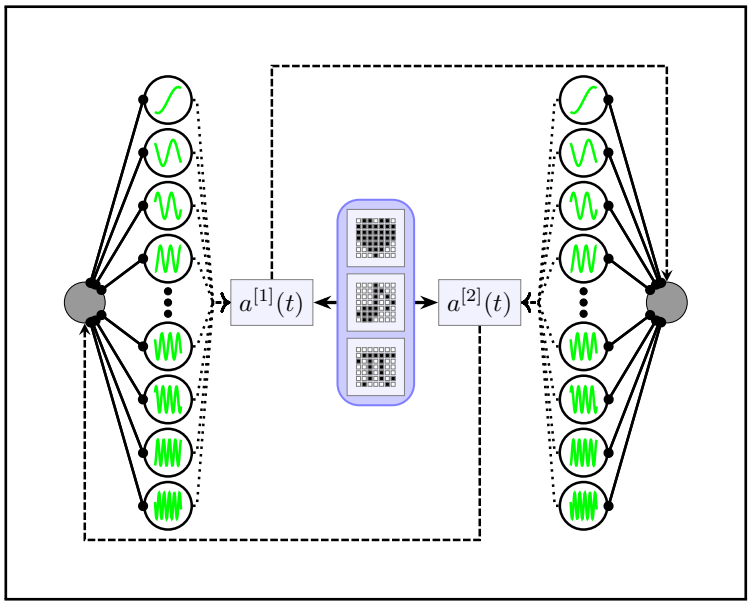
Technische Universität München
Fakultät für Physik
E19a

Pattern recognition with phase oscillator networks: Analysis and experimental verification of a novel architecture

Daniel Heger



Fakultät für Physik

Pattern recognition with phase oscillator networks: Analysis and experimental verification of a novel architecture

Daniel Heger

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines *Doktors der Naturwissenschaften* genehmigten Dissertation.

Vorsitzender: Prof. Dr. Reinhard Kienberger
Prüfer der Dissertation:
1. Prof. Dr. Katharina Krischer
2. Prof. Dr. Ulrich Gerland

Die Dissertation wurde am 25.06.2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 04.10.2018 angenommen.

Danksagung

Hier möchte ich allen Menschen danken, deren Unterstützung diese Promotion möglich gemacht hat: Meine Doktormutter Katharina Krischer hat mir von Anfang an die Möglichkeit gegeben, meine Forschung frei zu gestalten. Mit Ihrer Erfahrung brachte Sie viele Ideen und neue Blickwinkel in unsere Diskussionen ein, während Ihre Begeisterung für das Thema und Ihr Optimismus halfen, die Dinge in einem positiven Licht zu sehen. Vielen Dank dafür! Ich möchte auch Katharina, Toby, Felix, Elena und Stefan für Ihr sorgfältiges Korrekturlesen und Ihre konstruktiven Verbesserungsvorschläge danken. Danke auch an meine Freunde und Geschwister, dass ihr mich nicht vergessen habt, obwohl ich zeitweise so in die Promotion abgetaucht bin, dass ich mich nicht gemeldet habe. Meinen Kollegen am Lehrstuhl E19a möchte ich für die angenehme Arbeitsatmosphäre und die schöne gemeinsame Zeit danken. Meine Eltern haben sicherlich die Grundsteine für meine wissenschaftliche Karriere gelegt: Das Grundvertrauen, das ich aufgrund der Geborgenheit während meiner Kindheit aufbauen konnte, half mir, Phasen der Unsicherheit und Misserfolge zu überstehen. Vor allem aber haben meine Eltern mich immer in meinem Spaß am Entdecken und Verstehen der Natur bestärkt.

Ganz besonderer Dank gebührt meiner Frau Elena: Sie stand immer hinter meinen Entscheidungen und hatte Verständnis, wenn ich mich des Öfteren ausschließlich mit Physik beschäftigt habe. Wenn ich demotiviert war, hat Sie mich überzeugt, weiter zu machen und wenn ich angesichts unzähliger Schwierigkeiten die Hoffnung zu verlieren drohte, hat Sie mir Kraft gegeben. Elena, ich liebe dich über alles! Danke, dass du immer für mich da bist.♥

Contents

Chapter 1

Introduction

Swarms of male fireflies blinking in unison have been observed [1]. Similar pendulum clocks swing synchronously if mounted on a common support [2]. Everyone of us has an internal, chemical clock with a period of roughly 24 hours that constantly adapts to the light from the sun [3]. After long flights along longitudes, however, this circadian clock is out of sync and needs a few days to readjust, which is the reason for jetlag [4]. All these seemingly different examples are based on the same principle: Some autonomous system - a firefly, a pendulum clock or the circadian rhythm - shows self-sustained periodic behavior. These *nonlinear oscillators* then adjust frequency and phase of their oscillation to match other oscillators or a periodic signal: They *synchronize*.

Synchronization occurs in many scientific disciplines, with applications in engineering, physics, chemistry, biology, medicine, and even in social sciences. The study of synchronization has intensified in the last decades, with [5–12] summarizing the state of the art.

Particularly intriguing, synchronization phenomena occur in neuroscience as well [13]: Oscillations can be found at different length- and timescales in the nervous system [14]. The synchronization of neural oscillators controls vital functions, but is also responsible for some neural diseases such as epileptic seizures [15]. Synchronization phenomena are also involved in cognition tasks of the brain [16].

Early attempts to understand the human brain also inspired the development of *artificial neural networks*: Interaction of networks of small subunits, called "artificial neurons", have been studied for roughly 70 years now [17]. After recent improvements in both algorithms (e.g. [18–20]) and computational power (e.g. GPU computing), today these artificial neural networks are successfully applied in industry for e.g. image and speech recognition, speech generation, or forecasting. In contrast to nerve cells, conventional artificial neurons are simple functions of their inputs and are evaluated at discrete points in time, emitting a single value. Admittedly, these rough simplifications are necessary to allow for efficient computation on conventional computers. If specialized hardware is considered, however, artificial neurons are not limited to clocked evaluation on CPUs: Oscillators, for instance, can be regarded as artificial neurons. Any coupling to another oscillator's signal continuously promotes

synchronization at some specific phase difference. Consequently, coupling many oscillators can lead to complex dynamics. Can synchronization of oscillators thus be used for computation?

Indeed, inherent parallel architectures have been proposed for such networks of oscillators. Continuous processing of signals and low power consumption of some proposed nano-oscillators [21–23] promise fast and energy efficient hardware, which could support conventional processors in specialized tasks. However, many problems are still unsolved. This thesis specifically aims to construct a network of oscillators acting as *autoassociative memory*, while at the same time mitigating or removing shortcomings of earlier architectures.

When an *autoassociative memory* is presented with a defective and/or incomplete piece of data, it recognizes and retrieves the correct data from a set of correct candidates. From a different point of view, the defective input data are mapped onto the most similar of the candidates. The ability to "map" is also found in complex physical systems: The trajectory of a system state will converge to an attractor. If several attractors coexist, different sets of initial conditions, called *basins of attraction*, will end up on different attractors. Therefore, the system "maps" all initial conditions within one basin onto its attractor. Note that few physical systems are actually suitable as autoassociative memories: First, suitable mappings of the defective data onto the initial conditions and from the attractors back onto the correct patterns have to be found. Additionally, initial conditions as well as attractors of a system need to be controlled, with the latter usually being difficult. Finally, the initial defective data should be mapped onto the most similar correct data candidate, which requires that the basins of attraction actually conform with a sensible definition of similarity. The idea to use basins of attractions for pattern recognition has originally been proposed by Hopfield for use in conventional neural networks [24]. Contributions from mathematics, physics and neuroscience made it possible to merge his ideas with the studies of coupled nonlinear oscillators. (See the end of [25] for a summary.)

Networks of nonlinear oscillators have been shown to act as autoassociative memory devices for binary patterns [26–33] according to the above-mentioned principle. In the original architecture [26–28], identical Kuramoto oscillators [25] are fully interconnected via programmable connections that can change sign and strength of the coupling according to the Hebbian Rule [34]: With α_i^m being the i^{th} pixel of the m^{th} memorized pattern, the coupling between oscillators i and j is proportional to $S_{ij} = \sum_{m=1}^{M} \alpha_i^m \alpha_j^m$. Here, each of the M correct pattern candidates consists of N pixels. Specifically, the phase ϑ_i of the i^{th} oscillator evolves according to $\dot{\vartheta}_i = \Omega - \frac{\epsilon}{N} S_{ij} \sin(\vartheta_i - \vartheta_j)$ with constants ϵ and Ω representing the coupling strength and the angular frequency of the oscillator. In a frame of reference rotating with Ω, this simplifies to $\dot{\varphi}_i = \dot{\vartheta}_i - \Omega = -\sum_{j=1}^{N} \frac{\epsilon}{N} S_{ij} \sin(\varphi_i - \varphi_j)$. Here, fixed points are the only type of attractors, and defective input patterns as well as correct pattern candidates can be mapped on two synchronized groups of oscillators whose phases

differ by π. However, this design has two main disadvantages: First, no distinct, well-separated fixed points exist for the memorized patterns [35]. Instead, there is one global attractor consisting of lines of attractive fixed points with neutrally stable eigendirections that connect every memorized pattern with every other. On short timescales, pattern recognition still works: Starting at the defective pattern, the system state quickly relaxes onto the global attractor close to the most similar pattern. On the attractor, however, perturbations due to external noise or implementation inaccuracies dominate and the system state drifts away from the correct pattern on longer timescales. Additionally, recognition success cannot be guaranteed as no well-defined basin of attraction exists for any single output pattern.

Second, the number of connections scales quadratically with the number of oscillators, so no large networks can be implemented in hardware.

So far, no architecture that solves both issues has been proposed. However, separate solutions for each problem have been discussed: Nishikawa et al. [32] showed that the degeneracy of the attractor can be lifted by adding second order Fourier modes to the coupling, i.e. he explored dynamics $\dot{\varphi}_i = \sum_{j=1}^{N} \left(-S_{ij} \sin\left(\varphi_i - \varphi_j\right) - \frac{\gamma}{N} \sin\left(2\left[\varphi_i - \varphi_j\right]\right)\right)$ with γ being a small parameter. A similar network with third order Fourier modes has been proposed as well [33]. However, it is unclear how such a coupling could be implemented. Additionally, the number of physical connections still scales quadratically with the number of oscillators. A partial solution for the scaling problem has been proposed by Hoppensteadt and Izhikevich [29] and has been further advanced by Hölzel and Krischer [30] and Kostorz et al. [31]: Oscillators of different frequencies are coupled to the same global coupling that affects every oscillator differently. These architectures require an external input of complex time-dependent functions, but the number of connections scales with $\mathcal{O}(N)$. Although the coupling mechanism is different, the effective dynamics $\dot{\varphi}_i = \dot{\vartheta}_i - \Omega_i = -\sum_{j=1}^{N} \frac{\epsilon}{N} S_{ij} \sin\left(\varphi_i - \varphi_j\right)$ are identical to the original architecture and thus, this approach inherits the unfavorable valleys of neutrally stable fixed points.

In this thesis, a novel network architecture is designed that combines both isolated attractors and minimal scaling of connection complexity without the need for complex external input. In order to prove implementability in hardware, validate our theoretical predictions and explore the limits of network parameters, we implemented the architecture with electronic Van der Pol oscillators.

First, we will lay out some foundations of nonlinear dynamics and analog electronics in chapter 2. In chapter 3, the new network architecture is presented. It is built on the aforementioned previous studies [29, 30] of globally coupled oscillatory devices, but we introduced two peculiar features: Different temporal modulation of the coupling strength and a replacement of the single network by two interconnected subnetworks. The result is a robust autoassociative memory that can be readily read out. Additionally, we can predict recognition success analytically. The implementation of the network in electronic circuits as well as the measurement environment are

discussed in detail in chapter 4, before measurement results are outlined in chapter 5. Here, the automation of the network hardware allowed to gather statistics on recognition success and recognition times in the presence of noise and inaccuracies. Network performance, however, is close to ideal for a large range of parameters. Finally, we summarize our results in chapter 6 and outline possible future research: While both theoretical questions and unexplored improvement possibilities remain open for our architecture itself, other applications in e.g. combinatorics can be explored as well.

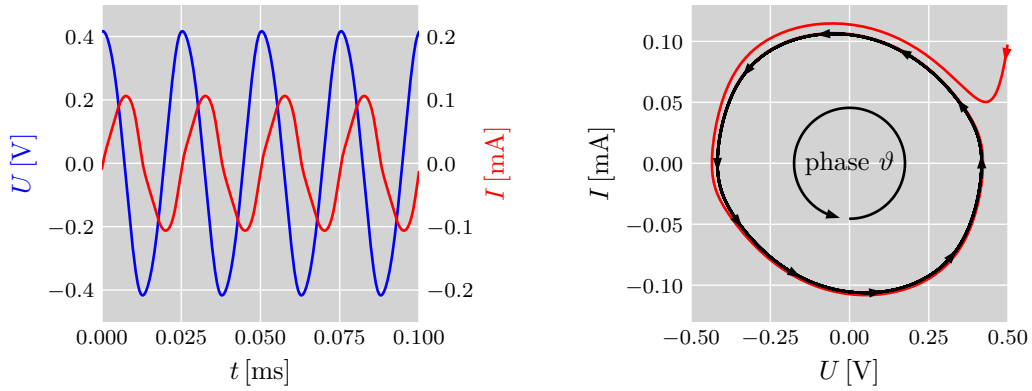
Chapter 2

Foundations

This thesis is interdisciplinary, thus touching or requiring selected topics from physics, information technology and electrical engineering. Consequently, the current chapter aims to communicate the most necessary foundations from nonlinear dynamics and analog electronics, convey intuition and provide resources for further reading. However, these excursions are kept simple on purpose and do not claim to be fully sufficient for understanding all details of the following chapters. Both the books *Nonlinear Dynamics and Chaos* [36] and *The art of electronics* [37] are suitable as an introduction, but also cover fine details, edge cases and exceptions.

2.1 Nonlinear dynamics

As mentioned in the introduction, we aim to construct a network of oscillators with several, separated attractors. Both internal dynamics of oscillators as well as their collaborative behavior can be described with methods of nonlinear dynamics: First, differential equations are rewritten as system of first order differential equations. If we consider e.g. a single electrical oscillator, variables could be a voltage U and a current I. Instead of focusing on individual trajectories as functions of time, the system is then described in the space spanned by its variables. Consequently, a single point in this *phase space* describes a possible state of the system. Over time, the system state "flows" through phase space, with its velocity vector determined by the time derivatives of the variables, which are in turn conveniently provided by the differential equations. Finally, after some transient, the system state settles on some structure in phase space and stays there for all future times. These structures, called *attractors*, come in different shapes: Points, which correspond to a constant system state, closed curves called *limit cycles*, which correspond to oscillations, and other, more complicated shapes like tori or even fractals. Note that a dynamical system can have several attractors, and all initial conditions that end up on one attractor belong to its *basin of attraction*.



(a) **Periodic long-term behavior**:
After an initial transient (not shown), variables describing an oscillator show periodic motion. In this example, variables are a voltage U and a current I, but generally, they depend on the oscillator: For e.g. a metronome, they could be position and velocity of its tip, while the variables for a firefly's blinking are concentrations of chemicals inside their bodies. Note that although we use oscillators with roughly sinusoidal voltages, oscillations need not be sinusoidal in general.

(b) **Limit cycle**:
Variables describing a nonlinear oscillator's internal dynamics are plotted against each other. Then, the periodic motion is situated on a closed curve, the so-called limit cycle.(black) Any displacement of the variables quickly settles on the limit cycle (red). If perturbations from the limit cycle are not too large, the dynamic of the oscillator can thus be described with a phase $\vartheta \in [0, 2\pi[$ only.

Figure 2.1: Exemplary dynamics of a nonlinear oscillator.
(Specifically, the data shown is from simulations of an electrical Van der Pol oscillator, which are also used in the prototype network. The simulated oscillator had a capacitance $C = 1\,\mathrm{nF}$, an inductance $L = 15{,}8\,\mathrm{mH}$ and an effective negative resistance $R'_{neg} = -33\,\mathrm{k\Omega}$ and oscillates with a frequency $f \approx 40\,\mathrm{kHz}$. For further simulation details, see Sec. 5.1.1.)

2.1.1 Nonlinear oscillators

The aforementioned principles will now be applied to a single nonlinear oscillator and illustrated with simulation data of the oscillators used in our experiment: The variables describing a nonlinear oscillator show periodic motion after a transient (Fig. 2.1a). In the two-dimensional phase space (Fig. 2.1b), which is spanned by these variables (e.g. U and I in our example), the dynamics of the system are much more visible: All system states that belong to the periodic motion lie on a limit cycle (black). Any trajectory, e.g. the red curve, does quickly end up on this attractor. Consequently, the amplitude of the oscillation is constant after a transient. Note that nonlinear oscillators strongly differ from linear, harmonic oscillators in this regard: The amplitude of undamped harmonic oscillators is not fixed, but depends on their total energy, which is conserved. Damped harmonic oscillations quickly die down - their attractor is thus a fixed point - and thus their oscillations are not a long-term, but only a transient phenomenon. On the other hand, nonlinear oscillators need an energy source to stabilize the limit cycle against perturbation. On the limit cycle, the system state can be described by a single variable: The phase $\vartheta \in [0, 2\pi[$ of the oscillation, which increases linearly with time. Actually, as close trajectories always approach the attractor, the definition of the phase can be extended to any points in the limit cycle's basin of attraction [38, 39]. Furthermore, interactions of oscillators can conveniently be described by their phases, if deviations from the limit cycle stay sufficiently small [38].

2.1.2 Phase description

While an oscillator's state can be described by a phase alone, any external signals first influence its internal variables. Imagine an external signal perturbing the voltage of the exemplary oscillator in Fig. 2.1b: If e.g. $\vartheta \approx 0$ or π, the perturbation in U is almost parallel to the limit cycle and directly increases or decreases phase. However, for $\vartheta \approx \frac{\pi}{2}$ or $\frac{3\pi}{2}$, the perturbation is perpendicular to the limit cycle and the phase remains the same. The phase change at an arbitrary phase to an internal variable X is generally described with a *phase response curve* $\frac{\mathrm{d}\vartheta}{\mathrm{d}X}$ and depends on the shape of the limit cycle and its neighborhood [38]. The phase dynamics of a single oscillator can then be summed up:

$$\dot{\vartheta} = \Omega + \sum_{X_i} \frac{\mathrm{d}\vartheta}{\mathrm{d}X_i} \Delta X_i$$

Here, $\Omega = 2\pi f$ is the angular frequency of the oscillator. A phase response curve for our exemplary oscillator is shown in Fig. 2.2. Simplifying, we restrict our further analysis to oscillators that emit a sinusoidal signal (e.g. their voltage) and have a sinusoidal phase response shifted by $\frac{\pi}{2}$ with respect to the oscillation. (Note that all

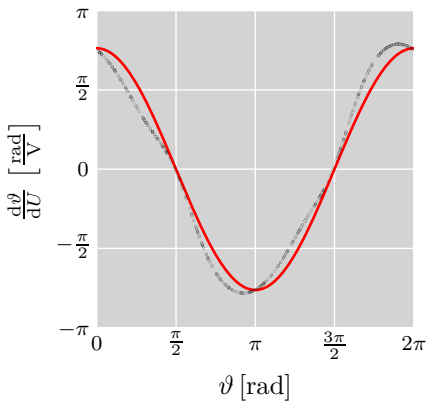


Figure 2.2: **Phase response curve** $\frac{\mathrm{d}\vartheta}{\mathrm{d}U}$ of the exemplary oscillator from Fig. 2.1. Black dots show the simulated phase response to voltage perturbations. The red curve is a cosine fit.

oscillators close to a Hopf-bifurcation [40] fulfill that requirement. Thus, the analysis remains valid for a large class of oscillators independent of their exact differential equations.)

2.1.3 Synchronization

Now, we couple two oscillators symmetrically, meaning that for each oscillator, an internal variable is perturbed proportional to the other's signal:

$$\dot{\vartheta}_1 = \Omega_1 + \epsilon \cos\vartheta_1 \cdot \sin\vartheta_2$$
$$\dot{\vartheta}_2 = \Omega_2 + \epsilon \cos\vartheta_2 \cdot \sin\vartheta_1$$

Here, ϑ_1 and ϑ_2 are the phases of the individual oscillators, Ω_1 and Ω_2 their angular frequencies and the coupling strength ϵ is a constant encapsulating amplitudes of signals and phase response curves as well as factors due to details of the coupling mechanism. If both oscillators synchronize, their frequencies align. Then, the difference between their phases $\Delta\vartheta = \vartheta_1 - \vartheta_2$ must thus become constant or at least be limited to a small range of values. Thus, we express the dynamics in $\Delta\vartheta$ and simplify with the trigonometric identity $\sin x \cos y = \frac{1}{2}\left[\sin(x-y) + \sin(x+y)\right]$:

$$\begin{aligned}\Delta\dot{\vartheta} &= \dot{\vartheta}_1 - \dot{\vartheta}_2 \\ &= \Delta\Omega + \epsilon\left(\cos\vartheta_1 \cdot \sin\vartheta_2 - \cos\vartheta_2 \cdot \sin\vartheta_1\right) \\ \Delta\dot{\vartheta} &= \Delta\Omega - \epsilon \sin\Delta\vartheta \end{aligned} \tag{2.1}$$

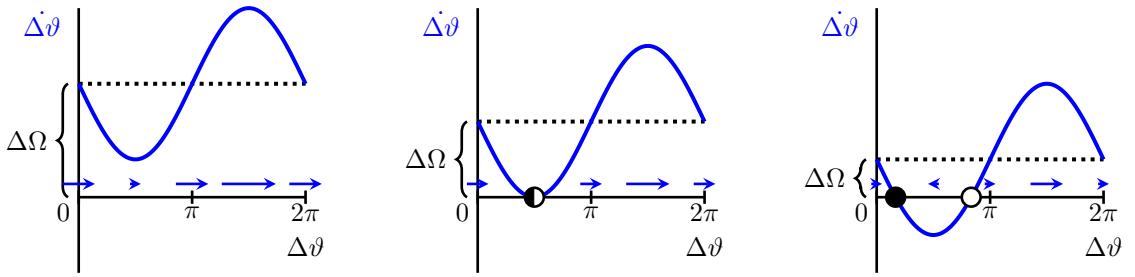
Now, the only variable describing our oscillator pair is the phase difference $\Delta\vartheta$ and the phase space of the system is thus on a line. If $\Delta\vartheta$ grows without bound, the frequencies cannot be equal. Consequently, for synchronization to occur an attractor must exist. In a one-dimensional system, this can only be a stable fixed point [36]: On a fixed point, the flow vanishes, so a trajectory starting on the fixed point remains there for all times:

- For $|\Delta\Omega| > \epsilon$, the flow $\Delta\dot{\vartheta}$ (Eq. (2.1)) is always larger than 0, so our oscillator pair cannot synchronize.
- For $|\Delta\Omega| < \epsilon$, the flow intersects with 0 at two points. Regarding stability, there is always one stable (attractive) and one unstable (repulsive) fixed point, as $\Delta\vartheta$ either flows towards the fixed point or away from it, as illustrated in Fig. 2.3c. As no other attractors exist, the system moves to the stable fixed point, where the oscillators are perfectly synchronized. Note that for $\Delta\Omega = 0$, the stable fixed point is at a phase difference of 0. If $|\Delta\Omega|$ increases, the fixed point moves away from zero till it annihilates with the unstable fixed point at $|\Delta\Omega| = \epsilon$. Similarly, if the coupling had a negative prefactor, the $\sin\Delta\vartheta$-term would be inverted as well. Then, the fixed points had inverted stability and the stable fixed point was close to π.

The synchronization of multiple oscillators with different frequencies - with every oscillator being coupled to every other - is substantially more complex and has been extensively studied by Kuramoto [41]. The main message, however, stays the same: As long as the coupling is sufficiently larger than the frequency differences, the oscillators synchronize.

2.1.4 Saddle-node bifurcations

In the last section, we discussed two symmetrically coupled oscillators. Particularly, when $\Delta\Omega \approx \epsilon$, only minuscule changes to the angular frequency difference $\Delta\Omega$ were necessary to switch from synchronization and a *phase-locked* state to an ever increasing phase difference. Such qualitative and substantial changes of a systems behavior due to small parameter changes are called *bifurcations*. The behavioral change coincides with a creation, change or annihilation of attractors, repellers or saddles in phase space. (Saddles are attractive in at least one direction and repulsive in at least another.) The bifurcation in our synchronization example is a *saddle-node bifurcation*, where a pair of fixed points annihilates. This is illustrated in Fig. 2.3, where one stable and one unstable fixed point are created if $|\Delta\Omega|$ is reduced to values smaller than ϵ, or annihilated if $|\Delta\Omega|$ is increased again. Note that saddle node bifurcations arise in many systems and the exact shape of the flow is not critical: As long as the flow close to the bifurcation can be approximated with a parabola - more precisely, as long as the flow is topologically equivalent to $\dot{X} = r + X^2$ close



(a) **No fixed points:** As $\Delta\Omega > \epsilon$, the flow of $\Delta\vartheta$ is positive in the whole phase space and one oscillator will always pass the other.

(b) **One half-stable fixed point:** At $\Delta\Omega = \epsilon$, there is exactly one fixed point $\Delta\vartheta^* = \frac{\pi}{2}$, where $\Delta\dot{\vartheta} = 0$. Everywhere else, the flow is still positive and $\Delta\vartheta$ will increase until it reaches the fixed point. The fixed point, however, is not stable: Any positive perturbation in $\Delta\vartheta$ will grow.

(c) **Fixed point pair:** For $\Delta\Omega < \epsilon$, there are two intersections of $\Delta\dot{\vartheta}$ with the $\Delta\vartheta$-axis. The flow on the phase space line strictly flows to the stable fixed point (black circle), while it flows away from the unstable fixed point (white circle) on both sides.

Figure 2.3: A fixed point pair emerges in a **saddle-node bifurcation**: Synchronization of two oscillators requires a stable fixed point in the phase space spanned by their phase difference $\Delta\vartheta$. As seen in Eq. (2.1), zeros of $\Delta\dot{\vartheta}$ occur only for $\Delta\Omega \leq \epsilon$ and correspond to a half-stable fixed point or a stable and an unstable fixed point, as can be seen in b) and c). The flow $\Delta\dot{\vartheta}$ on the one-dimensional phase space is additionally visualized with blue arrows.

to the bifurcation, with r being the bifurcation parameter - the structurally same annihilation or creation of a fixed point pair will take place. In this thesis, saddle-node bifurcations will appear more often, e.g. in the context of frequency annealing (Sec. 3.6.2 and 4.1.3) or removal of spurious coupling (Sec. 4.1.4). Note that in higher dimensions, saddle node bifurcations still occur. Stability of fixed points, however, must be determined for as many directions as there are dimensions.

2.1.5 Stability

Admittedly, we did not define stability rigorously yet. Actually, several definitions of stability exist [42]:

- Trajectories close to a *Ljapunov stable* attractor X^* stay in a neighborhood of X^* for $t \to \infty$.

- An *asymptotically stable* attractor is always Ljapunov stable, but trajectories $X(t)$ also approach it, so $|X(t) - X^*| \to 0$ for $t \to \infty$.
- *Exponentially stable* attractors are again a special case of asymptotic stability: Here, $|X(t) - X^*| < \mathrm{e}^{-at}$ in a neighborhood of X^* as $t \to \infty$.

Attractors that are only Ljapunov stable, but not asymptotically stable, are also called *neutrally stable*. One example of neutrally stable structures are the potential valleys in previous oscillatory neural networks [26–31], that were already mentioned in the introduction: In a unperturbed system, the system state would move to the closest attractor, which is part of a line of infinitely many neutrally stable fixed points. When deviations of the systems or perturbations are considered - e.g. a "tilt" of the potential function - we see that Ljapunov stable attractors are not desirable for an attractor network: The attractors loose their stability and the system state will move along the valleys.

In the oscillator network proposed in chapter 3, many fixed points exist, but only a few fixed points are attractors. Stability analysis for a fixed point X^* starts with a linearization around the fixed point: First, we determine all components $J_{ij} = \frac{\mathrm{d}\dot{X}_i}{\mathrm{d}X_j}|_{X=X^*}$ of the *Jacobian* matrix $\mathbf{J}$ at the fixed point. They describe, how the flow in the i^{th} coordinate changes if we perturb the j^{th} coordinate. In other words, we approximate the dynamics around the fixed point with

$$\dot{X}_i = \sum_j J_{ij}(X_j - X_j^*)\,.$$

In our one-dimensional synchronization example, this reduces to a single component: The derivative of the flow $\Delta\dot{\vartheta}$ with respect to $\Delta\vartheta$. If $\frac{\mathrm{d}\Delta\dot{\vartheta}}{\mathrm{d}\Delta\vartheta}|_{\Delta\vartheta=\Delta\vartheta^*} < 0$, the flow is negative for $\Delta\vartheta > \Delta\vartheta^*$ and positive for $\Delta\vartheta < \Delta\vartheta^*$ and thus the fixed point is stable and vice-versa, a positive slope of the flow indicates an unstable fixed point. For a general Jacobian matrix with more than one dimension, perturbations in one coordinate influence the flow in others, so the analysis is not as straightforward. However, remembering linear algebra, we diagonalize $\mathbf{J}$ by solving the eigenvalue equation $\mathbf{J}\boldsymbol{v} = \lambda\boldsymbol{v}$ for eigenvalues λ and eigenvectors $\boldsymbol{v}$, which corresponds to a coordinate change. (Specifically, eigenvalues are usually retrieved by solving the characteristic polynomial $\det(\mathbf{J} - \lambda\mathbb{1}) = 0$. Note that eigenvalues can be complex numbers.) As $\mathbf{J}$ is symmetric, we can always do this. Furthermore, it is guaranteed that the eigenvectors form an orthogonal base. In other words, any perturbations along one eigendirection can be treated independent of the others. If we express the dynamics around the fixed point in new coordinates $\tilde{X}_i$ along the eigendirections and centered around X^*, the dynamics simplify:

$$\dot{\tilde{X}}_i = \sum_j \tilde{J}_{ij}\tilde{X}_j = \lambda_i\tilde{X}_i$$

$$\Rightarrow \tilde{X}_i = \tilde{X}_i^0 \mathrm{e}^{\lambda_i t}$$

Consequently, the eigenvalues λ_i directly specify the stability along the corresponding eigendirection:

- $\mathrm{Re}(\lambda) < 0$: stable eigendirection (exponentially stable).
- $\mathrm{Re}(\lambda) > 0$: unstable eigendirection (exponentially unstable).
- $\mathrm{Re}(\lambda) = 0$: The linear term of the flow is zero along the eigendirection, so our linearization cannot decide on stability: Any non-zero higher order term will change the sign of the flow, and other methods must be applied to determine stability.

In chapter 3, a novel oscillator network for pattern recognition is presented, where N pairs of oscillators represent pixels of a pattern. The dynamics are expressed in phase differences $\Delta\vartheta_i$, attractors are exponentially stable fixed points, and analysis in this N-dimensional phase space is based on the principles presented here. Preparing for chapter 4, which covers the implementation of the network with electronic circuits, the next section outlines basics of analog electronics.

2.2 Electronic circuitry

Any circuits and circuit components used in the experiment are well-known in electronics literature [37], but are probably unknown to readers with a pure physics background. This section gives a short overview over electronic principles, circuit components and subcircuits.

2.2.1 Circuit elements

Before presenting individual subcircuits, we first introduce single components. We start with the operational amplifier, which is used in many circuits that utilize feedback:

Operational amplifiers

An operational amplifier (op-amp) has one output and two inputs and needs to be powered. The op-amp amplifies the difference of the input voltages with a very high amplification factor (typically $10^5 - 10^6$, ideally ∞), but without exceeding the positive or negative supply voltage. However, an op-amp is almost always used with feedback, which means that the surrounding circuitry allows the output voltage to influence the input voltages. Consequently, the feedback network must be considered for the analysis of an op-amp's behavior. Feedback to the inverting input (marked with "$-$" in circuits) diminishes the voltage difference $\Delta U = U_+ - U_-$ between the inputs and is thus called negative feedback. Positive feedback goes to the non-inverting input ("+") and enhances ΔU. If negative feedback is larger than positive feedback, behavior of an ideal op-amp is described by the **golden rules**, as stated in [37]:

1. "The output attempts to do whatever is necessary to make the voltage difference between the inputs zero."
2. "The inputs draw no current."

Op-amp circuits offer multiple advantages:

- First, the low output resistance of most op-amp circuits allows to consider different parts of circuitry independently: Voltages provided by the op-amp circuit do not depend on the input resistance of the following circuitry. Similarly, circuit components leading to the inputs lose no current to the inputs or lose a well-defined current to the feedback circuitry. Combining op-amp circuits is thus straightforward and predictable.
- Second, the output almost exclusively depends on the feedback network. Choosing precise components for the feedback network, combined with the first point, allows for precise output.

- Third, many different circuits are known and information on pitfalls and non-ideal behavior is easily available.
- Forth, op-amps are inexpensive and integrated circuits (ICs) housing several op-amps can be purchased, which diminishes layout complexity and space.

Consequently, op-amp circuits are used in abundance for the design of the experimental network.

JFETs as voltage controlled resistors

JFETs are 3-terminal-devices with inputs called source, drain and gate and can be used as voltage-controlled resistors with some additional circuitry. The gate draws no current and controls the current between source and drain. If the voltage difference V_{DS} between source and drain is small, the current is

$$I_{DS} = 2k[(U_G - U_S - V_{GS,off})V_{DS} - \frac{V_{DS}^2}{2}] \ .$$

Here, k and $V_{GS,off}$ are constants. Voltages V with two indices indicate voltage differences between two points, while U with a single index is measured relative to ground.
An ideal resistance would be constant for any change of drain or source voltage. In other words, I_{DS} should be exactly proportional to V_{DS}. If we add parts of source and drain voltage to a control voltage U_C at the gate, the quadratic term and the dependency on U_S can be removed:

$$\begin{aligned} U_G &\stackrel{!}{=} U_C + 0.5 \cdot U_S + 0.5 \cdot U_D \\ \Rightarrow \ I_D &= 2k \cdot (U_C - V_{GS,off})V_{DS} \\ \Rightarrow \ R_{JFET} &= \frac{V_{DS}}{I_S} = \frac{1}{2k \cdot (U_C - V_{GS,off})} \end{aligned}$$

In the experiment, we use the JFET 2N5486 (gate-source cutoff voltage $V_{GS,off} \in [-2\,\mathrm{V}, -6\,\mathrm{V}]$) as a voltage controlled resistor for frequency annealing of oscillator pairs. (See Sec. 4.1.3) Note that the gate forms a semiconductor junction with the channel connecting source and drain similar to a diode. In order to avoid unintended diode conduction and thus any gate current for the the JFET 2N5486, $U_G - U_S \leq +0{,}5\,\mathrm{V}$ and $U_G - U_D \leq +0{,}5\,\mathrm{V}$ must be ensured.

Integrated circuits

Frequently required circuits are produced on silicon wafers and can be bought as chips, also called integrated circuits (ICs). ICs usually save cost, space, and avoid

implementation errors and we consequently used some for building the experimental network:

- The ICs TL074 and TL072 contain 4 resp. 2 operational amplifiers which perform well at the frequencies used.
- The Multiplier IC AD633 offers precise analog multiplication for all different combination of input voltage signs. The AD633 offers two fully differential inputs, an additional summation input and an output range of [±10 V]. It's output voltage W follows the function

$$W = \frac{(X_1 - X_2)(Y_1 - Y_2)}{10\,\mathrm{V}} + Z \ .$$

  Here, X_1,X_2,Y_1 and Y_2 are the input voltages and Z is the voltage at an additional summation input. Note that the multiplication result is divided by 10 V.

  If resistors R_{WZ} and R_{Zg} form a voltage divider between the output W, the summation input Z, and ground, the output is amplified: Voltage at the summation input is $Z = \frac{R_{Zg}}{R_{Zg}+R_{WZ}}$, which infers

$$W = \frac{R_{Zg} + R_{WZ}}{R_{WZ}} \cdot \frac{(X_1 - X_2)(Y_1 - Y_2)}{10\,\mathrm{V}} \ .$$

- The LF398 contains a "sample and hold" circuit. During "sampling", the IC passes voltages unchanged, but voltage levels are "recorded" on a capacitor. When a logic signal switches the IC to "hold"-mode, the input is disconnected and the output is set to the voltage on the capacitor, mirroring the last sampled value. The capacitor is not part of the IC, but needs to be added on one pin of the IC and can be selected for holding time or speed of sampling. If a serial resistor is added to the capacitor, the sampled value is additionally low-passed.
- The ICs DG411 and DG412 contain 4 analog switches respectively. Each analog switch can be opened / closed by changing the digital signal applied.

2.2.2 Subcircuits

Numerous subcircuits are used in the implementation of the network and the experiment: Summations, amplifications, inversions and multiplications are necessary for the construction of signals between the oscillators. Supporting circuitry for readout of phase differences and frequency adjustment requires removal of frequency components. Finally, electrical oscillators need to be implemented as well, with subcircuits imitating coils and negative resistances. First, the inverting summer will be

analyzed to illustrate the application of the golden rules on op-amp circuits. Other simple op-amp circuits are only displayed together with their purpose, as accurate descriptions can easily found in the literature [37]. Afterwards, filters and finally Van der Pol oscillators will be discussed.

Inverting summer

The inverting summer (Fig. 2.4a) is used in most modules. It can be easily analyzed by applying the golden rules:
Golden rule 1:

$$U_- = U_+ = 0\,\mathrm{V}$$

As U_- is effectively grounded, any input voltage cannot influence the other inputs. Due to golden rule 2, currents at the inverting input must cancel out:

$$I_{in} = -I_{out} \qquad \text{(Ohm's Law)}$$

$$\sum_{i=1}^{N} \frac{U_i - U_-}{R_i} = -\frac{U_{out} - U_-}{R_{out}}$$

$$U_{out} = -\sum_{i=1}^{N} \frac{R_{out}}{R_i} U_i \tag{2.2}$$

Amplification factors for different inputs can be chosen with appropriate values for R_i and R_{out}.

If the circuit has only one input, it is called "inverting amplifier" or "inverter". Another variation of this circuit is the addition of input voltages at the non-inverting input, which are counted with positive sign. However, compensation resistors have to be chosen appropriately in order to guarantee stable circuit performance.

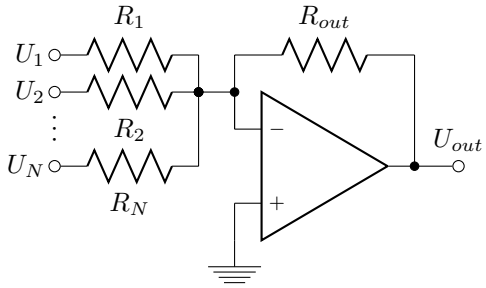
Other often used subcircuits are shown in Fig. 2.4, 2.5 and 2.6. The analysis for most of them is analog to the inverting summer.

Filters

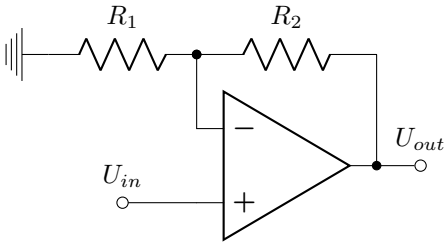
At first sight, a network of oscillators requires no removal of frequencies: After all, oscillators require the reception of the other oscillator's signals for synchronization. Frequencies sufficiently different from the oscillator's intrinsic frequency do not influence its phase and thus need not be filtered as well.

There is, however, supporting circuitry that requires low-pass filtering:

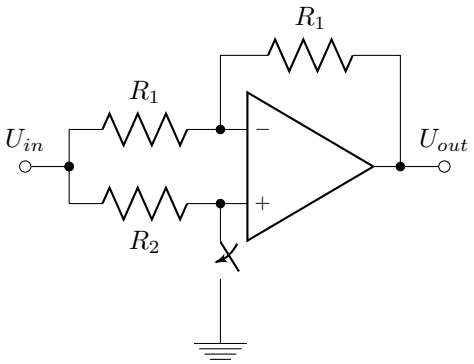
- Voltages proportional to $\cos \Delta\vartheta$ can be constructed with multiplications and low pass filters (Sec. 3.6.3 and 4.1.2). Compared to measurements of the oscillator voltages, they are easier to measure and provide the phase difference directly.



(a) **Inverting summer**: $U_{out} = -\sum_{i=1}^{N} \frac{R_{out}}{R_i} U_i$. Analysis of this circuit is shown in Sec. 2.2.2 as an example. If there is only one input, this circuit is called **inverting amplifier**.



(b) **Non-inverting amplifier:** $U_{out} = \frac{R_1+R_2}{R_1} U_{in}$

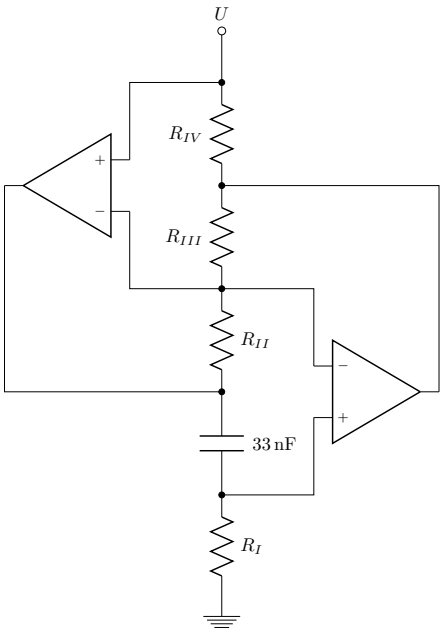


(c) **Conditional inverter / optional inverter**: If the switch is closed, this circuit is an inverting amplifier with gain 1, inverting the input. If the switch is open, $U_+ = U_{in}$ is mirrored by both U_- and U_{out}, passing the input unchanged.

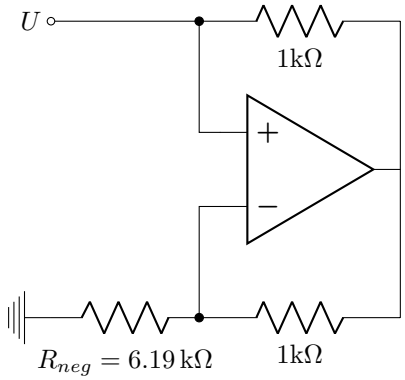
Figure 2.4: Circuits used for summation, amplification and inversion

- Frequencies of oscillators differ in any experiment and can also drift. The readout voltages can be fed into a feedback loop that adjusts frequencies automatically (Sec. 3.6.2 and 4.1.3). The precision of this frequency annealing can be further increased with a low pass filter.

Most importantly, a perfect filter - corresponding to a step function in frequency space - does not exist. In order to perfectly analyze and separate frequencies of a signal, the whole signal, including it's future evolution, must be known. Imagine physicists who Fourier transform a partially known signal in order to remove higher frequencies: They can approximate the correct solution with window functions or continuation of the known part to both past and future, but they will get an imper-



(a) **Gyrator**: The gyrator circuit emulates an inductance $L = R_I R_{II} R_{IV} C / R_{III}$. Compared to a coil, the gyrator has greater thermal stability and accuracy. Additionally, the inductance can be changed by varying a resistor. (See also [43], [37] chapter 5.10.)



(b) **Negative impedance converter (NIC)**: Acts as a negative resistance by supplying negative current corresponding to the supplied voltage U. The absolute value of the negative resistance is equal to the value of the resistor marked as R_{neg}.

Figure 2.5: Subcircuits of the Van der Pol oscillator

fect result. An electronic filter is similarly limited, but additionally it is supposed to output the filtered signal at once. In practice, low-pass filters show non-ideal characteristics and filter design is usually a trade-off for a specific use case:

- Real filters do not show a step-like frequency response at the cutoff-frequency f_{3dB}. Instead, they show a constant slope in a double-logarithmic plot of amplification("gain") versus frequency ("Bode plot") at high frequencies.
- At frequencies lower than f_{3dB} (passband), the signal is approximately undamped.
- f_{3dB} does not separate passband and the so-called stopband perfectly: There is a transition region around f_{3dB}, the "knee curve".
- Ripples can exist in passband or stopband depending on low-pass type.

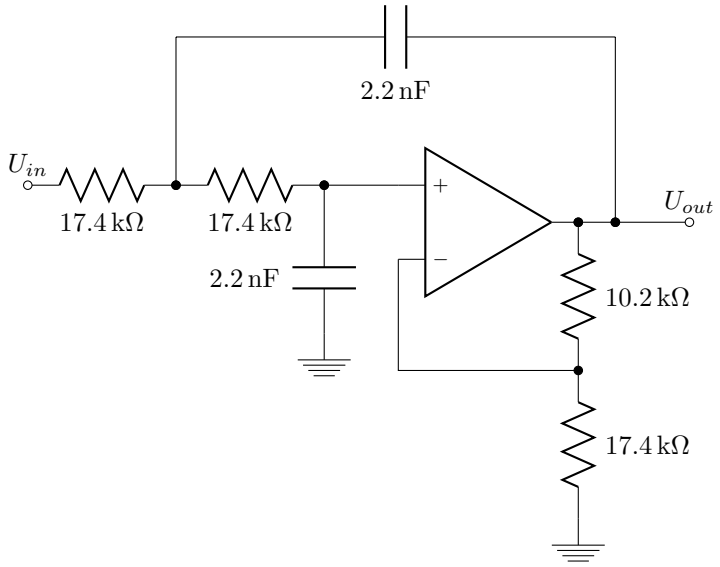


Figure 2.6: **Low pass**: Removes higher frequency components from the input signal with a cutoff frequency $f_{3dB} = 4{,}16\,\text{Hz}$. While many quantities like number of components, steepness of frequency response, delay behavior etc. can be adjusted with the filter design, this specific filter type is a solid compromise, a 2-pole Butterworth-filter. This specific low pass circuit is called a voltage-controlled voltage-source (VCVS) filter. The low-pass filter is optimized for the readout mechanism and the frequency annealing implementation shown in Sec. 4.1.2 and 4.1.3.

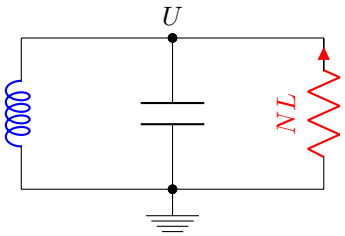
- The filtered frequencies are delayed as well. The delay or the respective phase shift depend on the frequency.
- Filters can be designed with a specific "response", where trade-offs between steepness of transition, ripples in pass- and/or stopband and delay behaviors tailor a filter to fit a specific use case.
- Filter response in the time domain varies between filters as well.
- Additionally, there is the "order" of a filter: Filters of higher order need more components, are more complex and have a larger signal delay, but they have a steeper slope in the Bode plot.

The active low-pass used is the experiment is shown in Fig. 2.6. More details can be found in [37] chapter 5, "active filters", but are not essential for understanding this thesis. Summing up, a low-pass cannot remove a high frequency perfectly. Additionally, the signal will be slightly delayed, and response to fast changes in time depends heavily on the filter. Finally, electric oscillator are required for any network:

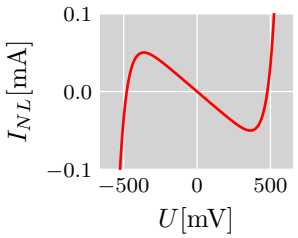
2.2.3 Van der Pol oscillators

Electrical Van der Pol oscillators consist of an LC circuit and an active nonlinear element.(Fig. 2.7a) The nonlinear element supplies energy at low voltages, which sustains the oscillations. See [44] for an explanation of the oscillation mechanism and Fig. 2.7b for a I-U curve of the nonlinear element used. Oscillations can be both sinusoidal or relaxational depending on L,C, and R_{neg}. An extensive analysis of parameter dependence is done in [45]. A circuit implementation with good frequency stability, tunable frequencies and sinusoidal oscillations was published in [31, 45] and is shown in Fig. 2.7c: The nonlinear element (red) is implemented with a negative impedance converter (NIC; see Fig. 2.5b and [37], Chapter 5.03) and two diodes. The NIC acts as a negative resistor, which means it supplies a current proportional to the voltage, but with opposing sign. The diodes limit the negative resistance, so the nonlinear element supplies power at low voltages, but damps at high voltages. As coils are prone to temperature drifts and their inductance does not have high precision, a gyrator circuit (blue) is used instead. Apart from a precisely defined inductance, manual frequency tuning is enabled by placing a potentiometer inside the gyrator circuit. The gyrator and its modifications will be discussed in more detail in Sec. 4.1.3.

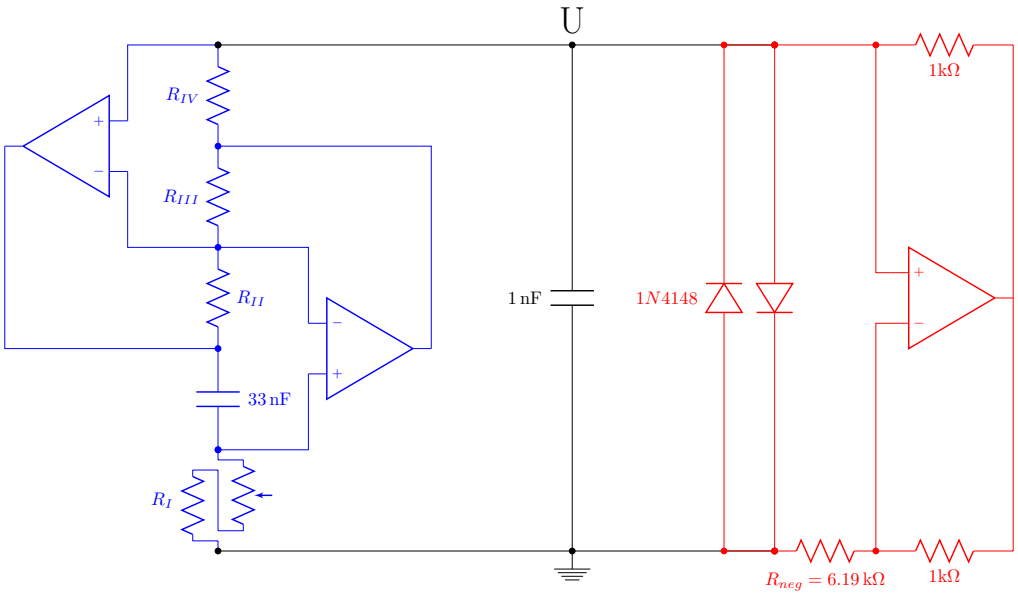
Now, with foundations in both nonlinear dynamics and analog circuitry, our novel network architecture is presented and analyzed in the next chapter.



