

Reservoir Computing Based on Spintronics Technology



Tomohiro Taniguchi, Sumito Tsunegi, Shinji Miwa, Keisuke Fujii,
Hitoshi Kubota, and Kohei Nakajima

Abstract Recent developments in reservoir computing based on spintronics technology are described here. The rapid growth of brain-inspired computing has motivated researchers working in a broad range of scientific field to apply their own technologies, such as photonics, soft robotics, and quantum computing, to brain-inspired computing. A relatively new technology in condensed matter physics called spintronics is also a candidate for application to brain-inspired computing because the small size of devices (nanometer order), their low energy consumption, their rich magnetization dynamics, and so on are advantageous for realization of highly integrated network systems. In fact, several interesting functions, such as a spoken-digit recognition and an associative memory operation, achieved using spintronics technology have recently been demonstrated. Here, we describe our recent advances in the development of recurrent neural networks based on spintronics auto-oscillators, called spin-torque oscillators, such as experimental estimation of the short-term memory capacity of a vortex-type spin-torque oscillator and numerical simulation

T. Taniguchi (✉) · S. Tsunegi · H. Kubota

National Institute of Advanced Industrial Science and Technology (AIST), Research Center for Emerging Computing Technologies, Central 2, 1-1-1, Umezono, Tsukuba, Ibaraki 305-8568, Japan

e-mail: tomohiro-taniguchi@aist.go.jp

S. Tsunegi

e-mail: tsunegi.sb@aist.go.jp

H. Kubota

e-mail: hit-kubota@aist.go.jp

S. Miwa

The Institute for Solid State Physics, The University of Tokyo, Kashiwa, Chiba 277-8581, Japan
e-mail: miwa@issp.u-tokyo.ac.jp

K. Fujii

Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan
e-mail: fujii@qc.ee.es.osaka-u.ac.jp

K. Nakajima

Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo, Tokyo 113-8656, Japan
e-mail: k_nakajima@mech.t.u-tokyo.ac.jp

© Springer Nature Singapore Pte Ltd. 2021

K. Nakajima and I. Fischer (eds.), *Reservoir Computing*, Natural Computing Series,
https://doi.org/10.1007/978-981-13-1687-6_14

of reservoir computing using several macromagnetic oscillators. The results demonstrate the potential high performance of spintronics technology and its applicability to brain-inspired computing.

1 Introduction

A new candidate element for brain-inspired computing has recently appeared: spintronics, which is the main topic of this chapter. In this section, we provide a brief introduction to this research field.

1.1 Recurrent Neural Network and Reservoir Computing

The brain has long been a fascinating research target in science. Nonlinear interactions between neurons in the brain via synapses carry and/or store information, and prompt the growth of living things. Human beings, as well as biological systems in general, show surprising computational ability due to their brains. The brain can perform various functions, including (visual and/or audio logic) recognition, associative memory, prediction from past experiences, action modification, and task optimization. Moreover, the human brain has notably low power consumption for computation. If we could replace current computing systems based on the von Neumann architecture by neural networks, our lives would be drastically changed. Brain-inspired computing aimed at achieving artificial neural networks has attracted much attention in a wide range of scientific research fields, such as physics, chemistry, biology, engineering, and nonlinear science.

A crucial task in the advancement of brain-inspired computing is developing an appropriate model (Gerstner et al. 2014; Goodfellow et al. 2017) and implementing it in a real system. A class of artificial neural network, called recurrent neural network (RNN), is an architecture exhibiting a dynamic response to a time sequence of an input data (Mandic and Chambers 2001). The response of an RNN depends not only on the input and the system's weights at a certain time but also on the input at some previous time. In other words, RNNs store past input information. Thus, RNNs enable a time sequence of input data such as data on spoken languages and movies to be classified and calculated.

A reservoir computing system is an RNN in which the internal weights of the network are not changed by learning; only the reservoir-to-output weights are trained (Maass et al. 2002; Jaeger and Haas 2004; Verstraeten et al. 2007; Appeltant et al. 2011; Nakajima 2020). This approach simplifies the tuning of the weights used for computing and enables any physical system to be used as a "reservoir." Various such systems have been developed, including a photonic architecture with delayed feedback (Brunner et al. 2013), a soft robotic reservoir using body dynamics generated

from a soft silicone arm (inspired by octopus's arm in water) (Nakajima et al. 2015), and a quantum reservoir consisting of several qubits (Fujii and Nakajima 2017).

The purpose of this chapter is to introduce a promising new candidate for reservoir computing: spintronics. Spintronics is a relatively young research field in condensed matter physics and has been developed by investigating spin-dependent electron transport in solids with nanometer (nm) scale. Spintronics technology has several advantages for reservoir computing, such as its applicability to an array structure with nanoscale dimensions, low energy consumption, and strong nonlinearity. However, the research targets in spintronics to date have been electronic devices, such as nonvolatile memory and microwave generators, and the ability of spintronics technology from the viewpoint of artificial neural networks has not yet been thoroughly investigated. What is required in reservoir computing is rich (or complex) dynamics of physical systems because the reservoir should show several and distinguishable responses to different sequences and/or inputs. The richness (or complexity) of a dynamical system is quantitatively characterized by its memory and nonlinear capacities, which characterize, respectively, the amount of information the RNN can store and its nonlinear computational capability. Here, the capacities are estimated by examining the response to binary input as done in Jaeger (2002). Short-term memory capacity is simply characterized how much past information fed into the system can be stored and reconstructed from the current output of the reservoir with trained weight. This task can be accomplished even in a linear system, in principle. Therefore, the nonlinear computational capability on the stored information is further characterized by the parity check capacity by predicting the parity of the past input sequence, which essentially requires nonlinearity. For example, the short-term memory capacity of a quantum reservoir consisting of several qubits (≥ 4) with virtual nodes (≥ 4) was reported to be of the order of 10 (Fujii and Nakajima 2017). In this chapter, we describe our recent measurements of memory capacity in spintronics devices (Tsunegi et al. 2018a; Furuta et al. 2018).

We start by describing the history of this research field, and discuss the applicability of spintronics technology to brain-inspired computing.

1.2 History of Spintronics and Key Technologies

An electron has two degrees of freedom, charge and spin. Research on electricity and magnetism has a long history, and it was unified at the end of the nineteenth century as classical electromagnetism (Jackson 1999). People began to notice that both electricity and magnetism originate from a charged particle, but it took time to widely accept the existence of elementary particles (or atoms and molecules) until the beginning of the twentieth century. The rapid growth of quantum mechanics provided a new picture of electrons and of elementary particles, in general, which resulted in a drastic change in our understanding of the nature. In particular, the discovery of spin led the transition from classical to quantum physics because it was the first finding of an internal degree of freedom in elementary particles (Dirac

1928). Spin has relativistic quantum mechanical origin (Weinberg 2005), determines the statistics of the elementary particles (Pauli 1940), and is related to many physics, such as the stimulated emission of light (Dirac 1927), the origin of ferromagnetism (Heisenberg 1928), superconductivity (Bardeen et al. 1957), and the fate of stars (Chandrasekhar 1984). Importantly, the quantum mechanics has also resulted in the development of applied physics. Electronics based on semiconductors has provided many devices useful in everyday life. Magnetic devices as well have become widely used for such things as a storage device. Charge and spin have been, however, used separately for electronics and magnetic devices, respectively.

Spintronics (or spin electronics) is both electronics and magnetism at nanoscale and aims to develop new electronics devices, such as nonvolatile memory [magnetoresistive random access memory (MRAM)] and microwave generators, in which spin plays a key role. Quantum mechanical interactions between the spins of conducting electrons in metals and semiconductors, magnetization in ferromagnets, magnons (spin-wave quanta) in insulating ferromagnets, light helicity, and so on provide interesting new phenomena in fundamental physics and have opened the door to practical applications. The growth of spintronics is due to advances in nanostructure fabrication processes. This is because the spin of electrons in condensed matter is not a conserved quantity. For example, the spin in metals relaxes over the length scale (the spin diffusion length) (Valet and Fert 1993), which is typically of the order of nanometers (Bass and Pratt 2007); therefore, nanostructures are necessary to observe spin-dependent phenomena.

A key phenomenon in spintronics is the tunneling magnetoresistance (TMR) effect, which was found by Julliere in a magnetic tunnel junction (MTJ) consisting of Fe/Ge/Co (Julliere 1975). The magnetoresistance effect is a physical phenomenon in which the resistance of an electrical circuit depends on the magnetization directions of the ferromagnets. The TMR effect originates from spin-dependent tunneling through an insulator, and so it is a purely quantum mechanical phenomenon. Although several other magnetoresistance effects, such as anisotropic magnetoresistance and the planar Hall effect, have been found since the nineteenth century and their origins were revealed to be quantum mechanical phenomena in the middle of the twentieth century (Thomson 1856; Kundt 1893; Pugh and Rostoker 1953; Karplus and Luttinger 1954; McGuire and Potter 1975; Sinitsyn 2008; Nagaosa et al. 2010), the discovery of the TMR effect in 1975 is often regarded as a beginning of spintronics. Since the electrons pass through the interfaces between multilayers, the structure of an MTJ is often called a “current-perpendicular-to-plane (CPP)” structure. The research target shifted from MTJs (Julliere 1975; Maekawa and Gafvert 1982; Maekawa and Shinjo 2002) to giant magnetoresistive (GMR) systems consisting of metallic ferromagnetic/nonmagnetic multilayers when the relatively large magnetoresistance effect, called the “GMR effect,” was found in 1988–1989 by Fert and Grunberg independently (Baibich et al. 1988; Binasch et al. 1989). The GMR effect was first observed in a current-in-plane (CIP) structure, in which the electrons move in a direction parallel to the metallic interface. It was, however, soon noticed that the CPP structure shows a larger GMR effect than the CIP structure (Pratt et al. 1991; Zhang and Levy 1991). Research interest shifted to back MTJs when a large TMR effect was found in

an Fe/AlO/Fe MTJ in 1995 (Miyazaki et al. 1995; Moodera et al. 1995) and a giant TMR effect was found in an MTJ with an Fe/MgO/Fe MTJ in 2004 (Yuasa et al. 2004a; Parkin et al. 2004; Yuasa et al. 2004b).

A magnetoresistance effect, such as the TMR effect, enables the magnetization direction in a ferromagnet to be detected by using an electrical circuit. Typical TMR and GMR devices include two ferromagnets, and the resistance of the circuit is low (high) when the alignment of their magnetizations is parallel (antiparallel). In addition, the theoretical prediction of a spin-transfer torque (STT) effect by Slonczewski and Berger provided a new technology controlling the magnetization direction (Slonczewski 1989, 1996; Berger 1996; Slonczewski 2005). After the discovery of magnetism, the magnetization direction of a ferromagnet has been controlled by applying an external magnetic field. The STT effect is a completely different phenomenon, where an electric current is applied to a ferromagnetic/nonmagnetic/ferromagnetic trilayer structure. One ferromagnet, the “reference layer,” acts as a polarizer of the spin angular momentums of the conducting electrons. When these spin-polarized electrons are injected into the other ferromagnet, the “free layer,” the spins of the conducting electrons are transferred to the magnetization of the ferromagnet via quantum mechanical interaction and excite STT (Ralph and Stiles 2008). When the magnitude of the electric current is sufficiently large, the STT can change the magnetization direction in the free layer. Such magnetization reversals were observed in a GMR device in 2000 (Katine et al. 2000) and in an MTJ in 2005 (Kubota et al. 2005).

In summary, there are two important phenomena in spintronics, the TMR and STT effects. The magnetization direction of a nanostructured ferromagnet can be detected electrically by using the TMR effect and can be changed by using the STT effect. These are the operation principles for any kind of spintronics device (Locatelli et al. 2014). In fact, as discussed below, brain-inspired computing based on spintronics technology utilizes these phenomena. There are many other interesting phenomena in spintronics, such as the spin pumping effect (Silsbee et al. 1979; Tserkovnyak et al. 2002; Mizukami et al. 2002), the spin Hall effect (Dyakonov and Perel 1971; Hirsch 1999; Kato et al. 2004), and the topological effect (Murakami et al. 2003). Readers who are interested in spintronics can easily find good textbooks and review articles, such as Maekawa (2006), Shinjo (2009), and Maekawa et al. (2012).

