Master's thesis

NTNU
Norwegian University of Science and Technology

Jens Anker Lyngaas

# Neuromorphic computing with ferromagnetic nano-oscillators

Master's thesis in MSPHYS
Supervisor: Alireza Qaiumzadeh
Co-supervisor: Verena Johanna Brehm
May 2023

NTNU
Norwegian University of
Science and Technology

Jens Anker Lyngaas

# Neuromorphic computing with ferromagnetic nano-oscillators

Master's thesis in MSPHYS
Supervisor: Alireza Qaiumzadeh
Co-supervisor: Verena Johanna Brehm
May 2023

Norwegian University of Science and Technology

NTNU
Norwegian University of
Science and Technology

# Abstract

In recent years, neural networks have risen to prominence through the field of machine learning. These networks are primarily implemented on commercially-available transistor-based devices, however such devices are not well-suited for emulating the brain in the way neural networks desire. Here is a natural intersection with the field of neuromorphic computing, or computing inspired by the human brain.

This thesis employs micromagnetic computer simulations to investigate the properties and synchronization of ferromagnetic oscillators. The oscillators are driven by the spin-Hall effects, and contained within a 2D plane using magnetic anisotropy. Once characterized, the synchronization and desynchronization are used to create an oscillatory neural network, drawing from the macroscopic oscillations within the brain. This neural network is additionally modeled with the Kuramoto model, thereby opening for it to be interfaced with existing training algorithms.

Finally, investigation of the voltage response to (de)synchronization reveals that global synchronization and the order parameter of the system can be determined by voltage fluctuations alone. This is of great use for experimental realizations of this device.

The presented oscillator is an extension of a ferromagnetic spiking neuron, being materially similar with only modifications to the electric setup.

# Sammendrag

I løpet av de siste årene har nevrale nettverk blitt fremtredende gjennom maskinlæring. Disse er ofte implementert på kommersielt tilgjengelige, transistorbaserte enheter, som ikke er godt egnet til å etterligne hjernen slik nevrale nettverk ønsker. Her er det et naturlig skjæringspunkt med nevromorfisk prosessering, eller prosessering inspirert av menneskehjernen.

Denne masteroppgaven benytter mikromagnetiske datasimuleringer for å undersøke egenskapene til, og synkroniseringen av, ferromagnetiske oscillatorer. Oscillasjonene drives av spin-Hall-effektene og begrenses til et 2D-plan ved hjelp av magnetisk anisotropi. Etter at synkroniseringen og desynkroniseringen er begge karakterisert, brukes de til å skape et oscillatorisk nevralt nettverk inspirert av de makroskopiske oscillasjonene i hjernen. Det kompliserte magnetiske systemet forenkles ved bruk av Kuramoto-modellen, og dermed åpnes det for bruk av eksisterende eksisterende treningsalgoritmer for maskinlæring.

Undersøkelse av spenningsresponsen til (de)synkronisering viser både global synkronisering og systemets ordenparameter kan bestemmes via forandring i spenningen. Dette er til stor nytte for eksperimentelle realiseringer av disse ferromagnetiske enhetene.

Den presenterte oscillatoren er en utvidelse av et kunstig ferromagnetisk nevron, og er materielt lik, dog med endringer i det elektriske oppsettet.

# Preface

This thesis is the final part of a two-year Master of Science in Physics program at the Norwegian University of Science and Technology (NTNU), lasting from the fall of 2022 to the spring of 2023. It has been carried out at the Center of Quantum Spintronics (QuSpin) under the supervision of Dr. Alireza Qaiumzadeh and the co-supervision of Verena Johanna Brehm.

I would like to extend my gratitude to Alireza and Verena for their excellent help and ideas with every aspect of this project. The pride I have in my work today would not be without their invaluable input. My deepest thanks also go to Dr. Serban Lepadatu, whose help with his software, BORIS computational spintronics, was pivotal in realizing this project. Finally, to my family, close and extended, home and abroad, for their immense support all these years.

Jens Anker Lyngaas
Trondheim, Norway
May 2023

# Contents

# 1 Glossary

- **AFM** - Antiferromagnet/-magnetism/-magnetic
- Demag - Demagnetization (demagnetizing field)
- **EP-SHNO** - Easy-Plane Spin Hall nano-oscillator
- **FM** - Ferromagnet/-magnetism/-magnetic
- **HM** - Heavy Metal
- **IP** - In-plane
- **(i)SHE** - (Inverse) Spin Hall effect
- **(i)SHA** - (Inverse) Spin Hall Angle
- **LLG** - Landau–Lifshitz–Gilbert
- **LLG-SA** - Landau–Lifshitz–Gilbert Spin Accumulation
- **ONN** - Oscillatory neural network
- **OOP** - Out-of-plane
- **SHNO** - Spin Hall nano-oscillator
- **sLLG** - stochastic Landau–Lifshitz–Gilbert
- **SMR** - Spin Hall Magnetoresistance
- **SNN** - Spiking neural network

# 2 Introduction

After nearly 60 years, Moore's law may still apply [1]. The law is an industry observation stating that the number of transistors in an integrated circuit doubles every two years, by the transistors themselves getting smaller. But it may not be long for this world [2]: As transistors come closer to atomic scales, issues begin to emerge. Quantum tunneling between the transistors becomes a larger problem [3], as does the thermal effects of an increasing power density, and many more [3]. Even if contemporary transistors were not pushing up against the limits of Moore's law, there are several other drawbacks to modern computer techniques: Most computers have the processing occurring in a different location from the memory [4], constraining the speed to that of the data transfer between these units. This is the so-called von Neumann bottleneck [5]. In addition, though modern computers are parallelized in the sense that there can be multiple processors operating at once, it is nevertheless an inherently serial operation: One instruction at a time [5].

The brain is able to perform all the necessary functions for life alongside real-time image recognition, abstract thinking, and the formulation of sentences, all at once and yet also in synchronicity. This is at a fraction of the power: AlphaGo, the first program able to beat a professional human player in a game of Go, required more than 1 MW of power [6]. The human brain requires about 20 W [7]. The gulf in efficiency is huge, and creating computers operating at human levels of efficiency is of great interest.

Neuromorphic computing is a field of computer science that aims to emulate the brain in either computer hardware or software, in order to increase the efficiency of computing [5]. In many cases, this comes in the form of specialized computer chips, but it does not have to: Artificial neural networks are a form of neuromorphics and are presently implemented on conventional computer components. However, creating and using these networks is computationally intensive in both time and resources [8], and they often employ clusters of graphical processing units (GPUs) in order to parallelize their operation as much as possible [6][9]. By mimicking the brain on the hardware itself, one opens for much greater efficiency in performing tasks where the brain outperforms a computer [5]. Computer vision is one such task. Unfortunately, transistors are not suited for the creation of neuromorphic devices, requiring potentially thousands of them for a single neurosynaptic unit [10]. Researchers have therefore been searching for other paradigms where neuromorphic computing can be realized [11].

Spintronics is a field of quantum mechanics that investigates and exploits the spin of an electron in addition to its charge. This is fundamentally different from modern electronics, and it opens for improvements on existing technology in addition to the creation of entirely new devices. Spintronics has already been used to create neuromorphic devices [12][13][11]. One example is the use of spintronic devices to create artificial neurons in a spiking neural network, relying on behavior very similar to that of the biological neurons within the brain

[14]. However, there are many other avenues in spintronic neuromorphics.

## 2.1 Outline

This thesis will first introduce four core concepts: Ferromagnetism and spin Hall effects (Section 3), the Kuramoto model (Section 4), neuromorphic computing Section 5, and the basics of neural networks (Section 5). An implementation for an easy-plane spiking ferromagnetic neuron is presented in Section 6. Sections 7 and 8 cover the implementation of an easy-plane spin Hall nano-oscillator (EP-SHNO) in a chain, studying the synchronization and controllability thereof. Section 9 shows the potential of an EP-SHNO in an oscillatory neural network, in line with existing implementations thereof.

# 3 Ferromagnetic materials

This chapter will introduce core concepts from both ferromagnetism and spintronics that partake in the creation of spin-Hall devices.

## 3.1 Contributions to the FM free energy

To study a ferromagnetic system, and more specifically the ground state, we need to investigate the contributions to its potential energy. Understanding the various energy contributions is also central to deriving the equations of motion inside a ferromagnetic system. We will first consider the atomistic expressions for these terms, and use these to derive expressions in the continuum (micromagnetic) limit.

Note that when we use the term 'spin', we are not referring to individual electron spins but rather the overall magnetic moment at a given lattice site.

### 3.1.1 Exchange coupling

The exchange interaction within a magnetic material is usually the most dominant interaction and an important contributor to the potential energy. Between individual spin lattice sites it can be described by the Heisenberg model [15] with the following Hamiltonian:

$$\mathcal{U} = \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j \qquad (i \neq j) \tag{3.1}$$

Here, $J_{ij}$ is a material constant and the $\mathbf{S}$-vectors are the spins at sites $i$ and $j$. This is a general form that takes into account coupling between all lattice sites. Within the isotropic Heisenberg model, it is often assumed that $J_{ij}$ is constant when $i, j$ are its nearest neighbors, and zero otherwise.

$$J_{ij} = \begin{cases} J & i, j \text{ neighbors} \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

When including next-nearest neighbors and beyond, $J_{ij}$ will normally become weaker for neighbors further apart. We will assume nearest-neighbor coupling unless otherwise specified. This equation states that the interaction between the spins is governed by a dot product and will energetically prefer a parallel or antiparallel alignment, depending on the sign of $J$. For $J > 0$, the energy will be minimized by $\mathbf{S}_i \parallel \mathbf{S}_j$, which is the ferromagnetic case. And conversely, $J < 0$ is an antiferromagnet. Some sources write $-\sum J_{ij}$, which switches this sign convention.

### 3.1.2 Magnetic anisotropy

With the Heisenberg model, it is assumed that the system is entirely isotropic, meaning that for an individual spin there is no energetically preferential orientation in space. Magnetizing a Heisenberg ferromagnet to an arbitrary direction is therefore equally as costly. However, real magnetic materials are not isotropic, and certain orientations of the magnetization are more costly than others [16].

The most important anisotropic contribution is the magnetocrystalline anisotropy, which arises from spin-orbit interaction: The electron orbitals are linked to the crystallographic structure. Due to their interaction with the spins, they make the latter prefer to align along a well-defined crystallographic axis (in the case of *uniaxial* anisotropy) or axes (in the case of many-axial anisotropy) [16][17]. While usually weaker than the exchange interaction, this contribution is important as it has a defined spatial direction, whereas the exchange coupling only seeks to align the spins with respect to each other [17].

We define a normalized magnetization in 3D space

$$\mathbf{m} = \frac{\mathbf{M}}{|\mathbf{M}|} = (\alpha_1, \alpha_2, \alpha_3)$$

$$\alpha_1 = \sin\theta\cos\phi \qquad \alpha_2 = \sin\theta\sin\phi \qquad \alpha_3 = \cos\theta$$

The energy can be considered a power series expansion of the magnetization components [17],

$$E = E_0 + \sum_i b_i\alpha_i + \sum_{ij} b_{ij}\alpha_i\alpha_j + \sum_{ijk} b_{ijk}\alpha_i\alpha_j\alpha_k + \mathcal{O}(\alpha^5) \tag{3.3}$$

where $b$ is the anisotropic interaction strength. As the energy is a scalar quantity, we can invert the system and leave the energy unchanged, $E(\mathbf{m}) = E(-\mathbf{m})$ and thus $E(\alpha_i) = E(-\alpha_i)$. Thus, all odd powers in the power series disappear [17]. We discard the higher-order $\mathcal{O}(\alpha^5)$ terms, and adjust the zero-point energy to get a more convenient expression:

$$E = \sum_{ij} b_{ij}\alpha_i\alpha_j \tag{3.4}$$

From a micromagnetic perspective, this can instead be expressed as

$$\mathcal{U}_{\text{ani}} = -K\sum_i (\mathbf{S}_i \cdot \hat{\mathbf{e}})^2 = -K\sum_i \mathbf{S}_i^2 \cos^2\theta, \tag{3.5}$$

where $K$ is a material constant and $\hat{\mathbf{e}}$ is the unit direction of the anisotropy. The latter equation simply used the definition of the dot product via angles between vectors. The fact that this quantity is squared means that $\cos^2\theta$ is minimized for $\theta = \pm\pi/2$ and maximized for $\theta = 0$ *and* $\theta = \pi$. The sign of $K$ decides which of these are energetically less expensive.

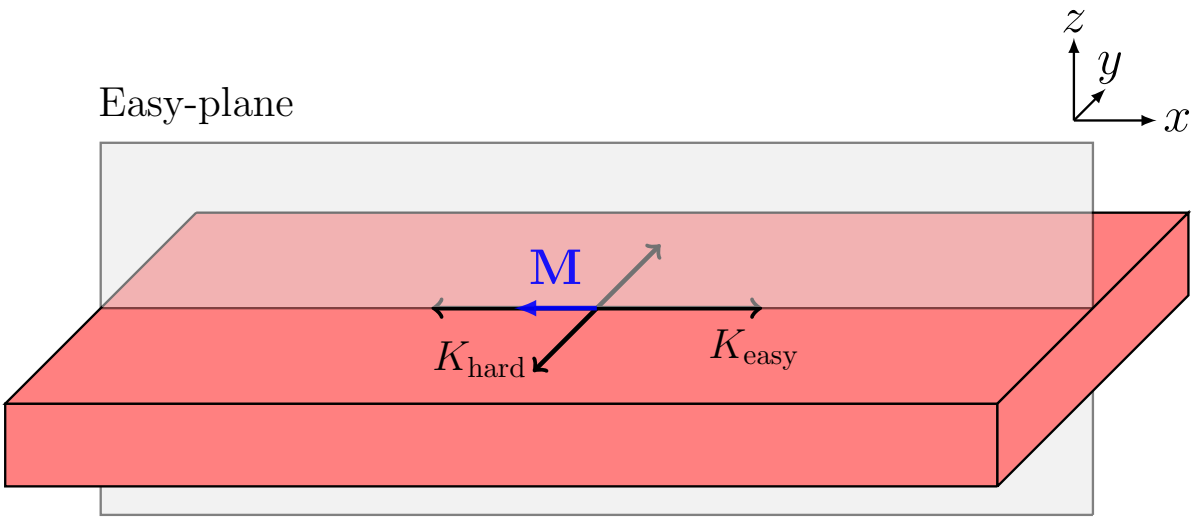



Figure 3.1: Example setup for a hard-axis anisotropy. Red mesh is a ferromagnet, $K_{\text{hard}} \parallel \hat{\mathbf{y}}$ is the hard-axis anisotropy direction forming a normal vector to the easy-plane, $K_{\text{easy}} \parallel \hat{\mathbf{x}}$ is the easy-axis within this easy-plane, and $\mathbf{M}$ is the magnetization is in one of its ground states.

For the most part, we will consider a version of (3.5), written on the form:

$$\mathcal{U}_{\text{ani}} = -K_{\text{easy}}(\mathbf{S} \cdot \hat{\mathbf{e}}_{\parallel})^2 - K_{\text{hard}}(\mathbf{S} \cdot \hat{\mathbf{e}}_{\perp})^2 \tag{3.6}$$

This form of anisotropy creates an energetically expensive *hard-axis*, with strength $K_{\text{hard}} < 0$, which is the normal vector for an energetically inexpensive *easy-plane* composed of a normal axis and an *easy-axis*. "Normal", in this case, means there is no associated energy correction with that axis, while the easy-axis has a strength of $K_{\text{easy}} > 0$. In our case, the absolute value of $K_{easy}$ is orders of magnitude smaller than $K_{hard}$. As a result of this, the magnetization is strongly discouraged from existing outside of the easy-plane, and within said plane, it prefers to exist along the easy-axis. This is exemplified in Figure 3.1.

### 3.1.3 External magnetic field (Zeeman coupling)

The Zeeman effect is the result of coupling between the spins and an *external* magnetic field. The expression of the energy contribution is

$$\mathcal{U}_{\text{Zeeman}} = -\mu_B \sum_i \mathbf{B}(\mathbf{r}, t) \cdot \mathbf{S}_i \tag{3.7}$$

where $\mu_B = 4\pi \times 10^{-7} \text{J/A}^2\text{m}$ is the magnetic field constant and $\mathbf{B}$ is the external magnetic field. The coupling is then summed over all lattice sites [18].

As $\mu_B$ is a defined positive quantity, the energy is minimized when $\mathbf{B}$ and $\mathbf{S}$ are aligned. In contrast to the anisotropy, the dot product in the Zeeman term is not quadratic and

therefore specifies a preferential direction for the spins.

### 3.1.4 Dipolar interaction and demagnetization

Even though a sample may not contain any inherent magnetocrystalline anisotropy as described in Section 3.1.2, it may still have a preferred direction that arises from the shape of the sample, through stray (external) and demagnetizing (internal) fields [19].

For a magnetic dipole in an external magnetic field, the dipole's magnetic field can be expressed as

$$\mathbf{B} = \frac{\mu_0}{4\pi r^3}[3(\mathbf{m}\cdot\hat{\mathbf{r}})\cdot\hat{\mathbf{r}} - \mathbf{m}] \tag{3.8}$$

Combining this with the energy of a dipole in a magnetic field, $E = -\mathbf{m}\cdot\mathbf{B}$, we consider the case where the magnetic field is caused by a nearby magnetic dipole, such that $\mathbf{m}_{i/j}$ represent the value of the dipole located at site $r_{i/j}$, and $\mathbf{R}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ [17]. Expanding the $\hat{\mathbf{r}}$'s, we get

$$E_{\mathrm{dip}}^{i,j} = -\frac{1}{4\pi\mu_{\mathrm{o}}}\frac{3\mathbf{m}_{\mathrm{i}}\cdot\mathbf{R}_{\mathrm{ij}}\mathbf{m}_{\mathrm{j}}\cdot\mathbf{R}_{\mathrm{ij}} - \mathbf{m}_{\mathrm{i}}\cdot\mathbf{m}_{\mathrm{j}}R_{\mathrm{ij}}^2}{R_{\mathrm{ij}}^5} \tag{3.9}$$

as an expression for the energy caused by the ij-dipolar interaction.

The surface of a finite sample exhibits magnetic dipoles which causes a stray field outside the sample. In turn, this stray field induces an internal magnetic field, called the demagnetization field [17]. We are most interested in the demagnetization field, however equation (3.9) is cumbersome and does not readily capture the full effect of either of these fields. We therefore introduce an explicit demagnetizing field, $\mathbf{H}_{\mathrm{demag}}$, and write the total energy due to self-interaction as [15]

$$E_{\mathrm{demag}} = -\frac{1}{2}\mu_0\int \mathbf{m}\cdot\mathbf{H}_{\mathrm{demag}}\mathrm{d}V \tag{3.10}$$

Note that this quantity operates in the continuous limit by it being an integral over the whole space rather than a summation over every lattice site. Regardless, the effect is still the same: This energy is minimized when the magnetization is aligned with the demagnetizing field. This field is dependent upon the shape and properties of the sample in question.

As it is shape-dependent, there is no succinct way of expressing the demagnetizing field for all geometries [15].

Though normally small, many ferromagnets rely entirely on the demagnetization to provide the anisotropy. Furthermore, demagnetization is a long-distance effect compared to the exchange coupling, and therefore it remains important to this work.

## 3.2 Continuum limit for the spin lattice

In some cases, it is useful to treat a material as though it has a continuous magnetization $\mathbf{m}$ rather than with individual, localized spin $\mathbf{S}_i$. Note that this $\mathbf{m}$ is not the same as $m_i$ from earlier. This has already been performed for the demagnetization interaction, but we wish to extend this treatment to the most relevant potential energy contributions for our work.

To perform this change, some tricks are employed. We will proceed term-by-term in the free energy.

### 3.2.1 Continuum exchange interaction

Recall that we expressed the exchange interaction as

$$\mathcal{U} = \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j \qquad (i \neq j) \tag{3.11}$$

To obtain the Hamiltonian in the continuum model, we only consider nearest-neighbor interactions,

$$\mathcal{U} = J \sum_{i} \mathbf{S}_i \cdot \mathbf{S}_{i+1} \tag{3.12}$$

which is a reasonable assumption as the strength of the coupling $J_{ij}$ is expected to decay with distance. This is the Heisenberg model, described previously.

We express the dot product term by term,

$$\mathcal{U} = J \sum_{i} S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + S_i^z S_{i+1}^z \tag{3.13}$$

and use the relation $(S_i^x - S_{i+1}^x)^2 = (S_i^x)^2 - 2S_i^x S_{i+1}^x + (S_{i+1}^x)^2$ to rewrite it as

$$\mathcal{U} = -\frac{J}{2} \sum_{i} \Big( (S_{i+1}^x - S_i^x)^2 + (S_{i+1}^y - S_i^y)^2 + (S_{i+1}^z - S_i^z)^2 \ - |S_i|^2 - |S_{i+1}|^2 \Big)$$

We assume all spins are of constant length, such that $|S_j| = S$, and can therefore remove the last two terms by adjusting the zero-point energy. Additionally, we introduce the notation $\mathbf{s} = \mathbf{S}/S$, where $\mathbf{s}$ is a unit vector and S is the magnitude of the spin. Thus we can write this as

$$\mathcal{U} = -JhS \sum_{i} \Big( \frac{(s_{i+1}^x - s_i^x)^2}{h} + \frac{(s_{i+1}^y - s_i^y)^2}{h} + \frac{(s_{i+1}^z - s_i^z)^2}{h} \Big) \tag{3.14}$$

where $h$ is the lattice spacing. Now we take this to the continuum limit by setting $h \to 0$, which yields

$$\mathcal{U} = J \int dV (\nabla \mathbf{m})^2 \tag{3.15}$$

Here we have also transformed the spin to its macroscopic equivalent, the magnetization $\mathbf{m}$.

### 3.2.2 Continuum anisotropy interaction

Recall that the uniaxial anisotropy contributes

$$\mathcal{U}_{\text{ani}} = -K \sum_i (\mathbf{S}_i \cdot \hat{\mathbf{e}})^2 \tag{3.16}$$

to the Hamiltonian, where K is the anisotropic energy. We will be focusing on the uniaxial anisotropy only, as future work will study the combination of an easy-axis and a hard-axis anisotropy, each represented uniaxially.

Without loss of generality, we set the anisotropy to be aligned with the $\hat{\mathbf{z}}$. It is then expressed as

$$\mathcal{U}_{\text{ani}} = -K \sum_i (S_i^z)^2 \tag{3.17}$$

The steps are similar to those for the exchange interaction, however we do not have any cross-terms and so it is much simpler.

$$\mathcal{U}_{\text{ani}} = -K \sum_i (S_i^z)^2 = -KS^2 \sum_i (s_i^z)^2 = -KS^2 h \sum_i \frac{(s_i^z)^2}{h}$$

where h is the lattice spacing and we have used the definition of S and $s_i$ from earlier.

Now taking the continuum limit $h \to 0$, we obtain

$$\mathcal{U}_{\text{ani}} = KS^2 \int dV (m_i^z)^2 \tag{3.18}$$

### 3.2.3 Continuum Zeeman interaction

The continuum version of the Zeeman interaction is very similar to that of the anisotropy interaction presented above. First recall that the Zeeman term contributes

$$\mathcal{U}_{\text{Zeeman}} = -\mu_B \sum_i \mathbf{B} \cdot \mathbf{S}_i \tag{3.19}$$

We write the magnetic field as $\mathbf{B} = \text{B}\hat{\mathbf{e}}_b$, where $\hat{\mathbf{e}}_b$ is a function of both space and time. This, in addition to earlier definitions, transform the above equation into

$$\mathcal{U}_{\text{Zeeman}} = -\mu_B BSh \sum_i \frac{\hat{\mathbf{e}}_b \cdot \mathbf{s}_i}{h} \tag{3.20}$$

where $\mathbf{s}_i$ is defined as above and h is the lattice constant.

We now take the continuum limit and obtain

$$\mathcal{U}_{\text{Zeeman}} = \int dV \mu_B B \hat{\mathbf{e}}_b \cdot \mathbf{m} \tag{3.21}$$

## 3.3 The Landau-Lifshitz-Gilbert equation

Thus far we have considered the contributions to the Hamiltonian in a ferromagnetic system. They will now be used to find the equation of motion for our system, the Landau-Lifshitz-Gilbert (LLG) equation.

The LLG equation is a differential equation that describes the dynamics of the magnetization in both ferromagnetic and antiferromagnetic materials. The equation is as follows [15][20]:

$$\dot{\mathbf{m}} = -\gamma \mathbf{m} \times \mathbf{H} + \alpha \mathbf{m} \times \frac{\partial \mathbf{m}}{\partial t} + \vec{\tau} \tag{3.22}$$

where $\mathbf{m}$ is the magnetization, $\gamma$ is the gyromagnetic ratio, $\mathbf{H}$ is the effective field, $\alpha$ is the Gilbert damping factor, and $\vec{\tau}$ is some external torque applied to the system [21]. This torque has not yet been characterized as it is necessarily system-dependent and arises from a variety of sources. The effective field $\mathbf{H}$ encapsulates all the physics within the systems; all the previously derived energy contributions regarding magnetic effects are related to it by

$$\mathbf{H}(\mathbf{r}, t) = -\frac{1}{\mu_0 M_s} \frac{\delta \mathcal{U}[\mathbf{m}(\mathbf{r}, t)]}{\delta \mathbf{m}}, \tag{3.23}$$

where $\mathcal{U}$ is the free energy of the system as a function of the magnetization, and $M_s$ is the saturation magnetization. Thus, the potential energy terms we derived above are related to the effective field by a functional derivative.

We will give some intuition by breaking the LLG equation down into parts.

### 3.3.1 Rotational term

For the sake of simplicity, assume that the effective field is constant in space and time. This can be the very basic system of a single spin in a uniform magnetic field. For this case we can, without loss of generality, write it as $\mathbf{H} = (0, 0, H)$. We perform the first cross product in equation (3.22), neglecting all other terms, and get

$$\begin{aligned} \dot{m}_x &= -\gamma H m_y \\ \dot{m}_y &= \gamma H m_x \\ \dot{m}_z &= 0 \end{aligned} \tag{3.24}$$

which is immediately recognizable as a rotation in the $xy$-plane. Therefore, if the spin is perturbed from its preferred alignment with the $\hat{\mathbf{z}}$-axis, there is no restoring force and it will begin to rotate in the $xy$-plane with a fixed amplitude, the speed of which is dependent on the size of the perturbation. Otherwise, the spin will remain stationary. Therefore, the first term in the LLG equation results in a spin precession around the direction of the effective field direction $\mathbf{H}$.

### 3.3.2 Damping term

We continue to assume that $\mathbf{H} \parallel \hat{z}$.

The second term itself is rather difficult to handle as it is implicit and contains the expression $\dot{\mathbf{m}} \propto \mathbf{m} \times \dot{\mathbf{m}}$, making the time-derivative of the magnetization depend on itself. We seek to find a more instructive way of expressing this equation by making it explicit.

We start with the LLG equation, (3.22), and for simplicity we drop the torque. We cross the remaining equation with the magnetization,

$$\mathbf{m} \times \dot{\mathbf{m}} = -\gamma \mathbf{m} \times \mathbf{m} \times \mathbf{H} + \alpha \mathbf{m} \times \mathbf{m} \times \dot{\mathbf{m}} \tag{3.25}$$

We focus our attention on the last term in this equation, as we wish to get rid of it. To simplify this, we employ the relation $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$, which yields

$$\mathbf{m} \times (\mathbf{m} \times \dot{\mathbf{m}}) = (\mathbf{m} \cdot \dot{\mathbf{m}})\mathbf{m} - (\mathbf{m} \cdot \mathbf{m})\dot{\mathbf{m}} \ = -\dot{\mathbf{m}} \tag{3.26}$$

where we have also used the relation $\mathbf{m} \cdot \dot{\mathbf{m}} = 0$, under the assumption that the magnitude of $\mathbf{m}$ is constant [21].

Inserting this back into (3.25), we get

$$\mathbf{m} \times \dot{\mathbf{m}} = -\gamma \mathbf{m} \times \mathbf{m} \times \mathbf{H} - \alpha \dot{\mathbf{m}} \tag{3.27}$$

Subsequently, we insert this equation into (3.22)

$$\dot{\mathbf{m}} = -\gamma \mathbf{m} \times \mathbf{H} + \alpha(-\gamma \mathbf{m} \times \mathbf{m} \times \mathbf{H} - \alpha \dot{\mathbf{m}}) + \vec{\tau} \tag{3.28}$$

$$= -\gamma \mathbf{m} \times \mathbf{H} - \alpha\gamma \mathbf{m} \times \mathbf{m} \times \mathbf{H} - \alpha^2 \dot{\mathbf{m}} + \vec{\tau} \tag{3.29}$$

$$(1 + \alpha^2)\dot{\mathbf{m}} = -\gamma \mathbf{m} \times \mathbf{H} - \alpha\gamma \mathbf{m} \times \mathbf{m} \times \mathbf{H} + \vec{\tau}, \tag{3.30}$$

which yields the explicit LLG-equation

$$\dot{\mathbf{m}} = -\gamma_L \mathbf{m} \times \mathbf{H} - \gamma_R \mathbf{m} \times \mathbf{m} \times \mathbf{H} + \vec{\tau} \tag{3.31}$$

where the index $i$ is still dropped, and where we have defined $\gamma_L = \gamma/1 + \alpha^2$ and $\gamma_R = \alpha\gamma/1 + \alpha^2$. This is also known as the Landau-Lifshitz (LL) equation, as it preceded the LLG equation (3.22) by 20 years. [21].

This is a much simpler equation to look at, as the time-dependent term has been isolated to one side of the equation. However, they are mathematically equivalent [22], and we have shown that we can go from one to the other by algebra alone, so we will be using both forms of the LLG equation interchangeably throughout. The equations are practically equivalent when the dissipation energy, that is to say the damping, is small. In this case, we express $\alpha$ as

$$\alpha = \frac{\gamma_R}{\gamma_L} \ll 1 \tag{3.32}$$

and $\alpha^2$ can be neglected in the definitions of $\gamma_L = \gamma$ and $\gamma_R = \alpha\gamma$, the two equations become physically equivalent as well [22].

For instructive purposes, we expand the damping term in (3.31) and get

$$\dot{m}_x = -\gamma_R m_x m_z H_z \tag{3.33}$$

$$\dot{m}_y = -\gamma_R m_y m_z H_z \tag{3.34}$$

$$\dot{m}_z = \gamma_R (m_x^2 + m_y^2) H_z \tag{3.35}$$

There are two equilibria of this double cross product, $\mathbf{m} \parallel \mathbf{H}$ and $\mathbf{m} \parallel -\mathbf{H}$. We have made the same assumption $\mathbf{H} \parallel \hat{\mathbf{z}}$ as above. Analysis of the rotational term, (3.24), showed that a perturbation from $\mathbf{m} \parallel \hat{\mathbf{z}}$ will turn into a rotation around the $\hat{\mathbf{z}}$-axis. The damping term will provide an additional force back to $\hat{\mathbf{z}}$. Together these will cause a spiral towards the $\hat{\mathbf{z}}$-axis [16][23].

### 3.3.3 Torque term

As stated previously, the torque term occurs due to external influences on the system, the nature of which are not contained within the LLG equation. We will discuss possible torques which can affect the system.

## 3.4 Spin torques

The two terms of the LLG equation, (3.22), that we have tackled thus far are contained within the effective field, tied directly to the Hamiltonian via a functional derivative. This subsection will explore the torque term, $\tau$, of the LLG equation (3.22) in more detail.

The torque itself is not provided by any one source, but from any source outside of those contributing to the effective field.

### 3.4.1 Spin-transfer torque

Spin-transfer torque (STT) is a torque occurring in a ferromagnet when a spin-polarized current (e.g. a polarized electron flow) is directed into a magnetic layer. In doing so,

the angular momentum of the polarized current can be transferred into the spins of the magnetic layer, changing the magnetization and potentially flipping it altogether. The torque is mutual, and the polarized electrons are also altered by flowing through the magnet, but we focus solely on the change within the ferromagnet [24].

The effect is usually only seen in very thin ferromagnetic layers. In larger layers, the stronger magnetization often causes the charge current to flip its spin instead.

STT enters the LLG equation via the torque term,

$$\vec{\tau} = \tau_{\parallel} \frac{\mathbf{m} \times (\mathbf{p}_c \times \mathbf{m})}{|\mathbf{p}_c \times \mathbf{m}|} + \tau_{\perp} \frac{\mathbf{p}_c \times \mathbf{m}}{|\mathbf{p}_c \times \mathbf{m}|}, \tag{3.36}$$

where $\mathbf{p}_c$ is the polarization of the spin-polarized charge current passing through the ferromagnetic layer with magnetization $\mathbf{m}$. This equation was derived by John Slonczewski in 1996. Another name for this term, often used in textbooks, is Slonczewski torque [25].

### 3.4.2 Spin-orbit torque

Spin-orbit torque (SOT) differs from the spin-transfer torque by the charge current: In STT, the torque is provided by a spin-polarized charge current passing through the magnetic material. In SOT, the torque arises entirely from spin-current flowing into the ferromagnet [26], which can emerge without the charge current entering the ferromagnet. A way to create such a spin current is with a bilayer consisting of a heavy metal and a ferromagnet, where within the heavy metal there is a time-dependent accumulation of spins near the interface [27]. Such an accumulation can occur via the spin-Hall effect, which will be discussed in detail later.

It is largely agreed that there are two distinct SOTs in a bilayer: Damping-like (DL) and field-like (FL). They enter the LLG equation as the torque term, expressed as follows:

$$\vec{\tau} = -|\tau_{FL}| \mathbf{m} \times \hat{p}_s - |\tau_{DL}| \mathbf{m} \times (\mathbf{m} \times \hat{p}_s) \tag{3.37}$$

where $\hat{p}_s$ is the spin-current polarization (not to be confused with the spin-polarized charge current mentioned above) and $\mathbf{m}$ is the direction of the magnetization. Comparing the DL term to the damping term of the LLG-equation (characterized by the Gilbert damping parameter $\alpha$), the naming becomes evident. As does the field-like term in comparison to the rotational term.

## 3.5 Spin-Hall effects

Two very important features necessary for the implementation of our neuromorphic devices are the spin-Hall effects. We give a detailed introduction.

As spin is an internal degree of freedom much like the electric charge, particles with spin can be characterized as a charge current density and spin current density. This

notion is something we have already touched upon when referring to spin current and its polarization, but which will now be explored further.

Spin is one of the two types of angular momentum in quantum mechanics, and as a result it can be represented as a vector quantity. The flow of spin-carriers it therefore becomes a tensorial quantity, written $\overleftrightarrow{q}$, with components $q_{ij}$, where i and j represent the cartesian coordinates of the carrier flow and the spin flow respectively. As an example, imagine we have an electron density $n$ moving in the $\hat{\mathbf{x}}$-direction with a velocity $v$, which is entirely spin-polarized in the $\hat{\mathbf{y}}$-direction. In this case, the only non-zero component is $q_{xy} = nv$ [28]. Going forward, we will write $\mathbf{q}^{(0)}$ and $\mathrm{q}_{ij}^{(0)}$ to represent the charge and spin current respectively. The governing equations can be written as follows [28]:

$$\mathbf{q}^{(0)} = -\mu n\mathbf{E} - D\nabla n \tag{3.38}$$

$$q_{ij}^{(0)} = -\mu E_i P_j - D\frac{\partial P_j}{\partial x_i}. \tag{3.39}$$

Here, $n$ is the charge current density, $\mu$ and $D$ are the mobility and diffusion constants respectively, $\mathbf{E}$ is an applied electric field and $\mathbf{P}$ is the spin polarization density.

Equation (3.38) is the general drift-diffusion model of electron flow. The polarization of this flow plays no role in this equation, as we expect from classical electrodynamics. Equation (3.39) describes the spin current of polarized electrons. This exists independently of spin-orbit interactions as the spins themselves are simply carried in the electron flow [28]. Here we have assumed that the mobility's dependence on the spin current is small, and that there are no other sources of current [28].

In most cases we will be treating a 1D charge current, $\mathbf{E} = (E_x, 0, 0)$. For simplicity, we also assume that it is entirely polarized in the y-direction, and that the electrons are only allowed to diffuse along the $x$-axis (i.e. purely 1D motion of the charge current). Then (3.39):

$$\mathbf{q}^{(0)} = -\mu n E_x - D\frac{\partial n}{\partial x} \tag{3.40}$$

$$q_{xy}^{(0)} = -\mu E_x P_y - D\frac{\partial P_y}{\partial x}. \tag{3.41}$$

This is a very simplistic case, and in a real system, the electrons may diffuse in all directions even though the field is unidirectional as we assumed here. There may therefore be many non-zero terms in $\overleftrightarrow{q}$. We will come back to this later.

As a flowing current implies an orbit, the spin-orbit interactions couple these two currents. The result is a correction to each equation. For an isotropic material with inversion symmetry, the corrected currents are [28]

$$q_i = q_i^{(0)} + \gamma \epsilon_{ijk} q_{jk}^{(0)} \tag{3.42}$$

$$q_{ij} = q_{ij}^{(0)} - \gamma \epsilon_{ijk} q_k^{(0)}, \tag{3.43}$$

where $\gamma$ is a parameter proportional to the spin-orbit coupling strength (**not** the gyromagnetic ratio), and $\epsilon_{ijk}$ is the Levi-Civita tensor. Note that we have written the charge current $\mathbf{q}$ in component form. As an example, these equations imply that a spin current defined by $q_{zy}$ (flow of spin carriers in the $z$-direction polarized in the $y$-direction) will induce a charge current in the $x$-direction, $q_x = q_x^{(0)} + \gamma q_{zy}^{(0)}$. And inversely, a charge current in the $z$-direction induces spin currents $q_{xy}$. A more illustrative, but unrelated, example is given in Figure 3.2, which highlights the directions of the SHE and iSHE in a conductor.

