

Superparamagnetic Tunnel Junctions - True Randomness, Electrical Coupling and Neuromorphic Computing

Leo Alexander Schnitzspan

born in Filderstadt, Germany

Dissertation for Attaining the Academic Degree of
“Doctor rerum naturalium” (Dr. rer. nat.)

of the Departments

08 – *Physics, Mathematics, and Computer Science,*

09 – *Chemistry, Pharmaceuticals, and Geoscience,*

10 – *Biology and University of Medicine*

of the Johannes Gutenberg-University

Mainz, November 2023

Reviewers:

1. Prof. Dr. Gerhard Jakob
Institute of Physics
Johannes Gutenberg-University Mainz
Mainz, Germany

2. ***

Examination date: January, 23th 2024

Declaration of Authorship

I hereby declare that I wrote the dissertation submitted without any unauthorized external assistance and used only sources acknowledged in the work. All textual passages which are appropriated verbatim or paraphrased from published and unpublished texts as well as all information obtained from oral sources are duly indicated and listed in accordance with bibliographical rules. In carrying out this research, I complied with the rules of standard scientific practice as formulated in the statutes of Johannes Gutenberg-University Mainz to insure standard scientific practice.

Signature: _____

Leo Schnitzspan
Mainz, November 2023

Abstract

This research deals with superparamagnetic tunnel junctions (SMTJs), encompassing their fabrication, characterization and potential applications in the context of neuromorphic computing and as a random number generator. Magnetic tunnel junctions (MTJs) based on a magnesium oxide barrier between cobalt-iron-boron alloys exhibit a significant tunnel magnetoresistance (TMR) effect, typically on the order of 100–200%. This characteristic high TMR signal has led to their widespread commercial use in sensing or as magnetoresistive random access memory (MRAM). Here, our developed TMR stack with in-plane magnetization exhibits a TMR ratio of over 200% at a resistance area product of $550 \Omega\mu\text{m}^2$. Within this work at JGU Mainz, the first successful MTJ nanopillar fabrication was accomplished in the group, and numerous optimization steps have been undertaken for the development of superparamagnetic tunnel junctions exhibiting nanosecond fluctuations. In the superparamagnetic regime, the magnetization of the ferromagnetic “free layer” fluctuates solely due to thermal excitation, resulting in a volatile MTJ. This inherent fluctuation, occurring naturally, can serve as an entropy source for random number generators, which makes stochastic MTJs attractive for applications with demanding requirements on random number generation, such as Monte Carlo simulations. In this work, it is demonstrated that a random number generator based on SMTJs exhibits true randomness of nanosecond time scale when combined with logic XOR gates. The quality of true random bit generation is assessed by evaluating all randomness tests of the statistical test suite provided by the National Institute of Standards and Technology (NIST). The fluctuation rate and state probability can be manipulated by external magnetic fields, applied currents or voltages, or by the temperature. Electrons that cross the tunnel barrier transfer a torque to the magnetization of the ferromagnetic free layer due to their spin (referred to as “spin-transfer-torque”) and thus significantly influencing the stochastic behavior of the SMTJ. In addition, the Joule heating affects the fluctuation rate at high

current densities. It is demonstrated that both contributions, the Joule heating and the effect of the “spin-transfer-torque”, can be determined from dwell time analysis. Furthermore, coupling in the switching behavior can arise when two or more stochastic MTJs are electrically connected. In this work, the coupling strength between two SMTJs has been analyzed using the cross-correlation of the voltage fluctuation as a function of the applied source voltage. This approach was both simulated and experimentally verified using time series measurements for two stochastic MTJs. A network of multiple SMTJs can also be used to generate a Gaussian probability distribution, which might potentially be beneficial for noise-based neuromorphic computing approaches. At the end of this thesis, a neuromorphic circuit implementation based on SMTJs, diodes and transistors is presented. This implementation allows for an analog computation of a noise-based local learning algorithm (node perturbation) in a neuromorphic hardware. The proposed approach represents a hardware-based alternative to the established backpropagation algorithm, since the neural network enables an analog and local calculation of the weight adjustment by a learning rule called “node perturbation”.

Zusammenfassung

Diese Forschungsarbeit befasst sich mit superparamagnetischen Tunnelkontakten (SMTJs), ihrer Herstellung, Charakterisierung und möglichen Anwendungen im Kontext des neuromorphen Rechnens und als Zufallszahlengenerator. Magnetische Tunnelkontakte (MTJs), die auf einer Magnesiumoxidbarriere zwischen Kobalt-Eisen-Bor-Legierungen basieren, weisen einen signifikanten Tunnelmagnetowiderstand (TMR) auf, der typischerweise in der Größenordnung von 100–200 % liegt. Dieses charakteristische hohe TMR-Signal führte zu ihrer weit verbreiteten kommerziellen Nutzung in der Sensorik oder als magnetoresistiver Massenspeichern (MRAM: Magnetic Random Access Memory). Der von uns entwickelte TMR-Schichtstapel mit in-plane-Magnetisierung weist ein TMR-Verhältnis von über 200 % bei einem Widerstandsflächenprodukt von $550 \Omega\mu\text{m}^2$ auf. Im Rahmen dieser Arbeit an der JGU Mainz wurde die erste erfolgreiche Herstellung von MTJ-Nanosäulen in der Arbeitsgruppe durchgeführt und zahlreiche Optimierungsschritte für die Entwicklung von superparamagnetischen Tunnelkontakten unternommen, um Fluktuationszeiten im Nanosekundenbereich zu erzielen. Im superparamagnetischen Regime fluktuiert die Magnetisierung der ferromagnetischen “free layer” Schicht allein durch thermische Anregung, wodurch ein volatiler Tunnelkontakt entsteht. Diese inhärente Fluktuation, die natürlich vorkommt, kann als Entropiequelle für Zufallszahlengeneratoren dienen, was sie für Anwendungen mit hohen Anforderungen an die Zufallszahlengenerierung, wie z.B. Monte-Carlo-Simulationen, attraktiv macht. In dieser Arbeit wird gezeigt, dass ein Zufallszahlengenerator, der auf SMTJs basiert, in Kombination mit logischen XOR-Gattern echte Zufälligkeit im Nanosekundenbereich aufweist. Die Qualität der Erzeugung echter Zufallsbits wird durch die Auswertung aller Zufallstests der Testbatterie des National Institute of Standards and Technology (NIST) bewertet. Die Fluktuationsrate und die Zustandswahrscheinlichkeit können durch externe Magnetfelder, angelegte Ströme oder Spannungen oder durch die Temperatur manipuliert werden. Elektronen, die den Tunnelkontakt durchqueren, über-

tragen aufgrund ihres Spins ein Drehmoment auf die Magnetisierung der ferromagnetischen freien Schicht (“Spin-transfer-torque”) und verändert damit maßgeblich das stochastische Verhalten des SMTJ. Zusätzlich beeinflusst die Joulesche Wärme bei hohen Stromstärken die Fluktuationsrate. Es wird gezeigt, dass beide Beiträge, die Joulesche Wärme und der Einfluss des “Spin-transfer-torques”, sich aus der Schaltzeitanalyse bestimmen lassen. Darüber hinaus kann eine Kopplung im Schaltverhalten auftreten, wenn zwei oder mehr stochastische MTJs elektrisch miteinander verbunden werden. In dieser Arbeit wurde die Kopplungsstärke zwischen zwei SMTJs anhand der Kreuzkorrelation der Spannungsschwankung in Abhängigkeit von der angelegten Quellspannung analysiert. Dieser Ansatz wurde sowohl simuliert als auch experimentell anhand von Zeitreihenmessungen für zwei stochastische MTJs verifiziert. Ein Netzwerk aus mehreren SMTJs kann auch zur Generierung einer Gaußschen Wahrscheinlichkeitsverteilung verwendet werden, was potentiell für rauschbasierte neuromorphe Anwendungskonzepte von Nutzen sein könnte. Am Ende dieser Arbeit wird eine neuromorphe Schaltungsimpementierung auf der Basis von SMTJs, Dioden und Transistoren vorgestellt. Diese Implementierung ermöglicht die analoge Berechnung eines rauschbasierten lokalen Lernalgorithmus (Node perturbation) in einer neuromorphen Hardware. Der vorgeschlagene Ansatz stellt eine hardwarebasierte Alternative zum etablierten Backpropagation-Algorithmus dar, da das neuronale Netz mit Hilfe von stochastischen Tunnelbarrieren eine analoge und lokale Berechnung der Gewichtsanzpassung durch einen Lernalgorithmus “Node perturbation” ermöglicht.

Contents

Declaration of Authorship	v
Abstract	vii
Zusammenfassung	ix
List of abbreviations	xv
1 Introduction	1
2 Theoretical background	5
2.1 Magnetism	5
2.1.1 Atomic magnetism	5
2.1.2 Heisenberg exchange interaction	6
2.1.3 Stoner interaction	8
2.1.4 Spontaneous magnetization	9
2.1.5 Micromagnetic approximation	9
2.1.6 Stoner-Wohlfarth model	14
2.2 Interface and interlayer effects in thin films	15
2.2.1 Orange-peel coupling	15
2.2.2 Exchange bias	16
2.2.3 Interlayer exchange coupling	17
2.3 Magnetic tunnel junction	18
2.3.1 Jullière model	19
2.3.2 Tunneling in CoFeB/MgO/CoFeB	20
2.3.3 Tunneling	22
2.3.4 Landau–Lifshitz–Gilbert equation	25
2.3.5 Spin-transfer-torque	27
2.3.6 Joule heating	30

2.4	Superparamagnetic tunnel junction	31
2.4.1	Néel-Brown model	31
2.4.2	Magnetic field dependence	32
2.4.3	Current and voltage dependence	34
2.4.4	SMTJ state probability	36
2.5	Random number generation	36
2.5.1	True random number generation	37
2.5.2	Randomness tests	38
2.5.3	Post-processing of random bits	39
2.5.4	Random telegraph noise	41
2.5.5	Time series measurements	42
2.6	Neuromorphic computing	44
2.6.1	Neural network architecture	45
2.6.2	Activation functions and synaptic weights	46
2.6.3	Learning	47
2.6.4	Node perturbation	48
2.6.5	Local learning in node perturbed neural networks	49
3	Experimental techniques	53
3.1	Characterization	53
3.1.1	Atomic force microscopy	53
3.1.2	Current in-plane tunneling	54
3.1.3	Scanning electron microscopy	56
3.1.4	Superconducting quantum interference device	58
3.1.5	X-ray reflectometry	60
3.2	Fabrication	61
3.2.1	Sputtering	61
3.2.2	TMR stack	63
3.2.3	Annealing	64
3.3	Patterning	65
3.3.1	Bottom electrode patterning	66
3.3.2	Optical lithography	66
3.3.3	Electron beam-lithography	67
3.3.4	Nanopillar resist collapse	71
3.3.5	Ion beam etching	72
3.3.6	Passivation	76
3.3.7	Nanopillar opening	77
3.3.8	Top electrode patterning	78
3.4	Electrical measurements	78

3.4.1	Probe station	79
3.4.2	High frequency measurements	79
3.5	Simulation	81
3.5.1	SPICE	81
4	Results	83
4.1	TMR characterization and optimization	83
4.1.1	TMR stack design	83
4.1.2	Magnetization hysteresis	86
4.1.3	Effect of annealing temperature	88
4.1.4	TMR and RA product	91
4.2	Nanopillar magnetic tunnel junctions	94
4.2.1	MTJ hysteresis	96
4.2.2	MTJ angular field dependence	97
4.3	Superparamagnetic tunnel junctions	99
4.3.1	Superparamagnetic hysteresis	99
4.3.2	Angular field dependence	101
4.3.3	Spin-transfer-torque	104
4.3.4	Nanosecond stochastic switching	105
4.3.5	True random number generation	107
4.3.6	Spin-transfer-torque and Joule heating	110
4.4	Electrical coupling of superparamagnetic tunnel junctions	114
4.4.1	Time series of two coupled SMTJs	114
4.4.2	Simulation	116
4.4.3	Experimental results	122
4.5	Neuromorphic computation with SMTJs	124
4.5.1	Probability distributions sampled by SMTJs	124
4.5.2	Node perturbation circuit	126
4.5.3	SPICE simulation of a noise perturbed perceptron	131
5	Conclusion and Outlook	135
5.1	Conclusion	135
5.2	Outlook	136
	Publications	139
A	Appendix	141
A.1	Passivation layer	141
A.2	Electron-beam lithography	141
A.3	Recipes	141

A.3.1	Nanopillar etch mask	141
A.3.2	Optical lithography	142
	Bibliography	145

List of abbreviations

AFM	Antiferromagnet or Atomic Force Microscopy
AI	Artificial Intelligence
AMR	Anisotropic Magnetoresistance
BP	Backpropagation
CIPT	Current In-Plane Tunneling
CMOS	Complementary Metal Oxide Semiconductor
CMP	Chemical Mechanical Polishing
CoFe	Co ₇₀ Fe ₃₀
CoFeB	Co ₄₀ Fe ₄₀ B ₂₀
COPI	Constrained Parameter Inference
EBL	Electron Beam Lithography
FM	Ferromagnet
FMR	Ferromagnetic Resonance
GND	Ground
GMR	Giant Magnetoresistance
IBE	Ion Beam Etching
JH	Joule Heating
LLG	Landau–Lifshitz–Gilbert
MR	Magnetoresistance
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
MRAM	Magnetic Random Access Memory
MTJ	Magnetic Tunnel Junction
NIST	National Institute of Standards and Technology
NIST STS	NIST Statistical Test Suite
NN	Neural Network
NP	Node perturbation
PRNG	Pseudorandom Number Generator
PtMn	Platinum Manganese (Pt ₃₈ Mn ₆₂)

Py	Permalloy ($\text{Ni}_{80}\text{Fe}_{20}$)
SOC	Spin-Orbit Coupling
SOT	Spin-Orbit Torque
SQUID	Superconducting Quantum Interference Device
STT	Spin-Transfer Torque
ReLU	Rectified Linear Unit
RKKY	Ruderman-Kittel-Kasuya-Yosida
SAF	Synthetic Antiferromagnet
SEM	Scanning Electron Microscopy
SIMS	Secondary Ion Mass Spectrometry
SMTJ	Superparamagnetic Tunnel Junction
TMR	Tunnel Magnetoresistance
TRNG	True Random Number Generator
UV	Ultraviolet
XOR	Exclusive OR Gate
XRR	X-ray Reflectivity

1 | Introduction

Spintronics [1, 2] is an emerging field within the realm of condensed matter physics and material science that aims to exploit the intrinsic spin of electrons as well as their charge in solid-state devices. It offers the potential for developing more efficient and advanced electronic devices with enhanced functionalities, such as increased data storage capacity, improved processing speeds, and reduced power consumption. A significant milestone in spintronics was accomplished by the discovery of the giant magnetoresistance (GMR) effect by Albert Fert [3] and Peter Grünberg [4] in 1988, leading to the award of the Nobel prize for their pioneering work. This effect has paved the way for diverse spintronic applications, such as magnetic field sensors and magnetic recording read heads [5]. While the GMR effect typically remains below 50 % at room temperature, it can surpass 100 % in superlattice structures at lower temperatures [6]. The next milestone was set by the development of devices based on the tunnel magnetoresistance (TMR) effect [7], which is based on spin-dependent tunneling of electrons across a thin insulating layer. Despite already being discovered in 1975 by Jullière [8], it did not attract much attention due to its initially small effect at low temperatures. With the development of MgO-based magnetic tunnel junctions (MTJs) reaching values up to 604 % at room temperature and 1144 % at 5 K [9], this technology is replacing GMR-based devices, such as read-heads in hard disk drives [10], due to its competitively high magnetoresistance effect. In addition, MTJs are exploited for other applications, such as a non-volatile magnetic random access memory (MRAM) [11], field sensors [12], strain sensors [13] and angle sensors [14] or neuromorphic computing and other non-conventional computing concepts [15, 16, 17]. The primary focus of this thesis is an MTJ nanodevice, renowned for its durability, reliability, and compatibility with CMOS technology. Its great compatibility with the highly advanced CMOS technology allows for a potential monolithic chip design of MTJs and CMOS circuits.

As the famous Moore law [18] describes the development of memory and comput-

ing power by the size of the fundamental transistor unit, it is commonly known that this law does not continue forever due to fundamental constraints in the physical nature [19]. While CMOS technology is excellent in precise, deterministic and serial computation, it lacks the ability of randomness, stochasticity and inherent parallel computation. As it requires more and more human and financial resources for further development, an alternative approach is the development of application-specific hardware design, such as in the emerging field of artificial intelligence (AI) [20, 21], in order to improve the computational performance for specific applications. Most prominent is the ongoing development of AI, where computation is carried out more similar to the human brain [22], a concept referred to as neuromorphic computation [17]. Neuromorphic computing [23, 24] is a new paradigm, which aims to model the processing of information like neurons in the brain. A combination with spintronics gives rise to neuromorphic spintronics [25], which allows for energy-efficient biologically-inspired computing. Different ideas for non-conventional computing based on TMR elements have been proposed, such as probabilistic networks [26], Ising and Boltzmann machines [27], spiking neural networks [28], stochastic computing architectures [29], Bayesian neural networks [30] or reservoir computing [31]. A fundamental building block is a noisy and probabilistic MTJ, which is typically employed as an artificial neuron in the network. Due to their inherent randomness, fast switching, robustness and energy efficiency, superparamagnetic tunnel junctions (SMTJs) are ideal for addressing neuromorphic concepts in which randomness and probability are required. In this work, a superparamagnetic tunnel junction is developed and characterized and the stochastic random fluctuations are studied. SMTJ fluctuations are controlled by magnetic fields and currents via spin-transfer-torques, allowing for the modification of the state probabilities. The use of SMTJs for a true random number generator and its scalability is demonstrated. The concept of a network of multiple stochastic MTJs, leads to a coupling effect based on voltage and STT fluctuations, which is studied in this work for a circuit of two SMTJs. Coupling between these MTJs has to be facilitated in order to enable “communication” between fundamental units in the network, which could result in complex network dynamics. This coupling mechanism is crucial for neuromorphic networks, such as Boltzmann machines [32], where each neuron can occupy a binary state with certain probability and where the system dynamics is then described by the network’s total energy. At the end of this thesis, the potential use of volatile magnetic tunnel junctions for a local-learning approach called “node perturbation” in analog neural networks is discussed and offers an alternative to the popular backpropagation algorithm. An implementation concept is proposed, which focuses on the hardware architecture of a circuit based on

SMTJs and CMOS components to implement an analog neural network with local learning abilities. SMTJs introduce the required noise for node perturbation in an energy- and area-efficient way and allow for noise-based learning, which is biologically more plausible and does not require additional circuitry for the execution of the backpropagation algorithm.

2 | Theoretical background

In this chapter, we delve into the fundamentals of magnetism, beginning at the atomic level and progressing to a description of magnetism in condensed matter. We elaborate the different contributions to the magnetic free energy, which govern a system's magnetic state at equilibrium. Next, we examine magnetic tunnel junctions and their charge carrier tunneling mechanisms, with a focus on CoFeB/MgO/CoFeB junctions. We then introduce the classical Landau-Lifshitz-Gilbert (LLG) equation to describe magnetization dynamics. Following this, the characteristics of superparamagnetic tunnel junctions (SMTJs) are highlighted and the prerequisites of true random number generation are presented. Finally, we discuss the possible use of these phenomena in neuromorphic architectures and introduce the basic principles of particular neural networks.

2.1 Magnetism

This section starts with an exploration of the quantum aspects of magnetism and subsequently advances towards describing macroscopic phenomena relevant to the results in this work. We begin by exploring the magnetism exhibited by electrons, followed by an introduction to the micromagnetic energy terms that play a pivotal role in defining the magnetic state of a system. Furthermore, we delve into the magnetization of nanometer-sized particles subjected to an external applied field.

2.1.1 Atomic magnetism

Magnetic properties in condensed matter arise primarily from the characteristics of electrons, under the consideration of negligible effects of the atomic nuclei. In particular, the magnetic moment of an electron has its origin in orbital angular momentum and spin angular momentum. If we consider the electron rotating

2. Theoretical background

around the nuclei, as a conductor loop, we are able to determine the magnetic moment $\boldsymbol{\mu}$ emerging from this electrical current I according to: $\boldsymbol{\mu} = I \cdot \mathbf{A}$, where \mathbf{A} represents the area of the conductor loop. By using the definition of angular momentum: $l = m\omega r^2$, with ω the orbital angular velocity and r the radius, we can rewrite the magnetic moment induced by the orbital angular momentum of the electron to:

$$\boldsymbol{\mu}_l = I\mathbf{A} = -\frac{1}{2}e\omega r^2\hat{\mathbf{n}} = -\frac{e}{2m_e}l \quad (2.1)$$

with $\hat{\mathbf{n}}$ the unit vector of the plane normal, e the modulus of the electron charge and m_e the electron mass. The orbital magnetic moment can also be represented with the gyromagnetic ratio $\gamma = e/2m_e$ or in terms of the Bohr magneton $\mu_B = e\hbar/2m_e$ with \hbar the reduced Planck constant and $g_l = 1$ the g -factor [33]:

$$\boldsymbol{\mu}_l = -g_l\gamma\mathbf{l} = -g_l\mu_B\mathbf{l}/\hbar \quad (2.2)$$

In addition, the electron exhibits an intrinsic angular momentum \mathbf{s} , called spin. The corresponding spin magnetic moment is analogous to the orbital angular magnetic moment given by [33]:

$$\boldsymbol{\mu}_s = -g_s\gamma\mathbf{s} = -g_s\mu_B\mathbf{s}/\hbar, \quad (2.3)$$

where g_s denotes the spin g -factor ($g_s \approx 2$). The total magnetic moment is then given by the coupled orbital and spin moments:

$$\boldsymbol{\mu}_j = \boldsymbol{\mu}_l + \boldsymbol{\mu}_s = -g_j\mu_B\mathbf{j}/\hbar \quad (2.4)$$

Here, \mathbf{j} is the combined angular momentum: $\mathbf{j} = \mathbf{l} + \mathbf{s}$ and the g -factor is given by $g_j = 1 + (j(j+1) + s(s+1) - l(l+1))/(2j(j+1))$. The coupling between spin and orbital moments is a relativistic effect, caused by the motion of the electron around the nucleus, leading to a magnetic field at the site of the electron which interacts with the magnetic moment of the electron. This is called spin-orbit-coupling [34]. The spin-orbit-coupling strength is approximately proportional to Z^4/n^6 [33], where Z refers to the atomic number and n the first quantum number.

2.1.2 Heisenberg exchange interaction

The coupling between individual magnetic moments of localized electrons can be explained by a purely quantum mechanical effect, called exchange interaction. The following derivation of the exchange interaction is based on Ref. [35, 36]. It originates from the Coulomb interaction [37], which describes the interaction

between two fermionic particles, and Pauli's exclusion principle [38], which states that two (or more) identical fermions can not occupy the same quantum state for the same location. In particular, Pauli's exclusion principle states that the whole many-particle wave function is antisymmetric for fermions. If we consider the whole wave function of a system with two localized electrons, it can be described in a non-relativistic approximation as a product of the spatial Ψ and spin wave function χ . This wavefunction is either symmetric (no sign change) or antisymmetric (sign change) under the exchange operation, referring to an exchange of electrons in space and spin coordinates. Two localized electrons can be described by their position \mathbf{r}_i and spin \mathbf{S}_i . To obey Pauli's exclusion principle, there exist only possible wave functions, which are either the product of a symmetric spatial wave function and an antisymmetric spin wave function or an antisymmetric spatial wave function and a symmetric spin wave function.

$$\Psi_s(\mathbf{r}_1) = \frac{1}{\sqrt{2}} [\Psi_1(\mathbf{r}_1)\Psi_2(\mathbf{r}_2) + \Psi_1(\mathbf{r}_2)\Psi_2(\mathbf{r}_1)] \chi_s \quad (2.5)$$

$$\Psi_t(\mathbf{r}_1) = \frac{1}{\sqrt{2}} [\Psi_1(\mathbf{r}_1)\Psi_2(\mathbf{r}_2) - \Psi_1(\mathbf{r}_2)\Psi_2(\mathbf{r}_1)] \chi_t \quad (2.6)$$

with

$$\chi_s = 1/\sqrt{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \quad (2.7)$$

$$\chi_t = |\uparrow\uparrow\rangle; |\downarrow\downarrow\rangle; 1/\sqrt{2}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \quad (2.8)$$

Here, χ_s represents the antisymmetric spin wave function for a singlet state ($S=0$), while χ_t represents the symmetric spin wave function for a triplet state ($S=1$). The energy difference for both possible states yields to the exchange integral J_{ex} , which refers to the electrostatic exchange energy for parallel and antiparallel spin configuration:

$$J_{ex} = E^s - E^t = \int \Psi_1^*(\mathbf{r}_1)\Psi_2^*(\mathbf{r}_2)H(\mathbf{r}_1, \mathbf{r}_2)\Psi_1(\mathbf{r}_2)\Psi_2(\mathbf{r}_1) d\mathbf{r}_1 d\mathbf{r}_2, \quad (2.9)$$

where H represents the Hamiltonian of the Coulomb potential, which describes the electrostatic interaction between electrons and between electrons and their atoms' nuclei. According to Heisenberg the (effective) exchange Hamiltonian is given by:

$$H_{Heis} = -2J_{ij}\mathbf{S}_i\mathbf{S}_j \quad (2.10)$$

The extension to the many-particle Heisenberg Hamiltonian leads to:

$$H_{Heis} = - \sum_{\langle i,j \rangle} J_{ij}\mathbf{S}_i\mathbf{S}_j \quad (2.11)$$

with the sum over all spin pairs $\langle i, j \rangle$. Depending on the wave function, a positive exchange integral J_{ex} results in a parallel alignment (ferromagnetic coupling), while a negative J_{ex} results in an antiferromagnetic alignment (antiferromagnetic coupling).

2.1.3 Stoner interaction

The magnetism of metals is characterized by delocalized electrons in the conduction band and is therefore also called itinerant magnetism or bandmagnetism. The Heisenberg Hamiltonian is not appropriate for these systems since the exchange interaction is facilitated through itinerant electrons. In addition, it is observed that magnetic moment per atom has a fractional/nonintegral value (e.g. $2.2 \mu_B$ for iron [34]), which indicates that the magnetism can not be explained by localized moments of the atoms. The idea, which has been proposed by Stoner [39], is that electrons at the Fermi-level of the conduction band can be redistributed such that an electron from one spin sub-band changes to the other spin sub-band, which is energetically shifted, thereby decreasing the potential energy. As long as this energy reduction is higher than the increase in kinetic energy due to the realignment of the spins, it results in a ferromagnetic state of the system. The prerequisite is that the exchange energy integral has to be positive, the density of states at the Fermi-level large (such as in the 3d bands), and the energy bands partially filled and preferably narrow. The redistributed electrons δN are then given by:

$$\delta N = \frac{1}{4} \mathcal{D}(E_F) \delta E, \quad (2.12)$$

where δE is the energy shift of the sub-bands (“exchange splitting”) and $\mathcal{D}(E_F)$ the density of states at the Fermi-level. A condition for band ferromagnetism, formulated by Stoner [39], states that the product of the Stoner exchange parameter I and the density of states at the Fermi-level $\mathcal{D}(E_F)$ has to be greater than one [34]:

$$I \mathcal{D}(E_F) > 1 \quad (2.13)$$

In non-magnetic materials, the Stoner criterion is not fulfilled, thus no ferromagnetic order is present and the spin sub-bands are equally populated, as shown in Figure 2.1a. Figure 2.1b demonstrates the band ferromagnetism with the energy shift of the spin sub-bands for the typical ferromagnets Ni, Fe and Co. When the density of states is sufficiently high, it becomes energetically favorable for the bands to split, thereby resulting in spontaneous ferromagnetism in the metal. Figure 2.1 is not drawn to scale, since the 3d-bands contain 10 electrons while the 4s-bands take up only two electrons.

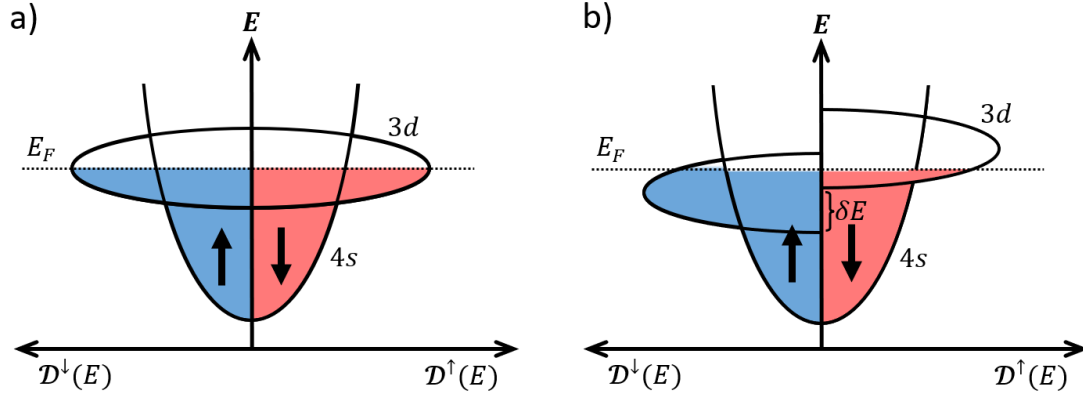


Figure 2.1: a) Band structure of a non-magnetic metal with 4s- and 3d-bands, which are partially filled. Both spin sub-bands (spin-up in blue and spin-down in red) are equally populated. b) Band structure of a ferromagnetic metal. The shift in the spin sub-bands results in a net magnetic moment. The figure is not drawn to scale, as the 3d-bands contain more electrons than the 4s bands.

2.1.4 Spontaneous magnetization

A ferromagnet (also a ferrimagnet) exhibits a magnetization M below the Curie temperature without any external magnetic field. At 0 K the maximum spontaneous magnetization M_s is observed for the ordered spin state, however, M_s decreases for increasing temperatures due to excitations of spin waves, called magnons. This dependence is given for low temperatures and in a first-order approximation by Bloch's $T^{3/2}$ law [40]:

$$M_s(T) = M_s(0)(1 - (T/T_c)^{3/2}), \quad (2.14)$$

where T is the temperature, T_c the Curie temperature and M_s the saturation magnetization. Near the Curie temperature, the temperature dependence of M_s can deviate from Bloch's law and is often described by:

$$M_s(T) \propto (T_c - T)^\beta, \quad (2.15)$$

where β is the material specific critical exponent [40]. For temperatures at T_c a second-order phase transition from the ferromagnetic to the paramagnetic state occurs.

2.1.5 Micromagnetic approximation

For the characterization of macroscopic magnetic systems, it is convenient to replace the atomistic Heisenberg model by a continuum description based on a

2. Theoretical background

micromagnetic approximation [41]. This micromagnetic approximation allows the computation of the spin configuration of any sample geometry. For convenience, a vector field notation of the magnetization \mathbf{M} can be chosen with a normalized magnetization vector $\mathbf{m} = \mathbf{M}/M_s$. In order to determine the magnetic configuration, the total free energy of the system has to be considered, which comprises intrinsic and extrinsic contributions. The fundamental energy terms, relevant for most magnetic systems, like exchange energy, shape anisotropy, crystalline anisotropy, or interface anisotropy are briefly described in the following. The direction of magnetization in a sample is generally not arbitrary, but has a preference in alignment due to magnetic anisotropy of the system. Magnetic anisotropy refers to the directional dependence of a material's magnetic properties, especially its energy [42]. The sample's preferred alignment direction, is known as the easy-axis direction and minimizes the total energy, whereas alignments in other directions which maximize the energy, are called hard-axes directions.

Exchange energy

The exchange anisotropy refers to the Heisenberg exchange energy, which, depending on the exchange integral, either leads to ferro- or antiferromagnetism. In the macroscopic view, the spin order of a sample deviates from this perfect spin order, which is based on the quantum-mechanical effect of exchange coupling. This happens when other anisotropy energies (e.g. magnetostatic energy) are present and prefer another spin alignment of the system. Therefore, the equilibrium spin order is achieved when the total energy is minimized. In the case of a ferromagnet, this can lead to an “energy cost” of exchange energy. In the continuum approximation, we consider a continuous magnetization, thus we can link the cost of exchange energy to the magnetization gradient $\nabla\mathbf{m}$. This yields to the expression of the exchange energy:

$$E_{ex} = A_{ex} \int_V (\nabla\mathbf{m})^2 dV, \quad (2.16)$$

where \mathbf{m} is the normalized magnetization, V the volume and A_{ex} the material specific exchange stiffness constant. The exchange stiffness $A_{ex} = J_{ex}c/a$ is a function of the exchange integral J_{ex} , the nearest neighbor distance a and a lattice dependent constant c , which depends on the crystal structure [43].

Dipolar interaction

Magnetostatics refers to magnetic fields, which are constant in time. Magnetic moments in a ferromagnetic material, such as 3d transition metals Ni, Fe and

Co, exhibit a parallel ordering due to exchange coupling. However, each magnetic moment also couples via stray fields to all other magnetic moments in the ferromagnetic material. This is called dipolar coupling, due to the dipole nature of a single magnetic moment. For two magnetic dipoles $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ at a distance \mathbf{r} , the potential energy can be calculated according to:

$$E_{dipol} = \frac{\mu_0}{4\pi r^3} \left(\boldsymbol{\mu}_1 \boldsymbol{\mu}_2 - \frac{3(\boldsymbol{\mu}_1 \mathbf{r})(\boldsymbol{\mu}_2 \mathbf{r})}{r^2} \right) \quad (2.17)$$

For an ellipsoid with homogeneous demagnetization field H_d the dipolar energy is:

$$E_d = -\frac{1}{2}\mu_0 \int_V H_d M dV \approx -\frac{1}{2}\mu_0 N_d M_s^2 \quad (2.18)$$

with $H_d = N_d M_s$ and N_d the demagnetization factor or demagnetization coefficient. N_d is a tensor and its trace is equal to 1 ($N_x + N_y + N_z = 1$) [36]. The demagnetization factor for an ellipsoid with semi-axes $l_x > l_y > l_z$ is then $N_x < N_y < N_z$, resulting in an easy-axis parallel to the x-axis.

Shape anisotropy

For a given shape of a sample (with a magnetic monodomain), the magnetostatic energy difference between the easy-axis and hard-axis configuration gives rise to anisotropy and is given by the so-called shape anisotropy energy:

$$E_{shape} = -\frac{1}{2}\delta N \mu_0 M_s^2 V = -K_{shape} V, \quad (2.19)$$

where δN is the difference in the demagnetization coefficient for easy- and hard-axis configuration [44].

For a uniformly magnetized infinitely extended thin film sample in the xy-plane (with $N_x = N_y = 0$ and $N_z = 1$), the shape anisotropy energy is simply given by:

$$E_{film} = -\frac{1}{2}\mu_0 M_s^2 V \quad (2.20)$$

For a circular disk with the z-axis along the plane normal, the magnet is isotropic in the plane, resulting in $N_x = N_y$ and therefore the shape anisotropy coefficient is given by $\delta N = N_z - N_x$. In general, the shape anisotropy coefficient of non-trivial structures has to be calculated numerically, however, in the case of a disk, we can

2. Theoretical background

approximate this form by an oblate spheroid with aspect ratio a to provide an analytical solution [44, 45].

$$\delta N_{spheroid} = 1 - \frac{3a^2}{2(1-a^2)} \left(\frac{1}{a\sqrt{1-a^2}} \cos^{-1}(a) - 1 \right) \quad (2.21)$$

Here, $a = t/d$, with the thickness t (semi-axes of spheroid) and the diameter d (semi-axes of spheroid). The dependence of δN as a function of the aspect ratio is illustrated in Figure 2.2. MTJ nanopillar (circular) free layer's aspect ratio is in the order of 0.05, which corresponds to an anisotropy coefficient δN of roughly 0.9 in the approximation of an spheroid.

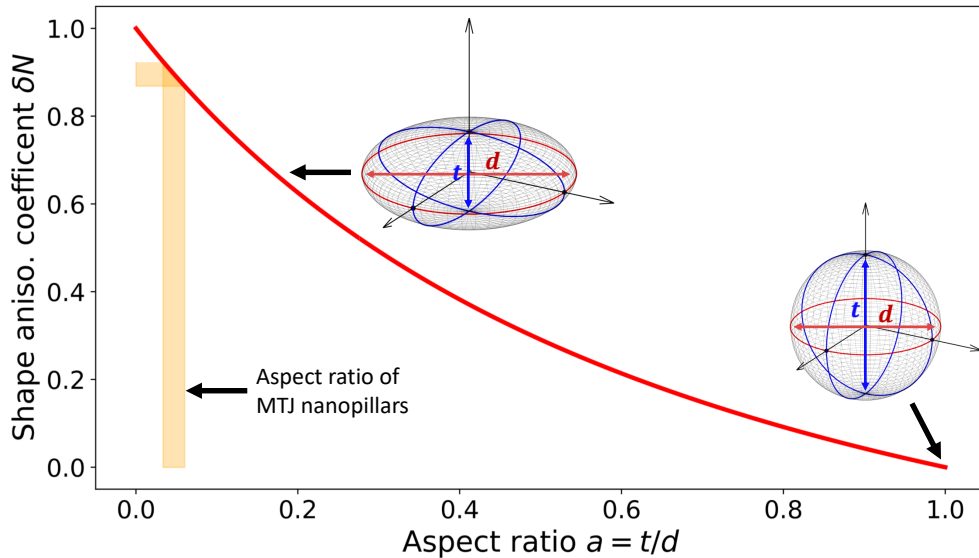


Figure 2.2: The shape anisotropy coefficient as a function of the aspect ratio of a spheroid. The free layer of an MTJ can be approximated by a spheroid. The typical aspect ratio for MTJ nanopillars is depicted in orange.

Interface anisotropy

For magnetic thin films, the interface to adjacent layers becomes significant and can give rise to interfacial anisotropy. For instance, in the bilayer structure of MgO/CoFe, interfacial anisotropy (with the anisotropy direction out of the plane) arises due to the overlap between oxygen p-orbitals and transition metal d-orbitals, resulting in an orbital hybridization and strong spin-orbit coupling [46]. In general, interface anisotropy scales inversely with the thickness t of the ferromagnet [47]. The effective anisotropy is then the combination of the

“bulk anisotropy” and surface/interface anisotropy, since the magnetic moments in the bulk and at the surface are strongly coupled by exchange interaction. The bulk anisotropy can be from crystalline and/or magnetostatic origin. The total anisotropy is given by [47]:

$$K = K_{bulk} + K_s/t \quad (2.22)$$

For MgO/CoFeB interfaces, K_s is typically of the order of 1-2 mJ/m² [48].

Magnetocrystalline anisotropy

In magnetocrystalline anisotropy the magnetic energy depends on magnetization alignment with respect to the crystallographic axes. The main origin of magnetocrystalline anisotropy is the interplay of spin-orbit coupling and the crystal field of the lattice [34, 49]. Stray field coupling energies are also present. The general equation for uniaxial anisotropy is given by [34]:

$$E_{crystal} = K_1^u V \sin^2(\theta) + K_2 V \sin^4(\theta) + \dots \quad (2.23)$$

with K_i being anisotropy constants and θ the angle between the magnetization and the easy axis. For different symmetries, in hexagonal, tetragonal, or cubic lattices, the higher-order terms differ from each other.

Zeeman energy

Zeeman energy refers to the energy of a magnetic moment in an external magnetic field H_{ext} . This energy reaches its minimum when the magnetization is aligned parallel to the magnetic field and reaches its maximum for the antiparallel alignment. The Zeeman energy can be calculated using the following equation:

$$E_Z = -\mu_0 \int_V \mathbf{M} \mathbf{H}_{ext} dV, \quad (2.24)$$

where \mathbf{M} is the magnetization, V the volume, \mathbf{H}_{ext} the external magnetic field and μ_0 the vacuum permeability.

In general, the anisotropy field H_a of any ferromagnet can be determined and is defined as the field necessary to saturate the magnetization of a sample with uniaxial anisotropy in its hard axis direction. The magnetic energy in the system is then given by: $E = K_u V \sin^2(\theta) - \mu_0 H_a M_s V$, where K_u is the anisotropy constant, V the volume and M_s the saturation magnetization. By minimizing

2. Theoretical background

$E \rightarrow \partial E / \partial \theta = 0$ with $\theta = \pi/2$ this leads to the definition of the anisotropy field, which is given by:

$$H_a = 2K_u / \mu_0 M_s \quad (2.25)$$

2.1.6 Stoner-Wohlfarth model

The Stoner-Wohlfarth model, originally proposed in the 1948s [50], is a theoretical framework used to describe the magnetization behavior of single-domain ferromagnets in the presence of an external magnetic field and anisotropy. In this model, the magnetization does not vary within the ferromagnet and is considered as a single macroscopic spin, represented by the vector \mathbf{M} . It is assumed that anisotropy is present in the ferromagnetic material and that thermal effects, like thermal excitation of magnetic moments, can be neglected. In general, the magnetization aligns parallel to the easy-axis. However, if an external magnetic field is present, the magnetization vector will be affected by the field strength and orientation. The equilibrium magnetization direction is reached where the total magnetization energy is minimized. This angle-dependent energy is described by the Stoner-Wohlfarth Equation [50]:

$$E = K_u V \sin^2(\phi) - \mu_0 M_s V H \cos(\theta - \phi), \quad (2.26)$$

where ϕ is the angle between the easy-axis and the magnetization and θ the angle between the easy-axis and the magnetic field direction, M_s the saturation magnetization, V the volume and K_u the uniaxial anisotropy constant. The first term describes the magnetic anisotropy energy and the second term the Zeeman energy. From Equation 2.26, the magnetic hysteresis loop can be determined by calculating the root of the first derivative of the energy with respect to the magnetization direction. The uniaxial anisotropy field H_k , which is the field required to align the magnetization along the hard axis direction, can also be derived from Equation 2.26:

$$H_k = \frac{2K_u}{\mu_0 M_s} \quad (2.27)$$

Next, we consider interface and interlayer effects in multilayer thin films.

2.2 Interface and interlayer effects in thin films

2.2.1 Orange-peel coupling

Magnetostatic coupling between two ferromagnetic layers separated by a thin non-magnetic spacer layer can either be caused by topological roughness or by magnetic roughness [51]. The first effect is called Néel-orange-peel coupling [52, 53]. During multilayer thin film growth, topological roughness typically propagates from layer to layer resulting in correlated interface roughness. This yields to uncompensated magnetic poles at the interface. Since the roughness is typically in-phase with the adjacent ferromagnetic layer, the magnetization of both layers tends to align with each other and thereby gives rise to a coupling effect [54]. The amplitude of this coupling can be of a few mT, depending on the growth, the thicknesses of each layer and the interface roughness [53, 54, 55]. While orange-

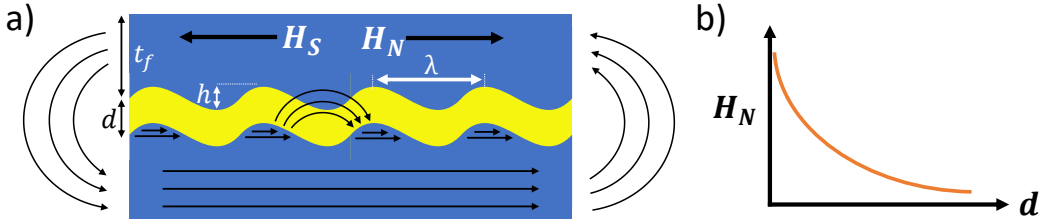


Figure 2.3: a) Illustration of the dominant interlayer coupling mechanisms, magnetostatic stray field coupling (H_S) and Néel orange-peel coupling (H_N). b) The dependence of orange-peel coupling field H_N on barrier thickness d is depicted.

peel coupling has been demonstrated in Al_2O_3 based tunnel junctions [55], in MgO-based MTJs this effect is often not so significant due to a crystalline smooth interface at the tunneling barrier. However, the thickness of the tunneling barrier strongly affects the orange-peel coupling strength, due to its exponential dependence. For thin MgO barriers (approximately < 2 nm) this effect can exceed the stray field coupling originating from the synthetic antiferromagnet (SAF) [56]. The amplitude of the Néel-orange-peel coupling field is given by [55]:

$$H_N = \frac{\pi^2 h^2}{\sqrt{2} \lambda t_f} M_s \exp\left(\frac{-2\pi\sqrt{2}d}{\lambda}\right), \quad (2.28)$$

where h and λ are the amplitude and the wavelength of the roughness profile, d the thickness of the spacer layer, t_f the thickness of the free layer, and M_s is the magnetization of the free layer. While orange-peel coupling prefers a parallel

2. Theoretical background

alignment of both ferromagnetic layers, interlayer stray field coupling, depicted as H_s in Figure 2.3a, originating from the uncompensated magnetic poles at the lateral edges of a sample leads to an antiferromagnetic coupling effect [57]. The degree of compensation of both coupling effects depends on the size, geometry and thicknesses of the thin film sample. In addition, in polycrystalline magnetic layers, local fluctuations of anisotropy at the interface can act as a source for magnetic charge accumulation, which is also called “magnetic roughness” [51], resulting in a coupling effect. It’s worth noting that a significant deviation from Equation 2.28 is present for thin ferromagnetic layers, resulting in the overestimation of Néel-orange-peel coupling fields H_N . This is due to the fact, that external interface roughness effects are not considered. According to Kools et al. [58], these interface roughnesses introduce an additional antiferromagnetic coupling effect, which can be accounted for by incorporating a correction factor dependent on the thicknesses t_f and d in Equation 2.28.

2.2.2 Exchange bias

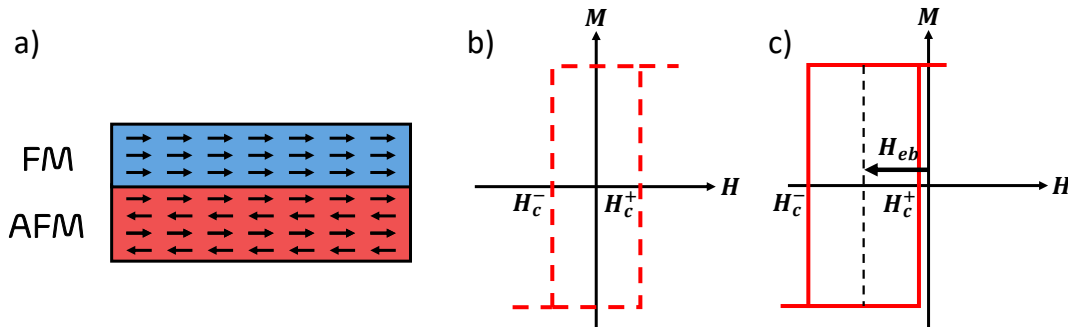


Figure 2.4: a) Illustration of a bilayer of a ferromagnetic thin film in contact with an antiferromagnet. b) Illustration of an hysteresis loop of a ferromagnet. c) Hysteresis loop of a ferromagnet in contact with an antiferromagnet. The induced shift H_{eb} of the hysteresis is called exchange bias.

Exchange bias [59, 60] is a phenomenon in magnetism observed when a ferromagnetic (FM) material is in contact with an antiferromagnetic (AFM) material, as depicted in Figure 2.4a. At the interface between the two materials, the spins are exchange coupled and thereby induce an unidirectional bias in the ferromagnetic layer. This effect is utilized in magnetic data storage devices and sensors to stabilize magnetization and improve their performance and reliability. Exchange bias can be observed experimentally by a shift in the hysteresis loop, as shown in Figure 2.4b-c. The center of the hysteresis loop $H_0 = (H_c^+ + H_c^-)/2$ (H_c : coercive field for pos. and neg. fields) is shifted from its normal position

$H = 0$ to a non-zero field value H_{eb} . This shift implies a bias towards a particular field direction, a phenomenon commonly referred to as exchange bias. In particular, strong exchange bias emerges when the bilayer system is field-cooled from a temperature above the Néel temperature T_N of the antiferromagnet to a temperature below T_N . If during this temperature transition, the ferromagnet is still below the Curie temperature T_C and a magnetic field is present for its alignment, it will induce an exchange coupling effect, whereby the exchange energy at the interface is minimized. This field-cooling procedure establishes unidirectional anisotropy, which is important for sensor or memory device applications, in which a reference magnetization direction is necessary. However, if the bilayer system is heated above the Néel temperature of the AFM, the effect of exchange bias is lost. In general, field-cooling, also known as annealing, can induce uniaxial anisotropy and the emergence of non-uniform properties, such as an increased interface roughness [61], resulting in the broadening of the hysteresis loop. This phenomenon is often observed experimentally and is primarily attributed to grain size formations [61]. In addition, exchange bias is dependent on the thickness of the AFM layer, since the grain size and thereby also the AFM domain structure, is affected by the AFM film thickness [59].

2.2.3 Interlayer exchange coupling

Interlayer exchange coupling is a coupling interaction between magnetic layers in a multilayer structure based on the exchange interaction mediated by electrons in the spacer layer [62]. Typically observed between two ferromagnetic layers, which are separated by a non-magnetic spacer layer, this interaction is notably weaker than the direct exchange coupling present in bulk single crystals [63]. The coupling can be explained by the Ruderman-Kittel-Kasuya-Yosida (RKKY model) [64], which is grounded in spin-dependent spacial oscillations in the density of states [65, 66, 67]. A modulation in spacer thickness leads to an oscillation of interlayer exchange coupling strength, which can either result in an overall ferromagnetic or antiferromagnetic coupling of the magnetic layers. In the following, a brief overview of the RKKY theory is given.

Ruderman-Kittel-Kasuya-Yosida coupling

The RKKY (Ruderman-Kittel-Kasuya-Yosida) [64, 68, 69] coupling is a phenomenon that describes the indirect exchange interaction between two ferromagnetic layers separated by a thin non-magnetic conductive layer. The long-range coupling is mediated by spin-polarized conduction electrons, which interact with

localized moments of the d- or f-shell of each ferromagnet. The coupling strength has an oscillatory nature and declines approximately with $1/t^3$, where t is the distance between the localized magnetic moments. This coupling mechanism is used for synthetic antiferromagnets, where two ferromagnetic layers are separated by a thin non-magnetic metallic spacer layer, like iridium or ruthenium. Due to the RKKY coupling, these ferromagnets can couple parallel or antiparallel depending on the distance t . Thereby, this interaction enables the implementation of a synthetic antiferromagnet, where two ferromagnetic layers are antiferromagnetically coupled. In this work, a synthetic antiferromagnet structure (CoFe/Ru/CoFeB) is employed in order to compensate for stray fields of the TMR stack. In general, the oscillatory nature of the coupling is a function of the Fermi wave vector k_F and the distance t . This oscillatory function is given by [34]:

$$F = (\sin(\theta) - \theta \cos(\theta))/(\theta)^4, \quad (2.29)$$

where $\theta = 2k_F t$. The effective coupling strength is also proportional to the number of conduction electrons per atom n and the Fermi energy E_F , therefore the RKKY coupling strength can be approximated by [34]:

$$J_{RKKY} \propto n^2/E_F \cdot F(2k_F t) \quad (2.30)$$

The interlayer exchange coupling is not only dependent on spacer thickness, as described in the RKKY model, but also on the thickness of the ferromagnetic layers [67, 70]. Since the coupling is governed by interferences of spin-polarized waves within the ferromagnets, the coupling is also expected to oscillate with the ferromagnetic layer thickness, as waves are reflected at each interface of the ferromagnet, all contributing to the overall interference [67]. Coupling is also present in structures with an insulating spacer layer, like MgO in magnetic tunnel junctions, where a strong decline in coupling strength with barrier thickness together with an oscillatory component is observed [71, 72, 73].

2.3 Magnetic tunnel junction

A magnetic tunnel junction (MTJ) refers to a device based on the tunnel magnetoresistance (TMR) effect, which exhibits a relatively large resistance change under a change of magnetization orientation. The essential component of an MTJ comprises two ferromagnetic layers separated by a thin insulator (called tunnel barrier), allowing for the tunneling of electrons from the first ferromagnetic electrode to the second. In general, additional ferromagnetic and non-ferromagnetic

layers are integrated into a multilayer stack, known as the TMR stack, to obtain desired properties of the MTJ device. The tuning and optimization are thoroughly studied in industry to achieve desired MTJ devices for different kinds of applications, including memory devices [11], field sensors [12], strain sensors [13] and angle sensors [14].

In the following, an overview of the TMR stack used for the fabrication of nanopillar MTJs is outlined, including a brief description with the most important characteristics. Afterwards, the TMR effect, the tunneling process and the magnetization dynamics are discussed in detail.

2.3.1 Jullière model

The effect of TMR, which is based on spin-dependent tunneling of electrons through a thin insulating layer, was originally discovered by Jullière on Fe/GeO/Co tunnel junctions at 4.2K [8]. The observed TMR ratio was only about 14%. Jullière interpreted these results in terms of a simple model, which is based on the following assumptions. The first assumption refers to the conservation of the electron spin during the tunneling process, such that spin-up and spin-down electrons can be treated as independent processes in two separate conduction channels [8]. Therefore, the total tunnel current is given by the spin-up and spin-down current: $I = I^\uparrow + I^\downarrow$. For each conducting spin channel, the current has to be proportional to the product of the density of states at the Fermi energy of electrodes one and two [8]. In the case of a parallel alignment of the FM electrodes, the majority spins of the first electrode tunnel to the majority states of the second electrode, while the minority spins tunnel to the minority states. However, in the case of an antiparallel alignment, the majority spins of the first electrode tunnel to the minority states of the second electrode and vice versa. As a second assumption, Jullière considers the conductance or tunnel current proportional to the tunneling density of states of the two FM electrodes. Thus, the tunnel current for the parallel and antiparallel configuration can be described as:

$$I_p \propto I_p^\uparrow + I_p^\downarrow = \mathcal{D}_1^{maj}(E_F)\mathcal{D}_2^{maj}(E_F) + \mathcal{D}_1^{min}(E_F)\mathcal{D}_2^{min}(E_F) \quad (2.31)$$

$$I_{ap} \propto I_{ap}^\uparrow + I_{ap}^\downarrow = \mathcal{D}_1^{maj}(E_F)\mathcal{D}_2^{min}(E_F) + \mathcal{D}_1^{min}(E_F)\mathcal{D}_2^{maj}(E_F), \quad (2.32)$$

where $D_{1,2}^{maj/min}$ is considered the majority or minority density of states of electrode 1 or 2 and $I_p^{\uparrow,\downarrow}$ the respective spin-polarized current channel. The spin polarization

P_i for a ferromagnetic layer i can be defined as:

$$P_i = \frac{\mathcal{D}_i^{maj}(E_F) - \mathcal{D}_i^{min}(E_F)}{\mathcal{D}_i^{maj}(E_F) + \mathcal{D}_i^{min}(E_F)} \quad (2.33)$$

From this, the TMR ratio can be calculated according to [8]:

$$TMR = \frac{G_p - G_{ap}}{G_{ap}} = \frac{I_p - I_{ap}}{I_{ap}} = \frac{2P_1P_2}{1 - P_1P_2} \quad (2.34)$$

Despite a good approximation of the TMR, the model of Jullière has various drawbacks. One of these is the fact that the characteristics of the tunnel barrier are not considered, thus the tunneling current is neither dependent on the barrier height nor on the barrier width. In addition, the band structure is not considered, and orbital hybridizations are neglected. Jullière's model gives reasonable approximations for amorphous AlO_x based tunnel junctions [74], however, in MgO -based tunnel junctions hybridizations and spin-dependent tunneling have to be considered.

2.3.2 Tunneling in $\text{CoFeB}/\text{MgO}/\text{CoFeB}$

Contrary to aluminum oxide-based tunnel junctions, the utilization of magnesium oxide (MgO) as a tunnel barrier significantly enhances the TMR ratio, reaching values up to 604% at room temperature and 1144% at 5 K [9]. The TMR phenomenon observed in MgO -based MTJs can not be explained by the simplified model of Jullière. In contrast to AlO_x -based MTJs, where incoherent tunneling is present, in $\text{Fe}(001)/\text{MgO}(001)/\text{Fe}(001)$ single crystals coherent tunneling is predicted [76]. Experimental evidence supporting this coherent tunneling phenomenon is also discernible through the oscillatory magnetoresistance behavior linked to the MgO layer thickness, as observed in $\text{Fe}(001)/\text{MgO}(001)/\text{Fe}(001)$ MTJs [77]. As shown in Figure 2.5, the MgO barrier acts as a spin filter. Notably, only electrons exhibiting Δ_1 symmetry significantly contribute to the tunnel conductance since the wave function for different bands will decay more rapidly in the barrier. When considering the minority and majority electrons in Co or CoFe electrodes, similarly the majority Δ_1 electrons dominate the tunneling process, whereas the minority electrons lack a state with Δ_1 symmetry, thus do not contribute to the tunneling current [75, 78]. This is the reason for high TMR ratios observed in $\text{CoFe(B)}/\text{MgO}$ -based MTJs compared to tunnel junctions with amorphous AlO_x barriers. However, the strong spin filtering effect is only present for crystalline interfaces and a crystalline MgO barrier. The importance of crystalline interfaces becomes clear when considering the TMR ratio as a function

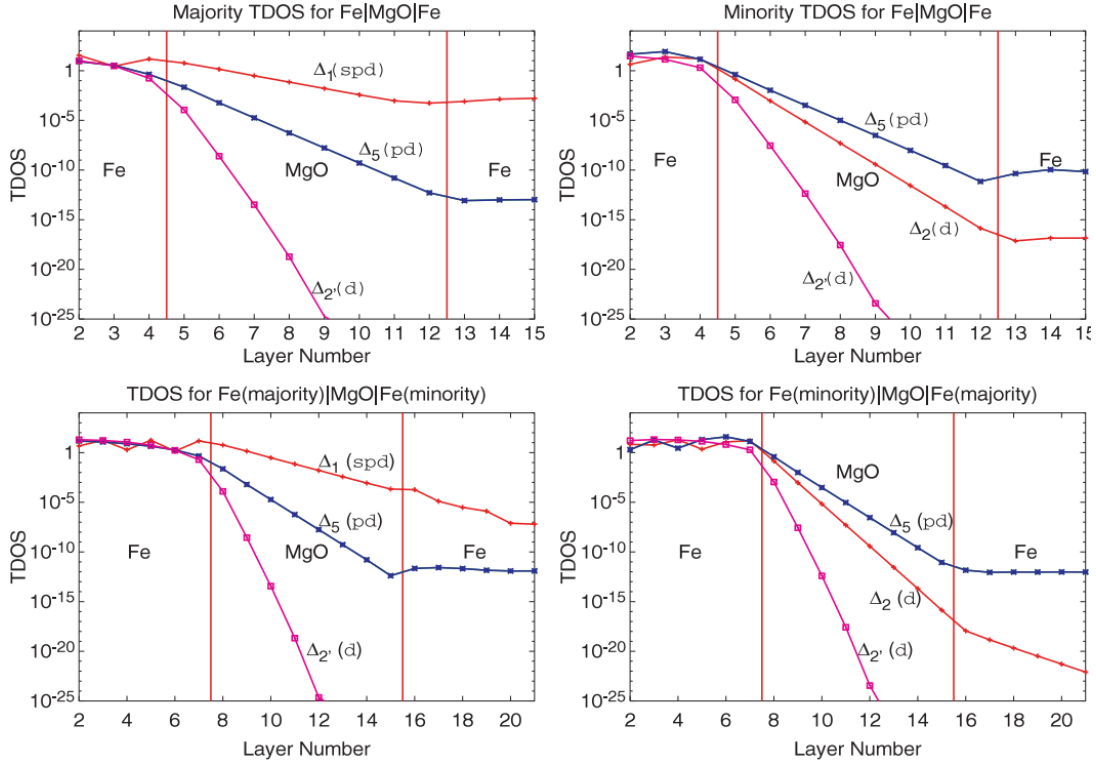


Figure 2.5: Tunneling density of states (TDOS) for majority and minority electrons traveling perpendicular to the plane for a crystalline Fe/MgO/Fe tunnel barrier with 8 monolayers of MgO. The labeling of the TDOS curves corresponds to the symmetry of the incident Bloch state in the left Fe electrode. The Δ_1 band primarily contributes to the tunnel conductance. Figure adapted from Ref. [75].

of MgO barrier thickness. The rapid decrease in TMR ratio for lower RA products with thinner MgO barrier thicknesses is most likely attributed to a change in the MgO grain structure [79], thus deteriorating the MgO barrier interface. The MgO barrier can demonstrate a polycrystalline nature after its deposition on amorphous CoFeB, especially for very thin barriers of a few monolayers [80]. Thus, the barrier exhibits grain boundaries that persist even after the annealing process, leading to an increase in the lateral MgO grain size [79] and thereby improving the TMR ratio. Significant degradation of the TMR ratio appears for RA products below $10 \Omega \mu\text{m}^2$ [79, 81, 82], since it implies a few monolayer thick barrier comprising MgO grain boundaries, which might deteriorate the coherent tunneling [79].

2.3.3 Tunneling

Quantum tunneling is a phenomenon in quantum mechanics where particles, such as electrons, atoms or molecules, can penetrate through energy barriers that would be classically considered impassable. This phenomenon occurs due to wave-like nature of particles. The probability of a particle to tunnel through a potential barrier, exceeding the particle's energy, depends on the potential barrier height and width as well as the momentum of the particle. Starting from the time independent Schrödinger equation $E\Psi(\vec{x}) = H\Psi(\vec{x})$ the Hamilton operator H of a particle in a potential φ is given by:

$$H = -\frac{\hbar^2}{2m_e}\Delta + \varphi(\vec{x}) \quad (2.35)$$

Here, \vec{x} denotes the position of the particle, \hbar the reduced Planck constant and m_e the mass of the electron. For the simplified 1D case and with the energy of the particle being lower than the potential barrier height, we can rewrite the Schrödinger equation to:

$$-\frac{\hbar^2}{2m_e}\frac{d^2\Psi(x)}{dx^2} + \varphi(x)\Psi(x) = E_x\Psi(x) \quad (2.36)$$

The solutions of this equation take different forms for different values of x , considering the wave inside or outside the potential. While the equation can be solved for a simple rectangular potential, it is generally difficult to solve the equation with an asymmetric potential. A solution can be derived under the semiclassical Wentzel-Kramers-Brillouin (WKB) approximation, which approximates non-rectangular potential barriers by multiple thin rectangular barriers. A prerequisite is that the de-Broglie wavelength outside the barrier, which is given by $\lambda = h/(2m_e(E - \varphi(x))^{1/2})$, has to be smaller than the barrier thickness d , such that the barrier is approximately constant with respect to the wavelength. The transmission coefficient T for an electron tunneling through a potential barrier φ , which is defined as the ratio of the transmitted to the incident tunneling current probability, is then given under the WKB approximation by [83]:

$$T = e^{-\frac{2}{\hbar}\int_0^d\sqrt{2m_e(\varphi(x)-E_x)}dx} \quad (2.37)$$

Here, E_x is energy component of the incident electron in the x direction, φ the potential barrier and d the barrier width.

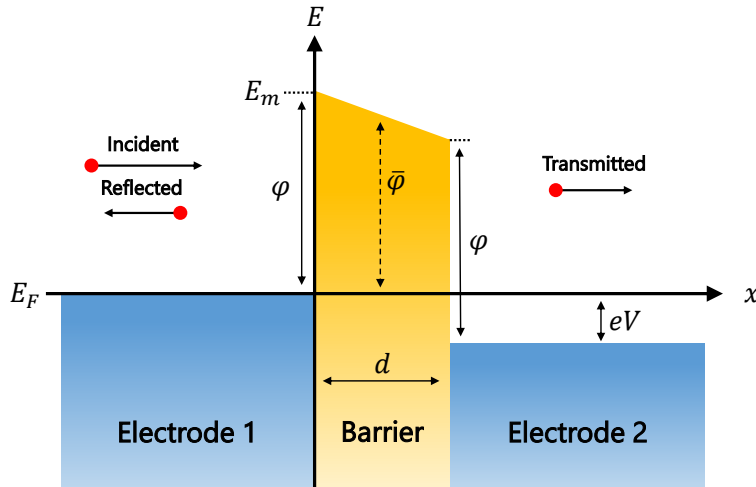


Figure 2.6: Tunneling model after Simmons et al. [83] with a metal/insulator/metal structure. Although the incident electron's energy is below the potential barrier φ (with a mean potential of $\bar{\varphi}$), it is still able to tunnel through a barrier characterized by a thickness (width) of d and a height of φ . The applied voltage to the electrodes leads to an overall energy shift eV of the Fermi energy E_F in electrode 2.

Simmons's model

In magnetic tunnel junctions, the resistance is strongly dependent on the applied voltage amplitude, since the tunneling current is a non-linear function of the applied voltage. The TMR ratio is also affected by the applied voltage bias. Since the understanding of the tunnel current is of importance for any TMR device, we will briefly discuss two models providing the theoretical framework for the observed I-V characteristics.

John Simmons demonstrated that the tunnel current in a magnetic tunnel junction through a symmetrical potential barrier of arbitrary shape and as a function of potential difference (eV) can be calculated by a WKB approximation [83]. An insulating thin film between two electrodes introduces a potential barrier of height φ between these electrodes such that it impedes the charge current flow. Electrons then can either surpass the potential barrier in case their energy is sufficiently high or tunnel through the barrier in case of a thin barrier width, which is equivalent to a thin insulating film. The generalized equation of Simmons, which also includes the effect of field emissions and in which the potential barrier is approximated by a mean potential $\bar{\varphi}$, describes the current voltage relationship

2. Theoretical background

according to [84]:

$$J(V) = (4\pi m_e e / h^3 B^2) [\pi B k_B T / \sin(\pi B k_B T)] \cdot \exp(-A\sqrt{\bar{\varphi}}) [1 - \exp(-BeV)] \quad (2.38)$$

with $B = A/2\sqrt{\bar{\varphi}}$ and $A = (4\pi\beta d/h)\sqrt{2m}$. Here, β describes the shape of the potential barrier, d the barrier thickness, k_B the Boltzmann constant m_e the mass of the electron, e the electron's charge and h Planck's constant. The temperature dependence of the tunnel current is given by [84]:

$$J(V, T) = J(V, 0) \cdot \pi B k_B T / \sin(\pi B k_B T) \quad (2.39)$$

with $B = (2\pi\beta d/h)\sqrt{2m\bar{\varphi}}$. Given the analytical solution of Equation 2.38, it cannot fully explain the observed parabolic voltage dependence or the zero-bias offset of some MTJ I-V measurements [85, 86].

Brinkman's model

Simmons's model of tunneling does not fully explain all effects observed in the tunneling process in a magnetic tunnel junction. Especially, the often observed parabolic current-voltage dependence is not in good agreement of the model. This effect can be explained in another model approach given by Brinkman et al. [87], which also provides an analytical expression for the IV-dependence.

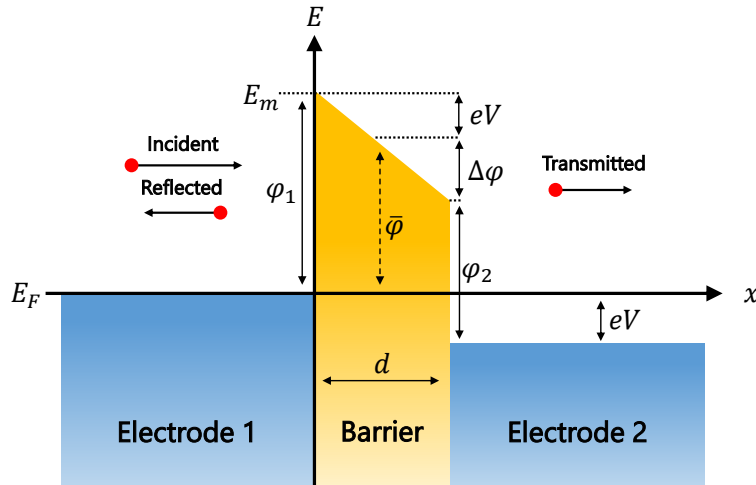


Figure 2.7: Tunneling model after Brinkman et al. [87], with left and right potential barrier heights φ_1 and φ_2 .