(a) Idealized circuit:
A electrical Van der Pol oscillator can be built by adding a non-linear element (NL) to a LC circuit. The non-linear element ideally has a cubic I-U-curve, thus showing differential negative resistance between the inflection points [44].



(b) I-U-curve of the nonlinear element used (N-type). It consists of a negative resistance limited by two diodes. ($R_{neg} = -6{,}19\,\mathrm{k\Omega}$. Model and constants for the diodes are from [45])



(c) Actual Van der Pol oscillator circuit used in [31, 45] and this thesis. The coil in the idealized circuit 2.7a is replaced by a gyrator circuit, which offers better precision and temperature stability and allows for frequency tuning. The negative resistance consists of a negative impedance converter (NIC; See [37], Chapter 5.03.) limited by two diodes. In [31, 45], $R_{II} = R_{III} = R_{IV} = 1\,\mathrm{k\Omega}$, R_I consisted of a $1\,\mathrm{k\Omega}$ potentiometer in series with a $100\,\Omega$ resistor. Necessary modifications of the gyrator are described in Sec. 4.1.3. Note that the oscillators lose energy to the coupling circuitry, so coupling resistors have to be comparable to get similar oscillation shapes and phase response curves (Sec. 4.1.1).

Figure 2.7: Electrical Van der Pol oscillator

Chapter 3

Theoretical results

3.1 A new scalable architecture

The proposed architecture consists of two identical networks of N oscillators each with equal frequency distribution. Oscillators within each of these "subnetworks" are globally coupled and the coupling strength is additionally modulated in time. For the first network, the coupling modulation[1] is constructed from products of signals of the second network's oscillators and vice versa. Due to its symmetrical layout, which is visualized in Fig. 3.1, we name the network the MONACO-Architecture: *Mirrored Oscillator Networks for Autoassociative COmputation*.

Motivated by experiments with networks of electrical Van der Pol oscillators [30, 31], we assume that the oscillators are weakly coupled in one variable, have sinusoidal signals and a phase response curve proportional to a cosine. Then, the recognition dynamics can be reduced to a phase description [46]:

$$\begin{aligned}
\dot{\vartheta}_i^{[1]} &= \Omega_i^{[1]} + \cos\vartheta_i^{[1]} \cdot a^{[2]}(t) \cdot \frac{\epsilon}{N} \sum_{j=1}^{N} \sin\vartheta_j^{[1]} \\
\dot{\vartheta}_i^{[2]} &= \Omega_i^{[2]} + \cos\vartheta_i^{[2]} \cdot a^{[1]}(t) \cdot \frac{\epsilon}{N} \sum_{j=1}^{N} \sin\vartheta_j^{[2]} \\
a^{[1]}(t) &= \sum_{k,l=1}^{N} S_{kl} \sin\vartheta_k^{[1]} \sin\vartheta_l^{[1]} \\
a^{[2]}(t) &= \sum_{k,l=1}^{N} S_{kl} \sin\vartheta_k^{[2]} \sin\vartheta_l^{[2]}
\end{aligned} \tag{3.1}$$

$\vartheta_i^{[1]}$ is the phase of the i^{th} oscillator in the first network and $\Omega_i^{[1]}$ its natural frequency. In the global signal $a^{[2]}(t) \cdot \epsilon/N \cdot \sum_{j=1}^{N} \sin\vartheta_j^{[1]}$, $a^{[2]}(t)$ denotes the coupling modulation

[1] Note that coupling modulations were named "coupling functions" by Hölzel [30], but this term is already used differently in the field.

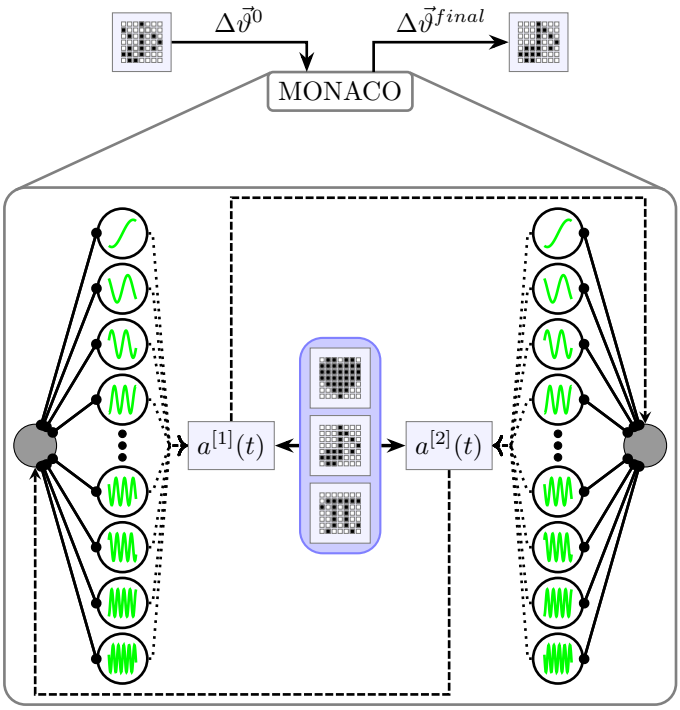


Figure 3.1: Schematics of our new MONACO-Architecture: Oscillators (depicted as black circles surrounding a green sine-wave) are divided into two networks with the same frequency distribution that are both globally coupled. Coupling strength of both global couplings is modulated in time with functions $a^{[1]}(t)$ / $a^{[2]}(t)$ that depend on physical signals of oscillators from the other network and patterns shown in the middle.

generated from the second network's signals and ϵ is a small parameter which will be shown to be the effective coupling strength of the averaged dynamics. The amplitude perturbation is converted into a change in phase by multiplying with the phase response function $\cos\vartheta_i^{[1]}$ and the coupling matrix $\mathbf{S}$ controls attractors of the system.

Note that the frequency distribution is ideally the same in both networks, so N pairs of oscillators with approximately equal angular frequencies $\Omega_i^{[1]} \approx \Omega_i^{[2]}$ exist. We will first assume $\Omega_i^{[1]} = \Omega_i^{[2]} = \Omega_i$ for the main analysis. Afterwards, the influence of frequency differences is outlined in Sec. 3.6.1. For sufficiently weak coupling and specifically chosen frequencies, S_{ij} only effectively connects oscillator pairs i and j and the architecture can act as an autoassociative memory: Apart from $\Omega_i \neq \Omega_j \; \forall i \neq j$, all frequencies Ω_i must be larger than $\Omega_{max}/3$ and all difference frequencies $\Delta\Omega_{ij} = \Omega_i - \Omega_j$ must be pairwise different as shown in Appendix A. As we demonstrate below(Eq. (3.5)), these conditions allow for further simplification of Eq. (3.1).

As the oscillator pairs of equal frequency synchronize at phase differences $\Delta\vartheta_i = \vartheta_i^{[1]} - \vartheta_i^{[2]}$ of either 0 or π ($\pm 2\pi n$) in this setup, the $\Delta\vartheta_i$ are easy to read out (e.g. with one signal multiplication and a low-pass filter) and will be our "system state" to be manipulated. The coupling matrix is chosen according to the Hebbian Rule [34]:

$$S_{ij} = \sum_{m=1}^{M} \alpha_i^m \alpha_j^m \text{ with } \alpha_i \in \{\pm 1\} \tag{3.2}$$

Then attractors will exist for each memorized pattern $\boldsymbol{\alpha}^m$ and its inverse $-\boldsymbol{\alpha}^m$ according to the following $\{\Delta\boldsymbol{\vartheta} \mapsto \boldsymbol{\alpha}\}$-mapping(see also Sec. 3.2):

$$\begin{aligned} 0 + 2\pi n &\mapsto +1 \\ \pi + 2\pi n &\mapsto -1 \end{aligned} \tag{3.3}$$

When we talk about patterns "being attractive", it is meant in the sense that attractors in $\Delta\boldsymbol{\vartheta}$ exist according to this mapping.

Assume a defective pattern $\boldsymbol{\alpha}^d$ should be recognized as a pattern $\boldsymbol{\alpha}^{m'}$, which is the most similar to $\boldsymbol{\alpha}^d$ out of M correct pattern candidates $\boldsymbol{\alpha}^m$. For the recognition, $\boldsymbol{\alpha}^d$ is set as initial condition of the network according to Eq. (3.3) and the coupling matrix S_{ij} contains all correct pattern candidates as memorized patterns according to Eq. (3.2). As the defective pattern is close to the correct pattern in phase space, the system state will move to an attractor representing $\boldsymbol{\alpha}^{m'}$ and can be read out. Note that setting initial conditions is fast and easy in the MONACO-architecture: As the system state is coded into phase differences, simply coupling oscillator pairs with negative or positive sign according to $\Delta\vartheta_i = -\alpha_i^d \cdot E \sin\Delta\vartheta_i$ and $E \gg \epsilon$ for a short time $T_{init} \ll 1/\epsilon$ ensures a correct initialization.

If instead of an erroneous pattern only a small correct part of a pattern is known, missing pixel in $\boldsymbol{\alpha}^d$ can be filled with $+1$ or -1 with equal probability. Afterwards, recognition is performed as above.

Phase differences $\Delta\vartheta_j$ from an exemplary simulation of the phase dynamics (Eq. (3.1)) are shown in Fig. 3.3 for $N = 49$ oscillator pairs and 6 defective pixels. The memorized patterns $\boldsymbol{\alpha}^m$ used are visualized in Fig. 3.2 and are not orthogonal in the sense that $\langle \boldsymbol{\alpha}^{m_1}, \boldsymbol{\alpha}^{m_2} \rangle \neq 0 \ \forall m_1, m_2$ and $m_1 \neq m_2$ ($\langle , \rangle$ denotes the standard scalar product.). The erroneous phase differences change to represent the correct ♪-shaped output pattern.

However, the recognition process can fail if the number of erroneous pixels is too large. A failed recognition is shown in Fig. 3.4: The system state moves to an unknown attractor which corresponds to none of the $\boldsymbol{\alpha}^m$. In order to predict recognition success, a simple criterion is derived and tested in Sec. 3.3. Before analyzing the dynamics, we want to point out that the coupling matrix $\mathbf{S}$ does not need to be wired explicitly, which would require $\mathcal{O}(N^2)$ connections. By rewriting both coupling modulations as squares of scalar products instead, they can be generated with

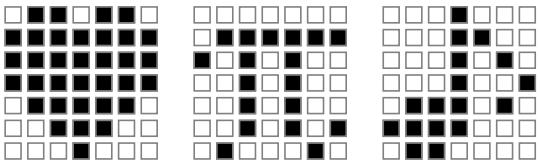


Figure 3.2: Non-orthogonal patterns with 49 pixels that were used as memorized patterns $\boldsymbol{\alpha}^m$ in simulations for Fig. 3.3, Fig. 3.4 and the statistics in Sec. 3.3. $\alpha_i^m = +1$ is visualized as a black pixel and white pixels correspond to $\alpha_i^m = -1$.

$\mathcal{O}(N \cdot M)$ connections only:

$$\begin{aligned} a^{[1/2]}(t) &= \sum_{k,l=1}^{N} \sum_{m=1}^{M} \alpha_k^m \alpha_l^m \sin \vartheta_k^{[1/2]} \sin \vartheta_l^{[1/2]} \\ &= \sum_{m=1}^{M} \Big(\sum_{j=1}^{N} \alpha_j^m \sin \vartheta_j^{[1/2]} \Big)^2 \end{aligned} \tag{3.4}$$

Whenever MONACO is used as an autoassociative memory as presented here, $a^{[1]}(t)$ and $a^{[2]}(t)$ should therefore always be constructed according to Eq. (3.4) instead of Eq. (3.1). Depending on usage, the $\boldsymbol{\alpha}^m$ can be hardwired or changed for each recognition process.

3.2 Analysis of the dynamics

3.2.1 Simplification of the evolution equations

Prior to determining attractors, we simplify the phase equations (Eq. (3.1)) with the technique of averaging [47]: The right hand sides of Eq. (3.1) consist of many different frequency components. If the coupling strength ϵ is sufficiently small, larger frequencies average out on times much smaller than the largest timescale and the smallest frequencies dominate the dynamics:

$$\begin{aligned} \dot{\vartheta}_i^{[1]} &\approx \Omega_i + \frac{\epsilon M}{8N} \sin\big(2\Delta\vartheta_i\big) \\ &\quad - \frac{\epsilon}{4N} \sum_{j=1}^{N} S_{ij} \left[\sin\big(\Delta\vartheta_i + \Delta\vartheta_j\big) + \sin\big(\Delta\vartheta_i - \Delta\vartheta_j\big) \right] \\ \dot{\vartheta}_i^{[2]} &\approx \Omega_i - \frac{\epsilon M}{8N} \sin\big(2\Delta\vartheta_i\big) \\ &\quad + \frac{\epsilon}{4N} \sum_{j=1}^{N} S_{ij} \left[\sin\big(\Delta\vartheta_i + \Delta\vartheta_j\big) + \sin\big(\Delta\vartheta_i - \Delta\vartheta_j\big) \right] \end{aligned}$$

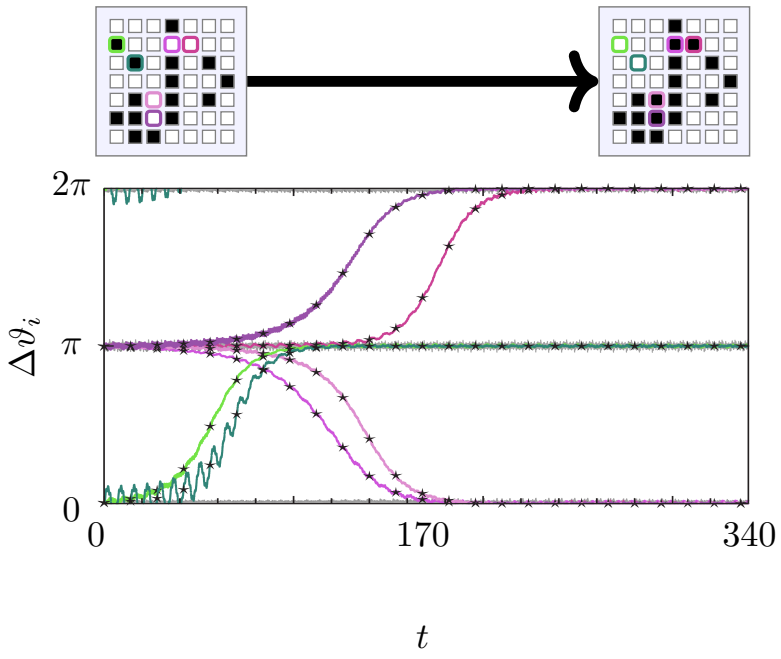


Figure 3.3: Successful Recognition: A binary pattern with 6 erroneous pixels (framed) shown on the top left is correctly recognized as one of 3 memorized patterns shown in Fig. 3.2. White pixels are mapped onto $\Delta\vartheta_i = \pi$ and black pixels correspond to $\Delta\vartheta_i = 0$ or 2π. Trajectories representing erroneous pixels (thick and marked with stars) successfully change by π, which corresponds to an inversion of the pixel. Trajectories corresponding to already correct pixels, however, do not change. For simulation details, see Sec. 3.4.

The lengthy averaging calculation is shown in Appendix A and includes restrictions on the frequency distribution of the oscillators. Using the trigonometric theorem $\sin x \cos y = 1/2\big(\sin(x+y) + \sin(x-y)\big)$, we can express our equation system with the phase differences $\Delta\vartheta_i$ only:

$$\begin{aligned}
\Delta\dot{\vartheta}_i &= \dot{\vartheta}_i^{[1]} - \dot{\vartheta}_i^{[2]} \approx \\
&\quad - \frac{\epsilon}{2N} \sum_{j=1}^{N} S_{ij} \left[\sin\left(\Delta\vartheta_i + \Delta\vartheta_j\right) + \sin\left(\Delta\vartheta_i - \Delta\vartheta_j\right)\right] \\
&\quad + \frac{\epsilon M}{4N} \sin\left(2\Delta\vartheta_i\right) \\
\Delta\dot{\vartheta}_i &= -\frac{\epsilon}{N} \sin\Delta\vartheta_i \Big(\sum_{j=1}^{N} S_{ij} \cos\Delta\vartheta_j - \frac{M}{2} \cos\Delta\vartheta_i \Big)
\end{aligned} \tag{3.5}$$

This is the main evolution equation that governs the dynamics of the architecture.

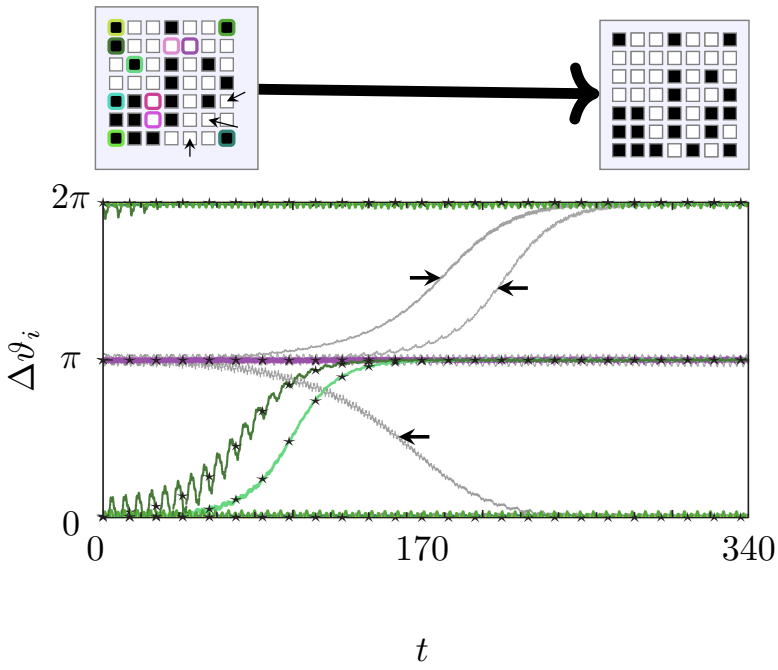


Figure 3.4: Recognition fails due to too many defects: A binary pattern with 11 erroneous pixels (framed) shown on the top left should be recognized as one of three memorized patterns shown in Fig. 3.2. Trajectories representing erroneous pixels (thick and marked with stars) are supposed to change. However, the recognition fails: No trajectories corresponding to erroneous white pixels change from $\Delta\vartheta_i = \pi$ to $\Delta\vartheta_i = 0$ (or 2π). Likewise, only two trajectories representing erroneous black pixels change by π, although seven should change to white. Additionally three trajectories, corresponding to already correct pixels, change to wrong values (Both pixels and trajectories are marked with arrows). The system settles at the pattern shown on the top right, which is none of the memorized patterns. For simulation details, see Sec. 3.4.

3.2.2 Fixed points and their stability

At fixed points $\Delta\boldsymbol{\vartheta}^*$ of the dynamics, all velocity components $\Delta\dot{\vartheta}_i$ must vanish. Depending on which factor in Eq. (3.5) vanishes, pixel indices can be sorted into two sets p and q:

$$\begin{aligned} &\bullet\ i \in p \ \Leftrightarrow\ \sin\Delta\vartheta_i^* = 0 \ \Leftrightarrow\ \Delta\vartheta_i^* \in \{0,\pi\} + 2\pi n \\ &\bullet\ i \in q \ \Leftrightarrow\ \sum_{j=1}^{N} S_{ij}\cos\Delta\vartheta_j^* - \frac{M}{2}\cos\Delta\vartheta_i^* = 0 \end{aligned} \tag{3.6}$$

We show in Appendix B that all fixed points with indices in q are unstable. Therefore, all attractors are well-separated fixed points with $i \in p\ \forall i$.

Only fixed points with $\sin \Delta\vartheta_i^* = 0\ \forall i$ and $\sum_{j=1}^{N} S_{ij} \cos \Delta\vartheta_j^* - M/2 \cos \Delta\vartheta_i^* \neq 0$ remain as candidates for attractors.

The stability of fixed points can generally be examined by linearizing the dynamics around the fixed point by evaluating the eigenvalues of the Jacobian $J_{ik} = \partial \Delta\dot{\vartheta}_i / \partial \Delta\vartheta_k$ at the fixed point $\Delta\boldsymbol{\vartheta}^*$:

$$\begin{aligned} J_{ik} = &- \frac{\epsilon}{N} \cos \Delta\vartheta_i \delta_{ik} \left(\sum_{j=1}^{N} S_{ij} \cos \Delta\vartheta_j - \frac{M}{2} \cos \Delta\vartheta_i \right) \\ &- \frac{\epsilon}{N} \sin \Delta\vartheta_i \left(- S_{ik} \sin \Delta\vartheta_k + \frac{M}{2} \delta_{ik} \sin \Delta\vartheta_i \right) \end{aligned}$$

As $i \in p\, \forall i$ implies $\sin \Delta\vartheta_i^* = 0\, \forall i$, the second term vanishes:

$$J_{ik}(\Delta\boldsymbol{\vartheta}^*) = -\delta_{ik} \frac{\epsilon}{N} \cos \Delta\vartheta_i^* \left(\sum_{j=1}^{N} S_{ij} \cos \Delta\vartheta_j^* - \frac{M}{2} \cos \Delta\vartheta_i^* \right)$$

$\mathbf{J}$ is a diagonal matrix, therefore eigenvectors $\hat{e}_i$ are the standard base with the following eigenvalues:

$$\lambda_i = -\frac{\epsilon}{N} \cos \Delta\vartheta_i^* \left(\sum_{j=1}^{N} S_{ij} \cos \Delta\vartheta_j^* - \frac{M}{2} \cos \Delta\vartheta_i^* \right)$$

We can simplify the analysis further by defining "pattern coordinates" $\boldsymbol{\alpha}$ with $\alpha_i = \cos \Delta\vartheta_i$ as generalization of Eq. (3.3) and inserting the definition of the coupling matrix $\mathbf{S}$:

$$\begin{aligned} \lambda_i &= -\frac{\epsilon}{N} \left(\sum_{m=1}^{M} \alpha_i^m \alpha_i^* \sum_{j=1}^{N} \alpha_j^m \alpha_j^* - \frac{M}{2} \alpha_i^{*2} \right) \\ \lambda_i &= -\frac{\epsilon}{N} \left(\sum_{m=1}^{M} \alpha_i^m \alpha_i^* \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^* \right\rangle - \frac{M}{2} \right) \end{aligned} \tag{3.7}$$

The signs of the eigenvalues determine the stability: Positive eigenvalues denote growing perturbations along the corresponding eigendirection, while negative eigenvalues indicate decay. Therefore, all fixed points with $\lambda_i < 0\, \forall i$ are isolated

attractors:

$$\sum_{m=1}^{M} \alpha_i^m \alpha_i^* \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^* \right\rangle > \frac{M}{2} \qquad \wedge \qquad \alpha_i^* \in \{\pm 1\} \tag{3.8}$$

Memorized patterns map to isolated attractors, if inter-pattern scalar products are sufficiently small. If patterns $\boldsymbol{\alpha}^m$ are orthogonal, inter-pattern scalar products vanish completely:

$$\begin{aligned}
\lambda_i(\boldsymbol{\alpha}^{m'}) &= -\frac{\epsilon}{N}\left(\sum_{m=1}^{M} \alpha_i^m \alpha_i^{m'} \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle - \frac{M}{2}\right) \\
&= -\frac{\epsilon}{N}\left(\sum_{m=1}^{M} \alpha_i^m \alpha_i^{m'} \delta_{mm'} N - \frac{M}{2}\right) \\
&= -\epsilon\left(1 - \frac{M}{2N}\right) < 0 \quad \forall i \\
\Leftrightarrow \quad & M < 2N
\end{aligned} \tag{3.9}$$

Not more than N orthogonal patterns can exist ($\text{span}(\boldsymbol{\alpha}^m) \leq N$, but $\text{span}(\boldsymbol{\alpha}^m) = M$ for linear independent patterns.), so $M < 2N$ is always fulfilled and orthogonal patterns are guaranteed to be stable.

For general $\boldsymbol{\alpha}^m$, we get

$$\begin{aligned}
\lambda_i(\boldsymbol{\alpha}^{m'}) &\overset{!}{<} 0 \\
-\epsilon - \frac{\epsilon}{N}\left(\sum_{m\neq m'}^{M} \alpha_i^m \alpha_i^{m'} \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle - \frac{M}{2}\right) &< 0 \\
-\sum_{m\neq m'}^{M} \alpha_i^m \alpha_i^{m'} \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle &< N - \frac{M}{2}.
\end{aligned}$$

As we want a criterion to ensure that all memorized patterns are attractors, we must exclude that any eigendirection of any pattern becomes unstable:

$$\begin{aligned}
\max_{i,m'}\left(-\sum_{m\neq m'}^{M} \alpha_i^m \alpha_i^{m'} \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle\right) &< N - \frac{M}{2} \\
\Sigma_{max} \overset{!}{=} \max_{m'}\left(\sum_{m\neq m'}^{M} \left|\left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle\right|\right) &< N - \frac{M}{2}
\end{aligned} \tag{3.10}$$

Additionally, if the $\boldsymbol{\alpha}^m$ are attractors, their *inverses will be attractors as well*

because their eigenvalues are identical:

$$\begin{aligned}\lambda_i(-\boldsymbol{\alpha}^{m'}) &= -\frac{\epsilon}{N}\left[\sum_{m=1}^{M}\alpha_i^m(-\alpha_i^{m'})\left\langle\boldsymbol{\alpha}^m,(-\boldsymbol{\alpha}^{m'})\right\rangle-\frac{M}{2}\right]\\ &= -\frac{\epsilon}{N}\left(\sum_{m=1}^{M}\alpha_i^m\alpha_i^{m'}\left\langle\boldsymbol{\alpha}^m,\boldsymbol{\alpha}^{m'}\right\rangle-\frac{M}{2}\right)\\ &= \lambda_i(\boldsymbol{\alpha}^{m'})\end{aligned}$$

Moreover, there are further spurious attractors that do not represent one of the $\boldsymbol{\alpha}^m$, but they are difficult to describe. If the initial pattern does not start in the basin of attraction of an $\boldsymbol{\alpha}^m$, the output of the system will be one of these attractors. Therefore, stability is not sufficient for recognition success and we have to derive a criterion from the basins of attraction. However, first we derive a more common criterion for the network capacity that can be compared in different network architectures.

3.2.3 Error-free capacity

The *error-free capacity* $M_{max}(N)/N$ is a measure for the amount of memorized patterns $\boldsymbol{\alpha}^m$ that can be stored in a given network while any pattern can still be retrieved without errors. Specifically, we determine the maximum number of patterns $M_{max}(N)$ so $P(\boldsymbol{\alpha}^{m'}$ is stable$)\to 1$ for $M < M_{max}(N)$ and $P(\boldsymbol{\alpha}^{m'}$ is stable$)\to 0$ for $M > M_{max}(N)$. Similar to approaches for other architectures, we derive $M_{max}(N)$ in a probabilistic manner for random memorized patterns with $P(\alpha_i^m = +1) = P(\alpha_i^m = -1) = 0.5\,\forall m,i$ in the limes $N\to\infty$.
First, we simplify the rescaled Jacobian $\tilde{\mathbf{J}} = \mathbf{J}/\epsilon$ at a memorized pattern $\boldsymbol{\alpha}^{m'}$:

$$\begin{aligned}\tilde{J}_{ik}(\boldsymbol{\alpha}^{m'}) &= -\delta_{ik}\frac{1}{N}\left(\sum_{m=1}^{M}\alpha_i^m\alpha_i^{m'}\sum_{j=1}^{N}\alpha_j^m\alpha_j^{m'}-\frac{M}{2}\alpha_i^{m'2}\right)\\ &= -\delta_{ik}\frac{1}{N}\left(\sum_{\substack{m=1\\m\neq m'}}^{M}\alpha_i^m\alpha_i^{m'}\sum_{j=1}^{N}\alpha_j^m\alpha_j^{m'}+N-\frac{M}{2}\right)\\ &= -\delta_{ik}\frac{1}{N}\left(\sum_{\substack{m=1\\m\neq m'}}^{M}\sum_{\substack{j=1\\j\neq i}}^{N}\alpha_i^m\alpha_i^{m'}\alpha_j^m\alpha_j^{m'}\right.\\ &\qquad\left.+(M-1)+N-\frac{M}{2}\right)\end{aligned}$$

$$\tilde{\mathbf{J}}(\boldsymbol{\alpha}^{m'}) = -\left(1 + \frac{M-2}{2N}\right)\mathbf{I} + \mathbf{D}$$

Here, $\mathbf{I}$ is the identity matrix and

$$D_{ik} = -\delta_{ik}\frac{1}{N}\sum_{\substack{m=1\\m\neq m'}}^{M}\sum_{\substack{j=1\\j\neq i}}^{N}\alpha_i^m\alpha_i^{m'}\alpha_j^m\alpha_j^{m'}.$$

As $\tilde{\mathbf{J}}$, $\mathbf{I}$ and $\mathbf{D}$ are diagonal,

$$\begin{aligned}\lambda_{max}(\tilde{\mathbf{J}}) &= -\left(1 + \frac{M-2}{2N}\right) + \lambda_{max}(\mathbf{D})\\ &= -\left(1 + \frac{M-2}{2N}\right) + \max_i D_{ii}.\end{aligned}$$

Then, the stability condition can be expressed as function of $\max_i D_{ii}$ alone:

$$\begin{aligned}&&\lambda_{max}(\mathbf{J}) &< 0\\ \Leftrightarrow\quad&&\lambda_{max}(\tilde{\mathbf{J}}) &< 0\\ \Leftrightarrow\quad&&\max_i D_{ii} &< 1 + \frac{M-2}{2N}\end{aligned} \tag{3.11}$$

The following lemma concerning this largest eigenvalue $\lambda_{max}(\mathbf{D}) = \max_i D_{ii}$ has been proven in [32] as Lemma 6 under the assumption that all pixels of all memorized patterns $\boldsymbol{\alpha}^m$ are randomly chosen with probability $P(+1) = P(-1) = 0.5$: (All occurring logarithms are natural.)

Lemma 1.
Let $x > 0$, and

$$\bar{\beta} = \limsup_{N\to\infty}\frac{M(N)\log(N)}{N}, \quad \underline{\beta} = \liminf_{N\to\infty}\frac{M(N)\log(N)}{N}.$$

If $\bar{\beta} < x^2/2$, then $P(max_i D_{ii} \geq x) \to 0$ as $N \to \infty$.
If $\underline{\beta} > x^2/2$, then $P(max_i D_{ii} \geq x) \to 1$ as $N \to \infty$.

According to Eq. (3.11), $\boldsymbol{\alpha}^{m'}$ is stable for $N \to \infty$ and $M > 1$ if $\max_i D_{ii} < 1$ is fulfilled. Therefore, we are interested in the probability $P(max_i D_{ii} \geq 1)$ and choose $x = 1$.
$P(\boldsymbol{\alpha}^{m'}$ stable$) = 1 - P(max_i D_{ii} \geq 1) \to 1$ if $\bar{\beta} < 1/2$ and $P(\boldsymbol{\alpha}^{m'}$ unstable$) =$

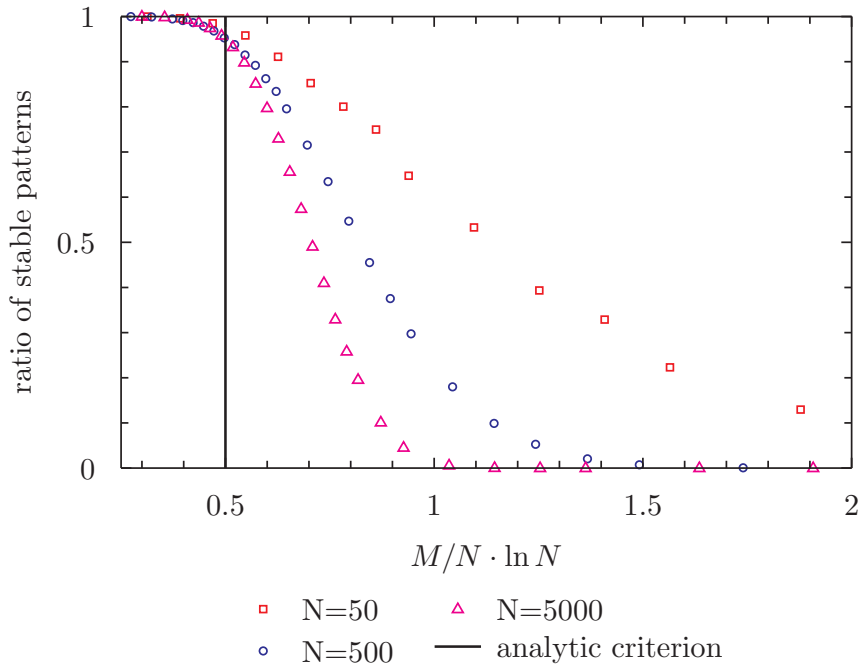


Figure 3.5: The analytic criterion for the error-free capacity Eq. (3.12) is compared to explicit evaluation of the patterns' eigenvalues. For each data point, 1000 sets of pattern were created randomly with $P(\alpha_i^m = +1) = P(\alpha_i^m = -1) = 0.5$ and the stability of each pattern was determined with Eq. (3.8).

$P(max_i D_{ii} \geq 1) \to 1$ if $\underline{\beta} > 1/2$, which implies

$$\begin{aligned} M_{max}(N) &= \frac{N}{2\log(N)} \\ \Leftrightarrow \frac{M_{max}(N)}{N} &= \frac{1}{2\log(N)} \end{aligned} \tag{3.12}$$

Note that other capacity measures exist, such as the *loading rate*, which describes the fraction M_{max}/N under the assumption that attractors for each pattern do exist, but might be shifted, so retrieved patterns might have some errors. Therefore, the *error-free capacity* always is a lower bound on the loading rate. While these probabilistic measures are useful for comparing architectures, their validity is constrained in reality: Real networks are of finite size and memorized patterns need not be chosen randomly. Bounds on guaranteed stability were derived in Eq. (3.10) and a criterion for guaranteed recognition is derived in Sec. 3.3.

3.2.4 Intuitive explanation of the recognition mechanism

The results of the fixed point analysis allow a more intuitive view of Eq. (3.5) by partially expressing the system state in pattern coordinates $\boldsymbol{\alpha}$ with $\alpha_i = \cos \Delta\vartheta_i$:

$$\begin{aligned}
\Delta\dot{\vartheta}_i &= -\frac{\epsilon}{N} \sin \Delta\vartheta_i \left(\sum_{j=1}^{N} S_{ij} \cos \Delta\vartheta_j - \frac{M}{2} \cos \Delta\vartheta_i \right) \\
&= -\frac{\epsilon}{N} \sin \Delta\vartheta_i \left(\sum_{j=1}^{N} \sum_{m=1}^{M} \alpha_i^m \alpha_j^m \alpha_j - \frac{M}{2} \alpha_i \right) \\
&= -\sin \Delta\vartheta_i \cdot \frac{\epsilon}{N} \left(\sum_{m=1}^{M} \alpha_i^m \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha} \right\rangle - \frac{M}{2} \alpha_i \right)
\end{aligned}$$

Let $\boldsymbol{\alpha}^{m'}$ be the memorized pattern the system state $\boldsymbol{\alpha}$ is closest to:

$$\begin{aligned}
\Delta\dot{\vartheta}_i = &-\sin \Delta\vartheta_i \cdot \alpha_i^{m'} \cdot \frac{\epsilon}{N} \cdot \\
&\cdot \left(\left\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha} \right\rangle + \sum_{\substack{m=1 \\ m \neq m'}}^{M} \alpha_i^{m'} \alpha_i^m \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha} \right\rangle - \frac{M}{2} \alpha_i^{m'} \alpha_i \right)
\end{aligned}$$

Now assume the system state $\boldsymbol{\alpha}$ is sufficiently close to $\boldsymbol{\alpha}^{m'}$: Then $\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha} \rangle$ is larger than the sum of all other terms in parentheses. Hence, the fixed points and their stability are the same as in $\mathrm{d}/\mathrm{d}t\, \Delta\vartheta_i = -\alpha_i^{m'} \sin \Delta\vartheta_i$. If $\alpha_i^{m'} = +1$, $\Delta\vartheta_i^* = 0$ is stable and $\Delta\vartheta_i^* = \pi$ is unstable and vice-versa for $\alpha_i^{m'} = -1$, so $\lim_{t\to\infty} \alpha_i = \cos \Delta\vartheta_i^* = \alpha_i^{m'}$.

From another point of view, the system "defines" "relative closeness" to memorized patterns by comparing their projections onto the system state $\boldsymbol{\alpha}$. This fails, however, if the scalar products are of comparable size: Then the distribution of the α_i^m matters for each pixel, which leads to spurious attractors unequal to all $\boldsymbol{\alpha}^m$. Note that the $-M\alpha_i/2$-term does not really contribute to the recognition mechanism. While it increases eigenvalues of all stable fixed points slightly, therefore reducing stability (see Eq. (3.7)), it does not influence the basins of attraction much, as we will illustrate in the next section.

3.3 Basins of attraction and guaranteed recognition

We have a firm understanding of the system now and can guarantee that the chosen patterns $\boldsymbol{\alpha}^m$ are attractive. However, we cannot guarantee recognition success yet: The system state might relax to the additional unwanted attractors described by

Eq. (3.8) or even worse, the basins of attraction of the $\boldsymbol{\alpha}^m$ might be malformed, leading to a $\boldsymbol{\alpha}^m$ whose projection on the defective pattern is not the largest.

3.3.1 Lower bound on the basins of attraction

Matching success is guaranteed if the defective starting pattern is in the basin of attraction of the correct memorized pattern $\boldsymbol{\alpha}^{m'}$. A lower bound on the basin of attraction can be derived by proofing the following lemmata:

1. Surfaces of constant projection on the correct memorized pattern $\boldsymbol{\alpha}^{m'}$ confine the system state to larger projections if the initial projection is sufficiently large.

2. $\boldsymbol{\alpha}^{m'}$ is the only attractor inside this confined space.

As the system state cannot leave the confined space, it has to settle on $\boldsymbol{\alpha}^{m'}$ as the only attractor. Therefore, the confined space is part of $\boldsymbol{\alpha}^{m'}$s basin of attraction.

Transformation to α-space

For our following discussion, we will transfer the $\Delta\boldsymbol{\vartheta}_i$-dynamics (Eq. (3.5)) completely into the " pattern coordinates " $\boldsymbol{\alpha}$ with $\alpha_i = \cos\Delta\vartheta_i$, which are a generalization of the mapping of the memorized patterns $\boldsymbol{\alpha}^m$.

$$
\begin{aligned}
\dot{\alpha_i} &= (\cos\dot{\Delta\vartheta_i}) \\
&= \frac{\partial\cos\Delta\vartheta_i}{\partial\Delta\vartheta_i}\frac{\partial\Delta\vartheta_i}{\partial t} \\
&= -\sin\Delta\vartheta_i\left[-\frac{\epsilon}{N}\sin\Delta\vartheta_i\left(\sum_{j=1}^{N} S_{ij}\cos\Delta\vartheta_j - \frac{M}{2}\cos\Delta\vartheta_i\right)\right] \\
&= \frac{\epsilon}{N}\sin^2\Delta\vartheta_i\left(\sum_{j=1}^{N}\sum_{m=1}^{M}\alpha_i^m\alpha_j^m\cos\Delta\vartheta_j - \frac{M}{2}\cos\Delta\vartheta_i\right) \\
&= \frac{\epsilon}{N}(1-\cos^2\Delta\vartheta_i)\left(\sum_{j=1}^{N}\sum_{m=1}^{M}\alpha_i^m\alpha_j^m\cos\Delta\vartheta_j - \frac{M}{2}\cos\Delta\vartheta_i\right) \\
&= \frac{\epsilon}{N}(1-\alpha_i^2)\left(\sum_{m=1}^{M}\alpha_i^m\sum_{j=1}^{N}\alpha_j^m\alpha_j - \frac{M}{2}\alpha_i\right) \\
\dot{\alpha_i} &= \frac{\epsilon}{N}(1-\alpha_i^2)\left(\sum_{m=1}^{M}\alpha_i^m\left\langle\boldsymbol{\alpha}^m,\boldsymbol{\alpha}\right\rangle - \frac{M}{2}\alpha_i\right) \qquad (3.13)
\end{aligned}
$$

Note that although the mapping between $\Delta\vartheta_i$ and α_i is *not* injective, the transformation is still valid: Eq. (3.5) is mirror-symmetric to $0 + \pi n$ with $n \in \mathbb{N}$, so space can be divided into regions separated by $\Delta\vartheta_i = [0, \pi] + 2\pi n$ or $\Delta\vartheta_i = [\pi, 2\pi] + 2\pi n$ in every i and flow lines in each region are mapped onto the same $\boldsymbol{\alpha}$-coordinates. As the flow across the boundaries of these hypercubes is zero, it is not necessary to consider the periodicity of the flow. From another point of view, the ambiguity of attractors in $\Delta\boldsymbol{\vartheta}$ is removed in the $\boldsymbol{\alpha}$-coordinates. As the dynamics of $\boldsymbol{\alpha}$ do not depend on the sign or periodicity of $\Delta\boldsymbol{\vartheta}$, it is a more natural coordinate for the autoassociative memory.

Confinement by hypersurfaces of constant projection

Let's consider a hypersurface of constant projection on the correct output pattern $\boldsymbol{\alpha}^{m'}$: In the pattern coordinates the equation $\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \rangle = C$ describes a hyperplane that divides the N-dimensional hypercube of all possible patterns into patterns with a projection larger or smaller than C. If projections on $\boldsymbol{\alpha}^{m'}$ do not decrease for all points on the surface, the system state can only move tangential to the hyperplane or towards larger projections. (Movement tangential to the hyperplane is in fact impossible with a slightly stricter condition, as shown further below.)

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{|\boldsymbol{\alpha}^{m'}|} \left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle &\geq 0 \\ \frac{\mathrm{d}}{\mathrm{d}t} \left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle = \left\langle \dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}^{m'} \right\rangle &\geq 0 \\ \sum_{i=1}^{N} (1 - \alpha_i^2) \alpha_i^{m'} \left(\sum_{m=1}^{M} \alpha_i^m \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha} \right\rangle - \frac{M}{2} \alpha_i \right) &\geq 0 \end{aligned} \tag{3.14}$$

If Eq. (3.14) is fulfilled for all $\boldsymbol{\alpha}$ on a hypersurface $\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \rangle = C$, it confines the system state. However, to exclude additional attractors besides $\boldsymbol{\alpha}^{m'}$ in the confined space is difficult with Eq. (3.14) and a good criterion for guaranteed recognition should neither depend on the hyperplanes nor on the specific pixels of $\boldsymbol{\alpha}$ or the memorized patterns $\boldsymbol{\alpha}^m$. Therefore, we employ a series of *worst-case approximations and upper bounds*:

Eq. (3.14) is fulfilled if all single summands are greater than zero. Note that this approximation also excludes movement tangential to the hypersurfaces: Without the possibility for summands to cancel each other, $\mathrm{d}/\mathrm{d}t \left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle = 0$ is only fulfilled if $\dot{\boldsymbol{\alpha}} = 0$, so all remaining solutions are fixed points. Then the following inequalities must hold $\forall i$ and $\forall \boldsymbol{\alpha}$ on the surface:

$$\underbrace{(1 - \alpha_i^2)}_{\geq 0} \alpha_i^{m'} \left(\sum_{m=1}^{M} \alpha_i^m \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha} \right\rangle - \frac{M}{2} \alpha_i \right) \geq 0$$

$$\alpha_i^{m'} \sum_{m=1}^{M} \alpha_i^m \langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle - \frac{M}{2} \alpha_i^{m'} \alpha_i \geq 0$$

$$\left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle \geq - \sum_{\substack{m=1 \\ m \neq m'}}^{M} \alpha_i^m \alpha_i^{m'} \langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle + \frac{M}{2} \alpha_i^{m'} \alpha_i$$

As the left hand side is constant on a hypersurface, the criterion needs to be evaluated for a maximized right hand side only and the criterion for the surface can be reduced to one single inequality:

$$\left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle \geq \max_{i, \boldsymbol{\alpha}} \left(- \sum_{\substack{m=1 \\ m \neq m'}}^{M} \alpha_i^m \alpha_i^{m'} \langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle + \frac{M}{2} \alpha_i^{m'} \alpha_i \right)$$

The sum is maximal in i for $\alpha_i^m = -\alpha_i^{m'} \operatorname{sgn}(\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle)\ \forall m \neq m'$, as all scalar products add up. (If such an i always exists is not relevant here, as we look for a worst case approximation independent of the $\boldsymbol{\alpha}^m$.) The second term is generally much smaller, but $M/2$ at most:

$$\max_{i, \boldsymbol{\alpha}} \left(- \sum_{\substack{m=1 \\ m \neq m'}}^{M} \alpha_i^m \alpha_i^{m'} \langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle + \frac{M}{2} \alpha_i^{m'} \alpha_i \right) \leq \quad \max_{\boldsymbol{\alpha}} \left(\sum_{\substack{m=1 \\ m \neq m'}}^{M} |\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle| + \frac{M}{2} \right)$$

As the maximum of one single $|\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle|$ is much easier to calculate, we approximate an upper bound:

$$\max_{\boldsymbol{\alpha}} \left(\sum_{\substack{m=1 \\ m \neq m'}}^{M} |\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle| + \frac{M}{2} \right) \leq \sum_{\substack{m=1 \\ m \neq m'}}^{M} \max_{\boldsymbol{\alpha}} \left(|\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle| \right) + \frac{M}{2}$$

In total, our criterion on the hypersurface has reduced to

$$C = \left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle \geq \sum_{\substack{m=1 \\ m \neq m'}}^{M} \max_{\boldsymbol{\alpha}} \left(|\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle| \right) + \frac{M}{2}. \tag{3.15}$$

While any hyperplane that fulfills Eq. (3.15) confines the system state to larger projections, it is still not trivial to evaluate due to the direct dependence on $\boldsymbol{\alpha}$.

Removing direct dependence on $\boldsymbol{\alpha}$

$\max_{\boldsymbol{\alpha}}(|\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle|)$ can be approximated as a function of $\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \rangle = C$ and inter-pattern scalar products.

First, $\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \rangle$ is expressed with the difference vector $\Delta\boldsymbol{\alpha} = \boldsymbol{\alpha} - \boldsymbol{\alpha}^{m'}$ between $\boldsymbol{\alpha}$ and the closest memorized pattern $\boldsymbol{\alpha}^{m'}$:

$$\begin{aligned}\left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle &= \left\langle \boldsymbol{\alpha} - \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^{m'} \right\rangle + \left\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^{m'} \right\rangle \\ &= \left\langle \Delta\boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle + N \\ &= \sum_{i=1}^{N} \Delta\alpha_i \alpha_i^{m'} + N\end{aligned}$$

With $\operatorname{sgn}(\Delta\alpha_i) = \operatorname{sgn}\big(\alpha_i^{m'}(\underbrace{\alpha_i \alpha_i^{m'}}_{\leq 1} - 1)\big) = -\alpha_i^{m'}$ we get:

$$\left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle = N - \sum_{i=1}^{N} |\Delta\alpha_i| \tag{3.16}$$

$$\begin{aligned}\Rightarrow \max_{\boldsymbol{\alpha}} \left(\left| \langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle \right| \right) &= \max_{\Delta\boldsymbol{\alpha}} \left(\left| \left\langle \Delta\boldsymbol{\alpha}, \boldsymbol{\alpha}^m \right\rangle + \left\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^m \right\rangle \right| \right) \\ &< \max_{\Delta\boldsymbol{\alpha}} \left(\sum_i \Delta\alpha_i \alpha_i^m \right) + \left| \left\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^m \right\rangle \right| \\ &= \sum_i \left| \Delta\alpha_i \right| + \left| \left\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^m \right\rangle \right| \\ &= N - \left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle + \left| \left\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^m \right\rangle \right|\end{aligned}$$

Volumes of growing projection

Finally, we can remove all direct dependence on $\boldsymbol{\alpha}$ from Eq. (3.15):

$$\left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle \geq \sum_{\substack{m=1 \\ m \neq m'}}^{M} \max_{\boldsymbol{\alpha}} \left(\left| \left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \right\rangle \right| \right) + \frac{M}{2}$$

$$\left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle \geq \sum_{\substack{m=1 \\ m \neq m'}}^{M} \left(N - \left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle + \left| \left\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^m \right\rangle \right| \right) + \frac{M}{2}$$

$$M \cdot \left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle \geq (M-1) \cdot N + \sum_{\substack{m=1 \\ m \neq m'}}^{M} \left| \left\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^m \right\rangle \right| + \frac{M}{2}$$

$$\left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle \geq \frac{M-1}{M} \cdot N + \frac{1}{M} \sum_{\substack{m=1 \\ m \neq m'}}^{M} \left| \left\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^{m} \right\rangle \right| + \frac{1}{2} \tag{3.17}$$

This final criterion for a confining hyperplane does not depend on a point on the surface.

Additionally, every surface $\left\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \right\rangle = C_{min}$ that fulfills Eq. (3.17) defines a *volume of growing projection* for larger C: As the right hand side of Eq. (3.17) is constant, all hyperplanes with $C > C_{min}$ fulfill the criterion as well.

If several attractors existed in the confined space, however, no conclusion could be made on the basins of attraction, as a confined system state could move to any of them. Therefore, we exclude that any attractor besides $\boldsymbol{\alpha}^{m'}$ exists in a volume of growing projection:

$\boldsymbol{\alpha}^{m'}$ being the only attractor enclosed

Assume an attractor $\boldsymbol{\alpha}^{a}$ exists inside the region defined by Eq. (3.17). Now consider a small perturbation around $\boldsymbol{\alpha}^{a}$ that increases $\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \rangle$, for example $\epsilon \alpha_i^{m'} \cdot \hat{e}_i$ if $\alpha_i^a \neq \alpha_i^{m'}$. As $\mathrm{d}/\mathrm{d}t \langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \rangle \geq 0$ in the confined space, the system cannot relax back to $\boldsymbol{\alpha}^{a}$. No non-isolated attractor exists (see Sec. 3.2), so $\boldsymbol{\alpha}^{a}$ has at least one unstable eigendirection which contradicts the assumption that $\boldsymbol{\alpha}^{a}$ is an attractor.

The only exception is the attractor $\boldsymbol{\alpha}^{m'}$ itself: As it has the largest projection on itself, all perturbations must lower $\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \rangle$.

Summing up: Every system state $\boldsymbol{\alpha}$ that obeys Eq. (3.17) must be in the basin of attraction of $\boldsymbol{\alpha}^{m'}$, as projection on $\boldsymbol{\alpha}^{m'}$ increases monotonically along the trajectory and $\boldsymbol{\alpha}^{m'}$ is the only attractor for larger projections.