The following combination of equations (3.38)-(3.39) and (3.42)-(3.43) is phenomenological, however it captures the spin-orbit coupling. We get

$$\mathbf{j}/e = \mu n \mathbf{E} + D \nabla \mathbf{n} + \beta \mathbf{E} \times \mathbf{P} + \delta \nabla \times \mathbf{P} \tag{3.44}$$

$$q_{ij} = -\mu E_i P_j - D \frac{\partial P_j}{\partial x_i} + \epsilon_{ijk} \left( \beta n E_k + \delta \frac{\partial n}{\partial x_k} \right). \tag{3.45}$$

where $\beta = \gamma\mu$ and $\delta = \gamma D$ are constants introduced which satisfy the Einstein relation, and $\mathbf{j} = -e\mathbf{z}$ [28]. The first two terms in each of these equations are the ones we know from equations (3.38-3.39) that fully describe the system in the absence of spin-orbit coupling, while the latter two are the corrections to account for spin-orbit coupling. Each of these terms are rich in new dynamics, and thus we will tackle each in a non-chronological order.

**Anomalous Hall effect**

The term $\beta \mathbf{E} \times \mathbf{P}$ describes the *anomalous Hall effect* [28], which is a direct coupling between the electric field and the magnetization $\mathbf{P}$, contributing to the charge current. In the regular Hall effect, an external magnetic field causes a deflection of the charge current, and subsequently an accumulation of charges at each side of the conductor. In the anomalous Hall effect, the magnetic field is now provided by the orientation of the magnetization. The permanent magnetization of a ferromagnet therefore ensures that this occurs even in the absence of an external magnetic field. It is important to note that this does not stem from the contribution of the magnetization to the magnetic field, but rather a result of spin-orbit interaction [28].

In certain magnetic materials, the anomalous Hall effect is far greater than the regular Hall effect (though strongly temperature-dependent) [29]. For this reason it is sometimes referred to as the extraordinary Hall effect.

### 3.5.1 Spin Hall effect

The last two terms of equation (3.45), $\beta n E_k + \delta \partial n / \partial x_k$ couple the electric field or charge current diffusion gradient to the spin current, and describe the *spin Hall effect* (SHE) [28]. Note the indices: An electric field or a charge diffusion in the **k**-direction will induce a spin current flowing in the **i**-direction polarized in the **j**-direction, that is entirely perpendicular to the direction of the charge current flow. In other words, a charge current induces a spin current transverse to itself. In a finite sample, this current will cause an accumulation of spins near the boundaries, similar to the charge accumulation caused by the regular Hall effect.

For future reference, it is worth considering a specific component of the $q_{ij}$-tensor, namely the $q_{zy}$, that is a spin current in the $\hat{\mathbf{z}}$-direction polarized in the $\hat{\mathbf{y}}$-direction. Expressing this in full yields

$$q_{zy} = -\mu E_z P_y - D\frac{\partial P_y}{\partial x} + \epsilon_{zyx}\Big(\beta n E_x + \delta \frac{\partial \mathbf{n}}{\partial x}\Big)$$
$$q_{zy} = -\mu E_z P_y - D\frac{\partial P_z}{\partial y} - \beta n E_x - \delta \frac{\partial \mathbf{n}}{\partial x}$$

where we have expanded the Levi-Civita tensor explicitly. There is no electric field in the $\hat{\mathbf{z}}$-direction, and the second term is not caused by the SHE and is thus unimportant for this discussion. This yields

$$q_{zy} = -\beta n E_x - \delta \frac{\partial \mathbf{n}}{\partial x} \tag{3.46}$$

In essence, our charge current in the $\hat{\mathbf{x}}$-direction causes a $\hat{\mathbf{y}}$-polarized spin current flowing in the $\hat{\mathbf{z}}$-direction.

### 3.5.2 Inverse Spin Hall Effect

The term $\delta \nabla \times \mathbf{P}$, which couples the polarization of the spins $\mathbf{P}$ to the current density $\mathbf{j}$, describes the so-called inverse spin Hall effect (iSHE). As suggested by the name, it is the reciprocal of the regular spin Hall effect: An inhomogeneous spin density, or a spin current, induces an electrical current $\mathbf{j}$.

Consider that we prepare a system with a $\hat{\mathbf{x}}$-directed charge current. From above, we know that this flow will introduce a non-zero $q_{zy}$, that is a $\hat{\mathbf{y}}$-polarized spin current flowing in the $\hat{\mathbf{z}}$-direction (there may be other components, but we will overlook them and only study this simplified case). In our finite sample, this spin current will cause an accumulation of spins at the boundaries: We will get a non-uniform polarization density in the system, as there is a higher concentration of one $\hat{\mathbf{y}}$-polarized charges at the $\hat{\mathbf{z}}$-boundary. Assuming that the spin accumulation is uniform in all other directions and that we only consider $\hat{\mathbf{y}}$-polarization, the fourth term of eq. (3.44) is

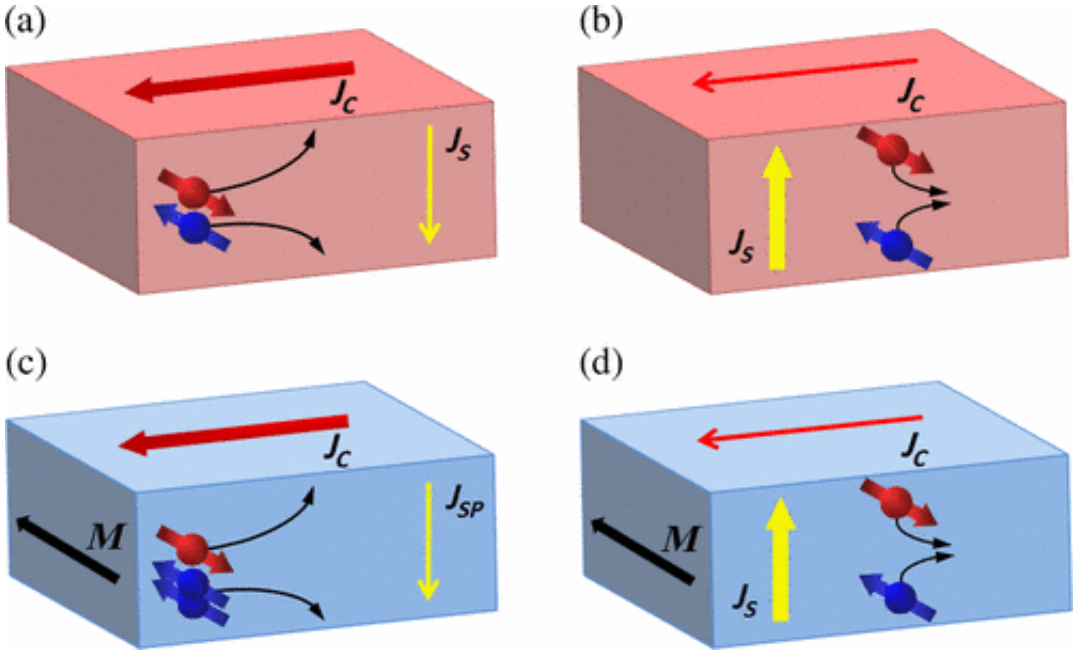



Figure 3.2: Schematic diagrams of (a) spin Hall effect, (b) inverse spin Hall effect in a nonmagnetic metal, as well as (c) anomalous Hall effect, (d) inverse spin Hall effect in a ferromagnetic metal. $\mathbf{J}_C$ is a charge current corresponding to the electron flow (red and blue, color-coded by spin direction), $\mathbf{J}_S$ is the spin current, and $\mathbf{J}_{SP}$ is the spin-polarized current. Figure from [30].

$$\nabla \times \mathbf{P} = (-\partial_z P_y, 0, 0). \tag{3.47}$$

In other words, the $\hat{\mathbf{y}}$-polarized spin current in the $\hat{\mathbf{z}}$-direction induces a charge current in the $\hat{\mathbf{x}}$-direction, perpendicular to itself. This is depicted in Figure 3.2 b), where a spin current $\mathbf{J}_s$ causes a charge current $\mathbf{J}_c \parallel -\hat{\mathbf{x}}$. This figure also illustrates the other (spin) Hall effects.

To better characterize the spin Hall effect and the inverse spin Hall effect, it is common to introduce Hall angles, $\theta_{SHA}$ and $\theta_{iSHA}$. The definition of these quantities are as follows

$$\mathbf{J}_s = \theta_{SHA}(\mathbf{P} \times \mathbf{J}_c) \tag{3.48}$$

$$\mathbf{J}_c = \theta_{iSHA}(\mathbf{P} \times \mathbf{J}_s) \tag{3.49}$$

where $\mathbf{J}_s$, $\mathbf{J}_c$ and $\mathbf{P}$ are the spin current, the charge current, and the charge current polarization respectively. The $\theta_{SHA}$ and $\theta_{iSHA}$ can therefore be regarded as material parameters that describe the efficiency of conversion from charge current to spin current, and vice versa, respectively [31][30]. These quantities are more explicit and useful for numerical implementation [32].

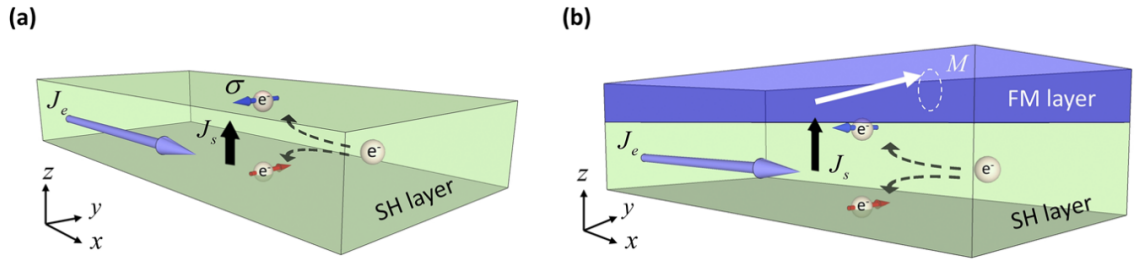



Figure 3.3: Illustration of SHE in (a) a spin Hall (SH) layer (or an HM layer) and (b) a SH/FM bilayer, showing the effect the interface accumulation has on the magnetization. Figure from [34].

### 3.5.3 Spin Hall magnetoresistance

When a current is applied to a(n) (anti)ferromagnetic material, the resistance may depend on the magnetization, a phenomenon called magnetoresistance. It can arise from anisotropy (AMR), from multilayer systems (GMR, CMR, EMR), and as will be discussed here, the two spin Hall effects. In this case, it is called spin Hall magnetoresistance (SMR) [33]. SMR is fundamentally different from the other forms of magnetoresistance: While for instance AMR relies on the flow of current inside the ferromagnet itself, SMR is a non-equilibrium proximity effect, allowing for magnetoresistance to occur within a conducting layer adjacent an isolating ferromagnet (i.e. a bilayer) [33].

We continue our example above, with a charge current flowing in the $\hat{\mathbf{x}}$-direction inside a heavy metal (HM), with an insulating ferromagnet (FM) atop it in the $\hat{\mathbf{z}}$-direction, often referred to as an HM/FM bilayer. Due to the spin Hall effect within the conductor, we will have an accumulation of $\hat{\mathbf{y}}$-polarized spins at the FM/HM interface. The electrons that cause this spin accumulation cannot penetrate into the insulating ferromagnet, however the polarization will be absorbed at the interface, acting as a torque that transfers angular momentum into the ferromagnet. An example of this is shown in Figure 3.3, where the spin current $\mathbf{J}_s$ causes the magnetization $\mathbf{M}$ to oscillate around its axis. For $\mathbf{M} \nparallel \mathbf{P}$, the absorption is greatest [35], and the subsequent torque seeks to drive it into alignment with the $\hat{\mathbf{y}}$-axis [33][35]. Mathematically, this enters the LLG equation, eq. (3.22), as part of the $\vec{\tau}$-term. For an accumulation of spins parallel to the magnetization, the incoming spin current is instead reflected back into the metal [33][35]. The result is that the resistance in the HM-layer changes with the magnetization of the FM layer, becoming smallest when $\mathbf{M} \parallel \mathbf{P}$. The magnitude of the resistance depends on the *instantaneous* direction and magnitude of the magnetization, and the polarization of the charge current. A result of this is that we can create dynamics in the magnetization, and subsequently measure the instantaneous phase of the magnetization, with the conductor: If a constant current

is applied through the HM layer, the effective change in resistivity due to SMR will be measurable in the resulting changes of the voltage.

# 4 Kuramoto model

The Kuramoto model is a mathematical model used to describe systems of coupled oscillators. The long-term dynamics are governed by the following equation in universal form:

$$\dot{\theta}_i = F_i + \sum_j^N \Gamma_{ij} \sin(\theta_j - \theta_i) \tag{4.1}$$

Here, $\theta_i$ is the phase and $F_i$ the natural frequency of oscillator i, and $\Gamma_{ij}$ is an general function describing the coupling strength between oscillators $i$ and $j$, dependent on their phase difference [36]. By natural frequency we mean the frequency that the oscillators would have in the absence of coupling. The oscillators are otherwise assumed to be weakly coupled [36].

These equations are difficult to analyze in general [36] due to the unspecific nature of $\Gamma_{ij}$. The simplest possible version considers all-to-all, equally weighted, sinusoidal coupling,

$$\dot{\theta}_i = F_i + \frac{k}{N} \sum_j^N \sin(\theta_j - \theta_i) \tag{4.2}$$

This is the governing equation for the Kuramoto model. $k$ is an associated coupling constant and N is the number of oscillators in the system [36]. The $1/N$ factor ensures well-behaved dynamics for a large number of oscillators [36]. Indeed, for $N \to \infty$, the Kuramoto model is exactly solvable [36].

For $k = 0$ or for $F_i = F_j, \theta_i = \theta_j \, \forall \, i, j$, the behavior is trivial. For $k \neq 0$, the frequency of oscillator $i$ is larger than its own natural frequency $F_i$ if oscillator $j$ is ahead of it, $\theta_j > \theta_i$, and smaller in the converse case. The effect is that each oscillator will depart its natural frequency based on its present phase compared to the other oscillators within the system. The governing equations seek to make all the instantaneous frequencies synchronized, either completely or with a phase difference (phase-locked). It should be noted that this model is not so much concerned with the instantaneous phases as it is about their instantaneous frequencies [36].

Noise can be added to the system by rewriting the equation as

$$\dot{\theta}_i = F_i + \zeta_i + \frac{k}{N} \sum_j^N \sin(\theta_j - \theta_i) \tag{4.3}$$

where $\zeta_i$ is the time-dependent fluctuation of oscillator $i$, entirely independent of all other oscillators. This can be modeled with a white noise process, i.e.

$$\langle \zeta_i(t) \rangle = 0 \qquad \langle \zeta_i(t) \zeta_j(t') \rangle = 2D\delta_{ij}\delta(t - t') \tag{4.4}$$

however for our purposes we will consider only a zero-noise system.

To further study the dynamics of the oscillators, we define an order parameter by imagining a 2D oscillation in the complex plane [36], and thus write our definition as

$$re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j} \tag{4.5}$$

where $r$ is defined as the radius or phase coherence, and $\psi$ is the mean phase of the system [36]. In this case, $r \in [0, 1]$, with $r = 0$ and $r = 1$ describing completely incoherent and coherent motion respectively.

Multiplying both sides by $e^{-i\theta_i}$, we get

$$re^{i(\psi-\theta_i)} = \sum_{j=1}^{N} e^{i(\theta_j-\theta_i)} \tag{4.6}$$

which, when considering the imaginary part only, yields [36]

$$r\sin(\psi - \theta_i) = \frac{1}{N}\sum_{j=1}^{N} \sin(\theta_j - \theta_i) \tag{4.7}$$

Thus, the Kuramoto model can be rewritten as [36]

$$\dot{\theta}_i = F_i + kr\sin(\psi - \theta_i) \tag{4.8}$$

This is remarkable: Rather than describing the dynamics of each oscillator's phase by a deviation from all the other phases, this reformulation allows it to be described by a deviation from the mean phase $\psi$. Moreover, because the order parameter $r$ grows as more and more oscillators synchronize, a positive feedback loop is formed in the system, given that the coupling strength $k > k_c$, where $k_c$ is a critical coupling strength [36]. Thus far, we have imposed no restrictions on r and $\psi$.

It is illustrative to determine the synchronization times in a large system of oscillators, and so we will convert $r$ to a more suitable and measurable quantity: One of our neuromorphic devices uses the SHE to oscillate, forced to do so within an easy-plane plane via anisotropy. This becomes a 2D oscillation well-suited for the Kuramoto model. We take the absolute value of the order parameter equation:

$$|re^{i\psi}| = \left|\frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j}\right|, \tag{4.9}$$

which reduces to

$$|r| = \frac{1}{N}\left|\sum_{j=1}^{N} e^{i\theta_j}\right| \tag{4.10}$$

$$|r| = \frac{1}{N}\left|\sum_{j=1}^{N} \cos\theta_j + i\sin\theta_j\right| \tag{4.11}$$

$$|r| = \frac{1}{N}\sqrt{\Big(\sum_{j=1}^{N}\cos\theta_j\Big)^2 + \Big(\sum_{j=1}^{N}\sin\theta_j\Big)^2} \tag{4.12}$$

$$\tag{4.13}$$

Explicitly expanding this:

$$|r| = \frac{1}{N}\sqrt{\cos^2\theta_1 + \cos^2\theta_2 + ... + \cos^2\theta_N + \sin^2\theta_1 + \sin^2\theta_2 + ... + \sin^2\theta_N + \text{cross terms}} \tag{4.14}$$

Using a trigonometric identity we reduce all the quadratic terms to $N$, and then write the cross terms as a sum:

$$|r| = \frac{1}{N}\sqrt{N + \sum_{i=1}^{N}\sum_{i\neq j}^{N}(\cos\theta_i\cos\theta_j + \sin\theta_i\sin\theta_j)} \tag{4.15}$$

Under the assumption that the magnetization is confined within the xz-plane, we define a normalized magnetization of a ferromagnet $i$ as $\mathbf{m}_i$, with $m_{ix} = \cos\theta_i$ and $m_{iz} = \sin\theta_i$, whereupon

$$|r| = \frac{1}{N}\sqrt{N + \sum_{i=1}^{N}\sum_{i\neq j}^{N}(m_{ix}m_{jx} + m_{iz}m_{jz})} \tag{4.16}$$

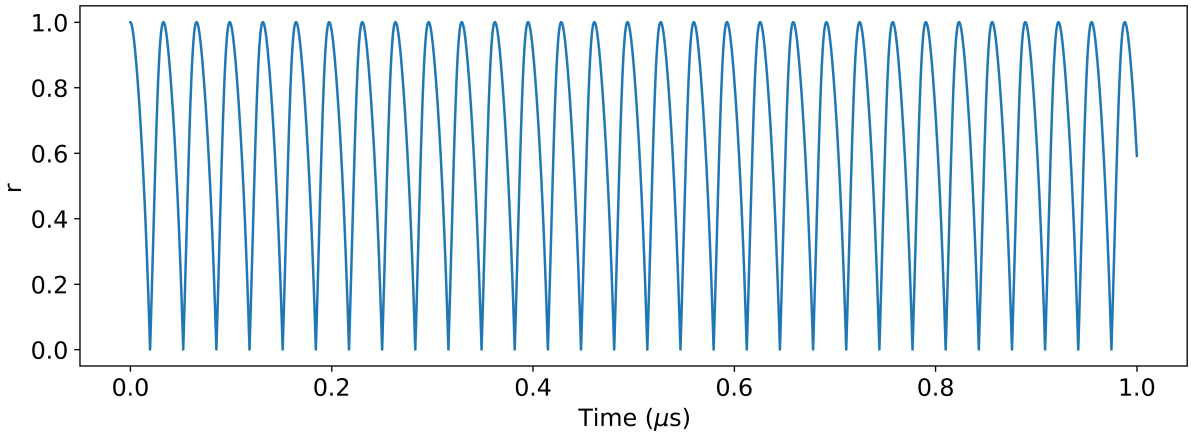
In this work we will often consider two-oscillator systems, for which the order parameter is expressed as

$$|r| = \frac{1}{2}\sqrt{2 + 2m_{1x}m_{2x} + 2m_{2z}m_{1z}} \tag{4.17}$$

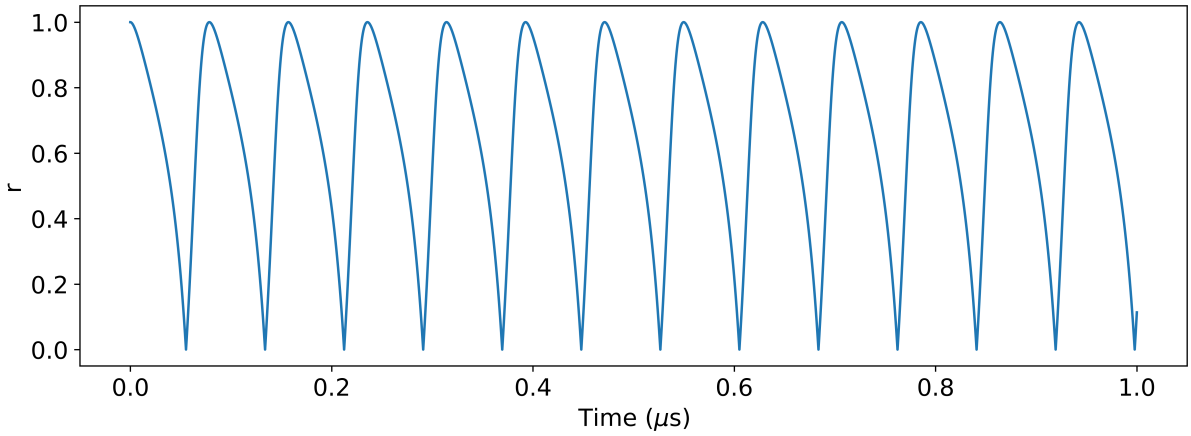
or, in trigonometric terms

$$|r| = \frac{1}{2}\sqrt{2 + 2\cos\theta_1\cos\theta_2 + 2\sin\theta_1\sin\theta_2} \tag{4.18}$$

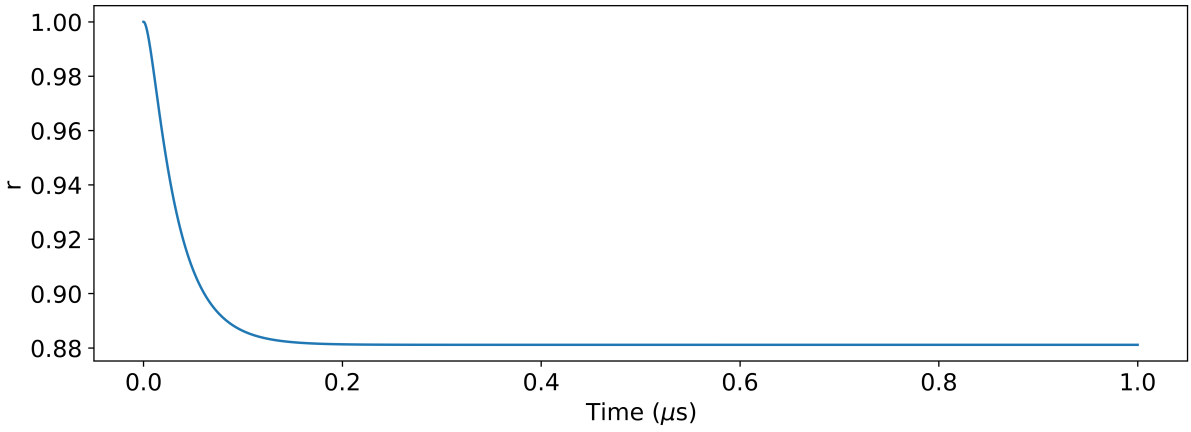
We consider a few example cases to give some intuition about this quantity: When the oscillators are entirely in-phase, $\theta_1 = \theta_2$ and thus $\cos\theta_1 = \cos\theta_2$ and $\sin\theta_1 = \sin\theta_2$. Inserted into (4.15), this yields $|r| = 1$. For antiparallel oscillators, $\theta_1 = \theta_2 + \pi$, then $\cos\theta_1 = -\cos\theta_2$ and $\sin\theta_1 = -\sin\theta_2$, which yields $|r| = 0$ when inserted. For oscillators




(a) $F_1 = 6.5$ GHz, $F_2 = 6.7$ GHz, k = 0.03 GHz.




(b) $F_1 = 6.6$ GHz, $F_2 = 6.7$ GHz, k = 0.03 GHz.




(c) $F_1 = 6.65$ GHz, $F_2 = 6.7$ GHz, k = 0.03 GHz.

Figure 4.1: Examples of the time evolution of the order parameter for the Kuramoto model. System parameters are listed below the figures. In a) and b), the oscillators are desynchronized. In c), the oscillators are synchronized with a phase difference, shown by the constant value $r < 1.0$. By increasing the coupling constant or bringing the natural frequencies closer, this $r$ will approach 1.0, eventually being completely constant at $r = 1.0$ for $F_1 = F_2$.

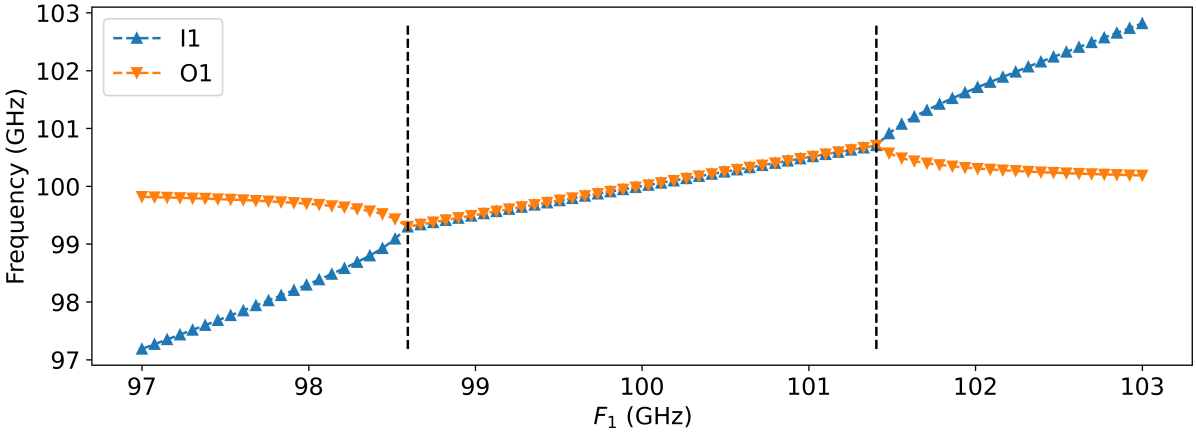



Figure 4.2: Instantaneous frequencies $f_1$ and $f_2$ as a function of natural frequency $F_1 \in [97, 103]$ GHz. $F_2 = 100$GHz, $k = 0.72$GHz. Dashed lines mark the beginning and end of the synchronization range, with synchronization defined as $|\hat{f}_1 - \hat{f}_2| < 0.0012$ GHz.

$\pi/2$ out of phase, $\theta_1 = \theta_2 \pm \pi/2$ yields $|r| = \sqrt{2}$, and so forth. If the phase difference is maintained over time, the oscillators are considered phase-locked and the order parameter becomes a constant in time $r \leq 1$ depending on the phase difference. For non-phase-locked oscillators, the order parameter will oscillate with the instantaneous phase difference.

Figure 4.1 shows the time evolution of the order parameter for three different setups of 2 oscillators. The coupling constant is fixed in all cases while the natural frequency gap is slowly closed, until synchronization is spontaneously attained. As the natural frequencies $F_1$ and $F_2$ approach each other, the variations in $|r|$ slow down, until they reach a critical natural frequency after which they synchronize to some $|r| < 1.0$.

As with earlier work [12], we define the **synchronization range** as the range of natural frequencies for which two oscillators synchronize. As the Kuramoto model is equivalent under an exchange of oscillators, we illustrate this by tuning the natural frequency of oscillator 1, $F_1$, while keeping $F_2$ fixed. Figure 4.2 plots the instantaneous frequencies $f_1$ and $f_2$ as a function of $F_1$. We note the gradual departure of the instantaneous frequencies from their corresponding natural frequencies, until they spontaneously synchronize for $F_1 \in [98.6, 101.4]$ GHz, a range approximately 4 times the coupling constant. This is a general relation for the 2-oscillator system: The synchronization range is always 4 times the coupling constant, regardless of the exact value of $F_2$. In Figure 4.1c where $k = 0.03$ GHz, the synchronization range of size 0.12 GHz, or 0.06 GHz in either direction centered on the fixed $F_2$.

# 5 Introduction to neuromorphics

## 5.1 Neuromorphic computing

Neuromorphic computing is a concept originating in the late 1980s from the seminal work of Carver Mead [10]. It is, in essence, encapsulated within its own title: Computation which aims to mimic the function of the human brain [5], rather than the much more approachable method of modern computers. For while it is approachable, it is also terribly inefficient by comparison. Estimating the floating-point operations per second (FLOPS) of the human brain is difficult as the architecture is vastly different from a conventional computer, but the lower bound is placed at $2 \times 10^{16}$ FLOPS [37]. By comparison, the fastest supercomputer as of November 2022 is the Frontier machine at Oak Ridge National Laboratory (ORNL), which at its peak operates at $1.6 \times 10^{18}$ FLOPS [38]. This is also the most efficient supercomputer, achieving approximately $62.68 \times 10^9$ FLOPS $\mathrm{W}^{-1}$. By comparison, the brain operates at approximately 20 W [7], meaning an efficency of $10^{15}$ FLOPS $\mathrm{W}^{-1}$, several orders of magnitude greater.

By emulating the brain's functionality in hardware, neuromorphic computing aims to close the gap in FLOPS and create more efficient computers. There are a number of key aspects that differentiate a conventional and neuromorphic computational device, which contribute to the speed.

Conventional computers for the most part abide by the von Neumann architecture. This was proposed by John Von Neumann in the paper "First draft of a report on the EDVAC" [4]. The EDVAC was among the first electronic computers ever designed, and in the report on it, von Neumann proposed, among other things, the separation of the processor unit and memory. This has proven widely successful and is adopted into practically every commercially-available computer today. There are, however, issues with this: First, computational speed is inevitably constrained by the rate of the data transfer between the memory and the processor unit. This is a resource-intensive operation [5], and much work has been spent mitigating the effect of the bottleneck in conventional computers, such as memory caching, branch prediction, etc. These only postpone the inevitable without solving the issue. The brain has a different structure altogether, wherein the data storage and the analysis thereof occur in the same place, or within the same device [5]; in the brain, the neuron itself is responsible for both storing memory and propagating the electrical impulses that operate on it. Neuromorphic hardware therefore avoids the Von Neumann bottleneck by its very design, opening for potentially large increases in computational speeds.

Other drawbacks of modern computers are their inherent serial in nature: A problem is broken into a series of instructions, which are executed sequentially on a single processor which changes the state of the program. While virtually all modern computers are parallel

from a hardware perspective, with for instance several graphical processors operating simultaneously, each of the individual devices remains serial. By combining several serial units, in addition to employing modern computer science tricks, it effectively creates a parallel computing scheme. The brain, by comparison, is inherently parallel: It is able to maintain life functions, recognize images in milliseconds, form sentences, visualize abstract concepts and perform many subconscious tasks all at once. Emulating the brain's structure may increase the potential for parallel computation, very beneficial for AI-related tasks such as computer vision [5].

As one can imagine, AI applications will benefit greatly from the adoption of neuromorphic principles, as their work already seeks to incorporate brain-like functionality on computing devices.

A drawback has been pointed out earlier: Conventional computer devices employ workarounds and mitigations of their limitations rather than dealing with them directly using new hardware. This makes it difficult to integrate them into neuromorphic computing principles. Research has therefore been spent on finding other hardware which is more applicable to neuromorphic computing. Spintronics is one of these examples. In order to understand the application of spintronics to neuromorphic computing, and especially to neural networks, an introduction to all three is necessary. We will first explore existing concepts in neural networks, followed by neuromorphic computing and their relation to these. This will lay the groundwork for a study of the isolating ferromagnet for use in spintronic neuromorphic.

We will give a qualitative and quantitative introduction to conventional neural networks. Additionally, we will introduce spiking and oscillatory neural networks.

## 5.2 Neural structure

In order to understand the basis of neuromorphic devices, it is important to first understand the operation for the brain and central nervous system (CNS). The key components in play are the neurons and the synapses. The neurons are specialized cells which perform the basic signal processing and transmission inside the CNS. Our brains contain roughly 100 billion neurons [39]. The neurons themselves are divided into three separate parts: The dendrites, soma and axon. Figure 5.1 depicts a single neuron, with every component inside of it. We have added labels to highlight the most important aspects.

The dendrites gather the input from neighboring neurons [39], branching out and opening for potentially tens of thousands of connections into a single neuron [40]. The resulting input is passed into the soma, the neuron's processing unit. The arrival timing here is of utmost importance: The soma has a resting potential where it remains without any stimulation. Arriving signals will temporarily raise the potential of the neuron until it hits a threshold, called the activation potential, at which point it generates an action

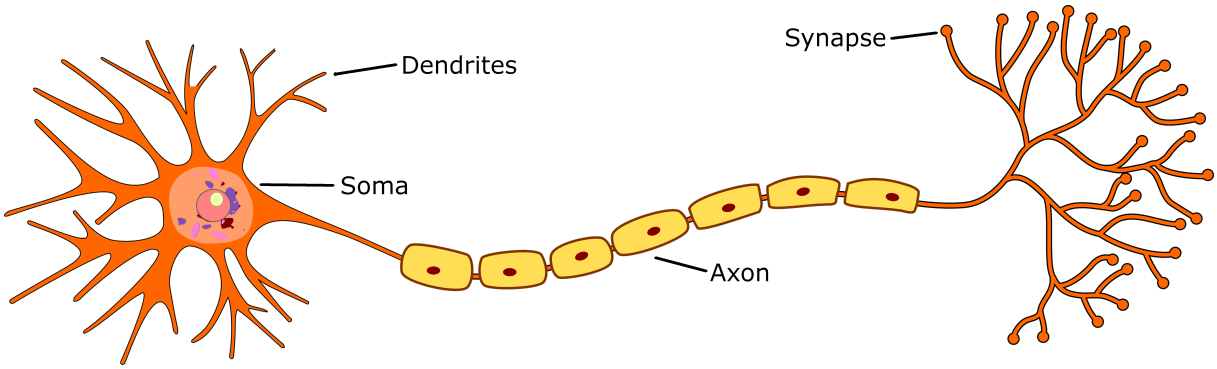



Figure 5.1: Schematic of a neuron, with important parts labeled. Figure adapted from [43], labels added.

potential, or a large and momentary increase in potential across the membrane. This is often called spiking or firing [41]. If the input signals arrive within a short enough time frame and are of sufficient strength, the sum of their inputs is enough to cause a firing [41]. Otherwise, the neuron "leaks" the accumulated energy and relaxes back to its resting potential [42].

The output is then passed to the axon, where it propagates into several axon terminals, which release neurotransmitters to the dendrites of connected neurons [39]. The actual connections from the axon terminals to neighboring dendrites are through the synapses, which act as weights $\omega$ on different signals, promoting some and downprioritizing others. This process then repeats itself in the next neuron, potentially causing a series of spiking events which propagate through the network of neurons.

The aforementioned weight $\omega$ of the synapse should be familiar to anyone who has knowledge on artificial neural networks (an introduction to which we present later). The adjustability of the synaptic weights makes certain pathways through the brain more conducive, which forms the cornerstone of learning and the formation of memories [44].

## 5.3 Neural oscillations

Neural oscillations, also known as brainwaves, come in two flavors: Within individual neurons it comes as oscillations within either the membrane or action potential. On a larger scale, it is representative of the synchronization of a number of neurons, giving rise to macroscopic oscillations that are visible on an electroencephalogram (EEG) [45]. The latter oscillations are of greater interest to us, and so we focus on them. It is to be noted that although these macroscopic oscillations show up on EEGs, the connection between these oscillations and cognitive activity is a subject of ongoing research [46], and their computational function is an ongoing debate among neuroscientists [47].

These macroscopic oscillations are characterized by their frequency, amplitude and phase, all of which change due to the properties of the individual neurons and, more

importantly, synchronization within neural ensembles (so-called local synchronization). Though EEGs usually have a broad spectrum, certain bands of frequencies have an increased oscillatory activity that is highlighted. To name a few of these, alpha wave (8-12 Hz), beta wave (13-30 Hz), delta wave (1-4 Hz), and two bands of gamma waves (up to 150 Hz).

Each wave is more prominent during certain activities. The alpha wave, for instance, is predominant during wakeful relaxation where the eyes are closed [48], while the high-frequency beta activity is linked to cognitive tasks [48]. The purpose of these oscillations may be to group neurons together into computational units in response to cognitive load. There is a dramatic change in EEG output during sleep compared to wakeful stages, suggesting a strong link to states of consciousness [45]. However, as mentioned, the exact nature of these oscillations is not fully understood. Other, more concrete suggestions for their purpose include feature binding (the combination of input and information into a cohesive experience) [49], neural coding (relationship between input stimulus and the subsequent response) [45][50] and the control of everyday rhythmic behavior like walking and breathing [45].