1.3 Brain-Inspired Computing Based on Spintronics

There are many advantages of applying spintronics technology to brain-inspired computing (Grollier et al. 2016). First of all, the small size of spintronics device and their applicability to an array structure enable development of high-density computing devices. The typical size of an MTJ is of the order of 1 nm in thickness and 10–100 nm in diameter (Fukushima et al. 2018; Watanabe et al. 2018). The non-volatility of the magnetization as memory results in low power consumption during operation (Dieny et al. 2016). The large output signal due to the giant TMR effect

is another advantage for electronics applications (Yuasa et al. 2004a; Parkin et al. 2004; Yuasa et al. 2004b; Djayaprawira et al. 2005; Ikeda et al. 2008; Yakushiji et al. 2010; Sukegawa et al. 2013; Kubota et al. 2013; Tsunegi et al. 2016a). The figure of merit for TMR and GMR devices is the magnetoresistance ratio, which is defined as $(R_{AP} - R_P)/R_P$, where R_P and R_{AP} are the resistances of the system for parallel (P) and antiparallel (AP) magnetization alignments. Currently, the TMR ratio is of the order of 100%, which results in a visible change in the output signal that is sufficient to experimentally detect by overcoming noise. Strong nonlinearity of the magnetization dynamics excited by the STT effect (Bertotti et al. 2005, 2007, 2009; Slavin and Tiberkevich 2009) is desirable for computation calculating a time sequence of input data. In fact, as mentioned below, we found that an MTJ with TMR and STT effects is applicable as an element for reservoir computing. The fast relaxation time of the magnetization dynamics (Zhou et al. 2010; Kudo et al. 2010; Suto et al. 2011; Nagasawa et al. 2011; Rippard et al. 2013; Jenkins et al. 2016; Taniguchi et al. 2017; Tsunegi et al. 2018b), of the order of 1–100 nanoseconds (ns), is suitable for fast computing. Application of these features to brain-inspired computing by using spintronics technology has recently been reported.

An associative memory operation in a network consisting of CIP devices was reported by Borders, Fukami, and their collaborators in 2017 (Borders et al. 2017). They used analog switching of the magnetization in a CoNi ferromagnet placed on a PtMn antiferromagnet as memory for storing the weight of computing. A total of 36 devices were used to demonstrate the associative memory operation of identifying a word consisting of three alphabetic letters.

Spoken-digit recognition competitive to that of state-of-the-art neural networks was demonstrated by Torrejon, Grollier, and their collaborators, including some of the present authors of this chapter, also in 2017 (Torrejon et al. 2017). A vortex spin-torque oscillator (STO) was used for the computation, and the information of a sequence of spoken words was stored using the nonlinear dynamics of the magnetic vortex. A high recognition rate ($\geq 95\%$) was obtained for spoken digits (0–9) pronounced by five speakers. The detail of this work is explained in the other chapter of this book. Vowel recognition with four coupled STOs was also achieved recently (Romera et al. 2018). In the following sections, we will revisit artificial neural networks based on an STO.

There have also been several reports of numerical simulation focused on brain-inspired computing based on a spintronics computing architecture (Furuta et al. 2018; Vodenicarevic et al. 2017; Kudo and Morie 2017; Huang et al. 2017; Nakane et al. 2018; Chen et al. 2018; Pinna et al. 2018; Arai and Imamura 2018; Nomura et al. 2018). For example, Kudo and Morie demonstrated pattern recognition using an array of STOs in 2017 (Kudo and Morie 2017). Nakane and his collaborators investigated the possibility of reservoir computing based on spin-wave excitation and detection in 2018 (Nakane et al. 2018). As can be seen, a wide variety of device designs have been proposed. This is due to the rich physics of nanostructured devices. The magnetization dynamics in a ferromagnet is usually highly nonlinear. Since spintronics devices consist of nanomagnets, magnetic interactions, such as dipole interaction and spin wave, can be used as an operation principle. In addition, electrical

interaction through the STT effect can also be an operation principle for such devices. Accordingly, spintronics provides interesting examples of dynamical systems and therefore can be considered a strong candidate as a fundamental element of brain-inspired computing, see, for example, a review (Grollier et al. 2020) including both experiments and theory.

1.4 Spin-Torque Oscillator (STO)

A key device for brain-inspired computing based on spintronics technology is the STO (Torrejon et al. 2017). An STO typically consists of an MTJ, for which the main structure is a trilayer (reference layer/insulating barrier/free layer), as illustrated in Fig. 1a. The typical material for the barrier is MgO (Yuasa et al. 2004a, b; Parkin et al. 2004). As mentioned above, when an electric current is injected into an MTJ, STT is excited on the magnetization in the free layer. The STT leads to magnetization dynamics when the magnitude of the electric current is sufficiently large. The magnetization dynamics is typically classified as reversal (Kubota et al. 2005) or oscillation (Kiselev et al. 2003). Magnetization reversal has been used as an operation principle for MRAM (Dieny et al. 2016). Magnetization oscillation has been widely studied for both GMR and TMR structures with the aim of developing practical applications such as microwave generator, highly sensitive sensors, and phased array radar (Kiselev et al. 2003; Rippard et al. 2004; Houssameddine et al. 2007). An STO device uses this magnetization oscillation in an MTJ for such practical applications.

Since STOs are nonlinear oscillator showing an auto-oscillation (a limit cycle), there are a dissipation due to friction and source of energy injection, as is any auto-oscillator in nature (Pikovsky et al. 2003). The source of the energy is the work

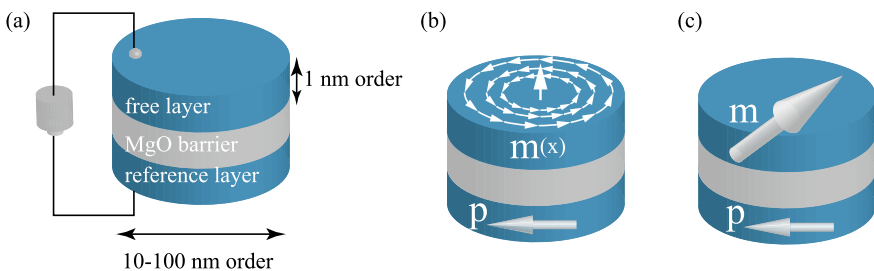


Fig. 1 **a** Schematic diagram of fundamental structure of MTJ connected to electric battery. **b** Schematic diagram of MTJ with vortex-type free layer. Each white arrow represents local direction of magnetic moment $\mathbf{m}(\mathbf{x})$. Magnetic moment at center points in the perpendicular direction with respect to film surface and is called a “vortex core.” Magnetic moments around core lie in film plane. **c** Schematic diagram of MTJ with macro-magnetic free layer. Almost all local magnetic moments point in same direction, so local variable $\mathbf{m}(\mathbf{x})$ can be replaced with macroscopic variable \mathbf{m} . In both **b**, **c**, direction of magnetization in reference layer, which is assumed to be macromagnetic, is represented as \mathbf{p}

done by the STT, while the dissipation is relaxation minimizing the magnetic energy of the ferromagnets. The oscillation frequency ranges from 100MHz to 100GHz depending on the device structure and magnetic configuration, see Sects. 2.1, 2.2, and 2.3 for more detail. The oscillation of the magnetization in an STO is detected through the TMR effect. The resistance of the MTJ is $R = 1/G$, where conductance G is given by

$$G = \frac{G_P + G_{AP}}{2} + \frac{G_P - G_{AP}}{2} \cos \theta, \quad (1)$$

where θ is the relative angle between the magnetizations of the free and reference layers, and $G_P = 1/R_P$ and $G_{AP} = 1/R_{AP}$ are the conductances for the parallel and antiparallel magnetization alignments. Therefore, the resistance, or voltage, of the STO oscillates over time when the magnetization in the free layer is in an auto-oscillation state. In other words, the oscillation state of the STO can be detected by electrical measurement.

Our motivation focusing on reservoir computing based on spintronics technology is as follows. Imagine that we fabricate an array of STOs. Each STO interacts with the other STOs through magnetic and/or electrical interactions. Such a system may experience phase synchronization (Slavin and Tiberkevich 2009), in which the phase differences between the magnetizations saturate to certain values. This synchronization phenomenon may be applicable to some portion of brain-inspired computing (Kudo and Morie 2017). It should be noted, however, that the coupling strength between the STOs normally cannot be changed after fabrication of the array. This means that the interconnections (weights) between the elements in the RNN cannot be tuned for computing. Reservoir computing solves this problem and thus opens the door to the applicability of STO devices to brain-inspired computing. The highly nonlinear, and thus complex, dynamics of STOs should provide rich dynamics to a reservoir and is unnecessary to be tuned for computing. To clarify the applicability of STO devices to reservoir computing it is necessary to evaluate the performance capabilities of STOs, which is the main objective of this chapter (Tsunegi et al. 2018a; Furuta et al. 2018).

2 Methods

In this section, we summarize the methods we used to evaluate the short-term memory and parity check capacities of reservoir computing with STOs.

2.1 Landau–Lifshitz–Gilbert Equation

We start by introducing a theoretical model of the magnetization dynamics frequently used in the fields of magnetism and spintronics. It is often assumed that the mag-

netization dynamics in a ferromagnet is described by the Landau–Lifshitz–Gilbert (LLG) equation with spin-transfer torque (Slonczewski 1996; Landau and Lifshitz 1935):

$$\frac{d\mathbf{m}}{dt} = -\gamma\mathbf{m} \times \mathbf{H} - \frac{\gamma\hbar P g(\theta) J}{2eM_s d} \mathbf{m} \times (\mathbf{p} \times \mathbf{m}) + \alpha\mathbf{m} \times \frac{d\mathbf{m}}{dt}, \quad (2)$$

where \mathbf{m} is the unit vector pointing in the local magnetization direction. Magnetic field \mathbf{H} consists of both internal and external fields; the internal fields include an exchange field, anisotropy field, and dipole field. Parameter γ , the “gyromagnetic ratio,” is related to Bohr magneton μ_B and Landé g -factor g_L via $\gamma = g_L \mu_B / \hbar$. Saturation magnetization M_s is the magnetic moment per volume, whereas d is the thickness of the free layer. Spin polarization P ($0 \leq P < 1$) characterizes the strength of the STT, and the vector \mathbf{p} ($|\mathbf{p}| = 1$), which is usually parallel to the magnetization of the reference layer, determines the direction of the STT. The current density flowing in the MTJ is denoted as J . Factor $g(\theta)$, a function of the relative angle between \mathbf{m} and \mathbf{p} ($\cos \theta = \mathbf{m} \cdot \mathbf{p}$), determines the angular dependence of the STT. The explicit form of $g(\theta)$ depends on the structure of the spintronics device (Xiao et al. 2004; Taniguchi et al. 2015). Dimensionless parameter α , the “damping constant,” which characterizes the strength of the dissipation due to friction of the magnetization dynamics (Gilbert 2004) and is of the order of $10^{-3} - 10^{-2}$ for typical ferromagnets used in spintronics devices (Oogane et al. 2006). The magnetization dynamics in a ferromagnet has been investigated by solving the LLG equation both analytically (Bertotti et al. 2009) and numerically (Lee et al. 2004).

The LLG equation is a nonlinear differential equation. The first term on the right-hand side of Eq. (2) is the field-torque term and describes a steady precession (oscillation) of the magnetization around the magnetic field. The second (STT) term moves the magnetization parallel or antiparallel to the direction of \mathbf{p} depending on the direction of the electric current. The third term (the damping term) describes the energy dissipation of the ferromagnet phenomenologically (Gilbert 2004). Auto-oscillation of magnetization \mathbf{m} in an STO is excited when the energy dissipation ($\propto \alpha$) is cancelled by the energy supplied by the work done by the STT (Bertotti et al. 2009). When this condition is satisfied, the auto-oscillation around the magnetic field, which is described by the field-torque term, is realized. The oscillation trajectory is almost on a constant energy curve (Bertotti et al. 2009).