3.3.2 Guaranteed recognition

Recognition criteria

As any defective initialized pattern is binary, it can be characterized by the number of defective pixels n^f in which defective input pattern and correct memorized pattern are different. Eq. (3.17) can be solved for n^f with Eq. (3.16), as n^f is a special case of $\sum_i |\Delta \alpha_i| / 2$:

$$N - 2n^f > \frac{M-1}{M} \cdot N + \frac{1}{M} \sum_{\substack{m=1 \\ m \neq m'}}^{M} \left| \left\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^{m} \right\rangle \right| + \frac{1}{2}$$

$$n^f < \frac{1}{2M}\left(N - \sum_{\substack{m=1 \\ m\neq m'}}^{M} \left|\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^{m}\rangle\right|\right) - \frac{1}{4} \tag{3.18}$$

(The equality in Eq. (3.17) must be dropped here, as perturbations and higher order terms neglected in Eq. (3.5) might push a defective pattern on the outermost hyperplane out of the confined space.)
For *pairwise orthogonal patterns*, $\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^{m}\rangle = 0 \; \forall m \neq m'$ and Eq. (3.18) becomes:

$$\boxed{n^f < \frac{N}{2M} - \frac{1}{4}} \tag{3.19}$$

We now treat *general patterns* with $\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^{m}\rangle \neq 0$. A criterion that does not depend on the correct memorized pattern $\boldsymbol{\alpha}^{m'}$ is obtained with the definition $\Sigma_{max} = \max_{\boldsymbol{\alpha}^{\tilde{m}}}(\sum_{m=1,m\neq\tilde{m}}^{M} |\langle \boldsymbol{\alpha}^{\tilde{m}}, \boldsymbol{\alpha}^{m}\rangle|) > \sum_{m=1,m\neq m'}^{M} |\langle \boldsymbol{\alpha}^{m'}, \boldsymbol{\alpha}^{m}\rangle|$ from Sec. 3.2. Then the worst case of Eq. (3.18) is

$$\boxed{n^f < \frac{N - \Sigma_{max}}{2M} - \frac{1}{4}}. \tag{3.20}$$

Eq. (3.20) guarantees successful recognition for arbitrary patterns.

Consistency check

The basin of attraction has to vanish when the fixed point looses stability. Therefore, we can regain stability criteria for the $\boldsymbol{\alpha}^m$ by minimizing the necessary extension of the basin of attraction in Eq. (3.19) and Eq. (3.20), which corresponds to $\lim_{n^f\to 0}$:

$$\lim_{n^f\to 0} \text{Eq. (3.19)}: \quad 0 < \frac{N}{2M} - \frac{1}{4}$$
$$M < 2N$$

This coincides with our calculation that pairwise orthogonal patterns are always stable: At most, N orthogonal patterns can exist, as they are linear independent and $\dim(\text{span}(\{\boldsymbol{\alpha}^m\})) \leq N$, so $M < 2N$ is always fulfilled.

$$\lim_{n^f\to 0} \text{Eq. (3.20)}: \quad 0 < \frac{N - \Sigma_{max}}{2M} - \frac{1}{4}$$
$$\Sigma_{max} < N - \frac{M}{2}$$

This again reproduces our result for the stability of non-orthogonal patterns.

3.4 Numerical simulations

In this section we validate our criterion for successful pattern recognition with simulations of the full phase dynamics Eq. (3.1).

3.4.1 Numerical methods and parameters

The equations have been implemented in C and integration was performed with the classical Runge-Kutta method. A timestep $dt = 1 \cdot 10^{-4}$ and a coupling strength $\epsilon = 0.1$ were used. The angular frequencies were distributed according to $\Omega_i = 1200 + 1800 \cdot G_i/G_N$, where G_i is the i^{th} element of a Golomb ruler [48]. (See also Appendix A.) The near optimal Golomb rulers used were both taken from [49]: {0, 17, 20, 86, 119, 140, 166, 227, 240, 255, 353, 430, 520, 559, 564, 565, 602, 675, 724, 781, 817, 833, 905, 929, 961, 970, 980, 1131, 1162, 1189, 1212, 1319, 1403, 1433, 1437, 1451, 1462, 1497, 1504, 1589, 1601, 1680, 1763, 1785, 1825, 1880, 1888, 1956, 1958} for $N = 49$ and {0, 34, 44, 91, 95, 147, 207, 278, 332, 364, 375, 405, 458, 520, 682, 698, 701, 710, 853, 868, 901, 946, 973, 1022, 1080, 1150, 1155, 1172, 1240, 1254, 1290, 1429, 1540, 1546, 1605, 1642, 1682, 1684, 1705, 1751, 1771, 1806, 1835, 1943, 1967, 2041, 2151, 2164, 2182, 2189, 2190, 2270} for $N = 52$.

For simulations in Fig. 3.3 and Fig. 3.4, defective patterns were chosen manually and memorized patterns are taken from Fig. 3.2. All pseudorandom numbers (necessary for random distribution of erroneous pixels and construction of random orthogonal patterns) were created using C's standard random number generator rand() from stdlib, which was seeded with the time in microseconds times the process ID.

3.4.2 Testing criteria for guaranteed recognition

In order to test criteria Eq. (3.19) and (3.20), simulations were performed for both the non-orthogonal patterns shown in Fig. 3.2 with N=49 pixels as well as for 3 random orthogonal patterns with N=52 pixels. Simulations started after setting the initial conditions to a defective pattern similar to one of the memorized patterns but different in exactly n^f randomly distributed erroneous pixels. In order to save simulation time, simulations were aborted if the system state reached one of the memorized patterns, as they are proven to be attractors. In all other cases, simulations were continued until $|\alpha_i| \geq 0.9\,\forall i$ for a period $t_{wait} = 500$. Recognition success was tested by projecting the $\boldsymbol{\alpha}$-coordinates of the final system state on the memorized patterns: If $\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^{m'} \rangle / N > 0.99$, recognitions were counted as successful.

For the non-orthogonal patterns with $M = 3$, $N = 49$ and $\Sigma_{max} = 10$, the recognition criterion Eq. (3.20) predicts recognition success for $n^f < (N - \Sigma_{max})/(2M) - 0.25 = 6.25$. 300 simulations were performed for $n^f \in \{6..16\}$ for each pattern and results are summed up in Table 3.1. All recognitions were successful for $n^f \leq 11$ and

n^f	6-11	12	13	14	15	16
failed recognitions ♥	0	0	0	1	5	13
failed recognitions ♪	0	2	2	3	3	15
failed recognitions π	0	1	2	6	6	22
failure rate (%)	0	0.3	0.4	1.1	1.6	5.6

Table 3.1: Failed Recognitions with non-orthogonal patterns as shown in Fig. 3.2. 300 recognitions were performed for each pattern and each number of erroneous pixels n^f. Erroneous pixels were distributed randomly for each simulation.

n^f	8-12	13	14	15	16	17
failed recognitions	0	1	1	4	13	29
failure rate (%)	0	0.1	0.1	0.4	1.3	2.9

Table 3.2: Failed Recognitions with random orthogonal patterns with $N = 52$ pixels. 1000 recognitions were performed for each number of erroneous pixels n^f. Random distribution of erroneous pixels and the construction of random orthogonal patterns was repeated for each simulation.

the rate of failed recognitions grows slowly for larger n^f. Obviously, our criterion seems to be too strict.

Similarly, 1000 simulations were performed with orthogonal random patterns with $N = 52$ and $M = 3$ for each $n^f \in \{8..17\}$. Here, $n^f < N/(2M) - 0.25 = 8.42$ is predicted by Eq. (3.19). Random orthogonal patterns were constructed by using the elementwise product $\circ$: As orthogonal patterns with $\alpha_i \in \pm 1$ differ in exactly $N/2$ pixels, a pattern $\boldsymbol{\alpha}^2$ orthogonal to any pattern $\boldsymbol{\alpha}^1$ can be easily found by creating a "difference vector" $\boldsymbol{d}^{1,2}$, where $N/2$ $+1$- and -1-entries are randomly distributed. Then $\boldsymbol{\alpha}^2 = \boldsymbol{\alpha}^1 \circ \boldsymbol{d}^{1,2}$.

For 3 orthogonal patterns, $\boldsymbol{\alpha}^1$, $\boldsymbol{d}^{1,2}$ and $\boldsymbol{d}^{1,3}$ were first chosen randomly. Then $|\langle \boldsymbol{\alpha}^2, \boldsymbol{\alpha}^3 \rangle| = |\langle \boldsymbol{d}^{1,2}, \boldsymbol{d}^{1,3} \rangle|$ was minimized by switching 2 randomly selected pixels in a randomly selected difference vector, if the absolute value of the scalar product diminished.

Results are summed up in Table 3.2. Similar to the simulations with non-orthogonal patterns, recognitions are always successful for $n^f \leq 12$, which is significantly larger than predicted by the criterion for guaranteed recognition. For even larger n^f, the rate of failed recognitions stays small.

3.4.3 Failed recognitions are rare events

One might expect that the criterion for guaranteed recognition is not optimal for both the orthogonal random patterns and our choice of non-orthogonal patterns, so that 7 respectively 9 erroneous pixels or even more can always be correctly recognized as well. However, failed recognitions are just rare for $n^f = 7$ / $n^f = 9$ instead. We

now construct problematic starting patterns with $n^f = 7$ for the non-orthogonal memorized patterns that fail in the recognition process:

According to Eq. (3.15), recognition will fail if the scalar products between the defective starting pattern and non-similar memorized patterns are extremized. Considering the scalar products $\langle \boldsymbol{\alpha}^{\heartsuit}, \boldsymbol{\alpha}^{\pi} \rangle = +5$, $\langle \boldsymbol{\alpha}^{\heartsuit}, \boldsymbol{\alpha}^{\eighthnote} \rangle = -5$, and $\langle \boldsymbol{\alpha}^{\pi}, \boldsymbol{\alpha}^{\eighthnote} \rangle = -1$, an erroneous heart-pattern is most likely to fail. Assume furthermore that the number of erroneous pixels n^f is fixed. Then the right hand side of Eq. (3.15) can be maximized by distributing the errors on positions where they increase the projection on the π- and decrease the projection on the ♪-pattern. 10 such "worst-case" positions can be found for the ♥-pattern and $\binom{10}{7} = 120$ possible combinations exist to distribute $n^f = 7$ erroneous pixels on the "worst-case" positions.

Simulations were performed for all of these "worst case patterns". Recognition failed for all simulations and the system state relaxed to an attractor with projections of 0.59, -0.51 and 0.51 on the ♥-, ♪- and π-pattern. Possible worst-case positions for erroneous pixels and the irregular output pattern are shown in Fig. 3.6. Indeed, simulations with randomly distributed errors could not recognize this:

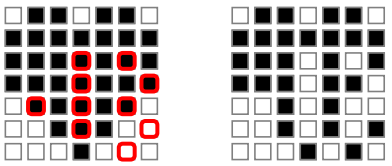


Figure 3.6: On the left side, an unperturbed ♥-pattern is shown. Erroneous pixels on red-circled positions extremize the sum of inter-pattern scalar products. All erroneous ♥-patterns with 7 erroneous pixels on marked locations fail the recognition process. Simulations of all such patterns resulted in the spurious attractor shown on the right.

As there are $\binom{49}{7} \approx 8,6 \cdot 10^7$ possibilities to distribute the erroneous pixel on the pattern and only $\binom{10}{7} = 120$ worst case distributions can be found, the chance to encounter a failing random starting pattern is almost negligible. Furthermore, all $\binom{10}{6} = 210$ worst-case-patterns for $n^f = 6$ were successfully recognized as the ♥-pattern in simulations, which again validates Eq. (3.20) as criterion for guaranteed recognition. Similar calculations can be performed for the orthogonal case. This is a good example that extracting basins of attractions in high-dimensional systems with simulations can only give an approximation on the success rate but no guaranteed criterion. From another point of view, failed recognitions are rare, so a higher n^f is acceptable if a non-perfect recognition rate is sufficient.

3.5 Architecture discussion and comparison

MONACO gains its distinctive properties from two design features:

1. *Two mirrored globally coupled subnetworks* are used.
   First of all, the use of two groups enables the internal generation of the coupling modulations. Second, the effective coordinates of the network are phase differences $\Delta\vartheta_i$ of oscillators of equal frequency. Values of phase differences can easily be read out by multiplying signals of an oscillator pair and using a low-pass filter, gaining $\cos\Delta\vartheta_i$. Similarly, setting the initial conditions requires only positive or negative coupling between two oscillators forming a pair. Third, the effective average coupling strength ϵ is doubled with two subnetworks, enabling faster recognition (compare with Appendix A). The fourth advantage is much subtle: In all architectures with externally generated coupling modulations, frequencies in the coupling modulation are fixed to the natural frequencies of the uncoupled oscillators. However, the so-called "acceleration effect" [50] changes the frequencies of even weakly coupled oscillators. Any mismatches between oscillator frequency and coupling modulation frequency components would further limit the coupling strength ϵ. As a higher coupling strength reduces recognition time, we decided to avoid the problem altogether: Since oscillators in both networks are affected symmetrically by the coupling, the acceleration effect will be equal and frequencies in the coupling modulations are adjusted automatically. It is noteworthy, that the coupling between single oscillator pairs is above the Kuramoto threshold and thus frequencies of the two oscillators adapt. Hence, the architecture allows for some tolerance in the frequency mismatch of an oscillator pair.

2. Novel *coupling modulations* are used.
   As shown in Eq. (3.4), the used coupling modulations can be constructed with $\mathcal{O}(N \cdot M)$ connections only. Note that there cannot be any better scaling, as patterns consist of $N \cdot M$ independent pixels. Additionally, this coupling modulations introduce novel effective dynamics Eq. (3.5), where the only existing attractors are isolated fixed points with $\cos\Delta\vartheta_i \in \{\pm 1\}$ (Section 3.2). As every pixel settles at these binary values, the output is inherently digital, which further simplifies readout and subsequent processing. All memorized patterns are attractive if inter-pattern scalar products are not too large (see Eq. (3.10) for guaranteed stability). As memorized patterns are no transient phenomenon, but long-term stable, readout does not need to be exactly timed and the output can be retrieved at a later time. Furthermore, the dynamics allow us to calculate a lower bound on the basins of attraction analytically (Sec. 3.3). This leads to a non-probabilistic criterion for guaranteed recognition that includes finite-size effects, Eq. (3.20).

Note that the mirrored subnetwork structure should not be confused with "layers" from "traditional" layered neural networks. MONACO is very similar to a continuous version of the Hopfield model [24] (compare with Eq. (3.13)): Each oscillator pair

corresponds to an artificial neuron that "stores" its phase difference $\Delta\vartheta_i$. The synchronization process can be seen as continuous updating of the $\Delta\vartheta_i$. MONACO's subnetworks, however, change the properties of the "neurons", while the ideal effective dynamics Eq. (3.5) remain unchanged except for the coordinates they are represented in. This is distinct from more "traditional" layered neural networks, where the layer structure is essential to the dynamics.

On the contrary, a design with two subnetworks is not necessary in order to obtain the described dynamics including isolated attractors: A multiplicative coupling modulation suffices; consider e.g. the following single-network-system:

$$\dot{\vartheta}_i = \Omega_i + \cos\vartheta_i \cdot a_{ext}(t) \cdot \frac{\epsilon}{N} \sum_{j=1}^{N} \sin\vartheta_j$$

$$a_{ext}(t) = \sum_{k,l=1}^{N} S_{kl} \sin\Omega_k t \sin\Omega_l t$$

Here, the averaged dynamics would be the same as for MONACO, but in coordinates $\varphi_i(t) = \vartheta_i(t) - \Omega_i t$ (compare with Appendix A):

$$\dot{\varphi}_i = -\frac{\epsilon}{2N} \left(\sum_{j=1}^{N} S_{ij} \sin\varphi_i \cos\varphi_j - \frac{M}{2} \sin\varphi_i \cos\varphi_i \right)$$

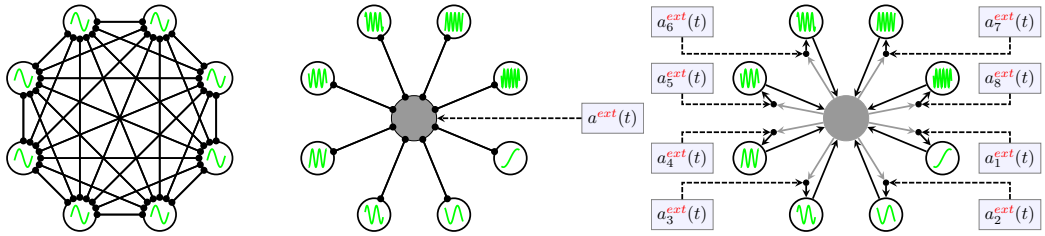
Phase shifts φ_i must be used, as no oscillators with equal frequencies exist in this setup and therefore, phase differences $\Delta\vartheta_i$ are no useful coordinate. As discussed below, tracking changes of the φ_i requires very precise frequency and time measurements, which renders readout difficult and error-prone. Therefore, this exemplary network is inferior to MONACO.

The MONACO-architecture will now be compared to other associative memories consisting of phase oscillators. Distinctive features are compared in Table 3.3, while schematics of are shown in Fig. 3.7. We discriminate between two types of networks: In *physically* all-to-all connected networks (architectures (I) [26–28] and (III) [32]), oscillators have the same frequencies and every oscillator is connected with every other(see Fig. 3.7 a). Therefore, the number of connections scales with $\mathcal{O}(N^2)$ in these networks, which limits the networks' size. As proposed in [29], oscillators of different frequency can be all-to-all connected *dynamically* with only one physical connection per oscillator if the oscillators' coupling is modulated in time. In architecture (IIA) [29, 30], the oscillators are globally coupled to a sum of the oscillators' signals with a single temporal modulation of the coupling (see Fig. 3.7 b). Due to the global coupling, the number of connections scales with $\mathcal{O}(N)$ connections only.

Architecture (IIB) [29, 31] follows a slightly more complicated scheme, where every oscillator receives the signals of all other oscillators, but each oscillator has its

network / property	all-to-all connected network (I) [26–28]	dynamically all-to-all connected (IIA) [29, 30]	dynamically all-to-all connected (IIB) [29, 31]	all-to-all connected with higher order Fourier modes (III) [32]	MONACO
Number of connections	$\mathcal{O}(N^2)$	$\mathcal{O}(N)$ + external coupling modulation $\geq \mathcal{O}(N \cdot M)$	$\mathcal{O}(N)$ + external coupling modulations $\geq \mathcal{O}(N \cdot M)$	$\geq \mathcal{O}(N^2)$; coupling function implementation unknown	$\boldsymbol{\mathcal{O}(N \cdot M)}$
frequency distribution	$\boldsymbol{\mathcal{O}(N^0)}$	$\mathcal{O}(N^2)$	$\mathcal{O}(N^{\ln 3/\ln 2})$	$\boldsymbol{\mathcal{O}(N^0)}$	$\mathcal{O}(N^2)$
initialization quality	ambiguous, fast	ambiguous [51], slow	ambiguous [51], slow	ambiguous, fast	**not ambiguous, fast**
effective recognition dynamics	$\dot{\vartheta}_i - \Omega = \epsilon \sum_{j=1}^{N} S_{ij} \sin(\vartheta_j - \vartheta_i)$	$\dot{\varphi}_i = \epsilon \sum_{j=1}^{N} S_{ij} \sin(\varphi_j - \varphi_i)$	$\dot{\varphi}_i = \epsilon \sum_{j=1}^{N} S_{ij} \sin(\varphi_j - \varphi_i)$	$\dot{\vartheta}_i - \Omega = 1/N\big(\sum_{j=1}^{N} S_{ij} \sin(\vartheta_j - \vartheta_i) + \bar{\epsilon} \sin 2(\vartheta_j - \vartheta_i)\big)$	$\Delta\dot{\vartheta}_i = -\epsilon/N\big(\sum_{j=1}^{N} S_{ij} \sin \Delta\vartheta_i \cos \Delta\vartheta_j - \frac{M}{2} \sin \Delta\vartheta_i \cos \Delta\vartheta_i\big)$
isolated attractors?	✗ [35]			✓	✓
loading rate	≤ 0.048 [52–54]			$\geq \frac{2\bar{\epsilon}^2}{\log N}$	$\geq \frac{1}{2 \log N}$
error-free capacity	$\frac{2}{N}$ [26, 32]			$\boldsymbol{\frac{2\bar{\epsilon}^2}{\log N}}$	$\boldsymbol{\frac{1}{2 \log N}}$

Table 3.3: Comparison of MONACO with other autoassociative memory architectures based on phase-oscillator networks. The best performances for every property are marked in bold.



(a) all-to-all connected networks (I) and (III)

(b) dynamically all-to-all connected (IIA)

(c) dynamically all-to-all connected (IIB)

Figure 3.7: Schematics of previous oscillatory neural network architectures that act as autoassociative memories

own coupling modulation (see Fig. 3.7 c). Nevertheless, the scaling of the number of connections is still $\mathcal{O}(N)$.

MONACO is a dynamically all-to-all connected network as well. Each subnetwork is globally coupled similar to (IIA), albeit with a different coupling modulation (see Fig. 3.1). The use of two mirrored subnetworks allows for the *internal* generation of the global coupling modulations. In contrast, the hardware implementation for architectures (IIA) and (IIB) introduced in [30, 31] was fed by computer-generated coupling modulations. In MONACO, the scaling of the number of connections is $\mathcal{O}(N \cdot M)$ (see Eq. (3.4) for the coupling modulations and consider that global coupling scales with $\mathcal{O}(N)$). This scaling is optimal if the generation of the coupling modulations is considered, as $N \cdot M$ pixels have to be incorporated.

However, the reduction in the number of spatial connections is not for free: The original complexity in space is transferred to a complexity in time with the number of frequencies contained in the coupling modulation growing like $\mathcal{O}(N^2)$ for architecture (IIA) and MONACO [29]. Frequency conditions for architecture (IIB) are less restrictive and the number of frequencies scales with $\mathcal{O}(N^{\ln 3/\ln 2})$ [31].

Now, coordinates of the network dynamics will be discussed, as they determine how initial conditions can be enforced as well as how the system state can be read out. In (I) and (III), the desired dynamics occur in oscillators' phases $\vartheta_i = \Omega t + \varphi_i$, so pixels of the same value have the same phase. An encoded pattern is then represented by two groups of oscillators whose phases differ by π. Note that this representation itself is ambiguous, as it is physically impossible to decide if a group follows or precedes the other one. In other words, the physical state represents a pattern as well as its inverse. In (IIA) and (IIB), equal pixels are represented by equal phase shifts φ_i and different pixels differ by a phase shift difference of π. Note that phase shifts are only unique up to a constant $\varphi_i^0 = \varphi_i(t = 0)$. As a consequence, only differences $\varphi_i(t) - \varphi_i(t')$ can be determined. In MONACO patterns are coded

into phase differences $\Delta\vartheta_i = \vartheta_i^{[1]} - \vartheta_i^{[2]}$ of oscillators of equal frequencies. Each pixel is mapped onto a phase difference with $\alpha_i = \cos\Delta\vartheta_i$, so $\alpha_i = +1$ corresponds to a synchronized oscillator pair and $\alpha_i = -1$ to a antisynchronized one. Therefore, MONACO's system state represents a pattern without ambiguity.

The different nature of the variables entail that also the setting of initial conditions differs radically between architectures: Phase differences in MONACO are easily manipulable: Oscillator pairs corresponding to $+1$ are directly coupled positively, while pixels with -1 receive a negative coupling, resulting in synchronized pairs with $\Delta\vartheta_i = 0$ or $\Delta\vartheta_i = \pi$. Phases ϑ_i change quickly in time, so they are difficult to control directly. However, initial conditions in (I) and (III) can be set similar to MONACO by coupling all N oscillators in a row, where oscillators representing equal pixels are coupled positively and unequal pixels interact via a negative coupling. In (IIA) and (IIB), two main problems must be overcome to set initial conditions: First, phase shifts cannot be manipulated directly and second, phase shifts are undefined without a temporal reference. Hoppensteadt and Izhikevich [29] proposed to use the same coupling circuitry as used for the recognition, but with a different coupling matrix $\mathbf{S}$: $S_{ij} = \alpha_i^d \alpha_j^d$ is used to initialize a defective pattern $\boldsymbol{\alpha}^d$. Then, recognition is performed with the usual coupling matrix $S_{ij} = \sum_{m=1}^{M} \alpha_i^m \alpha_j^m$. By evaluating phase shift changes between the introduced initial condition and the recognition, pixel changes can be retrieved without the constants φ_i^0. However, initialized patterns are ambiguous: As $S_{ij}(-\boldsymbol{\alpha}^d) = (-1)^2 \alpha_i^d \alpha_j^d = S_{ij}(\boldsymbol{\alpha}^d)$, the inverse pattern $-\boldsymbol{\alpha}^d$ is initialized half of the time. Additionally, as this method is limited by the averaging condition similar to the recognition, this method is of timescale $1/\epsilon$ and therefore considerably slower than the direct coupling used for (I), (III) and MONACO.

Similarly, readout of the final pattern is easy in MONACO: As mentioned above, $\cos\Delta\vartheta_i = \alpha_i$ can be read out directly from the corresponding oscillator pair. Readout in (I) and (III) is analogue, but phase differences between different pixels are determined, which again describes both a specific pattern and its inverse. For (IIA) and (IIB), phase shifts have to be determined by comparing the phase of an oscillator with an external reference[2]. Then, the difference of phase shifts between final state and the initial conditions needs to be evaluated[3]. In refs. [30, 31], this was done with a computer and analog-digital converter cards.

Ease of readout additionally depends on the effective dynamics of the architectures: Traditional Kuramoto-type networks (I) employ a coupling that depends only on the mutual phase differences of all oscillators ($\propto \sin(\vartheta_i - \vartheta_j)$). While (IIA) and (IIB) have a seemingly more complex structure due to their coupling modulations, dynamics are effectively the same as in (I) after averaging (Compare with Table

[2]This should in principle be possible with precise reference oscillators, a precise clock and a memory for the initialized phase shift values.

[3]As Readout in (IIA) and (IIB) only measures changes between initial values and thus the final pattern can be constructed non-ambiguously from the initial defective pattern, the ambiguity introduced in the setting of initial values is effectively removed.

3.3), albeit in different coordinates. In these dynamics, the individual patterns are not individual attractors, but part of one large attractor. More precisely, patterns are connected by lines of attractive non-isolated fixed points [35]. Consequently, recognition is only possible for short times, as the system state drifts on the attractor due to implementation inaccuracies and higher order terms and readout must occur immediately after the recognition is successful. Additionally, the system state does only settle close to the correct memorized pattern, so the output values are not inherently digital as the patterns are.

MONACO's dynamics (*cf.* Eq. (3.1)) take on a simple mathematical form after averaging (Eq. (3.5) for the formulation in phase differences, Eq. (3.13) in pattern space $\boldsymbol{\alpha}$). In these novel dynamics, binary memorized patterns are *individual* attractors. In [32], yet another dynamics was introduced with architecture (III) (see Table 3.3). Memorized patterns are isolated attractors here as well due to higher order Fourier modes in the coupling function. Due to the isolated attractors, readout does not need to be exactly timed and the output is inherently digital in MONACO as well as in architecture (III). Additionally, the dynamics of (III) enable the exclusion of spurious attractors for specific parameter ranges, while MONACO's dynamics allowed us to determine lower bounds on the basins of attraction, as discussed below and in Section 3.3.

Concerning quantitative measures for associative networks, often the capacity or loading rate of a network is used. It describes the maximum possible ratio of M and N, where the system state still settles close to the correct memorized pattern. Usually, it is computed for a set of random memorized patterns in the limes $N \to \infty$. This definition, however, includes deviations from the memorized patterns, so e.g. some bits may be erroneous at retrieval. Nishikawa et al. point out the importance of error-free retrieval for engineering applications [32] and remind of the error-free capacity (def. in Sec. 3.2.3) as a more meaningful quantity, as it is used for traditional neural networks [55]. The error-free capacity of MONACO (Eq. (3.12)) is on a par with architecture (III) [32] and equal to the error-free capacity of the Hopfield model [55] while memorized patterns are typically unstable in architectures (I), (IIA) and (IIB) with an error-free capacity of $2/N$ [26, 32]. The loading rate for architectures (I),(IIA) and (IIB) has been derived as 0.048 [52–54], while it has not been calculated for neither architecture (III) nor MONACO yet. However, the error-free capacities are lower bounds on the loading rates and may be larger than the value for (I) - (IIB) similar to the error-free capacities.

While the loading rate and the error-free capacity are useful for comparing architectures, their probabilistic nature and the derivation for $\lim N \to \infty$ impair their significance for real networks: Specific sets of memorized patterns are possibly not random and finite size effects might improve or impair pattern stability as well as recognition success. Non-probabilistic criteria valid for all network sizes allow to exactly evaluate performance of a network for a specific use case and enable the development of more complex algorithms using the recognition process repeatedly. We

derived such criteria for MONACO: Eq. (3.10) guarantees stability of all memorized patterns if scalar products between memorized patterns are not too large. Eq. (3.20) guarantees recognition success by giving a lower bound on the number of allowed erroneous pixels n^f. If a network stores a large number of patterns M, the minimal size of the basins of attraction will be quite small and few erroneous pixels n^f can be guaranteed to be corrected. In many applications, however, the number of patterns M is much smaller than N and the ability to correct larger errors is desired.

The last aspects to be discussed concern recognition time and oscillator accuracy. In physically all-to-all connected networks (I) and (III), oscillator frequencies are not restrictive, as long as they are similar enough to be well above the Kuramoto transition. In contrast, frequency conditions in dynamically all-to-all connected networks limit the network size: Since in practice there will be only a certain frequency interval available, the number of oscillators is limited by the accuracy of the frequencies [30]. Recognition times have not been calculated analytically for any of the oscillatory neural networks presented here. However, we assume that the frequency restrictions present in (IIA), (IIB) and MONACO lead to slower recognition times compared to (I) or (III). Nevertheless, the shift of frequency due to the acceleration effect [50] present in the real dynamics of (IIA) and (IIB) [29, 30, 35] does not interfere with the recognition process in MONACO since the change in frequency is identical in each oscillator pair due to its mirrored structure. Additionally, it is possible to introduce several coupling modulations per subnetwork similar to the transition from architecture (IIA) to (IIB) for the MONACO-architecture (Sec. 3.7). In this improved network the scaling of necessary frequencies is reduced to $\mathcal{O}(N^{\ln(3)/\ln(2)})$.

3.6 Implementation considerations

Before actually designing the network's circuitry, we outline difficulties and solutions for the implementation. First, frequency deviations are discussed:

3.6.1 Influence of frequency deviations

The MONACO architecture and our previous analysis is based on two major assumptions about the network's frequencies:

1. Angular frequencies $\Omega_i^{[1]} = \Omega_i^{[2]} = \Omega_i$ of oscillators in a pair are equal.

2. Angular frequencies Ω_i of different pairs fulfill the frequency conditions derived in Appendix A:
   - $\Omega_i - \Omega_j \neq \Omega_k - \Omega_l$ for pairwise different i,j,k,l or more precisely, frequency differences must be sufficiently different: $|(\Omega_i - \Omega_j) - (\Omega_k - \Omega_l)| \gg 0 \ \forall i, j, k, l$.

- $\Omega_i > \Omega_{max}/3 \; \forall i$

  An optimal distribution can be obtained by multiplying a minimal frequency difference $\Delta\Omega_{min}$ with a Golomb-ruler [48], a set of integers with non-equal differences [30], and adding $\Omega_{max}/3$.

In any real system, however, drifts or inaccuracies of components will lead to deviations of the frequencies.

Violating the frequency conditions would create unwanted resonant terms that would modify the coupling matrix S in the main evolution equation Eq. 3.5. In [30], Hölzel et al. quantified permitted frequency deviations: If the oscillators in a pair do not differ in frequency, deviations $\Delta\omega_i = \Omega_i - \Omega_{i,ideal}$ of their frequency from the optimal distribution do not create additional perturbation terms if $\Delta\omega_i \ll \Delta\Omega_{min} \propto 1/N^2$. If a network is sufficiently small, this condition can be easily met.

If frequency conditions are fulfilled, differences $\Delta\Omega_i = \Omega_i^{[1]} - \Omega_i^{[2]}$ of angular frequencies between oscillators in a pair still modify Eq. 3.5:

$$\Delta\dot{\vartheta}_i = \Delta\Omega_i - \frac{\epsilon}{N}\sin\Delta\vartheta_i\Big(\sum_{j=1}^{N} S_{ij}\cos\Delta\vartheta_j - \frac{M}{2}\cos\Delta\vartheta_i\Big) \tag{3.21}$$

This frequency mismatch between the oscillator pairs is much more critical: While small mismatches only shift $\Delta\vartheta_i^*$ slightly, larger mismatches destabilize the attractors till finally phase slips occur in the i^{th} oscillator pair. With the assumption that we start in the correct basin of attraction and worst-case approximations of Eq. (3.21), we gain $\Delta\Omega_i \ll \epsilon/(2N)$. (Note that Stefan Litzel analyzed synchronization of oscillator pairs with frequency differences according to a Gaussian distribution very thoroughly in his bachelor's thesis [56]. However, he explored an architecture variant with slightly different dynamics.)

As ϵ has to be sufficiently small to obtain Eq. (3.21) in the first place, the latter condition seems to be difficult to fulfill with precise manufacturing alone. Therefore, we chose to add an additional frequency annealing step before setting the initial conditions, which is illustrated in the next section.

3.6.2 Frequency annealing mechanism

As shown in the last section, the network is sensitive to frequency differences $\Delta\Omega_i$ between oscillator pairs. Manual adjustment of the frequencies with a potentiometer (or several stacked potentiometers) is tedious. More importantly, thermal drifts shift the oscillator frequencies, thus making recurrent adjustment necessary. Consequently, an automatic adjustment method is required.

We propose a frequency annealing mechanism based on nonlinear dynamics: Both oscillators of the i^{th} pair can be described by the dynamics of their phases $\vartheta_i^{[1]}$ and

$\vartheta_i^{[2]}$. If the phase difference between the oscillators does not change ($\mathrm{d}/\mathrm{d}t(\vartheta_i^{[1]} - \vartheta_i^{[2]}) = \Delta\dot{\vartheta}_i = 0$), the frequencies must be equal. If we can thus construct a system with stable fixed points in $\Delta\vartheta_i$, the phase difference must move to the fixed point and the frequencies anneal.
In detail, we implement the following scheme: One oscillator of a pair is modified, so that its frequency depends on a continuous, 2π-periodic function F of the pair's phase difference. (Note that e.g. any continuous function of our readout signal $\propto \cos(\Delta\vartheta_i)$ fulfills that requirement.) The other oscillator acts as a reference only, so its dynamics remain unchanged:

$$\begin{aligned} \dot{\vartheta}_i^{[1]} &= \Omega_i^{[1]} - F(\Delta\vartheta_i) & \dot{\vartheta}_i^{[2]} &= \Omega_i^{[2]} \\ \Delta\dot{\vartheta}_i &= \Delta\Omega_i - F(\Delta\vartheta_i) \end{aligned} \tag{3.22}$$

Eq. (3.22) is an ordinary differential equation of first order. No complicated behavior like periodic orbits or chaos can exist in such a system, only fixed points [36]. Frequencies of both oscillators are equal at any fixed points $F(\Delta\vartheta_i) = \Delta\Omega_i$, as $\Delta\dot{\vartheta}_i = 0$. Stability of these fixed points is decided by the sign of the derivative $\mathrm{d}\Delta\dot{\vartheta}_i/\mathrm{d}\vartheta_i = -\mathrm{d}F/\mathrm{d}\vartheta_i$ at the fixed point.

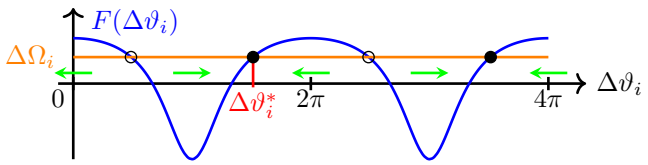


Figure 3.8: Frequency annealing mechanism: Any phase differences $\Delta\vartheta_i^*$ with $F(\Delta\vartheta_i^*) = \Delta\Omega_i$ are fixed points of Eq. (3.22). At any fixed point, the frequency difference $\Delta\dot{\vartheta}_i$ is zero and thus the frequencies are equal. The slope $\frac{\partial\Delta\dot{\vartheta}_i}{\partial\Delta\vartheta_i} = -\frac{\partial F}{\partial\Delta\vartheta_i} < 0$ indicates stable fixed points. If an unstable fixed point can be found, there must be another stable fixed point where F returns, as F is periodic and continuous. Consequently, if F intersects with $\Delta\Omega_i$ anywhere, a stable fixed point exists.

Fig. 3.8 illustrates that for every unstable fixed point another stable fixed point must exist: As F is bounded, continuous and periodic, it has to cross $\Delta\Omega$ again with positive slope before the next period. Note that these fixed point pairs are created in saddle-node bifurcations (Sec. 2.1.4). Summing up:

- If $F(\Delta\vartheta_i^*) = \Delta\Omega_i$ for any phase difference $\Delta\vartheta_i^*$, a stable fixed point exists. Then, both frequencies anneal and the phase difference is locked at the fixed point.
- If no point $F(\Delta\vartheta_i^*) = \Delta\Omega_i$ exists, there is no fixed point and frequency annealing cannot work. Therefore, $[\min(F), \max(F)]$ defines the possible annealing range. From another point of view, the bifurcation parameter $-\Delta\Omega_i$ must

move into the range of F to create fixed point pairs in saddle-node bifurcations. In order to facilitate annealing, either the range of F can be increased or the difference of natural frequencies can be diminished by other means, e.g. manual adjustment.

- This frequency adjustment leaves the phase difference $\Delta\vartheta_i$ locked at the fixed point. However, $\Delta\vartheta_i$ needs to move unhindered during initialization and recognition. Therefore, we replaced the feedback loop after annealing with a constant $C = F(\Delta\vartheta_i^*)$ before initialization.

Note that similar methods are used in electronic engineering under the name "phase locked loop" (See e.g. [37] chapter 9.27.), but with a different point of view, different applications and without cutting the feedback loop.

3.6.3 Phase difference readout mechanism

One issue of phase oscillator networks is the readout of the network state: If phases and thus the oscillator signals are directly measured, the readout needs to be significantly faster than the oscillator frequency, which in turn severely limits the oscillator frequencies. In contrast, phase differences $\Delta\vartheta_i$ change on the timescale of $\mathcal{O}(\epsilon^{-1}) \gg \mathcal{O}(\Omega^{-1})$ and knowledge of single phases is not required in our architecture.

Voltages which depend on $\Delta\vartheta_i$ only can be constructed by multiplying two oscillators' signals and filtering them with a low pass:

$$\begin{aligned} A\sin\left(\vartheta_i^{[1]}\right) \cdot A\sin\left(\vartheta_i^{[2]}\right) &= A^2 \cdot \frac{1}{2}\left[\cos\left(\vartheta_i^{[1]} - \vartheta_i^{[2]}\right) + \cos\left(\vartheta_i^{[1]} + \vartheta_i^{[2]}\right)\right] \\ &\approx A^2 \cdot \frac{1}{2}\left[\cos\left(\Delta\vartheta_i\right) + \cos\left(2\Omega_i t\right)\right] \end{aligned}$$

After an ideal low pass, only $\frac{A^2}{2}\cos\left(\Delta\vartheta_i\right)$ remains. As MONACO's dynamics can be expressed with $\alpha_i = \cos\left(\Delta\vartheta_i\right)$ only (Eq. (3.13)), we gain complete information on the current system state. With this direct method, readout circuitry needs to sample faster than $\mathcal{O}(\epsilon)$ only.

Note that the sign of the phase difference is lost due to the cosine, which might be relevant for other architectures. In that case, one of the oscillator signals can be integrated or differentiated prior to multiplication. Another possible readout method is the addition of oscillator signals, which is followed by an rectifier and a low pass, as proposed in [57]. This method is not used in this thesis, but might be interesting if the amount of analog multiplications should be reduced. Frequency range and accuracy pose another implementation challenge. In the next section, we will propose a variant of MONACO with more physical connections but less restrictive frequency conditions.

3.7 A variant with different coupling modulations: MONACO II

As mentioned in 3.6.1, any two frequency differences in the original MONACO architecture must not be the same to avoid unwanted coupling terms. The necessary frequency range then scales quadratically with the number of pixels N.

In [31], Kostorz et al. have demonstrated the use of separate coupling modulations for each oscillator. As the modulations contain less frequencies each, less unwanted coupling terms need to be negated with frequency conditions. We now combine this idea with our mirrored structure and construct $2N$ coupling modulations from oscillator signals:

$$\begin{aligned} \dot{\vartheta}_i^{[1/2]} &= \Omega_i^{[1/2]} + \cos\vartheta_i^{[1/2]} \cdot a_i^{[2/1]}(t) \cdot \frac{\epsilon}{N} \sum_{j=1}^{N} \sin\vartheta_j^{[1/2]} \\ a_i^{[1/2]}(t) &= \sum_{m=1}^{M} \alpha_i^m \sin\vartheta_i^{[1/2]} \sum_{k=1}^{N} \alpha_k^m \sin\vartheta_k^{[1/2]} \end{aligned} \tag{3.23}$$

Here, the i^{th} oscillator of the second group receives all signals $\sum_j \sin\vartheta_j^{[2]}$ in the second group modulated by its own private coupling modulation $a_i^{[1]}$, which is constructed from signals of the first group. For the construction of the $a_i^{[1]}$ and $a_i^{[2]}$, the sums $\sum_{k=1}^{N} \alpha_k^m \sin\vartheta_k^{[1/2]}$ can be reused between pixels.

The implementation of this variant, called MONACO II, was topic of the master's Thesis of Nicholas Pfifferling [58]. Averaging of Eq. (3.23) is analog to the original MONACO architecture (MONACO I), with frequency conditions being the same as Kostorz et al. derived in their work [31]:

$$\Omega_i \neq \Omega_j \qquad \forall i, j \tag{3.24}$$

$$\Omega_i \neq \frac{1}{2}\left(\Omega_j - \Omega_k\right) \qquad i, j, k \tag{3.25}$$

$$\Omega_i \neq \frac{1}{2}\left(\Omega_j + \Omega_k\right) \qquad i, j, k \tag{3.26}$$

Similar to MONACO I, frequencies in a group have to be pairwise different. The second condition can be fulfilled with $\Omega_{min} > \frac{1}{3}\Omega_{max}$, while the last criterion defines the scaling properties. As derived in [31], the criterion is fulfilled by all integers which do not contain any 2 in base-3 notation [59], which is a form of Cantor set. Angular frequencies Ω_i can be derived by multiplying the sequence with some minimal frequency difference $\Delta\Omega_{min}$. The frequency space required grows with $N^{\frac{\ln 3}{\ln 2}} \approx N^{1.58}$, which is a significant improvement over the factor N^2 in MONACO I. Finally, we

gain the averaged dynamic equation for MONACO II [58]:

$$\Delta\dot{\vartheta}_i = -\frac{\epsilon}{2N}\sin\Delta\vartheta_i\Big(\sum_{j=1}^{N} S_{ij}\cos\Delta\vartheta_j\Big) \tag{3.27}$$

Comparison of Eq. (3.27) with the averaged dynamics of MONACO I (Eq. 3.5) reveals two major differences:

- The coupling strength prefactor is halved for MONACO II. We expect longer recognition times at the same ϵ, but no statement about an upper limit of ϵ and thus minimal recognition times can be made.
- The term $\propto \frac{M}{2}\sin\Delta\vartheta_i\cos\Delta\vartheta_i$ is missing in MONACO II. In MONACO I, it destabilizes both $\Delta\vartheta_i = 0$ and $\Delta\vartheta_i = \pi$. As it is small compared to the other terms, the difference in dynamics should be small.

Derivations for stability or basins of attraction of MONACO II are not shown, as they are completely analog to MONACO I with the only difference being the missing term.

Focus of this thesis will remain on MONACO I, as only preliminary results exist for MONACO II yet. In the next chapter, implementation of electronic circuitry will be elaborated in detail for both architectures.

Chapter 4

Network Implementation

In this chapter, our implementation of MONACO will be discussed. The network was designed for the following requirements:

- The network is supposed to be a *proof-of-principle* for the proposed architecture. Therefore, it must mirror the full phase dynamics of MONACO (Eq. (3.1)) as closely as possible.
- The *influence of spurious attractors* on the recognition process should be examined.
- The *criterion for guaranteed recognition* Eq. (3.20) should be verified.
- Measurements should be conducted with *orthogonal* as well as *random* memorized patterns.
- The influence of different coupling strengths ϵ on both recognition times and recognition success should be examined.
- *External noise* should be added to oscillator pairs to study resistance to perturbations and deviations of the recognition times.
- In addition, we want to examine a variant of MONACO that has individual coupling modulations for each pixel, but less restricting frequency conditions. (MONACO II; see Sec. 3.7.) The *circuitry* for MONACO I should be *reusable* for MONACO II as much as possible.

All requirements can be fulfilled with a network of $N = 8$ oscillator pairs and circuitry for $M \leq 3$ patterns.

In contrast to previous approaches that include a time-dependent global coupling [30, 31], coupling modulations are generated inside the network. Consequently, more circuitry is required. In order to ease debugging, allow for changes of subcircuits and enable sharing of circuitry between MONACO I and II, the circuitry is split into modules:

- **8 oscillator modules** contain an oscillator pair of a specific angular frequency Ω_i as well as circuitry for the readout of $\alpha_i = \cos(\Delta\vartheta_i)$ and for frequency annealing of this pair.
- **8 pattern modules** are used to input both the defective pattern as well as the memorized, correct pattern candidates.
- **1 mainboard** which connects modules and distributes power.
- **1 coupling voltage module** for MONACO I. It generates the product of coupling modulations with the sum of the oscillator signals. This coupling voltage is coupled to oscillators during recognition.
- **1 summation module** for MONACO II, which creates sums of oscillator signals necessary for coupling voltage creation.
- **8 multiplication modules** for MONACO II that create coupling voltages for individual pixels.

The interaction between modules in MONACO I during the recognition step is summarized in Fig. 4.1: Amplified oscillator signals are transferred to the pattern modules, where they are multiplied with pattern pixels of memorized patterns. Those are combined in the coupling voltage module and coupling voltages connect to coupling resistors on the oscillator modules. The signal flow during the initialization step is similar: Oscillator signals are multiplied with pixels of the defective pattern on the pattern modules, then directly coupled to the other oscillator in a pair.

The network implementation is built to interface with external hardware for automation of measurements: Patterns for both initialization and recognition are supplied as logic signals. Similarly, steps of the recognition process are switched with logic signals as well. As the frequency annealing mechanism proposed in Sec. 3.6.2 needs additional steps prior to initialization, there are 4 steps in total:

- *Annealing step*: The feedback loop according to Eq. (3.22) is closed and the fixed point of equal frequencies is quickly reached. Coupling voltages for both initialization and recognition are not connected.
- *Hold step:* The feedback is replaced with the voltage at the fixed point. Frequencies are still annealed, but uncoupled. Remaining frequency differences can be measured and annealing failure can be identified. Coupling voltages are still not connected.
- *Initialization step:* Initialization coupling voltage is connected. Oscillator pairs synchronize at phase differences according to the pixels of the defective pattern.
- *Recognition step:* Initialization coupling is disconnected and recognition coupling is simultaneously connected.

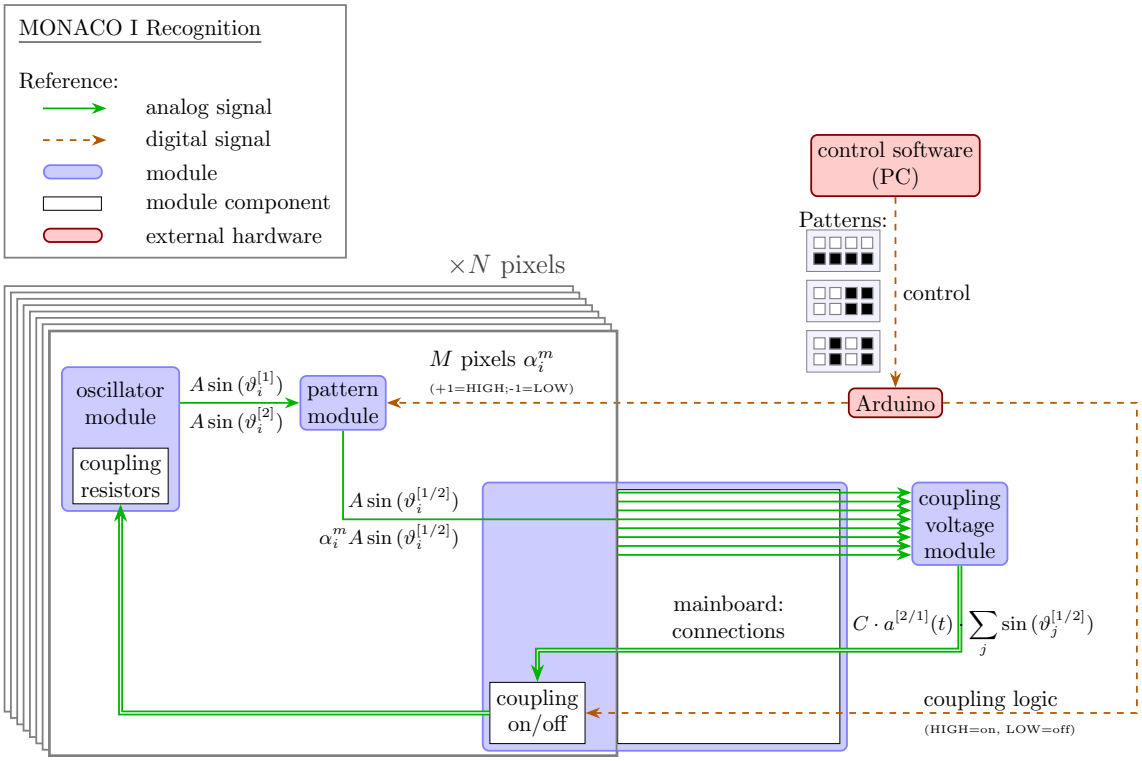


Figure 4.1: Diagram summarizing the signal flow between modules (blue) and external hardware (red) in MONACO I during the recognition process: Oscillators on the oscillator module are coupled to a coupling voltage provided by the coupling voltage module, which in turn receives oscillator voltages modulated by pattern pixels from the pattern module. External hardware provides the memorized patterns α_i^m as logic signals and enables/disables the recognition step by passing/blocking the coupling voltage.