## 5.4 Basics of neural networks

Artificial neural networks, as they are most commonly made today, consist primarily of two components: Nodes, or neurons, commonly arranged in layers, connected by adjustable weights $\omega$, modeling the synapses in a biological neural network. The first layer is called the *input layer*, where input data is introduced to the system. The last layer is called the *output layer*. Between these are *hidden layers* responsible for transforming the input data into some form of output classification. Although this is a very common neural network there are other forms. One example is perceptrons [51], which do not have hidden layers at all. Each layer is connected only to the nodes in the previous and subsequent layer, with each node taking input from the previous layer in the form of a weighted sum [52],

$$Y_j^{K+1} = \sum_i \omega_{ji} Y_i^K \tag{5.1}$$

where $\omega_{ji}$ are the synaptic weights associated with that node-node connection, $Y_i^K$ is data outputted by a node in the Kth layer, and $Y_j^{K+1}$ is the data received by node $j$ in the K+1 layer. This is known as the propagation function. The magnitude of the weight therefore specifies the strength of a specific connection. This input is transformed into an output via an activation function, $\sigma(x)$. The activation function is not standardized to any one specific function, though to compute non-trivial problems with a small amount of nodes this function must be non-linear. [53]. Common choices are the ReLu function,

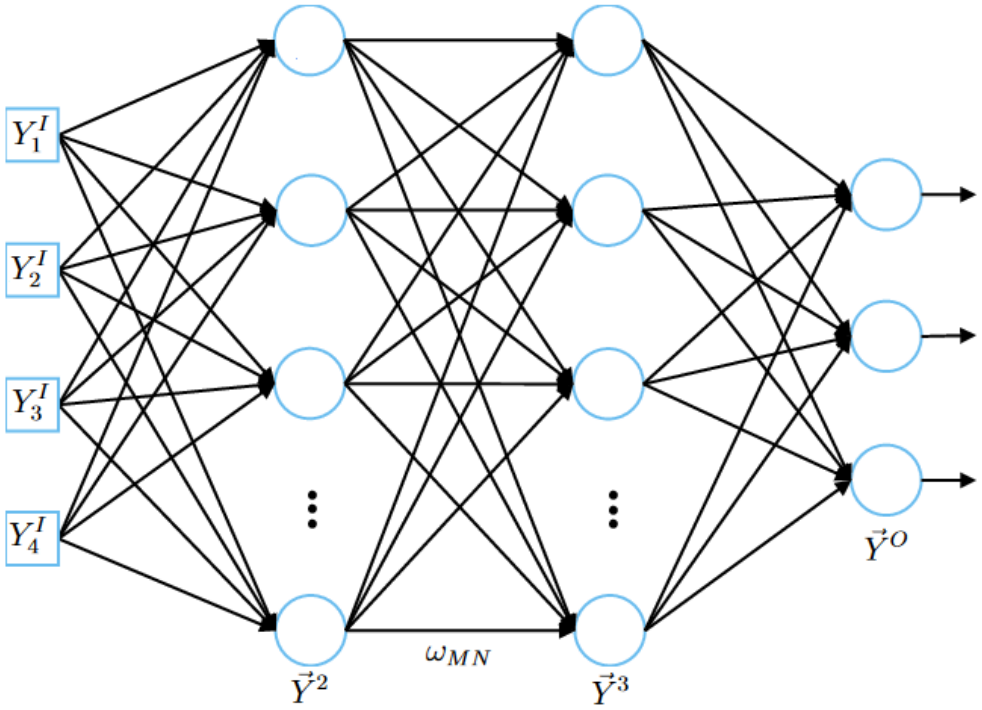



Figure 5.2: Conventional feedforward neural network in 4 layers, 2 of which are hidden. Adapted from [54] by modifying labels.

the hyperbolic tangent, or the Sigmoid function, respectively

$$\sigma_{\text{ReLu}}(x) = \max(0, x)$$
$$\sigma_{\text{hyp}}(x) = \tanh(x)$$
$$\sigma_{\text{sig}}(x) = \frac{1}{1 + e^{-x}}.$$

This output is then distributed as input in the next layer, in the same fashion as above. One can also introduce bias nodes, which are not connected to any preceding layers and instead only provide a bias input to a specific layer.

The second important aspect of a neural network is *training*, which is done by processing a large volume of examples. Initially, the weights of the system are randomized. Then, input data from this example data set is presented to the system through the input nodes. After passing through all the hidden layers, the output is compared to the expected output. The difference between the actual and expected output is defined as the error. At this point, a *learning rule* is used to adjust the weights of the system such that this error is minimized. With a large data set, the weights will tend towards values that minimize this error for all the examples [55]. At some point, one can stop the training process and use the trained network to classify unknown data accurately [55].

This whole process can be described in terms of matrix multiplication [52]: Consider N neurons in the $K$th layer, represented by the vector $\vec{Y}^K$, connected all-to-all to M neurons in the $K+1$th layer, $\vec{Y}^{K+1}$, where we borrow the notation for the neurons from eq. (5.1). A sketch of this system is given in Figure 5.2. As each of the $Y_j^{K+1}$ neurons are a weighted sum of $Y_i^K$, we can write this as an equivalent matrix multiplication

$$\vec{Y}^{K*} = W_K \vec{Y}^K + \vec{b} \tag{5.2}$$

where element $i$ of $\vec{Y}^K$ corresponds to $Y_i^K$, $W_K$ is an $M \times N$ matrix where the elements are $\omega_{ji}$, both from eq. (5.1) and we include a bias $\vec{b}$ which also connects to every neuron in $K+1$. $\vec{Y}^{K*}$ is a temporary value, because we now apply the activation function $\sigma$ on every neuron, that is on every element in the $\vec{Y}^{K*}$ vector, to determine which neuron is allowed to propagate. The input into the next layer, $\vec{Y}^{K+1}$, is then

$$\vec{Y}^{K+1} = h\sigma(W_k \vec{Y}^K + \vec{b}) \tag{5.3}$$

where we have additionally added a so-called learning parameter $h$.

This describes the propagation of input data $\vec{Y}^I$ through the network, through a series of transformations from input to output through all the layers. As mentioned, our initial input data will be training data and thus have corresponding expected output $\hat{\vec{Y}}$. We must now update our weights in such a way that input data will be transformed to match this expected data. To do this, we first define a loss function (or objective function). There is no single loss function, and developers can select any suitable function to fit their needs. An example is

$$\mathcal{J} = |\vec{Y}^O - \hat{\vec{Y}}|, \tag{5.4}$$

simply the difference between the output data $\vec{Y}^O$ and the expected data.

To summarize thus far: We have obtained how some input data will be transformed when propagating through our neural network. We have additionally defined a loss function for use during training to determine how the neural networks perform. Now it is time to implement the backpropagation, such that we can meaningfully update the synaptic weights in a way that best minimizes this error. To do so, we must find the gradient of $\mathcal{J}$ with respect to the weights, $\boldsymbol{\nabla}\mathcal{J} = \partial\mathcal{J}/\partial W_k$. One can also allow other parameters, such as the biases $\vec{b}$, to be tunable, at which point it will enter this gradient calculation.

From this, we obtain the iterative expression that will minimize the loss function by adjusting the system weights,

$$W_{k+1} = W_k - \tau\boldsymbol{\nabla}\mathcal{J}, \tag{5.5}$$

where $\tau$ is a step length deciding the impact each sample should have on the learning of the model [52].

Finding gradients is a very time-consuming process, and modern neural networks involve an ever growing number of nodes, drastically increasing the number of parameters and thus the number of elements in the $W_k$ matrix. The large language model GPT-3, which is, at time of writing, a generation behind the cutting edge large language model GPT-4, is known to have employed 175 billion parameters, all of which had to be iteratively trained [56]. The parallelizability of these computations is therefore paramount, and is what is employed in machine learning frameworks such as TensorFlow [9].

## 5.5 Spiking neural networks

The nodes in an artificial neural network employ a static propagation model where all neurons are powered at all times, which very loosely models a biological neural network. In a spiking neural network (SNN), the neurons are constructed in such a way that they more closely resemble biological neurons. An old, but still instructive model is the leaky integrate-and-fire (LIF) model for neurons, which provides a time-dependence in addition to a whole host of effects. We will come back to the LIF model, but it should be noted that there are other models for neurons used in SNNs, for instance the very important Hodgkin-Huxley model [57].

The neurons in an SNNs only transmit when their internal potential reaches a threshold value. Here, the neuron fires (spikes), releasing a large current pulse which travels into the next neuron. This current pulse is modulated by artificial synapses, similar to the weights the conventional neural network. When propagating, this pulse increases the potential of connected neurons, potentially causing a cascade of spikes propagating through the network. In addition, the original neuron in this consideration relaxes back from its threshold value.

The time-dependence is contained in how each neuron receives and potentially propagates the information. This is in part described by the LIF model: The arriving spikes $Y_i^K$ are multiplied by weights $\omega_i$ and integrated (summed) as prior. If this summation results in a value higher than a threshold, a spiking event occurs within the neuron and the information is transmitted as $Y_I^{K+1}$. If the potential remains below the threshold no spiking occurs. However, the neuron itself wants to relax back to its ground state, and will slowly leak the potential rather than store it indefinitely, ergo the summing is time-dependent. A series of spikes arriving from many neurons within a short period (called a spike train) may be necessary to cause the destination neuron to fire, and it is within these spike trains that the input data is encoded [58].

In addition to the time-dependence, SNNs often fulfill (or should fulfill) criteria that hold for biological neurons. The most important properties are [57][14]:

- Refractory periods: For a short period after firing, the neuron will not fire again, even if normally sufficient stimuli is applied.

- Latency: Upon receiving stimuli, biological neurons should exhibit delay, or latency, in their response.
- Bursting: A converse of a refractory periods, where the magnitude of the stimuli is great enough that the neuron fires repeatedly.
- Inhibition: The ability for neurons to act in an inhibitory rather than just excitory fashion when encountering a spike, thereby preventing propagation.

There are advantages of a SNN compared to its conventional counterpart: They are in theory faster than the neurons in a conventional network, and require less energy as the neurons are only powered when they are firing, unlike the static neurons of today that are continuously powered. Additionally, they may require fewer total neurons for the same level of accuracy as a conventional neural network. There are also obstacles in the way of creating a SNN: At present, there are a limited number of unoptimized methods of training a SNN, and the ones used in conventional NNs (such as gradient descent) are not suitable for the different paradigm that is SNN. Schemes such as spike-timing-dependent plasticity and Hebbian learning exist and have been implemented, but the details thereof are omitted in this introduction [11][59]. Furthermore, implementing SNNs with transistors is very hardware intensive [10].

## 5.6 Oscillatory neural network

Though it may not be the easiest to implement, the spiking neuron is intuitively related to the biological neuron: It obviously captures the spiking itself, but certain implementations also capture other properties of biological neurons such as bursting, refractory periods and latency. But the brain also experiences macroscopic oscillations that are observable in electroencephalography diagrams, so-called neural oscillations or brain waves [46]. The details of these oscillations have been discussed in Section 5.3. Oscillatory neural networks (ONN) try to emulate this behavior, exploiting the synchronization of oscillators for computation [46]. This reliance on synchronization make ONNs closely tied to the Kuramoto model [60].

It is possible to use the normal Hebbian learning rule in ONNs, making them similar to existing neural networks [60]. During Hebbian learning, it is the synaptic weights between the oscillators that are updated every iteration, for instance the coupling constant $K$ of the Kuramoto model [46] that was discussed in Section 4. Information is encoded in the phase of the oscillators. Setting one oscillator as the reference, we can use the relative phase of all the other oscillators in comparison to the reference to encode the desired data [46].

It is also possible to keep $K$ constant for all iterations of the system, and instead tune the natural frequency $F_i$ in the Kuramoto model. This paradigm has received less

attention than the former, but is more attractive for nano-scale applications [60]: On a nano-scale device, tuning the coupling constant is often difficult as it is tied directly to the material parameters themselves, while tuning the frequency is comparatively simple [60]. This approach effectively neglects the exact value of the phases for each oscillator and instead evaluates the frequencies: The oscillators are either synchronized or desynchronized to some reference frequency, encoding a binary state. This approach to ONNs has already been realized in nanoscale spintronic devices, both in theory [12][61][13] and experimentally [49], as well as on several other hardware implementations [62]. This method is key to our work. Therefore, we present a detailed introduction based on the work of Vodenicarevic et. al. [60], who developed the learning algorithm used in our work.

There are two key advantages of an ONN: Synchronized oscillators all operate in parallel, which can be exploited for parallel computation. Additionally, computing with oscillators in the frequency domain allows low-voltage operation compared to standard transistor-based computation. In short, ONNs are low-power and highly parallelized, exactly what is sought-after for training and implementing artificial intelligence on various devices [46].

### 5.6.1 Training an ONN

The training method by Vodenicarevic et. al. [60] is an extension of one presented by Vassilieva et. al. [63]. The latter is not used in this thesis, and so it will be introduced in broad strokes to highlight the differences.

The Vassilieva learning algorithm [63], was initially presented in 2011 through a mathematical context. The algorithm consists of presenting training input examples $\vec{Y}^I$, with known expected output $\hat{\vec{Y}}$, to the network through a set of input oscillators in a network consisting of both input and output oscillator (each of arbitrary number, though usually more output than input oscillators). Once done, it obtains the list of synchronized pairs of oscillators, and compares it to $\hat{\vec{Y}}$. The learning algorithm then operates on the natural frequencies of the oscillators, slightly pushing apart the unexpected synchronizations and at the same time promoting the desynchronized pairs that are supposed to be synchronized by pulling their natural frequencies closer [60]. Note that a "pair" of oscillators, in this case, refers to a pairing of the input oscillator to any other non-input oscillator in the system, which all function as output oscillators. All output oscillators synchronize to one or none of the input oscillators, and are not allowed to interact with each other. As pointed out in Vodenicarevic et. al. [60], this algorithm is greedy, only acting on IO-pair as though they are isolated and do not influence each other. For some systems this is an acceptable assumption [61]. But in a general system, all oscillators in the system will affect each other, and thus the algorithm may end in suboptimal results [60]. Furthermore, without inter-oscillator coupling, the size of a problem that can be presented to the

oscillators is limited [60].

The Vodenicarevic algorithm is, in broad terms, closer to what is expected from a conventional neural network: The output oscillators are now allowed to interact with each other in addition to the input oscillators, and we look at the synchronizations between various output oscillator pairs to determine the state of the system. Further similarities to conventional neural networks are that it is an offline learning algorithm which operates by actively minimizing, through gradient descent, the total error function:

$$E^{tot} = \frac{1}{M} \sum_{m=1}^{M} \sum_{\text{o-pairs}} E_{i,j}^{(m)} \tag{5.6}$$

where $M$ is the number of training samples, $E_{i,j}^{(m)}$ is the error of output pair, and o-pairs are pairs of output oscillators $(i, j)$ which we are summing over for sample $m$. To quantify this error further, we introduce a measure for the actual desynchronization

$$D_{i,j}^{m} = \frac{|\overline{f}_i^{(m)} - \overline{f}_j^{(m)}|}{k/2} \tag{5.7}$$

where $k$ is the coupling constant (which could be that of the Kuramoto model, but other suitable models can be used [60]), and $\overline{f}_i^{(m)}$ is the time-averaged instantaneous frequency of the $i$th oscillator, which is potentially different from its natural frequency $F_i$ due to interactions with other oscillators. If $D_{i,j}^{m} \leq 1$, the oscillators are considered synchronized, otherwise they are not. Note that $D_{i,j}^{(m)}$ is always a non-negative number.

Based on this, the error itself is defined differently for our two different states: If the oscillators $(i, j)$ are **expected** to be synchronized for sample $m$, the error is

$$E_{i,j}^{(m)} = \begin{cases} \frac{1}{2}\left(D_{i,j}^{(m)}\right)^2 & \text{if } D_{i,j}^{(m)} \leq 1, \\ \frac{1}{1+e^{-4\left(D_{ij}^{(m)}-1\right)}} & \text{if } D_{i,j}^{(m)} > 1 \end{cases} \tag{5.8}$$

while if they are expected to be desynchronized, the error is

$$E_{i,j}^{(m)} = \begin{cases} 1 - \frac{1}{2}\left(D_{i,j}^{(m)}\right)^2 & \text{if } D_{i,j}^{(m)} \leq 1 \\ 1 - \frac{1}{1+e^{-4\left(D_{ij}^{(m)}-1\right)}} & \text{if } D_{i,j}^{(m)} > 1. \end{cases} \tag{5.9}$$

Note that in both these equations, the cases considers the state of the actual oscillators. For all values of $D_{i,j}^{(m)}$, the error $E_{i,j}^{(m)} \in [0, 1)$. Figure 5.3 depicts the loss function as a function of the desynchronization, $D_{ij}^{(m)}$.

We now perform the gradient descent. From Figure 5.3, it is apparent that the loss function is minimzed by $D_{ij}^{(m)} \to 0$ for oscillators meant to be synchronized and $D_{ij}^{(m)} \to \infty$ for desynchronized. Just like with regular machine learning, we now compute the error gradient with respect to our tunable parameter, which are the natural frequencies

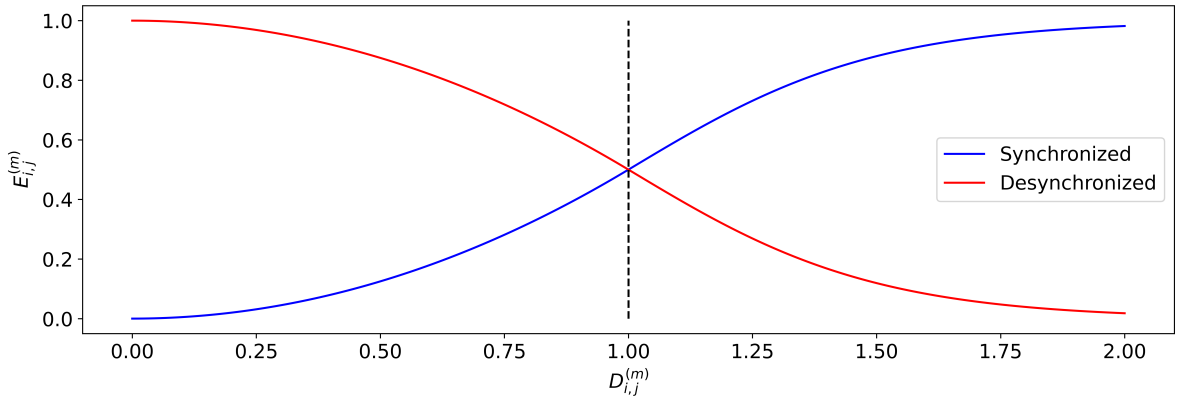



Figure 5.3: Vodenicarevic loss function, plotted for a range $D_{ij}^{(m)} \in [0, 2]$. The functions will asymptotically go to 1 and 0 respectively. Dashed line marks the transition from synchronized to desynchronized.

of the output oscillators ($D_{ij}^{(m)}$ depends on these implicitly, as the mean instantaneous frequencies $\overline{f}_i^{(m)}$ depend on the initial configuration of the natural frequencies $F_i$). Thus, we need to determine $\partial E_{i,j}^{(m)}/\partial \mathbf{F}$, which we can expand via the chain rule to be

$$\frac{\partial E_{i,j}^{(m)}}{\partial F_a} = \frac{\partial E_{i,j}^{(m)}}{\partial \bar{f}_i^{(m)}} \frac{\partial \bar{f}_i^{(m)}}{\partial F_a} + \frac{\partial E_{i,j}^{(m)}}{\partial \bar{f}_j^{(m)}} \frac{\partial \bar{f}_j^{(m)}}{\partial F_a}, \tag{5.10}$$

where $\frac{\partial E_{i,j}^{(m)}}{\partial \bar{f}_i^{(m)}}$ is found by differentiating eq. (5.8) and eq. (5.9).

Like Vodenicarevic et. al., we use the Kuramoto model both for this explanation and later in our implementation. The differential equations are given in eq. (4.3), which govern the instantaenous frequency and the phase, both relevant to our implementation. We differentiate both of these with respect to the natural frequencies and get

$$\begin{cases} \dfrac{\partial f_i(t)}{\partial F_a} = \delta_{i,a} + k \displaystyle\sum_j \left( \frac{\partial \theta_j(t)}{\partial F_a} - \frac{\partial \theta_i(t)}{\partial F_a} \right) \cos\left[\theta_j(t) - \theta_i(t)\right] \\ \dfrac{\partial \theta_i(t+dt)}{\partial F_a} = \dfrac{\partial \theta_i(t)}{\partial F_a} + 2\pi dt \dfrac{\partial f_i(t)}{\partial F_a} \end{cases} \tag{5.11}$$

where $\delta$ is the Kronecker delta function, and $\partial \boldsymbol{\theta}(t=0)/\partial \mathbf{F} = 0$. This differential of instantaneous frequency with respect to natural frequency is then combined into the learning rule, eq. (5.10), becoming a complicated expression omitted presently.

During implementation, undesired sensitivity to initial conditions are avoided by randomizing the initial phases $\theta \in [0, 2\pi)$, and the natural frequencies $F_i$ are clipped to stay within a certain region, dependent on the range of devices used.

All work presented in this section is attributed to Vodenicarevic et. al. [60]. This is an offline learning algorithm of complexity $\mathcal{O}(N^3)$, but it is highly parallelizable. As this algorithm is related to standard gradient descent common in conventional machine

learning schemes, it can and has been implemented as a neural layer in machine learning libraries such as TensorFlow, which ensure the use of GPU acceleration if available, drastically increasing the speed of training [60].

## 5.7 Summary

In this section, we have presented two novel methods of neural networks that more closely align with current models of the brain. The first are SNNs, which model the leaky integrate-and-fire behavior of a biological neuron, in addition to many other biological properties. These are complicated networks that do not have a readily-available learning algorithm nor integrate easily with existing hardware, but promise highly parallel, low-power computing. The second are ONNs, modeling a more macroscopic brainwave activity using the synchronization and phase-locking of coupled oscillators to perform parallelized cost-effective computation. These oscillators come in two flavors, the more hardware-friendly of which trains the system by updating the natural frequencies of output oscillators. This has already seen realizations, both experimentally and in simulation, and already has several learning algorithms which employ existing neural networking techniques, for instance gradient descent.

These are two examples of neuromorphic approaches to neural networks, but it's not a comprehensive list, and many more exist. An example is reservoir computing, which relies on a combination of the feedforward network presented in Section 5.4 and a recurrent neural network, which itself is a feedforward network that allows for cyclical connections between nodes. These recurrent networks are used as a reservoir, feeding into a regular feedforward network. A more detailed introduction to this and other forms of neuromorphics, in addition to applications thereof, are found in [11].

# 6 Spin-Hall magnetoresistance in a spiking artificial neuron

In this section, we aim to present a ferromagnetic spiking neuron driven by the spin-Hall effect (SHE). Additionally, we will investigate the use of spin-Hall magnetoresistance (SMR) for obtaining the state of the ferromagnet, thus detecting spiking events in an entirely self-consistent manner.

## 6.1 Background

Recent studies have investigated the use of of spintronics for the creation of nanoscale artificial neurons [11], exploiting for instance the spin-Hall effects [14][64], magnetic tunnel junctions [65], spin-transfer torque [66], and many more [11]. The results are promising, and some of these devices are able to capture many of the properties of biological neurons [14].

Spin-Hall magnetoresistance (SMR) is a type of magnetoresistance that occurs in a conductor when it is in direct contact a magnetic material, due to the magnetization of said magnetic material. It is described as the simultaneous action of the spin-Hall effect (SHE) and inverse spin-Hall effect (iSHE) [33], causing a change in magnetization and pumping a charge current respectively. A detailed description is given in Section 3.5.3. However, since the resistance is now dependent on the direction of the magnetization, we can observe changes in the adjacent ferromagnet by measuring the resistance across the conducting layer. Such a change can, for instance, be the spiking events occurring in an artificial neuron, where the magnetization (confined to an easy-plane in our case) changes its direction by 180°. This should produce a characteristic pulse which we can detect. In addition, we wish to see if this output voltage pulse due to SMR can be used to spike other artificial neurons. If this is the case, it opens for the creation of a spiking neural network (SNN) entirely driven by SMR.

| Parameter | Value |
|---|---|
| Lattice constant | 12.376 Å (25 °C) [67] |
| Density | $5.11\,\mathrm{g/cm^3}$ (25 °C) [67] |
| Exchange stiffness | $3.65 \times 10^{-12}\ \mathrm{J\,m^{-1}}$ [68] |
| Gilbert damping parameter ($\alpha$) | $(5.3 \pm 0.1) \times 10^{-4}$ [69] |
| Saturation magnetization | 137 mT [68] |
| $G_r$ | $(3.8 \pm 1) \times 10^{14}\mathrm{S/m^2}$ [70] |
| $G_i$ | $(1.7 \pm 0.7) \times 10^{13}\mathrm{S/m^2}$ [70] |
| $\sigma$ | $10\,\mu\mathrm{S\,m^{-1}}$ [71] |

Table 1: Selected material parameters of a YIG

Figure 6.1 depicts our proposed setup. We investigate a $24 \times 12 \times 2.4$nm HM/FM bilayer. We consider a YIG-like isolating ferromagnetic mesh (red, key parameters listed in Table 2), with a $\hat{\mathbf{y}}$-directed hard-axis, and a $\hat{\mathbf{x}}$-directed easy-axis, forming an easy xz-plane, which we have defined by an explicit magnetic anisotropy within our simulation parameters. In reality, the easy- and hard-axes of a YIG usually spring from shape anisotropy and demagnetization. Additionally, there are some parameter differences from the real material: We have increased the damping parameter from $\alpha$ $10^{-4}$ to $\alpha = 0.02$, to decrease the sensitivity and thus make it easier to handle. The electrical conductivity has also been increased from the real value of $10\,\mu\mathrm{S\,m^{-1}}$ to $0.1\,\mathrm{S\,m^{-1}}$, which is a requirement of the numerical solver in use by BORIS computational spintronics. Nevertheless, this is much lower than the conductivity of the Pt-layer, $7\,\mathrm{MS\,m^{-1}}$, so we consider it an adequate value to make the ferromagnet isolating. The parameters from the Pt layer (HM) are sourced from BORIS's material database [72], with important quantities listed in Table 3.

Our simulations run as follows: We apply two electrodes to the short faces of the platinum mesh, through which we initially run a small, constant bias current $I_b$ in $\hat{\mathbf{x}}$, or an in-plane (IP) current. Due to the SHE in the Pt layer, there will be a deflection of the electrons in the $\hat{\mathbf{z}}$-direction, causing a spin-accumulation at the interface between the Pt and the YIG, where they will transmit or reflect depending on the magnetization $\mathbf{M}$ of the YIG, as explained in Section 3.5.3. This current will want to align the magnetization along the $\hat{\mathbf{y}}$-direction, however due to the strong hard-axis anisotropy the magnetization will instead be deflected within the easy-plane. The reason for the bias current is to make the simulation run faster: For no applied current, small fluctuations in the magnetization due to e.g. rounding errors will introduce a small charge current in the Pt layer. Without a ground electrode to flow towards, these electrons propagate in a complicated manner which slows down the simulation drastically. In a real implementation of an artificial

| Parameter | Value |
|---|---|
| Cell size | 12x12x6 Å (25 °C) [67] |
| Density | 5.11 g/cm$^3$ (25 °C) [67] |
| Exchange stiffness | $3.65 \times 10^{-12}$ J m$^{-1}$[68] |
| Gilbert damping parameter ($\alpha$) | 0.02 [69] |
| Saturation magnetization | 137 mT [68] |
| $G_r$ | $3.8 \times 10^{14}$S/m$^2$ [70] |
| $G_i$ | $1.7 \times 10^{13}$S/m$^2$ [70] |
| $\sigma$ | 0.1 S m$^{-1}$ |

Table 2: Selected simulation parameters for the YIG-like material. Changes will be reasoned in the main text.

| Parameter | Value |
|---|---|
| $\theta_{\text{SHA}}$ | 0.19 |
| $\theta_{\text{iSHA}}$ | 0.19 |
| $\sigma$ | 7 MS/m |

Table 3: Selected material parameters of Pt [72]

neuron, a bias current may be beneficial as it can make spiking easier, but it is not a necessity.

At specific intervals we will provide a rectangular current pulse, $I_s$ of much greater magnitude than $I_b$, for a short time $t_s$. This will momentarily increase the torque applied to the magnetization, breaking the equilibrium. The resulting angular momentum provided to the magnetization will carry it past the $\hat{\mathbf{z}}$-axis, whereupon it will accelerate towards the other ground state, oriented antiparallel to the initial ground state (usually we start in $-\hat{\mathbf{x}}$, so it will flip to $+\hat{\mathbf{x}}$). This we refer to as a spiking event. These magnetization dynamics will, due to SMR, change the resistance in the adjacent conducting layer. As the simulation aims to keep the current constant, the effect will be a fluctuation in the voltage, which we can isolate from the input rectangular pulse. The Pt layer is therefore used to both drive the magnetization dynamics and measure the results.

## 6.2 Results

We present our results for a system where the ferromagnet is of size $24 \times 12 \times 2.4$nm, initially biased by $I_b = 35\,\mu\text{A}$ for 50 ps. We then apply a $I_s = 600\,\mu\text{A}$ current spike for $t_s = 10$ ps, then return the current to the bias value. We apply two more of these pulses

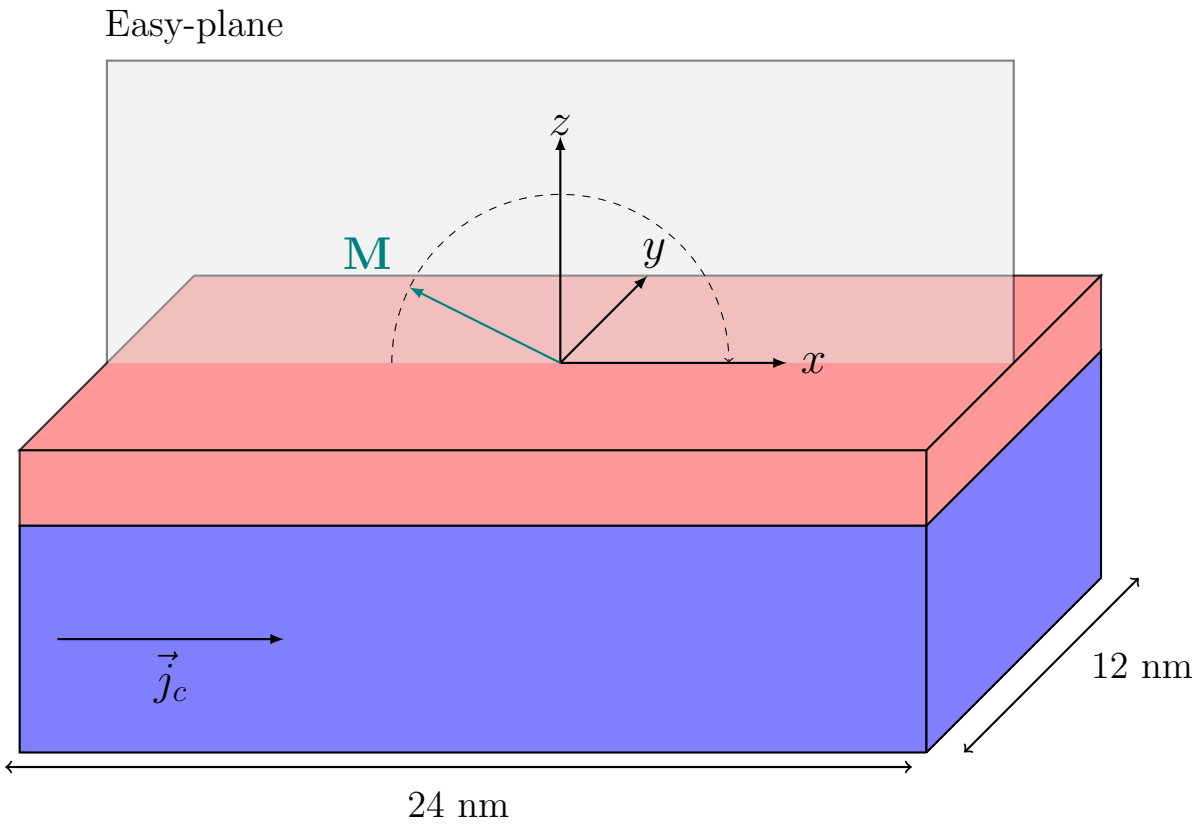



Figure 6.1: Simulation setup for the artificial neuron, consisting of a heavy-metal ferromagnetic (HM/FM, red and blue meshes respectively) bilayer, with the former having the properties of platinum (Pt) and the latter having parameters similar to a yttrium iron garnet (YIG) (actual parameters listed in Table 2). A charge current $\vec{j}_c$ is applied in the $\hat{\mathbf{x}}$-direction. The anisotropic easy-plane lies in the xz-plane, and the magnetization $\vec{M}$ oscillates within it.

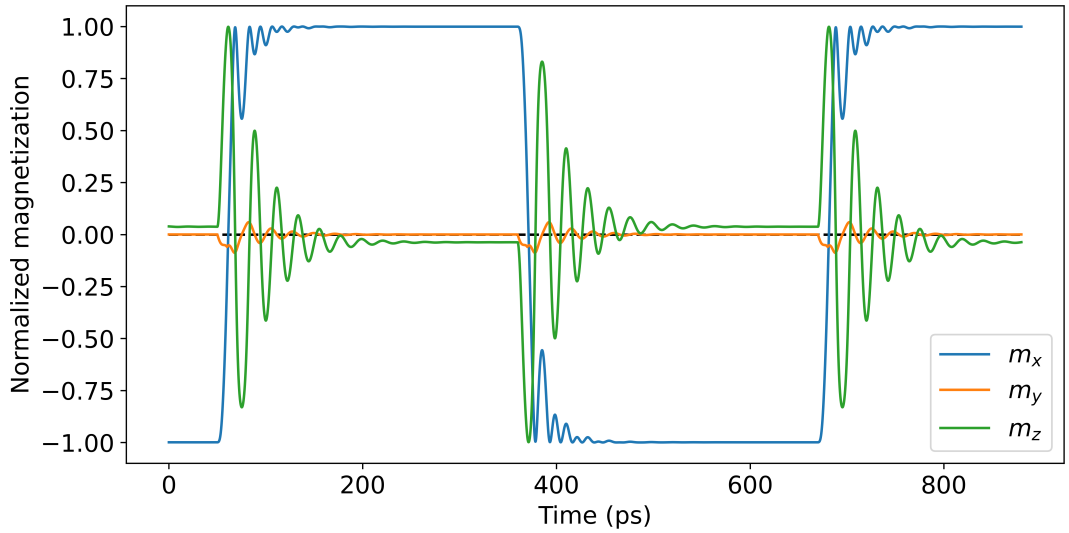
at 300 ps intervals, so that the magnetization is allowed to come to rest in its new ground state.

Figure 6.2a depicts the response of our artificial neuron to three current pulses, depicted by their normalized magnetization components. Figure 6.2b depicts the voltage output in red, overlaid the angle of the magnetization to the $\hat{\mathbf{x}}$-axis. The timing of the input current pulses, $I_s = 600\,\mu\text{A}$ over $t_s = 10\,\text{ps}$, are highlighted with black dashed lines. We observe that for the spiking events, the $M_x$-component almost immediately flips 180°. Once flipped, it overshoots the $\hat{\mathbf{x}}$-oriented ground state. It does not have sufficient angular momentum to pass the $\hat{\mathbf{z}}$-axis and spike again, and so despite overshooting it eventually relaxes into its new ground state, and we see that the oscillations disappear. After 300 ps, another pulse is applied, and we see once again that the magnetization flips 180° in a similar manner.

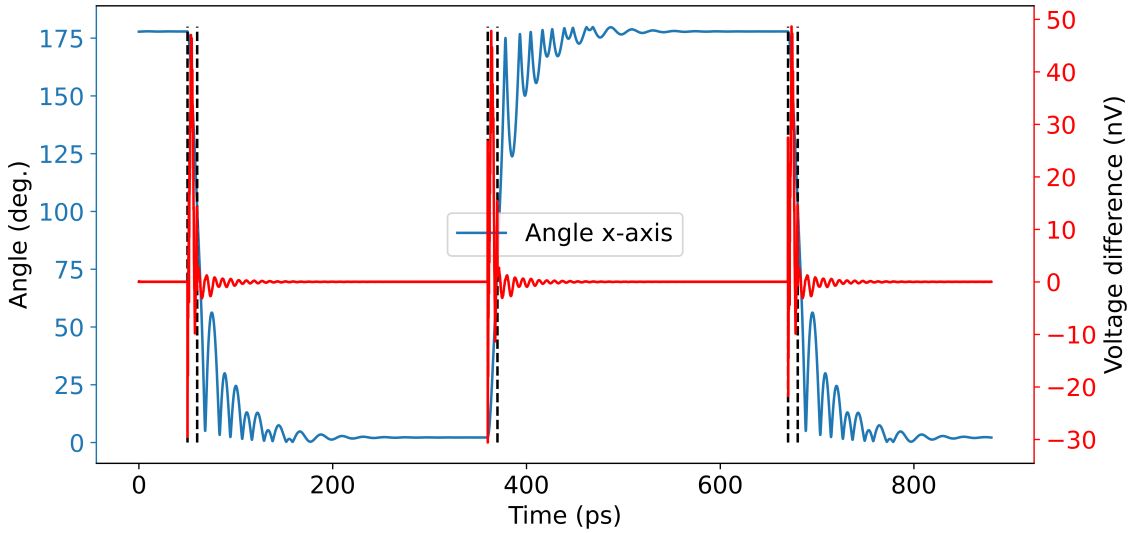
The spiking event itself occurs some time after the current pulse is applied: The magnetization responds immediately and starts moving, however it continues the spiking event even after the spiking current is removed. This is as expected, and a replication of the behavior presented in [14] and [64]. The delay exists simply because the torque itself is providing a large, momentary increase in the angular momentum, which carries the magnetization direction past the $\hat{\mathbf{z}}$-axis, at which point the spike occurs. This delay is referred to as **latency**, and is a key feature of biological neurons.

A side note is the behavior of the $\hat{\mathbf{y}}$-component. This is caused by the SHE, seeking to align the magnetization with the $\hat{\mathbf{y}}$-axis by pumping a spin current into the system. This alignment is prevented by the hard-axis anisotropy. Nevertheless, the current pulse momentarily increases this component, as one would expect. However, it is notable that this component responds in the exact same manner for $m_x = \pm\hat{\mathbf{x}}$, that is regardless of orientation.

The voltage response depicted in Figure 6.2b reaches a max of 50 nV for a brief moment during the application of the spiking voltage. After the application, the spiking event creates many smaller oscillations, before dropping to zero as the magnetization comes to rest. Note that, as mentioned prior, this voltage is isolated such that only the effect of SMR is displayed. By comparison, our input pulse is 600 µV, over 4 orders of magnitude larger than the output of this neuron. Furthermore, while the input is a single rectangular pulse, the output is a series of successively smaller peaks corresponding to the magnetization dynamics, each no longer than a few picoseconds. The output is therefore extremely small and short-lived compared to the input, and additionally of a different shape. This neuron is therefore unable to cause further spiking events downstream from itself. The magnitude of the output would need to be drastically increased, and the shape would also have to be altered. However, we are able to detect spiking events.