It has been observed experimentally that the voltage-dependent tunnel junction conductance appears to be approximately parabolic, with a minimum conductance (or maximum resistivity) occurring at a finite voltage ($V \neq 0$) [88, 89]. This can not be explained by Simmon's model, due to the approximation of an average potential barrier $\bar{\varphi}(V)$ instead of $\varphi(V, x)$. In the model of Brinkman et al., instead of a symmetric potential barrier an asymmetric barrier is considered, which results in an "offset" from $V \neq 0$ of the conductance parabola. This asymmetric trapezoidal barrier is then of the form of: $\varphi(x, V) = \varphi_1 + (x/d)(\varphi_2 - eV - \varphi_1)$, in which φ_1 and φ_2 are the barrier heights of electrode one and two at zero voltage, x the position and d the barrier thickness. Tunneling probability through this barrier is then given by Equation 2.37. However, the solution for the tunnel current for a trapezoidal barrier can only be calculated numerically and is given in the WKB approximation by the following expression [87]:

$$J = \frac{2e}{h} \sum_{k_t} \int_{-\infty}^{\infty} \exp\left(-\frac{2}{\hbar} \int_0^d \sqrt{2m[\varphi(x, V) - E_x]} dx\right) dE_x \quad (2.40)$$

Here, it is the sum over transverse wave vectors k_t of the integrated tunnel probability and not the density of states. Equation 2.40 can be expanded in powers of the voltage in order to obtain an approximation for the conductance in a tunnel junction [87]:

$$G(V) = G(0) \left(1 - \frac{eA_0\Delta\varphi}{16\bar{\varphi}^{3/2}} V + \frac{9e^2A_0^2}{128\bar{\varphi}} V^2\right) \quad (2.41)$$

Here, e is the elementary charge, $\Delta\varphi = \varphi_2 - \varphi_1$, $A_0 = 4\sqrt{2md}/3\hbar$ and $G(0) \approx (3.16 \cdot 10^{10} \sqrt{\bar{\varphi}}/d) \exp(-1.025d\sqrt{\bar{\varphi}})$. If the barrier thickness exceeds 1 nm, then this expression 2.41 is accurate to roughly 10%. Due to the asymmetric potential barrier, the offset voltage, at which the maximum resistivity or minimum conductance is found, is given by $V_{offset} \approx 0.649(\Delta\varphi/d\sqrt{\bar{\varphi}}e)$ [87]. Equation 2.41 explains the asymmetry often observed in IV-measurements of MTJs as well as the zero-bias offset, which is typically in the order of few mV [86].

2.3.4 Landau–Lifshitz–Gilbert equation

The Landau-Lifshitz-Gilbert (LLG) equation is a fundamental equation in magnetization dynamics, which describes the time evolution of a magnetic moment in response to an effective magnetic field \mathbf{H}_{eff} in the system, which is the local field perceived by the magnetic moment. It describes the precessional motion around \mathbf{H}_{eff} in combination with a damping effect, which aligns the magnetization vector \mathbf{m} parallel to \mathbf{H}_{eff} . The effective field results from a combination of

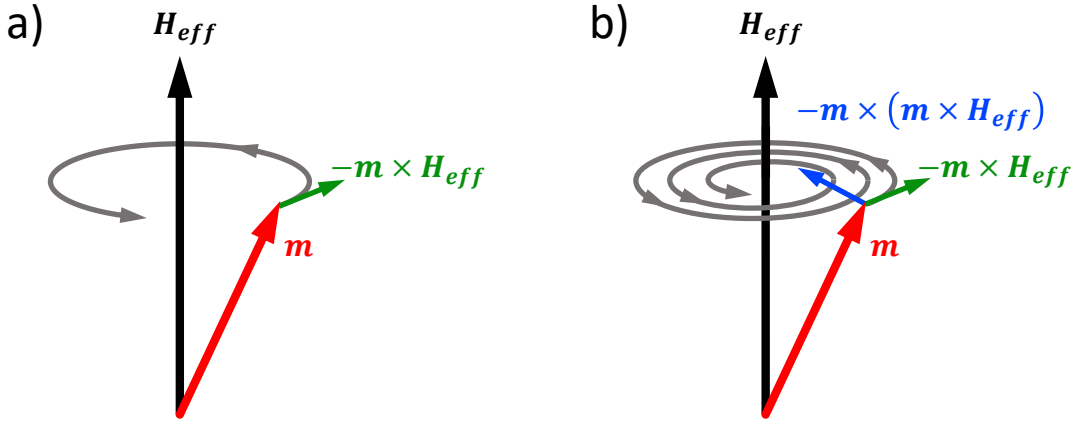


Figure 2.8: a) Illustration of the magnetization dynamics of the normalized magnetization vector \mathbf{m} in an effective field \mathbf{H}_{eff} . a) Magnetization precession without damping and b) with damping. The precession torque is depicted in green, while the damping torque is in blue.

the external magnetic field, dipole field, exchange stiffness, and anisotropy field stemming from the spin-orbit interaction. Mathematically, \mathbf{H}_{eff} can be determined from the gradient of the magnetic energy with respect to the orientation of the magnetization ($\nabla_{\mathbf{m}}$):

$$\mathbf{H}_{eff} = -\frac{1}{\mu_0 M_s V} \nabla_{\mathbf{m}} E = \frac{2A_{ex}}{\mu_0 M_s V} \nabla^2 \mathbf{m} - \frac{1}{\mu_0 M_s V} \nabla_{\mathbf{m}} E_{anis} + H_d + H_{ext}, \quad (2.42)$$

where μ_0 is the vacuum magnetic permeability, M_s the saturation magnetization, V is the volume and E the total magnetic energy of the system. The phenomenological LLG-equation [90], which describes the dynamic motion of the normalized magnetization vector $\mathbf{m} = \mathbf{M}/M_s$ under damping can then be formulated as [91]:

$$\frac{d\mathbf{m}}{dt} = -\gamma_0 \mathbf{m} \times \mathbf{H}_{eff} + \alpha \mathbf{m} \times \frac{d\mathbf{m}}{dt}, \quad (2.43)$$

where $\gamma_0 = g\mu_B/\hbar$ is the gyromagnetic ratio (positive defined), where μ_B is the Bohr magneton and g is the Landé g -factor and α the Gilbert damping constant. The first term on the right-hand side represents the precession of the magnetic moment around the effective magnetic field, while the second term accounts for the damping effect, aligning the magnetic moment with the effective field. This implicit equation can also be transformed into an explicit Landau-Lifshitz (LL) equation [92]:

$$\frac{d\mathbf{m}}{dt} = -\gamma \mathbf{m} \times \mathbf{H}_{eff} - \gamma\alpha \mathbf{m} \times (\mathbf{m} \times \mathbf{H}_{eff}), \quad (2.44)$$

where $\gamma = \gamma_0/(1+\alpha)^2$ is the phenomenological corrected gyromagnetic ratio. The LL equation indicates that both torque terms act perpendicular to each other and are proportional to \mathbf{H}_{eff} . In addition, Equation 2.45 allows for micromagnetic simulations of the time-dependent state of magnetization. Numerical solutions can provide good approximations of the magnetization dynamics, especially for lengths and time scales at which the magnetization stays approximately uniform. The magnetization dynamic in an effective field with the precession and damping torque is illustrated in Figure 2.8.

If temperature fluctuations are considered, a Gaussian stochastic (field) term \mathbf{h}_T can be added with a vanishing mean of $\langle \mathbf{h}_T(t) \rangle = \mathbf{0}$ and a correlation of $\langle \mathbf{h}_T(t)\mathbf{h}_T(t') \rangle = 2D\delta(t-t')$ [93, 94]. D denotes the strength of the thermal fluctuations and $\langle \cdot \rangle$ the average over all fluctuations of the fluctuating field. The LL equation is then given by:

$$\frac{d\mathbf{m}}{dt} = -\gamma \mathbf{m} \times (\mathbf{H}_{eff} + \mathbf{h}_T) - \gamma\alpha \mathbf{m} \times (\mathbf{m} \times (\mathbf{H}_{eff} + \mathbf{h}_T)) \quad (2.45)$$

However, this model of a Gaussian distributed isotropic field contribution due to temperature does not accurately represent a realistic spin wave spectrum of the free layer's magnetization. Limitations to this simplified approach arise from neglecting the isotropic spin-wave dispersion, influenced by the crystalline anisotropy in the free layer [94]. Moreover, it assumes a gapless spin-wave spectrum and does not account for the influence of external magnetic fields [94]. Especially, for a magnetization with a large anisotropy, at a low temperature and under an external applied field, the stochastic LL equation becomes inaccurate and solutions of micromagnetic simulations become questionable [94].

2.3.5 Spin-transfer-torque

A spin current transfers angular momentum, which can be transferred to the magnetization of the ferromagnet, which can lead to magnetization reorientation. This phenomenon is called spin-transfer torque (STT), which was first described by Slonczewski and Berger [95, 96]. In a nanometer sized magnetic tunnel junction, the magnetic free layer can be well described by the Stoner-Wohlfarth model (see Section 2.1.6). If we consider a spin polarized current I_s going through a magnetic layer it will exert a torque onto the free layer's magnetization, as illustrated in Figure 2.9. Here, an initially unpolarized charge current undergoes polarization due to the presence of the pinned reference layer's magnetization, denoted as M_{ref} . This polarized current then exerts a torque onto the free layer's magnetization, owing to the misalignment of magnetization and spin

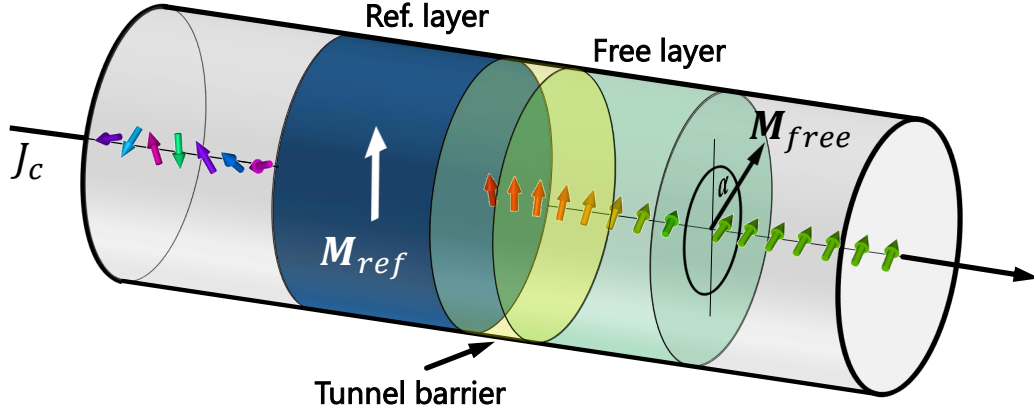


Figure 2.9: Illustration of the spin-transfer-torque (STT) in a cylindric magnetic tunnel junction. The charge current is spin-polarized by the ferromagnetic reference layer and exerts a torque onto the magnetization of the ferromagnetic free layer.

polarization directions. A spin polarized current here refers to an imbalance in the electron flow of up- and down-oriented spins. Under the assumption of no spin-flip scattering and for low temperatures, the spin transport can be regarded as diffusive transport with two separate spin channels. Thus the spin polarization can be expressed as:

$$P = \frac{I_{\uparrow} - I_{\downarrow}}{I_{\uparrow} + I_{\downarrow}} = \frac{\sigma_{\uparrow} - \sigma_{\downarrow}}{\sigma_{\uparrow} + \sigma_{\downarrow}}, \quad (2.46)$$

where $I_{\uparrow,\downarrow}$ is the spin current and $\sigma_{\uparrow,\downarrow}$ the conductance of each channel. Alternatively, the spin polarization can be expressed by the density of states at the Fermi energy: $P = (\mathcal{D}_{\uparrow} - \mathcal{D}_{\downarrow})/(\mathcal{D}_{\uparrow} + \mathcal{D}_{\downarrow})$, where $\mathcal{D}_{\uparrow,\downarrow}$ are the majority- and minority-density of states at the Fermi-level [95]. In the ballistic transport regime, the spin polarization is also dependent on the Fermi velocity $v_{F\uparrow,\downarrow}$ [97, 98]:

$$P_{ballistic} = \frac{\mathcal{D}_{\uparrow}v_{F\uparrow} - \mathcal{D}_{\downarrow}v_{F\downarrow}}{\mathcal{D}_{\uparrow}v_{F\uparrow} + \mathcal{D}_{\downarrow}v_{F\downarrow}} \quad (2.47)$$

In the case of a tunneling spin-polarized current, the spin polarization is not only dependent on the density of states but also on the tunneling matrix element of each spin channel, which is determined by the tunnel barrier:

$$P_{tunnel} = \frac{\mathcal{D}_{\uparrow}|M_{\uparrow}^2| - \mathcal{D}_{\downarrow}|M_{\downarrow}^2|}{\mathcal{D}_{\uparrow}|M_{\uparrow}^2| + \mathcal{D}_{\downarrow}|M_{\downarrow}^2|}, \quad (2.48)$$

where $M_{\uparrow,\downarrow}$ is the tunnel matrix element.

The torque that arises from this spin current can be explained through the principle of angular momentum conservation. A change in the direction of an electron

spin leads to a change of the same amount of angular momentum into the opposite direction and therefore to a change of magnetization of the ferromagnet. Throughout this angular momentum transfer, the total angular momentum is conserved. Therefore, the change in the magnetization of itinerant electrons and local magnetization is conserved as well [99]:

$$\frac{d\hat{M}}{dt} + \frac{d\hat{\mu}}{dt} = 0 \quad (2.49)$$

$$\Gamma_{STT} = \frac{d\hat{M}}{dt} = -\frac{d\hat{\mu}}{dt} \quad (2.50)$$

Here, \hat{M} refers to the unit vector of the local magnetization and $\hat{\mu}$ to the vector of the spin accumulation. In general, Γ_{STT} has two components, a field-like and a damping-like torque. In a metallic spin valve system, the field-like torque is typically relatively small compared to the damping-like torque, as transverse spins dephase rapidly [100]. For this reason, the magnetization switching in nanostructured MTJs is mainly determined by the competing damping and damping-like (STT) torque. Both STT contributions can be written in the form of:

$$\Gamma_{STT} = \Gamma_{dl} + \Gamma_{fl} = \mathbf{a} \cdot \mathbf{m} \times (\mathbf{m} \times \mathbf{p}) - \mathbf{b} \cdot \mathbf{m} \times \mathbf{p}, \quad (2.51)$$

where \mathbf{p} represents the unit vector of the spin momenta of the spin current polarized by the reference layer, \mathbf{m} to the unit vector of the magnetic free layer, and \mathbf{a} , \mathbf{b} refer to STT specific parameters. More explicitly, the STT torque can also be expressed in terms of charge current I and spin polarization unit vector \mathbf{p} of the incoming electrons according to [101]:

$$\Gamma_{dl} = \eta_{dl}(\theta) \frac{\mu_B I}{eV} \mathbf{m} \times (\mathbf{m} \times \mathbf{p}) \quad (2.52)$$

$$\Gamma_{fl} = \eta_{fl}(\theta) \frac{\mu_B I}{eV} \mathbf{m} \times \mathbf{p} \quad (2.53)$$

Here, μ_B is Bohr's magneton, V the volume of the free layer, e the elementary charge and $\eta(\theta)_{dl/fl}$ is a prefactor of the form of $\eta(\theta) = q/(a + b \cos(\theta))$, determining the relative strength and angle dependence of the torque, where θ is the angle between the polarisation of the incoming electrons and the free layer magnetisation [101]. The first term is usually called ‘‘Slonczewski torque’’ or ‘‘damping-like torque’’, while the second term is denoted as ‘‘field-like torque’’ due to its symmetry to the related components of the LL equation. In metallic multilayer systems, the field-like torque is typically in the range of 1-3% of the anti-damping-like torque [101]. To incorporate temperature fluctuations and the effect of STT,

the LLG can be expanded to the stochastic Landau-Lifshitz-Gilbert-Slonczewski (sLLGS) equation:

$$\frac{d\mathbf{m}}{dt} = -\gamma_0 \mathbf{m} \times (\mathbf{H}_{eff} + \mathbf{h}_T) + \alpha \mathbf{m} \times \frac{d\mathbf{m}}{dt} + \Gamma_{STT}, \quad (2.54)$$

where Γ_{STT} (field-like STT neglected) refers to the Slonczewski spin-transfer torque [95, 102], \mathbf{h}_T is thermal induced Gaussian stochastic fluctuation, which accounts for thermal effects, α the damping and γ_0 the gyromagnetic ratio. To describe how a spin-polarized current acts on a non-uniform magnetization, for example like in domain walls, adiabatic and non-adiabatic torques have to be considered [103].

2.3.6 Joule heating

Joule heating refers to the process in which the passage of an electric current through a conductor generates heat. According to Joule's first law, the heating power P equals the product of resistance R and electrical current I squared: $P = I^2 R$. This can be explained by a model, where under an applied voltage electrons are accelerated and inelastically scatter with phonons (quantized lattice oscillations), such that the energy of the electrons is transferred to the lattice yielding to an increase in its temperature. The effect of Joule heating is employed for some applications, like electrical heating or annealing (which is used for the annealing process of our TMR stacks). However, most often it is a by-product in any electrical device (besides superconductors) and has to be considered in the device's performance. Since superparamagnetic tunnel junctions are very sensitive to temperature changes, due to their thermal-activated switching mechanism, Joule heating in superparamagnetic MTJs is studied in detail in this work. Joule heating in a magnetic tunnel junction is dependent on the current direction. Since the tunneling itself is ballistic and lacking any energy dissipation, the hot electrons from one electrode side relax to the Fermi energy in the other electrode on the other side of the junction. Therefore, the energy dissipation depends on the direction of electron flow and heat is primarily generated at one of the ferromagnetic electrodes [104]. In addition, the insulating MgO barrier acts as a thin thermal barrier, leading to a temperature difference of both sides of the barrier [104]. Given that heat conduction in insulators is primarily facilitated by phonons, whereas in metals, it is primarily carried by electrons, an electron-phonon heat transfer at the tunnel junction is necessary [104]. Furthermore, simulations indicate that the temperature increase at the barrier of nanometer-sized MTJs is present after a few nanoseconds [105].

2.4 Superparamagnetic tunnel junction

A superparamagnetic tunnel junction (SMTJ) unites the characteristics of a magnetic tunnel junction and the phenomenon of superparamagnetism. An MTJ comprising a superparamagnetic free layer will exhibit large random resistance fluctuations induced by thermal excitations. SMTJs are of interest as building blocks for various potential applications such as true random number generation [106], stochastic computing [107], neuromorphic computing [108] or probabilistic computing [26]. In the following, the switching dynamics of SMTJs are discussed.

2.4.1 Néel-Brown model

The Néel-Brown model [93, 109], also called the Néel-Arrhenius model, describes the time-dependent dynamics of a magnetic nanoparticle under thermal agitation. A ferromagnetic single-domain nanoparticle can be described by a single magnetic moment/spin ("macrospin approximation") and shows magnetization fluctuations induced by thermal excitation. If the nanoparticle's anisotropy energy is comparable to the thermal energy, the magnetization undergoes random rotations within a timescale shorter than the measurement time ($\tau_m \approx 1$ s). This phenomenon is known as superparamagnetism. Despite being ferromagnetic, a superparamagnetic particle exhibits a field dependence similar to a paramagnet with a lack of hysteresis and zero coercivity, in the case for $\tau \ll \tau_m$.

In the presence of uniaxial anisotropy, a magnetic nanoparticle possesses two meta-stable states, corresponding to the lowest energy states within the system. However, due to finite temperature, the magnetization randomly switches between these two states. The average time between two consecutive switching events is referred to as the dwell time τ and its dependence on the anisotropy and temperature is described by the Néel-Arrhenius law [109]:

$$\tau = \tau_0 \exp\left(\frac{E_b}{k_B T}\right) = \tau_0 \exp\left(\frac{KV}{k_B T}\right), \quad (2.55)$$

where τ_0 is the attempt time of the order of 1 ns and depends on material characteristics and the particle's geometry [91, 109]. T is the ambient temperature, k_B the Boltzmann constant, E_b the energy barrier between both states, K the magnetic anisotropy and V the volume. Due to the exponential dependence on the energy barrier, dwell times can span a vast range, from picoseconds to millions of years or even longer. When the dwell time is significantly shorter than the measurement time, the measured magnetization tends to average out to zero.

Conversely, if the dwell time greatly exceeds the measurement time, the magnetization becomes “blocked”, thus no switching occurs during the measurement time. This case is then referred to as ferromagnetic rather than superparamagnetic. It also leads to the concept of the blocking temperature, which is defined as the temperature at which the measurement time τ_m equals the average dwell time τ .

2.4.2 Magnetic field dependence

Now we will consider a simplified model of the ferromagnetic free layer without coupling effects caused by adjacent layers within the TMR stack. Due to the typical thin film geometry, a large anisotropy ($H_k^{eff} \approx 1.5$ T) arises, causing the magnetization of the free layer to lie in the plane, neglecting interface anisotropy. In addition, we take into account an in-plane uniaxial anisotropy, induced by shape anisotropy, crystalline anisotropy or strain-induced anisotropy. When considering only the presence of uniaxial anisotropy and Zeeman energy, and assuming that the magnetization is primarily in-plane, the total magnetic energy can be described by the Stoner-Wohlfarth model as:

$$E = 1/2\mu_0 H_k^{ip} V M_s \sin^2(\phi) - \mu_0 H M_s V \cos(\theta - \phi), \quad (2.56)$$

where H_k^{ip} is the in-plane anisotropy field, H the applied magnetic field, M_s the saturation magnetization, V the volume of the free layer, ϕ the angle between the easy-axis and the magnetization and θ the angle between the easy-axis and the magnetic field direction.

If the applied field is parallel to the easy-axis direction, mainly the P- and AP-energy levels are tuned, as illustrated in Figure 2.10a, where the energy landscape is plotted as a function of the magnetization rotation angle ϕ . The local minima at 0° and 180° correspond to the P- and AP-state respectively. The energy barrier E_b between those states is given by the difference of energy maximum (E_{max}) and minimum (E_{min}). In general, the (modulated) energy barrier for one of the states is defined as: $E_b^{p,ap} = E_{max} - E_{min}^{p,ap}$, while the unmodulated energy barrier is defined as $E_b = E_b(H = 0)$, as indicated in Figure 2.10a. A small applied field will increase the energy of one of the states while it decreases the other. This modulation of the double-well potential affects the energy barrier and therefore also the dwell times according to [110]:

$$\tau_{p,ap} = \tau_0 e^{\frac{E_b}{k_B T} \left(1 \pm \frac{H}{H_k}\right)^n}, \quad (2.57)$$

where H is the external applied in-plane magnetic field, H_k the switching field at zero temperature and n a constant which is either 1.5 or 2 [110]. The exponent of

$n = 2$ can be derived from the Stoner-Wohlfarth model for a single domain with uniaxial anisotropy and a magnetic field applied along the easy-axis direction. However, if considering higher-order anisotropy terms [111] or a misalignment of the magnetic field, the exponent can vary and possibly tend towards 1.5 [112]. There has been a theoretical debate about the correct value of the exponent n , however, experiments indicate an exponent of 2 [112]. Therefore, Equation 2.57 changes to:

$$\tau_{p,ap} = \tau_0 e^{\Delta(1 \pm H/H_k)^2} \quad (2.58)$$

Here, Δ is the energy barrier in multiples of $k_B T$ between the two stable states. The energy barriers for both states $\Delta_{p,ap} = \Delta (1 \pm H/H_k)^2$ are plotted as a function of the applied easy-axis field, as shown in Figure 2.10b. The effective energy barrier, defined as $\Delta_{eff} = (\Delta_p + \Delta_{ap})/2$, is plotted in Figure 2.10c, which indicates a parabolic dependence. The dependence for an easy-axis field ($\theta = 0$) is given by:

$$\Delta_{eff} = \Delta \left(1 + \left(\frac{H_{\parallel}}{H_k} \right)^2 \right) \quad (2.59)$$

where Δ is the energy barrier without any applied field and H_{\parallel} the magnetic field applied parallel to the easy-axis.

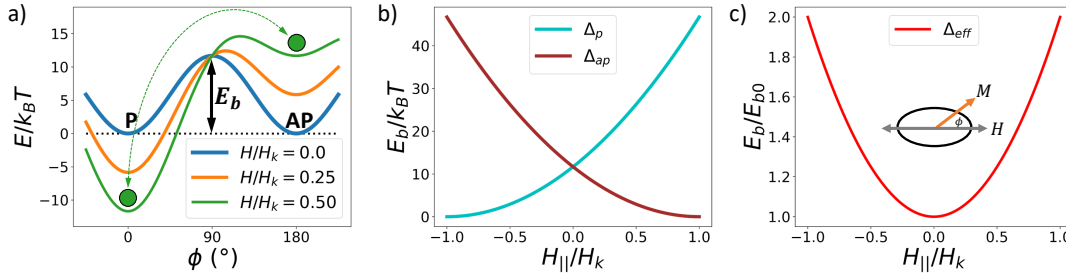


Figure 2.10: Illustration of the easy-axis field dependence for a superparamagnetic tunnel junction. a) The magnetic energy depending on the magnetization orientation in relation to the easy-axis direction is plotted for different relative fields H_{\parallel}/H_k , applied parallel to the easy-axis. b) Plotted is the energy barrier of the P-state (Δ_p) and AP-state (Δ_{ap}) as a function of the relative field. c) The effective energy barrier $\Delta_{eff} = (\Delta_p + \Delta_{ap})/2$ is shown as a function of the applied field. The inset demonstrates an ellipsoidal free layer (easy-axis direction along the long axis) with magnetization M and magnetic field direction H .

In case of a perpendicular applied hard-axis field ($\theta = \pi/2$), both states have the same energy level while the energy barrier is symmetrically tuned. Hence, applied fields in the order of H_k can significantly lower the energy barrier, which is illustrated in Figure 2.11a. While the position of the energy barrier is fixed

2. Theoretical background

at 90° , the positions for the energy minima (P- and AP-state) shift towards the barrier position for positive applied fields, as depicted in Figure 2.11b. When considering the effective energy barrier, the double well potential turns into a single well at an in-plane hard-axis field of H_k . The dependence of the effective barrier on the hard-axis field, as shown in Figure 2.11c, follows:

$$\Delta_{eff} = \Delta \left(1 - \frac{|H_\perp|}{H_k}\right)^2 \quad (2.60)$$

Here, Δ is the energy barrier without any applied field and H_\perp the magnetic field in hard-axis direction.

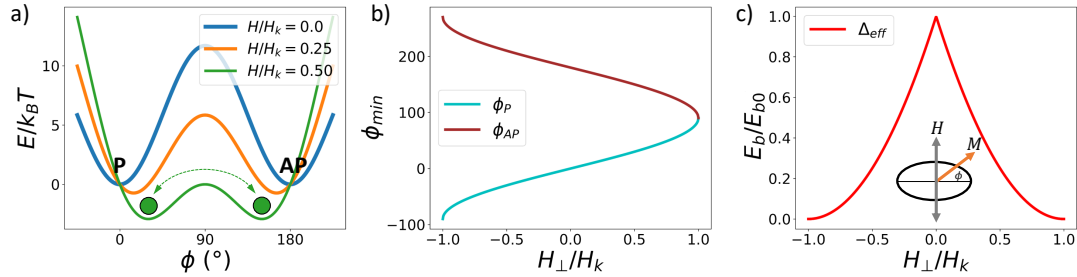


Figure 2.11: Hard-axis field ($\theta = \pi/2$) dependence for a superparamagnetic tunnel junction a) The magnetic energy depending on the magnetization orientation in relation to the easy-axis direction is plotted for different relative fields H_\perp/H_k . b) Angle ϕ for the local minima of the P- and AP-state. c) Effective energy barrier $\Delta_{eff} = (\Delta_p + \Delta_{ap})/2$ depending on the absolute of the relative field.

Following this, the average fluctuation rate Γ can be tuned via a reduction in effective energy barrier by:

$$\Gamma^{-1} = \tau = \sqrt{\tau_p \tau_{ap}} = \tau_0 e^{\Delta_{eff}} = \tau_0 e^{\Delta(1-|H_\perp|/H_k)^2} \quad (2.61)$$

In general, the outlined in-plane field dependencies allow the tuning of the superparamagnetic switching to a balanced 50/50-state by an easy-axis field component and the tuning of the overall switching rate by a hard-axis field component. Nonetheless, a shift of the energy minima is associated with a variation in the angle ϕ and this coincides with a reduced (stochastic) TMR signal, as the TMR resistance is a cosine function of the angle ϕ .

2.4.3 Current and voltage dependence

Under an applied current across the tunnel junction, dwell times can be manipulated through a spin-polarized current, which exerts a torque onto the magnetic

free layer, as discussed in Section 2.3.5. In general, a spin-polarized current exerts a field-like and a damping-like torque onto the magnetic free layer. The amplitude of each of the both torques depend on the current density, diffusive properties, structure, and geometry of the MTJ. The damping-like torque originates from the exchange interaction between the spin-polarized electrons and the local moments within the ferromagnet, aiming to align the electrons spins along the magnetization [113]. The field-like torque is based on spatial displacement of conduction electrons and local moments, momentum transfer and spin-flip scattering [113]. The field-like torque leads to a precessional motion around the effective field, while the damping-like torque results in an alignment of the magnetization and the electron spin. Typically, the damping-like torque is much stronger than the field-like torque [100] and is the dominating contribution in nanometer-sized MTJs. Therefore, the damping-like torque can stabilize one of the states, while destabilizing the other state, depending on the applied current direction. The effect on the energy barrier and the dwell times can be described by the following equation [110, 114, 115]:

$$\tau_{p,ap} = \tau_0 e^{\Delta(1 \pm I/I_c^{p/ap})n_I} \quad (2.62)$$

Here, Δ is the energy barrier, I the applied current and $I_c^{p/ap}$ the critical switching current at 0K for the P- to the AP- state and vice versa. The exponent n_I is typically assumed to be 1, however, it is still under discussion whether it is 1 [110, 116, 117] or 2 [112, 118]. An exponent of 1 has been theoretically derived by considering a fictitious temperature, while an exponent of 2 is obtained through the examination of the stochastic process, utilizing the Fokker-Planck equation as the basis [112]. In high-TMR MTJs the critical switching current for the magnetization switching of the free layer is typically lower for the switching of the AP- to the P-state than for the P- to the AP-state ($|I_c^{ap}| < |I_c^p|$) [119]. Consequently, a particular current polarity will stabilize one orientation while destabilizing another, which can be used in SMTJs in order to stabilize either the P- or AP-state. Similarly, the dwell times can be tuned by an applied voltage and the effect on the energy barrier is then determined by the ratio of applied voltage to the “critical” voltage V_c . Since the critical current defines the critical voltage V_c via the MTJ resistance, we can write the effective energy barrier as a function of V_c . Experimentally it has been observed that the critical switching voltage for the P-state and AP-state are similar [112, 120, 121]: $V_c^p \approx V_c^{ap}$, thus we consider $V_c = V_c^p = V_c^{ap}$ and modify the Néel-Arrhenius law to [116, 120]:

$$\tau_{p,ap} = \tau_0 e^{\Delta(1 \pm V/V_c)} \quad (2.63)$$

Here, τ_0 is the attempt time, Δ the energy barrier, V is the voltage across the junctions, and V_c is the critical voltage for deterministic switching at 0 K [110, 114]. Interestingly, V_c is independent of the junction size, since the MTJ resistance is proportional to $1/d^2$ and the critical current is proportional to d^2 .

2.4.4 SMTJ state probability

The state probability (P_{ap}) is defined as the probability to observe an AP-state in an SMTJ (at any given time) and corresponds to the time-averaged state probability. It can be determined from the dwell time ratio: $\tau_{ap}/(\tau_{ap} + \tau_p)$. With Equation 2.63 we can derive the characteristic voltage-dependent AP-state probability in the approximation of zero switching time between states:

$$\begin{aligned} P_{ap}(V) &= \frac{\tau_{ap}}{\tau_{ap} + \tau_p} = \frac{e^{\Delta(1-V/V_c)}}{e^{\Delta(1-V/V_c)} + e^{\Delta(1+V/V_c)}} \\ &= \frac{1}{e^{2\Delta V/V_c} + 1} \end{aligned} \quad (2.64)$$

Here, Δ is the energy barrier, V the applied voltage and V_c the critical voltage. Equation 4.12 is of the form of a sigmoidal function and can be used to model the probability of a stochastic MTJ for different applied voltages. Likewise, a sigmoidal relationship is found for a function of applied easy-axis magnetic fields:

$$P_{ap}(H) = \frac{1}{e^{-4\Delta H/H_k} + 1} \quad (2.65)$$

with Δ the energy barrier, H the applied magnetic field and H_k the anisotropy field. The sign of the exponent in Equation 2.65 depends on the definition of the field direction with respect to the ferromagnetic layers. Here, it denotes the stabilizing effect for the AP-state for positive field values.

2.5 Random number generation

This section elaborates the differences between pseudorandom and true random number generation, while also highlighting techniques for improving the randomness quality. We will emphasize the methods used for the testing the quality of randomness and discuss acquisition, autocorrelation, and cross-correlation for the analysis of SMTJ signals.

2.5.1 True random number generation

Random number generators (RNGs) find application in various technological devices, particularly for cryptographic encryption purposes. In the encryption process, the first crucial step often involves the creation of a sizable random number, which provides the cryptographic key for the encryption algorithm. Since most commonly employed encryption algorithms, renowned for their high security standards, are typically very robust to attacks other than brute force, a great degree of insecurity can arise through the use of pseudorandom numbers. Thus, an unpredictable true random number generator (TRNG) is the requirement for the generation of a secret (random) key, which prevents an attacker of being able to compute the key in any way.

While TRNGs are used in cryptographic applications, pseudorandom number generators (PRNGs) are not employed since they are based on mathematical deterministic functions. Although designed to generate sequences of numbers that exhibit characteristics similar to true randomness, they are deterministic in nature and exhibit periodicity in very long sequences. Pseudorandom numbers seem to be true random for short periods, but will fail to pass statistical randomness tests for very long periods. If the starting value, called seed, of a PRNG algorithm is known, every number can be easily computed. It is essential to note that while PRNGs offer practical convenience and computational efficiency, they are not suited for highly secure cryptographic applications.

In contrast, TRNGs are devices or processes that generate sequences of numbers with characteristics of “true randomness”, like unpredictability and uniformity. Unlike PRNGs, TRNGs are not deterministic and derive randomness from inherently unpredictable physical phenomena, which are often referred to as entropy sources. The foundation of true randomness lies in the principles of quantum mechanics, which exhibit inherent unpredictability. TRNGs leverage these phenomena to produce random sequences that cannot be precisely determined or reproduced, even with complete knowledge of the system’s initial conditions.

There are different possible entropy sources for TRNGs, such as quantum, chaotic, noise or oscillator systems. Classical TRNGs based on electrical noise utilize a high-gain differential amplifier for the purpose of amplifying noise to a certain level [122] and comparing it against a threshold within a comparator to generate a digital random signal. However, this process consumes a lot of energy and requires a large circuitry. Quantum TRNGs are based on quantum phenomena. Any TRNG system (e.g. flipping a coin) may be called “quantum”, however, this name typically refers to only those generators which utilize a single intrinsically random quantum effect, like a superposition of photons or superconducting

qubits [123, 124]. Unlike quantum TRNGs where each bit generation shares the same initial state and measurement, noise-based TRNGs differ in their approach due to the requirement of a steady change in their state.

If truly random, the output sequence of a TRNG will pass stringent statistical tests for randomness (to a certain probability). While TRNGs offer the advantage of producing truly random sequences, they may be less efficient, more expensive, and more energy consuming than PRNGs. This is the reason why TRNG based on superparamagnetic MTJs are suited for energy efficient true random number generation.

2.5.2 Randomness tests

In order to access the quality of randomness¹, various statistical tests can be used to analyse the probability of being true random. Several statistical test suites (STS), like “Diehard”, “TestU01” or “NIST STS” [122], are available for TRNG analysis. In this work, the National Institute of Standards and Technology Statistical Test Suite (NIST STST) [125] was employed for the evaluation of random sequences, generated by SMTJs. As depicted in Table 4.1, the “test battery” encompasses 15 different statistical tests and multiple sub-tests, which examine the statistics of bits or of blocks of bits. To derive meaningful statistical results for all tests, the total bit length n of the bitstream under examination should exceed 10^6 [126]. A true random number bitstream analyzed by this test battery is expected to pass all (or most of) these tests. Nevertheless, even truly random data has a probability of 80 % of not passing at least one of these tests [126].

Each statistical test can be regarded as a hypothesis test of a pattern in a bitstream, where the null hypothesis H_0 refers to the sequence being tested as truly random. Contrary, the alternative hypothesis H_1 refers to the sequence as not truly random. The significance level is typically chosen as $\alpha = 0.01$, which means that we expect to reject the null hypothesis in less than 1 % of the cases for a true random number. Each test compares the occurrence of a certain pattern in the data with the expected theoretical test statistic for a truly random (infinitely long) sequence. In order to evaluate a test, the respective probability distribution of the test statistic has to be known, which typically is a χ^2 or a normal distribution. From the known distribution and test statistic, the p -value can be calculated. If the p -value is smaller (greater) than α , the hypothesis is rejected (accepted). However, a wrong rejection (acceptation) of a hypothesis is also possible, though unlikely, and termed type I or type II error. A type I error occurs when the null hypothesis H_0 is rejected, although it is true, while a type

¹The quality of randomness refers to the likelihood of being of true random origin.

II error occurs when H_1 is rejected although it is true. The probability of a type I error is equal to the significance level α .

	Test name	n	m or M	# sub-tests
1.	Frequency	$n \geq 100$	-	1
2.	Block frequency	$n \geq 100$	$20 \leq M \leq n/100$	1
3.	Runs	$n \geq 100$	-	1
4.	Longest run	$n \geq 128$		1
5.	Binary matrix rank	$n > 38912$	-	1
6.	Spectral	$n \geq 1000$	-	1
7.	Non-overlapping T. M.	$n \geq 8m - 8$	$2 \leq m \leq 21$	148
8.	Overlapping T.M.	$n \geq 10^6$		1
9.	Maurer's Universal	$n > 387840$		1
10.	Linear complexity	$n \geq 10^6$	$500 \leq M \leq 5000$	1
11.	Serial		$2 < m < \lfloor \log_2 n \rfloor - 2$	2
12.	Approximate entropy		$m < \lfloor \log_2 n \rfloor - 5$	1
13.	Cumulative sum	$n \geq 100$		2
14.	Random excursions	$n \geq 10^6$		8
15.	Random excursions variant	$n \geq 10^6$		18

Table 2.1: Overview of the NIST Statistical Test Suite [125, 126] with the recommended bitstream size n for each particular test and the indication of the number of sub-tests. Some tests can be parameterized by a parameter m or M . In this work, the recommended values [125] for m and M have been used.

In order to summarize the testing results of multiple sub-tests, the particular p -values can be combined into a single p -value according to Fisher's method [127]. Fisher's method is frequently employed in meta-analysis to combine the p -values of separate hypothesis tests, each one being independent and based on the same hypothesis. This combination leads to a new null hypothesis, suggesting that all initial null hypotheses hold true. In the case of rejecting H_0 , it signifies the rejection of at least one null hypothesis. Given that the p -value of a hypothesis test adheres to a uniform distribution within the range $[0,1]$, it can be shown that $-2 \sum_i \log(p_i)$ for i from 1 to n follows to a chi-squared distribution χ_{2n}^2 with $2n$ degrees of freedom [127]. In this work, the p -values derived from various sub-tests are combined and assessed using Fisher's method, which applies to tests 7, 11, 13, 14, and 15 (see Tab. 4.1).

2.5.3 Post-processing of random bits

The output raw bits of an entropy source are often biased and not uniformly distributed for various reasons. Therefore, post-processing of the raw output can

2. Theoretical background

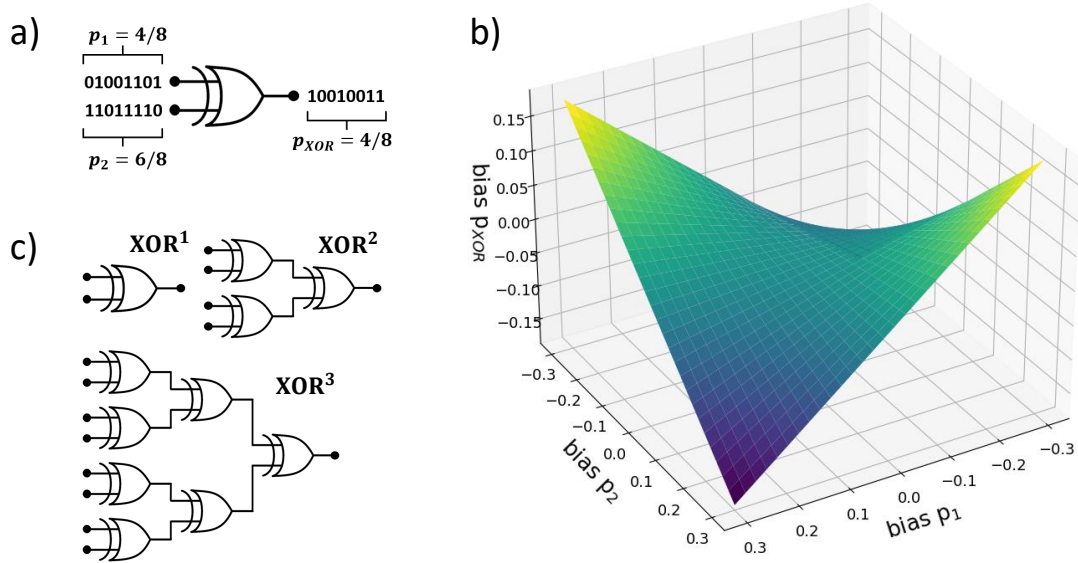


Figure 2.12: a) Demonstration of the stochastic computation of an XOR gate. Two binary bitstreams with probabilities (of “1”) $p_1 = 4/8$ and $p_2 = 6/8$ result in a more unbiased and “balanced” bitstream of $p_{XOR} = 4/8$. b) The de-biasing function for a single XOR operation on two biased stochastic bitstreams is illustrated. c) Different degrees of XOR operations. XOR^1 , XOR^2 and XOR^3 for one, two or three rounds of logical XOR operation.

be necessary, despite a true random entropy source. There are different methods of post-processing (also called “whitening”), which, in general, refers to the idea of sacrificing a few bits in order to arrive at a more unbiased and higher quality of randomness, thus passing more statistical randomness tests. For instance, the von Neumann whitening [128] is a well-known method to improve the quality of randomness of a biased sequence. By dividing a sequence into chunks of two bits, these bits are compared and if they are equal, no output is generated and if they are distinct (01 or 10) a corresponding output is either 1 or 0. This technique will balance the resulting bitstream to 1/2 and improve the quality of randomness. Other methods are based on sophisticated algorithms, such as cryptographic hash functions to de-bias a random sequence [129]. These often require a CMOS overhead and can be slow in performance. Thus, a hardware minimalistic and energy-efficient approach is the use of logic XOR gates, called XOR-whitening. A single XOR performs a logic exclusive OR operation on two input bitstreams. Given two bitstreams and their respective probabilities (to observe a “1”) p_1 and p_2 , the XOR output probability for the outgoing bitstream is given by [130]: $p_{XOR} = p_a + p_b - 2p_a p_b$, as depicted in Figure 2.12a. This equation is illustrated in Figure 2.12b, which displays the de-biasing effect of a single XOR operation,

where two biased bitstreams with bias b_1 and b_2 transform into a less biased bitstream (vertical axis). Here, the bias is defined as the difference of probabilities of 1's and 0's: $b = p(0) - p(1)$. Figure 2.12c depicts different degrees of XOR operations. The interconnection of multiple XOR gates leads to a greater debiasing effect and a potential improved quality of randomness. In an XOR^2 operation, two sets of XOR operations are applied to four input bitstreams, while in an XOR^3 operation, three sets of XOR operations are employed, working on eight input bit streams (see Fig. 2.12a).

2.5.4 Random telegraph noise

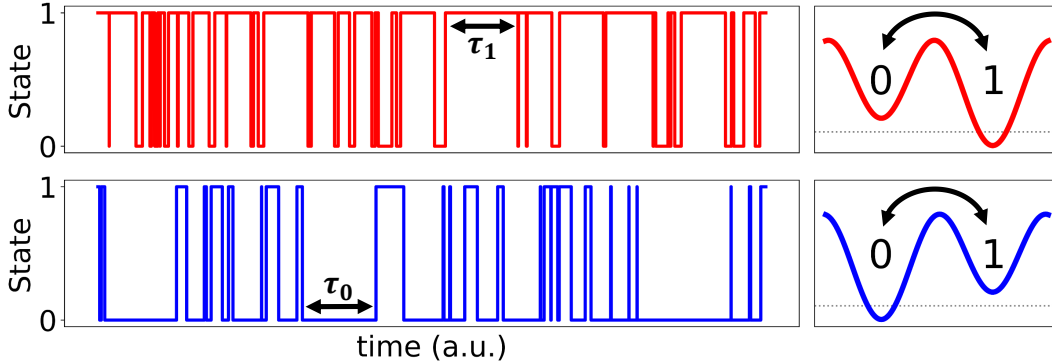


Figure 2.13: Illustration of random telegraph noise. The dwell times τ_0 and τ_1 for state 0 and state 1 are Poisson distributed. The signal in red demonstrates random switching with a preference for the 1-state while the signal in blue is slightly biased to the 0-state. This kind of telegraph noise can result from a thermally excited quasiparticle in a 1D potential well, hopping randomly between both local minima, as illustrated on the right.

Telegraph noise features a random stochastic process in which spontaneous transitions between two or more levels occur. It is observed in various physical systems, like flash memory [131] or field effect transistors [132, 133], and is characterized by an exponential distribution of dwell times τ . Random telegraph noise, which can be considered as a two-state Markov process, exhibits a Poisson distribution of dwell times. Hence, the Poisson process can describe the probability of a stochastic MTJ switching to another state [91]. Moreover, at low current densities in stochastic MTJs, the switching event follows a probability density function (PDF) given by [134]: $f^{sw}(t) = 1/\tau \cdot e^{-t/\tau}$ for a mean dwell time τ , which is the inverse of the average switching rate, and waiting time t . Therefore, the thermal induced switching probability that the free layer's magnetization switches within time t in an MTJ can be expressed as (Néel-Brown relaxation time formula [110]):

$$F^{sw}(t) = 1 - e^{-t/\tau} \approx t/\tau \quad (t \ll \tau) \quad (2.66)$$

2. Theoretical background

which is the cumulative distribution function (CDF) of $f^{sw}(t)$ [134]. As a consequence, the switching probabilities from the P to the AP state and vice versa can be expressed as:

$$P_{p \rightarrow ap}^{sw} = 1 - e^{-t/\tau_p} \quad (2.67)$$

$$P_{ap \rightarrow p}^{sw} = 1 - e^{-t/\tau_{ap}}. \quad (2.68)$$

2.5.5 Time series measurements

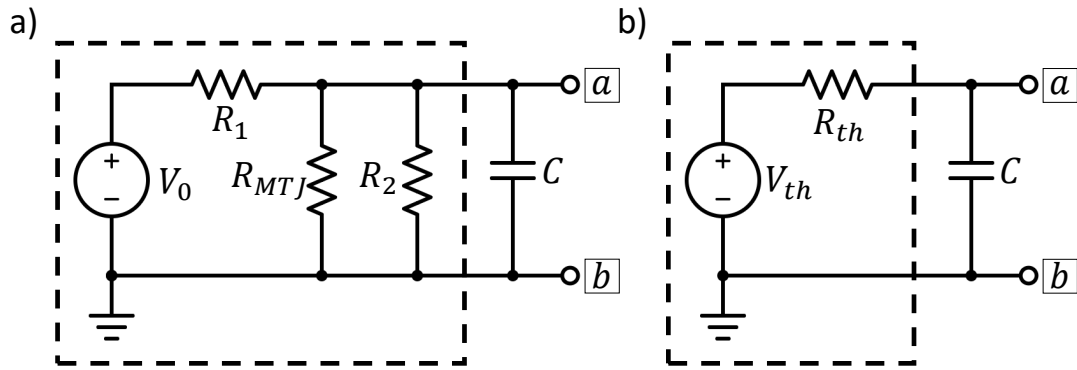


Figure 2.14: a) The electrical circuit of the measurement setup for the measurement of stochastic fluctuations of SMTJs is shown. A resistor denoted as R_1 is connected in series with the MTJ and a resistor R_2 is connected in parallel. The capacitor C accounts for the circuit's parasitic capacitance. b) The equivalent Thevenin circuit describes a simple RC-circuit with a time resolution of $\tau_{RC} = R_{th}C$, where R_{th} represents the Thevenin equivalent resistance.

The correct measurement of stochastic fluctuations, particularly in the case of nanosecond switching, relies heavily on the time resolution of the measurement setup and the oscilloscope. In electrical circuits, the presence of a small parasitic capacitance (typically below 1 nF) can impede the propagation of voltage signals. A sudden voltage rise (drop) will charge (discharge) this capacitance, which can be described by the characteristic RC time constant: $\tau_{RC} = RC$, which defines the time it takes to charge (discharge) the capacitor to approximately 63% ($\approx 1 - 1/e$) through the resistor. If a voltage signal is shorter than the RC time, then it cannot be acquired accurately, since the signal changes before the capacitor is fully charged or discharged, leading to a smoothed measured signal. To enhance time resolution and ensure accurate signal acquisition, it is desirable to keep the RC value low. Figure 2.14a illustrates the electrical circuit used for the measurement of stochastic switching. A DC voltage V_0 is applied to the voltage divider consisting of R_1 and $R_2 || R_{MTJ}$. The parasitic capacitance is represented

by the capacitor C connected in parallel to ground. Figure 2.14b depicts the equivalent Thevenin circuit diagram [135], where the resistances are replaced by R_{th} and the voltage by V_{th} . The equivalent Thevenin voltage V_{th} can be calculated according to:

$$V_{th} = V_0 \cdot \frac{R_{MTJ} || R_2}{R_{MTJ} || R_2 + R_1} = V_0 \cdot \frac{R_{MTJ} R_2}{R_{MTJ} R_2 + R_1 (R_{MTJ} + R_2)}, \quad (2.69)$$

The equivalent Thevenin resistance is then obtained by short-circuiting all voltage sources in the circuit, resulting in:

$$R_{th} = R_1 || R_2 || R_{MTJ} = \left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_{MTJ}} \right)^{-1} \quad (2.70)$$

The value of R_{th} is primarily dominated by the lowest resistance of R_1 , R_2 and R_{MTJ} . Thus, it can be lowered either by R_1 or R_2 (assuming R_{MTJ} can not be changed). By lowering the resistance R_{th} , the time constant ($\tau_{RC} = R_{th}C$) is decreased and the signal can be acquired more accurately at the outputs (a) and (b), as shown in Figure 2.14.

Considering the frequency domain, the cutoff frequency f_c , at which the signal has a -3 dB gain and which determines the bandwidth of the system, is given by: $f_c = (2\pi\tau_{RC})^{-1}$. Figure 2.15 demonstrates the impact of different resistance values, via the RC time constant, on the measured signal. In Figure 2.15a nanosecond telegraph signal with infinite rime resolution (bandwidth) is simulated. The observed measured signals (Gaussian noise added) for various RC time constants with their corresponding cutoff frequencies are highlighted in Figure 2.15b-d. The signal in Figure 2.15d does not indicate sharp transitions anymore, since the transition time is shorter than the RC time. In addition, when the switching rates approach or exceed f_c , the average amplitude decreases, resulting in smoothed transitions. It also has to be emphasized that by lowering the resistance to achieve a higher time resolution, the measured voltage drop across the voltage divider decreases as well, which might lead to the necessity of signal amplification. As demonstrated in Figure 2.15, observed transition rates in a two-level system with finite time resolution are always smaller than the true transition rates of the system [136]. The finite bandwidth of the measurement system limits the accurate observation of the stochastic fluctuation, thereby leading to an effective underestimation of transition rates [136].

2. Theoretical background

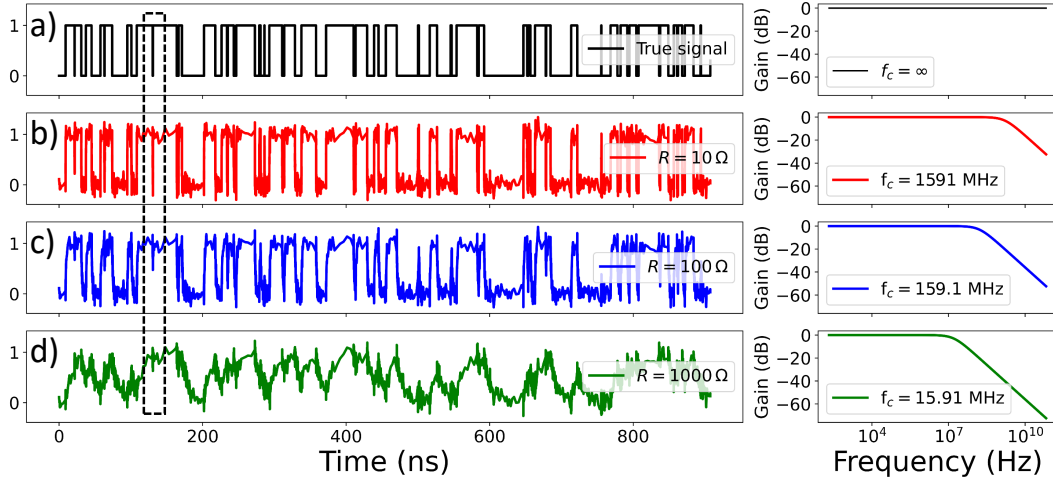


Figure 2.15: a) A (true) two-level random telegraph signal. b) - d) Observed signals by the measurement setup under finite measurement bandwidth, for different resistances in the RC circuit, and their respective Bode plots with the indicated cut-off frequency. Simulations are performed with *LTspice* for $C = 10$ pF and R as indicated in the plot. Gaussian noise is added to the time series in b) - d).

2.6 Neuromorphic computing

Neuromorphic computing [17, 23, 137], is a multidisciplinary research field that combines principles from neuroscience, computer science, physics and engineering to develop computing systems, which are inspired by the structure and function of the neural networks of human brains. It aims to create hardware and software architectures that mimic parallelism, plasticity², and energy efficiency of the brain's information processing. The biologically-inspired neural network architectures are in contrast to the well established von Neumann architectures, primarily due to parallel and analog instead of digital and sequential data processing. Thus, they can typically handle large amounts of data efficiently in terms of computational cost and energy. While classical computation is strictly deterministic, neuromorphic computation is able to incorporate stochasticity, randomness and noise similar to the human brain [22]. In the following, the basic principles of an artificial neural network are described.

²The ability of a neural network to adapt, reorganize and to learn from experience is called (neuro-)plasticity.

2.6.1 Neural network architecture

The typical structure of an artificial feed-forward neural network (NN) with two hidden layers is depicted in Figure 2.16a. Key components of an artificial neural network are neurons (also called “nodes”) and synaptic weights [20]. Neurons serve as the core processing units within a NN. As indicated in Figure 2.16b, each neuron computes the weighted sum of its inputs, applies an activation function to this sum and propagates the result to all following neurons in the next layer [20]. Often an adjustable bias value is added to the summed weight of each neuron in order to provide more degrees of freedom to the NN. A connection between neurons (also called “edge”) is associated with an adjustable synaptic weight (a real number) determining the strength of the connection. Each time a signal propagates (forward or backward) from one neuron to the next, the signal is multiplied by the corresponding weight value w_{ij} . These weight parameters are fine-tuned by a learning algorithm with the objective of achieving accurate predictions for a given input [20]. The learning algorithm operates based on a loss function, which quantifies the prediction error and is aimed to be minimized during the learning process through appropriate adjustments of the synaptic weights.

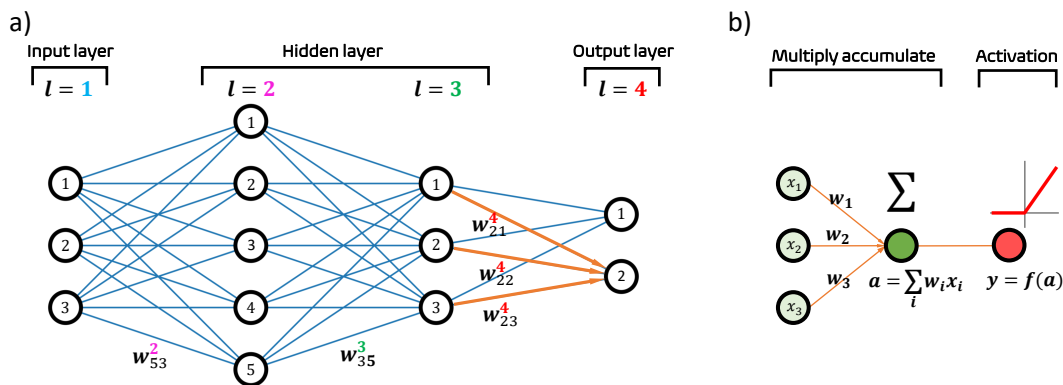


Figure 2.16: a) Illustration of an artificial neural network ($3 \times 5 \times 3 \times 2$) with two hidden layers. A connection between two neurons is associated with a weight w_{ij}^l , which is applied to the signal before propagating to the next neuron i of layer l . b) An illustration of a perceptron demonstrates the two-step computation involving three inputs. Initially, a multiply-and-accumulate (MAC) operation is executed on all inputs x_i to obtain the pre-activation a . The non-linear activation function then transforms the weighted-sum into an output y .

2.6.2 Activation functions and synaptic weights

Activation functions [138] are essential components in neural networks, serving as mathematical operations applied to the output of each neuron or node [138]. Their primary purpose is to introduce non-linearity into the network, enabling it to learn complex relationships and patterns in the data. Figure 2.17 illustrates a few of the most common and popular activation functions used in neural networks [138]. For instance, the sigmoid activation function maps input values to a range between 0 and 1, forming an S-shaped curve. This function found utility primarily in the early days of neural network development and proved valuable for addressing binary classification problems. Nonetheless, it encountered limitations due to the vanishing gradient problem [139]. In contrast, a simple and widely adopted activation function, the rectified linear unit (ReLU) function, exhibits gradients that are either 1 or 0. It helps to mitigate vanishing gradients and accelerates training, but it can suffer from the “dying ReLU” problem [140], which refers to the gradient being 0 and staying at 0, preventing any further weight updates. To address the “dying ReLU” issue, the leaky ReLU [141] allows for a small gradient for negative inputs, which helps to maintain some gradient flow for all neurons during training. An alternative activation is the exponential linear unit (ELU), which can be viewed as a smoothed version of the ReLU [142]. Here, mean activations are closer to zero, which speeds up learning in deep neural networks and can lead to higher accuracy in classification tasks [142]. Typically, the choice of the activation function depends on the specific problem, architecture, and training considerations. Modern neural networks often use variants of ReLU (like Leaky ReLU or ELU) as their default activation functions, due to their effectiveness in training deep networks.

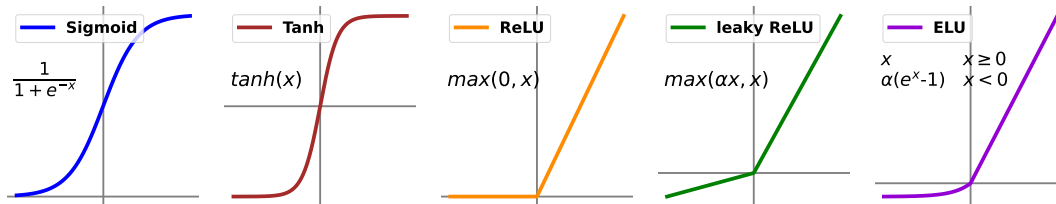


Figure 2.17: A subset of the most popular and common activation functions used in neural network architectures. The crossing of the coordinate axes denotes the origin (0,0) in the coordinate system.

During the forward pass of neural network computation, the input signals from connected neurons are multiplied by their respective synaptic weights and summed up. This weighted sum is then passed as an input to an activation function to

produce the neuron’s output. This operation is referred to as multiply and accumulate (MAC) operation. Typically, weights are real positive or negative floating point numbers (without main constraints) and a multiplication is performed in CMOS logic. In general, the MAC operation is time and energy intensive, thus an improvement in hardware design would accelerate the learning in a neural network. Considering the energy consumption, there are ideas proposed for storing weights in MTJs and performing the MAC operation in an analog manner instead of digitally, which would allow for “in-memory” computing [15, 16]. Depending on the architecture this might involve various constraints on the weight values, such as limited precision [143], a limitation on the sign (only positive weights) or a limited weight quantization (e.g. 4-bit weights) [144]. However, in this work the focus is on the combination of SMTJs and CMOS logic as an implementation of the activation function to accelerate learning in neuromorphic architectures.

2.6.3 Learning

In neurology, learning is the modification of biological synapses between neurons in the brain induced by a stimulus. In artificial neural networks, synaptic weights are represented by real numbers. Here, learning is defined as the adaption/adjustment of weights according to a specific learning rule. There are different approaches to learning algorithms [22], like target propagation, feedback alignment [145], Hebbian learning [146] or backpropagation [147]. However, the most popular and most established algorithm is called backpropagation (BP) [22, 147]. Its efficiency comes from the deterministic calculation of errors and gradients at each neuron of the neural network with the aim to minimize prediction errors. In BP the prediction error is propagated backwards through the network. In general, backpropagation demands weight symmetry in the forward and backward paths.

In a feed-forward NN a given input \mathbf{x} propagates to the next layer by weight multiplication. The “weighted sum”, also called “pre-activation”, $\mathbf{a} = \mathbf{W} \cdot \mathbf{x}$ provides the input for the neuron’s non-linear activation function f . The neuron’s output, called “activation”, is then computed by: $\mathbf{x} = f(\mathbf{a})$ and propagated to the next layer of neurons by applying the next weight multiplication. After the forward pass, the network’s output $\hat{\mathbf{y}}$ (NN prediction) is compared to the actual target values \mathbf{y} using a loss function. The loss function \mathcal{L} , sometimes called cost function, quantifies how far off the predictions are from the true values. For convenient reasons it is often defined as the mean squared error: $\mathcal{L} = \frac{1}{2}(\hat{\mathbf{y}} - \mathbf{y})^2 = \frac{1}{2}(\mathbf{x}^L - \mathbf{y})^2$, where $\hat{\mathbf{y}} = \mathbf{x}^L$ is the output of the last layer L in the NN and \mathbf{y} and \mathbf{x} are column-vectors. The loss \mathcal{L} has to be minimized, which is

2. Theoretical background

done by computing the gradients of the loss with respect to each parameter in the network, starting from the output layer L and moving backward layer by layer. This is called backpropagation. Once the gradients are computed, an optimization algorithm updates the network’s weight parameters in the direction that reduces the loss/cost. This process is repeated multiple times until the network’s performance converges to a certain level. The weight updating speed is determined by the “learning rate” η , which has to be chosen appropriately for efficient and fast learning. In BP the weight update at layer $l \in (1, \dots, L)$ is calculated according to:

$$\Delta w_{ij}^l = -\eta \frac{\partial \mathcal{L}}{\partial w_{ij}^l} \quad (2.71)$$

w_{ij}^l is defined as the weight between layer $l-1$ and l , where ij defines the connection between the j -th node in layer $l-1$ and the i -th node in layer l . The gradient of the loss function with respect to a weight at layer l can be determined by applying the chain rule:

$$\frac{\partial \mathcal{L}}{\partial w_{ij}^l} = \frac{\partial \mathcal{L}}{\partial \mathbf{x}^L} \cdot \frac{\partial \mathbf{x}^L}{\partial \mathbf{a}^L} \cdot \frac{\partial \mathbf{a}^L}{\partial \mathbf{x}^{L-1}} \cdot \frac{\partial \mathbf{x}^{L-1}}{\partial \mathbf{a}^{L-1}} \cdot \frac{\partial \mathbf{a}^{L-1}}{\partial \mathbf{x}^{L-2}} \cdots \frac{\partial \mathbf{x}^l}{\partial \mathbf{a}^l} \frac{\partial \mathbf{a}^l}{\partial w_{ij}^l} \quad (2.72)$$

Given the weighted input $\mathbf{a}^l = W^l \cdot \mathbf{x}^{l-1}$ and activation $\mathbf{x}^l = f(\mathbf{a}^l)$ at layer l the derivatives are:

$$\frac{\partial \mathbf{x}^l}{\partial \mathbf{a}^l} = f'(\mathbf{a}^l), \quad \frac{\partial \mathbf{a}^l}{\partial \mathbf{x}^{l-1}} = W^l \quad (2.73)$$

Here, $f'(\cdot)$ is the derivative of the activation function with element-wise operation and W the weight matrix. Thus, backpropagation is primarily the computation of gradients and matrix multiplications. By computing the error backward the calculated error of the last layer can be used for the previous layer, thereby saving computation time and accelerating learning. In addition, it simplifies the computation due to less matrix operations. Despite the efficiency of backpropagation, it becomes highly time-consuming when applied to large artificial neural networks, and it also lacks plausibility when compared to the learning process in biological neural networks. An alternative approach based on a local learning rule is described in the following.

2.6.4 Node perturbation

Contrary to the well established backpropagation, a more biologically plausible learning rule is called node perturbation [148, 149]. The approach is to perturb the neural activity and infer from the output error and local neuron activity the

weight updates. This is in contrast to backpropagation where the error is calculated and propagated backwards using gradients of the activation functions and symmetric weights. If a perturbation enhances performance, weight adjustments are made to reinforce the perturbation, and vice versa. The key advantage here is that it relies solely on a global error signal and (local) pre- and postsynaptic activity, therefore NP learning rule can speed up learning significantly. Similar to NP, weight perturbation (WP) [150] is a method to perturb weights instead of neurons. Despite performing better for limited precision weights, it has a drawback in computational complexity, which scales with N^4 for WP instead of N^3 for NP in a neural network with N neurons [150, 151].

In the human brain the neural activity is inherently stochastic [152] making stochastic MTJs well-suited for use in artificial neuromorphic hardware implementations. Backpropagation requires a feedback structure, which is not observed in neural circuits. In addition, it requires symmetric weights for the backward weight transport, which is also not present in biological neural networks. In order to avoid weight transport, a method called feedback alignment [153] was proposed, however this does not scale for large NN. More biologically plausible is node perturbation where the perturbations are typically random and are applied to all the nodes of a layer in a neural network. Each node (neuron) produces a noisy output according to: $\tilde{\mathbf{x}}_i = f(\mathbf{a})_i + \epsilon_i$, where ϵ_i is a value sampled from a Gaussian normal distribution $\mathcal{N}(0, 1)$. The gradient can then be estimated by the difference in loss function [148]: $\mathbf{g}_i \approx \tilde{\mathbf{g}}_i := (\tilde{\mathcal{L}}(\mathbf{x}^0, \tilde{\mathbf{x}}^L, \epsilon_i) - \mathcal{L}(\mathbf{x}^0, \mathbf{x}^L))\epsilon_i/\sigma$. Here, σ denotes the standard deviation, \mathcal{L} the loss function, ϵ the Gaussian noise, \mathbf{g} the gradient at node i and $\mathbf{x}^0, \mathbf{x}^L$ the input and output of the NN. Due to inherent stochasticity, easy read-out capability, energy-efficiency and CMOS compatibility, superparamagnetic tunnel junction are ideal components for hardware based analog learning via node perturbation.

2.6.5 Local learning in node perturbed neural networks

Here, we follow reference [149], to explain the principle of “activity”-based node perturbation. We consider a feed-forward (fully-connected) neural network comprising L layers. Consequently, the output (vector) \mathbf{x} of a specific layer l can be expressed as:

$$\mathbf{x}^l = f(\mathbf{a}^l) \quad (2.74)$$

Here, $\mathbf{a}^l = \mathbf{W}^l \mathbf{x}^{l-1}$ is the pre-activation, \mathbf{W}^l the weight matrix at layer l , $f(\cdot)$ the (element-wise) activation function. Consequently, input and output is defined by \mathbf{x}^0 and \mathbf{x}^L . The weight update under a credit assignment rule is called learning and is typically carried out via the backpropagation method [22].

2. Theoretical background

However, an alternative local-learning method is called node perturbation [148, 154]. When considering a perturbation ϵ of the pre-activation signal, the perturbed node output will be modified according to:

$$\tilde{\mathbf{x}}^l = f(\mathbf{a}^l + \epsilon^l) \quad (2.75)$$

The perturbation (vector) at layer l is generally sampled from a multivariate Gaussian distribution: $\epsilon^l \sim \mathcal{N}(\mathbf{0}, \mathbf{I}\sigma^2)$.

As in classical NP (see also Ref. [149]) we could consider a perturbation of a single node and derive the derivative of the loss function with respect to a pre-activation a_j^l for an element j in layer l . A perturbation introduced at a layer or also for a single node, will affect the loss function \mathcal{L} , thus we can define the NP loss difference from the perturbed network pass and from the unperturbed network pass: $\delta\mathcal{L} = \mathcal{L}(\tilde{\mathbf{x}}^l) - \mathcal{L}(\mathbf{x}^l)$. From the loss difference and the perturbation signal, it is possible to update the weights in a layer-wise fashion.

