# Silicon spintronics: Progress and challenges

Viktor Sverdlov, Siegfried Selberherr *

*Institute for Microelectronics, TU Wien Gußhausstraße 27-29, A-1040 Wien, Austria*

## ARTICLE INFO

## ABSTRACT

Electron spin attracts much attention as an alternative to the electron charge degree of freedom for low-power reprogrammable logic and non-volatile memory applications. Silicon appears to be the perfect material for spin-driven applications. Recent progress and challenges regarding spin-based devices are reviewed. An order of magnitude enhancement of the electron spin lifetime in silicon thin films by shear strain is predicted and its impact on spin transport in SpinFETs is discussed. A relatively weak coupling between spin and effective electric field in silicon allows magnetoresistance modulation at room temperature, however, for long channel lengths. Due to tunneling magnetoresistance and spin transfer torque effects, a much stronger coupling between the spin (magnetization) orientation and charge current is achieved in magnetic tunnel junctions. Magnetic random access memory (MRAM) built on magnetic tunnel junctions is CMOS compatible and possesses all properties needed for future universal memory. Designs of spin-based non-volatile MRAM cells are presented. By means of micromagnetic simulations it is demonstrated that a substantial reduction of the switching time can be achieved. Finally, it is shown that any two arbitrary memory cells from an MRAM array can be used to perform a logic operation. Thus, an intrinsic non-volatile logic-in-memory architecture can be realized.

© 2015 Elsevier B.V. All rights reserved.

## Contents

---

* Corresponding author.
  *E-mail address:* Selberherr@TUWien.ac.at (S. Selberherr).

## 1. Introduction

Continuous miniaturization of CMOS devices has made possible a tremendous increase in performance, speed, and density of modern integrated circuits. Numerous outstanding technological challenges have been resolved on this exciting journey. Among the most crucial technological changes lately adopted by the semiconductor industry was the introduction of a new type of multi-gate three-dimensional (3D) transistors [1]. This technology combined with strain techniques and high-*k* dielectrics/metal gates offers great performance and power advantages over planar structures and allows continuing scaling down to 14 nm feature size. There are good indications that device miniaturization with some technological adaptations will continue its pace down to 10 nm feature size. To continue with scaling further a possible modification in the channel material with improved characteristics, which until recently was kept pristine, is foreseen. Research on materials with high mobility [2] demonstrates their potential as channel material. Germanium is also widely considered a promising candidate for a channel material in upcoming devices. The introduction of new materials with higher mobility combined with a multi-gate 3D architecture for improved electrostatic control will potentially allow to proceed with scaling below 10 nm. Although single devices with gate length as short as a few nanometers, which is close to what is believed to be an ultimate limit, were long ago demonstrated [3], fabrication, control, integration, and reliability issues will slowly lead to unbearable hurdles. Growing critical technological challenges and soaring costs will gradually bring scaling to saturation.

A multi-gate 3D device architecture potentially allows device scaling below 10 nm, where transport in the channel becomes nearly ballistic. However, even though the transistor size is scaled down, the load capacitance per unit area of a circuit stops decreasing. This suggests that the on-current must stay constant in order to maintain appropriate high speed operation. In ultra-scaled MOSFETs with ballistic transport in the channel the conductance is determined by the number of transversal propagating modes in the channel and ceases to depend on the channel length. It then follows that to maintain the same current through a transistor the supply voltage cannot be decreased and must remain the same, even if the channel length is reduced. This results in an approximately constant power dissipation of a single MOSFET regardless of its channel length, which would lead to a rapid increase of generated heat with increased transistor density, resulting in a saturation of MOSFET scaling, even under the assumption that the excessive heat generated due to parasitic leakage currents can be controlled. The obvious saturation of MOSFET miniaturization puts clearly foreseeable limitations to the continuation of the increase in the performance of integrated circuits, and research for finding alternative technologies and computational principles becomes urgently needed.

Carriers in graphene are characterized by a mobility which can exceed silicon mobility by an order of magnitude [4], making this material very attractive for future electronic device applications. However, Dirac fermions can be transmitted through a classically forbidden region without any reflection. Therefore, applying a gate bias, which would normally switch off a traditional MOSFET, does not stop the current flow through a graphene sheet. A high on–off current ratio can be obtained, if the graphene sheet is used as a base of a bipolar-like transistor [5]; albeit, advantages steaming from the high graphene carrier mobility are not used. Nanoribbons may open a gap in the Dirac spectrum, thus making reflection from a forbidden region possible [6]. However, in order to avoid strong edge roughness scattering [7], the edges of nanoribbons must be atomistically perfect, which makes the practical use of nanoribbons hypothetical.

Recently, a single layer of molybdenum disulfide $MoS_2$, a purely two-dimensional material, has attracted interest [8]. In contrast to graphene, this is a direct gap semiconductor and is thus suitable for making a MOSFET, however, both electron and hole mobilities are limited by the carriers' interaction with homopolar phonons [9,10]. It remains to be seen, if the electron mobility and current in this material can be sufficiently boosted [11] to make it competitive with CMOS. Also, reliability issues such as bias temperature instability and hot carriers injection have to be further evaluated [12].

The principle of MOSFET operation is fundamentally based on the charge degree of freedom of an electron: The electron charge interacts with the gate induced electric field which can close the transistor by creating a potential barrier. Another intrinsic electron property, the electron spin, attracts at present much attention as a possible candidate for complementing or even replacing the charge degree of freedom in future devices [13–21]. The electron spin state is characterized by one of the two of its possible projections on a given axis and could be potentially used in digital information processing. In

**List of symbols**

| | |
|---|---|
| $a$ | long axis of a magnetic layer of elliptic form |
| $A$ | exchange constant |
| $\alpha$ | Gilbert damping |
| $\alpha_R$ | Rashba spin–orbit coupling constant |
| $b$ | short axis of a magnetic layer of elliptic form |
| $B$ | in-plane momentum-dependent gap between the conduction bands at the $X$-point |
| $\beta$ | interface asymmetry induced Dresselhaus-like spin–orbit coupling constant |
| $c$ | velocity of light |
| $C_{\uparrow(\downarrow)}$ | coefficients of up (down) spin projection wave functions in the channel |
| $\Delta_\Gamma$ | splitting between the relevant conduction bands at the $\Gamma$ point |
| $g$ | gyromagnetic splitting factor |
| $g_{1(2)}$ | Slonczewski coefficients for spin transfer torque |
| $\gamma$ | angle between magnetic field and transport direction |
| $\gamma_G$ | gyromagnetic ratio |
| $\Gamma$ | center of the Brillouin zone symmetry point |
| $d$ | thickness of magnetic recording layer |
| $D$ | acoustic deformation potential |
| $D_{xy}$ | shear strain deformation potential |
| $D_{SO}$ | deformation potential due to intrinsic spin–orbit interaction |
| $\mathbf{D}$ | deformation potential matrix |
| $D_{DIFF}$ | diffusion constant |
| $\Delta_{SO}$ | strength of effective spin–orbit interaction |
| $\Delta_\Gamma$ | gap between the relevant conduction bands at the $\Gamma$ symmetry point |
| $e$ | electron charge |
| $E_F$ | Fermi energy |
| $E_n^0$ | $n$th subband energy |
| $\underline{E_n}(\mathbf{K})$ | $n$th subband dispersion |
| $\overline{E}$ | error value |
| $f(E)$ | Fermi–Dirac function |
| $\varepsilon_{xy}$ | shear strain component |
| $\delta E_c$ | conduction band offset between the contacts and the channel |
| $\hbar$ | reduced Planck's constant |
| $h_0$ | exchange splitting energy in a ferromagnet |
| $H, \tilde{H}$ | effective Hamiltonian |
| $k_0$ | position of the valley minimum relative to the $X$-point |
| $k_B$ | Boltzmann constant |
| $k_F$ | Fermi wave vector |
| $K_{0\Gamma}$ | position of the valley minimum relative to the $\Gamma$-point |
| $k_x, k_y, k_z$ | electron wave vector projections |
| $k_D$ | inverse length defining $2\pi$ spin precession due to spin–orbit interaction |
| $\mathbf{K}_n$ | electron wave vector in $n$th subband |
| $L$ | channel length |
| $\Lambda_\Gamma$ | intervalley splitting in unstrained films |
| $m_0$ | electron mass |
| $m_l$ | longitudinal effective mass in silicon |
| $m_t$ | transversal effective mass in silicon |
| $m_f$ | effective mass in ferromagnet |
| $m*_n$ | transport mass in $n$th subband |
| $M$ | off-diagonal intervalley parameter |
| $\mu_0$ | vacuum permeability |
| $\mu_B$ | Bohr magneton |
| $P$ | spin polarization |
| $\mathbf{q}$ | scattering wave vector |
| $\rho$ | silicon density |
| $\rho_S$ | silicon resistivity |
| $\sigma_x, \sigma_y, \sigma_z$ | Pauli matrices |
| $t$ | silicon film thickness |
| $T$ | absolute temperature |

| | |
|---|---|
| $\tau_S$ | spin lifetime |
| $\tau_y$ | Pauli matrix in valley index space |
| $\Theta$ | Heaviside function |
| $\upsilon_{LA}$ | velocity of longitudinal acoustic phonons |
| $\upsilon_{TA}$ | velocity of transverse acoustic phonons |
| $\delta V$ | cell volume |
| $\Delta V$ | voltage signal |
| $V$ | crystal potential |
| $z$ | dimensionless tunnel barrier |

addition, it takes an amazingly small amount of energy to invert the spin orientation, which is necessary for low power applications. Even more, the electron spin as a vector may be pointed not only up or down but rather in any direction on a unit Bloch sphere. This opens the way to use the whole Bloch sphere of states to process and store information by initializing, manipulating, and detecting the spin orientation. Because the electron spin is a purely quantum mechanical property, the set of states on a Bloch sphere is called a quantum bit, or a qubit, as opposed to a bit of classical binary information. A computer using qubits for information processing is named a quantum computer. Due to their quantum mechanical nature, several qubits could form an entangled state, wherein a single qubit state cannot be described independently. The initially proposed quantum computation scheme [22] was based on spins in quantum dots. Single-spin systems in semiconductors can be broadly classified in two large categories: Atomic impurities [23] and spins in quantum dots [24]. Carriers in quantum dots are strongly confined in all three spatial directions. If the confinement potential possesses a special symmetry, the degenerate energy levels in quantum dots, like in real atoms, form shells which are well separated from each other. For this reason the electron system in a quantum dot is often called an artificial atom. Properties of artificial atoms can be tailored by adjusting confinement. It allows changing controllably the number of particles in a dot and the total spin [25]. A successful implementation of a quantum computer based on spins requires the possibility of efficient spin initiation, coherent manipulation, and reliable read-out. We note that a scheme for implementing a quantum-mechanical computer using the nuclear spins of donors in silicon was proposed nearly two decades ago [26]. A comprehensive review [27] describes group IV based solid-state proposals for quantum computation. An unprecedented advantage in these fields has been achieved by the researchers in the last decade [28]. The experiments on electron spins in semiconductors were performed at cryogenic temperatures, where in silicon a relaxation time of several seconds was demonstrated [29]. Although these results are encouraging, the development of a robust two-qubit gate becomes a pressing challenge [30] before proceeding to a larger computational network.

Silicon predominantly consists of $^{28}$Si nuclei with zero magnetic spin, with a few percent admixture of magnetic $^{29}$Si nuclei. Thus, dephasing of donors' electron spins in purified silicon cleaned from the magnetic isotope $^{29}$Si is significantly reduced [31], resulting in a coherence time of about 10 s at $T = 1.8$ K. The spin–orbit interaction is also weak in the conduction band of silicon, supporting a long coherence time at cryogenic temperatures. Regardless its unique properties, silicon, the main material used by modern microelectronics, was remaining aside from the main stream of spin-related applications. Certainly, the use of silicon for spin-driven devices would greatly facilitate their integration with MOSFETs on the same chip. However, this will necessarily require the capability of room temperature operation, which at the moment excludes the use of quantum-mechanical properties of localized electron spins in silicon due to fast dephasing, causing the loss of quantum information at (sub)nanosecond scale. More promising for room temperature quantum bit storage are the nuclear spins of the ionized shallow donors in purified $^{28}$Si, because they display an unprecedentedly long spin coherence time close to 40 min [32], which considerably exceeds the coherence time of 1 s demonstrated at room temperature in a quantum bit consisting of a single $^{13}$C nuclei/nitrogen vacancy in an isotopically purified diamond crystal [33].

There exists yet another opportunity to employ the spin of electrons in silicon. The spin relaxation time for conducting electrons describing the characteristic time of non-equilibrium spin relaxation towards its equilibrium value is in the range of nanoseconds [34,35]. Since the corresponding spin diffusion length is already in a micrometer range, the spin of conducting electrons can be used to encode, transfer, and process information. In this case the electron spin degree of freedom is used in its binary mode and, as in a standard bit, the information is stored as a positive or negative spin projection on a given axis, thus, in a classical sense. As the spin does not interact with the electric field and diffuses from the spin accumulation region regardless of the electric field applied, transferring the spin information does not require any charge current to flow. The ability to decouple spin from charge makes the spin degree of freedom potentially promising for low-power application.

Although the idea of complimenting or even replacing the charge degree of freedom by spin is around for almost three decades, even a demonstration of basic elements necessary for spin related applications, such as injection of spin-polarized currents in silicon, spin transport, spin manipulation, and detection, was missing until recently. The inability to inject spin-polarized electrons in silicon was a stumbling point preventing researchers from studying spin propagation in silicon. The main reason was that, contrary to III–V semiconductors, silicon is a semiconductor with an indirect gap, and techniques based on optical excitations with the absorption/emission of circularly polarized photons used to introduce/detect spins in GaAs cannot be applied [18,19].

Although it should be straightforward to inject spin-polarized carriers into silicon from a ferromagnetic contact, due to a fundamental conductivity mismatch [36] between a ferromagnetic metal contact and the semiconductor, the problem was without solution for a long time. A special technique [37] based on the attenuation of hot electrons with spins anti-parallel to the magnetization of the ferromagnetic film allowed creating an imbalance between the electrons with spin-up and spin-down in silicon, thus injecting spin-polarized current. The spin-coherent transport through the device was studied by applying an external magnetic field causing precession of spins during their propagation from source to drain. The detection is performed with a similar hot electron spin filter. Regardless of the fact that the drain current is fairly small due to the carriers' attenuation in the source and drain filters as compared to the current of injected spins, the experimental set-up represents a first spin-driven device which can be envisaged working at room temperature. Contrary to the MOSFET, however, the described structure is a two-terminal device. By using this all-electrical ballistic hot-electron injection and detection techniques it was also shown that band bending gives rise to highly non-Ohmic spin transport in n-type lightly phosphorus-doped silicon [38]. This phenomenon is due to the interaction of conduction electrons with shallow traps, which allows exploring the local interaction of spin information with an isolated impurity [39]. The oblique magnetic field results in a non-trivial behavior of spin precession and dephasing, making the Hanle peak to widen with the in-plane field component increased [40].

The first demonstration of coherent spin transport through an undoped 350 $\mu$m thick silicon wafer [41] has triggered a systematic study of spin transport properties in silicon [21].

In the following we briefly review the recent progress in spin-driven applications based on silicon and CMOS-compatible devices. Since we are interested in room temperature applications, the spin degree is used in its classical sense aiming at complementing/replacing the electron charge in digital applications. We begin with recent advances in spin injection and propagation in silicon, with a particular focus on spin lifetime enhancement techniques in an electron system in silicon-on-insulator structures widely considered as potential candidates for achieving ultimate MOSFET scaling. Long spin lifetime facilitates a silicon-based implementation of spin-based reconfigurable large-scale logic circuits [42] and all-spin logic devices [44]. At the same time, implementation of a Datta–Das silicon-based SpinFET [13] is still pending due to the lack of efficient approaches for spin manipulation in the transistor channel by purely electrical means.

The coupling between the spin (magnetization) and the charge (electric current) is stronger in magnetic tunnel junctions due to the tunneling magnetoresistance and spin-torque transfer effects [45,46]. The efficient conversion of the relative magnetization orientations to the resistance and magnetization reversal by current breath a new life into the magnetic random access memory concept (MRAM). The compatibility of MRAM with CMOS stimulated the fast development of new spin-transfer torque (STT) MRAM products which are already commercially available. We briefly discuss possible ways to reduce the yet quite large current value for magnetization switching. The availability of CMOS-compatible MRAM-based non-volatile arrays opens new horizons for developing and implementing intrinsic logic-in-memory architectures where the same elements are used for information storage and processing.

## 2. The spin field-effect transistor

The spin field-effect transistor (SpinFET) is a future semiconductor spintronic device promising a performance superior to what can be achieved with the present transistor technology. Complementing or replacing the charge degree of freedom used for computation in modern CMOS circuits with the electron spin promises to reduce in particular the energy dissipation [20]. A SpinFET is composed of two ferromagnetic contacts (source and drain), linked by a non-magnetic semiconductor channel region. The ferromagnetic contacts inject and detect spin-polarized electrons, in analogy to polarizer and analyzer as indicated already long ago by Datta and Das [13]. Because of the effective spin–orbit interaction in the channel, which depends on the perpendicular effective electric field, the spin of an electron injected from the source starts precessing. In order to distinguish this electric field-dependent spin–orbit interaction from the intrinsic electron spin–orbit interaction acting on any electron moving in a crystal potential, we term the electric field-dependent spin–orbit interaction as coupling. The electrons with spin, or to be more precise with the direction of the magnetic moment, aligned to the drain magnetization direction can easily leave the channel to the drain, thus contributing to the current. The total current through the device depends on the relative angle between the magnetization direction of the drain contact playing the role of an analyzer and the electron spin polarization at the end of the semiconductor channel. An additional current modulation is achieved by tuning the strength of the spin–orbit interaction in the semiconductor region, which depends on the effective electric field and can be controlled by purely electrical means applying a gate voltage. Although the SpinFET was proposed two decades ago [13], it has not been experimentally demonstrated up to now. In order to realize the SpinFET, the following requirements must be fulfilled [47]. First, an efficient spin injection into the channel (and detection) must be realized. Second, because the electron spin in the channel is not a conserved quantity and thus relaxes due to spin-flip processes, the corresponding scattering mechanisms must be detected and analyzed. It is important to identify the possibilities compatible with modern MOSFET technology, which can enhance the spin lifetime and spin diffusion length in the silicon channel. Finally, purely electrical means of spin manipulation in the channel must be identified to control the spin and thus the current flow to the drain. An example of such a manipulation is the gate voltage-dependent effective spin–orbit interaction defining the degree of the spin precession. Next we briefly discuss recent achievements and challenges for the practical realization of a SpinFET.

*2.1. Spin injection: Recent advances and open issues*

Spin injection into silicon and other semiconductors by purely electrical means from a ferromagnetic metal electrode was not very successful until recently. The fundamental reason has been identified as an impedance mismatch problem [36]. Even though there is a large spin imbalance between the majority and minority spins in a metal ferromagnet, both channels with spin-up and spin-down are equally populated in a semiconductor due to the relatively small density of states as compared to that for the minority spins in a ferromagnet. In other words, because of the large resistance of the semiconductor, the voltage applied to the contact between the ferromagnet and the semiconductor drops completely within the semiconductor. Therefore, the properties of the contact are dominated by the non-magnetic semiconductor, thus resulting in a current without spin polarization. One solution to overcome the impedance mismatch problem is the use of hot electron injection [37]; however, the efficiency of spin injection and detection is very limited.

Another solution to the impedance mismatch problem is the introduction of a potential barrier between the metal ferromagnet and the semiconductor [48]. In this case the influx of carriers from the ferromagnet into the semiconductor is reduced to such an extent that the majority spins supply just enough carriers to support the complete occupancy of the corresponding states in the semiconductor. Under such conditions the minority spin flow in semiconductors will be a fraction of that for the majority spins defined by the spin polarization in the ferromagnet. This guarantees the existence of a spin polarized current and the spin injection into the semiconductor.

Two schemes to analyze the spin injection into a semiconductor are available: The three-terminal method and the non-local one. The three-terminal method uses two contacts to introduce an electrical current into the semiconductor. One of the contacts is a tunnel contact with a ferromagnet. When current is flowing, a spin accumulation is created under this tunnel contact. This spin accumulation leads to a difference in the chemical potentials of spins aligned or anti-aligned to the contact magnetization direction. Because the ferromagnetic contact potential is that of the spins aligned with its magnetization, a detectable voltage is generated between the ferromagnetic electrode and the third contact located at a distance larger than the spin diffusion length, where the spin accumulation vanishes.

In a magnetic field perpendicular to the contact magnetization spins start to precess. Accordingly any excess spin density is reduced as well as the corresponding voltage signal. The half-width of the voltage dependence on the magnetic field is interpreted as the spin lifetime $\tau_s$. An estimate for the spin lifetime of about $0.1$–$1$ ns has been observed for $n$-doped silicon, in agreement with the electron spin resonance method [21]. An experimental signal obtained with the three-terminal method and interpreted as corresponding to spin accumulation was first reported in 2007 at low temperature, when the current was injected from an iron electrode through $Al_2O_3$ [49,50]. A signal at room temperature obtained for both n- and p-doped silicon was documented in 2009 [51] by using an $Ni_{80}Fe_{20}/Al_2O_3$ tunnel contact. In order to avoid an extended depletion layer the authors used heavily doped silicon samples. It was believed that it was actually this depletion layer and not the impedance mismatch problem, which long prevented the all-electric demonstration of a signal potentially corresponding to spin injection. Electrical signals believed to be corresponding to spin injection through silicon dioxide at temperatures as high as 500 K has also been reported [52].

It then follows that the tunnel contacts have to be optimized in order to facilitate spin injection. They must not be too thick to make the signal observable, but at the same time they must not be too transparent to avoid the impedance mismatch problem [21,53,54]. It turns out that nearly all tunnel contacts fabricated so far are characterized by a contact resistance larger than the optimal one [21], and the problem of making good contacts with low resistance per area still requires some attention. Recently, tunnel contacts made of a single layer graphene [55] have been shown to deliver a contact resistance close to the optimum [56]. Even more, the proper two-dimensional analysis demonstrates the important role of the contact width [57], predicting an optimal contact width to maximize the magnetoresistive effect in the lateral spin valve geometry used for spin injection.

The tunnel barrier in the contact of a ferromagnet to a semiconductor introduces spin-dependent interface resistances $R_\uparrow, R_\downarrow$ [21], which define the tunnel spin polarization $P = (R_\downarrow - R_\uparrow)/(R_\uparrow + R_\downarrow)$. Due to the additional spin filtering effect in MgO, the high quality tunneling stack made of a Fe and MgO $(001)$ crystalline barrier is expected to provide spin polarization up to 70% at room temperature [58]. A signal interpreted as corresponding to spin injection by using MgO as a tunnel barrier has been reported at room temperature [59,60]. It has been shown that the increase of the signal is caused by spin filtering in the MgO layer which must be of good quality to guarantee a high spin polarization.

Regardless several reports claiming demonstrating spin injection into silicon at room temperature with the three-terminal method, there are unsolved challenges which may put the obtained results in question or even compromise them. According to theory, in the three-terminal scheme the value of the voltage signal $\Delta V$ due to spin accumulation divided by the current density $j$ flowing through the injecting contact is proportional to

$$\Delta V/j = P^2 \rho_S \sqrt{D_{DIFF} \tau_s}, \tag{1}$$

where, because of the injection and detection, the tunnel spin polarization $P$ enters squared, and the silicon resistivity $\rho_S$ multiplied with the diffusion length $\sqrt{D_{DIFF} \tau_s}$ determines the additional area resistance of the contact due to spin accumulation under it. However, there is a several orders of magnitude discrepancy between the signal measured and the theoretical value (1). It turns out that the signal is stronger in three-terminal measurements, while it is weaker in the non-local scheme [21]. Similar observations were also made for germanium [61] as well as for other semiconductors [62]. The reasons for these discrepancies are heavily debated [21,63–65].

An additional source of concerns on the validity of the three-terminal methods to determine the spin accumulation is an unreasonably long, 0.3 ns spin lifetime extracted with this method in p-doped silicon as compared to that in the picosecond range or below expected for silicon. This indicates that the signal measured with the three-terminal method is not determined by the spin accumulation in the semiconductor. Indeed, in silicon as well as in germanium and GaAs scattering between the heavy-hole and light-hole bands leads to a very fast spin relaxation which is supported by optical measurements [66,67]. However, three-terminal measurements in germanium and silicon yield 50 ps and 300 ps spin lifetime for holes, correspondingly, strongly indicating that deep centers embedded in the Schottky tunnel barrier and the oxide contribute to the signal.

Recently, the large amplitude of the signal observed in the three-terminal injection method was attributed to resonant tunneling through deep impurities [64]. In contrast to [63], this model of resonance-tunneling magnetoresistance in a magnetic field seems to be consistent with extensive experimental data [68]. It explains on equal footing the inverted Hanle effect in a magnetic field parallel to the magnetization direction [64]. In confirmation to the theory [64], new experimental reports [69–71] clearly show that bias voltage and magnetic field dependences of the inelastic electron tunneling spectra account for the dominant contributions to the three-terminal magnetoresistance. These findings strongly support the claim [64] that it is inelastic tunneling through defects and deep centers at junction interfaces and within the barrier, rather than spin accumulation in silicon, which determines the magnetoresistance signal.