(a) Magnetization components as a function of time.




(b) IP voltage and $\hat{\mathbf{x}}$-axis angle as a function of time

Figure 6.2: Spiking of a $24 \times 12 \times 2.4$nm easy-plane ferromagnet interfacing a Pt layer, $I_b = 35\,\mu\text{A}$, $t_s = 10\,\text{ps}$, $I_s = 600\,\mu\text{A}$. a) components of the magnetization as a function of time, b) angle of the oscillator to the x-axis and voltage response due to SMR as a function of time. Black dashed bars indicate the timing of the input current pulses. The data in b) has been normalized, meaning that the bias value has been subtracted, and additionally the spiking pulse has been removed. Thus, all that is depicted is the response due to SMR.

## 6.3 Biological properties

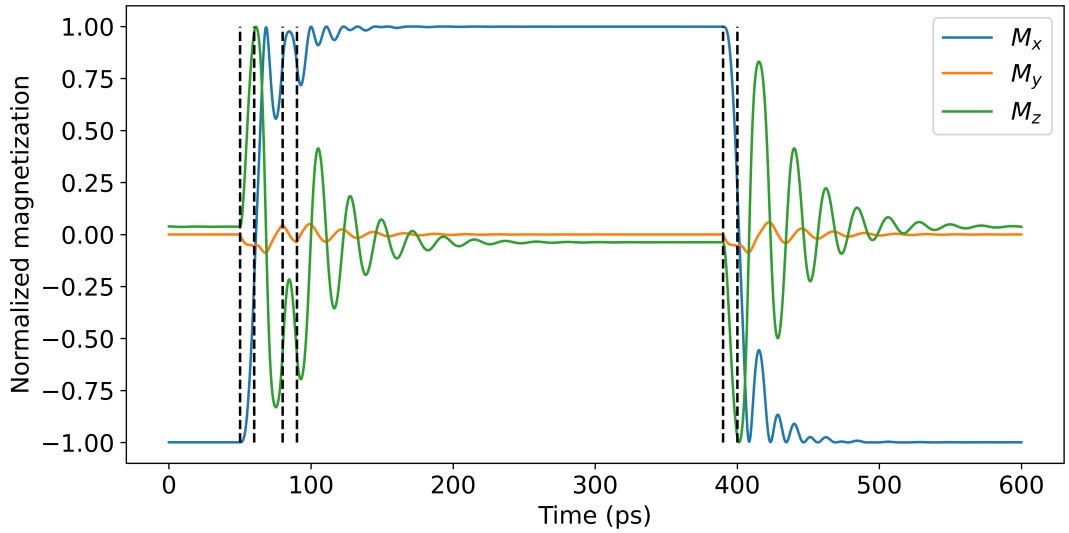
Having characterized the general spiking behavior, we now aim to relate key properties of the biological neuron to this spiking ferromagnetic setup and study the voltage response.

The latency property is immediately qualified by Figure 6.2, and we have discussed it previously: During the first spike, the magnetization has to be accelerated from the $-\hat{\mathbf{x}}$-direction and past the $\hat{\mathbf{z}}$-axis, at which point the anisotropy carries it over into the other ground state. As a result, the exact spiking event has some latency before it occurs. However, in Figure 6.2b, where the dashed black lines indicate the spiking event, we note that the majority of the dynamics and voltage output happen during the arrival of the pulse, and the aftermath is less pronounced. This means that while there is physical latency in how the magnetization responds to the pulse, it is not very prominent in the output.

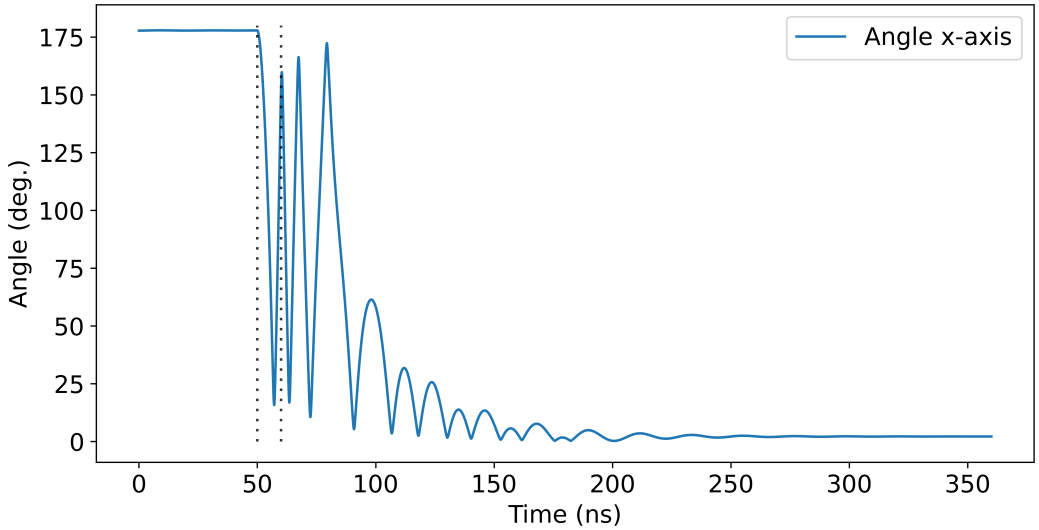
To tackle two more properties, namely bursting and refractory periods, we must create two new setups. Figure Figure 6.3 depicts the magnetization response in such simulations. We will tackle these one by one.

The refractory period is explained in Section 5.5. This is the property whereby a neuron is not susceptible to spike if it has recently spiked without the necessary time for relaxation. To test this, we apply a secondary pulse before the neuron is fully relaxed into the other ground state. There is no exact time for when this occurs, but by looking at the lifetime of the oscillations in Figure 6.2a, we decided on a 20 ps gap between the pulses. The applied pulses were otherwise of the exact same magnitude and time span. Figure 6.3a depicts our results, and they are positive: The neuron is still oscillating around its new ground state by the time the second pulse arrives. Compared to what appears to happen in Figure 6.2a, the oscillation is influenced by the spike and momentarily reversed, before continuing as prior with no spiking event occurring. In effect, the aftermath of the spiking event temporarily blocks any further events from occurring.

Bursting is effectively the opposite of a refractory period, whereby a single input pulse is sufficient to make the neuron fire repeatedly. As we have established that a 600 µA pulse over 10 ps is sufficient to cause a single spiking event, increasing either of these quantities should be sufficient to cause bursting. We choose to increase the magnitude of the pulse, as it can represent the arrival of several pulses from other neurons at the same time. More specifically, we increase the spiking current $I_s = 1800\,\mu\text{A}$. The results are depicted in Figure 6.3b, where we have elected to plot only the angle to the $\hat{\mathbf{x}}$-axis as it best illustrates what is otherwise very cluttered behavior. The thinly dotted line marks the arriving pulse. We see immediately that the angle oscillates greatly, between approximately 0° and 180° with sharp peaks. Due to the wrapped angles, these sharp peaks mean that the oscillator is making several full cycles. The large input current pulse also created a larger $\hat{\mathbf{y}}$-axis deflection of the magnetization, hence why the $\hat{\mathbf{x}}$-angle is not




(a) Magnetization as a function of time for the refractory case. Loosely dashed lines mark the two incoming pulses, separated by 20 ps in time, though of the same magnitude of $I_s = 600\,\mu\text{A}$ and duration $t_s = 10\,\text{ps}$.




(b) Angle between magnetization and $\hat{\mathbf{x}}$-axis. Loosely dashed lines mark the incoming pulse, $I_s = 1800\mu\text{A}$, $t_s = 10\,\text{ps}$.

Figure 6.3: Magnetization response illustrating the cases of refraction and bursting, both features of real neurons, implemented on our spiking neuron. Parameters are listed in the subcaptions.

reaching 0° or 180° fully. To recap, applying a large enough current creates several spiking events, compared to just a single spiking event. Therefore, this artificial neuron exhibits bursting.

The leakiness of this neuron is shown through all figures, but particularly Figure 6.2a: After a spiking event, the magnetization has enough angular momentum to overshoot the ground state, though not enough to cause a secondary spiking event. Subsequently, it slowly relaxes into the ground state, oscillating around it with a decreasing amplitude. Effectively, the angular momentum provided to it by the initial spiking event is leaked. This is no different from an insufficient input pulse momentarily exciting the magnetization from its ground state, and it subsequently relaxing once the pulse has passed. We have therefore implicitly shown that this neuron is leaky.

## 6.4 Outlook

Through the use of the two spin-Hall effects and an easy-plane anisotropy with a strong hard-axis, we have presented a ferromagnetic spiking neuron that responds to input current. If this current is sufficient, the neuron spikes, turning the magnetization 180° and creating an output current pulse. If it is insufficient, the neuron relaxes towards its ground state, thus acting like a leaky neuron. In addition to the leakiness, we have also verified that the artificial neuron exhibits bursting by spiking several times for a large input, and a refractory period wherein a recently-spiked neuron does not respond to an identical input. The latency property, wherein there is some latency between an input current and a spiking event, does exist, however it is very short and the neuron responds very quickly to the input current. Inhibition in this neuron was not tested.

However, there are several fundamental drawbacks. A first flaw is that there is nothing to stop a spiking pulse $I_s$ from propagating: If we apply an $I_s$ that is insufficient to cause a spiking event, the pulse will nevertheless continue down the conducting layer and potentially influence other neurons. This is not what we desire. To realize this neuron, a means of blocking the propagation is necessary. For a larger scale device, it may be simple to just manufacture a gap in the conductor and place the ferromagnet as a bridge between the two separate conductors, now called A and B for convenience. The input is now applied to A. This will ensure that the pulse is not allowed to propagate, as the ferromagnet is isolating. However, this workaround comes with a separate issue: Although the input pulse is no longer allowed to propagate, it can still cause magnetization dynamics within the ferromagnet. In accordance with the inverse spin-Hall effect, these dynamics will pump a current into conductor B, though smaller and of different size than that of a successful spike. In short, an insufficient $I_s$ will nevertheless cause an output, and the neuron is not all-or-nothing in terms of its propagation. We already see this effect in e.g. Figure 6.2b: The moment the spike is applied, the magnetization begins rotating rapidly,

and the result is a spike in the voltage. After the pulse has passed, the torque applied to the angle disappears and so the change in angle also drops, resulting in a lower voltage output after the spiking event. For a real neuron only when its membrane potential is higher than a certain threshold should there be an output. Therefore, to realize this neuron, a mechanism for blocking the propagation of unwanted signals is necessary. This will also be necessary for shaping the output pulse, as at present it is a series of pulses, each of small magnitude and duration, which does not match the input square pulse of large magnitude.

# 7 Spin Hall nano-oscillator

In this chapter, we will investigate simple systems of spin Hall nano-oscillators (SHNOs) as an extension of the spiking neuron presented above. More specifically, we will present an open chain of isolating, ferromagnetic easy-plane SHNOs (up to a total of 5 oscillators) placed on a common platinum layer through which we run a charge current. These oscillators will be placed close enough to interact and couple to each other. This interaction will be primarily due to the demagnetization field, but through the common platinum bar there will potentially be a secondary coupling through iSHE. As far as we are aware, the iSHE coupling has not been studied in an isolated manner.

## 7.1 Background and setup

Spin Hall nano-oscillators are a promising device within the field of spintronics, and especially for use in neuromorphic computing [64][12][61][13][73][74]. Their operation is similar to the spiking bilayer presented above above: A current density greater than some critical value is driven through a conducting layer, and the SHE creates a spin accumulation at that interfaces with a neighboring ferromagnet, causing magnetization dynamics. Unlike the spiking ferromagnet, we run this above-critical current continuously, turning the spiking into an oscillation. These devices are similar to spin torque nano-oscillators (STNOs). STNOs are driven by spin-transfer torque from a current going through the ferromagnet itself, distinct from the bilayer required for an SHNO. The bilayer setup has many benefits, including ease of manufacture [75] and synchronization strength strength. It has been shown that SHNOs can synchronize for an 8x8 array of oscillators [73] as well as for up to 50 oscillators in a 1D chain [75]. Additionally, simulations have shown that SHNOs can be used for neural network applications [12][61][49][13]. Recent research has shown an additional improvement in signal strength and quality with the use of an easy-plane SHNO (EP-SHNO) [76], but synchronization of this variant has, to our knowledge, not yet been studied.

Our basic system is depicted in Figure 7.1. It consists of a platinum (Pt) layer (blue mesh), on top of which we have placed two insulating ferromagnets with an easy-axis in the $\hat{\mathbf{x}}$-direction and a hard-axis in the $\hat{\mathbf{y}}$-direction (red meshes). A current is led through the conducting layer which creates a spin accumulation near the interface due to SHE. In addition to SHE, there are two other effects that play a pivotal role: First, the demagentization field (demag), the magnetic field generated by the ferromagnet's internal magnetization, will couple to all other magnets, attempting to align their magnetizations with each other. Secondly, the magnetization dynamics of SHNO1 will, via iSHE, pump a charge current into the conducting layer. This pumped charged current will flow in the $\hat{\mathbf{x}}$-direction and subsequently interact with SHNO2 alongside the applied $\vec{j}_c$. This

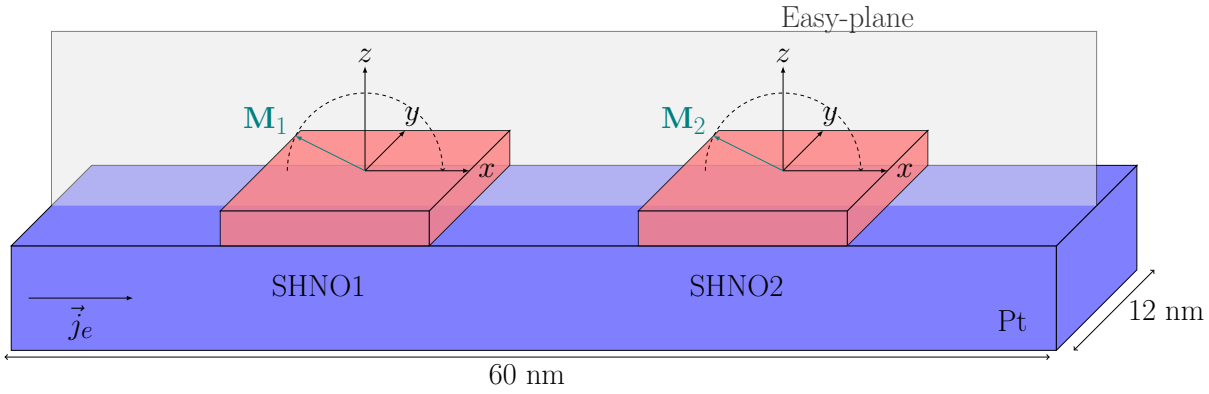



Figure 7.1: Simulation setup of the connect 2SHNO system. $\vec{M}_1$ and $\vec{M}_2$ represent the magnetization of SHNO1 and SHNO2 respectively, almost entirely confined within the anisotropic easy-plane, which has a normal vector along $\hat{\mathbf{y}}$. $\vec{j}_c$ is the applied charge current.

potentially provides an additional coupling between the oscillators. Out of these two couplings, only demag is known to provide synchronization [75][73]. The coupling due to iSHE will be asymmetric, as any current pumped by SHNO2 will not affect SHNO1 to the same degree.

The material parameters for the ferromagnets are the same as in Table 2, and for the platinum they are provided by the BORIS Computational Spintronics material database [72]. A set of relevant parameters are listed in Table 3. Additionally, we will also run systems of more than two oscillators, but they will be placed and numbered in accordance with Figure 7.1.

We define three measures: Angle difference, order parameter, and voltage difference. Angle difference refers to the angle between the averaged magnetization of each SHNOs compared to the others, averaged over a full period (due to the $\hat{\mathbf{x}}$-directed easy-axis, the torque varies throughout a full cycle, and thus so does the frequency). When this value is changing, the oscillators have different average frequencies. When it is aligned and non-zero, the oscillators are phase locked. When they are aligned and zero, they are entirely in-phase. We will sometimes refer to this as the phase difference. The order parameter has been introduced and discussed in Section 4. The voltage difference is our realistic observable, and the same as what is used with the spiking ferromagnet in Section 6.

Finally, we define the terms upstream and downstream as positions relative to the flow of current. As an example, in the 2SHNO case depicted in Figure 7.1, SHNO1 is the upstream oscillator and SHNO2 is the downstream oscillator. For systems with more than two oscillators, we will avoid using this terminology for anything but the first and last oscillators, and instead rely on the oscillator number.

## 7.2 Results

In the simulations presented in Figures 7.2 and 7.3, we show the angle difference between SHNOs (high and low damping respectively) in three geometries subject to three different combinations of effect: 2SHNO, 3SHNO and 4SHNO (rows), subject to either both demag and iSHE, only demag, or only iSHE (columns). The simulations are not run for the same length of time, as the behavior of some systems is slower than others. SHE is always enabled, as otherwise the oscillators do not have an applied torque. We initialize all oscillators to their $-\hat{\mathbf{x}}$-oriented ground state, aside from SHNO1 which is initialized 20° in the $+\hat{\mathbf{z}}$-direction, such that it leads the oscillation. The reasoning is that were we to start them in all the ground state, it would be more difficult to determine how the phases evolve to attain synchronization.

We first study the case of high damping in the ferromagnets, Figure 7.2, where $\alpha = 0.5$. For the all-effects 2SHNO case, Figure 7.2a, the angle difference drops from 20° to approximately $-7.5°$ within 30ns. The negative angle means that SHNO2, the downstream oscillator, leads the oscillation. This implies that though SHNO1 leads the oscillation initially, SHNO2 eventually takes the lead. When disabling the iSHE and subjecting the oscillators to only SHE and demag, shown in Figure 7.2b, the behavior is qualitatively preserved: The SHNOs synchronize with a phase difference. However, disabling demag and relying exclusively on iSHE yields a different, linear decrease in the angle difference, depicted in Figure 7.2c. The likeliest explanation is that, in this case, there is effectively no coupling trying to force the oscillators into synchronicity, and they are instead subject to different torques due to the combined spin-Hall effects. The linear decay implies that the downstream oscillator has a higher frequency on average.

We also observe the same behavior in the cases of more oscillators subject to iSHE only, Figure 7.2f and Figure 7.2i: All of the angle differences decrease linearly, eventually becoming negative and thus the downstream oscillator leads. Additionally, the darker lines like SHNO1-SHNO4 and SHNO2-SHNO4 have the most negative slopes, indicating that the further downstream the oscillator, the higher its frequency. For the case of high damping with $\alpha = 0.5$, iSHE introduces an extra torque on all downstream oscillators, causing an increase in frequency that grows in magnitude for every oscillator added.

It appears that the number of oscillators has an effect on the synchronization for high-damping: In the 2SHNO case, Figure 7.2a and 7.2b, the angle difference very simply drops from its initial value of 20° to its phase-locked value of $-7.5°$ within approximately 60 ns. In the 4SHNO case, Figure 7.2g and 7.2h, the magnetization initially oscillates wildly, before finding a configuration wherein the oscillators become mutually phase-locked, taking approximately 200 ns before the stability begins to emerge. However, in the 3SHNO case, Figure 7.2d and 7.2e, none of the oscillators synchronize at all, and the angle differences change throughout the whole 400 ns-simulation. It should be noted that there is a certain

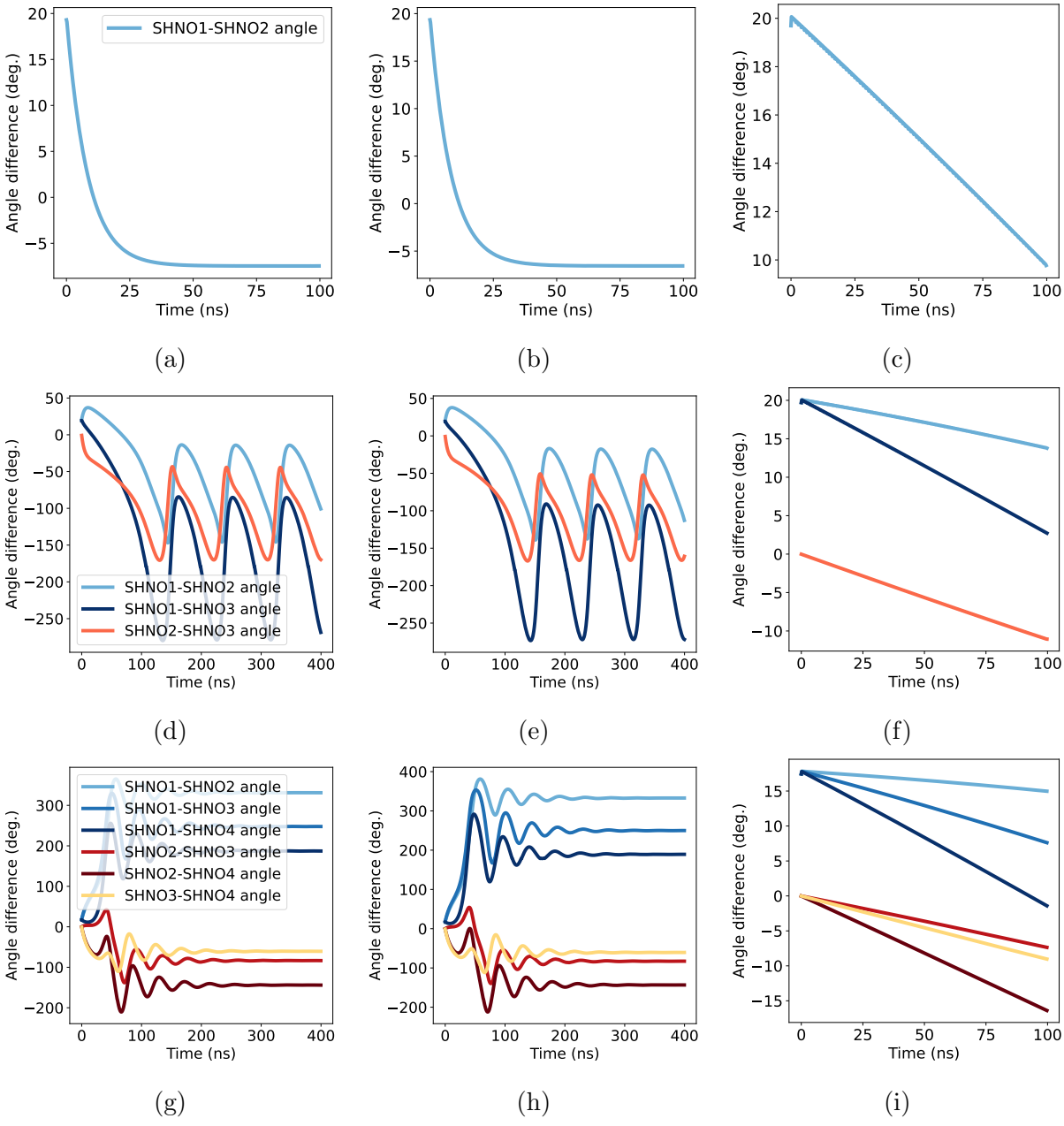



Figure 7.2: Angle difference as a function of time for the high-damping case, $\alpha = 0.5$. Left column: All effects. Middle column: Demag only. Right column: iSHE only. Top row: 2SHNO. Middle row: 3SHNO. Bottom row: 4SHNO.

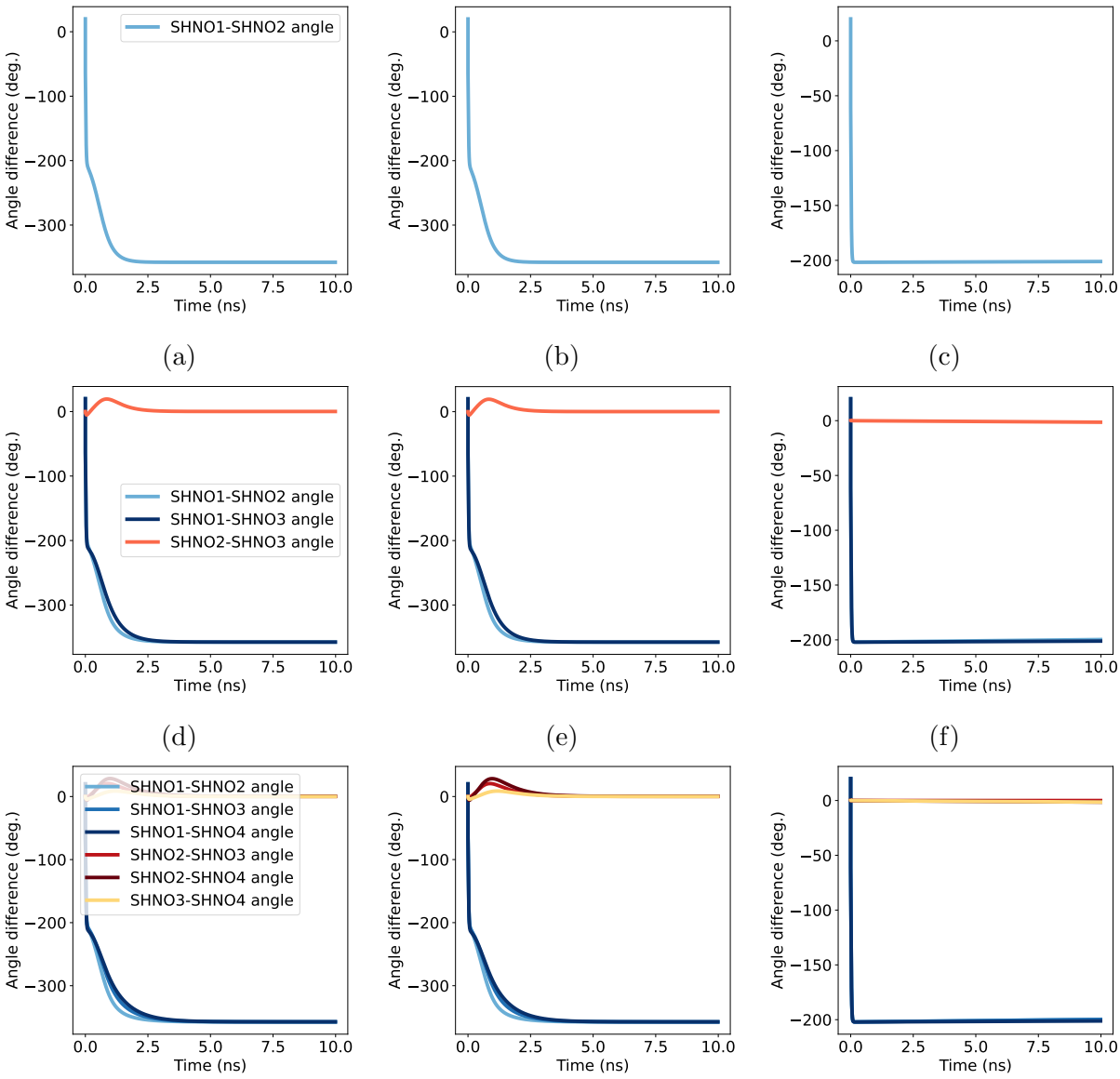



Figure 7.3: Angle difference as a function of time for the low-damping case, $\alpha = 0.02$. Left column: All effects. Middle column: Demag only. Right column: iSHE only. Top row: 2SHNO. Middle row: 3SHNO. Bottom row: 4SHNO.

degree of order: The long-term changes in phase for the 3SHNO system shows a periodicity. The SHNO1-SHNO3 angle, for instance, oscillates with a period of approximately 90 ns, much longer than the period of any single SHNO. Additionally, the SHNO1-SHNO2 has a magnitude and periodicity close to that of the SHNO2-SHNO3 angle, but seemingly delayed, with the peaks separated by around 20 ns. This is notable, as it implies the number of oscillators has a profound impact on the long-term synchronization of the system in the high-damping case, but also that orderly, yet desynchronized states can exist.

For highly dampened EP-SHNOs, demagnetization proves to be the most dominant effect, and for 2SHNO it serves to phase-lock the oscillators. For 4SHNO, it does the same, creating a phase-locked pattern for the oscillators. However, rigid synchronization is lost for three oscillators, as evidenced by Figures 7.2d, 7.2e. When considering the iSHE-only cases above, we noted that the linear behavior of the angle difference implied that the oscillators were not really interacting, and were simply subject to different, yet time-independent, average frequencies.

In the case of low damping, Figure 7.3, the dynamics are more predictable: A degree of synchronization is universal for all geometries, including 3SHNO, which was not the case in the high-damping case shown in Figure 7.2d. This is not clear in the subfigures of Figure 7.3, and so we will call attention to it: The angles differences are depicted without any wrapping, meaning that what is observed is the actual difference in angles after certain number of oscillations. What we observe is when demag is enabled (the left and middle column), the angle difference between SHNO1 and all other SHNOs drops rapidly from 20° to approximately −202°, then decays towards approximately −360°, whereupon it synchronizes. The oscillator therefore end up synchronized with nearly no phase difference, however SHNO1 attained this by lagging behind for 1 cycle. This occurs in all cases, but only for angles that exist with respect to SHNO1, and it is therefore likely tied to the initial angle of 20°: The torque from SHE counteracts that of the anisotropy, and so since SHNO1 is started with an angle to the $\hat{\mathbf{x}}$-axis, the magnetization experiences a smaller net torque for a shorter time compared to SHNO2, SHNO3 etc. Therefore, it accumulates less angular momentum in the first few oscillations, and the other SHNOs start to lead the oscillation. After several full cycles, the oscillators are in their steady states with the torque unable to accelerate the oscillations any further. Here, the long-term synchronization effects come into play and slowly synchronize the oscillators, which it achieves by further lagging SHNO1. This is also the case in both the 3SHNO and 4SHNO cases, Figure 7.3d, 7.3e, 7.3g and 7.3h, where all angles with respect to SHNO1 (blue lines) lag behind for one full cycle.

Additionally, the low-damping case synchronizes faster (approximately 3 ns for two EP-SHNOs compared to 30 ns for high damping). It is likely that the speed is tied directly to the damping, as the damping acts as a braking force on these oscillations, partially counteracting the long-term synchronizing force between the oscillators. It may also be

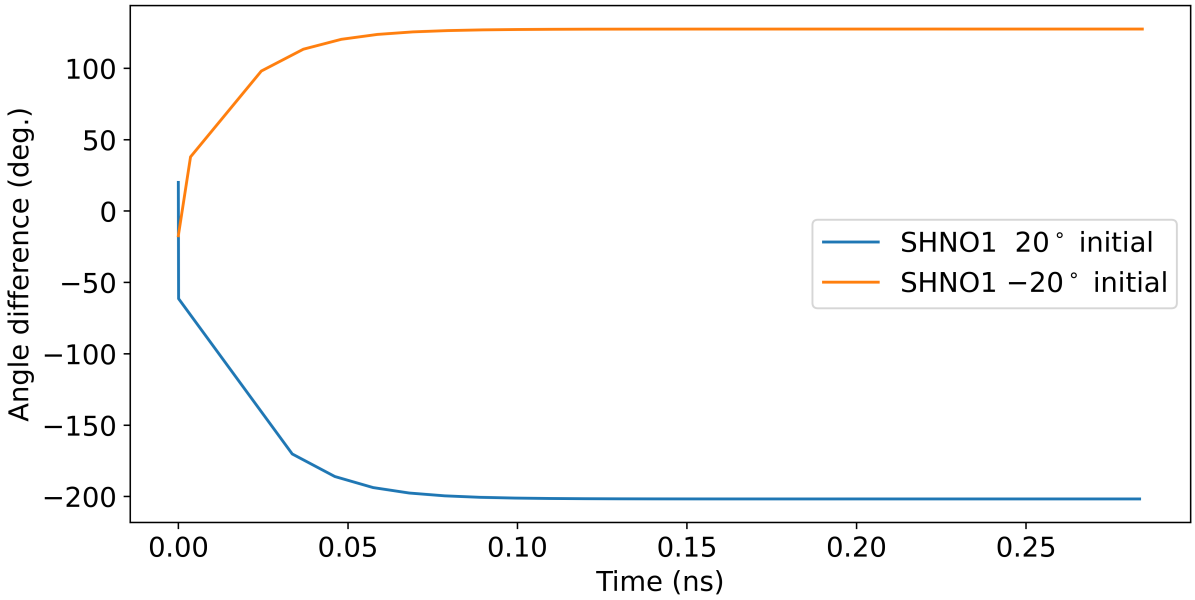



Figure 7.4: Angle difference as a function of time, 2SHNO with $\alpha = 0.02$, iSHE only (similar to Figure 7.3c), with SHNO1 initialized to 20° and −20°.

the cause of the unexpected final layout of the oscillator phases depicted in Figure 7.2g, as this has disappeared for low damping.

The iSHE-only cases of low damping, Figures 7.3c, 7.3f and 7.3i, show a different behavior from that depicted in the corresponding high damping cases: The leading oscillator, SHNO1, immediately and rapidly drops to having an angle of approximately −200° with all other oscillators, for all times. This is the same rapid drop we observe in all of Figure 7.3, however it does not continue to drop towards −360°. This is likely a result of the same effect observed in the corresponding high-damping case, e.g. Figure 7.2i, where there was no actual coupling seeking synchronization between the oscillators, but instead the presence of iSHE created larger frequencies downstream. In this case, the angle difference is caused by the aforementioned difference in initial torque, wherein SHNO1 experiences a smaller torque for a shorter time due to its angle with the $\hat{\mathbf{x}}$-axis. In this case, this difference in torque lasts until SHNO2 leads by approximately 200°, at which point the steady-state behavior is settling in and SHNO1 attains approximately the same frequency. By comparison, the oscillators which start in the same ground state, SHNO2, SHNO3 and SHNO4, do not show any significant angle difference, as they are subject to very similar torque.

To prove this hypothesis, we run a simulation where SHNO1 is set to −20°, that is below the $-\hat{\mathbf{x}}$-axis in the xz-plane. Figure 7.4 shows the results of this simulation, and it is as we anticipate: SHNO1, now subject to a larger torque for longer, leads by 127°. We note that this is asymmetric around the x-axis, likely a result of the different torques experienced: At −20°, SHNO1 experiences a momentary positive torque w.r.t the

oscillation direction due to the anisotropy while SHNO2 does not. In the 20° case, both oscillators are subject to only a negative torque due to anisotropy until they cross the $\hat{\mathbf{z}}$-axis.

This behavior underlines the importance of damping in minimizing the effect of the dynamics occurring prior to synchronization, as depicted in Figure 7.2i: Although we also have the initial 20° angle difference for SHNO1, there is no large drop early in the simulation caused by the difference in torque.

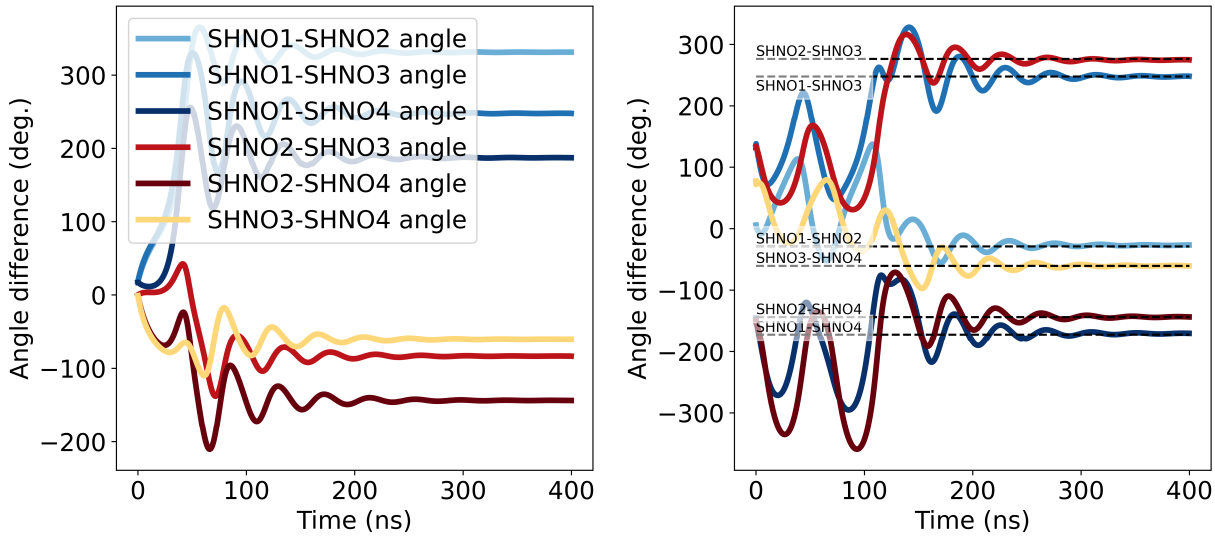
### 7.2.1 Sensitivity to initial conditions

In the high-damping 4SHNO case depicted in Figure 7.2g, we noted that the synchronization took a long time in comparison to other cases, including both the high-damping 2SHNO case (Figure 7.2a) and all the low-damping cases (Figure 7.3). In addition, the final synchronized state had the oscillators phase-locked with a range of phase differences. This begs the question of whether or not the synchronization is universal (that is, occurring for all initial configurations of the 4SHNO system), or if it only occurs for certain initial phase differences. If it is universal, are the final phase differences dependent upon the initial phase difference, or unique to the system and its properties (coupling constants, natural frequencies, damping etc.) and thus the same for all initial phase differences? Additionally, we want to see if the synchronization time is dependent on the initial order of the system.

We have tested the case where SHNO1 has a 20° angle to the $-\hat{\mathbf{x}}$-axis while all other oscillators are parallel to the $-\hat{\mathbf{x}}$-axis. To determine the dependence on initial conditions, we will initialize all the oscillators to random directions within the xz-plane, and then run the simulation for 400 ns. For brevity, we refer to these results as "baseline" and "randomized" respectively.

