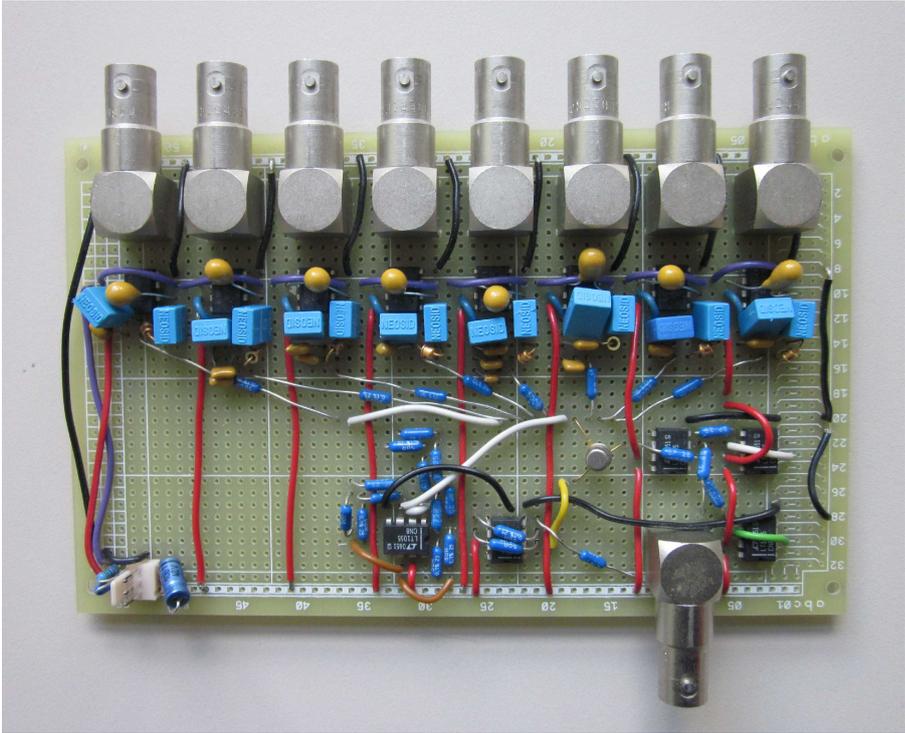
Fig. 38: Photograph of the circuit board for a network of TD-type oscillators used in the experiments described in chapter 6.
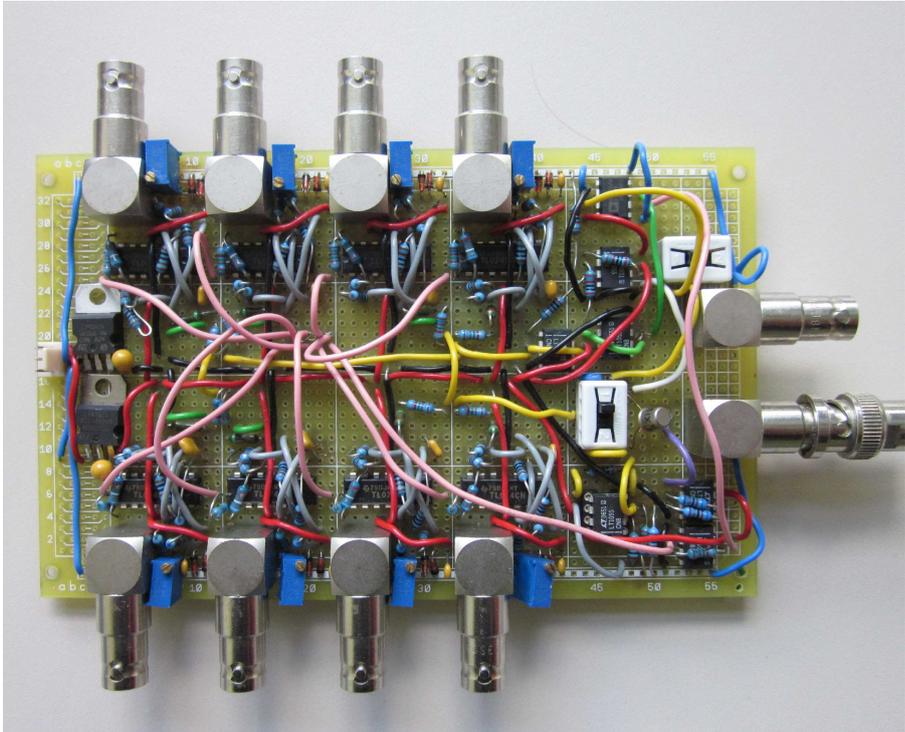
Fig. 39: Photograph of the circuit board for a network of NIC-type oscillators used in the experiments described in chapter 6.
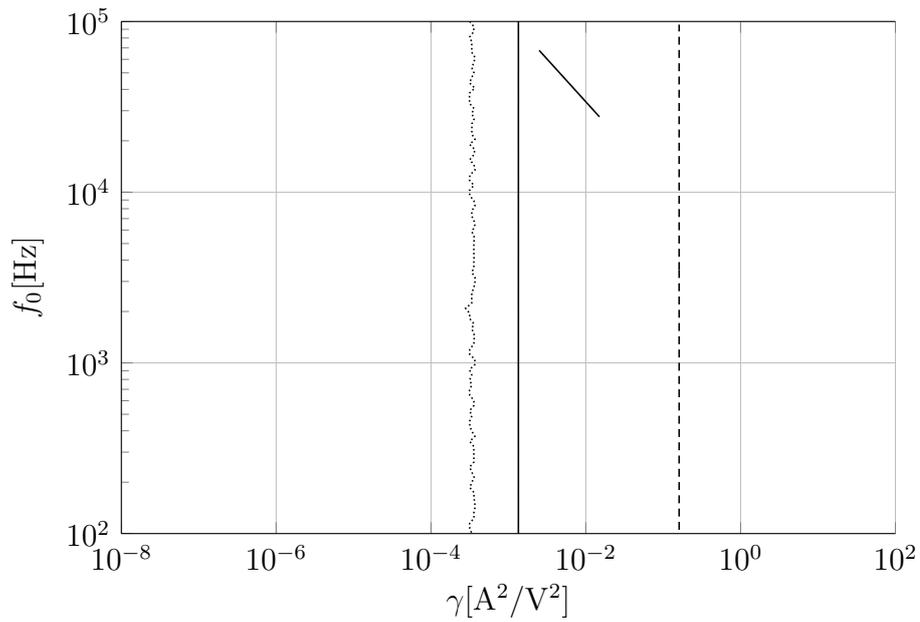
Fig. 40: Parameter space of the TD-type oscillators that were used in the experiments (i.e. with internal resistance $R_{\text{int}} = 500\,\Omega$). The figure is analog to Fig. 34, but additionally, the isolated line segment shows where the eight experimental TD-type oscillators were located in parameter space, namely on the line at $L = 47\,\mu\text{H}$ between frequencies of $f_0 = 29\,\text{kHz}$ and $f_0 = 73\,\text{kHz}$.
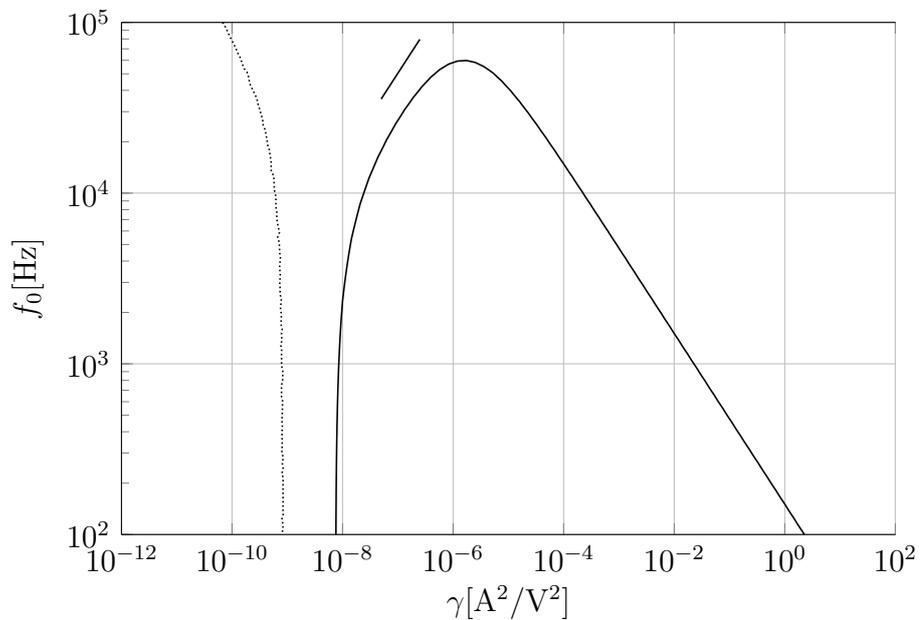
Fig. 41: f_0-γ-plane in the parameter space of the NIC-type oscillators that were used in the experiments (i.e. with negative differential resistance $R_{\text{neg}} = 6.19\,\text{k}\Omega$ and internal resistance $R_{\text{int}} = 8.25\,\text{k}\Omega$). The figure is analog to Fig. 35, but only for one specific set of parameters. The isolated line segment shows where the eight experimental NIC-type oscillators were located in parameter space, namely on the line at $C = 1\,\text{nF}$ between frequencies of $f = 35\,\text{kHz}$ and $f = 69\,\text{kHz}$.

potentiometer in the respective gyrator (and thereby adjusting the inductance L_i). The range of actual oscillation frequencies measured in the experiment was $f = 35\,\text{kHz}\ldots 69\,\text{kHz}$. The individual frequencies measured in one particular series of pattern recognition experiments are given in table 5.1. Since the actual values of the L_i were not measured, and since the measured frequency f was already somewhat below f_0 for the NIC-type oscillators, the exact range of γ and f_0 is not given here. However, the approximate range of experimental parameters is visualized in Fig. 41. Note that this range does not actually lie inside the region where $f/f_0 > 0.95$. However, this was still close enough to conveniently tune the frequencies over a wide range.

There are a few comments to be made concerning the data in table 5.1. The frequency range was selected with the condition (2.26) in mind, i.e. the minimum frequency was larger than a third of the maximum frequency. The order of magnitude of the frequencies was a compromise between the execution speed of the pattern recognition process on the one hand and the avoidance of any deteriorating non-ideal behavior of the TL071 and LT1055 op-amps at higher frequencies on the other hand. For example, the gain factor A of the LT1055, according to its datasheet, decreases from $A \approx 5 \times 10^5$ (or $115\,\text{dB}$) at zero frequency to $A \approx 1$ (or $0\,\text{dB}$) at $4\,\text{MHz}$. At $70\,\text{kHz}$, the gain is $A \approx 100$ (or $40\,\text{dB}$), which still seems sufficient for the LT1055 to operate like an ideal op-amp.

For both networks, the goal was to realize a Golomb ruler with as large a minimal frequency difference δ as possible. In the case of the NIC-type oscillators, this worked reasonably well. If only the first two digits of each frequency are considered, an optimal Golomb ruler with 8 marks and a minimal frequency difference of $\delta = \delta' = 2\pi \cdot 1\,\text{kHz}$ ensues (note that in the case of imperfect Golomb rulers, where the frequency differences are not multiples of a common step size, δ' as defined in equation (2.10) is the relevant fast time scale, rather than δ). Although over very long times, the oscillators tended to drift by as much as several $100\,\text{Hz}$, it was always possible to readjust the frequencies such that $\delta' \geq 2\pi \cdot 800\,\text{Hz}$. Unfortunately, the same cannot be said for the TD-type oscillators for which the frequency tuning was much more tedious. As a matter of fact, $\delta' \approx 2\pi \cdot 300\,\text{Hz}$ was the best value that could be achieved, and even that worked only for short times before the frequency drift (which was considerably faster than for NIC-type oscillators, as can be seen from table 5.1) reduced the value even more.

Another problem with the TD-type network was that all inductances had the same value, rather than all capacitances. In hindsight, this turned out to be a bad choice, because it meant that different oscillators experienced a different coupling strength.

The numerical values of the frequency deviations given in table 5.1 are not the relevant values for $\Delta\Omega$ in the sense of the discussion in 4.3, because in the experiments, the frequencies were determined anew for each experimental run, which means that the accuracy was higher (see also the first section in chapter 6.). They do, however, become

relevant if one wishes to use the same set of oscillators without continuously updating the frequency information, which puts even greater demands on the accuracy of the oscillators. Finally, note that the two circuit designs presented here still have room for improvement in terms of oscillator speed and oscillator frequency stability or flexibility and accuracy of the coupling mechanism; in the present form, they are straightforward implementations that do not make use of any sophisticated optimization techniques. For example, one could consider using more precise quartz oscillators instead of van der Pol oscillators. However, it is not entirely obvious how one should go on about this, because a standard quartz oscillator is an output-only black box that does not allow for the feedback needed to implement a two-way coupling. This is the equivalent of a zero phase response. Breaking up the black box to implement a feedback must be done with caution to not defeat the original purpose of frequency stability.

5.3.2   Circuit with improved coupling through an external resistor

The schematic of the circuit is shown in Fig. 42. The nominal values for R_{int}, R_{par} and R were $R_{\mathrm{int}} = 500\,\Omega$, $R_{\mathrm{par}} = 62\,\Omega$ and $R = 62\,\Omega$ for the TD-type network and $R_{\mathrm{int}} = 8.25\,\mathrm{k\Omega}$, $R_{\mathrm{par}} = 1\,\mathrm{k\Omega}$ and $R = 1\,\mathrm{k\Omega}$ for the NIC-type network. The tolerance of these resistances was 1%. Note that the values for the parallel negative impedance $Z_{\mathrm{par}} = -R_{\mathrm{par}}$ are not exactly equal to $-R_{\mathrm{int}}/8$; in particular, in the case of the NIC-type network, there is a deviation of 3% from the correct value. The distortion introduced by this deviation, however, is much smaller than the distortions introduced by the nonlinearity of the VCR2N.
$U_{\mathrm{DA}}(t)$, as already mentioned in the methods chapter, is a voltage signal given by

$$U_{\mathrm{DA}}(t) = U_{\mathrm{coup}}^{\mathrm{off}} + \frac{U_{\mathrm{coup}}^{\mathrm{amp}}}{N} \sum_{p=1}^{8} \sum_{q \neq p} w_{pq} \cos((\Omega_p - \Omega_q)t)$$

where the low cutoff value was fixed to $U_{\mathrm{DA}}^{\mathrm{min}} = 0\,\mathrm{V}$.
To implement the variable resistance $R_{\mathrm{var}}(t)$, a VCR2N transistor was used. To provide a potential difference equal to $-U_{\mathrm{DA}}(t)$ between gate and source of the VCR, a voltage follower and a differential amplifier were applied. The voltage follower is necessary because the differential amplifier draws a certain current which does not pass through the external impedance Z_{ext}. A buffer is needed to provide this current; otherwise the differential equations describing the system have to be modified. The differential amplifier subtracts $U_{\mathrm{DA}}(t)$ from the potential at the source of the VCR. Therefore, as $U_{\mathrm{DA}}(t)$ varies between $0\,\mathrm{V}$ and $3\,\mathrm{V}$ (which is the highest possible cutoff value $U_{\mathrm{DA}}^{\mathrm{max}}$ for the setup) , $R_{\mathrm{SD}}(t) = R_{\mathrm{var}}(t)$ varies roughly between $30\,\Omega$ and $200\,\Omega$ (see Fig. 37).
 For very small amplitudes of the coupling voltage, the relation between $U_{\mathrm{DA}}(t)$ and $R_{\mathrm{var}}(t)$

Fig. 42: Design of the network circuit with a VCR2N JFET voltage controlled resistor as the basis for the time-dependent coupling.

is almost linear, and therefore

$$R_{\mathrm{var}}(t) \approx R_{\mathrm{coup}}^{\mathrm{off}}(U_{\mathrm{coup}}^{\mathrm{off}}) + \frac{R_{\mathrm{coup}}^{\mathrm{amp}}}{N} \sum_{p=1}^{8} \sum_{q \neq p} w_{pq} \cos((\Omega_p - \Omega_q)t).$$

Based on an offset value of $U_{\mathrm{coup}}^{\mathrm{off}} = 1.5\,\mathrm{V}$ in $U_{\mathrm{DA}}(t)$, which results in an offset in $R_{\mathrm{var}}(t)$ close to $60\,\Omega$, as can be seen in Fig. 37, the serial negative impedance was chosen as $Z_{\mathrm{ser}} = -62\,\Omega$.

In the low amplitude limit, it is possible to relate the amplitude of a single component in the coupling signal to the coupling strength in the phase model using equation (5.5):

$$\varepsilon = \frac{20\,\Omega\,\mathrm{V}^{-1} \cdot U_{\mathrm{coup}}^{\mathrm{amp}}}{C_i R_{\mathrm{int}}^2}$$

Here, it was assumed (using Fig. 37) that at low amplitudes a variation of $1\,\mathrm{V}$ in $U_{\mathrm{DA}}(t)$ causes a variation of $20\,\Omega$ in $R_{\mathrm{var}}(t)$. In the NIC-type network, where all oscillators have the same capacitance $C_i = C = 1\,\mathrm{nF}$, this amounts to

$$\varepsilon \approx 3 \cdot 10^2\,\mathrm{V}^{-1}\,\mathrm{s}^{-1} \cdot U_{\mathrm{coup}}^{\mathrm{amp}}.$$

Fig. 43: Design of the network circuit with an inverting adder and an AD633 multiplier as the basis for the time-dependent coupling. The AD633 has 5 inputs, only two of which were used in the experiment to provide the desired output voltage. The rest of the inputs were connected to ground.

In the TD-type network, with capacitance values C_i ranging from $100\,\text{nF}$ to $660\,\text{nF}$, the effective coupling strength depends on the oscillator and is given by

$$\varepsilon \approx 1 \cdot 10^2\,\text{V}^{-1}\,\text{s}^{-1} \ldots 8 \cdot 10^2\,\text{V}^{-1}\,\text{s}^{-1} \cdot U_{\text{coup}}^{\text{amp}}.$$

5.3.3   Circuit with coupling through analog computing devices

This circuit, shown in Fig. 43 was designed in a straightforward way to feed back the product of the sum of all individual voltage signals U_i and the coupling signal $U_{\text{DA}}(t)$ to each oscillator. To achieve this, the U_i, each buffered by a voltage follower, are fed into an inverting adder. The output of the inverting adder is then once again inverted, multiplied with $U_{\text{DA}}(t)$, and divided by 10 in a AD633 multiplier. If the voltage signal

$$U_{\text{DA}}(t) = \frac{U_{\text{coup}}^{\text{amp}}}{N} \sum_{p=1}^{8} \sum_{q \neq p} w_{pq} \cos((\Omega_p - \Omega_q)t)$$

is applied (note that in this configuration, $U_{\text{coup}}^{\text{off}} = 0$), this amounts to the coupling strength (using equation 5.6)

$$\varepsilon = \frac{U_{\text{coup}}^{\text{amp}}}{10 \cdot 1\,\text{nF} \cdot 8.25\,\text{k}\Omega} \approx 1.2 \cdot 10^4\,\text{V}^{-1}\,\text{s}^{-1} \cdot U_{\text{coup}}^{\text{amp}}.$$

5.4 A physical perspective: Pattern recognition as the result of minimal entropy production

There is an interesting physical interpretation of the long term behavior of the electrical circuit, other than the fact that it is the long term solution of the differential equation it was designed to mirror: The system prefers states in which the average combined entropy production of all oscillators is minimal, in accordance with the minimum entropy production principle (Prigogine 1961).