It is worth mentioning that in a recent study of spin injection with the three-terminal method [72] it was argued that the theory [64] cannot explain the large voltage signal obtained in Ge with epitaxial contacts, because there is no oxide tunnel barrier and thus no deep center present. However, deep impurity centers could be embedded in the Schottky barrier during device fabrication [72]. It is plausible to assume that Mn atoms form deep energy defects in the gap of Ge and electron spins extracted from Ge get trapped at these deep centers, thus contributing to the signal in a similar fashion as the impurities embedded in an oxide tunnel barrier.

Since the electron spin does not interact with the electric field, the spin diffusion appears in all directions from the accumulation region. It then follows that one could add an auxiliary ferromagnetic electrode at a distance shorter than the spin diffusion length and measure the voltage difference between the auxiliary and the third electrode. The signal sign will then depend on whether the auxiliary electrode is parallel to the aligned or anti-aligned spin orientation injected through the current carrying ferromagnetic contact, which is determined by the relative orientations of the ferromagnetic contacts used for injecting and detecting spins. The amplitude of the signal will also be a function of the distance between the two ferromagnetic electrodes: The larger the distance, the smaller the signal. Because the spin accumulation is probed not under the spin injecting contact but at an arbitrary distance from it, this scheme is called non-local [73–82]. The spin signal observed with this scheme in degenerate Si at room temperature [81] is believed to be due to a genuine spin injection. The amplitude of the signal is an order of magnitude weaker [56] than predicted by (1), which is consistent with a possible spin polarization loss at the interfacial states in the ferromagnetic contacts. Spin transport in a non-degenerate n-doped Si at room temperature was successfully also achieved [83]. Interestingly, the magnitude of the spin signal exceeds 1 mV for a bias electric current of 1 mA, ten times larger than previously reported for degenerate Si [84].

Spin can be injected into silicon by other techniques as well. The injection of spins by heat [85] is one of them, for which a spin current through the contact exists without a charge current. Recently, the energy dispersion of the tunnel spin polarization of ferromagnetic tunnel contacts $Co_{70}Fe_{30}$/MgO to both silicon and germanium has been extracted from thermal and electrical spin currents [65]. It was demonstrated that the relative magnitude of the electrical spin current was determined by the value of the tunnel spin polarization, while the thermal spin signal is proportional to its energy derivative, which brings up a further contribution to the controversy discussed above. Interestingly, the application of a voltage can turn the thermal spin current off or even invert it [65].

Another technique is spin pumping by inducing magnetic excitations in a material in contact with a semiconductor [86,87], which can generate pure spin currents at room temperature [88]. The magnons are excited by the microwave. This technique is extremely useful, since it is free of the impedance mismatch problem, and a combination with conventional spin injection techniques will advance spintronics towards sophisticated device applications [89]. In [90] the spins were pumped in p-type silicon. The detection was performed non-locally by the inverse spin Hall effect in the Pd contact. Surprisingly, the signal was detected at a distance exceeding the spin diffusion length for holes. Because of the very short spin relaxation time of holes, the detected spin injection signal is most probably governed by the minority spin polarized electrons pumped to the inversion layer on the surface between the Schottky barrier and the FM, which then propagate to the Pd contact.

The introduction of a tunnel barrier is the standard but not the only possible option to overcome the impedance mismatch problem. The use of ferromagnetic contacts made of semiconductors would be another possible solution. Unfortunately, no semiconductors with ferromagnetic properties surviving up to room temperature are known. Attempts to rise the Curie temperature in (Ga, Mn)As by increasing the Mn concentration at carefully optimized synthesis conditions result in a $T_C$ saturation at about 200 K [91]. Several oxides can be transformed to dilute ferromagnetic semiconductors by doping with magnetic impurities. Although a high saturation magnetization and a finite coercivity at room temperature was recently demonstrated in an epitaxial film with $Sr_3SnO$ integrated on silicon [92], the application of the material as a contact may be overshadowed by a low conductivity due to variable range hopping.

An alternative approach to resolve the impedance mismatch problem is to use half-metallic ferromagnets. These materials possess metallic properties, but only with respect to the electrons of spin orientation up, and they act like insulators with respect to the electrons with spin orientation down. Thus, only the electrons with the spin orientation up can be

injected regardless of the differences in the densities of states in the half-metallic ferromagnet and the semiconductor. Many Heusler alloys and some transition metal oxides ($CrO_2$, $Fe_3O_4$, etc.) are half-metallic ferromagnets; however, the research is in its inceptive state. A recent report on the fabrication and structural study of high-quality $Co_2FeSi/SiO_x$ $N_y/Si(100)$ tunnel contacts [93] will stimulate the investigation of electrical and spin properties of half-metallic ferromagnetic contacts on silicon.

## 2.2. Spin propagation and relaxation in bulk silicon

For functionality of a SpinFET the possibility to transfer the excess spin injected from the source to the drain electrode is essential. The excess spin, or more precisely the spin projection on a given axis, is not a conserved quantity, in contrast to charge. While diffusing, it gradually relaxes to its equilibrium value which is zero in a non-magnetic semiconductor. In a ground breaking experiment it was demonstrated that spin can propagate through a 350 μm silicon wafer at liquid nitrogen temperatures. The lower estimation for the spin lifetime at room temperature obtained within the three-terminal injection scheme was of the order 0.1–1 ns [21]. This corresponds to a spin diffusion length, the length at which the spin relaxes, of 0.2–0.5 μm. The spin lifetime is determined by the spin-flip processes. Several important spin relaxation mechanisms are identified [18,19]. In silicon the spin relaxation due to the hyperfine interaction of spins with the magnetic moments of the $^{29}Si$ nuclei is important at low temperature. Because of the inversion symmetry in the silicon lattice the Dyakonov–Perel spin relaxation mechanism is absent in bulk systems, and at elevated temperatures the spin relaxation due to the Elliott–Yafet mechanism [18,19] becomes important.

The Elliott–Yafet mechanism is mediated by the intrinsic interaction between the orbital motion of an electron and its spin. Due to the spin dependence, the microscopic spin–orbit interaction does not conserve the electron spin, thus it generates spin flips, which is the Yafet process. When the microscopic spin–orbit interaction is taken into account, the Bloch function with a fixed spin projection is not an eigenfunction of the total Hamiltonian. Because the eigenfunction always contains a contribution with an opposite spin projection, even spin-independent scattering with phonons generates a small probability of spin flips, which is the Elliott process.

In order to analyze the spin relaxation in silicon, both, the Elliott and the Yafet processes must be taken on equal footing. In this way a good agreement between the experimentally observed and calculated spin life time as a function of temperature has been achieved, confirming that in bulk silicon the Elliott–Yafet mechanism is the dominant spin relaxation mechanism at ambient temperatures [34]. The spin lifetime in undoped silicon at room temperature is about 10 ns, which corresponds to a spin diffusion length of 2 μm. In case of heavily doped silicon the spin lifetime is determined by the Elliott–Yafet mechanism due to ionized impurity scattering and is expected to be around 1 ns at $N_D = 10^{19}$ $cm^{-3}$, in agreement with experiments. In contrast to mobility, the spin relaxation depends significantly on the doping atom type in heavily doped silicon [94]. As it has been revealed in [35,94,95], the spin relaxation is due to the scattering between the valleys along different crystallographic axes ($f$-processes) while the electron mobility is determined by intravalley elastic scattering. Therefore, the momentum relaxation time is governed by the tail of the ionized impurity potential at large distances from an impurity, which is identical for donors considered in [94], while the spin relaxation is due to the short range impurity potential behavior where spin–orbit interaction is the strongest, thus giving rise to the spin relaxation dependence on the doping atom type [94].

Because the intervalley $f$-processes are responsible for spin relaxation in bulk Si, it is advantageous to partly lift the degeneracy between the valleys in order to reduce the spin relaxation. This can be achieved by stress. Biaxial in-plane tensile stress or uniaxial out-of-plane compressive stress favors the two out-of plane valleys to lower their energies and become populated by electrons at the expense of the reduced occupation of the four in-plane valleys. If the energy splitting becomes larger than the energy of optical phonons needed for $f$-type intervalley scattering, the spin lifetime is predicted to increase by a factor between two and four, depending on the spin injection orientation relative to stress [96].

In confined electron systems of transistor channels the valley degeneracy, or rather the degeneracy between the two-dimensional subbands, is lifted due to different confinement energies for different valleys. It is well known that in (001) oriented silicon inversion channels the energy spectrum consists of primed and unprimed subbands. Primed subbands with higher subband energies originate in four in-plane valleys and are therefore four-fold degenerate, while the unprimed two-fold degenerate subbands are from the two out-of-plane valleys [97]. Despite partial degeneracy lifting in confined systems, the spin lifetime is significantly shorter in gated structures [98,99] due to the existence of the interface. Because of an ongoing shift from bulk field-effect transistors to transistors with the channel built on ultra-thin silicon-on-insulator (SOI) films and three-dimensional fin-like structures at the 22 nm and more recent 14 nm technology node and beyond allowing tighter confinement and thus better electrostatic control, it is expected that spin relaxation will further increase, and ways to boost the spin lifetime in SOI transistors are urgently needed. In the next section we show that uniaxial stress applied along [110] direction is very efficient in boosting the spin lifetime in (001) silicon films. The physical reason is also elucidated: It turns out that the remaining two-fold degeneracy of the unprimed subbands is successfully lifted by shear strain efficiently produced by [110] stress. This reduces significantly the principal contribution to the spin relaxation in thin films due to surface roughness and acoustic phonons mediated $g$-type scattering between the equivalent valleys along the same crystallographic axis, or, more precisely, the intersubband spin relaxation. Stress techniques and especially uniaxial stress along [110] oriented channels are routinely used by the semiconductor industry to enhance the transistor performance. It is therefore very attractive to use the same well developed techniques to increase the spin lifetime. It makes

the use of thin SOI films for spin applications very promising because, in contrast to the electron mobility, which is enhanced by a factor of two, the spin lifetime is boosted by an order of magnitude.

### 2.3. Spin lifetime in thin films and surface layers

The main contribution to spin relaxation in bulk silicon is due to optical phonon scattering between the valleys residing at different crystallographic axis, or $f$-phonons scattering [35,95]. An enhanced intervalley scattering due to accelerated $f$-phonon emission at high driving electric fields results in an unusual combination of reduction of the carrier transition time between the injector and the collector accompanied by a decrease in spin polarization [100]. This independently confirms the importance of $f$-phonon mediated intervalley scattering in bulk silicon spin relaxation. In electrically-gated lateral channel silicon systems a relatively large spin relaxation has been experimentally observed [98,99] indicating that the extrinsic interface induced spin relaxation mechanism becomes important. This may pose an obstacle in realizing spin-driven CMOS compatible devices, and a deeper understanding of fundamental spin relaxation mechanisms in silicon inversion layers, thin films, and fins is needed. The theory of spin relaxation must account for the most relevant scattering mechanisms. In confined electron systems these are the electron–phonon interaction, surface roughness and impurity scattering. In order to evaluate the corresponding scattering matrix elements, the wave functions must be provided. We briefly outline an approach based on an effective $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian, which allows to capture the important physics while keeping the simulation costs relatively low. The effective $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian must include the effective spin–orbit interaction [35], which, apart from scattering, is the main ingredient of the Elliott–Yafet spin relaxation mechanism. In addition, the confinement potential must be incorporated. It is also mandatory to have a sufficiently accurate model of the conduction band describing its non-parabolicity and warping under external stress [101].

### 2.3.1. Silicon conduction band and the two-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian

The conduction band in silicon consists of six equivalent valleys situated close to the edges of the Brillouin zone at the $X$-points. Each valley dispersion is anisotropic and described by the transversal effective mass $m_t = 0.19m_0$ in the plane perpendicular to the axis connecting the $\Gamma$ and the corresponding $X$-point and by the longitudinal effective mass $m_l = 0.91m_0$ along the same axis within the parabolic approximation [97]. At higher energies the non-parabolicity of the conduction band must be taken into account. In order to do so we resort to a more general method of obtaining the energy dispersion: The $\mathbf{k} \cdot \mathbf{p}$ method which is well established to describe the band structure analytically. It turns out that the $\mathbf{k} \cdot \mathbf{p}$ method with only two bands taken into account reproduces the band structure of silicon accurately below 0.7 eV, which is sufficient to describe the subband structure and transport properties of most silicon-based devices.

Let us consider a valley along [001] axis. In the diamond-like crystal structure of silicon, the lowest two conduction bands are $X_1$ and $X_{2'}$. In unstrained silicon their dispersions along the [001] symmetry axis are described as

$$H_{ii}^0(\mathbf{k}) = (-1)^{i-1} \frac{\hbar}{m_0} k_z p + \frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2 k_x^2}{2m_t} + \frac{\hbar^2 k_y^2}{2m_t}, \tag{2}$$

where $i = 1, 2$ numbers the bands, $k_z$ is the wave number relative to the $X$-point, and $p = (p_z)_{11} = -(p_z)_{22}$ defines the slope of the dispersion at the $X$-point. The dispersions (2) describe the two valleys with their minima located at $k_0 = |\pm 0.152\pi/a_0|$ symmetrically with respect to the $X$-point, where $a_0$ is the lattice constant in silicon. We note that the two lowest bands (2) are degenerate exactly at the edge of the Brillouin zone as prescribed by a symmetry due to the presence of the three glide reflection planes [102]. We therefore consider the application of the $\mathbf{k} \cdot \mathbf{p}$ method relative to this degeneracy point at the edge of the Brillouin zone rather than at the valley minimum usually considered in literature.

The degeneracy between the conduction bands at the $X$-point can be lifted, if uniaxial stress along [110] direction is applied [101,103]. In this case the strain tensor contains a shear component $\varepsilon_{xy}$ which couples the two bands. This coupling is described by the off-diagonal terms in the two-band basis [103]:

$$H_{12} = H_{12}^0 - D_{xy}\varepsilon_{xy}, \tag{3}$$

where $D_{xy} = 14$ eV denotes the shear deformation potential, and

$$H_{12}^0(\mathbf{k}) = \frac{\hbar^2 k_x k_y}{M}$$

$k_x$, $k_y$ are the transversal wave vector components. The parameter $M$ can be evaluated from the $\mathbf{k} \cdot \mathbf{p}$ perturbation theory, with its numerical value close to $M \approx m_t/(1 - m_t/m_0)$ reported in [103].

The two-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian we consider is therefore written in the form [104]: The Hamiltonian of the valley pairs along the [001]-axis is then given as [105]

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}. \tag{4}$$
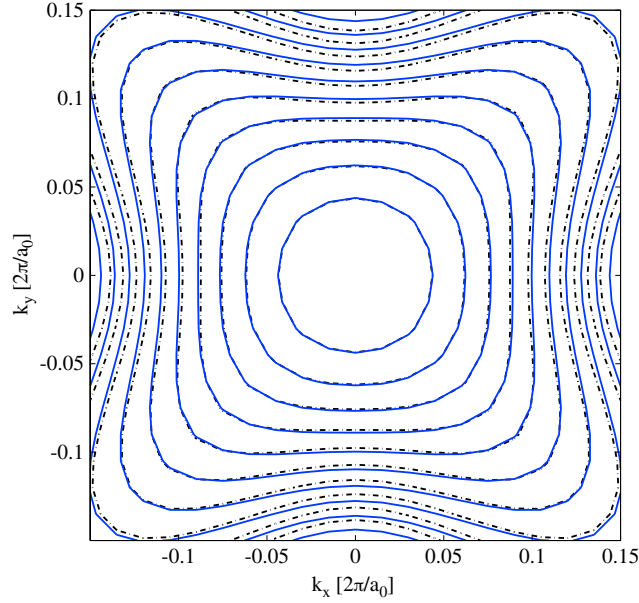
**Fig. 1.** Comparison of the dispersion relation (6) at the valley minimum (dashed–dotted contour lines) with the EPM results (solid lines). The distance between the equi-energy contour lines is 50 meV (cf. [104]).

The strain-induced shift of the conduction band minimum $\delta E_C$ has to be included in the diagonal elements for completeness by the substitution $H_{ii} = H_{ii} + \delta E_C$. Here, $\delta E_C$ is defined as [106]

$$\delta E_C = D_d(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) + D_u \varepsilon_{zz}, \qquad (5)$$

with $D_d$ denoting the dilation and $D_u$ the uniaxial deformation potentials for the conduction band and $\varepsilon_{jj}, j = x, y, z$ being the diagonal elements of the strain tensor.

Considering the off-diagonal elements $H_{12}$ as a perturbation, one obtains the following dispersion relations for the [001] valleys including the shear strain component:

$$E(\mathbf{k}) = \frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2(k_x^2 + k_y^2)}{2m_t} + \delta E_C \pm \sqrt{\left(\frac{\hbar}{m_0}k_z p\right)^2 + \left(D_{xy}\varepsilon_{xy} - \frac{\hbar^2 k_x k_y}{M}\right)^2}. \qquad (6)$$

In unstrained silicon $(\varepsilon_{xy} = 0)$ the dispersion relation (6) goes beyond the parabolic approximation. It allows to reproduce correctly the isotropic non-parabolic correction [104]. Even more, it describes correctly the conduction band warping seen in numerical band structure calculations. A comparison to the results of the empirical pseudopotential simulations of the in-plane dispersion displays good agreement as shown in Fig. 1.

With uniaxial stress along [110] direction introduced, the dependence of the effective mass on shear strain follows from (6). Defining a dimensionless variable $\eta = 2D_{xy}\varepsilon_{xy}/\Delta$ with $\Delta = 0.53$ eV being the energy difference between the two lowest conduction bands in unstrained silicon at the valley minimum, the dependence of the effective mass on shear strain is [107]:

$$m_{t1}(\eta) = m_t \begin{cases} \left(1 - \eta\frac{m_t}{M}\right)^{-1}, & |\eta| < 1 \\ \left(1 - \text{sgn}(\eta)\frac{m_t}{M}\right)^{-1}, & |\eta| > 1 \end{cases} \qquad (7)$$

$\text{sgn}(\eta)$ denotes the sign function. The shear strain component is positive for tensile and negative for compressive uniaxial [110] stress. The effective mass depends linearly on the shear strain component, in contrast to an assessment in [106] that within the linear deformation potential theory there should not exist any warping of the conduction band. A comparison of the theoretical effective mass dependences on strain with that extracted from the mobility measurement in strained samples [108] shows excellent agreement [104].

### 2.3.2. Conduction band Hamiltonian including spin

The two-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian (4) is sufficient to deal with charge transport properties, for instance, to explain a linear dependence of the effective mass on strain followed from the mobility enhancement measurements in uniaxially stressed silicon. In order to handle spin-dependent characteristics including the spin relaxation time, the spin degree of

freedom has to be taken into consideration. By choosing the spin projection axis along [001] direction, we augment the two lowest conduction band basis by including the two spin projections, positive and negative, on the given axis. The basis is conveniently chosen as $[X_{1\uparrow}, X_{1\downarrow}, X_{2'\uparrow}, X_{2'\downarrow}]$, where $\uparrow$ and $\downarrow$ indicate the spin projection at the quantization $z$-axis. The Hamiltonian of the valley pairs along the [001]-axis including spin is then written as [35,105]

$$H = \begin{bmatrix} H_1 & H_3 \\ H_3^\dagger & H_2 \end{bmatrix}, \tag{8}$$

where $H_1, H_2$, and $H_3$ are

$$H_i = \left[ \frac{\hbar^2 k_z^2}{2 m_l} - \frac{(-1)^i \hbar^2 k_0 k_z}{m_l} + \frac{\hbar^2 (k_x^2 + k_y^2)}{2 m_t} \right] I, \tag{9}$$

$$H_3 = \begin{bmatrix} D_{xy} \varepsilon_{xy} - \dfrac{\hbar^2 k_x k_y}{M} & (k_y - k_x \, \mathrm{i}) \, \Delta_{so} \\ (-k_y - k_x \, \mathrm{i}) \, \Delta_{so} & D_{xy} \varepsilon_{xy} - \dfrac{\hbar^2 k_x k_y}{M} \end{bmatrix}. \tag{10}$$

Here, $I$ is the identity $2 \times 2$ matrix. In the Hamiltonian (8) $k_0 = 0.15 \times 2\pi / a_0$ is the position of the valley minimum relative to the $X$-point in unstrained silicon.

The inclusion of the spin degree of freedom results in the appearance of the effective spin–orbit term $\tau_y \otimes \Delta_{so}(k_x\sigma_x - k_y\sigma_y)$ with

$$\Delta_{so} = \frac{\hbar^2}{2m_0^3 c^2} \left| \sum_n \frac{\langle X_1|p_j|n\rangle \langle n|[\nabla V \times \mathbf{p}]_j|X_{2'}\rangle}{E_n - E_X} \right|. \tag{11}$$

This effective spin–orbit term couples the states with the opposite spin projections from the opposite valleys. In (11) $E_n$ is the energy of the $n$th band at the $X$-point, $E_X$ is the energy of the two lowest conduction band $X_1$ (or $X_2'$) at the $X$-point, $\mathbf{p}$ is the momentum operator, $V$ is the bulk crystal potential, $\sigma_x, \sigma_y$, and $\sigma_z$ are the spin Pauli matrices, $\tau_y$ is the $y$-Pauli matrix in the valley degree of freedom, and $c$ is the speed of light. We stress that the spin–orbit term $\tau_y \otimes \Delta_{so}(k_x\sigma_x - k_y\sigma_y)$ is caused by the intrinsic spin–orbit interaction due to an electron spin moving in the crystal potential $V$ and must be distinguished from the electric field-dependent spin–orbit coupling of the Rashba type used to manipulate the electron spin under the gate in a SpinFET channel. Indeed, the spin–orbit interaction $\tau_y \otimes \Delta_{so}(k_x\sigma_x - k_y\sigma_y)$ does not lift the degeneracy between the energies of electrons moving with the same velocity but having opposite spin projections, in contracts to the electric field-dependent spin–orbit coupling responsible for spin manipulation.

The energy dispersion of the two lowest spin degenerate conduction bands is thus given by

$$E(k) = \frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2 (k_x^2 + k_y^2)}{2m_t} \pm \sqrt{\left( \frac{\hbar^2 k_z k_0}{m_l} \right)^2 + B^2}, \tag{12}$$

where

$$B = \sqrt{\Delta_{so}^2 \left(k_x^2 + k_y^2\right) + \left( D_{xy}\varepsilon_{xy} - \frac{\hbar^2 k_x k_y}{M} \right)^2}. \tag{13}$$

This expression generalizes the corresponding dispersion relation [35] by including shear strain [105].

The dispersion relation (12) allows to calculate the strength $\Delta_{so}$ of the effective spin–orbit interaction. Indeed, close to the $X$-point in an unstrained sample the gap between the $X_1$ and $X_2'$ conduction bands can now be opened by $\Delta_{so}$ alone, if one evaluates the dispersion for $k_x \neq 0$ but $k_y = k_z = 0$. The band splitting along the $x$-axis is then equal to $2 |\Delta_{so}k_x|$. By using the empirical pseudopotential method (EPM) [104,107] we found $\Delta_{so} = 1.27$ meV nm [105], in good agreement with the one reported in [35].

The Hamiltonian (8) is sufficiently rigorous to capture all physical phenomena needed. Indeed, apart from the conduction band non-parabolicity and warping under external stress, it includes the effective spin–orbit interaction [35]. The spin–orbit interaction is responsible for spin flips, even when the electrons scatter on non-magnetic potentials. It is the main ingredient of the Elliott–Yafet spin relaxation mechanism [18]. At the same time, the effective Hamiltonian (8) is relatively simple compared to those required for tight-binding or empirical pseudo-potential calculations. It can be used to evaluate the energy spectrum and the wave functions in confined systems while still allowing to keep the computational efforts bearable. The wave functions are the main ingredients to evaluate the scattering matrix elements and relaxation rates in confined systems.

### 2.3.3. Wave function calculations: Analytical solution for an infinite box potential

The Hamiltonian (8) is used to describe the electron spectrum and subbands in a confined electron system. After supplementing the diagonal parts with the confinement potential $U(\mathbf{r})$ as

$$H_i = H_i + U(\mathbf{r})I, \tag{14}$$

the corresponding Schrödinger equations for the wave functions $\Psi_\sigma(\mathbf{r})$ are obtained. For definitiveness let us consider a (001) oriented silicon film for which the confinement potential depends only on $z$ and the wave function can be written as

$$\Psi_\sigma(\mathbf{r}) = \exp(i\mathbf{K}\mathbf{R})\Psi_{n\mathbf{K}\sigma}(z), \tag{15}$$

where $\Psi_{n\mathbf{K}\sigma}(z)$ is the confined wave function describing the subband $n$ with the spin projection $\sigma$. Because of the non-parabolicity of the Hamiltonian (8) $\Psi_{n\mathbf{K}\sigma}(z)$ also depends on the in-plane wave vector $\mathbf{K}$.