Stripboard layouts for all modules can be found in Appendix C. A short introduction to electronic components and used subcircuits was given in Sec. 2.2. In the following sections, each module will be described in detail, starting with the oscillator modules:

4.1 Oscillator modules

Previous electronic implementations of Van der Pol oscillators in our group (Sec. 2.2.3) showed sinusoidal oscillations and sinusoidal phase response curves shifted by $\frac{\pi}{2}$ with respect to the oscillation [31]. Consequently, these oscillators qualify for a correct implementation of the phase dynamics Eq. (3.1). An external voltage signal can be coupled to such an oscillator by connecting it to the oscillator's internal

voltage via a *coupling resistor*. The resulting current then changes the internal voltage, which in turn shifts the phase of the oscillator.

However, the mirrored structure and the internal coupling modulation generation have additional requirements, which require modifications of the original circuit:

- The coupling strength ϵ needs to be varied between series of measurements. As voltages in the circuit cannot exceed the supply voltages and small voltage signals are sensitive to noise, we decided to vary coupling resistors instead. The influence of this resistance variation on the oscillator dynamics has to be compensated, as discussed in the next section.
- Circuitry for the readout of $\alpha_i = \cos(\Delta\vartheta_i)$ as proposed in Sec. 3.6.3 has to be included.
- Frequency differences between oscillator pairs have to be minimized. Manual adjustment is not sufficient and we implement an automatic frequency annealing scheme as conceived in Sec. 3.6.2, which requires a frequency dependency on some control voltage.

4.1.1 Coupling oscillators to external voltages via resistors

Let us consider connecting the idealized oscillator circuit 2.7a to an external coupling voltage U_{ext} with a resistor R_{ext} (Fig. 4.2):

$$\begin{aligned}
\dot{U} &= \frac{1}{C}\dot{Q} \\
&= \frac{1}{C}\left(-I_L - I_{diodes} + I_{R_{neg}} - I_{ext}\right) \\
&= -\frac{1}{C}\left(I_L + I_{diodes} - \frac{U}{R_{neg}} + \frac{U - U_{ext}}{R_{ext}}\right) \\
&= -\frac{I_L + I_{diodes}}{C} + \frac{U}{C}\left(\frac{1}{R_{neg}} - \frac{1}{R_{ext}}\right) + \frac{U_{ext}}{CR_{ext}}
\end{aligned} \tag{4.1}$$

Here, I_L, I_{diodes} and I_{ext} are the currents through the inductance, the diodes and the coupling resistor R_{ext}, while I_{neg} is the current from the negative resistance. Coupling to an external voltage U_{ext} should not directly depend on the internal voltage U. Additionally, the oscillator dynamics and thus the signal shape are changed even if U_{ext} is grounded. This strongly implies to split the influence of R_{ext} into a coupling part $\frac{U_{ext}}{CR_{ext}}$ and an integral part of the oscillator circuit: The integral part $-\frac{U}{CR_{ext}}$ acts as if the resistor was grounded, effectively decreasing the negative resistance in the circuit:

$$\frac{1}{R'_{neg}} = \frac{1}{R_{neg}} - \frac{1}{R_{ext}}$$

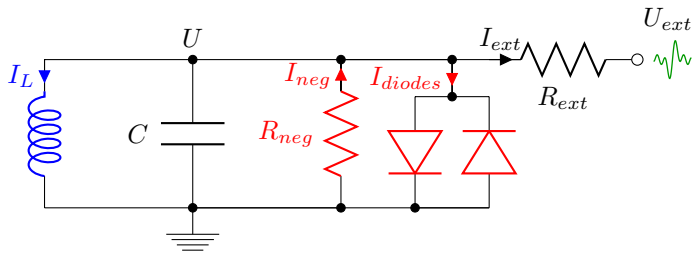


Figure 4.2: Oscillator coupled to external voltage U_{ext}. The oscillation is on average diminished by currents I_L and I_{diodes}, while the current I_{neg} from the negative resistance always increase the absolute value of U. The influence of R_{ext} can be divided into a coupling contribution proportional to U_{ext} and a reduction of the effective negative resistance. (See Eq. (4.1).)

In [31], Kostorz et al. had sinusoidal oscillations with a coupling resistor $R_{coupl} = 100\,\mathrm{k\Omega}$ and an additional resistor $R_s = 8{,}25\,\mathrm{k\Omega}$ to ground for an inverting summer, resulting in $R'_{neg,Kostorz} = -33\,\mathrm{k\Omega}$.
For MONACO, we need separate coupling resistors for both initialization and recognition, while any circuitry depending on the oscillator's signal is separated from the oscillator circuit with a non-inverting amplifier. Recognition coupling resistors R_{recog} are varied to study the effect of coupling strength variation. Nevertheless, R'_{neg} should be held constant for constant signal shape and comparison with [31]. Therefore, we introduce additional resistors R_{comp} to ground as compensation to fix R'_{neg} to $-33\,\mathrm{k\Omega}$:

$$\frac{1}{R'_{neg}} = -\frac{1}{33\,\mathrm{k\Omega}} = \frac{1}{R_{neg}} - \frac{1}{R_{recog}} - \frac{1}{R_{init}} - \frac{1}{R_{comp}} \tag{4.2}$$

A summary of sets R_{recog}, R_{init} and R_{comp} used is shown in Table 4.1.

4.1.2 Phase difference readout implementation

As mentioned in Sec. 3.6.3, we need to multiply signals of an oscillator pair and remove the double frequency with a low pass to get a voltage proportional to $\alpha_i = \cos(\Delta\vartheta_i)$.
We used the multiplier IC AD633 for signal multiplication. (See Sec. 2.2.1 for a description.) Note that we arranged inputs such that one input voltage is inverted and that we do not use the summation input on this module by grounding it. Summing

R_{recog} [kΩ]	R_{init} [MΩ]	calculated R_{comp} [kΩ]	available R_{comp} [kΩ]	R'_{neg} [kΩ]
10,2	10,0	30,2	30,0	-33,3
30,0	10,0	10,23	10,2	-33,3
62,0	10,0	8,70	8,66	-33,5
100	10,0	8,26	8,25	-33,1
243	10,0	7,87	7,87	-33,0
511	10,0	7,74	7,68	-34,2
1000	10,0	7,69	7,68	-33,1
243	0,10	8,54	8,45	-34,4

Table 4.1: Values of R_{recog}, R_{init} and R_{comp} used for the experiment as well as effective negative resistances R'_{neg} are shown. $R_{neg} = 6{,}19\,\mathrm{k\Omega}$ was constant for all experiments. R_{comp} was calculated according to Eq. (4.2), then the next available resistance value was chosen.

up, the voltage after the multiplier is

$$\begin{aligned}\frac{A\sin\vartheta_i^{[1]}\cdot\left(-A\sin\vartheta_i^{[2]}\right)}{10\,\mathrm{V}} &= \frac{-A^2}{20\,\mathrm{V}}\left[\cos\Delta\vartheta_i - \cos\left(\vartheta_i^{[1]}+\vartheta_i^{[2]}\right)\right]\\ &\approx \frac{-A^2}{20\,\mathrm{V}}\left[\cos\Delta\vartheta_i + \cos\left(2\Omega_i t\right)\right]\,.\end{aligned}$$

The selection of the low-pass filter turns out to be non-trivial: If we would only care about the readout of $\cos\Delta\vartheta_i$, the filter selection would not be critical, as phase differences change on the timescale of $\mathcal{O}(\epsilon^{-1})$ during recognition. However, we want to use a function of $\cos\Delta\vartheta_i$ as feedback voltage in the frequency annealing circuitry (See Sec. 4.1.3). As every causal filter induces delay and might show passband ripple, F is a function of both $\Delta\vartheta_i$, $\Delta\dot{\vartheta}_i$ and a phase shift τ, which violates our assumption $F = F(\Delta\vartheta_i)$. Additionally, the feedback mechanism moves to the fixed point quite fast, so the gain(F) is not sufficient to describe the loop dynamics: Responses in the time domain (e.g. response to pulses or step-functions) influence the dynamics substantially. In order to derive an ideal filter type by theory, the filter's transfer function would have to be included into Eq. (3.22) yielding an integro-differential equation that is difficult to solve and optimize.

In practice, frequency annealing performance of different filters was tested in SPICE by Nicolas Pfifferling during his master's thesis: A simple RC-circuit, which is a filter of first order, did not remove the double frequency sufficiently. Frequency annealing with a 2$^{\mathrm{nd}}$ order Chebyshev filter showed damped oscillations towards the fixed point, indicating a stable focus and thus complex eigenvalues. While the frequency annealing worked, complex eigenvalues can only exist in systems with 2 or more dimensions. As our assumption is clearly violated and we thus cannot

exclude other attractors, we rejected the Chebyshev out of caution. We finally chose a 2nd order Butterworth filter with a voltage gain $K = 1.586$ and a cutoff frequency $f_{3dB} = 4{,}16\,\mathrm{Hz}$, which showed a non-oscillating trajectory towards the fixed point. Fig. 2.6 shows the circuit used for the implementation of the Butterworth filter. We gain a voltage

$$U_\alpha = \frac{-A^2 K}{20\,\mathrm{V}} \cos \Delta\vartheta_i \,. \tag{4.3}$$

Finally, the output voltage must be readable by the external hardware, whose analog input is limited to $[0\,\mathrm{V}, 5\,\mathrm{V}]$. Therefore, we shift and scale the voltage with an inverting summer (Sec. 2.2.2) with $R_{out} = 10\,\mathrm{k\Omega}$, $U_1 = U_\alpha$, $R_1 = 820\,\Omega$, $U_2 = -15\,\mathrm{V}$ and $R_2 = 59\,\mathrm{k\Omega}$. With Eq. (2.2) and $K = 1.586$, we gain the final readout voltage U'_α:

$$\begin{aligned} U'_\alpha &= -\left(\frac{R_{out}}{R_2}(-15\,\mathrm{V}) + \frac{R_{out}}{R_1} U_\alpha \right) \\ &= \frac{R_{out}}{R_2}(15\,\mathrm{V}) + \frac{A^2 K}{20\,\mathrm{V}} \frac{R_{out}}{R_1} \cos\left(\Delta\vartheta_i\right) \\ &= 2{,}54\,\mathrm{V} + \frac{A^2}{1{,}034\,\mathrm{V}} \cos\left(\Delta\vartheta_i\right) \end{aligned}$$

As the value of A is not the same for all measurements, the prefactor in front of $\alpha_i = \cos\left(\Delta\vartheta_i\right)$ differs as well. Additionally, prefactor and shift vary between oscillator pairs due to limited accuracy of A, resistors and capacitors. However, the design considers sufficient margin so that the range $[0\,\mathrm{V}, 5\,\mathrm{V}]$ is not exceeded. (For e.g. $A = 1{,}4\,\mathrm{V}$, which was used most of the time, the prefactor is $1{,}9\,\mathrm{V}$.) In order to obtain precise values for $\alpha_i = \cos\left(\Delta\vartheta_i\right)$, calibration measurements are taken with fully synchronized pairs at $\Delta\vartheta_i \in (0, \pi)$, which allows precise mapping of readout voltages to $\alpha_i = \pm 1$.(See Sec. 5.4).

4.1.3 Frequency annealing implementation

Frequencies of the oscillator pairs in MONACO have to be very similar and manual adjustment is not sufficient due to thermal drifts. In order to implement the automatic frequency annealing method proposed in Sec. 3.6.2, the frequency of one oscillator must depend on a continuous, periodic function F of the phase difference $\Delta\vartheta_i$. After the frequencies are annealed, the frequency of the annealed oscillator must stay constant independent of $\Delta\vartheta_i$. Our implementation can be divided into three steps:

- Finding a feedback voltage $U_F(\Delta\vartheta_i)$:
  The readout voltage U'_α (Sec. 4.1.2) can be reused as feedback signal U_F: It is a linear function of $\cos\Delta\vartheta_i$ in the range $]0\,\mathrm{V}, 5\,\mathrm{V}[$ under ideal conditions.

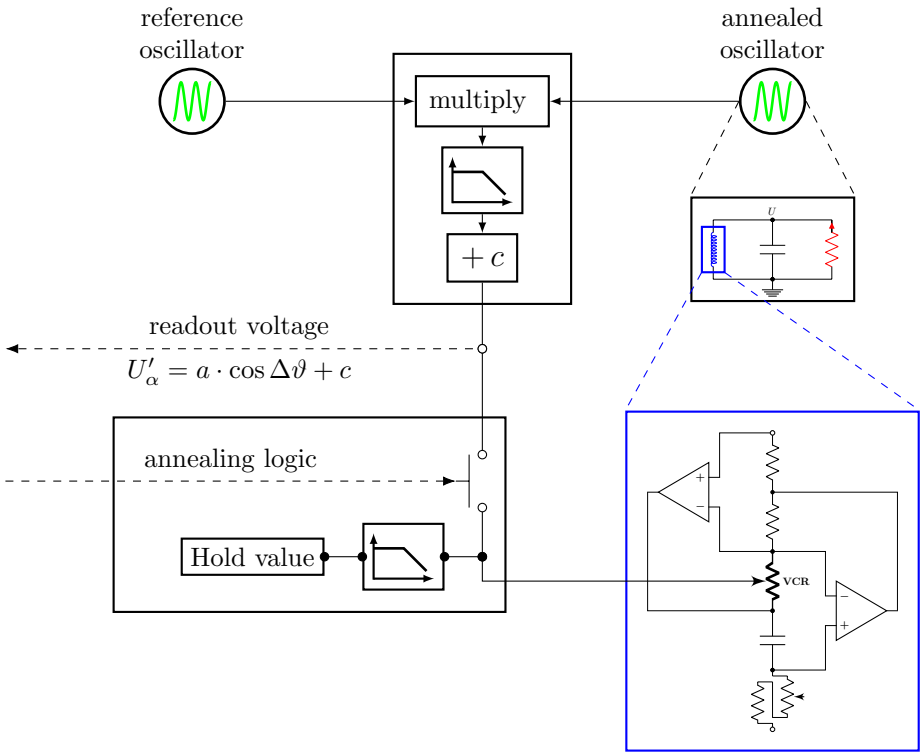


Figure 4.3: Scheme of the frequency annealing implementation: First, we require a feedback voltage that is a function of the phase difference $\Delta\vartheta$: We use the readout voltage U'_α, which is created by multiplying oscillator signals, low-passing the result and adding an offset.(See also Sec. 4.1.2) Second, the feedback voltage must influence the frequency of the annealed oscillator. For this purpose, we included a voltage controlled resistor (VCR) into the gyrator, which acts as the oscillator's inductance. Third, the feedback loop must be disconnected after the frequencies are annealed, as otherwise the phase difference $\Delta\vartheta$ would remain fixed as well. Additionally, the control voltage on the VCR must remain unchanged. This is facilitated by a Sample and Hold IC, which passes the voltage and continuously samples its value. When a logic signal is applied, the IC outputs the sampled value instead. This hold value is additionally low-passed to reduce influence of noise and remnants of the doubled frequency component from the creation of U'_α.

In reality, however, small remnants of the doubled base frequency are still present.

- Enabling/disabling feedback on the annealed oscillator:
  Phase difference and feedback voltage change until the frequency is annealed and stay constant afterwards. Replacing the feedback voltage with its constant value allows $\Delta\vartheta_i$ to change again, but keeps the frequency at the annealed value. Luckily, there are so-called sample-and-hold circuits which do exactly that: They allow an analog signal to pass until a logic signal changes. Then, they output the last sampled value. We used the IC LF398 with a hold capacitor 220 nF loaded through a serial 10 kΩ resistor.

- Finally, the frequency of the annealed oscillator must depend on U_F, in particular $\Delta\dot{\vartheta}_i = \Delta\Omega_i - F\left(U_F(\Delta\vartheta_i)\right)$.

Out of the three main elements of a Van der Pol oscillator shown in Fig. 2.7a, the inductance can be varied most easily: The nonlinear element does not change frequency substantially and a voltage-controlled capacitance (available as "varactor") disturbs oscillation shape. Our inductance is implemented as an Antoniou gyrator circuit (Fig. 2.5a) and a potentiometer is already used to roughly adjust the inductance manually. Therefore, we decided to change the gyrator's inductance L by including a voltage-controlled resistor.

The inductance of the Antoniou gyrator is given as $L = \frac{R_I R_{II} R_{IV} C}{R_{III}}$ for ideal opamps. However, changes in phase response curve and signal shape can be observed at high frequencies if the four resistor values are not sufficiently similar. These deviations are lowest for R_I, so the 1 kΩ potentiometer for manual adjustment with the 100 Ω serial resistor remains on this position. In order to keep the frequency annealing influence constant between pairs of different frequency, the voltage-controlled resistor was not combined with the manual adjustment, but replaced resistor R_{II}. The implemented voltage-controlled resistor needs to be close to $R_{VCR} \approx 1\,\text{k}\Omega$ in order to make distortion negligible. $R_{III} = R_{IV}$ were adjusted to 1,1 kΩ to better cover the frequency range.
The voltage controlled resistor was inserted into the gyrator of one oscillator, henceforth called *annealed oscillator*. The other oscillator of a pair, called *reference oscillator*, contains a fixed 1 kΩ resistor instead of the VCR.

An important design choice is the tradeoff between annealing range and annealing accuracy: Annealing range is given as the range of F. Our readout signal, however, contains remnants of the doubled frequency of the oscillators: Consequently, the feedback loop slightly oscillates around the fixed point for an ideal F. As the sample-and-hold circuit might pick any value during that oscillation, the amplitude of this oscillation is a major source of frequency annealing uncertainty. This frequency annealing error can be diminished by diminishing the influence of the feedback signal

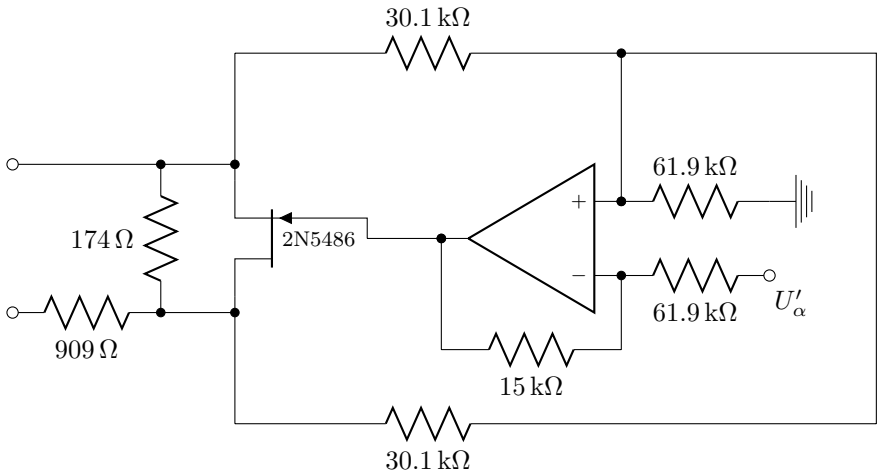


Figure 4.4: Circuitry for the voltage-controlled resistor used in the gyrator to facilitate frequency annealing. It's base is the JFET 2N5486, whose resistance is controlled by the voltage U'_α. A serial and a parallel resistor limit the total resistance to values $R_{VCR} \in [909\,\Omega, 909 + 174\,\Omega]$. As explained in Sec. 2.2.1, half of the JFET's drain and source voltages must be added to its gate to gain a well-defined resistance. This addition is done with a full summer, which can be seen as a combination of an inverting summer (Fig. 2.4a) and a non-inverting amplifier (Fig. 2.4b).

in the voltage controlled resistor which in turn also diminishes the range of F. Low-passing the hold value of the sample-and-hold IC further removes influence of the double base frequency without adding more delay to the feedback loop. The final circuit for the voltage controlled resistor is shown in Fig. 4.4: The main component is the JFET 2N5468, which is used as variable resistor as described in Sec. 2.2.1. The control voltage $U_C \approx -0.25\,U'_\alpha$ is added to the halved drain and source voltages in a full summer in order to obtain a resistor dependent on U_C only.
In order to limit the resistance variation of the JFET, it is accompanied by a serial resistor $R_{serial} = 909\,\Omega$ and a parallel resistor $R_{parallel} = 174\,\Omega$. Consequently, $R_{VCR} = R_{serial} + R_{parallel} = 1{,}08\,\text{k}\Omega$ for $R_{JFET} \to \infty$ and $R_{VCR} = R_{serial}$ for $R_{JFET} \to 0$. In the context of frequency annealing, the maximal resistance variation of $\pm 5\%$ translates to an inductance variation of $\pm 5\%$ which leads to an approximate frequency variation of $\pm 5\%$.

In Sec. 5.3.2, measurements verify that after annealing the remaining frequency difference is $\Delta f_i = \frac{\Delta\Omega_i}{2\pi} < 2\,\text{Hz}$, which corresponds to frequency deviations between oscillator pairs of less than 0.1‰.

4.1.4 Prevention of spurious coupling

In the first prototype circuits of oscillator pairs with frequency annealing, unwanted, "spurious" coupling between oscillator pairs existed, which disturbed both frequency annealing and recognition. As higher order terms average away, the relevant spurious coupling can be described as a function $G(\Delta\vartheta_i)$ of the phase difference.

In order to approximate the amount of spurious coupling, we observed otherwise uncoupled oscillator pairs: As $\Delta\dot{\vartheta}_i = \Delta\Omega_i + G(\Delta\vartheta_i)$, the pairs synchronize if $-\Delta\Omega_i$ is in the range of G, and pass each other otherwise. (Pairs of isolated fixed points emerge or annihilate at $-\Delta\Omega_i = \max(G)$ and $-\Delta\Omega_i = \min(G)$ in a saddle-node bifurcation; see Sec. 2.1.4.) Consequently, we manually adjusted pair frequencies until synchronization and noted the value of $\Delta f_i = \Delta\Omega_i/(2\pi)$ shortly before synchronization, which must be close to the range of G.

During his master's thesis, Nicolas Pfifferling measured $\Delta f_i \approx 20-60\,\mathrm{Hz}$ for single oscillator modules of the first generation, and $\Delta f_i \approx 40-150\,\mathrm{Hz}$ if they were plugged into the mainboard. (Note that the frequency difference is already diminished by G before full synchronization occurs, as $\Delta\dot{\vartheta}_i \ll \Delta\Omega_i$ for some values of $\Delta\vartheta_i$ and the phase difference thus spends a long time in this *bottleneck*. Consequently , the values we determined for the range of G can only be a lower limit on the real range of G.) These deviations modify frequency annealing and recognition:

- In the frequency annealing process, coupling is modified to $\Delta\dot{\vartheta}_i = \Delta\Omega_i - F(\Delta\vartheta_i) + G(\Delta\vartheta_i)$. In contrast to the ideal case $F(\Delta\vartheta_i{}^*) = \Delta\Omega_i$, fixed points $\Delta\tilde{\vartheta}_i{}^*$ are now determined by $F(\Delta\tilde{\vartheta}_i{}^*) = \Delta\Omega_i + G(\Delta\tilde{\vartheta}_i{}^*)$. Consequently, the additional error during frequency annealing will be $-G(\Delta\tilde{\vartheta}_i{}^*)$ up to the full range of G.

- Even with perfectly annealed oscillator pairs, a spurious coupling term favors synchronization at a specific phase difference. Recognition coupling must be increased to be much larger than the spurious coupling in order to make the modification negligible. The frequency difference from the incorrect frequency annealing, however, has to be overcome additionally and is much more critical to the recognition process.

In order to reduce the spurious coupling between pairs, we considered the following coupling causes:

- Signals from the oscillators can be carried on the power supply lines, as circuit components draw power depending on the phase of the oscillation. In the newer designs, we separated supply lines: Ferrite beads, which are inductors that also have increased ohmic resistance at high frequencies, were placed on the connections between common supply and oscillator supplies. Additionally, bypass capacitors between the supply rails and ground decrease the voltage

dips introduced by increased consumption and form a low pass with the Ferrite beads, further decreasing high frequency signals on the supply.

- Op-amps on the same IC had been shared between both oscillators, which might have induced a thermal or voltage coupling. In newer designs, each IC exclusively belongs to a single oscillator or other, unrelated circuitry.
- Signal lines running parallel to each other can transfer small amount of signals through induction. Consequently, oscillator signal lines were cut short and separated from the rest of the network with a non-inverting amplifier. This required to move coupling resistors close to the oscillators as well.
- Oscillator frequencies in the readout signal might be fed back into the annealed oscillator with the frequency annealing mechanism. We increased the filter order of the low-pass filter to remove higher frequencies more thoroughly.
- We found that significant radiation coupling existed between oscillators. Therefore, oscillators including coupling resistors, non-inverting amplifier and frequency annealing mechanism were mounted on extra circuit boards. These circuit boards were placed inside Faraday cages and connections to the network were implemented via socket and pin strips that were fed through small holes.

After implementing all standards and procedures mentioned above, uncoupled oscillator pairs cannot be synchronized with manual adjustment. Therefore, spurious coupling must now be below the oscillators' intrinsic frequency noise.

4.1.5 Oscillator module substructure

The final oscillator modules consist of a main module (Fig. C.2a) and two submodules shielded with a grounded iron box:
One submodule contains the reference oscillator (Fig. C.2b) including resistors R_{recog}, R_{init} and R_{comp} and a non-inverting amplifier for separating the internal oscillator voltage from the signal output. Potentiometers for adjustment of frequency and amplitude amplification can be tuned during the experiment through small holes in the Faraday cage. The other submodule (Fig. C.2c) contains the annealed oscillator. In addition to all components of the reference, it contains the Sample-and-Hold IC LF398 and the voltage controlled resistor (Fig. 4.4) for frequency annealing. The oscillator module itself contains the circuitry for construction of the readout voltage and routes voltages between the submodules and the mainboard. That includes (filtered) power supply, oscillator signals, coupling voltages for both initialization and recognition, readout voltage and annealing logic.

Note that copper lines without connection to a voltage source act as antenna. As coupling voltage connections for initialization or recognition are disconnected

during other steps of the experiment, induced voltages could influence oscillators via the coupling resistors. In order to avoid spurious coupling, coupling voltage connections are grounded with a resistor $R_\perp = 301\,\Omega$. When the coupling voltages are connected, this grounding resistor forms a voltage divider with the ON-resistance $r_{DS(ON)} \approx 25\,\Omega$ of the conducting switch, which in turn diminishes the coupling voltage by a factor $c_{vd} = \frac{R_\perp}{R_\perp + r_{DS(ON)}} \approx 0.92$.

In order to measure the influence of external noise, the grounding resistor $R_\perp$ for U_{init} of the reference oscillator can be connected to a BNC connector with a jumper. During experiments with external noise, the connector is set to ground during initialization as usual, but induces additional noise via R_{init} during the recognition step.

4.2 Pattern modules

In the coupling modulations Eq. (3.4) for the recognition process, correct pattern candidates $\boldsymbol{\alpha}^m$ are included as signal prefactors in sums $\sum_k \alpha_k^m \sin(\vartheta_k^{[1/2]})$. Similarly, oscillators $1/2$ in the i^{th} pair are coupled to the initialization coupling voltage $U_{i,init}^{[1/2]} = \alpha_i^d A \sin(\vartheta_i^{[2/1]})$ during initialization. As pattern pixels are either 1 or -1, the multiplication of pixel and signal is rather a *conditional inversion* of the oscillator signals. Each pattern module conducts all conditional inversions for one oscillator pair depending on external digital signals(Fig. 4.1). (Those are 8 in total: 3+1 pixels for the correct pattern candidates and the initial, defective pattern for both oscillators of a pair.)

The pattern modules contain the circuit in Fig. 2.4c (See also [37], Chapter 4.09, "optional inverter") for each conditional inversion with the stripboard layout shown in Fig. C.4. The operational amplifiers used are TL074 ICs, the analog switches are DG411. In order to avoid non-idealities due to the input resistance of the op-amp or the ON-resistance of the analog switch, we chose $R_1 = 20\,\text{k}\Omega$ and $R_2 = 10\,\text{k}\Omega$ for the resistors. Additional to the optional inverter, the modules contains the input resistors for the inverting summers on the coupling voltage module.

4.3 Coupling voltage module

This module constructs the final coupling voltages U_{recog} for the recognition coupling. Oscillator signals $A \sin(\vartheta_i^{[1/2]})$ and their product with memorized patterns $\alpha_i^m A \sin(\vartheta_i^{[1/2]})$ are supplied by the pattern modules. The coupling voltages are computed by combining inverting summers (Sec. 2.2.2 and Fig. 2.4a) and multiplication ICs AD633: (Sec. 2.2.1)

1a) Oscillator signals $A \sin(\Delta\vartheta_i^{[1/2]})$ are summed across all pixel indices with an

inverting summer. The inverted sums are also amplified by a scale factor $c = \frac{R_{out}}{R_{in}} = \frac{5{,}62\,\mathrm{k\Omega}}{10\,\mathrm{k\Omega}}$. Note that the input resistors R_{in} of the summers are actually on the pattern modules.

1b) For each memorized pattern $\boldsymbol{\alpha}^m$, products of oscillator signals and pattern pixels $\alpha_i^m A \sin\left(\Delta\vartheta_i^{[2/1]}\right)$ are summed analogously, but scaled with $c_\alpha = \frac{R_{out}}{R_{in}} = \frac{6{,}81\,\mathrm{k\Omega}}{10\,\mathrm{k\Omega}}$.

2) Simple high passes with $C = 1\,\mu\mathrm{F}$ and $R = 243\,\mathrm{k\Omega}$ remove any offsets from the signal sums 1a) and 1b). As the cutoff frequency $f_c = 1/(2\pi RC) \approx 0{,}65\,\mathrm{Hz}$ is very low, phase and amplitude of the oscillators' signals is preserved. While offsets proved to be negligible, the high passes were still kept to protect the multiplier ICs from shorts.

3) Sums of weighted oscillator signals 1b) are squared with multiplier ICs AD633.

4) Squares from 3) are summed over all patterns in another inverting summer with $c_m = \frac{R_{out}}{R_{in}} = \frac{1\,\mathrm{k\Omega}}{2\,\mathrm{k\Omega}}$, which creates a voltage proportional to the coupling modulations $a^{[1/2]}$.

5) Signal sums 1a) are modulated with the coupling modulation 4) through multiplication with another AD633.

$$U_{recog}^{[1/2]} = \left[-c_m \sum_{m=1}^{M} \frac{1}{10\,\mathrm{V}} \left[-c_\alpha \sum_{k=1}^{N} \alpha_k^m A \sin\left(\vartheta_k^{[2/1]}\right)\right]^2\right] \cdot \frac{1}{10\,\mathrm{V}} \cdot \left[-c \sum_{j=1}^{N} A \sin\left(\vartheta_j^{[1/2]}\right)\right]$$

$$U_{recog}^{[1/2]} = \frac{c c_m {c_\alpha}^2 A^3}{100\,\mathrm{V}^2} \sum_{m=1}^{M} \left[\sum_{k=1}^{N} \alpha_k^m \sin\left(\vartheta_k^{[2/1]}\right)\right]^2 \sum_{j=1}^{N} \sin\left(\vartheta_j^{[1/2]}\right) \tag{4.4}$$

A stripboard layout for the complete coupling voltage module is shown in Fig. C.5a.

4.4 Mainboard

The mainboard's purpose is the distribution of logic control voltages and power supply as well as providing necessary interconnections between modules. Modules are connected to the mainboard via pin-and-socket connectors and are arranged perpendicular to the mainboard. Analog switch ICs DG412 connect coupling voltages to oscillator modules depending on the logic signals for initialization and recognition step.

For MONACO I, oscillator modules, pattern modules and the coupling voltage module are plugged in.

The architecture variant MONACO II reuses oscillator modules and pattern modules, but constructs individual voltages for each pixel with 1 summation module and 8 multiplier modules. (See Sec. 4.6.) As multiplier modules are not needed for MONACO I, small "bridging" modules (Fig. C.5b) connect coupling voltages across the otherwise unused sockets. See Fig. C.3 for the mainboard layout.

4.5 Coupling strength derivation: MONACO I

The recognition coupling voltage Eq. (4.4) (diminished by c_{vd}, see Sec. 4.1.5) affects the internal voltage of single oscillators (Sec. 4.1.1) which in turn shifts their phases via the phase response curves(Sec. 2.1.2):

$$\begin{aligned}\dot{\vartheta}_i^{[1/2]} &= \Omega_i + A_{PRC}\cos\vartheta_i^{[1/2]} \cdot \frac{1}{C_{tot}R_{recog}} \cdot c_{vd} \cdot U_{recog}^{[1/2]} \\ \dot{\vartheta}_i^{[1/2]} &= \Omega_i + \frac{A_{PRC}c_{vd}}{C_{tot}R_{recog}} \frac{cc_m c_\alpha^2 A^3}{100\,\mathrm{V}^2} \\ &\qquad \cdot \cos\vartheta_i^{[1/2]} \sum_{m=1}^{M} \left[\sum_{k=1}^{N} \alpha_k^m \sin\left(\vartheta_k^{[2/1]}\right) \right]^2 \sum_{j=1}^{N} \sin\left(\vartheta_j^{[1/2]}\right)\end{aligned} \tag{4.5}$$

Note that the soldered capacitance $C = 1{,}0\,\mathrm{nF}$ is replaced with the total capacitance $C_{tot} = C + C_{par} = 1{,}2\,\mathrm{nF}$, as measurement of phase response curves (Sec. 5.1.2) revealed a parasitic capacitance C_{par} in parallel to C.

Comparison with the full phase description Eq. (3.1) and Eq. (3.4) yields:

$$\epsilon = N \frac{A_{PRC}c_{vd}}{C_{tot}R_{recog}} \frac{cc_m c_\alpha^2 A^3}{100\,\mathrm{V}^2} \tag{4.6}$$

Typical values for parameters in Eq. (4.6) are shown in Table 4.2. Table 4.3 shows all values of ϵ used as well as coupling resistance and oscillator signal amplitudes.

4.5.1 Coupling strength error calculation

Most parameters in Eq. (4.6) show non-negligible deviations, which are shown in Table 4.4.

Influence of these deviations Δv_i on ϵ are approximated with Taylor expansion and combined with linear error propagation, as all of them are assumed to be systematic and independent:

$$\Delta\epsilon = \sum_i \left| \frac{\mathrm{d}\,\epsilon}{\mathrm{d}v_i} \Delta v_i \right|$$

parameter	symbol	typical value(s)
number of pixels	N	8
amplitude of the phase response curve (Sec. 5.1.2)	A_{PRC}	$2{,}2\,\frac{\text{rad}}{\text{V}}$
soldered capacitance	C	1,0 nF
total capacitance (Sec. 5.1.2)	C_{tot}	1,2 nF
oscillator amplitude after amplification	A	1,40 V
amplification in pure signal sums	c	$\frac{5{,}62\,\text{k}\Omega}{10\,\text{k}\Omega} = 0.562$
amplification in weighted signal sums	c_α	$\frac{6{,}81\,\text{k}\Omega}{10\,\text{k}\Omega} = 0.681$
amplification in sum of squares	c_m	$\frac{1{,}0\,\text{k}\Omega}{2{,}0\,\text{k}\Omega} = 0.5$
amplification voltage divider	c_{vd}	$\frac{301\,\Omega}{301\,\Omega+25\,\Omega} \approx 0.92$
analog switch DG411DJZ ON-resistance	$r_{DS(ON)}$	$25\,\Omega$
grounding resistor	$R_\perp$	$301\,\Omega$

Table 4.2: Typical values for deriving the coupling strength ϵ. Values for the coupling resistance R_{recog}, which were varied to alter ϵ, are not shown. In a few measurements, the amplified oscillator amplitude A was additionally changed to 125 mV, 135 mV and 145 mV.

$R_{recog}\,[\text{k}\Omega]$	1000	511	511	511	511	243	100	62	30	10,2
$A\,[\text{V}]$	1,4	1,25	1,35	1,4	1,45	1,4	1,4	1,4	1,4	1,4
$\epsilon\,\left[\frac{\text{rad}}{\text{s}}\right]$	48,3	67,2	84,7	94,4	105	199	483	778	1608	4730

Table 4.3: List of all used coupling strengths ϵ including the respective coupling resistances R_{recog} and amplified oscillator amplitudes A.

$$\begin{aligned}\frac{\Delta\epsilon}{\epsilon} &= \left|\frac{\Delta A_{PRC}}{A_{PRC}}\right| + \left|\frac{\Delta C}{C}\right| + 3\underbrace{\left|\frac{\Delta A}{A}\right|}_{A\geq 125\,\text{mV}} + 9\left|\frac{\Delta R}{R}\right| + \underbrace{\left|\frac{r_{DS(ON)}}{R_\perp + r_{DS(ON)}}\cdot\frac{\Delta R_\perp}{R_\perp}\right|}_{\approx 0.1\left|\frac{\Delta R}{R}\right|\approx 0} + \left|\frac{\Delta r_{DS(ON)}}{R_\perp + r_{DS(ON)}}\right| \\ &= 0.09 + 0.04 + 0.12 + 0.09 + 0.05 \\ \frac{\Delta\epsilon}{\epsilon} &= 39\% \end{aligned} \tag{4.7}$$

Note that the calculation above is an approximation: As factors c_m, c and c_α in Eq. (4.5) each represent several resistor pairs in the summation circuitry, the deviations of those pairs ought to be considered independently with factors c_{mm}, c_j and $c_{\alpha k}$. In this case, however, Eq. (4.6) cannot be derived as shown, as the factors cannot

parameter v_i	error source	error size Δv_i
metal film resistors	production accuracy	$\Delta R = 1\%R$
$r_{DS(ON)}$	production accuracy	$r_{DS(ON)} \in [0\,\Omega, 35\,\Omega]$
A	thermal drifts	$\Delta A = 5\,\mathrm{mV}$
A_{PRC}	f-dependence	$\Delta A_{PRC} = 0{,}2\,\frac{\mathrm{rad}}{\mathrm{V}}$
C_{tot}	production accuracy; C_{par}	$\Delta C_{tot} = 4\%C_{tot}$

Table 4.4: List of parameter uncertainties for determination of the coupling strength uncertainty. Uncertainties for resistances were extracted from the respective data sheets. Extraction of uncertainties for A_{PRC} and C_{tot} is covered in Sec. 5.1.2 and Appendix D. Uncertainties for A are an upper bound on drifts observed during long measurement series.

be put in front of the corresponding sums and ϵ is only well-defined for single signal paths described by a tuple [m,k,j]. Eq. (4.7) describes the ϵ-deviation in such a signal path. As deviations of the resistors are quite small compared to other deviations and those are the same for all signal paths, Eq. (4.7) is a sufficient approximation of the combined error. Implementation and coupling strength derivation is similar for the architecture variant MONACO II (Sec. 3.7):

4.6 Architecture variant with 2N coupling modulations: MONACO II

In MONACO II, oscillator modules and pattern modules can be used without any changes. However, every pixel receives a different coupling modulation (Eq. (3.23)) and thus a different coupling voltage $U^{[1/2]}_{recog,i}$. Consequently, it is useful to split the construction of $U^{[1/2]}_{recog,i}$ into voltages required for all coupling voltages and the individual deviations.

4.6.1 Summation module

The summation module receives oscillator signals $A\sin(\vartheta_i^{[1/2]})$ and weighted signals $\alpha_i^m A\sin(\vartheta_i^{[1/2]})$ from the pattern modules and sums them over all pixels. These sums are required for each pixel's coupling voltages. Analogous to MONACO I, the sums are calculated with inverting summers, but with scaling factors $\tilde{c} = \frac{R_{out}}{R_{in}} = \frac{1\,\mathrm{k\Omega}}{10\,\mathrm{k\Omega}} = 0.1$ for pure signal sums and $\tilde{c}_\alpha = \frac{R_{out}}{R_{in}} = \frac{8{,}25\,\mathrm{k\Omega}}{10\,\mathrm{k\Omega}} = 0.825$ for sums of weighted oscillator signals. (See Fig. C.6a for a stripboard layout.) The calculated sum voltages are transferred to all multiplier modules via the mainboard.

4.6.2 Multiplier module

In the multiplier modules, individual coupling voltages are derived for each pixel:

1) First, weighted signal sums from the summation module and single weighted signals from the i^{th} pattern module are multiplied and summed to derive individual coupling modulations $a_i^{[1/2]}$: The summation input Z of $M = 3$ AD633 ICs is used to sum over the m multiplications and scale the output with $\tilde{c}_m = \frac{2\,\text{k}\Omega}{1\,\text{k}\Omega} = 2$ as described in Sec. 2.2.1, so no additional components are necessary.

2) The sum of oscillator signals is modulated with the coupling modulation via another multiplier IC without scaling, gaining the coupling voltages $U_{recog,i}^{[1/2]}$:

$$\begin{aligned} U_{recog,i}^{[1/2]} &= \left[\sum_{m=1}^{M} \tilde{c}_m \left[\alpha_i^m A \sin\left(\vartheta_i^{[2/1]}\right)\right] \cdot \frac{1}{10\,\text{V}} \cdot \left[-\tilde{c}_\alpha \sum_{k=1}^{N} \alpha_k^m A \sin\left(\vartheta_k^{[2/1]}\right)\right]\right] \\ &\quad \cdot \frac{1}{10\,\text{V}} \cdot \left[-\tilde{c} \sum_{j=1}^{N} A \sin\left(\vartheta_j^{[1/2]}\right)\right] \\ U_{recog,i}^{[1/2]} &= \frac{\tilde{c}_m \tilde{c}_\alpha \tilde{c} A^3}{100\,\text{V}^2} \sum_{m=1}^{M} \alpha_i^m \sin\left(\vartheta_i^{[2/1]}\right) \sum_{k=1}^{N} \alpha_k^m \sin\left(\vartheta_k^{[2/1]}\right) \cdot \sum_{j=1}^{N} \sin\left(\vartheta_j^{[1/2]}\right) \end{aligned} \tag{4.8}$$

Fig. C.6b shows the stripboard layout for the multiplier module.

4.6.3 Coupling strength derivation: MONACO II

Analog to Sec. 4.5, we first derive the phase dynamics in the circuitry:

$$\begin{aligned} \dot{\vartheta}_i^{[1/2]} &= \Omega_i + A_{PRC} \cos \vartheta_i^{[1/2]} \cdot \frac{1}{CR_{recog}} \cdot c_{vd} \cdot U_{recog,i}^{[1/2]} \\ \dot{\vartheta}_i^{[1/2]} &= \Omega_i + \frac{A_{PRC} c_{vd}}{CR_{recog}} \frac{\tilde{c}_m \tilde{c}_\alpha \tilde{c} A^3}{100\,\text{V}^2} \sum_{m=1}^{M} \alpha_i^m \sin\left(\vartheta_i^{[2/1]}\right) \sum_{k=1}^{N} \alpha_k^m \sin\left(\vartheta_k^{[2/1]}\right) \cdot \sum_{j=1}^{N} \sin\left(\vartheta_j^{[1/2]}\right) \end{aligned} \tag{4.9}$$

We derive the coupling strength by comparing with MONACO II's phase dynamics Eq. (3.23):

$$\epsilon = N \frac{A_{PRC} c_{vd}}{CR_{recog}} \frac{\tilde{c}_m \tilde{c}_\alpha \tilde{c} A^3}{100\,\text{V}^2} \tag{4.10}$$

Measurements for MONACO II have been made by Nicolas Pfifferling in the context of his master's thesis and are shown in Appendix E. As he had to use a previous

parameter	symbol	typical value(s)
amplitude of the PRC (dependence on f_i)	A_{PRC}	$1{,}0\,\frac{\text{rad}}{\text{V}} - 2{,}5\,\frac{\text{rad}}{\text{V}}$
oscillator amplitude after amplification	A	$\approx 1{,}6\,\text{V}$
amplification in pure signal sums	$\tilde{c}$	0.1
amplification in weighted signal sums	$\tilde{c}_\alpha$	0.825
amplification in coupling modulation construction	$\tilde{c}_m$	2

Table 4.5: Parameter values for MONACO II. Oscillator modules used were preliminary and thus phase response curves were not ideally sinusoidal and showed strong a strong frequency dependence in A_{PRC}. Additionally, amplified oscillator amplitudes A could not be adjusted during operation yet and thus have a larger, unknown error. All parameters not mentioned were the same as in Table 4.2.

generation of oscillator modules, phase response curves showed strong deviations from a cosine and amplitudes varied strongly. Additionally, oscillator amplitudes A could not be adjusted during network operation and were thus susceptible to drifts. Consequently, all results for MONACO II are only preliminary. Typical values for parameters in Eq. (4.10) are summarized in Table 4.5. As uncertainties are so high, coupling strengths are not calculated, and measurements are presented as function of coupling resistors R_{recog} instead. Both networks were measured and controlled with the same measurement environment, which also supplied the patterns as digital signals. It is presented in the next section.

4.7 Controlling the networks: Patterns and recognition procedure

This section explains the procedure of a recognition process and outlines digital signals and hardware necessary to control the network. In Fig. 4.5, a successful recognition process is shown. It is subdivided into 4 steps, which are controlled via 3 digital signals, that need to be either LOW (0 V) or HIGH (5 V) in each step:

- *Annealing step*: Frequencies are annealed with a feedback loop as described in Sec. 4.1.3. As the Sample-and-Hold ICs LF398 on the oscillator boards (Sec. 4.1) need to pass the feedback signal, the *annealing logic* needs to be HIGH. Both initialization coupling and recognition coupling are disconnected at the mainboard by the analog switches DG412, so *initialization logic* and *recognition logic* need to be LOW.
- *Hold step: Annealing logic* is switched to LOW between annealing and hold

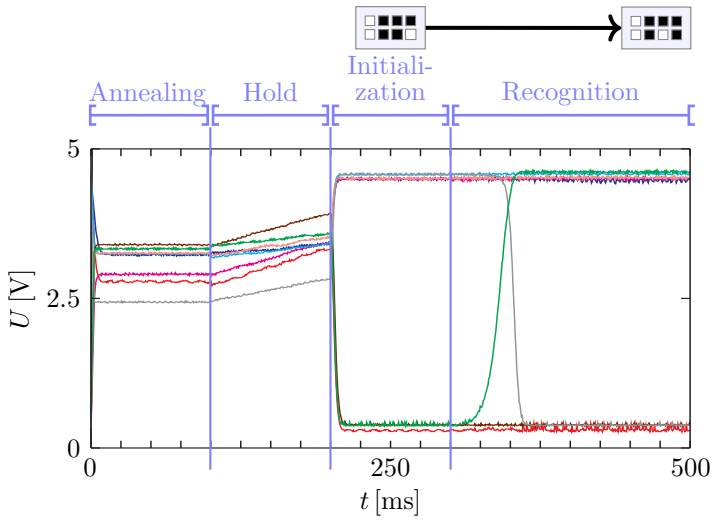


Figure 4.5: Exemplary measurement of the readout voltages (Sec. 4.1.2) during a successful recognition process of MONACO I ($\epsilon = 199\,\frac{\text{rad}}{\text{s}}$, random memorized patterns). Each colored trajectory corresponds to one of the $N = 8$ pixels. All steps of a recognition process are shown: In the *annealing* step, oscillator pairs are coupled in a feedback loop that decreases frequency differences Δf_i to approx. 0 at a fixed voltage. At the beginning of the *hold* step, the feedback loop is removed and uncoupled oscillators can be observed. Minor remaining frequency differences Δf_i show up as approx. linear drifts. With the onset of initialization coupling, oscillator pairs synchronize at phase differences 0 (large U) or π (small U) representing a pixel of the defective pattern α_i^d. When the recognition step starts, initialization coupling is removed and replaced with the recognition coupling, which has attractors for each memorized pattern $\boldsymbol{\alpha}^m$. Pixels move to the closest attractor, which is the memorized pattern most similar to $\boldsymbol{\alpha}^d$.

step. Thus, the Sample-and-Hold ICs output a constant voltage and the feedback is removed. All oscillators experience no coupling, and small remaining frequency differences Δf_i are visible as drifts.

- *Initialization step:* The DG412 connects the initialization coupling voltage by setting the *initialization logic* to HIGH. Consequently, oscillator pairs synchronize at phase differences according to the pixels of the defective pattern.

- *Recognition step:* Initialization coupling is disconnected and recognition coupling is simultaneously connected by setting the *initialization logic* to LOW again and the *recognition logic* to HIGH.

In addition to controlling the steps, patterns need to be provided to the conditional inverters on the pattern modules as digital signals as well: Here, a HIGH signal

represent $\alpha_i = +1$, or black pixels, while a LOW signal equals $\alpha_i = -1$ / white pixels. Summing up, digital signals have to be supplied for both the $(M+1)N = 32$ pattern pixels as well as the 3 control signals for the steps of the recognition process. Additionally, we need to measure the $N = 8$ analog voltage signals that represent α_i. As the phase differences change slowly compared with the oscillator frequencies and the patterns do only change between measurements, both do not need to be fast. The control signals need to be faster than the $\boldsymbol{\Delta\vartheta}$-dynamics, but are digital output.

We selected the microcontroller board Arduino Mega 2560 with the 8-bit microcontroller ATmega2560 from Atmel, which is clocked at 16 MHz. It provides 54 digital pins for input and output and 16 analog input pins with a 10-bit resolution between 0 and 5 Volt. Additionally, it is comparably cheap and can be programmed in C without additional equipment. Regarding limitations, the Arduino cannot directly output analog signals (except with pulse-width-modulation, which is slow). Additionally, analog sample rates are in the low kHz for standard settings and a single signal, and are further reduced if several signals are measured, as the Arduino has only a single, multiplexed AD-converter. Still, it is suitable for controlling the network.

Some experiments, like the measurement of the phase response curves $Z(\vartheta)$ of the oscillators, have more stringent requirements: A direct measurement of the oscillator signals and precise fabrication of sharp voltage peaks is required. For other measurements, high quality Gaussian white noise needs to be generated. All of these tasks are beyond the capabilities of the Arduino and thus we needed additional hardware.

4.7.1 Hardware for fast analog reads and analog signal generation

The AD-conversion card Spectrum M2i.4032 was used for direct measurements of fast analog signals, e.g. the oscillators' voltages. It offers 4 analog input channels with a 14-bit resolution, adjustment of the input voltage range and a sample rate up to 50 MS/s.

The generation of fast analog signals, like noise generation or sharp Gaussian peaks for the measurement of phase response curves, DA-conversion cards Spectrum M2i.6021, M2i.6011 and M2i.6012 were used. They offer 2 (M2i.6021 and M2i.6011) or 4 analog output channels (M2i.6012) with 14-bit resolution and 20 MS/s (M2i.6011 and M2i.6012) or 60 MS/s (M2i.6021) maximum sample rate.