Ideally, in the weakly coupled limit, each of the oscillators in Fig. 36b is a sinusoidal voltage source with a freely adjustable phase shift: $U_i = U_0 \sin \vartheta_i = U_0 \sin(\Omega_i t + \varphi_i)$. The power dissipated in the internal resistors R_{int} is then given by

$$P_{\text{int}} = \sum_{i=1}^{N}(U_0 \sin \vartheta_i - U_{\text{ext}})I_{\text{int},i} = U_0 \sum_{i=1}^{N} \sin \vartheta_i I_{\text{int},i} - U_{\text{ext}}I_{\text{ext}}$$

If Z_{ext} happens to be negative, part of this power is supplied by the negative impedance. On the other hand, if Z_{ext} is positive, it draws power from the oscillators. The power provided or consumed by the external impedance is given by

$$P_{\text{ext}} = U_{\text{ext}}I_{\text{ext}},$$

where a negative sign of P_{ext} means that the external impedance acts as a power supply. The overall power output of all oscillators is therefore

$$P = P_{\text{int}} + P_{\text{ext}} = U_0 \sum_{i=1}^{N} \sin \vartheta_i I_{\text{int},i} = \frac{U_0}{R_{\text{int}}} \sum_{i=1}^{N} \sin \vartheta_i (U_0 \sin \vartheta_i - U_{\text{ext}}).$$

In this equation, U_{ext} can be replaced by its functional dependence on the oscillator phases ϑ_i using (5.3):

$$
\begin{aligned}
P &= \frac{U_0^2}{R_{\text{int}}} \sum_{i=1}^{N} \sin \vartheta_i \left(\sin \vartheta_i - \frac{Z_{\text{ser}} + R_{\text{var}}(t)}{R_{\text{int}}} \sum_{j=1}^{N} \sin \vartheta_j \right) \\
&= \frac{U_0^2}{R_{\text{int}}} \left(\sum_{i=1}^{N} \sin^2 \vartheta_i - \frac{Z_{\text{ser}} + R_{\text{var}}(t)}{R_{\text{int}}} \sum_{i=1}^{N} \sum_{j=1}^{N} \sin \vartheta_i \sin \vartheta_j \right)
\end{aligned}
$$

Obviously, the momentary power output has a quite complicated dependence on time. However, a closer look on the average power output $\langle P \rangle$ associated with a certain distribution of phase shifts φ, given by

$$\langle P \rangle(\varphi) = \lim_{T \to \infty} \frac{1}{T} \int_0^T P(t, \varphi)\mathrm{d}t,$$

reveals that $\langle P \rangle(\varphi)$ is closely related to the potential function $E(\varphi)$ of the ideal phase equation discussed in section 2.2. Evaluating the integral over the first term and inserting the coupling resistance

$$R_{\mathrm{var}} = -Z_{\mathrm{ser}} + \frac{R_{\mathrm{coup}}}{N} \cdot a(t), \quad a(t) = \sum_{p=1}^{N} \sum_{q=1}^{N} w_{pq} \cos((\Omega_p - \Omega_q)t)$$

into the second term leads to

$$
\begin{aligned}
\langle P \rangle(\varphi) &= \frac{U_0^2}{R_{\mathrm{int}}} \left(\frac{N}{2} - \frac{R_{\mathrm{coup}}}{N R_{\mathrm{int}}} \lim_{T \to \infty} \frac{1}{T} \int_0^T a(t) \sum_{i=1}^{N} \sum_{j=1}^{N} \sin(\Omega_i t + \varphi_i) \sin(\Omega_j t + \varphi_j) \mathrm{d}t \right) \\
&= \frac{U_0^2}{R_{\mathrm{int}}} \left(\frac{N}{2} - \frac{R_{\mathrm{coup}}}{N R_{\mathrm{int}}} \lim_{T \to \infty} \frac{1}{T} \int_0^T \sum_{i,j,p,q} w_{pq} \cos((\Omega_p - \Omega_q)t) \sin(\Omega_i t + \varphi_i) \sin(\Omega_j t + \varphi_j) \mathrm{d}t \right).
\end{aligned}
$$

Consider a single term s_{ijpq} in the sum on the right hand side:

$$
\begin{aligned}
s_{ijpq} &= w_{pq} \cos((\Omega_p - \Omega_q)t) \sin(\Omega_i t + \varphi_i) \sin(\Omega_j t + \varphi_j) \\
&= \frac{w_{pq}}{2} \cos((\Omega_p - \Omega_q)t) \left(\cos((\Omega_i - \Omega_j)t + \varphi_i - \varphi_j) - \cos((\Omega_i + \Omega_j)t + \varphi_i - \varphi_j) \right)
\end{aligned}
$$

Due to the selection of frequencies, $\cos((\Omega_p - \Omega_q)t) \cos((\Omega_i + \Omega_j)t + \varphi_i - \varphi_j)$ is an oscillatory term, regardless of the values of i, j, p, and q and therefore, only the first term in parentheses may contribute to the average $\langle s_{ijpq} \rangle$ at all:

$$
\begin{aligned}
\langle s_{ijpq} \rangle &= \lim_{T \to \infty} \frac{1}{T} \int_0^T \frac{w_{pq}}{4} \left(\cos(\Omega_{qi}^{pj} t - \varphi_i + \varphi_j) + \cos(\Omega_{qj}^{pi} t + \varphi_i - \varphi_j) \right) \mathrm{d}t \\
&= \frac{w_{pq}}{4} \left(\delta_{pq} \delta_{ji} \cos(\varphi_j - \varphi_i) + \delta_{pi} \delta_{jq} \cos(\varphi_j - \varphi_i) + \delta_{pq} \delta_{ij} \cos(\varphi_i - \varphi_j) + \delta_{pj} \delta_{iq} \cos(\varphi_i - \varphi_j) \right)
\end{aligned}
$$

with the shorthand notation $\Omega_{qi}^{pj} = \Omega_p + \Omega_j - \Omega_q - \Omega_i$. Only if $\Omega_{qi}^{pj} = 0$ or $\Omega_{qj}^{pi} = 0$, $\langle s_{ijpq} \rangle$ is different from zero. This is the case exactly when the two lower indices are a permutation of the two upper indices.

Inserting $\langle s_{ijpq} \rangle$ into the equation for $\langle P \rangle(\varphi)$ and simplifying the sums over i, j, p and q gives

$$
\begin{aligned}
\langle P \rangle(\varphi) &= \frac{U_0^2}{R_{\mathrm{int}}} \left(\frac{N}{2} - \frac{R_{\mathrm{coup}}}{R_{\mathrm{int}}} \left(\frac{N}{2} + \frac{1}{2N} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij} \cos(\varphi_j - \varphi_i) \right) \right) \\
&= \frac{U_0^2}{R_{\mathrm{int}}} \left(\frac{N}{2} - \frac{R_{\mathrm{coup}}}{R_{\mathrm{int}}} \left(\frac{N}{2} - E(\varphi) \right) \right),
\end{aligned}
$$

where $E(\varphi)$ is the potential function of the ideal phase coupled network from 2.2. Another way to write this is

$$\langle P \rangle(\varphi) = \left(1 - \frac{R_{\mathrm{coup}}}{R_{\mathrm{int}}} \right) \langle P_0 \rangle + \frac{U_0^2 R_{\mathrm{coup}}}{R_{\mathrm{int}}^2} E(\varphi),$$

where $\langle P_0 \rangle$ is the average power output of the uncoupled set of oscillators (i.e for $R_{\mathrm{coup}} = 0$). Since the combined entropy production $\mathrm{d}S/\mathrm{d}t$ of all oscillators is related to the dissipated power via

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{P}{T},$$

where T is the ambient temperature, the average entropy production $\langle \mathrm{d}S/\mathrm{d}t \rangle$ is minimized for a minimum of the potential E, which is realized in the case of pattern recognition. The above reasoning can be applied any kind of network of van der Pol oscillators to predict and explain synchronization phenomena (see for example Matsumura, et al. (2012)).

6. Experimental results

In this chapter, results of experiments with the three different circuits presented in the design chapter are discussed, namely with

- a network of TD-Type oscillators with a VCR2N providing the coupling function as described in 5.3.2,

- a network of NIC-Type oscillators with a VCR2N providing the coupling function as described in 5.3.2 and

- a network of NIC-Type oscillators with an AD633 analog multiplier providing the coupling function as described in 5.3.3.

This order reflects the order in which the different designs were built, examined and improved during the thesis. Therefore, pattern recognition works best with the third type of network. The results of the pattern recognition experiments are discussed in the first section of the chapter.

The second section shows some examples of the long term behavior of the network during the recognition step, and compares them to the theoretical and numerical findings of section 4.1.

Finally, the third section deals with an interesting effect of a coupling of the form

$$a(t) = \sum_{p=1}^{N} \sum_{q=1}^{N} \cos((\Omega_p - \Omega_q)t)$$

(i.e. the coupling used for the initialization of the network to any pattern, after a suitable coordinate transform). This type of coupling makes oscillations more regular compared to the oscillations of an uncoupled set of oscillators.

6.1   Pattern recognition in a network with 8 oscillators

All experiments discussed below were conducted in the fashion described in detail in section 3.3.2. In brief, the experimental routine consisted of the following steps:

- The voltage signals of the oscillators were measured during a time interval of T_{record}, without any time-dependent coupling (i.e. $U_{\mathrm{DA}}(t) = 0$).

- The average frequencies Ω_i during the second half of this time interval were determined.

- The oscillators' voltage signals $U_i(t)$ were measured during a time interval of T_{record} again, this time in the presence of the time-dependent coupling signal

$$U_{\mathrm{DA}}(t) = U_{\mathrm{coup}}^{\mathrm{off}} + \frac{U_{\mathrm{coup}}^{\mathrm{amp}}}{N} \sum_{p=1}^{8} \sum_{q \neq p} \xi_i \xi_j \cos((\Omega_p - \Omega_q)t)$$

during the initialization in the first half of the interval and

$$U_{\mathrm{DA}}(t) = U_{\mathrm{coup}}^{\mathrm{off}} + \frac{U_{\mathrm{coup}}^{\mathrm{amp}}}{N} \sum_{p=1}^{8} \sum_{q \neq p} \sum_{k=1}^{3} \xi_i^k \xi_j^k \cos((\Omega_p - \Omega_q)t)$$

during the recognition in the second half of the interval. Here, $\boldsymbol{\xi}$ was the initialized, tested pattern, and the $\boldsymbol{\xi}^k$ were the memorized patterns. These patterns were chosen randomly for each measurement, but always such that the $\boldsymbol{\xi}^k$ were mutually orthogonal and $\boldsymbol{\xi}$ differed from $\boldsymbol{\xi}^1$ in exactly one bit. Whenever the value of $U_{\mathrm{DA}}(t)$ left a certain interval $[U_{\mathrm{DA}}^{\mathrm{min}}, U_{\mathrm{DA}}^{\mathrm{max}}]$, the signal was cut off and fixed at the boundary value, because of the limitations of the DA converter. By the choice of $U_{\mathrm{coup}}^{\mathrm{amp}}$, the coupling strength ε in the phase model was fixed as well. The relation between $U_{\mathrm{coup}}^{\mathrm{amp}}$ and ε depends on the implementation of the coupling circuitry and is explained in detail in the previous chapter (see 5.3).

- From the voltages $U_i(t)$, the time evolution of the phase shifts $\varphi_i(t)$ was determined. This was not done with respect to the frequencies Ω_i of the uncoupled oscillators, but with respect to the average frequencies Ω_i' the oscillators settled for during the second half of initialization. As long as the separation of time scales is large, this frequency change is an effect on a slower timescale than the evolution of the phase shifts due to pattern recognition (see section 4.2.5) and the results are not altered significantly.

- From the changes in phase shift $\Delta\varphi_i(t) = \varphi_i(t + T_{\mathrm{record}}/2) - \varphi_i(T_{\mathrm{record}}/2)$ during the recognition phase, the evolution of the proximity $P^1(t)$ of the network to the correctly recognized pattern $\boldsymbol{\xi}^1$ during the recognition phase was computed.

In the analysis given in section 4.2.1, it was discussed how a large uncertainty in the oscillator frequencies may influence the pattern recognition performance of the network negatively. Therefore, it is necessary to quantify the frequency fluctuations in the experiments. Since the frequencies were determined as a part of each single run of the experiment, the drifts given in table 5.1 are not the relevant values. A better measure is the difference in frequency $\Delta\Omega_{n+1} - \Delta\Omega_n$ that arises in an oscillator between two consecutive runs of an experiment (which were only seconds apart). Here, the index denotes the running number of the experiment. The time it takes to perform one pattern recognition run is

a)

b)
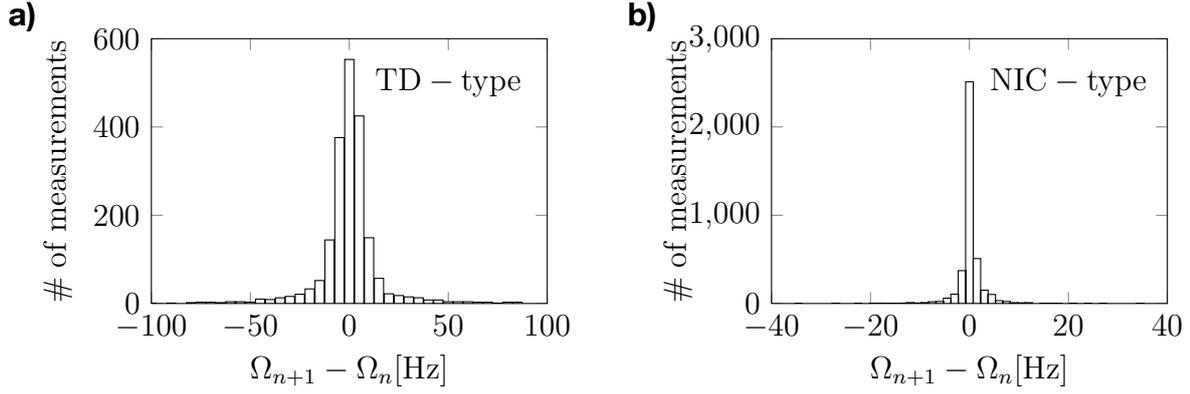
Fig. 44: a) Histogram of the differences in Ω measured for the uncoupled oscillators between two consecutive pattern recognition experiments for the series of 250 experiments with the TD-type network depicted in Fig. 45. **b)** Histogram of the differences in Ω measured for the uncoupled oscillators between two consecutive pattern recognition experiments for the series of 500 experiments with the NIC-type network depicted in Fig. 54.

also the time scale on which the relevant fluctuations in frequency occur that cannot be detected by the experimental setup. Fig. 44a shows the statistics of these fluctuations for the series of 250 pattern recognition experiments with the TD-type network shown further below in Fig. 45. Fig. 44b shows the statistics of the fluctuations for the series of 500 pattern recognition experiments with the NIC-type network also shown further below in Fig. 54. For the TD-type oscillators, most of the fluctuations lie in the window $\pm30\,$Hz, with a few single deviation values going as high as $\pm100\,$Hz. These strong fluctuations all occurred for the especially noisy oscillator around a frequency of $35.6\,$kHz. In the case of the NIC-type oscillators, fluctuations stay mainly within $\pm10\,$Hz, while the flanks of the distribution reach out to $\pm35\,$Hz. In the NIC-type network, there was no oscillator that was particularly noisier than the others. In the following, based on the distributions depicted in Fig. 44, the values $\Delta\Omega_{\mathrm{TD}} = 100\,$Hz and $\Delta\Omega_{\mathrm{NIC}} = 35\,$Hz were assumed for the size of the frequency windows for the two types of oscillators.

6.1.1   Network of TD-type oscillators with coupling through a VCR

The best success rate for pattern recognition that could be reached in a TD-type network was approximately 80%. Fig. 45 gives an overview over the results of the series of 250 experiments in which this rate was achieved. The measuring time interval of an individual experiment was $T_{\mathrm{record}} = 0.2\,$s, the amplitude of a single component in the coupling signal $U_{\mathrm{DA}}(t)$ was $U_{\mathrm{coup}}^{\mathrm{amp}} = 1500\,$mV and the offset was $U_{\mathrm{coup}}^{\mathrm{off}} = 1250\,$mV. The resulting coupling strength acting on the oscillators was therefore $\varepsilon \approx 1.5 \cdot 10^2\,\mathrm{s}^{-1} \dots 1.2 \cdot 10^3\,\mathrm{s}^{-1}$. Since the network does not settle for a steady state in the presence of noise (see theory chapter), pattern recognition, as expected, was only transient. The time of of the maximum success rate (i.e. the time at which the maximum number of experiments showed $P^1 > 0$) was $0.04\,$s after the coupling was switched from initialization to recognition mode.
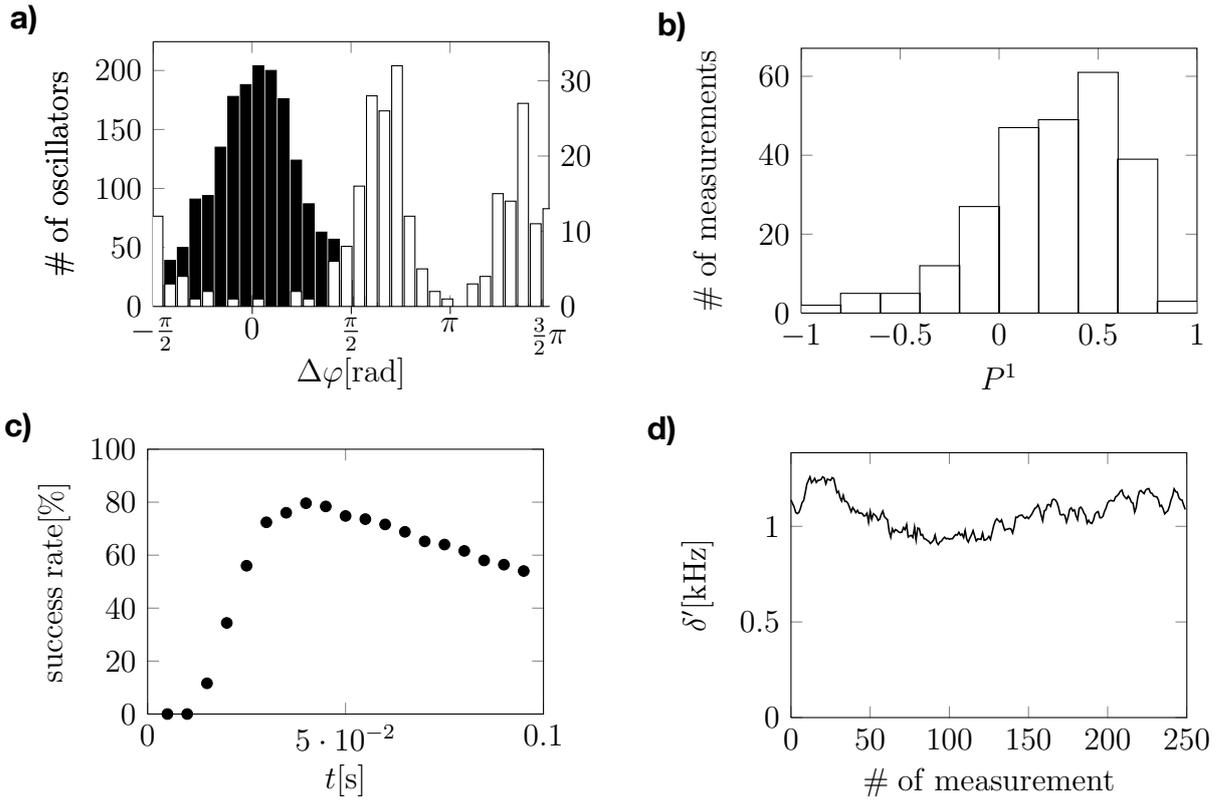
Fig. 45: Results of a series of 250 pattern recognition experiments with the circuit depicted in Fig. 42, using an array of TD-type oscillators. The parameters of the experiment were $T_{\mathrm{record}} = 0.2\,\mathrm{s}$, $U_{\mathrm{coup}}^{\mathrm{amp}} = 1500\,\mathrm{mV}$, $U_{\mathrm{coup}}^{\mathrm{off}} = 1250\,\mathrm{mV}$, $U_{\mathrm{DA}}^{\mathrm{min}} = 0\,\mathrm{mV}$, $U_{\mathrm{DA}}^{\mathrm{max}} = 2500\,\mathrm{mV}$ and $U_{\mathrm{TD}} = 300\,\mathrm{mV}$. The maximum recognition rate was $199/250 \approx 80\%$ after $0.04\,\mathrm{s}$ in recognition mode. **a)** Distribution of $\Delta\varphi(0.04\,\mathrm{s})$ for oscillators that are supposed to keep their phase shift (black bars, left vertical axis) and oscillators belonging to the defective bit (white bars, right vertical axis). **b)** Distribution of $P^1(0.04\,\mathrm{s})$. **c)** Success rate as measured by P^1 over time in recognition mode. **d)** Evolution of the minimal difference of frequency differences δ' during the series of experiments. The bin size in a) and b) was 0.2.

Fig. 45a shows the distribution of the changes $\Delta\varphi(0.04\,\mathrm{s})$ the phase shifts underwent up to this time for oscillators that were supposed to keep their phase shift (black bars) and oscillators that were supposed to change their phase shift by π for a correct recognition of $\boldsymbol{\xi}^1$, respectively. For the former, the values of $\Delta\varphi(0.04\,\mathrm{s})$ are arranged in a broad peak around $\Delta\varphi = 0$, while for the latter, there are two distinct peaks, one substantially below $\Delta\varphi = \pi$ and one substantially above. The peak below π is larger than its counterpart above π. For brevity, from here on the distribution of the $\Delta\varphi$ of oscillators that are supposed to keep their phase shift will be referred to as *zero-flip* peak or distribution, while the distribution of the $\Delta\varphi$ of oscillators corresponding to defective bits in the initialized pattern will be called π-*flip* distribution. Parts of the zero-flip distribution lie outside the desired range of $[-\pi/2, \pi/2]$; just as well, parts of the π-flip distribution lie outside the interval $[\pi/2, 3\pi/2]$. Those outliers account for the 20% failures. Further discussion of the shape of both distributions is postponed to section 6.1.4.