Figure 7.5 depicts the baseline and randomized systems side-by-side. We have added dashed lines Figure 7.5b to indicate the final (unwrapped) angles of the baseline system for ease of comparison. Some angles have been adjusted by a full cycle ($\pm 360°$) to illustrate that the system phase-locks with the exact same angle differences as discussed previously. However, the necessity of the adjustment illustrates the important point that the oscillators take different, likely lower-energy paths, to achieve the same synchronization pattern. Ergo, the system appears to be insensitive to the initial conditions and instead likely synchronizes on the basis of its parameters. The choice of path to synchronization is something that we have already observed in the case of the low-damping 2SHNO simulation, Figure 7.3a, wherein the angle difference immediately drops to $-200°$, whereupon the shortest path to its synchronization is by continuing to lag until the angle difference is $-360°$.

Additionally, the randomized system takes approximately 100 ns longer than the base-




(a) Initial angle of SHNO1 at 20°, all others 0° (baseline).

(b) All initial angles randomized. Dashed horizontal lines indicate the final angles (possibly adjusted) of 7.5a, which we are comparing to.

Figure 7.5: Comparison between the baseline and randomized high-damping 4SHNO case.

line system to reach the same level of synchronization, likely a result of the initial state being less ordered. It is reasonable to expect that any randomized initial condition will take approximately the same time to synchronize.

More simulations are necessary to conclude that the layout of the phases is decided by the system parameters alone rather than the initial configuration of the angles, but it seems likely that this is the case.

### 7.2.2 Five oscillators

The fact that both the high-damping 2SHNO and 4SHNO cases synchronize while the 3SHNO phase differences oscillate periodically, points to a phenomenon related to the number of oscillators. To be able to give any meaningful answers, we need to simulate another odd-numbered system. The next simplest is 5 oscillators on an open chain, 5SHNO. All other parameters are the same as in the case of the randomized high-damping 4SHNO system presented above, including the randomized initial angles of the oscillators.

Figure 7.6 depicts our results in the form of angle differences, as before. This simulation is run for 700 ns, though the first 550 ns are omitted. Additionally, these angles are wrapped between $-180°$ and $180°$ so they give an accurate picture of the final state without considering the evolution that led to it. This is unlike all cases depicted previously, however we are only interested in the long-term synchronization of this system. Additionally, some lines are made thinner and/or dashed so they are easier to differentiate.

What is immediately apparent is that even in the final state, the average phase dif-

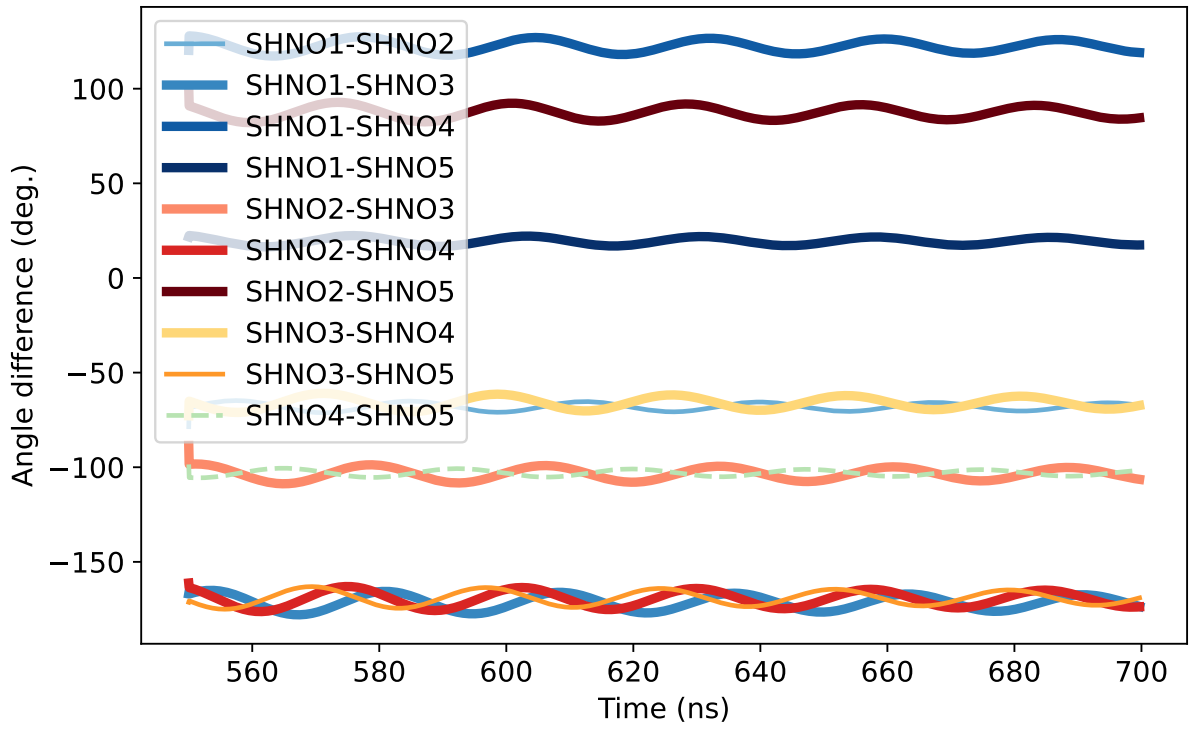



Figure 7.6: Angle difference as a function of time for the 5SHNO high-damping ($\alpha = 0.5$) case.

ferences oscillate. Normally there may be changes in frequency throughout one cycle causing small fluctuations in the angle difference. These average out and disappear when depicting the average phase difference. In this case, the phase differences exhibit long-term oscillations of small amplitude, lasting much longer than one cycle, on the order of 20 ns. This is important, as it implies the oscillators are not properly phase-locked, but nevertheless synchronized in a periodic manner.

The phase differences are also observed to cluster together. For instance, SHNO1-SHNO2 has almost the same phase difference as SHNO3-SHNO4, $-67°$. Other examples include SHNO2-SHNO3 and SHNO4-SHNO5, both of which have a phase difference of approximately $-103°$. It should be immediately apparent that both of these clusters are comprised of nearest neighbor oscillators. The cluster at the bottom of Figure 7.6 consists of SHNO1-SHNO3, SHNO2-SHNO4, and SHNO3-SHNO5 all of which are pairs of next-nearest neighbors with their phase differences clustered around $-170°$. The final phase differences are SHNO1-SHNO4, SHNO2-SHNO5 and SHNO1-SHNO5, (next-)-next-next-nearest neighbors, which show no sign of clustering. It is likely that they are too far away to meaningfully interact, and the final configuration is decided by the other oscillators.

It should be noted that due to the phase difference between SHNO1-SHNO3 and SHNO3-SHNO5 both being $-170°$, the phase difference between SHNO1-SHNO5 is nearly $360°$. In other words, the oscillators at the edge of this open chain are nearly in-phase.

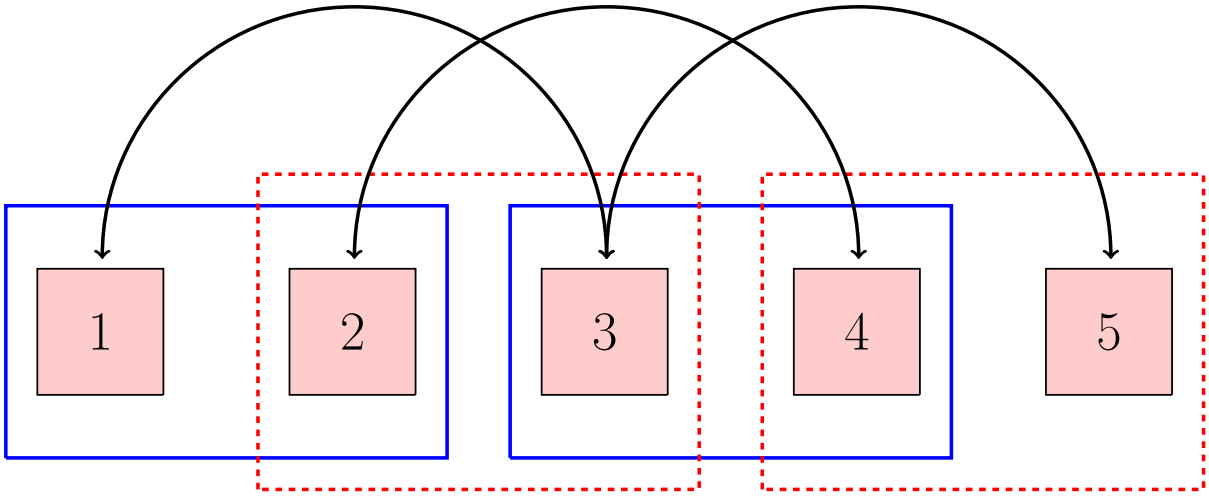



Figure 7.7: Synchronization groups of the 5SHNO case depicted in Figure 7.6. Blue solid lines indicate pairs synchronized with a $-67°$ phase difference, red dashed lines indicate groups with a $-103°$ phase difference, and arrows indicate synchronized next-nearest neighbors with $-170°$ phase difference.

For clarity, the layout of these clusters is depicted in Figure 7.7: All groupings of the same color have the same phase difference. For instance, as discussed previously, the SHNO1-SHNO2 phase difference is the same as SHNO3-SHNO4, ($-67°$) hence they are both blue in this diagram. The same goes for the red dotted lines. The arrows indicate the next-nearest coupled pairs like SHNO1-SHNO3. This diagram shows that the layout of the synchronization is dependent on the position: All possible next-nearest neighbor pairs synchronize with the same angle difference. However, we get two types of nearest-neighbor couplings, blue and red, each with its own characteristic phase difference. The system is therefore not symmetric in its synchronization pattern.

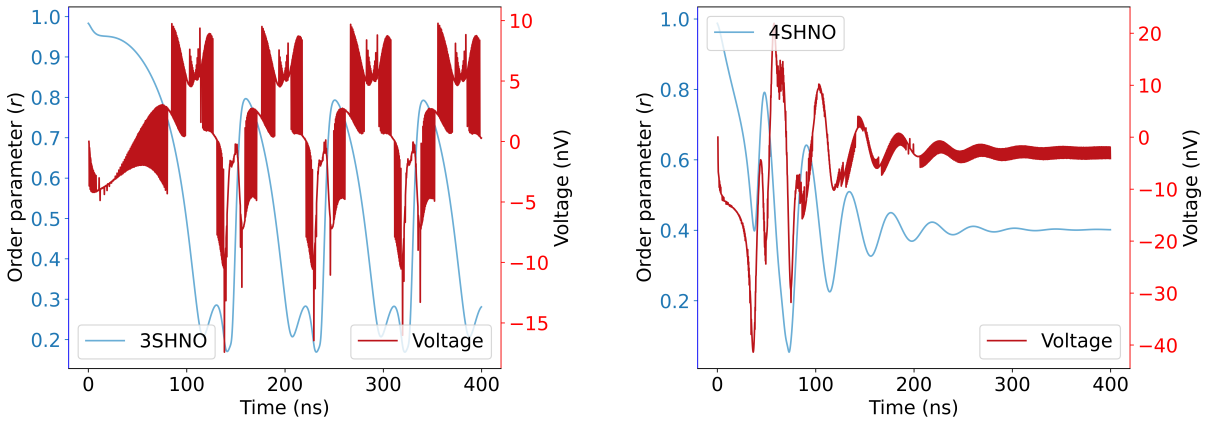
We note that the pairs within each cluster appear to respond to each other, causing a ripple in the phase differences. Take the bottom cluster in Figure 7.6: The SHNO2-SHNO4 phase difference (marked in dark red) lags behind the SHNO3-SHNO5 phase difference (thin orange), reaching its peak some nanoseconds later. And similarly, the SHNO1-SHNO3 phase difference lags the SHNO2-SHNO4. Furthermore, looking at the nearest-neighbor clusters (like SHNO2-SHNO3 and SHNO4-SHNO5), we note that they also respond to each other, each having a positive amplitude while the other is negative. Additionally, it appears this response shares the periodicity with that of the next-nearest neighbor cluster: Each full cycle of the SHNO3-SHNO5 phase difference corresponds to a full cycle of the SHNO4-SHNO5. The same goes for the SHNO1-SHNO2 and SHNO3-SHNO4 clusters. However, these nearest neighbor pairs do not lag each other in the same way the next-nearest neighbors do: The response is immediate and the phase differences are in anti-phase, with one phase difference increasing as another decreases correspondingly.

It is likely that this is an effect of the odd number of oscillators, as it is also what we observe in the high-damping 3SHNO case depicted in Figure 7.2d: The SHNO1-SHNO2 angle lags that of the SHNO2-SHNO3 by some nanoseconds, but it is nevertheless a periodic response. In that system, the amplitude of the phase difference oscillations is much greater than observed in this 5SHNO system, perhaps caused by the small number of oscillators.

The synchronization of weakly-coupled open chains of oscillators has been studied in [77]. Their work considered stochastic oscillators influenced by an explicit time delay. This time delay was applied directly to the coupling mechanism. They found that, in the case of a short time delay, the oscillators become phase-synchronized as the coupling strength increases. This means that they are periodically in-phase, either by the frequencies being the same or an integer multiple of each other [78]. Their amplitudes are unconstrained and may be very different. In contrast, when the time delay is large, alternate oscillators become phase-synchronized while neighboring oscillators become lag-synchronized [77]. Lag synchronization means that one oscillator shadows the exact behavior of another [78], very much like phase-locking. Furthermore, the number of oscillators plays a role in the behavior of the edge cases: For an even number of oscillators with a large time delay, the oscillators at the edges will be in lag synchronicity, while for an odd number the edges will be phase synchronized.

Though the SHNOs under consideration are neither intrinsically stochastic nor subject to explicit time delay, one can relate the time delay parameter to the damping: Lowering the damping increases the speed by which an oscillator responds to external torque, including the synchronizing forces. It has already been shown that the synchronization time for low damping oscillators is much lower than their high damping counterparts. Additionally, we have already made the observation that the low damping SHNOs always synchronize and do so in phase. By contrast, the synchronization of high damping SHNOs appears to end in phase-locked states that seemingly depend on the number of oscillators. Figure 7.2a and Figure 7.5b show that the edge oscillators of the 2SHNO and 4SHNO chains respectively enter phase-locked (i.e. lag synchronized) states. We also note that in the 4SHNO case depicted in e.g. Figure 7.5b, the final angle between SHNO1-SHNO4 is $-180°$. This is important, because we previously highlighted that in the 5SHNO case, SHNO1-SHNO5 are close to being in phase in comparison to all other pairs of oscillators in the system. The slight offset may be caused by iSHE, as it is shown to increase the frequency of downstream oscillators and may provide a considerable torque on SHNO5. Nevertheless, this is broadly in line with what was presented in [77].

The relationship between the number of oscillators and their synchronization behavior, and that between the time delay and the coupling, are by observation. For instance, the time delay is applied directly to the coupling (inter-oscillator), while the damping is a material parameter (intra-oscillator). Nevertheless, the behavior one expects from time




(a) 3SHNO

(b) 4SHNO

Figure 7.8: Order parameter (blue lines, left y-axes) and normalized voltages (red lines, right y-axes) of the high-damping ($\alpha = 0.5$) 3SHNO and 4SHNO cases, respectively.

delayed oscillators appears present in the case of highly damped SHNOs on a chain. Based on this, it is expected that the 6SHNO case will synchronize rapidly and in a more well-behaved manner than the 5SHNO case due to the symmetry. Additionally, the edge oscillators may be approximately $-180°$ out of phase due to the even number of oscillators, but effects like iSHE make the final phase difference considerably smaller.

## 7.3 Voltage fluctuations

In Figures 7.2 and 7.3, we detailed the behavior of our system using the magnetization directly or indirectly. As explained in previous sections, this is readily available in simulations, but using this as an observable is not suited for realizations due to experimental inaccessibility. In this section, we will investigate the behavior of the voltage in response to the oscillation and subsequent synchronization.

For this analysis, we consider two high-damping cases: 3SHNO and 4SHNO with all effects enabled, Figures 7.2d and 7.2g. These are respectively desynchronized (though with some periodicity to their phase difference) and synchronized. The reasoning is that these systems have many SHNOs and display very erratic behavior compared to their low-damping counterparts in Figure 7.3, while also characterizing both desynchronization and synchronization. Additionally, they run for a long time. Therefore, if we can characterize the synchronization for these, it is easy to justify an extension to all the other setups.

The voltage is measured in the same way as for the spiking neuron in Section 6: The IP current is supplied via two electrodes placed on the short faces of the conducting layer. These electrodes are set to maintain the current at 750 µA at all times. The effect of SMR will therefore be measurable as changes to the voltage.

Figure 7.8 depicts the measured voltages in the conducting layer overlaid the order

parameter $r$ for these systems. For Figure 7.8b, we know that the oscillators will synchronize with some phase difference, and this is immediately apparent for the order parameter, which eventually becomes a constant at $r \approx 0.4$. The voltage responds accordingly, oscillating erratically before becoming a constant in time as the oscillators reach their phase-locked orientation. The voltage fluctuations do not correspond exactly with those of the order parameter, but we see that the behavior is somewhat similar between them.

This is contrasted by Figure 7.8a, which we know to be desynchronized. Here, $r$ oscillates in a similar pattern to the angle difference in Figure 7.2d, as we would expect. The voltage oscillates too, and it is periodic, but its shape does not correspond exactly with the order parameter. Neither of these becomes constant in time.

This is promising: When the oscillators are synchronized, both the phase difference, the order parameter and the voltage fluctuations due to SMR all become constant. And conversely, when they do not synchronize, all these quantities fluctuate. It should be noted that there is no 1:1 behavior between the order parameter $r$ and the output voltage, especially for many-oscillator systems where the order parameter is a complicated trigonometric quantity. However, the voltage can be used to determine whether an entire system is synchronized or desynchronized. This will become important when using the SHNO as a real neuromorphic device, as this binary state (synchronized or not) is what will be used for a future neural network implementation. Potentially determining this in voltage alone is a huge benefit.

## 7.4 Outlook

We have observed that high-damping and low-damping SHNO systems (of order $\alpha = 0.5$ and $\alpha = 0.02$ respectively) of up to 5 oscillators in a chain exhibit very different behavior: The high-damping oscillators usually settle into phase-locked states that are seemingly insensitive to the initial orientation, while the low-damping oscillators synchronize with virtually no phase difference. Additionally, while the high-damping 2SHNO and 4SHNO systems will synchronize, the 3SHNO and 5SHNO do not strictly speaking do so. Rather, they enter a state wherein the phase differences oscillate, with either small or large amplitude, but in a periodic manner. Different sets of neighbors group together and attain the same phase differences. In essence, these systems become ordered, but not fully synchronized. These systems also have a degree of lag in their response to each other. This is not the case for the low-damping oscillators, which synchronize completely for all configurations simulated. Finally, there is also a large time-scale difference, with high-damping SHNOs taking potentially more than 200 ns to synchronize compared to 5 ns for a corresponding low-damping system.

Future work should look at the transition between these states, and the behavior as the damping is gradually tuned from high-damping to the low-damping regimes. More

oscillators can also be added to see how the clustering of phase differences changes. We have also shown that the iSHE is not necessary for the interesting layout of the high damping, and that the coupling by demag alone is sufficient. It should be investigated if the omission of the common Pt bar also preserves these phase-locked synchronizations in the high damping case. If this is the case, one could perhaps determine the phase of each oscillator (and thus the phase difference between them) by measuring the voltage alone, opening for greater experimental accessibility.

In our subsequent work, we will focus our attention on low damping. There are several reasons for this: The YIG-like parameters we have based our work on usually include a very small damping constant ($\alpha \sim 10^{-4}$ [69]), the synchronization occurs in more configurations and is overall more robust with very small phase differences, and as mentioned previously the dynamics are faster compared to the high-damping case. Furthermore, the low-damping 2SHNO case captures the synchronization behavior we wish to study with minimal extra effects.

# 8 Perturbing the spin Hall nano-oscillators

Thus far, we have covered the basic behavior of SHNOs driven by a common current. These "bare SHNOs" allowed us to look at the basic synchronization of SHNOs and additionally the impact of iSHE. However, these oscillators were, for all practical currents, always synchronized or phase-locked, and additionally shows that iSHE is weak compared to demagnetization. The next step towards creating an ONN with these EP-SHNOs is to reliably make and break synchronization for a given system. This sets the stage for our binary readout: Synchronized or desynchronized. To do this, we must make the frequency of one (or both) oscillator(s) tunable.

The new setup is depicted in Figure 8.1. This is slightly a modified system of the one in Section 7, as we have added a second layer of Pt atop both oscillators, through which we run either an IP or OOP current (or both).

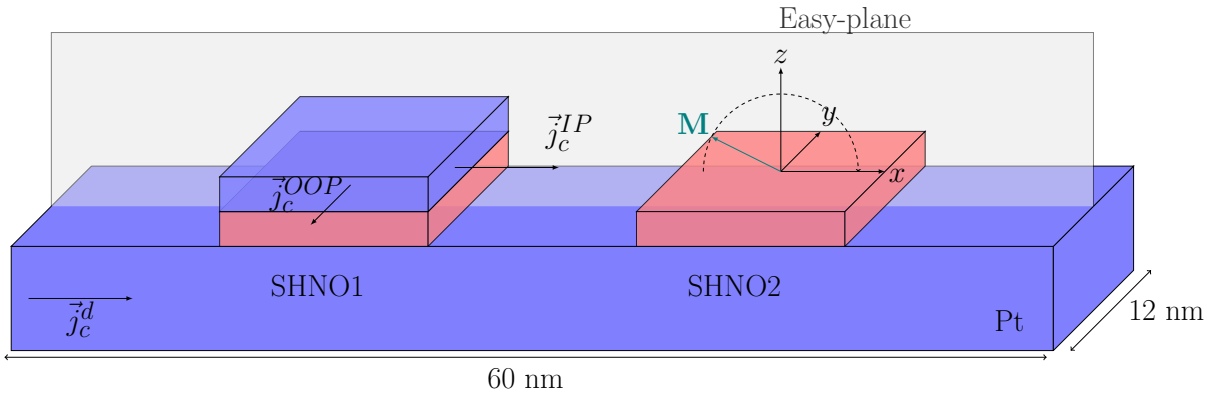
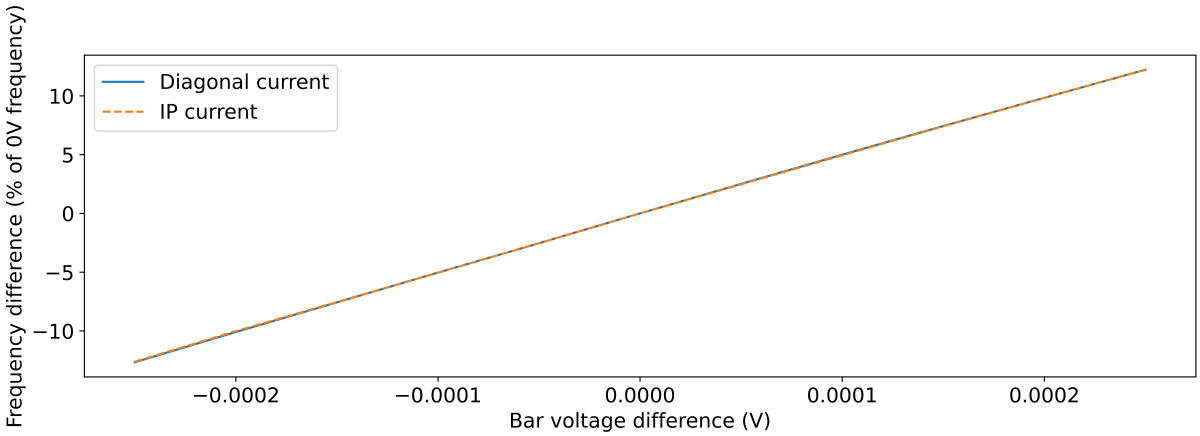
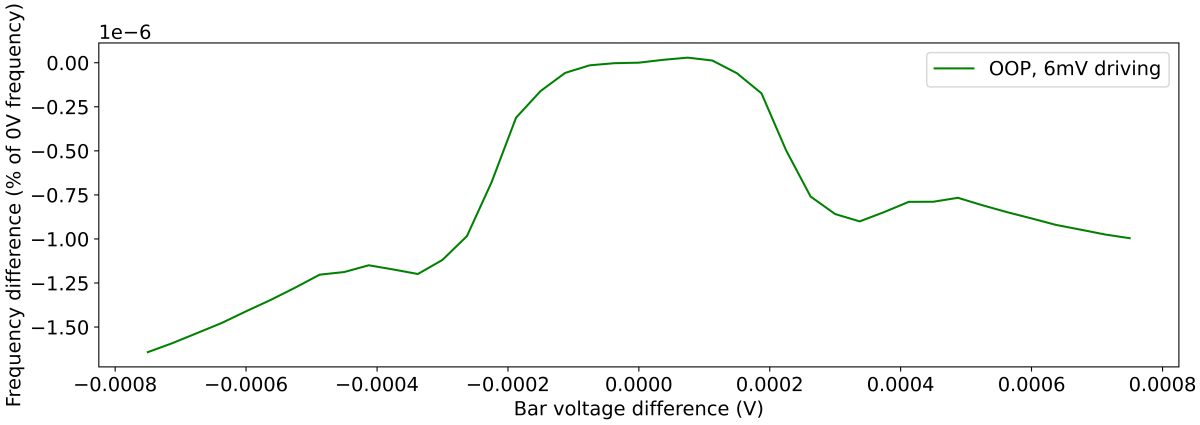



Figure 8.1: Setup for the individually controllable 2SHNO system. Blue meshes are conducting layers (Pt), and red meshes are ferromagnetic. The lower Pt bar is as before. The top Pt bar is materially equivalent to the lower Pt bar. Electrodes are applied to the xz- and yz-faces (IP and OOP respectively) of the top bars. In this diagram we have removed the top Pt bar on SHNO2 for clarity, but in all cases both SHNOs will have top bars, the reason being that we want the effect of the material's presence to be the same across all oscillators (even if they are not tuned by it).

We make a few definitions: The lower Pt bar refers to the Pt bar common for all the SHNOs, which carries the driving current $\vec{j}_c^d$ that drives the basic oscillations. The top (or upper) Pt bar refers to the smaller Pt mesh atop the SHNOs, each of which carries separate perturbing currents, $j_c^{IP}$ and $j_c^{OOP}$. We will mostly report these in terms of their corresponding voltage differences, $\Delta V^{\text{IP}}$ and $\Delta V^{\text{OOP}}$ respectively. This current is usually much smaller than $j_c^d$. We define that an out-of-plane (OOP) current is the current in the top bar which flows in the $\pm\hat{\mathbf{y}}$-direction, and in-plane (IP) current as flowing in the $\pm\hat{\mathbf{x}}$-direction. This is in relation to the defined xz-oriented easy-plane. Note that both




(a) Frequency response for an IP and diagonal current.




(b) Frequency response for an OOP current.

Figure 8.2: Oscillator frequency as a function of top bar potential in a 1SHNO case, as a % of the 0V case. Note that each plot is shown for different voltage ranges, as the more noteworthy behavior of the OOP case occurs at a greater potential difference.

a non-zero $j_c^{OOP}$ and $j_c^{IP}$ can be applied at once, causing a "diagonal" charge current. We will also flow the lower bar current in the $+\hat{\mathbf{x}}$-direction. We define "upstream" and "downstream" as positions in relation to this current, as previously. We also write the natural and instantaneous frequency of the $i$th oscillator as $F_i$ and $f_i$ respectively, and the corresponding upper bar voltage difference as $V_i^{IP}$ or $V_i^{OOP}$.

## 8.1 Controllability of the 1SHNO case

To introduce our methodology, we first show the controllability of an individual SHNO's frequency using the top bar. The setup for this is akin to the one shown in Figure 8.1, however in this case there is no second, downstream SHNO. We instead have one SHNO with one bar, which we tune while keeping the lower Pt voltage constant.

The frequency is measured by averaging the instantaneous frequency at every time

step for the whole series. The simulations are run for 500ps, however the first 200ps are neglected in this analysis to allow the oscillator to enter its steady state.

Figure 8.2 shows the (natural) frequency of SHNO1 in the 1SHNO case as a function of the potential in the top Pt layer, with the bottom fixed to a 6 mV potential difference. In Figure 8.2a, we compare both purely IP and diagonal charge current. The IP current affects the oscillator in same way as the current in the lower Pt, and so we both expect and observe a linear relationship between its magnitude and the frequency ratio, at a rate of approximately 0.05 %/µV for this driving frequency. We report the ratio to give a sense of scale of the impact caused by this perturbing frequency, though it is not preserved for very large voltages. This is due to the OOP component of the magnetization growing larger, breaking the linear relationship. The actual value of this relationship is, for the small perturbing voltages, approximately 0.0754 GHz/µV.

The OOP current Figure 8.2b causes entirely different behavior: The linear relationship between current and frequency is lost, and instead the oscillator responds with a lower frequency for both current directions. This relationship is asymmetric around 0 V, where the current direction changes sign. Furthermore, the magnitude of the response is on the order of $10^5$ times smaller compared to the case with IP current in Figure 8.2a. The likeliest cause of this asymmetry is the fact that the driving current creates a non-zero $\hat{\mathbf{y}}$-component of the magnetization, which we have explained in detail in Section 3.5.3.

The takeaway is that the IP current allows the greatest degree of control, as the response is linear and predictable, unlike that of the OOP case.

## 8.2 Synchronization of the 2SHNO case

In our study of the bare SHNO, we noted that the oscillators almost always became phase-locked, with the phase difference dependent on material parameters such as damping, anisotropy and inter-oscillator separation. The presence of iSHE did increase the frequency of downstream oscillators, but it was insufficient to break the synchronization. With an understanding of how we can control the oscillators, we now aim to reproducibly desynchronize the oscillators in the same manner employed by Garg et. al. [12], and thus obtain the synchronization range of our setup.

Unless otherwise specified, we will keep the lower Pt bar voltage difference fixed at 4 mV, with charge current flowing in the $+\hat{\mathbf{x}}$-direction, as before.

### 8.2.1 Phase difference for upstream perturbation

We first tune the upper bar voltage of the upstream oscillator, SHNO1. Figure 8.3 shows the period-averaged IP angle difference as a function of time. This is the same measure as used previously. We reiterate that a constant angle difference corresponds to oscillators being phase-locked, i.e. synchronized. We observe in Figure 8.3 that for voltages below

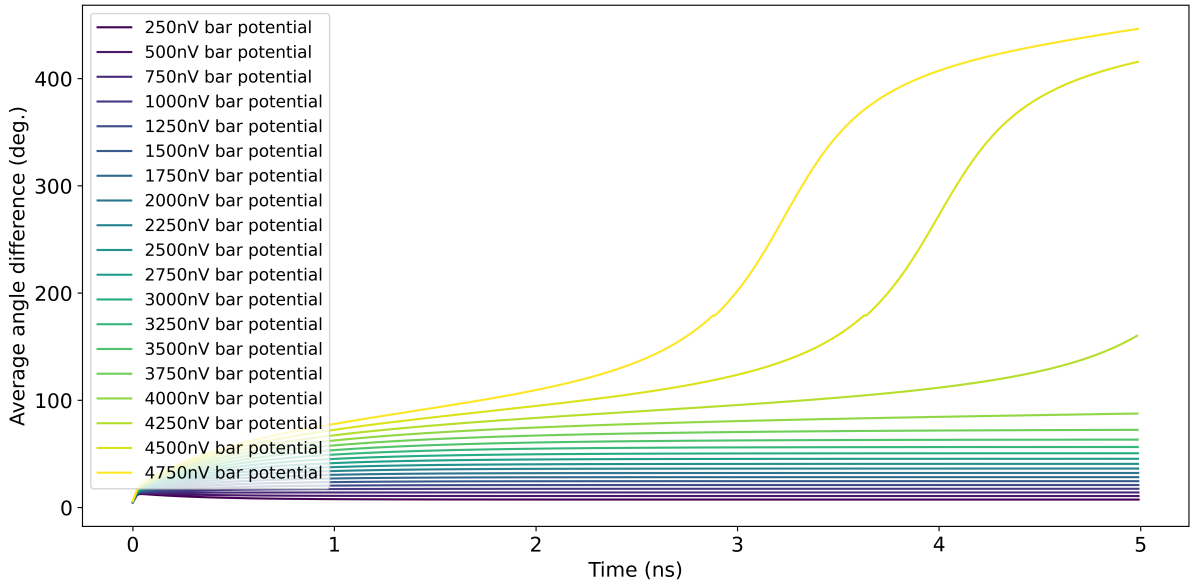
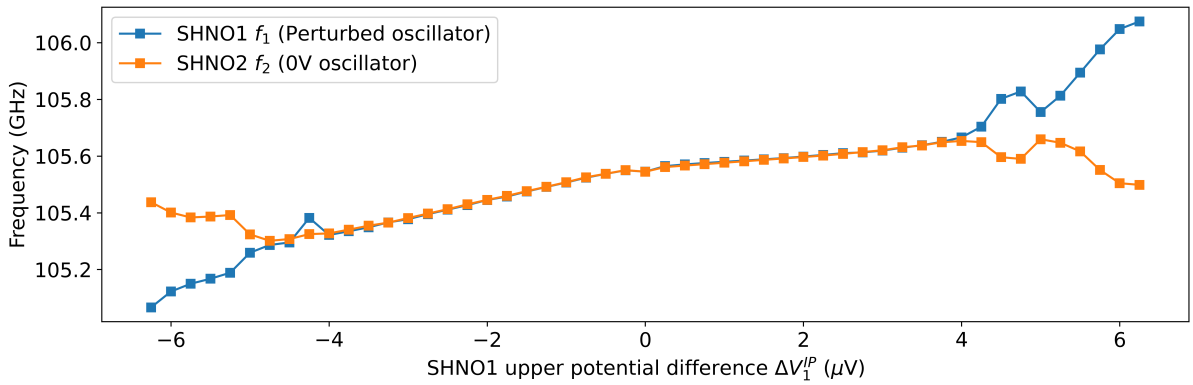



Figure 8.3: Angle difference as a function of time for a set of perturbing currents in SHNO1.

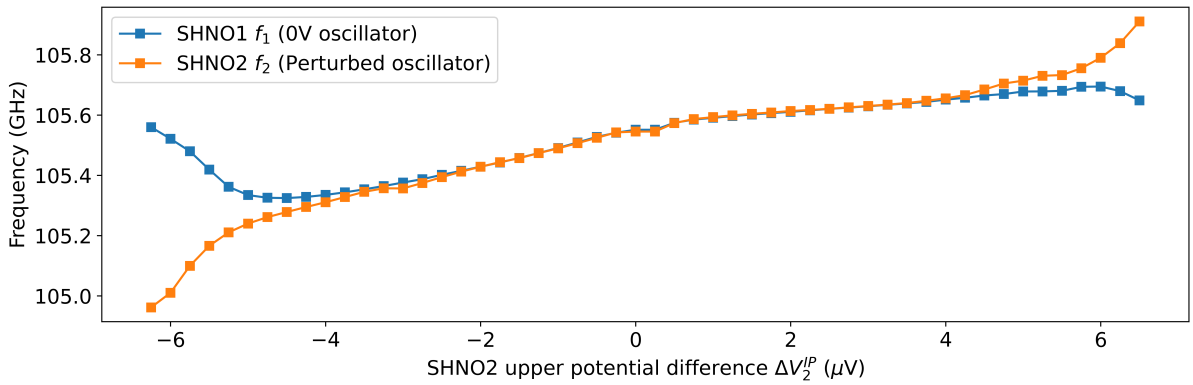
approximately 3.5 µV, the angle difference grows asymptotically towards a characteristic value which depends on the magnitude of the top Pt voltage, with higher voltages corresponding to a brighter color. This gives a spectrum of phase differences. Beyond 3.5 µV, the angle difference begins to plateau before continuing to grow. For these values, the perturbing voltage is large enough to overcome the restoring force provided by demagnetization, and a frequency difference occurs between the oscillators. The changing slope we see in these desynchronized oscillators (most prominently in the 4.75 µV case) is due to the coupling providing a negative torque until a 180° phase difference with the unperturbed oscillator. After that, the torque switches to positive, as it is the shortest path to align the oscillators, and the phase difference speeds up rapidly (these are still unwrapped angles). This behavior will continue in time.

### 8.2.2 Synchronization range for upstream and downstream perturbation

The period-averaged angle difference between oscillators gives the layout of the system in its (potentially) synchronized state, but it can also report to what degree oscillators are desynchronized. This is instructive, but in the implementation of the SHNO as a neuromorphic device we rarely consider the actual phases of the oscillators. Our focus is instead on the (instantaneous) frequencies. Therefore, we instead look at the **synchronization range** of the oscillators by depicting the time-averaged instantaneous frequency as a function of perturbing voltage. Figure 4.2 depicts an example of a synchronization range as a function of the natural frequency of one oscillator obtained via the Kuramoto model. However, in Section 8.1 we showed that the natural frequency is proportional




(a) Instantaneous frequency of the 2SHNO system as a function of input current in SHNO1.




(b) Instantaneous frequency of the 2SHNO system as a function of input current in SHNO2.

Figure 8.4: Instantaneous oscillator frequency as a function of the input current (i.e. natural frequency) in SHNO1 and SHNO2 (up- and downstream) respectively. The unperturbed oscillator is, in this case, held at 0V, while the perturbed oscillator is varied over the range specified.

to the perturbing voltage. As it is easier to obtain the voltage rather than the natural frequency, we choose to report our results in voltage.

A secondary motivation is that Garg et. al. [12] have reported, through micromagnetic simulations, the synchronization ranges for oscillators with completely separate conducting layers as part of their work. They used only one conducting layer per oscillator, combining the driving and perturbing voltages into one. This current was kept strictly unidirectional with respect to the oscillator. However, our setup allows for a bidirectional flow of the perturbing current also as shown in Section 8.1. Therefore, we allow a reversal of the perturbing current such that it can flow in the $\pm\hat{\mathbf{x}}$-directions, parallel or antiparallel the driving current. In this discussion, we are primarily concerned with the qualitative aspects of the synchronization range and the impact of iSHE. We are not concerned with the quantitative aspects, as those will be addressed in later sections.