2.2 Micro- and Macro-Structures of Nanomagnet

For the following discussions, we need to briefly review the microstructure of the free layer. As mentioned above and shown in Fig. 1a, the size of the free layer is typically 10–100 nm in diameter and 1 nm in thickness. Such a ferromagnet consists of numerous magnetic atoms, so its macroscopic magnetization consists of numerous magnetic moments. The direction of each magnetic moment in equilibrium is determined to minimize the magnetic energy E , which is related to the magnetic field \mathbf{H}

via $\mathbf{H} = -\delta E/\delta(M\mathbf{m})$ (Lifshitz and Pitaevskii 1980), and consists of the exchange energy, magnetic anisotropy energy, dipole field energy, and Zeeman energy (Hubert and Schäfer 1998).

When the free layer, in particular the thickness, is relatively large, the microstructure of each magnetic moment becomes a vortex so as to minimize the stray field energy, i.e., the magnetic moment at the center of the area points in the direction perpendicular to the cross-sectional area of the free layer whereas the magnetic moments around the center exhibit a chiral structure, as illustrated in Fig. 1b. In this case, \mathbf{m} is a function of location \mathbf{x} inside the free layer. Another micromagnetic structure called domain wall separates two uniformly magnetized domains. One way to investigate the magnetization dynamics in such systems is to use the LLG equations to compute the magnetization dynamics of all the magnetic moments simultaneously, i.e., by using numerical micromagnetic simulation (Brown and LaBonte 1965; Hayashi and Goto 1971; Nakatani et al. 1989, 2003; Dussaux et al. 2010). Another way to investigate the dynamics is to reduce the LLG equation to an equation of motion of the vortex core. The reduced equation is called Thiele equation, which describes the motions of the radius and rotation angle in the film plane of the vortex core (Thiele 1973; Guslienko 2006; Liu et al. 2007; Khvalkovskiy et al. 2009; Dussaux et al. 2012; Grimaldi et al. 2014).

On the other hand, when the free layer is relatively small, almost all magnetic moments point in the same direction so as to minimize the exchange and anisotropy energies. In this case, we can replace local magnetization \mathbf{m} in Eq. (2) with a macroscopic single moment as illustrated in Fig. 1c. This model is called a macromagnetic or macrospin model. The macromagnetic magnetization in the thin magnetic film used in spintronics devices usually points in the direction parallel to the film plane. This is because this in-plane magnetized configuration minimizes the stray field generated outside the ferromagnet and reduces the magnetic energy. The magnetization can be, however, oriented in the direction perpendicular to the film plane by controlling the perpendicular anisotropy energy. For example, adding an MgO barrier neighboring a CoFe ferromagnet induces an interface anisotropy effect (Hine et al. 1979; Yakata et al. 2009; Ikeda et al. 2010; Kubota et al. 2012). Applying an electrical voltage also changes the perpendicular anisotropy due to interface effect (Weisheit et al. 2007; Chiba et al. 2008; Maruyama et al. 2009; Nozaki et al. 2010; Shiota et al. 2012). The perpendicularly magnetized MTJ is of great interest for MRAM applications because it is suitable for high-density structures.

The number of variables in Eq. (2) for the vortex structure is $2\mathcal{N}$, where \mathcal{N} is the number of magnetic moments. In the macromagnetic case, the number of variables is 2. One might consider that, since vector \mathbf{m} is defined in a three-dimensional space, the number of variables should be $3\mathcal{N}$ and 3 for the vortex and macromagnetic case. However, the LLG equation conserves the norm of magnetization \mathbf{m} because the equation satisfies $d|\mathbf{m}|^2/dt = (1/2)\mathbf{m} \cdot (d\mathbf{m}/dt) = 0$, so there is a constraint condition $|\mathbf{m}| = 1$, which reduces the number of independent variables. This conservation of the magnetization norm means that the temperature of the system is sufficiently lower than the Curie temperature of the ferromagnet. According to the Poincaré–Bendixson theorem (Wiggins 1990), chaos is precluded in a macromagnetic model

without interaction with other ferromagnets and/or feedback. Magnetization \mathbf{p} in the reference layer is usually assumed to be macromagnetic.

The ferromagnets used in spintronics devices typically have two energy minima due to uniaxial anisotropy energy. Such a ferromagnet can be used as a bit of memory devices. The strengths of the STT and damping torque represented in Eq. (2) depend on the direction of the magnetization. Related to this fact, there are two thresholds, J_c and J_{sw} , of current density J in Eq. (2). When $J/J_c > 1$, the STT exceeds the damping torque near the equilibrium state and destabilizes the magnetization. When $J/J_{sw} > 1$, the magnetization moves from one equilibrium to another. Note that auto-oscillation of the magnetization is excited in range $J/J_c > 1$ and $J/J_{sw} < 1$. To satisfy these conditions, J_{sw}/J_c should be larger than one. If this condition is not satisfied, i.e., $J_{sw}/J_c < 1$, the MTJ exhibits the switching without showing an auto-oscillation (Taniguchi et al. 2013; Taniguchi 2015). This means that not all MTJs exhibit an auto-oscillation. The magnitude relationship between J_c and J_{sw} depends on the magnetic field and STT angular dependence.

The oscillation frequencies of a vortex and macromagnetic STO are typically of the order of 100 MHz and 1–10 GHz, respectively. A vortex-type STO was used for the experimental estimation of short-term memory capacity (Sect. 3) whereas the macromagnetic model was used for the numerical simulation (Sect. 4).

2.3 Recurrent Neural Network Based on STO

We developed two models of an STO-based RNN. One uses a single vortex-type STO, while the other uses a single macromagnetic STO or nine macromagnetic STOs as a reservoir. In both models, the output voltage from the STO is used as the dynamic response from the reservoir. According to Eq. (1), the resistance, or voltage, of an STO oscillates due to the oscillation of magnetization \mathbf{m} in the free layer ($\cos \theta = \mathbf{m} \cdot \mathbf{p}$). The output voltage from the STO is given by

$$v_{\text{out}}(t) = v_a(t) \sin [2\pi f t + \varphi_{\text{out}}(t)], \quad (3)$$

where f is the oscillation frequency of the TMR. Amplitude v_a and phase φ_{out} of the voltage are constant in an auto-oscillation state. On the other hand, when an additional current (voltage) pulse is applied to the STO, they vary over time. This is because the additional torque changes the balance between the STT and damping torque, causing the magnetization to move a different oscillation state. The relaxation process for the new oscillation state depends on the sequence of additional voltage pulses. In other words, the STO in the relaxation process stores past input information. Therefore, the STO is a candidate for reservoir computing. The ability of a reservoir is quantitatively characterized by its memory and nonlinear capacities, as mentioned in Sect. 1.1. Therefore, it is necessary to evaluate the memory and nonlinear capacities of STO devices.

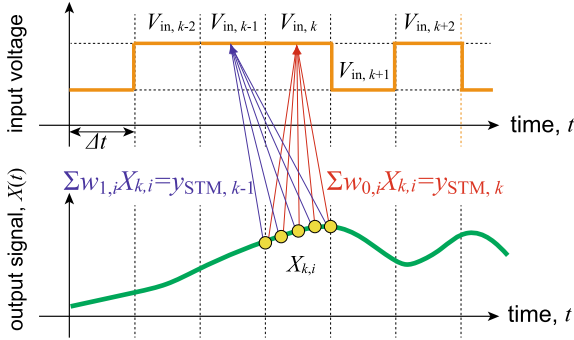


Fig. 2 Schematic illustration of input voltage V_{in} and output signal $X(t)$. The numbers such as k and $k + 1$ indicate the pulse numbers. The k th input voltage is $V_{in,k} = V_{low} + (V_{high} - V_{low})s_{in,k}$; the corresponding output signal at the i th node is $X_{k,i}$, where $s_{in,k}$ is input binary data. For evaluation of short-term memory capacity, training data $y_{STM,k}$ was directly related to input data via $y_{STM,k} = s_{in,k}$. Weight $w_{D,i}$ was determined so as to reproduce $y_{STM,k}$ ($V_{in,k-D}$) from $X_{k,i}$. Cases $D = 0$ and $D = 1$ are shown in figure. For evaluation of the parity check capacity, training data $y_{PC,k}$ was related to input binary data (voltage) via Eq. (12) (After Tsunegi et al. 2018a)

2.4 Methods for Evaluating Memory Capacities

Here, we describe methods for evaluating the short-term memory and parity check capacities.

We first explain, with the help of Fig. 2, how we evaluate short-term memory capacity. The figure schematically shows how the input is applied to the STO and how the output is obtained. As mentioned, memory capacity is evaluated by examining the response to binary input. We first sequentially apply Z -bit random voltage pulses (Z : integer) to the STO. The input voltage is expressed as $V_{in,k} = V_{low} + (V_{high} - V_{low})s_{in,k}$, where V_{low} and V_{high} are the lower and higher values of the input voltages, respectively, whereas $s_{in,k} = 0$ or 1 is the random binary data. The memory capacity is evaluated by examining the response of the reservoir to binary input (Jaeger 2002), and therefore masking procedure (Appeltant et al. 2011) is not applied to the input data in this work. The total length of the voltage pulses is $Z\Delta t$, where Δt corresponds to the duration of 1 bit. These input data are called training data. The k th ($k = 1, 2, \dots, Z$) pulse is applied from time $t = (k - 1)\Delta t$ to $t = k\Delta t$.

The STT excited by the input voltage changes the oscillation trajectory of the magnetization, resulting in a change in the voltage output from the STO, as schematically shown in Fig. 2. Not only the output voltage but also the STO variables can be used for the evaluation. For example, in Sect. 3, voltage amplitude v_a is used as the output signal, whereas resistance x is used in the simulation described in Sect. 4. In this section, we denote a general signal output from an STO as $X(t)$. We divide the output signal from $t = (k - 1)\Delta t$ to $t = k\Delta t$ into N -nodes, where N is the number of virtual nodes. These virtual nodes are used, along with linear regression, to obtain the output from the reservoir. For simplicity, we denote the output signal at the i th

($i = 1, 2, \dots, N$) node as $X_{k,i}$, i.e.,

$$X_{k,i} = X \left((k-1)\Delta t + i \frac{\Delta t}{n} \right). \quad (4)$$

Weight $w_{0,i}$ is then introduced to obtain the output from the reservoir;

$$\sum_{i=1}^{N+1} w_{0,i} X_{k,i}. \quad (5)$$

The learning task here is to find the weight that minimizes the error between the output from the reservoir and the training data, which is defined as

$$\sum_{k=1}^Z \left(\sum_{i=1}^{N+1} w_{0,i} X_{k,i} - y_{\text{STM},k} \right)^2, \quad (6)$$

where $y_{\text{STM},k}$ is the training data used for evaluating short-term memory capacity, and is related, specifically for the case without any delay, to the input binary data $s_{\text{in},k}$;

$$y_{\text{STM},k} = s_{\text{in},k}. \quad (7)$$

The meaning of the suffix “0” in $w_{0,i}$ in Eq. (6) will be explained below. The term with subscript $i = N + 1$ corresponds to the bias term used to tune the constant value, and $X_{k,N+1} = 1$ in Eq. (6).

Weight $w_{0,i}$ introduced above is determined to reproduce training data $y_{\text{STM},k}$ applied during time $(k-1)\Delta t \leq t \leq k\Delta t$. Since an RNN is an artificial neural network used for classifying and calculating a time sequence of input data, it should be able to reproduce the past input voltage from the current output signal. To enable this functionality, Eq. (6) is extended:

$$\sum_{k=1}^Z \left(\sum_{i=1}^{N+1} w_{D,i} X_{k,i} - y_{\text{STM},k-D} \right)^2, \quad (8)$$

where $D = 1, 2, 3, \dots$ is the delay. Weight $w_{D,i}$ is set so as to reproduce the $(k-D)$ th input binary data $y_{\text{STM},k-D}$ applied during time $(k-D-1)\Delta t \leq t \leq (k-D)\Delta t$ from output signal $X(t)$ obtained during time $(k-1)\Delta t \leq t \leq k\Delta t$.