This dependence, and a double integration with respect to the in-plane wave vector $\mathbf{K}$ present in the scattering rates, makes the calculation of the spin relaxation time with the numerically determined wave functions prohibitively expensive. It is mandatory to reduce the simulation time by obtaining the wave functions in an analytical form.

This can be achieved, if the confinement potential $U(z)$ is approximated by an infinite square well potential of width $t$. The wave functions are then localized within the well and vanish at the interfaces: $\Psi_{n\mathbf{K}\sigma}(\pm t/2) = 0$. First let us consider the unprimed subbands' ladder in a (001) silicon film originating from the [001] valleys described by (8). It is convenient to perform a unitary transformation on the basis functions [109]:

$$\Xi_{s\pm} = \frac{1}{2}\left[ \left( X_{1\uparrow} + X'_{2\uparrow} \right) \pm \left( X_{1\downarrow} + X'_{2\downarrow} \right) \frac{k_x - ik_y}{\sqrt{k_x^2 + k_y^2}} \right] \tag{16}$$

$$\Xi_{a\pm} = \frac{1}{2}\left[ \left( X_{1\uparrow} - X'_{2\uparrow} \right) \pm \left( X_{1\downarrow} - X'_{2\downarrow} \right) \frac{k_x - ik_y}{\sqrt{k_x^2 + k_y^2}} \right] \tag{17}$$

$$Y_{1,2} = \Xi_{s\pm} \cos\left(\frac{\Theta}{2}\right) \mp i\Xi_{a\pm} \sin\left(\frac{\Theta}{2}\right) \tag{18}$$

$$Y_{3,4} = \Xi_{a\pm} \cos\left(\frac{\Theta}{2}\right) \mp i\Xi_{s\pm} \sin\left(\frac{\Theta}{2}\right) \tag{19}$$

$$\tan(\Theta) = \frac{\Delta_{SO}\sqrt{k_x^2 + k_y^2}}{D\varepsilon_{xy} - \frac{\hbar^2 k_x k_y}{M}}. \tag{20}$$

The Hamiltonian (8) in the new basis transforms into

$$\tilde{H} = \begin{bmatrix} \tilde{H}_1 & \tilde{H}_3 \\ \tilde{H}_3 & \tilde{H}_2 \end{bmatrix}, \tag{21}$$

with

$$\tilde{H}_{1,2} = \left[ \frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2 \left( k_x^2 + k_y^2 \right)}{2m_t} + U(z) \pm B \right] I, \tag{22}$$

$$\tilde{H}_3 = \begin{bmatrix} \dfrac{\hbar^2 k_0 k_z}{m_l} & 0 \\ 0 & \dfrac{\hbar^2 k_0 k_z}{m_l} \end{bmatrix} \tag{23}$$

and splits into two equivalent $2\times2$ Hamiltonians describing the two pairs of coupled states. The coupled states within each pair are with opposite spins along the spin–orbit field axis $(k_x, -k_y)$. The states from different pairs with the same energy have opposite spin projections on $(k_x, -k_y)$. Thus, the spin degeneracy is preserved for arbitrary $(k_x, k_y)$ and in presence of strain $\varepsilon_{xy}$.

A linear combination of the two degenerate states of the same energy but opposite spins $\Psi_{j\mathbf{K}\sigma}$ and $\Psi_{j\mathbf{K}-\sigma}$ allows creating the wave function with a spin along any arbitrarily chosen axis. If the direction $\Uparrow$ of the axis is defined by the polar angles $\theta$ and $\phi$, the coefficients in the linear combination must be chosen such that the resulting wave function

$$\begin{aligned} |j \Uparrow\rangle &= \Psi_{j\mathbf{K}\Uparrow}, \\ |j \Downarrow\rangle &= \Psi_{j\mathbf{K}\Downarrow} \end{aligned} \tag{24}$$

satisfies [35]

$$\begin{aligned} \langle j \Uparrow |F| j \Downarrow\rangle &= 0, \\ \langle j \Uparrow |F| j \Uparrow\rangle &> 0, \end{aligned} \tag{25}$$
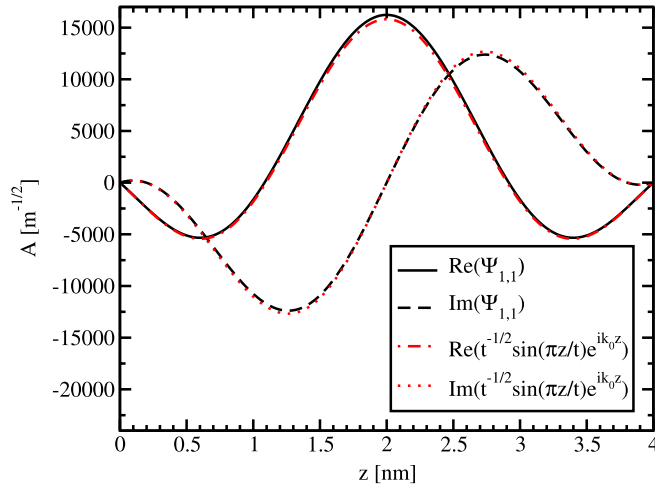
**Fig. 2.** The large component of the wave function of the lowest unprimed subband in an unstrained film located in the valley centered at $k_0$ (cf. [109]).

where

$$F = \sigma_z \cos(\theta) + (\sigma_x \cos(\phi) + \sigma_y \sin(\phi)) \sin(\theta). \tag{26}$$

For definitiveness let us consider the wave functions with the spin quantization axis along [001], or $z$ direction. The wave functions fulfilling the Schrödinger equation with (21) and satisfying the zero boundary conditions at the interfaces can be written in a closed form [104].

The four-component wave functions in the initial two valley-two spin projection basis are now straightforwardly obtained. These four functions from the lowest unprimed subband are ($k_x = 0$):

$$\Psi_{1\Uparrow} = \begin{pmatrix} \Psi_{1,1} \\ \Psi_{1,2} \\ \Psi_{1,1}^* \\ -\Psi_{1,2}^* \end{pmatrix}, \qquad \Psi_{1\Downarrow} = \begin{pmatrix} -\Psi_{1,2} \\ \Psi_{1,1} \\ \Psi_{1,2}^* \\ \Psi_{1,1}^* \end{pmatrix},$$

$$\Psi_{2\Uparrow} = \begin{pmatrix} \Psi_{2,2} \\ \Psi_{2,1} \\ -\Psi_{2,2}^* \\ \Psi_{2,1}^* \end{pmatrix}, \qquad \Psi_{2\Downarrow} = \begin{pmatrix} -\Psi_{2,1} \\ \Psi_{2,2} \\ -\Psi_{2,1}^* \\ -\Psi_{2,2}^* \end{pmatrix}. \tag{27}$$

In unstrained films the dominant components denoted as $\Psi_{1,1}$ and $\Psi_{2,2}$ for $\Psi_1$ and $\Psi_2$, respectively, are well described by

$$\Psi_{1,1(2,2)} = e^{ik_0 z} \sin\left(\frac{\pi z}{t}\right), \tag{28}$$

as shown in Fig. 2. The wave function (28) is the usual envelope quantization function located at the valley minimum $k_0 = 0.15 \frac{2\pi}{a}$. In the presence of shear strain $\varepsilon_{xy}$ and confinement the degeneracy between the two valleys is lifted. This results in slightly different envelope functions $\Psi_{1,1}$ and $\Psi_{2,2}$ (Fig. 3).

The small components of the four-components' wave function are proportional to the spin–orbit interaction strength. The amplitude of these components shown in Fig. 4 for an unstrained film of 4 nm thickness for $k_x = 0$ depends strongly on $k_y$. For $k_y = 1$ nm$^{-1}$ the small components of the wave functions are pronounced, while decreasing the $k_y$ value makes the small components vanishing.

Shear strain $\varepsilon_{xy}$ considerably suppresses the small components as shown in Fig. 5. $\Psi_{1,2}$ for a strain value of 1% is almost washed out, while in an unstrained film the wave function component is significant (Fig. 5). Vanishing values of the small components indicate the decrease in the spin mixing between the states with the opposite spin projections, which should result in a longer spin lifetime.

### 2.3.4. Valley splitting by shear strain

It is usually assumed that, because the unprimed subbands are originating from the two equivalent [001] valleys, they are double (or four-fold with spin included) degenerate [97]. However, this is true only in the parabolic band approximation, when the two valleys are independent. Due to the presence of the off-diagonal terms $H_3$, the Hamiltonian (8) couples the [001] valleys. This coupling results in an unprimed subband degeneracy lifting in confined systems, which is termed the valley splitting [97]. We note that the bulk valleys remain degenerate: However, the presence of the (001) confinement
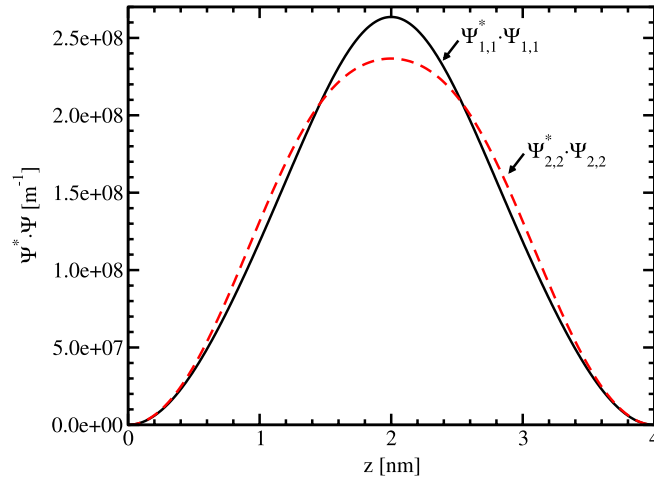
**Fig. 3.** The large components of the two unprimed subbands with $\varepsilon_{xy} = 0.05\%$ (cf. [109]).
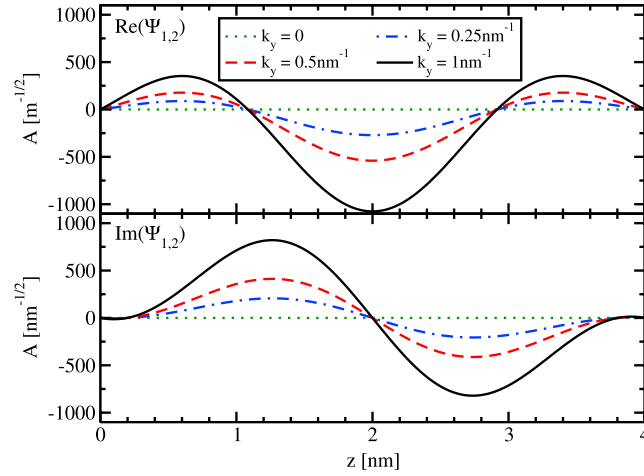


**Fig. 4.** The small components are proportional to the strength of the spin–orbit interaction (cf. [109]).

and/or interfaces combined with the off-diagonal valley coupling results in the degeneracy lifting of the unprimed subband ladder.

The $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian (8) allows to obtain the unprimed subband dispersion relations. In case of the infinite square well potential the dispersion relations can be cast in accurate analytical forms [104,105]. The valley splitting $\Delta E$ is approximately described by the following expression [104,105]:

$$\Delta E = \frac{2y^2 B}{k_0 t \sqrt{\left(1 - y^2 - \eta^2\right)\left(1 - y^2\right)}} \left| \sin\left( \sqrt{\frac{1 - y^2 - \eta^2}{1 - y^2}} k_0 t \right) \right|. \tag{29}$$

$$y = \frac{\pi}{k_0 t}, \tag{30}$$

$$\eta = \frac{m_l B}{k_0^2 \, \hbar^2}, \tag{31}$$

$t$ is the film thickness. To obtain (29) the theory [104] for valley splitting has been generalized by including the spin–orbit coupling [105]. The valley splitting is inversely proportional to the third power of the position of the minimum $k_0$, in agreement with earlier considerations [97].

The degeneracy between the subbands is exactly recovered, when the oscillation $\sin\left( \sqrt{\frac{1 - y^2 - \eta^2}{1 - y^2}} k_0 t \right)$ term is zero. However, this degeneracy is insignificant, because it does not result in any peculiar behavior of the spin relaxation scattering matrix elements. In recent experimental studies of the valley splitting and spin relaxation in quantum dots, however, an
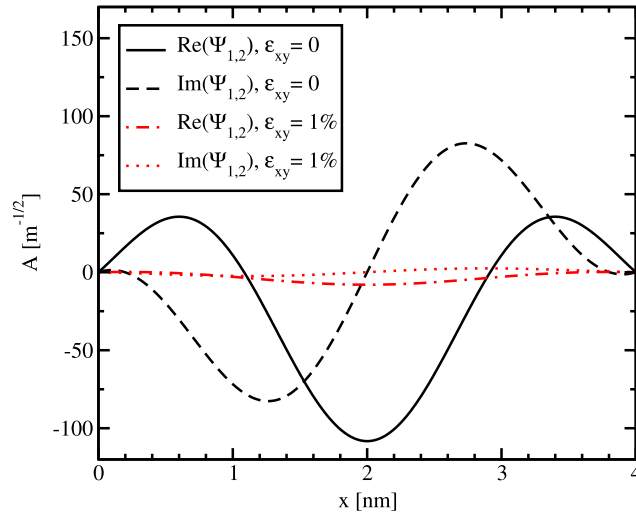
**Fig. 5.** The small components are considerably suppressed by tensile shear strain (cf. [109]).

enhancement of spin relaxation at the exact valley degeneracy point tuned by the magnetic field was observed [110]. As it has been shown, this can be explained by an additional spin–orbit coupling field not accounted in the model (8), the physical nature of which is not understood [110].

In contrast, the minimum of the $B$ term in (29) reveals a very strong increase of the intersubband spin relaxation. Under these conditions the subband splitting is purely determined by the effective spin–orbit interaction term and is linear in $\Delta_{SO}\sqrt{k_x^2 + k_y^2}$. This linear dependence of the splitting is similar to the Zeeman splitting in a magnetic field. Therefore, the spin–orbit interaction term $\Delta_{SO}\mathbf{k}$ with $\mathbf{k} = (k_x, -k_y)$ can be interpreted as an effective magnetic field, while the states within the pair $\Psi_{1\mathbf{k}+}$, $\Psi_{2\mathbf{k}-}$ and $\Psi_{1\mathbf{k}-}$, $\Psi_{2\mathbf{k}+}$ it couples have similarities with the Zeeman spin-up, spin-down states split by the effective field. If we now inject spin along the $z$-direction, it will precess in the in-plane effective magnetic field $\Delta_{SO}\mathbf{k}$. This results in a large mixing between the opposite spin states, which causes a large relaxation. Therefore, the minimum of $B$ in (29) defines the conditions for spin relaxation hot spots which should be contrasted against the hot spots in bulk silicon [35], appearing far away from the valley minima at the edge of the three-dimensional Brillouin zone. The origin of the spin relaxation hot spots in thin films lies in the unprimed subband degeneracy in a confined electron system which is lifted by the effective spin–orbit interaction alone. Because the hot spots are determined by the minimum of $B$, they are located in the middle of the two-dimensional Brillouin zone in an unstrained film. They contribute strongly to the spin relaxation, however, when shear strain is applied, the spin relaxation hot spots are moved towards higher energies. Pushing the hot spots above the Fermi energy outside the occupied states region will result in a considerably reduced spin relaxation and an increase of the spin lifetime with shear strain.

The expression for the valley splitting (29) accounts only for a contribution due to quantum interference of the wave functions belonging to different valleys in a confined system. The interference is most influenced by the coupling between the valleys through the $X$-point, which can be easily enhanced by shear strain $\varepsilon_{xy}$. According to (29), the splitting is proportional to the gap opened at the $X$-point between the otherwise degenerate states $X_1$ and $X'_2$ by the off-diagonal term $H_3$. Other contributions to the valley splitting, for instance, due to surface effects [114,115] can be taken into account. Let us consider the valley splitting in unstrained films attributed to the [001] equivalent valley coupling through the $\Gamma$-point [97].

The values of the valley splitting in an unstrained silicon film from a 30-band $\mathbf{k} \cdot \mathbf{p}$ model [111], an atomistic tight-binding model from [112], and from [113] are summarized in Fig. 6. Although looking irregular, the results follow a certain law. Fig. 7 demonstrates a good agreement of the results of the tight-binding calculations with an analytical expression for the subband splitting similar to (29):

$$\Lambda_\Gamma = \frac{2\pi \Delta_\Gamma}{(K_{0\Gamma}t)^3} \sin(K_{0\Gamma}t). \tag{32}$$

Here $\Delta_\Gamma$ is the splitting at the $\Gamma$-point, $K_{0\Gamma} = 0.85\frac{2\pi}{a_0}$. A good agreement is found for $\Delta_\Gamma = 2$ eV. Then the total valley splitting is given by [105]

$$\Delta E_\Gamma = \sqrt{\Delta E^2 + \Lambda_\Gamma^2} \tag{33}$$
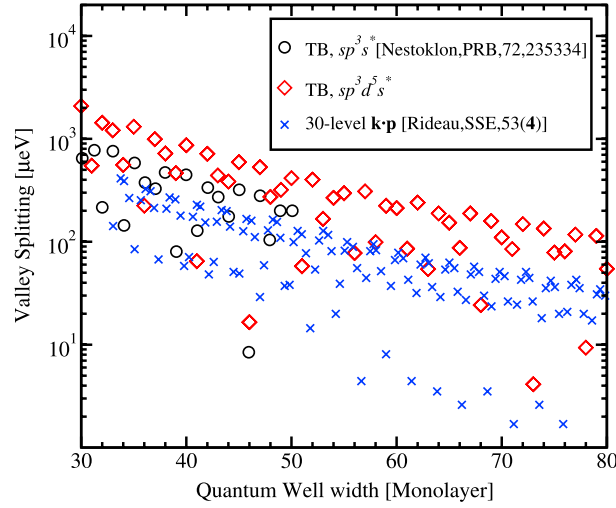
with $\Delta E$ from (29).

**Fig. 6.** Valley splitting in a Si quantum well at zero strain as a function of the quantum well from [111–113] (cf. [113]).
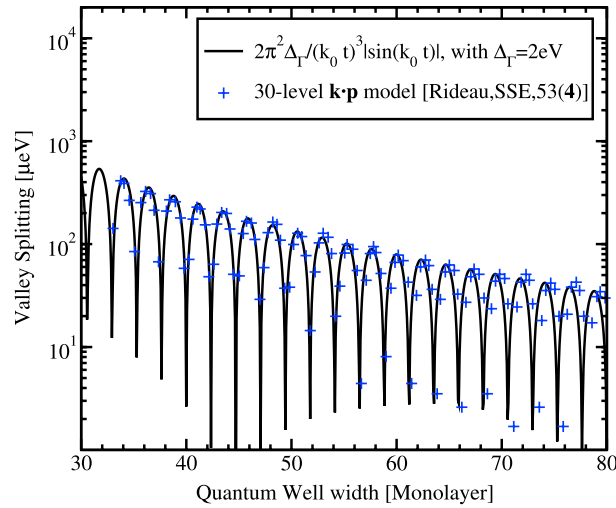


**Fig. 7.** Dependence of the valley splitting on the quantum well width from the 30 band $\mathbf{k} \cdot \mathbf{p}$ model [111] and the analytical expression with $\Lambda_\Gamma = 2.0$ eV (cf. [113]).

Fig. 8 shows the dependence of the lowest unprimed subbands' energies on shear strain with and without accounting for the $\Lambda_\Gamma$ term. The $\Lambda_\Gamma$ term lifts the subband degeneracy in unstrained films, while shear strain gives the major contribution to the splitting at higher strain values.

### 2.3.5. Electron–phonon and surface roughness induced spin relaxation rates

In the following, we briefly outline the calculation of the spin relaxation time. The spin and momentum relaxation times are evaluated by thermal averaging [35,95,116] as

$$\frac{1}{\tau_S} = \frac{\sum_n \int \frac{1}{\tau_n(\mathbf{K_n})} f(E)(1 - f(E)) \, d\mathbf{K_n}}{\int \int f(E) \, d\mathbf{K_n}}. \tag{34}$$

Here $f(E)$ is the Fermi–Dirac distribution function:

$$f(E) = \frac{1}{1 + \exp\left(\frac{E - \mu}{k_B T}\right)}. \tag{35}$$

By using the subband dispersion

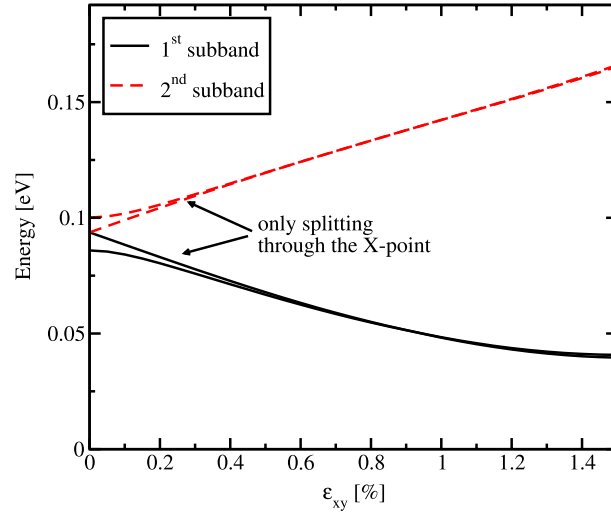$$E(\mathbf{K}_n) = E_n^0 + E_n(\mathbf{K}_n) \tag{36}$$

**Fig. 8.** Dependence of the energy of the 1st and the 2nd subbands together with the subband splitting on shear strain for the film thickness 2.1 nm (cf. [109]).

one obtains [116]

$$\int d\mathbf{K}_n = \int_0^{2\pi} \int_{E_n^0}^{\infty} \frac{K_n(E, \varphi)}{\left| \frac{\partial E_n}{\partial K_n} \right|_{K_n(E,\varphi)}} d\varphi dE, \tag{37}$$

where $K_n(E, \varphi)$ is the absolute value of the subband wave vector:

$$\mathbf{K}_n(E, \varphi) = (K_n(E, \varphi) \cos(\varphi), K_n(E, \varphi) \sin(\varphi)). \tag{38}$$

The partial derivative is taken along the direction of constant $\varphi$ for a given energy value $E$ (36).

We take into account the main mechanisms determining the mobility in thin silicon films, namely surface roughness and electron–phonon scattering, and analyze their role in spin relaxation. In contrast to the corresponding momentum relaxation rates in surface layers [97,116,117] and thin films, the spin relaxation rates are not available in the literature and are therefore summarized below. The surface roughness scattering matrix elements are proportional to the product of the subband function derivatives at the interface [97]. The surface roughness at the two interfaces is assumed to be equal and statistically independent. The surface roughness spin relaxation rate is calculated in the following way:

$$\frac{1}{\tau_n^{SR}(\mathbf{K}_n)} = \frac{4\pi}{\hbar (2\pi)^2} \sum_{ij} \int_0^{2\pi} d\varphi \pi \Delta^2 L^2 \frac{1}{\epsilon_{ij}^2 (\mathbf{K}_j - \mathbf{K}_n)} \frac{\hbar^4}{4m_l^2} \frac{K_j(E_j, \varphi)}{\left| \frac{\partial E_j}{\partial K_j} \right|_{K_j(E,\varphi)}} \left[ \left( \frac{d\Psi_{i\mathbf{K}_n\sigma}}{dz} \right)^* \left( \frac{d\Psi_{j\mathbf{K}_j - \sigma}}{dz} \right) \right]_{z=\pm\frac{t}{2}}^2$$

$$\times \exp \left( \frac{- (\mathbf{K}_j - \mathbf{K}_n) L^2}{4} \right) \Theta (E_j(\mathbf{K}_j) - E_j^0). \tag{39}$$

Here $E$ is the electron energy, $\mathbf{K}_n$, $\mathbf{K}_j$ are the in-plane wave vectors before and after scattering, $\varphi$ is the angle between $\mathbf{K}_n$ and $\mathbf{K}_j$, $\epsilon$ is the dielectric permittivity, $L$ is the autocorrelation length, $\Delta$ is the mean square value of the surface roughness fluctuations, $\Psi_{n\mathbf{K}\sigma}$ and $\Psi_{j\mathbf{K}\sigma}$ are the four components wave functions, $\sigma = \pm 1$ is the spin projection to the $\Uparrow$-axis chosen to be along [001], and $\Theta(x)$ is the Heaviside function [109].