Figure 8.4a shows the average instantaneous frequency of the oscillators as a function of $\Delta V_1^{\mathrm{IP}}$, while $\Delta V_2^{\mathrm{IP}} = 0$. As mentioned previously, the natural frequency $F_i$ is directly

proportional to $\Delta V_i^{\mathrm{IP}}$. These simulations are each run for 5 ns, though the first 1 ns is neglected to allow the oscillators to attain their steady-state behavior. For brevity, we will use the previously established notation of $f_1$ and $F_1$ being, respectively, the instantaneous and natural frequency of SHNO1.

There are several promising aspects of these results: First, at the edges where $|\Delta V_1^{\mathrm{IP}}|$ is greatest, $f_1$ differs from $f_2$ and grows as a function of $|\Delta V_1^{\mathrm{IP}}|$, which is indicative of desynchronization just as expected. As $F_1$ approaches $F_2$ by $V_1^{\mathrm{IP}} \to 0V$, both $f_1$ and $f_2$ depart their natural frequencies until they become equal and thus synchronize at approximately $-4\,\mu\mathrm{V}$. This synchronization continues to approximately $4\,\mu\mathrm{V}$, showing a degree of symmetry centered around 0V and giving a corresponding voltage range of $8\,\mu\mathrm{V}$. Within the synchronization range, the slope of $f_1$ is lower than outside, and we see that it starts growing faster for $|\Delta V_1^{\mathrm{IP}}| > 4\,\mu\mathrm{V}$, showing that the oscillators are mutually pulling on each others' frequencies. Additionally, when the $V_1^{\mathrm{IP}}$ moves far beyond the synchronization range, $f_2 \to F_2$, that is to say, it relaxes towards its natural frequency. This shows that two oscillators with vastly different frequencies do not interact to any significant degree, very much in line with what is predicted by the Kuramoto model.

Figure 8.4b shows the synchronization range for when $\Delta V_2^{\mathrm{IP}} \neq 0$ while the upstream has $\Delta V_1^{\mathrm{IP}} = 0V$. That is to say SHNO2 is subject to an upper bar voltage while SHNO1 is unperturbed. The unperturbed oscillator is now not impacted by the iSHE, which constantly alters the current in the conductor. The behavior is not changed in any qualitative degree, and we still observe a synchronization range of similar size. However, the change from a desynchronized state to a synchronized state is less spontaneous, and the frequencies appear to close with each other in a more gradual fashion. This may be an artifact of insufficient runtime of the system.

In both Figure 8.4a and 8.4b, there is a change in slope for the instantaneous frequency centered on 0V, and in both cases $f_1$ and $f_2$ grow slower for $\Delta V_1^{\mathrm{IP}} > 0$ and $\Delta V_2^{\mathrm{IP}} > 0$ respectively. This could be chalked up to the change in torque that arises from switching the current direction. However, although not shown in this thesis, the change in slope is also present in the case where fixed at $\Delta V_2^{\mathrm{IP}} = 6.5\,\mu\mathrm{V}$ and $\Delta V_1^{\mathrm{IP}}$ is varied in a strictly positive range centered on $6.5\,\mu\mathrm{V}$. In fact, performing this setup only shifts the behavior of the synchronization range further along the $\hat{\mathbf{x}}$-axis, while maintaining the same behavior as these simulations.

Additionally, it is also not present in the 1SHNO case shown in Figure 8.2a, where the frequency of a single SHNO was controlled by adjusting the voltages in a range symmetric about 0V. The response of the frequency is entirely linear with no change at any point in the range depicted, despite $\Delta V_1^{\mathrm{IP}}$ attaining both positive and negative values.

These two points rule out the possibility that the change in torque from a reversal of the current direction is the root cause, and the presence when $\Delta V_2 \neq 0$ while $\Delta V_1 = 0$ and fixed implies that it is also not caused by variations in iSHE due to the changing torques.

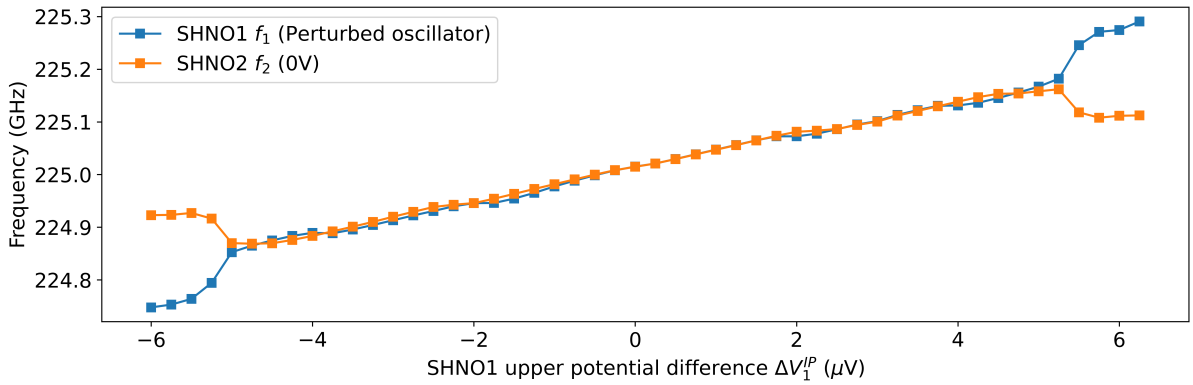



Figure 8.5: Instantaneous frequency of the 2SHNO system as a function of input current in SHNO1, where SHNO1 and SHNO2 are materially separated and only coupled via their demagnetization fields.

This kind of behavior is also not present in 2-oscillator Kuramoto model simulations as exemplified in Figure 4.2, nor in the work by Garg et. al. [12] who simulated pairs of SHNOs coupled by nothing but demagnetization. It must therefore be tied to something intrinsic to the common Pt geometry.

### 8.2.3 Synchronization of completely separate SHNOs

To isolate the effect causing the change in slope, we separate the SHNOs entirely. By separate, we mean that the oscillators are HM/FM/HM trilayers as before, but the common Pt bar is replaced with two materially identical bars, with the same thickness and width as before, but with the length reduced such that it is the same as the oscillator. These trilayers are subject to the same lower Pt bar voltage difference of 2 mV, with SHNO1 experiencing a perturbation in the top bar. This way, we mitigate the effect of both iSHE, but also the scattering and reflection due to SHE. The oscillators are now only coupled via demagnetization (which has been shown to be the dominant coupling factor). Similar work has already been undertaken by Garg et. al. [12]. We will later try to reproduce their results, so it is important to verify that we can reproduce their synchronization ranges first.

The reason for the reduction in voltage to 2 mV is that with the lower Pt layer reduced in size, the frequency of the oscillators increases dramatically. Earlier simulations for a common Pt bar, such as the ones depicted in Figure 8.4a, show frequencies of approximately 105 GHz for a 4 mV driving voltage. For these separate oscillators, even at 2 mV the frequency is approximately 225 GHz.

Figure 8.5 depicts the synchronization range of the separated oscillators. We immediately observe that there is no significant change in slope centered on 0V. The synchronization range is otherwise qualitatively the same, and of almost the same size despite

operating at over twice the frequency. The small difference is likely due to the changed properties of the simulation as a result of the separation. This is important: As mentioned previously, these oscillators operate at more than 200 GHz, yet the synchronization range is approximately 8 µV in size. This means that the synchronization range is independent of the natural frequencies, and entirely tied to the other parameters deciding the coupling strength.

There appears to be a much greater symmetry in this situation compared to the oscillators on a non-separated Pt bar, Figures 8.4a and 8.4b, not only in terms of the non-changing slope but also in terms of the local fluctuations in $f_1$ and $f_2$ and the behavior outside of the synchronization range. This, too, is to be expected from the Kuramoto model.

The likeliest root cause of the change in slope is the (de)coupling provided by iSHE: In the Kuramoto model, the coupling strength from SHNO1 onto SHNO2 is equal to that from SHNO2 onto SHNO1, $K_{12} = K_{21}$. This is usually represented by a symmetric 2x2 adjacency matrix. Materially separated SHNOs that are coupled only via demagnetization exhibit the same behavior as what is anticipated by the Kuramoto model [12], though small differences may occur due to the out-of-plane components of the magnetization and other effects not accounted for. Altering the Kuramoto model to account for asymmetric coupling only reduces the size and symmetry of the synchronization range without creating a changing slope, suggesting that additional features such as the implicitly time-dependent coupling (or state-dependent, as it depends on orientation) may be a root cause. Details on this time-dependent coupling are present in the discussion of voltage fluctuations as a result of synchronized or desynchronized SHNOs, presented in Section 7.3. Although the voltages provided by these fluctuations are orders of magnitude smaller than the input voltage, the effect adds up over time, especially with the phase difference previously observed.

## 8.3 Temperature robustness of the 2SHNO synchronization

Until now, our simulations have only considered the T = 0 K case. This provides much simpler dynamics at the expense of experimental realizability. Local fluctuations due to thermal noise may destroy the synchronization even at lower temperatures. On the other hand, as the temperature is a statistical noise, it may be averaged out over time. It's therefore of interest to study the EP-SHNO's robustness against temperature.

Our setup is the same 2SHNO case depicted in Figure 8.1. The simulations were performed using the stochastic variant of the LLG-SA equation, namely the sLLG-SA equation [32] with a 1 fs time step. To give a picture of the time evolution, we show the angle difference as a function of time without any period-averaging, such that every small fluctuation can be observed. These values will oscillate rapidly around their average

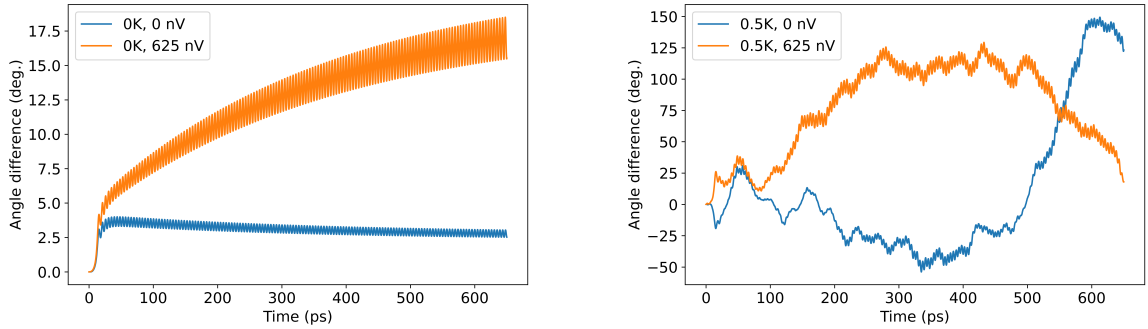
values.

Figure 8.6 contains 4 subfigures, each depicting the angle difference of the 2SHNO case subject to different temperatures. Within each figure there are two different values for the voltage, $\Delta V_1^{\mathrm{IP}} = 0\,\mathrm{V}$ and $\Delta V_1^{\mathrm{IP}} = 625\,\mathrm{nV}$. These voltages were selected because they both result in synchronization in the $\mathrm{T} = 0\,\mathrm{K}$ limit, falling well within the synchronization range depicted in Figure 8.4. Anything desynchronized in the $T = 0\,\mathrm{K}$ limit will still be desynchronized for finite temperature, and so it was considered unnecessary to immediately test desynchronized cases. More simulations need to be run to fully characterize the temperature dependence.

However, Figure 8.6 shows no synchronization between oscillators subject to finite temperature, neither for $\Delta V_1^{\mathrm{IP}} = 0\,\mathrm{V}$ or $\Delta V_1^{\mathrm{IP}} = 625\,\mathrm{nV}$ in any of $T = 500\,\mathrm{mK}$, $T = 1\,\mathrm{K}$ or $T = 1.5\,\mathrm{K}$. We first explain Figure 8.6a, which is the same situation as plotted in Figure 8.3, though this time for only two oscillators without any period-averaging. The behavior is as expected: Due to the constantly changing frequency as the direction of magnetization changes, the angle difference between the oscillators also fluctuates. However, it oscillates around a specific value, and we see that in the 0 V-case of Figure 8.6a this value eventually becomes constant in time. This illustrates why we perform a period-averaging in the first place, but it also characterizes the short-term oscillation over the course of one cycle. These short-term oscillations are present in all cases, and with little change in their periodicity or magnitude as the temperature grows. It therefore appears that the temperature plays a small role in these short-term fluctuations.

The loss of synchronization is characterized by the large, unconstrained changes in phase difference over time, as detailed in Section 7. For the finite-temperature simulations depicted in Figure 8.6, the magnetization of SHNO1 and SHNO2 switches between being parallel to nearly antiparallel several times within 100 ps to 200 ps. There is seemingly no regularity between the oscillations. Consider, for instance, the 0 V-case of Figure 8.6b, where the phase difference initially fluctuates between 25° an −50°, before suddenly going beyond 140°. This is also significant, as within this short period the angle difference of the $\Delta V_1^{\mathrm{IP}} = 0\,\mathrm{V}$-case becomes larger than that of the $\Delta V_1^{\mathrm{IP}} = 625\,\mathrm{nV}$-case. This does not occur for $\mathrm{T} = 0\,\mathrm{K}$, as the steady-state angle difference is a characteristic of the $\Delta V_1^{\mathrm{IP}}$ and thus the lines do not cross.

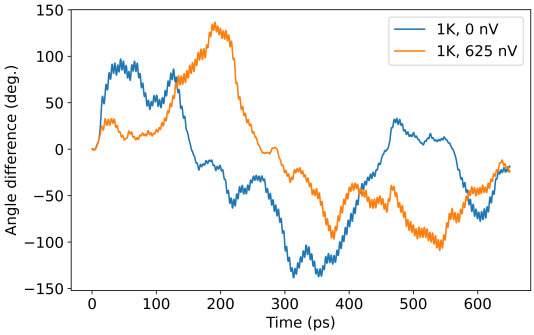
The simulations in Figure 8.6 were each run for 650 ps, a relatively short time compared to the expected synchronization dynamics. For instance, in the 625 nV case Figure 8.6a, the oscillators are tending towards their steady-state phase difference, but have yet to attain it. The synchronization ranges depicted in Figure 8.4a required 5 ns to be considered sufficiently accurate. The reason for this compromise is runtime, as enabling temperature requires much more computational resources. It may be the case that the oscillators come to rest after sufficient time. We therefore select the case where $\Delta V_1^{\mathrm{IP}} = 0\,\mathrm{V}$ and $\mathrm{T} = 0.5\,\mathrm{K}$, depicted in Figure 8.6, and run the simulation for 10 ns. This case has the



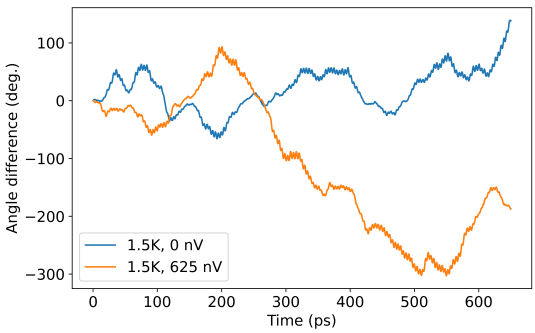

(a) $T = 0\,\mathrm{K}$




(b) $T = 0.5\,\mathrm{K}$




(c) $T = 1\,\mathrm{K}$




(d) $T = 1.5\,\mathrm{K}$

Figure 8.6: Instantaneous angle difference as a function of time for 2SHNO, with the temperature of both oscillators varied from 0 K t0 1.5 K. SHNO1 subject to either $\Delta V_1^{\mathrm{IP}} = 0\,\mathrm{V}$ or $\Delta V_1^{\mathrm{IP}} = 625\,\mathrm{nV}$.

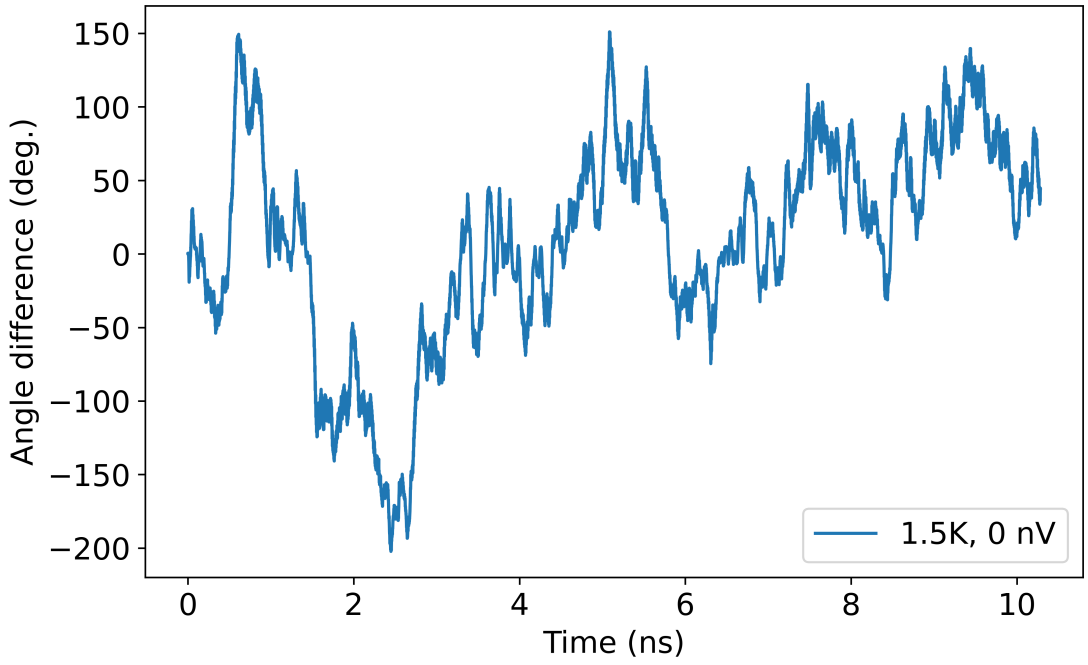



Figure 8.7: Angle difference as a function of time for 2SHNO, $\Delta V_1^{\mathrm{IP}} = 0\,\mathrm{V}$, $\mathrm{T} = 0.5\,\mathrm{K}$. Simulation run for 10 ns.

fewest alterations and synchronizes the most readily in the $T = 0\,\mathrm{K}$ case.

The results are shown in Figure 8.7. Even when run for longer than is expected necessary to establish steady-state behavior based on the $\mathrm{T} = 0\,\mathrm{K}$ case, the oscillators are still highly desynchronized: There are rapid, stochastic changes in the phase difference for short time scales, and there is seemingly no long-term synchronization effect kicking in.

The Curie temperature of a YIG, which shares some material parameters with our SHNOs, is $\mathrm{T_c} = 560\,\mathrm{K}$ [79], far beyond anything to be considered in this work. However, this is the bulk temperature, and for our thin film the actual Curie temperature may be wildly different. We therefore study the magnetization as a function of temperature for our sample. The results are not depicted, but we report that at 20 K, the magnetization is only 1.2 % weaker than the zero-temperature case. We consider this sufficient to rule out the loss of ferromagnetic ordering as a cause for this thermal desynchronization, as the desynchronization occurred for temperatures down to 500 mK.

The discussion presented here is insufficient to make any conclusive statements about the temperature sensitivity: It appears that the synchronization is highly dependent on temperature to the point where it is lost for temperature far below the Curie temperature, but no characterizing quantity has been found. Nor is this discussion sufficient to ascertain whether or not the oscillators *eventually* synchronize: We have run simulations up to 10 ns,

twice that which was previously considered sufficient to obtain the synchronization ranges in e.g. Figure 8.4, yet it may be that the oscillators simply take much longer when subject to finite temperature.

In later sections we will apply the Kuramoto model to a system on SHNOs in more detail, relating the synchronization range obtained via micromagnetic simulations to that of the perfect Kuramoto oscillators, and thus obtaining an effective coupling constant. As finite temperature is modeled via white noise, it would be reasonable to assume that the Kuramoto model with noise, eq. (4.4), could be used to model the finite-temperature micromagnetic system. Once the coupling constant has been obtained in the zero temperature limit, one could enable the noise and observe how (or if) the synchronization breaks down. These results could be used in a prediction of a characteristic temperature for the loss of synchronization, as well as deriving ways for increasing the temperature robustness. The methodology for the application of the Kuramoto model to a micromagnetic system will be shown in practice in Section 9.

## 8.4 Synchronization of the 3SHNO system

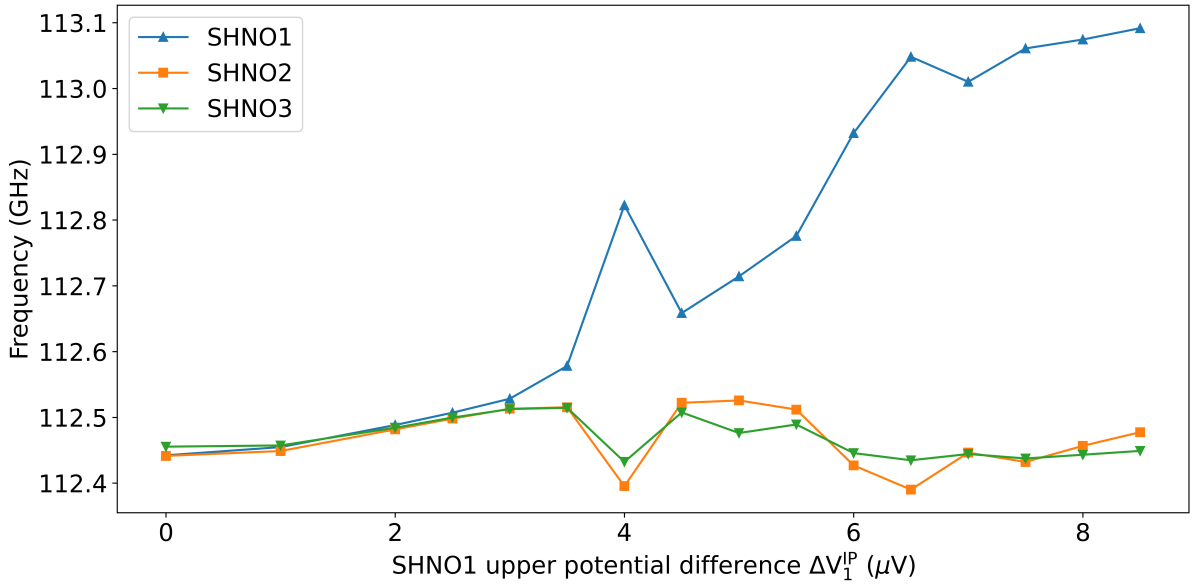
As discussed in [12], for binary classification we only need the synchronization of two output oscillators (disregarding the method of synchronizing the oscillators, which may require more). Present research has managed to synchronize up to 50 SHNOs in a chain [75], and an 8x8 arranged in a 2D array [73]. These have been performed with a different type of SHNO rather than the EP-SHNO; the long-range synchronization of the EP-SHNO has, to our knowledge, not yet been studied. Therefore it is illustrative to study the synchronization beyond the 2SHNO case when controlling the oscillators, and how the synchronization breaks down. We aim to study the effect of the 3SHNO system as a path toward more oscillators.

We will give consideration to two separate cases: A 3SHNO case in which SHNO1 is perturbed, and one where SHNO2 is perturbed. The motivation for the former is to see how the synchronization propagates when there are several oscillators with a larger spacing than in the 2SHNO case, while for the latter we are combining the behavior of both the 2SHNO upstream and downstream perturbation. We neglect the case of the last oscillator being perturbed, as the earlier discussion on the downstream perturbation showed that the behavior is broadly the same as the upstream perturbation.

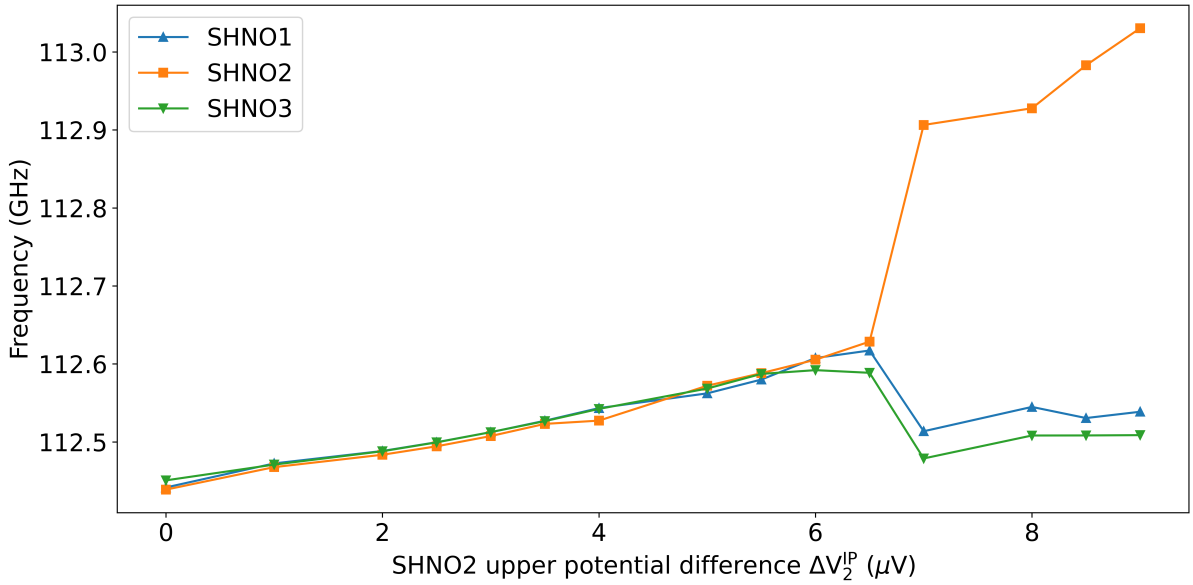
Note that, as previously mentioned, we are considering the low-damping case of the 3SHNO, which do synchronize naturally.

Figure 8.8a and Figure 8.8b depicts the instantaneous frequency of all three oscillators as a function of perturbing voltage $V_1^{IP}$ and $V_2^{IP}$ respectively. To properly characterize the system, we will use the labels for the oscillators (SHNO1 SHNO2 SHNO3) very often.

When the first oscillator is perturbed, Figure 8.8a, we see a familiar pattern: First,




(a) 3SHNO system where SHNO1 is subject to an upper bar voltage




(b) 3SHNO system where SHNO2 is subject to an upper bar voltage

Figure 8.8: Average instantaneous frequency of a 3SHNO system as a function of upper bar voltage of perturbed oscillator (SHNO1 or SHNO2 respectively). We only plot this for positive voltages, as the discussion above has dealt with the positive and negative voltages separately and shown that the synchronization range is more or less symmetric, though with a change in slope.

the synchronized region, lasting until approximately 3 µV. Here, the frequencies begin to part ways. At 4 µV we note a sudden desynchronization creating a local peak, before the frequency differences between SHNO1 and the other oscillators decrease slightly, then continues to grow. This mimics the behavior seen in the 2SHNO case where SHNO1 is perturbed, Figure 8.4a, though desynchronization occurs at a slightly lower frequency. Secondly, we see that after SHNO1 moves far beyond the synchronization range, both SHNO2 and SHNO3 stop being impacted by it to any meaningful degree, and more to the point they synchronize and drop towards their natural frequencies.

In the beginning, all oscillators lie within their respective synchronization range, and so as SHNO1 starts increasing its natural frequency, SHNO2 and SHNO3 follow suit and increase their instantaneous frequencies. The direct coupling between SHNO1 and SHNO3 is comparatively weak, and so it wants to desynchronize first. However, the coupling between SHNO1 and SHNO2 is strong, and so is the coupling between SHNO2 and SHNO3, meaning they are preventing desynchronization by proxy. Eventually, SHNO1 moves too far beyond the synchronization range with SHNO2, and it instead favors a synchronization with SHNO3, itself already beyond the synchronization range with SHNO1. The result is that SHNO2 and SHNO3 relax back towards their natural frequencies, while SHNO1 increases as one would expect. In effect, we have two separate 2SHNO systems in play, with SHNO2's preferred synchronization moving from one to the other: Initially it wants to follow SHNO1 due to the coupling, but as SHNO1 moves too far it switches to prefer synchronization with SHNO3. This switching is what causes the shortening of the synchronization range, as there is now an additional force coercing SHNO2 out of synchronization with SHNO1.

When the middle oscillator is perturbed, Figure 8.8b, the behavior is a combination of the upstream and downstream 2SHNO system, Figure 8.4a and Figure 8.4b. In this case, the coupling between SHNO1-SHNO2 and SHNO2-SHNO3 are both strong. We observe a comparatively strong synchronization with little splitting that lasts for over 6 µV, greater than when SHNO1 is perturbed in Figure 8.8a, and much improved over the corresponding 2SHNO cases in Figure 8.4 of approximately 4 µV. This is because there are two unperturbed oscillators in the system, SHNO1 and SHNO3, each of which impacts SHNO2 with approximately the same strength as before. Intuitively we can think of this as a 2SHNO system with much stronger coupling. The result is a larger synchronization range. When the perturbation of SHNO2 becomes too large, both SHNO1 and SHNO3 desynchronize effectively at once, and the result is a very sudden and large jump in frequency, while SHNO1 and SHNO3 relax toward their natural frequencies as before. From our data we cannot conclusively say that SHNO1 and SHNO3 will synchronize but there is a noticeable and persistent frequency difference in Figure 8.8b, suggesting that they may not readily synchronize.

In summary, perturbing the 3SHNO system results in a combination of 2SHNO sys-

tems: When SHNO1 in the 3SHNO system is perturbed, both SHNO2 and 3 will attempt to follow suit until SHNO2 prefers synchronization with SHNO3, and the full-system synchronization breaks. This shrinks the synchronization range. When SHNO2 is perturbed, both SHNO1 and SHNO3 are coupled strongly to it, and as a result they both follow suit before breaking synchronization effectively simultaneously. The result is a longer synchronization range.

## 8.5 Outlook

All of the simulations in Section 8 have been run for the low-damping case. In Section 7, we saw that these oscillators have the most predictable dynamics as they are prone to quick and reliable synchronization. However, the high-damping case showed interesting behavior, still synchronizing but with an unusual phase-locked configuration, or desynchronizing with a periodicity. Research has already been put into the effect of damping on the dynamics of a single SHNO [80], but none has been put into the characteristics of synchronization, or the breaking thereof (to the extent of our knowledge). It would therefore be of interest to see controlled (de)synchronization of a high-damping system.

For instance: While the high-damping 4SHNO system synchronizes strongly, the 3SHNO system is instead phase-synhronized, where the phase difference oscillates periodically with a large amplitude. This is shown in Figure 7.2. However, in Section 8.4, we showed that once synchronization in a 3SHNO system breaks down by SHNO1's frequency being too high, the remaining oscillators will be less impacted by its presence and instead act as though it is a 2SHNO system. This is especially evident in Figure 8.8a. It begs the question: If the first oscillator in a high-damping 4SHNO system is moved out of synchronization with all other oscillators, will the remaining oscillators act like a highly-dampened 3SHNO system and subsequently also enter a phase-synchronized state, as depicted in Figure 7.2d? And conversely, can we suddenly introduce an oscillator (by for instance tuning its frequency into the synchronization range) and create synchronizations as a 4SHNO system? Additionally, the 5SHNO depicted in Figure 7.6 shows clustering of the phase differences depending on position. By tuning the frequency of only one oscillator, one may be able to shift the whole cluster, or perhaps create entirely new clusters if the frequency becomes too large to counteract.

# 9 Neural network application of SHNOs

Traditional learning methods for neural networks usually operate on the synapses connecting the neurons: The weights in these synapses are adjusted by the learning algorithm as more training data is fed into the network, the theory for which has been discussed in Section 5.4. A similar methodology of adjusting weights can be applied to an ONN by adjusting the coupling constants between oscillators. However, for nanoscale devices, the coupling strengths are materially or spatially dependent, and as a result, altering the coupling for training requires heavy circuitry [60].

At time of writing, and to the best of our knowledge, there are two primary learning algorithms for ONNs implemented on SHNOs and STNOs. The first was initially presented by Vassilieva et. al. [63] in 2011 in a mathematical context (by which we mean that the algorithm is presented purely mathematically and is not tailored for a specific oscillator implementation). An overview of this algorithm is given in Section 5.6.1, but the synopsis is that this algorithm relies on adjusting the natural frequencies of the oscillators, promoting synchronization between certain pairs and disincentivizing between others in accordance with the expected output. Note that a "pair" of oscillators, in this case, refers to a pairing of the input oscillator to any other non-input oscillator in the system, all of which effectively function as output oscillators. There is no interaction between output oscillators. As pointed out by Vodenicarevic et. al. [60], this algorithm is greedy, only acting on IO-pair as though they are isolated and do not influence each other. For some systems this is an acceptable assumption [61]. But generally it is not a good assumption, and the algorithm may end up giving suboptimal results [60]. Furthermore, without inter-oscillator coupling, the size of a problem that can be presented to the oscillators is limited [60].

Vodenicarevic et. al. presented an algorithm more in line with existing machine learning architecture [60]. This algorithm considers coupling between all the oscillators, not just input-output pairs. A detailed mathematical summary of their learning algorithm is given in Section 5.6.1. The operating principle of adjusting the natural frequency of the output oscillators is the same as in Vassilieva et. al., but the algorithm is able to implement gradient descent by design.

In both these methods, the coupling strength is kept fixed, and thus it lends itself to the Kuramoto model. Indeed, Vodenicarevic et. al. rely on the Kurmaoto model in their paper [60], and previous work [81] has shown that it is a good approximation, despite its simplicity.

Both of these methods have been realized. Experimentally, Romera et. al. [49] successfully trained a network of four STNOs to recognize spoken vowels using the learning rule presented by Vassilieva et. al. In simulations, Singh et. al. [61] used the same learning algorithm for a classification scheme using SHNOs synchronizing to the frequencies of two

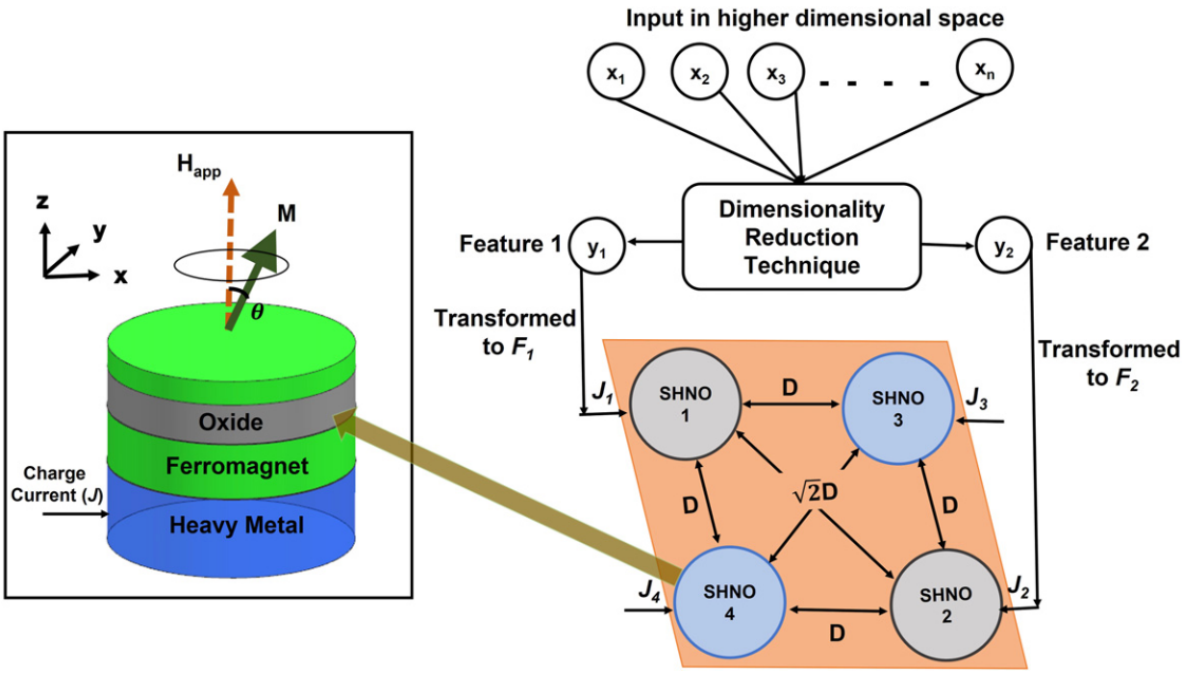



Figure 9.1: Micromagnetic setup of a SHNO neural network, presented by Garg et. al. [12]. The SHNO itself, depicted on the left, consists of heavy metal ferromagnetic (HM/FM) bilayer, where a charge current $J$ is applied to the ferromagnet. The easy axis is maintained by a magnetic field $H_{app}$. The right depicts the proposed neural network, where SHNO1 and SHNO2 are input oscillators to which currents $J_1$ and $J_2$ (corresponding to input data) are applied. SHNO3 and SHNO4 are output oscillators, subject to currents $J_3$ and $J_4$ corresponding to their training. The placement is such that the input-output pairs are a distance D apart, while the input-input and output-output are $\sqrt{2}$D apart. A dimensionality reduction technique converts the higher-dimensional dataset $\{x_1...x_n\}$ to $y_1$ and $y_2$, each input to one input oscillator.

RF magnetic fields containing the input data. Garg et. al. used the Vodenicarevic algorithm to train a simulated network of two input and two output oscillators for binary classification of flowers [12]. In the latter case, the binary state was based on whether or not the output oscillators were synchronized to each other.

Additionally, Garg et. al. used the Kuramoto model extensively in tandem with their micromagnetic simulations. This is encouraged by the Vodenicarevic algorithm, though it is not the only choice [60]. A secondary effect of the Kuramoto model is that a lot of the micromagnetic work can be outsourced, drastically reducing the necessary computational resources. In subsequent work, we will use the Kuramoto model for the same reason, however we will explore some aspects not yet discussed in preceding papers.

All of these realizations, both computational and experimental, have regarded the easy-axis SHNO (an example of which is depicted on the left in Figure 9.1) rather than the easy-plane SHNO (EP-SHNO). Having explored the various capabilities of EP-SHNO

synchronization, we will be extending the work of Garg et. al. to it.