After this learning, other Z' -random pulse sequences $s'_{\text{in},n}$ are applied to the STO as voltages, where Z' is the number of input data instances. The prime symbol is used to distinguish the quantities related to testing from those related to learning. Test data $y'_{\text{STM},n}$ is defined similarly to Eq. (7), i.e., $y'_{\text{STM},n} = s'_{\text{in},n}$. The output signal at the i th node responsive to the n th test data is denoted as $X'_{n,i}$. Using the weight determined by the learning, we define the reconstructed data as

$$y'_{R,n-D} = \sum_{i=1}^{N+1} w_{D,i} X'_{n,i}, \quad (9)$$

where $X_{n,N+1} = 1$. Whether the reconstructed data reproduces the test data with delay D is characterized by correlation coefficient $\text{Cor}(D)$ given by

$$\text{Cor}(D) = \frac{\sum_{n=1}^{Z'} (y'_{\text{STM},n-D} - \langle y'_{\text{STM},n-D} \rangle) (y'_{R,n-D} - \langle y'_{R,n-D} \rangle)}{\sqrt{\sum_{n=1}^{Z'} (y'_{\text{STM},n-D} - \langle y'_{\text{STM},n-D} \rangle)^2 \sum_{n=1}^{Z'} (y'_{R,n-D} - \langle y'_{R,n-D} \rangle)^2}}, \quad (10)$$

where $\langle \dots \rangle$ is the average value. The short-term memory capacity is estimated as

$$C_{\text{STM}} = \sum_{D=1}^{Z'} [\text{Cor}(D)]^2. \quad (11)$$

The short-term memory capacity characterizes the ability of the reservoir to store past input as is, which can be done by simply using a linear transformation. The nonlinear transformation ability for the stored information can be examined by using a parity check task, where the following task data are used for the learning and testing:

$$y_{\text{PC},k-D} = \sum_{i=0}^D s_{\text{in},k-D+i} \pmod{2}. \quad (12)$$

The ability of the nonlinear transformation is quantitatively evaluated by introducing parity check capacity C_{PC} , which is defined similarly to Eq. (11). In the evaluation of parity check capacity, the data input to the STO is again binary voltage data. However, when the weight is determined, $y_{\text{PC},k-D}$ given by Eq. (12) should be substituted into the right-hand side of Eqs. (6) and (8). Similarly, during testing, test data $y'_{\text{PC},n}$ are defined from the input binary data in accordance with Eq. (12) whereas the reconstructed data y'_R are obtained from the signal output from the STO and the weight determined by the learning, as shown in Eq. (9). The parity check capacity C_{PC} is evaluated from the correlation between the test and reconstructed data.

3 Reservoir Computing with STO (Experiment)

In this section, we report our experimental evaluation of the short-term memory capacity of a vortex-type STO.

3.1 Experimental Methods

The STO used in this study was made from an MTJ with a multilayer structure: sub/buffer/PtMn(15)/CoFe(2.5)/Ru(0.86)/CoFeB(1.6)/CoFe(0.8)/MgO(1.1)/CoFe(0.5)/FeB(7.0)/CoFe(0.5)/MgO(1.0)/Ta/Ru (thickness in nm). The films were patterned into an STO with a diameter of 425 nm. The CoFe/Ru/CoFeB/CoFe and CoFe/FeB/CoFe structures correspond to the free and reference layers, respectively. A magnetic vortex state, as schematically shown in Fig. 1b, appears in the free layer in this type of MTJ (Tsunegi et al. 2016a, 2018b).

Figure 3a schematically shows the circuit used for evaluating the short-term memory capacity of the STO. We first prepared random binary bits as input data. The data were converted into voltage pulses by using an arbitrary wave generator (AWG). Voltage pulses with duration Δt and a 5 ns lead edge were applied to the STO through a low-pass filter using a bias-tee with a cutoff frequency of 300 MHz. We used filters corresponding to a Gaussian filter and applied the inverse of the circuit transfer function to the input voltage to reduce the distortion in preprocessing. The waveform of the high-frequency output voltage was measured at room temperature using a real-time oscilloscope (5–10 Gsam/s) using a high-pass filter with a cutoff frequency of 400 MHz. The STO generated an oscillating voltage through auto-oscillation of the vortex core induced by the perpendicular component (film normal) of the STT (Dussaux et al. 2010). Therefore, we applied an out-of-plane magnetic field of 725 mT to the oscillator. The voltage output from the STO was separated from the input voltage by using the high-pass filter and measured with the oscilloscope. Amplitude v_a of the output voltage was estimated by using Hilbert transformation and used as the output signal X for evaluating short-term memory capacity.

Figure 3b, c shows examples of the input and output voltages for offset voltage $V_{\text{offset}} = 250$ mV and $\Delta t = 20$ ns. The input voltage was 200 or 300 mV, and the corresponding oscillation frequency of the STO was 540 or 555 MHz, respectively. The output voltage exhibited a temporal change in magnitude, in accordance with the change in the input pulse voltage. The amplitude is the envelope of the oscillating voltage, as shown by the red curve in Fig. 3c. It was used to estimate the short-term memory capacity.

In our experiments, we used 3500 random bits in total, where the first 500 bits were used for the washout, the following 2000 ($Z = 2000$) bits were used for learning, and the last 1000 ($Z' = 1000$) bits were used to estimate the short-term memory capacity. Hereafter, we call this sequence of 3500 bits a “single shot.”

Since the dynamics of the vortex core exhibited randomness due to thermal fluctuation, the initial conditions, as well as the amplitude noise in the output voltage, differed for every trial. Thus, the output voltage from the STO differed every run even when the same voltage pulses were used as input. Therefore, we repeated the single-shot experiment 60 times using fixed training and test data. The amplitude was averaged, with which the test data was also reconstructed. Using the average amplitude for both learning and testing should reduce the amplitude noise, leading to an increase in short-term memory capacity.

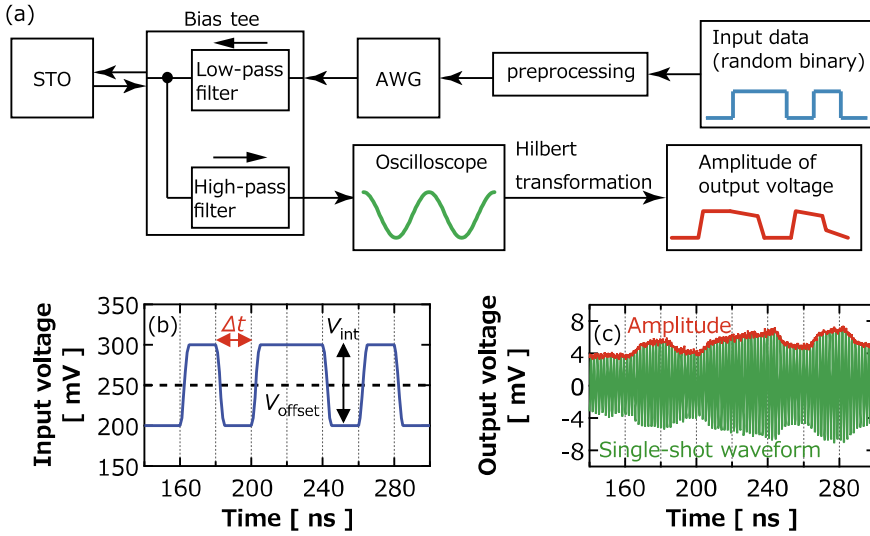


Fig. 3 **a** Measurement circuit used in experiments. Input data were preprocessed and injected into the STO through an arbitrary wave generator (AWG) and low-pass filter. Output voltage was measured with an oscilloscope after passing through the high-pass filter. Amplitude of output voltage was estimated using Hilbert transformation. Memory capacity was evaluated by comparing amplitude of output voltage with input data. **b** Example of sequential input-voltage pulses and **c** voltage output from STO. Intensity, offset, and pulse duration of input voltage were $V_{\text{int}} = 100$ mV, $V_{\text{offset}} = 250$ mV, and $\Delta t = 20$ ns, respectively (After Tsunegi et al. 2018a)

3.2 Short-Term Memory Capacity in Single STO

Figure 4a shows comparisons between the test (red solid line) and reconstructed (blue dotted line) data using 200 nodes in a single-shot experiment with delay $D = 1, 2,$ or 3. For $D = 1$, the reconstructed data reproduced the test data almost perfectly, indicating that the STO remembered the input voltage applied one bit before the present pulse. In contrast, the ability to reconstruct data was low for $D = 2$ and 3. Using the averaging technique improved the reconstruction data for $D = 2$, as can be seen in Fig. 4b. These reconstructed data and Eq. (10) were used to evaluate the correlation coefficients.

Figure 5a illustrates the relationship between delay D and correlation $\text{Cor}(D)$ for the single-shot (red solid line) and averaged data (blue dotted line). The short-term memory capacities estimated from these data and Eq. (11) were nearly 1.0 for the single-shot experiment and 1.8 for the “averaged 60 times” experiment. These results indicate that reducing the amplitude noise of the output voltage is important for increasing short-term memory capacity.

We performed the same experiment but with values of Δt other than 20 ns. The pulse duration dependency of the short-term memory capacity obtained from the averaged results is plotted in Fig. 5b by circles. The short-term memory capacity

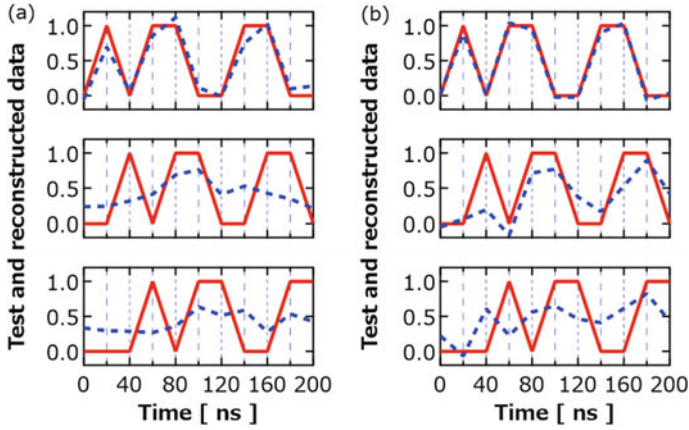


Fig. 4 Comparisons between test (red solid line) and reconstructed (blue dotted line) data for **a** single-shot and **b** averaged data. Delay D was 1, 2, and 3 from top to bottom (After Tsunegi et al. 2018a)

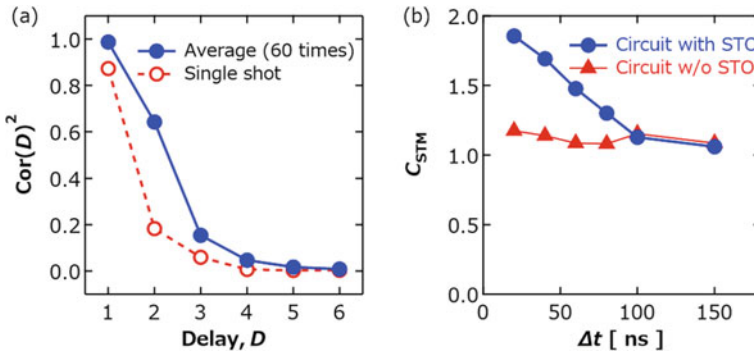


Fig. 5 **a** Square of correlation $\text{Cor}(D)^2$ between the test and reconstructed data for single-shot (red dotted) and averaged (blue solid) data as a function of delay D . **b** Dependence of memory capacity in circuit without (red triangles) and with (blue circles) STO on pulse duration Δt (After Tsunegi et al. 2018a)

decreased monotonically with an increase in the pulse duration for the following reason. When a voltage pulse is applied to the STO, the vortex core starts to relax to an oscillation orbit (limit cycle) determined by the voltage. Now let us assume that a voltage pulse is applied at $t = t_0$ and that the next pulse is applied at $t = t_0 + \Delta t$. If pulse width Δt is sufficiently greater than the relaxation time of the vortex core, the output voltage for $t > t_0 + \Delta t$ is not correlated to that for $t < t_0$. Therefore, the short-term memory capacity decreases with an increase in the pulse width. The pulse width should thus be less than the relaxation time of the STO for reservoir computing. The relaxation time of the STO strongly depends on the type of oscillator. Roughly speaking, the relaxation time is of the order of $1/(\alpha f)$, where α and f are the

damping constant and oscillation frequency, respectively (Taniguchi et al. 2017). The relaxation time of a vortex-type STO is of the order of 10–100 ns (Jenkins et al. 2016; Tsunegi et al. 2018b) whereas that of a macromagnetic STO is of the order of 1–10 ns (Kudo et al. 2010; Suto et al. 2011; Nagasawa et al. 2011). Evaluation of STO relaxation time will be a key issue for designing reservoir computing based on spintronics technology.