In a new approach, it is proposed to derive the directional derivative by perturbations, in which the derivative of a (activation) function is obtained based upon a random vector direction [149]. This random vector r at layer l is defined as:

$$\mathbf{r}^l = \begin{cases} \mathcal{N}(\mathbf{0}, \mathbf{I}) & \text{for } l = k \\ \mathbf{0} & \text{otherwise} \end{cases} \quad (2.76)$$

The directional derivative with respect to a specific layer via perturbation can be obtained for a deep neural network by [149]:

$$\nabla_{\mathbf{r}}\mathcal{L}(\mathbf{x}^l) = \lim_{h \rightarrow 0} \frac{\mathcal{L}(\tilde{\mathbf{x}}^l(h)) - \mathcal{L}(\mathbf{x}^l)}{h|\mathbf{r}|} = \lim_{h \rightarrow 0} \frac{\delta\mathcal{L}}{h|\mathbf{r}|}, \quad (2.77)$$

where \mathcal{L} is the loss function, h the noise magnitude and $\tilde{\mathbf{x}}^l$ the perturbed output and \mathbf{x}^l the unperturbed output. Equation 2.77 allows determining the gradient of each layer in the neural network, however, it is also inefficient considering the number of perturbations needed is proportional to the number of layers in the network. This is due to the fact that the perturbation has to be applied layer-wise one after another to acquire the directional derivative in each layer, which requires a high degree of noise control in the hardware circuitry.

To circumvent this problem, an alternative type of NP, called ‘‘activity node perturbation’’ approximates the directional derivatives across the whole network simultaneously [149]. This technique is not only more biologically plausible but also computationally efficient, faster and more hardware friendly. The weight

update in the “activity”-based NP approach is given by [149]:

$$\Delta \mathbf{W}^l = -\sqrt{N^l} \delta \mathcal{L} \frac{\tilde{\mathbf{a}}^l - \mathbf{a}^l}{|\tilde{\mathbf{a}}^l - \mathbf{a}^l|} (\mathbf{x}^{l-1})^\top \quad (2.78)$$

Here, instead of using the noise vector information for the weight update, the pre-activation difference is used. This has the advantage that the noise signal information is not necessary anymore, but instead we can use the change in pre-activation. A rescaling factor is given for the weight update, which comprises the square root of the layer size $\sqrt{N^l}$ and the absolute length of the activity difference. However, it turns out that the weight update is only dependent on the loss difference and local information, namely the pre-activation difference and previous layer outputs. This type of method, reliant solely on local information and a single global loss, is known as “direct feedback alignment” [145] and is one of the biologically plausible learning rules. It provides a great advantage over classical backpropagation in which the error propagates through the network, starting at the output layer and ending at the input layer. As an additional advantage, the NP learning algorithm is still efficient in networks where noise is inherently present and can not be turned off, like in biological nervous systems. Without an “unperturbed” network pass, but instead with two noisy passes, the network is able to perform approximately equally well [149]. This means that the activity difference from two different noise perturbations is sufficient for local learning. Consequently, this provides a great opportunity to incorporate stochastic MTJs in an analog node perturbed network, since SMTJs exhibit naturally uncorrelated and random noise of appropriate amplitude [106]. The key feature necessary to leverage the advantage of NP in hardware is random noise (like Gaussian-distributed noise), in resistance or conductance, which is typically inefficient to generate in classical CMOS architecture. Through the combination of multiple interconnected SMTJs, it is possible to obtain a Gaussian-like distribution in conductance, as it will be elaborated in the last chapter of this work. In addition, the NP learning algorithm can significantly improve under de-correlated inputs [149]. This method of decorrelating the inputs at each layer is called “constrained parameter inference” COPI [155]. However, this is out of the scope of this work and is not considered in greater detail.

2. Theoretical background

3

Experimental techniques

3.1 Characterization

Within this section, fundamental characterization techniques are described, which are required for the fabrication MTJ nanopillar devices. The characterization techniques encompass X-ray reflectometry to assess thin film properties and to calibrate sputter rates, atomic force and scanning electron microscopy, to ascertain nanopillar dimensions and to control individual fabrication steps, current in-plane tunneling measurements to characterize various TMR stacks and superconducting quantum interference device measurements to quantify magnetization and anisotropies.

3.1.1 Atomic force microscopy

Atomic force microscopy (AFM) [157, 158] is a widely employed technique utilized for the high-resolution visualization and characterization of sample surfaces. This method offers numerous advantages, notably the attainment of sub-nanometer resolution and its ability to investigate both conductive and non-conductive specimens, a limitation present in the first implemented scanning probe approach, Scanning Tunneling Microscopy (STM) [157]. AFM enables the examination of surfaces without necessitating a vacuum environment. AFM's fundamental configuration involves a probing mechanism, typically composed of silicon, silicon nitride, or silicon oxide. The probe, specifically a cantilever equipped with a sharp tip, interacts with the sample surface. As shown in Figure 3.1, a focused laser beam is directed onto the cantilever, and the light reflected from it is detected with spatial precision by a photodetector. Variations in surface height result in cantilever deflection, generating distinct outputs on the 4-quadrant photodetector. The detected signal is fed into a feedback loop, which actuates precise movements of the piezoelectric element and the cantilever along the z-direction,

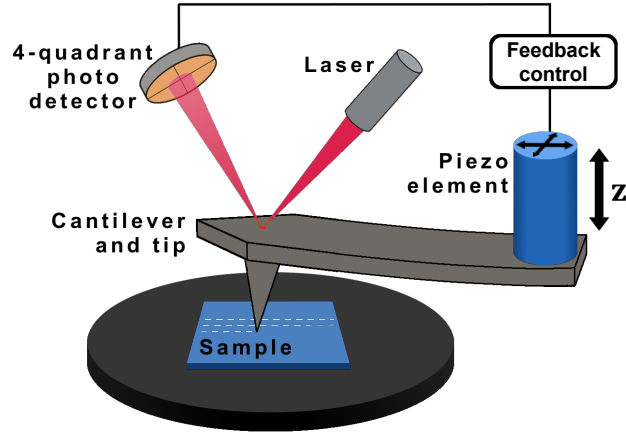


Figure 3.1: Schematic illustration of atomic force microscopy (AFM). A laser beam is directed onto the cantilever tip, and its reflection is captured by a photodetector. Any displacement of the cantilever causes alterations in the detected signal, which provides feedback on the piezo element attached to the cantilever, enabling precise adjustments of the tip height. Image taken from Ref. [156].

facilitating meticulous 3D mapping of the sample surface through line-by-line scanning. Upon bringing the cantilever tip close to the sample surface, the atoms in both the tip and the surface interact, resulting in either attractive or repulsive forces depending on the distance r . The potential of this interaction is called Lennard-Jones potential [159]. At large distances, the force exhibits an attractive nature ($\propto r^{-6}$) due to van der Waals interactions, and at smaller distances, there is a repulsive force ($\propto r^{-12}$) caused by the overlapping of atomic orbitals [159]. In the so-called “contact mode”, the AFM operates in the repulsive range of the interaction potential [158].

3.1.2 Current in-plane tunneling

Current-in-plane tunneling (CIPT) measurements allow for the calculation of TMR and resistance area product (RA) of an as-deposited wafer without any processing and patterning [160] and therefore provide a fast and reliable method to determine the most important parameters of a magnetic tunnel junction. The method is based on a series of four-point probe resistance measurements on the surface of the TMR stack under different magnetic fields. A CIPT probe consists of 12 tips with micrometer spacing, as shown in Figure 3.2a. For each probe spacing x , the resistance and the current-in-plane magnetoresistance MR_{cip} is measured by applying a current to two outer electrodes ($1=+I$ & $4=-I$) and

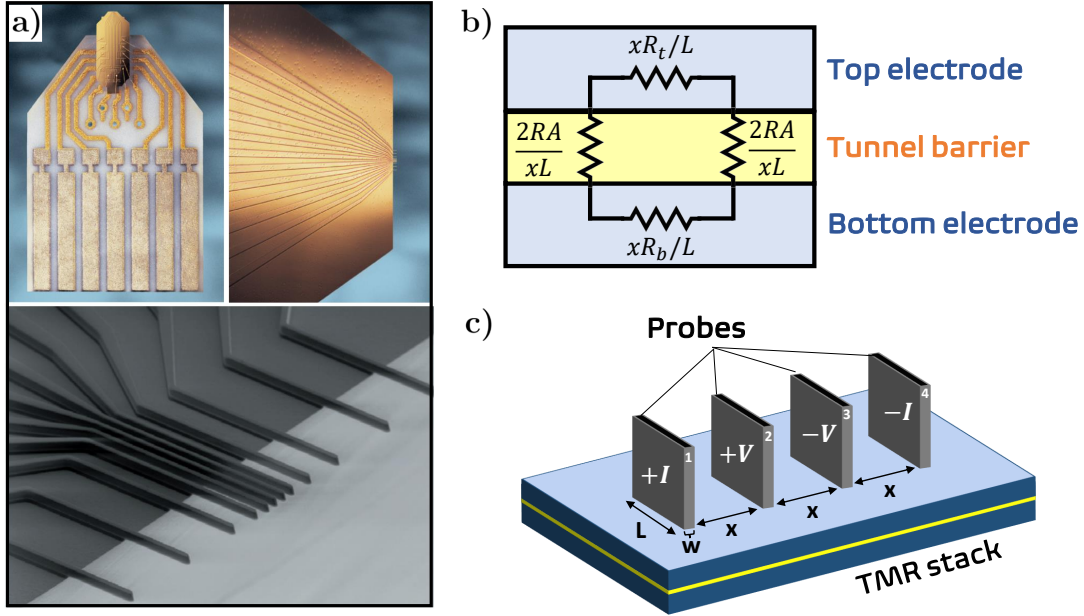


Figure 3.2: a) Photographs of the CIPT probe and an SEM image of the 12 tips, adapted from Ref. [160]. b) Circuit of the simplified resistor toy model. c) Schematic of the 4-point probe measurement where the tips touch the surface of the TMR stack.

sensing the voltage between two inner electrodes ($2=+V$ & $3=-V$), as illustrated in Figure 3.2c. Each tip has a length of L and a width of w . In Figure 3.2b a simplified description of CIPT is considered by a resistor toy model for two contacts placed on the surface. Here, the TMR stack is modeled by two horizontal resistors of the top and bottom layer (xR_t/L and xR_b/L) and two vertical resistors ($2RA/xL$ and $2RA/xL$), which model the tunnel barrier. R_t and R_b are sheet resistances of the top and bottom layers. For an appropriate spacing of the tips, a fraction of the current will flow through the tunnel junction, connecting the top and bottom resistors in parallel. For the simplified resistor model, the total resistance between two tips can be calculated according to [161]:

$$R_{cip} = \frac{x}{L} \frac{R_t R_b}{R_t + R_b} \left(1 + 4 \frac{R_t}{R_b} \frac{1}{4 + x^2/\lambda^2} \right) \quad (3.1)$$

Here, $\lambda = \sqrt{RA/(R_t + R_b)}$ is the characteristic length scale, where the tunneling current significantly contributes to the total measured resistance. For this reason, the tip spacing x has to match the top and bottom resistance as well as the RA product. If the spacing is too small, the current will only flow through the top layer and if the spacing is too large, the tunnel barrier resistance is negligible

3. Experimental techniques

compared to R_t and R_b , leading to a parallel connection of R_t and R_b , as illustrated in Figure 3.3a. If the tip spacing has an appropriate distance, the sheet resistance (R_{\square}) measurement is sensitive to the TMR resistance change arising through parallel and antiparallel alignments of the ferromagnets for positive and negative applied fields. For the *CIPTech-M200* tool from *Capres (kla)* [160], three types of probes (M12PP_004, M12PP_005 and M12PP_007) with different tip spacing are provided. The correct solution for the total resistance can be derived by solving a differential equation, as demonstrated in Ref. [161]. Under an applied external field, the parallel or antiparallel state of the MTJ exhibits a low or a high RA product, thus affecting the total (sheet) resistance R_{cip} . From the low (parallel) and high (antiparallel) resistance ($R_{cip}^{p/ap}$) the CIPT magnetoresistance is deduced from $MR_{cip} = (R_{cip}^{ap} - R_{cip}^p) / R_{cip}^p$. By performing a series of CIPT measurements for various probe spacings and fitting the resistance curve to extract R_t , R_b and RA it is possible to obtain the TMR value of the unpatterned stack according to: $TMR = (RA_{ap} - RA_p) / RA_p$. In Figure 3.3b a CIPT measurement for an annealed TMR stack is depicted, providing TMR ratio and RA product through a fit of the measurement data.

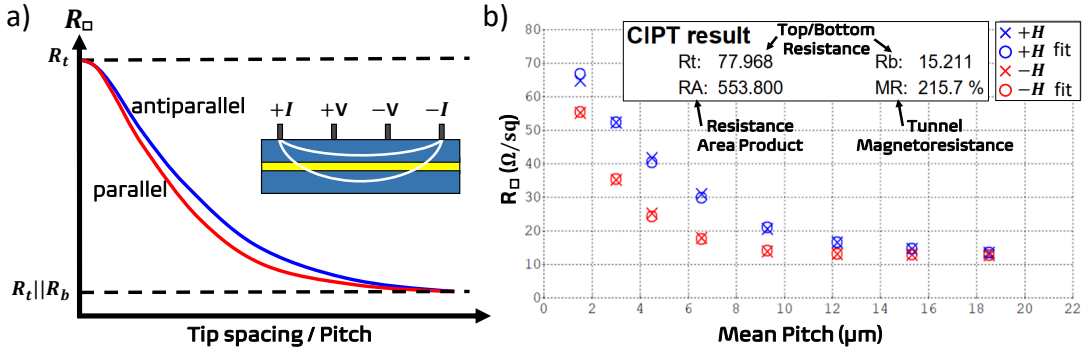


Figure 3.3: a) Sketch of a CIPT sheet resistance (R_{\square}) measurement of a TMR thin film. For a positive applied magnetic in-plane field, the MTJ is in the parallel configuration (red), whereas for a negative field, the MTJ is in the antiparallel configuration. b) CIPT measurement of a TMR thin film with 215.7% magnetoresistance (Figure generated by *Capres* software [161]).

3.1.3 Scanning electron microscopy

A scanning electron microscope (SEM) [162, 163] is a microscope that uses a beam of directed electrons to generate an image by scanning a sample surface and detecting scattered electrons. A raster scan of the sample is carried out, and the detected signal for each position is then transformed into a gray-scale pixel value. SEM is a powerful tool for analyzing nanomaterials due to its nanometer-scale

resolution of the order of 1 nm or even below [164]. Unlike optical microscopy, the resolution is not limited by the wavelength through Abbe's equation [165], but rather through an imperfect electron-optical system and the resulting electron spot size as well as the interaction volume of the electron beam with the specimen. In order to pattern nanopillars down to 50 nm, the commercial electron beam lithography system *Raith Pioneer* has been used [166]. This system combines a scanning electron microscope together with an electron beam lithography system. A more advanced and high-resolution SEM is the *Helios NanoLab 600i* [167], which allows the scanning of samples under specific angles. Typical settings involve a 20 kV electron beam acceleration voltage and the detection of secondary electrons by the so-called "through lens detector" (TLD) or "ion conversion and electron" (ICE) detector.

When a material is subjected to the impact of accelerated electrons with energies ranging from 1 to 30 keV, various phenomena occur, leading to the generation and detection of electrons and photons using distinct methods. These generated electrons and photons, known as scattered electrons or photons, are categorized based on their underlying physical processes. The subsequent sections provide concise explanations of these phenomena.

Backscattered Electrons

Backscattered electrons (BSE) are high-energy electrons with energies typically of a few keV (often simply defined as electrons with > 50 eV) that originate from the incident electron beam and are elastically scattered back (at wide angles) by interactions with the atoms' nuclei in the specimen. Since the BSE yield, namely the ratio of the number of emitted backscattered electrons to primary beam electrons, depends on the atomic number, these backscattered electrons carry information about the atomic number and weight of the elements in the specimen [164]. Thereby, an SEM can produce contrast-sensitive images, which indicate surface features and compositional variations in the specimen.

Secondary Electrons

Electrons can be in-elastically scattered at the specimen by a variety of interactions between the accelerated electrons and the atoms at the sample. Excitations and ionizations of the atoms can lead to the generation of secondary electrons (SE), which are most often used for the generation of an SEM image. Due to the low energy of these scattered electrons, typically secondary electrons are defined as electrons with energies below 50 eV [164]. Since these electrons only escape

from the upper part of the sample, a SE detector is mostly sensitive to the surface of the sample.

Auger electrons

Auger electrons [168] are generated when an atom undergoes ionization due to a primary electron, causing an outer shell electron to fall back into an inner shell vacancy [164]. During this transition, the excess energy is released, which can be carried away by an Auger electron [168, 169]. These electrons possess distinct energy levels and are valuable for obtaining chemical information about the specimen. Due to their low energy, only the emitted Auger electrons in proximity to the surface are detectable, and therefore these electrons are valuable for conducting surface analyses [164].

X-ray photons

The energy of the electron beam determines the generation of X-rays throughout the specimen's depth. This process involves inner-shell ionization, where a primary electron displaces a tightly bound inner-shell electron from an atom, creating an inner-shell vacancy. As a result, the atom undergoes a relaxation by having an outer-shell electron transition to fill the vacancy, thereby emitting an X-ray photon in the process. The emitted x-ray's energy corresponds to the element-specific energy difference between the involved electron shells, thus allowing for the identification of elements present in the scanned sample. This detection technique is called energy-dispersive X-ray spectroscopy (EDX).

3.1.4 Superconducting quantum interference device

A superconducting quantum interference device (SQUID) is a very sensitive magnetometer, allowing for the detection of extremely weak magnetic fields. A SQUID is based on the Josephson effect [170], a macroscopic quantum effect occurring in a superconductor, which is interrupted by a barrier, also called weak-link. The rf-SQUID consists of a superconducting loop interrupted by a single Josephson junction, whereas a dc-SQUID is based on a superconducting loop with two Josephson junctions [171]. Such a superconducting loop can only encompass multiples of a single flux quantum $\Phi_0 = h/2e$. Consequently, SQUIDs are sensitive to magnetic flux: $\Phi = \int B dA$ with B the magnetic flux density and A the area of the loop [171]. The critical current of the superconductor, which is the current at which the transition to superconductivity occurs, is strongly affected by the magnetic flux going through the loop. If a magnetic sample is

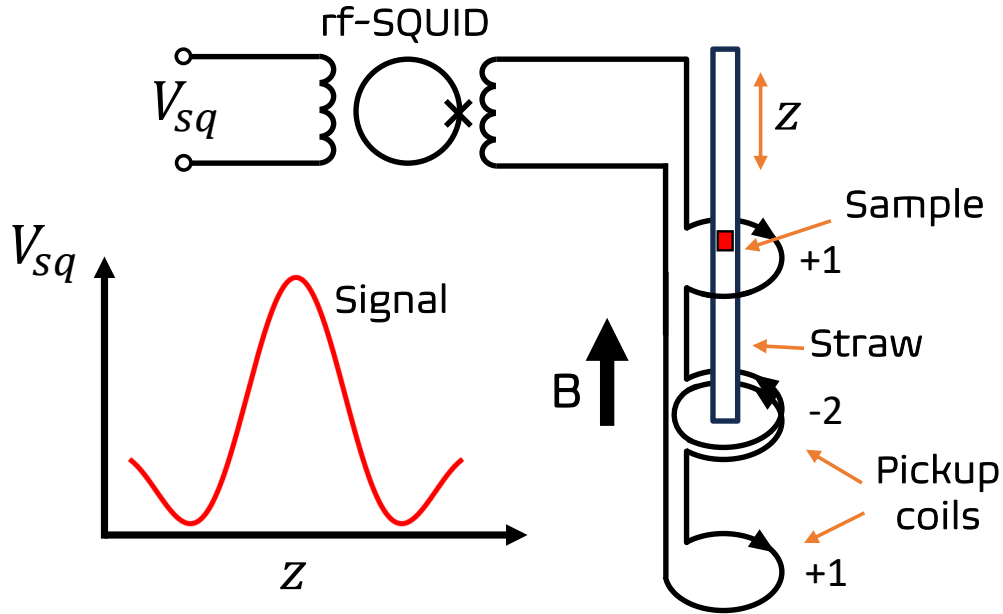


Figure 3.4: Schematic depiction of a SQUID magnetometer with a second-order gradiometer pickup coil and the illustration of the SQUID signal V_{sq} under axial movement of the sample.

placed close to the SQUID, then stray fields will flow through the superconducting loop, thereby affecting the critical current. The SQUID magnetometer operates by moving a magnetized sample through a pickup coil, typically a second-order gradiometer [171], which is a set of oppositely wound loops with winding numbers +1, -2, +1. This loop geometry will be insensitive to uniform magnetic fields in the z -direction as well as to the first derivative of the magnetic flux $\partial B_z / \partial z$, while a second order field gradient $\partial^2 B_z / \partial z^2$ will result in a non-zero net flux in the pickup coil [172]. By moving the sample, the magnetic moment of the sample induces an induction current in the pickup coil. The pickup coil then transfers the measured flux into the SQUID by inductive coupling. Furthermore, a compensation signal is coupled into the SQUID, which aims to compensate the magnetic flux through the SQUID loop [172]. The voltage of this compensation signal is then proportional to the magnetic moment of the sample. In general, a SQUID acts as a flux to voltage amplifier and if the sample is centered in the pick-up coil, it results in the maximum output voltage. The magnetization of a sample can be determined by the measured magnetic moment in relation to its volume. In this work, the SQUID magnetometer *MPMS XL* from *Quantum Design* was used.

3.1.5 X-ray reflectometry

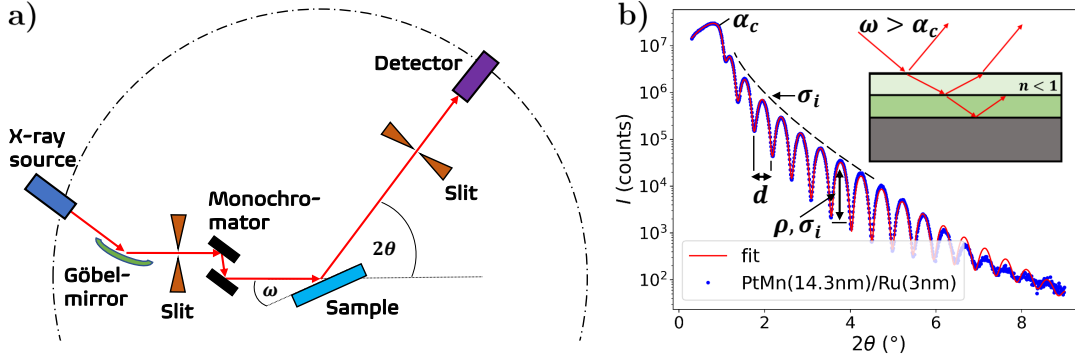


Figure 3.5: a) Schematic illustration of the basic principle of an XRR measurement. X-rays are generated at the copper cathode and are then collimated and monochromatized into a parallel beam using a Göbel mirror. The beam passes through a slit and a crystal monochromator before striking the sample at a grazing incidence angle ω . The reflection of the X-rays is measured at the detector. b) An XRR measurement showing the intensity on a log scale as a function of the angle. The oscillations indicate sample characteristics, like film thickness d , interface roughness σ_i , electron density ρ_i , and critical angle α_c .

X-ray reflectometry (XRR) [173] is a surface-sensitive measurement technique used to determine thin-film parameters including thickness, density, and surface or interface roughness. It is sensitive to film thicknesses in the range of approximately 1 nm to 100 nm. In this work, it was often used for the calibration of sputtering or etching rates. The basic principle of XRR is based on the interference of X-rays that are reflected off the surface and interfaces of thin films, at which the electron density changes. A schematic set-up is illustrated in Figure 3.5a, where X-ray source and X-ray photo-detector are placed on a circle in a goniometer geometry, whereas the sample is mounted in the center. The goniometer of the *Bruker D8 Discover* allows for a precise alignment of the incident angle ω and the reflection angle 2θ of the X-ray beam with respect to the surface plane. From a copper cathode X-rays are emitted to a parabolic shaped Göbel-mirror, which converts the divergent X-ray beam into a diffracted parallel beam. Due to Bragg diffraction, the X-ray radiation is monochromatized while the Bremsstrahlung is suppressed. A subsequent slit narrows the X-ray beam and an additional monochromator is used to obtain monochromatic X-rays with a wavelength of $\lambda = 0.15406$ nm (Cu- $K_{\alpha 1}$). Next, the incident X-ray beam is reflected at the sample surface and the intensity under an angle of 2θ is acquired by a radiation counter tube (detector). The intensity of the reflected X-rays is measured as a function of the incidence angle ω . For small

angles ($\omega < \alpha_c$) incident X-rays do not penetrate the sample and result in total reflection. For increasing incident angles, X-rays get reflected at the interfaces (see Fig. 3.5b) and lead to an interference pattern, called Kiessig fringes, due to a phase shift of the X-ray waves. This interference results in an oscillation of intensity, which is inversely related to the film thickness d , while the decline in intensity is related to the interface roughness σ_i . The amplitude of the oscillation also correlates with interface roughness and electron density ρ_i of the film, as indicated in Figure 3.5b. The oscillation pattern analysis was conducted using the open-source software *GenX* [174, 175]. The measurement data was fitted based on Parratt's algorithm [176] to extract film thickness, roughness, and electron density as fitting parameters. As an alternative, a single thin film thickness can be estimated from the angle difference between two adjacent oscillation-minima, according to: $d = \lambda / (2\theta_{i+1} - 2\theta_i)$. Figure 3.5b presents the XRR measurement of a $\text{SiO}_2/\text{Pt}_{80}\text{Mn}_{20}(14.3 \text{ nm})/\text{Ru}(3 \text{ nm})$ sample together with a *GenX*-fit.

3.2 Fabrication

In this section, the manufacturing process is detailed, starting with the deposition of the TMR stack and culminating in the patterning of MTJ nanopillar devices. The treatment of devices and challenges encountered during fabrication are also addressed and detailed.

3.2.1 Sputtering

Sputter deposition [178] is a physical vapor deposition method and is defined as the physical bombardment by ionized atoms onto a target material, thereby ejecting target atoms, which results in a deposition onto the substrate. By means of a high voltage, typically several hundred volts, between the target material and the anode a plasma ignites at the target cathode. Here, free electrons are accelerated in the electric field towards the anode and inelastically collide with the atoms of the sputtering gas (mostly inert Argon), resulting in the ionization of these atoms together with the release of secondary electrons. Positively charged ions of the sputtering gas are accelerated ballistically on approximately straight lines toward the negatively charged target, which yields to an erosion of the target material due to the bombardment with argon ions. The momentum of the Argon ions is transferred to the target atoms, resulting in the ejection of target atoms into the gaseous phase. In magnetron sputtering a magnetic field at the target yields to a confinement of the plasma close to the target surface and to

3. Experimental techniques

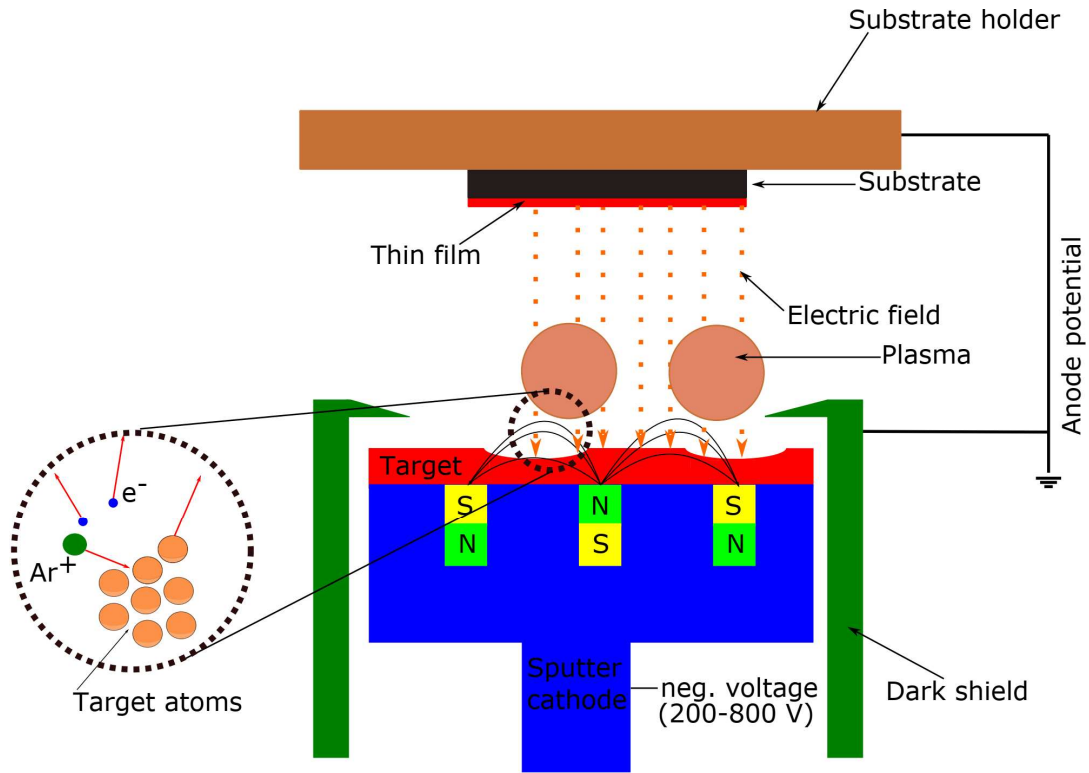


Figure 3.6: Illustration of sputtering, in which thin films are deposited on a substrate within a vacuum chamber. Argon gas is ionized by an applied voltage and argon ions are accelerated towards a negatively charged cathode. Under collision with the cathode (target) material, target atoms are ejected toward the substrate and lead to thin film growth. Image taken from [177].

spiral motions of the light-weight electrons due to the Lorentz force before being absorbed by the anode. This extension of the electron trajectory results in more ionizations per electron, thus increasing the sputter rate for the same applied power and pressure. An inert gas such as argon or krypton is used to prevent any chemical reaction during the sputtering process and the sputtering gas pressure is typically set to 0.05 mbar. For the growth of metallic films, a dc voltage is applied between the anode and cathode, however, for oxides and insulators rf (radio frequency) voltage is necessary to prevent the charging of the target, which would stop the sputtering process immediately. RF-sputtering is carried out at 13.56 MHz for the deposition of magnesium oxide and silicon nitride films. There are various different deposition parameters that can influence the growth process in magnetron sputtering, but the main parameters include: sputtering power, sputtering gas & gas pressure, target-substrate distance, sputtering angle and

temperature.

3.2.2 TMR stack

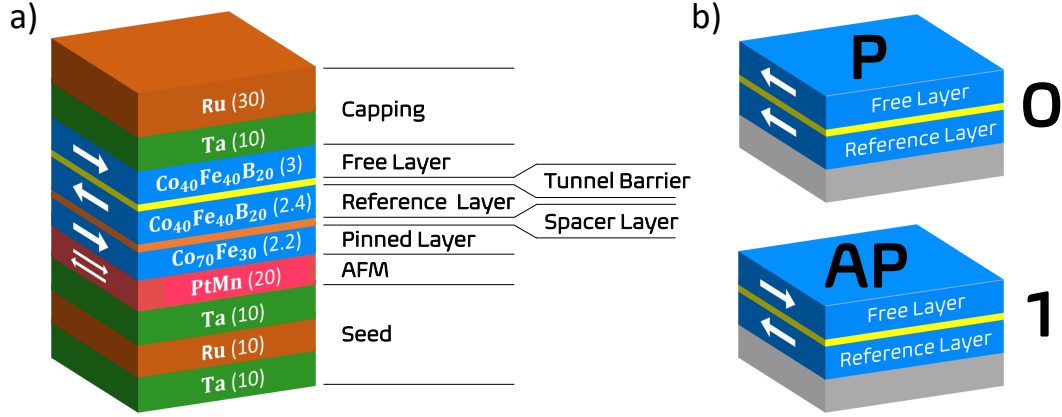


Figure 3.7: a) Illustration of the TMR stack, indicating the composition of each layer along with their respective film thickness (measured in nm). b) The parallel (P) and antiparallel (AP) configuration is depicted, providing the capability of storing 1 bit of information, representing either 0 or 1.

All TMR stacks, investigated in this work, consist primarily of an antiferromagnet, a synthetic antiferromagnet (SAF), a MgO tunnel barrier and a free-layer ferromagnet. The TMR films were deposited onto silicon wafers with a 1.5 μm oxide layer using rf- and dc-magnetron sputtering (*Singulus Rotaris* system). The used stack and its composition of different films with their respective film thicknesses in nanometer is: Ta(10)/Ru(10)/Ta(10)/PtMn(20)/CoFe(2.2)/Ru(0.8)/CoFeB(2.4)/MgO(1.1)/CoFeB(3.0)/Ta(10)/Ru(30), as illustrated in Figure 3.7a. The first CoFeB layer, called "free-layer", determines the MTJ configuration state (parallel or antiparallel), depending on its magnetization orientation with respect to the reference layer. Separated by an insulating MgO barrier, the next CoFeB layer, which is labeled "reference layer", is antiferromagnetically coupled with the CoFe ferromagnet (FM) layer underneath, called "pinned layer", and builds the SAF of the stack. The antiferromagnetic coupling of the SAF is accomplished by a thin (~ 0.8 nm) ruthenium spacer layer, which promotes interlayer exchange coupling. The in-plane magnetization orientation of the CoFe "pinned layer" is fixed through interlayer exchange coupling and by the PtMn antiferromagnet (AFM) layer through interfacial exchange coupling, resulting in an uniaxial anisotropy of the "pinned layer". The seed comprises a multilayer of Ta/Ru/Ta, which, due to its thickness, establishes a conductive

bottom layer in the stack. Tantalum acts as an amorphous and smooth seed, providing an excellent adhesion layer. The inclusion of a Ru layer prevents the formation of large grains or extensive columnar growth in the seed, due to different crystal growth directions of tantalum and ruthenium [179, 180] and provides a smooth surface of the seed underlayer [181]. This is crucial for subsequent layers, as any interface roughness would propagate throughout the entire stack. For the capping, a bilayer of Ta/Ru is chosen. Ta is necessary for the interface to the free layer, while Ru serves as good conductor, even in its oxidized state RuO₂. This characteristic of Ru is particularly advantageous for CIPT measurements, facilitating the characterization of the TMR stack (RA and TMR ratio) after the deposition. The TMR ratio of our stack is found to be approximately between 100 % and 215 % for resistance area (RA) products of 15 Ωμm² and 600 Ωμm², respectively. To introduce strong exchange bias, uniaxial anisotropy, and crystallization at the MgO barrier interface, an annealing procedure was carried out under an externally applied magnetic in-plane field. This procedure is described in the following section.

3.2.3 Annealing

Annealing is a process to achieve a crystalline interface [80, 181] at the tunneling barrier resulting in a high TMR ratio, and to set exchange bias of the TMR stack [182, 183]. During the annealing process, amorphous CoFeB crystallizes to a bcc crystal structure at the interface due to the neighboring crystalline MgO(001) texture, which serves as a crystal template for CoFeB [80]. In addition, field annealing of a ferromagnetic material can imprint a favored magnetization direction (parallel to the applied field) at elevated temperatures, due to the rearrangement of atoms. Mobile atoms can form pairs, align with the external magnetic field and thereby promote a short range directional ordering, which leads to uniaxial anisotropy in amorphous thin films [184]. In Figure 3.8 the annealing setup is shown. The sample is placed in a copper box, in which a cartridge heater (50 W) is embedded. The copper box is then mounted into a vacuum chamber. Annealing for TMR stacks is typically carried out at 300 °C in a vacuum under an external applied in-plane field of 300 mT. A turbo pump connected to the vacuum chamber provides a base pressure of 10⁻⁷ mbar and the water-cooled magnet is able to generate fields of up to 1 T (for an applied current of 45 A and a pole shoe distance of 0.104 m).

The annealing temperature is regulated by a PID controller (*Eurotherm* controller) and follows an annealing recipe, specified in the software program (*iTools*). The PID controller governs the power supply to the cartridge heater, thereby con-

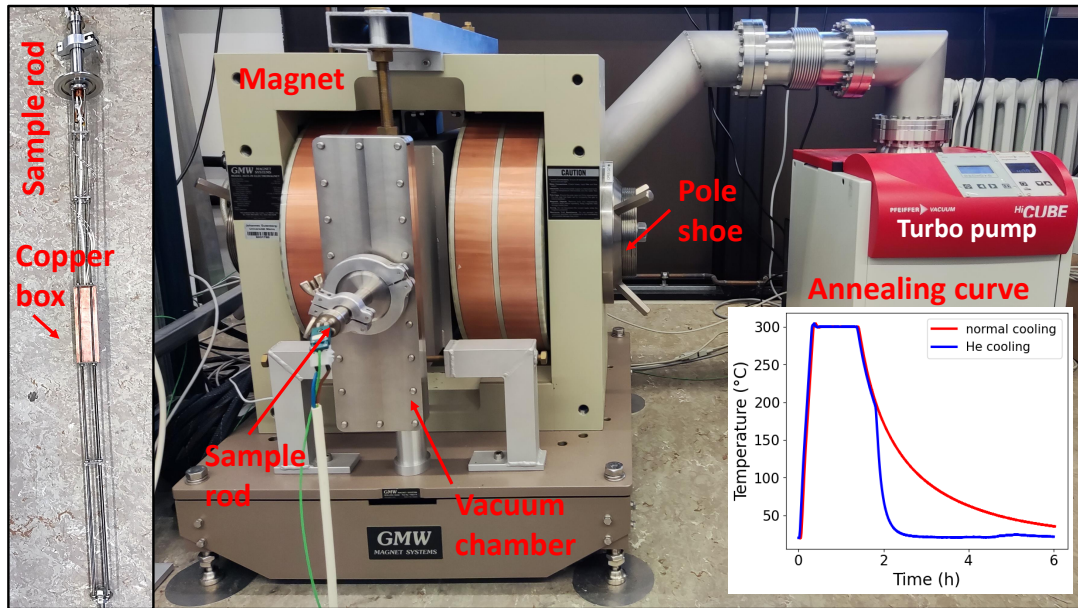


Figure 3.8: The setup of the annealing oven is illustrated, as well as the sample rod, which includes a copper box and a cartridge heater. The electromagnet provides an in-plane field during the annealing procedure. A turbo pump ensures a pressure of approximately 10^{-7} mbar in the vacuum chamber. The inset illustrates annealing temperature curves with and without He cooling.

trolling the dissipated power. During the annealing procedure, the temperature is gradually increased at a rate of $15\text{ }^{\circ}\text{C}/\text{min}$ to $300\text{ }^{\circ}\text{C}$ and maintained for 1 h or 2 h. Afterward, the heater is turned off, leading to exponential cool down, as depicted in Figure 3.8c. To expedite cooling, helium (He) gas can be introduced into the chamber, enhancing heat conduction from the heated copper box to the cooler chamber walls. In the subsequent step, the annealed TMR stack is transformed into MTJs of nanopillar dimensions through patterning facilitated by electron-beam lithography, which will be explained in the next section.

3.3 Patterning

In this section, the methods employed for structuring MTJ nanopillars with dimensions of 50-60 nm are laid out, along with an exploration of the challenges encountered during the patterning process.

3.3.1 Bottom electrode patterning

As a first processing step, a so-called “bottom electrode” is patterned by means of optical lithography and ion etching. The main purpose of this electrode is to have a well-defined electrical current path in the final MTJ device. The electrical current flow is then directed from the bottom electrode, through the MTJ into the top electrode. It is crucial for the top and bottom electrodes to remain electrically separated, and their lateral dimensions should largely avoid “overlap” to minimize the risk of a short circuit through the insulating passivation layer. In the process of fabrication, the requirement for a bottom electrode can be eliminated. However, this approach comes with certain disadvantages that need to be considered. First of all, such devices cannot be wire bonded due to the substantial forces generated by the ultrasound application of the wire bonder. These forces applied to the device can lead to either the breakage or the creation of pinholes in the insulating passivation material, thereby establishing a short circuit connection between the top and bottom electrodes. Another drawback is that multiple MTJ nanopillars are compelled to share the same electrical contact, like a shared ground (GND) connection. This shared electrical contact only allows interconnecting multiple MTJs in parallel and not in series. Another issue that arises in this configuration is the increased likelihood of short circuits caused by the presence of pinholes or impurities in the passivation layer. To mitigate the probability of short circuits, it is essential to minimize the intersection area between the top electrode and the bottom electrode. By reducing the overlap between these two electrodes, the risk of unintended electrical paths through the pinholes is minimized. A bottom electrode is patterned by optical lithography and mapped to the etch resist. Through IBE, the desired structure obtained by etching all non-covered regions of the sample. Bottom electrode designs are shown in Figure 3.9.

3.3.2 Optical lithography

In contrast to electron-beam lithography, optical lithography (or photolithography) [185] is an efficient and time-saving process, since all structures are exposed by ultraviolet (UV) light through a mask. The need for a mask and the limited resolution are slight disadvantages compared to EBL. Usually, UV light is utilized to expose structures in the micrometer range. Since the limiting resolution factor is the wavelength, a short wavelength in the UV spectrum is favored to enhance the resolution. In this work, a UV-lithography system, equipped with a mercury lamp, was used to define the bottom and top electrodes of the device. The lim-

ited resolution here is sufficient and compared to EBL this technique is much faster. The mask used for the optical lithography consists of a glass plate with chrome structures, which inhibit the UV exposure of the resist underneath. Two different recipes can be used for the optical lithography for patterning bottom or top electrodes. These are briefly described and outlined in detail in the Appendix A.3.2. The first recipe **A** is a single layer recipe, that does not provide a resist undercut, which might be necessary for the patterning of the bottom electrode. Recipe **B** is a two-layer recipe resulting in a large resist undercut, which results in better removability of redeposited sidewalls stemming from the etch process. This is important for further processing, since large sidewalls at the edge of the etch mask can lead to short circuits in the final device. Recipe **A** deploys a negative and recipe **B** two positive resists. In the context of lithography, a negative resist refers to a scenario where the segment of the photoresist exposed to light becomes less soluble to the developer solution. This is in contrast to a positive resist, where the exposed portions become soluble. The resist's thickness also influences the total development time. This thickness is primarily governed by the resist's viscosity and the maximum rotational speed ω of the spin coater. Through centrifugal force, the coating material is uniformly distributed over the substrate, resulting in a homogeneous film of a certain thickness. The duration of rotation is not crucial, as the resist solvent evaporates within the first seconds, but the critical factors are the maximum rotation frequency and the spin coater's acceleration speed. Detailed procedures for recipes A & B can be found in the Appendix A.3.2. Various optical masks used for the lithography of bottom and top electrodes are outlined in Figure 3.9.

3.3.3 Electron beam-lithography

Electron-beam lithography (EBL) [186] is a powerful tool, which enables the patterning of nanoscale electronic structures. This method allows for the precise design and placement of desired structures onto a specimen/sample. In EBL, a sharply focused electron beam is harnessed to inscribe well-defined patterns into an electron-sensitive polymer known as a resist. The interaction of the electron beam alters the resist's solubility, often achieved through polymer chain breakage or cross-linking. This alteration permits the selective dissolution of either the exposed (positive resist) or non-exposed (negative resist) regions of the resist when subjected to a solvent, called "developer". A key advantage over optical lithography lies in EBL's sub-10 nm resolution, enabling the creation of nanometer-scale structures without the need for a mask. The *Raith Pioneer* [166] EBL/SEM system was employed to carry out the patterning of all MTJ nanopillar devices. The

3. Experimental techniques

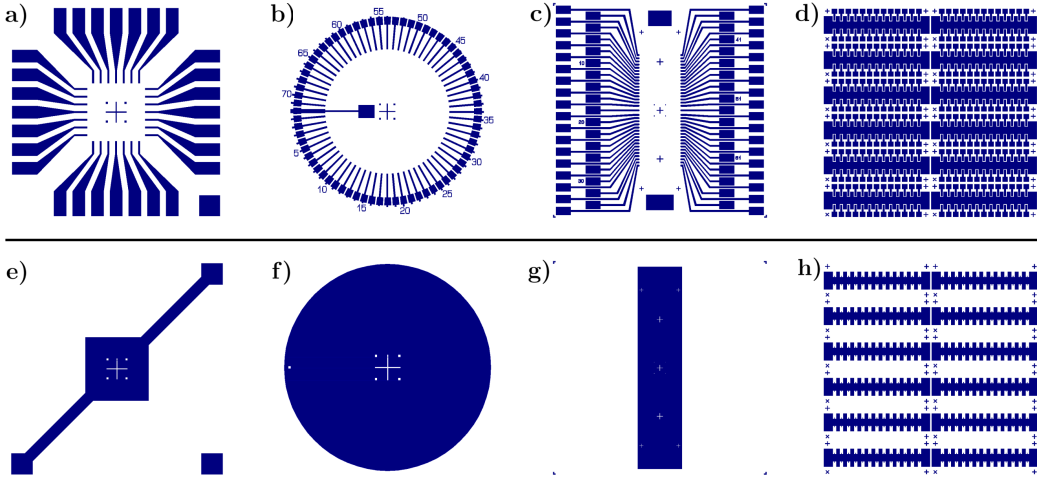


Figure 3.9: Development of optical mask designs. Each column represents an optical mask electrode design for MTJ nanopillars. The top row a)-d) refers to the masks for the top electrode, while the lower row e)-h) refers to the masks of the bottom electrodes. Masks d) and h) are suited for high-frequency measurements.

subsequent section will delve into the procedural steps, elucidated with the assistance of Figure 3.10. We start with a sample of a full deposited TMR stack on a substrate, as shown in Figure 3.10a. A single 110 nm thick layer of negative resist (*AR-N7520.07* [187]) is uniformly spun-coated onto the TMR stack, as shown in Figure 3.10b. The detailed recipe can be found in the Appendix A.3.2. Next, specific areas of the resist are exposed to an electron beam (Fig. 3.10c). This exposure results in a polymer cross-linking reaction to combine smaller polymers into larger, less soluble ones. This change in solubility enables the selective removal of polymer resist regions when subjected to the developer solvent, forming the desired pattern. After exposure, the exposed sample is immersed in a liquid developer to dissolve non-crosslinked regions of the resist. This development process is depicted in Figure 3.10d. In the next step, nanopillars are structured by ion beam etching (Fig. 3.10e). To obtain the final structured device, the (etch) resist can be removed by a remover solvent like acetone and under ultrasonic treatment. The etch resist is then removed by immersing it in a solution, thereby dissolving the resist, as illustrated in Figure 3.10. An alternative resist to *AR-N7520.07* is anorganic hydrogen silsesquioxane [188], also called “HSQ”. Due to its high resolution below 10 nm [189], excellent etch resistance and high sensitivity to electrons, it is one of the preferred resists for nanostructure patterning. HSQ has a cage-like structure and forms siloxane bonds under electron exposure HSQ [190] resulting in a strong etch resist. Despite all of these advantages, it

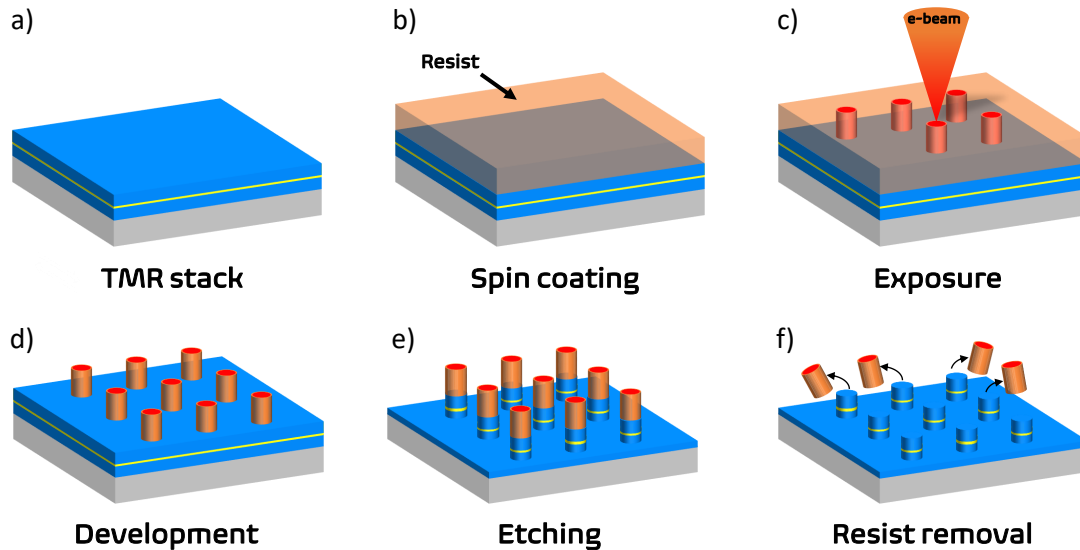


Figure 3.10: Illustration of the electron-beam lithography (EBL). a) The sample is a TMR stack, indicated by two ferromagnetic layers (blue) and the tunnel barrier (yellow), onto a substrate (grey). b) Spin coating of a resist polymer. c) Electron-beam exposure of defined structures, which chemically modifies the resist under electron bombardment. d) Development process, in which non-exposed regions are solved. e) Ion beam etching below the tunnel barrier. f) Resist removal procedure, where the resist capping is dissolved and removed.

is difficult to remove it again after the exposure due to the SiO_2 -like properties. For this reason, an organic e-beam resist, like *AR-N7520.07*, with lower spatial resolutions (30 nm), has been favored.

Alignment and exposure

For each lithography procedure, the system has to perform a new calibration of the beam deflection, which is called “write field alignment” [166]. In contrast to the automatic write field alignment, manual alignment provides more precision and control and has to be used for nanostructuring. However, for the write field alignment procedure, a reference point on the sample is required, ideally possessing sharp contrast and being only a few nanometers in size. With a laser interferometer-controlled stage that allows for position measurements with a precision of a few nanometres [166], the stage is moved by a few μm and the reference point is scanned (SEM) again. If the reference point is shifted during the scan, this deviation can be corrected. After a few repetitions, this offset becomes small enough to allow for the patterning of 50 nm-sized nanopillars.

3. Experimental techniques

The exposure dose D , is defined as the amount of charge deposited per unit area under an e-beam current I for a given dwell time τ and step size s : $D = \tau I/s^2$. For 110 nm thick *AR-N7520.07* resist structures with lateral dimensions below 100 nm, an exposure dose of 220 $\mu\text{C}/\text{cm}^2$ at 30 kV beam voltage is required. The high beam voltage of 30 kV is necessary to obtain sharp vertical defined edges, owing to reduced forward scattering. However, it requires a higher dose due to fewer secondary electrons in the exposed resist area. The minimum available aperture of 7.5 μm is used and corresponds to a beam current of approximately 8 pA. With a step size of 2 nm the dwell time is calculated to be approximately 1 μs . In order to expose a single circular nanopillar of a diameter of 50 nm this takes about 0.2 s. Comprehensive steps for this procedure are detailed in the Appendix A.3.1.

Distortion and resolution limit

Despite the fact that the wavelength of the accelerated electrons is several orders of magnitude smaller than visible light, and only limited by their kinetic energy, the feature resolution in EBL is limited by other factors. The fundamental limitation in practical EBL arises from the interaction range of (secondary) electrons with the resist and resist development mechanisms after the beam exposure, and is not limited by electron beam optics [191]. The EBL *Raith Pioneer* resolution limit is approximately 3 nm, while the resolution of the resist *AR-N7520.07* is around 30 nm. Electrons scatter due to electrostatic interactions, and this scattering causes the electron beam to spread out as it interacts with the resist on the sample. In addition, it can cause charge accumulation on the surface, which leads to electrostatic repulsion, causing further deflection from the charged area and widened pattern dimensions [192]. Depending on the acceleration voltage, forward scattering (inelastic small angle scattering) and backscattering (elastic large angle scattering) do also contribute to a limitation in resolution [191].

However, besides the resolution limit, a designed pattern structure may be distorted due to different effects such as aberration or astigmatism. Spherical aberration refers to a sharpness error, induced by the imperfect magnetic focus lens of the EBL system. It causes parallel incoming rays on the optical axis to not have the same focal point after passing through the system, due to different deflection. As a result, they do not converge at a single point. Chromatic aberration occurs when accelerated electrons possess different kinetic energies, thus different wavelengths, which results in different focal points and a blurring on the image plane. This leads to aberration in the axial (or longitudinal) as well as lateral (or transversal) direction. Astigmatism is characterized by divergent rays in two

orthogonal planes having distinct focal points. When the effect of astigmatism occurs, due to imperfections in the focus lenses, the shape of a nominally round beam becomes ellipsoidal, since the vertical and horizontal planes will come into clear focus at separate distances. To compensate for this effect, a stigmator [186], which is either electrostatic or magnetic, is used to force the beam back into its optimum shape.

3.3.4 Nanopillar resist collapse

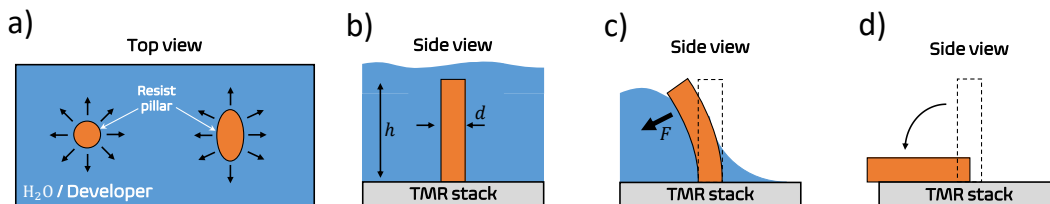


Figure 3.11: a) Illustration of the top view of a circular and an ellipsoidal etch resist in a developer solution bath. b) Side view the structured resist of height h and diameter d in the developer solution. c) An asymmetric evaporation of the solvent yields to an adhesive force F acting on the resist. d) The “collapsed” resist pillar after the evaporation of the solvent.

In the MTJ nanopillar patterning it is advantageous to have a thick etch resist in order to easily facilitate the “pillar opening” process (see Sec. 3.3.7). However, small diameters result in large aspect ratios of height to diameter. For a high aspect ratio patterned nanopillar, the development step after the e-beam exposure can lead to a fall over, called “collapse” of the pillar. This collapse arises due to capillary forces of the evaporating water (or developer liquid). It originates from the non-uniform evaporation of the liquid, which then exerts asymmetrical forces onto the resist structure [193]. This “collapse” effect is illustrated in Figure 3.11. Based on this effect, there exists capillary-force-induced collapse lithography (CCL) technique [194] with a high reproducibility and a high yield for nanogap fabrication. The collapse of nanostructures becomes important for high aspect ratios and lateral structure sizes below approximately 100 nm (since typical spin-coated resist heights are of the order of a few hundreds of nm). If an aspect ratio of a nanometer-sized pillar is greater than six, a nanopillar collapse is very likely. Possibilities to reduce the nanopillar collapse is to reduce the surface tension of the liquid. For instance, a methanol solution has lower surface tension than water. However, it is easier to reduce the resist thickness by increasing the spin coating frequency. Therefore, to avoid the resist collapse, the aspect ratio is kept between two and three. Figure 3.12 shows an SEM image of collapsed nanopillars, which

3. Experimental techniques

have been circular patterned but appear to be rectangular on the surface, due to the cylinder shape of the pillar.

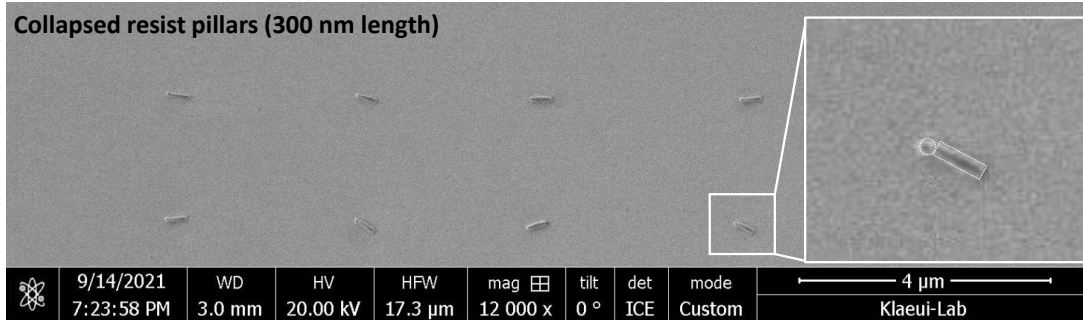


Figure 3.12: SEM image reveals collapsed resist nanopillars. The diameter is below 100 nm and the pillar height is approximately 300 nm. The inset provides an enlarged view, showcasing a resist pillar lying on the surface.

3.3.5 Ion beam etching

In order to structure MTJ nanopillars and electrode leads, ion beam etching (IBE) has been used. The ion beam etching system *IonSys500* [195] from *Meyer Burger* allows for homogeneous etching on samples of a diameter of 150 mm. Ion beam etching is a physical dry etching process in which ions are directed toward a sample surface under a specific angle. The selectivity of IBE is relatively low as it does not discriminate between individual material layers, like reactive IBE, which is often used by researchers and in the semiconductor industry. The *IonSys500* system is illustrated in Figure 3.13. The inert gas argon is injected into the discharge chamber, wherein a microwave electromagnetic field, generated by electron cyclotron resonance (ECR), is coupled to a microwave antenna to excite a plasma. A mass flow controller (MFC) regulates the flow of argon gas into the ignition chamber. Argon ions get accelerated towards the sample under an applied anode potential of $V_{beam} = +300\text{ V}$ and are subsequently de-ionized by a neutralizer in order to obtain a neutral beam of argon atoms to avoid electrical charging of the sample. The acceleration voltage at the inner grid results in a further acceleration of the ions and a subsequent de-acceleration behind the grid. The outermost grid is set to ground and thereby avoids further de-acceleration of the ions after the ejection of the plasma chamber. Depending on the beam voltage V_{beam} , the argon energy typically ranges from 300 eV to 1 keV and is able to effectively remove material from the sample surface upon impact under an angle α . The angle is here defined as the angle between the direction of the ion beam

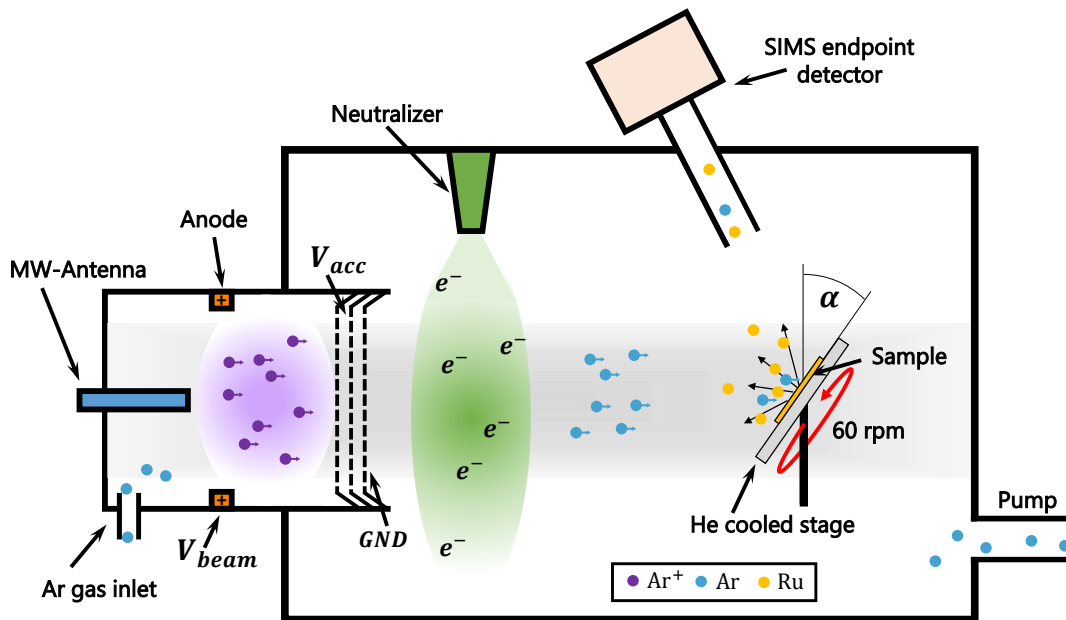


Figure 3.13: Sketch of the *IonSys500*, an ion beam etching (IBE) system with an integrated SIMS detector. Argon ions are accelerated and get neutralized before they collide with the sample under an angle α .

and the sample surface normal. The ejected material is exhausted by vacuum pumps, however, particles of the removed material can re-deposit on the sidewall of the resist and the desired structure. Thus, this so-called “redeposition” can lead to short circuits at the tunneling barrier. By altering the incidence angle, the rate of redeposition changes, as well as the etching rate [196, 197]. Another consequence of redeposition is called “fencing”, where redeposited material leads to undesired fence-like walls at the edges of an etch resist. To achieve a uniform and symmetric etching, the sample is rotated at 60 rpm.

The sputter yield Y , which is the mean number of emitted particles over incident particles, is a function of the ion mass, ion energy, target material and incidence angle and approximately follows [197]: $Y = \cos(\alpha) \exp(a \sin^2(\alpha))$, where α is the incidence angle and a is a material and ion-specific parameter. It is a result of two competitive effects of increased deposited energy and decreased traveling path by the sputtered atoms under an increased incidence angle. High sputter yields or etch rates are typically found for 40-60°. For low etching angles below 20°, etch trenches build at the edge of the mask, due to an increased flux density, as illus-

3. Experimental techniques

trated in Figure 3.14a. For too high etching angles (approximately $>50^\circ$), a long tail formation is present, which can lead to an increased effective size of the MTJ pillar, as depicted in Figure 3.14c. For sharp vertical etched walls, angles between $20\text{-}30^\circ$ are favored [198]. Another problem for nanopillar MTJs is redeposition at the tunnel barrier [199], as shown in Figure 3.14d-e, which can lead to short circuits or an electrical bypass at the tunnel junction, thereby greatly decreasing the TMR ratio [200, 201]. This electrical bypass due to redeposition is typically modeled by an additional shunt resistor R_s parallel to the MTJ resistance. The reduced TMR therefore can be calculated by [196]: $TMR' = TMR/(1 + R_{ap}/R_s)$. For angles below 30° this effect is strongly present [196]. In order to avoid redeposition at the tunnel barrier and to obtain almost vertical walls, an etching angle of 35° was chosen. The 35° etching was stopped at the PtMn layer, where a subsequent etch step under 70° was carried out for two minutes. This high-angle etching step reduces redeposited material, trims the nanopillar and smoothens the surface. With this recipe, an acceptable nanopillar yield is achieved.

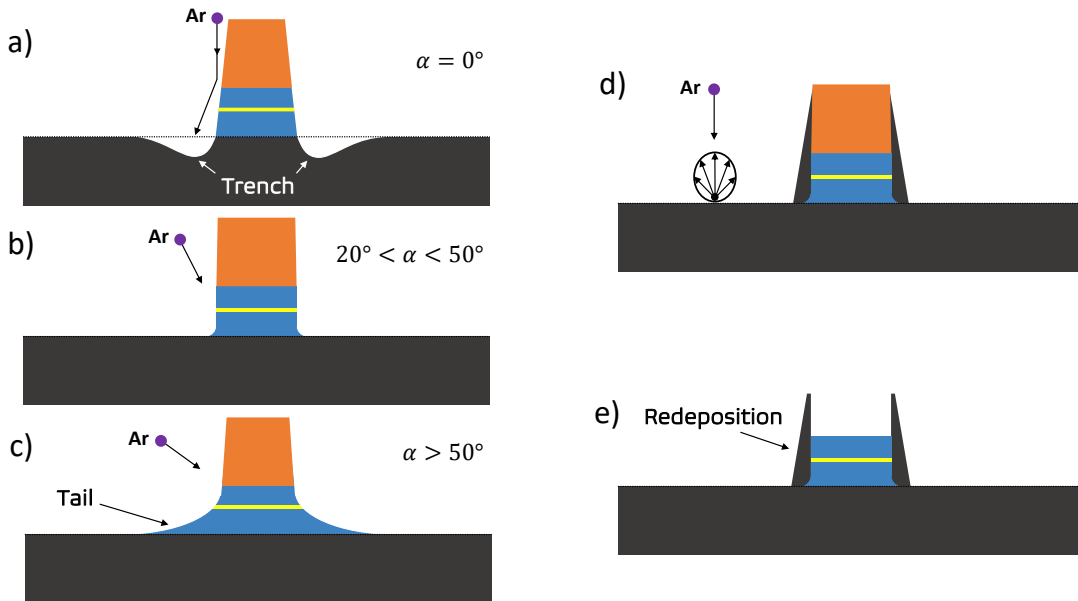


Figure 3.14: a) Argon ion etching under 0° leads to trenches being formed at the edges. b) A low-angle etch with sharp sidewalls. c) A high-angle etch results in long tails. d) Redeposited material at the tunnel barrier, which is also present after lift-off e).

Secondary ion mass spectrometry

Secondary ion mass spectrometry (SIMS) [202] is a surface-sensitive analysis technique, that enables the detection of chemical elements when subjected to an ion

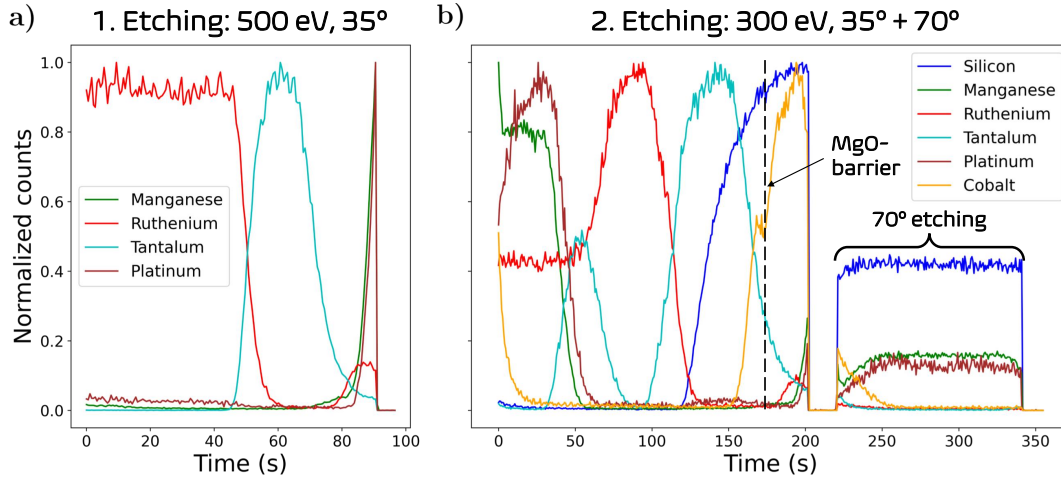


Figure 3.15: Normalized SIMS detector signal of the etching process of the MTJ stack. a) 1. Etch step to structure the bottom electrode of the device. The etching is stopped at the beginning of the PtMn layer. b) 2. Etch step to structure nanopillars. Etching is stopped at the PtMn layer. An additional etch step at 70° for two minutes is added, as indicated in the figure. The time at which the tunnel barrier is etched through is indicated by a dashed line.

beam. Isotopes and their ionized states of various elements are detected by a quadrupole mass spectrometer. This approach allows precise layer-specific etching within the TMR stack while simultaneously monitoring the etching process. Figure 3.15 shows the normalized signal of the SIMS detector of the following elements: Mn, Ru, Ta, Pt, Si, Co.

In Figure 3.15a the etching process employed for structuring the device’s bottom electrode is highlighted. Etching is terminated at the initial layer of PtMn. This, combined with subsequent etching (nanopillar etch), leads to a total etch depth down to the insulation substrate (SiO_2), thus defining the bottom electrode. The purpose of the second etch process is to define the MTJ nanopillar. After the patterning of a nanopillar etch mask by EBL, the etching is carried out down to the PtMn layer, guaranteeing the complete etching through the tunneling barrier layer MgO, as indicated in Figure 3.15b. A follow-up 70° etch for two minutes partially removes redeposited material, trims the nanopillars and enhances surface smoothness. After this procedure, a passivation layer has to be sputtered onto the sample to protect the device electrically as well as chemically.