## 9.1 System layout

At present, we have put most of our focus on an EP-SHNO on a common Pt bar in Section 7 and Section 8. This system is not symmetric, as the common Pt bar allows the iSHE to impact downstream SHNOs more than their upstream counterparts. The effect of this has been studied in Section 7, where we saw that the frequency of downstream oscillators was increased by iSHE. Additionally, with iSHE present, Section 8 shows a change in slope for the instantaneous frequency as a function of applied voltage. This disappeared when the common driving Pt bar was removed and the currents were applied completely independently, which is what is expected from the Kuramoto model (see Section 4 for more details). Going forward, we will consider only the case of separated oscillators. This is to reduce the number of effects impacting the system, in addition to making the synchronization range more predictable.

Figure 9.2 shows our proposed EP-SHNO neural network. This system consists of four separate insulating ferromagnetic meshes placed atop individual conducting meshes. The ferromagnets are oriented so that they face each other, all oscillating in the same $xz$-plane (coplanar). The hard axes are all oriented along the $\hat{\mathbf{y}}$-axis, perpendicular to the current, with the charge current flowing inside of the easy-plane. The SHNOs will thus oscillate within the same plane, though with a slight out-of-axis contribution as a result of the SHE. This SHE is provided by an applied current in the conducting layer, set to some value to keep the oscillator at a natural frequency. This natural frequency is defined as the frequency the oscillator would have in the absence of any other perturbing factor, such as other oscillators.

This orientation is not strictly necessary. The oscillators presented by Garg et. al., depicted in Figure 9.1, were oriented such that their easy axes were parallel, without facing each other. However, our work earlier in this thesis, for instance in Section 7 and Section 8 considered coplanar EP-SHNOs. A natural extension of this work is to consider a neural network with coplanar oscillators. It is not known if non-coplanar EP-SHNOs are able to synchronize to the same degree as considered earlier. This should be investigated.

In Figure 9.2, the left-right opposing pair is marked as I1 and I2, our input oscillators. Here, data encoded as current will set the natural frequency of these oscillators. The top-bottom opposing pair, O1 and O2, are the output oscillators. These will be provided a current to attain some natural frequency, which is tuned based on the training, moving iteratively towards the most optimal value for our dataset. We define a heterogeneous pair as any combination of input and output oscillators, and a homogeneous pair as the two input oscillators or the two output oscillators. By design, all pairings of one type (e.g. all heterogeneous oscillators) have the same coupling constant: They are spaced in

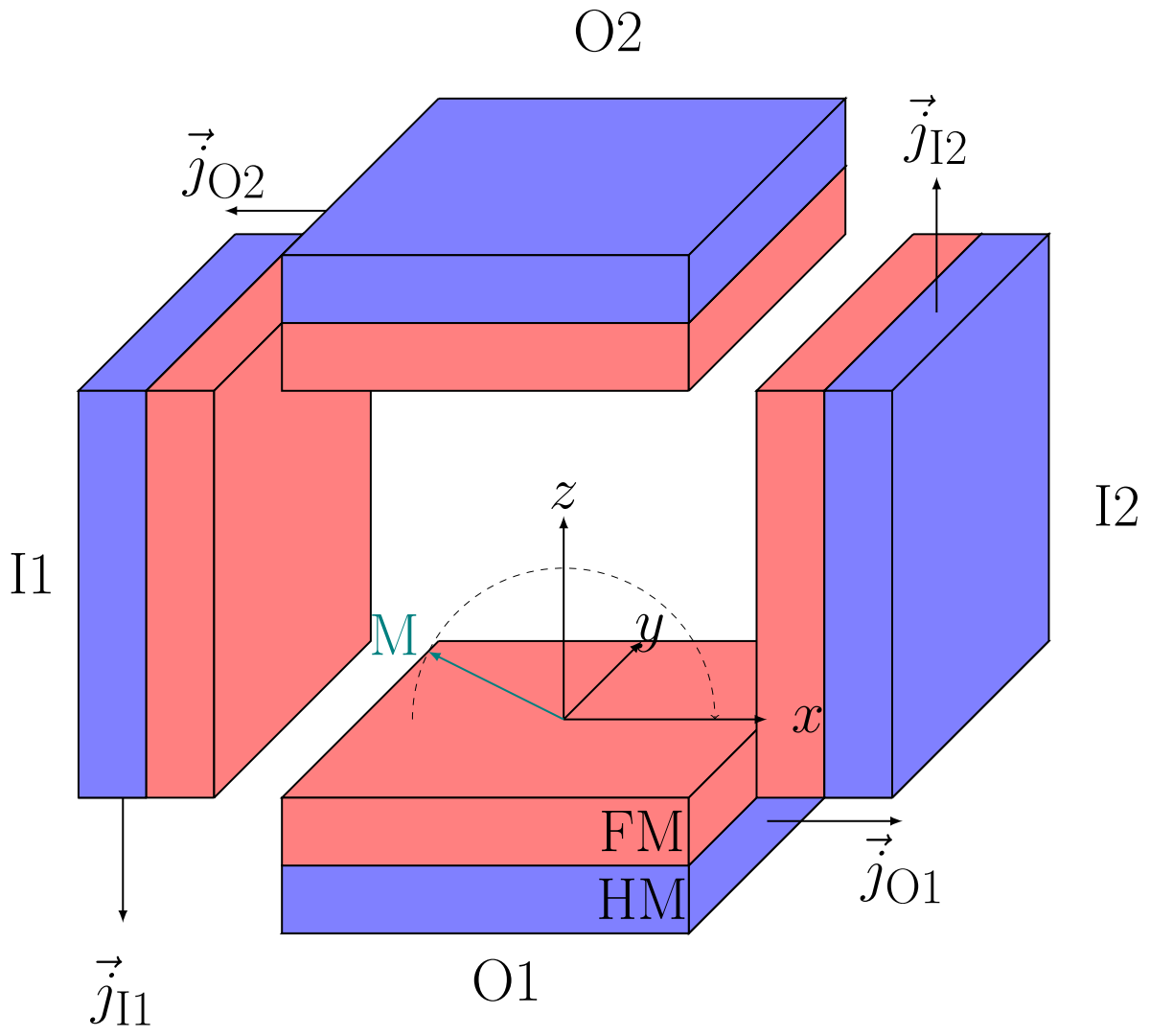



Figure 9.2: Diagram of the Vodenicarevic-style EP-SHNO neural network. Red meshes are ferromagnetic (FM) meshes, blue meshes are heavy metal (HM) meshes with the properties of platinum (Pt). I1 and I2 are the two input oscillators, O1 and O2 are the two output oscillators. The corresponding currents $\vec{j}_{I1}$, $\vec{j}_{I2}$ etc. flow through the HM layers.

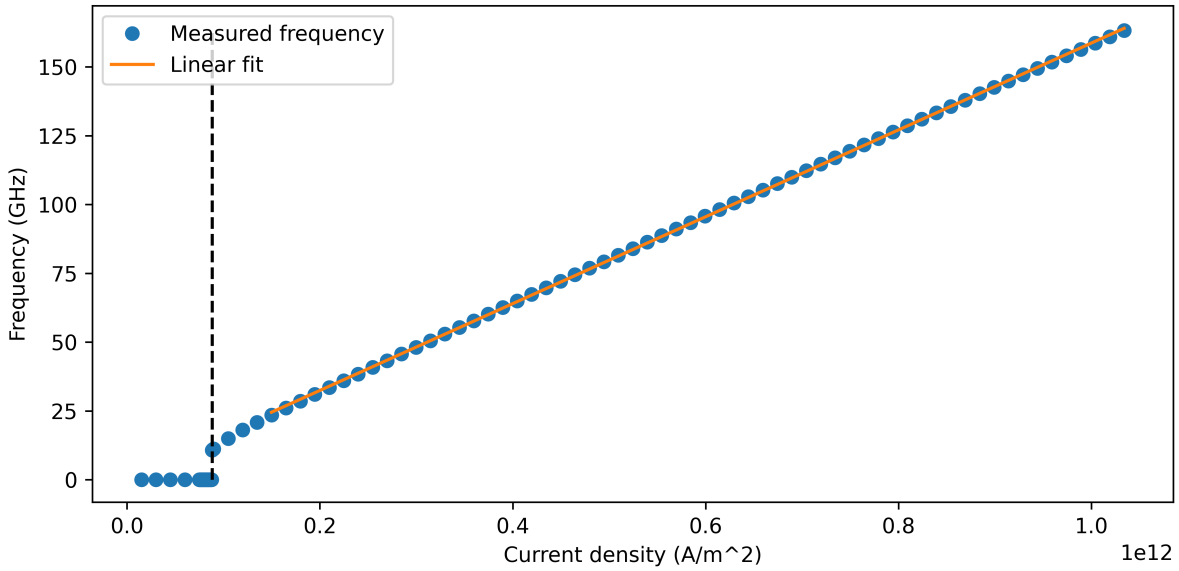



Figure 9.3: Oscillator frequency as a function of charge current density in its associated platinum layer. Prior to the activation current density ($(88 \pm 0.377) \times 10^9 \mathrm{A/m^2}$, corresponding to $(147.5 \pm 0.625)\mu\mathrm{V}$), marked with a dashed vertical line, the oscillator is stationary in an equilibrium. After the activation voltage, the frequency increases linearly with the current density, with a slope of $(0.158 \pm 0.0002)\mathrm{Hz/Am^2}$, corresponding to $(0.095 \pm 0.0001)\mathrm{GHz/\mu V}$.

such a way that the distance, i.e. coupling strength $k_{\mathrm{het}}$, is the same for any other pair of the same type [12]. The homogeneous oscillators also all have the same coupling strength, $k_{\mathrm{hom}} < k_{\mathrm{het}}$. Finally, we define the natural frequency for the oscillator as $F_{\mathrm{I1}}$, $F_{\mathrm{I2}}$, $F_{\mathrm{O1}}$ and $F_{\mathrm{O2}}$, and the instantaneous frequencies as $f_{\mathrm{I1}}$, $f_{\mathrm{I2}}$, $f_{\mathrm{O1}}$ and $f_{\mathrm{O2}}$.

## 9.2 Activation frequency

As we are no longer working with a separate driving and perturbing current, it is prudent to characterize this current (density) and its response in greater detail. Primarily, the activation current density, that is the current density at which the torque from the iSHE breaks equilibrium with the restoring forces such as anisotropy, and makes the SHNO oscillate. This will yield a minimum expected frequency, and a relationship between the applied charge current and frequency, which we will use to determine the correct current to attain a given natural frequency. In this setup, we remove all but one oscillator, I1, and apply a gradually increasing voltage difference across its adjacent platinum layer, just as before.

Figure 9.3 details the results of this. The voltage applied is slowly tuned from 0V to 1725 µV, with additional simulations run between 125 µV and 150 µV to more accurately establish the activation voltage. Initially, the torque from the applied current is

insufficient to cause oscillations, and the frequency remains at 0 Hz for all times. At approximately 147.5 µV, the torque from SHE is sufficient to overcome the anisotropy, and the SHNO begins to oscillate. This leads to a discontinuous jump from 0 Hz to approximately, 10 GHz. Beyond this activation value, the frequency grows linearly with the applied voltage. This is a non-linear activation function, which may be useful for other implementations of neural networks. In this work we are instead concerned with the slope in this relationship.

Note that Figure 9.3, and all subsequent plots, will work in voltage rather than current density, unlike Garg et. al. [12]. The reasons are manyfold: Primarily, it is easier to work with in the software we are using, BORIS computational spintronics [32], and thus easier to report. Secondly, the voltage and current density are linearly related for a conductor of fixed geometry, as is the case in our work. Thirdly, for such small devices as the ones we are using, reducing or increasing the current density by altering the conductor area does not necessarily reduce or increase the applied torque in a corresponding manner, due to the complicated dynamics of the electrons as a result of SHE and iSHE in such a thin conductor, making it difficult to control and thus a suboptimal way of measure.

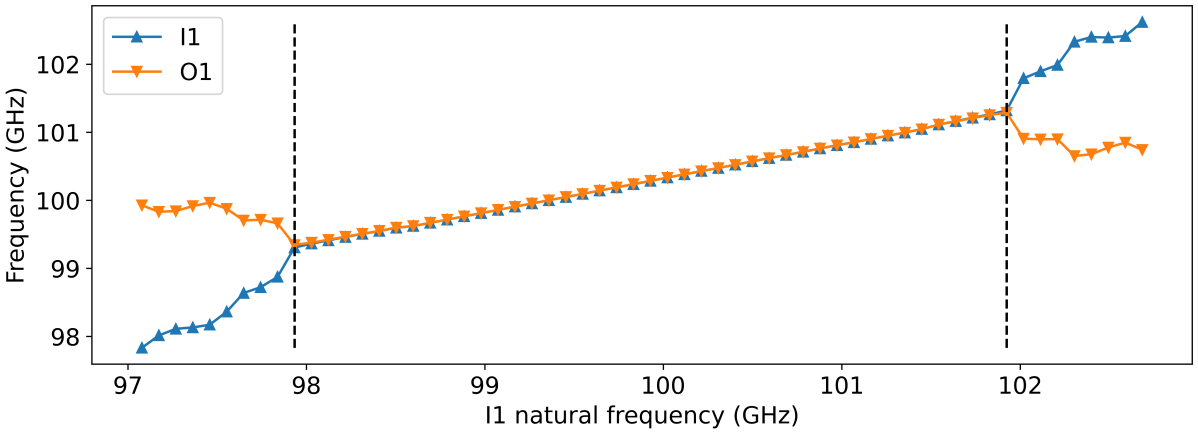
The linear relationship between the input voltage and the oscillator frequency will be exploited as a means of setting the natural frequency. This, in addition to the coupling constant, are the two unknowns needed to simulate the system via the Kuramoto model.

## 9.3 Determining coupling coefficients

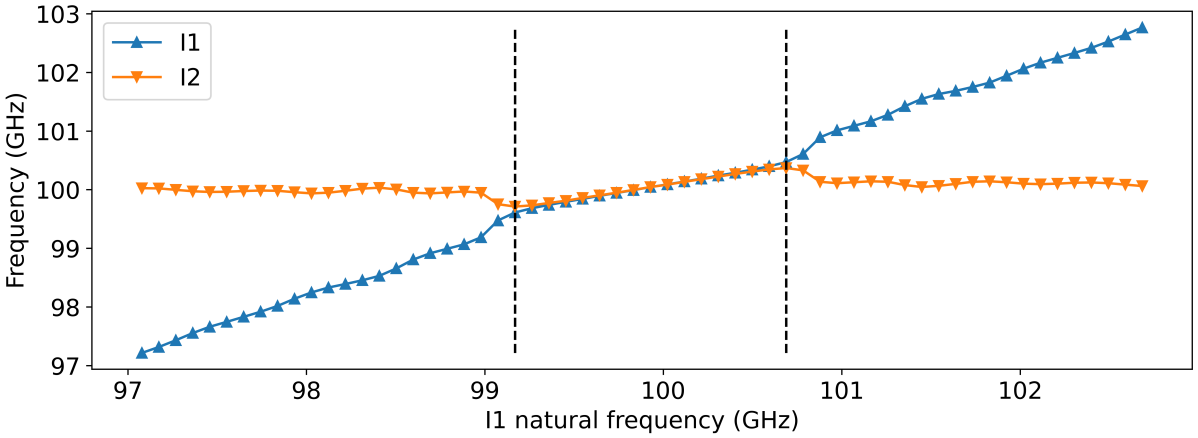
In the absence of altering the material parameters of an SHNO, either directly or via electrical means, the Kuramoto model coupling constant $k$ is proportional to oscillator separation [12]. This exclusive separation-dependence of the coupling is intrinsic to magnetic fields, with their strength decreasing over distance. Therefore, we wish to relate the coupling coefficient $k$ to the synchronization range using the Kuramoto model, and then extend it to the real system of oscillators. With this, we will be able to determine the coupling coefficient of both our heterogeneous and homogeneous pairs using the synchronization range, and thus simulate the system using the Kuramoto model.

Compared to Garg et. al, who mostly used frequencies in the GHz range [12], our oscillators are smaller with higher frequencies. This is seen in Figure 9.3, where the lowest oscillating frequency is approximately 10 GHz. Therefore, we specify that our oscillators operate on the order of 100 GHz to be sufficiently far from the activation frequency range. Our earlier micromagnetic simulations in Section 8.2 show that the synchronization range usually is around 1 GHz in size. In these simulations, the oscillators are closer and oriented differently, so we expect it to be on the order of a few GHz instead.

The synchronization ranges of our micromagnetic SHNOs are depicted in Figure 9.4. For the heterogeneous pairs, Figure 9.4a, the synchronization range is 3.9 GHz, while for




(a) Heterogeneous oscillator pair




(b) Homogeneous oscillator pair

Figure 9.4: Synchronization ranges of a homogeneous and a heterogeneous pair of micromagnetic oscillators. In both cases, the natural frequency of one oscillator (O1 and I2 respectively) is held at approximately 100 GHz, and the other (I1) is varied from 97 GHz to 103 GHz. Oscillator names correspond to the previous setup, which implies that the first setup has the oscillator oriented perpendicularly to each other, while in the latter they are face-to-face, though separated by a larger distance.

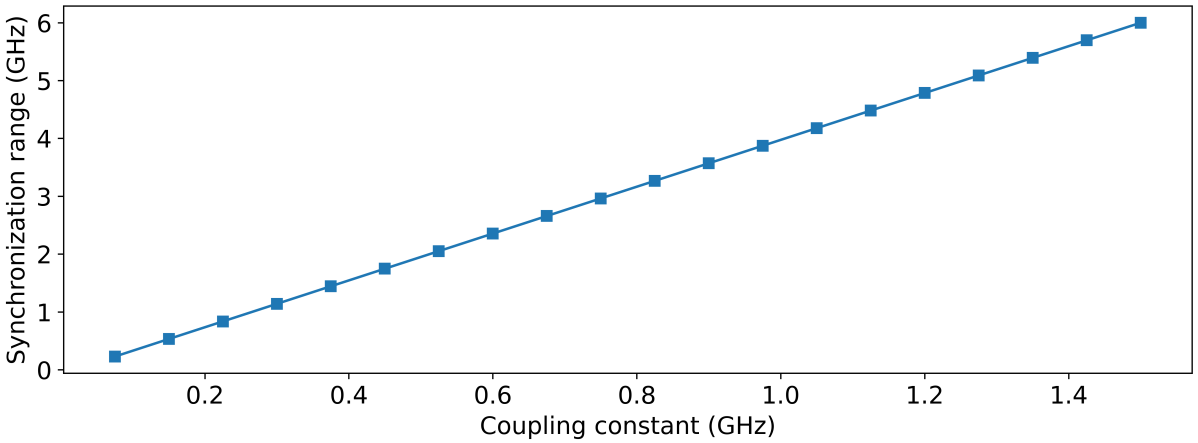



Figure 9.5: Synchronization range as a function of Kuramoto model coupling constant.

the homogeneous pairs, Figure 9.4b, it is 1.33 GHz.

Next, we run simulations with the Kuramoto model, steadily increasing the coupling constant $k$ for a fixed range of natural frequencies, to obtain a relationship between the synchronization range and the coupling constants. This is depicted in Figure 9.5. Here we immediately see a linear relationship between the coupling constant and synchronization range, and more importantly the synchronization range is 4 times larger than the coupling constant. This is something we expect from the Kuramoto model, discussed in Section 4. Thus, the synchronization ranges above correspond to $k_{\text{het}} = 0.97\,\text{GHz}$ and $k_{\text{hom}} = 0.33\,\text{GHz}$. The ratio between these is $\frac{0.97}{0.33} \approx 2.939$, much greater than the $\sqrt{2}$ value reported by Garg et. al. This is notable: The layout of our oscillators is such that the ratio between the center-center distance of homogeneous pairs is $\sqrt{2}$ greater than heterogeneous pairs, just as in Garg et. al. [12]. Yet the relative coupling strength is not $1/\sqrt{2}$ as they report, suggesting that the relative orientation of the oscillators also plays a role.

To verify that the obtained $k$ can accurately describe our system, we run a corresponding series of Kuramoto model simulations with $k = 0.97\,\text{GHz}$. The results are depicted in Figure 9.6. The behavior is quantitatively consistent with that of the heterogeneous micromagnetic pair depicted in Figure 9.4a: The synchronization range is nearly identical in size to that of the micromagnetic oscillators, with the slight offset likely caused by inaccuracies converting from desired frequency to input voltage.

## 9.4 Predictions via Kuramoto model

With the coupling coefficients for our system obtained, we can in theory model our ONN using the Kuramoto model. The goal is to obtain a set of system parameters, such as output oscillator initial frequencies, input oscillator range, in addition to the corresponding synchronization map of this setup. At present, only two parameters are currently char-

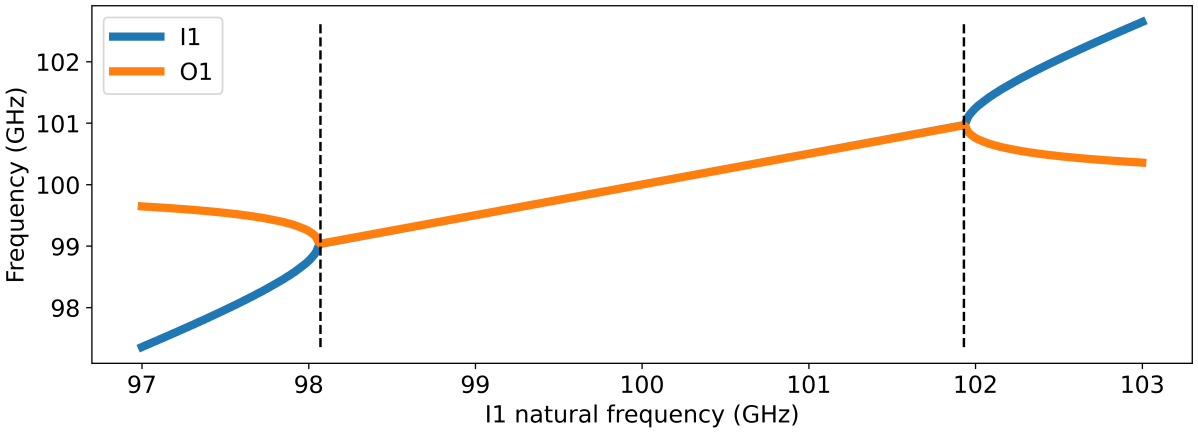



Figure 9.6: Synchroniztion range for $k = 0.97\,\mathrm{GHz}$.

acteristic of our system: $k_{\mathrm{het}} = 0.97\,\mathrm{GHz}$ and $k_{\mathrm{hom}} = 0.33\,\mathrm{GHz}$. These are, as explained previously, invariant of the magnitude of the natural frequencies.

We carry out these simulations using a Python-implementation of the Kuramoto model [82]. This library defines the interaction between the oscillators as an adjacency matrix, to which the coupling constant is multiplied. In our case, the adjacency matrix is of form

$$\begin{array}{c} \\ \mathrm{I1} \\ \mathrm{I2} \\ \mathrm{O1} \\ \mathrm{O2} \end{array} \begin{array}{c} \begin{array}{cccc} \mathrm{I1} & \mathrm{I2} & \mathrm{O1} & \mathrm{O2} \end{array} \\ \begin{pmatrix} 0 & 1/2.939 & 1 & 1 \\ 1/2.939 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1/2.939 \\ 1 & 1 & 1/2.939 & 0 \end{pmatrix} \end{array}$$

All diagonal elements are 0 to prevent self-coupling. The elements corresponding to heterogeneous pairs are set to 1, while the elements corresponding to homogeneous pairs are reduced by a factor of 2.939, the ratio $k_{\mathrm{het}}/k_{\mathrm{hom}}$. This matrix is multiplied by $0.97\,\mathrm{GHz}$, ensuring that all oscillators are coupled correctly. This adjacency matrix is constant throughout all simulations.

The other parameters which define the synchronization map are the natural frequencies of the output oscillators, as these values define the susceptibility for synchronization depending on the coupling constant. Less initial separation of the natural frequencies makes the oscillators more susceptible to synchronization, while for a sufficiently large separation the coupling is not great enough to cause any synchronization. Initially, we will also keep these fixed to show how they respond to variations in the input oscillator natural frequencies.

Figure 9.7 shows two types of plots side-by-side. The left columns are frequency maps, which depict the absolute frequency difference between the two output oscillators (encoded in the colormap) as a function of the two input oscillator natural frequencies,

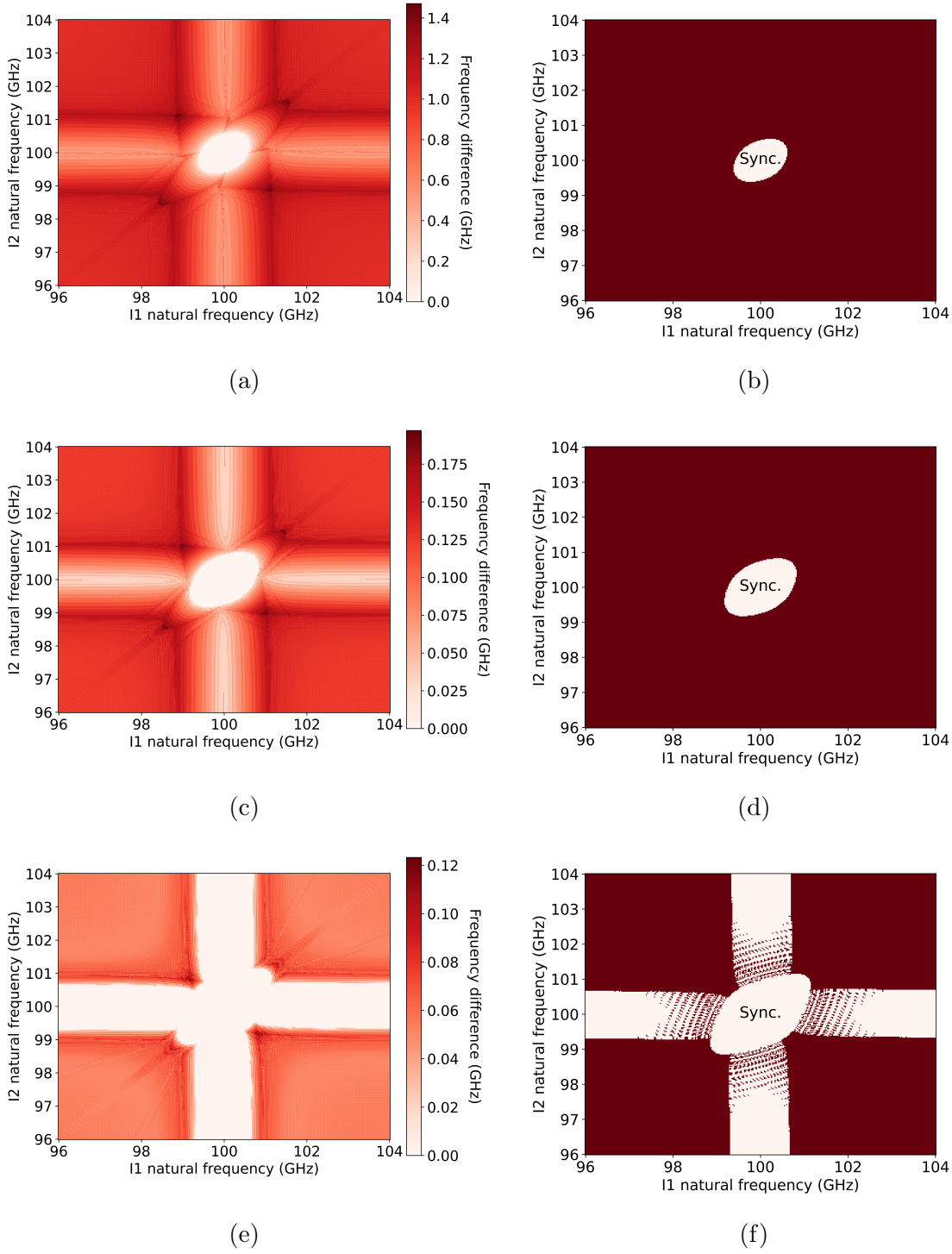



Figure 9.7: Frequency maps (left column) and synchronization maps (right column) for a 4SHNO network of 2I/2O oscillators. The natural frequencies of I1 and I2 correspond to the x- and y-axes, with $k_{\text{het}} = 0.97\,\text{GHz}$ and $k_{\text{hom}} = 0.33\,\text{GHz}$. The output oscillator natural frequencies are: a) b) 99.5 GHz and 100.5 GHz, c) d) 99.6 GHz and 100.4 GHz e) f) 99.8 GHz and 100.2 GHz. These are selected as they best illustrate the behavior. In the synchronization map, white corresponds to synchronization and red to desynchronization.

I1 and I2. This data is used in the creation of the synchronization maps, depicted on the right, which compare the absolute frequency difference to a threshold value of 0.0012 GHz [12] to obtain Boolean data: Synchronized or not. This threshold was chosen by Garg et. al. in their work, and at present we use the same value. It is possible to select a different threshold, so long as both the Kuramoto model predictions and the later micromagnetic simulations are subject to the same value.

The synchronization can be described as spontaneous, or happening very suddenly when slowly probing through the frequencies. This is not readily captured in the synchronization map due to the magnitude of the desynchronized frequencies. Consider the white desynchronized region found in the center of the plot: Inside of it, the frequency difference is on the order of 10 Hz, compared to the natural frequency difference itself which is on the order of 1 GHz. Simultaneously, one can achieve a frequency difference of 100 MHz by moving only a few data points outside of the synchronized region, growing into the GHz-range at the largest. This is the steady-state behavior that becomes apparent for very long simulations, in our case 100 ns per data point. This also allows a large degree of flexibility for the synchronization threshold: Previously it was set to 0.0012 GHz, but it can be set to a much higher value (e.g. 0.02 GHz) without altering neither the shape nor the size of the synchronization region.

All the frequency maps in Figure 9.7 show the emergence of horizontal and vertical regions, forming a cross, where the frequency difference is minimal and approximately constant for variations of a single input oscillator. For a sufficiently small separation, this transfers over to the corresponding synchronization map. We additionally note that, in the frequency map, there is a defined minimum and maximum value of the oscillators, even as the input natural frequencies grow. Both have the same root cause: Figure 9.4 reveals that, far beyond the synchronization range, the oscillators will relax towards their natural frequencies, and lower-frequency oscillators do not feel the presence of a nearby high-frequency oscillator. Once the input frequencies become either both too large, both too small, or a combination of the two, the output oscillators stop interacting with the input oscillators and instead relax towards their natural frequencies. Thus we obtain regions where the frequency difference is effectively a constant and equal to $F_{\mathrm{O1}} - F_{\mathrm{O2}}$.

The cross is explained by one input oscillator (for instance I1) having a frequency too large or small relative to either output oscillator, such that they do not couple to any significant degree. At the same time, the second input oscillator (for instance I2) has a frequency such that it individually influences both output oscillators, pulling their frequencies closer together. This natural frequency is $F_{\mathrm{I2}} \approx 100\,\mathrm{GHz}$, the midpoint between $F_{\mathrm{O1}}$ and $F_{\mathrm{O2}}$. For a sufficiently small initial separation of $F_{\mathrm{O1}}$ and $F_{\mathrm{O2}}$, this mutual tuning is sufficient to cause synchronization, $f_{\mathrm{O1}} = f_{\mathrm{O2}}$. This minimal region remains predominantly unchanged for variations of $F_{\mathrm{I1}}$ save for the central synchronized region, where all oscillators in the system become synchronized. This shows that the synchronization

along this cross is robust, and unaffected when $F_{\mathrm{I1}}$ is small or large compared to all other frequencies.

The system is invariant when swapping I1 and I2, and so we get the same behavior emerging when I1 is fixed while I2 varies, yielding a cross. For a system of many output oscillators, there will be many overlapping crosses in the full synchronization map. An example of this is seen in Romera et. al. [49]. In our case, we are concerned only with two output oscillators, and we will focus primarily on situations that do not have crosses, for instance Figure 9.7b.

Although there exists a minimal cross in Figure 9.7a, the resulting frequency difference is not small enough to qualify as synchronized in Figure 9.7b. The reason is likely that the coupling between the I2 and the output oscillators is not strong enough to close the 1 GHz natural frequency gap between O1 and O2, unlike in the case of smaller natural frequency gaps.

## 9.5 Micromagnetic simulations

Thus far we have assumed that our micromagnetic system could be modeled via the Kuramoto model. This is the baseline for the work by Garg et. al. [12] and, in part, Vodenicarevic et. al. [60]. If this assumption is true, we should be able to replicate the synchronization maps in Figure 9.7 using micromagnetic simulations. This figure is generated by running 300x300 simulations with the Kuramoto model for 4 oscillators. Running that many micromagnetic simulations with the necessary level of resolution will require far too many computational resources. This was the reason for switching to the Kuramoto model in the first place. Therefore, we simulate a subset of the input frequency combinations and verify them against the synchronization map obtained via the Kuramoto model. The frequency map is not obtained, as there will be too few micromagnetic simulations to make make it meaningful. In theory, and for enough simulations, it should be very similar to the frequency map obtained from the Kuramoto model.

We first apply the linear relationship for voltage and frequency found in Section 9.2 to obtain the correct voltage difference for our oscillators. Following that, we let the micromagnetic system run for 2 ns, and use the same method as for the Kuramoto model to obtain the synchronization map.

Figure 9.8 depicts the micromagnetic simulation overlaid the Kuramoto model prediction. The backing Kuramoto data is the same as described above. For the micromagnetic cases, we set $F_{\mathrm{O1}} = 99.5\,\mathrm{GHz}$ and $F_{\mathrm{O1}} = 100.5\,\mathrm{GHz}$, and let $F_{\mathrm{I1}}$ and $F_{\mathrm{I2}}$ vary according to the green and yellow points. These points, green and yellow, correspond to the output oscillators being synchronized and desynchronized respectively, just as with the Kuramoto model prediction. One change from our earlier work is that we have changed the threshold for synchronization: Our previous value of 0.0012 GHz was based on the

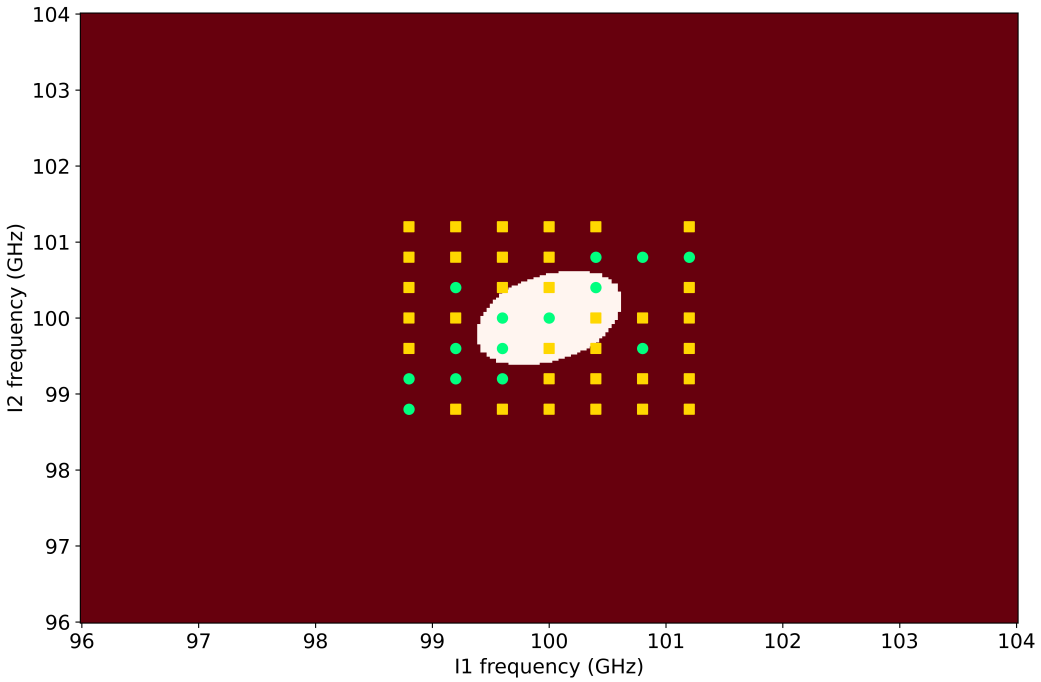



Figure 9.8: Micromagnetic simulations overlaid the synchronization map obtained via the Kuramoto model. The micromagnetic output oscillator natural frequencies are 99.5 GHz and 100.5 GHz, selected based on Figure 9.7b. Yellow squares represent desynchronized output oscillators and green circles represent synchronized output oscillators, for corresponding input oscillator natural frequencies.

work of Garg et. al. [12], performed with micromagnetic oscillators in the 6 GHz range. In addition, their predicted synchronization ranges were on the order of 0.2 GHz. In our case, we are working with oscillators in the 100 GHz range with synchronization ranges of 4 GHz. As we discussed previously, one can alter the synchronization threshold within a large range without causing significant changes to the synchronization region. We exploit this to rescale our threshold by 100/6, such that it becomes $|f_{O1} - f_{O2}| < 0.02\,\text{GHz}$. This is still within the aforementioned range. This is applied to both the Kuramoto model data and to the micromagnetic data.

If the Kuramoto model is a good fit, we expect the green points to more or less overlap with the white (synchronized) region of the Kuramoto model, and the yellow points to lie in the red (desynchronized) region.

We see in Figure 9.8 that there are both yellow squares in the synchronization region, and green circles outside of it. In addition, the synchronization does not appear as rigorous, with yellow squares embedded within what is supposed to be a wholly green synchronization region. This may in part be due to the simulations being run for insufficient time: We used a 2 ns run time for each simulation, however for obtaining the synchronization ranges a 5 ns run time was considered sufficient for materially separated oscillators. If the system were to be run for longer, the long-term synchronization dynamics may properly separate the space into two distinct regions, just like the underlying Kuramoto paper.