The spin relaxation rate due to the electron–phonon interaction with acoustic phonons is considered in the deformation potential. By analogy to the momentum relaxation the corresponding spin relaxation rates can be written as [118]

$$\frac{1}{\tau_n^{PH}(\mathbf{K}_n)} = \sum_{\nu} \frac{2k_BT}{\hbar \rho \upsilon_\nu^2} \sum_j \int_0^{2\pi} \frac{d\varphi}{2\pi} \int_{\infty}^{\infty} \frac{dq_z}{2\pi} \frac{K_j(E_j, \varphi)}{\left| \frac{\partial E_j}{\partial K_j} \right|_{K_j(E,\varphi)}} \sum_{\alpha\beta} \left| \frac{q_\alpha}{q} e_\beta^{(\nu)}(\mathbf{q}) \right|^2$$

$$\times \left[ \int_0^t dz \exp (-iq_z z) \Psi_{j\mathbf{K}_j - \sigma}^\dagger (z) \mathbf{D}_{\alpha\beta} \Psi_{n\mathbf{K}_n\sigma} (z) \right]^2 \Theta (E_j(\mathbf{K}_j) - E_j^0), \tag{40}$$

where $\rho = 2329 \frac{kg}{m^3}$ is the silicon density, $\upsilon_\nu$ and $e^\beta$ are the $\nu$-branch phonon velocity and polarization, $(q_x, q_y) = \mathbf{K}_n - \mathbf{K}_j$, and $\mathbf{D}$ is the 4×4 deformation potential matrix in the basis of the states at the $X$-point. In the case of the isotropic deformation

potential approximation which is valid for low field electron mobility calculations, only the interaction with the longitudinal acoustic phonons described by the phonon polarization

$$e^{\beta} = q_{\beta}/q, \quad \beta = x, y, z \tag{41}$$

is considered. In this case

$$\mathbf{D}_{\alpha\beta} = \delta_{\alpha\beta}\mathbf{D}, \tag{42}$$

and the matrix $\mathbf{D}$ is diagonal in the basis of the states at the $X$-point,

$$\mathbf{D} = \begin{bmatrix} D & 0 & 0 & 0 \\ 0 & D & 0 & 0 \\ 0 & 0 & D & 0 \\ 0 & 0 & 0 & D \end{bmatrix}. \tag{43}$$

In order to reproduce the low field electron mobility in surface layers and thin films, the value $D$ of the isotropic acoustic deformation potential is taken slightly above its value in bulk silicon, usually around 12–14 eV.

After performing the integration with respect to $q_z$ one gets:

$$\frac{1}{\tau_{LA}(\mathbf{K}_n)} = \frac{4\pi k_B T}{\hbar \rho v_{LA}^2} \sum \int_0^{2\pi} d\varphi \frac{1}{4\pi} \frac{K_j(E_j, \varphi)}{\left| \frac{\partial E_j}{\partial K_j} \right|_{K_j(E,\varphi)}} \int_0^t dz \left| \Psi_{j\mathbf{K}_{j-\sigma}}^{\dagger}(z) \mathbf{D} \Psi_{n\mathbf{K}_n\sigma}(z) \right|^2 \Theta(E_j(\mathbf{K}_j) - E_j^0). \tag{44}$$

Another important contribution is due to the shear deformation potential $D_{xy}$, which couples the opposite valleys. In this case the deformation potential matrix is in the form

$$\mathbf{D}_{xy} = \begin{bmatrix} 0 & 0 & \dfrac{D_{xy}}{2} & 0 \\ 0 & 0 & 0 & \dfrac{D_{xy}}{2} \\ \dfrac{D_{xy}}{2} & 0 & 0 & 0 \\ 0 & \dfrac{D_{xy}}{2} & 0 & 0 \end{bmatrix}. \tag{45}$$

Performing an integration over $q_z$ one obtains for the acoustic longitudinal phonons:

$$\frac{1}{\tau_{LA}(\mathbf{K}_n)} = \frac{4\pi k_B T}{\hbar \rho v_{LA}^2} \sum_j \int_0^{2\pi} \frac{d\varphi}{4\pi} \frac{K_j(E_j, \varphi)}{\left| \frac{\partial E_j}{\partial K_j} \right|_{K_j(E,\varphi)}} \int_0^t dz \int_0^t dz' \exp\left(-\sqrt{q_x^2 + q_y^2} |z - z'|\right)$$

$$\times \left[ \Psi_{j\mathbf{K}_{j-\sigma}}^{\dagger}(z) \mathbf{D}_{xy} \Psi_{n\mathbf{K}_n\sigma}(z) \right]^* \left[ \Psi_{j\mathbf{K}_{j-\sigma}}^{\dagger}(z') \mathbf{D}_{xy} \Psi_{n\mathbf{K}_n\sigma}(z') \right]$$

$$\times \frac{4q_x^2 q_y^2}{\left(\sqrt{q_x^2 + q_y^2}\right)^3} \left[ \sqrt{q_x^2 + q_y^2} |z - z'| + 1 \right] \Theta(E_j(\mathbf{K}_j) - E_j^0), \tag{46}$$

where $v_{LA} = 8700 \frac{\text{m}}{\text{s}}$ is the speed of the longitudinal phonons.

For the two branches of the transversal acoustic phonons the corresponding expression is:

$$\frac{1}{\tau_{TA}(\mathbf{K}_n)} = \frac{4\pi k_B T}{\hbar \rho v_{TA}^2} \sum_j \int_0^{2\pi} \frac{d\varphi}{4\pi} \frac{K_j(E_j, \varphi)}{\left| \frac{\partial E_j}{\partial K_j} \right|_{K_j(E,\varphi)}} \int_0^t dz \int_0^t dz' \exp\left(-\sqrt{q_x^2 + q_y^2} |z - z'|\right)$$

$$\times \left[ \Psi_{j\mathbf{K}_{j-\sigma}}^{\dagger}(z) \mathbf{D}_{xy} \Psi_{n\mathbf{K}_n\sigma}(z) \right]^* \left[ \Psi_{j\mathbf{K}_{j-\sigma}}^{\dagger}(z') \mathbf{D}_{xy} \Psi_{n\mathbf{K}_n\sigma}(z') \right]$$

$$\times \left[ \sqrt{q_x^2 + q_y^2} - \frac{8q_x^2 q_y^2 - \left(q_x^2 + q_y^2\right)^2}{q_x^2 + q_y^2} |z - z'| \right] \Theta(E_j(\mathbf{K}_j) - E_j^0), \tag{47}$$

where $v_{TA} = 5300 \frac{\text{m}}{\text{s}}$ is the speed of the transversal phonons.

The Yafet type contribution due to the spin–orbit interaction term modified by stress has to be included [95]. In the isotropic approximation the corresponding term reads:

$$\frac{1}{\tau_Y(\mathbf{K}_n)} = \frac{4\pi k_B T}{\hbar \rho v_{LA}^2} \sum_j \int_0^{2\pi} \frac{d\varphi}{4\pi} \frac{K_j(E_j, \varphi)}{\left| \frac{\partial E_j}{\partial K_j} \right|_{K_j(E,\varphi)}} \int_0^t dz \left| \Psi_{j\mathbf{K}_{j-\sigma}}^{\dagger}(z) \mathbf{D}_Y \Psi_{n\mathbf{K}_n\sigma}(z) \right|^2 \cdot \Theta(E_j(\mathbf{K}_j) - E_j^0). \tag{48}$$
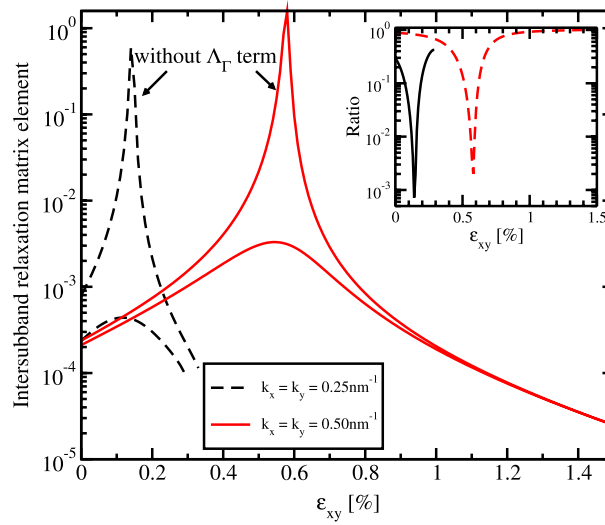
**Fig. 9.** Dependence of the normalized intersubband relaxation matrix elements on shear strain for a film thickness of 2.1 nm. The inset shows the ratio of the matrix elements with the $\Lambda_\Gamma$ term to the matrix elements without the $\Lambda_\Gamma$ term (cf. [109]).

Here the matrix $\mathbf{D}_Y$ is written as:

$$\mathbf{D}_Y = \begin{bmatrix} 0 & D_Y^\dagger \\ D_Y & 0 \end{bmatrix} \tag{49}$$

with $D_Y$

$$D_Y = \begin{bmatrix} 0 & D_{SO}\left(-r_y + i r_x\right) \\ D_{SO}\left(r_y + i r_x\right) & 0 \end{bmatrix}. \tag{50}$$

$(r_y, r_x) = \mathbf{K_1} + \mathbf{K_2}$ and the value of $D_{SO} = 15$ meV/$k_0$ is determined from the pseudopotential calculations describing the dependence of the splitting due to spin–orbit interaction between the conduction bands at the $X$-point under homogeneous stress. The value deduced is close to the one reported by [95].

The remaining entries of the deformation potential are zero: They correspond to the opposite spin coupling within the same valley.

### 2.3.6. Spin lifetime enhancement by shear strain

The scattering rates due to surface roughness scattering and electron–phonon interactions are evaluated in the following. In contrast to the mobility calculations, for which it is sufficient to evaluate the corresponding rates by using the wave functions only at the subband minima at $\mathbf{K}_j = 0$ [116], the in-plane momentum of the wave functions' dependence must be preserved for the spin relaxation calculations. Indeed, at $\mathbf{K}_j = 0$, the small components in the subband wave functions (27) are zero, and spin-flips are absent. The knowledge of the scattering rates on a fine mesh in the momentum space needed for an accurate double integration makes the spin lifetime computations (34) extremely time consuming even when the wave functions are found analytically, and an extensive code parallelization is required [109].

Fig. 9 shows the surface roughness scattering induced intersubband spin relaxation matrix elements for $\mathbf{K}_i = -\mathbf{K}_j$. Without the $\Delta_\Gamma$ term responsible for the valley splitting in unstrained films the matrix elements display a sharp maximum which visualizes the spin relaxation hot spots appearing at the condition

$$D_{xy}\varepsilon_{xy} - \frac{\hbar^2 K_x K_y}{M} = 0. \tag{51}$$

Under the condition (51) the valley splitting (29) is minimal. It is determined by the spin–orbit field alone, which results in a fairly strong spin relaxation.

The $\Delta_\Gamma$ term effectively "screens" and therefore reduces the pure spin–orbit contribution to the spin relaxation matrix elements at the hot spots. This results in an increase of the spin relaxation time in unstrained samples as displayed in Fig. 10 [109]. When a (001) film is stressed along [110] direction, a shear strain component leads to a strong increase of the spin lifetime demonstrated in Fig. 10 [109]. With the valley splitting in a relaxed film included this increase is somewhat less pronounced. However, even in this case the spin lifetime is boosted by almost two orders of magnitude.

The reason for the strong spin lifetime increase with strain lies in an effective reduction of the hot spots for spin relaxation. Indeed, with shear strain increased, the spin relaxation hot spots defined by (51) are pushed by high energy states above the
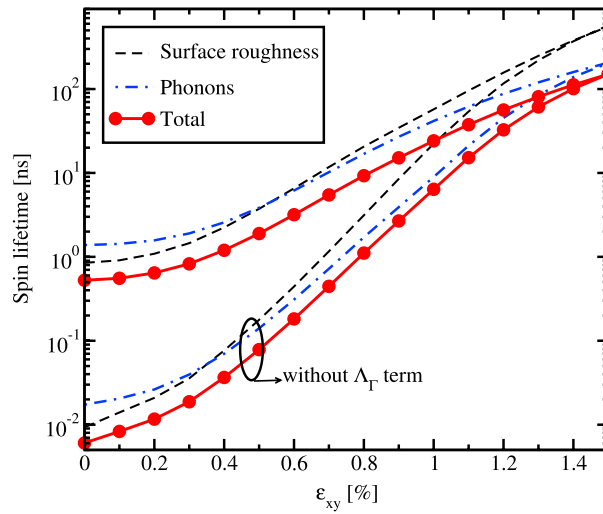
**Fig. 10.** Dependence of spin lifetime on shear strain for $T = 300$ K and film thickness 2.1 nm (cf. [109]).
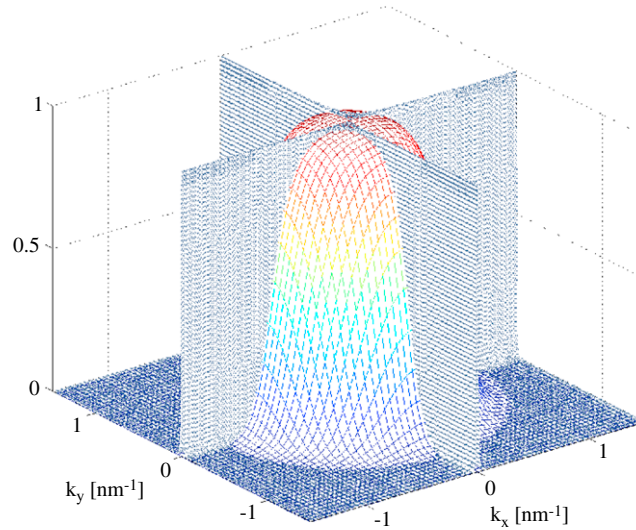


**Fig. 11.** Normalized intrasubband spin relaxation matrix elements in an unstrained film and the Fermi distribution (300 K).

Fermi level as shown in Fig. 11. Since these states are empty, they do not contribute to the spin relaxation time as follows from (34) (see Fig. 12).

## 2.4. Electric spin manipulation

Utilizing spin properties of electrons for future microelectronic devices may complement or even replace charge-based operated electric switches and memories. A textbook example of a device, where spin is used to manipulate the charge current, is the SpinFET [13]. The ferromagnetic source contact injects spin-polarized electrons into the semiconductor region. At the drain contact only the electrons with spin aligned to the drain magnetization can easily leave the channel and contribute to the current. Therefore, the total current through the device depends on the relative angle between the magnetization direction of the drain contact and the electron spin polarization at the end of the semiconductor channel. The current modulation is achieved by tuning the strength of the spin–orbit coupling in the semiconductor region. In gated systems the spin–orbit coupling is usually taken in the Rashba form [119,120], with the corresponding effective Hamiltonian

$$H_R = \alpha_R(k_x\sigma_y - k_y\sigma_x), \tag{52}$$

where $\alpha_R$ is strength of the spin–orbit coupling, $k_{x(y)}$ is the in-plane electron wave vector projection, and $\sigma_{x(y)}$ is the corresponding Pauli matrix. The spin–orbit coupling strength $\alpha_R$ depends on the effective electric field created by the gate voltage [119]. The electric field-dependent spin–orbit coupling splits the degeneracy between the spin-up and spin-down
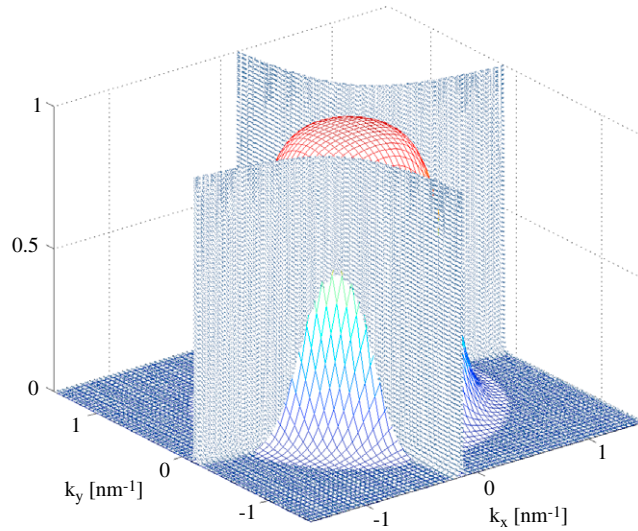
**Fig. 12.** Normalized intrasubband spin relaxation matrix elements in a film under strain, $\varepsilon_{xy} = 0.5\%$.

states which move with the same velocity in the same direction. The spin–orbit coupling can be therefore visualized as an effective magnetic field, the strength and direction of which depend on the in-plane momentum. The dependence of the spin–orbit magnetic field on the effective electric field opens a way to manipulate the degree of spin precession in the channel by purely electrical means. Indeed, the spin precession angle $\Delta\theta$ defined as the difference between the orientation of the spin of the electron at the end and at the beginning of the semiconductor region is [47]

$$\Delta\theta = \frac{2\alpha_R m^*}{\hbar^2} L, \tag{53}$$

where $m^*$ is the effective mass of the electron, and $L$ is the length of the semiconductor channel. In the absence of the spin–orbit coupling and the external magnetic field the electrons propagate with their spin orientation conserved. The strength of the spin–orbit coupling determines the minimum length of the semiconductor channel necessary to invert the orientation of the injected spin. In the case of a material with strong spin–orbit coupling such as InAs the semiconductor channel is shorter than that for a material with the weaker spin–orbit coupling such as silicon.

### 2.4.1. Electric field-dependent spin–orbit coupling

Silicon is characterized by weak intrinsic spin–orbit interaction in the conduction band and, as a consequence, it has a long spin life time. It is therefore an attractive material for spin applications. However, because of the weak spin–orbit interaction, silicon was not considered as a candidate for a SpinFET channel material. Indeed, the electric field-dependent spin–orbit coupling (52) is also expected to be small. Recently, however, it has been shown [121] that thin silicon films inside SiGe/Si/SiGe structures exhibit relatively large values of spin–orbit coupling. The reason is that in a (001) silicon film with an odd number of atomic layers the inversion symmetry is broken, and, therefore an appearance of the terms with the structure (52) or the structure similar to the Dresselhaus form with a corresponding effective Hamiltonian [121]

$$H_D = \beta(k_x\sigma_x - k_y\sigma_y) \tag{54}$$

is allowed. We note that the form (54) is similar to the effective spin–orbit coupling in surface layers of III–V materials, although the physics behind the terms is quite different: in III–V materials this Dresselhaus term is due to the absence of the inversion symmetry in the crystal lattice, while in silicon films it is due to the interface induced inversion symmetry breaking.

In actual samples with rough interfaces the inversion symmetry is always broken, even if in average the film contains an even number of atomic layers. Atomistic simulations taking into account the interfacial roughness must be carried out in order to obtain the values of the electric field dependences of the spin–orbit coupling constants $\alpha_r$ and $\beta$ in silicon [122]. It has been discovered that in Si/SiGe heterostructure the main spin orbit coupling is of the Dresselhaus type (54), with the constant $\beta$ almost linearly depending on the effective electric field [122]. It has been demonstrated that for an assumed built-in field of 50 kV/cm, the strength of the Dresselhaus spin–orbit coupling is found to be $\beta \approx 2\,\mu\text{eV}$ nm, which is in agreement with the value reported experimentally [123], while $\alpha_R \approx 0.1\,\mu\text{eV}$ nm. The value of $\beta$ is sufficient for realizing a silicon SpinFET.

The spin–orbit coupling (52) or (54) leads to an enhanced spin relaxation. The Dyakonov–Perel mechanism is the main spin relaxation mechanism in systems, where the electron dispersion curves for the two spin projections are non-degenerate [18,19]. In quasi-one-dimensional electron structures, however, a suppression of this spin relaxation mechanism

is expected [124]. Indeed, in case of elastic scattering only back-scattering is allowed. Reversal of the electron momentum results in the inversion of the effective magnetic field direction. Therefore, the precession angle does not depend on the number of scattering events along the carrier trajectory in the channel, but is a function of the channel length alone. Thus, spin-independent elastic scattering does not result in an additional spin relaxation. In the presence of an external magnetic field, however, spin-flip processes become possible, and the Elliott–Yafet spin relaxation mechanism will be activated [125].

### 2.4.2. A fin-based quasi-ballistic spinFET: model expressions for the magnetoresistance

A simple analytical expression for the magnetoresistance can be obtained for a quasi-ballistic one-dimensional SpinFET [125,126]. The effective Hamiltonian in the ferromagnetic regions has the following form in the one-band effective mass approximation

$$\hat{H}_F^L = \frac{\hat{p}_x^2}{2m_f^*} + h_0 \hat{\sigma}_z, \quad x < 0, \tag{55}$$

$$\hat{H}_F^R = \frac{\hat{p}_x^2}{2m_f^*} \pm h_0 \hat{\sigma}_z, \quad x > L, \tag{56}$$

where $m_f^*$ is the effective mass in the contacts, $h_0 = 2PE_F/(P^2 + 1)$ is the exchange splitting energy with $P$ defined as the spin polarization in the ferromagnetic regions, $E_F$ is the Fermi energy, and $\hat{\sigma}_z$ is the Pauli matrix; $\pm$ in (56) stands for the parallel and anti-parallel configurations of the contact magnetization. For the one-dimensional semiconductor channel region the Hamiltonian has been formulated in [125–127]. The spin–orbit coupling for silicon is taken in the form (54). The Hamiltonian in the channel reads

$$\hat{H}_S^{[100]} = \sum_n \frac{\hat{p}_x^2}{2m_n^*} + \delta E_n - \frac{\beta}{\hbar} \hat{\sigma}_x \hat{p}_x + \frac{1}{2} g \mu_B B \hat{\sigma}^*, \tag{57}$$

for [100] oriented fins and

$$\hat{H}_S^{[110]} = \sum_n \frac{\hat{p}_x^2}{2m_n^*} + \delta E_n - \frac{\beta}{\hbar} \hat{\sigma}_y \hat{p}_x + \frac{1}{2} g \mu_B B \hat{\sigma}^*, \tag{58}$$

for [110] oriented fins. Here $m_n^*$ is the transport subband effective mass, $\delta E_n$ is the subband mismatch between the ferromagnetic region and the channel, and $\beta$ is the strength of the spin–orbit coupling, $g$ is the Landé factor, $\mu_B$ is the Bohr magneton, $B$ is the magnetic field, and

$$\hat{\sigma}^* \equiv \hat{\sigma}_x \cos\gamma + \hat{\sigma}_y \sin\gamma \tag{59}$$

with $\gamma$ defined as the angle between the magnetic field and the transport direction.

To calculate the dependence of the transport properties on the spin–orbit coupling one needs the electron eigenfunctions in the various regions. By performing an appropriate unitary transformation on a four-component wave functions and taking into account that without spin–orbit coupling the subband energies are degenerate and the subband wave functions are equivalent, one can reduce the subband wave function in the transversal subband basis to a two-component spinor. For the ferromagnetic regions spin-up and spin-down eigenstates have the form $(1, 0)^\dagger$ and $(0, 1)^\dagger$, respectively. In these notations the wave function in the left contact for incoming spin-up electrons has the following form:

$$\Psi_L^\uparrow(x) = (e^{ik_\uparrow x} + R_\uparrow e^{-ik_\uparrow x}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + R_\downarrow e^{-ik_\downarrow x} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{60}$$

Correspondingly, $k_{\uparrow(\downarrow)} = \sqrt{2m_f^*(E \mp h_0)/\hbar^2}$ is the wave vector of the spin-up (spin-down) electron and $R_{\uparrow(\downarrow)}$ is the amplitude of the reflected wave. For the right contact the wave function is given by

$$\Psi_R(x) = C_\uparrow e^{ik_\uparrow x} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + C_\downarrow e^{ik_\downarrow x} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

For the semiconductor region the wave function for a [100]-oriented fin can be written as

$$\psi_S^{[100]}(x) = A_+ e^{ik_{x1}^{(+)} x} \begin{pmatrix} k_1 \\ 1 \end{pmatrix} + B_+ e^{ik_{x2}^{(+)} x} \begin{pmatrix} k_2 \\ 1 \end{pmatrix} + A_- e^{ik_{x1}^{(-)} x} \begin{pmatrix} k_3 \\ -1 \end{pmatrix} + B_- e^{ik_{x2}^{(-)} x} \begin{pmatrix} k_4 \\ -1 \end{pmatrix}, \tag{61}$$

where $k_{x1(x2)}^{(+)}$ and $k_{x1(x2)}^{(-)}$ are the two solutions for the wave vectors obtained from the equations

$$\frac{\hbar^2 k^2}{2m_s^*} + \delta E_c \pm \sqrt{\left(\frac{Bg\mu_B \cos(\gamma) - \beta k}{2}\right)^2 + \left(\frac{Bg\mu_B \sin(\gamma)}{2}\right)^2} = E,$$

with $+$ and $-$, correspondingly. The coefficients $k_1$, $k_2$, $k_3$, $k_4$ are calculated as:

$$k_1 = -\frac{\frac{iBg\mu_B\sin(\gamma)}{2} - \frac{Bg\mu_B\cos(\gamma)}{2} + \beta k_{x1}^{(+)}}{\sqrt{\left(\frac{Bg\mu_B\cos(\gamma)}{2} - \beta k_{x1}^{(+)}\right)^2 + \left(\frac{Bg\mu_B\sin(\gamma)}{2}\right)^2}},\tag{62}$$

$$k_2 = -\frac{\frac{iBg\mu_B\sin(\gamma)}{2} - \frac{Bg\mu_B\cos(\gamma)}{2} + \beta k_{x2}^{(+)}}{\sqrt{\left(\frac{Bg\mu_B\cos(\gamma)}{2} - \beta k_{x2}^{(+)}\right)^2 + \left(\frac{Bg\mu_B\sin(\gamma)}{2}\right)^2}},\tag{63}$$

$$k_3 = \frac{\frac{iBg\mu_B\sin(\gamma)}{2} - \frac{Bg\mu_B\cos(\gamma)}{2} + \beta k_{x1}^{(-)}}{\sqrt{\left(\frac{Bg\mu_B\cos(\gamma)}{2} - \beta k_{x1}^{(-)}\right)^2 + \left(\frac{Bg\mu_B\sin(\gamma)}{2}\right)^2}},\tag{64}$$

$$k_4 = \frac{\frac{iBg\mu_B\sin(\gamma)}{2} - \frac{Bg\mu_B\cos(\gamma)}{2} + \beta k_{x2}^{(-)}}{\sqrt{\left(\frac{Bg\mu_B\cos(\gamma)}{2} - \beta k_{x2}^{(-)}\right)^2 + \left(\frac{Bg\mu_B\sin(\gamma)}{2}\right)^2}}.\tag{65}$$

The reflection and transmission coefficients are determined by applying the boundary conditions at the ferromagnet/semiconductor interfaces.