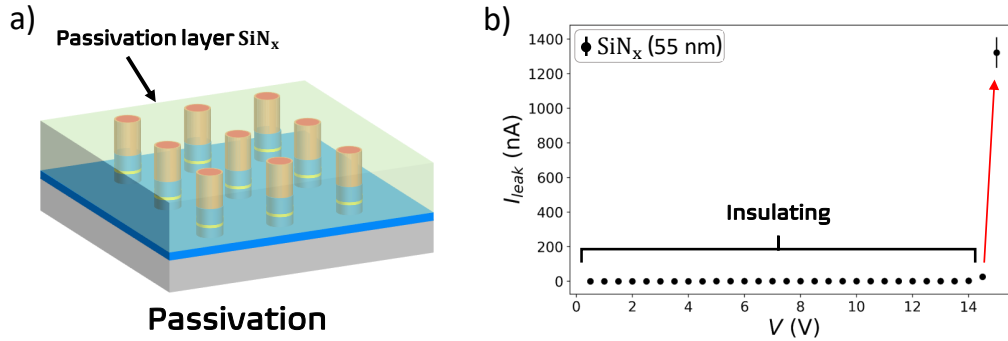


Figure 3.16: a) Passivated MTJ nanopillars covered with a SiN_x passivation layer. b) Measurement of the leak current through 55 nm sputtered SiN_x thin film. The breakdown voltage is reached for 15 V.

3.3.6 Passivation

The main purpose of passivation is the chemical as well as electrical protection of the tunnel barrier from its environment. A highly-insulating material with a large band gap alongside chemically inert or corrosion-resistant properties is favored. In addition, mechanical robustness is important for further processing steps, especially for the “pillar opening”. The silicon nitride Si₃N₄ was chosen for its excellent insulating properties, its large bandgap [203], high breakdown voltage [204], thermal stability, chemical inertness and mechanical strength [205]. The bandgap for sputtered thin films is of the order of 5 eV [203]. SiN films can reach very high hardness values of 9 or even greater on Mohs scale [206]. In order to obtain highly insulating silicon nitride thin films, a mixture of argon and nitrogen gas was chosen for the sputtering process. It has been shown that a 50/50 gas mixture leads to the highest resistivity of around $3 \cdot 10^{14} \Omega\text{cm}$ in sputtered films [207]. The relative permittivity is observed to be approximately 6 for SiN_x thin films [207]. In addition, the surface is often very smooth depending on the sputtering parameters [208], which also can indicate a low density of pinholes for amorphous SiN_x [209]. Pinholes are undesired electrical connections through the insulator and need to be avoided for MTJ nanopillar devices. The breakdown voltage for SiN_x is of the order of 2 – 10 MV/cm and for our samples found to be approximately 3 MV/cm, as shown in Figure 3.16. As an alternative, TaO_x has similar insulating properties and has been tested as well. TaO_x insulator sputtered in the *Singulus Rotaris* deposition system under additional oxygen flow was tested for the nanopillar fabrication, however, due to poorer breakdown voltages compared to SiN_x it was not considered any further. SiN_x is sputtered at 50 W at 0.05 mbar (5 Pa) at 13.56 MHz (used power supply: *AJA-100/300*) from a silicon

nitride target (Si_3N_4). The nitrogen and argon flow was adjusted to 77.7 sccm. The whole procedure can be found in the Appendix A.1. Following the deposition of approximately 55 nm of amorphous silicon nitride through sputtering, the entire sample becomes uniformly covered by this insulator material, as shown in Figure 3.16a. As a result, the structured nanopillars can not be measured electrically. To overcome this limitation, the SiN_x film, along with the etch resist, needs to be removed above the MTJ pillar. This process is referred to as “nanopillar opening”.

3.3.7 Nanopillar opening

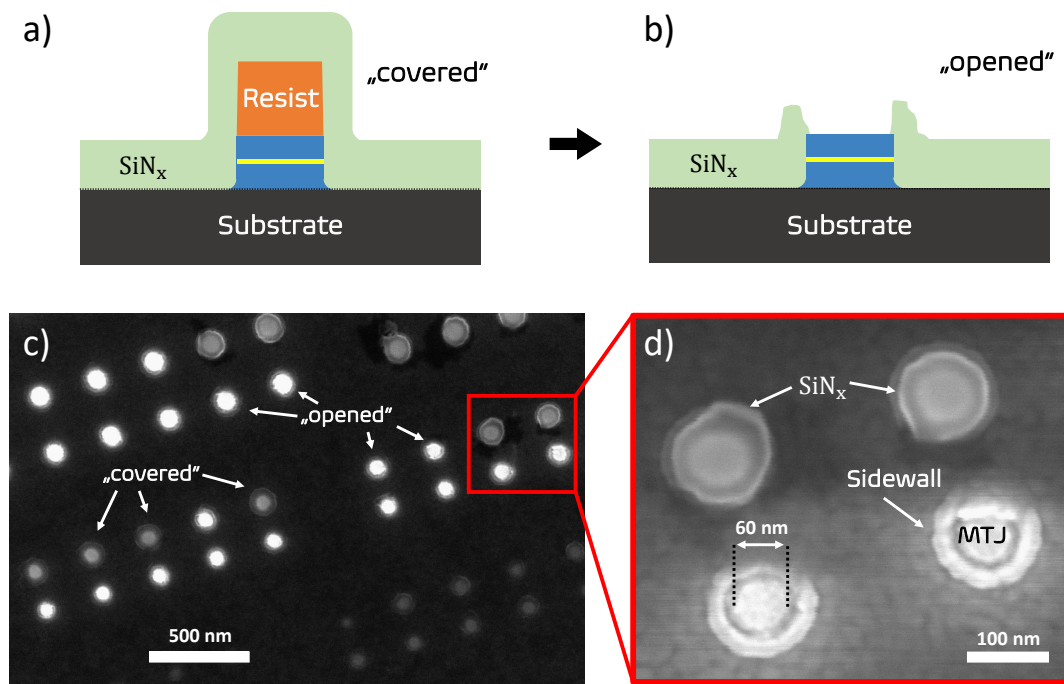


Figure 3.17: a) Illustration of a structured nanopillar, which is covered by the SiN_x passivation layer. b) The “opened” nanopillar is freed of resist and SiN_x above the pillar, thereby enabling an electrical connection to the MTJ cap in the next processing step. c) SEM image of “covered” (dark spots) and “opened” (bright spots) nanopillars. d) Zoomed SEM image of four nanopillars.

After the deposition of the passivation layer, the MTJ nanopillars (+ etch mask) are covered by the insulating material. In order to contact a nanopillar from the top, the passivation material together with the etch-resist above the MTJ has to be removed without destroying neither the tunnel junction nor the surround-

ing passivation material. Here, this process is called “nanopillar opening”. For this processing step (also called “planarization”), various methods are applied in the integrated circuits manufacturing industry or in academic research. These methods include etch-back [210, 211], chemical mechanical polishing or chemical mechanical planarization (CMP) [212, 213]. The goal of these methods is to create a highly planar and smooth surface. The “Q-tip method” can be regarded as the poor-man’s CMP method. Here, a Q-tip in an acetone bath is manually pressed onto the sample surface and moved manually over the surface in order to remove texture inhomogeneities and thereby remove the insulating capping of the nanopillar MTJs. This method typically is carried out for a few minutes and the results are checked on larger reference pillars under the optical microscope. In the following step, electrode leads (“top electrode”) have to be patterned in order to establish an electrical connection with the MTJ nanopillars.

3.3.8 Top electrode patterning

The top electrode can either be structured by EBL or by optical lithography, due to the micrometer size of these structures. The top electrode has to cover the MTJ nanopillar and has to facilitate electrical contact with the capping layer of the MTJ. Various optical masks for the structuring of the top electrode are depicted in Figure 3.9. After the lithography process of a defined top electrode structure, a bilayer of 5 nm chromium and 60 nm gold is sputtered onto the sample. By means of the lift-off method, the desired metal structure is obtained as a conductive top electrode.

3.4 Electrical measurements

In order to measure the TMR ratio of a magnetic tunnel junction, a voltage is applied between the top and the bottom electrode and the current is measured for the parallel and antiparallel state configuration. Typically, the bottom electrode is set to potential ground and the applied voltage is on the order of a few millivolts. The nanopillars are patterned in size such that the mean resistance is in the order of a few $k\Omega$ for a given RA product. Thereby, it is easy to distinguish between a short circuit at the tunnel barrier and a non-contacted MTJ nanopillar. A first characterization of a sample with several hundreds of nanopillars is carried out by a probe station. Individual MTJ nanopillars or interconnected MTJs are then further characterized by time series measurements.

3.4.1 Probe station

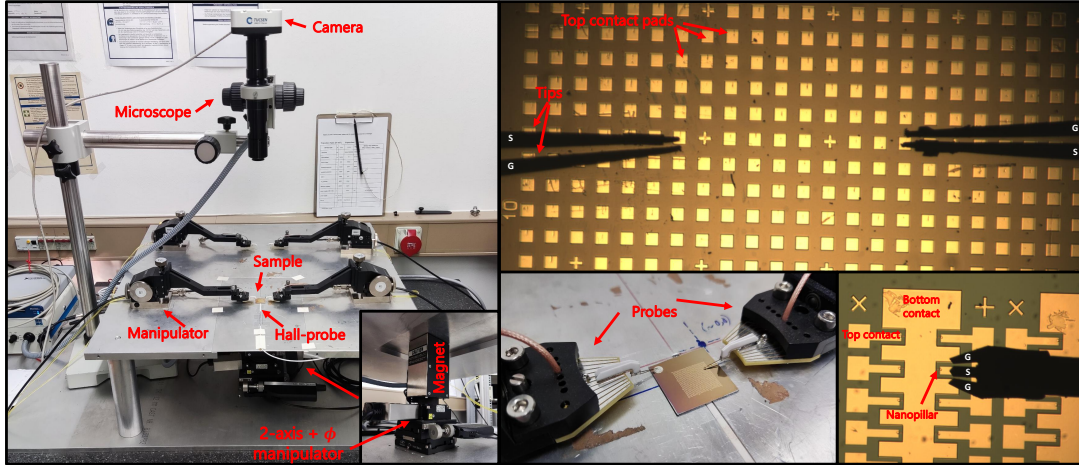


Figure 3.18: Nanopillar measurement setup of the probe station and microscope images of the measured sample. A water-cooled electromagnet below the sample provides an in-plane magnetic field, which can be aligned in azimuth direction ϕ . The micromanipulators contain two (G,S) or three (G,S,G) tips for dc- or hf-measurements.

As depicted in Figure 3.18, the mechanical probe station provides two or more manipulators, allowing for the precise positioning of thin needles (tips) onto the gold pads (top electrodes of the MTJs) in order to establish an electrical contact. The manipulators allow for a 3-axis control with micrometer precision. As soon as the top and bottom electrodes are in touch with the respective tips for +V & GND, a magnetic field can be applied and the change in resistance is measured. Further measurements with a lock-in amplifier or by an oscilloscope can also be carried out. It is important to note that the tunnel barrier of these MTJ nanopillars breaks for applied voltages larger than approximately 1 V.

Typically, during a measurement multiple devices are connected to the same power line, thereby these are inductively coupled. If a device with a large current is switched on or off, this yields an induced voltage propagating to all other devices, especially to the measurement device. An induced input voltage spike can then lead to an output voltage spike at the measurement device, thereby destroying the tunnel barrier at the MTJ.

3.4.2 High frequency measurements

In high-frequency measurements, the limiting factor can be the resolution limit of the measurement setup, due to parasitic capacitance of the circuit. Electrical con-

3. Experimental techniques

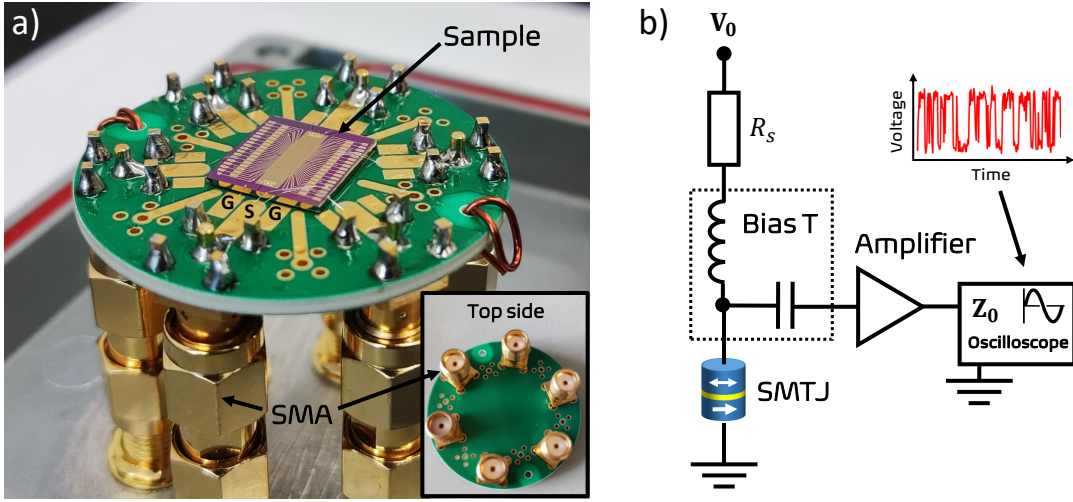


Figure 3.19: a) A picture of the rf-sample holder with six SMA connectors. Top and bottom electrodes are connected via bonded aluminum wires. The GSG design is compatible for high-frequency measurements. b) Illustration of the setup of a voltage divider circuit with an MTJ, shunt resistor, bias-T, amplifier, and oscilloscope for measurements of random telegraph noise signals.

nections are supposed to match their impedance, since impedance mismatch leads to reflections and signal modifications. All utilized cables are coaxial $50\ \Omega$ BNC- or SMA-cables. The sample is mounted on a sample holder with SMA connectors and the MTJ nanopillar electrodes are wire-bonded to the SMA connector gold pads, as depicted in the picture in Figure 3.19a. The characteristic wavelength for a 1 GHz signal is approximately $0.3\ \text{m}$ ($\lambda = c/f$). As a consequence, non-coaxial connections are supposed to be well below λ to avoid additional reflections. The limiting resolution time is determined by the RC time constant, which is given by the circuit's capacitance C and resistance R . The RC time constant is related to the 3 dB cutoff frequency f_c and defined according to: $\tau_{RC} = RC = 1/(2\pi f_c)$. An electrical signal propagating in the circuit is linked to the charging (or discharging) of the circuit's capacitance. The charging (discharging) is described by the exponential dependence: $\exp(-t/\tau_{RC})$. The time τ_{RC} can be considered as the time to charge the capacitance by $1 - 1/e$. As a consequence, AC signals with periods shorter than τ_{RC} or frequencies higher f_c can not be resolved accurately. Nonetheless, τ_{RC} is also proportional to resistance, which can often be reduced partially at the expense of signal amplitude. The signal, originating from the voltage drop between R_{MTJ} and R_s (Fig. 3.19b), is further amplified by a 20 dB low-noise amplifier (*ZFL-1000LN+* from *mini-circuits*). The stochastic fluctuation signal is acquired by an oscilloscope (*Tektronix DPO7543*).

3.5 Simulation

3.5.1 SPICE

SPICE (Simulation Program with Integrated Circuit Emphasis) is a software program for simulations of electrical circuits. *LTspice* [214] is a free simulator version from *Analog Devices*, which has been used in this work to simulate various circuits. In general, SPICE solves non-linear differential-algebraic equations by finding approximate solutions of simulated circuits and is an industry standard to verify circuits before going into manufacturing. A circuit typically encompasses several electrical components (e.g., diodes, field-effect-transistors, capacitors, etc.), each described by a model, based on physical characteristics. However, a new component can also be designed and defined manually, allowing for wide possibility of different circuit simulations. In addition, time-resolved or frequency-resolved simulations are possible. SPICE is a useful tool to verify and simulate circuits, in which MTJs are connected with CMOS logic.

3. Experimental techniques

4 | Results

This chapter comprises experimental and simulated results of various TMR stacks and patterned MTJs. Optimization and characterization measurements of the TMR stack are presented, and the analysis of superparamagnetic stochastic switching of nanometer-sized MTJs along with the discussion of electrical coupling of multiple stochastic MTJs is elaborated. Finally, prospects for neuromorphic computing based on superparamagnetic MTJs are presented.

4.1 TMR characterization and optimization

Today's TMR stacks are often optimized for a specific application and therefore comprise multiple magnetic and non-magnetic layers to achieve a desired magnetoresistive response. In the following, the TMR stack design is described and the properties and purpose of individual layers are outlined.

4.1.1 TMR stack design

The key components of a TMR stack, used in this work, consist of two cobalt-iron-boron ($\text{Co}_{40}\text{Fe}_{40}\text{B}_{20}$) ferromagnetic layers separated by a thin insulating magnesium-oxide (MgO) barrier, as shown in Figure 4.1a. When conducting a hysteresis measurement in which the tunnel magnetoresistance is measured as a function of an in-plane magnetic field, starting with the minimum field value, two distinct switching events are observed in the positive as well as negative field range (under neglect of stray fields). These switching events arise due to the differing coercivities of the free and reference ferromagnetic (FM) layers [215]. This difference in coercive fields can stem from variations in the FM layer thickness [215, 216] or distinctions in the composition of the ferromagnetic material. In order to pin the reference layer into a specific direction, an antiferromagnetic layer

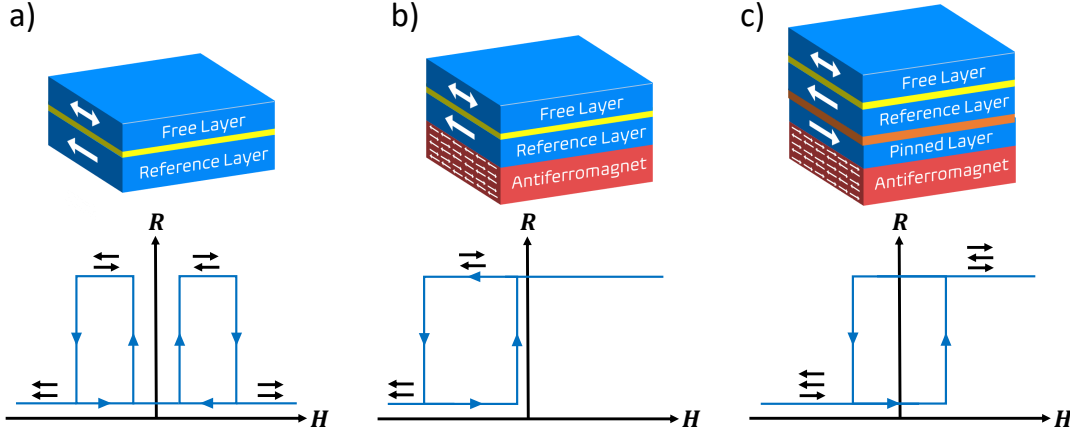


Figure 4.1: a) Illustration of a magnetic tunnel junction with two ferromagnetic layers (blue) separated by a thin insulating layer (yellow). The respective magnetoresistance as a function of an in-plane field sweep (starting with the minimum field value) is shown below. The arrows indicate the direction of the two FM layers. b) Illustration of an exchange biased tunnel junction with respective hysteresis curve. c) Illustration of an exchange biased tunnel junction with an incorporated synthetic antiferromagnet and the respective hysteresis curve.

(PtMn or IrMn) is placed in contact to the reference layer, as depicted in Figure 4.1b. This introduces a unidirectional anisotropy, called exchange bias [59, 217]. Here, a (horizontal) shift of the hysteresis occurs, which is not due to the exchange bias field but rather due to the stray field, which arises from the reference layer and is present at the free layer site. To compensate this stray field, a synthetic antiferromagnet (SAF) can be introduced, which consists of two ferromagnets (“pinned layer” + “reference layer”) separated by a thin conductive spacer layer, such as ruthenium. Under an appropriate spacer layer thickness, both FM layers can exhibit a strong antiferromagnetic coupling (see also interlayer coupling in Sec. 2.2.3). The layer next to the antiferromagnet (AFM) is called “pinned layer”, while the coupled FM layer is called “reference layer”. By adjusting the FM layer thickness, it is possible to compensate the stray field at the “free layer” site to almost 0, which is indicated by the centered hysteresis curve in Figure 4.1c. The relative resistance change from the parallel to the antiparallel alignment of the free and reference layer is defined as the TMR ratio according to: $TMR = R_{ap}/R_p - 1$.

In the following, each component of the full TMR stack employed in this work is briefly described, along with the compositions and characteristics of the respective layers.

Substrate

As a substrate, 4 inch, 1-side polished, p-type (boron) doped silicon wafers of 525 μm thickness and 1.5 μm oxide layer have been used [218]. The surface is polished and enables a smooth growth of the TMR stack. The 1.5 μm oxide layer is thick enough to resist the stress induced by wire bonding and thereby insulates the TMR stack from the doped silicon semiconductor wafer material.

Seed layer

A seed layer is important in order to guarantee a smooth growth and well-defined interfaces. A smooth and crystalline interface at the tunnel barrier site is key to accomplish a high magnetoresistance ratio. For this purpose, tantalum (Ta, bcc crystal growth [219]) in combination with ruthenium (Ru, hcp crystal growth [220]) has been chosen as a seed. A multilayer combination of Ta and Ru increases the total thickness, thereby enhancing the conductivity of the bottom electrode. Additionally, this approach prevents columnar growth [221] of large grains, which would lead to rougher interfaces and potential deterioration of the TMR ratio.

Antiferromagnet

The antiferromagnetic layer of the TMR stack is required in order to introduce exchange bias [217]. Exchange bias coupling leads to a unidirectional anisotropy in the adjacent ferromagnet (see Sec. 2.2.2), thereby pinning its magnetization direction. This pinning is necessary in order to define a reference magnetization direction in the device. The effect of exchange bias is studied in this work, and it is shown that the coupling field can exceed 300 mT for a bilayer structure of platinum-manganese and cobalt-iron ($\text{Pt}_{38}\text{Mn}_{62}/\text{Co}_{70}\text{Fe}_{30}$) [182].

Synthetic antiferromagnet

Since each ferromagnetic layer leads to stray fields and thereby affects the magnetization of all other layers of the stack, especially the free layer, it is beneficial to compensate for these stray fields by introducing a synthetic antiferromagnet (SAF) [222]. Here, the SAF is built by the combination of $\text{Co}_{70}\text{Fe}_{30}/\text{Ru}/\text{Co}_{40}\text{Fe}_{40}\text{B}_{20}$.

Tunnel barrier

The most sensible part of a magnetic tunnel junction is its tunnel barrier. Magnesium oxide (MgO) is used as a tunnel barrier in combination with $\text{Co}_{40}\text{Fe}_{40}\text{B}_{20}$ magnetic layers in a structure of: $\text{Co}_{40}\text{Fe}_{40}\text{B}_{20}/\text{MgO}/\text{Co}_{40}\text{Fe}_{40}\text{B}_{20}$. In one of the CoFeB layers the magnetization direction is pinned, therefore this layer is called “pinned layer”, while the second CoFeB layer can easily change its magnetization direction and is called “free layer”. If the MTJ incorporates a SAF structure, the ferromagnetic layer adjacent to the tunnel barrier is referred to as the reference layer, while the ferromagnetic layer next to the antiferromagnetic layer is designated as the pinned layer.

Magnetic free layer

The magnetic free layer is able to freely rotate its magnetization under relatively small external applied magnetic fields or for nanometer-sized tunnel junctions due to thermal excitations. Due to the uniaxial anisotropy (crystalline and/or shape anisotropy) in the free layer, its magnetization mostly aligns with the effective easy-axis direction, either in parallel or antiparallel configuration. The alloy used for the free layer is $\text{Co}_{40}\text{Fe}_{40}\text{B}_{20}$.

Capping

The top electrode ensures the protection of the tunnel junction from the environment, especially from oxidation. In addition, the top electrode has to be conductive in order to be able to characterize the stack by current in-plane tunneling (CIPT) measurements [161] and to be able to establish an electrical connection. For the capping, a bilayer of Ta/Ru is employed. The Ta layer, which is in contact to the magnetic free layer, absorbs boron during the annealing process and thereby promotes CoFe(B) crystallization [80, 223, 224].

4.1.2 Magnetization hysteresis

The TMR stack comprises multiple magnetic and non-magnetic layers, each serving distinct functionality to culminate in the desired attributes of the final MTJ device. In Figure 4.2a the TMR stack is illustrated, which primarily encompasses the CoFeB/MgO/CoFeB tunnel junction, a synthetic antiferromagnet (SAF) CoFeB/Ru/CoFe and an antiferromagnet PtMn to exert exchange bias onto the adjacent CoFe ferromagnet. The TMR stack has been deposited by a *Singulus Rotaris* deposition system at room temperature on a

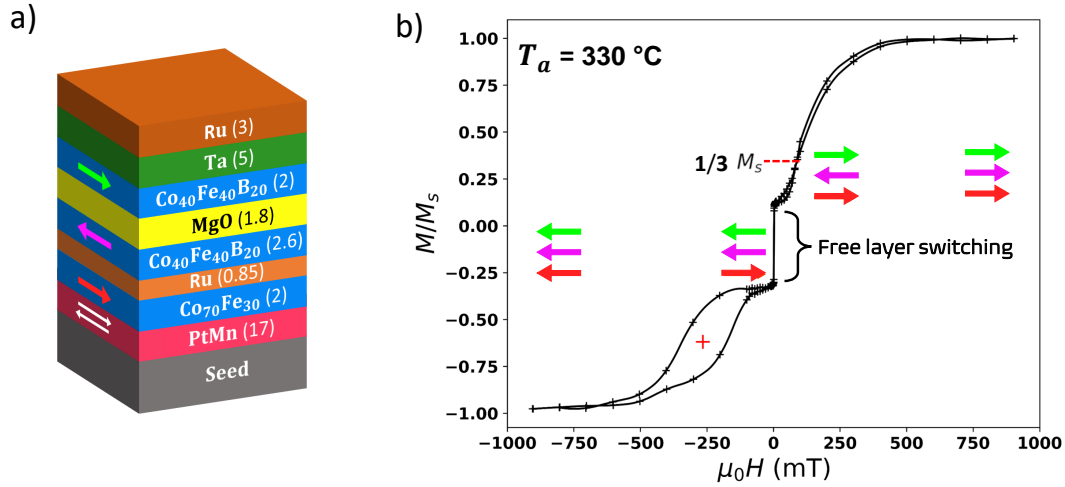


Figure 4.2: a) An illustration of the TMR stack, which comprises ferromagnetic layers (blue) with a respective magnetization orientation indicated with an arrow for the free layer (green), reference layer (violet) and pinned layer (red). b) Hysteresis loop measured in-plane with a SQUID magnetometer for a stack annealed at 330 °C. The colored arrows highlight the magnetization orientation of each ferromagnet. The exchange bias field H_{ex} is indicated as a red cross and the stability of the reference layer (in the SAF) H_{saf} is determined at $1/3$ of M/M_s as indicated by the red dashed line. (Modified figure reprinted with permission from [182] © 2023 IEEE.)

Si/SiO₂ substrate using rf- and dc-magnetron sputtering. CoFeB, MgO, PtMn, CoFe, Ta, and Ru have been sputtered with argon at room temperature at a base pressure of approximately $5 \cdot 10^{-5}$ Pa. The full stack composition with the respective thickness in nanometer is: Ta(5)/Ru(5)/Ta(5)/Ru(5)/Pt₃₈Mn₆₂(17)/Co₇₀Fe₃₀(2)/Ru(0.85)/Co₄₀Fe₄₀B₂₀(2.6)/MgO(1.8)/Co₄₀Fe₄₀B₂₀(2)/Ta(5)/Ru(3). The seed is composed of a multilayer Ta(5)/Ru(5)/Ta(5)/Ru(5), promoting a conductive bottom electrode and providing a smooth surface for the following layers. The SAF interlayer exchange coupling strength strongly depends on the spacer quality and thickness, as described in the theory chapter (Sec. 2.2.3). The SAF structure in the stack is optimized to compensate for stray fields, thereby centering the hysteresis close to 0 as it can be seen from Figure 4.2b.

Figure 4.2b depicts a SQUID measurement of the normalized magnetic moment of the TMR stack annealed at 330 °C. The external field is applied in the plane along the easy-axis direction. A saturation of magnetization is found for approximately 750 mT. For very low absolute fields (< 3 mT) the free layer aligns along the external field direction, thereby setting the MTJ state. When subjected to applied fields below -100 mT, the magnetization direction of the pinned layer ini-

tiates a switching, leading to a small hysteresis loop. The mean of the points of inflection derived from two Langevin fits applied to this loop defines the exchange bias field, denoted as H_{ex} and is indicated as a red cross in Figure 4.2b. This value is a measure of the pinned layer's pinning strength.

For positive applied fields, a gradual increase of magnetic moment is observed, which is due to the reorientation of the reference layer. The coupling strength of the reference layer in the SAF configuration here is defined as the positive magnetic field value at which the combined magnetization reaches 1/3 of the total magnetization, as depicted in Figure 4.2b. The selection of the 1/3 fraction is designed to enable enhanced sensitivity to detect changes in the hysteresis curve across various annealing temperatures, which will be affected by the strength of the interlayer exchange coupling. For sensor applications such as angle sensors, the stability of the SAF structure as well as the pinning strength of the pinned layer via exchange bias is of utmost importance since a weak pinning of the reference layer results in a (field strength dependent) deviation of the magnetoresistance from the ideal value. For this reason optimization of not only TMR and RA product is important but also of exchange bias and SAF coupling. Since all of these parameters are strongly dependent on the annealing temperature, we will delve deeper into the impact of annealing in the subsequent analysis.

4.1.3 Effect of annealing temperature

The primary parameters crucial for characterization typically encompass the TMR ratio and the RA product. The effect of annealing temperature on the TMR stack, in the range from 190 °C to 370 °C, has been investigated and the effect on TMR ratio, RA and interlayer coupling analyzed. Annealing was performed in an in-plane magnetic field of 350 mT for 3 h at a vacuum pressure of approximately 10^{-5} Pa for a specific temperature. Various annealing temperatures were investigated in order to observe a maximum MTJ performance as well as strong performance deterioration that might occur at low and high annealing temperatures. TMR and RA values were determined by means of current in-plane tunneling (CIPT) measurements of unpatterned TMR stacks, as elucidated in Section 3.1.2. Results of these measurements are highlighted in Figure 4.3a and 4.3b. For potential applications, not only TMR and RA values are important, but equally the (field) stability of the MTJ states. In general, the state stability is also strongly dependent on other layers in the stack, such as the antiferromagnet, which is responsible for the exchange bias. In order to access the information of stability, namely the coupling strength of exchange bias H_{ex} and interlayer coupling of the synthetic antiferromagnetic H_{saf} , SQUID hystere-

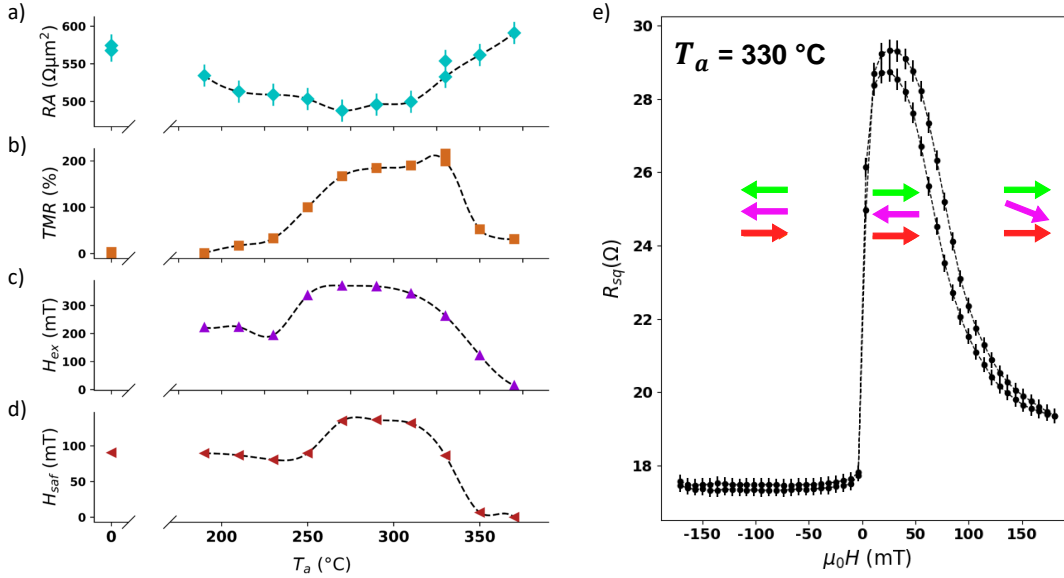


Figure 4.3: a) The resistance area (RA) product, b) tunnel magnetoresistance ratio (TMR), c) exchange bias field (H_{ex}), and d) SAF bias field H_{saf} are outlined as functions of the annealing temperature. A dashed line (cubic spline interpolation) indicates the respective trends. e) CIPT measurement of the sheet resistance R_{sq} of the 330°C -annealed TMR stack. The magnetization directions for free, reference and pinned layer are indicated by the colored arrows. (Reprinted, with permission, from [182] © 2023 IEEE.)

sis measurements have been performed for the TMR stacks, annealed at different temperatures.

Given the TMR temperature dependence as outlined in Figure 4.3b, it can be observed that the maximum TMR ratio is achieved for 330°C and a relatively high TMR ratio of approximately 200% is found in the range of 270°C to 310°C . In this temperature range, also the RA product is at its minimum, with the minimum value at 270°C and showing an approximately inverse relation with TMR ratio. A similar trend in RA product and TMR as a function of annealing temperature is also observed for out-of-plane MTJs [225]. For higher temperatures, the increase in RA can be explained by a possible formation of boron oxide (diffusion of B into MgO) at the MgO interface, which would deteriorate the coherent tunneling process. In addition, the effective barrier thickness might increase beyond 330°C , due to the migration of Mn atoms towards the barrier [226], which could lead to an increase in RA for higher temperatures. For annealing temperatures below 270°C , the decline in RA is predominantly attributed to the poor crystallization of the amorphous CoFeB layer [227]. This crystallization process results

4. Results

in a more well-defined MgO/CoFeB interface, ultimately leading to higher TMR ratios [228]. At 300 °C, the low RA value coincides with a higher TMR ratio. As the annealing temperature rises beyond this point, the accumulation of Mn atoms within the MgO barrier becomes pronounced [229], subsequently triggering a swift degradation of the TMR, as depicted in Figure 4.3b. This interpretation gains further support from TMR-temperature dependencies reported by other research groups [227, 229, 230, 231]. These observations underscore the pivotal role of the annealing temperature in shaping the quality of the MTJ and, consequently, the TMR ratio. A comparable trend has also been identified for MTJs featuring an Al₂O₃ barrier. In this case, the TMR ratio experiences a gradual enhancement with increasing annealing temperatures, up to 300 °C, followed by a rapid deterioration beyond this point [226], indicating that this effect does not depend on the barrier material type.

Figure 4.3c presents the trend of exchange bias H_{ex} with annealing temperature, determined from SQUID hysteresis loops (see also Fig. 4.2). Exchange bias is caused by the pinning antiferromagnetic (PtMn) onto the adjacent ferromagnet (CoFe) and exhibits a maximum for an annealing temperature of around 270 °C. For higher temperatures, the biasing effect declines due to enhanced diffusion of Ru atoms towards the AFM and FM layers. The migration of Mn atoms from the PtMn AFM layer towards the MgO barrier [229] results in a diminished strength of AFM pinning [61] and a decrease in the effective magnetic thickness of the CoFe layer [232]. Additionally, a reduction in the exchange bias induced by the AFM PtMn has been demonstrated, particularly at elevated temperatures above 300 °C, at which the blocking temperature distribution of PtMn exhibits a peak maximum [233]. This agrees well with the observed sharp decline in H_{ex} above 300 °C, as shown in Figure 4.3b.

The coupling strength of the SAF is accessed by H_{saf} via SQUID measurements and is outlined in Figure 4.3d. H_{saf} indicates at which magnetic field strength the reference layer will switch its magnetization orientation. Similar to H_{ex} , H_{saf} exhibits a “plateau” in the range of 270-310 °C, and then declines for higher temperatures, which is mainly caused by the diffusion of Mn, Ru and B atoms and by the formation of pinholes in the Ru spacer, resulting in ferromagnetic coupling of the reference and the pinned FM layers [230]. This deterioration of the Ru spacer can lead to a parallel alignment of all FM layers even for very low fields (≈ 10 mT). For annealing temperatures below 250 °C the exchange bias weakens and coincides with an overall symmetrical hysteresis loop (see also publication [182]), unlike for temperatures above 250 °C which are highly asymmetrical. This phenomenon can be explained by considering the behavior of PtMn

layer. Initially, after room temperature deposition, PtMn possesses a paramagnetic disordered face-centered-cubic (fcc) state without exhibiting any exchange bias effect [61, 234]. However, as the annealing temperature rises, a transition from its disordered fcc state to an ordered face-centered-tetragonal (fct) phase occurs, leading to the emergence of the exchange bias effect [235]. Therefore, annealing in a magnetic field (also called “field cooling”) is necessary to promote long-range chemical ordering in the PtMn layer and to induce strong exchange bias [61, 236]. The impact of the coupling strength H_{saf} can also be observed in CIPT sheet resistance measurements, as shown in Figure 4.3e. Here, the sheet resistance R_{sq} is shown for a TMR stack annealed at 330 °C. In the negative field range up to -150 mT, the MTJ remains in the parallel configuration, thus a low and constant resistance value is measured. Transitioning into the positive field range, the free layer magnetization (green arrow) abruptly switches at low field values to the antiparallel state. For higher positive field values, the reference layer is rotated out of its initial alignment, thereby passing over into a more parallel alignment, which results in an overall lower resistance R_{sq} . The reference layer stability observed from this CIPT measurement approximately matches with the SAF coupling strength (H_{saf} see Fig. 4.3d) of about 100 mT determined from SQUID measurements.

In conclusion, this analysis highlights a clear difference in annealing temperature requirements for achieving the highest TMR (330 °C) and strongest pinning of the pinned layer by exchange bias (290 °C). Accordingly, the choice of annealing temperature allows for optimization of the MTJ stack for either sensor stability, high TMR ratio, or a balance between both aspects, as these temperature-related effects primarily stem from interdiffusion and crystallization.

4.1.4 TMR and RA product

For the fabrication of MTJ nanopillars, the TMR stack was modified primarily in order to obtain a lower RA product, consequently leading to k Ω resistance for MTJ diameters below 100 nm. The TMR ratio and the RA product were determined from TMR thin films by CIPT measurements (see also Sec. 3.1.2), before the fabrication of MTJ nanopillars. An MTJ featuring a low RA product offers the advantage of being compatible with high current densities (increased tunneling probability for thin MgO barriers), at low applied voltages, consequently leading to stronger spin-transfer torque effects for low voltages. For this reason, the MgO barrier thickness was fine-tuned to 1.1 nm resulting in an RA product of $15 \pm 1 \Omega\mu\text{m}^2$ and a TMR ratio of about 160%. Figure 4.4 demonstrates the

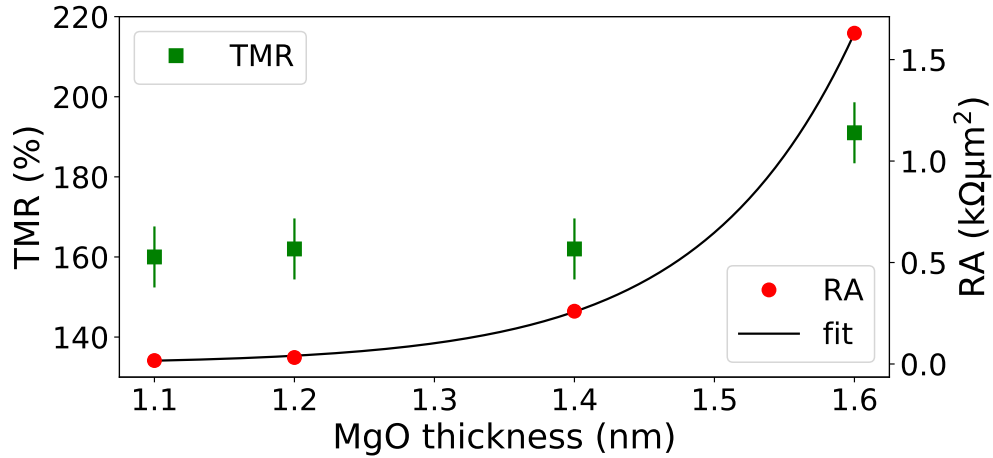


Figure 4.4: TMR and RA product as a function of MgO layer thickness. The TMR stacks were annealed at 300 °C under an applied field of 300 mT and measured via CIPTech-M200 [160]. The error bars in TMR indicate the standard deviation of the measured CIPT values. The fit to the RA product exhibits the exponential relation between tunnel barrier thickness and tunnel resistance.

exponential relation between MgO barrier thickness and RA product. The data is fitted by an exponential function of the form: $RA = c_1 \exp(c_2 t_{MgO})$, where c_1 and c_2 are fit parameters and t_{MgO} the MgO layer thickness. The TMR ratio is approximately constant for the thickness range, besides a higher TMR ratio of 191% for the MgO barrier thickness of 1.6 nm. However, it is important to highlight that a significant decline in the TMR ratio occurs when the RA product falls below the 10-15 $\Omega\mu\text{m}^2$ range [79], which is often attributed to a difference in MgO grain formation [79]. Therefore, the selection of a 1.1 nm MgO layer thickness is optimized to maintain a relatively high TMR ratio even for a low RA product, ensuring a favorable TMR-to-RA ratio.

The utilization of a thinner MgO barrier also has an impact on the free layer via stray field coupling and orange-peel coupling and therefore will lead to an offset in the hysteresis. Thus, to account for the shift in the free layer position caused by the thinner MgO barrier (from 1.6 nm to 1.1 nm), it was necessary to adjust the thicknesses of the reference and pinned FM layers in order to compensate via stray fields at the free layer site.

Figure 4.5 illustrates the variation of TMR and RA product within a 1 mm² area of the TMR stack with a 1.2 nm thick MgO barrier. Measurements are carried out with a CIPT tool (see also Sec. 3.1.2) and for a 300 °C field-annealed sample. The mean value of the TMR ratio is 161.6% and the corresponding standard variation

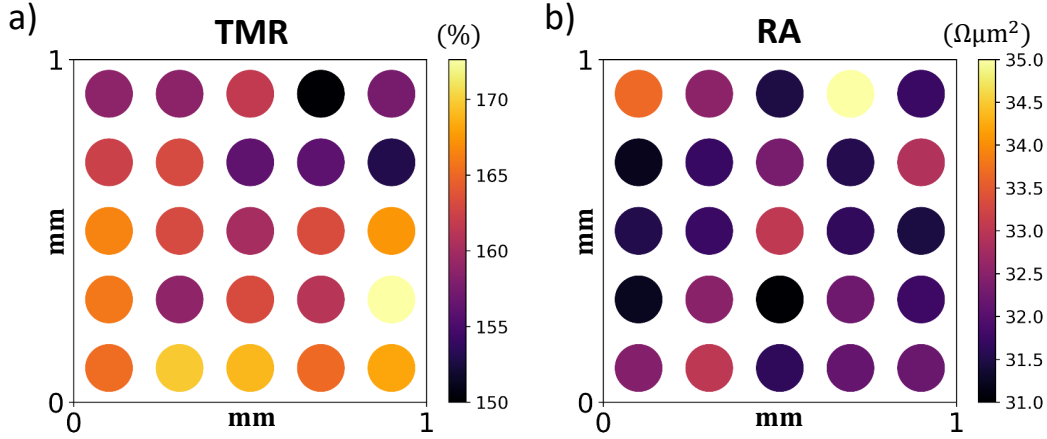


Figure 4.5: CIPT measurements of an unpatterned TMR stack with a 1.2 nm thick MgO barrier carried out with *CIPTech-M200* [160]. a) TMR measurements at various positions within a 1 mm² area of the TMR stack are visualized as a heatmap. b) Measured RA product at the same positions as in a). These variations are likely attributed to structural differences within the tunnel barrier.

is 7.6 %, whereas the mean RA product is $32.2 \Omega\mu\text{m}^2$ with a standard variation of $0.9 \Omega\mu\text{m}^2$. The variations are likely attributed to structural disparities within the tunnel barrier. For instance, these differences could arise from variations in grain structure within the MgO layer [79] or due to different crystallization of the CoFe(B) layer at the interface [223, 227]. However, as demonstrated in Figure 4.5, the fluctuations in both TMR and RA appear to be random, at least within an area of 1 mm². It is found that when measurements are taken at the edge of a wafer, the RA values will exhibit a bias towards lower values [82]. This bias results from a thinner MgO layer, which is a consequence of the angle-dependent sputter yield present during the deposition of the TMR stack. Consequently, a variation of TMR and resistance of the final nanometer-sized MTJ device is observed, as we will see in the following. However, it is unclear whether this variation stems from the variation within the TMR stack itself or if it arises from random characteristics in other processing steps, such as ion etching and EBL. In the following, we will discuss the characteristics of patterned nanopillar magnetic tunnel junctions.

4.2 Nanopillar magnetic tunnel junctions

For any device application, a TMR stack has to be patterned into a single or into multiple structures of a specific size. In this work MTJs were predominantly shaped into approximately circular pillars with lateral dimensions ranging from 50 nm to 100 nm. However, due to astigmatism or imperfections in the aperture adjustment of the EBL system, the electron beam may deviate from its intended circular shape, potentially resulting in slightly elliptical pillar structures. The development and fabrication of nanopillar MTJs is elaborated in detail in Section 3.3. Figure 4.6 presents different processing steps for the fabrication of the final MTJ device. The SEM image in Figure 4.6a shows a structured bottom electrode strip, showcasing a resist nanopillar precisely patterned at the location indicated by the arrow. In Figure 4.6b, another SEM image displays the resist nanopillar, boasting an approximate diameter of 60 nm and captured at a tilt angle of 47° . The resist provides an etch resistance in the ion etching process, thus the full TMR structure is maintained beneath the resist, while the surrounding area is etched down to the seed layer. Moving to Figure 4.6c, an SEM image of the etched bottom electrode featuring an MTJ nanopillar appearing as a white dot is shown. The strip has been etched down to the seed layer, and the entire sample has been covered by a 55 nm thick SiN_x passivation layer. Lastly, Figure 4.6d presents an optical microscope image of the final structure and highlights the sputtered gold top and bottom electrode pads, facilitating the electrical connection of the MTJ nanopillar. Measurements of this device structure can be easily performed by probing the electrode pads or by wire bonding the sample to a sample holder and thereby connecting it to a measurement instrument. In the following figure, a statistical analysis of the resistance distribution for 63 MTJ nanopillars and the associated variation in TMR is presented.

Figure 4.7 illustrates a statistical assessment of a sample with 136 patterned MTJ nanopillars. Approximately half of these nanopillars exhibit the desired TMR ratio of more than 100% and meet the desired resistance criteria. In this particular sample, the TMR stack comprises a MgO barrier of 1.2 nm, which relates to a CIPT-measured RA product of $37.5 \Omega \mu\text{m}^2$. Half of the MTJs were patterned in a circular shape with 50 nm diameter, while the other half were patterned in an ellipsoidal shape of $36 \text{ nm} \times 70 \text{ nm}$. Due to variations in the fabrication process, which may originate from EBL, resist development, ion etching and inconsistencies in the TMR stack, the resistance of MTJ pillars varies significantly between the parallel and antiparallel states for both circular and ellipsoidal nanopillars, as depicted in Figure 4.7a. In Figure 4.7b the TMR ratio as a function of P-state

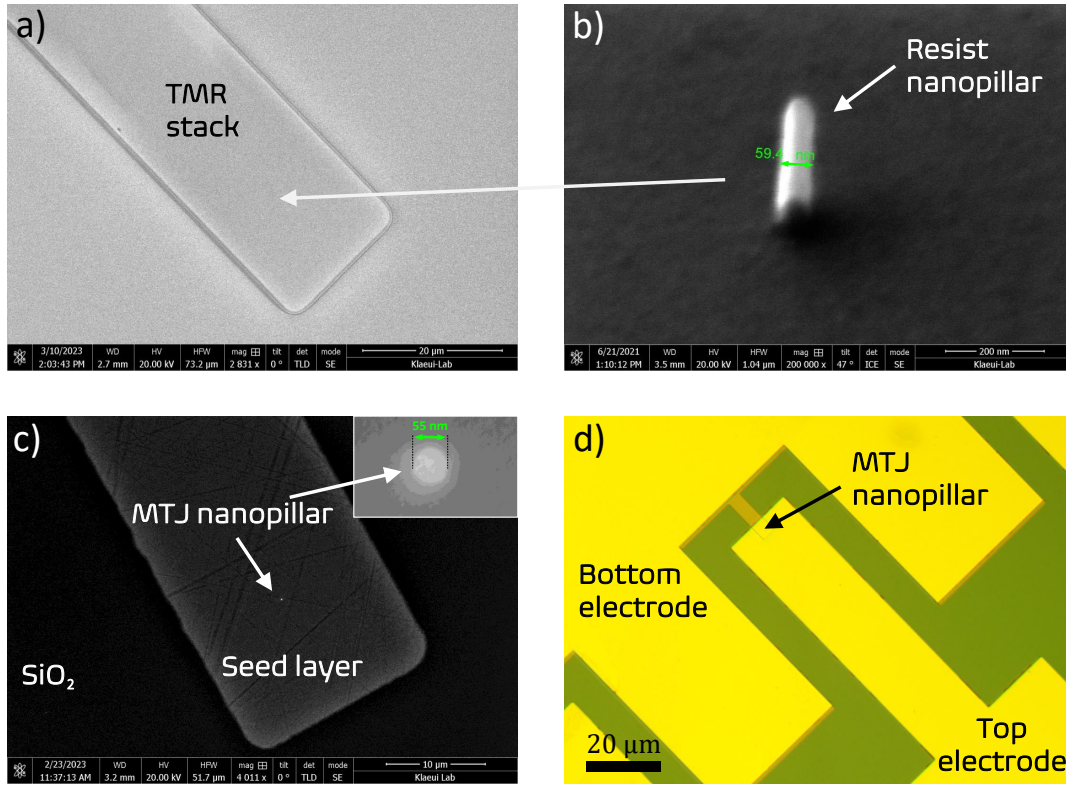


Figure 4.6: Various processing steps of the MTJ nanopillar fabrication process are outlined. a) An SEM image, taken from a structured bottom electrode strip. On this strip, a resist nanopillar is patterned at the position at which the arrow points. b) SEM image of the resist nanopillar of approximately 60 nm diameter under a tilt angle of 47°. c) After nanopillar etching, the SEM image shows the etched bottom electrode with an MTJ nanopillar (white dot). The strip is etched down to the seed layer and as a next step the whole sample is covered by a 55 nm thick SiN_x passivation layer. d) Optical microscope image of the final structure with top and bottom electrodes, connecting the MTJ pillar.

resistance is outlined. Despite the nearly identical patterned areas for circular and ellipsoidal pillars, ellipsoidal pillars exhibit larger areas due to their lower overall resistance. This phenomenon is likely a result of the electron beam exposure during the EBL, where proximity effects tend to round the ellipses into more circular shapes, thereby increasing the area. In addition, the highest TMR is observed within a specific resistance range, approximately aligning with the resistance values with the highest counts in the histogram (indicated by vertical lines in Fig. 4.7b). From the resistance and the RA product (obtained by CIPT measurements), the area of the MTJs can be determined by: $A = RA/R$. For the circular nanopillars, the patterned area closely matches the calculated area

4. Results

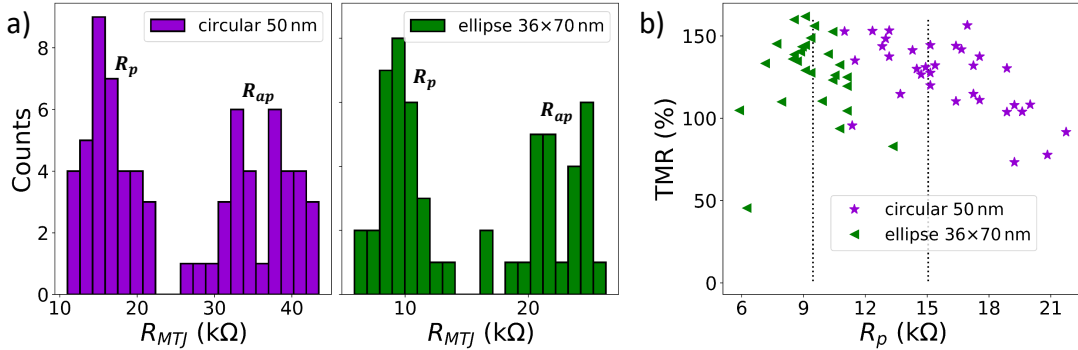


Figure 4.7: a) Distributions of the MTJ resistances for 63 circular and ellipsoidal patterned nanopillars of a sample are shown. b) The distributions of the TMR ratio for circular and ellipsoidal MTJs are plotted as a function of the low-resistance P-state. The dotted lines depict the resistances of maximum count in the histograms in a).

based on the RA product. However, for the ellipsoidal nanopillars, the calculated area is approximately 50% larger than the patterned area. As a consequence, the aspect ratio for ellipsoidal nanopillars is not maintained and the minor axis (of the ellipse) is elongated, therefore leading to lower MTJ resistances as anticipated. Next, we will study the magnetic field dependence of the MTJ resistance for nanometer-sized pillars.

4.2.1 MTJ hysteresis

In a magnetic tunnel junction, the free layer can easily be switched by applying an external magnetic field, due to its low coercivity and the absence of pinning. Figure 4.8a demonstrates three measurements of the MTJ resistance as a function of the applied field. The measurement demonstrates sharp switching and “squareness” in the hysteresis, indicating the presence of a monodomain state in the magnetic free layer. For this particular MTJ, with a diameter of approximately 80 nm, an RA product of $15 \Omega \mu\text{m}^2$ and a TMR ratio of 100%, the coercive field is found at roughly -1.9 mT and $+2.9 \text{ mT}$, indicating that the hysteresis exhibits an offset of 0.5 mT . This offset likely originates from the stray field, caused by an uncompensated magnetic moment of the reference and the pinned magnetic layers. In general, it is very difficult to fully compensate the stray field at the free layer site, thus a small offset is present for almost all measured MTJs. As we will see in the following, this stray field has a strong influence on the state-probability for superparamagnetic MTJs. In Figure 4.8b, the switching probability is determined by 184 hysteresis loops. The likelihood of a switching event primarily depends on the energy barrier, the applied field and the time spent in a given

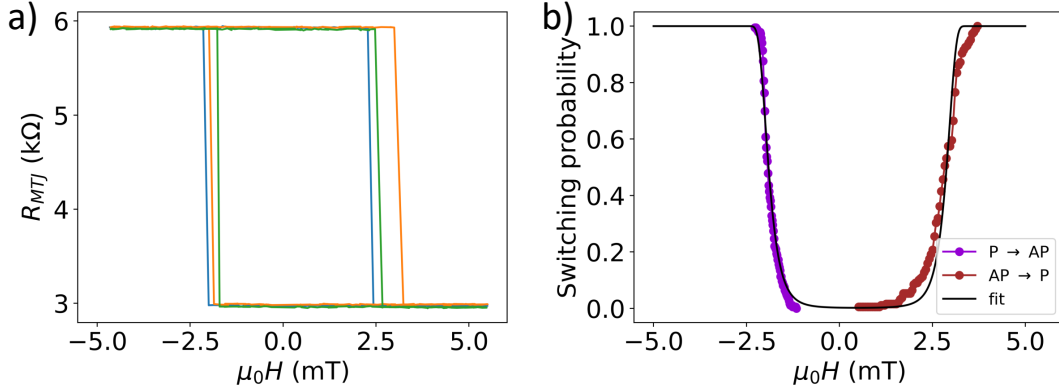


Figure 4.8: a) MTJ resistance as a function of applied magnetic field parallel to the easy-axis. As an illustration, three out of 184 loops are plotted (230 data points per loop). b) Switching probability, from the P- to the AP-state (violet) and from the AP- to the P-state (red) obtained from 184 $R_{MTJ} - H$ loops. The data is fitted according to Equation 4.1. (Data was acquired by Jonas Köhler).

state. The switching probability from parallel to anti-parallel or vice versa can be described by [237]:

$$P_{p,ap} = 1 - \exp \left(-\frac{t}{\tau_0} \exp \left[-\frac{E_b}{k_B T} \left(1 - \left(\frac{H_{\parallel} - H_{off}}{H_k^{ip}} \right)^2 \right) \right] \right), \quad (4.1)$$

where τ_0 is the attempt time, t is the time spent in a given state (measurement time), E_b the energy barrier, H_{off} the offset field, H_{\parallel} the applied magnetic field in easy-axis direction, T the temperature and H_k^{ip} the in-plane anisotropy. The fit reveals an offset field of 0.5 mT, an energy barrier of approximately $27 k_B T$ (at room temperature) and an in-plane anisotropy field of 5.1 mT.

4.2.2 MTJ angular field dependence

Here, the angular dependence of the MTJ resistance is measured as a function of the angle ϕ , defining the azimuth angle between the reference and the free layer magnetization directions. To overcome the inherent anisotropy within the free layer and ensure complete alignment of the free layer's magnetization with the applied magnetic field, a field of 90 mT was applied within the plane. As shown in Figure 4.9a, the resistance of the MTJ exhibits a characteristic cosine-like behavior. This dependence of the MTJ resistance on the azimuth angle ϕ between the free and reference layer orientations is mathematically described by

4. Results

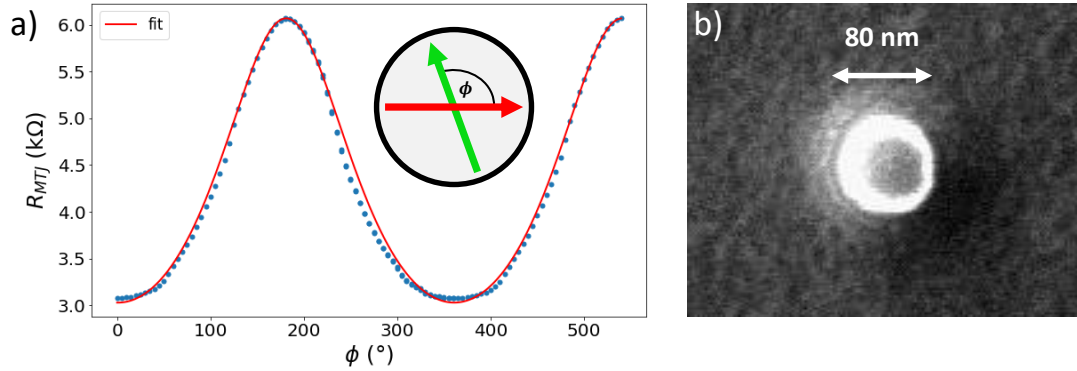


Figure 4.9: a) Measured MTJ resistance (data points) as a function of applied magnetic field direction angle. The red line is a fit applied to the data. A magnetic field of 90 mT was applied. The inset indicates the “fixed” reference layer magnetization as a red arrow, while the green arrow denotes the magnetization of the free layer, which aligns along the direction of the applied field. b) An SEM image of the etched MTJ nanopillar indicates a diameter of about 80 nm.

the following relationship [7]:

$$R(\phi) = \frac{2}{G_p + G_{ap} + (G_p - G_{ap}) \cos(\phi - \phi_0)} \quad (4.2)$$

$$= \frac{2R_{ap}}{1 + R_{ap}/R_p + TMR \cos(\phi - \phi_0)},$$

where $G_{p,ap}$ is the conductance and $R_{p,ap}$ the resistance of the P- and AP-state, TMR is the TMR ratio and ϕ_0 an angular offset between 0° field direction and the reference layer direction. The fit according to this equation is in good agreement with the experimental data, as illustrated in Figure 4.9. However, the measured resistance deviates from the theory, especially for angles in-between the parallel and antiparallel state (e.g. at 90° and 270°). This deviation arises due to the external magnetic field acting on the reference layer. It induces a torque, which is maximized when the field is applied perpendicular to the reference layer ($\tau = m_{ref} \times \mu_0 H$). Consequently, this leads to a more parallel state and thereby to an overall lower measured resistance. This effect is not present for the parallel alignment and negligible for the antiparallel alignment, since no torque is acting for $m_{ref} \parallel H$. The fitting results in a TMR ratio of 100%, a parallel resistance of 3.02 kΩ and an antiparallel resistance of 6.07 kΩ and an angular offset of 1.1° .

In general, the reference layer direction of an MTJ is not exactly known up to a few degrees. Therefore, by analyzing the angular resistance dependence of patterned MTJ nanopillars, we can precisely determine the reference layer direction relative to the magnetic field coordinate system after the patterning of

MTJ nanopillars. This analysis also serves to demonstrate the pinning strength of the reference layer within a structured nanopillar MTJ. However, for an accurate quantification of the pinning strength it is necessary to measure the resistance or magnetization as a function of an in-plane magnetic field perpendicular to the pinning direction until saturation is reached.

4.3 Superparamagnetic tunnel junctions

A superparamagnetic MTJ (SMTJ) possesses a superparamagnetic free layer, resulting in unpredictable switching between both (meta-)stable P- and AP-states. In this section, we comprehensively examine the characteristics of SMTJs, commencing with an exploration of field-dependent resistance. We then delve into the impact of spin-transfer torque, investigate nanosecond stochastic switching phenomena, explore the utility of a true random number generator based on SMTJs, and conclude by assessing the consequences of spin-transfer torque and Joule heating.

4.3.1 Superparamagnetic hysteresis

Industrial development and research often focus on the development of magnetic tunnel junctions as a memory device, called magnetic random access memory (MRAM), due to its inherent energy efficiency, non-volatility, fast read and write access times and robustness [11]. For a memory device, high thermal stability is required in order to store information reliably over a long time. On the other side, if an MTJ state is thermally unstable, meaning the magnetic free layer stochastically switches its orientation due to thermal activation, it results in superparamagnetic switching, thus the device is called superparamagnetic tunnel junction. Typically, the fluctuation/dwell times are in the range of milliseconds down to nanoseconds at room temperature, depending on the energy barrier E_b between the low energy states (parallel and antiparallel state) with respect to the thermal energy $k_B T$ [91].

The state probability and dwell times of the superparamagnetic MTJ can be affected by an external applied field [238] or by a current via spin-transfer torque [239]. In order to significantly manipulate the MTJ state by spin-transfer torque, the current going through the junction has to be high enough without applying a too high voltage, which would lead to breakdown of the barrier [240]. For this reason, the TMR stack, as shown in Figure 4.10a, was modified with a

4. Results

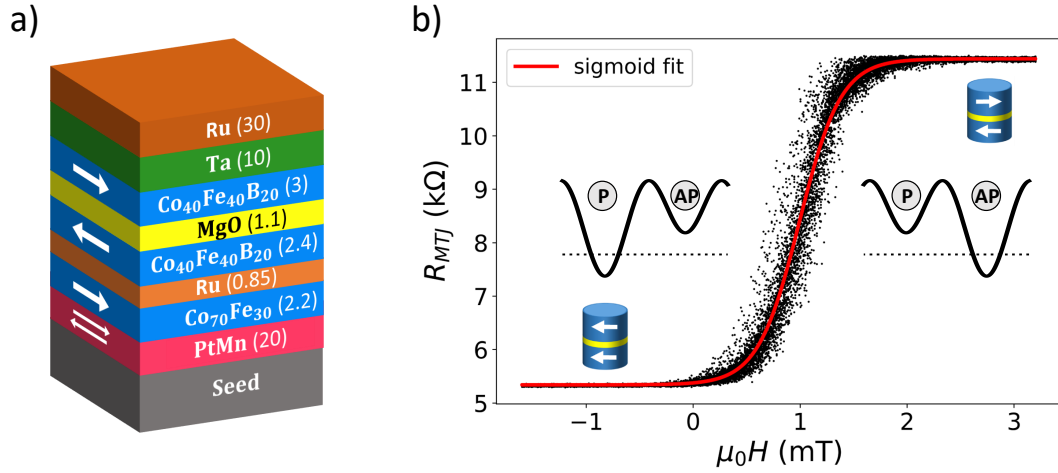


Figure 4.10: a) A TMR stack with an RA product of $15 \mu\text{m}^2$ and a TMR ratio of 150 %, used for SMTJs. b) The MTJ resistance versus the external applied easy-axis field of a SMTJ, plotted for 20 repetitions. The red line indicates the sigmoidal relation of the resistance and the magnetic field. Negative fields stabilize the parallel state, while positive fields stabilize the antiparallel state. Each data point refers to a time-averaged resistance, comprising multiple switching events.

thinner MgO barrier of 1.1 nm (compared to the stack in Fig. 4.2), which coincides with a low RA product of approximately $15 \Omega\mu\text{m}^2$, which was determined by CIPT measurements. The TMR ratio was found to be approximately 150 %. Figure 4.10a illustrates the TMR stack and with the following composition: Seed/PtMn(20)/CoFe(2.2)/Ru(0.85)/CoFeB(2.4)/MgO(1.1)/CoFeB(3.0)/Ta(10)/Ru(30). A multilayer of Ta(10)/Ru(10)/Ta(10) is used as the seed.

The superparamagnetic hysteresis for a 50 nm diameter SMTJ nanopillar is depicted in Figure 4.10b, in which the resistance is measured depending on the applied magnetic field aligned with the easy-axis direction. Notably, each data point refers to a time-averaged resistance, since the typical acquisition time of the instrument (*Keithley 2400*) is larger than the average fluctuation time: $\tau_{acq} > \tau_{SMTJ}$ (see also Sec. 2.4). This is also the reason for the relatively wide resistance variation around the inflection point ($\approx +1.1$ mT) in the plot. The sigmoid fit (red) illustrates the tunability of the SMTJ state, where a positive field stabilizes the antiparallel state while a negative field stabilizes the parallel state according to the theory. The sigmoid relation for the probability to observe the AP-state is given by (see also Sec. 2.4.4):

$$P_{ap}(H) = \frac{\tau_{ap}}{\tau_{ap} + \tau_p} = \frac{e^{4\Delta H/H_k}}{e^{4\Delta H/H_k} + 1} = \frac{1}{1 + e^{-4\Delta H/H_k}} \quad (4.3)$$

Here, $\tau_{p/ap}$ are dwell times, H is the applied in-plane field, H_k the in-plane anisotropy and Δ the energy barrier in multiples of $k_B T$. The change in probability can also be explained by a change in the energy landscape where the P- and AP-state are energetically shifted due to the additional Zeeman energy contribution, as shown in the figure inset. However, the inflection point of the sigmoid fit is not centered at 0, which indicates an “offset” of the hysteresis mainly caused by uncompensated stray fields at the free layer site. Despite the integration of a SAF within the stack, a complete compensation of the magnetic stray field at the free-layer position is hard to achieve [241]. In addition, orange-peel coupling (see Sec. 2.2.1) contributes to an offset, which is based on correlated interface roughness and present in particular for thin tunnel barriers, and favors the parallel alignment of the free layer with respect to the reference layer [53, 55, 56]. This offset field may differ between devices due to variations in the sample fabrication process. Thus, an in-plane easy-axis field up to a few millitesla is typically applied in order to compensate for this offset and to balance the SMTJ P- and AP-state probabilities.

4.3.2 Angular field dependence

Figure 4.11 illustrates (superparamagnetic) hysteresis measurements for different angular directions at room temperature. For each angle 15 hysteresis loops have been measured, and the resistance is fitted by a combination of a straight line (for the positive and negative field range) and a sigmoid function to approximate the trend. Here, the MTJ nanopillar resistance is measured as a function of an applied in-plane field. The nanopillar is of circular shape and has a diameter of approximately 60 nm. The magnetization of the reference layer (red arrow in Fig. 4.11) is aligned horizontally along the x-axis, however, the magnetization of the free layer prefers to align along the easy-axis (diagonal lines of the circular pillar in Fig. 4.11), which is found at an angle of 130° likely due to shape anisotropy. This can qualitatively be verified in the superparamagnetic resistance curve in Figure 4.11e, as for high absolute field values ($|\mu_0 H| > 3 \text{ mT}$) the resistance remains roughly constant. This is due to the free layer aligning with the external magnetic field, which is parallel to the easy-axis. However, if the field angle is higher (lower) like in Figure 4.11f (4.11d), the external field tends to align (misalign) the free layer along the reference layer direction. In general, the resulting free layer orientation for a given field can be calculated by the Stoner-Wohlfarth model for homogeneous magnetization [50]. An alignment parallel to the reference layer yields the lowest and highest observed resistances for the P- and AP-state respectively. It is worth noting, that the hysteresis curve for 25° in Fig.