However, it may not be sufficient for treating the shape and size of the region: We note that the upper-left side and lower-right side are uniformly desynchronized, as expected. In other words, there is great overlap along the semi-minor axes of the synchronization ellipse. This, in addition to the overlap with synchronized states in the expected synchronization region, are the most positive attribute on the micromagnetic synchronization map. Along the semi-major axis (upper-right and lower-left), the synchronized states leak over into the desynchronized region, stretching the micromagnetic synchronized region. Without characterizing the extent of the synchronization region along the semi-major axes, we cannot state whether or not they are centered on the same point, $F_{I1} = F_{I2} = 100\,\text{GHz}$. If they are, one could achieve some success in tuning the synchronization threshold, in addition to running the simulations for longer. However, it may not solve the shape issue entirely, as at some point this will start excluding more states along the semi-minor axes too.

It should be noted that the semi-major axis coincides with the linear $F_{O1} = F_{I2}$. This is to be expected of the Kuramoto model, as $k_{het} > k_{hom}$, and so when the input oscillators have the same natural frequency, they pull at each output oscillator with twice the force in the same direction. This stretches the synchronization region in that direction. If the micromagnetic synchronization region is centered at $F_{I1} = F_{I2} = 100\,\text{GHz}$, it indicates that our $k_{het}$, obtained via the Kuramoto model, is far lower than the micromagnetic

system. This cannot be stated conclusively until a full picture of the micromagnetic system is obtained.

The aforementioned issues regarding frequency-to-voltage conversion may play a role in this offset: With our choice of $F_{O1} = 99.5\,\text{GHz}$ and $F_{O2} = 100.5\,\text{GHz}$, we can use the case of $F_{I1} = F_{I2} = 100.0\,\text{GHz}$ as a calibration case. This case is supposed to be synchronized, and with the input oscillator frequencies being equal and located between that of the output oscillators it stands to reason that they should all synchronize to 100 GHz given enough time. Our simulation output reports the following instantaneous frequencies after 2 ns: $f_{I1} = 100.197\,\text{GHz}$, $f_{I2} = 100.180\,\text{GHz}$, $f_{O1} = 100.02\,\text{GHz}$, $f_{O2} = 100.018\,\text{GHz}$. The output is indeed synchronized, however the frequencies of the input oscillators are notably higher than both output oscillators and very close to the edge of what is defined by synchronization. The fact that both oscillators overshoot is an indication that the conversion is wrong and in need of refining, and it may also create a leftwards offset of the synchronization region.

Our simulations were too centered on the Kuramoto model synchronization range to characterize the apparent range of the micromagnetic oscillators. Therefore, we cannot state with confidence that the Kuramoto model is sufficient or insufficient to model this system of EP-SHNOs. We will highlight two positive observations: First, the micromagnetic synchronization region shows an indication of an elliptical shape like the underlying Kuramoto model data, and it appears that the semi-minor axis of the elliptical region corresponds to the micromagnetic model. It is the semi-major axis that is much larger. Second, we have previously shown that the Kuramoto model can be used to accurately characterize the synchronization ranges of two EP-SHNOs. Based on these observations, we believe that, with refinement, our presented methodology can be applied to a neural network of EP-SHNOs.

## 9.6 Discussion and outlook

Based on our results, we will discuss the continuation of this project, both completing the work already done and continuing it into future projects. First, the issues and work that remain on our present attempt at an EP-SHNO neural network, and then the next step that goes beyond our setup.

### 9.6.1 Remaining work

Previous work by Vodenicarevic et. al. has created an offline learning algorithm for oscillator-based classifiers that is compatible with the constraints of nanotechnology [60]. Garg et. al. have implemented it in and successfully trained a network of four SHNOs [12] in a similar setup to what is presented here. In this work, we have emulated the results leading up to the implementation of a network of four coplanar EP-SHNOs. First,

we ran several micromagnetic simulations of a 1SHNO system to obtain an activation voltage (and thus frequency), in addition to a relation between the applied voltage and the resulting natural frequency. For our system, we obtained a value of $0.095\,\mathrm{GHz/\mu V}$. With this in mind, we created a micromagnetic 2SHNO system of heterogeneous and homogeneous oscillators separately, and ran simulations to obtain their synchronization ranges. We verified these results against those obtained by the Kuramoto model, and then used the Kuramoto model to obtain a relation between the synchronization range and coupling coefficient of two oscillators. Having obtained the coupling constant, we were able to model the micromagnetic system using the Kuramoto model, significantly reducing the computational resources necessary in addition to opening for using the Vodenicarevic et. al. learning algorithm.

With this simplified model, we used the Kuramoto model to obtain detailed synchronization and frequency maps for this simplified system. We then employed these synchronization maps to predict suitable values for the micromagnetic input frequencies (i.e. voltages), and attempted to verify the maps' validity by running micromagnetic simulations with a selection of natural frequencies. This would lay the groundwork for neural networks based on EP-SHNOs: The learning algorithm would be implemented using predictions by the Kuramoto model alone, and the behavior of the EP-SHNOs is close to that of the regular SHNOs, of which a full neural network has been realized in Garg et. al. [12]. Unfortunately, comparing the results of the micromagnetic system with that of its Kuramoto model representation yielded unsatisfactory overlap: The synchronization range is an ellipse, and while there is good overlap along the semi-minor axes of this ellipse, the micromagnetic simulations exhibit a much larger semi-major axis than its Kuramoto model counterparts.

The Kuramoto model is not perfect for this system: The presence of an easy-axis means that the natural frequency is dependent on the orientation of the magnetization. The magnetization prefers alignment with the easy-axis, and thus there is an additional torque depending on the orientation and magnitude of the magnetization. Additionally, increasing the input current will increase the OOP deflection of the magnetization. In effect, this reduces the oscillatory part of the magnetization for higher frequencies by placing a larger component of the magnetization out-of-plane. The oscillators used by Garg et. al. [12] were the more common, uniaxial type of SHNO similar to those that have been used experimentally [75], with the axis set by a magnetic field rather than anisotropy, which reduced the impact of both of these effects. Such oscillators therefore lend themselves more naturally to the Kuramoto model.

Another issue related to the OOP deflection is that the frequency as a function of input voltage, shown in Figure 9.3, is not strictly linear. This was assumed to be negligible based on previous work. We tested this with a calibration case of in Figure 9.8 where $F_{\mathrm{I1}} = F_{\mathrm{I2}} = 100\,\mathrm{GHz}$ while $F_{\mathrm{O1}} = 99.5\,\mathrm{GHz}$ $F_{\mathrm{I2}} = 100.4\,\mathrm{GHz}$. The symmetry of this

case should ensure that all oscillators synchronize with $f = 100\,\mathrm{GHz}$. However, we noted that all frequencies were above 100 GHz, and worse that the input oscillators both had higher frequencies than the output oscillator. This implies, at the bare minimum, that the conversion from voltage and frequency is wrong, and that the assumption of linear conversion is insufficient.

We recommend that the work of Garg et. al. is replicated more closely and in greater detail within BORIS Computational Spintronics, as it is a simpler system with known results with which one can compare. Once done, it can be extended to EP-SHNO in a more rigorous manner.

### 9.6.2 Novel work

This work has been following the steps of Garg et. al. [12] in its setup and methodology. This is to be expected, as we have been experimenting with a variation of the same device in an attempt at creating similar results. Future work should explore new aspects of the SHNO ONN. Scalability is one avenue: The setups presented both here in Figure 9.2 and in Garg et. al. [12] do not easily lend themselves to expansion. The result of the square setup in both Garg et. al. and in our work is that heterogeneous pairs of oscillators have the same coupling constant $k_{\mathrm{het}}$, and all homogeneous pairs have the same coupling constant $k_{\mathrm{hom}} < k_{\mathrm{het}}$. However, adding more output oscillators will not maintain this: The coupling constants are entirely separation-dependent, and the addition of more oscillators will cause certain pairs of the two types to have intrinsically weaker coupling. Consider adding two more output oscillators, O3 and O4, to the system shown in Figure 9.2 along the y-axis, such that we have two input and four output oscillators forming an open cube. In this case, O1 and O2 will couple as usual with a homogeneous strength, while O1 and O3 will couple with heterogeneous strength. This will have a profound impact on the system.

The presented system is also not easy to manufacture: We have neglected the presence of an insulating substrate which would be necessary for the manufacturing process, and additionally made the oscillators face each other, creating a hole on the scale of 20 nm width. At present, manufactured SHNOs primarily come in the form of nanoconstrictions [75][73], often in a chain similar to what was presented in Section 8. There are several issues with using this as a neural network, some of which have been discussed previously: First, each SHNO must be uniquely tuned to a specific natural frequency, which was solved by the addition of a perturbing conducting layer in addition to the common conducting layer upon which all SHNOs were deposited. This requires very precise manufacturing, especially when interfaced with other hardware to provide the current. Secondly, the same issue as presented above arises, as there is no layout of e.g. four SHNOs which results in the present requirement of uniform coupling constants between the different types of pairs.

Thirdly, the presence of iSHE will, as examined in Sections 7 and 8, cause a position- and time-dependent coupling. It is therefore prudent to investigate these effects, both micromagnetically and as modifications to the Kuramoto model.

Some of these scalability issues have been tackled in Bhotla et. al. [13]. This paper builds upon the work of Garg et. al. [12]. The authors suggest an additional restriction on the output oscillator natural frequencies, which would require that the natural frequencies of the output oscillators, $F_{O1}$ and $F_{O2}$, be separated by a frequency $\Delta F$ for every update. The reason for this is that we want to maintain the shape of the synchronization region: We see in the synchronization maps, on the right in Figure 9.7, that the shape and size of the synchronized region are dependent on the separation between $F_{\mathrm{O1}}$ and $F_{\mathrm{O2}}$. Secondly, we know that the size of the synchronization *range* is independent of the actual natural frequency (if O1 oscillates at 100 GHz and synchronizes to I1 when it is within $\pm 3$ GHz, then the same range of $\pm 3$ GHz applies if O1 is at 200 GHz). The result is that by maintaining a certain separation $\Delta F$, we can move the synchronization region around in space without destroying its shape. With this rather simple modification in place, Bhotla et. al. go on to suggest several additional modifications that can improve the scalability, incorporating multi-output systems. This would allow for more than binary classification, but the details are omitted here.

In Section 8, we discussed the prospect of thermal noise applied to the system and the issues arising for our micromagnetic oscillators. To date, there appears to be no work for an implementation of even a strictly Kuramoto model-based implementation of a neural network for $T \neq 0$. As the synchronization in the micromagnetic implementation proved to be very sensitive to thermal noise, it would be prudent to explore the effect of thermal noise on an otherwise idealized ONN implemented with the Kuramoto model only, before it is extended to SHNOs, especially EP-SHNOs.

# 10 Conclusion and future work

## 10.1 Conclusion

This thesis has chiefly concerned itself with the spin-Hall effects and their presence in computer simulations of two different spintronic devices: The spiking ferromagnetic neuron and the spin-Hall nano-oscillator. In the former, we showed that a heavy metal/ferromagnetic bilayer can be made to flip its direction by the application of a short, high-magnitude current pulse. The subsequent flip is called a spiking event, and due to spin-Hall magnetoresistancem it will create a characteristic current output. We can therefore create and measure spiking events with the same conductor and current. However, our presented device has a fundamental drawback in that it is not successfully blocking signals that do not cause spiking events. In addition, the output voltage is very low compared to the input and the shape is not right, meaning that our presented device can not cause self-consistent propagation.

The easy-plane spin-Hall nano-oscillator (EP-SHNO) came as an extension of the setup for the spiking neuron, by placing more spiking units on the same conductor and increasing the current to create continuous spiking - auto-oscillation. This type of oscillator has received less attention than the easy-axis spin-Hall nano-oscillator, and has not been shown to synchronize. Additionally, previous work has primarily considered the coupling SHNOs via their demagnetization fields. The inverse spin-Hall effect (iSHE) due to a common conducting layer has not been explicitly studied.

It is shown herein that EP-SHNOs synchronize in both low-damping and high-damping regimes. For low-damping, the synchronization is robust and with a negligible phase difference. For high-damping, the synchronization is dependent on the number of oscillators, and synchronizes with large phase differences. This system may also phase-synchronize, creating periodic oscillations in the phase difference that propagates from one oscillator to another in a time-dependent manner. These phase differences are due to the material properties of the oscillators, and do not depend on the initial configuration of phases. iSHE is shown to be a desynchronizing factor, increasing the frequencies of oscillators downstream of the charge current in a common conducting layer. It is shown that in the case of synchronization, the voltage fluctuations due to spin-Hall magnetoresistance (SMR) are minimized, showing its potential use for determining the synchronization of a whole system. However, the synchronization of these oscillators appears to be highly sensitive to temperature.

The synchronization of EP-SHNOs is shown to reliably break by tuning the natural frequency. This is much the same as the Kuramoto model of coupled oscillators. With this in mind, this thesis presents a setup for an oscillatory neural network (ONN) implemented on four interacting EP-SHNOs, relying on a learning algorithm previously used for a

similar implementation.

## 10.2 Future work

The attempted implementation of an ONN has issues regarding synchronization that needs to be resolved before future work can be attempted. In particular, an evaluation of the use of the Kuramoto model for this novel type of SHNO needs to be undertaken. Extending the Kuramoto model to account for anisotropic effects, damping, and temperature may improve its accuracy and thus obtain better results. Once completed, this system should be made more experimentally viable, for instance by reintroducing a common conducting layer between all the oscillators. A more detailed outlook is given at the end of Section 9.

In the implementation of spintronic devices, this thesis primarily considered the low-damping case. This was a choice of material parameters that stuck, and was beneficial due to the fast nature of their dynamics compared to their high-damping counterparts. However, simulations of the high-damping EP-SHNO show interesting behavior for the synchronization, differing wildly from the low-damping case. A thorough investigation of its synchronization, desynchronization, and phase layout, with a subsequent evaluation of its potential as an ONN device, should follow. Additionally, this thesis focused entirely on EP-SHNOs that were coplanar, that is oscillating in the same $xz$-plane with parallel hard-axes in the $\hat{\mathbf{y}}$-direction. Future work should investigate the synchronization of non-coplanar oscillators, either with perpendicular hard-axes or parallel hard-axes but offset easy-planes. This would open for more complex system geometries and thus more flexibility in their setup, allowing a possible ONN to consider more complex problems.

# References


[1] Luke James. Moore's law in 2023: What's the status quo?, Apr 2023. URL: https://www.power-and-beyond.com/moores-law-in-2022-whats-the-status-quo-a-dc63a87e669b554d4d33d2a5ba73692a/.

[2] Tim Cross. After moore's law — technology quarterly, Feb 2016. URL: https://www.economist.com/technology-quarterly/2016-03-12/after-moores-law.

[3] R.K. Ratnesh, A. Goel, G. Kaushik, H. Garg, Chandan, M. Singh, and B. Prasad. Advancement and challenges in MOSFET scaling. *Materials Science in Semiconductor Processing*, 134:106002, November 2021. doi:10.1016/j.mssp.2021.106002.

[4] J. von Neumann. First draft of a report on the EDVAC. *IEEE Annals of the History of Computing*, 15(4):27–75, 1993. doi:10.1109/85.238389.

[5] Techslang. Neuromorphic computing: Artificial intelligence's next phase?, Nov 2020. URL: https://www.techslang.com/neuromorphic-computing-artificial-intelligences-next-phase/.

[6] Jacques Mattheij. Another way of looking at lee sedol vs alphago, Mar 2016. URL: https://jacquesmattheij.com/another-way-of-looking-at-lee-sedol-vs-alphago/.

[7] Danijela Marković, Alice Mizrahi, Damien Querlioz, and Julie Grollier. Physics for neuromorphic computing. *Nature Reviews Physics*, 2(9):499–510, July 2020. doi:10.1038/s42254-020-0208-2.

[8] Or Sharir, Barak Peleg, and Yoav Shoham. The cost of training nlp models: A concise overview, 2020. URL: https://arxiv.org/abs/2004.08900, doi:10.48550/ARXIV.2004.08900.

[9] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org. URL: https://www.tensorflow.org/.

[10] Steve Furber. Large-scale neuromorphic computing systems. *Journal of Neural Engineering*, 13(5):051001, August 2016. doi:10.1088/1741-2560/13/5/051001.

[11] Catherine D. Schuman, Thomas E. Potok, Robert M. Patton, J. Douglas Birdwell, Mark E. Dean, Garrett S. Rose, and James S. Plank. A survey of neuromorphic computing and neural networks in hardware, 2017. arXiv:1705.06963.

[12] Neha Garg, Sri Vasudha Hemadri Bhotla, Pranaba Kishor Muduli, and Debanjan Bhowmik. Kuramoto-model-based data classification using the synchronization dynamics of uniform-mode spin hall nano-oscillators. *Neuromorphic Computing and Engineering*, 1(2):024005, November 2021. doi:10.1088/2634-4386/ac3258.

[13] Sri Vasudha Hemadri Bhotla, Neha Garg, Tanmay Aggarwal, Pranaba Kishor Muduli, and Debanjan Bhowmik. An oscillator-synchronization-based off-line learning algorithm, with on-chip inference on an array of spin hall nano-oscillators. *IEEE Transactions on Nanotechnology*, 22:136–148, 2023. doi:10.1109/tnano.2023.3250261.

[14] Hannah Bradley, Steven Louis, Cody Trevillian, Lily Quach, Elena Bankowski, Andrei Slavin, and Vasyl Tyberkevych. Artificial neurons based on antiferromagnetic auto-oscillators as a platform for neuromorphic computing, 2022. URL: https://arxiv.org/abs/2208.06565, doi:10.48550/ARXIV.2208.06565.

[15] Ralph Skomski. *Simple Models of Magnetism*. Oxford University Press, 01 2008. doi:10.1093/acprof:oso/9780198570752.001.0001.

[16] Amikam Aharoni. *Introduction to the theory of ferromagnetism*. International Series of Monographs on Physics. Clarendon Press, Oxford, England, 2 edition, November 2002.

[17] Mathias Getzlaff. *Fundamentals of magnetism*. Springer, Berlin, Germany, 1 edition, October 2008. URL: libgen.li/file.php?md5=94087c4f24d630bdd7612d02caa359b9.

[18] Alex Hubert and Rudolf Schäfer. Material parameters for domain analysis. In *Magnetic Domains*, pages 337–372. Springer Berlin Heidelberg, 1998. doi:10.1007/978-3-540-85054-0_4.

[19] David J. Griffiths. *Introduction to Electrodynamics*. Cambridge University Press, hardcover edition, 7 2017.

[20] Soshin Chikazumi. *Physics of Ferromagnetism (International Series of Monographs on Physics, 94)*. Oxford University Press, paperback edition, 6 2009.

[21] Massimiliano D'Aquino. *Nonlinear Magnetization Dynamics in Thin-Films and Nanoparticles*. PhD thesis, Università degli Studi di Napoli Federico II, Naples, Italy, 2005. URL: http://www.fedoa.unina.it/148, doi:10.6092/UNINA/FEDOA/148.

[22] S. Iida. The difference between gilbert's and landau-lifshitz's equations. *Journal of Physics and Chemistry of Solids*, 24(5):625–630, May 1963. doi:10.1016/s0022-3697(63)80004-9.

[23] Mike A. Lund, Akshaykumar Salimath, and Kjetil M. D. Hals. Spin pumping in noncollinear antiferromagnets. *Physical Review B*, 104(17), 2021. doi:10.1103/physrevb.104.174424.

[24] D.C. Ralph and M.D. Stiles. Spin transfer torques. *Journal of Magnetism and Magnetic Materials*, 320(7):1190–1216, April 2008. doi:10.1016/j.jmmm.2007.12.019.

[25] J.C. Slonczewski. Current-driven excitation of magnetic multilayers. *Journal of Magnetism and Magnetic Materials*, 159(1-2):L1–L7, June 1996. doi:10.1016/0304-8853(96)00062-5.

[26] Sina Mehboodi. Could you explain the mechanism of spin transfer torque and spin-orbit torque in ferromagnetic materials? when it is happening and how ?, 03 2022. URL: https://www.researchgate.net/post/could_you_explain_the_mechanism_of_spin_transfer_torque_and_spin-orbit_torque_in_ferromagnetic_materials_when_it_is_happening_and_how.

[27] Theory of spin-orbit torque, Jan 2023. URL: https://www.nist.gov/programs-projects/theory-spin-orbit-torque.

[28] Mikhail I. Dyakonov, editor. *Spin physics in semiconductors*. Springer Series in Solid-State Sciences. Springer, Berlin, Germany, 2008 edition, August 2008.

[29] Robert Karplus and J. M. Luttinger. Hall effect in ferromagnetics. *Physical Review*, 95(5):1154–1160, September 1954. doi:10.1103/physrev.95.1154.

[30] B. F. Miao, S. Y. Huang, D. Qu, and C. L. Chien. Inverse spin hall effect in a ferromagnetic metal. *Physical Review Letters*, 111(6), August 2013. doi:10.1103/physrevlett.111.066602.

[31] Yongcheng Deng, Meiyin Yang, Yang Ji, and Kaiyou Wang. Estimating spin hall angle in heavy metal/ferromagnet heterostructures. *Journal of Magnetism and Magnetic Materials*, 496:165920, February 2020. doi:10.1016/j.jmmm.2019.165920.

[32] Dr. Serban Lepadatu. *Boris Computational Spintronics User Manual*. University of Central Lancashire, Jeremiah Horrocks Institute, February 2023. doi:10.13140/RG.2.2.31496.88322/1.

[33] H. Nakayama, M. Althammer, Y.-T. Chen, K. Uchida, Y. Kajiwara, D. Kikuchi, T. Ohtani, S. Geprägs, M. Opel, S. Takahashi, R. Gross, G. E. W. Bauer, S. T. B. Goennenwein, and E. Saitoh. Spin hall magnetoresistance induced by a nonequilibrium proximity effect. *Physical Review Letters*, 110(20), May 2013. doi:10.1103/physrevlett.110.206601.

[34] Minh-Hai Nguyen and Chi-Feng Pai. Spin–orbit torque characterization in a nutshell. *APL Materials*, 9(3):030902, March 2021. doi:10.1063/5.0041123.

[35] A BRATAAS, G BAUER, and P KELLY. Non-collinear magnetoelectronics. *Physics Reports*, 427(4):157–255, April 2006. doi:10.1016/j.physrep.2006.01.001.

[36] Steven H. Strogatz. From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1-4):1–20, September 2000. doi:10.1016/s0167-2789(00)00094-4.

[37] Anonymous Says: and Eureka Logic says:. Brain performance in flops, Jul 2019. URL: https://aiimpacts.org/brain-performance-in-flops/.

[38] TOP500. November 2022, November 2022. URL: https://www.top500.org/lists/top500/2022/11/.

[39] Dr. Alan Woodruff. What is a neuron?, Aug 2019. [Online; accessed 22-May-2023]. URL: https://qbi.uq.edu.au/brain/brain-anatomy/what-neuron.

[40] Mark Wanner. 600 trillion synapses and alzheimers disease, Dec 2018. URL: https://www.jax.org/news-and-insights/jax-blog/2018/december/600-trillion-synapses-and-alzheimers-disease.

[41] Queensland Brain Institute University of Queensland. How do neurons work?, Aug 2022. [Online; accessed 22-May-2023]. URL: https://qbi.uq.edu.au/brain-basics/brain/brain-physiology/how-do-neurons-work/.

[42] K.N. LEIBOVIC. Chapter 2 - the resting potential. In K.N. LEIBOVIC, editor, *Nervous System Theory*, pages 7–11. Academic Press, 1972. URL: https://www.sciencedirect.com/science/article/pii/B9780124412507500064, doi:https://doi.org/10.1016/B978-0-12-441250-7.50006-4.

[43] Sanu N. English: Structure of neuron (unlabeled). https://commons.wikimedia.org/w/index.php?title=File:Structure_of_Neuron.png&oldid=504154484, 2020. [Online; accessed 22-May-2023].

[44] How are memories formed? [Online; accessed 22-May-2023]. URL: https://qbi.uq.edu.au/brain-basics/memory/how-are-memories-formed.

[45] Peter Uhlhaas. Neural synchrony in cortical networks: history, concept and current status. *Frontiers in Integrative Neuroscience*, 3, 2009. doi:10.3389/neuro.07.017.2009.

[46] Madeleine Abernot, Thierry Gil, Manuel Jiménez, Juan Núñez, María J. Avellido, Bernabé Linares-Barranco, Théophile Gonos, Tanguy Hardelin, and Aida Todri-Sanial. Digital implementation of oscillatory neural network for image recognition applications. *Frontiers in Neuroscience*, 15, August 2021. doi:10.3389/fnins.2021.713054.

[47] Xiao-Jing Wang. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews*, 90(3):1195–1268, July 2010. doi:10.1152/physrev.00035.2008.

[48] Scientific American. What is the function of the various brainwaves?, December 1997. [Online; accessed 22-May-2023]. URL: https://www.scientificamerican.com/article/what-is-the-function-of-t-1997-12-22/.

[49] Miguel Romera, Philippe Talatchian, Sumito Tsunegi, Flavio Abreu Araujo, Vincent Cros, Paolo Bortolotti, Juan Trastoy, Kay Yakushiji, Akio Fukushima, Hitoshi Kubota, Shinji Yuasa, Maxence Ernoult, Damir Vodenicarevic, Tifenn Hirtzlin, Nicolas Locatelli, Damien Querlioz, and Julie Grollier. Vowel recognition with four coupled spin-torque nano-oscillators. *Nature*, 563(7730):230–234, October 2018. doi:10.1038/s41586-018-0632-y.

[50] Miguel Romera, Philippe Talatchian, Sumito Tsunegi, Kay Yakushiji, Akio Fukushima, Hitoshi Kubota, Shinji Yuasa, Vincent Cros, Paolo Bortolotti, Maxence Ernoult, Damien Querlioz, and Julie Grollier. Binding events through the mutual synchronization of spintronic nano-neurons. *Nature Communications*, 13(1), February 2022. doi:10.1038/s41467-022-28159-1.

[51] Mayank Banoula. What is perceptron? a beginners guide [updated]: Simplilearn, Apr 2023. URL: https://www.simplilearn.com/tutorials/deep-learning-tutorial/perceptron.

[52] Kanani, Bhavika Panara. Mathematics behind the neural network. Accessed: 2023-05-22. URL: https://studymachinelearning.com/mathematics-behind-the-neural-network/.

[53] Hinkelmann, Knut. Neural networks (lecture slides), 2015. Accessed: 2023-05-22. URL: http://didattica.cs.unicam.it/old/lib/exe/fetch.php?media=didattica:magistrale:kebi:ay_1718:ke-11_neural_networks.pdf.

[54] S. C. Hou and X. X. Yi. Quantum lyapunov control with machine learning, 2018. URL: https://arxiv.org/abs/1808.02516, doi:10.48550/ARXIV.1808.02516.

[55] Javaid Nabi. Machine learning -fundamentals, May 2019. URL: https://towardsdatascience.com/machine-learning-basics-part-1-a36d38c7916.

[56] Nefi Alarcon. Openai presents gpt-3, a 175 billion parameters language model, Mar 2023. URL: https://developer.nvidia.com/blog/openai-presents-gpt-3-a-175-billion-parameters-language-model/.

[57] E.M. Izhikevich. Which model to use for cortical spiking neurons? *IEEE Transactions on Neural Networks*, 15(5):1063–1070, September 2004. doi:10.1109/tnn.2004.832719.

[58] Amirhossein Tavanaei, Masoud Ghodrati, Saeed Reza Kheradpisheh, Timothée Masquelier, and Anthony Maida. Deep learning in spiking neural networks. *Neural Networks*, 111:47–63, March 2019. doi:10.1016/j.neunet.2018.12.002.

[59] H. Markram, W. Gerstner, and P. J. Sjöström. Spike-timing-dependent plasticity: A comprehensive overview. *Frontiers in Synaptic Neuroscience*, 4, 2012. doi:10.3389/fnsyn.2012.00002.

[60] Damir Vodenicarevic, Nicolas Locatelli, Julie Grollier, and Damien Querlioz. Nano-oscillator-based classification with a machine learning-compatible architecture. *Journal of Applied Physics*, 124(15):152117, October 2018. doi:10.1063/1.5042359.

[61] Utkarsh Singh, Neha Garg, Saurabh Kumar, Pranaba Kishor Muduli, and Debanjan Bhowmik. Learning of classification tasks with an array of uniform-mode spin hall nano-oscillators. *AIP Advances*, 11(4):045117, April 2021. doi:10.1063/9.0000192.

[62] Andrei Velichko, Maksim Belyaev, and Petr Boriskov. A model of an oscillatory neural network with multilevel neurons for pattern recognition and computing. *Electronics*, 8(1):75, January 2019. doi:10.3390/electronics8010075.

[63] E Vassilieva, G Pinto, José Acacio de Barros, and P Suppes. Learning pattern recognition through quasi-synchronization of phase oscillators. *IEEE Transactions on Neural Networks*, 22(1):84–95, January 2011. doi:10.1109/tnn.2010.2086476.

[64] Danijela Marković, Matthew W. Daniels, Pankaj Sethi, Andrew D. Kent, Mark D. Stiles, and Julie Grollier. Easy-plane spin hall nano-oscillators as spiking neurons for

neuromorphic computing. *Physical Review B*, 105(1), January 2022. doi:10.1103/physrevb.105.014411.

[65] Davi R. Rodrigues, Rayan Moukhader, Yanxiang Luo, Bin Fang, Adrien Pontlevy, Abbas Hamadeh, Zhongming Zeng, Mario Carpentieri, and Giovanni Finocchio. A spintronic huxley-hodgkin-analogue neuron implemented with a single magnetic tunnel junction, 2023. arXiv:2304.06343.

[66] Abhronil Sengupta and Kaushik Roy. Spin-transfer torque magnetic neuron for low power neuromorphic computing. In *2015 International Joint Conference on Neural Networks (IJCNN)*. IEEE, July 2015. doi:10.1109/ijcnn.2015.7280306.

[67] Daniel D Stancil and Anil Prabhakar. *Spin waves*. Springer, New York, NY, 2009 edition, April 2009.

[68] Therese Frostad, Hans L. Skarsvåg, Alireza Qaiumzadeh, and Arne Brataas. Spin-transfer-assisted parametric pumping of magnons in yttrium iron garnet. *Physical Review B*, 106(2), July 2022. doi:10.1103/physrevb.106.024423.

[69] V. Lauer, D. A. Bozhko, T. Brächer, P. Pirro, V. I. Vasyuchka, A. A. Serga, M. B. Jungfleisch, M. Agrawal, Yu. V. Kobljanskyj, G. A. Melkov, C. Dubs, B. Hillebrands, and A. V. Chumak. Spin-transfer torque based damping control of parametrically excited spin waves in a magnetic insulator. *Applied Physics Letters*, 108(1):012402, January 2016. doi:10.1063/1.4939268.

[70] Tobias Kosub, S. Vélez, Juan Gomez-Perez, Luis Hueso, Jürgen Fassbender, Felix Casanova, and Denys Makarov. Anomalous hall-like transverse magnetoresistance in au thin films on y 3 fe 5 o 12. *Applied Physics Letters*, 113:222409, 11 2018. doi:10.1063/1.5053902.

[71] R. Metselaar and PK Larsen. Electrical properties of yttrium iron garnet. *Journal of Solid State Chemistry - J SOLID STATE CHEM*, 70, 01 1978.

[72] Lepadatu, Dr. Serban. Online materials database. https://www.boris-spintronics.uk/online-materials-database/, 2023. [Online; accessed 22-May-2023].

[73] Mohammad Zahedinejad, Ahmad A. Awad, Shreyas Muralidhar, Roman Khymyn, Himanshu Fulara, Hamid Mazraati, Mykola Dvornik, and Johan Åkerman. Two-dimensional mutually synchronized spin hall nano-oscillator arrays for neuromorphic computing. *Nature Nanotechnology*, 15(1):47–52, December 2019. doi:10.1038/s41565-019-0593-9.

[74] Mohammad Zahedinejad, Himanshu Fulara, Roman Khymyn, Afshin Houshang, Mykola Dvornik, Shunsuke Fukami, Shun Kanai, Hideo Ohno, and Johan Åkerman. Memristive control of mutual spin hall nano-oscillator synchronization for neuromorphic computing. *Nature Materials*, 21(1):81–87, November 2021. doi:10.1038/s41563-021-01153-6.

[75] Akash Kumar, Himanshu Fulara, Roman Khymyn, Mohammad Zahedinejad, Mona Rajabali, Xiaotian Zhao, Nilamani Behera, Afshin Houshang, Ahmad A. Awad, and Johan Åkerman. Robust mutual synchronization in long spin hall nano-oscillator chains, 2023. URL: https://arxiv.org/abs/2301.03859, doi:10.48550/ARXIV.2301.03859.

[76] Eric Arturo Montoya, Amanatullah Khan, Christopher Safranski, Andrew Smith, and Ilya N. Krivorotov. Easy-plane spin hall oscillator, 2023. URL: https://arxiv.org/abs/2301.09228, doi:10.48550/ARXIV.2301.09228.

[77] Amitabha Nandi and Ram Ramaswamy. Synchronization of coupled stochastic oscillators: The effect of topology. *Pramana*, 70(6):1165–1174, June 2008. doi:10.1007/s12043-008-0121-1.

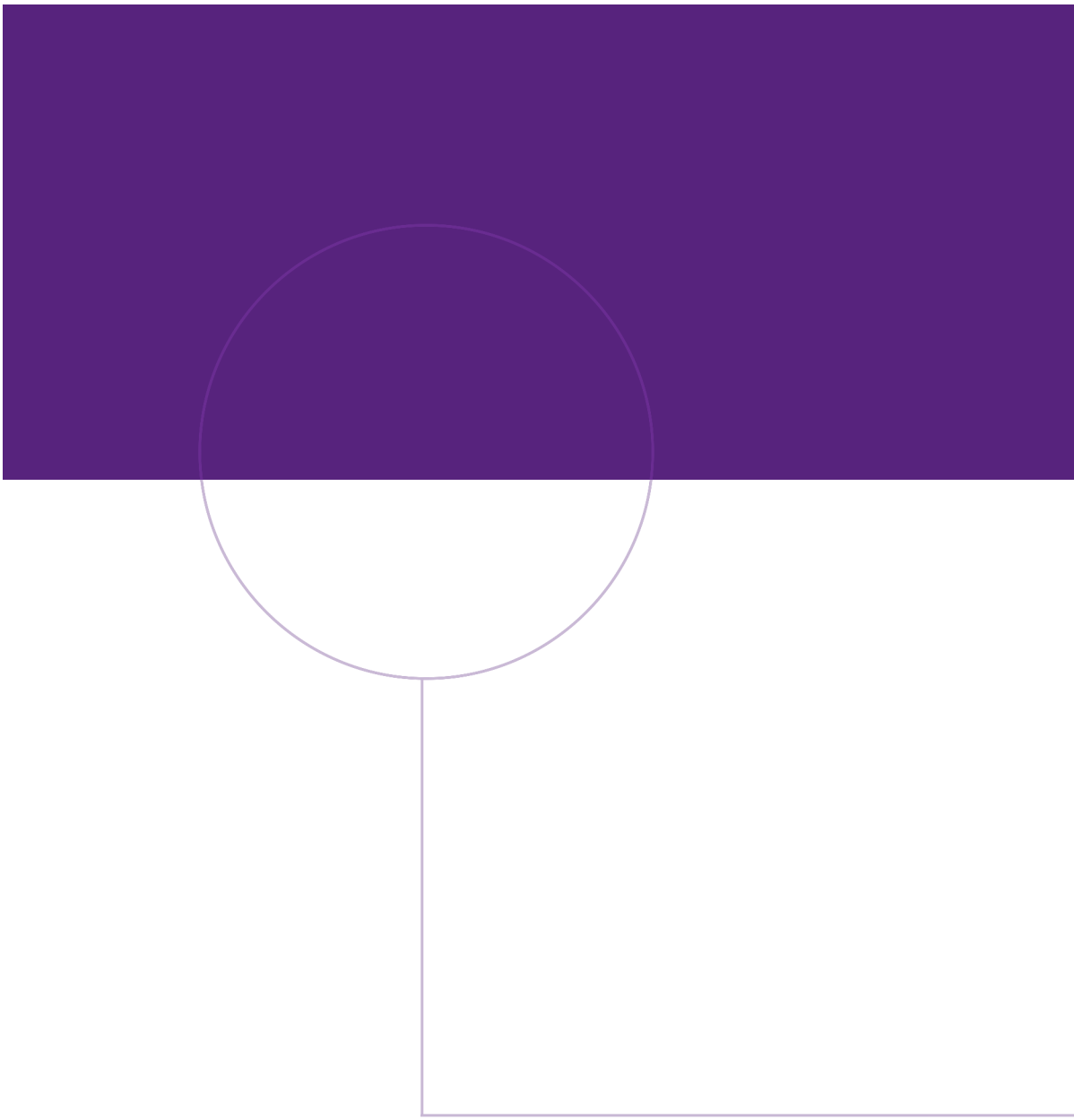
[78] Chuandong Li, Xiaofeng Liao, and Kwok wo Wong. Chaotic lag synchronization of coupled time-delayed systems and its applications in secure communication. *Physica D: Nonlinear Phenomena*, 194(3-4):187–202, July 2004. doi:10.1016/j.physd.2004.02.005.

[79] Vladimir Cherepanov, Igor Kolokolov, and Victor L'vov. The saga of YIG: Spectra, thermodynamics, interaction and relaxation of magnons in a complex magnet. *Physics Reports*, 229(3):81–144, July 1993. doi:10.1016/0370-1573(93)90107-o.

[80] Yuli Yin, Philipp Dürrenfeld, Mykola Dvornik, Martina Ahlberg, Afshin Houshang, Ya Zhai, and Johan Åkerman. Damping's effect on the magnetodynamics of spin hall nano-oscillators, 2018. arXiv:1802.05548.

[81] Damir Vodenicarevic, Nicolas Locatelli, Flavio Abreu Araujo, Julie Grollier, and Damien Querlioz. A nanotechnology-ready computing scheme based on a weakly coupled oscillator network. *Scientific Reports*, 7(1), March 2017. doi:10.1038/srep44772.

[82] Fabrizio Damicelli. Python implementation of the kuramoto model, 2019. URL: https://github.com/fabridamicelli/kuramoto.



NTNU

Norwegian University of
Science and Technology