Fig. 45b shows the histogram of the proximity $P^1(0.04\,\mathrm{s})$ of the network to the pattern $\boldsymbol{\xi}^1$. As expected, instances of near-perfect pattern recognition are very rare, with the maximum

of the distribution occurring around $P^1(0.04\,\mathrm{s}) = 0.5$. The left part of the distribution (i.e. $P^1(0.04\,\mathrm{s}) \leq 0$) account for the 20% failures. Since the zero-flip distribution and the π-flip distribution in Fig. 45a overlap, there is also no criterion other than $P^1 > 0$ that reliably separates defective from correct bits in 100% of the cases.

Fig. 45c shows the evolution of the success rate (measured by the ratio of experiments with $P^1(t) > 0$ to all 250 experiments) during the recognition phase. Immediately after the Hebbian coupling is switched on, the phase shifts had not yet adjusted and therefore the success rate was zero. Then, there is a quick rise to the maximum value of $199/250 \approx 80\%$ after $0.04\,\mathrm{s}$. For longer times, the rate goes down again, due to the transient nature of pattern recognition in the presence of noise.

Finally, Fig. 45c shows how the minimal difference δ' of two frequency differences present in the system evolved during the series of experiments. This quantity, which varied around $1\,\mathrm{kHz}$, puts an upper limit on the coupling strength, namely $\varepsilon \leq 0.3 \cdot \delta' \approx 3 \cdot 10^2\,\mathrm{s}^{-1}$ (compare section 4.2.5).

On the other hand, there is also a lower limit for ε established by the frequency inaccuracy $\Delta\Omega_{\mathrm{TD}} = 100\,\mathrm{Hz}$, namely $\varepsilon \geq \Delta\Omega_{\mathrm{TD}}/0.15 \approx 7 \cdot 10^2\,\mathrm{s}^{-1}$ (compare section 4.2.1).

Since these two conditions cannot be fulfilled simultaneously, the network in the configuration hat hand cannot be expected to operate at a 100% success rate at all. In fact, especially given the inhomogeneous distribution of the coupling strength

$$\varepsilon \approx 1.5 \cdot 10^2\,\mathrm{s}^{-1} \ldots 1.2 \cdot 10^3\,\mathrm{s}^{-1},$$

the network seems to be working better than it has any right to. However, if one keeps in mind that $\Delta\Omega_{\mathrm{TD}} = 100\,\mathrm{Hz}$ was a very conservative estimate for the frequency inaccuracy, a success rate of 80% is less surprising.

Fig. 46 shows the full time evolution of all phase shifts φ_i during an individual, successful pattern recognition experiment. The average phase shift dynamics is overlain by distinct small amplitude oscillations, indicating that the separation of time scales is at or even below the limit for reliable pattern recognition (the amplitude is comparable with the oscillations in Fig. 29 for $\varepsilon = 0.5 \cdot \delta$).

Fig. 47 shows the evolution of the phase shifts during another, this time unsuccessful, experiment. This particular experiment failed because the oscillator represented by the cyan curve still was in a transient during the second half of the initialization interval and therefore its frequency was determined incorrectly, thus derailing the whole recognition process. The two approaches for dealing with this type of problem, namely either increasing the coupling strength or choosing a longer time interval for one measurement were both not feasible for this particular network, because this would have meant to have even less separation of time scales or even larger frequency inaccuracies, respectively.
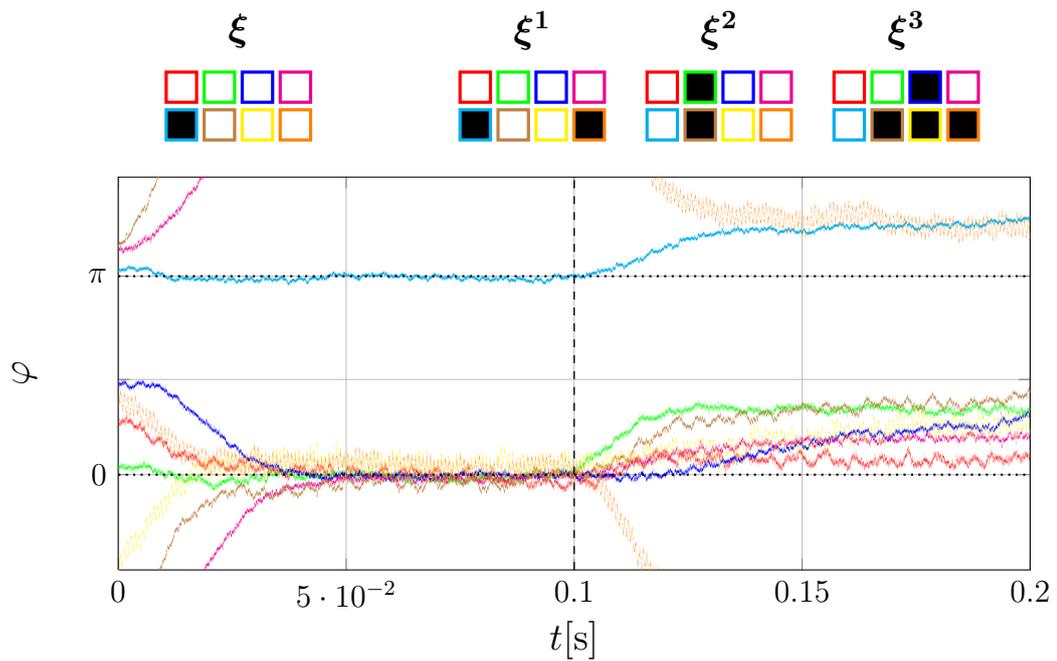
Fig. 46: Time evolution of the phase shifts φ_i during a successful pattern recognition experiment in the series shown in Fig. 45. The initial pattern ξ and the memorized patterns ξ^k are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot.
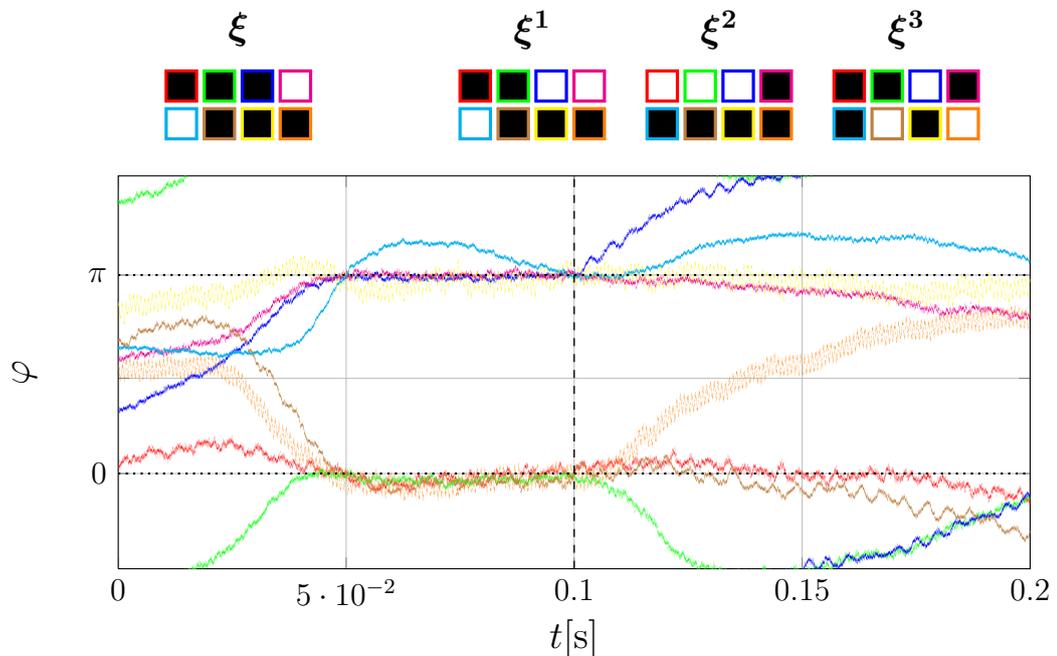
Fig. 47: Time evolution of the phase shifts φ_i during an unsuccessful pattern recognition experiment in the series shown in Fig. 45. The initial pattern ξ and the memorized patterns ξ^k are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot.

6.1.2 Network of NIC-type oscillators with coupling through a VCR

To obtain a higher success rate, the TD-type oscillators were replaced by the more frequency-stable NIC-type oscillators. Also, the inductances were realized with gyrators which simplified the tuning of the frequencies, facilitating a larger value of $\delta' \approx 5\,\mathrm{kHz}$ (instead of $\delta' \approx 1\,\mathrm{kHz}$ for the TD-type network, see Fig. 48). Moreover, as described earlier, the frequency was tuned by changing the value of L with the NIC-type oscillators and therefore the coupling strength was the same for all oscillators. The upper and lower limits for ε are now $\varepsilon \leq 0.3 \cdot \delta' \approx 1.5 \cdot 10^3\,\mathrm{s}^{-1}$ and $\varepsilon \geq \Delta\Omega_{\mathrm{TD}}/0.15 \approx 2.3 \cdot 10^{-2}\,\mathrm{s}^{-1}$, respectively.

As a result, there was actually a range of values for the coupling strength ε in which the system performed reasonably well, albeit not perfectly. The highest success rate for a NIC-type network with a VCR2N providing the time-dependent coupling was 95%. Four series of experiments were conducted, each consisting of 500 pattern recognition runs. The experimental parameters are given in table 6.1, together with the success rates. The values of the coupling strength ranged from $\varepsilon \approx 3 \cdot 10^2\,\mathrm{s}^{-1}$ to $\varepsilon \approx 8 \cdot 10^2\,\mathrm{s}^{-1}$. Note that for two of the series, the coupling strength was increased by a factor of 4/3 during recognition. The reason for this will become apparent below.

T_{record}	$U_{\mathrm{coup}}^{\mathrm{amp}}$	ε	N_{success}/N	success rate
$0.5\,\mathrm{s}$	$2000\,\mathrm{mV}$	$6 \cdot 10^2\,\mathrm{s}^{-1}$	448/500	90%
$0.5\,\mathrm{s}$	$2000\,\mathrm{mV}$ / $2660\,\mathrm{mV}$	$6 \cdot 10^2\,\mathrm{s}^{-1}$ / $8 \cdot 10^2\,\mathrm{s}^{-1}$	460/500	92%
$1\,\mathrm{s}$	$1000\,\mathrm{mV}$	$3 \cdot 10^2\,\mathrm{s}^{-1}$	433/500	87%
$1\,\mathrm{s}$	$1000\,\mathrm{mV}$ / $1330\,\mathrm{mV}$	$3 \cdot 10^2\,\mathrm{s}^{-1}$ / $4 \cdot 10^2\,\mathrm{s}^{-1}$	473/500	95%

Table 6.1: Parameter sets and success rates for the four series of experiments conducted with a NIC-type network with a VCR-based coupling. If two parameter values are given, the first one applies to the initialization phase and the second one to the recognition phase of the experiment. For all four series, the values $U_{\mathrm{DA}}^{\mathrm{min}} = 0\,\mathrm{mV}$, $U_{\mathrm{DA}}^{\mathrm{max}} = 2500\,\mathrm{mV}$ and $U_{\mathrm{coup}}^{\mathrm{off}} = 1250\,\mathrm{mV}$ were the same.

Since the values of ε are well within the favorable value region for ε, the fail rates of up to 13% cannot be explained by a small separation of time scales, nor by frequency inaccuracies. There are at least two other effects that hinder the pattern recognition, namely the distortion of the coupling signal by the VCR2N and the cutoff of the coupling signal at $U_{\mathrm{DA}}^{\mathrm{min}}$ and $U_{\mathrm{DA}}^{\mathrm{max}}$. These effects certainly play a role - in particular, as will be described in the next section, the network employing the (almost) distortion-free AD633 multiplier as main coupling element works considerably better than with the coupling discussed here. Still, there is another detrimental effect to pattern recognition that is not present in the phase-only description of the network discussed in the theory chapter.

To understand this effect, it is instructive to look at the statistics of, for instance, the series of experiments described in the third row of table 6.1, depicted in Fig. 48. In this series, the maximum number of successful pattern recognitions (87%) occurred $0.2\,\mathrm{s}$ after the coupling
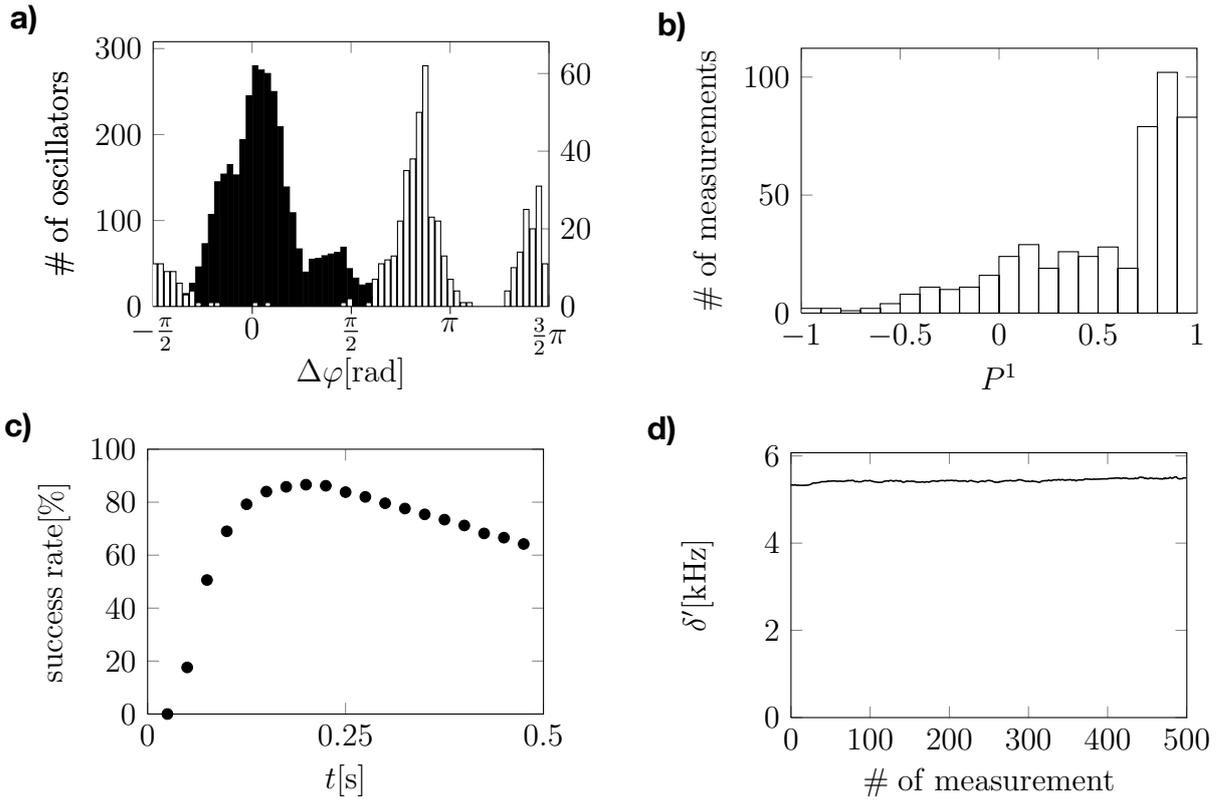
Fig. 48: Results of a series of 500 pattern recognition experiments with the circuit depicted in Fig. 42, using an array of NIC-type oscillators. The parameters of the experiment were $T_{\mathrm{record}} = 1\,\mathrm{s}$, $U_{\mathrm{coup}}^{\mathrm{amp}} = 1000\,\mathrm{mV}$, $U_{\mathrm{coup}}^{\mathrm{off}} = 1250\,\mathrm{mV}$, $U_{\mathrm{DA}}^{\mathrm{min}} = 0\,\mathrm{mV}$ and $U_{\mathrm{DA}}^{\mathrm{max}} = 2500\,\mathrm{mV}$. The maximum recognition rate was $433/500 \approx 87\%$ after $0.2\,\mathrm{s}$ in recognition mode. **a)** Distribution of $\Delta\varphi(0.2\,\mathrm{s})$ for oscillators that are supposed to keep their phase shift (black bars, left vertical axis) and oscillators belonging to the defective bit (white bars, right vertical axis). **b)** Distribution of $P^1(0.2\,\mathrm{s})$. **c)** Success rate as measured by P^1 over time in recognition mode. **d)** Evolution of the minimal difference of frequency differences δ' during the series of experiments. The bin size in a) and b) was 0.1.

was switched to recognition mode. Fig. 48a shows that the zero-flip peak in the distribution of $\varphi(0.2\,\mathrm{s})$ is strongly asymmetric, as opposed to its shape for the TD-type network. The π-flip distribution, i. e. the distribution of $\Delta\varphi(0.2\,\mathrm{s})$ for oscillators corresponding to a defective bit, again has two peaks. The difference in size between the two peaks is even more pronounced than in the TD-type network. Also it appears that both the zero-flip and the π-flip distribution are slightly shifted to the right. As a result, the histogram of $P^1(0.2\,\mathrm{s})$ shows proportionately more measurements close to $P^1(0.2\,\mathrm{s}) = 1$, than for the TD-type network, because there are more defective oscillators actually undergoing a phase shift close to π. The shift is also responsible for the fact that a disproportionately large part of the right peak in the π-flip distribution lies above $\Delta\varphi = 3\pi/2$ (or $-\pi/2$). As a result of this, oscillators corresponding to defective bits in the initial pattern that slow down instead of speeding up to adjust their phase shift to the correct value are especially likely to cause a failed pattern recognition.