4.7.2 Measurement software

The measurement of pattern recognition processes is coordinated via a C++/Qt-program, which directly controls all spectrum cards and communicates with the

Arduino via a serial connection: The control program sends a "request", and a C-program on the Arduino reacts and sends a "response". Different commands change step durations, transfer new patterns, set digital pins according to the patterns or start triggered or untriggered measurements. Additional commands for debugging include setting a constant Hold or Annealing mode. As these commands can be read from a script, measurement processes could be adjusted as necessary.

For the measurement of phase response curves, another C++/QT program controls the spectrum cards M2i.4032 and M2i.6021. As further elaborated in Sec. 5.1.2, the M2i.6021 creates sharp Gaussian peaks to perturb an oscillator. The oscillators' signal as well as the peaks are measured with the M2i.4032.

Both programs were based on software written by Robert Hölzel.

Chapter 5

Experiments

We already presented a typical measurement of readout voltages and different steps of a recognition process in Sec. 4.7. Now, we will discuss several experiments analyzing our networks and their components: First, basic properties of oscillators and network circuitry, namely phase response curves $Z(\vartheta)$, frequency differences Δf_i and coupling strength ϵ will be discussed. Second, we will evaluate the performance of the network for different coupling strengths and pattern compositions: Recognition success as well as recognition times are derived and analyzed. Finally, we will uncover the influence of noise on network dynamics and performance.

5.1 Phase response curves

In our implementation of MONACO, oscillators are coupled to the coupling voltages via resistors, which in turn changes the internal voltages of the oscillators. The oscillator's phase changes as a reaction to the perturbations of its internal variables (e.g. the voltage), which is described by phase response curves (PRCs, see Sec. 2.1.2).

Consequently, we need to determine the oscillators' phase response to voltage perturbations in order to derive the coupling strength ϵ in our network. Furthermore, we assumed a sinusoidal PRC shifted by $\frac{\pi}{2}$ with respect to the oscillation in Chapter 3 and need to review that assumption for our implementation of Van der Pol oscillators.

We slightly modify the method used in [30, 31]: The PRC $Z(\vartheta) = \frac{\mathrm{d}\vartheta}{\mathrm{d}U}$ can be measured one point at a time by perturbing the internal voltage U with an instantaneous, small shift ΔU and measuring the resulting phase shift $\Delta\vartheta$. Instead of directly manipulating U, however, we couple an oscillator to an external voltage U_{ext} through a resistor R_{ext}(See Sec. 4.1.1.):

$$\dot{\vartheta} = \Omega + Z(\vartheta) \cdot \frac{1}{R_{ext}C_{tot}} U_{ext}(t) \tag{5.1}$$

Instantaneous voltage shifts ΔU can be induced by applying a delta-function in U_{ext}

at $t = t_p$, which can be approximated by a sharp Gaussian peak:

$$U_{ext} = A_p \mathrm{e}^{-\frac{(t-t_p)^2}{2\sigma^2}}$$

Integrating Eq. (5.1) between times t_1 and t_2 before and after a perturbation, we can assume $Z(\vartheta) = Z(\vartheta_p)$ is constant during the short perturbation at $t = t_p$:

$$\begin{aligned}
\vartheta(t_2) - \vartheta(t_1) &= \Omega(t_2 - t_1) + \int_{t_1}^{t_2} Z(\vartheta) \cdot \frac{1}{R_{ext}C_{tot}} U_{ext}(t)\,\mathrm{d}t \\
&= \Omega(t_2 - t_1) + Z(\vartheta_p) \int_{t_1}^{t_2} \frac{1}{R_{ext}C_{tot}} U_{ext}(t)\,\mathrm{d}t \\
&= \Omega(t_2 - t_1) + Z(\vartheta_p)\Delta U
\end{aligned}$$

The voltage perturbation ΔU is independent of the phase of the oscillator:

$$\begin{aligned}
\Delta U &= \int_{t_1}^{t_2} \frac{1}{R_{ext}C_{tot}} U_{ext}(t)\,\mathrm{d}t \\
&= \frac{\sqrt{2\pi}\sigma A_p}{R_{ext}C_{tot}} \\
\Delta U &= \frac{\sqrt{\pi} \cdot \mathrm{FWHM} \cdot A_p}{2\sqrt{\ln 2} R_{ext}C_{tot}}
\end{aligned} \tag{5.2}$$

Here, $\mathrm{FWHM} = 2\sqrt{2\ln 2} \cdot \sigma$ is the Gauss peak's full width at half maximum, A_p the Gauss peak amplitude, R_{ext} is the coupling resistor and C_{tot} the total capacitance of the oscillator including soldered as well as parasitic capacitances.

After deriving ΔU, we need to determine the phases ϑ_p at the time t_p of the perturbation and the resulting phase shifts $\Delta\vartheta$.

It is advantageous to determine times when the voltage crosses 0 V with positive slope and use them as reference for phase determination. Thus, we measured times t_1,t_2 at zero-crossings before and after t_p.

Then, $\vartheta(t_2) - \vartheta(t_1) = 2\pi \cdot n$:

$$2\pi \cdot n = \underbrace{\Omega}_{\frac{2\pi}{T}}(t_2 - t_1) + \underbrace{Z(\vartheta_p)\Delta U}_{\Delta\vartheta} \tag{5.3}$$

$$\Delta\vartheta = -\left(\frac{t_2 - t_1}{T} - n\right) \cdot 2\pi \qquad \text{with} \qquad \vartheta_p = \left(\frac{t_p - t_1}{T} - \tilde{n}\right) \cdot 2\pi \tag{5.4}$$

Here, n is the number of actual traversed periods between t_1 and t_2, and $\tilde{n}$ is the

number of full periods before the perturbation. The period T is similarly derived by comparing times between zero-crossings before the perturbation.

Note that it is unfavorable to choose zero-crossings directly before or after the perturbation for t_1 and t_2 (as used in [30, 31]), as the Gaussian peak can disturb the measurement of the crossings: In such measurements with $n = 1$ and $\tilde{n} = 0$ and a positive Gaussian peak, the PRC shows a narrow, deep dip at 0 / 2π and a large peak directly before the dip. The dip originates from Gaussian peaks slightly after a zero-crossing, where the left tail of the peak had already increased phase before the crossing occurred. If this crossing before the peak is used as time reference t_1, $\Delta\vartheta$ is smaller and the PRC shows the dip. Similar deductions can be made for peaks slightly before a zero-crossing, where the charge induced by the peak has not been completely transferred yet, which would result in a dip as well. However, another effect dominates the region before 2π: If a peak occurs shortly before a crossing, the induced voltage perturbation briefly results in positive values of the oscillator voltage, which is erroneously detected as a crossing. As the Gaussian occurs before the real zero-crossing, the time difference in Eq. 5.4 is reduced, which results in the PRC's erroneous peak. Note that the PRC is shifted horizontally due to signal delay, so the dip and the peak can be observed at lower phases than expected. By using an earlier crossing as time reference t_1, it cannot be shifted and the dip can be avoided. Similarly, a crossing several periods after the Gaussian peak cannot be confused with the perturbation. Note that the number of zero-crossings occurring between t_1 and t_2 needs to be rectified, so crossings are counted only once if they differ by a fraction of the period only.

We chose the 3$^{\text{rd}}$ positive zero-crossing after the perturbation and the 2$^{\text{nd}}$ one before for the determination of $\Delta\vartheta$. T was derived from the 5$^{\text{th}}$ and the 2$^{\text{nd}}$ zero-crossing before the perturbation. Additionally, $R_{ext} = 8{,}25\,\text{k}\Omega$ had to be matched with an additional compensation resistor $R_{comp} = 100\,\text{k}\Omega$ parallel to the negative resistance $R_{neg} = 6{,}19\,\text{k}\Omega$ to measure at the same effective negative resistance $R'_{neg} = 34\,\text{k}\Omega$ as used in the experiment. The applied Gaussian peaks in U_{ext} had an amplitude $A_p = 1500\,\text{mV}$ and a full width at half maximum FWHM $= 0{,}1\,\mu\text{s}$.

In order to discern behavior of the ideal circuit from non-ideal deviations we conducted simulations in addition to direct measurements of the soldered oscillators.

5.1.1 PRC simulations

Phase response curves deviate from the ideal, sinusoidal shape depending on frequency, non-ideal behavior of circuit components and measurement uncertainties.

We conducted simulations of voltage and current in order to identify the ideal behavior of a Van der Pol oscillator with the non-linear element used: The capacitor

and gyrator were modeled as ideal capacitance C_{tot} / inductance L:

$$\dot{U} = \frac{1}{C_{tot}}\dot{Q} = -\frac{1}{C}I$$
$$\dot{I_L} = \frac{1}{L}U$$

Here, I is the total current flowing away from U including current from the nonlinear element, resistances and the inductance:

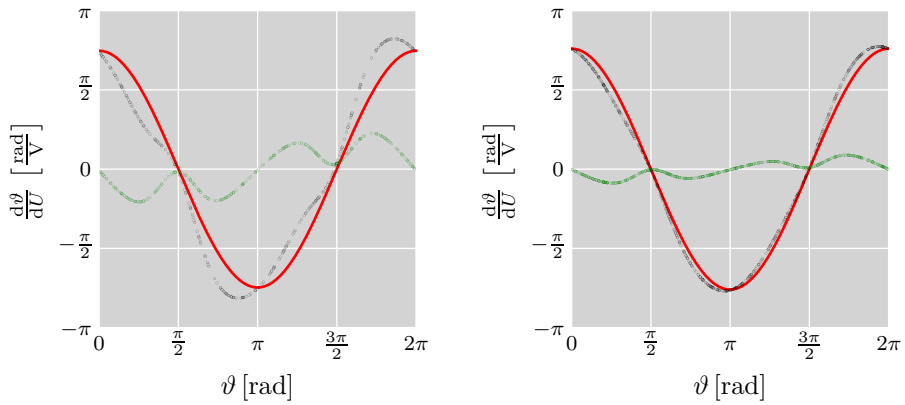
- The NIC-circuit in the nonlinear element is treated as ideal ohmic resistance providing inverted current $I_{neg} = -U/R_{neg}$.
- Both diodes of the nonlinear element are modeled with the Shockley diode equation $I_D(U) = I_S\left(\mathrm{e}^{U/U_T} - 1\right)$ [60]. Parameters $I_S = 8{,}235\,\mathrm{nA}$ and $U_T = 52{,}39\,\mathrm{mV}$ for the $1N4148$ diode were taken from measurements in [45].
- Resistors R_{comp}, R_{ext} and R_{coupl} were ideal ohmic resistances.

$$I = I_L + I_D(U) - I_D(-U) + U \cdot \left(\frac{1}{R_{coupl}} + \frac{1}{R_{comp}} + \frac{1}{R_{ext}} - \frac{1}{R_{neg}}\right) - \frac{U_{ext}}{R_{coupl}}$$

Simulated voltages, currents and times were tracked in mV, nA and ms. I_L and U are integrated with a simple Euler algorithm with a timestep $\mathrm{d}t = 10^{-7}$.

Exemplary simulation results for $C_{tot} = 1{,}0\,\mathrm{nF}$ and a cosine fit are shown in Fig. 5.1:

- Phase response curves (black dots) have roughly the shape of a cosine as expected.
- Deviations from the cosine increase with inductance L. As oscillation waveforms start to deviate from being purely sinusoidal at large $\frac{L}{R'_{neg}}$ [45], accompanying deviations in the PRC are not surprising.
- Amplitudes of the fitted cosines are roughly constant at $A_{PRC} = 2{,}4\,\frac{\mathrm{rad}}{\mathrm{V}}$, but deviate and fluctuate at large L, where the cosine does not fit well. (Fig. 5.2a)
- Zeros of the PRC are not shifted, but are always at $\vartheta = \frac{\pi}{2}n$.
- Deviations from the cosine fit(green dots) are rotational symmetric to π. Additionally, they are negative for $\vartheta \leq \pi$ and positive for $\vartheta \geq \pi$.



(a) $L = 41{,}0\,\mathrm{mH}$, $f = 24{,}7\,\mathrm{kHz}$ (b) $L = 5{,}17\,\mathrm{mH}$, $f = 69{,}9\,\mathrm{kHz}$

Figure 5.1: Exemplary simulations of phase response curves with $C_{tot} = 1{,}0\,\mathrm{nF}$ and a cosine fit. Simulation results are plotted as black dots. The fit $A_{PRC}\cos\vartheta$ is shown as red line. Green dots show the difference between the fit and the simulated data.

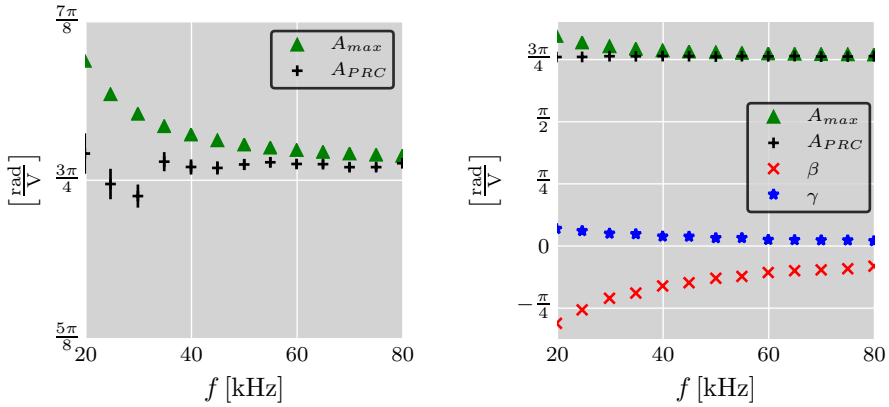
- Integrating deviations over either $\left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$ (where the cosine is negative) or $\left[\frac{3\pi}{2}, \frac{\pi}{2}\right]$ ($\cos\vartheta \geq 0$) yields 0 due to the rotational symmetry. Therefore, we can expect that the deviations do not influence dynamics much and we can choose the prefactor of the cosine as the correct amplitude for the calculation of the coupling strength.

Unfortunately, using a cosine-fit for the measurement data is difficult: As delays will introduce a shift in the ϑ-axis (See Sec. 5.1.2), the fit function needs to fit the shift as parameter as well. The fits of the shifted cosine, however, align with the main peaks and ignore the invariance of zeros of the PRC.

Consequently, we use a fit function $g(\vartheta)$ that correctly describes the cosine component and preserves the position of zeros in the PRC while being capable to correctly fit any shift s induced by the measurements:

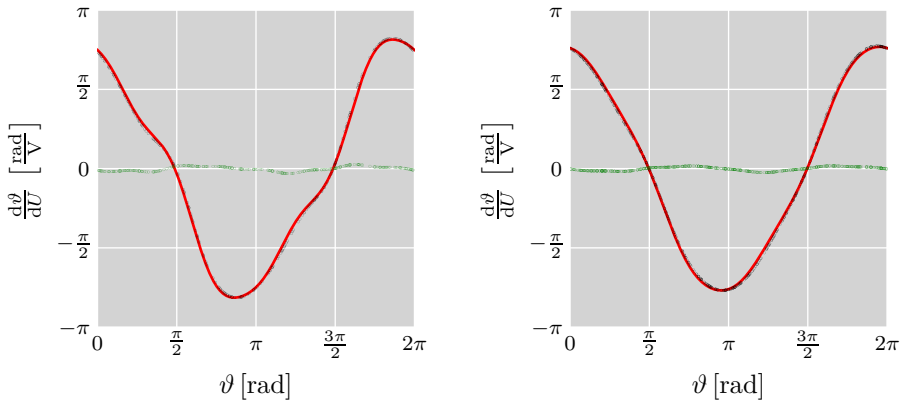
$$g(\vartheta) = \cos(\vartheta - s)\Big[A_{PRC} + \beta\sin(2[\vartheta - s]) + \gamma\sin(4[\vartheta - s])\Big] \tag{5.5}$$

$\cos\vartheta \cdot \sin 2\vartheta$ has zeros at $n\frac{\pi}{2}$ and concurs in both sign and rough peak shape with the remainder of the cosine fit. While $\cos\vartheta \cdot \sin 4\vartheta$ also preserves the position of the zeros, it can fit slight asymmetries of deviation peaks. Exemplary fits of the measurement data are shown in Fig. 5.3 and show excellent agreement. Additionally, cosine components are reliably fitted at $A_{PRC} = 2{,}4\,\frac{\mathrm{rad}}{\mathrm{s}}$ as seen in Fig. 5.2b with significantly less variance than the amplitude of the pure cosine fit. Note, however, that $g(\vartheta)$ was chosen based on observed features of the simulated PRCs and has no theoretical motivation known to us.



(a) Cosine fit (b) Heuristic fit

Figure 5.2: Fit parameters and the maximum value A_{max} of simulated phase response curves with $C_{tot} = 1{,}0\,\text{nF}$ are shown for a cosine fit (a) as well as the heuristic fit function Eq. (5.5) (b). Square roots of the covariance matrix are shown as error bars if they are larger than symbol size. Shifts $s \in [-\frac{1}{50}\pi, 0]$ have different units and are thus omitted in Fig. 5.2b.



(a) $L = 41{,}0\,\text{mH}$, $f = 24{,}7\,\text{kHz}$ (b) $L = 5{,}17\,\text{mH}$, $f = 69{,}9\,\text{kHz}$

Figure 5.3: Exemplary simulations of phase response curves with $C_{tot} = 1{,}0\,\text{nF}$ and a heuristic fit. Simulation results are plotted as black dots. The fit $g(\vartheta) = \cos(\vartheta - s)\Big[A_{PRC} + \beta\sin(2[\vartheta - s]) + \gamma\sin(4[\vartheta - s])\Big]$ is shown as red line. Green dots show the difference between the fit and the simulated data.

5.1.2 PRC measurement results

Oscillator voltages U were measured with the Spectrum M2i.4032 AD-conversion card at 50 $\frac{\mathrm{MS}}{\mathrm{s}}$. U_{ext} was generated with a a Spectrum M2i.6021 DA-conversion card at the same sampling rate. Both cards were controlled from a software written in Qt/C++. Phase response curves were then derived as described in Sec. 5.1. Additional considerations avoid systematic measurement uncertainties and automate measurements:

- Reference oscillators were measured while being plugged into the oscillator boards. Coupling resistors R_{ext} were connected in place of the resistors for initialization coupling and U_{ext} was connected directly in front of R_{ext}. Compensation resistors were placed as for normal recognition; resistors for recognition coupling and grounding resistors on the coupling line were removed. The measurements were conducted without closing the Faraday cage, but the annealed oscillator was removed and its output voltage grounded. (Preliminary measurements had shown no difference between the PRCs of reference and annealed oscillators.)

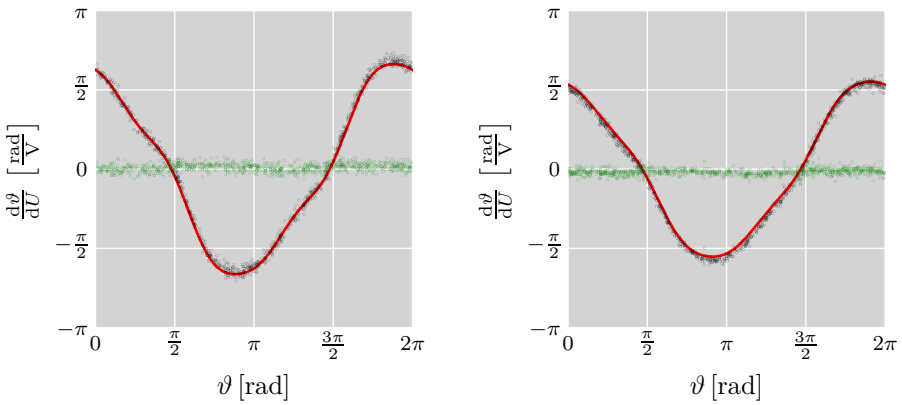
- In order to check the generation of U_{ext}, it was measured with the M2i.4032 cards as well. As both U and U_{ext} are measured with the same card, delay between both signals is minimal and times of the perturbations were extracted from the measurement of U_{ext} instead of relying on identical time zero points for M2i.4032 and M2i.6021. Local maxima $U_{ext,i}$ in U_{ext} were detected as peaks if $U_{ext,i} > 0.9A_p$ and $U_{ext,i-1} < U_{ext,i} > U_{ext,i+1}$. Note that in the presence of noise, this simple algorithm may detect a single peak multiple times. However, all but the first peak between two zero crossings are ignored by the subsequent computations. As the Gaussian peaks are very narrow, any deviations are negligible.

- Capacitance and serial resistance of connector and cable change the oscillator's dynamics if connected directly to U. Consequently, U was measured after the non-inverting amplifier, where the oscillator's dynamic cannot be perturbed. However, the non-inverting amplifier introduces a small delay which has to be considered in the fit of the PRC.

First, phase response curves were derived under the assumption $C_{tot} = 1{,}0\,\mathrm{nF}$. Exemplary measurements including fits to Eq. (5.5) are shown in Fig. 5.4, while fitted parameters are shown in Fig 5.5a. While single PRCs match the simulations qualitatively, fit parameters deviate significantly:

- Shifts s are non-zero as expected, as U is slightly delayed due to the non-inverting amplifier.

- Amplitudes A_{PRC} are significantly smaller than simulations predict.



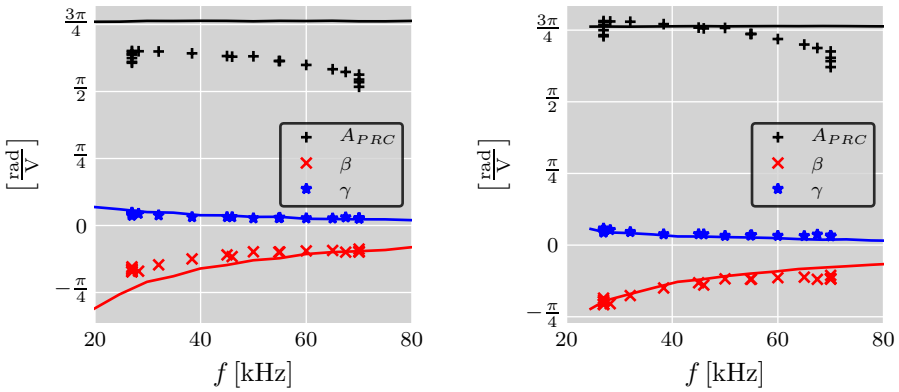
(a) $f = 27\,\mathrm{kHz}$ (b) $f = 70\,\mathrm{kHz}$

Figure 5.4: Exemplary measurements of phase response curves derived with the assumption $C_{tot} = 1{,}0\,\mathrm{nF}$. Measurements are plotted as black dots. The fit $g(\vartheta) = \cos(\vartheta - s)\Big[A_{PRC} + \beta \sin(2[\vartheta - s]) + \gamma \sin(4[\vartheta - s])\Big]$ is shown as red line. Green dots show the difference between the fit and the measured data.

- Amplitudes A_{PRC} are roughly constant at low frequencies, but decrease at high frequencies.

However, deviations have to be expected: Although we soldered a capacitance $C = 1{,}0\,\mathrm{nF}$ into the circuit boards, Appendix D suggests the existence of a significant parasitic capacitance on soldered stripboards. If the capacitance C_{tot} is underestimated, $Z(\vartheta) \propto C_{tot}$ is underestimated by the same factor. As A_{PRC} is roughly constant at low f and average deviations from simulations are roughly a factor 1.2 at $f = 27\,\mathrm{kHz}$, it seems likely that a parasitic capacitance $C_{par} = 0{,}2\,\mathrm{nF}$ exists in parallel to C. Consequently, we evaluated the measurements for $C_{tot} = 1{,}2\,\mathrm{nF}$ and compared them with new simulations with the same capacitance in Fig. 5.5b and Fig. 5.6:

- At low frequencies, fit parameters A_{PRC}, β and γ do now match the fit parameters of the simulations.
- At high frequencies, deviations occur in all parameters: The measured A_{PRC} is still lower than the value predicted by simulations. Parameters β and γ, however, are larger than expected, so PRCs deviate more strongly from the cosine at large f.
- Shifts of the simulated PRCs are smaller than $\frac{\pi}{90}$ and thus agree with our assumption of zeros at $\vartheta = \frac{\pi}{2}$ and $\frac{3\pi}{2}$.



(a) $C_{tot} = 1{,}0\,\text{nF}$ (b) $C_{tot} = 1{,}2\,\text{nF}$

Figure 5.5: Fit parameters A_{PRC}, β and γ for the heuristic fit function Eq. (5.5). Measured phase response curves (symbols) are compared to simulated PRCs (lines). Measurements were conducted for different oscillators at 27 kHz and 70 kHz, while measurements of the whole frequency range were conducted with a single oscillator only.

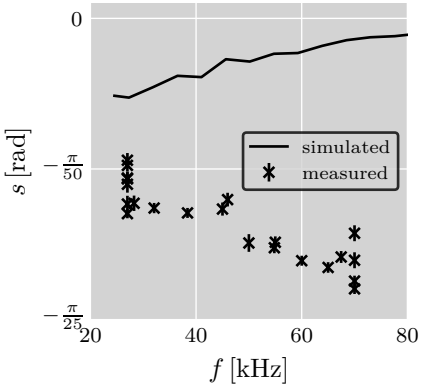


Figure 5.6: Fit parameters s for the heuristic fit function Eq. (5.5) are shown for measured PRCs (symbols) and simulated PRCs (lines) at $C_{tot} = 1{,}2\,\text{nF}$. Measurements were conducted for different oscillators at 27 kHz and 70 kHz, while measurements of the whole frequency range were conducted with a single oscillator only.

- As expected, shifts of measurements and simulations differ, because measurements are additionally delayed by the non-inverting amplifier.

The good agreement of fit parameters at low frequencies confirms, that the deviations can indeed be attributed to a parasitic capacitance. The origin of the deviations at high frequencies is currently unknown, but might be due to non-ideal behavior of op-amps in either the gyrator or the negative resistance. In any case, further research is required to fully explain the measured PRCs.

The derived PRCs differ between different oscillators or frequencies, thus coupling strengths ϵ are not the same as well. We will consider these deviations by assigning uncertainties to both C_{tot} and A_{PRC}:

As seen in Fig. 5.5b, PRCs deviate between oscillators of the same frequency. Specifically, measured A_{PRC}s at both $f = 27\,\mathrm{kHz}$ and $f = 70\,\mathrm{kHz}$ deviate at most 4% from the average over all oscillators. As these relative deviations are the same for both frequencies, we can attribute them to deviations of the total capacitance, which suggests $\Delta C_{tot} \approx 0{,}05\,\mathrm{nF}$.

Average values of A_{PRC} are in the range $\left[2{,}37\,\frac{\mathrm{rad}}{\mathrm{V}}, 2{,}04\,\frac{\mathrm{rad}}{\mathrm{V}}\right]$, decreasing from 27 kHz to 70 kHz. For the deviation of coupling strengths ϵ, we use the mean $A_{PRC} = 2{,}2\,\frac{\mathrm{rad}}{\mathrm{V}}$ of both extrema. Then, deviations amount to $\Delta A_{PRC} = 0{,}2\,\frac{\mathrm{rad}}{\mathrm{V}}$.

Still, several questions remain unsolved: (For difficulties in previous measurements, see [61].)

- Both ΔC_{tot} and ΔA_{PRC} are systematic uncertainties. C_{tot} could be derived for each oscillator independently. After additionally uncovering the frequency dependence of A_{PRC}, future implementations could compensate deviations in both variables.

- Although a significant parasitic capacitance certainly exists, other, minor effects might additionally affect A_{PRC}. As the phase response measurement alone is not sufficient to separate the influence of the capacitance from other error sources - e.g. in the coupling resistance - it is desirable to measure C_{tot} independently.

The amplitude of $Z(\vartheta)$ and the total capacitance C_{tot} affect the coupling strength ϵ. As the coupling strength determines recognition times and deviations from the ideal dynamics, it seems advisable to verify our values for ϵ. In addition, low frequency differences Δf_i of the oscillator pairs after annealing are essential for the performance of our network, but are difficult to calculate. Fortunately, we can obtain both ϵ and the Δf_i from direct measurements of the recognition dynamics.

5.2 Network preparation for recognition experiments

Before any recognition measurements can be conducted, we must ensure that parameters of our MONACO prototype - e.g. the amplitude A or the frequencies f_i - are adjusted and remain stable. Here, we describe the exact procedure for measurements of MONACO I: (For MONACO II, only preliminary measurements with non-ideal oscillators have been conducted so far and the procedure could not be followed exactly.)

1. Coupling resistors R_{recog} and R_{init} as well as compensation resistors R_{comp} cannot be changed while the network is operating, as they are positioned inside the Faraday cages. Consequently, they were changed before assembling the network.

2. After assembly, the power supply was connected and we waited at least ten minutes for thermal drifts to settle. In order to ease the following manual adjustment of frequencies f_i, the feedback loop of the automatic frequency adjustment circuit was activated (The network was set to the "Annealing Step"; Sec. 4.7.).

Frequencies f_i and amplitudes A can be adjusted during operation via potentiometers, which are accessible through holes in the Faraday cages that surround the oscillator submodules. Here, the first potentiometer is part of the gyrator, which changes the inductance and thus the frequency of the oscillator (Sec. 4.1.3). The second potentiometer is part of a non-inverting amplifier, thus varying the amplification factor for the oscillator amplitude A. Specifically, we followed the following procedure for each oscillator pair:

1. A Teledyne Lecroy HDO6034 oscilloscope is connected to the oscillator voltages. The oscilloscope is first configured to display frequencies of the oscillators and average them over time for greater accuracy.

2. The frequency of the *reference oscillator* is adjusted first according to Table 5.1. The frequencies in Table 5.1 fulfill the necessary frequency conditions and

pixel	1	2	3	4	5	6	7	8
f_i for MONACO I [kHz]	27.00	28.26	32.06	38.38	45.97	54.82	67.47	70.00
f_i for MONACO II [kHz]	27.00	30.31	36.92	40.23	56.77	60.08	66.69	70.00

Table 5.1: Oscillator frequencies used in the experiment for both architecture variants MONACO I and MONACO II. The f_i fulfill the frequency conditions Eq. (A.1) for MONACO I and Eq. (3.26) for MONACO II.

can be handled by the Van der Pol oscillators we implemented. The frequency accuracy $|f_i^{ideal} - f_i| \approx 0{,}05\,\mathrm{kHz}$ is mainly limited by the friction inside the potentiometer, which then limits the minimal angle it can be turned.

3. The frequency of the *annealed oscillator* is adjusted afterwards. Although the automatic frequency annealing readjusts this oscillator before each measurement, the manual adjustment can maximize available annealing range: As the frequency annealing feedback is active, frequency and phase difference $\Delta\vartheta_i$ lock as soon as the intrinsic frequency of the annealed oscillator is inside the annealing range. However, the frequency was tuned further, till $\Delta\vartheta_i \approx \frac{\pi}{2}$. At this point, $\cos\Delta\vartheta_i \approx 0$, centered between its extreme values. Future annealing steps - necessary due to e.g. thermal drifts - have thus an optimally symmetric annealing range.

4. Finally, the oscilloscope is configured to average over the oscillators' amplitudes, which are adjusted with the potentiometers up to $\Delta A \approx 0{,}01\,\mathrm{V}$.

Drifts shift both the oscillator amplitudes A and the frequencies of both reference and annealed oscillators. For measurements of MONACO I, parameters were checked after each measurement set and readjusted if necessary. In total, amplitudes did never deviate more than $\Delta A = 0{,}05\,\mathrm{V}$, frequency of reference oscillators was exact up to $|f_i^{ideal} - f_i| \approx 0{,}2\,\mathrm{kHz}$.

Preliminary measurements for MONACO II were conducted with a previous generation of oscillators, whose amplitudes could not be adjusted for the assembled network. Consequently, deviations are much higher.

5.3 Frequency deviations and coupling strength

The trajectory of a pixel depends on the coupling strength ϵ, the memorized patterns $\boldsymbol{\alpha}^m$ and the frequency differences Δf_i that remain after the frequency annealing process. From another point of view, both ϵ and Δf_i can be obtained from the slope of such a trajectory if we carefully select the memorized patterns.

In order to acquire precise phases, we measure the oscillator signals during recognitions with the M2i.4032 AD-card:

5.3.1 Mapping measured waveforms to phases

The M2i.4032 card returns times τ and voltages U_τ as integers with the τ being indices of the 16-bit integer array describing the U_τ. Units of τ were 1 µs and the U_τ correspond to voltages $\frac{U_\tau}{2^{13}} \cdot 5\,\mathrm{V}$. The zero-crossings of the waveforms constitute good reference points for the derivations of phases:

1. All times t_j^0 when the waveform crosses 0 V with positive slope were recorded: If $U_{\tau-1} < 0$ and $U_\tau > 0$, we linearly approximate the intersection of the t-axis (intercept theorem):

$$\begin{aligned} t_j^0 &= \tau - \frac{\tau - (\tau - 1)}{U_\tau - U_{\tau-1}} U_\tau \\ &= \tau - \frac{U_\tau}{U_\tau - U_{\tau-1}} \end{aligned}$$

2. Multiple crossings of the t-axis due to noise were excluded: First, the period of the oscillator was roughly approximated as $\tilde{T} = \frac{t^0_{last} - t^0_{first}}{\text{number of } t_j^0}$. Then, each t_j^0 was considered too close and excluded if $t_j - t_{j-1} < 5‰\tilde{T}$.

3. Phases at times τ were derived by comparing with the last and next unique intersections of the t-axis:

$$\varphi(\tau) = 2\pi \frac{\tau - t^0_{before}}{t^0_{next} - t^0_{before}}$$

5.3.2 Dynamics of a single changing pixel

In Sect. 4.5 , we calculated the coupling strength ϵ from the circuitry components of the network. Here, we show that ϵ can be derived from the slope of trajectories, which allows to verify our formula for ϵ and independently measure frequency differences Δf_i between the oscillator pairs. The averaged dynamic equation including differences of angular frequencies was already presented as Eq. (3.21):

$$\Delta\dot{\vartheta}_i = \Delta\Omega_i - \frac{\epsilon}{N} \sin\Delta\vartheta_i \Big(\sum_{j=1}^{N} S_{ij} \cos\Delta\vartheta_j - \frac{M}{2} \cos\Delta\vartheta_i \Big) \tag{5.6}$$

We restrict our analysis to trajectories with the following properties:

- Only the i^{th} pixel is defective.
- All other pixels are stable and therefore will not switch.
- Noise is low compared to frequency differences of oscillator pairs.

In order to analyze the slope of the transitions, we must first modify the dynamic equation for $\alpha_i = \cos(\Delta\vartheta_i)$ (Eq. (3.13)) to include frequency differences. Afterwards, we simplify Eq. (3.13) for transitions of a single pixel.

Influence of frequency differences

First, we consider the influence of frequency differences $\Delta f_i = \Delta\Omega_i/2\pi$ on the $\Delta\vartheta_i$-dynamics: The system starts on a fixed point in the ideal case $\Delta\Omega_i = 0$, and consequently the system state will move into regions with $\Delta\dot{\vartheta}_i > 0$ or $\Delta\dot{\vartheta}_i < 0$ depending on the perturbation and stay in that domain. For $\Delta\Omega_i \neq 0$, that symmetry is broken: $\Delta\Omega_i$ is the main perturbation for low noise and consequently decides on the $\Delta\vartheta_i$-region the system moves into. For $\Delta\Omega_i > 0$, $\Delta\vartheta_i$ grows till the pixel is switched and vice-versa for $\Delta\Omega_i < 0$. Consequently, $\Delta\Omega_i$ **always accelerates a pixel switch!** We derive the contribution $\Delta\tilde{\Omega}_i$ to $\dot{\alpha}_i$ with the chain rule:

$$\begin{aligned}\Delta\tilde{\Omega}_i &= \Delta\Omega_i \cdot \sin(\Delta\vartheta_i) \\ &= |\Delta\Omega_i|\,|\sin(\Delta\vartheta_i)| \cdot \alpha_i^{m'} \\ &= |\Delta\Omega_i| \sqrt{1-\alpha_i^2} \cdot \alpha_i^{m'}\end{aligned}$$

Trajectory slope at the turning point

The dynamics of α_i with inclusion of the non-zero frequency difference can then be summarized as follows: (See also Sec. 3.3.1.)

$$\dot{\alpha}_i = |\Delta\Omega_i| \sqrt{1-\alpha_i^2} \cdot \alpha_i^{m'} + \frac{\epsilon}{N}(1-\alpha_i^2)\left(\sum_{m=1}^{M} \alpha_i^m \langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}\rangle - \frac{M}{2}\alpha_i\right) \tag{5.7}$$

As only the i^{th} pixel of the correct pattern $\boldsymbol{\alpha}^{m'}$ is defective, we can simplify the coupling term as well:

$$\begin{aligned}\alpha_i(t=0) = \alpha_i^d &= -\alpha_i^{m'} \\ \boldsymbol{\alpha} &= \boldsymbol{\alpha}^{m'} + (\alpha_i - \alpha_i^{m'}) \cdot \hat{e}_i \\ \langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}\rangle &= \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle + \alpha_i^m(\alpha_i - \alpha_i^{m'})\end{aligned}$$

In Eq. (5.7):

$$\begin{aligned}\dot{\alpha}_i &= \Delta\tilde{\Omega}_i + \frac{\epsilon}{N}(1-\alpha_i^2)\left(\sum_{m=1}^{M} \alpha_i^m\Big(\left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle + \alpha_i^m(\alpha_i - \alpha_i^{m'})\Big) - \frac{M}{2}\alpha_i\right) \\ &= \Delta\tilde{\Omega}_i + \frac{\epsilon}{N}(1-\alpha_i^2)\left(\alpha_i^{m'} N + \sum_{m\neq m'}^{M} \alpha_i^m \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle + \sum_{m=1}^{M}\left(\alpha_i - \alpha_i^{m'}\right) - \frac{M}{2}\alpha_i\right) \\ &= |\Delta\Omega_i| \sqrt{1-\alpha_i^2} \cdot \alpha_i^{m'} + \frac{\epsilon}{N}(1-\alpha_i^2)\left(\alpha_i^{m'}(N-M) + \frac{M}{2}\alpha_i + \sum_{m\neq m'}^{M} \alpha_i^m \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle\right)\end{aligned}$$

Halfway through the transition, the slope is approximately constant and depends on angular frequency difference $\Delta\Omega_i$, coupling strength ϵ and patterns $\boldsymbol{\alpha}^m$ only:

$$\dot{\alpha}_i(\alpha_i = 0) = \alpha_i^{m'} \left[|\Delta\Omega_i| + \frac{\epsilon}{N} \underbrace{\left(N - M + \sum_{m \neq m'}^{M} \alpha_i^{m'} \alpha_i^m \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle \right)}_{X} \right] \tag{5.8}$$

If we can select suitable patterns with only the i^{th} pixel switching, we can derive $|\Delta\Omega_i|$ and ϵ:

1. Transitions of a pixel are measured with different sets of patterns that correspond to different values of the "pattern factor" $X = N - M + \sum_{m \neq m'}^{M} \alpha_i^{m'} \alpha_i^m \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle$. As the Δf_i are randomly distributed, several measurements are necessary for each X.

2. A linear fit extracts the slopes $\dot{\alpha}_i(\alpha_i = 0)$ in each measurement.

3. The slopes are again sorted as a function of X, and fitted with $s(X) = |\Delta\Omega_i| + \frac{\epsilon}{N} \cdot X$.

Selection of suitable memorized patterns for single transitions

In order to get many different values of $\dot{\alpha}_i(0)$, X has to be varied, which is done by choosing suitable sets of memorized patterns. Additionally, all pixels of the correct memorized pattern as well all non-switching pixels $k \neq i$ of the defective pattern have to be stable.

The stability of $\boldsymbol{\alpha}^{m'}$ is known from Sec. 3.2.2:

$$\sum_{m \neq m'}^{M} \alpha_l^{m'} \alpha_l^m \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle > -\left(N - \frac{M}{2} \right) = -6.5 \quad \forall l$$

Additionally, all non-switching pixels $k \neq i$ of the defective pattern must be stable as well:

$$\begin{aligned} \lambda_k(\boldsymbol{\alpha}^d) &< 0 \\ -\frac{\epsilon}{N} \left(\sum_{m=1}^{M} \alpha_k^m \alpha_k^d \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^d \right\rangle - \frac{M}{2} \right) &< 0 \\ \sum_{m \neq m'}^{M} \alpha_k^m \alpha_k^{m'} \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle + N - 2 \underbrace{\sum_{m=1}^{M} \alpha_k^m \alpha_k^{m'} \alpha_i^m \alpha_i^{m'}}_{\leq M} - \frac{M}{2} &> 0 \end{aligned}$$

$$\sum_{m \neq m'}^{M} \alpha_k^m \alpha_k^{m'} \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle > -\left(N - \frac{5M}{2}\right)$$
$$\sum_{m \neq m'}^{M} \alpha_k^m \alpha_k^{m'} \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle > -0.5$$

This last condition is much more restrictive than the criterion for stable memorized patterns. Measurements were conducted with some sets of memorized patterns that fulfill the equation and correspond to $X \in [5, 21]$:

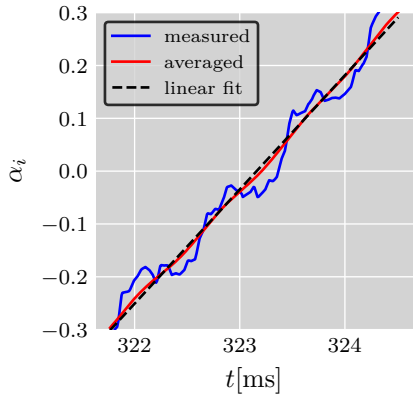
- 3 orthogonal patterns: $X = N - M = 5$.
- 3 identical patterns $\boldsymbol{\alpha}^{m'}$: $X = 3N - M = 21$
- 2 identical patterns $\boldsymbol{\alpha}^{m'}$ + 1 pattern with adjustable projection $\rightarrow X = 2N - M + \alpha_i^{m'} \alpha_i^m \left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle$. For simplicity, we choose $\alpha_i^{m'} \alpha_i^m = +1$. $\left\langle \boldsymbol{\alpha}^m, \boldsymbol{\alpha}^{m'} \right\rangle$ can then take values from -6 to +6.

Extracted coupling strengths and frequency differences

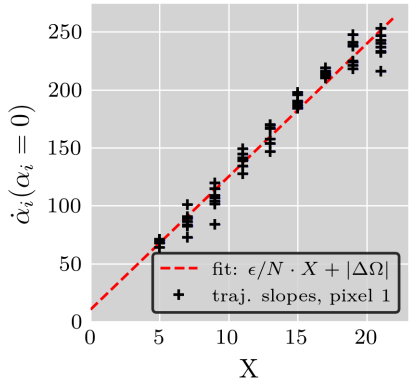
Measurements were performed at a calculated coupling strength $\epsilon = 94{,}4\,\frac{\text{rad}}{\text{s}}$ with one set of memorized patterns for each pixel and value of $X \in [5, 21]$. Several measurements were performed for each set of pixels and patterns, as the frequency differences Δf vary and some statistics is needed for a correct fit. The phases and α_i were extracted from measurements of the waveforms conducted with the Spectrum card M2i.4032 as described in Sec. 5.3.1.

In practice, higher order phase terms oscillate around the average dynamics described by Eq. (3.13). Consequently, these oscillations complicate useful fits of the $\dot{\alpha}_i(\alpha_i = 0)$ and should thus be removed. At the coupling strength $\epsilon = 94{,}4\,\frac{\text{rad}}{\text{s}}$, a moving average with a window of 2 ms removed most of the oscillations without decreasing the slope of the transition. The averaged trajectories were fitted with a straight line in the range $\alpha_i \in [-0.3, +0.3]$ to obtain the slope $\dot{\alpha}_i(\alpha_i = 0)$.(See Fig. 5.7a.) Note that the standard error of the fit parameters is negligible compared to the variance of $\Delta\Omega_i$. The $\dot{\alpha}_i(\alpha_i = 0)$ in turn were then fitted as a function of the "pattern factor" X, as shown in Fig. 5.7b, obtaining the slope $\frac{\epsilon}{N}$ and the offset $|\Delta\Omega_i|$. $|\Delta f_i| = \frac{1}{2\pi}|\Delta\Omega_i|$ and ϵ are shown in Fig. 5.7c and 5.7d.

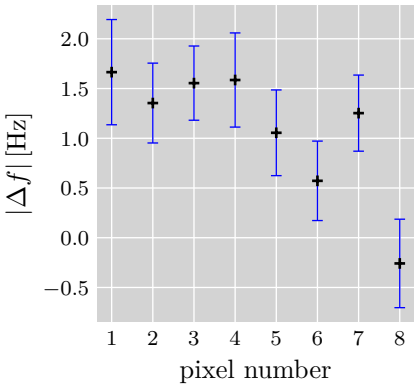
The Δf_i are small and confirm the correct operation of the frequency annealing circuitry. The measured coupling strengths ϵ, however, are strictly below the calculated value for the coupling strength. While the additional decline of ϵ at high frequencies reminds of the decline in A_{PRC} (Sec. 5.1.2), the decline is roughly two times larger than expected from A_{PRC} alone. One possible error source might be



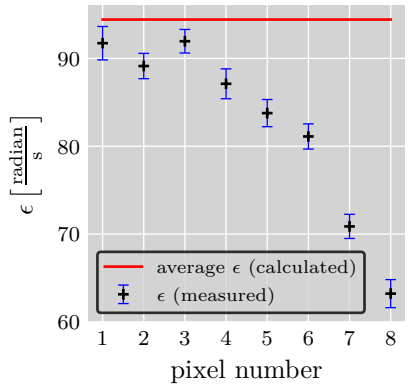
(a) **Exemplary fit of a trajectory** ($X = 27$, first pixel): The measured values of α_i show oscillations around the averaged dynamics. Removing these higher order terms with a running average over 2 ms allows for a precise linear fit of the slope in $\alpha_i \in [-0.3, +0.3]$, which depends on the coupling strength ϵ, the phase velocity $\Delta\Omega_i$ and the memorized patterns.



(b) **Exemplary fit of the slopes at $\alpha_i = 0$**: Slopes obtained as depicted in Fig. 5.7a are fitted as function of X, which in turn can be varied with the memorized patterns. The coefficients $\frac{\epsilon}{N}$ and $|\Delta\Omega_i|$ of the linear fit are important parameters of the network. Standard errors were derived as square-roots of the covariance matrix's diagonal elements.



(c) **Frequency differences**: $|\Delta f_i| = \frac{|\Delta\Omega_i|}{2\pi}$ are shown for each pixel. The low values prove the effectiveness of the frequency annealing circuitry.



(d) **Coupling strengths**: Measured coupling strengths ϵ are shown for each pixel and compared with the value expected from calculation. Deviations might originate in the oscillator amplitudes A, as $\epsilon \propto A^3$ is strongly influenced even by small ΔA.

Figure 5.7: Derivation of individual coupling strengths ϵ and frequency differences Δf_i from the slope of transitions. Fig. 5.7a and 5.7b present the analysis with examples. Results are shown in Fig. 5.7c and 5.7d.

the oscillator signal amplitude A. As it contributes to ϵ in the third power, even small ΔA change ϵ significantly. While the *maximum amplitude* was adjusted to the mentioned value $A = 1{,}4\,\mathrm{V}$, the *amplitude of the sinusoidal contribution* actually influences the other oscillators. Differences between these amplitude definitions could be large enough to explain the difference in scale. The decline of ϵ up to higher pixel numbers is difficult to analyze: As the deviations from an ideal sinus depend on frequency and components, each oscillator emits a different effective amplitude A. Consequently, each summand in the phase equation of the circuitry (Eq. (4.5)) has a different combination of amplitudes. Due to these differences between summands, the pattern pixels in each summand strongly influence the result, as they decide how these terms add up. Our measurements have been conducted with one set of memorized patterns for each pixel and value of X, which were not chosen randomly. Future measurements could improve our results by measuring different sets of memorized patterns for each X, which would randomize the effect of the patterns and return a more meaningful, average coupling strength for each oscillator.

In the end, the deviations in ϵ cannot be explained completely and remain a valid target for future research. The performance of the network, however, is high despite the deviations, as we will illustrate in Sec. 5.6. First, however, we will conclude our analysis of network components by verifying and calibrating readout voltages.

5.4 Calibration and verification of readout voltage measurements

Phases and $\alpha_i = \cos \Delta\vartheta_i$ can be derived from measurements of the oscillator voltages. (Sec. 5.3.1) However, high sampling rates and precision are necessary for this direct measurement, and future integrated oscillatory neural networks cannot afford this overhead. In contrast, readout voltages $U'_{\alpha,i}$ (See Sec. 4.1.2.) are created inside the network, contain all necessary phase information and can be read cheaply at low sampling rates. In our experiment, the $U'_{\alpha,i}$ are read with an Arduino Mega 2560 with 10 bit resolution. Voltage levels can be clearly separated and are viable for practical applications. (See Fig. 5.8b.) As we want to extract $\cos \Delta\vartheta_i$ precisely for our analysis, we add an additional calibration step:

The $U'_{\alpha,i}$ are supposed to be linear functions $a \cdot \cos \Delta\vartheta_i + c$, so system state variables $\alpha_i = \cos \Delta\vartheta_i$ are easy to obtain:

$$\cos \Delta\vartheta_i = \frac{U'_{\alpha,i} - c}{a} = s \cdot (U'_{\alpha,i} - c)\,, \tag{5.9}$$

In order to derive the α_i, the scaling factor s and the shift c need to be precisely known. However, these constants vary between pixels and drift with temperature, so recurrent recalibration is necessary for an exact mapping:

We measure two initializations with all negative and all positive pixels. As initialization coupling is strong enough that phase differences synchronize to phase differences 0 or π with negligible error, we can assume $\cos\Delta\vartheta_i = \pm 1$ and we read voltages $U'_{\alpha,i}(0)$ and $U'_{\alpha,i}(\pi)$ with the Arduino:

$$+1 = s(U'_{\alpha,i}(0) - c) \qquad -1 = s(U'_{\alpha,i}(\pi) - c)$$
$$s = \frac{2}{U'_{\alpha,i}(0) - U'_{\alpha,i}(\pi)} \qquad c = \frac{U'_{\alpha,i}(0) + U'_{\alpha,i}(\pi)}{2}$$

These *calibration measurements* were conducted before or after each set of recognition measurements. $U'_{\alpha,i}(0)$ and $U'_{\alpha,i}(\pi)$ were averaged over time to increase precision of s and c. Measurement data was then mapped according to Eq. (5.9). See Fig. 5.8 for an exemplary application of the calibration procedure.

In Fig. 5.9, α_i measurements with both Arduino and the AD-card Spectrum M2i.4032 are compared: On average, both curves align. The measurement with the AD-card, however, shows small regular oscillations in $\Delta\vartheta_i$ around the averaged trajectory, which correspond to higher order terms of the phase description. These cannot be seen in the Arduino data due to the Arduino's limited sampling rate. The measurements of the Arduino are nevertheless sufficiently precise to derive statistics about recognition success and recognition times, as is shown the next sections.