3.3 Contributions from Other Circuit Components

The definition used for short-term memory capacity in this work needs to be mentioned. The memory function of reservoir computing arises from the nonlinear response of the system to the input data. In the present system, not only the STO but also the other components in the circuit, such as the AWG and low-pass filter (Fig. 3a), exhibit nonlinear responses to input data. Therefore, the short-term memory capacity evaluated above should be, strictly speaking, regarded as that of the whole system, and includes the contributions from the other components of the circuit.

To validate the existence of the memory function in the STO, we also evaluated the short-term memory capacity of the system without the oscillator. The results are plotted in Fig. 5b by triangles. The short-term memory capacity of the circuit with the STO exceeded that of the circuit without the STO, proving that the STO has memory functionality.

3.4 Future Directions

Although the finite value of memory capacity found in this work supports the potential application of spintronics devices to artificial neural networks, the short-term memory capacity is still low compared with that of other systems such as a quantum reservoir with several qubits and virtual nodes (Fujii and Nakajima 2017). The low memory capacity of the present system is due to the trade-off between noise reduction of noise and capacity enhancement. As explained above, the memory of the past input voltage is stored in the amplitude of the output voltage. Noise in the amplitude of the output voltage reduces memory capacity. Therefore, the changes in the output voltage amplitude between neighboring nodes should be large enough to be distinguished from the noise. However, a large change in the output voltage amplitude between the neighboring nodes means fast relaxation, leading to a reduction in memory capacity. Therefore, solutions that further enhance memory capacity are needed.

A potential solution is to use an STO with delayed feedback (Khalsa et al. 2015; Tsunegi et al. 2016b), where the information of past input is naturally stored in the delayed feedback loop, resulting in an increase in short-term memory capacity (Riou et al. 2019; Yamaguchi et al. 2020). Another is to use frequency and/or phase locking of the STO by using forced or mutual synchronization (Marković et al. 2019; Tsunegi

et al. 2019). This is because locking by synchronization leads to a reduction in the amplitude and/or phase noise and thus stabilizes the output voltage (Kaka et al. 2005; Mancoff et al. 2005; Zhou and Akerman 2009; Locatelli et al. 2015; Urazhdin et al. 2016; Awad et al. 2017; Tsunegi et al. 2018c). Investigating such solution is a future research direction in reservoir computing using spintronics technology.

4 Reservoir Computing with STO (Simulation)

In this section, we discuss the quantitative analysis of the figures-of-merit for reservoir computing using MTJ devices by employing macromagnetic simulation. Specifically, not only a reservoir computing system with single MTJ but also one with multiple MTJs is discussed (Furuta et al. 2018).

4.1 Simulation Methods

Figure 6a, b shows schematics of the system used for simulating reservoir computing using MTJs. As mentioned above, an MTJ contains an insulating tunneling barrier layer with two ferromagnetic layers. For ferromagnetic layer-1 (the reference layer), the magnetization direction is designed to be fixed. This can be done by the exchange bias effect using antiferromagnetic materials, such as PtMn and IrMn (Meiklejohn and Bean 1956), or magnetic anisotropy energy. Here, the magnetization direction of layer-1 was fixed perpendicular to the film plane. For ferromagnetic layer-2 (the free layer), the magnetization direction was not fixed and thus could be controlled by the current (Katine et al. 2000) or voltage (Shiota et al. 2012). The MTJ device resistance reflects the magnetization direction.

The magnetization dynamics in ferromagnetic layer-2 follows the LLG equation with STT given by Eq. (2), without thermal fluctuation in the ferromagnetic layers (Brown 1963). Here, \mathbf{p} and \mathbf{m} in Eq. (2) correspond to the unit magnetization vectors for ferromagnetic layers 1 and 2 in Fig. 6(a) and 6(b), respectively. The effective magnetic field in the present study is given by

$$\mathbf{H} = -H_{azz}m_z\mathbf{e}_z, \quad (13)$$

where H_{azz} is a uniaxial perpendicular anisotropy field. The z -component of \mathbf{m} is denoted as m_z , and \mathbf{e}_z is the unit vector parallel to the z -axis. In our definition, $H_{azz} > 0$ ($H_{azz} < 0$) shows in-plane (perpendicular) magnetic anisotropy. The angular dependence of the STT in Eq. (2) is $g(\theta) = 1/(1 + P^2 \cos \theta)$ (Slonczewski 2005). The resistance of the MTJ varies with the relative angle between the spins in the free and pinned layers, $R = 1/G$ [see also Eq. (14)]

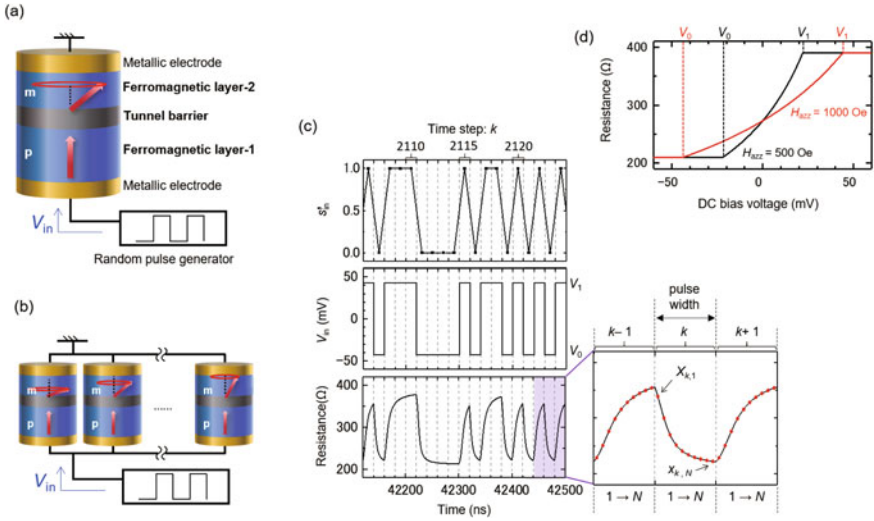


Fig. 6 **a** Schematic of reservoir system using magnetization dynamics in an MTJ. **b** System with multiple MTJs. Magnetization direction of ferromagnetic layer-2 (s_2) could be controlled by input bias voltage V_{in} ; Magnetization direction of ferromagnetic layer-1 (p) is fixed. **c** Input s_{in} , bias voltage V_{in} input to MTJ device, and MTJ device resistance as function of time, which are typical characteristics during learning and evaluation processes. Virtual nodes [$x_{k,1}$ to $x_{k,N}$ during k th binary input] were defined as shown in figure. **d** MTJ device resistance as function of static input DC bias voltage. Black and red plots indicate resistance, when uniaxial anisotropy fields were 500 Oe and 1000 Oe, respectively. V_0 and V_1 are voltages that rendered device resistance constant (After Furuta et al. 2018)

$$R = \frac{2R_{AP}R_P}{(R_{AP} + R_P) + (R_{AP} - R_P)(\mathbf{m} \cdot \mathbf{p})}. \tag{14}$$

The time evolution of the MTJ resistance is characterized by sequential calculation using the fourth Runge–Kutta method. For evaluating the short-term memory and parity check capacities, an input pulse voltage, V_{in} , corresponding to the computational input, $s_{in}(= 0 \text{ or } 1)$, was applied to the MTJs, as depicted in Fig. 6c. Figure 6(a) and 6(b) shows the schematics of circuits with single and multiple MTJs, respectively. In this section, the physical parameters are $H_{azz} = 1000 \text{ Oe}$, $\alpha = 0.009$, $R_P = 210 \Omega$, and $R_{AP} = 390 \Omega$. These values almost follow our previous experimental research (Miwa et al. 2014).

Figure 6c shows an example of MTJ device resistance under input voltage V_{in} . We used a pulse voltage with binary values of $V_0(= -44 \text{ mV})$ and $V_1(= +44 \text{ mV})$ as V_{in} . These binary values correspond to 0 and 1 in s_{in} , respectively, in the reservoir computing learning and evaluation processes. The pulse width (20 ns in Fig. 6c, for instance) corresponds to the discrete unit time step. Because the device resistance is scalar, the node dimension is simply one. However, the number of nodes can be increased by using virtual nodes (Appeltant et al. 2011; Nakajima et al. 2018). As

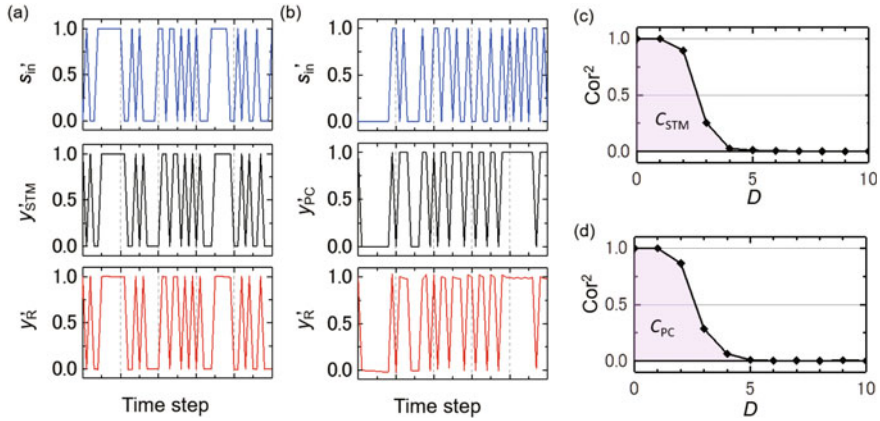


Fig. 7 **a** Input s_{in}' , test data y_{STM}' [Eq. (8) with $D = 1$], and reconstructed data y_R' used for evaluating the short-term memory task. **b** Input s_{in}' , test data y_{PC}' , (Eq. (9) with $D = 1$), and reconstructed data y_R' used for evaluating parity check task. **c** Correlation using Eq. (10); integrated values are defined as short-term memory capacity (C_{STM}). **d** Correlation using Eq. (10); integrated values are defined as the parity check capacity (C_{PC}). Input-voltage pulse width was 20 ns, and number of virtual nodes was 50 (After Furuta et al. 2018)

shown in Fig. 6c, the virtual nodes $x_{k,1}$ to $x_{k,N}$ during the k th binary input are defined, where output signal $x_{k,i}$ in the figure is the resistance with suffix $i = 1, 2, \dots, N$ corresponding to the node number. These virtual nodes are further defined as a node vector \mathbf{x}_k .

Figure 6d depicts the DC bias voltage dependence of the static MTJ device resistance. Under DC bias voltage conditions, the resistance was measured after the magnetization dynamics were damped. Under positive bias voltage conditions, the spin-polarized current flowed from the free layer to the pinned layer, and the STT induced auto-oscillation in \mathbf{m} . The relative magnetization angle between \mathbf{p} and \mathbf{m} increased, and an antiparallel-like magnetization configuration was realized. This means that the device resistance increases when a positive bias voltage is applied. Under negative bias voltage conditions, a parallel-like magnetization configuration was induced, and the device resistance decreases. For the input pulse voltage depicted in Fig. 6c, the binary values V_0 and V_1 are defined as voltages that rendered device resistance constant. As shown in Fig. 6d, V_0 and V_1 depended on uniaxial magnetic anisotropy H_{azz} .