The reason for the shift to the right is that all oscillators collectively slow down slightly during initialization mode (compared to their frequencies in the uncoupled state) and then
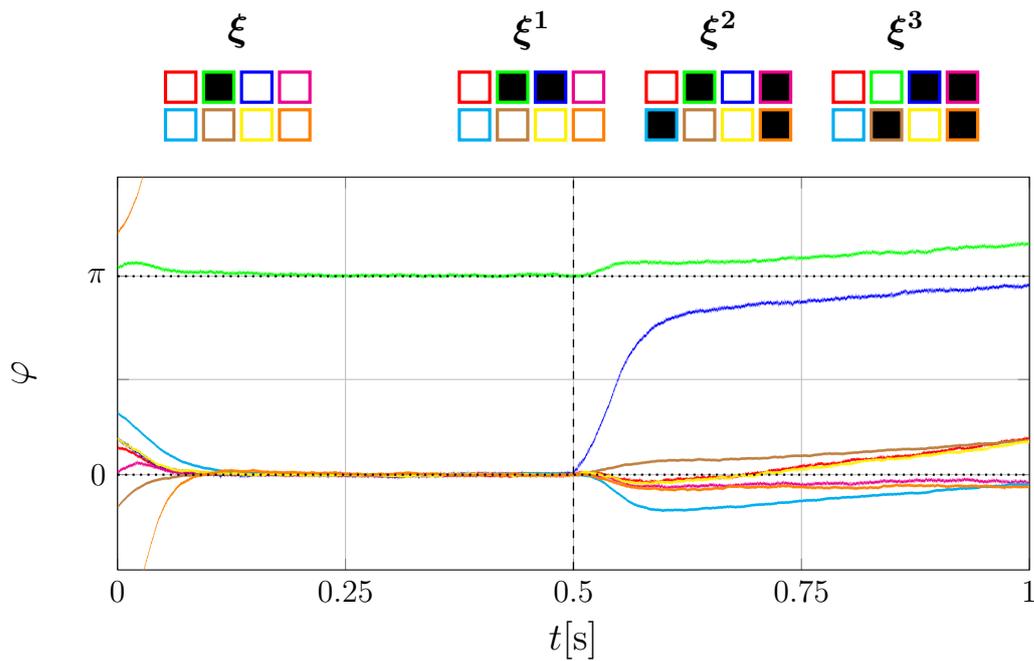
Fig. 49: Time evolution of the phase shifts φ_i during a successful pattern recognition experiment in the series shown in Fig. 48. The initial pattern ξ and the memorized patterns ξ^k are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot.
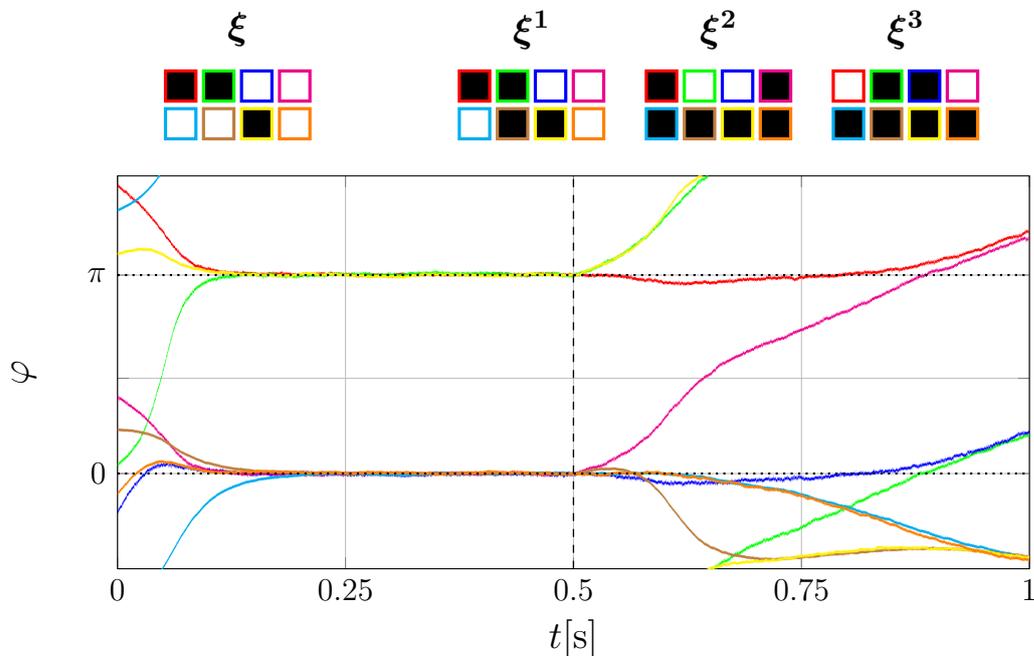
Fig. 50: Time evolution of the phase shifts φ_i during an unsuccessful pattern recognition experiment in the series shown in Fig. 48. The initial pattern ξ and the memorized patterns ξ^k are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot.

speed up again a little (still not recovering their original frequencies), once the coupling is changed from initialization mode to recognition mode. The slowdown effect increases with the coupling strength. In the series shown in Fig. 48, all oscillators slowed down by about $0.5\,\mathrm{Hz}\ldots1.5\,\mathrm{Hz}$ during initialization. This behavior can be observed indirectly for example in Fig. 49, which shows the results of an individual, successful pattern recognition experiment from the series depicted in Fig. 48. Only a speedup in the second half is visible, because the phase shifts were measured with respect to the frequencies during initialization. In Fig. 50, which shows the results of another, this time unsuccessful pattern recognition experiment from the same series, one can see, how a collective speedup (on top of other frequency influencing effects), can derail the pattern recognition: It appears that, whenever the oscillator belonging to the defective bit slows down in order to switch branches, other oscillators are particularly prone to speed up considerably (as the oscillators represented by the green, yellow and magenta curves do in Fig. 50), thus endangering a successful pattern recognition. This is intuitive, because to achieve an increase in the average speed of all oscillators, other oscillators must counteract a single oscillator that is slowing down. A switching oscillator that is speeding up causes less problems.

As of now, no completely satisfactory explanation for this behavior has been found. The nonlinearity of the VCR2N can be ruled out quite safely, because the effect is also present for a different coupling mechanism. It appears that the effect is related to the global frequency decrease due the very slow time scale dynamics described in section 4.2.5. This effect leads to a slightly slower average frequency during initialization as well. However, it does not cause a significant amount of recognition failures in an otherwise ideal system, nor does it suppress the right hand peak in the π-flip distribution; the effect must therefore be enhanced by a mechanism that is not captured by the phase model (the last two statements are corroborated in the discussion section further below).

To counteract the speedup between initialization and recognition mode, experiments were made with different values of ε before and after the coupling was switched to the Hebbian learning rule. Fig. 51 shows the results of the series of experiments described in the last row of table 6.1. The only difference with respect to the series in Fig. 48 was that the value of ε was increased by a factor of 4/3 during recognition. Increasing ε had the effect of eliminating the shift in the distribution of $\Delta\varphi$. As a consequence, the success rate, as measured by P^1, went up considerably, namely from 87% to 95%. Also, the decrease of the success rate in Fig. 51c is slower than in Fig. 48c. Of course, one could have achieved the same increase in the success rate by subtracting a suitable constant value from all $\Delta\varphi$ values before calculating the proximity.

Unfortunately, increasing ε did not have the desired effect of eliminating all false recognitions due to the collective speedup. Moreover, there is still some asymmetry in the system, favoring oscillators that speed up to switch branches: of 27 failed pattern recognition
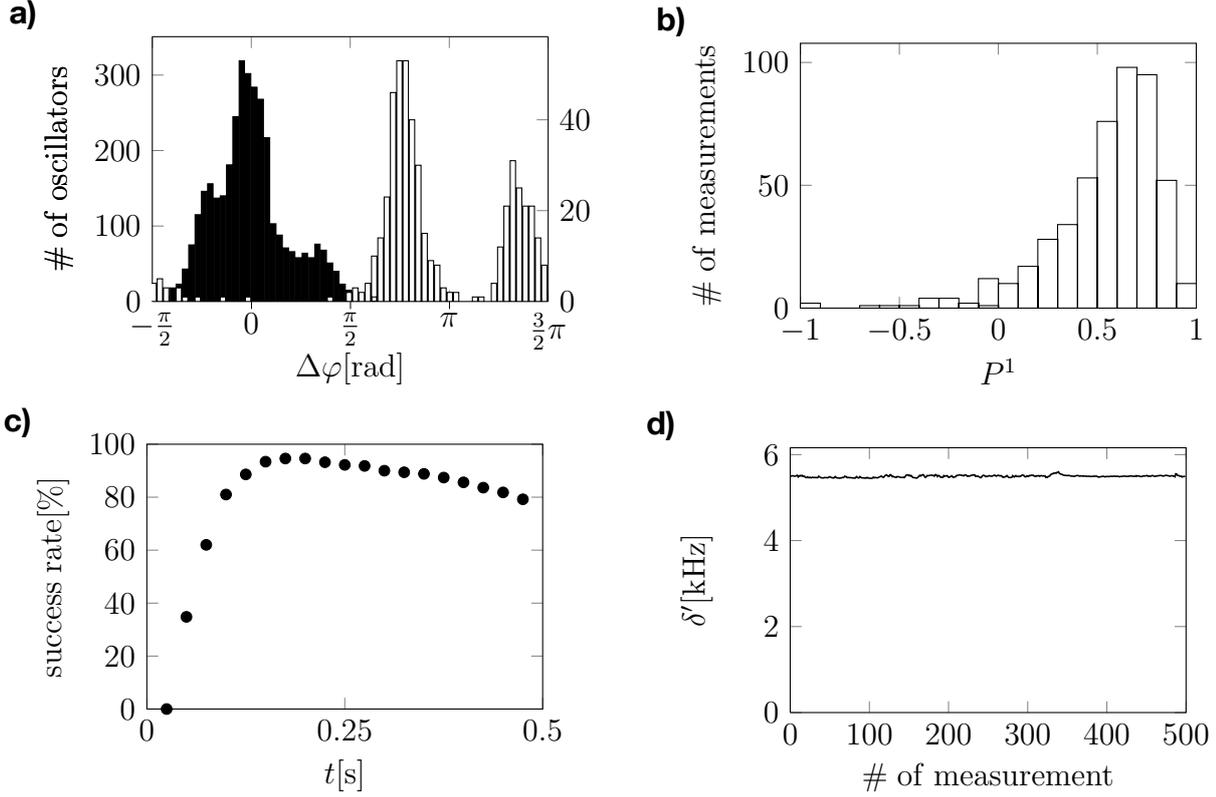
Fig. 51: Results of a series of 500 pattern recognition experiments with the circuit depicted in Fig. 42 using an array of NIC-type oscillators. The time interval of a single run was $T_{\text{record}} = 1\,\text{s}$. During initialization, a coupling signal with $U_{\text{coup}}^{\text{amp}} = 1000\,\text{mV}$ and $U_{\text{coup}}^{\text{off}} = 1250\,\text{mV}$ was applied, whereas during recognition $U_{\text{coup}}^{\text{amp}} = 1330\,\text{mV}$ and $U_{\text{coup}}^{\text{off}} = 1250\,\text{mV}$ were used. $U_{\text{DA}}^{\text{min}} = 0\,\text{mV}$ and $U_{\text{DA}}^{\text{max}} = 2500\,\text{mV}$ were the same in both cases. The maximum recognition rate was $473/500 \approx 95\%$ after $0.2\,\text{s}$ in recognition mode. **a)** Distribution of $\Delta\varphi(0.2\,\text{s})$ for oscillators that are supposed to keep their phase shift (black bars, left vertical axis) and oscillators belonging to the defective bit (white bars, right vertical axis). **b)** Distribution of $P^1(0.2\,\text{s})$. **c)** Success rate as measured by P^1 over time in recognition mode. **d)** Evolution of the minimal difference of frequency differences δ' during the series of experiments. The bin size in a) and b) was 0.1.

experiments, 22 featured an oscillator corresponding to the defective bit that slowed down during the recognition transition although (given the relative sizes of the two peaks in the π-flip distribution) experiments in which the oscillator in question slows down are less than half in number compared to experiments in which it speeds up. Fig. 52 and Fig. 53 show the results of a successful and an unsuccessful experimental run contained in the series shown in Fig. 51. Note how, due to the change in coupling strength, the average frequency is roughly the same during initialization and recognition in both cases.
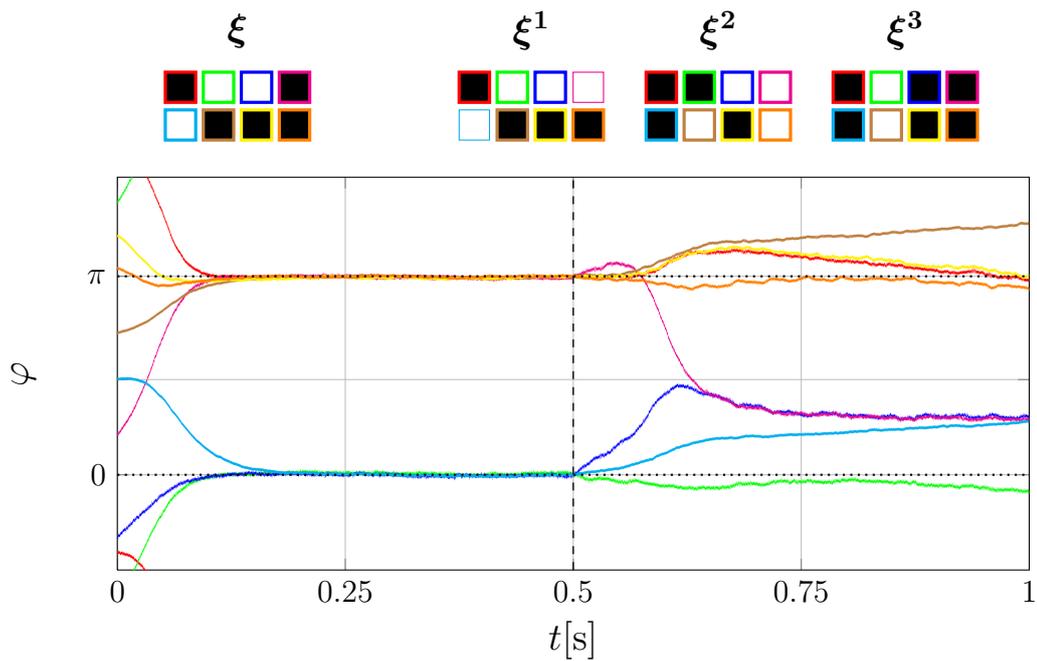
Fig. 52: Time evolution of the phase shifts φ_i during a successful pattern recognition experiment in the series shown in Fig. 51. The initial pattern $\boldsymbol{\xi}$ and the memorized patterns $\boldsymbol{\xi}^k$ are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot. The recognition is successful despite the fact that the oscillator belonging to the defective bit slows down during the transition (see text).
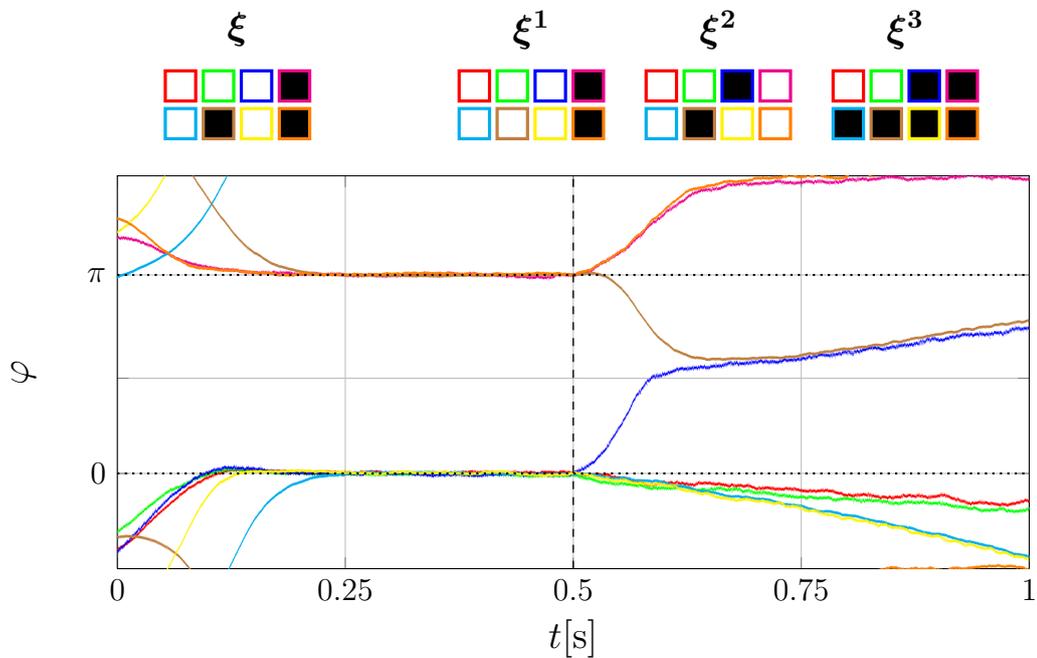
Fig. 53: Time evolution of the phase shifts φ_i during an unsuccessful pattern recognition experiment in the series shown in Fig. 51. The initial pattern $\boldsymbol{\xi}$ and the memorized patterns $\boldsymbol{\xi}^k$ are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot.

6.1.3 Network of NIC-type oscillators with coupling through analog computing devices

To avoid the distortion that the nonlinear characteristic of the VCR2N introduces to the coupling function, the coupling mechanism in the NIC-type network was also implemented using an inverting adder and a multiplier. This approach provided the best experimental results; table 6.2 shows the parameters and success rates of 6 series of 500 pattern recognition experiments each. The minimal pattern recognition success rate was 96%. In one series, every single one of the 500 experiments worked.

T_{record}	$U_{\text{coup}}^{\text{max}}, U_{\text{coup}}^{\text{min}}$	$U_{\text{coup}}^{\text{amp}}$	ε	N_{success}/N	success rate
$0.2\,\text{s}$	$\pm 750\,\text{mV}$	$180\,\text{mV}$	$2.2\cdot 10^3\,\text{s}^{-1}$	$479/500$	96%
$0.2\,\text{s}$	$\pm 750\,\text{mV}$	$180\,\text{mV} / 240\,\text{mV}$	$2.2\cdot 10^3\,\text{s}^{-1} / 2.9\cdot 10^3\,\text{s}^{-1}$	$489/500$	98%
$0.5\,\text{s}$	$\pm 375\,\text{mV}$	$90\,\text{mV}$	$1.1\cdot 10^3\,\text{s}^{-1}$	$499/500$	99.8%
$0.5\,\text{s}$	$\pm 375\,\text{mV}$	$90\,\text{mV} / 120\,\text{mV}$	$1.1\cdot 10^3\,\text{s}^{-1} / 1.4\cdot 10^3\,\text{s}^{-1}$	$500/500$	100%
$1\,\text{s}$	$\pm 190\,\text{mV}$	$45\,\text{mV}$	$5.4\cdot 10^2\,\text{s}^{-1}$	$483/500$	97%
$1\,\text{s}$	$\pm 190\,\text{mV}$	$45\,\text{mV} / 60\,\text{mV}$	$5.4\cdot 10^2\,\text{s}^{-1} / 7.2\cdot 10^2\,\text{s}^{-1}$	$491/500$	99%

Table 6.2: Parameter sets and success rates for the six series of experiments conducted with a NIC-type network with a coupling based on analog computing devices. If two parameter values are given, the first one applies to the initialization phase and the second one to the recognition phase of the experiment. For all six series, $U_{\text{coup}}^{\text{off}} = 0\,\text{mV}$ was used.

Since the upper and lower limits for ε are $\varepsilon \leq 1.5\cdot 10^2\,\text{s}^{-1}$ and $\varepsilon \geq 2.3\cdot 10^3\,\text{s}^{-1}$, respectively, it is surprising that pattern recognition in the first two series in table 6.2 worked so well. Once again, this is most likely because of the rather conservative estimate for the frequency accuracy. Still, the pattern recognition performance is best right in the middle between the upper and lower limits.

Fig. 54 shows the statistics of the series of experiments with 99.8% success rate with the parameter values shown in the third row of table 6.2. Compared to the NIC-type network coupled through a VCR, the asymmetry in the zero-flip and π-flip distributions is even greater. In fact, both the shoulder on the right of the zero-flip peak and the right hand side peak of the π-flip distribution of $\Delta\varphi$ have disappeared entirely (see Fig. 54a). Also, the overall shift of both distributions towards the right is even larger than in Fig. 48a. This is a strong indicator that the asymmetry of the zero-flip and π-flip distributions is indeed caused by the oscillators' speedup during recognition.

The positive side of the asymmetry effect is the fact that it tends to narrow both the zero-flip and the π-flip distributions, such that they are well separated (with only a few exceptions); this means that in general, pattern recognition works very well. Fig. 55 shows an example of a successful pattern recognition experiment from the series in Fig. 54. The speedup effect is much more prominent than in Fig. 49. Here, the slowdown during initialization was

a)

b)

c)

d)
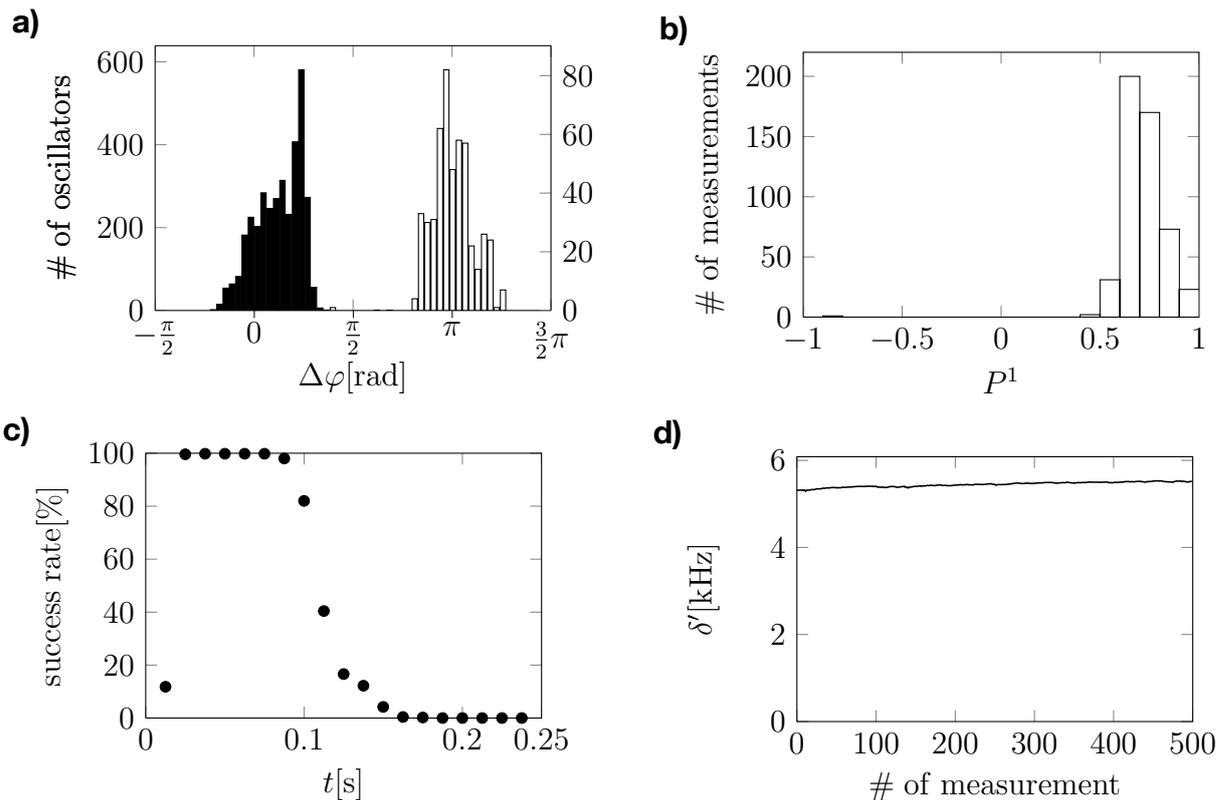
Fig. 54: Results of a series of 500 pattern recognition experiments with the circuit depicted in Fig. 43 using an array of NIC-type oscillators. The parameters of the experiment were $T_{\mathrm{record}} = 0.5\,\mathrm{s}$, $U_{\mathrm{coup}}^{\mathrm{amp}} = 90\,\mathrm{mV}$, $U_{\mathrm{coup}}^{\mathrm{off}} = 0\,\mathrm{mV}$, $U_{\mathrm{DA}}^{\mathrm{min}} = -375\,\mathrm{mV}$ and $U_{\mathrm{DA}}^{\mathrm{max}} = 375\,\mathrm{mV}$. The maximum recognition rate was $499/500 \approx 100\%$ during a time interval of roughly $0.04\,\mathrm{s}$ in recognition mode. **a)** Distribution of $\Delta\varphi(0.05\,\mathrm{s})$ for oscillators that are supposed to keep their phase shift (black bars, left vertical axis) and oscillators belonging to the defective bit (white bars, right vertical axis). **b)** Distribution of $P^1(0.05\,\mathrm{s})$. **c)** Success rate as measured by P^1 over time in recognition mode. **d)** Evolution the minimal difference of frequency differences δ' during the series of experiments. The bin size in a) and b) was 0.1.