4. Results

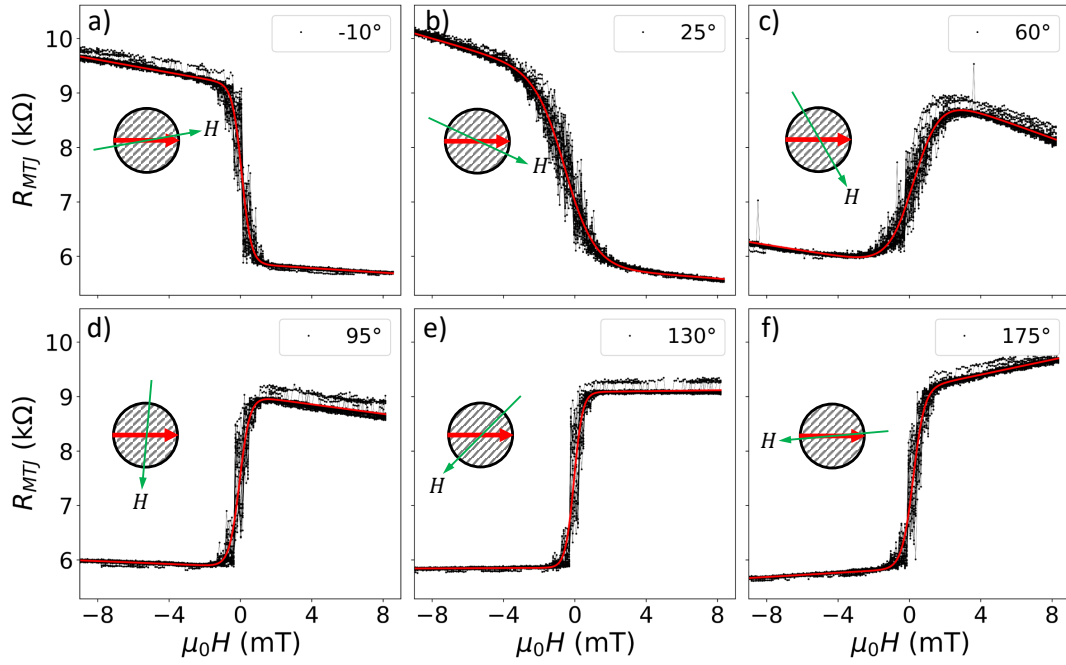


Figure 4.11: Hysteresis measurements are shown for a 60 nm MTJ for angles of -10° , 25° , 60° , 95° , 130° and 175° in a)-f). Each data point (depicted in black) is a time-averaged resistance state. The data is fitted (red curves) by a combination of a line and a sigmoid function: $f = a_{+/-} \cdot H + 1/(1 + \exp(-b \cdot H)) + c$, with $a_{+/-}$, b , c fitting parameters (a_+ for the positive and a_- for the negative field range). The top view of the circular pillar is given in the insets. Here, the magnetization direction of the reference layer is indicated by a red arrow, the applied field direction by a green arrow and the easy-axis direction of the free layer as diagonal lines. (This measurement was carried out by Jonas Köhler.)

4.11b is broadened. This is likely due to the fact that the external field is roughly perpendicularly applied to the easy-axis direction and the field component along the easy-axis is very small and only a fraction of the field amplitude, thus not stabilizing one of the MTJ states. However, as a conclusion, this particular MTJ is superparamagnetic without any applied field and has a sharp transition from the parallel to the antiparallel within a field range of approximately 4 mT.

Given Equation 4.3, an energy barrier of $\Delta = 1/2\mu_0 H_k V M_s / k_B T$ and the volume of the free layer $V = t\pi r^2$, the superparamagnetic resistance for a field applied along the easy-axis direction can be described by the following equation:

$$R_{MTJ}(H) = R_p \cdot \left(1 + \frac{TMR}{1 + e^{-2\mu_0 H M_s t\pi r^2 / k_B T}} \right), \quad (4.4)$$

where TMR is the TMR ratio, R_p the parallel MTJ resistance, V the volume

of the free layer, M_s the saturation magnetization of the free layer, t the free layer thickness, r the free layer radius, T the temperature, and k_B the Boltzmann constant. Interestingly, this equation is independent of the anisotropy of the free layer and the scaling factor of this sigmoid function is primarily determined by the magnetic moment of the free layer. Thus, by fitting the superparamagnetic hysteresis curve it is possible to determine the magnetic moment or, for a given M_s , the volume of the free layer. As the main uncertainty stems from the lateral size of the tunnel junction, this method provides a reasonable estimate of the MTJ radius. In this work, the saturation magnetization of $\text{Co}_{40}\text{Fe}_{40}\text{B}_{20}$ has been determined in SQUID measurements to be 1.07 ± 0.05 MA/m. The fit for a given superparamagnetic resistance measurement, as shown in Figure 4.11e, results in a radius of 34.7 ± 1.3 nm, which comes close but is larger than the calculated radius from the resistance area product of approximately 29 nm. As a result, this approach can be viewed as an alternative method for estimating the area of a patterned SMTJ through resistance measurements.

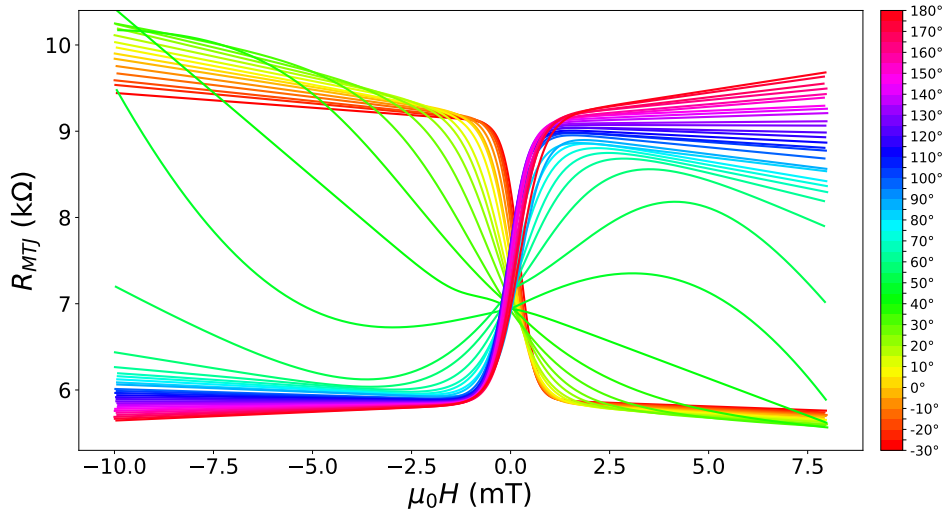


Figure 4.12: The characteristic superparamagnetic resistance trend is plotted for different applied field angles, as outlined in the colorbar. The curves are derived from fits of superparamagnetic hysteresis curves.

Figure 4.12 illustrates an overview of the angular dependent superparamagnetic resistance curve for field angles ranging from -30° to 180° . These curves are extracted from fits of the superparamagnetic hysteresis measurements, as shown in Figure 4.11. A rectangular-shaped superparamagnetic hysteresis is observed for 130° along the easy-axis direction. For angles $< 130^\circ$, the hysteresis starts tilting and at around 40° , which corresponds to an in-plane hard-axis field, the

hysteresis curve begins to “flip” and demonstrates the same symmetry for the opposite field orientation. In addition, all curves are centered around 0, therefore highlighting that this MTJ is superparamagnetic without any applied magnetic field.

4.3.3 Spin-transfer-torque

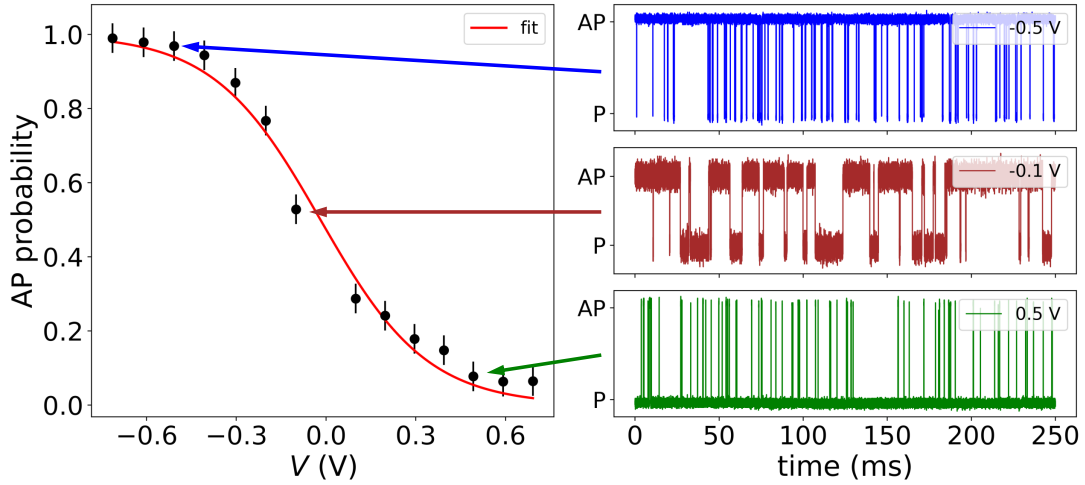


Figure 4.13: AP-state probability is plotted as a function of applied voltage. The fit in red outlines a sigmoid fit. On the right, the respective time series are shown for -0.5, -0.1 and 0.5 V.

Similarly to an external applied magnetic field, the SMTJ state can also be tuned by a current through the phenomenon known as spin-transfer-torque (STT), as described in Section 2.3.5. A spin-polarized current, exerting a torque onto the magnetic free layer, can stabilize either a P- or AP-state. Figure 4.13 demonstrates the effect of STT, where the probability of the AP-state is plotted as a function of the applied voltage. Each data point corresponds to a time series as indicated in the figure. For voltages close to 0 V, the stochastic time series represents an approximately 50/50 state, while for ± 0.5 V either the P- or AP-state is dominating. For the -0.1 V time series the noise is relatively large since the measured voltage signal is relatively low compared to the noise. In general, the dwell times for each voltage can be described by the following equation [108, 116]:

$$\tau_{p,ap} = \tau_0 e^{\Delta(1 \pm V/V_c)}, \quad (4.5)$$

where τ_0 is the attempt time, Δ the energy barrier, V is the applied voltage across the junctions, and V_c is the critical voltage for deterministic switching at 0 K [110].

The sigmoid relation, namely the state probability, is then given by:

$$P_{ap}(V) = \frac{1}{1 + e^{2\Delta V/V_c}} \quad (4.6)$$

The critical voltage for this MTJ (with $\tau_0 = 1$ ns and $\Delta = 14.2$) is determined to be 5.1 ± 1.4 V, which would exceed the breakdown voltage for MgO-based MTJs. The corresponding critical current density of 34 MA/cm² approximately matches with other reported values in literature [242]. In non-volatile MTJs, the effect of STT enables the deterministic switching of the free layer of a specific MTJ, thereby the possibility to write one bit of information. In stochastic MTJs, STT also leads to switching, however, due to inherent volatility it is observed as a manipulation of the state probability. This tunability via STT together with the effect of TMR allows for the electrical coupling of multiple SMTJs in a potential network and will be elaborated later in this work.

4.3.4 Nanosecond stochastic switching

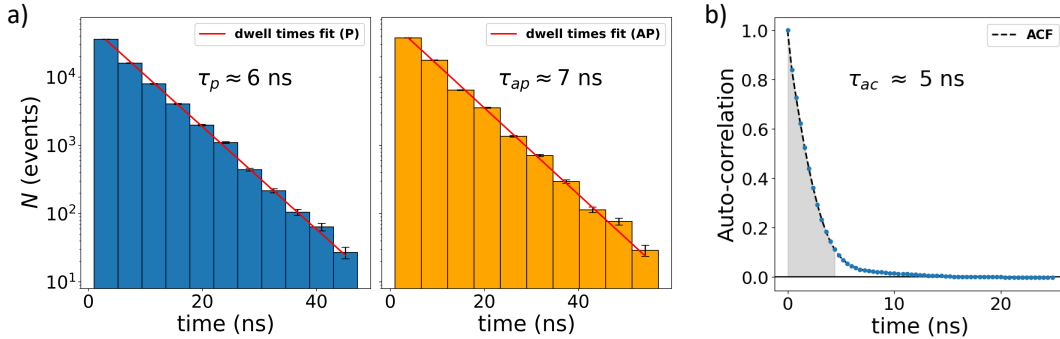


Figure 4.14: a) The histograms display dwell time distributions within a 1-ms time series for both the P and AP states. The bins represent waiting-time intervals in a Poisson process, indicating an expectation of a single switching event occurring in each interval. By fitting the data (red line), the average dwell times τ_p and τ_{ap} are determined. b) The autocorrelation function (ACF) is plotted for the time-series measurement. τ_{ac} is defined as the 99% integral under the ACF curve. (Reprinted figure with permission from [106]. Copyright 2023 by the American Physical Society.)

The following description of work and results is based on Ref. [106]. A magnetic tunnel junction, patterned in a size of 50 nm diameter, possesses a free layer magnetization, which can be regarded in the macrospin approximation as a single magnetization vector. According to the Néel-Arrhenius law, the stochastic fluctuation in an MTJ is primarily determined by the energy barrier height

4. Results

in relation to the (ambient) thermal energy, which we define as $\Delta = E_b/k_B T$. Thus, for in-plane SMTJs, the in-plane anisotropy H_k^{ip} has to be small since it determines the (in-plane) energy barrier according to: $E_b = 1/2 H_k^{ip} M_s V$. The out-of-plane anisotropy or the effective field of the free layer can be large for an in-plane switching process. As a matter of fact, a higher out-of-plane anisotropy results in enhanced fluctuation rates due to the influence on the attempt time $\tau_0 \propto 1/\alpha\gamma H_k^{eff}$ [91]. Thus, for ip-SMTJs, a large out-of-plane and a low in-plane anisotropy is desirable. A low in-plane anisotropy is achieved by circular-shaped nanopillars, where shape anisotropy energy of a disk results in an easy-plane configuration, thus H_k^{ip} of 0. However, even for circular-shaped nanopillars, uniaxial anisotropy is found after post-annealing, which is dependent on annealing time as well as temperature [243]. It promotes crystalline anisotropy by CoFe crystallization at the interface [244] and likely also bond-orientational ordering in the (bulk) amorphous CoFeB layer [245], resulting in uniaxial anisotropy fields of a few mT. Noteworthy, the (uniaxial) anisotropy present in the free layer is significantly influenced by the shape of the free layer. In particular, the extent of deviation from a circular form, determines the orientation of the easy-axis (elongated axis), due to the dominating shape anisotropy. During the nanopillar fabrication process, deviations from a perfectly circular shape are inevitable and are primarily caused by EBL and ion beam etching.

In order to acquire nanosecond fluctuations, a voltage divider setup with signal amplification, as described in Section 3.4.2 has been used. An in-plane offset field was applied in order to stabilize a balanced state with 50/50 state probability. The acquired time series for a 50 nm diameter SMTJ is then analyzed in terms of dwell times, which describes the time spent in either the P- or AP-state, and plotted in Figure 4.14a. Since dwell times are Poisson distributed, the number of observed dwell times in a specific time interval is plotted logarithmic in a histogram. P- and AP-dwell times are then derived by fitting the histogram according to: $N_{p/ap} = N_0 \exp(-t/\tau_{p/ap})$. The average dwell time τ is calculated by $\sqrt{\tau_p \tau_{ap}}$ and is found to be 6.7 ± 0.1 ns, which is one or two orders of magnitude shorter than most reported dwell times of oop-SMTJs [246, 247, 248, 249] or ip-SMTJs [239, 242, 250, 251] at room temperature. This originates from the small free layer volume and low in-plane uniaxial anisotropy ($H_k^{ip} \approx 7$ mT) resulting in a low energy barrier.

Given the random two-level telegraph signal, it is of interest to determine the autocorrelation time, which represents the average duration over which the signal loses its correlation with its previous states. This factor commonly sets the boundary for the rate at which random numbers can be generated by a random

number generator (RNG). The autocorrelation function (ACF) of the binarized SMTJ signal is plotted in Figure 4.14b and is calculated according to:

$$ACF(t) = \frac{\sum_{i=1}^{N-t} [s_i - \bar{s}_i][s_{i+t} - \bar{s}_i]}{\sum_{i=1}^N (s_i - \bar{s}_i)} \quad (4.7)$$

where \mathbf{s} is the binarized signal, \bar{s}_i the mean of it and t the lag of the time series of length N . After the autocorrelation time of $\tau_{ac} = 5.1 \pm 0.3$ ns, the signal is uncorrelated to any signal of its past. Here, it is defined as the maximum time by which the integral of the ACF reaches 99%. In the following, we will discuss the possibility of using superparamagnetic MTJs as a true random number generator (TRNG).

4.3.5 True random number generation

The following description of work and results is based on Ref. [106]. As discussed in Section 2.5, a random bitstream generated by a TRNG has to pass various statistical tests in order to be considered as a random number generator for cryptographic applications. Here, the amplified SMTJ signal was acquired by an oscilloscope and transformed into a digital sequence of 0s and 1s. The transformation involved comparing the analog signal with its median to ensure an equal chance of 0s and 1s, creating a balanced Bernoulli sequence, which is a prerequisite for generating a random bit stream. The resultant random bitstream can be compared to the outcome of flipping a fair coin, where heads represent 0 and tails represent 1, each occurring with a 50% chance. If true random, these coin flips are entirely independent, meaning that the outcome of a previous flip has no influence on the outcomes of future flips (Markovian process). For rigorous randomness analysis, the acquired nanosecond random telegraph signal is tested according to the National Institute of Standards and Technology Statistical Test Suite (NIST STS) [125], comprising 15 diverse tests (including subtests). A minimum bitstream length about 10^6 bits is required. The results for the “raw” bitstream are shown in the following table and indicate that the bitstream does not pass all tests for sampling times of 3 ns, 5 ns and 10 ns (the sampling time here refers to the time period at which the recorded signal is resampled).

A method to improve the quality of randomness is the use of logic XOR gates and multiple random bitstreams (described in Sec. 2.5). Two SMTJ bitstreams and a single XOR are not sufficient to pass the NIST STS, while bitstream after an XOR² operation passes all, thus providing true randomness as long as the sampling time is greater than the autocorrelation time. This is indicated in

4. Results

	Type Sampling time	Raw			XOR ¹			XOR ²		
		3ns	5ns	10ns	3ns	5ns	10ns	3ns	5ns	10ns
1.	Frequency	0	0	0	0	0.141	0	0.373	0.739	0.904
2.	Block frequency (m = 128)	0	0	0	0	0	0	0	0.703	0.236
3.	Runs	0	0	0	0	0	0	0	0.069	0.817
4.	Longest-run	0	0	0	0	0.019	0.001	0.001	0.990	0.185
5.	Binary matrix rank	0.335	0.028	0.897	0.507	0.251	0.072	0.517	0.692	0.632
6.	Discrete Fourier transform (spectral)	0	0.871	0	0	1	0.026	0.765	0.448	0.508
7.	Non-overlapping template matching	0	0	0	0	0	0	0	0.039	0.313
8.	Overlapping template matching	0	0	0	0	0	0	0	0.623	0.944
9.	Maurer’s “universal statistical”	0	0	0	0	0.074	0	0.182	0.472	0.761
10.	Linear complexity (M = 500)	0.467	0.137	0.737	0.967	0.929	0.986	0.174	0.085	0.673
11.	Serial (m = 16)	0	0	0	0	0	0	0	0.687	0.900
12.	Approximate entropy (m = 10)	0	0	0	0	0	0	0	0.118	0.151
13.	Cumulative sum	0	0	0	0	0	0	0.162	0.172	0.010
14.	Random excursion	0	0.027	0.068	0.068	0.073	0.042	0.073	0.074	0.074
15.	Random excursion variant	0	0	0	0	0.015	0	0.014	0.015	0.028

Table 4.1: NIST STS randomness test results with p-values for different XOR operations and sampling times. Instances where the p-values exceeded the significance threshold of 0.01 are indicated as passed and are highlighted in green. In cases involving multiple subtests, the respective p-values were amalgamated and evaluated using Fisher’s method [127]. The selection of sampling times at 3, 5, and 10 ns was made to be close to the autocorrelation time τ_{ac} . (Reprinted figure with permission from [106]. Copyright 2023 by the American Physical Society.)

Table 4.1, where the p-values, that exceeded the significance threshold of 0.01, are highlighted in green. For sampling times greater than the autocorrelation time of 5 ns, the same quality of randomness is achieved, whereas, for shorter sampling times, each state would be sampled too frequently, thus resulting in non-true randomness. By utilizing four SMTJs as inputs for an XOR² gate and considering an average autocorrelation time of 5 ns, it’s possible to achieve highly energy-efficient true random bit generation at a rate of 200 Mbit/s. Excluding any additional CMOS circuitry, the energy consumption per bit for a single SMTJ is generally on the order of nanojoules (depending on the RA product and voltage used). However, by employing a “precharge-sense-amplifier” for the state read-out, the energy consumption can be reduced to as little as a few femtojoules per bit [252].

Scalability of the true random number generator

XOR operations are efficient with respect to energy and area within CMOS technology. In our specific scenario, creating a TRNG would necessitate the use of

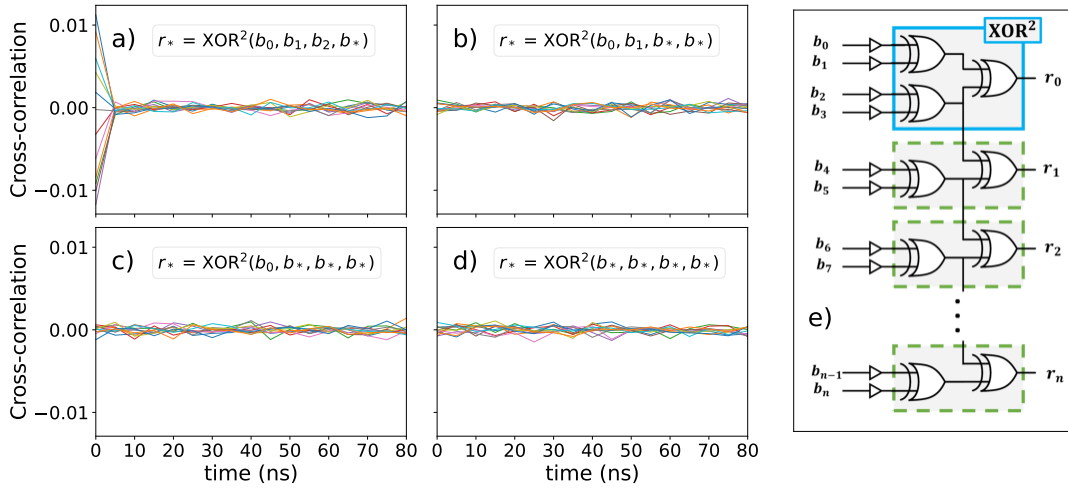


Figure 4.15: a) - d) demonstrates the cross-correlations of $r_0 = \text{XOR}^2(b_0, b_1, b_2, b_3)$ with $\text{XOR}^2(b_0, b_1, b_2, b_*)$, $\text{XOR}^2(b_0, b_1, b_*, b_*)$ and $\text{XOR}^2(b_*, b_*, b_*, b_*)$ for 20 different raw SMTJ bitstreams b_* . e) Raw bitstreams from SMTJs b_i are combined with XOR gates to generate random bitstreams r_i . The proposed circuit reduces the required number of SMTJs and XOR gates for generating multiple true random numbers. (Reprinted figure with permission from [106]. Copyright 2023 by the American Physical Society.)

three XOR gates and four SMTJs. However, in the context of scalability, a reduced amount of hardware is needed to generate additional true random numbers. This is due to the possibility of extracting randomness from an adjacent TRNG source. The concept is illustrated in Figure 4.15e, the combination of raw bitstreams b_i from SMTJs is processed through XOR gates to produce random bitstreams denoted as r_i . The XOR^2 operation for the generation of a true random sequence r_0 is indicated in the blue box. For the generation of an additional true random number, only two more SMTJs and two more XOR gates are required, as shown in the circuit. The new sequence (e.g. r_1) will be truly random and uncorrelated to r_0 .

This is evaluated in Figure 4.15a-d, where the cross-correlation of any new random number r_* with r_0 is calculated for 20 different bitstreams (b_0 – b_{19}) of length $5 \cdot 10^6$. In Figure 4.15a, a significant cross-correlation of r_0 and r_* at a time lag of 0 is present. Here, r_* is the result of the XOR operation $\text{XOR}^2(b_0, b_1, b_2, b_*)$, with b_* representing the additional new SMTJ bitstream. However, when multiple new SMTJs are combined, any subsequent cross-correlation diminishes, indicating that the output becomes uncorrelated with r_0 . The input sequence of bitstreams to the XOR^2 circuit does not affect the output, since the XOR operation is permutation invariant. Therefore, as long as two new bitstreams are

4. Results

integrated with two previously generated bitstreams (from b_2 and b_3), the result will be a true random number without correlation. This assertion is supported by a comparison with the cross-correlation displayed in Figure 4.15d. In this case, there's no noticeable qualitative difference and since all SMTJs are interchanged, r_* is compelled to be uncorrelated with r_0 .

Furthermore, the proposed circuit exhibits a significant advantage in its ability to accommodate SMTJs that are not precisely set to a balanced 50% equilibrium state. It appears that even if a bias up to 10% (deviation of 50%) is present in the bitstreams, the generated XOR² output still remains true random, due to the balancing effect of the XOR gate. This implies that the probabilities of each SMTJ can fall within the 40–60% range, providing flexibility to account for inconsistencies in device fabrication. Furthermore, it's worth noting that if biases are symmetrically opposed, they could potentially surpass the 10% bias threshold, as they partially compensate each other. This favorable characteristic arises due to the balancing effect of the XOR logic (see also Sec. 2.5) and is based on the XOR-function $p_{XOR} = -2p_1p_2 + p_1 + p_2$ for two bitstreams with probabilities p_1 and p_2 . In a cascade of XOR gates, like XOR², this balancing effect is enhanced.

4.3.6 Spin-transfer-torque and Joule heating

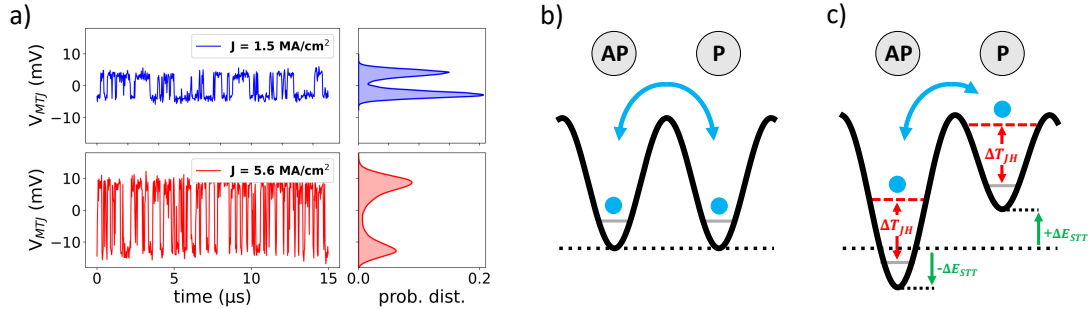


Figure 4.16: a) Acquisition of the SMTJ fluctuation signal for low ($J = 1.5 \text{ MA/cm}^2$) and high ($J = 5.6 \text{ MA/cm}^2$) current densities, together with their corresponding probability distributions. b) Quasiparticle model of the free layer's magnetization confined within a one-dimensional symmetrical potential well. The metastable states of the parallel and antiparallel configurations are separated by an energy barrier. c) The influence of STT introduces an antisymmetric shift in the energy levels of the P- and AP-state. Joule heating symmetrically elevates the energy of the quasiparticle for both states. (Reprinted figure with permission from [106]. Copyright 2023 by the American Physical Society.)

The following description of work and results is based on Ref. [106]. Applying

a current to an SMTJ has an impact on various factors, including its switching speed, state probability, and dwell times [242, 247, 248, 251]. This impact arises not solely from spin-transfer torque (STT) effects, but also from Joule heating localized at the position of the free layer. Joule heating has so far been primarily explored in nonvolatile MTJs, where the objective is to decrease the critical switching current to accomplish energy-efficient free layer switching [105, 253]. The current passing through an MTJ in response to an applied voltage is mainly dictated by the RA product via the MgO thickness. The current density is given by: $J = V/RA$. For an RA product of $15 \Omega\mu\text{m}^2$ and an applied voltage in the range of a few hundred millivolts, the typical current density corresponds to approximately $1\text{-}5 \text{ MA/cm}^2$. This magnitude of current density can become significant in the context of Joule heating. According to the Néel-Arrhenius law, a temperature increase at the free layer has a strong impact on the dwell times of an SMTJ. Figure 4.16a demonstrates the effect of different current densities on the SMTJ fluctuation characteristics. For a low current density ($J = 1.5 \text{ MA/cm}^2$) the average fluctuation rate and voltage amplitude is lower than compared to the high current density measurement ($J = 5.6 \text{ MA/cm}^2$). In addition, a change in the state probability distribution (density plot in Fig. 4.16a) from a P-state preference to an AP-state preference is observed, which is due to the presence of STT. To explain both, STT and Joule heating, a quasiparticle model is considered as illustrated in Figure 4.16b.

Initially, without any applied current, the P- and AP-state are on the same energy level (offset field compensated). This setup envisions the free-layer macrospin as a quasiparticle within a one-dimensional potential well. Ambient thermal energy, approximately 25 meV , raises the quasiparticle's energy, enhancing the likelihood of overcoming the energy barrier. Under a positive applied current, STT lowers the energy level of the AP state while increasing the P-state energy level. This energy shift, denoted as $+\Delta E_{STT}$, leads to a positive shift at the P-state (Fig. 4.16c). The energy shift affects the dwell times according to:

$$\tau_{p,ap} = \tau_0 \exp\left(\frac{E_b \mp \Delta E_{STT}}{k_B T}\right) \quad (4.8)$$

where $\tau_{p,ap}$ are the P- and AP-dwell times, τ_0 the attempt time, E_b the energy barrier at room temperature T and ΔE_{STT} the energy shift due to STT and k_B the Boltzmann constant. Considering the effect of Joule heating, it symmetrically raises the energy of the quasiparticle in the valleys, facilitating enhanced switching overall and shorter dwell times τ . The measured dwell times as a function of current density are plotted in Figure 4.17a, while the average fluctuation rate is shown in Figure 4.17b. The contribution of Joule heating and STT are depicted

4. Results

in Figure 4.17c-d. The temperature change with respect to room temperature ΔT_{JH} , demonstrates a quadratic relation to J , which is consistent with Joule heating power dissipation proportional to J^2 . This increase in temperature also explains the quadratic increase in the average fluctuation rate, which translates into an elevated bit generation rate in a TRNG.

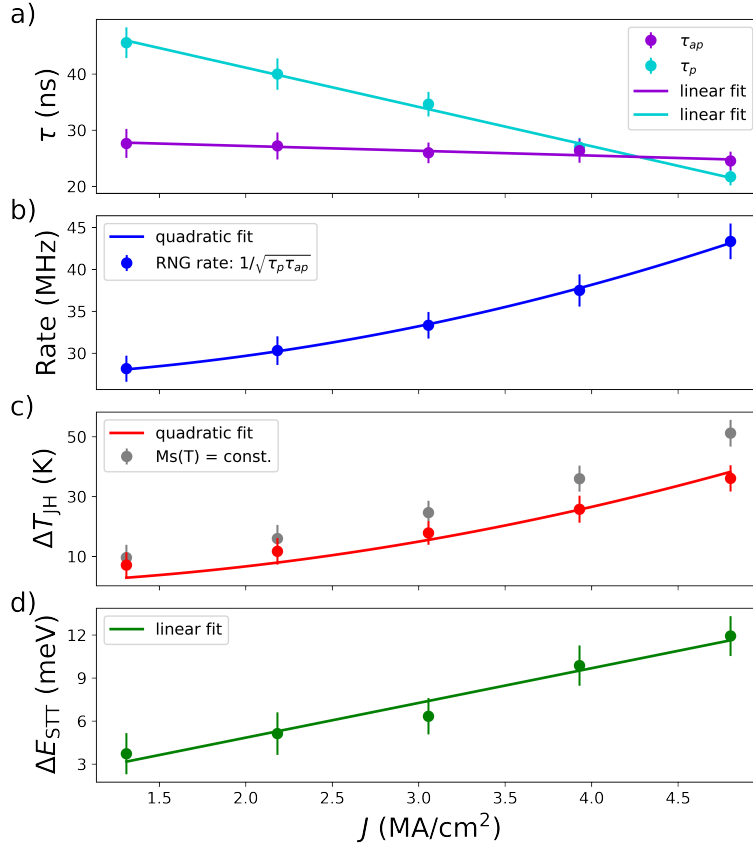


Figure 4.17: a) SMTJ dwell times for current densities ranging from 1.5 to 5.6 MA/cm^2 . Linear fits are applied to indicate the trends. b) The average fluctuation rate (blue), fitted with a quadratic function. c) The temperature increases due to Joule heating (red). The nonlinear increase is fitted with a quadratic function. The gray data points indicate the calculated temperature increase without the consideration of temperature-dependent M_s . d) The energy shift due to the STT (green). Here, a positive energy shift destabilizes the P state. (Reprinted figure with permission from [106]. Copyright 2023 by the American Physical Society.)

Determination of ΔT_{JH} and ΔE_{STT}

The calculation of the energy shift and increase in temperature due to Joule heating can be derived from the dwell time measurements. The energy barrier E_b for given dwell times can be determined by:

$$E_b = 1/2k_B T \log(\tau_p \tau_{ap} / \tau_0^2) \quad (4.9)$$

since the product of τ_p and τ_{ap} is independent of the STT energy-shift. For elevated temperatures T^* (here denoted with a “*”) the dwell times product independent of STT is also given by:

$$\tau_p^* \tau_{ap}^* = \tau_0^2 \exp\left(\frac{2E_b^*}{k_B T^*}\right) \quad (4.10)$$

Here, $T^* = T + \Delta T_{JH}$ and E_b^* are considered as the elevated temperature and energy barrier with dwell times τ_p^* and τ_{ap}^* . Since the energy barrier is a function of the temperature-dependent saturation magnetization M_s of the free layer (CoFeB), the temperature dependent change in E_b has to be considered according to Bloch’s $T^{3/2}$ law (see Sec. 2.1.4). Under the assumption of a constant uniaxial crystalline anisotropy in the free layer and in the case that the switching mechanism is mediated by domain wall nucleation and propagation, the energy barrier is then governed by the domain wall energy density (energy per area): $\sigma = 4\sqrt{A_{ex}K_u^{ip}}$, with A_{ex} the exchange stiffness and K_u^{ip} the anisotropy. The phenomenological exchange stiffness constant is given by: $A_{ex} = nJS^2/a$, where S is the magnitude of the spin at each atom in the crystal lattice with lattice constant a , J is the exchange integral, and n is the number of the nearest neighbor atoms. Since M_s is proportional to S , the exchange stiffness is proportional to M_s^2 and therefore: $A_{ex}(T) \propto M_s(T)^2$ [254]. When assuming that the in-plane anisotropy is dominated by shape anisotropy, due to non-ideal circular patterning, then the anisotropy K_u^{ip} is proportional to M_s^2 . In this case, the energy barrier would also be proportional to M_s^2 . However, due to an optimized MTJ nanopillar patterning process, it is assumed that nanopillar is circular and not dominated by in-plane shape anisotropy. For the calculation of the contributions of STT and Joule heating, the temperature dependent energy barrier as a function of the temperature dependent saturation magnetization is considered: $E_b(M_s(T))$. The resulting temperature-dependent energy barrier under the effect of Joule heating E_b^* is then described as:

$$E_b^* = E_b \frac{1 - (T^*/T_c)^{3/2}}{1 - (T/T_c)^{3/2}} = E_b \frac{1 - ((T + \Delta T_{JH})/T_c)^{3/2}}{1 - (T/T_c)^{3/2}}, \quad (4.11)$$

where E_b refers to the energy barrier at room temperature, E_b^* to the modified energy barrier due to Joule heating at temperature T^* and ΔT_{JH} to the temperature increase due to Joule heating. For $\text{Co}_{40}\text{Fe}_{40}\text{B}_{20}$ the Curie temperature T_c of 895 K [255] was used. Considering Equation 4.8, 4.10 and 4.11, the temperature increase induced by Joule heating can be determined by fitting the data and solving the equations numerically for different applied current densities. The asymmetric energy shift, due to STT, is then calculated by: $\Delta E_{STT} = 1/2k_B(T + \Delta T_{JH}) \log(\tau_{ap}/\tau_p)$ [106]. However, it's important to consider that the initial P- and AP-states at low current densities are not in equilibrium, as indicated by $\tau_p \neq \tau_{ap}$ (see Fig. 4.17). Shorter AP-dwell times are a consequence of an unbalanced stray field stemming from either the TMR stack itself or an externally applied in-plane field. Through linear fits applied to the dwell times, we can extrapolate theoretical dwell-time values at zero applied current, although not measurable. At $J = 0$, neither Joule heating increase ($T_{JH} = 0$) nor an STT ($E_{STT} = 0$) occurs, allowing for the determination of the energy barrier (at 293 K) to be 95.3 ± 1.1 meV. Given an asymmetric energy landscape at $J = 0$, an initial energy shift of approximately 11 meV is present due to the Zeeman energy. In summary, this analysis shows that Joule heating can not be neglected in superparamagnetic tunnel junctions, as it strongly affects the overall dwell times. In the next section, the electrical coupling of two interconnected stochastic SMTJs is elaborated.

4.4 Electrical coupling of superparamagnetic tunnel junctions

In this section, we consider the content of Ref. [256], which covers the electrical coupling of two stochastic magnetic tunnel junctions and how this coupling can be manipulated by an applied voltage. Experimental data will be elaborated, and related simulations will highlight the mechanism of the coupling effect and thereby provide a better understanding of the underlying phenomena.

4.4.1 Time series of two coupled SMTJs

Electrical coupling of multiple stochastic MTJs mediated by spin-transfer-torque offers advantages in terms of scalability, rapid mediation of coupling, and ease of device implementation. In this section, the stochasticity and coupling effect of two in-series electrical connected SMTJs, as shown in Figure 4.18a, is evaluated [256].

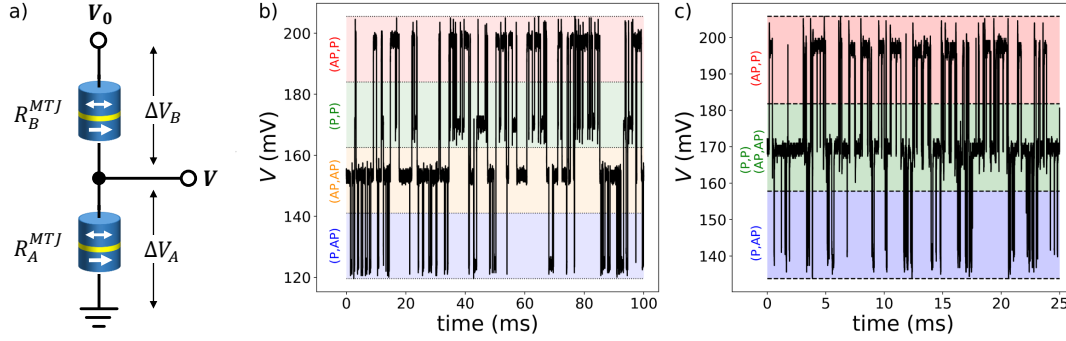


Figure 4.18: a) Sketch of the electrical circuit of two SMTJs in series. A source voltage V_0 is applied and the voltage drop V between both SMTJs is measured. A voltage drop across an MTJ is indicated by ΔV_i . b) & c) Oscilloscope measurements of two coupled SMTJs are shown. The resistance range of three or four separate states (P,P), (AP,P), (P,AP) and (AP,AP) is illustrated by background colors. (Reproduced from [256], with the permission of AIP Publishing.)

In a voltage divider circuit of two SMTJs, with different resistance states ($k\Omega$), the voltage drop V is measured by an oscilloscope (*Tektronix DPO7543*) for different applied DC voltages. Figure 4.18b presents the acquired stochastic voltage drop for an applied DC voltage of 0.3 V. Four resistance ranges (colored bands) appear as a consequence of four distinct SMTJ states: (P,P), (AP,P), (P,AP) and (AP,AP). Here, a bracket notation is used in order to describe the coupled mixed states of two SMTJs A and B : (S_A, S_B) , where S_i stands for the state (P or AP) of the i^{th} MTJ. However, if two states result in approximately the same voltage drop, then only three states are effectively distinguishable. In particular, the combination of SMTJ B and C (see also Fig. 4.19c) exhibits a similar voltage output for the (P,P) and (AP,AP) configuration, leading to a total of three states, as depicted in Figure 4.18c.

Figure 4.19 demonstrates different characteristics of used SMTJs. Figure 4.19a displays the voltage characteristic for SMTJ A for a “fixed” (by a magnetic ip-field) free layer. The decline in resistance for a non-volatile MTJ is attributed to the non-linear tunneling current. The resistance dependence can be described by the model of Brinkman et al. [87] and is approximated by a second order polynomial: $R(V) = a/(1 + bV^2) + d$, where a , b and c are fitting parameters. In Figure 4.19b each data point corresponds to a time series, for which the AP state probability was determined. The applied fit is a sigmoid function based on the following equation:

$$P_{ap}(V) = \frac{1}{e^{2\Delta V/V_c} + 1} \quad (4.12)$$

4. Results

Here, Δ is the energy barrier, V the applied voltage and V_c the critical voltage. Equation 4.12 represents the typical sigmoid relation and is used to model the state probability (called ‘‘PV-transfer function’’) for both SMTJs in the coupled circuit. In Figure 4.19c different time series are illustrated for an applied volt-

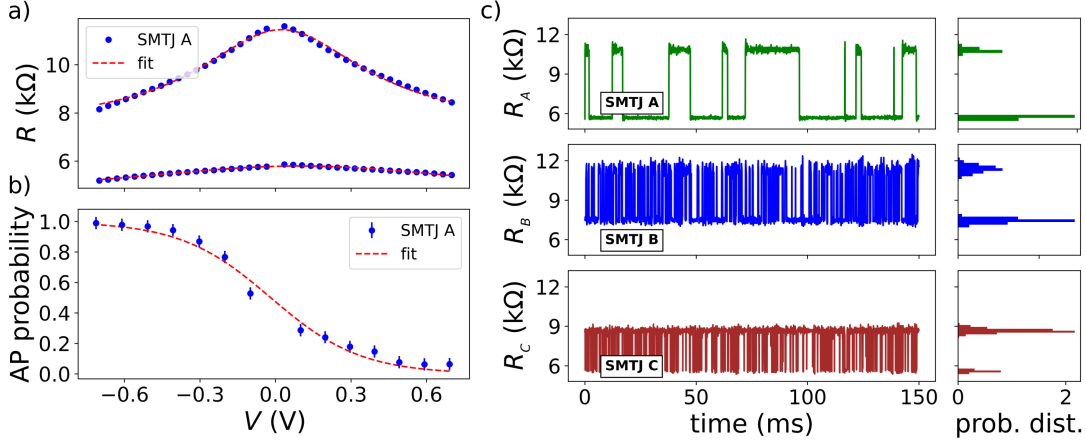


Figure 4.19: a) Bias voltage dependence of an MTJ for a magnetically fixed free layer. b) Typical AP probability transfer curve for an SMTJ with a sigmoid fit. c) Stochastic fluctuation of three different SMTJs *A*, *B* and *C* along with their respective probability distributions. (Reproduced from [256], with the permission of AIP Publishing.)

age of 0.2 V, showing a variation in MTJ resistance, fluctuation rate and state probability. This is likely due to differences in nanopillar size and variations in the tunnel barriers caused by the fabrication process. The dwell times for each SMTJ can be described by the Néel-Arrhenius model and can be tuned via STT, by modifying the energy barrier. In a good approximation the change in the energy barrier is given by [120]: $\tau_{p,ap} = \tau_0 e^{\Delta(1 \pm V/V_c)}$, with τ_0 the attempt time, $\Delta = E_b/k_B T$ the (unitless) energy barrier. The combination of SMTJ *A* and *B* results in a total of four distinguishable states, while the combination of *B* and *C* results in three different states.

4.4.2 Simulation

With the known characteristics, namely the voltage dependent resistance, called ‘‘RV-transfer function’’ and the voltage dependent state probability, called ‘‘PV-transfer function’’, it is possible to simulate the coupling of two stochastic SMTJs. For the generation of artificial two-level random telegraph noise, we can consider a Markov process with four Markov states, as illustrated in Figure 4.20a. The transition probabilities between these states are indicated by arrows and can be

calculated when the MTJ states and the applied voltage is known. The thermal switching probability that the magnetization switches within time t in an MTJ with modified energy barrier can be expressed as [110]:

$$\begin{aligned} P_{p,ap}^{sw}(t) &= 1 - \exp\left(-t/\tau_0 e^{-\Delta(1\pm V/V_c)}\right) \\ &= 1 - e^{-t/\tau_{p,ap}} \\ &\approx t/\tau_{p,ap} \quad (t \ll \tau_{p,ap}) \end{aligned} \quad (4.13)$$

where t is the time spent in a given state, $\tau_{p,ap}$ the dwell time of the parallel/antiparallel state, V the voltage across the MTJ and V_c the critical voltage. As a consequence, the switching probabilities from the P to the AP state and vice versa can be expressed as:

$$P_{p \rightarrow ap} = 1 - e^{-t/\tau_p} \quad (4.14)$$

$$P_{ap \rightarrow p} = 1 - e^{-t/\tau_{ap}} \quad (4.15)$$

These two switching probabilities are relevant for a single stochastic MTJ. However, for two coupled SMTJs there are 12 different switching probabilities, due to the electrical coupling. In general, for n SMTJs in series there exist 2^n distinct Markov states with $4^n - 2^n$ transition probabilities.

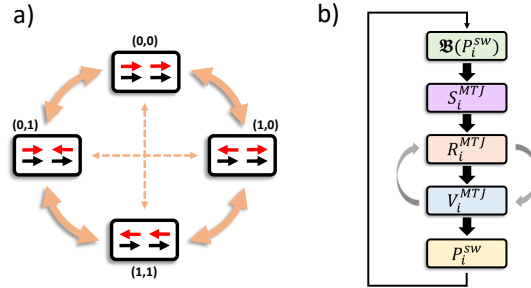


Figure 4.20: a) Four possible Markov states for two interconnected SMTJs are illustrated with the respective free layer (red arrow) and reference layer (black arrow) orientation. At each state, there is a certain probability to transition into another state. b) Illustration of the simulation procedure. Random variables are drawn from a Bernoulli distribution for a specific switching probability P^{sw} . If the sampled variable is true, then the MTJ state switches, which affects its resistance and voltage and thereby also the switching probability for both SMTJs.

To simulate the stochastic fluctuation of two electrically coupled SMTJs, we consider each SMTJ as a Bernoulli distributed random variable with (AP-) probability p . At a given voltage and temperature, an SMTJ is associated with a specific

probability, denoted as p_i of being in the AP-state. When the bias voltage applied to the junction is modified, this probability p_i undergoes changes due to the STT effect and the relationship is the above-mentioned PV-transfer function. Additionally, it's crucial to consider that the resistance, whether in parallel or antiparallel configuration, also exhibits non-linear behavior in response to the bias voltage, which is caused by the non-linear tunneling (see also model of Brinkman for Chap. 2, Sec. 2.3.3), which we consider as the RV-transfer function and has to be taken into account for the simulations.

For simulating a time series for two interconnected SMTJs, we initiate the process by selecting average dwell times $\tau = \sqrt{\tau_p \tau_{ap}}$ for each SMTJ and a time increment dt for each iteration. During the time step dt , the switching probability P^{sw} for each SMTJ can be calculated according to Equation 4.13 (for $dt \rightarrow 0$ thus $P^{sw} \rightarrow 0$). Next, we draw a binary random variable from the Bernoulli distribution: $\mathcal{B}(P_i^{sw})$ for SMTJ i . If the random variable equals 1, then a switching of the MTJ state ($\mathcal{S}_i \rightarrow \overline{\mathcal{S}}_i$) is simulated, which will alter the switching probability P_i^{sw} of both SMTJs, due to a change in the voltage drop ΔV_i caused by a change in the resistance states R_i . Given the known MTJ states at each time step, the resistance states can be deduced. However, owing to the interplay between MTJ resistance and voltage drop, influenced by tunneling (RV-transfer function), both variables mutually influence one another until an equilibrium state is reached. To address this, a recursive function of V and R is implemented in the simulation. The (equilibrium) voltage drop ΔV_i is then used to calculate the new switching probabilities P_i^{sw} for each SMTJ, according to the characteristic PV-transfer functions. Using the updated switching probabilities, new MTJ states can be sampled, and the simulation process starts from the beginning, which ensures a sequential series of Markov states. In order to obtain good statistical results, stochastic time series were simulated for $5 \cdot 10^6$ time steps (iterations), which corresponds to almost 10^5 switching events for each SMTJ, thus the likelihood for a switching event to occur in a single time step is around 1%.

Cross-correlation

Typically, cross-correlation serves as measure of similarity between two signals, like two time series, and helps to identify patterns or relationships between two signals. Thus, it is commonly used in signal processing for pattern recognition [257] or as a measure of similarity of two series (X and Y) as a function of the relative shift of one series with respect to the other. Here, cross-correlation is used as a measure to obtain the coupling strength between two stochastic

MTJs. For stochastic MTJs, time series can be cross-correlated in order to obtain a cross-correlation function, which is a function of time lag (shift) of both series. A positive time lag corresponds to a positive shift of time series Y with respect to X and a negative time shift corresponds to a negative shift of time series Y with respect to X . The cross-correlation function is defined as:

$$C(t) = \frac{\sum_i \frac{1}{N} (X_i - \bar{X}) \cdot (Y_{i-t} - \bar{Y})}{\sigma_X \sigma_Y} \quad (4.16)$$

where t is the time lag ($t \in [-N/2, \dots, N/2]$), \bar{X} , \bar{Y} the mean and σ_X , σ_Y the standard deviation of time series x and y with length N . The values X_i and Y_i with i outside the range of X and Y are 0 and do not contribute to the cross-correlation. The normalized cross-correlation (see Eq. 4.16), is equivalent to the normalized Pearson correlation coefficient with time lag t . By the cross-correlation function, the direction of influence can be accessed. If a correlation peak appears on the positive side of the cross-correlation function, then the effect of series Y on X is obtained, while for the negative side, the influencing effect of X on Y is accessed. This is a consequence of the positive and negative shift in the cross-correlation function.

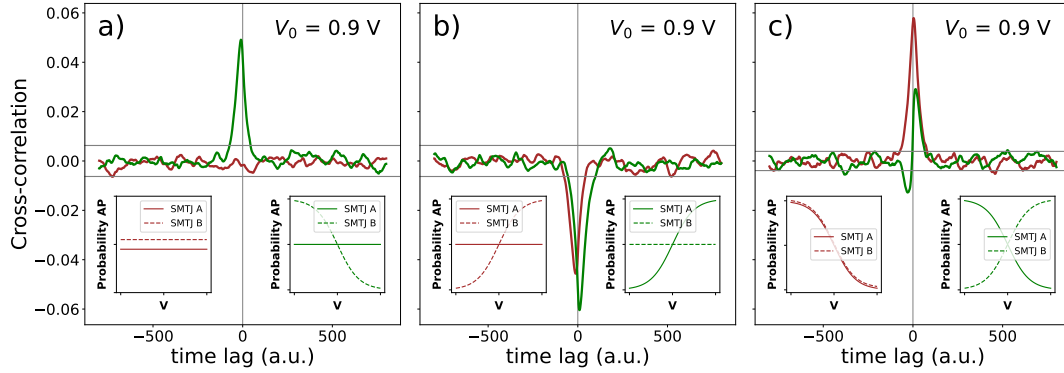


Figure 4.21: A simulation of the cross-correlation for various PV-transfer curves. a) The cross-correlation function is highlighted in red for two constant PV-transfer functions, as illustrated in the left inset. In green, the correlation is plotted for a sigmoid transfer curve of SMTJ A and a constant transfer curve of SMTJ B, as depicted in the right inset. b) & c) Illustration of the impact on cross-correlation due to different PV-transfer curves of SMTJ A and B. (Reproduced from [256], with the permission of AIP Publishing.)

Figure 4.21 illustrates the results of the simulation, considering instantaneous state transitions, equivalent average switching rates, and the absence of additional noise. In general, the cross-correlation function demonstrates a peak at a time

lag of 0, as anticipated in the context of two interdependent time series. This behavior emerges due to the impact of switching events in one MTJ on the state of the other MTJ, triggered by variations in STT. The amplitude of the correlation primarily relies on the TMR ratios and the applied voltage to the interconnected SMTJs, as these SMTJs are most “responsive” to voltage drops in the region where the gradient of the PV-transfer function is maximal. The correlation peaks can be observed as either positive or negative, and they may be shifted to positive or negative time lags. When a peak is located on the positive segment of the cross-correlation function, it signifies the impact of series B on A , whereas on the negative segment, it reflects the influence of A on B . Consequently, we can assess the influence of both SMTJs on each other. A positive correlation would indicate a tendency for both SMTJs to remain in the same state (referred to as the “similarity effect”), while a negative correlation would imply a preference for antiparallel alignment (known as the “dissimilarity effect”). The results are depicted in Figure 4.21 and outlined in Table 4.2. In Figure 4.21a, the cross-correlation is simulated for two SMTJs, with the coupling artificially disabled (represented by the red curve), which corresponds to constant PV-curves. The simulation is also conducted for the scenario where solely the PV-curve of SMTJ A remains constant, while the one for SMTJ B demonstrates the sigmoid function (illustrated by the green curve). The initial scenario demonstrates an absence of any cross-correlation peak, while in the subsequent situation, a positive peak becomes apparent within the negative time lag range. This indicates that SMTJ A impacts SMTJ B , with B demonstrating a preference for adopting the same state as A . Figure 4.21b illustrates the consequences of interchanging the PV-transfer curves of A and B , as depicted in the insets. This interchange results in a shift of the correlation peak from the negative (red curve) to the positive (green curve) time lag range. A more realistic simulation, outlined in Figure 4.21c, represents two interconnected SMTJs with sigmoid PV-transfer functions and bilateral coupling. If both PV-transfer functions follow the same trend, such as both monotonically increasing, it leads to a symmetric correlation peak centered at 0. However, if one of the PV-transfer functions is reversed ($f(x) \rightarrow f(-x)$), it results in an antisymmetric peak. In this case, there are positive as well as negative correlations near the 0 time lag, each corresponding to an SMTJ influencing the other.

In the previous simulation, it has been assumed that both SMTJs possess the same average dwell times: $\tau_A = \tau_B$. However, typically for different SMTJs the dwell times are different, as observed in Figure 4.19c. Therefore, for various dwell time ratios of SMTJ A and B the effect on the cross-correlation is simulated. As

Table 4.2: Cross-correlation summary

Lag	Correlation	Direction	Effect
+	+	$B \rightarrow A$	similarity
+	-	$B \rightarrow A$	dissimilarity
-	+	$A \rightarrow B$	similarity
-	-	$A \rightarrow B$	dissimilarity

it can be seen in Figure 4.22a, the width of the cross-correlation peak is primarily determined by the longest average dwell time of both SMTJs. If one of the dwell times is much greater than the other, it will lead, on average, to a coupling effect present over longer time periods, as the slow switching SMTJ is present for longer in a specific state and during this time the second SMTJ has a greater chance to react to the altered voltage.

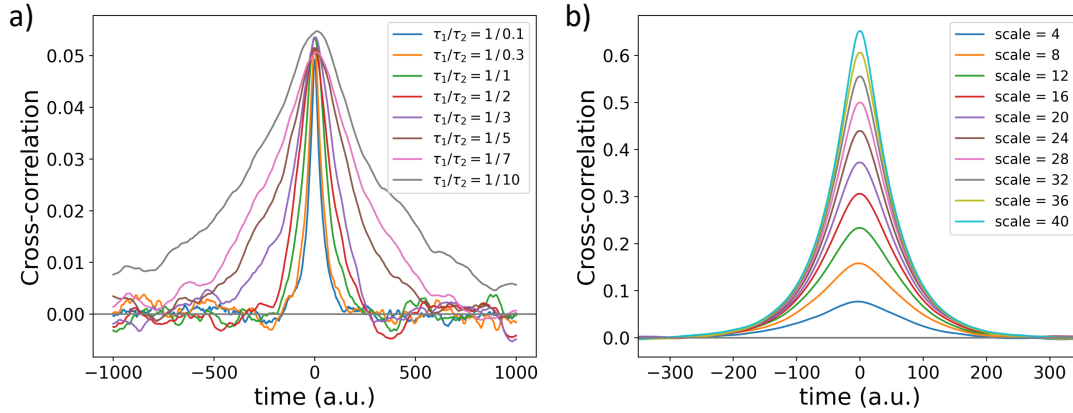


Figure 4.22: a) Cross-correlation for SMTJs with different dwell times τ_A and τ_B and dwell time ratios τ_A/τ_B . b) Cross-correlation for two SMTJs ($\tau_A = \tau_B$) with for different scaled PV-transfer curves. Here, “scale” denotes the exponential parameter s of the sigmoid function.

In order to enhance the coupling strength between the SMTJs, PV-transfer curves with a sharp transition are necessary, such that a small change in voltage leads to a substantial change in SMTJ probability. Figure 4.22b illustrates the effect of scaled PV-transfer curves, of the form: $P_{ap}(V) = 1/(1 + e^{s \cdot V})$, where V is the voltage and s is a scaling factor determining the steepness of the sigmoid function. The simulation demonstrates, that for very sharp PV-transfer curves a high cross-correlation, thus also a high coupling effect is expected. This cross-correlation can exceed 50%, if the PV-curves are “sharp” and centered at half of the applied voltage V_0 . In order to modify the PV-transfer curves, it would

be necessary to change the RA product of the TMR stack, which would enable higher current densities and thereby stronger spin-transfer-torques. This can be achieved by depositing the same stack with a thinner MgO barrier.

4.4.3 Experimental results

The analysis of the time series, as shown in Figure 4.18b, reveals coupling strength and mechanism of two SMTJs interconnected in series. Figure 4.23 demonstrates the experimental results of state probability and correlation of two coupled SMTJs, namely SMTJ *A* and *B*. Figure 4.23a illustrates the manipulation of states via the source voltage V_0 , showcasing the influence of voltage control on the interconnected system. Notably, the upper limit of the applied voltage, $V_0 = 1$ V, corresponds to a maximum MTJ current density of approximately 3 MA/cm^2 . For highly negative voltages, the dominant state is the (AP,AP) state, whereas in the positive voltage range, this state is less frequently occupied, primarily favoring the (P,P) state. This suggests that both SMTJs exhibit a similar probability trend with respect to V , with positive voltages stabilizing the parallel state and negative voltages stabilizing the antiparallel state.

In order to explore the coupling between the two SMTJs, the cross-correlation of the time series *A* and *B* was examined. The states of each SMTJ are transformed into a binary sequence of -1 and +1, denoting the antiparallel and parallel states, respectively. The normalized cross-correlation (see Eq. 4.16), analogous to the normalized Pearson correlation coefficient with time lag t , is employed to evaluate the strength of coupling. The outcomes are presented in Figure 4.23b alongside the derived autocorrelations. The cross-correlation signal manifests a minor peak of approximately 5% at 0 time lag, decreasing exponentially for larger time lags, as highlighted in the inset figure. Figure 4.23d displays the same cross-correlation over a time lag range of ± 4 s, revealing a distinct peak at 0 that highlights a significant positive coupling. Because of the distinctive sigmoid PV-transfer curves, the cross-correlation is anticipated to vary with the applied voltage. Consequently, the maximum correlation close to 0 time lag depending on V_0 is assessed. The findings are depicted in Figure 4.23c. Positive correlations can be observed on the positive voltage side, attributed to the “reversed” sigmoid curve, resulting in positive correlations, as explained for the simulation. Elevated voltages V_0 lead to increased correlations due to amplified voltage drops at the junctions, exerting a stronger influence on the state probabilities. Consequently, at low absolute voltages, the coupling strength diminishes, approaching zero. However, for highly negative voltages, the correlation also approaches zero, given that there is minimal switching due to both SMTJs primarily remaining in the AP state (see

also Fig. 4.23a). While the coupling effect is still present for negative voltages, quantifying it via cross-correlation techniques becomes more challenging. Additionally, the correlation trend undergoes a sign change at $V_0 = 0$, transitioning from negative to positive correlations. This phenomenon can be explained by the opposing influence of negative voltages on the PV-transfer function.

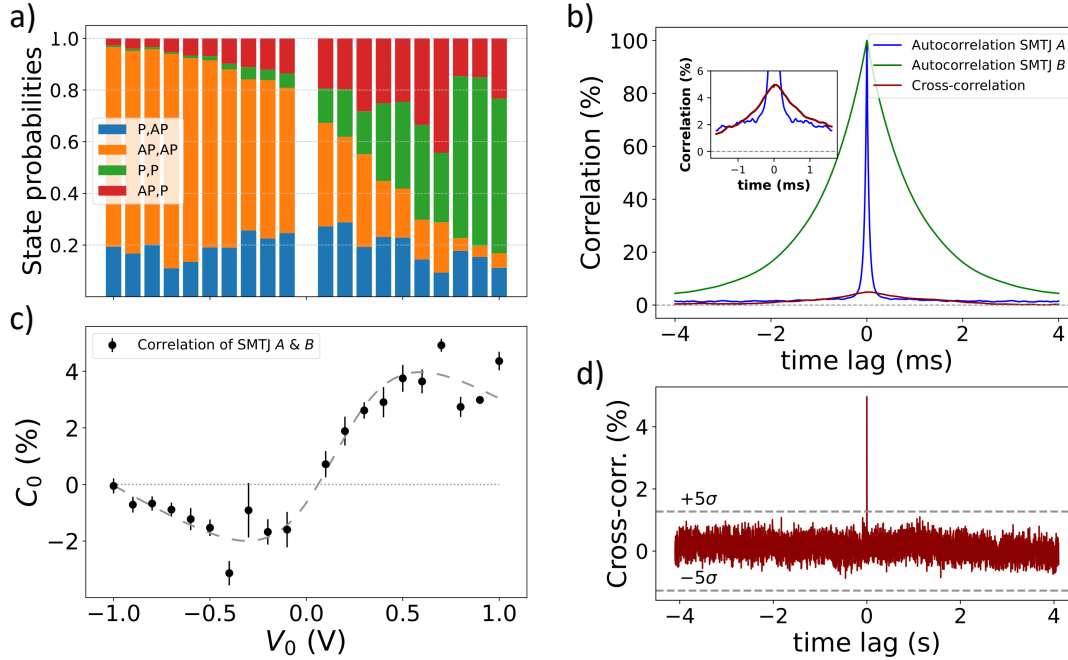


Figure 4.23: a) The dependence of state probabilities on the source voltage V_0 for two coupled SMTJs is illustrated. b) Autocorrelation functions of SMTJ A and B as well as their cross-correlation function are plotted ($V_0 = 0.7$ V). c) Cross-correlation at a time lag 0 is plotted as a function of applied voltage for two coupled SMTJs. The dashed line indicates the correlation trend. d) Cross-correlation ($V_0 = 0.7$ V) for a time lag range of ± 4 s. A $\pm 5\sigma$ level is indicated as a dashed line. (Reproduced from [256], with the permission of AIP Publishing.)

In conclusion, this analysis demonstrates the successful coupling of two series-connected SMTJs controlled by an applied voltage. Simulations underscore the significant influence of each SMTJ's probability transfer function, modulated via spin-transfer torques and the TMR effect. In the next section, we will discuss how multiple interconnected SMTJs can provide necessary randomness in a neural network to allow for an alternative learning approach, called node perturbation.

4.5 Neuromorphic computation with SMTJs

Various concepts of neuromorphic [17, 23, 25, 137] and non-conventional computing [258, 259] based on magnetic tunnel junctions have been proposed, such as probabilistic computing [26, 260], Boltzmann machines [27, 32], spin-torque and spin-Hall nano-oscillators [261, 262], spiking neural networks [28, 152, 263, 264], stochastic computing [29, 107] or reservoir computing [31, 265]. The demonstrated concepts often present either how an activation function can be established by a non-volatile MTJ or how an artificial synapse can be employed as a signal weight. However, the learning process, often reliant on the well known backpropagation [22], is typically not considered in detail, although it would require special circuitry in a hardware device. Here, an alternative concept is demonstrated based on a local learning approach, called node perturbation (NP) [148]. The recently proposed extension of the node perturbation algorithm, called “activity-based node perturbation” [149], allows for an efficient integration of superparamagnetic MTJs into an analog circuitry, thereby potentially leveraging the advantage of local learning through noise in hardware. The noise, necessary for the node perturbation algorithm, should be uncorrelated, random and of specific amplitude. Typically, in node perturbation, the perturbation signal is sampled from a normal distribution. In the following, we consider probability distributions generated through the interconnection of multiple stochastic MTJs.

4.5.1 Probability distributions sampled by SMTJs

A single SMTJ exhibits random fluctuations between the P- and AP-state, and therefore also a fluctuation in the resistance R_{MTJ} and conductance $G_{MTJ} = 1/R_{MTJ}$. If only two states (P- and AP-state) are prominent and the transition between these states is considered to be infinitesimally small (as a good approximation for average dwell times of μs or ms), then the MTJ resistance or conductance at a given time can be interpreted as a sampled Bernoulli random variable x . The Bernoulli distribution \mathcal{B} describes a discrete probability distribution for a random variable with only two outcomes, 0 and 1. The probability to observe 1 (or the AP-state) is given by p , while the probability to observe 0 (or the P-state) is given by $1-p$. Thus, the volatile state of an SMTJ can be simulated by drawing a random variable x from a Bernoulli distribution: $x \sim \mathcal{B}(p)$. The probability mass function of the Bernoulli distribution is given by:

$$f(s, p) = p^s(1 - p)^{1-s} \quad s \in \{0, 1\}, \quad (4.17)$$

where $s = 1$ for the AP-state and $s = 0$ for the P-state. A connection of n stochastic MTJs in series, will lead to the sum of random resistance states, thus the sum of Bernoulli random variables. In the case of identical and independent (Bernoulli) random variables, this sum corresponds to a Binomial random variable:

$$\sum_{i=1}^n x_i \sim \mathfrak{B}(k, n, p), \quad (4.18)$$

where \mathfrak{B} is the Binomial distribution, p refers to the AP-probability, which is the same or in good approximation similar for each SMTJ, n the number of MTJs and k the number of MTJs in the AP-state. The probability mass function is given by:

$$f(k, n, p) = \binom{n}{k} p^k (1-p)^{n-k}, \quad (4.19)$$

with the binomial coefficient defined as: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. In the case for $n \rightarrow \infty$ the Binomial distribution approximates the Gaussian distribution: $\mathfrak{B}(n, p) \approx \mathcal{N}(np, \sqrt{np(1-p)}) = \mathcal{N}(\mu, \sigma)$, with mean μ and standard deviation σ [266]. More general, according to the central limit theorem the sum of a large number of independent and identically distributed random variables, regardless of their original distribution, will approximate a normal distribution. For the sum of non-identical distributed random variables (as in the case for multiple SMTJs), it will still tend to follow a normal distribution. As a consequence, through the combination of multiple SMTJs, it is possible to generate ‘‘Gaussian-like’’ distributions in resistance or conductance. In Figure 4.24 a simulated distribution of 8 interconnected SMTJs is illustrated, as a series circuit and as a parallel circuit. Figure 4.24a describes a series circuit of SMTJs, which results in a Gaussian-like distribution in resistance R , whereas in Figure 4.24b a parallel circuit of SMTJs is illustrated, which leads to a Gaussian-like distribution in conductance G . In Figure 4.24c the probability distributions for R and G (green plot) are sampled for 8 SMTJs connected in series, with each equal is resistance ($R_p = 5 \text{ k}\Omega$ and $R_{ap} = 10 \text{ k}\Omega$) and probability ($p = 50\%$). The distribution exhibits a total of $n+1$ states, indicated by 9 bins in the histogram. The finite bin width here, leads to missing counts for a certain conductance range and indicates that the conductance values are not equidistant distributed. A more realistic case considers also a variation in resistance and probability for each individual MTJ. Here, a variation is modeled by using random values for the resistance as well as the probability, generated by a normal distribution with a relative standard variation of 15% ($\sigma_{R_{p,ap}} = 0.15R_{p,ap}$, $\sigma_p = 0.15p$). The probability distribution is plotted in red in Figure 4.24c. From this, it is evident that for the series circuit the resistance dis-

4. Results

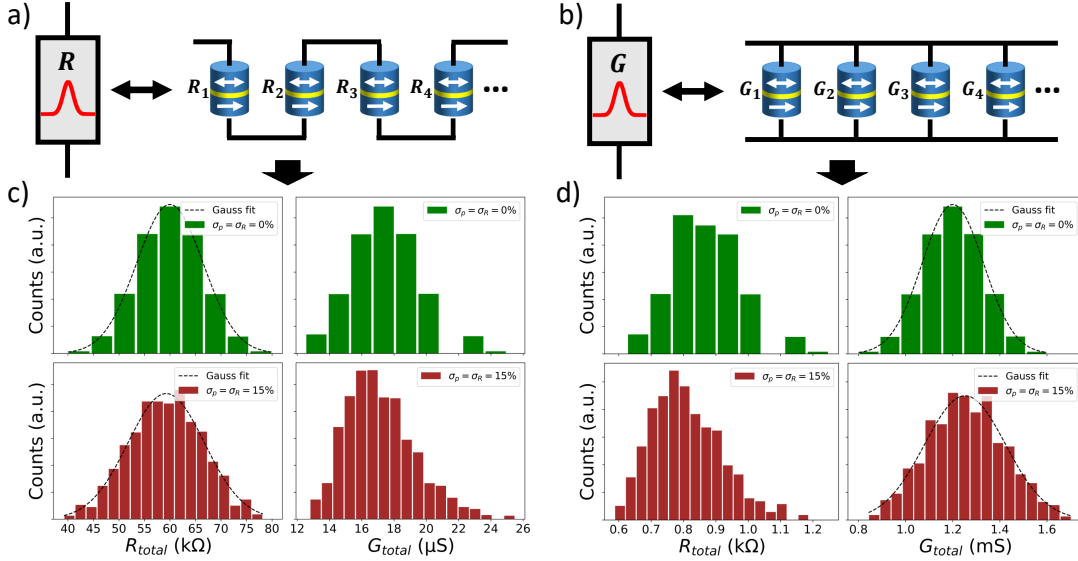


Figure 4.24: a) Illustration of a series circuit of multiple stochastic MTJs, exhibiting a Gaussian-like distribution in resistance. b) Illustration of the equivalent parallel circuit, resulting in a Gaussian-like distribution of conductance. c) Sampled distributions for series circuit for $R_p = 5\text{ k}\Omega$, $R_{ap} = 10\text{ k}\Omega$ and $p = 0.5$ without (green) and with 15% variation (red) in resistance states and probability. d) Similarly, the sampled distribution is plotted for the parallel circuit, without (green) and with 15% variation (red) of resistance states and probability. Each distribution is generated from 10^5 random variables.

tribution is Gaussian-like while the distribution in conductance is non-Gaussian and skewed. Due to the random variation in the MTJ characteristic, the resulting probability distribution can range from a perfect Gaussian to an almost uniform or skewed distribution. However, on average, it tends to approximate a Gaussian distribution. Another advantage of the intrinsic MTJ variation is the increased amount of sampled states, as indicated by the number of bins (18).