4.2 Short-Term Memory and Parity Check Capacities in Single STO

In this section, we present figures-of-merit for reservoir computing using a single MTJ device. The uniaxial magnetic anisotropy of the free layer \mathbf{m} was fixed: $H_{azz} =$

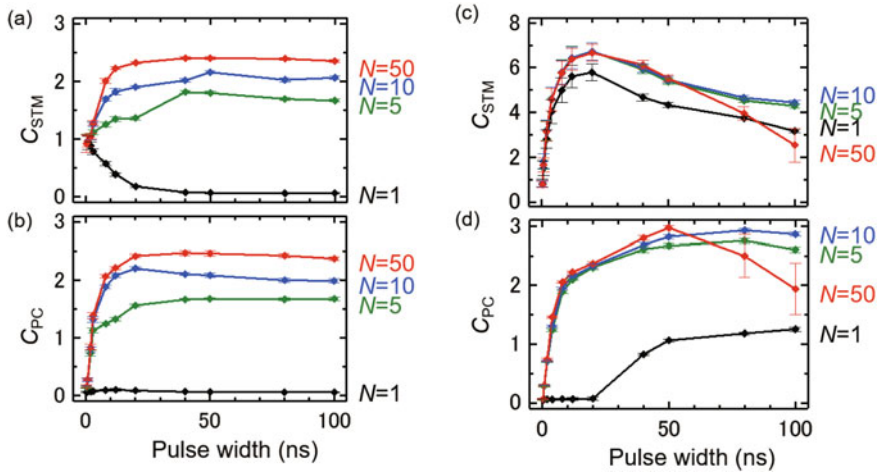


Fig. 8 **a, b** Results of reservoir computing using single MTJ for short-term memory capacity (C_{STM}), and parity check capacity (C_{PC}) as a function of input-voltage pulse width; N is number of virtual nodes in MTJ. **c, d** Results of reservoir computing using multiple MTJs for C_{STM} and C_{PC} with seven MTJs ($M = 7$) as function of input-voltage pulse width. Uniaxial magnetic anisotropy ratio ($H_{azz,k+1}/H_{azz,k}$) was 1.6 (After Furuta et al. 2018)

1000 Oe. Recall that a positive value of H_{azz} means that the magnetic cell in the MTJ is in-plane magnetized. Figure 7 shows the simulated data used for evaluating the short-term memory and parity check capacities for a single MTJ. The input-voltage pulse width was 20 ns and the number of virtual nodes, N , was 50. Figure 7a shows typical simulation results for input s'_{in} , test data for short-term memory task y'_{STM} , and reconstructed data y'_R as a function of time. Similarly, Fig. 7b shows input s'_{in} , test data for parity check task y'_{PC} , and trained output y'_R . Delay D in both figure was 1. The training and test data for the short-term memory task and parity check task are defined by Eqs. (7) and (12), respectively. The output was calculated using the simulated MTJ resistance (see Fig. 6c) and the weight. The weight was trivially calculated using the definitions given by Eq. (8). Figure 7c, d depicts the correlations [Eq. (10)] between the test and reconstructed data as a function of D . C_{STM} and C_{PC} are defined as the numerical integration of the correlation [see Eq. (11)] and as the capacity using training data for the short-term memory and parity check, respectively.

Figure 8(a) and 8(b) shows the short-term memory C_{STM} and parity check C_{PC} capacities, respectively, as functions of the input-voltage pulse width with virtual nodes N . Both C_{STM} and C_{PC} increased with the pulse width until it reached ~ 20 ns. They then remained nearly constant. When the pulse width is less than 20 ns, the change in the magnetization direction is very small, so the magnetization dynamics cannot work as a reservoir.

4.3 Short-Term Memory and Parity Check Capacities in Multiple STOs

When multiple MTJs are used for reservoir computing, higher figures-of-merit can be obtained. A schematic of a multiple-MTJ circuit for reservoir computing is depicted in Fig. 6b. MTJs are placed in parallel, and an identical pulse voltage is applied to all of them. Spatial multiplexing (Nakajima et al. 2019) is used to construct the nodes for reservoir computing. The node vector has $M \times N$ elements, where M is the number of MTJs and N is the number of virtual nodes in an MTJ:

$$\mathbf{x}_{L,k} = \begin{pmatrix} x_{L,k,1} \\ \vdots \\ x_{L,k,N} \end{pmatrix}, \quad (L = 1, 2, \dots, M), \quad (15)$$

$$\mathbf{x}_k \equiv (x_{1,k,1} \cdots x_{1,k,N} \ x_{2,k,1} \cdots x_{2,k,N} \cdots x_{M,k,1} \cdots x_{M,k,N})^t, \quad (16)$$

where “t” indicates a transposed vector, and $x_{L,k,i}$ is the signal output from the L th MTJ ($L = 1, 2, \dots, M$) at the i th node ($i = 1, 2, \dots, N$) in response to the k th binary data input ($k = 1, 2, \dots, Z$).

We tested various sets of uniaxial anisotropy H_{azz} of ferromagnetic layer-2 in each MTJ (see Table 1). For instance, when four MTJs were used and $H_{\text{azz},k}/H_{\text{azz},k+1} = 2$, the uniaxial anisotropies of the MTJs were 1000 Oe, 500 Oe, 250 Oe, and 125 Oe. Such variations in anisotropy can be obtained by voltage-controlled magnetic anisotropy in the MTJs (Maruyama et al. 2009). In this study, thermal fluctuation in ferromagnetic layer-2 was neglected. For instance, thermal fluctuation energy at room temperature (26 meV) is negligibly small compared to magnetization energy $M_s H_{\text{azz}} \mathcal{V}/2$ (~ 10 eV) when $H_{\text{azz}} = 1000$ Oe. Here, M_s and \mathcal{V} are the saturation magnetization and volume of the free layer and were assumed to be 1375 emu/c.c. and 23500 nm³, respectively (Miwa et al. 2014). Therefore, thermal fluctuation is comparable to or less than the magnetization energy of ferromagnetic layer-2 for $H_{\text{azz}} < 3$ Oe. In this region, simulation assuming the ground state is not very accurate, so a random magnetic field to reproduce the thermal fluctuation (Brown 1963) should be included in the simulation. Similar to the procedure for a single MTJ, the binary values V_0 and V_1 , the voltage input to the MTJs, were determined as shown in Fig. 6d. Note that V_0 and V_1 vary as a function of the uniaxial anisotropy field, and the smallest absolute values of the saturation voltages are used for as V_0 and V_1 for reservoir computing with multiple MTJs, i.e., V_0 and V_1 are determined for the MTJ with the smallest uniaxial magnetic anisotropy field.

We characterized C_{STM} and C_{PC} as functions of anisotropy ratio $H_{\text{azz},k}/H_{\text{azz},k+1}$. In the simulation, the input-voltage pulse width was 20 ns, and the number of virtual nodes for each MTJ was 50 for all MTJs. The maximum value of C_{STM} increased with the number of MTJs (M). Because each MTJ had a different uniaxial magnetic

Table 1 Sets of uniaxial magnetic anisotropy used for reservoir computing with multiple MTJs

$H_{\text{azz}}/H_{\text{azz},k+1}$	$H_{\text{azz},1}$	$H_{\text{azz},2}$	$H_{\text{azz},3}$...	$H_{\text{azz},7}$
1.0	1000 Oe	1000 Oe	1000 Oe	...	1000 Oe
1.1	1000 Oe	909.1 Oe	826.4 Oe	...	564.5 Oe
1.2	1000 Oe	833.3 Oe	694.4 Oe	...	334.9 Oe
...
2.9	1000 Oe	344.8 Oe	118.9 Oe	...	1.7 Oe
3.0	1000 Oe	333.3 Oe	111.1 Oe	...	1.4 Oe

anisotropy field H_{azz} , it had a different response speed to an external voltage/current. This variation in the response speed increased the C_{STM} of the system. In contrast, the increase in C_{PC} was insignificant compared to that in C_{STM} because there was no electric and/or magnetic interaction between the free layers of the MTJs. For instance, we found that $H_{\text{azz},k}/H_{\text{azz},k+1} = 1.6$ was the best condition for maximizing C_{STM} for $M = 7$. As shown in Fig. 8c, the C_{STM} was maximum around a pulse width of 20 ns. When the pulse width was smaller, the change in the magnetization direction due to the STT is too small for performing as a reservoir. When the pulse width was greater than 20 ns, the magnetization dynamics was almost completely damped during each unit time step, and such a condition is not preferable for reservoir computing. As shown in Fig. 8d, the best conditions are not the same for C_{STM} and C_{PC} . This is because a relatively long pulse is required to induce nonlinearity in the magnetization dynamics, in multiple MTJs.

4.4 Comparison with Echo-State Network

We used an echo-state network for comparison (Jaeger and Haas 2004; Jaeger 2001). An echo-state network is also a kind of RNN, in which the input-to-reservoir weight \mathbf{W}_{in} and internal weights \mathbf{W} between nodes are random, whereas the reservoir-to-output weight is optimized. The node vector at time step k of the echo-state network is determined in terms of the node vector at the previous step and the input at time step k . The following function is used as a node vector of the echo-state network,

$$\mathbf{x}_{\text{ESN},k} = \tanh(\mathbf{W}\mathbf{x}_{k-1} + \mathbf{W}_{\text{in}}s_{\text{in},k}). \quad (17)$$

The hyperbolic tangent function is used for component-wise projection. Weights \mathbf{W} and \mathbf{W}_{in} are given in a matrix and vector, respectively, in which the components are time-independent random values from -1 to $+1$. We normalized \mathbf{W} by dividing each component of \mathbf{W} by the spectral radius, which is the largest absolute value of the eigenvalues (singular value) of the weight matrix (Verstraeten et al. 2007). Weight \mathbf{W}_{in} is also normalized by its spectral radius. The C_{STM} and C_{PC} obtained

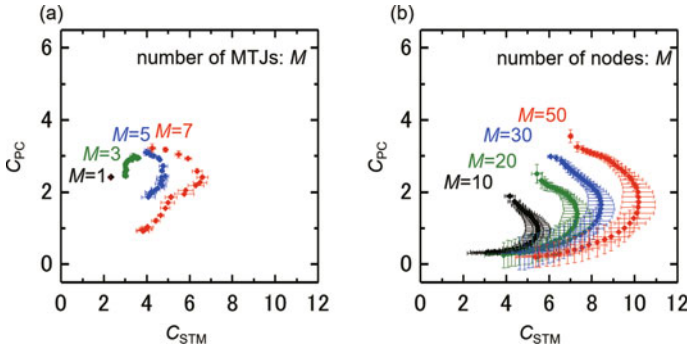


Fig. 9 Plots showing C_{STM} versus C_{PC} in **a** MTJ system and **b** echo-state network. In MTJ system, pulse width and number of virtual nodes (N) of each MTJ were fixed to 20 ns and 50, respectively (After Furuta et al. 2018)

using a multiple MTJ system are plotted in Fig. 9a; the pulse width was 20 ns and the number of virtual node was 50 for each MTJ. The data points from top to bottom correspond to increase in $H_{az,z,k}/H_{az,z,k+1}$ from 1.1 to 3.0. The C_{STM} and C_{PC} obtained using the echo-state network are shown in Fig. 9b. The data points from top to bottom correspond to increase the spectrum radius of the weight from 0.05 to 2.0. The results shown in Fig. 9 indicate that high-performance reservoir computing, similar to that of an echo-state network with 20–30 nodes, can be obtained for reservoir computing using 5–7 MTJs. In terms of the total number of virtual nodes in the system ($M \times N$), 35 nodes (7×5) in an MTJ system are comparable to 20–30 nodes in an echo-state network. Although C_{PC} increased slightly with M , we can obtain a large C_{PC} if there are magnetic and/or electrical interactions between the free layers in each MTJ.

5 Conclusion

In this chapter, we have described advances in brain-inspired computing devices based on spintronics technology. We have reviewed recent experimental and numerical work and reported our efforts in investigating the applicability of spintronics auto-oscillators (spin-torque oscillators, STOs) to recurrent neural networks. Spintronics devices have several attractive features for brain-inspired computing, such as low power consumption, applicability to high-density structures, a large output signal, and highly nonlinear and fast magnetization dynamics. Therefore, spintronics technology is promising for further development of artificial neural networks. However, the basic properties of spintronics devices, from the viewpoint of brain-inspired computing, have not been fully revealed yet. To prove the applicability of spintronics technology to computing, we need a deep understanding of magnetization dynamics in nanomagnets.

We have experimentally evaluated the short-term memory capacity of a reservoir computing with an STO. We have demonstrated the feasibility of learning and testing from output voltage by applying a time sequence of random voltage pulses to the STO. The short-term memory capacity obtained from the average output voltage was maximized to 1.8 under optimum conditions. Although this value includes the contribution not only from the STO but also from the other circuit components, the oscillator was shown to have a finite memory functionality. This indicates that it is important to reduce the amplitude noise in an STO applied to reservoir computing. The short-term memory capacity was increased by using a short pulse duration of 20 ns, which is comparable to the relaxation time (10–100 ns) of the oscillator. This indicates that it is also important to set the duration of the input pulses appropriately.