$3\,\mathrm{Hz} \ldots 4\,\mathrm{Hz}$.

Still, there is the possibility that the oscillator corresponding to the defective bit in ξ slows down instead of speeding up during recognition. This happened exactly once in the series depicted in Fig. 54 and, unfortunately, pattern recognition failed in this case (see Fig. 56). The oscillator in question is represented by the brown curve. Note that the information about the correct pattern in this instance is completely unsalvageable from the phase shift data, even after an arbitrary common offset is subtracted from all phase shift values to cancel out the speedup.

Like it was already done for the NIC-type network with VCR based coupling, the series of experiments was repeated with an increased coupling strength (factor 4/3) during the recognition phase. The results are shown in Fig. 57. In this series, every single experiment was successful (Fig. 58 shows an example). Again, the higher coupling strength during the recognition phase was able to cancel out the speedup effect. During this series of experiments, there was no oscillator corresponding to a defective bit that slowed down during recognition, which explains the 100% success rate.
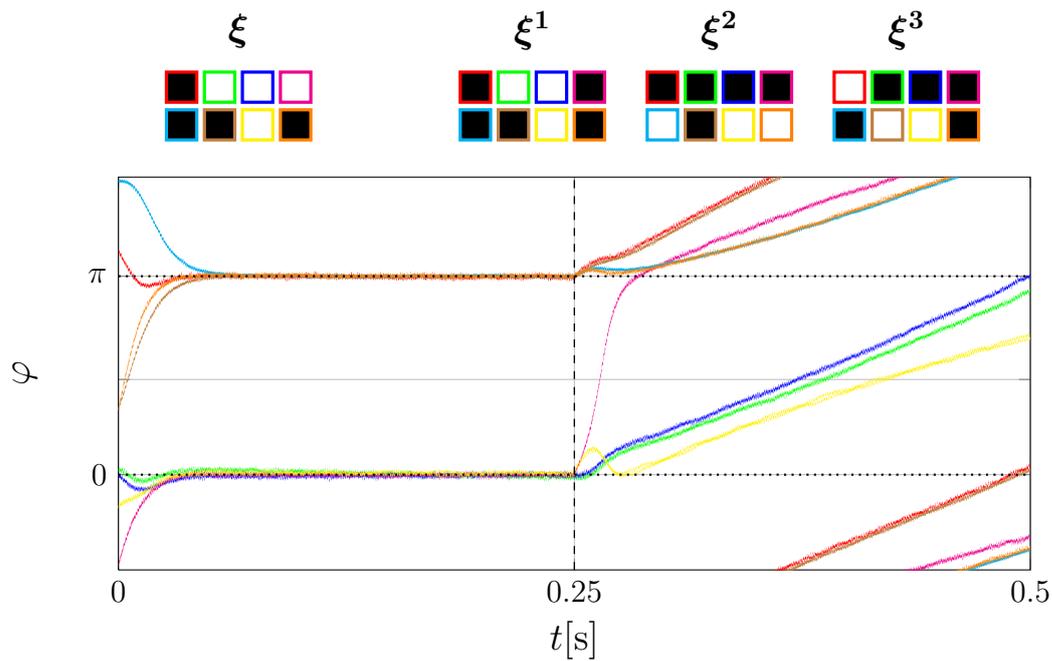
Fig. 55: Time evolution of the phase shifts φ_i during a successful pattern recognition experiment in the series shown in Fig. 54. The initial pattern ξ and the memorized patterns ξ^k are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot.
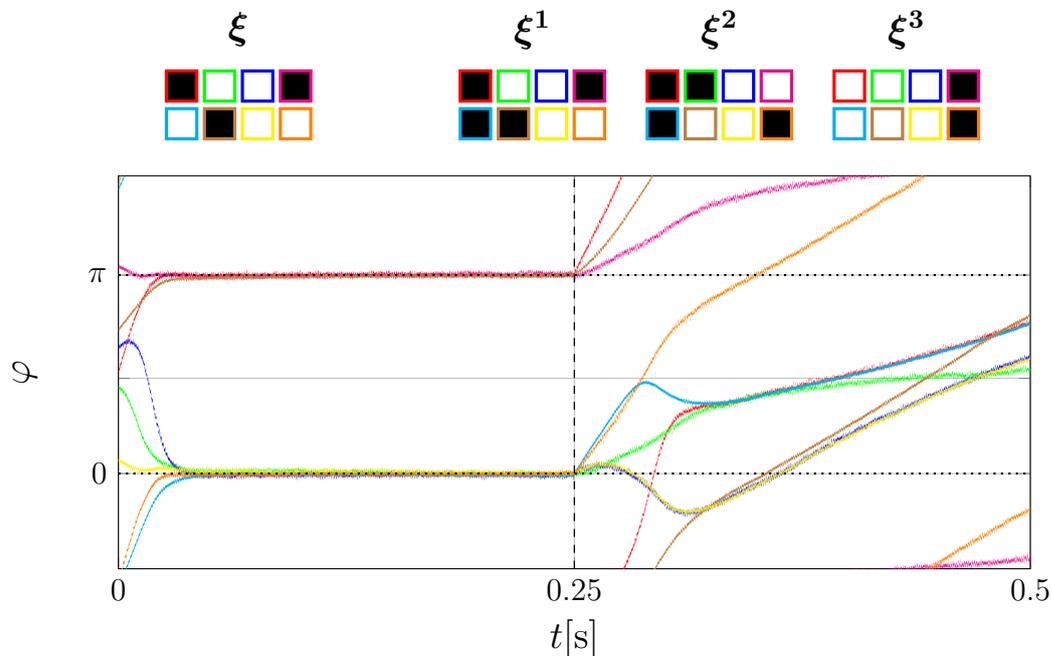
Fig. 56: Time evolution of the phase shifts φ_i during the single unsuccessful pattern recognition experiment in the series shown in Fig. 54. The initial pattern ξ and the memorized patterns ξ^k are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot.
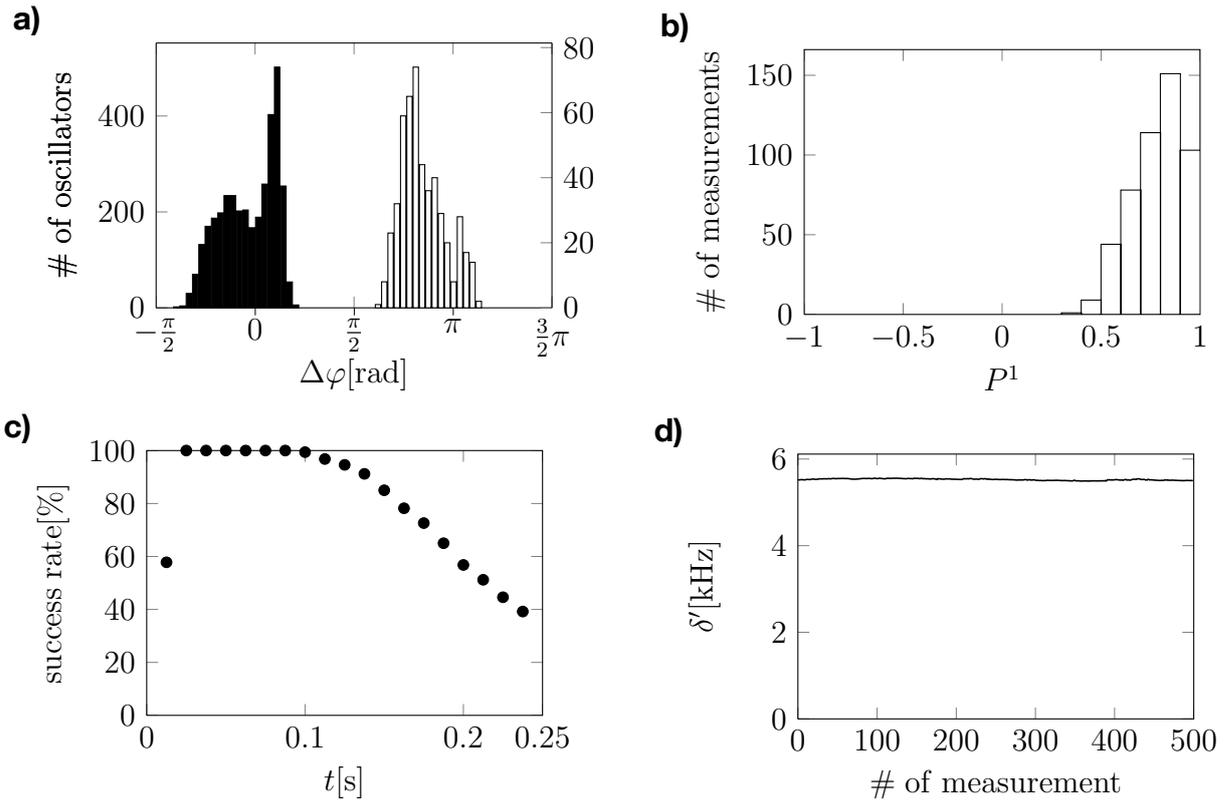
Fig. 57: Results of a series of 500 pattern recognition experiments with the circuit depicted in Fig. 43 using an array of NIC-type oscillators. The time interval of a single run was $T_{\mathrm{record}} = 0.5\,\mathrm{s}$. During initialization, a coupling signal with $U_{\mathrm{coup}}^{\mathrm{amp}} = 90\,\mathrm{mV}$ and $U_{\mathrm{coup}}^{\mathrm{off}} = 0\,\mathrm{mV}$ was applied, whereas during recognition $U_{\mathrm{coup}}^{\mathrm{amp}} = 120\,\mathrm{mV}$ and $U_{\mathrm{coup}}^{\mathrm{off}} = 0\,\mathrm{mV}$ were used. $U_{\mathrm{DA}}^{\mathrm{min}} = -375\,\mathrm{mV}$ and $U_{\mathrm{DA}}^{\mathrm{max}} = 375\,\mathrm{mV}$ were the same in both cases. The maximum recognition rate was $500/500 = 100\%$ during a time interval of roughly $0.06\,\mathrm{s}$ in recognition mode. **a)** Distribution of $\Delta\varphi(0.05\,\mathrm{s})$ for oscillators that are supposed to keep their phase shift (black bars, left vertical axis) and oscillators belonging to the defective bit (white bars, right vertical axis). **b)** Distribution of $P^1(0.05\,\mathrm{s})$. **c)** Success rate as measured by P^1 over time in recognition mode. **d)** Evolution of the minimal difference of frequency differences δ' during the series of experiments. The bin size in a) and b) was 0.1.
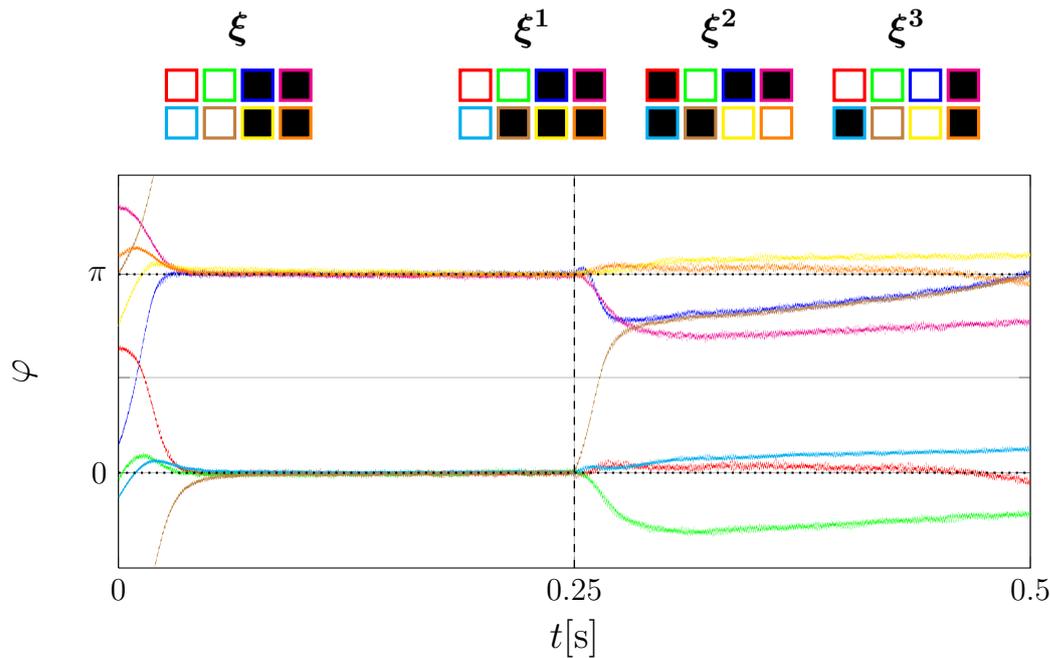
Fig. 58: Time evolution of the phase shifts φ_i during a pattern recognition experiment in the series shown in Fig. 57. The initial pattern ξ and the memorized patterns ξ^k are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot.

6.1.4   Discussion of the results

The most important result of the experiments presented above is the fact that a globally coupled network of NIC-type van der Pol oscillators could be used reliably for pattern recognition with 3 memorized mutually orthogonal 8 bit patterns if an adder and a multiplier were used to provide the time-dependent coupling.

Other than for the network of TD-type oscillators, for the network of NIC-type oscillators there is a favorable range of the coupling strength given by $\varepsilon = 1.5 \cdot 10^2 \, \mathrm{s}^{-1} \ldots 2.3 \cdot 10^3 \, \mathrm{s}^{-1}$ in which the main obstacle for pattern recognition is neither the frequency inaccuracy of the oscillators, nor the small separation of time scales.

Instead, there is another dynamical effect with a potentially negative influence on pattern recognition that does not appear in the phase model investigated in the theory chapter.

Before this effect is discussed, however, the experimental results are compared to the behavior of a hypothetical ideal pattern recognition device which exhibits phase jumps of exactly $\Delta\varphi = 0$ for oscillators belonging to correct bits in ξ and $\Delta\varphi = \pi$ for oscillators belonging to defective bits in ξ). A histogram of experiments with this device, in the presence of some Gaussian noise, should show two bell shaped distributions centered around 0 and π respectively. The observed statistics deviates from this behavior in the following points:

- There appears to be an overall increase in the oscillator frequencies, when the coupling function is changed from pattern initialization to pattern recognition. As a

result, the entire histogram of phase jumps appears shifted towards the right.

- Choosing a larger coupling strength for the recognition phase than for the initialization phase (ratio roughly 4:3) cancels out this shift effect.

- The zero-flip peak has a pronounced shoulder or even another smaller peak on the left, and a less pronounced shoulder on the right. The shoulder on the right (almost) disappears for the more successful experimental runs with less noise and distortion in the network dynamics.

- The π-flip distribution exhibits two peaks that are clearly separated. The peak on the right hand side is less pronounced and (almost) disappears for the more successful experimental runs with less noise and distortion in the network dynamics. This peak corresponds to oscillators that slow down during pattern recognition to adjust their phase shift.

To find an explanation for these findings it is instructive to look at comparable statistics for the ideal model of globally coupled phase oscillators obeying (4.18), as depicted in Fig. 59, which shows the statistics for a series of 10000 simulations using $\varepsilon = 5.4 \cdot 10^2\,\mathrm{s}^{-1}$ and an ideal Golomb ruler between $f_1 = 35\,\mathrm{kHz}$ and $f_8 = 69\,\mathrm{kHz}$. Like in the experiments, three randomized orthogonal patterns $\boldsymbol{\xi}^k$ were used, along with an initial pattern $\boldsymbol{\xi}$ that differed from ξ^1 in one bit.
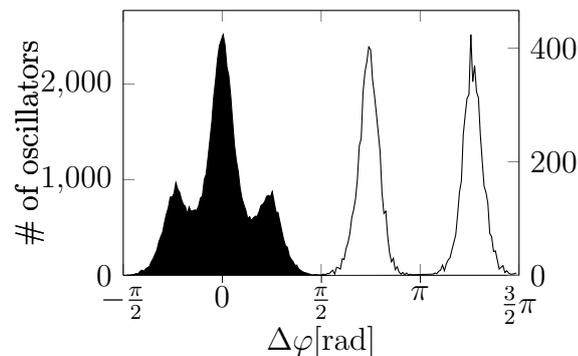
Fig. 59: Results of a series of 10000 numerical integrations simulating pattern recognition with the ideal phase equation (4.18) for 8 oscillators with frequencies that formed a Golomb ruler between $f_1 = 35\,\mathrm{kHz}$ and $f_8 = 69\,\mathrm{kHz}$. The time interval of a single integration run was $T_{\mathrm{run}} = 0.5\,\mathrm{s}$. The coupling strength was set to $\varepsilon = 5.4 \cdot 10^2\,\mathrm{s}^{-1}$. The choice of initial and memorized patterns was the same as in the experiments. Simulations were started at random initial phases. The graphic shows the zero-flip distributions (black area, left vertical axis) and the π-flip distribution (white area, right vertical axis) of $\Delta\varphi(0.075\,\mathrm{s})$. The bin size of the histogram was 0.005.

Fig. 59 shows that the zero-flip distribution has two distinct, fairly symmetrical side peaks next to the main peak. These are mirrored by the two distinct peaks in the π-flip distribution. The reasons for these features are the following:

First, the phase shift φ of the oscillator representing the defective bit does not change by

π during recognition, but by a little less, since the other oscillators are pushed in the other direction a bit (albeit not much, because of their larger number); after all, the sum of all phase shifts is conserved (on the slow, but not on the very slow time scale; this is true because the sum over the right hand side of (2.2) for all i is zero). This explains the two peaks in the π-flip distribution, formed by oscillators that either speed up (corresponding to the left peak) or slow down (corresponding to the right peak) to adjust to a correct phase shift. Second (which explains the two smaller side peaks for in the zero-flip distribution), the entry $w_{ij} = w_{ji}$ describing the coupling strength between oscillators i and j has an absolute value of either 1 or 3 during the recognition stage (for three memorized patterns). If the memorized patterns are chosen orthogonal, it turns out that each oscillator has exactly one partner oscillator to which it is coupled with relative strength 3 while it is coupled to the rest with relative strength 1. The side peaks are formed by the partners of the oscillators corresponding to defective bits, which experience a bigger backlash from the phase jump than the rest. Note that these are small size effects which should not play a role for larger systems with random patterns.