Contrary to that is the case of the parallel circuit in Figure 4.24d, where the Gaussian distribution is found for the conductance while the resistance distribution is skewed. The parallel circuitry has the advantage of an easy implementation due to the ease of interconnection of top and bottom electrodes, while also offering advantages in utilizing conductances for neuromorphic circuits, as elaborated below.

4.5.2 Node perturbation circuit

For the NP local-learning (see also Sec. 2.6.4), the weight update at layer l has the following relation: $\Delta \mathbf{W}^l \propto -\delta \mathcal{L}(\tilde{\mathbf{a}}^l - \mathbf{a}^l)(\mathbf{x}^{l-1})^\top$. Thus, in the NP circuit

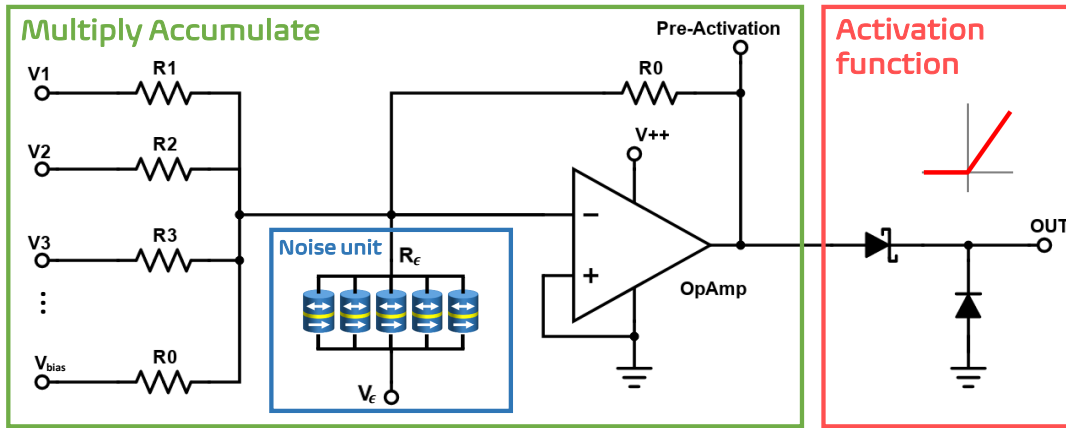


Figure 4.25: The NP perceptron circuit is illustrated as a combination of resistors, stochastic MTJs, diodes and an operational amplifier (OpAmp). Input voltages V_1, V_2, V_3, \dots are weighted by the respective inverse resistances. A noise unit of multiple MTJs provides a Gaussian-distributed perturbation. The multiply accumulate (MAC) operation is carried out by the OpAmp (summing amplifier configuration) and its output, called activation, is used as the input for the analog ReLU activation function, which is implemented through two diodes.

the pre-activation signal \mathbf{a} , the node input signal \mathbf{x} (previous layer output) and the global loss difference $\delta\mathcal{L}$ are required. All signals can be considered as analog voltages, which allows for fast signal transmission, ease of measurement and proper controllability. The following proposed circuit of a perceptron is primarily based on three parts: 1. Multiply and accumulate (MAC) unit, 2. Noise unit, 3. Activation function unit. The MAC performs the summed weight for given (analog) input voltages V_i through the use of a single operational amplifier (OpAmp). The non-inverting input (+) of the OpAmp is connected to ground, while the inverting input (-) is connected to the output via R_0 and to the input signals. The OpAmp aims to minimize the voltage difference at the inverting and non-inverting input, developing a virtual ground at the inverting input. The NP circuit for a simple perceptron is illustrated in Figure 4.25.

The operational amplifier is used in a summing amplifier configuration, in which multiple input voltages are linearly combined to produce an output voltage. Each input voltage V_i is weighted by its corresponding input resistor R_i . The output voltage is the sum of these weighted input voltages, where the OpAmp amplifies the resulting voltage to produce a scaled output. In addition, the Gaussian-like noise in the form of a Gaussian-distributed current is generated by stochastic MTJs and injected to the summed weight. A DC-voltage V_ϵ in combination with Gaussian-distributed MTJ conductance G_ϵ results in a normal distributed (noise)

4. Results

current, which can simply be scaled by tuning V_ϵ . This allows for an easy control of the injected noise amplitude, which is important to fine-tune the NP-learning performance. The summing amplifier MAC operation with added perturbation of node i at layer l can be calculated according to:

$$\begin{aligned}\tilde{a}_i^l &= -R_0 \left(V_1^{l-1} \frac{1}{R_{i1}^l} + V_2^{l-1} \frac{1}{R_{i2}^l} + V_3^{l-1} \frac{1}{R_{i3}^l} + \dots + V_{i,\epsilon}^l \frac{1}{R_{i,\epsilon}^l} + V_{i,bias}^l \frac{1}{R_0} \right) \\ &= -R_0 (V_1^{l-1} G_{i1}^l + V_2^{l-1} G_{i2}^l + V_3^{l-1} G_{i3}^l + \dots + V_\epsilon^l G_{i,\epsilon}^l) - V_{i,bias}^l\end{aligned}\quad (4.20)$$

Here, R_1, R_2, R_3, \dots are resistances, G_1, G_2, G_3, \dots are the respective conductances representing the weights of the perceptron, while V_1, V_2, V_3, \dots represent the previous layer voltage outputs. R_0 is a scaling factor and determines the magnitude of amplification. The perturbation amplitude can be controlled by V_ϵ , while random and uncorrelated noise is generated in the “noise unit” by a parallel circuit of superparamagnetic tunnel junctions and is defined as G_ϵ . This parallel circuit generates a Gaussian-like distribution of conductances and through $\epsilon = V_\epsilon G_\epsilon$ the perturbation signal. Depending on the resistance (conductance) states of each SMTJ, it is possible to exploit a number of $N + 1 < n < 2^N$ different states for N MTJs in the circuit. In case all MTJs exhibit the same parallel and antiparallel resistances, the number of distinct states equals $N + 1$. However, if each SMTJ has unique (and non-similar) resistance states, it can result in a total of 2^N distinct states. Since it is typically difficult to fabricate MTJs with exactly the same resistance levels due to the sophisticated processing steps involved, there is often a large variation in resistance states observed. Consequently, we expect to observe more distinct states (e.g. $\approx 2N$) in the parallel circuit, which is of advantage in the NP algorithm, since more states are sampled in the parameter space.

As illustrated in Figure 4.25, the node perturbed pre-activation voltage can easily be accessed as the output of the OpAmp. This is then used as the argument for the activation function, which here is proposed either as a ReLU function or as a sigmoid/tanh function. There are different possibilities in order to implement a non-linear function with a combination of diodes, resistors and MOSFETs (metal oxide semiconductor field-effect transistors). The implementation of an analog ReLU function is here proposed as a very simple circuit, which is exact, fast, energy efficient and has a low-area cost. The popular ReLU function can be modeled by a voltage divider circuit of two diodes. Here, the non-linearity originates from the non-linear I-V dependence of a diode, which is typically mathematically described by the Shockley equation [267]. Since a positive input at the OpAmp results in a negative OpAmp output, the perceptron is primarily sensitive to neg-

active input voltages, because for positive values the ReLU equals 0. Through negative biasing of the OpAmp, it is also possible to enable a dependency on positive input voltages.

The requirement for the non-linear analog activation functions is not only the accuracy, speed, energy and area-efficiency but also its scalability in terms of the delivered (forward) voltage. The implementation of the activation function ideally is an active (CMOS) component, comprising its own voltage source, able to inject voltage into the next layer of neurons. This ensures that the activation function output is not affected by the connected load resistance. If the activation function would be a passive component, it would lead to voltage drops through the connection of the neuron with neurons from the next layer, which would result in a node output voltage, which is strongly affected by the connected weights. For the diode ReLU implementation in combination with the OpAmp though this effect is present, it can be neglected since the resistance of a diode is either very high ($> 1\text{ M}\Omega$) in the reverse bias or relatively low (typically $< 1\text{ }\Omega$) in the forward bias mode compared to the total load resistance. The activation function would only be affected for load resistances comparable to the internal diode resistance (in forward bias). By choosing an appropriate diode type (e.g. *SMCJ33A*) and high weight resistances, the ReLU output can be considered independent of the connected weights, if $R_{diode}^{fw} \ll R_{load} = (\sum_i 1/R_i)^{-1}$.

As an alternative, the activation function can also be implemented using two MOSFET's as a sigmoid/tanh-like function. An additional power supply at the n-MOS and p-MOS transistor feeds the signal to the next nodes in the network.

Weight implementation

Utilizing a material property, such as conductance, enables the representation of a synaptic weight in a neural network. Within an arrayed structure, it becomes possible to design a weight matrix comprising individually adjustable weights at each intersection. In order to allow for parallel information processing, analog weight multiplication has several advantages over classical computation. Crossbar arrays enable massive parallelism, allowing for the simultaneous processing of multiple signals via Ohm's law. A voltage vector \mathbf{V} is transformed into a current vector \mathbf{I} by: $\mathbf{I} = \mathbf{G}\mathbf{V}$, where \mathbf{G} represents the conductance weight matrix of a certain layer. This is illustrated in Figure 4.26b. In addition, it can result in lower power consumption through low currents and minimal effort in terms of computational steps. Also, the ability to process data in parallel in real time with a quick response makes it attractive for real-world applications. Furthermore, crossbar

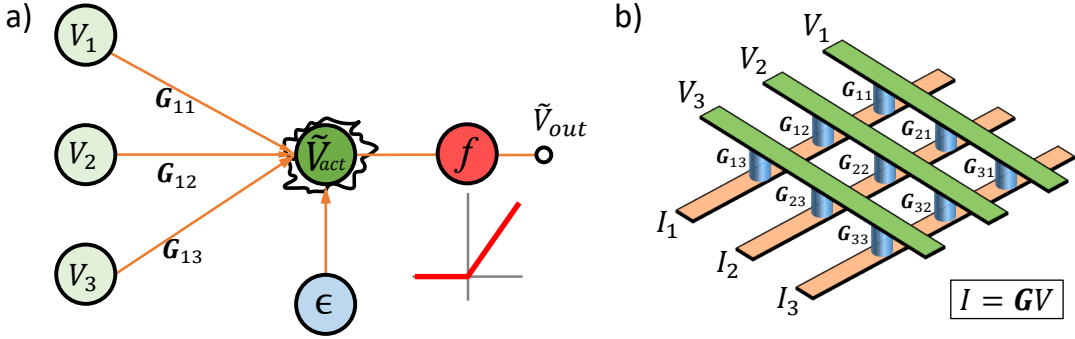


Figure 4.26: a) A node perturbation perceptron schematic with input voltages V_1, V_2, V_3 , conductance weights G_{ij} and perturbation ϵ is shown. The activation function transforms the pre-activation voltage V_{act} in a non-linear way into the output voltage. b) A crossbar array performs an analog weight multiplication by $\mathbf{I} = \mathbf{G}\mathbf{V}$. Input voltage lines are shown in green, while current output lines are outlined in orange.

arrays are well-suited for processing continuous signals and data, enabling the efficient processing of real-world data, like audio or sensor signals. Yet, a challenge in analog computing is the unpredictability of weight values within a crossbar array. Analog implementations often exhibit fluctuations in weight values, which may even differ after attempting to set the same target value multiple times. Consequently, the precise configuration of learned weights within a crossbar array is unattainable.

Several different types of crossbar arrays have been proposed [268, 269], which are typically based on a memristor device, capable of storing and maintaining a resistance state after being set by a current or voltage [270]. Typically, the requirement of the memristor is a wide resistance range to represent a large number of resistance states, equivalent to the number of weights, low resistance fluctuation or drift and high endurance. An example of such a memristor, which can be scaled on a hardware chip, is HfO_x [271, 272].

An alternative way for a physical crossbar array implementation has been proposed based on MTJs [15]. Since a single MTJ has only two stable resistance states, a strong weight constraint is given by using a binary MTJ crossbar array [16]. Crossbar arrays based on 1 transistor and 1 MTJ (1T-1MTJ) [15] or based on 1 diode and 1 MTJ (1D-1MTJ) [273, 274] have been studied, and allow for improved controllability and accuracy compared to purely passive crossbar arrays. Moreover, implementing negative weight values is challenging since the system is confined to positive resistance/conductance states. Consequently, these stringent weight constraints often result in a decline in network performance. Despite this limitation, the neural network is capable of operating on negative

signals resulting from negative inputs or negative node bias. However, detailed exploration of various possibilities for implementing a crossbar array for analog computing is not addressed here. The impact on the performance of an NP neural network circuit needs to be investigated in future research. Next, we consider a simulation of the node perturbed perceptron circuit.

4.5.3 SPICE simulation of a noise perturbed perceptron

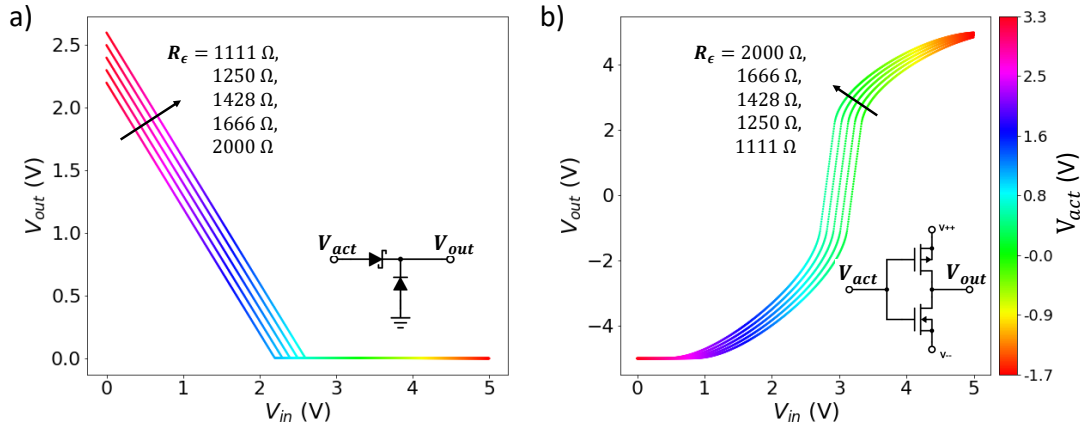


Figure 4.27: *LTspice* simulation of the node perturbation circuit. The output voltage V_{out} is plotted as a function of an input voltage V_{in} of the perceptron, and the activation voltage V_{act} that is present at the input of the activation function is outlined in the colorbar. The perturbation is simulated by five different resistance values for R_ϵ , which results in shifts of the activation function. a) illustrates the output for a ReLU activation function implemented by two diodes. b) The output is illustrated for a sigmoid-like activation function, implemented by two MOSFETs (inverter circuit).

Node perturbation requires measuring the loss of the neural network before and after the perturbation of the node activation. Figure 4.27 illustrates a *LTspice* simulation of the perceptron (shown in Fig. 4.25) for two different activation functions. For the simulation the following parameters have been chosen: three input voltages ($V_1 = 1\text{ V}$, $V_2 = 2\text{ V}$, $V_3 = 3\text{ V}$) and three corresponding weights ($R_1 = 10\text{ k}\Omega$, $R_2 = 50\text{ k}\Omega$, $R_3 = 100\text{ k}\Omega$) and ($R_0 = 10\text{ k}\Omega$), a bias voltage ($V_{bias} = -4.5\text{ V}$) and bias resistance ($R_{bias} = 10\text{ k}\Omega$) and a noise voltage ($V_\epsilon = 0.1\text{ V}$). A variation of an (arbitrary) input voltage (e.g. $V_{in} \hat{=} V_1$) at the perceptron together with the noise perturbation induced by superparamagnetic MTJs modifies the pre-activation $\tilde{\mathbf{a}}$. Consequently, the node output $f(\tilde{\mathbf{a}})$ is modified as well. Figure 4.27a demonstrates the output for the ReLU implementation as a function of input voltage for different noise resistance values R_ϵ generated by SMTJs. An

additional negative bias voltage is set in order to shift the activation into the positive voltage range. A higher total MTJ resistance results in a shift in the curve, as indicated by the arrow in Figure 4.27a. The pre-activation value for a given input is indicated by the color.

In Figure 4.27a, the ReLU function is observed to have a horizontal offset, as the output does not begin to rise at 0 of the activation voltage (green area), but rather at a positive offset of approximately 0.7 V. This is due to the forward voltage drop, an intrinsic property of the diode, caused by the depletion region of the p-n junction. A diode utilized in forward bias mode, with a positive voltage applied at the p-doped region and the negative voltage at the n-doped region, conducts current if the voltage is sufficiently high to close the depletion region at the p-n junction. This minimum voltage required to accomplish the collapse of the depletion region is called forward voltage. This forward voltage, which is typically of the order of 0.7 V, causes a shift in the ReLU function and primarily depends on the diode type and design. For our perceptron, simulated with a normal diode, an offset in the activation function is not critical, since it can be counteracted by the bias voltage V_{bias} in the circuit. However, it is also possible to choose a diode with a low forward bias voltage, which is typically the case for Schottky diodes (Schottky diode symbol in the inset in Fig. 4.27a).

In order to ensure the scalability of the circuit, it is necessary to consider the (activation function) output voltage under various resistive loads, which depend on the connected resistances (weights) connecting the next layer, as previously explained. Since the implementation of the ReLU function is a “passive” circuit, it will be affected by the connected resistive load to a certain degree. In combination with the “active” component of the summing amplifier before the diodes, the impact of the connected resistive load onto the output value can be neglected if the load is high-resistive and does not vary (under weight update) by approximately more than four orders of magnitude. As an alternative, the implementation of a sigmoid-like function by two MOSFETs as an active circuit, can circumvent this problem. Theoretically, the function output here is not affected by the connected resistive load. Similarly, Figure 4.27b illustrates the output for sigmoid-like CMOS implementation with noise induced shift of the output curve. In conclusion, the analysis demonstrates that the proposed SMTJ-based perceptron offers an area and energy efficient analog implementation of a perceptron, aligning well with node perturbed learning due to the inherent randomness provided by SMTJs. The compatibility of the proposed circuitry with extensively scaled neural networks provides potential appeal for industrial devices, particularly as the development of artificial intelligence continues to advance. The

incorporation of an analog computation scheme and integration with CMOS circuits could potentially result in enhanced energy efficiency within the proposed neuromorphic computing architecture. A significant advantage arises from the absence of any additional backpropagation circuitry. This absence enables a simplified neural network architecture, attributable to the local learning rule of node perturbation. The specific number of stochastic MTJs required for each node may vary based on performance and area requirements, which will require further analysis in the future. Nevertheless, the unique and promising aspect of stochastic noise-based local learning combined with SMTJs and CMOS represents an alternative approach to traditional implementations of neural networks for artificial intelligence. Through this work a unique approach to hardware design is given as a basis for neuromorphic devices rooted in local learning concepts.

4. Results

5

Conclusion and Outlook

5.1 Conclusion

In this thesis, magnetic tunnel junctions have been investigated with a focus on superparamagnetic switching at room temperature. The TMR stack used in this work has been optimized and its properties, such as TMR ratio, RA product and exchange bias, have been analyzed as a function of the annealing temperature. The optimal annealing temperature for a high TMR ratio of over 200% with an RA product of $550 \Omega \mu\text{m}^2$ is found for approximately 330°C . TMR stacks have been characterized and optimized via SQUID and CIPT measurements. For the realization of superparamagnetic tunnel junctions, a fabrication process has been developed to achieve 50–100 nm sized junctions. With time series measurements of these MTJ nanopillars it has been shown that superparamagnetic fluctuations on a time scale below 10 ns can be achieved. It is demonstrated that these random fluctuations provide an efficient source for true randomness, if combined with logic XOR gates and thereby passing all randomness tests of the NIST Statistical Test Suite. A proposed circuitry for the generation of multiple independent and true random bitstreams has been outlined, which indicates a high random bit generation rate of approximately 100 Mbit/s per SMTJ and robustness for biased (deviation of 50/50-state) stochastic MTJs. Furthermore, the manipulation of state probability and fluctuation rate has been demonstrated and the analysis of dwell time measurements has allowed one to quantify the effect of spin-transfer-torque as well as Joule heating. At high current densities, Joule heating results in a significant increase in the average switching rate.

The electrical interconnection of two stochastic MTJs revealed a coupling effect, which is based on the TMR and STT effect. It is demonstrated, in simulation as well as experimentally, that the coupling effect depends on the characteristic PV-transfer curve (state probability as a function of applied voltage) of each SMTJ and on the applied source voltage. From stochastic time series measurements,

the coupling strength can be quantified by means of cross-correlation, which may manifest as both positive and negative values. A positive correlation results in a behavior where one MTJ tends to align with the state of the second MTJ, while a negative correlation leads to the opposite effect with a preference of antiparallel alignment. The controllability of this electrical coupling effect has been demonstrated through a variation of the applied source voltage. Experimentally a cross-correlation of 5% is demonstrated and simulations indicate that under sharp PV-transfer curves with high STT effects the correlation can exceed 50%. Furthermore, cross-correlation measurements have confirmed that the coupling remains present for unequal average switching rates both SMTJs.

In the final part of this thesis, a neuromorphic approach with SMTJs is proposed and partially simulated. A neural network based on a perceptron circuit comprising SMTJs, transistors and diodes can perform analog computation and learn through local learning rules known as node perturbation. This provides an alternative to the well established backpropagation, which is biologically implausible and requires distinct computation paths for inference and error propagation. Node perturbation introduces noise via SMTJs to the activation of nodes in the network and thereby solely relies on forward passes without any the backward pass, which typically requires significant computation time and additional circuitry in large neural networks. Thus, a local learning approach via stochastic MTJs might significantly accelerate learning in scaled NNs. The principles for NP via SMTJs together with the circuit has been outlined and discussed and electrical simulations (*LtSpice*) of the proposed perceptron demonstrate the feasibility of this approach and thereby provide a basis for further development of hardware-based learning.

5.2 Outlook

It has been shown that superparamagnetic tunnel junctions are able operate at nanosecond time scales and are a source of uncorrelated high quality randomness. Their intrinsic natural noise generation allows the implementation of SMTJs in hardware, also for largely scaled NP networks. Due to great compatibility of MTJs and CMOS, an area and energy efficient monolithic neuromorphic chip can be envisioned as the ultimate goal for this concept. The efficiency, both in terms of power consumption and learning speed, remains unknown for a real hardware implementation and awaits future analysis. Areas of particular interest include exploring the probability distribution of the required noise, the impact of factors such as the number of SMTJs, fluctuation time scale, and MTJ characteristics

on overall learning dynamics. Introducing local learning in hardware with MTJs as an alternative to traditional backpropagation could be a significant milestone in neuromorphic spintronics. This approach holds great promise for large-scale neural networks, allowing for the simultaneous adjustment of all weights in the network, potentially leading to a drastic acceleration in learning speed. Moreover, the biologically plausible approach to learning in hardware may offer valuable insights into how the brain manages noise during the learning process.

The demonstrated coupling between two MTJs further suggests the potential for a larger network with improved coupling strength, making it intriguing for the development of a potential Boltzmann machine. Enhancing coupling through the amplification of the voltage signal is also a viable option, potentially resulting in very high coupling between SMTJs. In summary, the exploration of these concepts holds promise for their impact on non-conventional computing architectures and might lead to significant advancements in this field.

Publications

- **L. Schnitzspan**, A. Tries, M. Kläui, Electron transport and the effect of current annealing in a two-point contacted hBN/graphene/hBN heterostructure device. *J. Appl. Phys.* **128**, 124302 (2020)
- **L. Schnitzspan**, J. Cramer, J. Kubik, M. Tarequzzaman, G. Jakob, M. Kläui, Impact of annealing temperature on tunneling magnetoresistance multilayer stacks. *IEEE Magn. Lett.* **11**, 1-5 (2020)
- **L. Schnitzspan**, M. Kläui, G. Jakob, Nanosecond true-random-number generation with superparamagnetic tunnel junctions: Identification of Joule heating and spin-transfer-torque effects. *Phys. Rev. Appl.* **20**, 024002 (2023)
- **L. Schnitzspan**, M. Kläui, G. Jakob, Electrical coupling of superparamagnetic tunnel junctions mediated by spin-transfer-torques. *Appl. Phys. Lett.* **132**, 232403 (2023)
- H. Meer, O. Gomonay, C. Schmitt, R. Ramos, **L. Schnitzspan**, F. Kronast, M.-A. Mawass, S. Valencia, E. Saitoh, J. Sinova, L. Baldrati, and M. Kläui, Strain-induced shape anisotropy in antiferromagnetic structures. *Phys. Rev. B* **106**, 094430 (2022)

A | Appendix

A.1 Passivation layer

Procedure for the deposition of the passivation layer SiN_x :

- Heating the sample for 10 min at 120°C in vacuum
- Set the turbo pump to “speed setting mode” (50% of rotation speed)
- Set nitrogen and argon flow respectively to 77.7 sccm to reach a chamber pressure of 0.05 mbar
- Apply 50 W rf-power (13.56 MHz) to the SiN_x target
- Presputter for two minutes
- Deposit 55 nm (≈ 60 min) of SiN_x

A.2 Electron-beam lithography

Electron-beam lithography (*Raith Pioneer*) settings for MTJ nanopillars:

- 30 kV acceleration voltage
- 7.5 μm aperture
- 8 pA beam current
- 8 mm vertical distance
- 10 \times 10 μm writefield for each nanopillar
- Detector: Secondary electron (SE) detector (contrast: 100%)

A.3 Recipes

A.3.1 Nanopillar etch mask

Spin-Coating:

- Dropcasting resist “AR-N7520.07” [187] (0.1 ml)

- Prespin 1 s at 500 rpm
- Spin 60 s at 3000 rpm
- Softbake at 85 °C for 60 s

Exposure & Development:

- E-beam exposure 220 $\mu\text{C}/\text{cm}^2$ at 30 kV
- Development: “AR300-47” [275] for 60 s
- Development break: H₂O rinsing for 20 s
- Drying with nitrogen

A.3.2 Optical lithography

Recipe A

Spin-Coating:

- Dropcasting resist “AR-N7520.17” [187] (0.1 ml)
- Prespin 1 s at 500 rpm
- Spin 60 s at 3000 rpm
- Softbake at 85 °C for 60 s

Exposure & Development:

- UV-exposure for 10 s (contact printing)
- Development: “AR300-47” [275] for 150 s
- Development break: H₂O rinsing for 20 s
- Drying with nitrogen

Removal:

- Acetone 1 min
- Acetone and ultrasound for 1 min
- Isopropanol for 1 min
- H₂O rinsing for 20 s
- Drying with nitrogen

Recipe B

Spin-Coating:

- Dropcasting resist “LOR” [276] (0.1 ml)
- Prespin 1 s at 500 rpm
- Spin 60 s at 3000 rpm

- Softbake at 130 °C for 5 min
- Dropcasting resist “*S1813*” [277] (0.1 ml)
- Prespin 1 s at 500 rpm
- Spin 60 s at 3000 rpm
- Softbake at 100 °C for 2 min

Exposure & Development:

- UV-exposure for 18 s (contact printing)
- Development: “*MF-319*” [278] for 25 s
- Development break: H₂O rinsing for 20 s
- Drying with nitrogen

Removal:

- PG-remover [279] at 130 °C for 30 min
- Ultrasound for 1 min
- Isopropanol for 1 min
- H₂O rinsing for 20 s
- Drying with nitrogen

Bibliography

- [1] Hirohata, A. *et al.* Review on spintronics: Principles and device applications. *J. Magn. Magn. Mater.* **509**, 166711 (2020).
- [2] Awschalom, D. D. & Flatté, M. E. Challenges for semiconductor spintronics. *Nat. Phys.* **3**, 153–159 (2007).
- [3] Baibich, M. N. *et al.* Giant magnetoresistance of (001)Fe/(001)Cr magnetic superlattices. *Phys. Rev. Lett.* **61**, 2472 (1988).
- [4] Binasch, G., Grünberg, P., Saurenbach, F. & Zinn, W. Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange. *Phys. Rev. B* **39**, 4828 (1989).
- [5] Parkin, S. *et al.* Magnetically engineered spintronic sensors and memory. *Proc. IEEE* **91**, 661–680 (2003).
- [6] Parkin, S., Li, Z. & Smith, D. J. Giant magnetoresistance in antiferromagnetic Co/Cu multilayers. *Appl. Phys. Lett.* **58**, 2710–2712 (1991).
- [7] Zhu, J.-G. J. & Park, C. Magnetic tunnel junctions. *Mater. Today* **9**, 36–45 (2006).
- [8] Julliere, M. Tunneling between ferromagnetic films. *Phys. Lett. A* **54**, 225–226 (1975).
- [9] Ikeda, S. *et al.* Tunnel magnetoresistance of 604% at 300K by suppression of Ta diffusion in CoFeB/MgO/CoFeB pseudo-spin-valves annealed at high temperature. *Appl. Phys. Lett.* **93**, 082508 (2008).
- [10] Mao, S. *et al.* Commercial TMR heads for hard disk drives: Characterization and extendibility at 300 Gbit/in². *IEEE Trans. Magn.* **42**, 97–102 (2006).

- [11] Apalkov, D., Dieny, B. & Slaughter, J. M. Magnetoresistive random access memory. *Proc. IEEE* **104**, 1796–1830 (2016).
- [12] Oogane, M. *et al.* Sub-pT magnetic field detection by tunnel magnetoresistive sensors. *Appl. Phys. Express* **14**, 123002 (2021).
- [13] Tavassolizadeh, A. *et al.* Tunnel magnetoresistance sensors with magnetostrictive electrodes: Strain sensors. *Sensors* **16**, 1902 (2016).
- [14] Deak, J. & Jin, I. High-field tunneling magnetoresistive angle sensor. *IEEE Trans. Magn.* **55**, 1–4 (2019).
- [15] Jung, S. *et al.* A crossbar array of magnetoresistive memory devices for in-memory computing. *Nature* **601**, 211–216 (2022).
- [16] Goodwill, J. M. *et al.* Implementation of a binary neural network on a passive array of magnetic tunnel junctions. *Phys. Rev. Appl.* **18**, 014039 (2022).
- [17] Marković, D., Mizrahi, A., Querlioz, D. & Grollier, J. Physics for neuro-morphic computing. *Nat. Rev. Phys.* **2**, 499–510 (2020).
- [18] Moore, G. E. *et al.* Progress in digital integrated electronics. In *Electron devices meeting*, vol. 21, 11–13 (Washington, DC, 1975).
- [19] Waldrop, M. M. More than moore. *Nature* **530**, 144–148 (2016).
- [20] Aggarwal, C. C. *Neural networks and deep learning* (Springer, 2018).
- [21] LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
- [22] Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J. & Hinton, G. Backpropagation and the brain. *Nat. Rev. Neurosci.* **21**, 335–346 (2020).
- [23] Schuman, C. D. *et al.* Opportunities for neuromorphic computing algorithms and applications. *Nat. Comput. Sci.* **2**, 10–19 (2022).
- [24] Furber, S. Large-scale neuromorphic computing systems. *J. Neural Eng.* **13**, 051001 (2016).
- [25] Grollier, J. *et al.* Neuromorphic spintronics. *Nat. Electron.* **3**, 360–370 (2020).
- [26] Borders, W. A. *et al.* Integer factorization using stochastic magnetic tunnel junctions. *Nature* **573**, 390–393 (2019).

- [27] Niazi, S. *et al.* Training deep Boltzmann networks with sparse Ising machines. *arXiv:2303.10728* (2023).
- [28] Sengupta, A., Panda, P., Wijesinghe, P., Kim, Y. & Roy, K. Magnetic tunnel junction mimics stochastic cortical spiking neurons. *Sci. Rep.* **6**, 1–8 (2016).
- [29] Mondal, A. & Srivastava, A. Energy-efficient design of MTJ-based neural networks with stochastic computing. *ACM J. Emerg. Technol. Comput. Syst.* **16**, 1–27 (2019).
- [30] Liu, S. *et al.* Bayesian neural networks using magnetic tunnel junction-based probabilistic in-memory computing. *Front. Nanotechnol.* **4**, 1021943 (2022).
- [31] Ganguly, S., Camsari, K. Y. & Ghosh, A. W. Reservoir computing using stochastic p-bits. *arXiv:1709.10211* (2017).
- [32] Kaiser, J. *et al.* Hardware-aware in situ learning based on stochastic magnetic tunnel junctions. *Phys. Rev. Appl.* **17**, 014016 (2022).
- [33] Haken, H. & Wolf, H. C. *Atom-und Quantenphysik: Einführung in die experimentellen und theoretischen Grundlagen* (Springer-Verlag, 2013).
- [34] Coey, J. M. D. *Magnetism and magnetic materials* (Cambridge university press, Cambridge, 2010).
- [35] Gross, R. & Marx, A. *Festkörperphysik* (Oldenbourg Wissenschaftsverlag, München, 2012).
- [36] Bergmann, L., Schaefer, C. & Blügel, S. *Festkörper* (de Gruyter, Berlin, 2005).
- [37] Coulomb, C. A. Premier mémoire sur l'électricité et le magnétisme. *Hist. Acad. Roy. Sci.* **569** (1785).
- [38] Pauli, W. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. *Zeitschrift für Physik* **31**, 765–783 (1925).
- [39] Stoner, E. C. Collective electron ferromagnetism. *Proc. Math. Phys. Eng. Sci.* **165**, 372–414 (1938).
- [40] Blundell, S. *Magnetism in condensed matter* (OUP Oxford, 2001).

- [41] Hubert, A. & Schäfer, R. *Magnetic domains. The analysis of magnetic microstructures* (Springer, 2009).
- [42] Cullity, B. D. & Graham, C. D. *Introduction to magnetic materials* (John Wiley & Sons, 2011).
- [43] Cochran, J., Heinrich, B. & Arrott, A. Ferromagnetic resonance in a system composed of a ferromagnetic substrate and an exchange-coupled thin ferromagnetic overlayer. *Phys. Rev. B* **34**, 7788 (1986).
- [44] Watanabe, K., Jinnai, B., Fukami, S., Sato, H. & Ohno, H. Shape anisotropy revisited in single-digit nanometer magnetic tunnel junctions. *Nat. Commun.* **9**, 663 (2018).
- [45] Stoner, E. C. Xcvii. the demagnetizing factors for ellipsoids. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **36**, 803–821 (1945).
- [46] Yang, H. *et al.* First-principles investigation of the very large perpendicular magnetic anisotropy at Fe—MgO and Co—MgO interfaces. *Phys. Rev. B* **84**, 054401 (2011).
- [47] Dieny, B. & Chshiev, M. Perpendicular magnetic anisotropy at transition metal/oxide interfaces and applications. *Rev. Mod. Phys.* **89**, 025008 (2017).
- [48] Sato, H. *et al.* Perpendicular-anisotropy CoFeB-MgO magnetic tunnel junctions with a MgO/CoFeB/Ta/CoFeB/MgO recording structure. *Appl. Phys. Lett.* **101** (2012).
- [49] Darby, M. & Isaac, E. Magnetocrystalline anisotropy of ferro- and ferrimagnetics. *IEEE Trans. Magn.* **10**, 259–304 (1974).
- [50] Stoner, E. C. & Wohlfarth, E. A mechanism of magnetic hysteresis in heterogeneous alloys. *Philos. Trans. R. Soc. Lond. A* **240**, 599–642 (1948).
- [51] Tiusan, C., Hehn, M. & Ounadjela, K. Magnetic-roughness-induced magnetostatic interactions in magnetic tunnel junctions. *Eur. Phys. J. B* **26**, 431–434 (2002).
- [52] Néel, L. Sur un nouveau mode de couplage entre les aimantations de deux couches minces ferromagnétiques. *C. R. Acad. Sci. Paris ou CRAS* **255**, 1676–1681 (1962).
- [53] Tegen, S., Mönch, I., Schumann, J., Vinzelberg, H. & Schneider, C. Effect of Néel coupling on magnetic tunnel junctions. *J. Appl. Phys.* **89**, 8169–8174 (2001).

- [54] Zhang, J. & White, R. Topological coupling in magnetic multilayer films. *J. Appl. Phys.* **79**, 5113–5115 (1996).
- [55] Schrag, B. *et al.* Néel “orange-peel” coupling in magnetic tunneling junction devices. *Appl. Phys. Lett.* **77**, 2373–2375 (2000).
- [56] Costa, J. *et al.* Impact of MgO thickness on the performance of spin-transfer torque nano-oscillators. *IEEE Trans. Magn.* **51**, 1–4 (2015).
- [57] Anguelouch, A. *et al.* Two-dimensional magnetic switching of micron-size films in magnetic tunnel junctions. *Appl. Phys. Lett.* **76**, 622–624 (2000).
- [58] Kools, J., Kula, W., Mauri, D. & Lin, T. Effect of finite magnetic film thickness on Néel coupling in spin valves. *J. Appl. Phys.* **85**, 4466–4468 (1999).
- [59] Nogués, J. & Schuller, I. K. Exchange bias. *J. Magn. Magn. Mater.* **192**, 203–232 (1999).
- [60] Stiles, M. D. & McMichael, R. D. Model for exchange bias in polycrystalline ferromagnet-antiferromagnet bilayers. *Phys. Rev. B* **59**, 3722 (1999).
- [61] Chang, H. *et al.* Exchange bias in Co/MnPt polycrystalline films on Si(100)/SiO₂ substrates with Ta underlayer. *Thin Solid Films* **660**, 834–839 (2018).
- [62] Stiles, M. D. Interlayer exchange coupling. *J. Magn. Magn. Mater.* **200**, 322–337 (1999).
- [63] Duine, R., Lee, K.-J., Parkin, S. S. & Stiles, M. D. Synthetic antiferromagnetic spintronics. *Nat. Phys.* **14**, 217–219 (2018).
- [64] Kasuya, T. A theory of metallic ferro- and antiferromagnetism on Zener’s model. *Prog. Theor. Phys.* **16**, 45–57 (1956).
- [65] Parkin, S., More, N. & Roche, K. Oscillations in exchange coupling and magnetoresistance in metallic superlattice structures: Co/Ru, Co/Cr, and Fe/Cr. *Phys. Rev. Lett.* **64**, 2304 (1990).
- [66] Parkin, S. & Mauri, D. Spin engineering: Direct determination of the Ruderman-Kittel-Kasuya-Yosida far-field range function in ruthenium. *Phys. Rev. B* **44**, 7131 (1991).
- [67] Bruno, P. Theory of interlayer magnetic coupling. *Phys. Rev. B* **52**, 411 (1995).

- [68] Ruderman, M. A. & Kittel, C. Indirect exchange coupling of nuclear magnetic moments by conduction electrons. *Phys. Rev.* **96**, 99 (1954).
- [69] Yosida, K. Magnetic properties of Cu-Mn alloys. *Phys. Rev.* **106**, 893 (1957).
- [70] Bloemen, P. *et al.* Magnetic layer thickness dependence of the interlayer exchange coupling in (001)Co/Cu/Co. *Phys. Rev. Lett.* **72**, 764 (1994).
- [71] Faure-Vincent, J. *et al.* Interlayer magnetic coupling interactions of two ferromagnetic layers by spin polarized tunneling. *Phys. Rev. Lett.* **89**, 107206 (2002).
- [72] Katayama, T. *et al.* Interlayer exchange coupling in Fe/MgO/Fe magnetic tunnel junctions. *Appl. Phys. Lett.* **89** (2006).
- [73] Nistor, L. E. *et al.* Oscillatory interlayer exchange coupling in MgO tunnel junctions with perpendicular magnetic anisotropy. *Phys. Rev. B* **81**, 220407 (2010).
- [74] Tsymbal, E. Y., Mryasov, O. N. & LeClair, P. R. Spin-dependent tunnelling in magnetic tunnel junctions. *J. Condens. Matter Phys.* **15**, R109 (2003).
- [75] Zhang, X. & Butler, W. Band structure, evanescent states, and transport in spin tunnel junctions. *J. Condens. Matter Phys.* **15**, R1603 (2003).
- [76] Butler, W., Zhang, X.-G., Schulthess, T. & MacLaren, J. Spin-dependent tunneling conductance of Fe—MgO—Fe sandwiches. *Phys. Rev. B* **63**, 054416 (2001).
- [77] Matsumoto, R. *et al.* Oscillation of giant tunneling magnetoresistance with respect to tunneling barrier thickness in fully epitaxial Fe/MgO/Fe magnetic tunnel junctions. *Appl. Phys. Lett.* **90** (2007).
- [78] Zhang, X.-G. & Butler, W. Large magnetoresistance in bcc Co/MgO/Co and FeCo/MgO/FeCo tunnel junctions. *Phys. Rev. B* **70**, 172407 (2004).
- [79] Isogami, S. *et al.* In situ heat treatment of ultrathin MgO layer for giant magnetoresistance ratio with low resistance area product in CoFeB/MgO/CoFeB magnetic tunnel junctions. *Appl. Phys. Lett.* **93** (2008).
- [80] Yuasa, S., Suzuki, Y., Katayama, T. & Ando, K. Characterization of growth and crystallization processes in CoFeB/MgO/CoFeB magnetic tunnel junc-

- tion structure by reflective high-energy electron diffraction. *Appl. Phys. Lett.* **87** (2005).
- [81] Ikeda, S. *et al.* Dependence of tunnel magnetoresistance in MgO based magnetic tunnel junctions on Ar pressure during MgO sputtering. *Jpn. J. Appl. Phys.* **44**, L1442 (2005).
- [82] Costa, J. *et al.* High power and low critical current density spin transfer torque nano-oscillators using MgO barriers with intermediate thickness. *Sci. Rep.* **7**, 7237 (2017).
- [83] Simmons, J. G. Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film. *J. Appl. Phys.* **34**, 1793–1803 (1963).
- [84] Simmons, J. G. Generalized thermal J-V characteristic for the electric tunnel effect. *J. Appl. Phys.* **35**, 2655–2658 (1964).
- [85] Yuan, L., Liou, S.-H. & Wang, D. Temperature dependence of magnetoresistance in magnetic tunnel junctions with different free layer structures. *Phys. Rev. B* **73**, 134403 (2006).
- [86] Drewello, V. *et al.* Inelastic electron tunneling spectra of MgO-based magnetic tunnel junctions with different electrode designs. *Phys. Rev. B* **79**, 174417 (2009).
- [87] Brinkman, W., Dynes, R. & Rowell, J. Tunneling conductance of asymmetrical barriers. *J. Appl. Phys.* **41**, 1915–1921 (1970).
- [88] Hauser, J. & Testardi, L. Tunneling and band structure in semimetals. *Phys. Rev. Lett.* **20**, 12 (1968).
- [89] Oepts, W., Gillies, M., Coehoorn, R., Van de Veerdonk, R. & De Jonge, W. Asymmetric bias voltage dependence of the magnetoresistance of Co/Al₂O₃/Co magnetic tunnel junctions: Variation with the barrier oxidation time. *J. Appl. Phys.* **89**, 8038–8045 (2001).
- [90] Gilbert, T. L. A phenomenological theory of damping in ferromagnetic materials. *IEEE Trans. Magn.* **40**, 3443–3449 (2004).
- [91] Kanai, S., Hayakawa, K., Ohno, H. & Fukami, S. Theory of relaxation time of stochastic nanomagnets. *Phys. Rev. B* **103**, 094423 (2021).

- [92] Landau, L. & Lifshitz, E. On the theory of the dispersion of magnetic permeability in ferromagnetic bodies. In *Phys. Perspect.*, 51–65 (Elsevier, 1992).
- [93] Brown Jr, W. F. Thermal fluctuations of a single-domain particle. *Phys. Rev.* **130**, 1677 (1963).
- [94] Oezelt, H. *et al.* Full-spin-wave-scaled stochastic micromagnetism for mesh-independent simulations of ferromagnetic resonance and reversal. *npj Comput. Mater.* **8**, 35 (2022).
- [95] Slonczewski, J. C. Current-driven excitation of magnetic multilayers. *J. Magn. Magn.* **159**, L1–L7 (1996).
- [96] Berger, L. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B* **54**, 9353 (1996).
- [97] Dowben, P. A., Wu, N. & Binek, C. When measured spin polarization is not spin polarization. *J. Condens. Matter Phys.* **23**, 171001 (2011).
- [98] Velev, J. P., Dowben, P. A., Tsymbal, E. Y., Jenkins, S. J. & Caruso, A. Interface effects in spin-polarized metal/insulator layered structures. *Surf. Sci. Rep.* **63**, 400–425 (2008).
- [99] Shpiro, A., Levy, P. M. & Zhang, S. Self-consistent treatment of nonequilibrium spin torques in magnetic multilayers. *Phys. Rev. B* **67**, 104430 (2003).
- [100] Brataas, A., Kent, A. D. & Ohno, H. Current-induced torques in magnetic materials. *Nat. Mater.* **11**, 372–381 (2012).
- [101] Ralph, D. C. & Stiles, M. D. Spin transfer torques. *J. Magn. Magn. Mater.* **320**, 1190–1216 (2008).
- [102] Slonczewski, J. Excitation of spin waves by an electric current. *J. Magn. Magn.* **195**, L261–L268 (1999).
- [103] Zhang, S. & Li, Z. Roles of nonequilibrium conduction electrons on the magnetization dynamics of ferromagnets. *Phys. Rev. Lett.* **93**, 127204 (2004).
- [104] Gapihan, E. *et al.* Heating asymmetry induced by tunneling current flow in magnetic tunnel junctions. *Appl. Phys. Lett.* **100** (2012).
- [105] Kan, J. J. *et al.* A study on practically unlimited endurance of STT-MRAM. *IEEE Trans. on Electron Devices* **64**, 3639–3646 (2017).

- [106] Schnitzspan, L., Kläui, M. & Jakob, G. Nanosecond true-random-number generation with superparamagnetic tunnel junctions: Identification of Joule heating and spin-transfer-torque effects. *Phys. Rev. Appl.* **20**, 024002 (2023).
- [107] Daniels, M. W., Madhavan, A., Talatchian, P., Mizrahi, A. & Stiles, M. D. Energy-efficient stochastic computing with superparamagnetic tunnel junctions. *Phys. Rev. Appl.* **13**, 034016 (2020).
- [108] Mizrahi, A. *et al.* Neural-like computing with populations of superparamagnetic basis functions. *Nat. Commun.* **9**, 1–11 (2018).
- [109] Néel, L. Théorie du traînage magnétique des ferromagnétiques en grains fins avec application aux terres cuites **5**, 99–136 (1949).
- [110] Li, Z. & Zhang, S. Thermally assisted magnetization reversal in the presence of a spin-transfer torque. *Phys. Rev. B* **69**, 134416 (2004).
- [111] Kitakami, O., Shimatsu, T., Okamoto, S., Shimada, Y. & Aoi, H. Sharrock relation for perpendicular recording media with higher-order magnetic anisotropy terms. *Jpn. J. Appl. Phys.* **43**, L115 (2003).
- [112] Funatsu, T., Kanai, S., Ieda, J., Fukami, S. & Ohno, H. Local bifurcation with spin-transfer torque in superparamagnetic tunnel junctions. *Nat. Commun.* **13**, 4079 (2022).
- [113] Meo, A. *et al.* Spin transfer torque switching dynamics in CoFeB/MgO magnetic tunnel junctions. *Phys. Rev. B* **103**, 054426 (2021).
- [114] Sun, J. Z. Spin-current interaction with a monodomain magnetic body: A model study. *Phys. Rev. B* **62**, 570 (2000).
- [115] Koch, R., Katine, J. & Sun, J. Time-resolved reversal of spin-transfer switching in a nanomagnet. *Phys. Rev. Lett.* **92**, 088302 (2004).
- [116] Zink, B. R., Lv, Y. & Wang, J.-P. Independent control of antiparallel- and parallel-state thermal stability factors in magnetic tunnel junctions for telegraphic signals with two degrees of tunability. *IEEE Trans. Electron Devices* **66**, 5353–5359 (2019).
- [117] Zhao, H. *et al.* Low writing energy and sub nanosecond spin torque transfer switching of in-plane magnetic tunnel junction for spin torque transfer random access memory. *J. Appl. Phys.* **109** (2011).

- [118] Rizzo, N. *et al.* Thermally activated magnetization reversal in submicron magnetic tunnel junctions for magnetoresistive random access memory. *Appl. Phys. Lett.* **80**, 2335–2337 (2002).
- [119] Diao, Z. *et al.* Spin transfer switching and spin polarization in magnetic tunnel junctions with MgO and AlOx barriers. *Appl. Phys. Lett.* **87** (2005).
- [120] Li, Z. *et al.* Perpendicular spin torques in magnetic tunnel junctions. *Phys. Rev. Lett.* **100**, 246602 (2008).
- [121] Sun, J. & Ralph, D. Magnetoresistance and spin-transfer torque in magnetic tunnel junctions. *J. Magn. Magn. Mater.* **320**, 1227–1237 (2008).
- [122] Gong, L., Zhang, J., Liu, H., Sang, L. & Wang, Y. True random number generators using electrical noise. *IEEE Access* **7**, 125796–125805 (2019).
- [123] Kollmitzer, C., Petscharnig, S., Suda, M. & Mehic, M. *Quantum random number generation* (Springer, 2020).
- [124] Li, Y. *et al.* Quantum random number generator using a cloud superconducting quantum computer based on source-independent protocol. *Sci. Rep.* **11**, 23873 (2021).
- [125] Bassham III, L. E. *et al.* *Sp 800-22 rev. 1a. a statistical test suite for random and pseudorandom number generators for cryptographic applications* (National Institute of Standards & Technology, 2010).
- [126] Marton, K. & Suciú, A. On the interpretation of results from the NIST statistical test suite. *Sci. Technol.* **18**, 18–32 (2015).
- [127] Mosteller, F. & Fisher, R. A. Questions and answers. *Am. Stat.* **2**, 30–31 (1948).
- [128] Von Neumann, J. Various techniques used in connection with random digits. *Appl. Math Ser.* **12**, 3 (1951).
- [129] Stipčević, M. & Koç, Ç. K. True random number generators. In *Open Problems in Mathematics and Computational Science*, 275–315 (Springer, 2014).
- [130] Alaghi, A., Qian, W. & Hayes, J. P. The promise and challenge of stochastic computing. *IEEE Trans. CAD* **37**, 1515–1531 (2017).
- [131] Compagnoni, C. M. *et al.* Statistical model for random telegraph noise in flash memories. *IEEE Trans. Electron Devices* **55**, 388–395 (2007).

- [132] Hung, K. K., Ko, P. K., Hu, C.-M. & Cheng, Y. C. Random telegraph noise of deep-submicrometer MOSFET's. *IEEE Electron Device Lett.* **11**, 90–92 (1990).
- [133] Grasser, T. Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities. *Microelectron. Reliab.* **52**, 39–70 (2012).
- [134] Vincent, A. F. *et al.* Analytical macrospin modeling of the stochastic switching time of spin-transfer torque devices. *IEEE Trans. Electron Devices* **62**, 164–170 (2014).
- [135] Weißgerber, W. *Elektrotechnik für Ingenieure 1* (Springer, Wiesbaden, 2018).
- [136] Naaman, O. & Aumentado, J. Poisson transition rates from time-domain measurements with a finite bandwidth. *Phys. Rev. Lett.* **96**, 100201 (2006).
- [137] Christensen, D. V. *et al.* 2022 roadmap on neuromorphic computing and engineering. *Neuromorphic Comput. Eng.* **2**, 022501 (2022).
- [138] Rasamoelina, A. D., Adjailia, F. & Sinčák, P. A review of activation function for artificial neural network. In *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMII)*, 281–286 (IEEE, 2020).
- [139] Hochreiter, S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *Int. J. Uncertain. Fuzziness Knowledge-Based Syst.* **6**, 107–116 (1998).
- [140] Lu, L., Shin, Y., Su, Y. & Karniadakis, G. E. Dying ReLU and initialization: Theory and numerical examples. *arXiv:1903.06733* (2019).
- [141] Maas, A. L., Hannun, A. Y., Ng, A. Y. *et al.* Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml*, vol. 30, 3 (Atlanta, GA, 2013).
- [142] Clevert, D.-A., Unterthiner, T. & Hochreiter, S. Fast and accurate deep network learning by exponential linear units (ELUs). *arXiv:1511.07289* (2015).
- [143] Pietrolaj, M. & Blok, M. Neural network training with limited precision and asymmetric exponent. *J.Big Data* **9**, 63 (2022).

- [144] Sun, X. *et al.* Ultra-low precision 4-bit training of deep neural networks. *Adv. Neural Inf. Process.* **33**, 1796–1807 (2020).
- [145] Nøkland, A. Direct feedback alignment provides learning in deep neural networks. *Adv. Neural Inf. Process.* **29** (2016).
- [146] Do, H. The organization of behavior. *New York* (1949).
- [147] Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
- [148] Hiratani, N., Mehta, Y., Lillicrap, T. & Latham, P. E. On the stability and scalability of node perturbation learning. *Adv. Neural Inf. Process.* **35**, 31929–31941 (2022).
- [149] Dalm, S., van Gerven, M. & Ahmad, N. Effective learning with node perturbation in deep neural networks. *arXiv:2310.00965* (2023).
- [150] Jabri, M. & Flower, B. Weight perturbation: An optimal architecture and learning technique for analog VLSI feedforward and recurrent multilayer networks. *IEEE Trans. Neural Netw.* **3**, 154–157 (1992).
- [151] Flower, B. & Jabri, M. Summed weight neuron perturbation: An $O(n)$ improvement over weight perturbation. *Adv. Neural Inf. Process.* **5** (1992).
- [152] London, M., Roth, A., Beeren, L., Häusser, M. & Latham, P. E. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature* **466**, 123–127 (2010).
- [153] Lillicrap, T. P., Cownden, D., Tweed, D. B. & Akerman, C. J. Random synaptic feedback weights support error backpropagation for deep learning. *Nat. Commun.* **7**, 13276 (2016).
- [154] Cauwenberghs, G. A fast stochastic error-descent algorithm for supervised learning and optimization. *Adv. Neural. Inf. Process.* **5** (1992).
- [155] Ahmad, N., Schrader, E. & van Gerven, M. Constrained parameter inference as a principle for learning. *Transact. Mach. Learn. Res.* (2023). URL <https://openreview.net/forum?id=CUDdbTT1QC>.
- [156] Schnitzspan, L. *Charge Transport in Two Dimensional Carbon Heterostructures*. Master’s thesis, Universitätsbibliothek Mainz (2019).
- [157] Binnig, G., Quate, C. F. & Gerber, C. Atomic force microscope. *Phys. Rev. Lett.* **56**, 930 (1986).

- [158] Eaton, P. & West, P. *Atomic force microscopy* (Oxford university press, Oxford, 2010).
- [159] Jones, J. E. On the determination of molecular fields.—II. from the equation of state of a gas. *Proc. R. Soc. Lond. A* **106**, 463–477 (1924).
- [160] Capres, K. Capres CIPTech-M200 (2023). URL www.capres.com/Default.aspx?ID=23.
- [161] Worledge, D. & Trouilloud, P. Magnetoresistance measurement of unpatterned magnetic tunnel junction wafers by current-in-plane tunneling. *Appl. Phys. Lett.* **83**, 84–86 (2003).
- [162] Von Ardenne, M. Das Elektronen-Rastermikroskop: Theoretische Grundlagen. *Zeitschrift für Physik* **109**, 553–572 (1938).
- [163] Oatley, C. W. The early history of the scanning electron microscope. *J. Appl. Phys.* **53**, R1–R13 (1982).
- [164] Zhou, W. & Wang, Z. L. *Scanning microscopy for nanotechnology: techniques and applications* (Springer, New York, 2007).
- [165] Abbe, E. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für mikroskopische Anatomie* **9**, 413–468 (1873).
- [166] Raith GmbH. Electron-beam lithography system (2023). URL <https://raith.com/product/pioneer-two/>.
- [167] Fisher Scientific. *Helios NanoLab 650/600i User Operation Manual* (2023). URL www.thermofisher.com.
- [168] Auger, P. Sur les rayons β secondaires produits dans un gaz par des rayons X. *CR Acad. Sci.(F)* **177**, 169 (1923).
- [169] Meitner, L. Über die β -Strahl-Spektren und ihren Zusammenhang mit der γ -Strahlung. *Zeitschrift für Physik* **11**, 35–54 (1922).
- [170] Josephson, B. D. Possible new effects in superconductive tunnelling. *Phys. Lett.* **1**, 251–253 (1962).
- [171] Buchner, M., Höfler, K., Henne, B., Ney, V. & Ney, A. Tutorial: Basic principles, limits of detection, and pitfalls of highly sensitive squid magnetometry for nanomagnetism and spintronics. *J. Appl. Phys.* **124** (2018).
- [172] Clarke, J. & Braginski, A. I. *The SQUID handbook*, vol. 1 (Wiley Online Library, 2004).

- [173] Birkholz, M. *Thin film analysis by X-ray scattering* (John Wiley & Sons, 2006).
- [174] Björck, M. & Andersson, G. GenX: an extensible X-ray reflectivity refinement program utilizing differential evolution. *J. Appl. Crystallogr.* **40**, 1174–1178 (2007).
- [175] Glavic, A. & Björck, M. Genx 3: the latest generation of an established tool. *J. Appl. Crystallogr.* **55** (2022).
- [176] Parratt, L. G. Surface studies of solids by total reflection of X-rays. *Phys. Rev.* **95**, 359 (1954).
- [177] Pfeiffer, A. *Pure spin current transport and magnetic state manipulation in lateral spin valves*. Ph.D. thesis, Universitätsbibliothek Mainz (2020).
- [178] Wasa, K., Kanno, I. & Kotera, H. *Handbook of sputter deposition technology: fundamentals and applications for functional thin films, nano-materials and MEMS* (William Andrew, 2012).
- [179] Saha, R. & Barnard, J. A. Effect of structure on the mechanical properties of Ta and Ta(N) thin films prepared by reactive DC magnetron sputtering. *J. Cryst. Growth* **174**, 495–500 (1997).
- [180] Alagoz, A. S., Kamminga, J.-D., Grachev, S. Y., Lu, T.-M. & Karabacak, T. Residual stress reduction in sputter deposited thin films by density modulation. *MRS Online Proceedings Library (OPL)* **1224** (2009).
- [181] Hayakawa, J., Ikeda, S., Matsukura, F., Takahashi, H. & Ohno, H. Dependence of giant tunnel magnetoresistance of sputtered CoFeB/MgO/CoFeB magnetic tunnel junctions on MgO barrier thickness and annealing temperature. *Jpn. J. Appl. Phys.* **44**, L587 (2005).
- [182] Schnitzspan, L. *et al.* Impact of annealing temperature on tunneling magnetoresistance multilayer stacks. *IEEE Magn. Lett.* **11**, 1–5 (2020).
- [183] Feng, G. *et al.* Annealing of CoFeB/MgO based single and double barrier magnetic tunnel junctions: Tunnel magnetoresistance, bias dependence, and output voltage. *J. Appl. Phys.* **105** (2009).
- [184] Chikazumi, S. & Oomura, T. On the origin of magnetic anisotropy induced by magnetic annealing. *J. Phys. Soc. Jpn.* **10**, 842–849 (1955).
- [185] Ito, T. & Okazaki, S. Pushing the limits of lithography. *Nature* **406**, 1027–1031 (2000).

- [186] Mohammad, M. A., Muhammad, M., Dew, S. K. & Stepanova, M. Fundamentals of electron beam exposure and development. In *Nanofabrication: techniques and principles*, 11–41 (Springer, 2011).
- [187] Allresist. Negative E-beam resist: AR-N7520.17 (2023). URL www.allresist.com/wp-content/uploads/sites/2/2022/03/Allresist_Product-information-E-Beamresist-AR-N-7520new-English-web.pdf.
- [188] Min, Z., Baoqin, C., Changqing, X., Ming, L. & Jiebing, N. Study of process of HSQ in electron beam lithography. In *2010 IEEE 5th International Conference on Nano/Micro Engineered and Molecular Systems*, 1021–1024 (IEEE, 2010).
- [189] Tavakkoli KG, A., Piramanayagam, S., Ranjbar, M., Sbiaa, R. & Chong, T. Path to achieve sub-10-nm half-pitch using electron beam lithography. *J. Vac. Sci. Technol. B* **29** (2011).
- [190] Guerfi, Y., Carcenac, F. & Larrieu, G. High resolution HSQ nanopillar arrays with low energy electron beam lithography. *Microelectron. Eng.* **110**, 173–176 (2013).
- [191] Vieu, C. *et al.* Electron beam lithography: resolution limits and applications. *Appl. Surf. Sci.* **164**, 111–117 (2000).
- [192] Broers, A. N., Hoole, A. & Ryan, J. M. Electron beam lithography—resolution limits. *Microelectron. Eng.* **32**, 131–142 (1996).
- [193] Duan, H., Yang, J. K. & Berggren, K. K. Controlled collapse of high-aspect-ratio nanostructures. *Small* **7**, 2661–2668 (2011).
- [194] Kim, I., Mun, J., Hwang, W., Yang, Y. & Rho, J. Capillary-force-induced collapse lithography for controlled plasmonic nanogap structures. *Microssyst. Nanoeng.* **6**, 65 (2020).
- [195] Meyer Burger. IonSys - Model 500 - Ion Beam Deposition System (2023). URL www.energy-xprt.com/downloads/ionsys-model-500-ion-beam-deposition-system-brochure-656413.
- [196] Ji, M. *et al.* Study on the effect of re-deposition induced by ion beam etching on MTJ performances. *AIP Adv.* **9** (2019).

- [197] Wei, Q., Li, K.-D., Lian, J. & Wang, L. Angular dependence of sputtering yield of amorphous and polycrystalline materials. *J. Phys. D: Appl. Phys.* **41**, 172002 (2008).
- [198] Lee, R. E. Microfabrication by ion-beam etching. *J. Vac. Sci. Technol.* **16**, 164–170 (1979).
- [199] Chun, S.-w. *et al.* Multi-step ion beam etching of sub-30 nm magnetic tunnel junctions for reducing leakage and MgO barrier damage. *J. Appl. Phys.* **111** (2012).
- [200] Islam, R., Cui, B. & Miao, G.-X. Dry etching strategy of spin-transfer-torque magnetic random access memory: A review. *J. Vac. Sci. Technol. B* **38** (2020).
- [201] Sugiura, K. *et al.* Ion beam etching technology for high-density spin transfer torque magnetic random access memory. *Jpn. J. Appl. Phys.* **48**, 08HD02 (2009).
- [202] Benninghoven, A., Rudenauer, F. & Werner, H. W. *Secondary ion mass spectrometry: basic concepts, instrumental aspects, applications and trends* (John Wiley and Sons, New York, 1987).
- [203] Guruvenket, S., Ghatak, J., Satyam, P. & Rao, G. M. Characterization of bias magnetron-sputtered silicon nitride films. *Thin Solid Films* **478**, 256–260 (2005).
- [204] Rothmund, W. & Fritzsche, C. Dielectric breakdown of reactively sputtered silicon nitride. *Thin Solid Films* **15**, 199–205 (1973).
- [205] Vila, M., Cáceres, D. & Prieto, C. Mechanical properties of sputtered silicon nitride thin films. *J. Appl. Phys.* **94**, 7868–7873 (2003).
- [206] Hu, S. & Gregor, L. Silicon nitride films by reactive sputtering. *J. Electrochem. Soc.* **114**, 826 (1967).
- [207] Vila, M., Prieto, C. & Ramirez, R. Electrical behavior of silicon nitride sputtered thin films. *Thin Solid Films* **459**, 195–199 (2004).
- [208] Kim, J. H. & Chung, K. W. Microstructure and properties of silicon nitride thin films deposited by reactive bias magnetron sputtering. *J. Appl. Phys.* **83**, 5831–5839 (1998).
- [209] Surana, V. *et al.* Realization of high quality silicon nitride deposition at low temperatures. *J. Appl. Phys.* **126** (2019).

- [210] Ding, Y. *et al.* Fabrication of current-induced magnetization switching devices using etch-back planarization process. *J. Appl. Phys.* **97**, 10C702 (2005).
- [211] Marathe, A. P., Meng, X., Hebert, D. F., Nagai, Y. & Van Duzer, T. Planarization techniques for multilevel HTS integrated circuit process. *IEEE Trans. Appl. Supercond.* **3**, 2373–2376 (1993).
- [212] Krishnan, M., Nalaskowski, J. W. & Cook, L. M. Chemical mechanical planarization: slurry chemistry, materials, and mechanisms. *Chem. Rev.* **110**, 178–204 (2010).
- [213] Zhao, D. & Lu, X. Chemical mechanical polishing: theory and experiment. *Friction* **1**, 306–326 (2013).
- [214] Analog Devices. LTspice information center. www.analog.com/en/design-center/design-tools-and-calculators/ltspice-simulator.html (2023).
- [215] Loong, L. M. *et al.* Strain-enhanced tunneling magnetoresistance in MgO magnetic tunnel junctions. *Sci. Rep.* **4**, 6505 (2014).
- [216] Almasi, H. *et al.* Perpendicular magnetic tunnel junction with W seed and capping layers. *J. Appl. Phys.* **121** (2017).
- [217] Meiklejohn, W. H. & Bean, C. P. New magnetic anisotropy. *Phys. Rev.* **102**, 1413 (1956).
- [218] MicroChemicals GmbH. Si/SiO₂ wafer (2023). URL www.shop.microchemicals.com/produkte/shop/wafer/beschichtete-wafer/si-thermisches-sio2/si-sio2-wafer-wwd40525250b1314sxx3?c=376.
- [219] Ren, H. & Sosnowski, M. Tantalum thin films deposited by ion assisted magnetron sputtering. *Thin Solid Films* **516**, 1898–1905 (2008).
- [220] Morrow, P. *et al.* Texture of Ru columns grown by oblique angle sputter deposition. *J. Vac. Sci. Technol. A* **24**, 235–245 (2006).
- [221] Thornton, J. A. High rate thick film growth. *Annu. Rev. Mater. Sci.* **7**, 239–260 (1977).
- [222] Duine, R., Lee, K.-J., Parkin, S. S. & Stiles, M. D. Synthetic antiferromagnetic spintronics. *Nat. Phys.* **14**, 217–219 (2018).