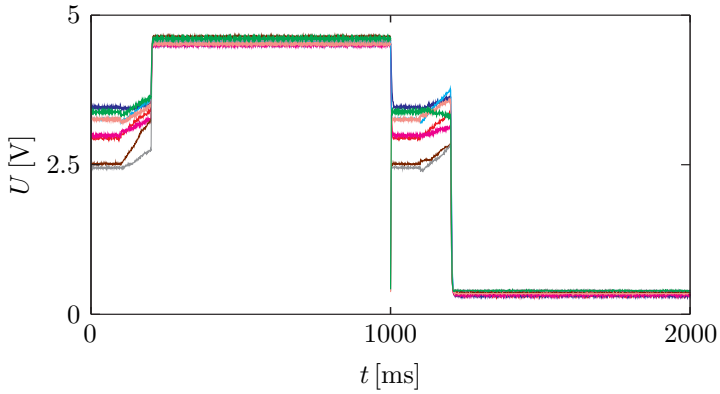
5.5 Removing noise from trajectories

Trajectories at high ϵ perform pattern recognition, but show a lot of delta-peak-shaped noise deviating from an ideal trajectory. (See e.g. Fig. 5.10 and 5.11.) In order to correctly extract statistics on recognition success and recognition times, trajectories were smoothed with a *median filter* [62, 63]: Smoothed trajectories $\tilde{\alpha}_i(t)$ are derived by applying a median to all values within a symmetrical time window $t_w = 30\,\text{ms}$ on the system state data $\boldsymbol{\alpha}$ after calibration (Sec. 5.4):

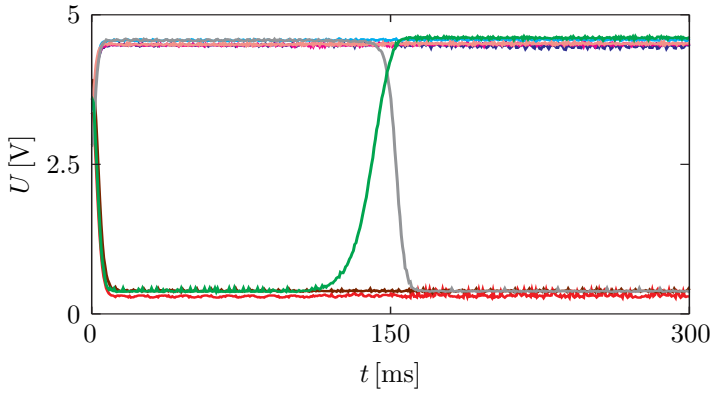
$$\tilde{\alpha}_i(t) = \text{median}([\alpha_i(t - 0.5t_w), \alpha_i(t + 0.5t_w)])$$

Smoothed data points were derived each $\Delta t = \frac{t_w}{300} = 100\,\mu\text{s}$. Median filtering was chosen over a running average for two reasons:

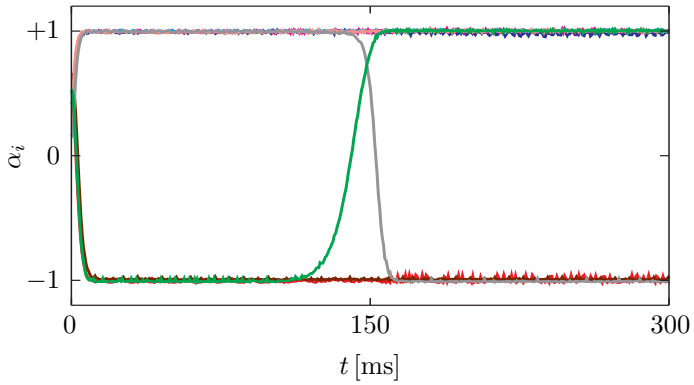
- Outliers are rejected. Peaks with a width $d < 0.5t_w$ are removed completely, while a running average would still be shifted considerably.
- Edges are preserved. Specifically, all monotone curves are invariant under median filtering [64]. As an ideal switching transition is monotone, the smoothed transition is not broadened and transition times can still be derived accurately



(a) **Exemplary calibration measurement**: Two initializations with positive / negative coupling map $\Delta\vartheta_i = 0$ / $\Delta\vartheta_i = \pi$ to corresponding voltage levels.



(b) **Uncalibrated measurement** (initialization and recognition): Voltage levels of different pixels do not align due to temperature dependencies and limited component precision.



(c) **Calibrated measurement**: The voltage data from Fig. 5.8b has been mapped to $\alpha_i = \cos\Delta\vartheta_i$ by using initialization voltage levels from Fig. 5.8a.

Figure 5.8: Calibration of Arduino readout voltage data. (Example at $\epsilon = 199\frac{\text{rad}}{\text{s}}$ and $M = 3$ orthogonal patterns.)

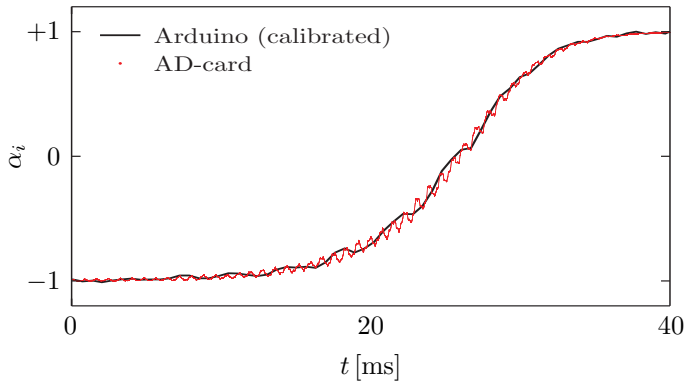


Figure 5.9: A defective pixel is shown during the transition to its correct value at a coupling strength $\epsilon = 199\frac{\text{rad}}{\text{s}}$. $\alpha_i = \cos(\Delta\vartheta_i)$ is compared for measurements with both the Spectrum AD-card M2i.4032 and the Arduino Mega 2560.
The AD-card measurement shows small oscillations around the averaged dynamics, which corresponds to higher order terms in the phase description. The measurement of the readout voltages with the Arduino Mega displays the averaged dynamics correctly, while the higher order phase terms cannot be resolved due the Arduino's limited sampling rate.

at high coupling strengths. (See Fig. 5.10 for an example of median filtering at high ϵ.)

We define the final system state $\boldsymbol{\alpha}^{\boldsymbol{final}}$ as the last value of the smoothed trajectories $\tilde{\boldsymbol{\alpha}}$ and use it as a reference for the derivation of recognition times T_{recog} (Sec. 5.7). Further, $\boldsymbol{\alpha}^{\boldsymbol{out}}$ with $\alpha_i^{out} = \operatorname{sgn}\alpha_i^{final}$ will be interpreted as the network's output for the derivation of recognition success (Sec. 5.6). In order to guarantee that the recognition is finished when $\boldsymbol{\alpha}^{\boldsymbol{final}}$ is extracted, we chose measurement times far larger than typical recognition durations. Histograms of measured recognition times are well inside the measured time frame despite their long tails, and measured final states match theoretical expectations. Note that any practical application of MONACO will probably use a similar procedure to detect the output pattern: Coupling strengths would be high for short recognition times and noise removal would thus be required. Additionally, the output pixels would be converted to digital signals for further usage.

5.6 Recognition success

After measuring properties of single oscillator pairs and validating the readout voltages in previous sections, we can finally assess the performance of MONACO I. (For MONACO II, only preliminary measurements exist yet, which are shown in

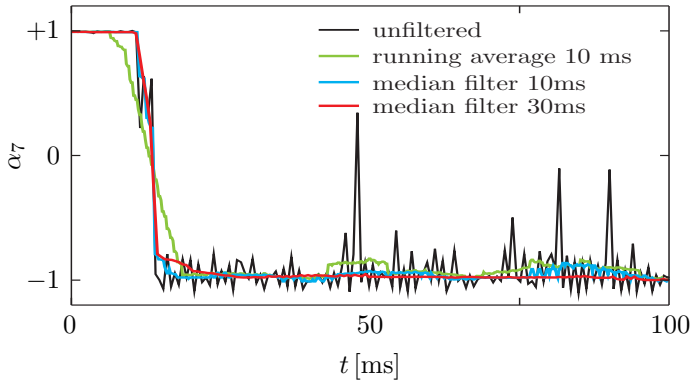


Figure 5.10: At high coupling strengths, noise consisting of sharp peaks dominates trajectories. This plot shows a single trajectory of a measurement at $\epsilon = 4730 \frac{\text{rad}}{\text{s}}$ with $M = 2$ random patterns. Applying a median filter with a symmetric time window $t_w = 10\,\text{ms}$ removes most noise, but does not shift or broaden the transition of the pixel. A median with a larger window $t_w = 30\,\text{ms}$ recovers an ideal transition curve. In contrast, a running average with $t_w = 10\,\text{ms}$ is strongly influenced by noise and modifies the transition shape.

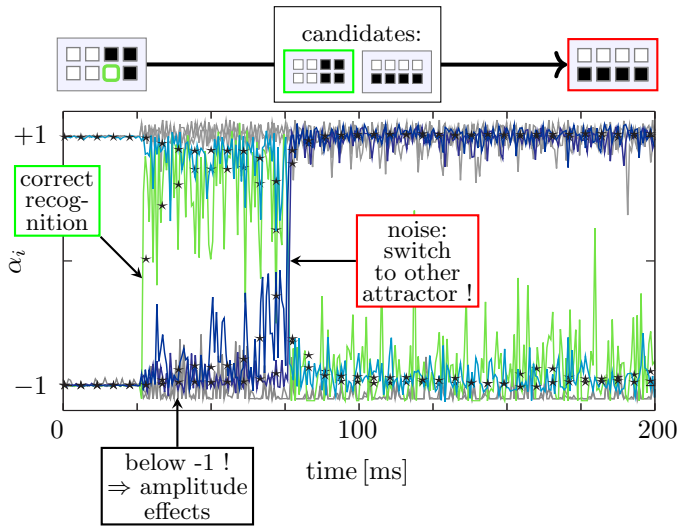


Figure 5.11: Example of a switching between attractors at very high coupling strengths $\epsilon = 4731 \frac{\text{rad}}{\text{s}}$ and $M = 2$ patterns. Only the green trajectory is supposed to switch for a correct recognition, which it first does successfully. However, noise induces a second switch at $t \approx 75\,\text{ms}$, when 4 pixels (blue and green trajectories; marked with stars) switch and change the correct output pattern to the other attractor erroneously. Additionally, readout voltages (and thus, calculated values for α_i) exceed the values from initialization. As the cosine is already at extreme values at initialization, the phase description is not sufficient to explain readout and dynamics during recognition and amplitude effects are relevant.

Appendix E.) In this section, we will focus on the success of the recognition process, meaning that the final pattern is the most similar to the erroneous input pattern. As discussed in Sec. 3.2.2, the network's potential attractors are the memorized patterns as well as their inverses and possibly unwanted, "spurious" attractors. We will now present different sets of measurements with either 2 or 3 orthogonal or random memorized patterns.

Let us first look at $M = 2$ orthogonal memorized patterns. The orthogonal patterns (00001111) and (00110011) and their inverses are always attractors (Sec. 3.2.2), and no spurious attractors were found. We conducted measurements for all $2^N = 256$ possible input patterns. If the final output pattern $\boldsymbol{\alpha}^{\boldsymbol{out}}$ was one of the memorized patterns and the erroneous input pattern $\boldsymbol{\alpha}^d$ was more similar to $\boldsymbol{\alpha}^{\boldsymbol{out}} = \boldsymbol{\alpha}^{m_1}$, i.e. $|\langle \boldsymbol{\alpha}^d, \boldsymbol{\alpha}^{m_1} \rangle| \geq |\langle \boldsymbol{\alpha}^d, \boldsymbol{\alpha}^{m_2} \rangle|$, we counted that recognition as successful. Note that for every attractive pattern, the inverse patterns $-\boldsymbol{\alpha}^{m_1}$ and $-\boldsymbol{\alpha}^{m_2}$ are also attractors (See Sec. 3.2.2), and are handled equal to the original memorized patterns.

In the whole range $\epsilon \in [48{,}3\,\frac{\text{rad}}{\text{s}}, 1608\,\frac{\text{rad}}{\text{s}}]$, the network successfully recognizes every input pattern in all measurements (See Table 5.2), even if they were not predicted to succeed by our *criterion for guaranteed recognition* (Eq. (3.19)). Qualitatively, these measurements are very similar to Fig. 5.8c.

At very high coupling strengths $\epsilon = 4731\,\frac{\text{rad}}{\text{s}}$, recognition success drops, but an upper limit of ϵ has to be expected: When we derived the ideal dynamics Eq. (3.5), we neglected higher order phase terms, amplitude effects and assumed that oscillators abide by the frequency conditions Eq. (A.1). At high coupling strengths ϵ, none of these assumptions remain valid and deviations from the ideal dynamics must exist. Inaccuracies further complicate an analysis of the dynamics. In total, several deviation origins have to be considered:

- The phase dynamics Eq. (3.5) determined with averaging are only valid at low ϵ. At high ϵ, *higher order phase terms* of the full, unaveraged phase equations (Eq. (3.1)) become large.

- At even higher ϵ, the phase of each oscillator is not sufficient to describe the dynamics of the oscillator: Perturbations of an oscillators internal dynamics cannot be considered independently anymore, so the *deviations of the oscillator's amplitude* have to be considered as well.

- Deviations from the frequency conditions Eq. (A.1) are not critical at low coupling strength, but might become important.

- Any electronic circuit carries some noise, which can have multiple origins, e.g. thermal motion of charge carriers or induced external signals. This noise is always present in the coupling voltage.

ϵ [$\frac{\text{rad}}{\text{s}}$]	success rate [%] (input: all patterns)	success rate [%] (input: theoretically guaranteed only)
48.3	100.0	100.0
94.4	100.0	100.0
198.6	100.0	100.0
482.5	100.0	100.0
778.2	100.0	100.0
1608.4	100.0	100.0
4730.5	89.5	75.0

Table 5.2: Success rates for 2 orthogonal patterns. For each coupling strength ϵ, 256 measurements were conducted with the memorized patterns (00001111) and (00110011). No spurious attractors were found. Recognitions were regarded as successful, if the output was the closest memorized pattern. 14 % of these measurements are predicted to succeed according to Eq. (3.19).

Due to multiple possible origins of deviations, a theoretical analysis seems daunting. First work has been done by Alexander Sparber in his bachelor's thesis [65]: He analyzed the full phase dynamics for $M = 1$ memorized pattern and perfect frequencies. At high ϵ, he found two competing effects: Oscillations around the averaged dynamics occur in $\Delta\vartheta_i$, increase in amplitude and distribute in Fourier space. Finally, phase slips occur due to these noisy oscillations, meaning that the phase differences jump by 2π. Another effect is the synchronization between different oscillator pairs: Frequencies of different pairs align, so frequency conditions are strongly violated. As coupling terms are then received from additional, unintended oscillators, the network's attractors are modified. Both effects depend heavily on the number of pixels, the memorized patterns and even initial conditions.

In our experiment, however, amplitude effects, electronic noise, frequency deviations and the existence of several memorized patterns further modify the dynamics. Consequently, we expect an upper limit for the coupling strength ϵ, but neither the critical value of ϵ nor the limiting effect can be deduced beforehand and might change with parameters.

Returning to our measurements with $M = 2$ orthogonal patterns, recognition success drops only at very high coupling strengths $\epsilon = 4731\,\frac{\text{rad}}{\text{s}}$: Here, the measurements predicted to succeed (Eq. (3.19)) are affected as well, which proves that Eq. (3.5) is not valid anymore.

Fig. 5.11 shows an example of noisy dynamics at very high coupling strengths. Specifically, large irregular deviations from the fixed points induce a switching between the two memorized patterns in this example, which invalidates the recognition process. Additionally, the readout voltages during recognition exceed values

ϵ $[\frac{\text{rad}}{\text{s}}]$	1st memorized pattern (or inverse)	2nd memorized pattern (or inverse)
48.3	110	146
94.4	107	149
198.6	123	133
482.5	157	99
778.2	159	97
1608.4	148	108
4730.5	214	42

Table 5.3: Distribution of output patterns for 2 orthogonal memorized patterns. For each coupling strength ϵ, 256 measurements were conducted with the memorized patterns (00001111) and (00110011).

from initialization, when the cosine was already extremized by phase differences 0 and π. Consequently, the phase description is not sufficient at this value of ϵ and amplitude deviations affect both readout and dynamics of the network.

Note that the Arduino cannot resolve the noise well enough for an accurate analysis. Future research will be required to separate different origins of deviations and estimate their magnitude.

Next, the distribution of output patterns should be analyzed. Note that for some erroneous input patterns $\boldsymbol{\alpha}^d$, both candidates $\boldsymbol{\alpha}^1$ and $\boldsymbol{\alpha}^2$ are correct output patterns: If the distances of $\boldsymbol{\alpha}^d$ to each output pattern are equal, meaning $\langle \boldsymbol{\alpha}^d, \boldsymbol{\alpha}^1 \rangle = \langle \boldsymbol{\alpha}^d, \boldsymbol{\alpha}^2 \rangle$, both output patterns are the closest to $\boldsymbol{\alpha}^d$. In an ideal network, this case would be decided by initial, random perturbations on $\boldsymbol{\alpha}^d$. In contrast, any real network has deviations in coupling strength, frequency or oscillator amplitudes, which influence the departure from ambiguous initial patterns and finally lead to an imbalance in pattern distribution.

The distribution of final patterns for the two orthogonal patterns at different coupling strengths is shown in Table 5.3. The counts for both memorized patterns (including their inverses) differ as expected, but are roughly equal. At very high coupling strengths $\epsilon = 4731 \frac{\text{rad}}{\text{s}}$, however, the first memorized pattern is significantly more frequent as final pattern than the second.

This change of stability can also be observed if we increase the number of orthogonal patterns to $M = 3$: We measured orthogonal memorized patterns (00001111), (00110011) and (01010101) and show the distribution of output patterns in Table 5.4. Similarly to $M = 2$, the distribution of memorized patterns shows some variance due to ambiguous input patterns, but each pattern is roughly equally frequent. At high coupling strength, stability changes again, but for $M = 3$, only one of the memorized patterns remains stable. This abnormal behavior might be due to synchronization of different oscillator pairs: Oscillators inside a group have differ-

ϵ $[\frac{\text{rad}}{\text{s}}]$	mem. pat. [%]	spur. pat. [%]	1st mem.	2nd mem.	3rd mem.	1st spur.	2nd spur.	3rd spur.	4th spur.
48.3	85.5	14.5	65	85	68	8	11	9	9
67.2	78.0	22.0	56	77	66	12	10	21	13
84.7	83.1	16.9	61	82	69	9	13	13	8
104.9	82.4	17.6	68	81	62	10	10	13	12
198.6	82.0	18.0	68	73	69	11	13	13	9
482.5	87.1	12.9	90	66	67	7	8	8	10
778.2	83.9	16.1	89	57	68	8	10	13	10
1608.4	91.0	9.0	110	66	56	6	10	0	7
4730.5	100.0	0.0	256	0	0	0	0	0	0

Table 5.4: Distribution of output patterns for 3 orthogonal patterns. For each coupling strength ϵ, 256 measurements were conducted with the memorized patterns (00001111), (00110011) and (01010101). Spurious attractors were (00010111), (01110001), (01001101), (11010100) and their inverses. Listed counts represent both a pattern and its inverse.

ϵ $[\frac{\text{rad}}{\text{s}}]$	success rate [%] (input: all patterns; correct: memorized)	success rate [%] (input: all patterns; correct: all attractors)	success rate [%] (input: theoretically guaranteed only)
48.3	85.5	100.0	100.0
67.2	78.0	100.0	100.0
84.7	83.1	100.0	100.0
104.9	82.4	100.0	100.0
198.6	82.0	99.6	100.0
482.5	87.1	100.0	100.0
778.2	83.9	100.0	100.0
1608.4	91.0	96.1	100.0
4730.5	67.2	57.0	33.3

Table 5.5: Success rates for 3 orthogonal patterns. For each coupling strength ϵ, 256 measurements were conducted with the memorized patterns (00001111), (00110011) and (01010101). Spurious attractors were (00010111), (01110001), (01001101), (11010100) and their inverses. Two definitions of recognition success are shown in the table: Either recognition was only counted as successful, if the output was the closest memorized pattern (or one of their inverses), or the output being closest attractor to the initial pattern was sufficient. 21 % of these measurements are predicted to succeed according to Eq. (3.19).

ent frequencies and only phase differences to the other group are supposed to be influenced by the coupling modulations. However, oscillators can synchronize to other frequency components in the coupling voltage intended for another oscillator, if the coupling is stronger than the frequency difference. As the minimal angular frequency difference $\Delta\Omega_{min} = 2\pi f_{min} = 2\pi \cdot 1{,}26\,\mathrm{kHz} = 7{,}92\,\frac{\mathrm{krad}}{\mathrm{s}}$ is exceeded by some coupling contributions with $\epsilon \cdot M \approx 14\,\frac{\mathrm{krad}}{\mathrm{s}}$, this seems viable. Concomitantly to the frequency adjustment of an oscillator, the coupling voltage components for its partner in the other group also change. At the same time, the partner oscillator and the neighboring oscillator pair see an increasingly narrower frequency difference, leading to a total synchronization of both pairs. Although it is not visible as prominently, these synchronization effects probably also influence the network with $M = 2$ orthogonal patterns, which sees an increase of stability for the first pattern as well.

In contrast to $M = 2$, however, several spurious attractors occur; namely (00010111), (01110001), (01001101), (11010100) and their inverses, which constitute a considerable percentage of final patterns. Looking closer at each pixel of the first spurious pattern $\boldsymbol{\alpha}^{sp_1}$, we see that $\alpha_i^{sp_1}$ concurs with the most frequent value of the i^{th} s pixels of the memorized patterns α_i^m. Explicitly, the first pixels of all 3 memorized patterns are 0, so $\alpha_1^{sp_1} = 0$ as well. The second pixels are 0,1 and 0, and $\alpha_2^{sp_1} = 0$ coincides with the most frequent value again. This is true for all pixels of $\boldsymbol{\alpha}^{sp_1}$.

Similarly, we gain the other spurious attractors by inverting some of the memorized patterns and then taking the most frequent pixel values for each index. These types of spurious attractors, called "symmetric mixture solutions" [66], also occur in other attractor networks, like the Hopfield network [24]. Specifically, these attractors are described by the *majority rule*:

$$\alpha_i^{sp} = \mathrm{sgn}\left(\sum_{m\in A} s^m \alpha_i^m\right) \tag{5.10}$$

Here, prefactors $s^m \in \{\pm 1\}$ denote that either a pattern or its inverse is used for the construction of a spurious attractor $\boldsymbol{\alpha}^{sp}$. Any subset of $A \subset \{1, ..., M\}$ can be used to construct a candidate for a spurious attractor, but stability is not guaranteed and needs to be verified with our stability criterion Eq. (3.10). In other networks, for example, using any even number of memorized patterns for construction results in unstable fixed points [66] and we could not detect any spurious patterns for $M = 2$ as well.

Unsurprisingly, spurious attractors influence recognition success as well. As seen in Table 5.5, the success rate for $M = 3$ orthogonal patterns is diminished, if we allow only memorized patterns (00001111), (00110011) and (01010101) and their inverses as possible correct candidates. If we, however, add spurious attractors as

ϵ [$\frac{\text{rad}}{\text{s}}$]	success rate [%] (input: all patterns; correct: memorized)	success rate [%] (input: theoretically guaranteed only)
48.3	98.1	99.1
198.6	99.4	100.0
482.5	98.9	100.0
778.2	98.2	100.0
1608.4	96.4	100.0
4730.5	65.8	57.5

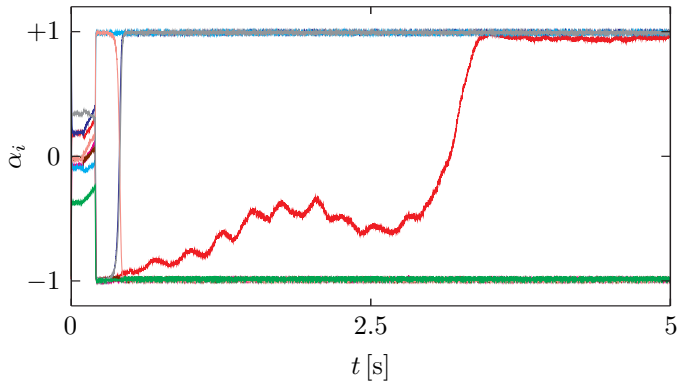
Table 5.6: Success rates for 2 random patterns. A list of 1000 sets of random memorized patterns with random input patterns was constructed and measured for each coupling strength ϵ. No spurious attractors exist for 2 patterns. Recognitions were regarded as successful, if the output was the closest memorized pattern. 11 % of these measurements are predicted to succeed according to Eq. (3.20).

acceptable final patterns and define success as $\boldsymbol{\alpha}^{\boldsymbol{out}}$ being the closest *attractor* to the erroneous input pattern, we regain our high success rates.

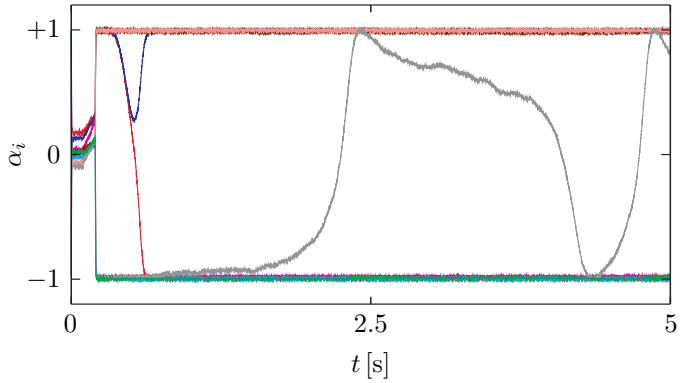
Depending on the application, the memorized patterns may not be orthogonal. Therefore, we discuss 2 random memorized patterns next: Pixels of the initial, erroneous pattern as well as the memorized patterns were selected with 50% probability for each value. $M = 2$ random patterns are not accompanied by spurious attractors, as suggested by the majority rule Eq. (5.10). Rates of recognition success are close to 100% again, as shown in Fig. 5.6, and drop at very high coupling strength. (Probably, reasons for the decline are the same as for $M = 2$ and $M = 3$ orthogonal patterns.) Additionally, a slight decline in the success rate of predicted recognitions occurs at very small coupling strengths $\epsilon = 48{,}3\,\frac{\text{rad}}{\text{s}}$. This hints at the lower limit of ϵ, where frequency differences Δf_i become relevant. (See Eq. (3.21).) At medium values of the coupling strength, the coupling term is much larger than differences of angular frequencies $\Delta\Omega_i = 2\pi\Delta f_i$ and the dynamics remain unaffected. If the coupling between oscillators in a pair gets close to $\Delta\Omega_i = 2\pi\Delta f_i$, stability of pixels will be diminished. This can lead to noise-induced transitions or even loss of synchronization in a pair, so phase slips occur. Exemplary measurements showing these phenomena are shown in Fig. 5.12, but they are still rare at this value of the coupling strength.

Finally, we measured $M = 3$ random patterns. All spurious attractors found are predicted by the majority rule Eq. (5.10). On the other hand, neither spurious nor memorized patterns need to be attractors (See e.g. Eq. (3.7): If two memorized patterns are similar, they can destabilize the third memorized pattern.).

Success statistics are summarized in Table 5.7. Obviously, $M = 3$ random patterns exceed the capacity of the small network: While success rates for memorized patterns



(a) Diffusive transition



(b) Phase slips

Figure 5.12: At low coupling strengths, the coupling sometimes cannot compensate frequency differences Δf_i and electronic noise. Exemplary measurements of a noise-induced transition and phase slips are shown in a) and b) for $\epsilon = 48{,}3\,\frac{\text{rad}}{\text{s}}$ and $M = 2$ random patterns. However, these measurements are still rare at $\epsilon = 48{,}3\,\frac{\text{rad}}{\text{s}}$ and recognitions are successful most of the time.

are at roughly 70 %, the success rates for all attractors are lower, and guaranteed recognitions are not all successful.

The decline in the rate of predicted success shows that some non-ideal deviation must contribute to the low success rates. As the rate is diminished for all values of ϵ, frequency differences Δf_i, higher order phase terms or amplitude effects can be excluded as main contribution: They are only relevant for especially low or high values of the coupling strength. Electronic noise scales with ϵ and is thus relevant for all coupling strengths. If electronic noise was the main non-ideality, however, measurements for $M = 2$ should probably be diminished more. Whatever the underlying cause, the capacity of our network implementation is exceeded for $M = 3$ memorized patterns with $N = 8$ pixels. Summing up, the network maps erroneous pattern onto the most similar attractor very reliably, if its capacity is not exceeded or

ϵ $[\frac{\text{rad}}{\text{s}}]$	success rate [%] (input: all patterns; correct: memorized)	success rate [%] (input: all patterns; correct: memorized + spurious cand.)	success rate [%] (input: theoretically guaranteed only)
48.3	73.7	60.1	90.6
94.4	74.9	62.1	90.6
198.6	76.4	66.0	96.2
482.5	72.9	61.1	93.8
778.2	73.3	60.5	93.8
1608.4	67.6	55.5	90.6

Table 5.7: Success rates for 3 random patterns. A list of 1000 sets of random memorized patterns with random input patterns was constructed and measured for each coupling strength ϵ. For $\epsilon = 199\,\frac{\text{rad}}{\epsilon}$, additional 2000 sets of random patterns were measured. Spurious patterns exist according to the majority rule, but spurious patterns and even memorized patterns can be unstable. On the other hand, any attractor found was either a memorized pattern or a spurious pattern. Two definitions of recognition success are shown in the table: Either recognition was only counted as successful, if the output was the closest memorized pattern, or the output being closest memorized or spurious pattern to the initial pattern was sufficient. Only 3 % of these measurements are predicted to succeed according to Eq. (3.20).

extreme values of the coupling strength are chosen. In any implementation, the time necessary for the recognition process is an important property as well. Recognition times T_{recog} are thus discussed in the next sections.

5.7 Recognition times

After a time T_{recog}, the system state will come to rest in an isolated minimum of the potential function and the recognition will be finished. This recognition time T_{recog}, however, is not constant and dependencies on ϵ, patterns and noise are unknown. The following algorithm was used to derive T_{recog} even for noisy measurements:

1. First, smoothed trajectories $\tilde{\boldsymbol{\alpha}}$ are derived with a median filter as described in Sec. 5.5.

2. We use the criterion

$$\tilde{\alpha}_i(t) \cdot \operatorname{sgn}\left(\alpha_i^{final}\right) \geq 0.5 \cdot |\alpha_i^{final}| \tag{5.11}$$

   to determine if the i^{th} pixel has switched. Here, $\boldsymbol{\alpha}^{\boldsymbol{final}}$ is the last smoothed system state of the measurement. In other words, the pixels α-value needs the right sign and be significantly different from 0. At low ϵ and low external noise, $|\alpha_i^{final}| = |\alpha_i^{out}| = 1$ and the criterion recognizes pixels whose switching process is 75% complete. This is not necessarily true for very high ϵ, however, as α_i^{final} is shifted due to amplitude effects and has a high uncertainty due to higher order terms and electronic noise, which is corrected by the smoothing procedure and the scaling to $|\alpha_i^{final}|$.

3. The recognition time T_{recog} is then defined as the time difference between the beginning of the recognition process and the time $t^{\{N\}}$ when all N pixels fulfill Eq. (5.11) the first time.

Note that pixels do not switch simultaneously every time: Often, a small eigenvalue might prolong the departure of a pixel from the input pattern compared to others (See e.g. Fig. 5.13) or some pixel might become unstable only during the other pixels' transitions. On the other hand, scalar products $\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}^m \rangle$ are the same for each pixel. If pixels have equal prefactors for the scalar products, they show very similar, basically simultaneous transitions. (Compare with the dynamic equation for the system state Eq. (3.13) and again Fig. 5.13.) This peculiar separation between different switching pixels often carries the system state $\boldsymbol{\alpha}$ close to unstable fixed points, where a description with eigenvalues is viable. These "intermediate patterns" can thus be used to infer from the eigenvalues of the intermediates on the dynamics of the system state. (This will be further discussed in Sec. 5.7.5.)

We keep track of intermediate patterns $\boldsymbol{\alpha}^{int}$ by deriving $\alpha_i^{int} = \operatorname{sgn} \tilde{\alpha}_i$ each time the switching criterion Eq. (5.11) is fulfilled for a pixel. If the intermediate changes by several pixels at evaluation of the switching criterion, this means pixels switching simultaneously and we count it as one joint switching transition.

In Fig. 5.13, detection of the last and the second last pixel switches is marked in an exemplary measurement. Recognition times can differ greatly and are spread

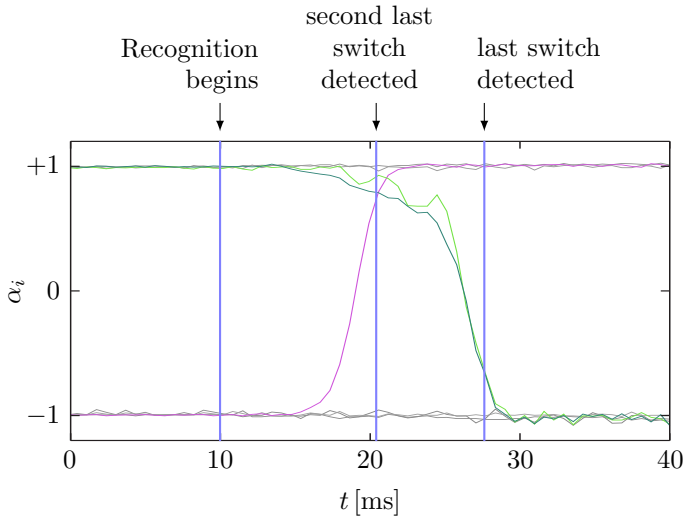


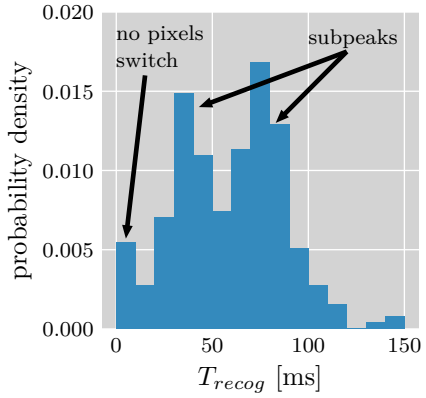
Figure 5.13: Detection of recognition times T_{recog}. Vertical lines mark important points in time as detected by the algorithm for recognition time detection (Sec. 5.7): The start of the recognition step, the end of the last, simultaneous switch of two pixels, and the second last pixel switch are shown. As the last switch ends the recognition process, the time difference between the start of the recognition step and the last switch equals the recognition time. The second last pixel switch is important for the derivation of intermediate patterns. The measurement was conducted at a coupling strength $\epsilon = 778 \frac{\text{rad}}{\text{s}}$ with $M = 3$ random memorized patterns.

over a large range of different values. First, the typical shape of recognition time distributions are discussed.

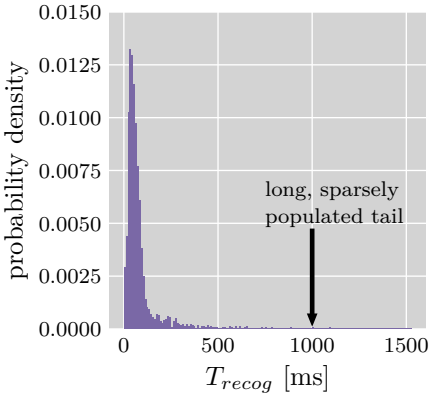
5.7.1 Distribution characteristics

Exemplary histograms of the distribution of the recognition times T_{recog} (Fig. 5.14 and 5.15) show characterizing features:

- Distributions span a large range of values.
- Distributions may have distinct subpeaks.
- Distributions may have a long, sparsely populated tail.
- Recognitions with an already correct initial pattern form a peak at $T_{recog} = 0\,\text{ms}$.

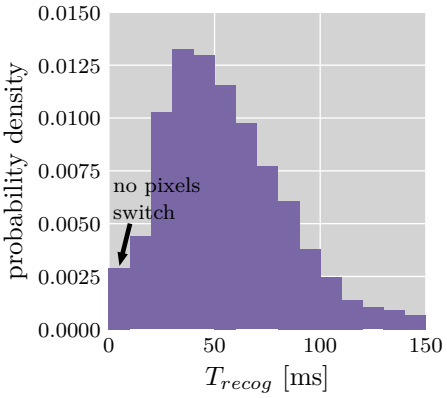


(a) Orthogonal output patterns, all 256 possible input patterns

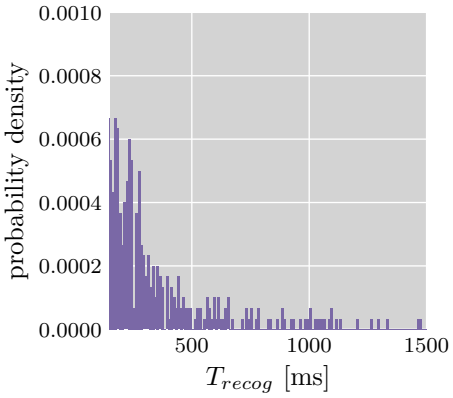


(b) Random input and output patterns, 3000 measurements

Figure 5.14: Characterizing features of the distributions of the recognition time T_{recog} for both orthogonal and random output patterns. ($M = 3$ output patterns, coupling strength $\epsilon = 199\,\frac{\text{rad}}{\text{s}}$. Bins have a size of 10 ms.)

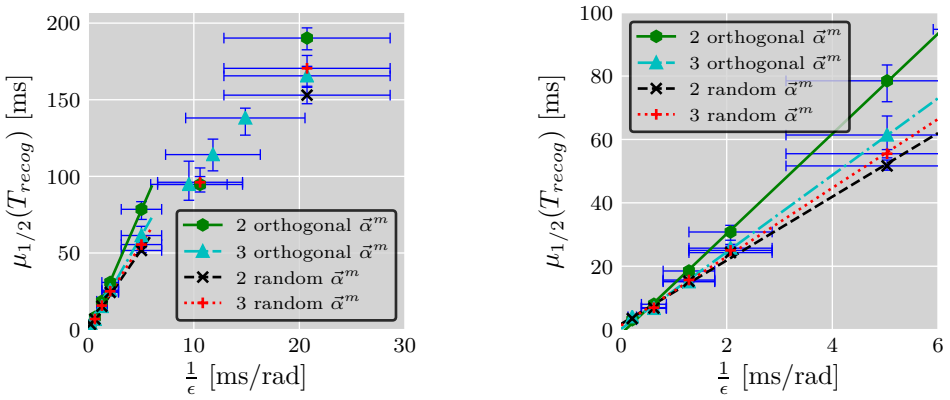


(a) Main peak



(b) Long, sparsely populated tail

Figure 5.15: Enlarged features of the distribution of the recognition time T_{recog} for 3 random output patterns from Fig. 5.14b. (Coupling strength $\epsilon = 199\,\frac{\text{rad}}{\text{s}}$. Bins have a size of 10 ms.)



(a) All measurements

(b) Zoom on linear region

Figure 5.16: Medians $\mu_{1/2}$ of the recognition times T_{recog} plotted against the inverse coupling strength $\frac{1}{\epsilon}$. Error bars of the median denote a confidence interval of 95%.

5.7.2 Typical recognition times

Typical recognition times are contained in the head of a distribution f, which can be characterized by measures of central tendency: The mean is skewed by large outliers of a distribution and the mode depends on the binning and might not be unique. Therefore, the median $\mu_{1/2}$ of the recognition time distributions is suited best as measure of typical recognition times. Additionally, confidence intervals of the median can be derived without assumptions about the distribution [67].

We expect recognition times T_{recog} to be inversely proportional to the coupling strength ϵ in an ideal system.(See Sec. 3.1.) Fig. 5.16 shows the medians $\mu_{1/2}$ of the recognition times for all combinations of 2 or 3 random or orthogonal output patterns $\boldsymbol{\alpha}^m$:

- $\mu_{1/2}(T_{recog})$ is approximately proportional to $\frac{1}{\epsilon}$ as predicted. Deviations occur at small ϵ only, where frequency differences $\Delta\Omega_i$ affect the dynamics more strongly. (See Sec. 3.6.1 and 5.8.)
- Medians are similar for orthogonal/random patterns and different M.

Most importantly, typical recognition times can be reduced by increasing the coupling strength ϵ.

5.7.3 Influence of the number of switching pixels

Inspired by the peak at $T_{recog} = 0$, which corresponds to already correct patterns, we examine the influence of the number of switching pixels n_{switch} on the recognition

times. n_{switch} is defined as the number of pixels, in which the final pattern and the initial defective pattern differ.

In Fig. 5.17 and 5.18, histograms for measurements with specific numbers of switching pixels n_{switch} are shown for $\epsilon = 199 \frac{\text{rad}}{\text{s}}$:

- Distributions shift to higher values of T_{recog} with growing n_{switch}. Consequently, the median increases as well, as shown in Fig. 5.19.
- Differences of n_{switch} partially explain the distribution substructure present for orthogonal output patterns: Contributions from $n_{switch} \in [1, 2]$ form a first peak , while the second peak is mainly consists of measurements with $n_{switch} \in [3, 4]$, but also some with $n_{switch} = 2$.
- The distributions for individual n_{switch}, however, show a substructure themselves for orthogonal output patterns. Ultimately, the origin of these substructures may stem from the stability of single pixels: If eigenvalues for several pixels are positive and equal, they switch together at roughly the same speed. In contrast, some pixels may only loose stability after some other pixels have already switched, which results in the second and possibly more peaks.
- For random output patterns, distributions are much broader and subpeaks are not distinctive. While the order and concurrence of pixel switches probably still affects recognition, there are much more possible absolute values of the eigenvalues of initial patterns, which broadens subpeaks.

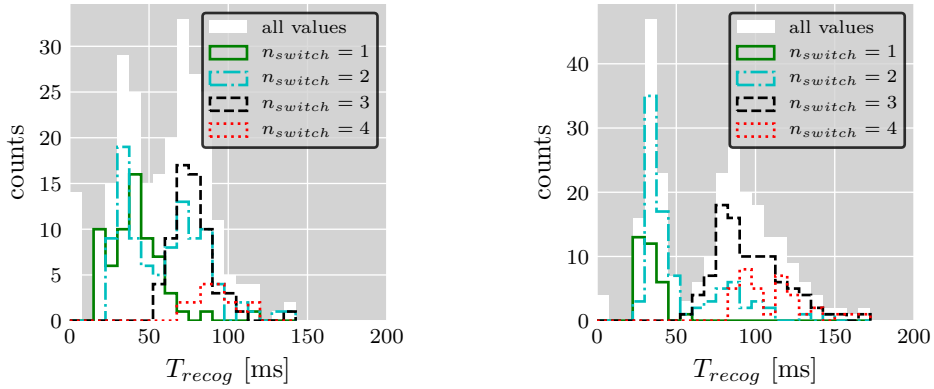
The role of eigenvalues of intermediate patterns will be more closely discussed in Sec. 5.7.5. First, tails of the distributions will be examined.

5.7.4 Rare events: Analysis of sparsely populated tails

In practical applications of MONACO, readout of the output pattern will likely be performed at a fixed point in time. However, recognition times show a long, sparsely populated tail for some sets of parameters. Then, readout will either cut the tail of the distribution - possibly reading a pattern before the recognition is finished - or the recognition process might take a long time. Consequently, analysis of the tails is essential.

Standard deviations of the distributions are shows in Fig. 5.20 for different coupling strengths ϵ. As the standard deviation is larger for distributions well spread out, long tails coincide with large standard deviation.

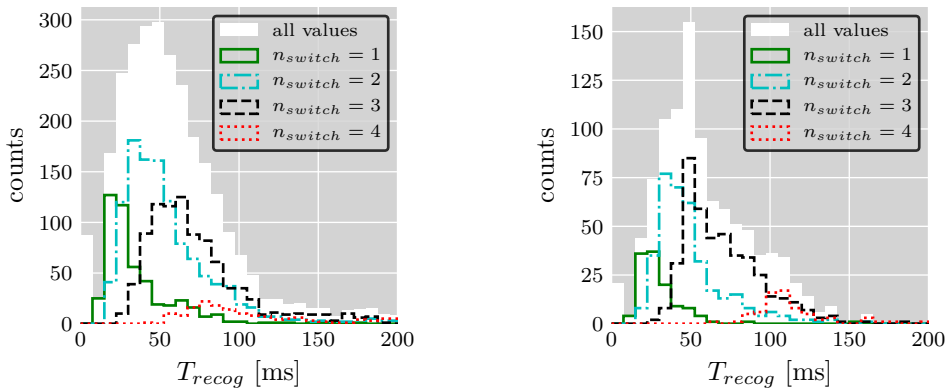
- σ is larger for random patterns than for orthogonal patterns. This reflects the fact that T_{recog} of random output patterns show a tail at all measured ϵ, while the recognition time distributions with orthogonal output patterns miss a tail for all but the highest coupling strengths. See Appendix F for plots of all measured distributions.



(a) 3 orthogonal patterns

(b) 2 orthogonal patterns

Figure 5.17: Histograms of the recognition times T_{recog} for different number of switching pixels n_{switch} and the combined histogram for all values of n_{switch}. ($\epsilon = 199\,\frac{\text{rad}}{\text{s}}$, orthogonal output patterns, bin size is 7,5 ms.)



(a) 3 random patterns

(b) 2 random patterns

Figure 5.18: Histograms of the recognition times T_{recog} for different number of switching pixels n_{switch} and the combined histogram for all values of n_{switch}. The sparsely populated tails are cut at $T_{recog} = 200$ ms. ($\epsilon = 199\,\frac{\text{rad}}{\text{s}}$, random output patterns, bin size is 7,5 ms.)

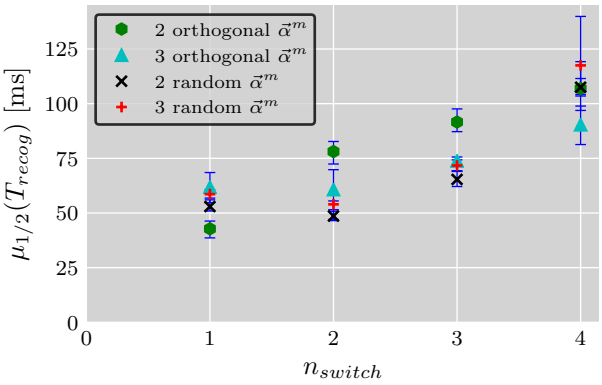
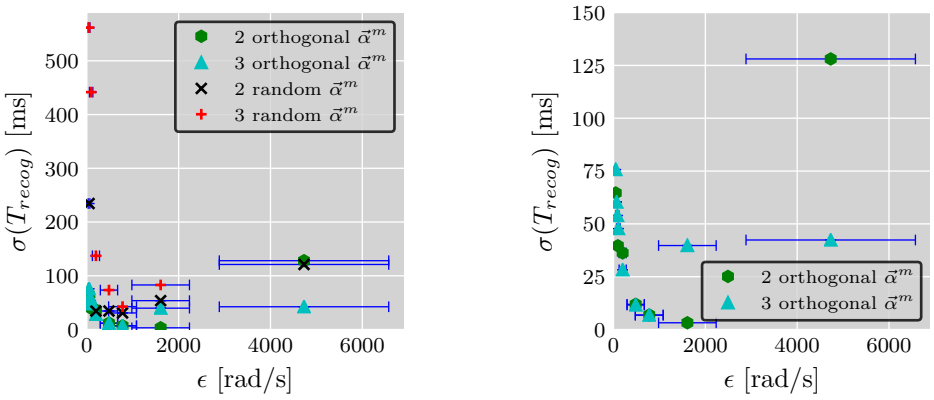


Figure 5.19: Median of the recognition times T_{recog} for a subsets of measurements characterized by the number of switching pixels n_{switch}. All measurements were conducted at $\epsilon = 199\,\frac{\text{rad}}{\text{s}}$. Error bars of the median denote a confidence interval of 95%.



(a) All measurements

(b) Zoom on orthogonal patterns

Figure 5.20: Standard deviation σ of the recognition times T_{recog} plotted against the coupling strength ϵ. (No error bars are shown for σ, as calculations of confidence intervals usually assume normal distributions and thus cannot be applied.)

- For orthogonal output patterns, σ decreases roughly proportional to $\frac{1}{\epsilon}$ for small till medium coupling strength. At high coupling strength, tails appear ($\epsilon \geq 1608\,\frac{\text{rad}}{\text{s}}$ for $M = 3$, $\epsilon \geq 4730\,\frac{\text{rad}}{\text{s}}$ for $M = 2$), which are accompanied by a sharp rise in σ and coincide with drops in recognition success shown in Sec. 5.6.

- For random output patterns, σ has a minimum at medium coupling strengths as well.

Most importantly, σ increases with ϵ only at very large coupling strengths. The increase in σ thus coincides with the decline of recognition success. Consequently, optimizing for both recognition success and diminished tails is possible without tradeoffs. From another point of view, the increase σ and the loss of stability might originate from the same deviations from the ideal dynamics. As illustrated in the next section, eigenvalues of intermediate patterns are connected to the distribution tails as well.

5.7.5 Long tails revisited: Diffusion over potential barriers

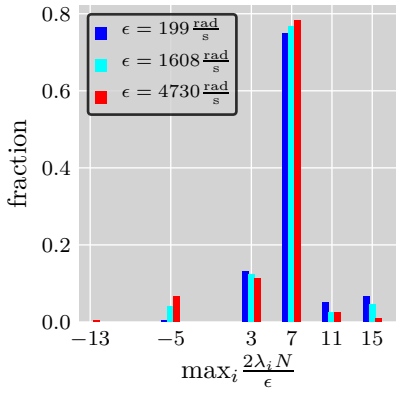
As proposed in Sec. 5.7.3, the eigenvalue spectrum of initial and intermediate patterns might define the distribution of recognition times T_{recog}.

Eigenvalues describe the linearized behavior around a fixed point, so a pixel should leave an unstable fixed point faster for a larger corresponding eigenvalue. As the deviation from the fixed point is in turn the slowest part of the switching process, recognition times should depend on the eigenvalues of the initial, defective pattern, which is an unstable fixed point.