Another small size effect is the fact that there were actually 8 simulations that failed to deliver the correctly recognized pattern *even in an absolutely noiseless* phase model. The ratio of time scales is $\varepsilon/\delta \approx 0.18$, which in this simulation is sufficient to cause a fail rate of 0.08%. This is a small size effect that should disappear in larger networks, because, if the ratio of invariant eigendirections to phase space dimensions were significantly lower than 3/8 (namely equal to $\alpha \leq 0.042$ in a higher dimensional network with random memorized patterns) any drift effect is expected to be proportionately smaller and less prone to cause pattern recognition errors. Note that the experiments were also subject to this effect, but it was overshadowed by other detrimental effects

In the simulations, the overall shift towards larger phase shift values seems not to be present at all. Also, the right hand side peak of the π-flip distribution does not seem to be suppressed compared to the left hand side peak. As mentioned before, this indicates that the shift of the whole distribution of the $\Delta\varphi$ to the right and the asymmetry between the two peaks in the π-flip distribution (entailing the asymmetry of the shoulders of the zero-flip distribution) in the experiments are related effects, meaning that if there is an overall tendency for oscillators to speed up due to the recognition coupling, the symmetry between the two peaks is broken as well. If one looks closely, there is a slight dominance of the peak on the left over the peak on the right ($n_{\mathrm{speedup}} = 5139$, $n_{\mathrm{slowdown}} = 4859$ simulations (note that the missing two oscillators belong to simulations with failed recognition)). However, this effect is far to small to explain the huge asymmetry in the experimental data.

In conclusion, it seems clear that the symmetry breaking observed in the experiments cannot be explained by a phase-only dynamical model. On the other hand, if the experiments are

modeled by the realistic equations

$$
\begin{aligned}
\dot{U}_i &= \frac{1}{C_i}\left(-I_{\mathrm{NL}}(U_i) - I - \frac{U_i}{R_{\mathrm{int}}}\right) + \frac{U_{\mathrm{DA}}(t)}{10 \cdot C_i R_{\mathrm{int}}}\sum_{i=1}^{8} U_i \\
\dot{I}_i &= \frac{U_i}{L_i},
\end{aligned}
\tag{6.1}
$$

where $C_i = C = 1 \cdot 10^{-9}\,\mathrm{nF}$, the L_i are chosen appropriately to implement the Golomb ruler between $f_1 = 35\,\mathrm{kHz}$ and $f_8 = 69\,\mathrm{kHz}$ and $U_{\mathrm{DA}}(t)$ is chosen to provide a coupling strength of $\varepsilon = 1.1 \cdot 10^3\,\mathrm{s}^{-1}$, the asymmetry is close to that observed in the experiments (see Fig. 60a). Interestingly, the histogram can be altered quite significantly by choosing a different characteristic I_{NL} than the one used in the NIC-type van der Pol oscillator (Fig. 60b). This is shown in Fig. 61a. All other parameters being equal, simulating the system (6.1) with the nonlinear characteristic I_{NL} shown in Fig. 61b yields results that are much closer to the phase model.

Apparently, the dynamics of the system away from the limit cycle, which is influenced by the shape of the nonlinearity $I_{\mathrm{NL}}(U)$, is critical for the asymmetry effect. This is reminiscent of the so called acceleration/deceleration effect investigated by Aonishi, et al. (2002). This effect leads to an increase or decrease of the average frequency of locked coupled oscillators of similar frequency, depending on the curvature of the isochrones in the vicinity of the limit cycle. It is not self-evident however, how this mechanism can be transferred to the system of weakly coupled oscillators far below the Kuramoto threshold.

As a final remark, it should be noted that, like the very distinct features of the zero-flip and π-flip distributions in the phase model, the asymmetry effect should be strongest for a small number of oscillators and play less of a role for a system with a larger number of oscillators and random patterns.
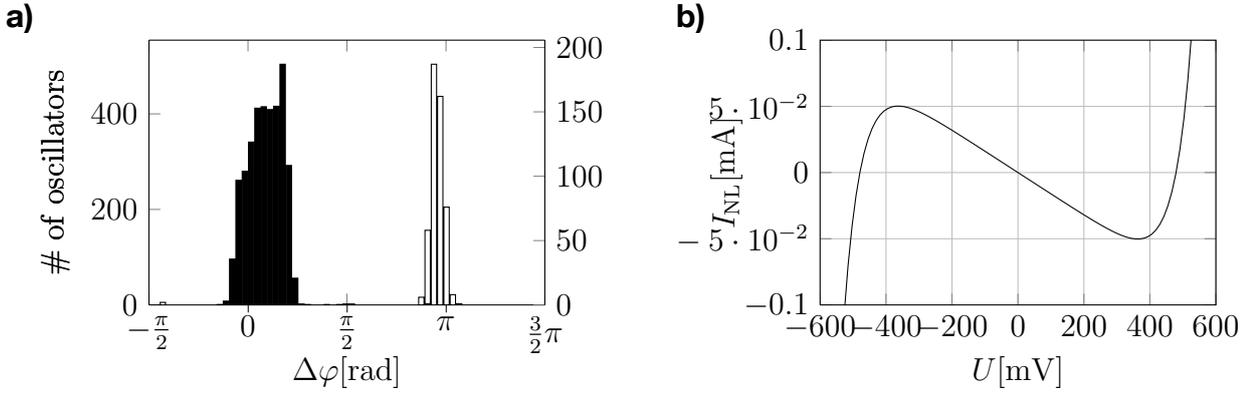
Fig. 60: a) Results of a series of 500 numerical integrations of a network of 8 van der Pol oscillators each described by the equations (6.1). The parameters were chosen to model a series of experiments with the coupling based analog computing devices and the parameters $U_{\text{coup}}^{\text{amp}} = 90\,\text{mV}$, $U_{\text{coup}}^{\text{off}} = 0\,\text{mV}$, $U_{\text{DA}}^{\text{min}} = -375\,\text{mV}$ and $U_{\text{DA}}^{\text{max}} = 375\,\text{mV}$. The choice of initial and memorized patterns was the same as in the experiments. Simulations were started at random initial phases. The graphic shows the zero-flip distribution (black bars, left vertical axis) and the π-flip distribution (white bars, right vertical axis) of $\Delta\varphi(0.05\,\text{s})$. The bin size of the histogram was 0.1. **b)** Shape of the nonlinear characteristic $I_{\text{NL}}(U)$ that was used for the simulations (compare Fig. 15b).
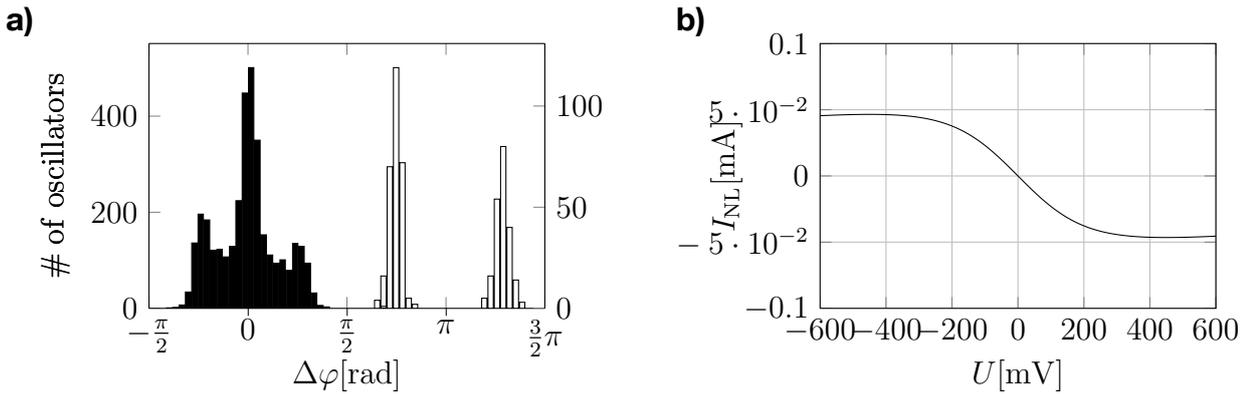
Fig. 61: a) Results of a series of 500 numerical integrations of a network of 8 van der Pol oscillators each described by the equations (6.1). The parameters were the same as for Fig. 60, except for the nonlinear characteristic I_{NL}. The graphic shows the zero-flip distribution (black bars, left vertical axis) and the π-flip distribution (white bars, right vertical axis) of $\Delta\varphi(0.05\,\text{s})$. The bin size of the histogram was 0.1. **b)** Shape of the nonlinear characteristic $I_{\text{NL}}(U) = -5 \cdot 10^{-5}\,\text{A} \cdot \tanh(U \cdot 5\,\text{V}^{-1}) \cdot e^{-U^2/4\,\text{V}^2}$ that was used for the simulations.

6.2 Long term behaviour during pattern recognition

From the experimental data presented in section 6.1, for instance the phase shift evolution depicted in Fig. 52, it can be seen that the system did not settle for a stationary state after the defective oscillator in the initial pattern underwent the phase shift flip, just as expected based on the reasoning in section 4.1. However, the measuring intervals were too short for the movement patterns in phase space predicted by the theory to emerge. Therefore, some longer measurements were performed. Two examples are shown in Fig. 62 and Fig. 63, respectively.

Fig. 62 shows a measurement in which moderate frequency inaccuracies lead to a very regular movement pattern in phase space. After the recognition process (which seems to happen almost instantly on the pictured time scale), the system cyclically followed the pattern sequence $\xi^1 \to \xi^3 \to \xi^2 \to \xi^3 \to \ldots$ in a fashion very similar to Fig. 19.

Fig. 63 shows a measurement with larger fluctuations than Fig. 62, where the transition between patterns was much faster. Also, the phase shift evolution of any single oscillator was somewhat erratic, with stretches of regular behavior in between. Therefore, there did not seem to be a clear periodic sequence of pattern states any more, after the system started out with patterns $\xi^1 \to \xi^3 \to \xi^2 \to \ldots$; still, the network always returned to representations of one of the memorized patterns (for example pattern ξ^2 around $t = 1.8\,\mathrm{s}$). The important point illustrated by Fig. 62 and Fig. 63 is the fact that, while the network is in a state representing the "pattern recognition phase" rather than the "glassy phase" in the sense discussed in 4.3, no single one of the memorized patterns is actually exclusively realized in the network. Rather than being in a state, where all phase shifts are locked, representing a memorized pattern, only subsets of oscillators exhibit a lock in their phase shift values. At certain junctions in phase space (namely at the points where the potential valleys in Fig. 18a meet, i.e. whenever the network represents a memorized pattern), those subsets may change.

Note that, although initial pattern recognition worked in both experiments shown in Fig. 62 and Fig. 63, respectively, this is no necessary condition for the long term pattern switching to occur. For example, in Fig. 53, the process of pattern switching starts at pattern ξ^2 instead of the correctly recognized pattern ξ^1.
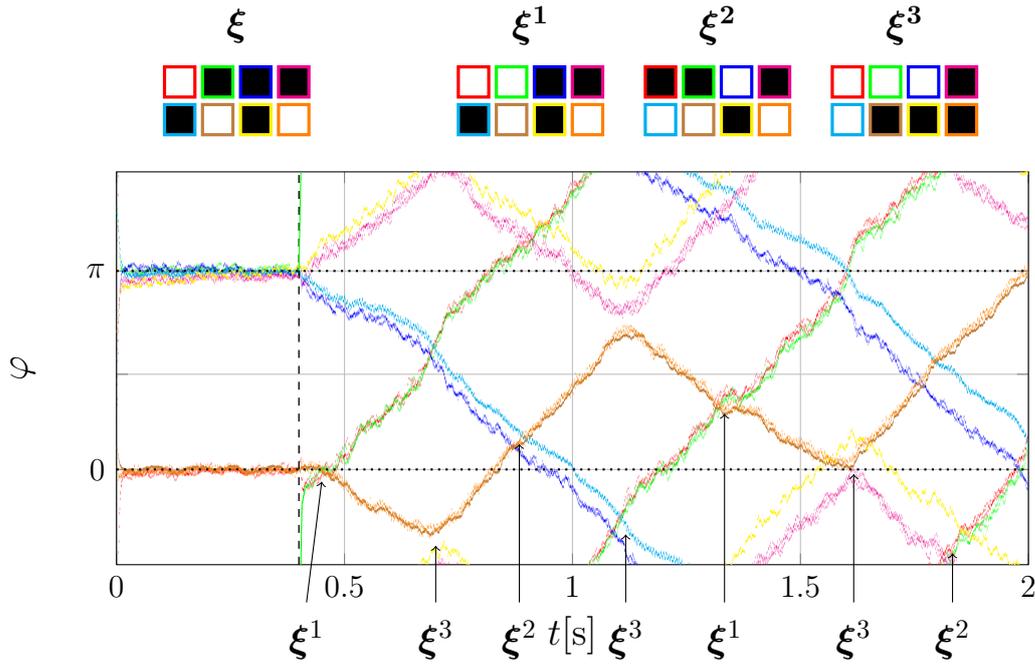
Fig. 62: Time evolution of the phase shifts φ_i during a pattern recognition experiment with the circuit depicted in Fig. 43 using an array of NIC-type oscillators. The time interval of a single run was $T_{\text{record}} = 2\,\text{s}$. During initialization, which in this case took only $0.4\,\text{s}$, a coupling signal with $U_{\text{coup}}^{\text{amp}} = 180\,\text{mV}$ and $U_{\text{coup}}^{\text{off}} = 0\,\text{mV}$ was applied, whereas during recognition $U_{\text{coup}}^{\text{amp}} = 240\,\text{mV}$ and $U_{\text{coup}}^{\text{off}} = 0\,\text{mV}$ were used. $U_{\text{DA}}^{\text{min}} = -375\,\text{mV}$ and $U_{\text{DA}}^{\text{max}} = 375\,\text{mV}$ were the same in both cases. The initial pattern $\boldsymbol{\xi}$ and the memorized patterns $\boldsymbol{\xi}^k$ are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot.
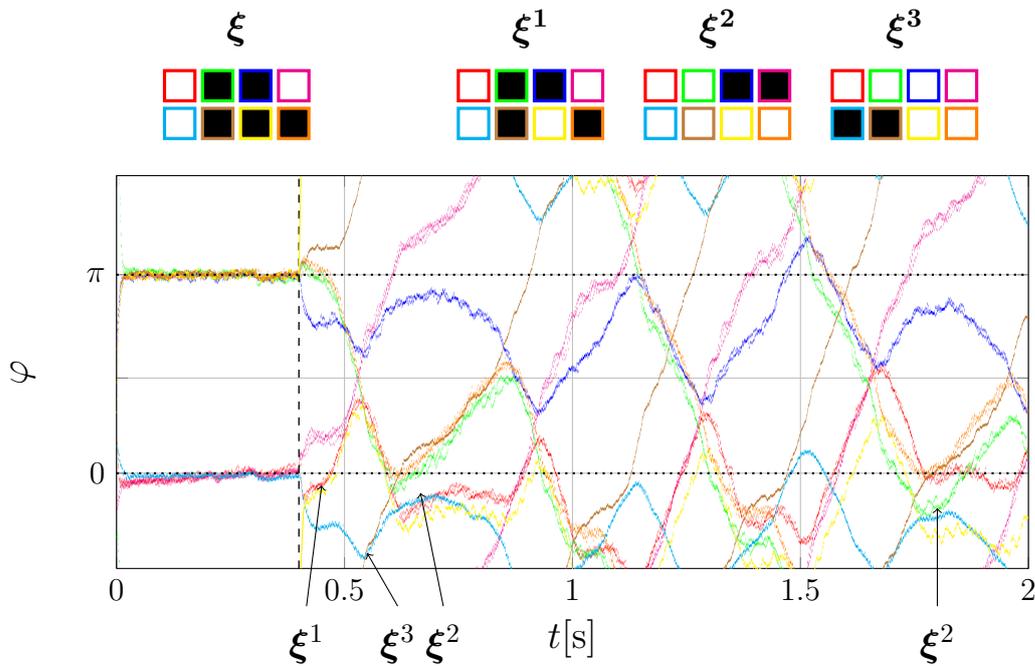
Fig. 63: Time evolution of the phase shifts φ_i during another pattern recognition experiment with the same circuit and parameters as in Fig. 62. The initial pattern $\boldsymbol{\xi}$ and the memorized patterns $\boldsymbol{\xi}^k$ are depicted above. Black squares correspond to $\xi_i = -1$, white squares correspond to $\xi_i = 1$. The color of the border around each square corresponds to the color of the phase shift curve in the plot.

6.3 The stabilizing effect of pattern initialization on frequency drift

It has been pointed out repeatedly that during the initialization phase of the coupling, the dynamics of the phase shifts in the weakly coupled network is equivalent to the dynamics in the classic Kuramoto network, which is a simple description for a set of nonlinear oscillators with a constant, synchronizing coupling.

Systems of noisy oscillators of equal frequency, when subjected to such a coupling, are known to exhibit smaller fluctuations in the coupled, synchronized state than in the uncoupled state (see for example Kori, et al. (2012) and references therein). For small networks with white, Gaussian phase noise, the standard deviation σ_T in the period of a single oscillator decreases with the number N of coupled oscillators according to the square root law
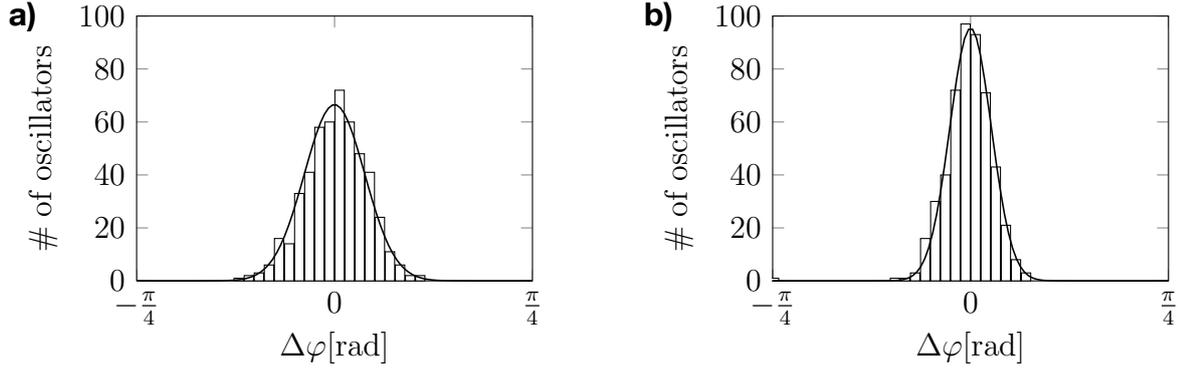
$$\sigma_T = \frac{1}{\sqrt{N}} \cdot \sigma_{T,\text{uncoupled}}$$

The intuitive explanation for this fact is that the coupling works to enforce an equal average value for the period over all oscillators, thus evening out deviations of individual periods in opposite direction.

In the weakly coupled network of this thesis, a similar noise reducing effect is to be expected, even though the oscillators have now frequencies far apart from each other. However, there is an important difference: There are defined fluctuations in the oscillation periods that are brought about by the coupling mechanism, which manifest themselves in small oscillations of the phase shifts φ_i around the value predicted by the average dynamics. This effect causes a variation in oscillation periods that is not related to noise in the system. Since these fluctuations are part of the ideal dynamics of the coupled network, they cancel out in the long time average and will never lead to the accumulation of a substantial deviation of any of the φ_i. Therefore, the accumulated phase shift deviation from the average dynamics in a certain time interval is a good measure for the level of undesired fluctuations present in a particular oscillator.

To measure the fluctuations in the period of the oscillators, the same setup as for the pattern recognition experiments was used, in the version with NIC-type oscillators and coupling through active circuit elements. The measurement consisted of two steps: In the first step, no coupling (i.e. $U_{\text{DA}} = 0$) was applied for a time of $T_{\text{record}} = 1\,\text{s}$. In the second step, a constant positive coupling between all 8 oscillators (i.e. $w_{ij} = 1$) with $U_{\text{coup}}^{\text{amp}} = 90\,\text{mV}$, $U_{\text{coup}}^{\text{off}} = 0\,\text{mV}$, $U_{\text{DA}}^{\text{min}} = -375\,\text{mV}$ and $U_{\text{DA}}^{\text{max}} = 375\,\text{mV}$ was applied for the same time interval. The accumulated deviation $\Delta\varphi$ in the phase shift of each oscillator over the second half of the respective time interval was monitored in both steps.