We compute the current through the device as

$$I^{P(AP)}(V) = \frac{e}{h}\int_{\delta E}^{\infty}\left[T_\uparrow^{P(AP)}(E) + T_\downarrow^{P(AP)}(E)\right]\left\{\frac{1}{1 + e^{\frac{E-E_F}{k_BT}}} - \frac{1}{1 + e^{\frac{E-E_F+eV}{k_BT}}}\right\}dE,\tag{66}$$

where $\Delta V$ is the voltage. The spin-up ($T_\uparrow^P$) and spin-down ($T_\downarrow^P$) transmission probabilities for the parallel configuration of the contact magnetization are defined as

$$T_\uparrow^P = |C_\uparrow|^2 + \frac{k_\downarrow}{k_\uparrow}|C_\downarrow|^2,\tag{67}$$

$$T_\downarrow^P = \frac{k_\uparrow}{k_\downarrow}|C_\uparrow|^2 + |C_\downarrow|^2.\tag{68}$$

For the anti-parallel configuration of the contact magnetization the transmission probabilities are given by

$$T_\uparrow^{AP} = \frac{k_\downarrow}{k_\uparrow}|C_\uparrow|^2 + |C_\downarrow|^2,\tag{69}$$

$$T_\downarrow^{AP} = |C_\uparrow|^2 + \frac{k_\uparrow}{k_\downarrow}|C_\downarrow|^2.\tag{70}$$

The conductance is defined as

$$G^{P(AP)} = \lim_{V\to 0}\frac{I^{P(AP)}}{V}.\tag{71}$$

In the limit of low temperature the conductance must coincide with the one obtained from the Landauer–Büttiker formula:

$$G^{P(AP)} = \frac{e^2}{h}\left(T_\uparrow^{P(AP)}(E_F) + T_\downarrow^{P(AP)}(E_F)\right).\tag{72}$$

Finally, the tunneling magnetoresistance (TMR) is defined as

$$\text{TMR} \equiv \frac{G^P - G^{AP}}{G^{AP}}.\tag{73}$$

Fig. 13 shows the dependence of the TMR on the value of the band mismatch $\delta E_c$ between the ferromagnetic source contact and the semiconductor channel. Here we assume the effective mass for the ferromagnetic region $m_f^* = m_0$ and for the semiconductor region $m_n^* = 0.036m_0$, where $m_0$ is the electron rest mass. The TMR oscillates between positive and negative values. As the length of the semiconductor channel decreases, the period of the oscillations increases roughly proportionally to the inverse length of the semiconductor channel.

Temperature exerts a significant influence on the device characteristics as shown in Fig. 14. For a channel length $L = 0.05\ \mu\text{m}$ the oscillatory amplitude of the TMR decreases for $T = 77$ K and completely vanishes for $T = 180$ K. The

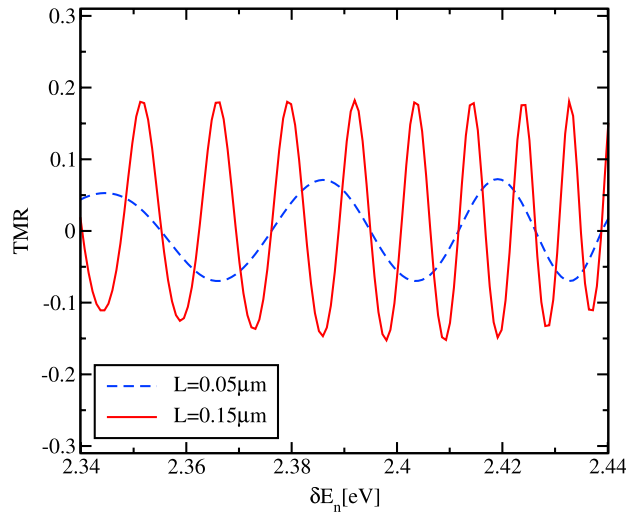**Fig. 13.** TMR dependence on the value of $\delta E_n$, for $E_F = 2.47$ eV, $P = 0.4$, $B = 0$ T, $z = 0$, $\alpha_R = 42.3$ meV nm (cf. [127]).
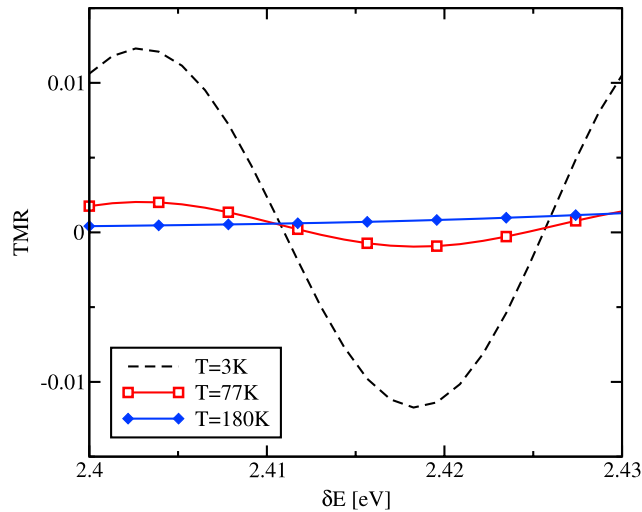


**Fig. 14.** TMR dependence on the value of $\delta E_n$, for $\alpha_R = 31.7$ meV nm (value typical for InAs), $E_F = 2.47$ eV, $P = 0.4$, $B = 0$ T, $z = 0$, $L = 0.05$ μm (cf. [127]).

reason for the oscillatory behavior to disappear at $T = 180$ K is a relatively short period of the conductance oscillations (and correspondingly TMR oscillations) shown in Fig. 13 with respect to $\delta E_n$. When the temperature becomes comparable with the period, it averages out the TMR oscillations making the signal very small. However, for a shorter channel the amplitude of the oscillations should be sufficient to modulate the current in the SpinFET at elevated temperatures.

### 2.4.3. Confinement effects on the transport effective mass in wires and fins

Due to the fact that the parabolic band approximation for the band structure in silicon is not sufficient to accurately obtain the subband structure in thin and narrow silicon fins, a more accurate subband description [104] based on the bulk $\mathbf{k} \cdot \mathbf{p}$ dispersion (8) [101,105] must be employed. We consider square silicon fins of [100] and [110] orientations, with (001) horizontal faces. The resulting Schrödinger equation with the confinement potential appropriately added to the Hamiltonian [101] is discretized the box integration method and solved for each value of the conserved momentum $p_x$ along the current direction using efficient numerical algorithms available through the Vienna Schrödinger–Poisson framework (VSP) [129].

Fig. 15 shows the dependence of the subband energies as function of the fin thickness $t$ for the lowest subbands. The fin orientation is along the [110] direction. The increase of the splitting between the unprimed subbands which would be degenerate in the effective mass approximation for thinner fins is seen. The splitting between the subbands in a [100] fins and nanowires is negligible [128]. However, the dependence of the effective mass of the ground subband in [100] fins on $t$
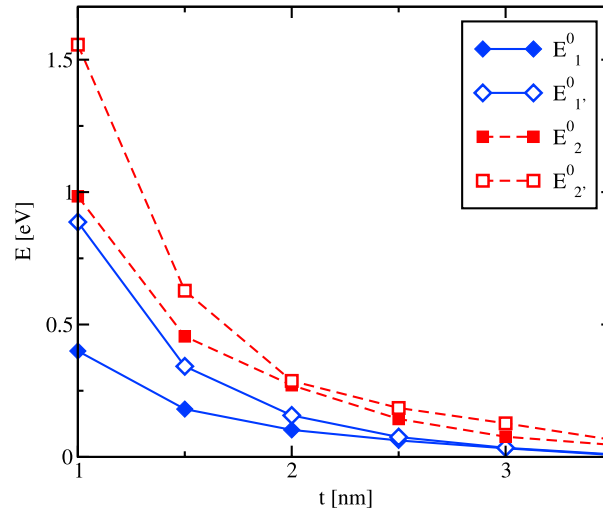
**Fig. 15.** Subband minima as a function of thickness $t$ in a [110]-oriented fin (cf. [127]).
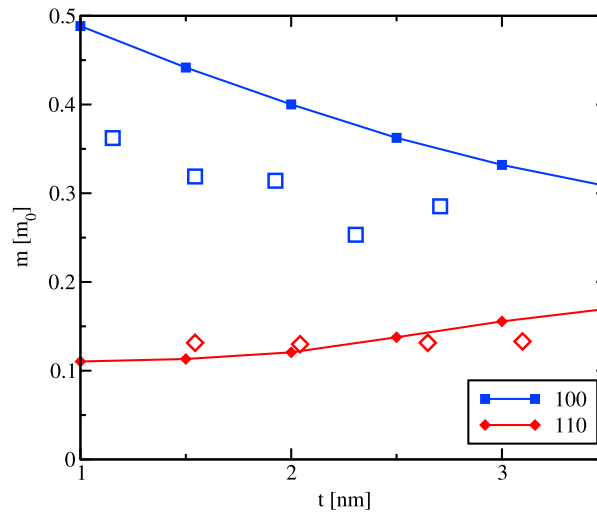


**Fig. 16.** Ground subband effective mass dependence on $t$ in [100] and [110] fins. Open symbols are from [128] (cf. [127]).

shown in Fig. 16 is more pronounced as compared to [110] fins. Results of these $\mathbf{k} \cdot \mathbf{p}$ calculations are in good agreement with density-functional calculations [128] also shown in Fig. 16.

### 2.4.4. Magnetoresistance and fin orientation

Fig. 17 shows the dependence of the TMR for [100] and [110] oriented fins with $t = 1.5$ nm on the value of the spin–orbit interaction. Fins with [100] orientation show a stronger dependence on $\beta$ compared to [110] oriented fins. Thus [100] oriented fins are to prefer for silicon SpinFETs. The reason of the stronger dependence is that the characteristic length, on which the spin–orbit interaction produces the full spin precession is defined by the inverse of the wave vector $k_D = m_n^* \beta / \hbar^2$. As shown in Fig. 16, the effective mass for the [110] oriented fins is smaller compared to the one in [100] oriented fins, hence for the same variation of $k_D$ in case of the [110] oriented fins a larger variation of $\beta$ is required to achieve the same TMR value modulation.

### 2.4.5. Tunnel barriers and magnetoresistance at elevated temperature

As it is demonstrated in Fig. 14, temperature exerts detrimental effects on the TMR signal. To facilitate the injection of the spin-polarized current into the channel, one has to introduce, following [125,126], tunnel barriers at the interfaces between the contacts and the channel with the dimensionless strength
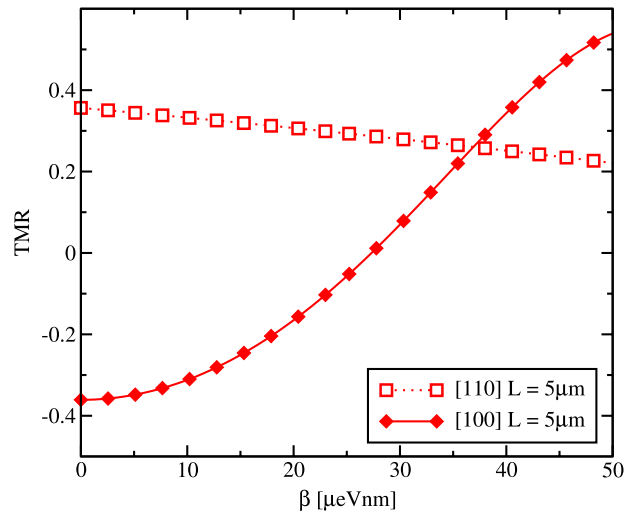
$$z = 2m_f^* U / h^2 k_F,$$ (74)

**Fig. 17.** TMR dependence on the value of the Dresselhaus spin–orbit interaction for $t = 1.5$ nm, $B = 4$ T, $P = 0.6$, $z = 3$, $\gamma = 0$ (cf. [127]).
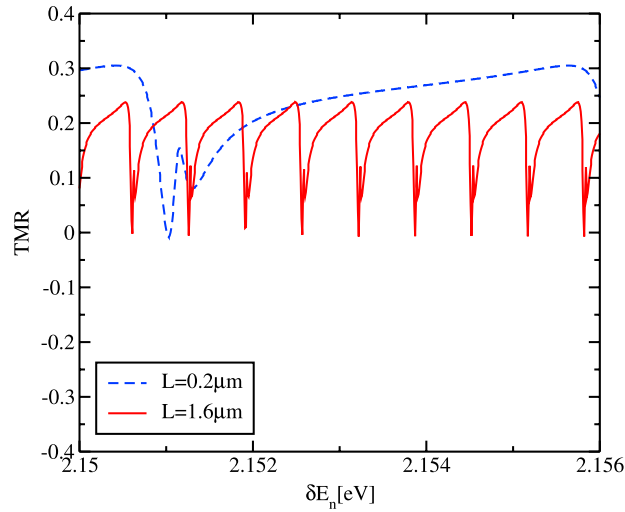


**Fig. 18.** TMR dependence on the value of the $\delta E_n$ for $E_F = 2.47$ eV, $P = 0.4$, $z = 3$, $\beta = 42.3\ \mu$eV nm (cf. [127]).

where

$$k_F = \sqrt{2m_f^* E_F / \hbar^2}. \tag{75}$$

Fig. 18 shows the dependence of the oscillations of the TMR on the value of the conduction band mismatch $\delta E_n$ for $t = 1.5$ nm. The period of the oscillations is roughly inversely proportional to the length of the semiconductor channel, in agreement with Fig. 13. The presence of the barriers at the interfaces between the contact and the channel exerts a significant influence on the oscillation shape. For higher and thicker barriers, the TMR, although being a periodic function of the conductance band mismatch, becomes positive (or negative) in a broad range of the conduction band mismatch. This sign definitiveness leads to the preserved TMR modulation as a function of the spin–orbit interaction strength even at high temperatures as shown in Fig. 19. This opens a practical possibility to modulate the TMR by changing the value of $\beta$ even at room temperature. Unfortunately, because the spin–orbit interaction in silicon is weak, the channel length needed to achieve the TMR modulation is close to a micron [127].

As efficient means of all-electric spin manipulation in silicon are yet to be discovered, the long spin lifetime and spin diffusion length allow to transfer the injected spin at large distances. This makes silicon attractive for use as efficient spin interconnects [130] in lateral spin valves and logic gates [42,43] known as all-spin logic [44]. Spin interconnects are critical for emerging spin-driven devices as they allow avoiding the spin-charge signal conversion thus reducing energy, delay, and circuitry.
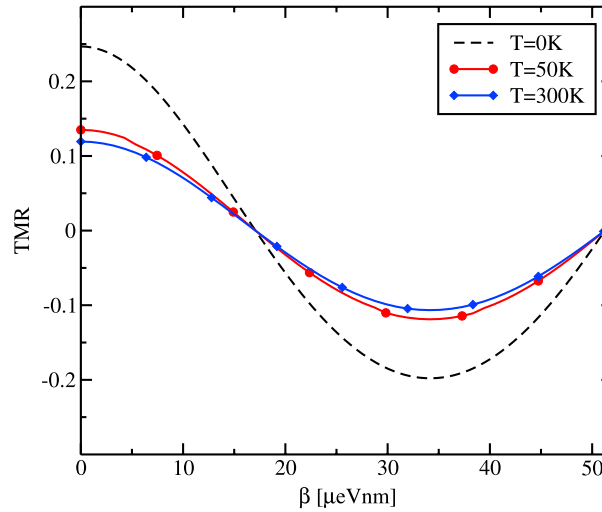
**Fig. 19.** TMR dependence on the value of the Dresselhaus spin–orbit interaction for $E_F = 2.47$ eV, $\delta E_n = 2.154$ eV, $P = 0.4$, $z = 3$, $L = 8$ μm, $V = 1$ meV (cf. [127]).

In contrast to metallic spin channels [44] suffering from the short spin relaxation length at room temperature, silicon can preserve the spin accumulation over distances of a few micrometers along a strong electric field [131]. The giant enhancement of the spin lifetime by strain in thin silicon films makes silicon-on-insulator and fin-FET technologies currently employed for 22 and 14 nm nodes CMOS fabrication a promising candidate for spin communication in reconfigurable logic circuits [42].

For short silicon channels, the only option to realize a SpinFET is to exploit the relative magnetic orientation of the source and drain ferromagnetic contacts [47]. This adds a possibility to develop reprogrammable logic where the information is additionally stored in the relative magnetization orientation of the source and drain. In these reprogrammable transistors different on-currents are expected under the same drain and gate voltage by changing the drain magnetization orientation relative to the source. It is important that, once modified, the drain magnetization remains preserved infinitely long without any extra power applied. The magnetization of the drain contact can be altered by an external magnetic field or by spin flow. The large spin flow can be conceptually achieved by purely electrical means, when the spin current impinging the drain is a large fraction of the charge current. Practically, however, the spin injection efficiency from a ferromagnetic contact into silicon achieved up to date is still too small to influence the drain magnetization. The introduction of a tunnel barriers to boost the spin injection efficiency and magnetoresistance ratio for parallel/anti-parallel source/drain magnetization orientation reduces the on-current. However, the magnetoresistance ratio of 10% demonstrated at 12 K for a 10 μm long silicon [132] channel is not sufficient for applications. It is far inferior to the magnetoresistance ratio in magnetic tunnel junctions, where the two ferromagnetic metals are separated by a thin tunnel barrier. The magnetization of one of the metal contacts can be switched fast purely electrically by pushing the current through the structure. Once switched, the magnetization is preserved without external energy sources. In addition to the high magnetoresistance ratio at room temperature and fast switching, magnetic tunnel junctions are CMOS compatible and are considered as building blocks for emerging magnetic non-volatile memory discussed in the next section.

## 3. Spin transfer torque magnetic RAM

Spin field-effect transistor research is still at the fundamental level. Indeed, several important issues related to spin injection, detection, and spin manipulation by purely electrical means are not completely resolved or even unknown. At the same time the introduction of non-volatility helps significantly reducing the heat generation, especially at stand-by, booting, and resuming stages. With memory cells based on electric charge storage rapidly approaching the physical limits of scalability the demand for new devices exploring innovative principles for information storage is rapidly increasing. Spin or magnetization projection was successfully used to store the information in magnetic carriers and hard disk drives.

However, spins do not interact efficiently with electric fields and charges, the elements laying at the heart of modern microelectronics. The use of magnetic fields for spin and magnetization manipulation is undesirable as it becomes impossible to maintain the high currents needed to generate magnetic fields with all circuits elements scaled down, including the wire cross-sections in magnetic coils.

An efficient coupling between the magnetic and electric degree of freedom is achieved at the quantum mechanical level. Indeed, in magnetic tunnel junctions (MTJs) a high resistance difference between the parallel and anti-parallel relative magnetization orientation is achieved. These structures are thus attractive for information storage as they allow reading the binary information by efficiently converting different relative magnetization orientations into very distinct voltage signals.
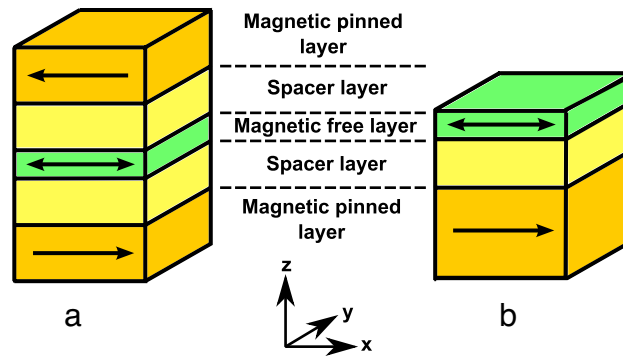
**Fig. 20.** Schematic illustration of different MTJ types: (a) penta-layer MTJ; (b) three-layer MTJ (cf. [133]).

Importantly, a memory cell based on an MTJ is non-volatile, scalable, simple in structure, requires relatively low operating voltages, and exhibits low power consumption, high operation speed, and high endurance.

The external magnetic field was used to alter the magnetization orientation in earlier versions of magnetic RAM. As it cannot be used for manipulation in scaled memory cells, the link between the electric charge/current and the magnetization switching was missing. The prediction [45,46] of purely electrical switching of the magnetization by means of the spin transfer torque (STT) makes STT magnetic random access memory (MRAM) a promising candidate for future universal memory. Indeed, STT-MRAM is characterized by small cell size ($4F^2$) and high density inherent to DRAM, fast access time (less than 10 ns) intrinsic to SRAM, non-volatility and long retention time subject to flash as well as high endurance ($10^{14}$). Most importantly, STT-MRAM is compatible with CMOS technology as MTJs are typically grown on the top of a silicon wafer. Several companies are working on developing STT-MRAM commercially, with first products already on the market.

The simple three-layer MTJ (Fig. 20(b)) represents a sandwich of two magnetic layers separated by a thin spacer which forms a tunnel barrier. While the magnetization of the pinned layer is fixed, the magnetization orientation of the free layer, which acts as the recording layer, can be switched between the two stable states parallel and anti-parallel to the fixed magnetization direction. Switching between the two states is induced by spin-polarized current flowing through the MTJ. Because the spin-polarized current is only a fraction of the total charge current passing through the cell, high currents are required to switch the magnetization direction of the recording layer. Because of an expected limit of a MTJ cross-section scaling needed to ensure an appropriate thermal stability, the value of the switching current is decreased by decreasing the current density passing through an MTJ. The reduction of the current density required for switching and the increase of the switching speed are the most important challenges in STT-MRAM developments [134].

### 3.1. In-plane and perpendicular magnetization MTJs for STT-MRAM

The spin torque enhancement in a penta-layer MTJ shown in Fig. 20(a) results in a significantly lower switching current density at a switching delay comparable to that in three-layer structures [135], which makes the penta-layer cells attractive for low power high performance memory applications. In order to improve the READ signal, the magnetization of one of the fixed layers has to be flipped [136], possibly with a help of thermally assisted switching [137]. The penta-layer structure has been investigated [138] by using the ballistic Green's function formalism combined with the magnetic layer dynamics based on the Landau–Lifschitz–Gilbert equation. The spin torque enhancement has been reported in the anti-parallel configuration under dual barrier resonance tunneling conditions, when the current is high.

Depending on the orientation of the layer magnetizations, the magnetic elements can be divided into two categories: Perpendicular with out-of-plane magnetization and in-plane with the magnetization lying in the plane of the magnetic layers. The problem of a relatively high switching current density in in-plane structures is rooted in the specificity of the switching dynamics, when the magnetization must turn out of plane, which results in an additional penalty contribution to the switching barrier. In in-plane structures the thermal barrier separating the two stable states is due to the shape anisotropy alone and is therefore lower than the switching barrier.

The solution of this problem is to resort to perpendicular MTJs. Indeed, in perpendicular MTJs the switching path is the same as the path for magnetization reversal due to thermal agitations. Therefore, the switching barrier is equal to the thermal barrier. However, the reduction of the Hilbert damping and thus the switching current density [139] and the increase of thermal stability in perpendicular structures [140,141] are substantial challenges for the known materials and prompts effortful search for new materials with superior characteristics.

Below we present a different solution to the switching time and current reduction without sacrificing the thermal stability. The idea is to use an in-plane structure with a composite free layer as recording layer, which is made of two half-elliptic parts separated by a narrow gap [133]. The superior performance of an MTJ with such a recording layer is proven by micromagnetic simulations briefly outlined next.