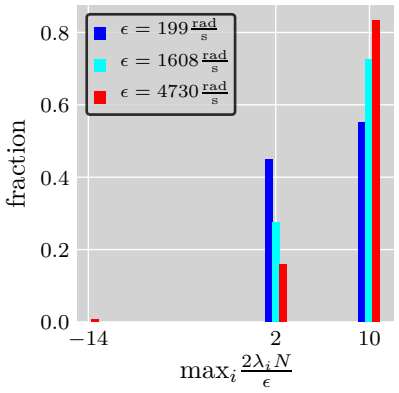
As mentioned already at the beginning of Sec. 5.7, the system gets close to other unstable fixed points during the recognition process. As the phase space close to these "intermediate patterns" does not change abruptly, dynamics of $\boldsymbol{\alpha}$ should depend on the eigenvalues of the intermediates. Specifically, both the dynamic equation in system coordinates Eq. (3.13) and the eigenvalues Eq. (3.7) contain a weighted sum of scalar products of the current system state with the memorized patterns. While the sum in Eq. (3.13) changes depending on the distance to the fixed point, the overall change is small, as few of the N pixels change. Consequently, large eigenvalues of the intermediate imply a large accelerating factor in the dynamics, and only eigenvalues of small absolute value allow a stability change of the pixel anywhere near the intermediate. Consequently, we expect the largest eigenvalue of an intermediate pattern to be positive, as that seems to be a prerequisite for switching. (From another point of view, the intermediate would not be an intermediate if it would attract the system state in all eigendirections. It would then be the final output pattern.) If a pixel takes a long time to switch, it will probably be the last to finish the switching process. (If the switch renders other pixels unstable upon completion, it will not be the last switch. However, this is rare.) Therefore, the largest eigenvalue of the last switch should be strongly correlated with recognition time. Consequently, we examine the largest eigenvalues for the last intermediate pattern.

Note that the system state might be far from a detected intermediate in a few measurements. In these rare cases, the calculated eigenvalues have little meaning for the dynamics and might be negative.

Fig. 5.21 and 5.22 show the the occurrence of largest eigenvalues before the final pixel switch of a measurement. The eigenvalues at the intermediate are calculated

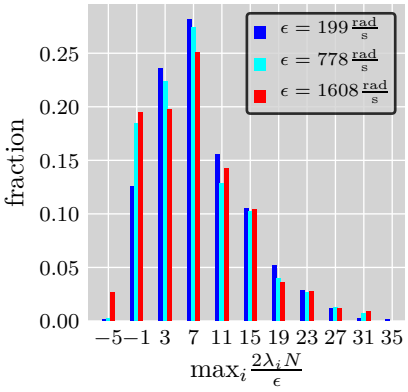


(a) 3 orthogonal output patterns

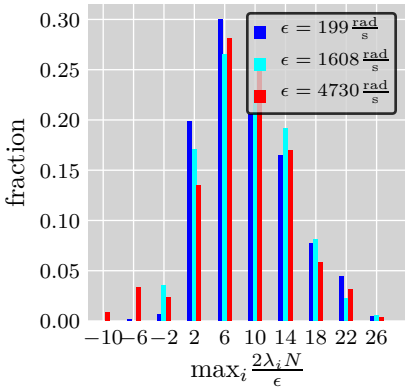


(b) 2 orthogonal output patterns

Figure 5.21: Distribution of largest eigenvalues $\max_i \lambda_i$ before the last pixel switch. (Orthogonal patterns, 256 measurements for each coupling strength and pattern number.)



(a) 3 random output patterns



(b) 2 random output patterns

Figure 5.22: Occurrence of largest eigenvalues $\max_i \lambda_i$ before the last pixel switch.(Random patterns, 3000 measurements for $M = 3$ random patterns at $\epsilon = 199\,\frac{\text{rad}}{\text{s}}$ and 1000 measurements for all other coupling strengths and pattern numbers.)

with Eq. (3.7), thus assuming an ideal network. (Eigenvalues λ_i are rescaled to emphasize the discrete nature of the eigenvalue spectrum: $\frac{2\lambda N}{\epsilon}$ are even or odd whole numbers depending on M. See Eq. (3.7) in Sec. 3.2.2.)

- The spectrum of largest eigenvalues is restricted to few values for orthogonal patterns. At $\epsilon = 199\,\frac{\text{rad}}{\text{s}}$, almost all eigenvalues are positive, with the negative ones belonging to measurements without well-defined intermediate patterns. At high coupling strengths, however, a significant fraction of the calculated largest eigenvalues is negative.

- For random patterns, the eigenvalue spectrum contains many possible values, which explains the absence of a substructure for recognition time histograms, where the multitude of eigenvalues lead to significant broadening of possible subpeaks.

- For a significant fraction of measurements for random patterns, $\max_i \lambda_i < 0$. While this only occurs at high coupling strengths for $M = 2$, measurements with $M = 3$ random patterns show negative largest eigenvalues at all coupling strengths. In an ideal system described by Eq. (3.5), such intermediates should be stable.

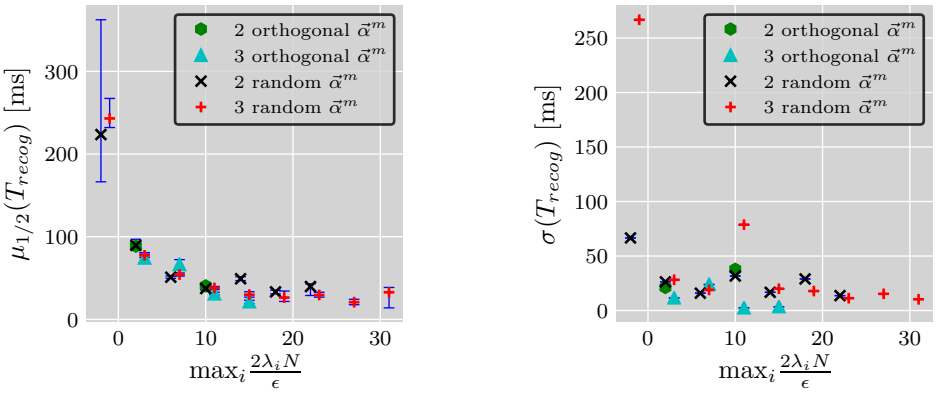
- In Fig. 5.23, median and standard deviation sum up the dependence on the largest calculated eigenvalue of the last intermediate: Both median and standard deviation are much larger for negative largest eigenvalues of the last intermediate patterns, so these measurements contribute substantially more to the tail of the full distribution.

Several possible deviations from Eq. (3.5) might be responsible for the switching despite negative eigenvalues:

- Inaccuracies $\Delta\Omega_i$ of the pair frequencies in principle destabilize all eigendirections.(See Sec. 3.6.1) At low ϵ, $\Delta\Omega_i$ affect the dynamics the most, so they seem irrelevant for the measurements shown.

- Coupling strength deviations due to component uncertainties alter the real eigenvalues present in the system (Sec. 5.1.2). As this effect would be constant $\forall\epsilon$, it does not explain the rise at high ϵ

- At high ϵ, higher order terms in the phase description lead to oscillations around the fixed points of the averaged dynamics and might destabilize eigendirections with negative eigenvalues of small absolute value. At even higher ϵ, the phase description is no longer sufficient and amplitude perturbations disturb the ideal dynamics.



(a) Median

(b) Standard deviation

Figure 5.23: Median and standard deviation as function of largest eigenvalues $\max_i \lambda_i$ before the last pixel switch ($\epsilon = 199\,\frac{\text{rad}}{\text{s}}$). (No error bars are shown for σ, as calculations of confidence intervals usually assume normal distributions and thus cannot be applied.)

- Additionally, electronic noise on the signal can enable a diffusive behavior over the remaining potential barrier. As coupling strength is varied with different coupling resistors directly in front of the oscillators, the signal's prefactor and any noise on the signal will be scaled as well. Consequently, the influence of coupling strength deviations and electronic noise should be independent of coupling strength.

All effects described either decrease the potential barrier, or induce a diffusive behavior that enables crossing of that barrier. Fig. 5.24 shows a typical measurement with $\max_i \lambda_i = -0.5\frac{\epsilon}{N}$ for the last switch. The diffusive behavior shown is visible in all manually observed measurements with extraordinary long recognition times.

The electronic noise and the inaccuracies in Ω_i and ϵ are not sufficient to destabilize negative eigenvalues for orthogonal patterns, as no tails occur for low and medium coupling strength. Whether the emergence of tails at high ϵ is due to higher order phase terms or amplitude effects remains unclear and might be tackled in future research. As recognition time distributions for random patterns show tails at all coupling strengths and an increase of tails for both low and high ϵ, none of the described effects seems completely negligible for the case of random output patterns.

5.8 External noise can improve performance

At the beginning of the recognition step, the system state must first depart from the input pattern, which is an unstable fixed point. As perturbations are small, this

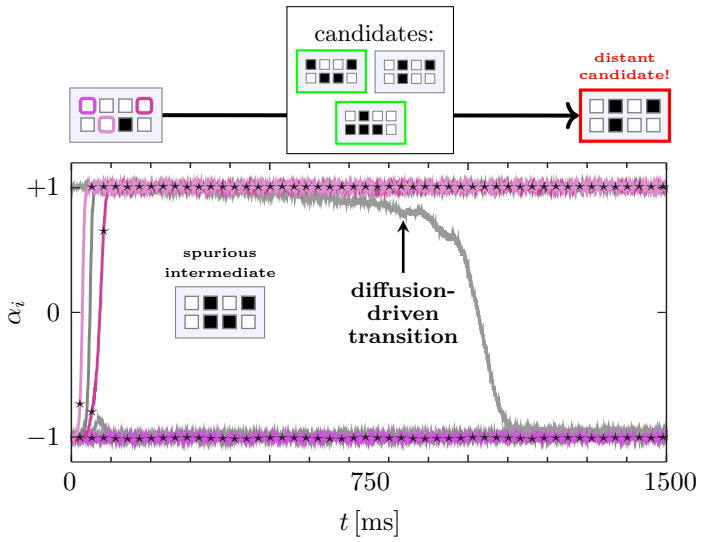
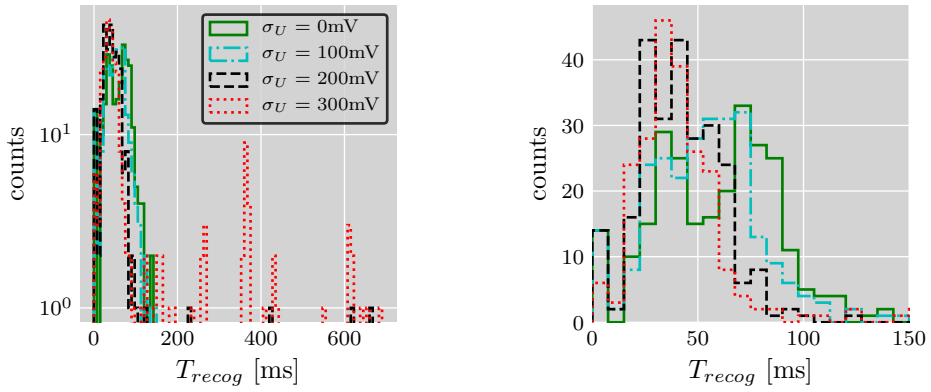


Figure 5.24: Noise can induce transitions from seemingly stable patterns. In this measurement, the system first settles on one of the spurious attractors. Before the last switch, largest eigenvalue $\max_i \lambda_i = -0.5\frac{\epsilon}{N}$ was calculated to be negative, but close to zero. Then, the pixel changes very slowly on a noisy trajectory. This seemingly diffusive transition might originate from a noise-induced diffusion over a potential barrier or a loss of stability due to non-ideal effects. The measurement was done at $\epsilon = 199\,\frac{\text{rad}}{\text{s}}$ with $M = 3$ random memorized patterns.

departure adds significantly to the recognition times. From another point of view, recognition times should diminish if perturbations were increased, which in practice could be realized by adding noise. Possibly, the noise might additionally destabilize spurious attractors, while memorized patterns remain stable. This destabilization could in turn lead to a better control over the recognition results. Thus, we artificially induced noise to explore system behavior.

Gaussian noise was applied to the reference oscillator in order to apply a noise component to the phase differences $\Delta\vartheta_i$. The noise was created with the Spectrum DA converter boards M2i.6011, M2i.6012 and M2i.6021: Emitted voltages consist of Gaussian random variables with variance σ_U that were changed every $\Delta t = 1\,\mu\text{s}$. These voltages were applied to the initialization resistor of the reference, but only during the recognition step. Histograms of recognition times with different σ_U of the applied noise are compared in Fig. 5.25, while medians and standard deviation of the distributions are shown in Fig. 5.26: The heads of the distributions are of similar shape. At high noise, however, the distributions gain sparsely populated tails similar to measurements at high coupling strengths or 3 random patterns. Consequently, the standard deviation rises sharply for $\sigma_U \geq 200\,\text{mV}$. As expected, medians decrease significantly with σ_U. However, it is not evident at first sight why



(a) Full histograms, logarithmic scale

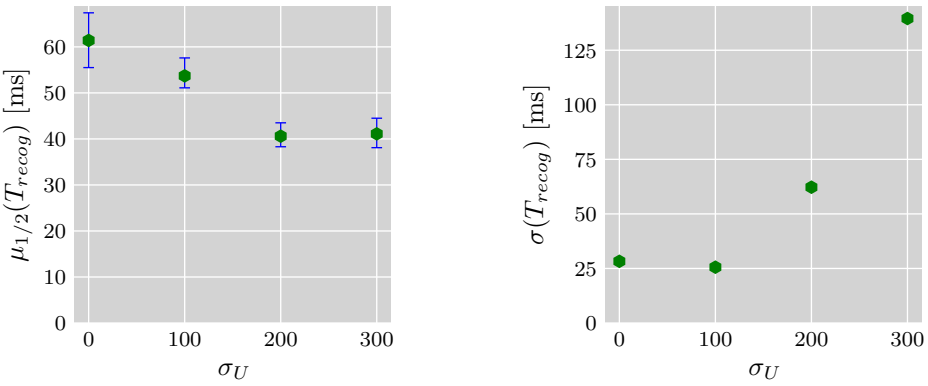
(b) Histogram heads, linear scale

Figure 5.25: Histograms of the recognition times T_{recog} for different standard deviations σ_U of externally applied Gaussian white noise. ($\epsilon = 199\,\frac{\text{rad}}{\text{s}}$, 3 orthogonal output patterns, bin size is 7,5 ms.)

the effect is not even larger: Trajectories spend a lot of time close to the initial fixed point, and noise in the α_i-variables should carry the system state away from the fixed point quickly. The noise in the α_i, however, is unfortunately strongly dependent on the average value of phase differences $\Delta\vartheta_i$: At the fixed point, $\frac{\mathrm{d}\alpha_i}{\mathrm{d}\Delta\vartheta_i} = 0$, so any noise in ϑ_i^{ref} (and thus $\Delta\vartheta_i$) will influence α_i in second order only. Still, supplying noise with $\sigma_U = 100\,\text{mV}$ reduces recognition times by more than 10%. Finally, we examine the influence of Gaussian noise on recognition success and pattern stability. As visible in Table 5.8 and 5.9, stability of spurious attractors is reduced slightly. Concomitantly, recognition rates of the memorized patterns rise slightly. Notably, the success rate for memorized patterns is almost 95% at the highest noise level. Summing up, external Gaussian noise can increase recognition success and reduce recognition time. In the future, different types of noise could be examined to further increase the effect: By using a noise distribution with larger tails, e.g. a Laplace distribution, noise should affect the dynamics more strongly at the fixed points, without leading to increased deviations during the transition.

5.9 Discussion

Summing up, our implementation of MONACO recognizes $M = 2$ orthogonal or random patterns without error. For $M = 3$ orthogonal memorized patterns, spurious attractors occur according to the majority rule Eq. (5.10), but more than 80 % of all possible input patterns are still mapped to the most similar memorized patterns. Indeed, the system state always moves to the closest attractor, and spurious attract-



(a) Median

(b) Standard deviation

Figure 5.26: Median and standard deviation of recognition times T_{recog} for 3 orthogonal patterns, $\epsilon = 199\,\frac{\text{rad}}{\text{s}}$ and varying standard deviations σ_U of externally applied Gaussian white noise. (No error bars are shown for the standard deviations, as calculations of confidence intervals usually assume normal distributions and thus cannot be applied.)

σ [rad]	success rate [%] (input: all patterns; correct: memorized)	success rate [%] (input: all patterns; correct: all attractors)	success rate [%] (input: theoretically guaranteed only)	unknown final pattern
0.0	82.0	99.6	100.0	0
480.3	84.4	99.6	100.0	0
960.6	84.6	99.2	98.1	0
1440.9	94.9	88.3	94.4	1

Table 5.8: Success rates for 3 orthogonal patterns at $\epsilon = 199\,\frac{\text{rad}}{\epsilon}$ with influence of externally applied Gaussian white noise. For different standard deviations σ of the noise, 256 measurements were conducted with the memorized patterns (00001111),(00110011) and (01010101). Spurious attractors were (00010111),(01110001),(01001101),(11010100) and their inverses. Two definitions of recognition success are shown in the table: Either recognition was only counted as successful, if the output was the closest memorized pattern, or the output being closest attractor to the initial pattern was sufficient. 21 % of these measurements are predicted to succeed according to Eq. (3.19).

σ [rad]	mem. pat. [%]	spur. pat. [%]	1st mem.	2nd mem.	3rd mem.	1st spur.	2nd spur.	3rd spur.	4th spur.
0.0	82.0	18.0	68	73	69	11	13	13	9
480.3	84.4	15.6	79	59	78	8	12	11	9
960.6	84.6	15.4	72	66	77	6	16	9	8
1440.9	96.5	3.1	79	83	85	1	5	0	2

Table 5.9: Distribution of output patterns for 3 orthogonal patterns at $\epsilon = 199\,\frac{\text{rad}}{\epsilon}$ with influence of externally applied Gaussian white noise. 256 measurements were conducted with the memorized patterns (00001111),(00110011) and (01010101). Spurious attractors were (00010111),(01110001),(01001101),(11010100) and their inverses.

ors might be used to denote input patterns which are similar to several memorized patterns. The capacity of the small network is only exceeded at $M = 3$ random pattern, where recognition rates drop to 70 %.

Typical recognition times can be reduced by increasing the coupling strength ϵ. Recognition times additionally vary with memorized and initial patterns, so the resulting distributions may show subpeaks and long, sparsely populated tails. As large recognition times are rare, however, reading output patterns directly after a distribution's head would influence recognition success only slightly.

At the upper limit of the coupling strength ϵ, recognition success drops and the tails of recognition time distributions increase in length and density. Stability of patterns is modified and noise promotes transitions between remaining attractors. While these deviations could originate from several different sources, both effects might be induced by higher order phase terms, which can create noisy oscillations around the averaged dynamics and promote unfavorable synchronization between different oscillator pairs. If the higher order terms are indeed the main cause for the deviations, the architecture variant MONACO II should provide even higher coupling strengths without errors, as less higher order terms are created and frequency conditions allow a more even distribution of oscillator frequencies.

Finally, both recognition rates for memorized patterns as well as recognition times of the network can be improved by adding Gaussian noise to the oscillators: Spurious attractors destabilize, which leads to an increase in correctly recognized memorized patterns. Additionally, the noise assists in the system state's departure from the input pattern, which reduces recognition times. As the effect of Gaussian noise is strongly diminished at fixed points, future research may test different noise distributions for a stronger effect.

Chapter 6

Summary and outlook

In this thesis, we devised a network of oscillators that acts as an autoassociative memory. Main differences to previous architectures [29, 30] are a novel, multiplicative temporal modulation of the coupling and a symmetric structure containing two oscillator groups. Oscillators in each group are globally coupled and the coupling is modulated with a function of the other group's signals. As a result, memorized patterns are isolated attractors of the network and coordinates are phase differences, which can easily be read out. In contrast to previous networks with modulated coupling, no external, time-dependent signals need to be supplied. Inspired by the symmetrical layout, we named the new architecture MONACO: *Mirrored Oscillator Networks for Autoassociative COmputation*. By deriving lower bounds on the basins of attraction, an analytic criterion for recognition success was derived and simulations validated our analysis. Additionally, we presented an architecture variant with a different frequency-connection tradeoff that assigned each oscillator an individual coupling modulation.

As any theoretical proposal can miss substantial practical obstacles for implementation, we built a network of 8 oscillator pairs in hardware as proof of principle: The implementation is based on previous networks of electronic Van der Pol oscillators [30, 31]. One major challenge was spurious coupling of oscillator pairs, which was mitigated with a mixture of various techniques, e.g. low-pass filtering of the voltage supply, spatial separation, and small Faraday cages. Frequency differences between oscillators of a pair proved to be critical as well: As manual adjustment takes too long to compensate for drifts, an automatic frequency annealing method based on nonlinear dynamics was conceived. Implementation required major changes to the oscillators, but frequency differences were reduced to less than 0.1 ‰.

The final network matches theoretical predictions very well: Isolated attractors as well as oscillations around the averaged dynamics occur as expected. Automation of the measurement process allowed to derive statistics on recognition events and recognition times for different coupling strengths, pattern numbers and both orthogonal and random patterns: The small network could perfectly discriminate 2 orthogonal or random memorized patterns. For 3 orthogonal patterns, spurious attractors could be observed. Still, more than 80 % of recognitions retrieved mem-

orized patterns correctly. If spurious attractors were included as possible correct output states, the perfect recognition rates of 100 % could be regained. In applications, the spurious attractors could e.g. be used to sort out patterns that are similar to several memorized patterns. If we exceed the capacity of the network by memorizing 3 random patterns, recognition rates drop, but are still larger than 70 %.

Deviations of the network dynamics are expected at extreme values of the coupling strength ϵ: At very low $\epsilon < 48{,}3\,\frac{\mathrm{rad}}{\mathrm{s}}$, the network is limited by remaining frequency differences. However, this lower bound was not found in the measured data. As large coupling strengths in principle decrease recognition times, however, the upper limit at very high $\epsilon \geq 1608\,\frac{\mathrm{rad}}{\mathrm{s}}$ is much more important: Here, changes in stability as well as a reduction in recognition success can be observed. While higher order phase terms as well as amplitude effects occur there, synchronization of different oscillator pairs is also possible. Over a large range of coupling strengths, however, the network performs nearly perfect.

Recognition times decrease with ϵ and typical recognition times of 16 ms are reached at $\epsilon = 778\,\frac{\mathrm{rad}}{\mathrm{s}}$ without impairing recognition success. Generally, the recognition times follow a non-Gaussian distribution that has a sparsely populated tail for random patterns and shows significant substructure for orthogonal patterns. However, tails are so sparsely populated that retrieving the final pattern early does not decrease recognition success recognizably.

Trajectories remain close to the initial, defective pattern for a considerable portion of the recognition time, as the pattern is an unstable fixed point. In order to quicken the departure of the system state, the network was exposed to Gaussian white noise at medium ϵ and 3 orthogonal memorized patterns: Some amount of noise was able to reduce recognition times and slightly destabilize spurious patterns, leading to increased recognition success.

Regarding MONACO, many other questions are worth exploring: First, synchronization of a single oscillator pair with a frequency difference is well-understood. In our network, however, already synchronized oscillator pairs promote synchronization in other pairs. The critical coupling strength for synchronization might be derived with the same approach as presented in [56]. Second, the architecture itself might be improved further: We already presented the architecture variation MONACO II with less strict frequency conditions, whose physical connections scale similarly to the original architecture with $\mathcal{O}(NM)$. In both MONACO I and II, frequency conditions are used to remove undesirable resonant terms. If another architecture variant would avoid creating such unnecessary coupling terms, frequency conditions could be further simplified. Third, destabilization of spurious attractors with noise probably depends heavily on the distribution. Switching from a Gaussian noise source to a heavy-tailed distribution - therefore increasing the diffusion away from the fixed points - promises more effective suppression of spurious attractors without loss of already correct recognitions. Finally, regarding implementation, recognition

times and power consumption can be greatly reduced by switching to other oscillator types like e.g. transition metal oxide (TMO) nano-oscillators [21]. At these high frequencies, signal delays might influence network performance and compensation of these effects might be necessary.

In addition, our improved dynamics open the way to more diverse oscillatory neural network architectures and applications: Mixing usual Kuramoto-like dynamics with MONACO's dynamics, the terms stabilizing both phase differences 0 and π can be tuned. This tradeoff might be another path to destabilize spurious attractors, as Nishikawa et. al. present a similar approach in [32]. Additionally, MONACO's dynamics seem very similar to continuous Hopfield networks [17]. It might be possible to directly map the dynamics of MONACO or a similar architecture to the dynamics of the continuous Hopfield network, or at least transfer some results. Such a mapping would e.g. allow different learning algorithms proposed for the Hopfield network to be investigated: Our current architecture uses constant coefficients according to the Hebbian rule and "real", incremented learning of patterns has not been used. While this is possible in Hopfield networks, application of learning is essential for other neural network types. Furthermore, as Hopfield networks can be used to solve combinatorics like the traveling salesman problem [17], oscillatory neural networks could solve those as well. Finally, Boltzmann machines are neural networks closely related to Hopfield networks. In addition to "visible" units, where patterns are initialized or retrieved, they also have "hidden" units not directly related to input or output. Furthermore, noise is added to the system to avoid local potential minima resp. spurious attractors. Both differences can be easily included in oscillatory neural networks. If incremental learning can be implemented in hardware (or weights can be transferred from mapped Boltzmann machines), oscillatory neural networks can contribute to the commercial success of neural networks in industry, where restricted Boltzmann machines and deep Boltzmann machines are used for feature representation or object and speech recognition.

In conclusion, we provide ways to remedy several major flaws of oscillatory neural networks. While some questions remain open, the dream of powerful, yet energy-saving neural network hardware seems to come in reach.

Appendix A

Averaging and frequency restrictions

In this Appendix, we apply the method of averaging [47] to the phase description Eq. (3.1). Therefore, we first expand the products in Eq. (3.1) with the trigonometric equalities $\sin x \cos y = [\sin(x-y) + \sin(x+y)]/2$ and $\sin x \sin y = [\cos(x-y) - \cos(x+y)]/2$ to obtain all frequency components:

$$\begin{aligned}
\dot{\vartheta}_i^{[1]} - \Omega_i^{[1]} &= \tfrac{\epsilon}{N} \cos\vartheta_i^{[1]} \cdot \sum_{j=1}^{N} a^{[2]}(t) \sin\vartheta_j^{[1]} \\
&= \tfrac{\epsilon}{N} \sum_{j,k,l=1}^{N} S_{lk} \sin\vartheta_l^{[2]} \cos\vartheta_i^{[1]} \cdot \sin\vartheta_j^{[1]} \sin\vartheta_k^{[2]} \\
&= \tfrac{\epsilon}{4N} \sum_{j,k,l=1}^{N} S_{lk} \Big[\sin\left(\vartheta_l^{[2]} - \vartheta_i^{[1]}\right) + \sin\left(\vartheta_l^{[2]} + \vartheta_i^{[1]}\right) \Big] \\
&\qquad \cdot \Big[\cos\left(\vartheta_j^{[1]} - \vartheta_k^{[2]}\right) - \cos\left(\vartheta_j^{[1]} + \vartheta_k^{[2]}\right) \Big] \\
&= \tfrac{\epsilon}{8N} \sum_{j,k,l=1}^{N} S_{lk} \Big[\quad \sin\left(\vartheta_l^{[2]} - \vartheta_i^{[1]} - \vartheta_j^{[1]} + \vartheta_k^{[2]}\right) \\
&\qquad + \sin\left(\vartheta_l^{[2]} - \vartheta_i^{[1]} + \vartheta_j^{[1]} - \vartheta_k^{[2]}\right) \\
&\qquad - \sin\left(\vartheta_l^{[2]} - \vartheta_i^{[1]} - \vartheta_j^{[1]} - \vartheta_k^{[2]}\right) \\
&\qquad - \sin\left(\vartheta_l^{[2]} - \vartheta_i^{[1]} + \vartheta_j^{[1]} + \vartheta_k^{[2]}\right) \\
&\qquad + \sin\left(\vartheta_l^{[2]} + \vartheta_i^{[1]} - \vartheta_j^{[1]} + \vartheta_k^{[2]}\right) \\
&\qquad + \sin\left(\vartheta_l^{[2]} + \vartheta_i^{[1]} + \vartheta_j^{[1]} - \vartheta_k^{[2]}\right) \\
&\qquad - \sin\left(\vartheta_l^{[2]} + \vartheta_i^{[1]} - \vartheta_j^{[1]} - \vartheta_k^{[2]}\right) \\
&\qquad - \sin\left(\vartheta_l^{[2]} + \vartheta_i^{[1]} + \vartheta_j^{[1]} + \vartheta_k^{[2]}\right) \Big]
\end{aligned}$$

As $\vartheta_i^{[1/2]} \approx \Omega_i t + \mathcal{O}(\epsilon)$, each of the sin-terms might oscillate with frequencies of $\mathcal{O}(\Omega_i)$ or $\mathcal{O}(\Delta\Omega_{ij})$.($\Delta\Omega_{ij} = \Omega_i - \Omega_j$) As the characteristic timescales $\mathcal{O}({\Omega_i}^{-1})$ and $\mathcal{O}({\Delta\Omega_{ij}}^{-1})$ are much smaller than ϵ^{-1}, the time average of these oscillating terms vanishes on times $\mathcal{O}(\epsilon^{-1}) \gg \mathcal{O}({\Delta\Omega_{ij}}^{-1})$. If frequencies in the argument cancel each other out, however, the argument is constant on timescales $\mathcal{O}(\epsilon^{-1})$ and all

oscillating terms are negligible. Depending on the signs in the sin-argument, there can be different possibilities how constant terms can arise:

In the first term, for example, frequencies cancel if $\Omega_l + \Omega_k = \Omega_i + \Omega_j$. That is always true for $l = i \wedge k = j$ or $l = j \wedge k = i$, imposing an interaction between the i^{th} and j^{th} oscillators in both networks depending on S_{ij}. However, frequencies might also cancel if the frequency distribution is chosen poorly, which would wrongly connect oscillators with different numbers i, j, k, l only dependent on S_{lk}. Therefore, we require $\Omega_l + \Omega_k \neq \Omega_m + \Omega_n \,\forall$ pairwise different l, k, m, n.

Similarly, the lowest order is obtained in the third term for $\Omega_l = \Omega_i + \Omega_j + \Omega_k$. In order to avoid interaction between the i^{th} and j^{th} oscillators based on S_{lk} again, the frequency distribution must obey $\Omega_l \neq \Omega_m + \Omega_n + \Omega_k \,\forall l, k, m, n$ and the third term becomes negligible as well as the fourth, fifth and sixth term.

While the eighth term averages out without further conditions, we get identical contributions from the second and the seventh term. This can be seen by renaming indices l and k and using $S_{lk} = S_{kl}$:

$$\begin{aligned}
\dot{\vartheta}_i^{[1]} - \Omega_i^{[1]} \approx & \frac{\epsilon}{8N} \sum_{j,k,l=1}^{N} S_{lk} \Big[(\delta_{il}\delta_{kj} + \delta_{ik}\delta_{lj} - \delta_{ik}\delta_{kl}\delta_{lj}) \sin\left(\vartheta_l^{[2]} - \vartheta_i^{[1]} - \vartheta_j^{[1]} + \vartheta_k^{[2]}\right) \\
& + 2\delta_{il}\delta_{kj} \sin\left(\vartheta_l^{[2]} - \vartheta_i^{[1]} + \vartheta_j^{[1]} - \vartheta_k^{[2]}\right) \Big] \\
= & \frac{\epsilon}{8N} \sum_{j=1}^{N} \Big[- S_{ij} \sin\Big(\big(\vartheta_i^{[1]} - \vartheta_i^{[2]}\big) + \big(\vartheta_j^{[1]} - \vartheta_j^{[2]}\big)\Big) \\
& - S_{ji} \sin\Big(\big(\vartheta_i^{[1]} - \vartheta_i^{[2]}\big) + \big(\vartheta_j^{[1]} - \vartheta_j^{[2]}\big)\Big) \\
& + 2S_{ij} \sin\Big(\big(\vartheta_i^{[1]} - \vartheta_i^{[2]}\big) - \big(\vartheta_j^{[1]} - \vartheta_j^{[2]}\big)\Big) \Big] \\
& + \frac{\epsilon}{8N} S_{ii} \sin\Big(2\big(\vartheta_i^{[1]} - \vartheta_i^{[2]}\big)\Big)
\end{aligned}$$

Final simplifications can be obtained by introducing the phase difference of oscillators with identical frequency $\Delta\vartheta_i = \vartheta_i^{[1]} - \vartheta_i^{[2]}$ and using $S_{ji} = S_{ij}$ as well as $S_{ii} = M$. For $\dot{\vartheta}_i^{[2]}$, the calculation is the same with inverted upper indices:

$$\dot{\vartheta}_i^{[1]} = \Omega_i^{[1]} + \frac{\epsilon M}{8N} \sin\big(2\Delta\vartheta_i\big) - \frac{\epsilon}{4N} \sum_{j=1}^{N} S_{ij} \left[\sin\big(\Delta\vartheta_i + \Delta\vartheta_j\big) + \sin\big(\Delta\vartheta_i - \Delta\vartheta_j\big)\right]$$

$$\dot{\vartheta}_i^{[2]} = \Omega_i^{[2]} - \frac{\epsilon M}{8N} \sin\big(2\Delta\vartheta_i\big) + \frac{\epsilon}{4N} \sum_{j=1}^{N} S_{ij} \left[\sin\big(\Delta\vartheta_i + \Delta\vartheta_j\big) + \sin\big(\Delta\vartheta_i - \Delta\vartheta_j\big)\right]$$

Remark:
As shown in [30], both conditions on the frequency distribution can be simplified further:(l, k, m, n are still pairwise different.)

$$\begin{aligned} \Omega_l + \Omega_k &\neq \Omega_m + \Omega_n \\ \Omega_l - \Omega_n &\neq \Omega_m - \Omega_k \\ \Delta\Omega_{ln} &\neq \Delta\Omega_{mk} \end{aligned} \tag{A.1}$$

All difference frequencies have to be different to each other. This can be fulfilled by multiplying the minimal difference frequency $\Delta\Omega_{min}$ with a Golomb-ruler [48], a set of integers with non-equal differences. Similarly, the second condition can be simplified to $\Omega_l - \Omega_m = \Delta\Omega_{lm} \neq \Omega_n + \Omega_k \, \forall l, k, m, n$. This last inequality is always fulfilled if $\Omega_{min} > \Omega_{max}/3$.

Appendix B

Ljapunov function and unstable fixed point sets

In this Appendix, we derive a Ljapunov function for Eq. (3.5). We use it to show that all fixed points with at least one index i that fulfills Eq. (3.6) are unstable.

Ljapunov function

First, we express Eq. (3.5) as a gradient system with potential U, where $\Delta\dot{\vartheta}_i = -\partial U/\partial\Delta\vartheta_i \,\forall i$:

$$U = -\frac{\epsilon}{2N}\sum_{l=1}^{N}\left(\sum_{k=1}^{N} S_{kl}\cos\Delta\vartheta_k\cos\Delta\vartheta_l - \frac{M}{2}\cos^2\Delta\vartheta_l\right) \tag{B.1}$$

This is equivalent to the overdamped motion of a particle in an energy landscape, where $\dot{\boldsymbol{v}} \propto -\boldsymbol{\nabla}E$. Therefore, U decreases along trajectories and is a Ljapunov-function, which ensures that fixed points are the only attractors possible in Eq. (3.5).

Unstable fixed points

In order to prove that all fixed points with some $i \in q$ are unstable, we express the system state in pattern coordinates $\boldsymbol{\alpha}$ with $\alpha_i = \cos\Delta\vartheta_i$ and insert the coupling matrix $S_{ij} = \sum_{m=1}^{M}\alpha_i^m\alpha_j^m$ into our potential function U:

$$\begin{aligned} U &= -\frac{\epsilon}{2N}\sum_{l=1}^{N}\left(\sum_{k=1}^{N} S_{kl}\cos\Delta\vartheta_k\cos\Delta\vartheta_l - \frac{M}{2}\cos^2\Delta\vartheta_l\right) \\ &= -\frac{\epsilon}{2N}\sum_{k,l=1}^{N}\left(\sum_{m=1}^{M}\alpha_k^m\alpha_l^m\alpha_k\alpha_l - \frac{M}{2}\delta_{kl}\alpha_l^2\right) \end{aligned}$$

$$= -\frac{\epsilon}{2N}\left(\sum_{m=1}^{M}\langle\boldsymbol{\alpha}^m,\boldsymbol{\alpha}\rangle^2 - \frac{M}{2}\langle\boldsymbol{\alpha},\boldsymbol{\alpha}\rangle\right)$$

Now consider a small perturbation $\gamma\hat{e}_i$ from a fixed point $\boldsymbol{\alpha}^*$ where $i \in q$:

$$\begin{aligned}
&U(\boldsymbol{\alpha}^* + \gamma\hat{e}_i) = \\
&= -\frac{\epsilon}{2N}\left[\sum_{m=1}^{M}\big(\langle\boldsymbol{\alpha}^m,\boldsymbol{\alpha}^*\rangle + \underbrace{\langle\boldsymbol{\alpha}^m,\gamma\hat{e}_i\rangle}_{=\gamma\alpha_i^m}\big)^2 - \frac{M}{2}\big(\langle\boldsymbol{\alpha}^*,\boldsymbol{\alpha}^*\rangle + 2\langle\boldsymbol{\alpha}^*,\gamma\hat{e}_i\rangle + \langle\gamma\hat{e}_i,\gamma\hat{e}_i\rangle\big)\right] \\
&= U(\boldsymbol{\alpha}^*) - \frac{\epsilon\gamma}{N}\underbrace{\left(\sum_{m=1}^{M}\alpha_i^m\langle\boldsymbol{\alpha}^m,\boldsymbol{\alpha}^*\rangle - \frac{M}{2}\alpha_i^*\right)}_{=0,\text{ as } i\in q.\text{(see Eq. (3.6))}} - \frac{\epsilon}{2N}\left(\sum_{m=1}^{M}\gamma^2\underbrace{(\alpha_i^m)^2}_{=+1} - \frac{M}{2}\gamma^2\right) \\
&= U(\boldsymbol{\alpha}^*) - \frac{\epsilon M}{4N}\gamma^2
\end{aligned}$$

As U decreases close to $\boldsymbol{\alpha}^*$, there must be an unstable eigendirection and the fixed point must be unstable if at least one i with $\sum_{j=1}^{N} S_{ij}\cos\Delta\vartheta_j^* - \frac{M}{2}\cos\Delta\vartheta_i^* = 0$ exists.(i.e. $i \in q$)

Therefore, only the isolated fixed points with $\sin\Delta\vartheta_i^* = 0\,\forall i$ and $S_{ik}\sin\Delta\vartheta_l^* - M/2\delta_{ik}\sin\Delta\vartheta_i^* \neq 0\,\forall i$ can be attractors.

Appendix C

Stripboard layouts

In this Appendix we show the stripboard layouts of the final modules used for measurements of MONACO I. Common components are labeled in Fig. C.1. In addition to showing bare components, some color coding was used: Red, black and blue wires or copper strips correspond to the power supply's 15 V, 0 V and −15 V. On the oscillator submodules, some pins are colored gray: These pins do not transfer power or signal, but provide mechanical support only.

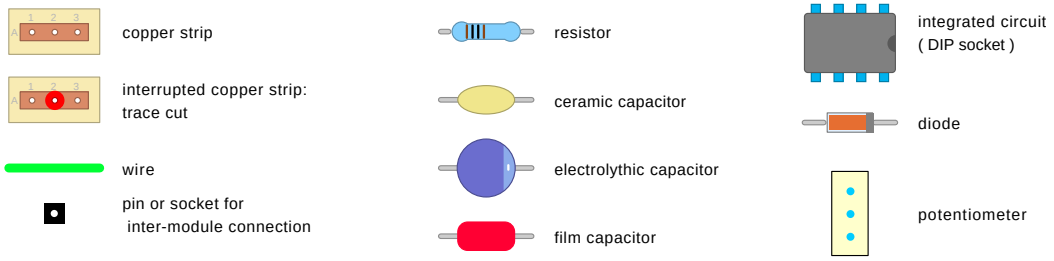
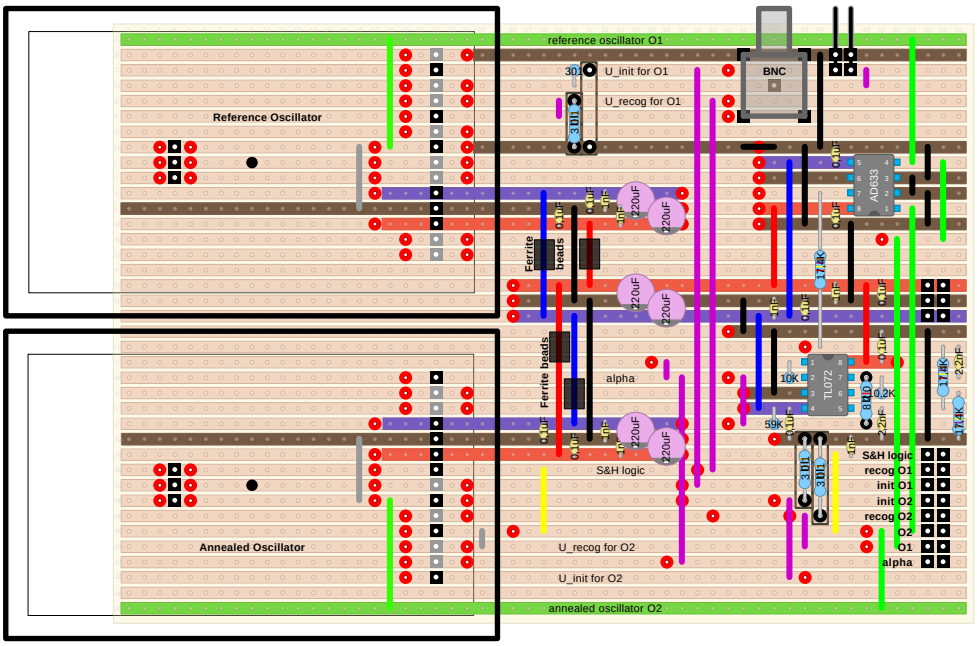
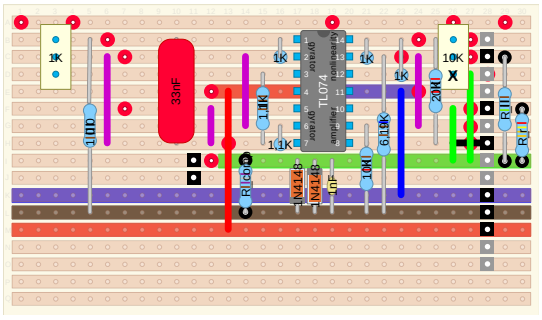


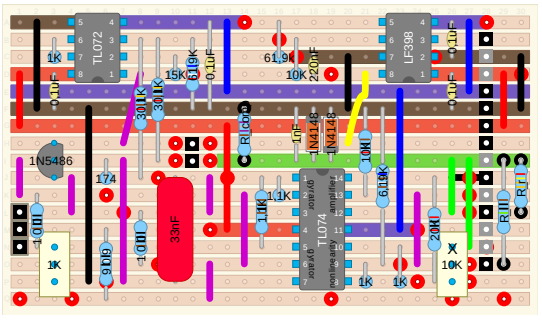
Figure C.1: Legend for most frequently used components.



(a) **Oscillator module**
Copper strips below the multiplier AD633 are broken between all pin pairs except 6-3. All copper strips below the operational amplifier IC TL072 are broken between pin pairs.



(b) **Reference oscillator submodule**
Copper strips below the operational amplifier IC TL074 are broken between all pin pairs except 3-12. The copper strip is also broken below the southernmost pin of the 1 kΩ-Potentiometer and the pin is not connected.



(c) **Annealed oscillator submodule**
Copper strips below all ICs are broken between all pin pairs except below pin pair 5-10 of the operational amplifier IC TL074. The copper strip is also broken below the northernmost pin of the 1 kΩ-Potentiometer and the pin is not connected.

Figure C.2: Stripboard layouts for the oscillator modules and the submodules for reference oscillators and annealed oscillators.

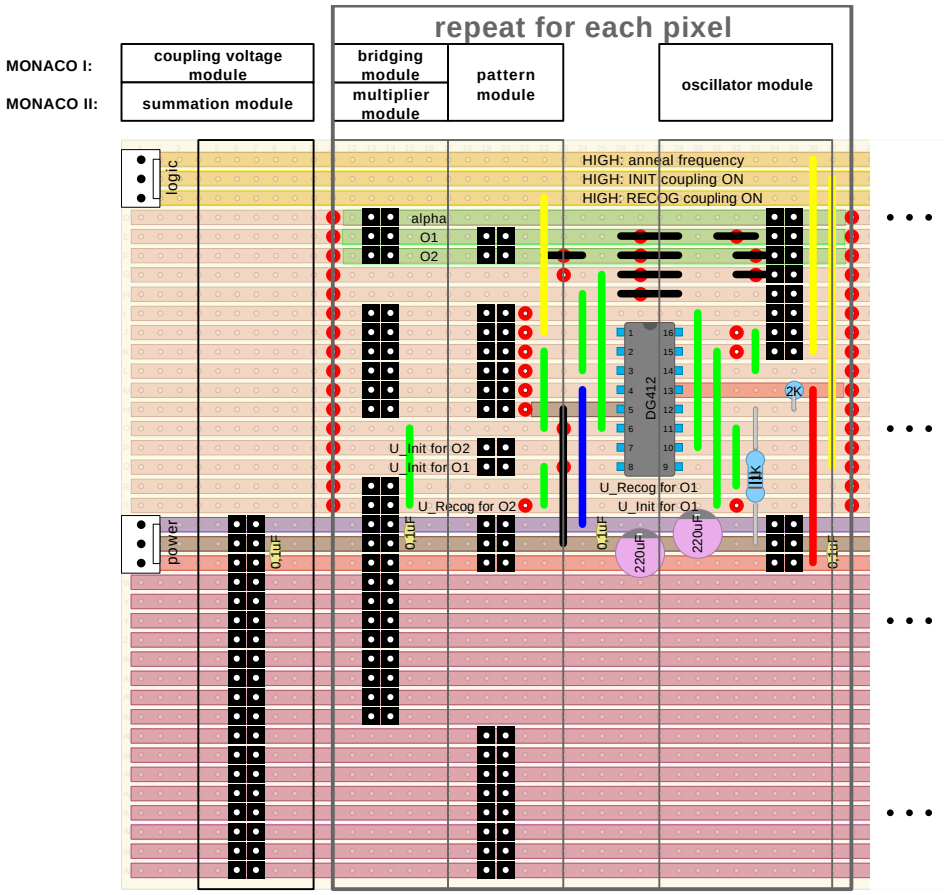


Figure C.3: Stripboard layout for the **mainboard**. Copper strips below the analog switch DG412 are broken between all pin pairs except 1-16 and 8-9. Module placement for the two architecture variations MONACO I and MONACO II is shown at the top.

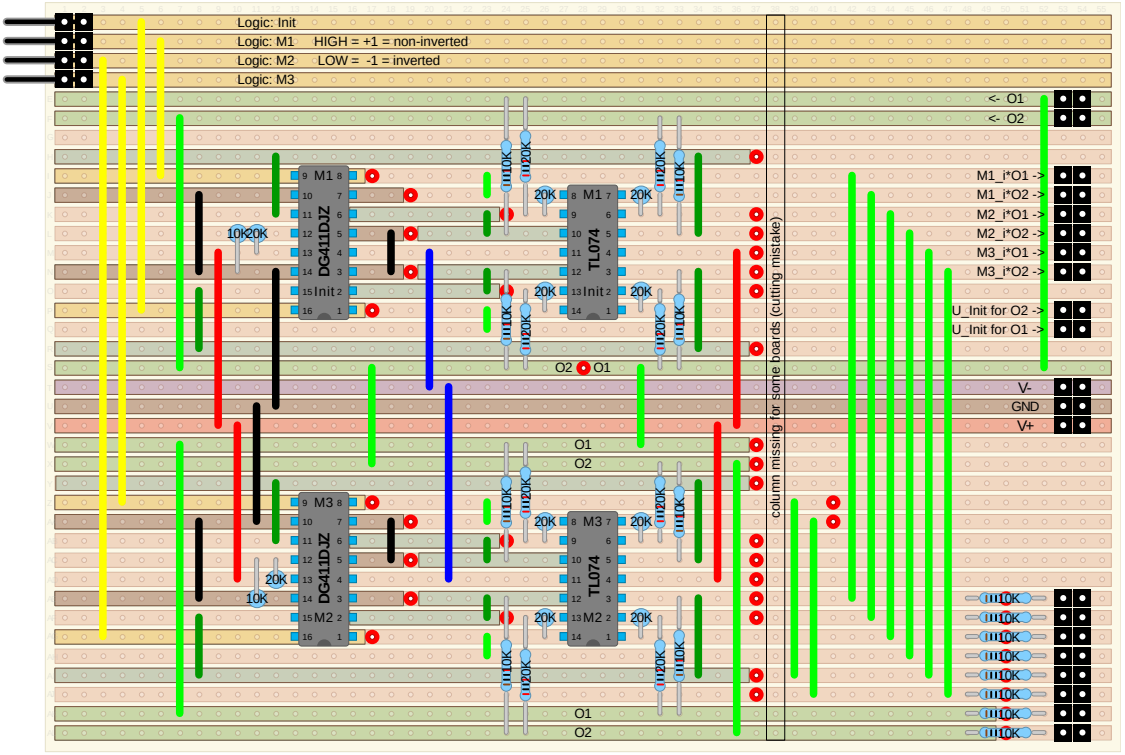
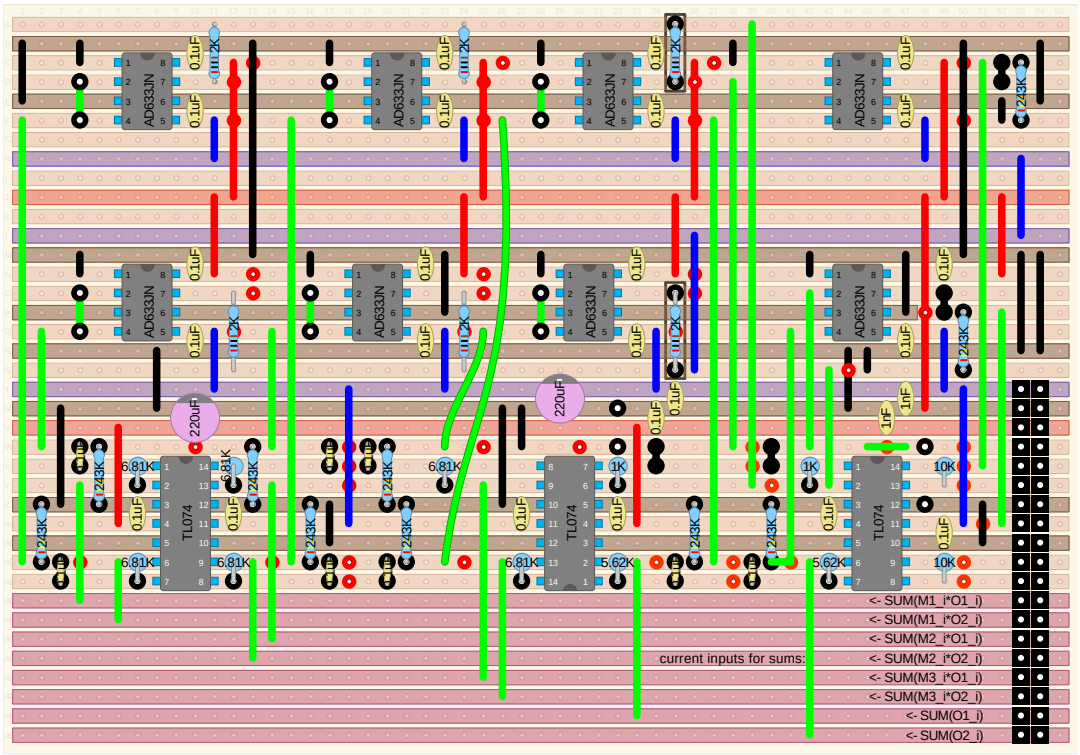
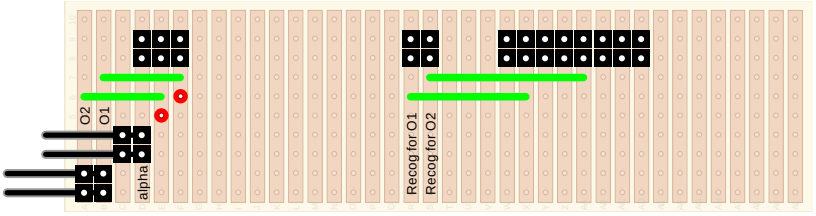


Figure C.4: Stripboard layout for the **pattern modules**. Copper strips below the analog switches DG411 are broken between all pin pairs except 1-16, 3-14, 7-10 and 8-9. Copper strips below operational amplifier ICs TL074 are all broken between pin pairs.