- [223] Mukherjee, S. S. *et al.* Crystallization and grain growth behavior of CoFeB and MgO layers in multilayer magnetic tunnel junctions. *J. Appl. Phys.* **106** (2009).
- [224] Parkin, S. S. *et al.* Giant tunnelling magnetoresistance at room temperature with MgO(100) tunnel barriers. *Nat. Mater.* **3**, 862–867 (2004).
- [225] Tomczak, Y. *et al.* Thin Co/Ni-based bottom pinned spin-transfer torque magnetic random access memory stacks with high annealing tolerance. *Appl. Phys. Lett.* **108** (2016).
- [226] Cardoso, S., Freitas, P., De Jesus, C., Wei, P. & Soares, J. Spin-tunnel-junction thermal stability and interface interdiffusion above 300°C. *Appl. Phys. Lett.* **76**, 610–612 (2000).
- [227] Lee, Y. M., Hayakawa, J., Ikeda, S., Matsukura, F. & Ohno, H. Giant tunnel magnetoresistance and high annealing stability in CoFeB/MgO/CoFeB magnetic tunnel junctions with synthetic pinned layer. *Appl. Phys. Lett.* **89** (2006).
- [228] Bae, J. *et al.* Compositional change of MgO barrier and interface in CoFeB/MgO/CoFeB tunnel junction after annealing. *J. Appl. Phys.* **99** (2006).
- [229] Hayakawa, J., Ikeda, S., Lee, Y. M., Matsukura, F. & Ohno, H. Effect of high annealing temperature on giant tunnel magnetoresistance ratio of CoFeB/MgO/CoFeB magnetic tunnel junctions. *Appl. Phys. Lett.* **89** (2006).
- [230] Schmalhorst, J., Brückl, H., Reiss, G., Gieres, G. & Wecker, J. Thermally induced changes of magnetic coupling in a pinned artificial antiferromagnet used in magnetic tunnel junctions. *J. Appl. Phys.* **94**, 3268–3270 (2003).
- [231] Gan, H. *et al.* Origin of the collapse of tunnel magnetoresistance at high annealing temperature in CoFeB/MgO perpendicular magnetic tunnel junctions. *Appl. Phys. Lett.* **99** (2011).
- [232] Kim, Y. K. *et al.* Interface and microstructure evolutions in synthetic ferrimagnet-based spin valves upon exposure to postdeposition annealing. *J. Appl. Phys.* **93**, 7924–7926 (2003).
- [233] Anderson, G. W., Huai, Y. & Pakala, M. Spin-valve thermal stability: The effect of different antiferromagnets. *J. Appl. Phys.* **87**, 5726–5728 (2000).

- [234] Rickart, M. *et al.* Exchange bias of MnPt/CoFe films prepared by ion beam deposition. *J. Appl. Phys.* **95**, 6317–6321 (2004).
- [235] Soares, M. M. *et al.* Chemically ordered MnPt ultrathin films on Pt(001) substrate: Growth, atomic structure, and magnetic properties. *Phys. Rev. B* **85**, 205417 (2012).
- [236] Saito, M., Hasegawa, N., Koike, F., Seki, H. & Kuriyama, T. Ptmn single and dual spin valves with synthetic ferrimagnet pinned layers. *J. Appl. Phys.* **85**, 4928–4930 (1999).
- [237] Fukami, S., Anekawa, T., Zhang, C. & Ohno, H. A spin–orbit torque switching scheme with collinear magnetic easy axis and current configuration. *Nat. Nanotechnol.* **11**, 621–625 (2016).
- [238] Hayakawa, K. *et al.* Nanosecond random telegraph noise in in-plane magnetic tunnel junctions. *Phys. Rev. Lett.* **126**, 117202 (2021).
- [239] Kim, T. *et al.* Demonstration of in-plane magnetized stochastic magnetic tunnel junction for binary stochastic neuron. *AIP Adv.* **12**, 075104 (2022).
- [240] Lv, H. *et al.* Barrier breakdown mechanism in nano-scale perpendicular magnetic tunnel junctions with ultrathin MgO barrier. *AIP Adv.* **8** (2018).
- [241] Jenkins, S. *et al.* Magnetic stray fields in nanoscale magnetic tunnel junctions. *J. Phys. D. Appl. Phys.* **53**, 044001 (2019).
- [242] Bapna, M. & Majetich, S. A. Current control of time-averaged magnetization in superparamagnetic tunnel junctions. *Appl. Phys. Lett.* **111**, 243107 (2017).
- [243] Wang, Y.-H. *et al.* Interfacial and annealing effects on magnetic properties of CoFeB thin films. *J. Appl. Phys.* **99** (2006).
- [244] Zhu, W. *et al.* Annealing influence on magnetic anisotropies of CoFe thin films. *AIP Adv.* **9** (2019).
- [245] Hindmarch, A., Rushforth, A., Champion, R., Marrows, C. & Gallagher, B. Origin of in-plane uniaxial magnetic anisotropy in CoFeB amorphous ferromagnetic thin films. *Phys. Rev. B* **83**, 212404 (2011).
- [246] Bapna, M. *et al.* Magnetostatic effects on switching in small magnetic tunnel junctions. *Appl. Phys. Lett.* **108**, 022406 (2016).

- [247] Parks, B. *et al.* Superparamagnetic perpendicular magnetic tunnel junctions for true random number generators. *AIP Adv.* **8**, 055903 (2018).
- [248] Reiss, G., Ludwig, J. & Rott, K. Superparamagnetic dwell times and tuning of switching rates in perpendicular CoFeB/MgO/CoFeB tunnel junctions. *arXiv:1908.02139* (2019).
- [249] Kobayashi, K. *et al.* Sigmoidal curves of stochastic magnetic tunnel junctions with perpendicular easy axis. *Appl. Phys. Lett.* **119**, 132406 (2021).
- [250] Zink, B. R., Lv, Y. & Wang, J.-P. Telegraphic switching signals by magnet tunnel junctions for neural spiking signals with high information capacity. *J. Appl. Phys.* **124**, 152121 (2018).
- [251] Parks, B., Abdelgawad, A., Wong, T., Evans, R. F. & Majetich, S. A. Magnetoresistance dynamics in superparamagnetic Co-Fe-B nanodots. *Phys. Rev. Appl.* **13**, 014063 (2020).
- [252] Vodenicarevic, D. *et al.* Low-energy truly random number generation with superparamagnetic tunnel junctions for unconventional computing. *Phys. Rev. Appl.* **8**, 054045 (2017).
- [253] Lee, D. & Lim, S. H. Increase of temperature due to Joule heating during current-induced magnetization switching of an MgO-based magnetic tunnel junction. *Appl. Phys. Lett.* **92**, 233502 (2008).
- [254] Asti, G. *et al.* Nucleation of weak stripe domains: Determination of exchange and anisotropy thermal variation. *Phys. Rev. B* **76**, 094414 (2007).
- [255] Lee, K.-M., Choi, J. W., Sok, J. & Min, B.-C. Temperature dependence of the interfacial magnetic anisotropy in W/CoFeB/MgO. *AIP Adv.* **7**, 065107 (2017).
- [256] Schnitzspan, L., Kläui, M. & Jakob, G. Electrical coupling of superparamagnetic tunnel junctions mediated by spin-transfer-torques. *Appl. Phys. Lett.* **123**, 232403 (2023).
- [257] Tsai, D.-M. & Lin, C.-T. Fast normalized cross correlation for defect detection. *Pattern Recognit. Lett.* **24**, 2625–2631 (2003).
- [258] Camsari, K. Y., Sutton, B. M. & Datta, S. P-bits for probabilistic spin logic. *Appl. Phys. Rev.* **6**, 011305 (2019).
- [259] Camsari, K. Y., Faria, R., Sutton, B. M. & Datta, S. Stochastic p-bits for invertible logic. *Phys. Rev. X* **7**, 031014 (2017).

- [260] Kaiser, J. & Datta, S. Probabilistic computing with p-bits. *Appl. Phys. Lett.* **119** (2021).
- [261] Torrejon, J. *et al.* Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017).
- [262] Zahedinejad, M. *et al.* Two-dimensional mutually synchronized spin hall nano-oscillator arrays for neuromorphic computing. *Nat. Nanotechnol.* **15**, 47–52 (2020).
- [263] Sengupta, A., Parsa, M., Han, B. & Roy, K. Probabilistic deep spiking neural systems enabled by magnetic tunnel junction. *IEEE Trans. Electron Devices* **63**, 2963–2970 (2016).
- [264] Farcis, L. *et al.* Spiking dynamics in dual free layer perpendicular magnetic tunnel junctions. *Nano Lett.* **23**, 7869–7875 (2023).
- [265] Edwards, A. J. *et al.* Passive frustrated nanomagnet reservoir computing. *Commun. Phys.* **6**, 215 (2023).
- [266] Stocker, T. C. & Steinke, I. *Statistik: Grundlagen und Methodik* (de Gruyter, Oldenbourg, 2022).
- [267] Shockley, W. The theory of p-n junctions in semiconductors and p-n junction transistors. *Bell Syst. Tech. J.* **28**, 435–489 (1949).
- [268] Burr, G. W. *et al.* Neuromorphic computing using non-volatile memory. *Adv. Phys. X* **2**, 89–124 (2017).
- [269] Hasan, R., Taha, T. M. & Yakopcic, C. On-chip training of memristor crossbar based multi-layer neural networks. *Microelectron. J.* **66**, 31–40 (2017).
- [270] Xia, Q. & Yang, J. J. Memristive crossbar arrays for brain-inspired computing. *Nat. Mater.* **18**, 309–323 (2019).
- [271] Li, C. *et al.* Analogue signal and image processing with large memristor crossbars. *Nat. Electron.* **1**, 52–59 (2018).
- [272] Li, C. *et al.* Efficient and self-adaptive in-situ learning in multilayer memristor neural networks. *Nature Commun.* **9**, 2385 (2018).
- [273] Dorrance, R. *et al.* Diode-MTJ crossbar memory cell using voltage-induced unipolar switching for high-density MRAM. *IEEE Electron Device Lett.* **34**, 753–755 (2013).

- [274] Khan, M. N. I. & Ghosh, S. Multi-bit read and write methodologies for diode-MTJ crossbar array. In *2020 21st International Symposium on Quality Electronic Design (ISQED)*, 93–98 (IEEE, 2020).
- [275] Allresist. Developer: AR-300 (2023). URL www.allresist.com/wp-content/uploads/sites/2/2020/03/AR300-40_english_Allresist_product_information.pdf.
- [276] Kayaku Advanced Materials, I. Resist: LOR (2023). URL <https://kayakuam.com/wp-content/uploads/2023/08/KAM-LOR-PMGI-Data-Sheet.8.21.23-final.pdf>.
- [277] Kayaku Advanced Materials, I. Resist: S1813 (2023). URL <https://kayakuam.com/wp-content/uploads/2019/09/S1800-G2.pdf>.
- [278] Kayaku Advanced Materials, I. Developer: MF319 (2023). URL https://kayakuam.com/wp-content/uploads/2019/10/MF_319_Data_Sheet.pdf.
- [279] Kayaku Advanced Materials, I. Remover: PG (2023). URL <https://kayakuam.com/wp-content/uploads/2020/11/KAM-Remover-PG-TDS.10.29.20-final.pdf>.

Impact of annealing temperature on tunneling magnetoresistance multilayer stack

Leo Schnitzspan¹, Joel Cramer¹, Jan Kubik², Tareq Tarequzzaman², Gerhard Jakob¹, and Mathias Kläui¹

¹Institute of Physics, Johannes Gutenberg-University Mainz, 55099 Mainz, Germany

²Analog Devices, Limerick, V94 RT99, Ireland

The effect of annealing temperatures on the tunnel magnetoresistance of MgO-based magnetic tunnel junctions (MTJs) has been investigated for annealing between 190 °C and 370 °C. The tunnel magnetoresistance (TMR) shows a maximum value of 215 % at an annealing temperature of 330 °C. A strong sensitivity of TMR and the exchange bias of the pinned ferromagnetic layers on the annealing temperature was observed. Depending on sensor application requirements, the MTJ can be optimized either for stability and pinning strength or for a high TMR signal by choosing the appropriate annealing temperature. The switching mechanism of the ferromagnetic layers in the MTJ and the influence of annealing on the layer properties including the MgO barrier are discussed.

Index Terms—Spin Electronics, magnetic tunnel junctions, annealing, exchange bias, tunneling magnetoresistance.

I. INTRODUCTION

MAGNETIC tunnel junctions (MTJs) attract great interest in science due to their extraordinary high tunnel magnetoresistance (TMR) at room temperature [1], [2], high field sensitivity [3], [4], reliability, and low power operation [5], which pave the way in a variety of applications such as magnetic sensors [4], [5] or magnetic non-volatile memories [6], [7]. MgO-based MTJs with a pinned and a free ferromagnetic (FM) layer composed of CoFeB reveal magnetoresistances over 200 % at room temperature [8] for the antiparallel with respect to the parallel configuration of the FM layers. The effect is based on spin-dependent tunneling of electrons through a thin insulating film, first observed by Juliere for Fe/Ge/Co tunnel junctions at 4.2 K [9]. A very high TMR ratio for a sputtered CoFeB/MgO/CoFeB structure was reported as 361 % with $\text{Co}_{40}\text{Fe}_{40}\text{B}_{20}$ electrodes [10] and 472 % for structures with $\text{Co}_{20}\text{Fe}_{60}\text{B}_{20}$ electrodes [11] at room temperature. The combination of large room temperature magnetoresistance, high magnetic field sensitivity [4], thermal stability [12], and fast current-induced switching [13] makes MTJs based on MgO barriers unique and applicable for a variety of applications, such as memory and in particular sensors. It is well known that different parameters, such as barrier thickness [14], ferromagnetic layer thickness and composition, exchange bias, crystal structure [15], and annealing temperature, have great impact on specific properties of the MTJ, namely on the tunnel magnetoresistance (TMR), resistance area product (RA), which characterizes primarily the resistivity of the tunnel barrier and on the pinning of the fixed/pinned layer by the (synthetic) antiferromagnet. Therefore, it is of importance to examine the relation between a parameter change and the change of TMR, RA and pinning strength of the MTJ pinned layer as for different applications, one needs to optimize one or the other property.

Here, we report the impact of the annealing temperature (T_a) on a synthetic antiferromagnet (SAF) CoFeB/MgO/CoFeB stack with in-plane magnetization and its influence on TMR and RA values as well as the effect on the pinning strength of FM layers in our MTJ stack. The relation of T_a with the exchange bias and switching of the synthetic antiferromagnet (SAF) is shown. These results are prerequisites to develop improved MTJ stacks with TMR values of over 200 % at room temperature combined with strong pinning of the pinned layer and provide a better understanding of different effects originating from the annealing process. Annealing is crucial for the robustness and thermal stability of a TMR stack. For sensor applications and in MRAM technology thermal stability is still a challenge, since the magnetic anisotropy, responsible for data retention, decreases rapidly with temperature but is required for devices with a wide range of operation temperatures [16]. For this reason the relation of annealing temperature and stack stability is studied in this work. With sensor stability as a critical MR sensor parameter, declared in the IEEE Magnetoresistive Sensor Development Roadmap [17], we aim with this work to contribute to the development of promising MR sensor applications in the field of TMR sensor stability.

II. EXPERIMENT

The MTJ stack, investigated in this work, consists primarily of a synthetic antiferromagnet (SAF), a MgO tunnel barrier and a free-layer ferromagnet. The MTJ films were deposited on a Si/SiO₂ substrate using rf- and dc-magnetron sputtering (Singulus Rotaris). CoFeB was sputtered at room temperature and with an argon pressure of $5.8 \cdot 10^{-5}$ Pa, whereas MgO, PtMn, CoFe, Ta, and Ru are sputtered at $3.5 \cdot 10^{-5}$ Pa. All target materials have purities of at least 99.9 % and each layer was sputtered using a single target at a time. The used stack and its composition of different films with their respective film thicknesses in nanometer (in brackets) is: Ru(3)/Ta(5)/Co₄₀Fe₄₀B₂₀(2)/MgO(1.8)/Co₄₀Fe₄₀B₂₀(2.6)/Ru(0.85)/Co₇₀Fe₃₀(2)/Pt₃₈Mn₆₂(17)/Ru(5)/Ta(5)/Ru(5)

/Ta(5)/SiO₂/Si. The stack is illustrated in Fig. 1 (a). The first CoFeB layer, called "free-layer", determines the MTJ configuration state (parallel or antiparallel), depending on its magnetization orientation. Separated by an insulating MgO barrier, the next CoFeB layer, which is labeled "reference layer", is antiferromagnetically coupled with the CoFe ferromagnet (FM) layer underneath, called "pinned layer", and builds the SAF of the stack. The in-plane magnetization orientation of the CoFe "pinned layer" is fixed through interlayer exchange coupling and by the PtMn antiferromagnet (AFM) layer through interfacial exchange coupling, leading to an uniaxial anisotropy of the "pinned layer".

During the sputtering process an Argon flow of 55–100 sccm was used. Each MTJ annealing process was carried out in the following way: An in-plane homogenous magnetic field of approximately 350 mT was applied during the process to define the uniaxial magnetic anisotropy of the magnetic pinned layer. The temperature was increased continuously at a rate of 5 K/min until the target temperature was reached. By means of a proportional–integral–derivative (PID) controller, the target temperature was held for 3 h and thereafter the stack was cooled down at the same rate of 5 K/min. Annealing temperatures between 190 °C and 370 °C were investigated in order to observe maximum sensor performance as well as strong performance deterioration that might occur at low and high annealing temperatures. The stack was mounted into a copper sample holder enclosing the sample in order to provide efficient heat transfer and was placed in a vacuum tube at approximately 10⁻⁵ Pa. TMR and *RA* values were determined through current in-plane tunneling (CIPT) measurements [18], carried out with a CIPTEch-M200 device from Capres, which is capable to determine these values for unpatterned MTJ film stacks. By doing a series of four point probe resistance measurements at various probe spacings on the surface of an unpatterned stack for a positive and a negative applied in-plane magnetic field, the magnetoresistance and resistance area product are obtained. For potential applications, not only TMR and *RA* values are important, but equally the stability of the MTJ state (parallel or antiparallel configuration). This is limited by the stability of the pinned layer, as for very large fields all layers will eventually point in the same direction and thus the sensor will fail. For this reason the stability of the SAF FM layer (reference layer) is analyzed by SQUID magnetization measurements, where the stability of the SAF is characterized by the bias field (H_{saf}) which is necessary to obtain the magnetization of 1/3 with respect to the saturation magnetization M_s . In this case, the reference layer is in process of switching its orientation, hence this is an indicator of a failure of the sensor device operation. The fraction of 1/3 has been chosen because the form change of the hysteresis curve for different annealing temperatures (see. Fig. 1) can be sensed at 1/3 of M_s better than for other fractions of magnetization. So in order to gauge the performance of the sensor, one needs to ascertain the switching behavior of the individual FM layers in the MTJ stack.

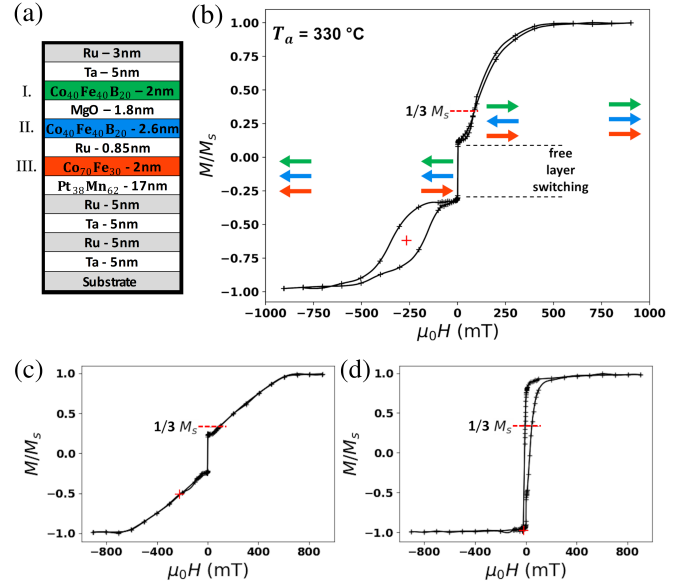


Fig. 1: (a) MTJ stack with FM layers depicted as free (green or I.), reference (blue or II.) and pinned (red or III.) layers. (b) Hysteresis loop of the relative magnetization of the TMR stack at room temperature after annealing at 330 °C measured by means of a SQUID magnetometer (Quantum Design MPMS II). At around ± 750 mT the magnetization is saturated. The red cross represents the exchange field of this sample, determined by the mean of two Langevin fits, respectively of the magnetization curve measured for increasing or decreasing applied field. One third of the saturation magnetization for positive fields is depicted as "1/3 M_s ", at which H_{saf} , a value for stability of the reference layer, is obtained. The colored arrows denote the magnetization direction of the free layer (green), reference layer (blue) and pinned layer (red). The free layer switches abruptly at low fields, whereas the the reference and pinned layers switch gradually at higher fields. In (c) and (d) the hysteresis curves for TMR stacks with the lowest (190 °C) and highest (370 °C) annealing temperatures are plotted.

III. RESULTS AND DISCUSSION

To analyze the switching mechanism of the FM layers in the MTJ stack, magnetization hysteresis loops were measured for annealed samples between 190 °C and 370 °C. A measurement for the 330 °C sample is shown in Fig. 1(b), where the switching of the FM layers is depicted as arrows of the in-plane magnetization direction in green for the free layer, in blue for the reference layer and in red for the antiferromagnetic pinned layer. For small applied fields, below 3 mT in magnitude, the free layer switches immediately, whereas for positive fields greater +20 mT the reference layer is deflected from its zero field alignment. For applied fields below -100 mT, the pinned layer starts to switch the magnetization direction and due to the pinning it results in a small hysteresis loop. The mean of the inflection points of two Langevin fits of this loop is defined as the exchange bias field H_{ex} , which is a measure for the pinning strength of the pinned layer and is indicated

as a red cross. In Fig. 1 (c) and (d) the hysteresis curves for the lowest and highest annealing temperatures are shown. It can be seen that the form of the hysteresis is strongly affected by T_a and that for 370 °C there is no individual layer switching observable, hence all layers switch as one entity. The values obtained for H_{ex} and H_{saf} for the lowest and highest annealing temperature should also be taken with caution, since the definition of these values from the hysteresis curve is difficult and not that clear compared to the 330 °C annealed sample. The most important result here is that the maximum stability for this TMR stack, considering H_{ex} and H_{saf} , is reached for the annealing temperature range of 270 °C to 310 °C.

In order to obtain the dependence of RA and TMR on the annealing temperature, current in-plane tunneling (CIPT) measurements of the MTJ stack were carried out and are shown in Fig. 2. Error bars are plotted for RA and TMR, but for TMR are smaller than the symbols. Error bars were determined as standard deviations by multiple CIPT measurements at different positions in a 1 mm² area of the 330 °C annealed TMR stack. In Fig. 2 (a), we show that the RA has a minimum at 270 °C and increases for lower and higher temperatures. A possible formation of boron oxide (diffusion of B into MgO) at the MgO interface [19] can lead to an increased tunneling resistance or RA value through a disruption of the coherent tunneling process for high annealing temperatures [20]. In addition, the effective barrier thickness could increase above 330 °C, due to migration of Mn atoms towards the barrier [21], contributing to the increase in RA for higher temperatures. For low annealing temperatures, below 270 °C, the decrease of RA with T_a is governed likely by crystallization of the amorphous CoFeB [10], resulting in a better defined MgO/CoFeB interface and higher TMR ratios [19]. The low RA value at around 300 °C coincides with a high TMR ratio. For higher T_a , an accumulation of Mn atoms in the MgO barrier [11] induces a rapid degradation of TMR, see Fig. 2 (b). Our explanation is also supported by TMR- T_a -dependences reported by other groups [1], [10], [11], [22], [23], which underpins the importance of annealing temperature onto the MTJ quality and thus the TMR ratio. A very similar trend was also found for a TMR structure, based on an Al₂O₃ barrier, where the TMR ratio gradually increases with annealing temperature up to 300 °C and rapidly deteriorates for annealing temperatures above [21], showing that this effect does not depend on the barrier material type. Contrary to this Al₂O₃ based TMR stack, where the RA product is almost constant for annealing temperatures up to 400 °C [21], our MgO based TMR stack shows an approximately inverse relation of TMR ratio and RA product. The exchange bias fields H_{ex} (for the pinned layer) for all annealing temperatures are plotted in Fig. 2 (c). It shows a maximum of the exchange bias at 270 °C and decreasing H_{ex} for temperatures above this. The exchange bias is caused by the pinning of the antiferromagnet, which is weakened during the annealing process at high temperatures, where thermally activated Ru atoms diffuse [10] into the AFM and FM pinned layers, deteriorate the lattice structure and lead to pinholes in the Ru spacer [22]. This results in a ferromagnetic coupling of FM reference and pinned layer and a reduction of the AFM-

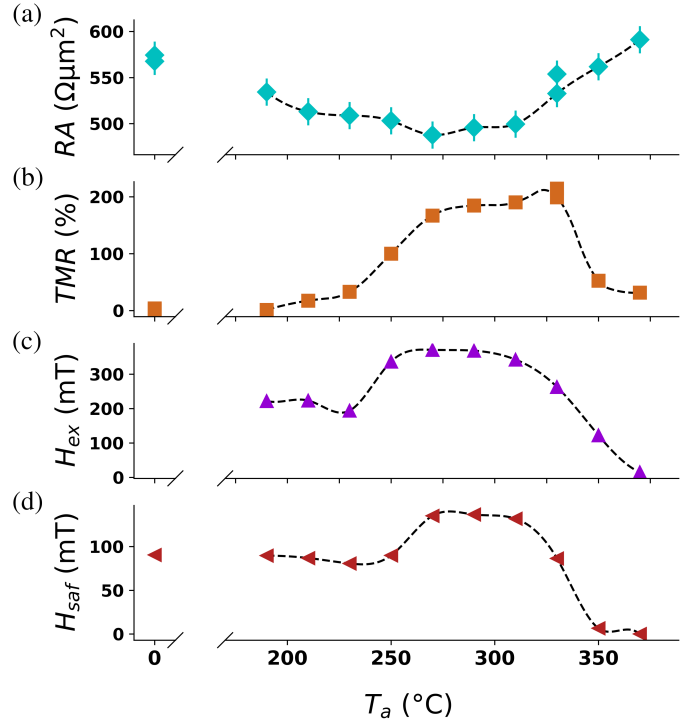


Fig. 2: Depending on the annealing temperature T_a , the (a) resistance area product (RA), (b) tunnel magnetoresistance (TMR), (c) exchange bias field (H_{ex}) and (d) SAF bias field H_{saf} is plotted. The dashed line is a cubic spline interpolation as a guide for the eye, with the assumption of a continuous change of the parameters with temperature. Error bars for RA are depicted but not for TMR, since they are smaller than the data points. The maximum TMR value is 215 % and the two data points (for RA and TMR) at 330 °C indicate two samples of the same sputtered stack, one measured at the edge and one at the center showing good agreement. Data points at 0 °C are from unannealed MTJ samples.

exchange coupling [22]. The diffusion of Mn atoms of the AFM towards the MgO barrier [11] leads to a weakening of the AFM pinning and a reduced net magnetic thickness of CoFe layer [24]. Furthermore, it has been shown that the PtMn AFM induced exchange bias decreases quickly for high temperatures, especially above 300 °C, where the blocking temperature distribution for PtMn has a sharp peak [25] and where we see a rapid decline of H_{ex} . The SAF bias field H_{saf} , plotted in Fig. 2 (d), is a measure for the pinning strength of the reference layer and indicates when the reference layer switches its orientation. It is defined as the positive magnetic field value where the magnetization of the whole MTJ stack reaches 1/3 of the total magnetization, as indicated in Fig. 1 (b). Consistent with H_{ex} , H_{saf} decreases rapidly for higher annealing temperatures and shows its maximum at around 290 °C. Both bias fields show a similar "plateau" in the range of 270–310 °C where the TMR stack has the most robust and most stable behavior and which should be favored for sensor applications.

The decrease of H_{saf} for high T_a , beginning at 310 °C,

originates mostly from the diffusion of Mn, Ru and B atoms into different layers together with the formation of pinholes in the Ru spacer [22], leading to a weakening of the inter-layer coupling between reference and pinned layer. Due to high thermal energy at high temperatures, Mn atoms diffuse from the PtMn AFM layer, deteriorate the AFM structure, leading to a significant decrease of exchange bias or pinning strength of the PtMn layer [26]. This deterioration results in H_{ex} and H_{saf} values close to 0 mT at 370 °C. For the low temperature range < 250 °C, exchange bias and SAF bias fields are approximately constant, especially for the SAF bias field where the not-annealed stack has the same H_{saf} value as for the 200 °C annealed stack. Also SQUID hysteresis loops in this range are almost symmetrical, unlike for temperatures above 250 °C which are highly unsymmetrical. This can be explained by the PtMn AFM which appears to be paramagnetic and disordered and shows no exchange bias effect after room temperature deposition [26], [27], but will emerge after a structural transition of PtMn from unordered face-centered-cubic (fcc) to face-centered-tetragonal (fct) ordered $L1_0$ phase at higher temperatures [28]. In order to obtain a strong exchange bias, thermal annealing with an applied magnetic field is necessary to promote this long-range chemical ordering in the PtMn AFM layer [26], [29]. The threshold therefore lies around at 250 °C, where both H_{ex} and H_{saf} strongly increase and form a small "plateau", see Fig. 2(c) and (d). For low annealing temperatures < 250 °C or without annealing, the AFM pinning effect is absent and pinned and reference layer favor to align antiparallel. By applying a magnetic in-plane field (positive or negative) both layers tend to align parallel while one of them switches, which yields to a constant field value for H_{saf} (where 1/3 of the magnetization is reached). For high temperatures, however, the Ru spacer deteriorates, leading to a parallel alignment of all FM layers for very low fields close to 10 mT. The optimized annealing temperature for high pinning strengths (290 °C) differs from the temperature for the maximum TMR (330 °C) by 40 °C for our MTJ stack. Depending on the choice of annealing temperature, this MTJ stack can thus either be optimized for sensor stability or for high TMR ratio or for a compromise of both properties. Since the observed annealing temperature effects are mainly caused by interdiffusion and crystallization, similar effects and results are expected for perpendicularly magnetized TMR structures.

IV. CONCLUSION

In this work we have shown that, depending on the requirements, for instance, of a sensor application, an MTJ can be optimized for stability and pinning strength or for a high TMR signal by choosing an appropriate annealing temperature. We demonstrate that the annealing temperature is critical and affects the MTJ properties for potential application. For low annealing temperatures up to a threshold of 250 °C the PtMn AFM layer shows no pinning effect, resulting in unstable MTJ stack, in which the FM layers can switch more easily for low magnetic fields. However, for high annealing temperatures at around 350 °C, interdiffusion deteriorates the stability of the MTJ stack. As a consequence, too high temperatures result

in a degradation of the tunnel barrier, exchange bias and a reduction of TMR and RA values. Thermally activated diffusion of Mn, Ru and B and the formation of pinholes in the Ru SAF spacer during the annealing process lead to a lower pinning strength of the pinned and reference FM layer. Optimized MTJ stability with strong AFM pinning effect is found for annealing temperatures between 270 °C and 310 °C. This shows the impact of annealing temperature on the MTJ performance and as a consequence an appropriate annealing temperature should be chosen. For our MgO-based SAF stack, we find that we can either optimize it for pinning strength and thus stability or for a maximum TMR ratio.

ACKNOWLEDGMENT

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - TRR 173 - 268565370, the Graduate School of Excellence Materials Science in Mainz (DFG/GSC 266). R.L. acknowledges the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement FAST number 752195 and the Horizon 2020 FET OPEN s-Nebula project 863155. This work is also supported in part by BMBF within the funding (16ES0938) for the programme Forschungslabore Mikroelektronik Deutschland (ForLab). We acknowledge the Joint Research Project with Analog Devices for funding and the ForLab MagSens Center of Competence for research and development on high performance magnetic sensor systems for providing the infrastructure.

REFERENCES

- [1] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," *IEEE Electron Device Lett.*, vol. 20, pp. 569–571, Nov. 1999.
- [2] S. Ikeda, K. Miura, H. Yamamoto, K. Mizunuma, H. Gan, M. Endo, S. Kanai, J. Hayakawa, F. Matsukura, and H. Ohno, "A perpendicular-anisotropy CoFeB–MgO magnetic tunnel junction," *Nat. Mater.*, vol. 9, no. 9, p. 721, 2010.
- [3] M. Tondra, J. M. Daughton, D. Wang, R. S. Beech, A. Fink, and J. A. Taylor, "Picotesla field sensor design using spin-dependent tunneling devices," *J. Appl. Phys.*, vol. 83, no. 11, pp. 6688–6690, 1998.
- [4] R. Chaves, P. Freitas, B. Ocker, and W. Maass, "MgO based picotesla field sensors," *J. Appl. Phys.*, vol. 103, no. 7, p. 07E931, 2008.
- [5] W. Zhao, C. Chappert, V. Javerliac, and J.-P. Noziere, "High speed, high stability and low power sensing amplifier for MTJ/CMOS hybrid logic circuits," *IEEE Trans. Magn.*, vol. 45, no. 10, pp. 3784–3787, 2009.
- [6] H. Yoda, T. Kishi, T. Nagase, M. Yoshikawa, K. Nishiyama, E. Kitagawa, T. Daibou, M. Amano, N. Shimomura, S. Takahashi *et al.*, "High efficient spin transfer torque writing on perpendicular magnetic tunnel junctions for high density MRAMs," *Curr. Appl. Phys.*, vol. 10, no. 1, pp. e87–e89, 2010.
- [7] J. Slaughter, N. Rizzo, J. Janesky, R. Whig, F. Mancoff, D. Housameddine, J. Sun, S. Aggarwal, K. Nagel, S. Deshpande *et al.*, "High density ST-MRAM technology," in *2012 International Electron Devices Meeting*. IEEE, 2012, pp. 29–3.
- [8] X. Liu, D. Mazumdar, W. Shen, B. Schrag, and G. Xiao, "Thermal stability of magnetic tunneling junctions with MgO barriers for high temperature spintronics," *Appl. Phys. Lett.*, vol. 89, no. 2, p. 023504, 2006.
- [9] M. Juliere, "Tunneling between ferromagnetic films," *Phys. Lett. A*, vol. 54, no. 3, pp. 225–226, 1975.
- [10] Y. M. Lee, J. Hayakawa, S. Ikeda, F. Matsukura, and H. Ohno, "Giant tunnel magnetoresistance and high annealing stability in CoFeB/MgO/CoFeB magnetic tunnel junctions with synthetic pinned layer," *Appl. Phys. Lett.*, vol. 89, no. 4, p. 042506, 2006.


- [11] J. Hayakawa, S. Ikeda, Y. M. Lee, F. Matsukura, and H. Ohno, "Effect of high annealing temperature on giant tunnel magnetoresistance ratio of CoFeB/MgO/CoFeB magnetic tunnel junctions," *Appl. Phys. Lett.*, vol. 89, no. 23, p. 232510, 2006.
- [12] C. Park, J. Kan, C. Ching, J. Ahn, L. Xue, R. Wang, A. Kontos, S. Liang, M. Bangar, H. Chen *et al.*, "Systematic optimization of 1 Gbit perpendicular magnetic tunnel junction arrays for 28 nm embedded STT-MRAM and beyond," in *2015 IEDM*. IEEE, 2015, pp. 26–2.
- [13] T. Kishi, H. Yoda, T. Kai, T. Nagase, E. Kitagawa, M. Yoshikawa, K. Nishiyama, T. Daibou, M. Nagamine, M. Amano *et al.*, "Lower-current and fast switching of a perpendicular TMR for high speed and high density spin-transfer-torque MRAM," in *2008 IEDM*. IEEE, 2008, pp. 1–4.
- [14] J. Hayakawa, S. Ikeda, F. Matsukura, H. Takahashi, and H. Ohno, "Dependence of giant tunnel magnetoresistance of sputtered CoFeB/MgO/CoFeB magnetic tunnel junctions on MgO barrier thickness and annealing temperature," *Jpn. J. Appl. Phys.*, vol. 44, no. 4L, p. L587, 2005.
- [15] S. Yuasa, Y. Suzuki, T. Katayama, and K. Ando, "Characterization of growth and crystallization processes in CoFeB/MgO/CoFeB magnetic tunnel junction structure by reflective high-energy electron diffraction," *Appl. Phys. Lett.*, vol. 87, no. 24, p. 242503, 2005.
- [16] J. M. Iwata-Harms, G. Jan, H. Liu, S. Serrano-Guisan, J. Zhu, L. Thomas, R.-Y. Tong, V. Sundar, and P.-K. Wang, "High-temperature thermal stability driven by magnetization dilution in CoFeB free layers for spin-transfer-torque magnetic random access memory," *Sci. Rep.*, vol. 8, no. 1, pp. 1–7, 2018.
- [17] C. Zheng, K. Zhu, S. C. De Freitas, J.-Y. Chang, J. E. Davies, P. Eames, P. P. Freitas, O. Kazakova, C. Kim, C.-W. Leung *et al.*, "Magnetoresistive sensor development roadmap (non-recording applications)," *IEEE Trans. Magn.*, vol. 55, no. 4, pp. 1–30, 2019.
- [18] D. Worledge and P. Trouilloud, "Magnetoresistance measurement of unpatterned magnetic tunnel junction wafers by current-in-plane tunneling," *Appl. Phys. Lett.*, vol. 83, no. 1, pp. 84–86, 2003.
- [19] J. Bae, W. Lim, H. Kim, T. D. Lee, K. Kim, and T. Kim, "Compositional change of MgO barrier and interface in CoFeB/MgO/CoFeB tunnel junction after annealing," *J. Appl. Phys.*, vol. 99, no. 8, p. 08T316, 2006.
- [20] X. Liu, D. Mazumdar, W. Shen, B. Schrag, and G. Xiao, "Thermal stability of magnetic tunneling junctions with MgO barriers for high temperature spintronics," *Appl. Phys. Lett.*, vol. 89, no. 2, p. 023504, 2006.
- [21] S. Cardoso, P. Freitas, C. De Jesus, P. Wei, and J. Soares, "Spin-tunnel-junction thermal stability and interface interdiffusion above 300°C," *Appl. Phys. Lett.*
- [22] J. Schmalhorst, H. Brückl, G. Reiss, G. Gieres, and J. Wecker, "Thermally induced changes of magnetic coupling in a pinned artificial antiferromagnet used in magnetic tunnel junctions," *J. Appl. Phys.*, vol. 94, no. 5, pp. 3268–3270, 2003.
- [23] H. Gan, H. Sato, M. Yamanouchi, S. Ikeda, K. Miura, R. Koizumi, F. Matsukura, and H. Ohno, "Origin of the collapse of tunnel magnetoresistance at high annealing temperature in CoFeB/MgO perpendicular magnetic tunnel junctions," *Appl. Phys. Lett.*, vol. 99, no. 25, p. 252507, 2011.
- [24] Y. K. Kim, G.-H. Park, S.-R. Lee, S.-H. Min, J. Y. Won, and S. A. Song, "Interface and microstructure evolutions in synthetic ferrimagnet-based spin valves upon exposure to postdeposition annealing," *J. Appl. Phys.*, vol. 93, no. 10, pp. 7924–7926, 2003.
- [25] G. W. Anderson, Y. Huai, and M. Pakala, "Spin-valve thermal stability: The effect of different antiferromagnets," *J. Appl. Phys.*, vol. 87, no. 9, pp. 5726–5728, 2000.
- [26] H. Chang, C. Shen, F. Yuan, P. Pan, Y. Chien, C. Wang, L. Horng, and S. Jian, "Exchange bias in Co/MnPt polycrystalline films on Si(100)/SiO₂ substrates with Ta underlayer," *Thin Solid Films*, vol. 660, pp. 834–839, 2018.
- [27] M. Rickart, P. Freitas, I. Trindade, N. Barradas, E. Alves, M. Salgueiro, N. Muga, J. Ventura, J. Sousa, G. Proudfoot *et al.*, "Exchange bias of MnPt/CoFe films prepared by ion beam deposition," *J. Appl. Phys.*, vol. 95, no. 11, pp. 6317–6321, 2004.
- [28] M. M. Soares, M. De Santis, H. C. Tolentino, A. Y. Ramos, M. El Jawad, Y. Gauthier, F. Yildiz, and M. Przybylski, "Chemically ordered MnPt ultrathin films on Pt(001) substrate: Growth, atomic structure, and magnetic properties," *Phys. Rev. B*, vol. 85, no. 20, p. 205417, 2012.
- [29] R. Farrow, R. Marks, S. Gider, A. Marley, S. Parkin, and D. Mauri, "Mn_xPt_{1-x}: A new exchange bias material for permalloy," *J. Appl. Phys.*, vol. 81, no. 8, pp. 4986–4988, 1997.

Nanosecond True-Random-Number Generation with Superparamagnetic Tunnel Junctions: Identification of Joule Heating and Spin-Transfer-Torque Effects

Leo Schnitzspan^{1,2}, Mathias Kläui^{1,2,*} and Gerhard Jakob^{1,2}

¹*Institute of Physics, Johannes Gutenberg-Universität Mainz, Mainz 55122, Germany*

²*Max Planck Graduate Center Mainz, Mainz 55122, Germany*

 (Received 13 January 2023; revised 24 May 2023; accepted 28 June 2023; published 1 August 2023)

This work investigates nanosecond superparamagnetic switching in 50-nm-diameter in-plane magnetized magnetic tunnel junctions (MTJs). Due to the small in-plane uniaxial anisotropy, dwell times below 10 ns and autocorrelation times down to 5 ns are measured for circular superparamagnetic tunnel junctions (SMTJs). SMTJs exhibit probabilistic switching of the magnetic free layer, which can be used for the generation of true random numbers. The quality of the random bit streams generated by our SMTJ is evaluated with a statistical test suite (NIST STS, sp 800-22) and shows true randomness after three exclusive OR (XOR) operations of four random SMTJ bit streams. A low-footprint CMOS circuit is proposed for fast and energy-efficient random-number generation. We demonstrate that the probability of a 1 or 0 can be tuned by spin-transfer torque (STT), while the average bit-generation rate is mainly affected by the current density via Joule heating. Although both effects are always present in MTJs, most often Joule heating is neglected. However, with a resistance-area (RA) product of $15\Omega\ \mu\text{m}^2$ and current densities of the order of $1\ \text{MA}/\text{cm}^2$, an increasing temperature at the tunneling site results in a significant increase in the switching rate. As Joule heating and STT scale differently with current density, the device design can be optimized based on our findings.

DOI: [10.1103/PhysRevApplied.20.024002](https://doi.org/10.1103/PhysRevApplied.20.024002)

I. INTRODUCTION

A superparamagnetic tunnel junction (SMTJ), acting as a random stochastic noise source, is the counterpart to nonvolatile random access memory (MRAM) [1], where high temporal stability (switching energy barrier $E_b > 40 k_B T$) of the state is required to store the information of a bit. The volatile behavior of the superparamagnetic junction originates from the ambient thermal energy acting on the magnetization of the free layer. This thermal energy is high enough to overcome the low energy barrier of a few $k_B T$, resulting in a superparamagnetic state with fluctuation times of milliseconds down to nanoseconds at room temperature. The time spent in each state, called the dwell time, can be controlled by an applied external magnetic field, by spin-transfer-torque, or by spin-orbit torque [2]. Due to the inherent probabilistic nature of SMTJs and their state controllability and energy efficiency, SMTJs have been proposed for various computational concepts [3], such as invertible logic [4], Boltzmann machines [5], reservoir computing [6], spiking neural networks [7,8], or stochastic computing [9]. These concepts can provide advantages over pure digital complementary metal-oxide-semiconductor (CMOS) based computational

logic, since the fundamental building block is a stochastic bit. Despite the tremendous development of conventional CMOS-based deterministic computers over recent decades, there are still classes of problems that cannot be addressed efficiently, due to the deterministic nature of von Neumann computers. Many numerical-computation techniques are based on the Markov chain Monte Carlo (MCMC) methods [10] and require many random numbers. Since SMTJs are inherently probabilistic, as opposed to deterministic CMOS circuits, they can provide random signals with low power and a low areal footprint, which can be transformed to random bit streams using a few CMOS transistors. MTJs are already CMOS compatible on scale for different applications such as MRAM [1]. Through the combination of CMOS logic and a set of probabilistic SMTJs (called p-bits [4]), it has been shown that computationally hard problems such as the “traveling-salesman problem” can be solved efficiently [11]. The performance of such a “p-computer” will then be a consequence of the p-bit hardware density and the average fluctuation rate, described by the Néel-Arrhenius law. Fast fluctuations and dwell times in the nanosecond range are desired. For in-plane easy-axis MTJs (ip MTJs), dwell times of the order of milliseconds [12], microseconds [13–15], and down to nanoseconds [16,17] have been reported. Compared to out-of-plane MTJs (oop MTJs), ip

*mathias.klaui@klaui-lab.de

MTJs often exhibit shorter dwell times, due to a different contribution of in-plane and out-of-plane anisotropy energies [18]. Out-of-plane MTJs often comprise larger time scales of milliseconds [19,20] to microseconds [21–23] and for this reason, we have decided to design our MTJ stack with in-plane ferromagnetic layers. We demonstrate that two phenomena occur when a current is applied to an MTJ nanopillar: spin-transfer torque (STT) and Joule heating. STT refers to the transfer of angular momentum from the flowing electrons to the magnetic moments in the ferromagnetic layer. Joule heating, on the other hand, refers to the generation of heat induced by the flow of charge current through the device. Both effects are always present in an MTJ and have to be considered to understand the dependence of the fluctuation times and the current density. So far, Joule heating and STT have been studied together on superparamagnetic nanoislands with spin-polarized scanning tunneling microscopy [24] but Joule heating has mostly been ignored or neglected for SMTJs, where only STT effects have been considered.

In this work, we separate and extract the contributions of STT and Joule heating for different current densities to understand the mechanism of the superparamagnetic switching dependence. We show that STT affects the state energy linearly with applied current, while Joule heating has a quadratic dependence on the current. Due to the low energy barrier between the two states, Joule heating results in an overall faster switching, while STT tunes the state probability differently. In the case of a random-number generator, it is important to control the state probability to a desired value. An implementation idea based on logic exclusive OR (XOR) gates is proposed in order to generate a stream of true random bits. The quality of the randomness is quantified by the Statistical Test Suite from NIST [25]. We show that true randomness, with a bit-generation rate of 200 Mbit/s, is achieved by using four SMTJs and a combination of three XOR gates. It is also demonstrated that for the generation of multiple independent true random bit streams, two XOR gates and two SMTJs are sufficient.

This circuit design saves energy and space and therefore has the potential for applications where many random bits are required.

II. METHODS

A. MTJ sample preparation

TMR stacks were deposited at room temperature on oxidized Si substrates using rf- and dc-magnetron sputtering (Singulus Rotaris) with the composition Ta(10 nm)/Ru(10 nm)/Ta(10 nm)/(Pt-Mn)(20 nm)/(Co-Fe)(2.2 nm)/Ru(0.8 nm)/(Co-Fe-B)(2.4 nm)/MgO(1.1 nm)/(Co-Fe-B)(3.0 nm)/Ta(10 nm)/Ru(20 nm) and is based on an optimized stack, developed earlier [26]. The stack is illustrated in Fig. 1(a), where the function of each layer is specified. The tunnel magnetoresistance (TMR) ratio of a continuous film of our stack is found to be approximately 150% and for patterned nanopillars it is in the range of 100–150%.

The state of the MTJ (parallel or antiparallel) is determined by the magnetization orientation of the free layer with respect to the reference layer. The Co-Fe-B reference layer together with a Co-Fe pinned layer forms a synthetic antiferromagnet (SAF), in order to compensate stray fields at the free-layer site. In addition, the pinned Co-Fe layer is exchange biased [27] by a Pt-Mn antiferromagnet (AFM). Even though a SAF is integrated into the stack, the stray field at the free-layer position will never be compensated exactly to zero [28]. For this reason, an in-plane easy-axis field of a few millitesla is typically applied in order to compensate for this “offset” and to set the SMTJ into equal state probability. This offset is approximately 1.1 mT in Fig. 1(b) but can vary from device to device due to variations in the sample fabrication process. In general, in the case of a stochastic MTJ, an in-plane offset field is selected to achieve an equal occupancy of both the parallel (P) and antiparallel (AP) states, aiming for an approximately 50% state probability. As shown in Fig. 1(b), the state probability can be tuned by the applied in-plane easy-axis field. A

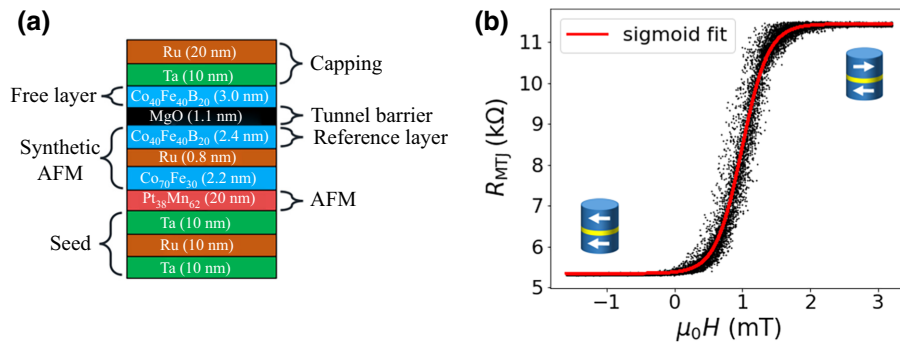


FIG. 1. (a) A TMR stack with an RA product of $15\Omega\ \mu\text{m}^2$ and a TMR ratio of 150%, used for SMTJs. (b) The MTJ resistance versus the external applied easy-axis field of a SMTJ, plotted for 20 repetitions. The red line indicates the sigmoidal relation of the resistance and the magnetic field. Negative fields stabilize the parallel state, while positive fields stabilize the antiparallel state.

change in the field strength results in a shift of the energy levels, thus stabilizing either the P or the AP state. The relative increase in resistance from the P to the AP state is defined as the TMR ratio: $TMR = R_{AP}/R_P - 1$. To obtain a high TMR ratio, annealing was carried out in a 300-mT in-plane field at 300 °C for 1 h. The current in-plane tunneling (CIPT) measurements [29] of our unpatterned stack exhibit a resistance-area (RA) product of $15 \pm 2 \text{ } \Omega \text{ } \mu\text{m}^2$ and a TMR ratio of $150 \pm 3\%$. The MTJ nanopillars are patterned using 30-kV electron-beam lithography, to define a circular etch mask. Argon-ion etching under 55° and 20° is used to structure the MTJ pillar. As a passivation layer, SiN_x was chosen for its hardness and good insulating properties. It was sputtered under a 50:50 mixed argon-nitrogen atmosphere at 5 Pa. Above the MTJ nanopillar, a 60-nm-thick gold pad was sputtered as a top-electrode lead in order to be able to bond it to a measurement sample holder. By using the top electrode and the conducting seed layer as a bottom electrode, a voltage was applied across the junction and the switching behavior was studied.

B. Setup and measurement

Electrical measurements were carried out at room temperature with an MTJ in series with a shunt resistor R_s . The amplified signal of the SMTJ fluctuation was measured with a 3.5-GHz-bandwidth digital oscilloscope (Tektronix DPO7543, 40 Gs/s). A 50-Ω low-noise amplifier (ZFL-1000LN+, mini-circuits) with a gain of 20 dB was used to amplify the signal. To ensure that the time-series measurement has a high time resolution and is not limited by the RC time constant of the setup, the parasitic capacitance and resistances of the setup have to be small. The 50-Ω input impedance of the amplifier results in a small signal voltage drop of the order of 1 mV at the amplifier. Since the MTJ resistor (5–10 kΩ) is in parallel with the 50-Ω input impedance of the amplifier, the equivalent resistance is approximately also 50 Ω (neglecting the impedance of C_i for high frequencies). Forming a voltage divider with R_s leads to a small ac voltage under stochastic fluctuation of R_{MTJ} [see Eq. (VI) in Appendix A 1]. In addition, a low

resistance lowers the RC time constant of the setup, thus leading to a high time resolution of approximately 1 ns ($R \approx 50 \text{ } \Omega$ and $C \approx 20 \text{ pF}$). Therefore, the time series of the fluctuating voltage signal are typically recorded with a sampling rate of 1 GS/s for a few milliseconds. The measurement setup is illustrated in Fig. 2(a) and the recorded nanosecond time series is plotted in Fig. 2(b).

III. DWELL TIMES AND AUTOCORRELATION TIMES

A high random bit-generation rate is primarily limited by the average fluctuation rate or dwell time. The dwell time in the macrospin approximation for $E_b/k_B T \geq 1$ follows the Néel-Arrhenius law [30]: $\tau_{P,AP} = \tau_0 \exp(E_b/k_B T)$, where E_b is the energy barrier between the states, T is the temperature, and τ_0 is the attempt time. The assumption of the Néel-Arrhenius model is switching between two distinct energy minima, i.e., two distinct resistance states. For very low energy barriers with respect to the thermal energy, a Fokker-Planck–based uniaxial low barrier solution as formulated by Brown will be needed [31]. As we can be close to that regime, our extracted barriers should be considered as the effective barriers for a Néel-Arrhenius–model description. In order to generate true random bits, the prerequisite is the independence of each generated bit with respect to the previously generated bit. An important metric for this is the autocorrelation time τ_{ac} of the SMTJ signal, which can be interpreted as the time scale for randomness and describes for which time interval the signal is still correlated to its past. As shown theoretically [32], MTJs with in-plane magnetic anisotropy exhibit shorter autocorrelation times than MTJs with out-of-plane anisotropy due to the difference in precessionlike fluctuations, attributed to the large demagnetization field for ip MTJs. In the macrospin approximation with a single normalized magnetization vector $m = M/M_s$ fluctuating randomly on a Bloch sphere, the anisotropy energies will restrict the fluctuation to a specific region in the sphere. The demagnetization energy constrains the fluctuation to the equator of the Bloch sphere and the uniaxial anisotropy

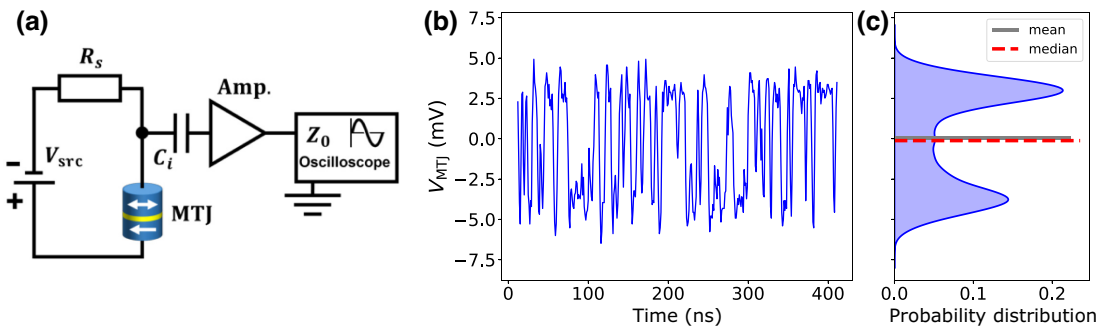


FIG. 2. (a) The circuit diagram of the measurement setup. (b) An example of the nanosecond stochastic switching of the amplified SMTJ signal. (c) The probability distribution (kernel-density estimate) of the SMTJ signal for 10^5 data points.

to regions near the easy-axis directions. In order to measure these superparamagnetic fluctuations, a circular nanopillar MTJ with 50 nm in diameter was patterned by electron-beam lithography and structured with argon-ion etching. Next, a small in-plane offset field of a few millitesla was applied to tune the stochastic MTJ to a 50:50 state probability. With the measurement setup shown in Fig. 2(a), stochastic fluctuations have been acquired, where the dc component of the signal is filtered out by the capacitance in the circuit. The time between two consecutive switching events is defined as the dwell time and is measured for the parallel as well as the antiparallel state. The dwell times can be determined in two ways. First, a time series measurement of 0.5 s (sampling rate of 1 GS/s) is binarized, by using the median as a reference threshold, and then the series is divided into an array of AP (= 1) and P (= 0) dwell-time intervals. The mean of all AP and P time intervals then corresponds to the average dwell time for the AP and P states. The average dwell time τ is calculated by $\sqrt{\tau_P \tau_{AP}}$, as suggested in the literature [17]. For the average dwell time, we find $\tau = 7.3 \pm 0.2$ ns. Second, the switching mechanism can be described by a Poisson process, where the probability for a switching event follows the Poisson distribution, similarly to what happens in a radioactive decay, where the decay probability is described by a Poisson process. By fitting the number of switching events for the AP and P states with $N = N_0 \exp(-t/\tau_{(p,ap)})$, the average dwell times τ_P , τ_{AP} can be extracted in a more robust manner, as shown in Fig. 3. All of the calculated (average) dwell times in this work are determined by the second method, through fitting a Poisson distribution to the dwell times. Figure 3 indicates that the distribution of dwell times is in agreement with the theory, as shown by a linear decrease in the logarithmic plot. The average dwell time is found to be $\tau = 6.7 \pm 0.1$ ns, which is orders of magnitude shorter than most measured dwell times of oop-SMTJs or ip-SMTJs. We further determine the autocorrelation time of the binarized signal, which is the limiting factor for the generation of random numbers. The autocorrelation function (ACF) is calculated by the following equation:

$$\text{ACF}(t) = \frac{\sum_{i=1}^{N-t} [s_i - \bar{s}_i][s_{i+t} - \bar{s}_i]}{\sum_{i=1}^N (s_i - \bar{s}_i)^2}, \quad (1)$$

where s is the binarized signal, \bar{s}_i the mean of the signal, and t is the time lag of the time series of length N . The ACF is plotted in Fig. 4(a) and shows an exponential decrease. After the autocorrelation time τ_{ac} , the signal is uncorrelated with any signal of its past. It is defined as the maximum time by which the integral of the ACF reaches 99%. This results in an autocorrelation time of $\tau_{ac} = 5.1 \pm 0.3$ ns. Next, we consider the rise time of the signal, which was determined by averaging 100 rising edges in a time series (sampling rate 2.5 GHz). Typically, the definition of the

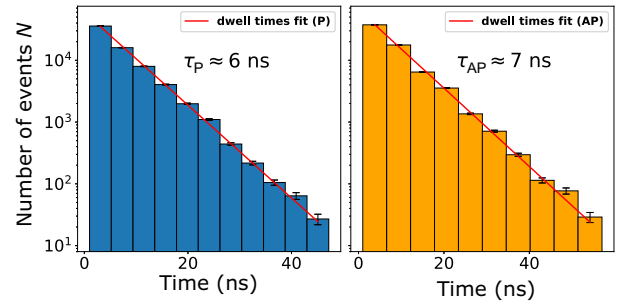


FIG. 3. A histogram of the dwell times for a 1-ms time series for the P and AP states. The bins can be interpreted as a waiting-time interval in a Poisson process, for which exactly one switching event will occur in this interval. From a fit (red line), the average dwell times τ_P and τ_{AP} are determined, according to a Poisson process.

rise time τ_{rise} is a 10%–90% change of amplitude on the rising edge of a signal. However, this stochastic telegraph noise has a noise contribution, coming from the setup, as well as a small stochastic variation in the signal for both the P and AP states due to the thermal excitation of the magnetization vector. For this reason, we define the rise time as the time between the midrange point minus one standard deviation to the midrange point plus one standard deviation: $\tau_{rise} = t_{mid+\sigma} - t_{mid-\sigma}$. The midrange point is defined as $1/2(\max(s) + \min(s))$. Of 100 rising edges, the average rise time is determined to be $\tau_{rise} = 1.5 \pm 0.6$ ns, as shown in Fig. 4(b). The RC time constant has to be shorter than the rise time and is approximately of the order of 1 ns for $R \approx 50\Omega$ and $C \approx 20$ pF. The rising edge (or falling edge) does not indicate an exponential increase (decrease), suggesting that the time resolution of our setup is sufficiently high to acquire the actual signal and that it is not affected by parasitic capacitances in the circuit, which would result in an exponential charge (discharge) characteristics of the signal. For this reason, we only show the shortest measured rise times.

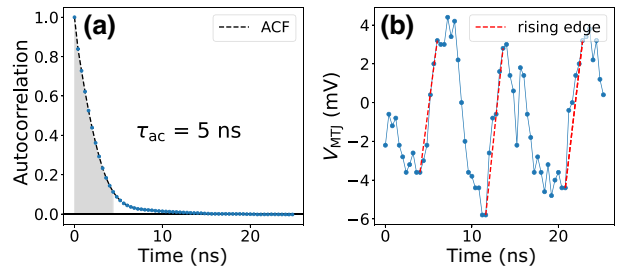


FIG. 4. (a) The autocorrelation function (ACF), plotted for an SMTJ time-series measurement. τ_{ac} is defined as the 99% integral under the ACF curve. (b) A detailed view of the rising and falling edges of the time series. The three red dashed lines show three rising edges, which are approximately linear and of the order of 2 ns.

IV. TRUE-RANDOM-NUMBER GENERATION

Due to the intrinsic stochastic nature of SMTJs, the potential application as an energy-efficient true-random-number generator is of interest and has been proposed and studied by various groups [34–38]. To evaluate the quality of randomness of our stochastic MTJ signal and to clarify the use for potential cryptographic applications as a true-random-number generator, the National Institute of Standards and Technology Statistical Test Suite (NIST STS) [25] was applied to the binarized MTJ signal. The amplified analog output signal was acquired by an oscilloscope and was converted into a Bernoulli sequence of 0s and 1s by comparing the output signal with the median of the acquired time series. This ensures an equal probability of 0 and 1 of this Bernoulli sequence and is a prerequisite for a random bit stream. The random bit stream can then be interpreted as the result of flipping an unbiased coin (head 0, tail 1) with a probability of 50% yielding to a head or a tail, where for true randomness the flips are independent of each other, such that any past result of a coin flip does not affect any future results. The NIST STS provides 15 different tests (188 tests including all subtests) and requires a minimum bit-stream length of the order of 10^6 . There is a considerable amount of work that has been published recently regarding the quality of random-number generation by SMTJs [14,16,22,34]. It has been shown that cryptographic quality randomness, by passing all NIST STS tests, is not reached by a single SMTJ bit

stream but, rather, by combining multiple bit streams with XOR gates (also called “XOR whitening”). This is a common method to improve the quality of randomness and is easy to implement by using standard CMOS-based XOR gates. To further improve the quality of the randomness, nested XOR operations can be applied to multiple bit streams. In an XOR² operation, two rounds of XOR operations are applied to four input bit streams [see blue rectangle in Fig. 5(e)], whereas in an XOR³ operation three rounds of XOR operations are applied to eight input bit streams. With this method, the resulting bit stream of an XOR² operation passes all NIST STS tests, thus providing true randomness, which can be used for cryptographic applications. With an average autocorrelation time of 5 ns and four SMTJs as an input for an XOR² gate, true random bit generation of 200 Mbit/s can be achieved in a very energy-efficient way. Without the consideration of any additional CMOS circuitry, the energy dissipation of a single SMTJ can be of the order of nJ/bit or fJ/bit [34], depending on the applied voltage and resistance of the tunnel-junction device. For the evaluation, 5×10^8 data points were recorded, transformed into a binarized sequence, and divided into eight bit streams of equal size, in order to be able to feed the inputs to the XOR³ gate. Other works achieve cryptographic randomness by passing all NIST STS tests for a bit stream after a XOR² operation [12] or XOR³ operation [34]. Here, we also observe true randomness after an XOR² operation for a sampling time of 5 ns, as shown in Table I. This

TABLE I. The results of the p values for NIST STS tests after applying different XOR operations to the binarized SMTJ telegraph signal. A randomness test is passed and highlighted in green for p values above the significance level of 0.01. For multiple subtests, the particular p values were combined and tested according to Fisher’s method [33]. The sampling times of 3, 5, and 10 ns are chosen to be close to τ_{ac} .

Type	Raw			XOR ¹			XOR ²		
	3 ns	5 ns	10 ns	3 ns	5 ns	10 ns	3 ns	5 ns	10 ns
(1) Frequency	0	0	0	0	0.141	0	0.373	0.739	0.904
(2) Block frequency ($m = 128$)	0	0	0	0	0	0	0	0.703	0.236
(3) Runs	0	0	0	0	0	0	0	0.069	0.817
(4) Longest run	0	0	0	0	0.019	0.001	0.001	0.990	0.185
(5) Binary matrix rank	0.335	0.028	0.897	0.507	0.251	0.072	0.517	0.692	0.632
(6) Discrete Fourier transform (spectral)	0	0.871	0	0	1	0.026	0.765	0.448	0.508
(7) Nonoverlapping template matching	0	0	0	0	0	0	0	0.039	0.313
(8) Overlapping template matching	0	0	0	0	0	0	0	0.623	0.944
(9) Maurer’s “universal statistical”	0	0	0	0	0.074	0	0.182	0.472	0.761
(10) Linear complexity ($M = 500$)	0.467	0.137	0.737	0.967	0.929	0.986	0.174	0.085	0.673
(11) Serial ($m = 16$)	0	0	0	0	0	0	0	0.687	0.900
(12) Approximate entropy ($m = 10$)	0	0	0	0	0	0	0	0.118	0.151
(13) Cumulative sum	0	0	0	0	0	0	0.162	0.172	0.010
(14) Random excursion	0	0.027	0.068	0.068	0.073	0.042	0.073	0.074	0.074
(15) Random-excursion variant	0	0	0	0	0.015	0	0.014	0.015	0.028

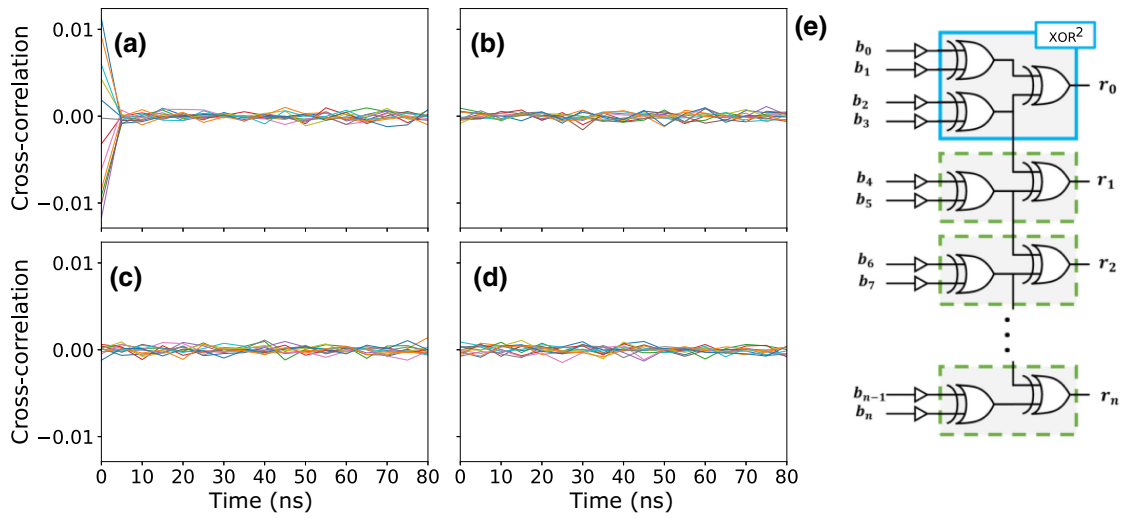


FIG. 5. (a)–(d) Cross-correlations of $r_0 = \text{XOR}^2(b_0, b_1, b_2, b_3)$ with (a) $\text{XOR}^2(b_0, b_1, b_2, b_*)$, (b) $\text{XOR}^2(b_0, b_1, b_*, b_*)$, (c) $\text{XOR}^2(b_0, b_*, b_*, b_*)$, and (d) $\text{XOR}^2(b_*, b_*, b_*, b_*)$ for 20 different raw SMTJ bit streams b_* . (e) Raw bit streams from SMTJs b_i are combined with XOR gates to generate random bit streams r_i . This circuit reduces the required number of SMTJs and XOR gates for generating multiple true random numbers.

sampling time matches the autocorrelation time measured for this time series. Larger sampling times provide the same quality of randomness, whereas for sampling times below 5 ns, each state would be sampled too frequently, thus resulting in nontrue randomness with artificial and nonrandom chunks of only 0 s and only 1 s.

Randomness is a probabilistic property; therefore, a random sequence can be evaluated in terms of the probabilities of the occurrence of a specific pattern in the sequence by means of hypothesis tests. The resulting p values give the probabilities of obtaining the observed pattern assuming that the underlying sequence is random. A common level of significance in cryptography for statistical hypothesis tests is $\alpha = 0.01$, which was also chosen in this work. For each of the 15 different NIST STS tests, the p value was calculated after applying different XOR operations to the binarized SMTJ telegraph signal. All results are summarized in Table I. The raw time-series data in our case never pass all of the NIST STS tests, independent of the chosen sampling rate. If a p value is greater than 0.01 (significance level) for a particular test, then the bit stream is characterized as random and passes this test, whereas for a p value below 0.01, the null hypothesis (of true randomness) is rejected and fails this test.