Fig. 64 shows the results of a series of 500 such measurements, depicting the phase shift deviation accumulated by the $50\,\text{kHz}$ oscillator between $t_1 = 0.5\,\text{s}$ and $t_2 = 0.9\,\text{s}$ both for
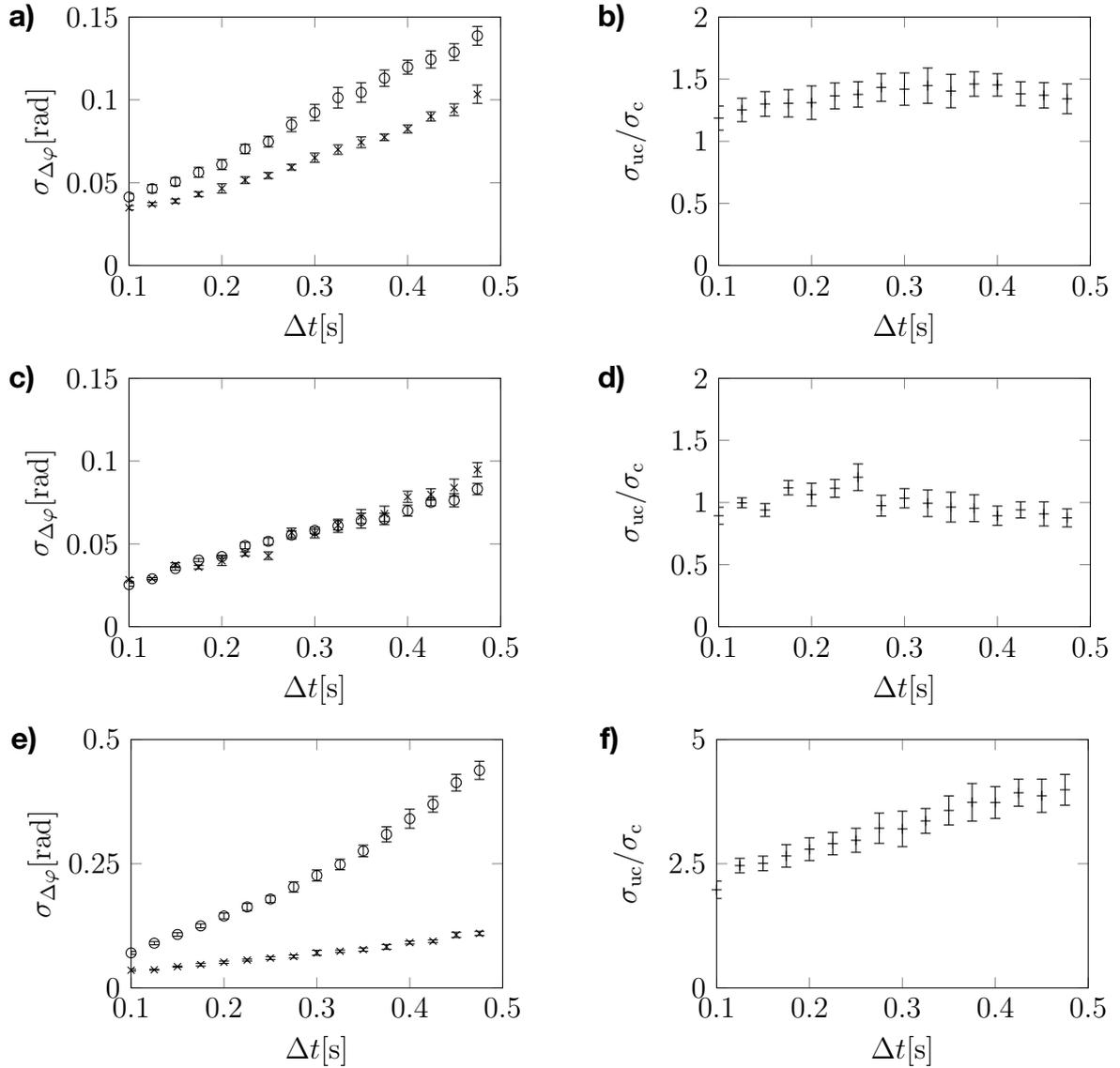
Fig. 64: **a)** Distribution of phase shift deviations accumulated in an uncoupled (i.e. $U_{\text{coup}} = 0$) NIC-type oscillator with a frequency of roughly $50\,\text{kHz}$ without coupling in the time interval of $0.4\,\text{s}$. **b)** Distribution of phase shift deviations accumulated in the same oscillator over the same time interval in the 8-oscillator network with completely synchronizing coupling (i.e. $w_{ij} = 1$ and $U_{\text{coup}}^{\text{amp}} = 90\,\text{mV}$, $U_{\text{coup}}^{\text{off}} = 0\,\text{mV}$, $U_{\text{DA}}^{\text{min}} = -375\,\text{mV}$ and $U_{\text{DA}}^{\text{max}} = 375\,\text{mV}$). Each histogram shows the data of 500 measurements. The interval represented by a single column is 0.04. The solid curves are Gaussian fits to the data.

the uncoupled (Fig. 64a) and the coupled case (Fig. 64b). Both histograms show a normal distribution; for the uncoupled case, the distribution is almost 1.5 times wider than for the coupled case.

Fig. 65a shows the evolution in the standard deviation of both distributions over time, Fig. 65b shows the ratio of the two standard deviations. After an initial increase, the this value remains nearly constant around 1.4. The initial increase is expected, because for very low times, the accumulated deviation in phase shift is mainly determined by the inaccuracy of the frequency determination algorithm, which is equal for the coupled and uncoupled case.

The effect of the coupling on oscillation fluctuations seems to vary greatly between oscillators. For example, the fluctuations of the second slowest oscillator ($36\,\text{kHz}$) remain almost unaffected (see Fig. 65c,d), while the fluctuations of the oscillator at $57\,\text{kHz}$ are reduced by a factor of up to 3.5 (see Fig. 65e,f). For the other oscillators, the ratio was in between. Interestingly, the strength of the effect does not seem to be perfectly correlated with the frequency; from the measurements one can only conclude that the coupling has a certain stabilizing effect on every oscillator (with the exception of the $36\,\text{kHz}$ oscillator). A precise theoretical derivation of the expected effects, however, is yet to be formulated.

Fig. 65: a), c), e) Evolution of the standard deviation $\sigma_{\Delta\varphi}$ of phase shift deviations plotted over the time interval Δt in which the deviations were accumulated for the uncoupled $50\,\text{kHz}$ / $36\,\text{kHz}$ / $57\,\text{kHz}$ NIC-type oscillator (circles) and for the same oscillator in the 8-oscillator network with completely synchronizing coupling (i.e. $w_{ij} = 1$ and $U_{\text{coup}}^{\text{amp}} = 90\,\text{mV}$, $U_{\text{coup}}^{\text{off}} = 0\,\text{mV}$, $U_{\text{DA}}^{\text{min}} = -375\,\text{mV}$ and $U_{\text{DA}}^{\text{max}} = 375\,\text{mV}$) (crosses). Each circle or cross corresponds to a histogram similar to Fig. 64a and Fig. 64b respectively; the error bars show the error of the fitting algorithm. **b), d), f)** Ratio of both standard deviations σ_{uc} (uncoupled oscillator) and σ_{c} (coupled oscillator) over the time interval for the $50\,\text{kHz}$ / $36\,\text{kHz}$ / $57\,\text{kHz}$ oscillators.

7. Summary

In the course of the work on this thesis, a mathematical model for a neural network of coupled phase oscillators proposed by Hoppensteadt & Izhikevich (1999) was theoretically investigated for its use in real-life applications. Based on the results of this investigation, a prototype network of 8 electronic oscillators was designed and built, successfully proving the feasibility of the mathematical concept. The task performed by this neural network is the recognition of a defective pattern as one of a set of given memorized patterns. The unique property of the model is that the synaptic coupling between any pair of neurons is dynamically governed by a single external coupling parameter with a complex time-dependence, rather than by a large set of static individual coupling parameters. This is made possible by subjecting a set of Kuramoto oscillators to a weak, time-dependent coupling. If the oscillators and the coupling function are suitably chosen, the average dynamics is equivalent to that of a set of strongly coupled Kuramoto oscillators with equal frequencies, which is known to be capable of pattern recognition, if the coupling between pairs of oscillators is chosen according to a Hebbian learning rule (Arenas & Pérez-Vicente 1994). Note that "weak coupling" here denotes a coupling strength far below the synchronization threshold, while "strong coupling" means that the coupling strength is above this threshold but still small enough to allow a phase-only description.

In the theoretical part of the thesis, the oscillatory network was analyzed on three different levels:

First, the stability of the stationary states corresponding to the recognized patterns of a Hebbian network of strongly coupled Kuramoto oscillators with equal frequency was examined. For the case of mutually orthogonal learned patterns (i.e. patterns that have maximum distance in pattern space), it was proven that the solutions corresponding to the memorized patterns are part of a larger set of degenerate steady states. Numerical evidence was found that suggests there is a degenerate attractor in the general case, too. The consequence of these findings is that pattern recognition in the neural network at hand is only transient in the presence of noise, giving way to a more complicated long term behavior. The different types of possible long term behavior in the presence of noise were categorized.

Second, to quantify the effects of noise, a series of numerical simulations was conducted to determine the noise threshold at which reliable pattern recognition breaks down. This was done for the phase description of the weakly coupled network with the global time dependent coupling. The main result was that frequency inaccuracies of the oscillators as well as the time-dependent coupling function are the most detrimental type of noise, because the effect of frequency inaccuracies increases quadratically with the number of oscillators.

Third, a phase equation mirroring the experimental dynamics, based on the assumption of experimental oscillators with perfectly sinusoidal waveforms and phase response curves, was derived and it was shown that the average dynamics on the relevant slow time scale equals that of the original weakly coupled phase model. On an even slower time scale, however, there is an additional common drift that is introduced by using realistic oscillators instead of phase oscillators. Using the multiple time scale analysis of this more realistic phase equation, together with the results on the effects of noise and some general considerations about the efficiency of the experimental setup, a rule of thumb was derived which relates the number of oscillators to the necessary frequency accuracy needed for a network of this size.

For the experiment, two different electrical oscillators, both of the van der Pol type (Kanamaru 2007), were used. The parameter space for both types of oscillator was explored in order to find the region with sinusoidal waveform and phase response curve. Also, two different mechanisms for implementing a weak time-dependent global coupling in a network of electrical oscillators were developed. The main challenge was to find a way to realize a seamless transition between positive and negative values of the coupling. To achieve this, the first mechanism uses a voltage controlled resistor in conjunction with two negative impedance converters. The second mechanism employs several active electronic devices like adders and multipliers to model the dynamical equation directly. Relating the phase model based on ideal waveform and phase response curve mentioned above to a realistic two-variable description of the oscillators, the coupling strength in the mathematical model was expressed as a function of experimental parameters. Also, the process of pattern recognition in the electric circuit could be related to the tendency of the network to find a state of minimal entropy production.

To test both coupling mechanisms, two networks with 8 van der Pol oscillators each were built, with different designs for the individual oscillators. Using the two networks in combination with the two coupling variants, several series of experiments were made in which the network recognized randomized 8 bit patterns with one defect out of three given mutually orthogonal patterns. The best experimental configuration was a network featuring NIC-circuits as nonlinear elements in the van der Pol oscillators, gyrators as inductors and a coupling through active electronic devices. Its success rate was 100% (i.e. 500 in a series of 500 pattern recognition experiments were successful). Other configurations had success rates of up to 80% and up to 95%, respectively. In the experiments, a symmetry breaking was observed. For a successful pattern recognition, oscillators corresponding to incorrect bits in the tested pattern must change their phase shift by π by either speeding up or slowing down. In the experiments, especially for a network with low noise level and low distortion of the coupling function, the speeding up of these oscillators was heavily favored. This effect cannot be observed in a pure phase model of the network and has yet

to be fully understood. However, some numerical simulations revealed that the strength of this effect depends on the shape of the nonlinearity of the van der Pol oscillator.

Additionally to the pattern recognition experiments, the long term behavior predicted by the theoretical analysis was observed after the transient pattern recognition and a positive effect of the weak coupling on the regularity of oscillations was demonstrated.

With these experiments, in particular the series with 100% successful pattern recognition, it has been established that the original mathematical model can be implemented successfully as an electronic circuit. To the knowledge of the author, this was also the first time that the phase shifts of coupled heterogeneous oscillators were selectively and accurately manipulated in an experiment with a very weak coupling far below the Kuramoto threshold, even without the pattern recognition aspect.

Based on this foundation, further research can now concentrate on improving on the disadvantages of the approach. In particular, it is unlikely that the network will be used in a real world application in its current form for the following reason: Due to the fact that the coupling strength scales inversely with the square of the number of oscillators, the convergence time of the network increases quadratically with the number of oscillators. This simply means that there is no computational speed gained by the parallel architecture compared to a sequential computer, even for arbitrarily large network sizes. Still, the weak global coupling approach for neural networks can be useful for applications when combined with more conventional ways of neural coupling. In such systems, there is a trade-off between the complexity of the coupling topology and the execution speed of the network. First, promising steps in this direction have already been made by Kostorz, et al. (2013).

Another intriguing direction of future research is the investigation of more modular, clustered network designs, which suffer less of the restrictions of an all-to-all coupled network. Finally, it seems promising to study networks in which individual oscillators act both as neurons and as generators of the time-dependent coupling function. Systems adhering to this concept seem closer to the actual inner workings of the human brain; in particular they would be self-contained, not depending on an external global coupling function.

References

Abbott, L. F. (1990): A network of oscillators, Journal of Physics A: Mathematical and General **23**(16): 3835.

Acebrón, Juan A., Bonilla, L. L., Pérez Vicente, Conrad J. & Ritort, Félix (2005): The Kuramoto model: A simple paradigm for synchronization phenomena, Reviews of Modern Physics **77**(1): 137.

Amit, Daniel J., Gutfreund, Hanoch & Sompolinsky, H. (1985): Storing infinite numbers of patterns in a spin-glass model of neural networks., Physical Review Letters **55**: 1530.

Aonishi, Toru (1998): Phase transitions of an oscillator neural network with a standard Hebb learning rule., Physical Review E **58**: 4865.

Aonishi, Toru, Kurata, Koji & Okada, Masato (1999): Statistical mechanics of an oscillator associative memory with scattered natural frequencies, Physical review letters **82**(13): 2800.

Aonishi, Toru, Kurata, Koji & Okada, Masato (2002): Acceleration effect of coupled oscillator systems, Physical Review E **65**(4): 046223.

Aoyagi, Toshio & Kitano, Katsunori (1997): Effect of random synaptic dilution in oscillator neural networks, Physical Review E **55**(6): 7424.

Aoyagi, Toshio & Kitano, Katsunori (1998): Retrieval dynamics in oscillator neural networks, Neural Computation **10**(6): 1527.

Apostolos, Dimitromanolakis (2002): Analysis of the Golomb ruler and the Sidon set problems and determination of large, near-optimal Golomb rulers., Diploma thesis, Technical University of Crete: Retrieved on June 26, 2012, from `http://www.cs.toronto.edu/~apostol/golomb/`.

Arenas, Alex, Díaz-Guilera, Albert, Kurths, Jürgen, Moreno, Yamir & Zhou, Changsong" (2008): Synchronization in complex networks, Physics Reports **469**(3): 93.

Arenas, Alex & Pérez-Vicente, Conrad J. (1994): Phase locking in a network of neural oscillators, EPL (Europhysics Letters) **26**: 79.

Backus, John (1978): Can programming be liberated from the von Neumann style?: a functional style and its algebra of programs, Commun. ACM **21**(8): 613.

Baldi, Pierre & Meir, Ronny (1990): Computing with arrays of coupled oscillators: An application to preattentive texture discrimination, Neural Computation **2**(4): 458.

Chua, Leon (1971): Memristor-the missing circuit element, Circuit Theory, IEEE Transactions on **18**(5): 507.

Cohen, Avis H., Holmes, Philip J. & Rand, Richard H. (1982): The nature of the coupling between segmental oscillators of the lamprey spinal generator for locomotion: A mathematical model, Journal of Mathematical Biology **13**: 345.

Cook, J (1989): The mean-field theory of a q-state neural network model, Journal of Physics A - Mathematical and General **22**(12): 2057.

Daido, Hiroaki (1987): Population dynamics of randomly interacting self-oscillators. i, Progress of Theoretical Physics **77**(3): 622.

Daido, Hiroaki (1992): Quasientrainment and slow relaxation in a population of oscillators with random and frustrated interactions, Physical Review Letters **68**: 1073.

Drakakis, Konstantinos (2009): A review of the available construction methods for Golomb rulers, Advances in Mathematics of Communications **3**(3): 235.

Esaki, Leo (1958): New phenomenon in narrow germanium $p - n$ junctions, Physical Review **109**: 603.

Ferrucci, David, Brown, Eric, Chu-Carroll, Jennifer, Fan, James, Gondek, David, Kalyanpur, Aditya A., Lally, Adam, Murdock, J.William, Nyberg, Eric, Prager, John, Schlaefer, Nico & Welty, Chris (2010): Building Watson: An overview of the DeepQA project, AI Magazine **31**(3): 59.

Golomb, Samuel (1997): The use of combinatorial structures in communication signal design, in C. Mitchell (ed.), Applications of combinatorial mathematics, Clarendon Press, S. 59.

Guckenheimer, John & J., Holmes Philip (1983): Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Springer.

Hebb, Donald O. (1949): The organization of behavior, Wiley & Sons.

Hemmen, J. Leo van & Wreszinski, Walter F. (1993): Lyapunov function for the Kuramoto model of nonlinearly coupled oscillators, Journal of Statistical Physics **72**(1): 145.

Holler, Mark A., Tam, Simon M., Castro, Hernan A. & Benson, Ronald. G. (1989): An electrically trainable artificial neural network (etann) with 10240 'floating gate' synapses, Neural Networks, 1989. IJCNN., International Joint Conference on, IEEE, S. 191.

Hölzel, Robert W. (2007): Mustererkennung in einem System global gekoppelter elektrochemischer Oszillatoren, Diploma thesis, Technische Universität München.

Hölzel, Robert W. & Krischer, Katharina (2011): Pattern recognition with simple oscillating circuits, New Journal of Physics **13**(7): 073031.

Hong, H., Um, TI, Shim, Y. & Choi, M. Y. (2001): Temporal association in a network of neuronal oscillators, Journal of Physics A: Mathematical and General **34**(24): 5021.

Hopfield, John J. (1982): Neural networks and physical systems with emergent collective computational abilities., P. natl. acad. sci. USA **79**: 2554.

Hoppensteadt, Frank C. & Izhikevich, Eugene M. (1999): Oscillatory neurocomputers with dynamic connectivity., Physical Review Letters **82**: 2983.

Horowitz, P. & Hill, W. (1989): The Art of Electronics, Cambridge University Press.

Kanamaru, Takashi (2007): Van der Pol oscillator, Scholarpedia **2**(1): 2202.

Kiss, István Z., Wang, Wen & Hudson, John L. (1999): Experiments on arrays of globally coupled periodic electrochemical oscillators, The Journal of Physical Chemistry B **103**(51): 11433.

Kiss, István Z., Zhai, Yumei. & Hudson, John L. (2002): Emerging coherence in a population of chemical oscillators, Science **296**(5573): 1676.

Kiss, István Z., Zhai, Yumei & Hudson, John L. (2005): Predicting mutual entrainment of oscillators with experiment-based phase models, Physical Review Letters **94**(24): 248301.

Knutson, Allen & Tao, Terence (2001): Honeycombs and sums of hermitian matrices, Notices of the American Mathematical Society **48**: 171.

Kopell, Nancy & Ermentrout, George B. (1988): Coupled oscillators and the design of central pattern generators, Mathematical biosciences **90**(1): 87.

Kori, Hiroshi, Kawamura, Yoji & Masuda, Naoki (2012): Structure of cell networks critically determines oscillation regularity, Journal of Theoretical Biology **297**(0): 61.

Kostorz, Kathrin, Hölzel, Robert W. & Krischer, Katharina (2013): A novel type of a weakly coupled oscillatory network with associative properties, submitted .

Kowalski, Jacek M., Ansari, Ali, Prueitt, Paul S., Dawes, Robert L. & Gross, Gunther (1988): On synchronization and phase locking in strongly coupled systems of planar rotators, Complex Systems **2**(4): 441.

Krischer, Katharina, Varela, Hamilton, Bîrzu, Adrian, Plenge, Florian & Bonnefont, Antoine (2003): Stability of uniform electrode states in the presence of ohmic drop compensation, Electrochimica Acta **49**: 103.

Kuramoto, Yoshiki (1975): Self-entrainment of a population of coupled non-linear oscillators, Mathematical Problems in Theoretical Physics, Bd. 39, S. 420.

Kuramoto, Yoshiki (1984): Chemical oscillations, waves and turbulence., Springer.

Lohr, Steve (2011): I.B.M announces brainy computer chip, The New York Times .

Marquardt, Donald W. (1963): An algorithm for least-squares estimation of nonlinear parameters, Journal of the Society for Industrial and Applied Mathematics **11**(2): 431.

Matsumura, Kosuke, Nagai, Takahiro, Uwate, Yohiro & Nishio, Yoshifumi (2012): Analysis of synchronization phenomenon in coupled oscillator chains, Circuits and Systems (ISCAS), 2012 IEEE International Symposium on, IEEE, S. 620.

Mazouz, Nadia, Flätgen, Georg & Krischer, Katharina (1997): Tuning the range of spatial coupling in electrochemical systems: From local via nonlocal to global coupling, Physical Review E **55**(3): 2260.