Using macromagnetic simulation, we demonstrated reservoir computing using the magnetization dynamics in MTJs. With reservoir computing using 5–7 MTJs, we can obtain performance similar to that of an echo-state network using hyperbolic tangent function with 20–30 nodes. Higher performance can be obtained by enabling magnetic and/or electrical interactions between the free layers in each MTJ.

Acknowledgements The authors are thankful to Julie Grollier at CNRS/Thales for valuable discussions. S. T., T. T., and H. K. are grateful to Kay Yakushiji, Akio Fukushima, and Shinji Yuasa at the National Institute of Advanced Industrial Science and Technology (AIST) for their supports. S. M., K. F., S. T., and H. K. are grateful to Taishi Furuta, Eiti Tamura, Minori Goto, Shota Hasebe, Koki Shimose, and Yoshishige Suzuki at the Osaka University for their discussions. T. T. is grateful to Takahide Kubota, Atsushi Sugihara, Satoshi Iba, Aurelie Spiesser, Hiroki Maehara, and Ai Emura for their support and encouragement. K. N., S. T., and T. T. are supported from a project subsidized by the New Energy and Industrial Technology Development Organization (NEDO). S. M. is supported by JSPS KAKENHI (No. JP18H03880, JP26103002). K. F. is supported by JSPS KAKENHI No. 16H02211, JST ERATO Grant No. JPMJER1601, JST CREST Grant No. JPMJCR1673, and JST PRESTO Grants No. JPMJPR15E7 and No. JPMJPR1668. K. N. is supported by JST PRESTO Grant Number JPMJPR15E7, Japan, and JSPS KAKENHI Grant Numbers JP18H05472, JP16KT0019, and JP15K16076.

References

- L. Appeltant, M.C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C.R. Mirasso, I. Fischer, *Nat. Commun.* **2**, 468 (2011)
- H. Arai, H. Imamura, *Phys. Rev. Appl.* **10**, 024040 (2018)
- A.A. Awad, P. Dürrenfeld, A. Houshang, M. Dvornik, E. Iacocca, R.K. Dumas, J. Akerman, *Nat. Phys.* **13**, 292 (2017)
- M.N. Baibich, J.M. Broto, A. Fert, F. Nguyen van Dan, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, J. Chazelas, *Phys. Rev. Lett.* **61**, 2472 (1988)
- J. Bardeen, L.N. Cooper, J.R. Schrieffer, *Phys. Rev.* **108**, 1175 (1957)
- J. Bass, W.P. Pratt Jr., *J. Phys.: Condens. Matter* **19**, 183201 (2007)
- L. Berger, *Phys. Rev. B* **54**, 9353 (1996)
- G. Bertotti, C. Serpico, I.D. Mayergoyz, A. Magni, M. d'Aquino, R. Bonin, *Phys. Rev. Lett.* **94**, 127206 (2005)
- G. Bertotti, C. Serpico, I.D. Mayergoyz, R. Bonin, M. d'Aquino, *J. Magn. Magn. Mater.* **316**, 285 (2007)

- G. Bertotti, I. Mayergoyz, C. Serpico, *Nonlinear Magnetization Dynamics in Nanosystems* (Elsevier, Oxford, 2009)
- G. Binasch, P. Grünberg, F. Saurenbach, W. Zinn, *Phys. Rev. Lett.* **63**, 664 (1989)
- W.A. Borders, H. Akima, S. Fukami, S. Moriya, S. Kurihara, Y. Horio, S. Sato, H. Ohno, *Appl. Phys. Express* **10**, 013007 (2017)
- W.F. Brown Jr., *Phys. Rev.* **130**, 1677 (1963)
- W.F. Brown Jr., A.E. LaBonte, *J. Appl. Phys.* **36**, 1380 (1965)
- D. Brunner, M.C. Soriano, C.R. Mirasso, I. Fischer, *Nat. Commun.* **4**, 1364 (2013)
- S. Chandrasekhar, *Rev. Mod. Phys.* **56**, 137 (1984)
- M.-C. Chen, A. Sengupta, K. Roy, *IEEE Trans. Magn.* **54**, 1500270 (2018)
- D. Chiba, M. Sawicki, Y. Nishitani, Y. Nakatani, F. Matsukura, H. Ohno, *Nature* **455**, 515 (2008)
- B. Dieny, R.B. Goldfarb, K.-J. Lee (eds.), *Introduction to Magnetic Random-Access Memory* (Wiley-IEEE Press, Hoboken, 2016)
- P.A.M. Dirac, *Proc. Roy. Soc. Lond. A* **114**, 243 (1927)
- P.A.M. Dirac, *Proc. Roy. Soc. Lond. A* **117**, 610 (1928)
- D.D. Djayaprawira, K. Tsunekawa, M. Nagai, H. Maehara, S. Yamagata, N. Watanabe, S. Yuasa, Y. Suzuki, K. Ando, *Appl. Phys. Lett.* **86**, 092502 (2005)
- A. Dussaux, B. Georges, J. Grollier, V. Cros, A.V. Khvalkovskiy, A. Fukushima, M. Konoto, H. Kubota, K. Yakushiji, S. Yuasa, K.A. Zvezdin, K. Ando, A. Fert, *Nat. Commun.* **1**, 8 (2010)
- A. Dussaux, A.V. Khvalkovskiy, P. Bortolotti, J. Grollier, V. Cros, A. Fert, *Phys. Rev. B* **86**, 014402 (2012)
- M.I. Dyakonov, V.I. Perel, *Phys. Lett. A* **35**, 459 (1971)
- K. Fujii, K. Nakajima, *Phys. Rev. Appl.* **8**, 024030 (2017)
- A. Fukushima, T. Taniguchi, A. Sugihara, K. Yakushiji, H. Kubota, S. Yuasa, *AIP Adv.* **8**, 055925 (2018)
- T. Furuta, K. Fujii, K. Nakajima, S. Tsunegi, H. Kubota, Y. Suzuki, S. Miwa, *Phys. Rev. Appl.* **10**, 034063 (2018)
- W. Gerstner, W.M. Kistler, R. Naud, L. Paninski, *Neuronal Dynamics* (Cambridge University Press, Cambridge, 2014)
- T.L. Gilbert, *IEEE Trans. Magn.* **40**, 3443 (2004)
- I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, Cambridge, 2017)
- E. Grimaldi, A. Dussaux, P. Bortolotti, J. Grollier, G. Pillet, A. Fukushima, H. Kubota, K. Yakushiji, S. Yuasa, V. Cros, *Phys. Rev. B* **89**, 104404 (2014)
- J. Grollier, D. Querlioz, M.D. Stiles, *Proc. IEEE* **104**, 2024 (2016)
- J. Grollier, D. Querlioz, K.Y. Camsari, K. Everschor-Sitte, S. Fukami, M.D. Stiles, *Nat. Electron.* **3**, 360 (2020)
- K.Y. Guslienko, *Appl. Phys. Lett.* **89**, 022510 (2006)
- N. Hayashi, E. Goto, *Jpn. J. Appl. Phys.* **10**, 128 (1971)
- W.J. Heisenberg, *Z. Phys.* **49**, 619 (1928)
- S. Hine, T. Shinjo, T. Takada, *J. Phys. Soc. Jpn.* **47**, 767 (1979)
- J.E. Hirsch, *Phys. Rev. Lett.* **83**, 1834 (1999)
- D. Houssameddine, U. Ebels, B. Delaët, B. Rodmacq, I. Firastrau, F. Ponthenier, M. Brunet, C. Thirion, J.-P. Michel, L. Prejbeanu-Buda, M.-C. Cyrille, O. Redon, B. Dieny, *Nat. Mater.* **6**, 447 (2007)
- Y. Huang, W. Kang, X. Zhang, Y. Zhou, W. Zhao, *Nanotechnology* **28**, 08LT02 (2017)
- A. Hubert, R. Schäfer, *Magnetic Domains* (Springer, Berlin, 1998)
- S. Ikeda, J. Hayakawa, Y. Ashizawa, Y.M. Lee, K. Miura, H. Hasegawa, M. Tsunoda, F. Matsukura, H. Ohno, *Appl. Phys. Lett.* **93**, 082508 (2008)
- S. Ikeda, K. Miura, H. Yamamoto, K. Mizunuma, H.D. Gan, M. Endo, S. Kanai, J. Hayakawa, F. Matsukura, H. Ohno, *Nat. Mater.* **9**, 721 (2010)
- J.D. Jackson, *Classical Electrodynamics*, 3rd edn. (John Wiley and Sons Inc, Hoboken, 1999)
- H. Jaeger, *GMD Report* **148**, 13 (2001)
- H. Jaeger, *GMD Report* **152**, 60 (2002)

- H. Jaeger, H. Haas, *Science* **304**, 78 (2004)
- A.S. Jenkins, R. Lebrun, E. Grimaldi, S. Tsunegi, P. Bortolotti, H. Kubota, K. Yakushiji, A. Fukushima, G. de Loubens, O. Klein, S. Yuasa, V. Cros, *Nat. Nanotechnol.* **11**, 360 (2016)
- M. Lulliere, *Phys. Lett. A* **54**, 225 (1975)
- S. Kaka, M.R. Pufall, W.H. Rippard, T.J. Silva, S.E. Russek, J.A. Katine, *Nature* **437**, 389 (2005)
- R. Karplus, J.M. Luttinger, *Phys. Rev.* **95**, 1154 (1954)
- J.A. Katine, F.J. Albert, R.A. Buhrman, E.B. Myers, D.C. Ralph, *Phys. Rev. Lett.* **84**, 3149 (2000)
- Y.K. Kato, R.C. Myers, A.C. Gossard, D.D. Awschalom, *Science* **306**, 1910 (2004)
- G. Khalsa, M.D. Stiles, J. Grollier, A. Dussaux, K.A. Zvezdin, V. Cros, *Phys. Rev. B* **80**, 140401 (2009)
- A.V. Khvalkovskiy, J. Grollier, A. Dussaux, K.A. Zvezdin, V. Cros, *Phys. Rev. B* **80**, 140401 (2009)
- S.I. Kiselev, J.C. Sankey, I.N. Krivorotov, N.C. Emley, R.J. Schoelkopf, R.A. Buhrman, D.C. Ralph, *Nature* **425**, 380 (2003)
- H. Kubota, A. Fukushima, Y. Ootani, S. Yuasa, K. Ando, H. Maehara, K. Tsunekawa, D.D. Djayaprawira, N. Watanabe, Y. Suzuki, *Jpn. J. Appl. Phys.* **44**, L1237 (2005)
- H. Kubota, S. Ishibashi, T. Saruya, T. Nozaki, A. Fukushima, K. Yakushiji, K. Ando, Y. Suzuki, S. Yuasa, *J. Appl. Phys.* **111**, 07C723 (2012)
- H. Kubota, K. Yakushiji, A. Fukushima, S. Tamaru, M. Konoto, T. Nozaki, S. Ishibashi, S. Saruya, S. Yuasa, T. Taniguchi, H. Arai, H. Imamura, *Appl. Phys. Express* **6**, 103003 (2013)
- K. Kudo, T. Morie, *Appl. Phys. Express* **10**, 043001 (2017)
- K. Kudo, T. Nagasawa, K. Mizushima, H. Suto, R. Sato, *Appl. Phys. Express* **3**, 043002 (2010)
- A. Kundt, *Ann. Phys. (Berlin)* **285**, 257 (1893)
- L. Landau, E. Lifshitz, *Phys. Z. Sowjet* **8**, 153 (1935)
- K.-J. Lee, A. Deac, O. Redon, J.-P. Nozières, B. Dieny, *Nat. Mater.* **3**, 877 (2004)
- E.M. Lifshitz, L.P. Pitaevskii, *Statistical Physics Part 2, Course of Theoretical Physics*, vol. 9 (Butterworth-Heinemann, Oxford, 1980)
- Y. Liu, H. He, Z. Zhang, *Appl. Phys. Lett.* **91**, 242501 (2007)
- N. Locatelli, V. Cros, J. Grollier, *Nat. Mater.* **13**, 11 (2014)
- N. Locatelli, A. Hamadeh, F.A. Araujo, A.D. Belanovsky, P.N. Skirdkov, R. Lebrun, V.V. Naletov, K.A. Zvezdin, M. Munoz, J. Grollier, O. Klein, V. Cross, G. de Loubens, *Sci. Rep.* **5**, 17039 (2015)
- W. Maass, T. Natschläger, H. Markram, *Neural Comput.* **14**, 2531 (2002)
- S. Maekawa (ed.), *Concepts in Spin Electronics* (Oxford Science Publications, Oxford, 2006)
- S. Maekawa, U. Gäfvart, *IEEE Trans. Magn.* **18**, 707 (1982)
- S. Maekawa, T. Shinjo (eds.), *Spin Dependent Transport in Magnetic Nanostructures* (CRC Press, Boca Raton, 2002)
- S. Maekawa, S.O. Valenzuela, E. Saitoh, T. Kimura (eds.), *Spin Current* (Oxford Science Publications, Oxford, 2012)
- F.B. Mancoff, N.D. Rizzo, B.N. Engel, S. Tehrani, *Nature* **437**, 393 (2005)
- D.P. Mandic, J.A. Chambers, *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability* (Wiley, 2001)
- D. Marković, N. Leroux, M. Riou, F.A. Araujo, J. Torrejon, D. Querlioz, A. Fukushima, S. Yuasa, J. Trastoy, P. Bortolotti, J. Grollier, *Appl. Phys. Lett.* **114** (2019)
- T. Maruyama, Y. Shiota, T. Nozaki, K. Ohta, N. Toda, M. Mizuguchi, A.A. Tulapurkar, T. Shinjo, M. Shiraishi, S. Mizukami, Y. Ando, Y. Suzuki, *Nat. Nanotechnol.* **4**, 158 (2009)
- T.R. McGuire, R.I. Potter, *IEEE Trans. Magn.* **11**, 1018 (1975)
- W.H. Meiklejohn, C.P. Bean, *Phys. Rev.* **102**, 1413 (1956)
- S. Miwa, S. Ishibashi, H. Tomita, T. Nozaki, E. Tamura, K. Ando, N. Mizuochi, T. Saruya, H. Kubota, K. Yakushiji, T. Taniguchi, H. Imamura, A. Fukushima, S. Yuasa, Y. Suzuki, *Nat. Mater.* **13**, 50 (2014)
- T. Miyazaki, N. Tezuka, *J. Magn. Magn. Mater.* **139**, L231 (1995)
- S. Mizukami, Y. Ando, T. Miyazaki, *Phys. Rev. B* **66**, 104413 (2002)
- J.S. Moodera, L.R. Kinder, T.M. Wong, R. Meservey, *Phys. Rev. Lett.* **74**, 3273 (1995)
- S. Murakami, N. Nagaosa, S.-C. Zhang, *Science* **301**, 1348 (2003)