### 3.2. Micromagnetic modeling of STT-MRAM

Our simulations are based on the magnetization dynamics described by the Landau–Lifschitz–Gilbert–Slonczewski equation:

$$\frac{d\mathbf{m}}{dt} = -\frac{\gamma_G}{1+\alpha^2} \cdot ((\mathbf{m} \times \mathbf{h}_{eff}) + \alpha \cdot [\mathbf{m} \times (\mathbf{m} \times \mathbf{h}_{eff})]$$
$$+ \frac{g\mu_B j}{e\gamma_G M_s d} \cdot (g_1(\Theta_1) \cdot (\alpha \cdot (\mathbf{m} \times \mathbf{p}_1) - [\mathbf{m} \times (\mathbf{m} \times \mathbf{p}_1)])$$
$$- g_2(\Theta_2) \cdot (\alpha \cdot (\mathbf{m} \times \mathbf{p}_2) - [\mathbf{m} \times (\mathbf{m} \times \mathbf{p}_2)]))). \tag{76}$$

$\gamma_G$ is the gyromagnetic ratio, $\alpha$ is the Gilbert damping parameter, $g$ is the gyromagnetic splitting factor, $\mu_B$ is Bohrs magneton, $j$ is the current density, $e$ is the electron charge, $d$ is the thickness of the recording layer, $\mathbf{m} = \mathbf{M}/M_s$ is the position-dependent normalized vector of the magnetization in the recording layer, $\mathbf{p}_1 = \mathbf{M}_{p1}/M_{sp1}$ and $\mathbf{p}_2 = \mathbf{M}_{p2}/M_{sp2}$ are the normalized magnetizations in the first and second pinned layers, respectively. $M_s$, $M_{sp1}$, and $M_{sp2}$ are the saturation magnetizations of the recording layer, the first pinned layer, and the second pinned layer, correspondingly. We use Slonczewski's expressions for the torque in the MTJ with a dielectric layer [142]

$$g_1(\Theta_1) = 0.5 \cdot \Pi[1 + \Pi^2 \cdot \cos(\Theta_1)]^{-1} \tag{77}$$

and with a metal layer [45]

$$g_2(\Theta_2) = [-4 + (1 + \Pi)^3 (3 + \cos(\Theta_2))/4\Pi^{3/2}]^{-1} \tag{78}$$

between the ferromagnetic contacts, respectively, where $\Pi$ is a polarizing factor [45]. In the penta-layer structure the two spin torques are acting independently on the two opposite interfaces of the recording ferromagnetic layer, provided its thickness is larger than the scale on which the electron spins entering into the ferromagnet become aligned to the ferromagnet's magnetization. The local effective field is calculated as:

$$\mathbf{h}_{eff} = \mathbf{h}_{ext} + \mathbf{h}_{ani} + \mathbf{h}_{exch} + \mathbf{h}_{demag} + \mathbf{h}_{th} + \mathbf{h}_{amp} + \mathbf{h}_{ms}. \tag{79}$$

$\mathbf{h}_{ext}$ is external field, $\mathbf{h}_{ani}$ is anisotropic field, $\mathbf{h}_{exch}$ is a exchange field, $\mathbf{h}_{demag}$ is a demagnetizing field, $\mathbf{h}_{th}$ is a thermal field, $\mathbf{h}_{amp}$ is the Ampere field, and $\mathbf{h}_{ms}$ is the magnetic coupling between the pinned layers and the recording layer.

In the uniaxial anisotropy case the anisotropic field is [143]:

$$\mathbf{h}_{ani} = \frac{2K_1}{\mu_0 M_s} (\mathbf{m} \cdot \mathbf{u})\mathbf{u}, \tag{80}$$

where $\mathbf{u}$ is a unit vector along an easy anisotropy axis. An easy axis is an energetically favorable direction of spontaneous magnetization. For the cubic anisotropy the projection $i$ of the field is calculated as:

$$h_{ani,i} = -\frac{2D_{ii}}{\mu_0 M_s} m_i. \tag{81}$$

Herewith

$$D_{11} = K_1(m_y^2 + m_z^2) + K_2 m_y^2 m_z^2, \tag{82}$$

$$D_{22} = K_1(m_x^2 + m_z^2) + K_2 m_x^2 m_z^2, \tag{83}$$

$$D_{33} = K_1(m_x^2 + m_y^2) + K_2 m_x^2 m_y^2, \tag{84}$$

$K_1$ and $K_2$ are the material-dependent anisotropy coefficients, and $\mu_0$ is the magnetic constant.

The exchange field is calculated as [143]:

$$\mathbf{h}_{exch}(\mathbf{r}) = \frac{2A}{\mu_0 M_s} \sum_j ((\mathbf{m}(\mathbf{r} + \delta\mathbf{r}_j) - \mathbf{m}(\mathbf{r}))/|\delta\mathbf{r}_j|^2). \tag{85}$$

$A$ is the exchange constant, and $\mathbf{r} + \delta\mathbf{r}_j$ specifies all $j$ nearest neighbors to $\mathbf{r}$. For calculating the demagnetization field we used the method proposed in [144]. The thermal field is calculated as [145]:

$$h_{thj} = \sigma \cdot \sqrt{\frac{\alpha}{1+\alpha^2} \cdot \frac{2k_B T}{\gamma_G \delta V \Delta t M_s}}. \tag{86}$$

Here $j = x, y, z$, $\sigma$ is a Gaussian random uncorrelated function, $\delta V$ is the volume of cell, and $\Delta t$ is the time step.
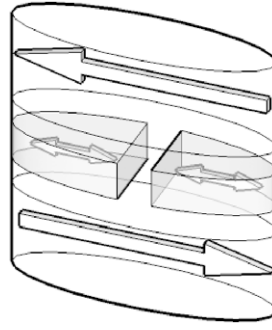
**Fig. 21.** Illustration of the proposed MRAM cell with a composite recording layer.
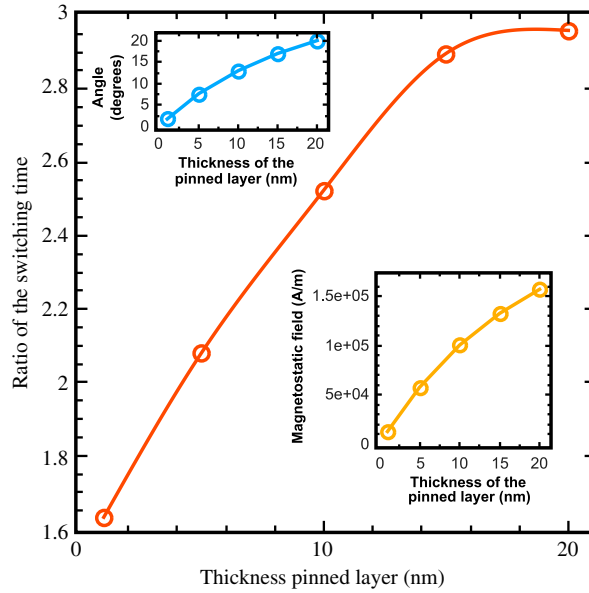


**Fig. 22.** Ratio of the switching time in the monolithic structure vs. in the composite structure as function of the thickness of the pinned layer. The top insets show the averaged initial angle, the bottom inset: The dependence of the absolute values of the $z$-component of the averaged magnetic field (cf. [133]).

The eddy currents field is [146]:

$$\mathbf{h}_{amp,i} = \sum_{j=1..N} \frac{\mathbf{J}_j}{4\pi} \times \int_j \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} dv. \tag{87}$$

Here, $\mathbf{J}_j$ is the current induced on every cell ($j : 1..N$).

### 3.3. Fast switching in a composite recording layer

Next we discuss the simulation results for the structure CoFe/MgO/Py/MgO/CoFe, where Py is $Ni_{81}Fe_{19}$ with an elliptical cross-section. The system with a composite ferromagnetic layer is obtained by removing a central stripe of a certain width from the monolithic recording layer (Fig. 21). When the current starts flowing through the structure with the monolithic recording layer, the spin torque remains marginal in the central region, where the magnetization is along the $x$-axis. As the amplitude of the end domains, precession increases, the central region experiences almost no spin torque and preserves its initial orientation along the $x$-axis, thus preventing the whole layer from alternating its magnetization orientation. This is albeit not the case, when the central region is removed in the composite structure and the end domains become virtually independent. Fig. 22 demonstrates a substantial decrease of the switching time in the penta-layer structure with the composite recording layer, for the same current density, as a function of the thickness of the pinned ferromagnetic layers. Due to the removal of the central region which is the "bottleneck" for switching in the monolithic structure, the switching time decreases. The reduction of the switching time depending on geometry parameters is investigated in detail in [147,148].
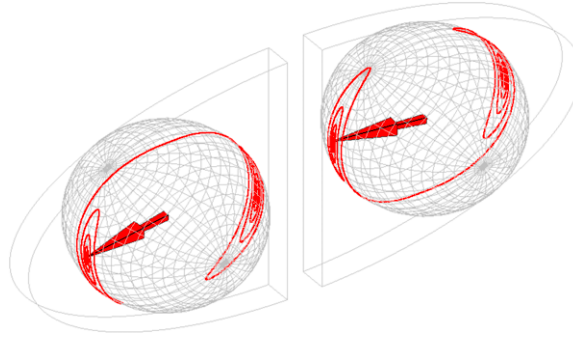
**Fig. 23.** Magnetization components vs. time for an elliptical $52.5 \times 25$ nm$^2$ MTJ with a composite recording layer. The magnetizations of the left and right halves shown separately stay in-plane explaining the acceleration of switching (cf. [147]).
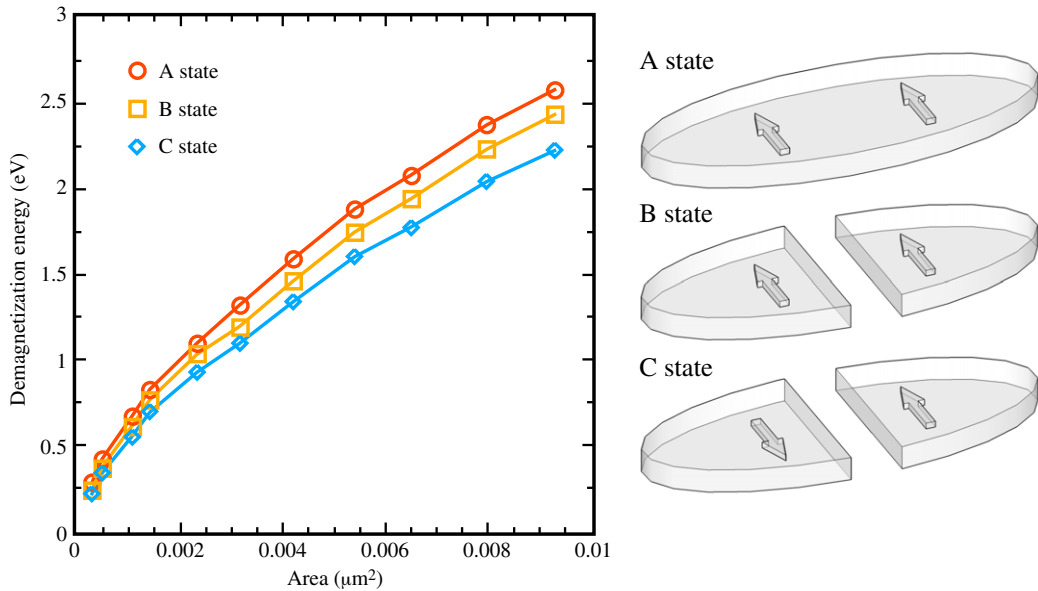


**Fig. 24.** Dependence of the demagnetization energy for MTJs with (A) monolithic and (B, C) composite recording layers as a function of the cross-sectional area. The state (C) with anti-parallel alignment of the magnetizations between the left and right halves is characterized by the smallest energy and thus by the smallest switching barrier (cf. [147]).

### 3.3.1. Switching energy barrier reduction

In order to reveal the reason for fast switching, we investigate the magnetization dynamics of the left and right parts of the composite structure separately (Fig. 23). A structure with an elliptical $52.5 \times 25$ nm$^2$ cross-section and the following layer sequence CoFeB(5 nm)/MgO(1 nm)/CoFeB(2 nm)/MgO(1 nm) /CoFeB(5 nm) is considered. The central 2.5 nm stripe is removed from the middle CoFeB layer. Fig. 23 shows that the switching processes of the left and right parts of the composite recording layer occur in opposite senses to each other. Most importantly, the magnetization of each half stays almost in-plane. This switching behavior leads to a decrease of the switching energy barrier.

To find the switching path due to thermal agitations, it is necessary to determine the state for which the barrier separating the two stable magnetization states in the recording layer is minimal. Fig. 24 shows that the switching path due to thermal agitations must go through the state with magnetizations of the halves opposite to each other (the C state in Fig. 24). Since during the switching the composite layer proceeds through the same state, we conclude that, as in perpendicular MTJs, the switching barrier and the thermal barrier are the same [147] as shown in Fig. 25. Reduction of the switching barrier in the composite structure leads to a reduction of the switching time compared to the conventional MTJ, provided the current densities are the same.

### 3.3.2. Optimization of thermal stability and switching uniformity

Next we analyze the influence of scaling the dimensions of the recording layer on the thermal stability factor. Due to the removal of the central region in the monolithic structure and anti-parallel alignment of the two parts of the composite layer, the thermal barrier is reduced by 30%. However, to boost the thermal stability factor, it is sufficient to increase the thickness
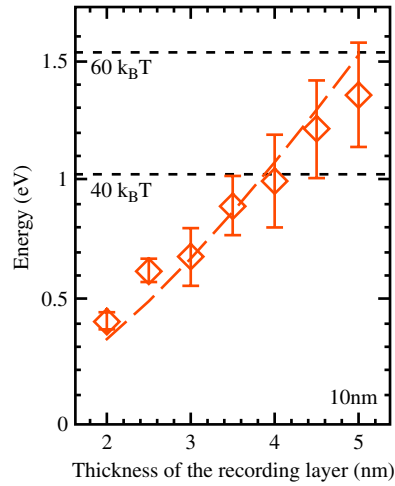
**Fig. 25.** Thermal energy barrier (lines) vs. switching energy barrier (symbols) for the composite structure. The long axis is 52.5 nm, the short axis is 10 nm, the thickness of the recording layer is 5 nm. Each point is a result of statistical averaging with respect to 50 different realizations of the switching process (cf. [133]).
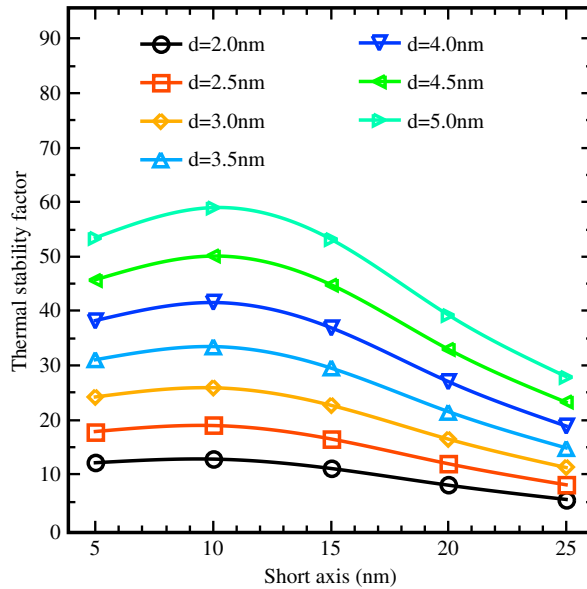


**Fig. 26.** Thermal stability barrier for MTJs with composite recording layer as a function of the short axis length for several recording layer thicknesses *d*. The long axis is fixed at 52.5 nm; the fixed layer thickness layer is 5 nm (cf. [147]).

of the recording layer and/or the aspect ratio. Fig. 26 shows the thermal stability factors of MTJs with a composite recording layer as a function of the recording layer thickness *d*. An MTJ with $52.5 \times 10$ nm$^2$ cross-section and $d = 5$ nm recording layer thickness has a thermal stability factor $\sim 60 k_B T$ comparable to the one of a perpendicular MTJ with a monolithic layer [149].

Because of the magnetizations of each part of the composite layer moving in opposite senses, the magnetic field generated by an opposite part helps switching the magnetization of the given part. This makes the switching not only faster but also more homogeneous as compared to that of the structure with the monolithic layer. Indeed, the width of the switching time distribution for MTJs with a composite recording layer can be almost $\sim 2000$ times narrower than that for MTJs with a monolithic layer (Fig. 27).

Finally, a structure with a recording layer composed of ellipses with the axes $a/2$ and $b$ inscribed into a rectangle $a \times b$ (Fig. 28) is investigated. By means of micromagnetic simulations it is shown that the structure is characterized by a switching time similar to the one in a previously considered structure (Fig. 21). The switching time is also similar to that in a smaller MTJ shown in Fig. 28(b) with a monolithic layer of an elliptical cross-section corresponding to one-half of the recording layer shown in Fig. 28(a). However, the smaller MTJ possesses the thermal stability factor of nearly two times smaller than the composite structure Fig. 28(a). It demonstrates the potential of the in-plane structures with the composite recording layer
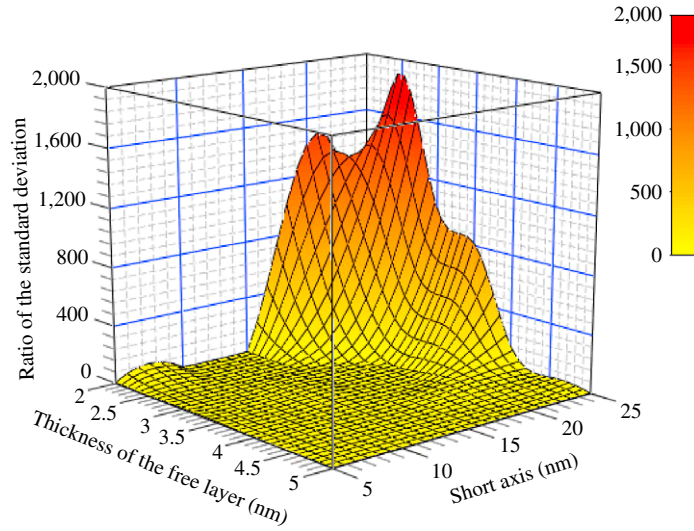
**Fig. 27.** Ratio of the standard deviation of the switching time in the monolithic structure and composite structure as function of thickness of the recording layer and short axis length. The long axis is fixed at 52.5 nm and the thickness of the fixed layers is 10 nm (cf. [148]).
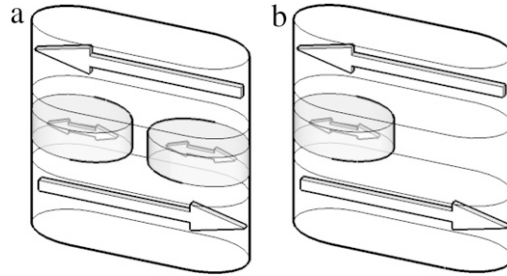


**Fig. 28.** Schematic illustration of penta-layer MTJs (cf. [133]).

for STT-MRAM applications. Importantly, since the MTJ in Fig. 28(a) does not require a narrow gap between the two parts of the composite recording layer, it can be easily fabricated.

### 3.4. New architectures and materials for efficient magnetization switching

MTJs with a perpendicular magnetization orientation are promising for the STT-MRAM application. In order for STT-MRAM to become a competitive alternative to DRAM, it is mandatory to have MTJs with a large thermal stability barrier $E$ over 80 $k_BT$, a small junction size (less than 20 nm in diameter), and a higher than 100% TMR ratio [150]. Many materials for the recording magnetic layer and the ferromagnetic electrodes have been explored. It appears that a composite structure [151] of the recording layer is needed to fulfill the requirements. It is shown [150] that a carefully designed recording layer containing a Co/Pt multilayer with a CoFeB/Ta insertion layer provide MTJs with desired characteristics.

To generate spin-polarized currents needed for the magnetization switching of the recording layer in MTJs ferromagnetic electrodes are required. Recently, it has been discovered that a current passing through a non-magnetic material with high values of the spin Hall angle may flip the magnetization in a nanomagnet fabricated on it [152]. Based on the phenomenon, a three-terminal MTJ has been suggested, where the magnetization of the recording layer is switched by the torque generated by the spin current appearing due to the Rashba effect and/or the spin Hall effect from the current passing in a non-magnetic material. If a VLSI compatible material, Cu doped with 10% Ir, is used as a channel material, the switching happens at a current density less than $10^6$ A cm$^{-2}$ [153], comparable to that reported in two-terminal MTJs. Magnetization reversal in an MTJ with perpendicular magnetization due to the spin torques produced by the current in a non-magnetic material has also been recently reported [154].

An electric field alters the properties of a ferromagnet, in particular the magnetization anisotropy, through the modulation of the carrier density. A combination of STT and electric field mediated switching allows to reverse the magnetization faster and more reliably [155]. An electric field also helps to invert the magnetization in a three-terminal MTJ [156].

Materials with high spin Hall angle or Rashba torques are required for efficient switching in three-terminal MTJs. Due to the spin-momentum locking in the conducting states at the interface resulting in a large spin polarization, when the current
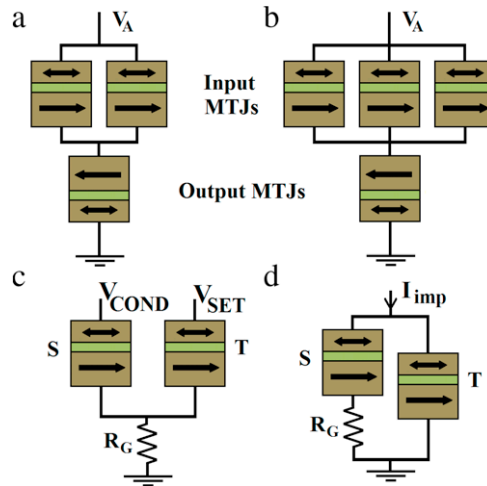
**Fig. 29.** MTJ-based two-input (a) and three-input (b) reprogrammable logic gates [160,161]. (c) Voltage-controlled (VC-IMP) and (d) current-controlled (CC-IMP) implication logic gates [162]. The resistance states of the target ($T$) and the source ($S$) MTJs act as the inputs ($t$ and $s$) and the final logic state of $T$ ($t'$) is the output of the implication gates.

is flowing through the states [157], a relatively new class of materials, topological insulators, are promising for spintronic applications. Indeed, a very large spin transfer torque [158] and a magnetization switching by giant spin–orbit torque [159] were recently reported in a magnetic heterostructure containing a topological insulator, prompting for a swift appearance of devices with topological insulators on the market.

## 4. STT-MRAM based logic-in-memory

### 4.1. Non-volatile logic

The spin field-effect transistors research reviewed in Section 2 is still at a fundamental level because of several important issues related to spin injection, detection, and manipulation by purely electrical means. At the same time the introduction of non-volatile logic could help significantly reducing heat generation, especially at stand-by, booting, and resuming stages. It is extremely attractive to use the same elements as memory and latches, which reduce the time delay and energy waste while transferring data between CPU and memory blocks in conventional computer architectures. As demonstrated in Section 3 STT-MTJ-based memory has all the characteristics of a universal memory [163]. Furthermore, the MTJ technology is attractive for building logic configurations which combine non-volatile memory cells and logic circuits (so-called logic-in-memory architecture) to overcome the leakage power issue [164–166]. Several MTJ/MOS hybrid logic-in memory circuits have so far been proposed [167]. A six-input lookup table circuit using non-volatile logic in memory architecture is a standby-power free field-programmable array which also allows to reduce the transistor count by 62% [168].

Recently, the realization of MTJ-based non-volatile logic gates has been successfully demonstrated, for which the MTJ devices are used simultaneously as non-volatile memory cells and main computing elements [160–162]. In [160,161] reprogrammable logic gates (Fig. 29(a) and (b)) realize the basic Boolean logic operations AND, OR, NAND, NOR, and the Majority operation. All basic Boolean logic operations are executed in two sequential steps including an appropriate preset operation in the output MTJ and then applying a voltage pulse ($V_A$) with a proper amplitude to the gate. Depending on the logic states of the input MTJs, the preset in the output MTJ, and the voltage level applied to the gate, a conditional switching behavior in the output MTJ is provided, which corresponds to a particular logic operation [161].

### 4.2. Material implication based logic-in-memory

The MTJ-based logic gates reported in [162] are designed on a voltage-controlled implication circuit topology (Fig. 29(c)) and on a beneficial current-controlled topology (Fig. 29(d)) to realize Boolean implication (IMP) operations. Compared to the TiO$_2$ memristive switches [170], MTJs provide a higher endurance, which is still a major challenge for memristors to be used as universal memory cells or computing elements [171,172]. Furthermore, the bistable resistance state of the MTJs eliminates the need for refreshing circuits, which are required in memristive computations due to the state drift phenomenon as explained before. According to the simulation results from [173,174], without refreshing the states in a memristive implication gate, state drift can cause a one-bit error in less than ten sequential logic steps.