McCulloch, Warren S. & Pitts, Walter (1943): A logical calculus of the ideas immanent in nervous activity, Bulletin of mathematical biology **5**(4): 115.

Mertens, David & Weaver, Richard (2011): Individual and collective behavior of vibrating motors interacting through a resonant plate, Complexity **16**(5): 45.

Mikhailov, Alexander S., Zanette, D.H., Zhai, Yumei M., Kiss, István Z. & Hudson, John L. (2004): Cooperative action of coherent groups in broadly heterogeneous populations of interacting chemical oscillators, Proceedings of the National Academy of Sciences of the United States of America **101**(30): 10890.

Nagel, Laurence W. & Pederson, Donald O. (1973): SPICE (Simulation Program with Integrated Circuit Emphasis), Technical Report UCB/ERL M382, EECS Department, University of California, Berkeley.

Neu, John C. (1980): Large populations of coupled chemical oscillators, SIAM Journal on Applied Mathematics **38**(2): 305.

Neumann, John von (1945): First draft of a report on the EDVAC, Technical report: reprinted in the Ann. Hist. Comput. **15**: 25 (1993).

Niebur, Ernst, Schuster, Heinz G., Kammen, Daniel M. & Koch, Christof (1991): Oscillator-phase coupling for different two-dimensional network connectivities, Physical Review A **44**(10): 6895.

Nishikawa, Takashi, Lai, Ying-Cheng & Hoppensteadt, Frank C. (2004): Capacity of oscillatory associative-memory networks with error-free retrieval, Physical review letters **92**(10): 108101.

Nobre, F. D. & Sherrington, D. (1986): The infinite-range clock spin glass model: an investigation of the relevance of reflection symmetry, Journal of Physics C: Solid State Physics **19**(8): L181.

Park, Kibeom & Choi, M. Y. (1995): Synchronization in a network of neuronal oscillators with finite storage capacity, Physical Review E **52**: 2907.

Pérez-Vicente, Conrad J., Arenas, Alex & Bonilla, L. L. (1996): On the short-time dynamics of networks of Hebbian coupled oscillators, Journal of Physics A: Mathematical, Nuclear and General **29**(1): 9–16.

Pikovsky, Arkady, Rosenblum, Michael & Kurths, Jürgen (2001): Synchronization: A Universal Concept in Nonlinear Sciences, Cambridge University Press.

Prigogine, Ilya (1961): Introduction to Thermodynamics of Irreversible Processes, Second Edition, Interscience Wiley.

Radhakrishnan, Krishnan & Hindmarsh, Alan C. (1993): Description and use of LSODE, the Livermore Solver for Ordinary Differential Equations, Technical report.

Rojas, Raul (1996): Neural Networks., Springer.

Sakaguchi, Hidetsugu, Shinomoto, Shigeru & Kuramoto, Yoshiki (1987): Local and global self-entrainments in oscillator lattices, Prog. Theor. Phys **77**(5): 1005.

Schmidhuber, Jürgen, Cireşan, Dan, Meier, Ueli, Masci, Jonathan & Graves, Alex (2011): On fast deep nets for AGI vision, Artificial General Intelligence S. 243.

Senani, R., Bhaskar, D. R., Gupta, S. S. & Singh, V. K. (2009): A configuration for realizing floating, linear, voltage-controlled resistance, inductance and FDNC elements, International Journal of Circuit Theory and Applications **37**(5): 709.

Shim, Y., Hong, H. & Choi, M. Y. (2002): Noise-enhanced temporal association in neural networks, Physical Review E **65**(3): 036114.

Shmilovitz, Doron (2005): On the definition of total harmonic distortion and its effect on measurement interpretation, Power Delivery, IEEE Transactions on **20**(1): 526.

Shockley, William B. (1949): The theory of $p-n$ junctions in semiconductors and $p-n$ junction transistors, Bell. Syst. Tech. J. **28**: 435.

Sompolinsky, H., Golomb, D. & Kleinfeld, D. (1991): Cooperative dynamics in visual processing, Physical Review A **43**(12): 6990.

Strogatz, Stephen H. (1994): Nonlinear Dynamics and Chaos., Westview Press.

Strogatz, Stephen H. (2000): From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators, Physica D **143**: 1.

Strogatz, Steven H. & Mirollo, Renato E. (1988): Phase-locking and critical phenomena in lattices of coupled nonlinear oscillators with random intrinsic frequencies, Physica D: Nonlinear Phenomena **31**(2): 143.

Strukov, Dmitri B., Snider, Gregory S., Stewart, Duncan R. & Williams, R. Stanley (2008): The missing memristor found, Nature **453**(7191): 80.

Sutter, Herb (2005): The free lunch is over : A fundamental turn toward concurrency in software, Dr. Dobb's Journal **30**(3): 202.

Taylor, Annette F., Kapetanopoulos, Panagiotis, Whitaker, Benjamin J., Toth, Rita, Bull, Larry & Tinsley, Mark R. (2008): Clusters and switchers in globally coupled photo-chemical oscillators, Physical review letters **100**(21): 214101.

Temirbayev, Amirkhan A., Zhanabaev, Zeinulla Z., Tarasov, Stanislav B., Ponomarenko, Vladimir I. & Rosenblum, Michael (2012): Experiments on oscillator ensembles with global nonlinear coupling, Physical Review E **85**(1): 015204.

Tietze, Ulrich., Schenk, Christoph & Gamm, Eberhard (2008): Electronic circuits: Handbook for design and application, Springer.

Yamana, Michiko, Shiino, Masatoshi & Yoshioka, Masahiko (1999): Oscillator neural network model with distributed native frequencies, Journal of Physics A: Mathematical and General **32**: 3525.

Yoshioka, Masahiko & Shiino, Masatoshi (2000): Associative memory storing an extensive number of patterns based on a network of oscillators with distributed natural frequencies in the presence of external white noise, Physical Review E **61**(5): 4732.

A Appendix

A1   Table of repeatedly used symbols and abbreviations

- abbreviations

    FWHM full width at half maximum

    JFET junction field effect transistor

    NIC negative impedance converter

    op-amp operational amplifier

    TD tunnel diode

    VCR voltage controlled resistor

- capital Latin letters

    $E(\varphi)$ potential function of the strongly coupled network

    $\boldsymbol{J}, J_{ij}$ Jacobian of the strongly coupled network, entries of the Jacobian

    M number of memorized patterns for a given network

    N number of oscillators in a given network

    P proximity: measure for the closeness of a state φ to a pattern $\boldsymbol{\xi}$

    R_{int} internal resistance, connecting each oscillator to the coupling node

    T_{record} length of the measuring time interval for a single pattern recognition experiment

    $U_{\mathrm{coup}}^{\mathrm{amp}}$ amplitude of a single frequency component in the coupling voltage $U_{\mathrm{DA}}(t)$

    $U_{\mathrm{coup}}^{\mathrm{off}}$ amplitude of a single frequency component in the coupling voltage $U_{\mathrm{DA}}(t)$

    $U_{\mathrm{DA}}(t)$ output of the DA converter used to provide the coupling signal in the experiment

    $U_{\mathrm{DA}}^{\mathrm{min}}$ lower cutoff value of U_{DA}

    $U_{\mathrm{DA}}^{\mathrm{max}}$ upper cutoff value of U_{DA}

    $Y(\vartheta)$ waveform of an oscillator

    $Z(\vartheta)$ phase response curve of an oscillator

- small Latin letters

    $a(t)$ functional dependence of the time-dependent coupling a weakly coupled network

$f(t), g(t), h(t)$ time dependent expressions in the order of magnitude of 1

i, j, p, q summation index that usually runs from 1 to N, indicates the number of an oscillator

k, l, m summation index that usually runs from 1 to M, indicates the number of a memorized pattern

m overlap: measure for the closeness of a state φ to a pattern $\boldsymbol{\xi}$

w_{ij} entries of the coupling matrix in the strongly coupled network, coefficients of the coupling function in a weakly coupled network.

- capital Greek letters

  Ω undisturbed frequency of an oscillator; frequency of a noise signal

  $\Omega_{\min}, \Omega_{\max}$ bounds of the frequeny interval used by the network

  $\Delta\Omega_i$ deviation of a single oscillator from its undisturbed frequency

  $\Delta\Omega$ all frequency deviations $\Delta\Omega_i$ fall within the interval $[-\Delta\Omega/2, \Delta\Omega/2]$; also called absolute frequency accuracy in the thesis.

- small Greek letters

  α load rate of the network; $\alpha = M/N$

  δ minimal difference of two oscillator frequencies in a weakly coupled network

  δ' minimal difference of the frequency differences of two pairs of oscillators in a weakly coupled network

  ε coupling strength in the ideal weakly coupled network; coupling strength before averaging in the realistic weakly coupled network

  $\varepsilon_{\mathrm{eff}}$ effective coupling strength after averaging in the realistic weakly coupled network; $\varepsilon_{\mathrm{eff}} = \varepsilon/2$

  $\eta(t)$ small, undesired perturbation of the dynamics of the realistic weakly coupled network

  $\vartheta(t)$ phase of an oscillator

  λ eigenvalue of $\boldsymbol{J}$

  μ size of an undesired perturbation of the dynamics of the realistic weakly coupled network; $\eta(t) = \mu h(t)$

  μ_{eff} effective size of an undesired perturbation of the dynamics of the realistic weakly coupled network after averaging

  $\boldsymbol{\xi}$ initialized pattern vector; N entries $\xi_i \in [-1, 1]$

$\boldsymbol{\xi}^k$ memorized pattern vector

$\boldsymbol{\xi}_{\text{final}}$ pattern vector represented (in general only approximately) by the network when the recognized pattern is read out

$|\xi\rangle$ state vector in complex pattern space; N entries $|\xi\rangle_i = e^{\mathrm{i}\varphi_i}$

$\varphi(t)$ phase of an oscillator

$\boldsymbol{\varphi}(t)$ state vector of the network in phase space; N entries $\varphi_i(t)$

$\boldsymbol{\varphi}^*$ steady state in the phase space of the strongly coupled network

$\boldsymbol{\varphi}^{*k}$ steady state in the phase space of the strongly coupled network approximately representing $\boldsymbol{\xi}^k$

$\boldsymbol{\varphi}_{\text{initial}}$ state vector of the network when the initialization has ended

$\boldsymbol{\varphi}_{\text{final}}$ state vector of the network when the recognized pattern is read out

$\boldsymbol{\varphi}_{\text{perfect}}$ state vector of the network after a hypothetical perfect recognition

ω_i deviation of a single oscillator from its undisturbed frequency; similar to $\Delta\Omega_i$, however $\sum_i \omega_i = 0$

A2   Numerical integration example

To run the example, create a folder containing the files `opkda1.f`, `opkda2.f` and `opkdmain.f` obtained from `http://www.netlib.org/odepack` as well as a the file `integrate.cpp` containing the following program:

```
using namespace std;

#include <iostream>
#include <fstream>
#include <vector>
#include <cmath>

#define N_OSZI 16

#define T_END 100.
#define N_SAMPLES 500
#define DT (T_END/N_SAMPLES)
#define T_SWITCH (T_END/4)

#define N_PATTERN 3

double pattern[][N_OSZI]={
{ 1,-1,1,-1,-1,-1,-1,1,1,-1,1,1,-1,-1,-1,1},
{ 1,1,1,1,-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1},
{ 1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1},
{ 1,-1,1,-1,-1,1,-1,1,1,-1,1,-1,-1,1,-1,1}
};
```

```
double initCoeffs[N_OSZI*N_OSZI];
double recCoeffs[N_OSZI*N_OSZI];

// variables and arrays required by LSODE
int NEQ_ = N_OSZI;
double T_;
double Y_[N_OSZI];
int ITOL_ =1;
double RTOL_ = 1e-7;
double ATOL_ = 1e-7;
int ITASK_ = 1;
int ISTATE_ = 1;
int IOPT_ = 0;
int LRW_=22 + 9*N_OSZI + N_OSZI*N_OSZI;
double RWORK_[22 + 9*N_OSZI + N_OSZI*N_OSZI];
int LIW_=20+ NEQ_;
int IWORK_[20+ N_OSZI];
int MF_=22; // stiff method with internally generated Jacobian
int MU_=0;
int ML_=0;

// wrapper for fortran ode solving routine
extern "C" void dlsode_(void F(int*,double*,double*,double*), int* NEQ,\
        double* Y, double * T , double *TOUT,int* ITOL, double *RTOL,\
        double *ATOL, int* ITASK, int* ISTATE,int* IOPT,double* RWORK,\
        int* LRW,int* IWORK, int* LIW, \
        void JAC(int*,double*,double*,int*,int*,double*,int*),int* MF);\

// right hand side of the differential equation, is passed to the solver
void func(int* NEQ, double* T, double* Y, double* Ydot);

// fills coupling matrix array with w_ij
void compute_coefficients()
{
        for(int i = 0;i<N_OSZI;++i)
        {
            for(int j = 0;j<N_OSZI;++j)
            {
                initCoeffs[i*N_OSZI+j]=pattern[0][i]*pattern[0][j];
                recCoeffs[i*N_OSZI+j]=0;
                for(int k=1;k<=N_PATTERN;++k)
                recCoeffs[i*N_OSZI+j]+=pattern[k][i]*pattern[k][j];
            }
        }
}

int main(int argc, char** args)
{
        compute_coefficients();

        // set initial conditions
        T_=0;
        srand(time(NULL));
        for(int i=0;i<NEQ_;++i)
            Y_[i] = (double)rand()/(double)RAND_MAX*2*M_PI-0.5*M_PI;
```

```
        int percentStep=T_END/DT/100.;
        double dt = DT;

        // data structures that store the results
        vector<double> time;
        vector<vector<double> > phases(N_OSZI);

        // integration
        int j = 0;
        for(double t = 0;t<=T_END;t+=dt)
        {
                dlsode_(func, &NEQ_, Y_, &T_ , &t ,&ITOL_,&RTOL_, &ATOL_,&ITASK_,\
                        &ISTATE_,&IOPT_,RWORK_,&LRW_,IWORK_,&LIW_,NULL,&MF_);
                time.push_back(T_);
                for(int i = 0;i<N_OSZI;++i)
                        phases[i].push_back(Y_[i]);

                if(j%percentStep==0)
                        cerr<<"\b \b \b \b"<<j/percentStep<<"%";
                ++j;
        }

        // write out results "to phases.dat",
        fstream file;
        file.precision(15);
        file.open ("phases.dat", ios::out);

        // but before, perform a global phase shift
        // such that branches are at 0 and pi respectively
        double shift = phases[0][N_SAMPLES/(T_END/T_SWITCH)]-M_PI*(pattern[0][0]+1)/2;
        for(int i=0;i<time.size();++i)
        {
                file<<time[i]<<"\t";
                // and restrict the values of phi to [-0.5*pi,1.5*pi[
                for(int j=0;j<N_OSZI;++j)
                {
                        double d = phases[j][i]-shift;
                        while(d>1.5*M_PI)
                                d-=2*M_PI;
                        while(d<-.5*M_PI)
                                d+=2*M_PI;

                        file<<d<<"\t";
                }
                file<<endl;
        }

        file.close();
                cerr<<"\b\b \b \b \b100%"<<endl;

        return 0;
}

// right hand side of the differential equation, is passed to the solver
void func(int* NEQ, double* T, double* Y, double* Ydot)
{
```

```
    if(*T<T_SWITCH)
    {
        for(int i = 0;i<N_OSZI;++i)
        {
            Ydot[i]=0;
            for(int j = 0;j<N_OSZI;++j)
            {
                Ydot[i]+=initCoeffs[i*N_OSZI+j]*sin(Y[j]-Y[i]);
            }
            Ydot[i]/=(double)N_OSZI;
        }
    }
    else
    {
        for(int i = 0;i<N_OSZI;++i)
        {
            Ydot[i]=0;
            for(int j = 0;j<N_OSZI;++j)
            {
                // .999999:  slight distortion to move the system
                // away from the unstable equilibrium
                Ydot[i]+=recCoeffs[i*N_OSZI+j]*sin(Y[j]-Y[i]*.999999);
            }
            Ydot[i]/=(double)N_OSZI;
        }
    }
}
```

At a shell prompt change the directory to the folder; then compile and link the program with

```
gfortran opkda1.o opkda2.o opkdmain.o -c
g++ opkda1.o opkda2.o opkdmain.o integrate.cpp -o integrate -lgfortran
```

and run it with

```
./integrate
```

or just

```
integrate
```

depending on your shell. The output file `phases.dat` should then contain the data for Fig. 1.

A3   Netlists for LTSpiceIV

To run spice with one of the netlists below create a text file (e.g `VDP.cir`) containing one of the netlists below. Now go to the directory containing the file at the command prompt and type

```
scad3 -b VDP.cir -ascii
```

which starts LTspiceIV in batch mode, performs a transient analysis of the circuit and stores a time series of the oscillator signal in the file `VDP.raw` in ASCII format. The starting and stopping time of the series as well as the time step are set in the `.tran` line. The parameters of the oscillator itself are set in the `.param` lines.

A3.1 Van der Pol oscillator with tunnel diode

```
*********************************************
.param C={1.59155e-011}
.param L={0.159155}
.param RINT={500}
.param U={0.3}
.tran 0 0.0021 0.0011 1e-007
.print tran v(v_osc)
.ic v(v_osc)=0.000001
*********************************************
XU1 P001 0 40571
V1 P001 v_osc {U}
C1 v_osc 0 {C}
L1 v_osc 0 {L}
R1 v_osc 0 {RINT}
* tunnel diode 40571
.subckt 40571 plus minus
G1 plus minus value={(V(plus)-V(minus))*.143461-(V(plus)-V(minus))**2*1.482+(V(plus)
+ -V(minus))**3*6.13939-(V(plus)-V(minus))**4*11.7777+(V(plus)-V(minus))**5*8.65729}
.ends 40571
.end
```

A3.2 Van der Pol oscillator with NIC and gyrator

```
*************************************************
.param C={5.30516e-012}
.param L={5.30516}
.param RL={1000}
.param RNEG={10000}
.param RINT={1e+010}
.tran 0 0.00766667 0.00433333 3.33333e-007
.print tran V(V_OSC)
.ic V(V_OSC)=0.0000001
*****************************************************
R1 V_OSC N002 1k
R2 N002 N005 1k
R3 N005 N004 1k
R5 N006 0 {RL}
C2 N004 N006 {L/1e6}
C1 V_OSC 0 {C}
R7 V_OSC N001 1k
R8 N003 N001 1k
R9 0 N003 {RNEG}
D1 V_OSC 0 1N4148
D2 0 V_OSC 1N4148
V1 0 VN 15
V2 VP 0 15
R4 V_OSC 0 {RINT}
XU1 V_OSC N005 VP VN N004 TL071
XU2 N006 N005 VP VN N002 TL071
XU3 V_OSC N003 VP VN N001 TL071
.model D D
.model 1N4148 D(Is=2.52n Rs=.568 N=1.752 Cjo=4p M=.4 tt=20n IBV=200m BV=75)
*****************************************************
* TL071 OPERATIONAL AMPLIFIER "MACROMODEL" SUBCIRCUIT
* CREATED USING PARTS RELEASE 4.01 ON 06/16/89 AT 13:08
```

```
* (REV N/A) SUPPLY VOLTAGE: +/-15V
* CONNECTIONS: NON-INVERTING INPUT
* | INVERTING INPUT
* | | POSITIVE POWER SUPPLY
* | | | NEGATIVE POWER SUPPLY
* | | | | OUTPUT
* | | | | |
.SUBCKT TL071 1 2 3 4 5
*
C1 11 12 3.498E-12
C2 6 7 15.00E-12
DC 5 53 DX
DE 54 5 DX
DLP 90 91 DX
DLN 92 90 DX
DP 4 3 DX
EGND 99 0 POLY(2) (3,0) (4,0) 0 .5 .5
FB 7 99 POLY(5) VB VC VE VLP VLN 0 4.715E6 -5E6 5E6 5E6 -5E6
GA 6 0 11 12 282.8E-6
GCM 0 6 10 99 8.942E-9
ISS 3 10 DC 195.0E-6
HLIM 90 0 VLIM 1K
J1 11 2 10 JX
J2 12 1 10 JX
R2 6 9 100.0E3
RD1 4 11 3.536E3
RD2 4 12 3.536E3
RO1 8 5 150
RO2 7 99 150
RP 3 4 2.143E3
RSS 10 99 1.026E6
VB 9 0 DC 0
VC 3 53 DC 2.200
VE 54 4 DC 2.200
VLIM 7 8 DC 0
VLP 91 0 DC 25
VLN 0 92 DC 25
.MODEL DX D(IS=800.0E-18)
.MODEL JX PJF(IS=15.00E-12 BETA=270.1E-6 VTO=-1)
.ENDS
*****************************************************
.end
```