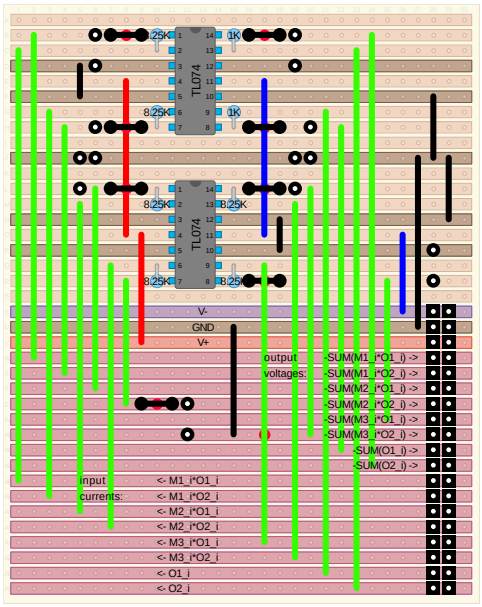


(a) **Coupling voltage module**
Copper strips are broken below the all ICs except below pin pairs 3-6 of multipliers AD633JN and pin pairs 3-12 and 5-10 of operational amplifier ICs TL074.

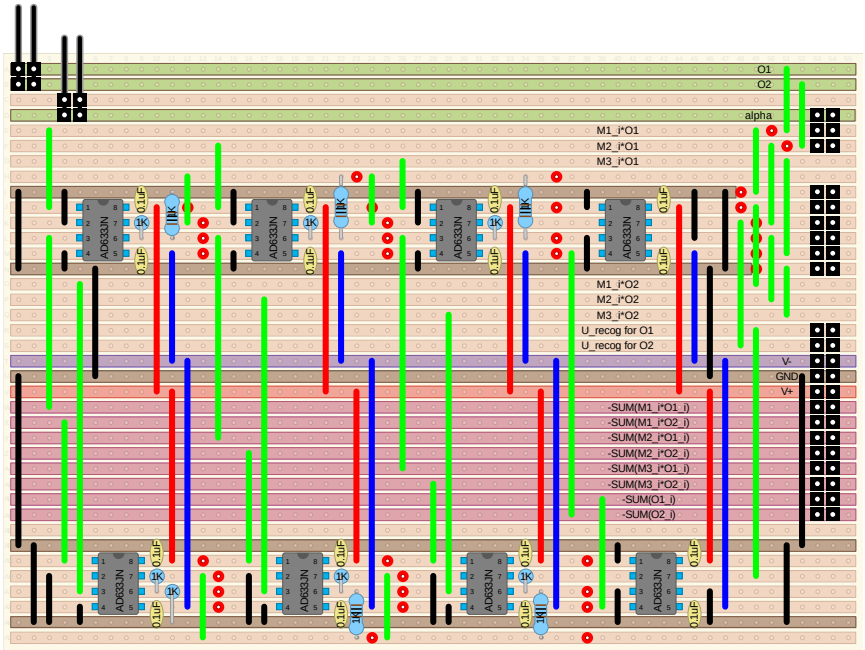


(b) **Bridging module**

Figure C.5: Stripboard layouts for modules exclusively used for architecture variant MONACO I: coupling voltage module and bridging modules.



(a) **Summation module**
Copper strips are broken below all operational amplifier ICs TL074 except between pin pairs 3-12 and 5-10.



(b) **Multiplier module**
Copper strips are broken below all ICs.

Figure C.6: Stripboard layouts for modules exclusively used for architecture variant MONACO II: multiplier modules and summation module.

Appendix D

Deviations of the capacitors used in the oscillators

As the capacitors used as capacitance of the Van der Pol oscillators did not have a tolerance marking, the capacitance deviations were measured manually.

The measurement circuitry shown in figure D.1 can be described by the voltage over in the combined capacitance $C_{tot} = C + C_{par}$:

$$\begin{aligned} C_{tot}U &= Q \\ \dot{U} &= \frac{1}{C_{tot}}\dot{Q} \\ &= \frac{1}{C_{tot}}\left(\frac{U_{in} - U}{R_{in}} - \frac{U}{R_{in,osci}}\right) \\ &= \frac{1}{C_{tot}}\left(\frac{U_{in}}{R_{in}} - \left[\frac{1}{R_{in}} + \frac{1}{R_{in,osci}}\right]U\right) \\ &\stackrel{!}{=} \frac{1}{C_{tot}}\frac{U_{t=\infty} - U}{R'_{in}} \end{aligned}$$

This is a simple capacitor charge equation with $\frac{1}{R'_{in}} = \frac{1}{R_{in}} + \frac{1}{R_{in,osci}}$ and an effective input voltage $U_{t=\infty} = U_{in}\frac{R_{in,osci}}{R_{in,osci}+R_{in}}$ with the solution $U = U_{t=\infty}(1 - \exp(-\frac{t}{R'_{in}C_{tot}}))$ if $U(t = 0) = 0\,\mathrm{V}$.

The time constant of the whole circuit $\tau = R'_{in}C_{tot}$ was measured as time difference between the positive slope of a slow square wave($f = 100\,\mathrm{Hz}$, $U_{in} \in [0\,\mathrm{V}, 10\,\mathrm{V}]$) and the time when $(1 - 1/e)U_{t=\infty}$ was reached in the measured voltage response. As $C \approx 1\,\mathrm{nF}$, a large $R_{in} = 243\,\mathrm{k\Omega}$ was used to get a large enough τ. First, the parasitic capacitance C_{par} was determined by omitting C, resulting in $C_{par} = \frac{\tau}{R'_{in}} = \frac{23{,}12\,\mu\mathrm{s}}{195{,}5\,\mathrm{k\Omega}} = 0{,}118\,\mathrm{nF}$. Afterwards, 18 capacitors were measured as $C = \frac{\tau}{R'_{in}} - C_{par}$. The resulting mean 1,001 nF is close to the ideal value of 1 nF which verifies our measurement method. The standard deviation is 0,018 nF. For the propagation of

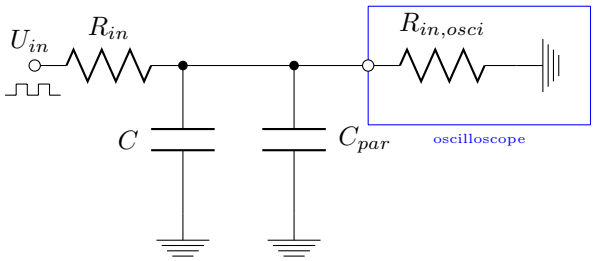


Figure D.1: Circuitry for determining capacitance C by measuring the time constant τ. $R_{in} = 243\,\mathrm{k\Omega}$ is a large resistance and capacitors were in the range of $1\,\mathrm{nF}$. Additionally, the parasitic capacitance of the circuitry $C_{par} = 0{,}118\,\mathrm{nF}$ and the input resistance $R_{in,osci} = 1\,\mathrm{M\Omega}$ had to be considered in the derivation of C. A slow square wave with $f = 100\,\mathrm{Hz}$ with $U_{in} \in [0\,\mathrm{V}, 10\,\mathrm{V}]$ was applied at R_{in}, and an effective time constant τ was taken as time difference between the positive slope of the square wave and the time when the measured voltage reached the fraction $1 - 1/e \approx 63,21\%$ of $U_{t=\infty} = \frac{R_{in,osci}}{R_{in,osci}+R_{in}} = 8{,}05\,\mathrm{V}$. C_{par} was determined with the same method, but without connecting an additional capacitance C.

uncertainty of ϵ, we will use the 95%-confidence interval / 2 standard deviations as deviation of the used capacitors, which results in deviations $\frac{\Delta C}{C} \approx 4\%$.

Appendix E

Preliminary measurements of MONACO II

Measurements of the architecture variant MONACO II (Sec. 3.7 and 4.6.) have been conducted by Nicolas Pfifferling as part of his master's thesis. As older oscillator modules were used, phase response curve amplitudes vary strongly with frequency and amplified oscillator amplitudes could not be adjusted during operation, which makes them susceptible to drifts. Consequently, measurements of MONACO II are only preliminary. Success rates for 3 memorized orthogonal patterns and are shown in Fig. E.1. Overall, the preliminary measurements strongly indicate improved performance compared to MONACO I:

- Success rates for memorized patterns surpass MONACO I.
- No upper limit of network performance could be found, although coupling resistors were decreased down to the limits of the circuitry. Consequently, significantly smaller recognition times can be reached.
- The absence of a performance drop at high frequencies also implies that stability of patterns is not shifted as strongly at high coupling strengths.

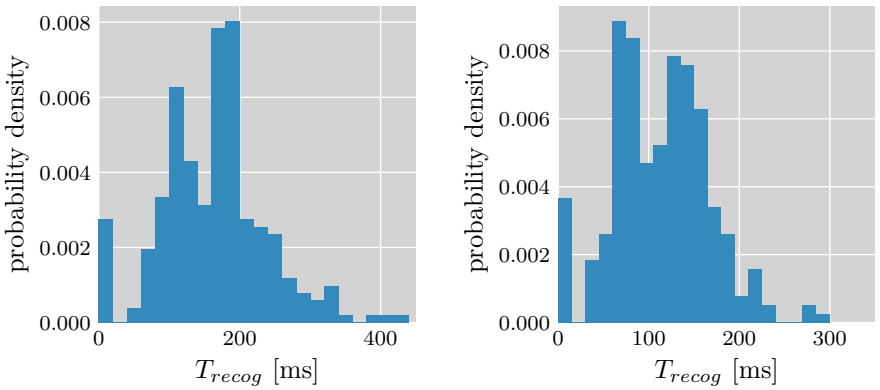
Summing up, MONACO II promises to be a lot more stable and performant than MONACO I, as even preliminary measurements with large inaccuracies surpass measurements of MONACO I with optimized oscillator modules.

R_{recog} [kΩ]	success rate [%] (memorized)	mem. pat.	spur. pat.	number of measurements
1000	86.2	227	19	246
511	87.4	226	29	255
243	89.8	230	26	256
133	88.6	217	28	245
100	91.4	234	22	256
82.5	92.6	237	19	256
62	92.2	236	20	256
30	93.3	239	17	256
10.2	92.6	237	19	256
7.68	91.8	235	21	256

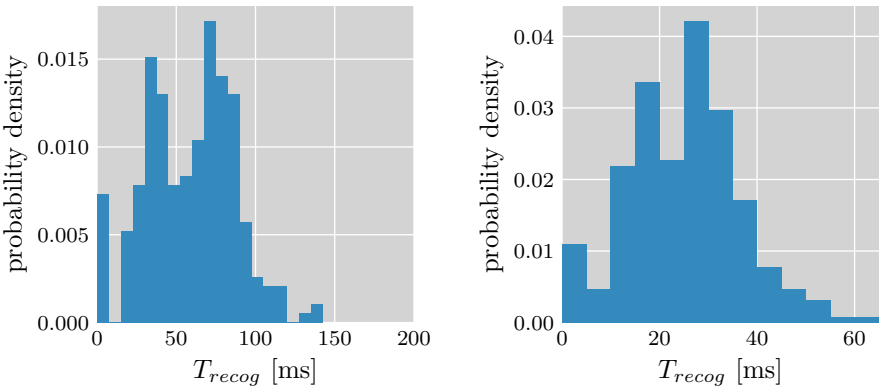
Table E.1: Success rates for three orthogonal patterns (MONACO II). For each coupling strength ϵ, 256 measurements were conducted with the memorized patterns (00001111), (00110011) and (01010101). Spurious attractors were (00010111), (01110001), (01001101), (11010100) and their inverses. Recognitions were counted as successful, if the output was the closest memorized pattern (or one of their inverses). For some recognition resistances (1 MΩ, 511 kΩ and 133 kΩ), some measurements had to be excluded as manual frequency adjustment was not precise enough and led to failure of the automatic frequency annealing.

Appendix F

Recognition time histograms

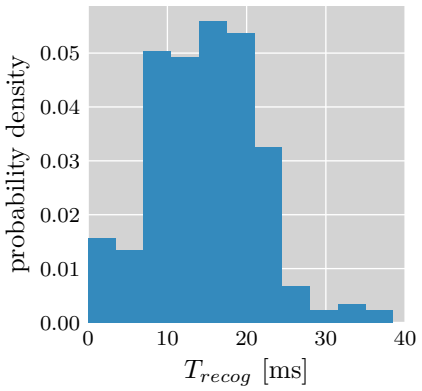


(a) $\epsilon = 48{,}3\ \frac{\text{rad}}{\text{s}}$ (b) $\epsilon = 84{,}7\ \frac{\text{rad}}{\text{s}}$

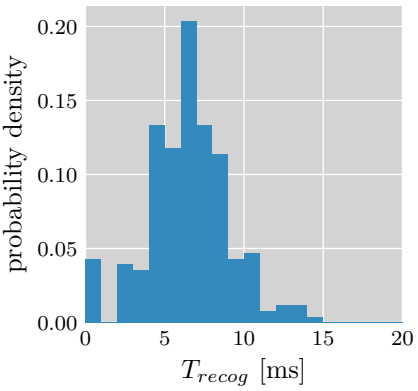


(c) $\epsilon = 199\ \frac{\text{rad}}{\text{s}}$ (d) $\epsilon = 483\ \frac{\text{rad}}{\text{s}}$

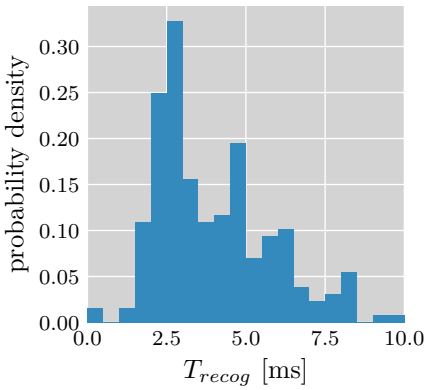
Figure F.1: Recognition time distributions for *three orthogonal patterns* at low coupling strengths ϵ. Measurements were conducted for all 256 possible input patterns.



(a) $\epsilon = 778\,\frac{\text{rad}}{\text{s}}$

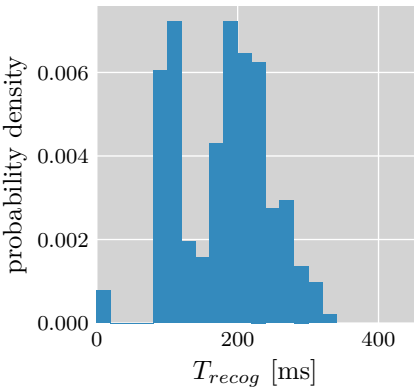


(b) $\epsilon = 1608\,\frac{\text{rad}}{\text{s}}$, head

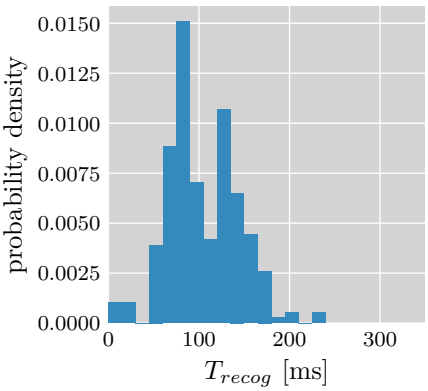


(c) $\epsilon = 4730\,\frac{\text{rad}}{\text{s}}$, head

Figure F.2: Recognition time distributions for *three orthogonal patterns* at high coupling strengths ϵ. Measurements were conducted for all 256 possible input patterns. For $\epsilon = 1608\,\frac{\text{rad}}{\text{s}}$ and $\epsilon = 4730\,\frac{\text{rad}}{\text{s}}$, the distributions have a sparsely populated tail, which is not shown.

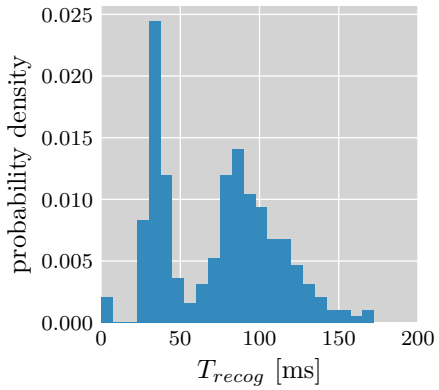


(a) $\epsilon = 48{,}3\,\frac{\text{rad}}{\text{s}}$

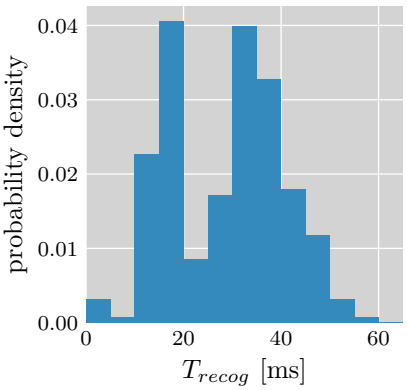


(b) $\epsilon = 94{,}4\,\frac{\text{rad}}{\text{s}}$

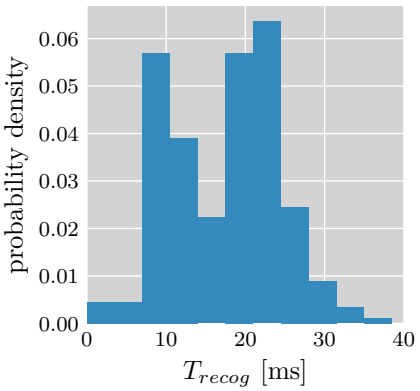
Figure F.3: Recognition time distributions for *two orthogonal patterns* at low coupling strengths ϵ. Measurements were conducted for all 256 possible input patterns.



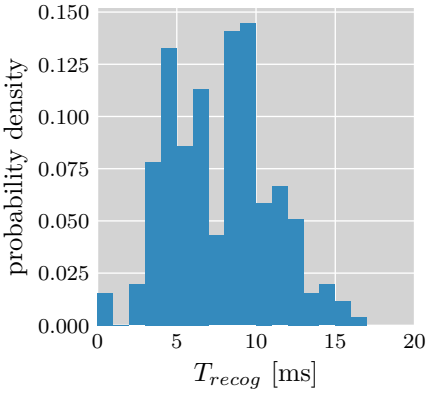
(a) $\epsilon = 199\,\frac{\text{rad}}{\text{s}}$



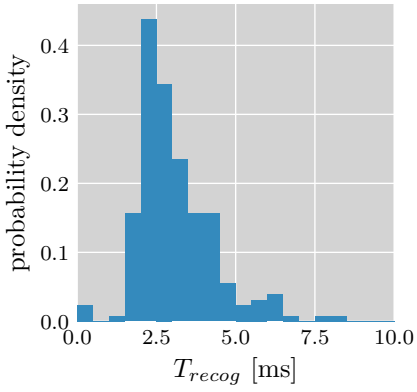
(b) $\epsilon = 483\,\frac{\text{rad}}{\text{s}}$



(c) $\epsilon = 778\,\frac{\text{rad}}{\text{s}}$

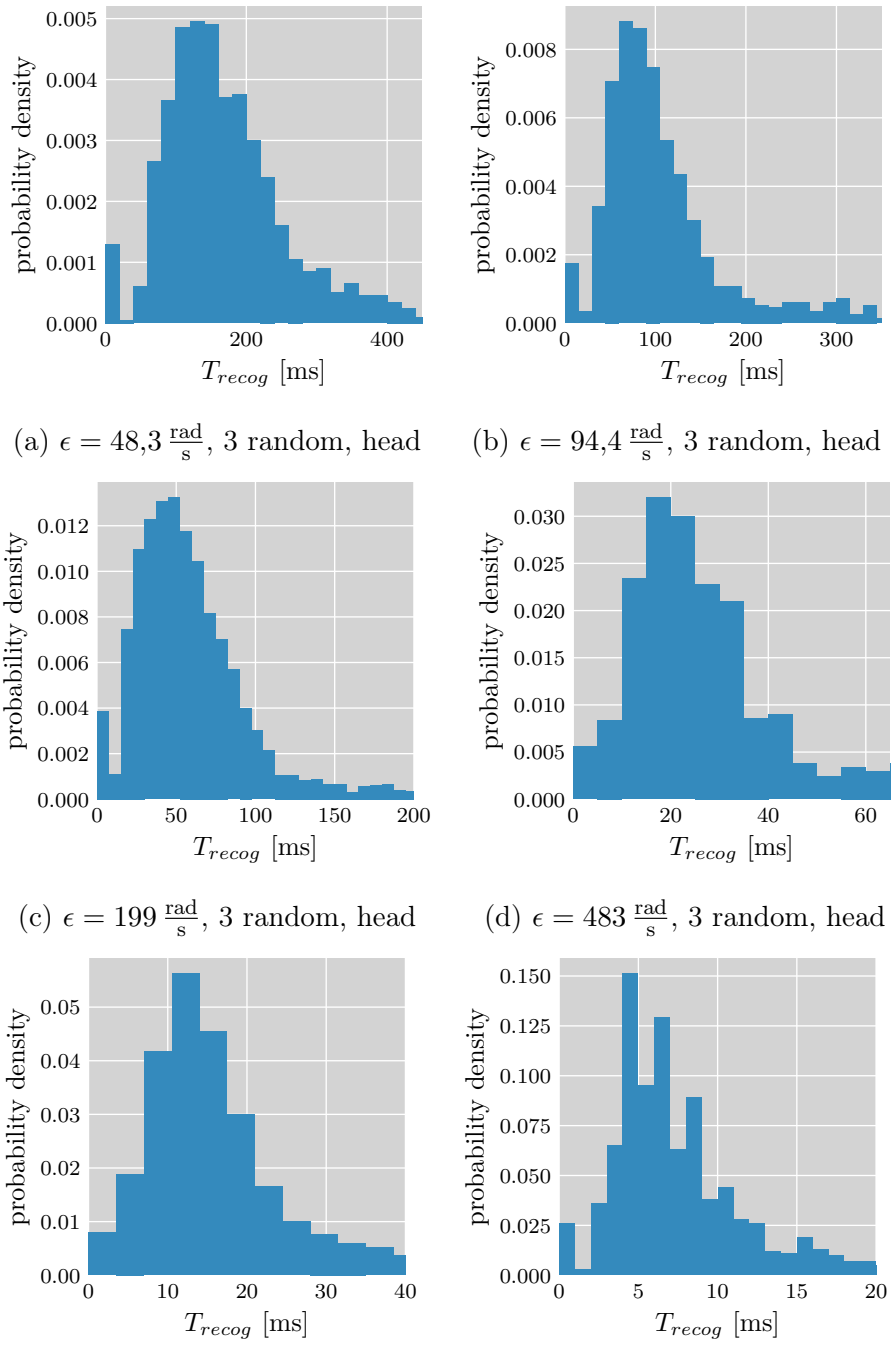


(d) $\epsilon = 1608\,\frac{\text{rad}}{\text{s}}$



(e) $\epsilon = 4730\,\frac{\text{rad}}{\text{s}}$, head

Figure F.4: Recognition time distributions for *two orthogonal patterns* at high coupling strengths ϵ. Measurements were conducted for all 256 possible input patterns. For $\epsilon = 4730\,\frac{\text{rad}}{\text{s}}$, the distribution has a sparsely populated tail, which is not shown.



(a) $\epsilon = 48{,}3\,\frac{\text{rad}}{\text{s}}$, 3 random, head

(b) $\epsilon = 94{,}4\,\frac{\text{rad}}{\text{s}}$, 3 random, head

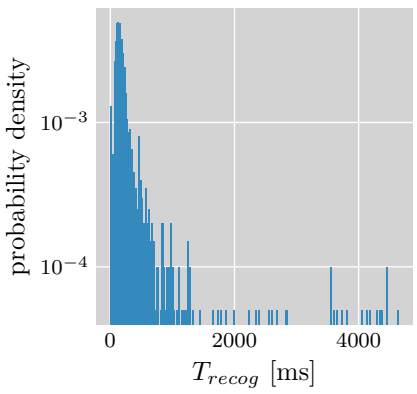
(c) $\epsilon = 199\,\frac{\text{rad}}{\text{s}}$, 3 random, head

(d) $\epsilon = 483\,\frac{\text{rad}}{\text{s}}$, 3 random, head

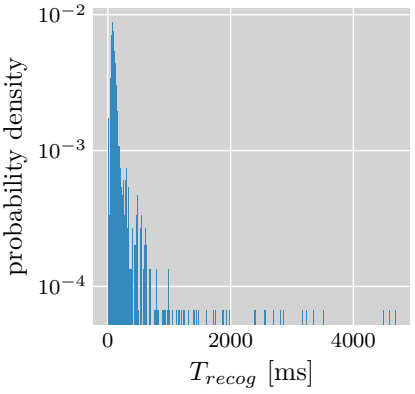
(e) $\epsilon = 778\,\frac{\text{rad}}{\text{s}}$, 3 random, head

(f) $\epsilon = 1608\,\frac{\text{rad}}{\text{s}}$, 3 random, head

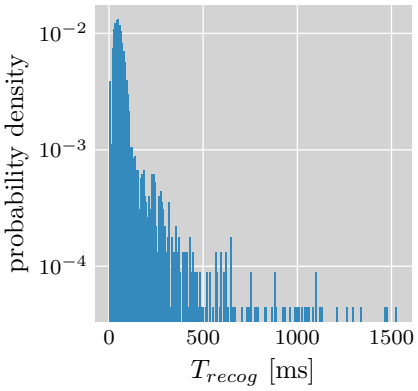
Figure F.5: Three random patterns: Recognition time distributions heads. For $\epsilon = 199\,\frac{\text{rad}}{\text{s}}$, 3000 measurements were used, for all other coupling strengths 1000 measurements were conducted. Each distribution has a sparsely populated tail which is not shown.



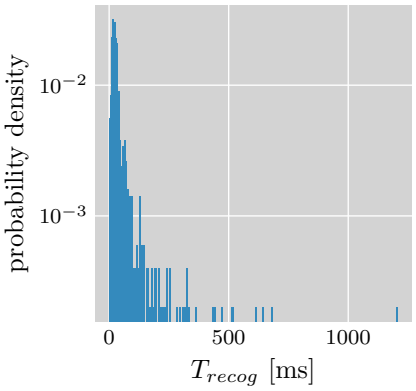
(a) $\epsilon = 48{,}3\ \frac{\text{rad}}{\text{s}}$, 3 random, full



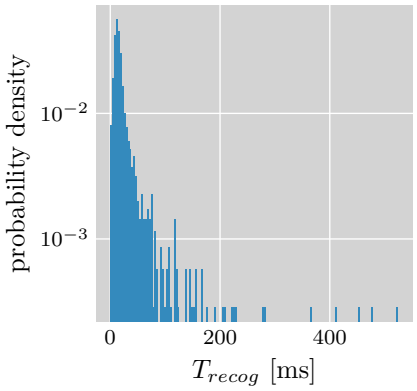
(b) $\epsilon = 94{,}4\ \frac{\text{rad}}{\text{s}}$, 3 random, full



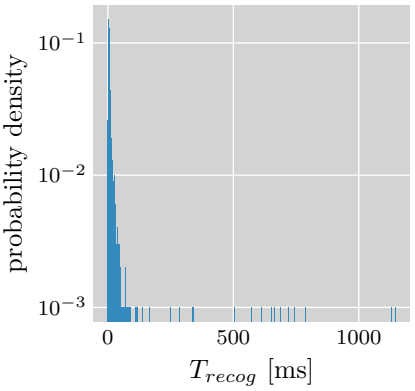
(c) $\epsilon = 199\ \frac{\text{rad}}{\text{s}}$, 3 random, full



(d) $\epsilon = 483\ \frac{\text{rad}}{\text{s}}$, 3 random, full

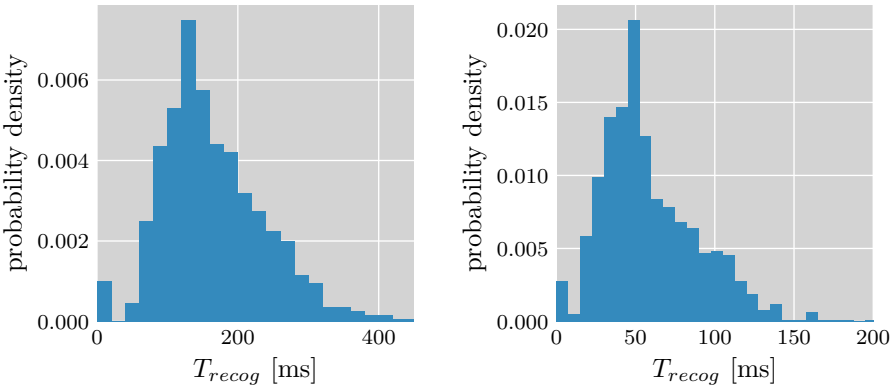


(e) $\epsilon = 778\ \frac{\text{rad}}{\text{s}}$, 3 random, full



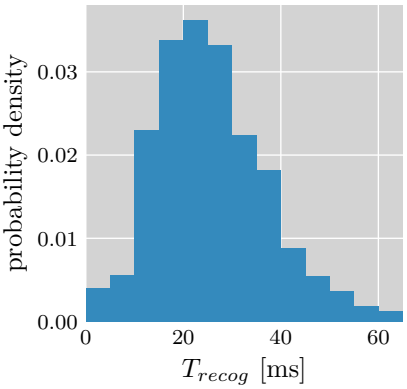
(f) $\epsilon = 1608\ \frac{\text{rad}}{\text{s}}$, 3 random, full

Figure F.6: Three random patterns: Full recognition time distributions, probability axis log-scaled for better visibility of long tails. For $\epsilon = 199\ \frac{\text{rad}}{\text{s}}$, 3000 measurements were used, for all other coupling strengths 1000 measurements were conducted.

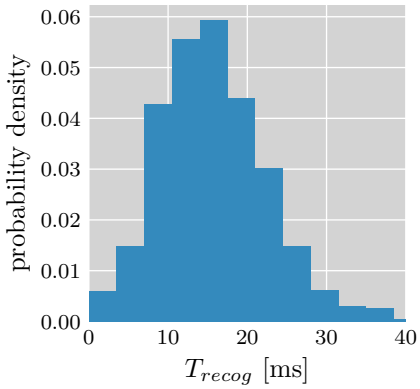


(a) $\epsilon = 48{,}3\ \frac{\text{rad}}{\text{s}}$, 2 random, head

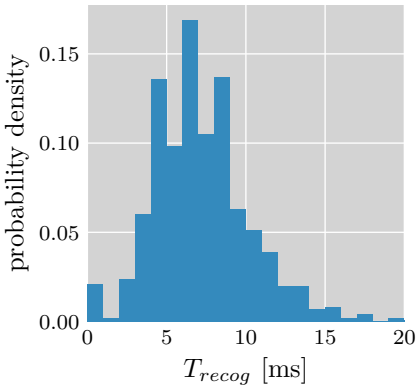
(b) $\epsilon = 199\ \frac{\text{rad}}{\text{s}}$, 2 random, head



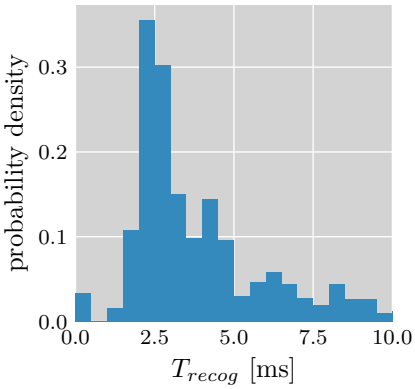
(c) $\epsilon = 483\ \frac{\text{rad}}{\text{s}}$, 2 random, head



(d) $\epsilon = 778\ \frac{\text{rad}}{\text{s}}$, 2 random, head

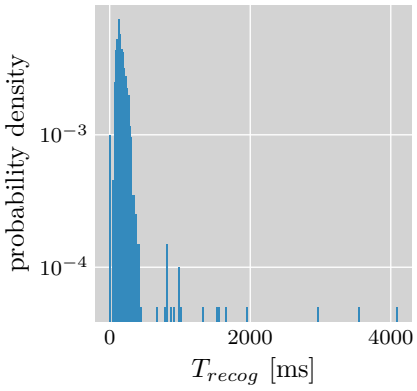


(e) $\epsilon = 1608\ \frac{\text{rad}}{\text{s}}$, 2 random, head

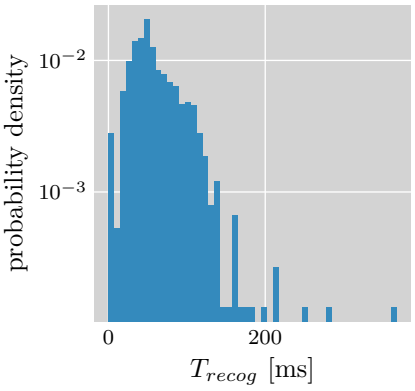


(f) $\epsilon = 4730\ \frac{\text{rad}}{\text{s}}$, 2 random, head

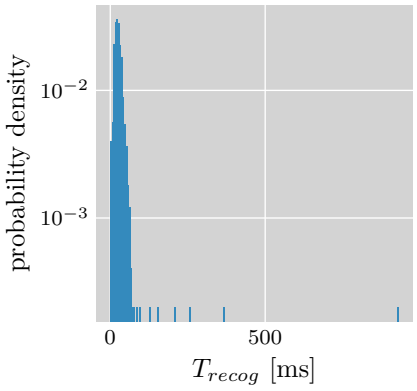
Figure F.7: Two random patterns: Recognition time distributions heads. 1000 measurements were conducted for each histogram shown. Each distribution has a sparsely populated tail which is not shown.



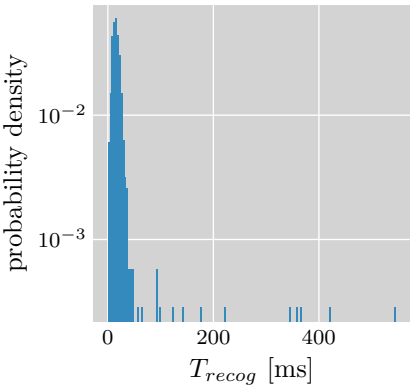
(a) $\epsilon = 48{,}3\ \frac{\text{rad}}{\text{s}}$, 2 random, full



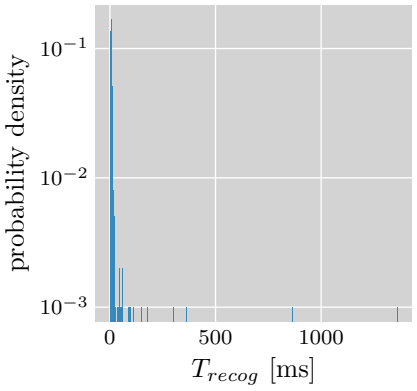
(b) $\epsilon = 199\ \frac{\text{rad}}{\text{s}}$, 2 random, full



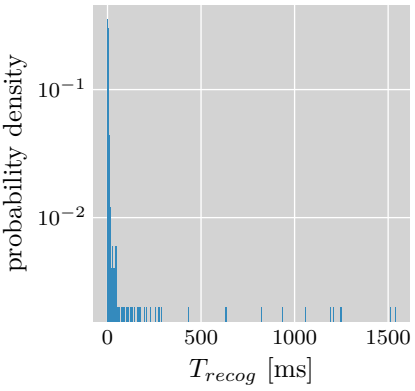
(c) $\epsilon = 483\ \frac{\text{rad}}{\text{s}}$, 2 random, full



(d) $\epsilon = 778\ \frac{\text{rad}}{\text{s}}$, 2 random, full



(e) $\epsilon = 1608\ \frac{\text{rad}}{\text{s}}$, 2 random, full



(f) $\epsilon = 4730\ \frac{\text{rad}}{\text{s}}$, 2 random, full

Figure F.8: Two random patterns: Full recognition time distributions, probability axis log-scaled for better visibility of long tails. 1000 measurements were conducted for each histogram shown.

Bibliography

[1] John Buck. »Synchronous rhythmic flashing of fireflies. II.« In: *The Quarterly review of biology* 63.3 (1988), pp. 265–289.

[2] Christiaan Huygens and Horologium Oscillatorium. »The pendulum clock«. In: *Trans RJ Blackwell, The Iowa State University Press, Ames* (1986).

[3] Steven H Strogatz, Richard E Kronauer and Charles A Czeisler. »Circadian regulation dominates homeostatic control of sleep length and prior wake length in humans«. In: *Sleep* 9.2 (1986), pp. 353–364.

[4] Steven H Strogatz, Richard E Kronauer and Charles A Czeisler. »Circadian pacemaker interferes with sleep onset at specific times each day: role in insomnia«. In: *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 253.1 (1987), R172–R178.

[5] Arkady Pikovsky, Michael Rosenblum and Jürgen Kurths. *Synchronization: a universal concept in nonlinear sciences*. Vol. 12. Cambridge university press, 2003.

[6] Grigory V Osipov, Jürgen Kurths and Changsong Zhou. *Synchronization in oscillatory networks*. Springer Science & Business Media, 2007.

[7] Susanna C Manrubia, Alexander S Mikhailov et al. *Emergence of dynamical order: synchronization phenomena in complex systems*. Vol. 2. World Scientific, 2004.

[8] Stefano Boccaletti. *The synchronized dynamics of complex systems*. Vol. 6. Elsevier, 2008.

[9] Andrzej Stefański. *Determining Thresholds of Complete Sychronization and Application*. Vol. 67. World Scientific, 2009.

[10] Takashi Nishikawa, ed. *Synchronization and Cascading Processes in Complex Networks [Focus Issue]*. Vol. 21. 2. AIP Publishing, 2011.

[11] Johan AK Suykens and Grigory V Osipov. *Synchronization in complex networks [Focus Issue]*. Vol. 18. 3. AIP Publishing, 2008.

[12] Juergen Kurths et al., eds. *Control and synchronization in chaotic dynamical systems [Focus Issue]*. Vol. 13. 1. AIP Publishing, 2003.

[13] Ramon Guevara Erra, Jose Luis Perez Velazquez and Michael Rosenblum. »Neural synchronization from the perspective of nonlinear dynamics«. In: *Frontiers in computational neuroscience* 11 (2017), p. 98.

[14] Francisco Varela et al. »The brainweb: phase synchronization and large-scale integration«. In: *Nature reviews neuroscience* 2.4 (2001), p. 229.

[15] Klaus Lehnertz et al. »Synchronization phenomena in human epileptic brain networks«. In: *Journal of neuroscience methods* 183.1 (2009), pp. 42–48.

[16] Frank C Hoppensteadt and Eugene M Izhikevich. *Weakly connected neural networks*. Vol. 126. Springer Science & Business Media, 2012.

[17] Egill Masson and Yih-Jeou Wang. »Introduction to computation and learning in artificial neural networks«. In: *European Journal of Operational Research* 47.1 (1990), pp. 1–28.

[18] David E Rumelhart, Geoffrey E Hinton and Ronald J Williams. »Learning representations by back-propagating errors«. In: *Nature* 323.6088 (1986), p. 533.

[19] Sepp Hochreiter and Jürgen Schmidhuber. »Long short-term memory«. In: *Neural computation* 9.8 (1997), pp. 1735–1780.

[20] Diederik P. Kingma and Jimmy Ba. »Adam: A Method for Stochastic Optimization«. In: *CoRR* abs/1412.6980 (2014). arXiv: 1412.6980. URL: http://arxiv.org/abs/1412.6980.

[21] Thomas C Jackson et al. »Oscillatory Neural Networks Based on TMO Nano-Oscillators and Multi-Level RRAM Cells«. In: *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on* 5.2 (2015), pp. 230–241.

[22] Thomas Barois et al. »Ultra low power consumption for self-oscillating nano-electromechanical systems constructed by contacting two nanowires«. In: *Nano letters* 13.4 (2013), pp. 1451–1456.

[23] V V Perminov et al. »Synchronization in the system of coupled oscillators based on VO2 switches«. In: *Journal of Physics: Conference Series*. Vol. 929. 1. IOP Publishing. 2017, p. 012045.

[24] John J Hopfield. »Neural networks and physical systems with emergent collective computational abilities«. In: *Proceedings of the National Academy of Sciences* 79.8 (1982), pp. 2554–2558.

[25] Juan A. Acebrón et al. »The Kuramoto model: A simple paradigm for synchronization phenomena«. In: *Rev. Mod. Phys.* 77 (1 Apr. 2005), pp. 137–185.

[26] Toru Aonishi. »Phase transitions of an oscillator neural network with a standard Hebb learning rule«. In: *Physical Review E* 58.4 (1998), p. 4865.

[27] Toshio Aoyagi and Katsunori Kitano. »Effect of random synaptic dilution in oscillator neural networks«. In: *Physical Review E* 55.6 (1997), p. 7424.

[28] Alex Arenas and Conrad J Pérez Vicente. »Exact long-time behavior of a network of phase oscillators under random fields«. In: *Physical review E* 50.2 (1994), p. 949.

[29] Frank C. Hoppensteadt and Eugene M. Izhikevich. »Oscillatory Neurocomputers with Dynamic Connectivity«. In: *Phys. Rev. Lett.* 82 (14 Apr. 1999), pp. 2983–2986.

[30] Robert W Hölzel and Katharina Krischer. »Pattern recognition with simple oscillating circuits«. In: *New Journal of Physics* 13.7 (2011), p. 073031.

[31] Kathrin Kostorz, Robert W Hölzel and Katharina Krischer. »Distributed coupling complexity in a weakly coupled oscillatory network with associative properties«. In: *New Journal of Physics* 15.8 (2013), p. 083010.

[32] Takashi Nishikawa, Frank C Hoppensteadt and Ying-Cheng Lai. »Oscillatory associative memory network with perfect retrieval«. In: *Physica D: Nonlinear Phenomena* 197.1 (2004), pp. 134–148.

[33] Rosangela Follmann et al. »Phase oscillatory network and visual pattern recognition«. In: *Neural Networks and Learning Systems, IEEE Transactions on* 26.7 (2015), pp. 1539–1544.

[34] Donald Olding Hebb. *The organization of behavior: A neuropsychological theory*. Psychology Press, 2005.

[35] Robert W. Hölzel and Katharina Krischer. »Stability and Long Term Behavior of a Hebbian Network of Kuramoto Oscillators«. In: *SIAM Journal on Applied Dynamical Systems* 14.1 (2015), pp. 188–201.

[36] Steven H Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Westview press, 2014.

[37] Paul Horowitz and Winfield Hill. *The art of electronics*. Cambridge Univ. Press, 1989.

[38] E. M. Izhikevich and B. Ermentrout. »Phase model«. In: *Scholarpedia* 3.10 (2008). revision #91649, p. 1487. DOI: 10.4249/scholarpedia.1487.

[39] K. Josic, E. T. Shea-Brown and J. Moehlis. »Isochron«. In: *Scholarpedia* 1.8 (2006). revision #137879, p. 1361. DOI: 10.4249/scholarpedia.1361.

[40] Y. A. Kuznetsov. »Andronov-Hopf bifurcation«. In: *Scholarpedia* 1.10 (2006). revision #90964, p. 1858. DOI: 10.4249/scholarpedia.1858.

[41] Steven H Strogatz. »From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators«. In: *Physica D: Nonlinear Phenomena* 143.1-4 (2000), pp. 1–20.

[42] P. Holmes and E. T. Shea-Brown. »Stability«. In: *Scholarpedia* 1.10 (2006). revision #137538, p. 1838. DOI: 10.4249/scholarpedia.1838.

[43] A Antoniou. »New gyrator circuits obtained by using nullors«. In: *Electronics Letters* 4.5 (1968), pp. 87–88.

[44] T. Kanamaru. »Van der Pol oscillator«. In: *Scholarpedia* 2.1 (2007). revision #138698, p. 2202. DOI: 10.4249/scholarpedia.2202.

[45] Robert Hölzel. »A Neural Network of Weakly Coupled Nonlinear Oscillators with a Global, Time-Dependent Coupling-Theory and Experiment«. PhD thesis. Technical University of Munich, 2013.

[46] Yoshiki Kuramoto. *Chemical oscillations, waves, and turbulence*. Vol. 19. Springer Science & Business Media, 2012.

[47] Ferdinand Verhulst. *Nonlinear Differential Equations and Dynamical Systems*. Hochschultext / Universitext. Springer Berlin Heidelberg, 1996. ISBN: 9783540609346.

[48] Solomon W Golomb. »The use of combinatorial structures in communication signal design«. In: *Applications of Combinatorial Mathematics*. Vol. 60. Institute of Mathematics and its Applications Conference Series. 1997, pp. 59–78.

[49] Mike D Atkinson and Anne-Lise Hassenklover. »Sets of integers with distinct differences«. In: *Rep. SCS-TR-63*. School of Comp. Sci., Carleton Univ Ottawa, Ont., Canada, 1984.

[50] Toru Aonishi, Koji Kurata and Masato Okada. »Acceleration effect of coupled oscillator systems«. In: *Phys. Rev. E* 65 (4 Apr. 2002), p. 046223.

[51] *As Readout in (IIA) and (IIB) only measures changes between initial values and thus the final pattern can be constructed non-ambiguously from the initial defective pattern, the ambiguity introduced in the setting of initial values is effectively removed.*

[52] Toshio Aoyagi and Katsunori Kitano. »Retrieval dynamics in oscillator neural networks«. In: *Neural computation* 10.6 (1998), pp. 1527–1546.

[53] J Cook. »The mean-field theory of a Q-state neural network model«. In: *Journal of Physics A: Mathematical and General* 22.12 (1989), p. 2057.

[54] Satoki Uchiyama and Hirokazu Fujisaka. »Stability of oscillatory retrieval solutions in the oscillator neural network without Lyapunov functions«. In: *Physical Review E* 65.6 (2002), p. 061912.

[55] Robert J McEliece et al. »The capacity of the Hopfield associative memory«. In: *Information Theory, IEEE Transactions on* 33.4 (1987), pp. 461–482.

[56] Stefan Litzel. »Synchronisation und Entropieproduktion von Oszillatoren jenseits von Kuramoto-Kopplung«. Bachelor's Thesis. Technical University of Munich, 2015.

[57] Damir Vodenicarevic, Nicolas Locatelli and Damien Querlioz. »A neural network based on synchronized pairs of nano-oscillators«. In: *Nanotechnology (IEEE-NANO), 2017 IEEE 17th International Conference on*. IEEE. 2017, pp. 512–514.

[58] Nicolas Pfifferling. »Eine neuartige Implementierung von gekoppelten Oszillatorpaaren zur Mustererkennung«. MA thesis. Technical University of Munich, 2017.

[59] The OEIS Foundation Inc 2011. *The online encyclopedia of integer sequences*. [Online; accessed 21-April-2018]. 2018. URL: `http://oeis.org/A005836`.

[60] William Shockley. »The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors«. In: *Bell Labs Technical Journal* 28.3 (1949), pp. 435–489.

[61] *Our measurements of the phase response curves do not match previous measurements of phase response curves in our group: In [31],* $A_{PRC} \approx 1{,}75\,\frac{\mathrm{rad}}{\mathrm{V}}$ *at* $61{,}5\,\mathrm{kHz}$ *and in [45],* $A_{PRC} \approx 2{,}55\,\frac{\mathrm{rad}}{\mathrm{V}}$ *at* $67\,\mathrm{kHz}$*. Examination of old measurement software in our group revealed an erroneous factor* $\sqrt{2}$ *and deviations due to a parasitic capacitance are not mentioned in either [45] or [31]. Hoelzel's result matches our result if we correct for the erroneous factor and correct for a parasitic capacitance of similar size. Kostorz might have either manually corrected the factor* $\sqrt{2}$ *or measured* U *directly with a soldered connector, thus modifying the capacitance of the oscillator. Comparison with our own measurements suggests both explanations are equally viable and explain the published values.*

[62] John W Tukey. *Exploratory data analysis*. Vol. 2. Reading, Mass., 1977.

[63] David C Stone. »Application of median filtering to noisy data«. In: *Canadian Journal of chemistry* 73.10 (1995), pp. 1573–1581.

[64] N Gallagher and Gı Wise. »A theoretical analysis of the properties of median filters«. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 29.6 (1981), pp. 1136–1141.

[65] Alexander Sparber. »Einfluss starker Kopplung auf ein Oszillatornetzwerk zur Mustererkennung«. Bachelor's Thesis. Technical University of Munich, 2015.

[66] Daniel J Amit. *Modeling brain function: The world of attractor neural networks*. Cambridge university press, 1992.

[67] Steven J. Wilson. *Confidence Intervals for Percentiles and Medians*. [Online; accessed 30-April-2018]. URL: http://www.milefoot.com/math/stat/ci-medians.htm.