In the voltage- and current-controlled implication gates (Fig. 29(c) and (d)) the logic operation (N)IMP is realized based on a conditional switching in the target MTJ ($T$). Depending on the initial resistance states of the source and the target MTJs,
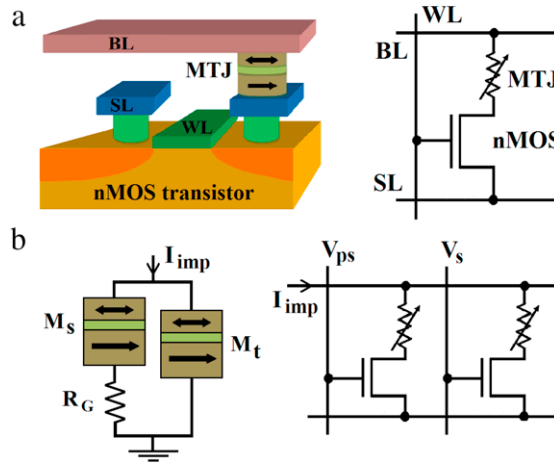
**Fig. 30.** The common STT-MRAM architecture [169] based on the one-transistor/one-MTJ (1T/1MTJ) structure (cf. [162]).

a switching event is enforced in the target MTJ only, when both MTJs are initially at high resistance states. For the other input patterns, the resistance states of the MTJs are left unchanged. In the MTJ-based voltage-controlled implication (VC-IMP) gate (Fig. 29(c)), the logic operation is executed by simultaneously applying the voltage pulses $V_{COND}$ and $V_{SET}$. As $|V_{COND}| < |V_{SET}|$, the voltage drop on S is smaller than the critical voltage level required for STT switching and thus S is left unchanged. The resistance state of S provides a voltage modulation across T through $R_G$. Due to this modulation, T switches, when S is in the high resistance state, but remains unchanged, when S is in the low resistance state.

In the current-controlled implication (CC-IMP) gate (Fig. 29(d)) the logic operation is performed by applying the current pulse $I_{IMP}$ to the gate [162]. $I_{IMP}$ is applied in a direction which tends to enforce high to low resistance state switching events on both MTJs. The current $I_{IMP}$ is split between S and T inversely proportional to the total resistance of each branch. The current split depends on the input pattern as the resistance value of each branch is a function of the logic state of its MTJ. The only desired switching event is the high to low resistance switching in T when both S and T are in the high resistance state. Due to $R_G$, a majority part of the current $I_{IMP}$ flows through T. This current is higher than the critical current required for switching. When S is in the low resistance state, the current flowing through T is decreased below the critical level required for switching as the current in the other branch increases. Because of $R_G$, the current flowing through S is always below the critical current level required for switching and thus S is left unchanged.

The MTJ-based CC-IMP gate (Fig. 29(d)) provides superior performance [175]. However, its structural asymmetry caused by $R_G$ makes the generalization of the CC-IMP gate to a large-scale implication logic circuit more problematic. Indeed, as $R_G$ is connected in series to S, S (T) can be used only as source (target) MTJ for the implication operations and the logic result stored in T cannot be used as a source input for the next implication operation. Therefore, intermediate read/write operations are required to read the data stored in a target MTJ and to write it to a source MTJ, which increases complexity, energy consumption, and delay.

The problem of asymmetry is addressed by an innovative solution [175] using the access transistor as a voltage-controlled resistor. In the MRAM array, the structural asymmetry required for the CC-IMP is provided, when the select and pre-select voltage signals ($V_s$ and $V_{ps}$) are applied to two arbitrary world lines (WLs) as shown in Fig. 30. As $V_{ps} < V_s$, the transistors exhibit different channel resistances and the required structural asymmetry is provided by the pre-selected transistor showing a higher resistance which acts as $R_G$. The logic operation is performed by applying simultaneously the current $I_{imp}$ to the common bit line (BL) and $V_s$ and $V_{ps}$ to the WLs of the target and the source 1T/1MTJ cells, respectively. The logic result is stored as the final resistance state of the selected (target) MTJ, which can be used now as a source input by pre-select in the next operations.

The logic implementation using MTJ-based logic gates relies on a conditional switching behavior provided by state-dependent current modulations on the output (target) MTJs. These modulations are caused by the changes in the MTJs' resistances for different initial logic states. The resistance modulation between the high and low resistance states in the MTJ is proportional to the TMR ratio of the MTJs. Therefore, from a device point of view, the average error probabilities of all MTJ-based operations are expected to decrease with increasing TMR ratio which is considered as the most important device parameter for the reliability [176].

## 5. Summary and conclusion

Because of the recent ground-breaking experimental and theoretical findings silicon is now gaining momentum to be used in electronic applications involving spin. All-electrical spin injection and propagation at room temperature in lightly- and heavily-doped silicon has been successfully demonstrated. However, the several orders of magnitude larger than the

expected amplitude of the spin accumulation signal obtained within the three-terminal method casts reasonable doubts on the signal nature. A recently offered explanation of the large magnitude of the observed signal based on spin-dependent trap-assisted tunneling suggests that it is not caused by the spin accumulation in silicon and, therefore, a more sophisticated non-local detection scheme must be applied to quantify spin injection. Based on the non-local detection schemes it is proven that spins are injected in n-type silicon, while an unambiguous demonstration of electrical spin injection into p-type silicon is still pending due to an extremely short spin lifetime and diffusion length.

Silicon is a perfect material for spin-driven application due to the weak strength of the spin–orbit interaction in the conduction band and the non-magnetic atoms it is composed of, which results in a long electron spin diffusion length even at room temperatures. In silicon-on-insulator and fin-based structures additional spin relaxation due to surface roughness and phonons causes spin lifetime shortening, by analogy to the mobility degradation in ultra-thin silicon body MOSFETs. However, in contrast to the momentum relaxation determined by intra-valley scattering, the spin relaxation is determined by intra-valley transitions. In (001)-oriented ultra-thin films the valley degeneracy is partly lifted, so the main spin-flip processes defining the spin relaxation are the *g*-type transitions between the two low-energy valleys along the same [001] direction. The remaining degeneracy is lifted due to the valley–orbit interaction which can be greatly enhanced by applying uniaxial [110] stress. This suppresses the most efficient spin-flips due to inter-valley scattering, which results in a giant, almost exponential with shear strain, spin lifetime enhancement. Therefore, mechanical stress routinely applied to enhance the electron mobility can also be used to boost the spin lifetime. This makes strained silicon thin films and nanowires potentially strong candidates for spin interconnects in all-spin logic devices and circuits.

The weak spin–orbit interaction in the conduction band of silicon renders the all-electric spin manipulation by means of Rashba-type electric field-dependent spin–orbit fields inefficient. The typical channel length needed to achieve the spin direction inversion is about a micron in strong effective field. A lack of an efficient way to couple spins with the electric field makes the realization of silicon-based SpinFETs highly improbable, and the search for an all-electric way to manipulate spins in silicon must be continued.

At the same time, an efficient coupling between the electrical and the magnetic degrees of freedom is achieved in magnetic tunnel junctions at the quantum-mechanical level. This makes magnetic memory cells with all-electric switching, STT-MRAM, a viable candidate for future universal memory, which is fast, non-volatile, and CMOS compatible. While the first MRAM arrays are already commercially available, the reduction of the switching current, while preserving the high thermal stability barrier, is the main challenge in the area. A composite structure of the recording layer helps reducing the critical current density without compromising the thermal stability in MRAM cells with perpendicular and in-plane magnetization orientation.

Finally, STT-MRAM arrays are perfectly suited for implementing an intrinsic logic-in-memory architecture, where the same elements are employed to store and to process information. Indeed, an implication logic gate can be built by using any of the two MRAM cells, which opens the way to store and to process the data on the same elements.

Regardless of an ultimate progress and indisputable advantages in understanding room temperature spin properties in silicon many exciting challenges are lying ahead. They include but are not limited to finding efficient ways of manipulating spins in silicon and MTJs by means of electric fields and voltages alone, as well as inventing new energy efficient computational architectures with high performance. Modeling methods and simulation tools, which, according to the ITRS, help reducing R&D costs by 40% already now, are indispensable instruments for assisting to resolve these exciting challenges on a thrilling path of progress.

## Acknowledgments

## References

[1] M. Bohr, The evolution of scaling from the homogeneous era to the heterogeneous era, in: International Electron Devices Meeting, IEDM, 2011, pp. 1.1.1–1.1.6.
[2] M. Hudait, G. Dewey, S. Datta, J. Fastenau, J. Kavalieros, W.K. Liu, D. Lubyshev, R. Pillarisetty, W. Rachmady, M. Radosavljevic, T. Rakshit, R. Chau, Heterogeneous integration of enhancement mode $In_{0.7}Ga_{0.3}As$ quantum well transistor on silicon substrate using thin ($\leq 2 \mu m$) composite buffer architecture for high-speed and low-voltage (0.5V) logic applications, in: International Electron Devices Meeting, IEDM, 2007, pp. 625–628.
[3] B. Doris, M. Ieong, T. Kanarsky, Y. Zhang, R.A. Roy, O. Documaci, Z. Ren, F.-F. Jamin, L. Shi, W. Natzle, H.-J. Huang, J. Mezzapelle, A. Mocuta, S. Womack, M. Gribelyuk, E.C. Jones, R.J. Miller, H.-S.P. Wong, W. Haensch, Extreme scaling with ultra-thin Si channel MOSFETs, in: International Electron Devices Meeting, IEDM, 2002, pp. 267–270.
[4] E.H. Hwang, S. Adam, S. Das sarma, Carrier transport in two-dimensional graphene layers, Phys. Rev. Lett. 98 (2007) 186806.
[5] W. Mehr, J. Dabrowski, J. Christoph Scheytt, G. Lippert, Y.-H. Xie, M.C. Lemme, M. Ostling, G. Lupina, Vertical graphene base transistor, IEEE Electron Device Lett. 33 (5) (2012) 691–693.
[6] A.H. Castro Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, A.K. Geim, The electronic properties of graphene, Rev. Modern Phys. 81 (2009) 109–162.
[7] M.V. Fischetti, J. Kim, S. Narayanan, Z.-Y. Ong, C. Sachs, D.K. Ferry, S.J. Aboud, Pseudopotential-based studies of electron transport in graphene and graphene nanoribbons, J. Phys. Condens. Matter 25 (47) (2013) 473202.
[8] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, A. Kis, Single-layer $MoS_2$ transistors, Nat. Nanotechnol. 6 (2011) 147–150.
[9] R. Fivaz, E. Mooser, Mobility of charge carriers in semiconducting layer structures, Phys. Rev. 163 (1967) 743–755.
[10] Y. Song, H. Dery, Transport theory of monolayer transition-metal dichalcogenides through symmetry, Phys. Rev. Lett. 111 (2013) 026601.
[11] B. Radisavljevic, A. Kis, Mobility engineering and metal–insulator transition in monolayer $MoS_2$, Nat. Mater. 12 (2013) 815–820.

[12] W. Park, Y. Lee, J. Kim, S. Lee, C. Kang, C. Cho, S. Lim, U. Jung, W. Hong, B. Lee, Reliability characteristics of MoS$_2$ FETs, in: International Conference on Solid State Devices and Materials, SSDM, 2013, pp. 684–685.
[13] S. Datta, B. Das, Electronic analog of the electro-optic modulator, Appl. Phys. Lett. 56 (7) (1990) 665–667.
[14] G.A. Prinz, Magnetoelectronics, Science 282 (5394) (1998) 1660–1663.
[15] S.A. Wolf, D.D. Awschalom, R.A. Buhrman, J.M. Daughton, S. von Molnar, M.L. Roukes, A.Y. Chtchelkanova, D.M. Treger, Spintronics: A spin-based electronics vision for the future, Science 294 (5546) (2001) 1488–1495.
[16] A. Ney, C. Pampuch, R. Koch, K.H. Ploog, Programmable computing with a single magnetoresistive element, Nature 425 (2003) 485–487.
[17] M. Johnson, Magnetoelectronics, Elsevier, 2004.
[18] I. Zutic, J. Fabian, S. Das Sarma, Spintronics: fundamentals and applications, Rev. Modern Phys. 76 (2004) 323–410.
[19] J. Fabian, A. Matos-Abiaguea, C. Ertler, P. Stano, I. Zutic, Semiconductor spintronics, Acta Phys. Slovaca 57 (2007) 565–907.
[20] S. Bandyopadhyay, M. Cahay, Electron spin for classical information processing: A brief survey of spin-based logic devices, gates and circuits, Nanotechnology 20 (41) (2009) 412001.
[21] R. Jansen, Silicon spintronics, Nat. Mater. 11 (2012) 400–408.
[22] D. Loss, D.P. DiVincenzo, Quantum computation with quantum dots, Phys. Rev. A 57 (1998) 120–126.
[23] J.J. Pla, K.Y. Tan, J.P. Dehollain, W.H. Lim, J.J.L. Morton, D.N. Jamieson, A.S. Dzurak, A. Morello, A single-atom electron spin qubit in silicon, Nat. Lett. 489 (2012) 541–545.
[24] J.J.L. Morton, D.R. McCamey, M.A. Eriksson, S.A. Lyon, Embracing the quantum limit in silicon computing, Nature 467 (2011) 687–691.
[25] R. Hanson, L.P. Kouwenhoven, J.R. Petta, S. Tarucha, L.M.K. Vandersypen, Spins in few-electron quantum dots, Rev. Modern Phys. 79 (2007) 1217–1265.
[26] B.E. Kane, A silicon-based nuclear spin quantum computer, Nature 393 (1998) 133–137.
[27] P.S. Fodor, J. Levy, Group IV solid state proposals for quantum computation, J. Phys.: Condens. Matter 18 (21) (2006) S745.
[28] F.A. Zwanenburg, A.S. Dzurak, A. Morello, M.Y. Simmons, L.C.L. Hollenberg, G. Klimeck, S. Rogge, S.N. Coppersmith, M.A. Eriksson, Silicon quantum electronics, Rev. Modern Phys. 85 (2013) 961–1019.
[29] A. Morello, J.J. Pla, F.A. Zwanenburg, K.W. Chan, K.Y. Tan, H. Huebl, M. Mottonen, C.D. Nugroho, C. Yang, J.A. van Donkelaar, A.D.C. Alves, D.N. Jamieson, C.C. Escott, L.C.L. Hollenberg, R.G. Clark, A.S. Dzurak, Single-shot readout of an electron spin in silicon, Nature 479 (2010) 345–353.
[30] D.D. Awschalom, L.C. Bassett, A.S. Dzurak, E.L. Hu, J.R. Petta, Quantum spintronics: engineering and manipulating atom-like spins in semiconductors, Science 339 (6124) (2013) 1174–1179.
[31] A.M. Tyryshkin, S. Tojo, J.J.L. Morton, H. Riemann, N.V. Abrosimov, P. Becker, H.-J. Pohl, T. Schenkel, M.L.W. Thewalt, K.M. Itoh, S.A. Lyon, Electron spin coherence exceeding seconds in high-purity silicon, Nat. Mater. 11 (2012) 143–147.
[32] K. Saeedi, S. Simmons, J.Z. Salvail, P. Dluhy, H. Riemann, N.V. Abrosimov, P. Becker, H.-J. Pohl, J.J.L. Morton, M.L.W. Thewalt, Room-temperature quantum bit storage exceeding 39 minutes using ionized donors in silicon-28, Science 342 (6160) (2013) 830–833.
[33] P.C. Maurer, G. Kucsko, C. Latta, L. Jiang, N.Y. Yao, S.D. Bennett, F. Pastawski, D. Hunger, N. Chisholm, M. Markham, D.J. Twitchen, J.I. Cirac, M.D. Lukin, Room-temperature quantum bit memory exceeding one second, Science 336 (6086) (2012) 1283–1286.
[34] J.L. Cheng, M.W. Wu, J. Fabian, Theory of the spin relaxation of conduction electrons in silicon, Phys. Rev. Lett. 104 (2010) 016601.
[35] P. Li, H. Dery, Spin–orbit symmetries of conduction electrons in silicon, Phys. Rev. Lett. 107 (2011) 107203.
[36] G. Schmidt, D. Ferrand, L.W. Molenkamp, A.T. Filip, B.J. van Wees, Fundamental obstacle for electrical spin injection from a ferromagnetic metal into a diffusive semiconductor, Phys. Rev. B 62 (2000) R4790–R4793.
[37] I. Appelbaum, B. Huang, D.J. Monsma, Electronic measurement and control of spin transport in silicon, Nature 447 (2007) 295–298.
[38] H.-J. Jang, J. Xu, J. Li, B. Huang, I. Appelbaum, Non-ohmic spin transport in $n$-type doped silicon, Phys. Rev. B 78 (2008) 165329.
[39] Y. Lu, J. Li, I. Appelbaum, Spin-polarized transient electron trapping in phosphorus-doped silicon, Phys. Rev. Lett. 106 (2011) 217202.
[40] J. Li, B. Huang, I. Appelbaum, Oblique hanle effect in semiconductor spin transport devices, Appl. Phys. Lett. 92 (14) (2008) 142507.
[41] B. Huang, D.J. Monsma, I. Appelbaum, Coherent spin transport through a 350 micron thick silicon wafer, Phys. Rev. Lett. 99 (2007) 177209.
[42] H. Dery1, P. Dalal, L. Cywinki, L. Sham, Spin-based logic in semiconductors for reconfigurable large-scale circuits, Nature 447 (2007) 573–576.
[43] Y. Song, H. Dery, Spin transport theory in ferromagnet/semiconductor systems with noncollinear magnetization configurations, Phys. Rev. B 81 (2010) 045321.
[44] B. Behin-Aein, D. Datta, S. Salahuddin, S. Datta, Proposal for an all-spin logic device with built-in memory, Nat. Nanotechnology 5 (2010) 266–270.
[45] J. Slonczewski, Current-driven excitation of magnetic multilayers, J. Magn. Magn. Mater. 159 (1–2) (1996) L1–L7.
[46] L. Berger, Emission of spin waves by a magnetic multilayer traversed by a current, Phys. Rev. B 54 (1996) 9353–9358.
[47] S. Sugahara, J. Nitta, Spin-transistor electronics: An overview and outlook, Proc. IEEE 98 (12) (2010) 2124–2154.
[48] E.I. Rashba, Theory of electrical spin injection: Tunnel contacts as a solution of the conductivity mismatch problem, Phys. Rev. B 62 (2000) R16267–R16270.
[49] B.T. Jonker, G. Kioseoglou, A.T. Hanbicki, C.H. Li., P.E. Thompson, Electrical spin-injection into silicon from a ferromagnetic metal/tunnel barrier contact, Nat. Phys. 3 (2007) 542–546.
[50] O.M.J. van't Erve, A.T. Hanbicki, M. Holub, C.H. Li, C. Awo-Affouda, P.E. Thompson, B.T. Jonker, Electrical injection and detection of spin-polarized carriers in silicon in a lateral transport geometry, Appl. Phys. Lett. 91 (21) (2007) 212109.
[51] S.P. Dash, S. Sharma, R.S. Patel, M.P. de Jong, R. Jansen, Electrical creation of spin polarization in silicon at room temperature, Nature 462 (2009) 491–494.
[52] C. Li, O. van't Erve, B. Jonker, Electrical injection and detection of spin accumulation in silicon at 500 K with magnetic metal/silicon dioxide contacts, Nat. Commun. 2 (2011) 245.
[53] A. Fert, H. Jaffrès, Conditions for efficient spin injection from a ferromagnetic metal into a semiconductor, Phys. Rev. B 64 (2001) 184420.
[54] K.M.B.-C. Min, C. Lodder, R. Jansen, Tunable spin tunnel contacts to silicon using low-work-function ferromagnets, Nat. Mater. 5 (2006) 817–822.
[55] O.M.J. van't Erve, A.L. Friedman, E. Cobas, C.H. Li, J.T. Robinson, B.T. Jonker, Low-resistance spin injection into silicon using graphene tunnel barriers, Nat. Nanotechnology 7 (2012) 737–742.
[56] R. Jansen, S.P. Dash, S. Sharma, B.C. Min, Silicon spintronics with ferromagnetic tunnel devices, Semicond. Sci. Technol. 27 (8) (2012) 083001.
[57] H. Dery, L. Cywiński, L.J. Sham, Lateral diffusive spin transport in layered structures, Phys. Rev. B 73 (2006) 041306.
[58] S. Yuasa, D.D. Djayaprawira, Giant tunnel magnetoresistance in magnetic tunnel junctions with a crystalline MgO (001) barrier, J. Phys. D 40 (21) (2007) R337.
[59] K.-R. Jeon, B.-C. Min, I.-J. Shin, C.-Y. Park, H.-S. Lee, Y.-H. Jo, S.-C. Shin, Electrical spin accumulation with improved bias voltage dependence in a crystalline CoFe/MgO/Si system, Appl. Phys. Lett. 98 (26) (2011) 262102.
[60] A. Spiesser, S. Sharma, H. Saito, R. Jansen, S. Yuasa, K. Ando, Electrical spin injection in $p$-type Si using Fe/MgO contacts, in: Proc. SPIE, vol. 8461, 2012, p. 84610K.
[61] S. Sharma, Electrical creation of spin polarization in silicon devices with magnetic tunnel contacts (Dissertation), Rijksuniversiteit Groningen, 2013.
[62] M. Tran, H. Jaffres, C. Deranlot, J.-M. George, A. Fert, A. Miard, A. Lemaitre, Enhancement of the spin accumulation at the interface between a spin-polarized tunnel junction and a semiconductor, Phys. Rev. Lett. 102 (2009) 036601.
[63] R. Jansen, A.M. Deac, H. Saito, S. Yuasa, Injection and detection of spin in a semiconductor by tunneling via interface states, Phys. Rev. B 85 (2012) 134420.
[64] Y. Song, H. Dery, Magnetic-field-modulated resonant tunneling in ferromagnetic–insulator–nonmagnetic junctions, Phys. Rev. Lett. 113 (2014) 047205.
[65] K.-R. Jeon, H. Saito, S. Yuasa, R. Jansen, Energy dispersion of tunnel spin polarization extracted from thermal and electrical spin currents, Phys. Rev. B 91 (2015) 155305.
[66] D.J. Hilton, C.L. Tang, optical orientation and femtosecond relaxation of spin-polarized holes in GaAs, Phys. Rev. Lett. 89 (2002) 146601.

[67] F. Pezzoli, F. Bottegoni, D. Trivedi, F. Ciccacci, A. Giorgioni, P. Li, S. Cecchi, E. Grilli, Y. Song, M. Guzzi, H. Dery, G. Isella, Optical spin injection and spin lifetime in Ge heterostructures, Phys. Rev. Lett. 108 (2012) 156603.

[68] S. Sharma, A. Spiesser, S.P. Dash, S. Iba, S. Watanabe, B.J. van Wees, H. Saito, S. Yuasa, R. Jansen, Anomalous scaling of spin accumulation in ferromagnetic tunnel devices with silicon and germanium, Phys. Rev. B 89 (2014) 075301.

[69] O. Txoperena, Y. Song, L. Qing, M. Gobbi, L.E. Hueso, H. Dery, F. Casanova, Impurity-assisted tunneling magnetoresistance under a weak magnetic field, Phys. Rev. Lett. 113 (2014) 146601.

[70] A.G. Swartz, S. Harashima, Y. Xie, D. Lu, B. Kim, C. Bell, Y. Hikita, H.Y. Hwang, Spin-dependent transport across $Co/LaAlO_3/SrTiO_3$ heterojunctions, Appl. Phys. Lett. 105 (3) (2014) 032406.

[71] H.N. Tinkey, P. Li, I. Appelbaum, Inelastic electron tunneling spectroscopy of local devices, Appl. Phys. Lett. 104 (23) (2014) 232410.

[72] A. Spiesser, H. Saito, R. Jansen, S. Yuasa, K. Ando, Large spin accumulation voltages in epitaxial $Mn_5Ge_3$ contacts on ge without an oxide tunnel barrier, Phys. Rev. B 90 (2014) 205213.

[73] T. Sasaki, T. Oikawa, T. Suzuki, M. Shiraishi, Y. Suzuki, K. Tagami, electrical spin injection into silicon using MgO tunnel barrier, Appl. Phys. Express 2 (5) (2009) 053003.

[74] Y. Ando, K. Hamaya, K. Kasahara, Y. Kishi, K. Ueda, K. Sawano, T. Sadoh, M. Miyao, Electrical injection and detection of spin-polarized electrons in silicon through an $Fe_3Si/Si$ Schottky tunnel barrier, Appl. Phys. Lett. 94 (18) (2009) 182105.

[75] T. Sasaki, T. Oikawa, T. Suzuki, M. Shiraishi, Y. Suzuki, K. Noguchi, Temperature dependence of spin diffusion length in silicon by hanle-type spin precession, Appl. Phys. Lett. 96 (12) (2010) 122101.

[76] Y. Ando, K. Kasahara, K. Yamane, K. Hamaya, K. Sawano, T. Kimura, M. Miyao, Comparison of nonlocal and local magnetoresistance signals in laterally fabricated $Fe_3Si/Si$ spin-valve devices, Appl. Phys. 3 (9) (2010) 093001.