XOR operations can easily be implemented in CMOS and in our case, a simple random-number generator would require three XOR gates and four SMTJs. However, for further random-number generators, less hardware is necessary, since it is possible to tap the adjacent source randomness and XOR it with a new bit stream coming from the next two SMTJs. This is depicted schematically in Fig. 5(e), where raw bit streams from SMTJs b_i are combined with

XOR gates to generate random bit streams r_i . If r_0 , which is generated by $b_0 - b_3$, is truly random, then $r_1 - r_n$ will provide true random numbers too, due to the same XOR^2 operation. Nonetheless, an additional true random number only requires two more SMTJs and two more XOR gates (for r_1 , b_4 and b_5 are added). A further condition requires the true random numbers $r_0 - r_n$ to be uncorrelated. Therefore, we determined the cross-correlation of any new random number r_* with r_0 , as shown in Figs. 5(a)–5(d). Cross-correlations are calculated for 20 different bit streams ($b_0 - b_{19}$) of length 5×10^6 . Figure 5(a) points out a significant cross-correlation at 0 time lag for r_0 and $r_* = \text{XOR}^2(b_0, b_1, b_2, b_*)$, where b_* is the added new SMTJ bit stream. When two or more new SMTJs are combined, there is no further cross-correlation observed, meaning that the output will be uncorrelated with r_0 . The sequence of SMTJ bit streams to the XOR^2 circuit does not affect the outcome, since the XOR operation is permutation invariant. As long as two new bit streams (b_* and b_*) are combined with two already used bit streams (b_2 and b_3), it will result in a true random and uncorrelated number. It is also confirmed by the comparison with the cross-correlation in Fig. 5(d) with $r_* = \text{XOR}^2(b_*, b_*, b_*, b_*)$, where no qualitative difference is recognizable. Here, all SMTJs are exchanged; thus r_* has to be uncorrelated with r_0 . In addition, we wish to emphasize that the permutation-invariance property of the XOR operation renders possible cross-correlations of the form (r_a, r_b) , with $a \neq b$, to become uncorrelated. Our proposed circuit also offers a significant advantage in its ability to utilize SMTJs that are not precisely set to a balanced 50% equilibrium state. Even if there is a biased probability or deviation from the equilibrium state

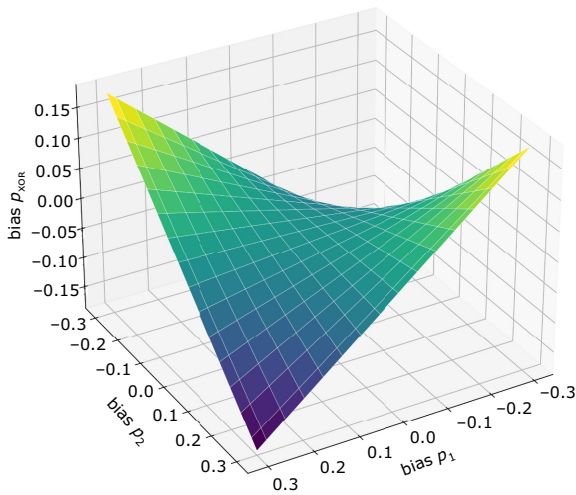


FIG. 6. The stochastic XOR function for two input bit streams with probabilities p_1 and p_2 , which are biased. The bias is the deviation from the “balanced” 0.5, or 50% equilibrium, state.

of SMTJs, a true-random-output bit stream can still be generated as long as the bias (deviation from 50%) is below roughly 10%. This means that the probabilities for each SMTJ can be in the range of 40–60%, which can accommodate fabrication inhomogeneities in the devices and simplify the fabrication process. The balancing effect of the XOR logic is responsible for this property. In stochastic computing, an XOR function performs the following computation on two bit streams with probabilities p_1 and p_2 [39]: $p_{\text{XOR}} = -2p_1p_2 + p_1 + p_2$. When the input probabilities are biased, the XOR logic tends to balance the output toward 50%. This balancing property is demonstrated in Fig. 6 for two biased input probabilities. In the case of our XOR² operation logic, this effect is even greater, allowing for large biases of the SMTJs. Additionally, it should be noted that if biases are antisymmetric, they can even be greater than 10%, since they partially cancel each other out. This feature further demonstrates the versatility and robustness of our proposed circuit. Next, we evaluate how an applied MTJ current controls the bit-stream generation rate and modifies the state probability. Both features are important for the performance of a true-random-number generator.

V. JOULE HEATING AND STT EFFECTS

The application of a current to an SMTJ can affect its switching speed, state probability, and dwell times not only by STT but also by Joule heating at the site of the free layer at the tunneling barrier. Joule heating has most often been studied in nonvolatile MTJs, where a reduction of the critical switching current is desired in order to perform fast and energy-efficient free-layer switching [40,41]. The current through an MTJ for an applied voltage is mainly determined by the RA product and the areal size of the

junction. The RA product, which is strongly dependent on the thickness of the tunneling barrier, defines the current density of any nanopillar size because of the relation $J = V/RA$. For an RA product of $15\Omega \mu\text{m}^2$ and a few hundred millivolts, the typical current density is of the order of 1 MA/cm^2 , where Joule heating can be significant. A simulation for a 50-nm-diameter MTJ suggests that the barrier temperature could increase by more than 100 K after applying 8 MA/cm^2 for a few hundred nanoseconds [41] and according to the Néel-Arrhenius law this will significantly decrease the dwell time of an SMTJ. Here, we have measured dwell times for different current densities and determined the contributions of STT and Joule heating.

Figures 7(e) and 7(f) illustrate the effect of STT and Joule heating. Initially, without STT, both states (P and AP) are tuned to the same energy level, where the free-layer macrospin is considered as a quasiparticle in a one-dimensional potential well. The ambient thermal energy (approximately 25 meV) raises the energy of the quasiparticle, such that the likelihood of overcoming the energy barrier is enhanced. By applying a current to the MTJ, the energy landscape is modified as illustrated in Fig. 7(f), where the AP-state energy level is decreased and the P-state energy level is enhanced by STT in our circuit. In the quasiparticle picture, this results in an energy shift ΔE_{STT} in the energy landscape and is defined here as a positive energy shift at the P state [see Fig. 7(f)]. Higher current densities result in a higher energy level for the P state yielding shorter dwell times, as can be seen for τ_P in Fig. 7(a). The energy shift ΔE_{STT} is positive for lower current densities, which means that the AP energy state is decreased, whereas the P energy state is increased. Furthermore, Joule heating, originating from electron-phonon interaction, raises the energy level of the quasiparticle in the valleys, symmetrically leading to overall faster switching. The temperature increase with respect to room temperature, ΔT_{JH} , reveals a quadratic relation to J [see Fig. 7(c)], as expected for the heating power, which scales with J^2 . Figure 7(b) reveals a similar trend of the average fluctuation rate, which is caused by the Joule-heating effect. This would correspond to a higher bit-generation rate for a random-number generator. The maximum temperature increase at 5.6 MA/cm^2 reaches $36 \pm 5 \text{ K}$ and corresponds to an applied voltage at the tunnel junction of approximately 0.6 V. In order to validate this temperature increase caused by Joule heating, a temperature-dependent dwell-time measurement would provide valuable insights and would allow for a detailed comparison. However, this measurement is part of a future project currently in preparation. The time series for the lowest and highest applied current densities are plotted in Fig. 8. Faster fluctuations with larger amplitude can be observed at high current densities, whereas for low current densities the fluctuation rate and amplitude are lower. By this measurement, we demonstrate that Joule heating has a significant effect on the dwell

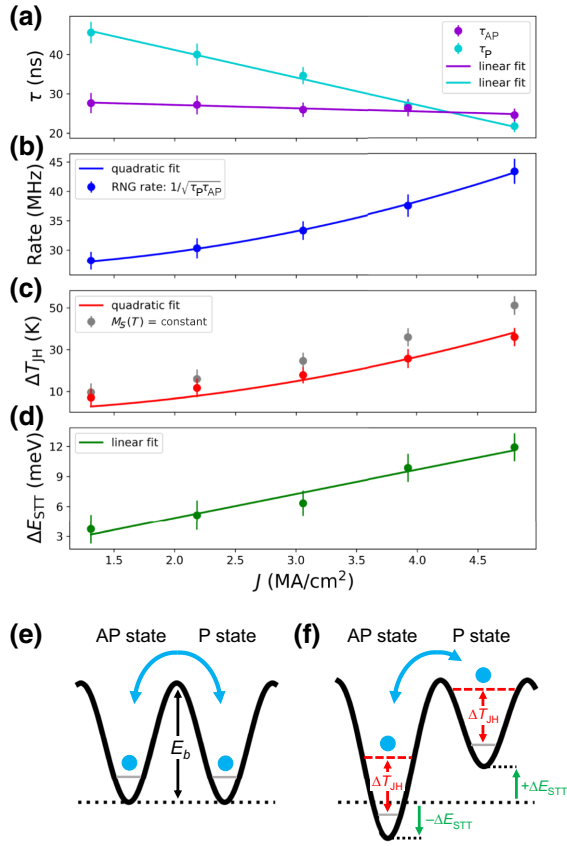


FIG. 7. (a) SMTJ dwell times for current densities ranging from 1.5 to 5.6 MA/cm². Linear fits are applied to indicate the trends. (b) The average fluctuation rate (blue), fitted with a quadratic function. (c) The temperature increase due to Joule heating (red). The nonlinear increase is fitted with a quadratic function. The gray data points indicate the calculated temperature increase without the consideration of temperature-dependent M_s . (d) The energy shift due to the STT (green). Here, a positive energy shift destabilizes the P state. (e) A schematic of the free-layer magnetization orientation described as a quasiparticle in a one-dimensional symmetrical potential well, with two metastable states separated by an energy barrier. (f) The effect of STT antisymmetrically shifts the energy level of the P and AP states and the effect of Joule heating raises the energy level of the quasiparticle symmetrically for both states.

times of an SMTJ and cannot be neglected. For a potential application as a random-number generator or as a p-bit for a probabilistic computer, the random-number-generation rate can be tuned by the current density. When considering an SMTJ as a spiking neuron for neuromorphic computing [42], the current through the junction will modify the spike rate, thus enabling, e.g., rate coding in a neuromorphic network.

A. Calculation of ΔT_{JH} and ΔE_{STT}

The energy shift of the states ΔE_{STT} affects the dwell times according to

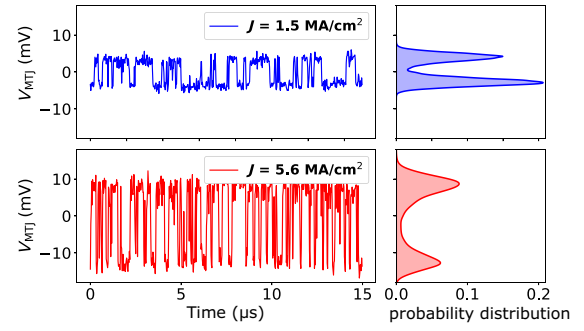


FIG. 8. Time-series measurements of the amplified fluctuating MTJ signal at low (1.5 MA/cm²) and high (5.6 MA/cm²) current density together with the respective probability distribution of each signal.

$$\tau_{P,AP} = \tau_0 \exp\left(\frac{E_b \mp \Delta E_{STT}}{k_B T}\right), \quad (2)$$

where $\tau_{P,AP}$ are the dwell times, T is the temperature, τ_0 is the attempt time, and E_b is the energy barrier between the metastable states. The energy barrier E_b for given dwell times can be determined by

$$E_b = 1/2 k_B T \log(\tau_P \tau_{AP} / \tau_0^2), \quad (3)$$

since the product of τ_P and τ_{AP} is independent of the energy-barrier modification by STT. For this reason, the product of the dwell times for an enhanced current density is also independent of STT:

$$\tau_P^* \tau_{AP}^* = \tau_0^2 \exp\left(\frac{2E_b^*}{k_B T^*}\right). \quad (4)$$

Here, T^* and E_b^* are considered as the altered (in our case, increased) temperature and energy barrier with dwell times τ_P^* and τ_{AP}^* . The increased temperature can be described as $T^* = T + \Delta T_{JH}$ with $T = 293$ K. The energy barrier at elevated temperature, $E_b^* = E_b(T^*) = E_b(M_s(T^*))$, is a function of the temperature-dependent saturation magnetization M_s of the free layer made of Co₄₀Fe₄₀B₂₀, which approximately follows Bloch's $T^{3/2}$ law: $M_s(T) = M_s(0)(1 - (T/T_c)^{3/2})$. For the Curie temperature T_c of Co₄₀Fe₄₀B₂₀, we have chosen 895 K [43]. The in-plane crystalline anisotropy is considered to be constant since a temperature increase of a few kelvin is unlikely to affect the polycrystalline phase of the free layer. The resulting temperature-dependent energy barrier E_b^* is then described as

$$E_b^* = E_b \frac{1 - (T^*/T_c)^{3/2}}{1 - (T/T_c)^{3/2}}. \quad (5)$$

The detailed equation between the dwell times and the temperature for the determination of ΔT_{JH} can be found in

Appendix A 2 [Eq. (A1)]. The asymmetric contribution to the energy barrier, due to STT, is calculated by $\Delta E_{\text{STT}} = 1/2k_B(T + \Delta T_{\text{JH}}) \log(\tau_{\text{AP}}/\tau_{\text{P}})$. However, one has to take into account that the initial P and AP states at low current densities are not in equilibrium, because $\tau_{\text{P}} \neq \tau_{\text{AP}}$ [see Fig. 7(a)]. Shorter AP dwell times arise due to an uncompensated stray field from the TMR stack itself and/or from an external applied in-plane field. With linear fits to the dwell times, we can estimate the theoretical dwell-time values at zero applied current, $\tau_{\text{P}}(J = 0)$ and $\tau_{\text{AP}}(J = 0)$, which are not measurable. However, at $J = 0$ there is neither a temperature increase ($\Delta T_{\text{JH}} = 0$) nor an STT effect ($\Delta E_{\text{STT}} = 0$) and the energy barrier at 293 K can be determined by Eq. (3) to be 95.3 ± 1.1 meV. Since the estimated dwell times at $J = 0$ are not equal and the energy landscape is asymmetric at this point, this shows an initial energy shift of the states due to the Zeeman energy, which is here approximately 11 meV. From Eq. (4), the temperature increase can be calculated numerically [see also Eq. (A1) in Appendix A 2].

VI. CONCLUSIONS

In this work, nanosecond superparamagnetic switching in in-plane MTJs with average dwell times of 6.7 ± 0.1 ns at room temperature is demonstrated and the potential application as a true random bit-stream generator is evaluated by the NIST STS. By employing an XOR² operation on four independent SMTJ bit streams, we demonstrate the successful generation of true random bits with a significantly high throughput rate of 200 Mbit/s. The scalability of our true-random-number generator is demonstrated by a circuit design allowing for multiple uncorrelated true random numbers with fewer stochastic MTJs and a smaller CMOS-area footprint. Further analysis reveals that even SMTJs with biased state probabilities can yield unbiased bit streams due to the balancing effect induced by the XOR operation, effectively passing all NIST tests. The influence of an applied current on the bias or state probability and switching rate of an SMTJ is investigated in detail by the analysis of dwell-time measurements. In particular, we show that the dwell times are linearly dependent on the applied current density and that Joule heating and STT both have a significant effect on the dwell times as well as the dwell-time ratio $\tau_{\text{P}}/\tau_{\text{AP}}$. STT shifts the energy levels of the MTJ states antisymmetrically and has a linear relation to the current density. Joule heating raises both energy levels equally and has a quadratic dependence on the current density. At the free layer, it can reach up to 329 ± 5 K for 5.6 MA/cm² at ambient room temperature. Since the dwell time is strongly dependent on the temperature, the shortest dwell times are observed for the highest current densities. This effect is not negligible and therefore should always

be considered in superparamagnetic MTJs with a low RA product.

The data that support the findings of this study are available from the corresponding author upon request.

ACKNOWLEDGMENTS

This work was supported by the Max Planck Graduate Center with the Johannes Gutenberg-Universität Mainz (MPGC) and used infrastructure provided by For-Lab MagSens. We acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project No. 268565370 (SFB TRR173 Projects A01 and B02) as well as TopDyn and the Zeiss Foundation through the Center for Emergent Algorithmic Intelligence and the Horizon Europe Project No. 101070290 (NIM-FEIA). We would also like to thank T. Reimer for technical support during the development of the samples.

APPENDIX: FURTHER INFORMATION

1. SMTJ voltage signal

For the calculation of the voltage drop at the amplifier input, the circuit can be considered as a voltage divider circuit with shunt resistance R_s , MTJ resistance R_{MTJ} , in either low (parallel) R_{P} or high (antiparallel) resistance state, and input impedance Z_i . The measured voltage drop V can then be determined by:

$$V = V_{\text{src}} \frac{R_{\text{MTJ}} || Z_i}{R_s + R_{\text{MTJ}} || Z_i} = \frac{\frac{R_{\text{MTJ}} Z_i}{R_{\text{MTJ}} + Z_i}}{R_s + \frac{R_{\text{MTJ}} Z_i}{R_{\text{MTJ}} + Z_i}},$$

where V_{src} is the applied dc source voltage. The peak-to-peak voltage signal V_{pp} , which results from the MTJ resistance change, can be calculated by:

$$\begin{aligned} V_{\text{pp}} &= V_{\text{AP}} - V_{\text{P}}, \\ &= V_{\text{src}} \left[\frac{\frac{R_{\text{AP}} Z_i}{R_{\text{AP}} + Z_i}}{R_s + \frac{R_{\text{AP}} Z_i}{R_{\text{AP}} + Z_i}} - \frac{\frac{R_{\text{P}} Z_i}{R_{\text{P}} + Z_i}}{R_s + \frac{R_{\text{P}} Z_i}{R_{\text{P}} + Z_i}} \right]. \end{aligned}$$

For an MTJ resistance of 5–10 k Ω , $R_s = 50$ Ω and $V_{\text{src}} = 500$ mV, the fluctuating MTJ voltage signal (V_{pp}) is in the order of a millivolt.

2. Derivation of the effect of Joule heating

The product of dwell times at temperature T is given by $\tau_{\text{P}} \tau_{\text{AP}} = \tau_0^2 \exp(2E_b/k_B T)$. For an elevated temperature $T^* = T + \Delta T_{\text{JH}}$, the energy barrier is modified according to Bloch's $T^{3/2}$ law [see Eq. (5)]. The following equations

represent the dependence of the Joule heating ΔT_{JH} on the dwell times and can be solved numerically for ΔT_{JH} :

$$\tau_P^* \tau_{AP}^* = \tau_0^2 \exp\left(\frac{2E_b^*}{k_B T^*}\right), \log\left(\frac{\tau_P^* \tau_{AP}^*}{\tau_0^2}\right) = \frac{2E_b}{k_B T^*} \times \frac{1 - (T^*/T_c)^{3/2}}{1 - (T/T_c)^{3/2}} = \frac{T \log(\tau_P \tau_{AP} / \tau_0^2) (1 - ((T + \Delta T_{JH})/T_c)^{3/2})}{(T + \Delta T_{JH}) (1 - (T/T_c)^{3/2})},$$

$$\frac{\log(\tau_P^* \tau_{AP}^* / \tau_0^2)}{\log(\tau_P \tau_{AP} / \tau_0^2)} = \frac{T(1 - ((T + \Delta T_{JH})/T_c)^{3/2})}{(T + \Delta T_{JH})(1 - (T/T_c)^{3/2})}. \quad (\text{A1})$$

Here, τ_P^* and τ_{AP}^* refer to the dwell times at an increased current density, while τ_P and τ_{AP} consider (reference) dwell times at a current density close to 0, where no Joule-heating effect is present.

3. Fisher's method

Fisher's method [33] is often used in meta-analysis to combine the results, namely, the p values, of independent hypothesis tests (each having the same hypothesis) to give a new null hypothesis suggesting that all null hypotheses are true. For a rejection H_0 , at least one of the null hypotheses is rejected. Since the p value of a hypothesis test follows a uniform distribution on the interval $[0,1]$, it can be shown that $-2 \sum_{i=1}^n \log(p_i)$ follows a chi-squared distribution χ_{2n}^2 with $2n$ degrees of freedom. In this work, the p values of several subtests are combined and evaluated according to Fisher's method and the resulting p value is noted in Table I. This holds for tests 11, 13, 14, and 15.

-
- [1] D. Apalkov, B. Dieny, and J. M. Slaughter, Magnetoresistive random access memory, *Proc. IEEE* **104**, 1796 (2016).
- [2] V. Ostwal and J. Appenzeller, Spin-orbit torque-controlled magnetic tunnel junction with low thermal stability for tunable random number generation, *IEEE Magn. Lett.* **10**, 1 (2019).
- [3] J. Grollier, D. Querlioz, K. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, Neuromorphic spintronics, *Nat. Electron.* **3**, 360 (2020).
- [4] K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, Stochastic p-Bits for Invertible Logic, *Phys. Rev. X* **7**, 031014 (2017).
- [5] J. Kaiser, W. A. Borders, K. Y. Camsari, S. Fukami, H. Ohno, and S. Datta, Hardware-Aware *In Situ* Learning Based on Stochastic Magnetic Tunnel Junctions, *Phys. Rev. Appl.* **17**, 014016 (2022).
- [6] S. Ganguly, K. Y. Camsari, and A. W. Ghosh, Reservoir computing using stochastic p-bits (2017), *ArXiv:1709.10211*.
- [7] A. Sengupta, P. Panda, P. Wijesinghe, Y. Kim, and K. Roy, Magnetic tunnel junction mimics stochastic cortical spiking neurons, *Sci. Rep.* **6**, 1 (2016).
- [8] A. Sengupta, M. Parsa, B. Han, and K. Roy, Probabilistic deep spiking neural systems enabled by magnetic tunnel junction, *IEEE Trans. Electron Devices* **63**, 2963 (2016).
- [9] M. W. Daniels, A. Madhavan, P. Talatchian, A. Mizrahi, and M. D. Stiles, Energy-Efficient Stochastic Computing with Superparamagnetic Tunnel Junctions, *Phys. Rev. Appl.* **13**, 034016 (2020).
- [10] W. R. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice* (CRC Press, Chapman & Hall/CRC, London, 1995).
- [11] K. Y. Camsari, B. M. Sutton, and S. Datta, p-bits for probabilistic spin logic, *Appl. Phys. Rev.* **6**, 011305 (2019).
- [12] T. Kim, H. Park, K.-H. Han, Y.-J. Nah, H. C. Koo, B.-C. Min, S. Hong, and O. Lee, Demonstration of in-plane magnetized stochastic magnetic tunnel junction for binary stochastic neuron, *AIP Adv.* **12**, 075104 (2022).
- [13] B. R. Zink, Y. Lv, and J.-P. Wang, Telegraphic switching signals by magnet tunnel junctions for neural spiking signals with high information capacity, *J. Appl. Phys.* **124**, 152121 (2018).
- [14] M. Bapna and S. A. Majetich, Current control of time-averaged magnetization in superparamagnetic tunnel junctions, *Appl. Phys. Lett.* **111**, 243107 (2017).
- [15] B. Parks, A. Abdelgawad, T. Wong, R. F. Evans, and S. A. Majetich, Magnetoresistance Dynamics in Superparamagnetic Co-Fe- B Nanodots, *Phys. Rev. Appl.* **13**, 014063 (2020).
- [16] C. Safranski, J. Kaiser, P. Trouilloud, P. Hashemi, G. Hu, and J. Z. Sun, Demonstration of nanosecond operation in stochastic magnetic tunnel junctions, *Nano Lett.* **21**, 2040 (2021).
- [17] K. Hayakawa, S. Kanai, T. Funatsu, J. Igarashi, B. Jinnai, W. Borders, H. Ohno, and S. Fukami, Nanosecond Random Telegraph Noise in In-Plane Magnetic Tunnel Junctions, *Phys. Rev. Lett.* **126**, 117202 (2021).
- [18] S. Kanai, K. Hayakawa, H. Ohno, and S. Fukami, Theory of relaxation time of stochastic nanomagnets, *Phys. Rev. B* **103**, 094423 (2021).
- [19] G. Reiss, J. Ludwig, and K. Rott, Superparamagnetic dwell times and tuning of switching rates in perpendicular CoFeB/MgO/CoFeB tunnel junctions (2019), *ArXiv:1908.02139*.
- [20] M. Bapna, S. K. Piotrowski, S. D. Oberdick, M. Li, C.-L. Chien, and S. A. Majetich, Magnetostatic effects on switching in small magnetic tunnel junctions, *Appl. Phys. Lett.* **108**, 022406 (2016).

- [21] W. A. Borders, A. Z. Pervaiz, S. Fukami, K. Y. Camsari, H. Ohno, and S. Datta, Integer factorization using stochastic magnetic tunnel junctions, *Nature* **573**, 390 (2019).
- [22] B. Parks, M. Bapna, J. Igbokwe, H. Almasi, W. Wang, and S. A. Majetich, Superparamagnetic perpendicular magnetic tunnel junctions for true random number generators, *AIP Adv.* **8**, 055903 (2018).
- [23] K. Kobayashi, W. A. Borders, S. Kanai, K. Hayakawa, H. Ohno, and S. Fukami, Sigmoidal curves of stochastic magnetic tunnel junctions with perpendicular easy axis, *Appl. Phys. Lett.* **119**, 132406 (2021).
- [24] S. Krause, G. Herzog, A. Schlenhoff, A. Sonntag, and R. Wiesendanger, Joule Heating and Spin-Transfer Torque Investigated on the Atomic Scale Using a Spin-Polarized Scanning Tunneling Microscope, *Phys. Rev. Lett.* **107**, 186601 (2011).
- [25] L. E. Bassham III, A. L. Rukhin, J. Soto, J. R. Nechvatal, M. E. Smid, E. B. Barker, S. D. Leigh, M. Levenson, M. Vangel, and D. L. Banks, *et al.*, *Sp 800-22 rev. 1a. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications* (National Institute of Standards and Technology, Gaithersburg, MD, 2010).
- [26] L. Schnitzspan, J. Cramer, J. Kubik, M. Tarequzaman, G. Jakob, and M. Kläui, Impact of annealing temperature on tunneling magnetoresistance multilayer stacks, *IEEE Magn. Lett.* **11**, 1 (2020).
- [27] R. Stamps, Mechanisms for exchange bias, *J. Phys. D: Appl. Phys.* **33**, R247 (2000).
- [28] S. Jenkins, A. Meo, L. E. Elliott, S. K. Piotrowski, M. Bapna, R. W. Chantrell, S. A. Majetich, and R. F. Evans, Magnetic stray fields in nanoscale magnetic tunnel junctions, *J. Phys. D: Appl. Phys.* **53**, 044001 (2019).
- [29] D. Worledge and P. Trouilloud, Magnetoresistance measurement of unpatterned magnetic tunnel junction wafers by current-in-plane tunneling, *Appl. Phys. Lett.* **83**, 84 (2003).
- [30] P. Hänggi, P. Talkner, and M. Borkovec, Reaction-rate theory: Fifty years after Kramers, *Rev. Mod. Phys.* **62**, 251 (1990).
- [31] W. F. Brown Jr, Thermal fluctuations of a single-domain particle, *Phys. Rev.* **130**, 1677 (1963).
- [32] O. Hassan, R. Faria, K. Y. Camsari, J. Z. Sun, and S. Datta, Low-barrier magnet design for efficient hardware binary stochastic neurons, *IEEE Magn. Lett.* **10**, 1 (2019).
- [33] F. Mosteller and R. A. Fisher, Questions and answers, *Am. Stat.* **2**, 30 (1948).
- [34] D. Vodenicarevic, N. Locatelli, A. Mizrahi, J. S. Friedman, A. F. Vincent, M. Romera, A. Fukushima, K. Yakushiji, H. Kubota, and S. Yuasa, *et al.*, Low-Energy Truly Random Number Generation with Superparamagnetic Tunnel Junctions for Unconventional Computing, *Phys. Rev. Appl.* **8**, 054045 (2017).
- [35] A. Fukushima, T. Seki, K. Yakushiji, H. Kubota, H. Imaura, S. Yuasa, and K. Ando, Spin dice: A scalable truly random number generator based on spintronics, *Appl. Phys. Express* **7**, 083001 (2014).
- [36] X. Chen, J. Zhang, and J. Xiao, *et al.*, Magnetic-Tunnel-Junction-Based True Random-Number Generator with Enhanced Generation Rate, *Phys. Rev. Appl.* **18**, L021002 (2022).
- [37] N. Rangarajan, A. Parthasarathy, and S. Rakheja, A spin-based true random number generator exploiting the stochastic precessional switching of nanomagnets, *J. Appl. Phys.* **121**, 223905 (2017).
- [38] Z. Fu, Y. Tang, X. Zhao, K. Lu, Y. Dong, A. Shukla, Z. Zhu, and Y. Yang, An overview of spintronic true random number generator, *Front. Phys.* **9** (2021).
- [39] A. Alaghi and J. P. Hayes, in *Proceedings of the 25th edition on Great Lakes Symposium on VLSI* (2015), p. 331.
- [40] D. Lee and S. H. Lim, Increase of temperature due to joule heating during current-induced magnetization switching of an MgO-based magnetic tunnel junction, *Appl. Phys. Lett.* **92**, 233502 (2008).
- [41] J. J. Kan, C. Park, C. Ching, J. Ahn, Y. Xie, M. Pakala, and S. H. Kang, A study on practically unlimited endurance of STT-MRAM, *IEEE Trans. Electron Devices* **64**, 3639 (2017).
- [42] A. Mizrahi, T. Hirtzlin, A. Fukushima, H. Kubota, S. Yuasa, J. Grollier, and D. Querlioz, Neural-like computing with populations of superparamagnetic basis functions, *Nat. Commun.* **9**, 1 (2018).
- [43] K.-M. Lee, J. W. Choi, J. Sok, and B.-C. Min, Temperature dependence of the interfacial magnetic anisotropy in W/CoFeB/MgO, *AIP Adv.* **7**, 065107 (2017).

Electrical coupling of superparamagnetic tunnel junctions mediated by spin-transfer-torques

Cite as: Appl. Phys. Lett. **123**, 232403 (2023); doi: 10.1063/5.0169679

Submitted: 27 July 2023 · Accepted: 15 November 2023 ·

Published Online: 4 December 2023



View Online



Export Citation



CrossMark

Leo Schnitzspan,^{1,2}  Mathias Kläui,^{1,2}  and Gerhard Jakob^{1,2,a)} 

AFFILIATIONS

¹Institute of Physics, Johannes Gutenberg-University Mainz, 55122 Mainz, Germany

²Max Planck Graduate Center Mainz, 55122 Mainz, Germany

^{a)}Author to whom correspondence should be addressed: jakob@uni-mainz.de

ABSTRACT

In this work, the effect of electrical coupling on stochastic switching of two in-plane superparamagnetic tunnel junctions (SMTJs) is studied, using experimental measurements as well as simulations. The coupling mechanism relies on the spin-transfer-torque effect, which enables the manipulation of the state probability of an SMTJ. Through the investigation of time-lagged cross-correlation, the strength and direction of the coupling are determined. In particular, the characteristic state probability transfer curve of each SMTJ leads to the emergence of a similarity or dissimilarity effect. The cross-correlation as a function of applied source voltage reveals that the strongest coupling occurs for high positive voltages for our SMTJs. In addition, we show state tunability as well as coupling control by the applied voltage. The experimental findings of the cross-correlation are in agreement with our simulation results.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0169679>

Magnetic tunnel junctions (MTJs) consist of two ferromagnetic layers, separated by an insulating layer, and exhibit a large resistance change from the parallel to the antiparallel state, caused by the tunnel magnetoresistance (TMR) effect. Due to low energy consumption for the nonvolatile configuration and CMOS compatibility, MTJs are well suited for memory devices, such as magnetic random access memory (MRAM).¹ However, if the energy barrier between the parallel (P) and antiparallel (AP) state is small, thermal excitation can induce spontaneous fluctuations between these states.² In recent studies, random fluctuations in superparamagnetic tunnel junctions (SMTJs) have been demonstrated across time scales ranging from milliseconds to nanoseconds.^{3–7} Moreover, the potential of SMTJs for generating true random numbers has been recognized,^{7–10} which holds particular significance for cryptographic applications that require a high quality of randomness. Additionally, SMTJs are promising candidates in neuromorphic spintronics, as artificial synapses or neurons.¹¹ In larger spintronic systems, coupling between individual elements can occur due to spin torques acting on the magnetization of these devices. For instance, previous investigations have demonstrated the electrical coupling of spin-torque nano-oscillators (STNO).^{12,13} Understanding and investigating the coupling behavior of stochastic MTJs is of paramount importance for the implementation of effective probabilistic systems, such as Boltzmann machines.¹⁴ The coupling between SMTJs refers to the interaction and influence of one junction on another, which can be

established through various effects, including dipole coupling,¹⁵ strain,¹⁶ electrical interaction via spin-transfer-torques (STT),^{17,18} or by a CMOS circuit.¹⁹

Here, we build upon prior work on electrical coupling of SMTJs.^{17,18} Our focus is on investigating the effect of coupling on stochastic switching in a series circuit under the influence of DC voltage control. This approach provides several notable benefits, including scalability, rapid mediation of coupling, and ease of device implementation. So far, electrical coupling of two-level superparamagnetic tunnel junctions in parallel circuitry has only been studied under the influence of external magnetic fields¹⁷ or via an applied current source.¹⁸ However, for practical applications, these approaches might not be ideal in terms of feasibility, and therefore, we focus on the tunability of two SMTJs in series under an applied DC voltage. Electrical coupling might also be promising for neuromorphic computing approaches²⁰ and probabilistic networks^{21,22} based on stochastic MTJs, such as Boltzmann¹⁴ or Ising machines.²³

TMR stacks were deposited at room temperature on oxidized Si substrates using rf- and dc-magnetron sputtering (Singulus Rotaris) with the following composition (film thickness in nanometer): Ta(10)/Ru(10)/Ta(10)/PtMn(20)/CoFe(2.2)/Ru(0.8)/CoFeB(2.4)/MgO(1.1)/CoFeB(3.0)/Ta(10)/Ru(30) and is based on an optimized stack, developed earlier.²⁴ The TMR ratio of our stack is found to be $\sim 100\%$ and the resistance area (RA) product around $15 \Omega \mu\text{m}^2$. Annealing was carried out in a

300 mT in-plane magnetic field at 300 °C for 1 h. An in-plane field of a few milli-Tesla is typically applied in order to compensate for stray fields from other ferromagnetic layers in the stack and to tune the stochastic switching. Nanopillars were patterned in circular shapes of diameters of ~ 60 nm. By structuring circular MTJs with a nanometer scale, the low in-plane anisotropy results in a low energy barrier, which can lead to superparamagnetic fluctuation caused by thermal excitations in the free layer's magnetization.^{5–7} It is worth emphasizing that appropriate compensation of stray fields at the free layer is crucial to ensure volatility and not a pinning of a particular state.

After nanopillar patterning, two stochastic MTJs were interconnected in series, as illustrated in Fig. 1(a). The MTJ resistance states are in the kilohm range and the resulting volatile voltage drop, as shown in Figs. 1(a) and 1(b), is acquired by an oscilloscope (Tektronix DPO7543, used sampling rate: 400 kHz). Depending on the resistance states of both SMTJs, either three or four states are effectively distinguishable, since two different states can result in almost the same voltage drop $R_B/(R_A + R_B)$. In particular, the combination of SMTJ B and C provides similar voltage outputs for the states (P, P) and (AP, AP), leading to a total of three states, as shown in Fig. 1(c). To describe the coupled mixed states of two SMTJs A and B, we use a bracket notation: (S_A, S_B) , where S_i stands for the current state (P or AP) of the i th MTJ.

We conducted simulations to investigate the stochastic fluctuations of two electrically coupled superparamagnetic tunnel junctions arranged in series. We consider that each SMTJ on its own behaves like a Bernoulli distributed random variable of probability p . For a specific voltage and temperature, an SMTJ i has a certain probability p_i to occupy the AP state. A modification in the bias voltage across the junction will result in an altered probability p_i due to the STT effect. The relation between AP probability and the applied voltage is here called a PV-transfer function. However, the (parallel or antiparallel) resistance also exhibits non-linear behavior in response to the bias voltage, which is denoted as the RV-transfer function and must be considered in the simulation. Due to the superparamagnetic property of the MTJ free layer, random resistance fluctuations (also called telegraph noise) occur, which can be characterized by dwell times τ of the P- and AP-state. The dwell time in the macrospin approximation for $E_b/k_B T \geq 1$, follows the Néel–Arrhenius law:²⁵ $\tau = \tau_0 \exp(E_b/k_B T)$,

where E_b is the energy barrier between the states, k_B is the Boltzmann constant, T is the temperature, and τ_0 is the attempt time. The assumption of the Néel–Arrhenius model is switching between two distinct energy minima, i.e., two distinct resistance states. For simplicity, we use $\Delta = E_b/k_B T$ in the following. Dwell times can be manipulated by applying a current or voltage to the tunnel junction, which affects the effective energy barrier through the effect of spin-transfer torque,^{26,27}

$$\tau_{p,ap} = \tau_0 e^{\Delta(1 \pm I/I_c^{p,ap})}. \quad (1)$$

In general, in high-TMR MTJs, the absolute of the critical switching current for the antiparallel to parallel switching I_c^p is lower than the switching current of the parallel to antiparallel switching I_c^{ap} .²⁸ However, considering the critical switching voltage, which is defined via the MTJ resistance and the critical switching current, it is found to be similar for the parallel and antiparallel state. Therefore, we consider $|V_c^p| \approx |V_c^{ap}| = V_c$ and rewrite the equation for the dwell times to^{26–28}

$$\tau_{p,ap} = \tau_0 e^{\Delta(1 \pm V/V_c)}. \quad (2)$$

Here, τ_0 is the attempt time, Δ is the (unitless) energy barrier, V is the voltage across the junctions, and V_c is the critical voltage for deterministic switching at 0 K.^{26,27} The dwell time ratio $\tau_{ap}/(\tau_{ap} + \tau_p)$ then describes the probability to observe the AP state. With Eq. (2), we can derive the characteristic voltage-dependent AP probability

$$P_{ap}(V) = \frac{\tau_{ap}}{\tau_{ap} + \tau_p} = \frac{e^{\Delta(1-V/V_c)}}{e^{\Delta(1-V/V_c)} + e^{\Delta(1+V/V_c)}},$$

$$= \frac{1}{e^{2\Delta V/V_c} + 1}. \quad (3)$$

Here, Δ is the energy barrier, V is the applied voltage, and V_c is the critical voltage. Equation (3) represents the typical sigmoid relation and is used to model the PV-transfer function for both SMTJs.

For the generation of artificial random telegraph noise, which can be considered as a two-state Markov process, the Poisson distribution of dwell times has to be considered. Telegraph noise is observed in various physical systems, like flash memory²⁹ or field effect transistors,³⁰ and is characterized by an exponential distribution of dwell times. Hence, the Poisson process can describe the probability of a stochastic MTJ switching to another state.² Moreover, at low current densities in

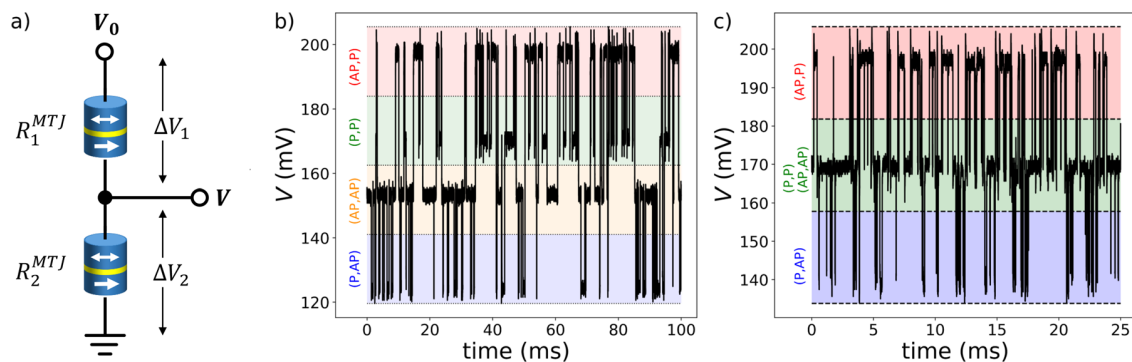


FIG. 1. (a) Sketch of the measurement setup of two SMTJs in series. The source voltage V_0 is applied and the voltage drop V between both SMTJs is measured. A voltage difference across an MTJ resistance R_i^{MTJ} is indicated by ΔV_i . (b) and (c) Oscilloscope measurements of two coupled SMTJs ($V_0 = 0.3$ V). The resistance band of three or four separate states (P, P), (AP, P), (P, AP), and (AP, AP) is depicted by background colors.

stochastic MTJs, the switching event follows a probability density function (PDF) given by Ref. 31: $f^{sw}(t) = 1/\tau \cdot e^{-t/\tau}$ for a mean dwell time τ , which is the inverse of the average switching rate. Therefore, the thermal switching probability that the magnetization switches within time t in an MTJ with modified energy barrier can be expressed as²⁶

$$F_{p,ap}^{sw}(t) = 1 - \exp(-t/\tau_0 e^{-\Delta(1 \pm V/V_c)}),$$

$$= 1 - e^{-t/\tau_{p,ap}},$$

$$\approx t/\tau_{p,ap} \quad (t \ll \tau_{p,ap}), \quad (4)$$

which is the cumulative distribution function (CDF) of $f^{sw}(t)$.³² As a consequence, the switching probabilities from the P to the AP state and vice versa can be expressed as $P_{p \rightarrow ap}^{sw} = 1 - e^{-t/\tau_p}$, $P_{ap \rightarrow p}^{sw} = 1 - e^{-t/\tau_{ap}}$. These two switching probabilities are relevant for a single stochastic MTJ. However, for two coupled SMTJs, there are 12 different switching probabilities, due to the electrical coupling. In general, for n SMTJs in series, there exist 2^n distinct Markov states with $4^n - 2^n$ transition probabilities.

To simulate a time series of these Markov states for two coupled SMTJs, we start by choosing average dwell times $\tau = \sqrt{\tau_p \tau_{ap}}$ for both SMTJs and a time step dt for each iteration. During the time step dt , the switching probability P^{sw} for each SMTJ can be calculated according to Eq. (4) (for $dt \rightarrow 0$ also $P^{sw} \rightarrow 0$). Next, we draw a binary random variable from a Bernoulli distribution: $\mathcal{B}(P_i^{sw})$ for SMTJ i . If switching occurs ($S_i \rightarrow \bar{S}_i$), then this will affect the switching probability P_i^{sw} of both SMTJs, due to a change in the voltage drop ΔV_i caused by a change in the resistance states R_i . Since for each time step the MTJ states are known, the resistance states can be derived. However, because the MTJ resistance depends on the voltage (RV-transfer function) due to the tunneling, the MTJ resistance and voltage drop will both affect each other until an equilibrium is found. This problem is solved by a recursive function of V and R in the simulation. The (equilibrium) voltage drop ΔV_i is then used to calculate the modified switching probabilities P_i^{sw} for each SMTJ, according to the characteristic PV-transfer functions. From the updated switching probabilities, the next MTJ states can be sampled, and the simulation process starts from the beginning, which ensures a sequential series of Markov states.

In order to obtain good statistical results, stochastic time series were simulated for 0.5×10^7 time steps (iterations), which corresponds to almost 10^5 switching events for each SMTJ; thus, the likelihood for a switching event to occur in a single time step is around 1%. For the simulation, the PV-transfer curves are modeled by a sigmoidal function [see Eq. (3)], whereas the RV-transfer curves are described by an approximation according to Brinkman *et al.*³³ The reduction in resistance for higher absolute bias voltages is due to the non-linear tunneling current and can be represented by the following function:³³ $R(V) = a/(1 + bV^2) + c$, where a , b , and c are (fitting) parameters.

In order to evaluate the coupling strength, a cross-correlation function of simulated (or measured) time series data can be determined. Cross-correlation is commonly used in signal processing for pattern recognition³⁴ or as a measure of similarity of two series (A and B) as a function of the relative shift of one series with respect to the other. In our case, the cross-correlation of two SMTJ time series is calculated and can be interpreted as a “switching correlation.” A high correlation value implies that a switching event of one SMTJ significantly impacts the switching likelihood of the second SMTJ. A positive (negative) time lag corresponds to a positive (negative) shift of time series B with respect to A . The cross-correlation function is defined as

$$C(t) = \frac{\sum_i \frac{1}{N} (A_i - \bar{A})(B_{i-t} - \bar{B})}{s_A s_B}, \quad (5)$$

where t is the (time) lag ($t \in [-N/2, \dots, N/2]$ in multiples of the incremental time step dt), \bar{A} , \bar{B} are the means, and s_A and s_B are the standard deviations of time series A and B with (time) length N . The values A_i and B_i with $i \notin [0, 1, \dots, N - 1]$ are 0 and do not contribute to the cross-correlation.

Figure 2 highlights the simulation results, assuming instantaneous state transitions, the same average switching rates, and the absence of additional noise. The cross-correlation function reveals a peak at 0 time lag, which is expected in the context of two interacting time series. This behavior arises due to the influence of switching events in one MTJ on the state of the other MTJ, caused by changes in STT. The correlation amplitude primarily depends on the TMR ratios and the applied voltage to the coupled SMTJs, since the SMTJs are

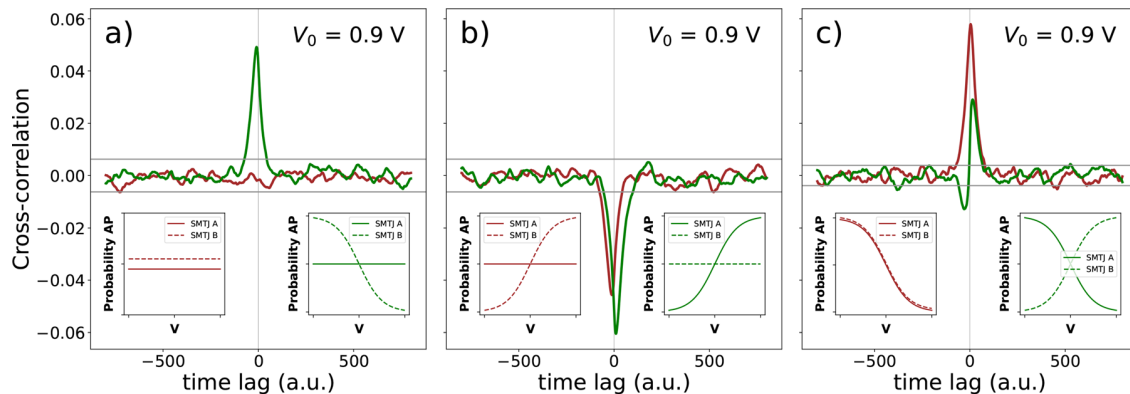


FIG. 2. A simulation of the cross-correlation depending on different PV-transfer curves. (a) In red, the cross-correlation function is shown for two constant PV-transfer curves, as depicted by the left inset. In green, the correlation is shown for a sigmoid transfer curve for SMTJ A and a constant transfer curve for SMTJ B, as depicted by the right inset. (b) and (c) demonstrate the effect on cross-correlation for different PV-transfer curves of SMTJ A and B.

TABLE I. Cross-correlation summary.

Time Lag	Correlation	Direction	Effect
+	+	$B \rightarrow A$	Similarity
+	-	$B \rightarrow A$	Dissimilarity
-	+	$A \rightarrow B$	Similarity
-	-	$A \rightarrow B$	Dissimilarity

most “sensitive” to voltage drops in regions where the gradient of the PV-transfer function is maximal. The correlation peaks are positive or negative, and shifted to positive or negative time lags. If a peak is on the positive side of the cross-correlation function, then the effect of series B on A is obtained, while for the negative side, the effect of A on B is accessed. Therefore, we are able to measure the effect of both SMTJs on each other. A positive correlation would coincide with a preference of both SMTJs to stay in the same state (“similarity effect”), while a negative correlation coincides with a preference for opposite states (“dissimilarity effect”). These findings are illustrated in Fig. 2 and summarized in Table I. In Fig. 2(a), the cross-correlation is simulated for two SMTJs, where the coupling is artificially turned off (red curve), which corresponds to constant PV-curves, and for the case where only the PV-curve of SMTJ A is constant, while the one for SMTJ B shows the sigmoid function (green curve). In the first case, no cross-correlation peak is observed, while for the second case, a positive peak emerges in the negative time lag range, which indicates that SMTJ A affects SMTJ B and B prefers to adopt the same state as A . Figure 2(b) illustrates the effect of switching the PV-transfer curves of A and B , as shown in the insets, where as a consequence also the correlation peak shifts from negative (red curve) to positive (green curve) time lag range. A more realistic simulation is depicted in Fig. 2(c) for two coupled SMTJs with sigmoid PV-transfer functions and bilateral coupling. If both PV-transfer functions exhibit the same trend, such as both monotonically increasing, it will lead to one symmetric correlation peak centered at 0. However, if one of the PV-transfer functions is reversed [$f(x) \rightarrow f(-x)$], it results in an

anti-symmetrical peak, with positive as well as negative correlations close to 0 time lag, each peak, respectively, corresponding to an SMTJ.

In order to verify our simulation results, we conducted time series measurements for different combinations of SMTJs, as illustrated in Fig. 1. Depending on the resistance states of each MTJ, this results in four possible resistance states, leading to four different voltage drops. Figure 3(c) displays the characteristic time series of SMTJs A , B , and C for 0.2 V. Due to differences in nanopillar size and variations of the tunnel barriers caused by the fabrication process, the P and AP resistance levels of A , B , and C differ from each other. In addition, the average switching rate of B and C is higher due to lower energy barriers in these junctions. In Fig. 3, the characteristic RV- and PV-transfer functions are plotted. The lowering of the resistance for a nonvolatile MTJ state in Fig. 3(a) is attributed to the non-linear tunneling current and the resistance curve is fitted by the following approximation:³³ $R(V) = a(1 + bV^2) + c$, where a , b , and c are fitting parameters. In Fig. 3(b), each data point corresponds to a time series, for which the AP state probability was determined. The applied fit is a sigmoid function based on Eq. (3). For two SMTJs in series, there exist four different states, which also depend on the applied source voltage, since each state probability is influenced by the characteristic PV-transfer curve of each SMTJ. Figure 4(a) highlights the state manipulation through the source voltage V_0 , demonstrating the impact of voltage control on the coupled system. Notably, the maximum applied voltage of $V_0 = 1$ V here corresponds to a maximum MTJ current density of ~ 3 MA/cm². For large negative voltages, the dominant state is the (AP, AP) state, however, in the positive voltage range, this state is less often occupied mainly for the sake of state (P, P). This indicates that both SMTJs have the same probability trend with V , in which positive voltages stabilize the parallel while negative voltages stabilize the anti-parallel state. However, the coupling strength cannot be easily quantified from this plot.

To investigate the coupling between both SMTJs, the cross-correlation of the time series A and B was analyzed. For this, the states of SMTJ A and B were derived from the measured (mixed) time series, as shown in Fig. 1(b). For instance, the P state of SMTJ A corresponds to the measured voltage of the coupled states of (P, AP) and (P, P),

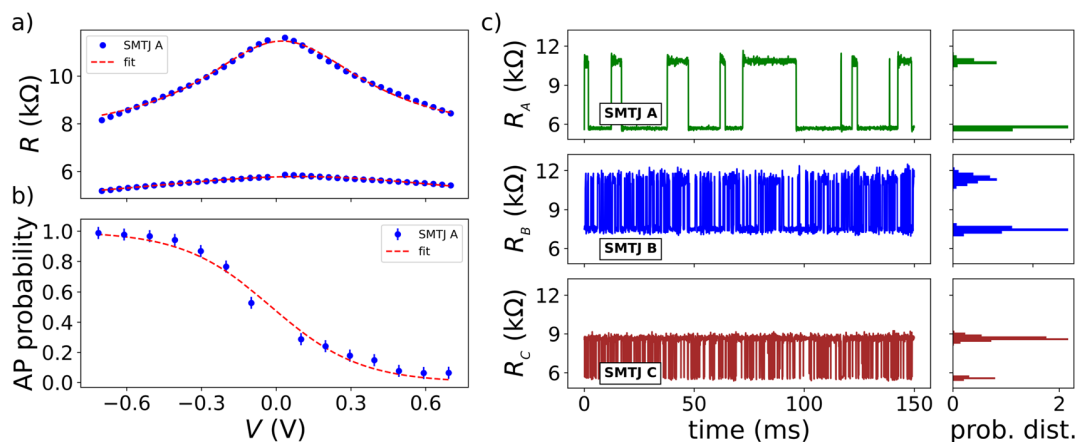


FIG. 3. (a) MTJ resistance as a function of applied bias voltage for a fixed free layer. (b) AP probability transfer curve for the stochastic MTJ with the sigmoid fit in red. (c) Time series of three different SMTJs A , B , and C along with histograms illustrating their respective probability distributions.

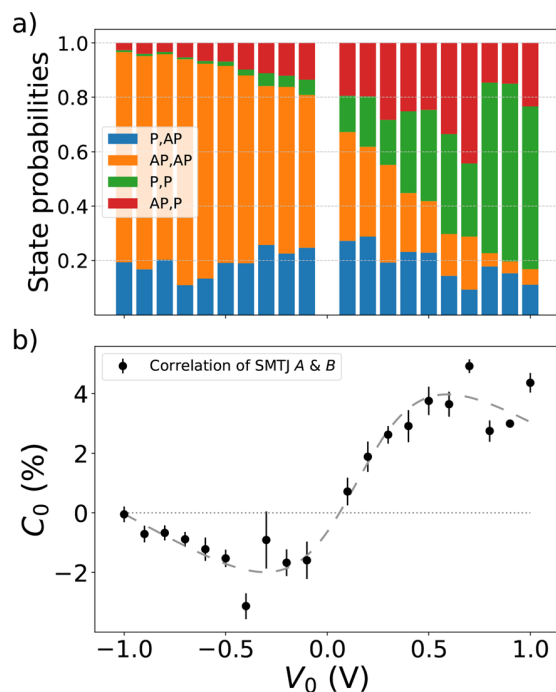


FIG. 4. (a) The dependence of state probability on the source voltage V_0 for two coupled SMTJs is illustrated. (b) Cross-correlation at a time lag 0 is plotted as a function of applied voltage for two coupled SMTJs. The dashed line indicates the correlation trend.

whereas the AP state corresponds to all states of (AP, AP) and (AP, P), as indicated in Fig. 1(b). The states of the derived binary sequence of each SMTJ are converted to a sequence of -1 and $+1$, representing the parallel and antiparallel states, respectively. The normalized cross-correlation [see Eq. (5)], which is equivalent to the normalized Pearson correlation coefficient with time lag t , is used to assess the strength of coupling. Results are shown in Fig. 5(a) together with the derived autocorrelations. The cross-correlation signal exhibits a small peak of around 5% at 0 time lag and declines exponentially for increasing time lags as outlined in the inset figure. In Fig. 5(b), the same cross-correlation is illustrated over a time lag range of ± 4 s. Here, the sharp peak at 0 indicates a significant positive coupling. Due to the characteristic sigmoidal PV-transfer curves, the cross-correlation is expected to change with applied voltage. Therefore, we measured the maximum correlation close to 0 time lag depending on V_0 . The results are shown in Fig. 4(b). The asymmetry with respect to $V_0 = 0$ in Fig. 4 stems from the fact that the PV-transfer curves are not symmetric with respect to $V = 0$, but rather shifted due to stray fields present at the magnetic free layer. For this reason, the 50/50 probability state is not found at zero voltage but rather at an offset voltage, leading to a biased probability, which can be seen in the probability distribution in Fig. 4(a) and cross-correlation in Fig. 4(b). On the positive voltage side, positive correlations can be observed, which can be attributed to the “reversed” sigmoid curve, which results in positive correlations, as elucidated in the simulation section. Higher voltages V_0 lead to higher correlations due to the higher voltage drops across the junctions, thereby exerting a greater influence on the state probabilities.

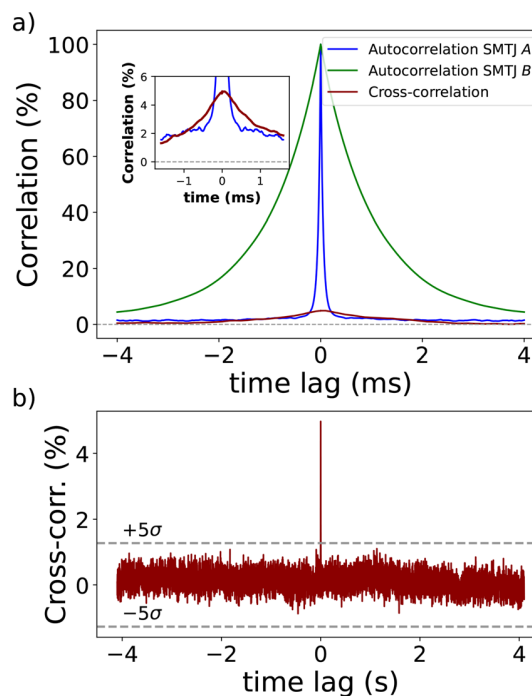


FIG. 5. (a) Autocorrelation functions of SMTJ A and B as well as their cross-correlation function are plotted ($V_0 = 0.7$ V). (b) Cross-correlation ($V_0 = 0.7$ V) for a time lag range of ± 4 s. A $\pm 5\sigma$ level is indicated as a dashed line.

Consequently, at low absolute voltages, the coupling strength is reduced, falling close to zero. However, for high negative voltages, the correlation also approaches zero, since here almost no switching occurs because both SMTJs predominantly remain in the AP state [see Fig. 4(a)]. Although the coupling effect persists for negative voltages, it becomes challenging to quantify it using cross-correlation techniques. Furthermore, the trend of correlations undergoes a sign change at $V_0 = 0$, changing from negative to positive correlations. This can be explained by the “reversed” impact of negative voltages on the PV-transfer function. Considering the impact of dwell times on cross-correlation, it becomes evident that the average dwell time determines the coupling timescale and, consequently, the width of the cross-correlation peak. In cases where two SMTJs with significantly different average dwell times are coupled, such as in our experiment, the SMTJ with slower switching characteristics exerts a more substantial influence on the faster-switching SMTJ than vice versa. As a result, the cross-correlation peak is primarily dictated by the “slower” SMTJ, which has a more pronounced influence on the other SMTJ, which behaves more sensitive due to a faster response on modified voltages.

In conclusion, this study highlights the effective coupling of two serially-connected superparamagnetic tunnel junctions (SMTJs). The controllability of this electrical coupling is demonstrated through the manipulation of the applied voltage. Through comprehensive simulations, we reveal that the coupling effect is significantly influenced by the probability transfer function of each individual SMTJ. Here, spin-transfer torques play a pivotal role, especially when modulated by the MTJ resistance through the TMR effect, which governs the change in

charge current and thereby also the change in spin-current. The switching cross-correlation function is shown to be an effective tool for quantifying both the strength and the nature of the coupling. Our results reveal that SMTJs can exhibit coupling tendencies toward the same state, the opposite state, or a combination of both. Experimental measurements confirm the existence of coupling, even in cases with different average switching rates, and indicate a preference for SMTJs to align into the same states under positive voltages. Notably, the switching cross-correlation undergoes a sign change for negative voltages, and a peak value of $\sim 5\%$ is observed at 0.7 V. The coupling strength can be enhanced by increasing the TMR ratio or by stronger spin-transfer torques. Overall, we demonstrated the ability to control the stochasticity of electrically connected SMTJs, enabling the generation of a tunable probability distribution, which is a main feature in Bayesian neural networks. In particular, the simple electrical coupling effect can be harnessed easily to enable stochastic MTJ-based Boltzmann machines.

This work was supported by the Max Planck Graduate Center with the Johannes Gutenberg-Universität Mainz (MPGC) and used infrastructure provided by ForLab MagSens. We acknowledge the support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project No. 268565370 (SFB TRR173 Projects A01 and B02), by TopDyn and the Zeiss Foundation through the Center for Emergent Algorithmic Intelligence, and the Horizon Europe Project No. 101070290 (NIMFEIA) as well as ERC-2019-SyG No. 856538 (3D MAGiC). We would also like to thank T. Reimer for his technical support during the development of the samples.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Leo Schnitzspan: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Project administration (lead); Software (lead); Validation (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (equal). **Mathias Kläui:** Funding acquisition (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal). **Gerhard Jakob:** Funding acquisition (equal); Resources (equal); Supervision (lead); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon request.

REFERENCES

- D. Apalkov, B. Dieny, and J. M. Slaughter, “Magnetoresistive random access memory,” *Proc. IEEE* **104**, 1796–1830 (2016).
- S. Kanai, K. Hayakawa, H. Ohno, and S. Fukami, “Theory of relaxation time of stochastic nanomagnets,” *Phys. Rev. B* **103**, 094423 (2021).
- G. Reiss, J. Ludwig, and K. Rott, “Superparamagnetic dwell times and tuning of switching rates in perpendicular CoFeB/MgO/CoFeB tunnel junctions,” [arXiv:1908.02139](https://arxiv.org/abs/1908.02139) (2019).
- M. Bapna and S. A. Majetich, “Current control of time-averaged magnetization in superparamagnetic tunnel junctions,” *Appl. Phys. Lett.* **111**, 243107 (2017).
- C. Safranski, J. Kaiser, P. Trouilloud, P. Hashemi, G. Hu, and J. Z. Sun, “Demonstration of nanosecond operation in stochastic magnetic tunnel junctions,” *Nano. Lett.* **21**, 2040–2045 (2021).
- K. Hayakawa, S. Kanai, T. Funatsu, J. Igarashi, B. Jinnai, W. Borders, H. Ohno, and S. Fukami, “Nanosecond random telegraph noise in in-plane magnetic tunnel junctions,” *Phys. Rev. Lett.* **126**, 117202 (2021).
- L. Schnitzspan, M. Kläui, and G. Jakob, “Nanosecond true-random-number generation with superparamagnetic tunnel junctions: Identification of Joule heating and spin-transfer-torque effects,” *Phys. Rev. Appl.* **20**, 024002 (2023).
- D. Vodenicarevic, N. Locatelli, A. Mizrahi, J. S. Friedman, A. F. Vincent, M. Romera, A. Fukushima, K. Yakushiji, H. Kubota, S. Yuasa *et al.*, “Low-energy truly random number generation with superparamagnetic tunnel junctions for unconventional computing,” *Phys. Rev. Appl.* **8**, 054045 (2017).
- A. Fukushima, T. Seki, K. Yakushiji, H. Kubota, H. Imamura, S. Yuasa, and K. Ando, “Spin dice: A scalable truly random number generator based on spintronics,” *Appl. Phys. Express* **7**, 083001 (2014).
- X. Chen, J. Zhang, J. Xiao *et al.*, “Magnetic-tunnel-junction-based true random-number generator with enhanced generation rate,” *Phys. Rev. Appl.* **18**, L021002 (2022).
- J. Grollier, D. Querlioz, K. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, “Neuromorphic spintronics,” *Nat. Electron.* **3**, 360–370 (2020).
- K. Yogendra, M. Koo, and K. Roy, “Energy efficient computation using injection locked bias-field free spin-hall nano-oscillator array with shared heavy metal,” in *IEEE International Symposium on Nanoscale Architectures (NANOARCH)* (IEEE, 2017), pp. 89–94.
- J. Torrejon, M. Riou, F. A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima *et al.*, “Neuromorphic computing with nanoscale spintronic oscillators,” *Nature* **547**, 428–431 (2017).
- J. Kaiser, W. A. Borders, K. Y. Camsari, S. Fukami, H. Ohno, and S. Datta, “Hardware-aware *in situ* learning based on stochastic magnetic tunnel junctions,” *Phys. Rev. Appl.* **17**, 014016 (2022).
- M. T. McCray, M. A. Abeer, and S. Bandyopadhyay, “Electrically programmable probabilistic bit anti-correlator on a nanomagnetic platform,” *Sci. Rep.* **10**, 12361 (2020).
- A. Welbourne, A. Levy, M. Ellis, H. Chen, M. Thompson, E. Vasilaki, D. Allwood, and T. Hayward, “Voltage-controlled superparamagnetic ensembles for low-power reservoir computing,” *Appl. Phys. Lett.* **118**, 202402 (2021).
- P. Talatchian, M. W. Daniels, A. Madhavan, M. R. Pufall, E. Jué, W. H. Rippard, J. J. McClelland, and M. D. Stiles, “Mutual control of stochastic switching for two electrically coupled superparamagnetic tunnel junctions,” *Phys. Rev. B* **104**, 054427 (2021).
- N.-T. Phan, L. Soumah, A. Sidi El Valli, L. Hutin, L. Anghel, U. Ebels, and P. Talatchian, “Electrical coupling of perpendicular superparamagnetic tunnel junctions for probabilistic computing,” in *IEEE International Symposium on Nanoscale Architectures* (IEEE, 2022), pp. 1–6.
- P. Debashis, R. Faria, K. Y. Camsari, S. Datta, and Z. Chen, “Correlated fluctuations in spin orbit torque coupled perpendicular nanomagnets,” *Phys. Rev. B* **101**, 094405 (2020).
- A. Mizrahi, T. Hirtzlin, A. Fukushima, H. Kubota, S. Yuasa, J. Grollier, and D. Querlioz, “Neural-like computing with populations of superparamagnetic basis functions,” *Nat. Commun.* **9**, 1533 (2018).
- W. A. Borders, A. Z. Pervaiz, S. Fukami, K. Y. Camsari, H. Ohno, and S. Datta, “Integer factorization using stochastic magnetic tunnel junctions,” *Nature* **573**, 390–393 (2019).
- K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, “Stochastic p-bits for invertible logic,” *Phys. Rev. X* **7**, 031014 (2017).
- N. A. Aadit, A. Grimaldi, M. Carpentieri, L. Theogarajan, J. M. Martinis, G. Finocchio, and K. Y. Camsari, “Massively parallel probabilistic computing with sparse Ising machines,” *Nat. Electron.* **5**, 460–468 (2022).
- L. Schnitzspan, J. Cramer, J. Kubik, M. Tarequzzaman, G. Jakob, and M. Kläui, “Impact of annealing temperature on tunneling magnetoresistance multilayer stacks,” *IEEE Magn. Lett.* **11**, 4503705 (2020).

- ²⁵P. Hänggi, P. Talkner, and M. Borkovec, "Reaction-rate theory: Fifty years after Kramers," *Rev. Mod. Phys.* **62**, 251 (1990).
- ²⁶Z. Li and S. Zhang, "Thermally assisted magnetization reversal in the presence of a spin-transfer torque," *Phys. Rev. B* **69**, 134416 (2004).
- ²⁷J. Z. Sun, "Spin-current interaction with a monodomain magnetic body: A model study," *Phys. Rev. B* **62**, 570 (2000).
- ²⁸Z. Li, S. Zhang, Z. Diao, Y. Ding, X. Tang, D. Apalkov, Z. Yang, K. Kawabata, and Y. Huai, "Perpendicular spin torques in magnetic tunnel junctions," *Phys. Rev. Lett.* **100**, 246602 (2008).
- ²⁹C. M. Compagnoni, R. Gusmeroli, A. S. Spinelli, A. L. Lacaita, M. Bonanomi, and A. Visconti, "Statistical model for random telegraph noise in flash memories," *IEEE Trans. Electron Devices* **55**, 388–395 (2008).
- ³⁰K. K. Hung, P. K. Ko, C.-M. Hu, and Y. C. Cheng, "Random telegraph noise of deep-submicrometer MOSFET's," *IEEE Electron Device Lett.* **11**, 90–92 (1990).
- ³¹A. F. Vincent, N. Locatelli, J.-O. Klein, W. S. Zhao, S. Galdin-Retailleau, and D. Querlioz, "Analytical macrospin modeling of the stochastic switching time of spin-transfer torque devices," *IEEE Trans. Electron Devices* **62**, 164–170 (2015).
- ³²Y. Higo, K. Yamane, K. Ohba, H. Narisawa, K. Bessho, M. Hosomi, and H. Kano, "Thermal activation effect on spin transfer switching in magnetic tunnel junctions," *Appl. Phys. Lett.* **87**, 082502 (2005).
- ³³W. Brinkman, R. Dynes, and J. Rowell, "Tunneling conductance of asymmetrical barriers," *J. Appl. Phys.* **41**, 1915–1921 (1970).
- ³⁴D.-M. Tsai and C.-T. Lin, "Fast normalized cross-correlation for defect detection," *Pattern Recognit. Lett.* **24**, 2625–2631 (2003).