- N. Nagaosa, J. Sinova, S. Onoda, A.H. MacDonald, N.P. Ong, *Rev. Mod. Phys.* **82**, 1539 (2010)
- T. Nagasawa, H. Suto, K. Kudo, K. Mizushima, R. Sato, *J. Appl. Phys.* **109**, 07C907 (2011)
- K. Nakajima, *Jpn. J. Appl. Phys.* **59**, 060501 (2020)
- K. Nakajima, H. Hauser, T. Li, R. Pfeifer, *Sci. Rep.* **5**, 10487 (2015)
- K. Nakajima, H. Hauser, T. Li, R. Pfeifer, *Soft Robot.* **5**, 339 (2018)
- K. Nakajima, K. Fujii, M. Negoro, K. Mitarai, M. Kitagawa, *Phys. Rev. Appl.* **11**, 034021 (2019)
- R. Nakane, G. Tanaka, A. Hirose, *IEEE Access* **6**, 4462 (2018)
- Y. Nakatani, Y. Uesaka, N. Hayashi, *Jpn. J. Appl. Phys.* **28**, 2485 (1989)
- Y. Nakatani, A. Thiaville, J. Miltat, *Nat. Mater.* **2**, 521 (2003)
- H. Nomura, T. Furuta, K. Tsujimoto, Y. Kuwabiraki, F. Peper, E. Tamura, S. Miwa, M. Goto, R. Nakatani, Y. Suzuki, *Jpn. J. Appl. Phys.* **58**, 070901 (2018)
- T. Nozaki, Y. Shiota, M. Shiraishi, T. Shinjo, Y. Suzuki, *Appl. Phys. Lett.* **96**, 022506 (2010)
- M. Oogane, T. Wakitani, S. Yakata, R. Yilgin, Y. Ando, A. Sakuma, T. Miyazaki, *Jpn. J. Appl. Phys.* **45**, 3889 (2006)
- S.S.P. Parkin, C. Kaiser, A. Panchula, P.M. Rice, B. Hughes, M. Samant, S.-H. Yang, *Nat. Mater.* **3**, 862 (2004)
- W. Pauli, *Phys. Rev.* **58**, 716 (1940)
- A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2003)
- D. Pinna, F.A. Araujo, J.-V. Kim, V. Cros, D. Querlioz, P. Bessiere, J. Droulez, J. Grollier, *Phys. Rev. Appl.* **9**, 064018 (2018)
- W.P. Pratt Jr., S.F. Lee, J.M. Slaughter, R. Loloee, P.A. Schroeder, J. Bass, *Phys. Rev. Lett.* **66**, 3060 (1991)
- E.M. Pugh, N. Rostoker, *Rev. Mod. Phys.* **25**, 151 (1953)
- D.C. Ralph, M.D. Stiles, *J. Magn. Magn. Mater.* **320**, 1190 (2008)
- M. Riou, J. Torrejon, B. Garitain, F.A. Araujo, P. Bortolotti, V. Cros, S. Tsunegi, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, D. Querlioz, M.D. Stiles, J. Grollier, *Phys. Rev. Appl.* **12**, 024049 (2019)
- W.H. Rippard, M.R. Pufall, S. Kaka, S.E. Russek, T.J. Silva, *Phys. Rev. Lett.* **92**, 027201 (2004)
- W. Rippard, M.R. Pufall, A. Kos, *Appl. Phys. Lett.* **103**, 182403 (2013)
- M. Romera, P. Talatchian, S. Tsunegi, F.A. Araujo, V. Cros, P. Bortolotti, J. Trastoy, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. Ernoult, D. Vodenicarevic, T. Hirtzlin, N. Locatelli, D. Querlioz, J. Grollier, *Nature* **563**, 230 (2018)
- T. Shinjo (ed.), *Nanomagnetism and Spintronics* (Elsevier, Amsterdam, 2009)
- Y. Shiota, T. Nozaki, F. Bonell, S. Murakami, T. Shinjo, Y. Suzuki, *Nat. Mater.* **11**, 39 (2012)
- R.H. Silsbee, A. Janossy, P. Monod, *Phys. Rev. B* **19**, 4382 (1979)
- N.A. Sinitsyn, *J. Phys.: Condens. Matter* **20**, 023201 (2008)
- A. Slavin, V. Tiberkevich, *IEEE Trans. Magn.* **45**, 1875 (2009)
- J.C. Slonczewski, *Phys. Rev. B* **39**, 6995 (1989)
- J.C. Slonczewski, *J. Magn. Magn. Mater.* **159**, L1 (1996)
- J.C. Slonczewski, *Phys. Rev. B* **71**, 024411 (2005)
- H. Sukegawa, S. Mitani, T. Ohkubo, K. Inomata, K. Hono, *Appl. Phys. Lett.* **103**, 142409 (2013)
- H. Suto, T. Nagasawa, K. Kudo, K. Mizushima, R. Sato, *Appl. Phys. Express* **4**, 013003 (2011)
- T. Taniguchi, *Phys. Rev. B* **91**, 104406 (2015)
- T. Taniguchi, H. Arai, S. Tsunegi, S. Tamaru, H. Kubota, H. Imamura, *Appl. Phys. Express* **6**, 123003 (2013)
- T. Taniguchi, J. Grollier, M.D. Stiles, *Phys. Rev. Appl.* **3**, 044001 (2015)
- T. Taniguchi, T. Ito, S. Tsunegi, H. Kubota, Y. Utsumi, *Phys. Rev. B* **96**, 024406 (2017)
- A.A. Thiele, *Phys. Rev. Lett.* **30**, 230 (1973)
- W. Thomson, *Proc. R. Soc. A* **8**, 546 (1856)
- J. Torrejon, M. Riou, F.A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M.D. Stiles, J. Grollier, *Nature* **547**, 428 (2017)
- Y. Tserkovnyak, A. Brataas, G.E.W. Bauer, *Phys. Rev. Lett.* **88**, 117601 (2002)

- S. Tsunegi, K. Yakushiji, A. Fukushima, S. Yuasa, H. Kubota, *Appl. Phys. Lett.* **109**, 252402 (2016a)
- S. Tsunegi, E. Grimaldi, R. Lebrun, H. Kubota, A.S. Jenkins, K. Yakushiji, A. Fukushima, P. Bortolotti, J. Grollier, S. Yuasa, V. Cros, *Sci. Rep.* **6**, 26849 (2016b)
- S. Tsunegi, T. Taniguchi, S. Miwa, K. Nakajima, K. Yakushiji, A. Fukushima, S. Yuasa, H. Kubota, *Jpn. J. Appl. Phys.* **57**, 120307 (2018a)
- S. Tsunegi, T. Taniguchi, K. Yakushiji, A. Fukushima, S. Yuasa, H. Kubota, *Appl. Phys. Express* **11**, 053001 (2018b)
- S. Tsunegi, T. Taniguchi, R. Lebrun, K. Yakushiji, V. Cros, J. Grollier, A. Fukushima, S. Yuasa, H. Kubota, *Sci. Rep.* **8**, 13475 (2018c)
- S. Tsunegi, T. Taniguchi, K. Nakajima, S. Miwa, K. Yakushiji, A. Fukushima, S. Yuasa, H. Kubota, *Appl. Phys. Lett.* **114**, 164101 (2019)
- S. Urazhdin, V.E. Demidov, R. Cao, B. Divinskiy, V. Tyberkevych, A. Slavin, A.B. Rinkevich, S.O. Demokritov, *Appl. Phys. Lett.* **109**, 162402 (2016)
- T. Valet, A. Fert, *Phys. Rev. B* **48**, 7099 (1993)
- D. Verstraeten, B. Schrauwen, M. D'Haene, D. Stroobandt, *Neural Netw.* **20**, 391 (2007)
- D. Vodenicarevic, N. Locatelli, F.A. Araujo, J. Grollier, D. Querlioz, *Sci. Rep.* **7**, 44772 (2017)
- K. Watanabe, B. Jinnai, S. Fukami, H. Sato, H. Ohno, *Nat. Commun.* **9**, 663 (2018)
- S. Weinberg, *The Quantum Theory of Fields*, vol. 1 (Cambridge University Press, Cambridge, 2005)
- M. Weisheit, S. Fähler, A. Marty, Y. Souche, C. Poinignon, D. Givord, *Science* **315**, 349 (2007)
- S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (Springer, New York, 1990)
- J. Xiao, A. Zangwill, M.D. Stiles, *Phys. Rev. B* **70**, 172405 (2004)
- S. Yakata, H. Kubota, Y. Suzuki, K. Yakushiji, A. Fukushima, S. Yuasa, K. Ando, *J. Appl. Phys.* **105**, 07D131 (2009)
- K. Yakushiji, K. Noma, T. Saruya, H. Kubota, A. Fukushima, T. Nagahama, S. Yuasa, K. Ando, *Appl. Phys. Express* **3**, 053003 (2010)
- T. Yamaguchi, N. Akashi, S. Tsunegi, H. Kubota, K. Nakajima, T. Taniguchi, *Phys. Rev. Res.* **2**, 023389 (2020)
- S. Yuasa, T. Nagahama, A. Fukushima, K. Ando, Y. Suzuki, *Jpn. J. Appl. Phys.* **43**, L588 (2004a)
- S. Yuasa, T. Nagahama, A. Fukushima, Y. Suzuki, K. Ando, *Nat. Mater.* **3**, 868 (2004b)
- S. Zhang, P.M. Levy, *J. Appl. Phys.* **69**, 4786 (1991)
- Y. Zhou, J. Akerman, *Appl. Phys. Lett.* **94**, 112503 (2009)
- Y. Zhou, V. Tiberkevich, G. Consolo, E. Iacocca, B. Azzerboni, A. Slavin, J. Akerman, *Phys. Rev. B* **82**, 012408 (2010)