[77] T. Sasaki, T. Oikawa, M. Shiraishi, Y. Suzuki, K. Noguchi, Comparison of spin signals in silicon between nonlocal four-terminal and three-terminal methods, Appl. Phys. Lett. 98 (1) (2011) 012508.

[78] Y. Ando, K. Kasahara, K. Yamane, Y. Baba, Y. Maeda, Y. Hoshi, K. Sawano, M. Miyao, K. Hamaya, Bias current dependence of spin accumulation signals in a silicon channel detected by a schottky tunnel contact, Appl. Phys. Lett. 99 (1) (2011) 012113.

[79] M. Shiraishi, Y. Honda, E. Shikoh, Y. Suzuki, T. Shinjo, T. Sasaki, T. Oikawa, K. Noguchi, T. Suzuki, Spin transport properties in silicon in a nonlocal geometry, Phys. Rev. B 83 (2011) 241204.

[80] Y. Ando, Y. Maeda, K. Kasahara, S. Yamada, K. Masaki, Y. Hoshi, K. Sawano, K. Izunome, A. Sakai, M. Miyao, K. Hamaya, Electric-field control of spin accumulation signals in silicon at room temperature, Appl. Phys. Lett. 99 (13) (2011) 132511.

[81] T. Suzuki, T. Sasaki, T. Oikawa, M. Shiraishi, Y. Suzuki, K. Noguchi, Room-temperature electron spin transport in a highly doped Si channel, Appl. Phys. Express 4 (2011) 023003.

[82] Y. Ando, K. Kasahara, S. Yamada, Y. Maeda, K. Masaki, Y. Hoshi, K. Sawano, M. Miyao, K. Hamaya, Temperature evolution of spin accumulation detected electrically in a nondegenerate silicon channel, Phys. Rev. B 85 (2012) 035320.

[83] T. Sasaki, Y. Ando, M. Kameno, T. Tahara, H. Koike, T. Oikawa, T. Suzuki, M. Shiraishi, Spin transport in nondegenerate si with a spin MOSFET structure at room temperature, Phys. Rev. Appl. 2 (2014) 034005.

[84] T. Sasaki, T. Suzuki, Y. Ando, H. Koike, T. Oikawa, Y. Suzuki, M. Shiraishi, Local magnetoresistance in Fe/MgO/Si lateral spin valve at room temperature, Appl. Phys. Lett. 104 (5) (2014) 052404.

[85] J.-C.L. Breton, S. Sharma, H. Saito, S. Yuasa, R. Jansen, Thermal spin current from a ferromagnet to silicon by seebeck spin tunnelling, Nature 475 (2011) 82–85.

[86] Y. Tserkovnyak, A. Brataas, G.E.W. Bauer, Enhanced gilbert damping in thin ferromagnetic films, Phys. Rev. Lett. 88 (2002) 117601.

[87] Y. Tserkovnyak, A. Brataas, G.E.W. Bauer, Spin pumping and magnetization dynamics in metallic multilayers, Phys. Rev. B 66 (2002) 224403.

[88] K. Ando, S. Takahashi, J. Ieda, H. Kurebayashi, T. Trypiniotis, C.H.W. Barnes, S. Maekawa, E. Saitoh, Electrically tunable spin injector free from the impedance mismatch problem, Nat. Mater. 10 (2011) 655–659.

[89] Y. Ando, K. Ichiba, S. Yamada, E. Shikoh, T. Shinjo, K. Hamaya, M. Shiraishi, Giant enhancement of spin pumping efficiency using $Fe_3Si$ ferromagnet, Phys. Rev. B 88 (2013) 140406.

[90] E. Shikoh, K. Ando, K. Kubo, E. Saitoh, T. Shinjo, M. Shiraishi, Spin-pump-induced spin transport in *p*-type si at room temperature, Phys. Rev. Lett. 110 (2013) 127201.

[91] P. Nemec, V. Novak, N. Tesarova, E. Rozkotova, H. Reichlova, D. Butkovicova, F. Trojanek, K. Olejnik, P. Maly, R. Campion, B. Gallagher, J. Sinova, T. Jungwirth, The essential role of carefully optimized synthesis for elucidating intrinsic material properties of (Ga, Mn)As, Nat. Commun. 4 (2013) 1422.

[92] Y.F. Lee, F. Wu, R. Kumar, F. Hunte, J. Schwartz, J. Narayan, Epitaxial integration of dilute magnetic semiconductor $Sr_3SnO$ with Si (001), Appl. Phys. Lett. 103 (11) (2013) 112101.

[93] Y. Takamura, K. Hayashi, Y. Shuto, S. Sugahara, Fabrication of high-quality $Co_2FeSi/SiO_xN_y/Si(100)$ tunnel contacts using radical-oxynitridation-formed $SiO_xN_y$ barrier for Si-based spin transistors, J. Electron. Mater. 41 (5) (2012) 954–958.

[94] Y. Song, O. Chalaev, H. Dery, Donor-driven spin relaxation in multivalley semiconductors, Phys. Rev. Lett. 113 (2014) 167201.

[95] Y. Song, H. Dery, Analysis of phonon-induced spin relaxation processes in silicon, Phys. Rev. B 86 (2012) 085201.

[96] H. Dery, Y. Song, P. Li, I. Zutic, Silicon spin communication, Appl. Phys. Lett. 99 (8) (2011) 08250.

[97] T. Ando, A.B. Fowler, F. Stern, Electronic properties of two-dimensional systems, Rev. Modern Phys. 54 (2) (1982) 437–672.

[98] J. Li, I. Appelbaum, Modeling spin transport in electrostatically-gated lateral-channel silicon devices: Role of interfacial spin relaxation, Phys. Rev. B 84 (2011) 165318.

[99] J. Li, I. Appelbaum, Lateral spin transport through bulk silicon, Appl. Phys. Lett. 100 (16) (2012) 162408.

[100] J. Li, L. Qing, H. Dery, I. Appelbaum, Field-induced negative differential spin lifetime in silicon, Phys. Rev. Lett. 108 (2012) 157201.

[101] G.L. Bir, G.E. Pikus, Symmetry and Strain-induced Effects in Semiconductors, John Willey & Sons, 1974.

[102] P. Yu, M. Cardona, Fundamentals of Semiconductors, Springer, 2003.

[103] J.C. Hensel, H. Hasegawa, M. Nakayama, Cyclotron resonance in uniaxially stressed silicon. II. Nature of the covalent bond, Phys. Rev. 138 (1A) (1965) A225–A238.

[104] V. Sverdlov, Strain-induced Effects in Advanced MOSFETs, Springer, 2011.

[105] D. Osintsev, O. Baumgartner, Z. Stanojevic, V. Sverdlov, S. Selberherr, Subband splitting and surface roughness induced spin relaxation in (001) silicon SOI MOSFETs, Solid-State Electron. 90 (2013) 34–38.

[106] I. Balslev, Influence of uniaxial stress on the indirect absorption edge in silicon and germanium, Phys. Rev. 143 (1966) 636–647.

[107] E. Ungersboeck, S. Dhar, G. Karlowatz, V. Sverdlov, H. Kosina, S. Selberherr, The effect of general strain on band structure and electron mobility of silicon, IEEE Trans. Electron Devices 54 (2007) 2183–2190.

[108] K. Uchida, T. Krishnamohan, K.C. Saraswat, Y. Nishi, Physical mechanisms of electron mobility enhancement in uniaxial stressed MOSFETs and impact of uniaxial stress engineering in ballistic regime, in: International Electron Devices Meeting, IEDM, 2005, pp. 129–132.

[109] D. Osintsev, Modeling spintronic effects in silicon (Dissertation), Institute for Microelectronics, TU Wien, 2014.

[110] C.H. Yang, A. Rossi, R. Ruskov, N.S. Lai, F.A. Mohiyaddin, S. Lee, C. Tahan, G. Klimeck, A. Morello, A.S. Dzurak, Spin-valley lifetimes in a silicon quantum dot with tunable valley splitting, Nature Commun. 4 (2013) 2069.

[111] D. Rideau, M. Feraille, M. Michaillat, Y. Niquet, C. Tavernier, H. Jaouen, On the validity of the effective mass approximation and the luttinger $\mathbf{k} \cdot \mathbf{p}$ model in fully depleted SOI MOSFETs, Solid-State Electron. 53 (4) (2009) 452–461.

[112] M.O. Nestoklon, L.E. Golub, E.L. Ivchenko, Spin and valley–orbit splittings in SiGeSi heterostructures, Phys. Rev. B 73 (2006) 235334.

[113] D. Osintsev, V. Sverdlov, N. Neophytou, S. Selberherr, Valley splitting and spin lifetime enhancement in strained silicon heterostructures, in: International Mauterndorf Winterschool on New Developments in Solid State Physics, Mauterndorf, Austria, 2014, pp. 88–89.

[114] K. Takashina, Y. Ono, A. Fujiwara, Y. Takahashi, Y. Hirayama, Valley polarization in Si (100) at zero magnetic field, Phys. Rev. Lett. 96 (2006) 236801.

[115] A. Dusko, A.L. Saraiva, B. Koiller, Splitting valleys in Si/SiO$_2$: identification and control of interface states, Phys. Rev. B 89 (2014) 205307.
[116] M.V. Fischetti, Z. Ren, P.M. Solomon, M. Yang, K. Rim, Six-band $\mathbf{k} \cdot \mathbf{p}$ calculation of the hole mobility in silicon inversion layers: dependence on surface orientation, strain, and silicon thickness, J. Appl. Phys. 94 (2) (2003) 1079–1095.
[117] C. Jacoboni, L. Reggiani, The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials, Rev. Modern Phys. 55 (3) (1983) 645–705.
[118] M.V. Fischetti, S.E. Laux, Monte Carlo study of electron transport in silicon inversion layers, Phys. Rev. B 48 (4) (1993) 2244–2274.
[119] Y. Bychkov, E. Rashba, Properties of a 2D electron gas with lifted spectral degeneracy, JETP Lett. 39 (2) (1984) 78–81.
[120] S. Giglberger, L.E. Golub, V.V. Bel'kov, S.N. Danilov, D. Schuh, C. Gerl, F. Rohlfing, J. Stahl, W. Wegscheider, D. Weiss, W. Prettl, S.D. Ganichev, Rashba and Dresselhaus spin splittings in semiconductor quantum wells measured by spin photocurrents, Phys. Rev. B 75 (2007) 035327.
[121] J.-M. Jancu, J.-C. Girard, M.O. Nestoklon, A. Lemaître, F. Glas, Z.Z. Wang, P. Voisin, STM images of subsurface Mn atoms in GaAs: Evidence of hybridization of surface and impurity states, Phys. Rev. Lett. 101 (2008) 196801.
[122] M. Prada, G. Klimeck, R. Joynt, Spin–orbit splittings in Si/SiGe quantum wells: from ideal Si membranes to realistic heterostructures, New J. Phys. 13 (1) (2011) 013009.
[123] Z. Wilamowski, W. Jantsch, Suppression of spin relaxation of conduction electrons by cyclotron motion, Phys. Rev. B 69 (2004) 035328.
[124] A. Bournel, P. Dollfus, P. Bruno, B. Hesto, Gate-induced spin precession in an In$_{0.53}$Ga$_{0.47}$As two dimensional electron gas, Eur. Phys. J. Appl. Phys. 4 (1) (1998) 1–4.
[125] M. Cahay, S. Bandyopadhyay, Phase-coherent quantum mechanical spin transport in a weakly disordered quasi-one-dimensional channel, Phys. Rev. B 69 (2004) 045303.
[126] K.-M. Jiang, R. Zhang, J. Yang, C.-X. Yue, Z.-Y. Sun, Tunneling magnetoresistance properties in ballistic spin field-effect transistors, IEEE Trans. Electron Devices 57 (8) (2010) 2005–2012.
[127] D. Osintsev, V. Sverdlov, Z. Stanojeviç, A. Makarov, S. Selberherr, Temperature dependence of the transport properties of spin field-effect transistors built with InAs and Si channels, Solid-State Electron. 71 (2012) 25–29.
[128] H. Tsuchiya, H. Ando, S. Sawamoto, T. Maegawa, T. Hara, H. Yao, M. Ogawa, Comparisons of performance potentials of silicon nanowire and graphene nanoribbon MOSFETs considering first-principles bandstructure effects, IEEE Trans. Electron Devices 57 (2) (2010) 406–414.
[129] M. Karner, A. Gehring, S. Holzer, M. Pourfath, M. Wagner, H. Kosina, T. Grasser, S. Selberherr, VSP-A Multi-Purpose Schrödinger-Poisson Solver for TCAD applications, in: International Workshop on Computational Electronics, IWCE, 2006, pp. 255–256.
[130] S.-C. Chang, S. Manipatruni, D. Nikonov, I. Young, A. Naeemi, Design and analysis of Si interconnects for all-spin logic, IEEE Trans. Magn. 50 (9) (2014) 1–13.
[131] Z.G. Yu, M.E. Flatté, Spin diffusion and injection in semiconductor structures: electric field effects, Phys. Rev. B 66 (2002) 235302.
[132] T. Marukame, T. Inokuchi, M. Ishikawa, H. Sugiyama, Y. Saito, Read/write operation of spin-based MOSFET using highly spin-polarized ferromagnet/MgO tunnel barrier for reconfigurable logic devices, in: International Electron Devices Meeting, IEDM, 2009, pp. 1–4.
[133] A. Makarov, V. Sverdlov, D. Osintsev, S. Selberherr, Reduction of switching time in pentalayer magnetic tunnel junctions with a composite-free layer, Phys. Status Solidi - Rapid Res. Lett. 5 (12) (2011) 420–422.
[134] A.V. Khvalkovskiy, D. Apalkov, S. Watts, R. Chepulskii, R.S. Beach, A. Ong, X. Tang, A. Driskill-Smith, W.H. Butler, P.B. Visscher, D. Lottis, E. Chen, V. Nikitin, M. Krounbi, Basic principles of STT-MRAM cell operation in memory arrays, J. Phys. D 46 (7) (2013) 074001.
[135] G.D. Fuchs, I.N. Krivorotov, P.M. Braganca, N.C. Emley, A.G.F. Garcia, D.C. Ralph, R.A. Buhrman, Adjustable spin torque in magnetic tunnel junctions with two fixed layers, Appl. Phys. Lett. 86 (15) (2005) 152509.
[136] B. Dieny, Spin transfer torque random access memories, in: CMOS Emerging Technologies Research, CMOSETR, 2014, p. 18.
[137] I.L. Prejbeanu, S. Bandiera, J. Alvarez-Hérault, R.C. Sousa, B. Dieny, J.-P. Nozières, Thermally assisted MRAMs: Ultimate scalability and logic functionalities, J. Phys. D 46 (7) (2013) 074002.
[138] N.N. Mojumder, C. Augustine, D.E. Nikonov, K. Roy, Effect of quantum confinement on spin transport and magnetization dynamics in dual barrier spin transfer torque magnetic tunnel junctions, J. Appl. Phys. 108 (10) (2010) 104306.
[139] S. Ikeda, K. Miura, H. Yamamoto, K. Mizunuma, H.D. Gan, M. Endo, S. Kanai, J. Hayakawa, F. Matsukura, H. Ohno, A perpendicular-anisotropy CoFeB–MgO magnetic tunnel junction, Nat. Mater. 9 (9) (2010) 721–724.
[140] H. Sato, M. Yamanouchi, S. Ikeda, S. Fukami, F. Matsukura, H. Ohno, Perpendicular-anisotropy CoFeB–MgO magnetic tunnel junctions with a MgO/CoFeB/Ta/CoFeB/MgO recording structure, Appl. Phys. Lett. 101 (2) (2012) 022414.
[141] D.-S. Wang, S.-Y. Lai, T.-Y. Lin, C.-W. Chien, D. Ellsworth, L.-W. Wang, J.-W. Liao, L. Lu, Y.-H. Wang, M. Wu, C.-H. Lai, High thermal stability and low gilbert damping constant of CoFeB/MgO bilayer with perpendicular magnetic anisotropy by Al capping and rapid thermal annealing, Appl. Phys. Lett. 104 (14) (2014) 142402.
[142] J.C. Slonczewski, Currents, torques, and polarization factors in magnetic tunnel junctions, Phys. Rev. B 71 (2005) 024411.
[143] J.E. Miltat, M.J. Donahue, Numerical micromagnetics: Finite difference methods handbook of magnetism and advanced magnetic materials, in: Magnetism and Advanced Magnetic Materials, John Wiley & Sons, 2007, pp. 1–22.
[144] A. Kàkay, Numerical investigations of micromagnetic structures (Dissertation), Hungarian Academy of Science, 2005.
[145] K. Ito, T. Devolder, C. Chappert, M.J. Carey, J.A. Katine, Probabilistic behavior in subnanosecond spin transfer torque switching, J. Appl. Phys. 99 (8) (2006) 08G519.
[146] L. Torres, L. Lopes-Dias, E. Martinez, O. Alejos, Micromagnetic dynamic computations including eddy currents, IEEE Trans. Magn. 39 (2003) 5.
[147] A. Makarov, V. Sverdlov, S. Selberherr, Magnetic tunnel junctions with a composite free layer: a new concept for future universal memory, in: Future Trends in Microelectronics, John Wiley & Sons, 2013, pp. 93–101.
[148] A. Makarov, Modeling of emerging resistive switching based memory cells (Dissertation), Institute for Microelectronics, TU Wien, 2014.
[149] H. Sato, M. Yamanouchi, K. Miura, S. Ikeda, H.D. Gan, K. Mizunuma, R. Koizumi, F. Matsukura, H. Ohno, Junction size effect on switching current and thermal stability in CoFeB/MgO perpendicular magnetic tunnel junctions, Appl. Phys. Lett. 99 (4) (2011) 042501.
[150] S. Ishikawa, H. Sato, M. Yamanouchi, S. Ikeda, S. Fukami, F. Matsukura, H. Ohno, Co/Pt multilayer-based magnetic tunnel junctions with a CoFeB/Ta insertion layer, J. Appl. Phys. 115 (17) (2014) 17C719.
[151] H. Sato, E.C.I. Enobio, M. Yamanouchi, S. Ikeda, S. Fukami, S. Kanai, F. Matsukura, H. Ohno, Properties of magnetic tunnel junctions with a MgO/CoFeB/Ta/CoFeB/MgO recording structure down to junction diameter of 11 nm, Appl. Phys. Lett. 105 (6) (2014) 062403.
[152] L. Liu, C.-F. Pai, Y. Li, H.W. Tseng, D.C. Ralph, R.A. Buhrman, Spin-torque switching with the giant spin hall effect of tantalum, Science 336 (6081) (2012) 555–558.
[153] M. Yamanouchi, L. Chen, J. Kim, M. Hayashi, H. Sato, S. Fukami, S. Ikeda, F. Matsukura, H. Ohno, Three terminal magnetic tunnel junction utilizing the spin hall effect of iridium-doped copper, Appl. Phys. Lett. 102 (21) (2013) 212408.
[154] C. Zhang, M. Yamanouchi, H. Sato, S. Fukami, S. Ikeda, F. Matsukura, H. Ohno, Magnetization reversal induced by in-plane current in Ta/CoFeB/MgO structures with perpendicular magnetic easy axis, J. Appl. Phys. 115 (17) (2014) 17C714.
[155] S. Kanai, Y. Nakatani, M. Yamanouchi, S. Ikeda, H. Sato, F. Matsukura, H. Ohno, Magnetization switching in a CoFeB/MgO magnetic tunnel junction by combining spin-transfer torque and electric field-effect, Appl. Phys. Lett. 104 (21) (2014) 212406.
[156] L. Liu, C.-F. Pai, D.C. Ralph, R.A. Buhrman, Gate voltage modulation of spin-hall-torque-driven magnetic switching, ArXiv e-prints arXiv:1209.0962.
[157] C. Li, O. van't Erve, J.T. Robinson, Y. Liu, L. Li, B. Jonker, Electrical detection of charge-current-induced spin polarization due to spin-momentum locking in Bi$_2$Se$_3$, Nat. Nanotechnology 9 (2014) 218.
[158] A.R. Mellnik, J.S. Lee, A. Richardella, J.L. Grab, P.J. Mintun, M.H. Fischer, A. Vaezi, A. Manchon, E.-A. Kim, N. Samarth, D.C. Ralph, Spin-transfer torque generated by a topological insulator, Nature 511 (2014) 449–451.
[159] Y. Fan, P. Upadhyaya, X. Kou, M. Lang, S. Takei, Z. Wang, J. Tang, L. He, L.-T. Chang, M. Montazeri, G. Yu, W. Jiang, T. Nie, R.N. Schwartz, Y. Tserkovnyak, K.L. Wang, Magnetization switching through giant spin–orbit torque in a magnetically doped topological insulator heterostructure, Nat. Mater. 13 (2014) 699–704.

[160] A. Lyle, J. Harms, S. Patil, X. Yao, D.J. Lilja, J.-P. Wang, Direct communication between magnetic tunnel junctions for nonvolatile logic fan-out architecture, Appl. Phys. Lett. 97 (15) (2010) 152504.

[161] A. Lyle, S. Patil, J. Harms, B. Glass, X. Yao, D. Lilja, J. Wang, Magnetic tunnel junction logic architecture for realization of simultaneous computation and communication, IEEE Trans. Magn. 47 (10) (2011) 2970–2973.

[162] H. Mahmoudi, T. Windbacher, V. Sverdlov, S. Selberherr, Implication logic gates using spin-transfer-torque-operated magnetic tunnel junctions for intrinsic logic-in-memory, Solid-State Electron. 84 (2013) 191–197.

[163] C. Augustine, N. Mojumder, X. Fong, H. Choday, S.P. Park, K. Roy, STT-MRAM for future universal memories: Perspective and prospective, in: 2012 28th International Conference on Microelectronics, MIEL, 2012, pp. 349–355.

[164] T. Endoh, STT-MRAM technology and its NV-logic applications for ultimate power management, in: CMOS Emerging Technologies Research, CMOSETR, 2014, p. 14.

[165] W. Zhao, E. Belhaire, C. Chappert, F. Jacquet, P. Mazoyer, New non-volatile logic based on spin-MTJ, Phys. Status Solidi A 205 (6) (2008) 1373–1377.

[166] M. Natsui, D. Suzuki, N. Sakimura, R. Nebashi, Y. Tsuji, A. Morioka, T. Sugibayashi, S. Miura, H. Honjo, K. Kinoshita, S. Ikeda, T. Endoh, H. Ohno, T. Hanyu, Nonvolatile logic-in-memory array processor in 90 nm MTJ/MOS achieving 75% leakage reduction using cycle-based power gating, in: International Solid-State Circuits Conference, ISSCC, 2013, pp. 194–195.

[167] T. Hanyu, Challenge of MTJ/MOS-hybrid logic-in-memory architecture for nonvolatile VLSI processor, in: International Symposium on Circuits and Systems, ISCAS, 2013, pp. 117–120.

[168] D. Suzuki, M. Natsui, T. Endoh, H. Ohno, T. Hanyu, Six-input lookup table circuit with 62% fewer transistors using nonvolatile logic-in-memory architecture with series/parallel-connected magnetic tunnel junctions, J. Appl. Phys. 111 (7) (2012) 07E318.

[169] M. Hosomi, H. Yamagishi, T. Yamamoto, K. Bessho, Y. Higo, K. Yamane, H. Yamada, M. Shoji, H. Hachino, C. Fukumoto, H. Nagao, H. Kano, A novel nonvolatile memory with spin torque transfer magnetization switching: spin-RAM, in: International Electron Devices Meeting, IEDM, 2005, pp. 459–462.

[170] J. Borghetti, G. Snider, P. Kuekes, J. Yang, D. Stewart, R. Williams, Memristive switches enable stateful logic operations via material implication, Nature 464 (2010) 873–876.

[171] J.J. Yang, M.-X. Zhang, J.P. Strachan, F. Miao, M.D. Pickett, R.D. Kelley, G. Medeiros-Ribeiro, R.S. Williams, High switching endurance in $TaO_x$ memristive devices, Appl. Phys. Lett. 97 (23) (2010) 232102.

[172] J. Yang, D. Strukov, D. Stewart, Memristive devices for computing, Nature Nano 8 (2013) 13–24.

[173] H. Mahmoudi, V. Sverdlov, S. Selberherr, State drift optimization of memristive stateful IMP logic gates, in: International Workshop on Computational Electronics, IWCE, 2012, pp. 243–244.

[174] S. Kvatinsky, A. Kolodny, U. Weiser, E. Friedman, Memristor-based imply logic design procedure, in: International Conference on Computer Design, ICCD, 2011, pp. 142–147.

[175] H. Mahmoudi, T. Windbacher, V. Sverdlov, S. Selberherr, MRAM-based Logic array for large-scale non-volatile logic-in-memory applications, in: International Symposium on Nanoscale Architectures, NANOARCH, 2013, pp. 26–27.

[176] H. Mahmoudi, T. Windbacher, V. Sverdlov, S. Selberherr, Reliability analysis and comparison of implication and reprogrammable logic gates in magnetic tunnel junction logic circuits, IEEE Trans. Magn. 49 (12) (2013) 5620–5628.