Non-Volatility by Spin in Modern Nanoelectronics

V. Sverdlov, A. Makarov, J. Weinbub, and S. Selberherr

Abstract - Continuous miniaturization of semiconductor devices has been the main driver behind the outstanding increase of speed and performance of integrated circuits. In addition to a harmful active power penalty, small device dimensions result in rapidly rising leakages and fast growing stand-by power. The critical high power consumption becomes incompatible with the global demands to sustain and accelerate the vital industrial growth, and an introduction of new solutions for energy efficient computations becomes paramount.

A highly attractive option to reduce power consumption is to introduce non-volatility in integrated circuits. Preserving the data without power eliminates the need for refreshment cycles and related leakages as well as the necessity to initialize the data in temporarily unused parts of the circuit. Spin transistors are promising devices, with the charge-based functionality complemented by the electron spin. The non-volatility is introduced by making the source and drain ferromagnetic. Recent advances in resolving several fundamental problems including spin injection from a metal ferromagnet to a semiconductor, spin propagation and relaxation, as well as spin manipulation by the electric field, resulted in successful demonstrations of such devices. However, the small relative current ratio between parallel/anti-parallel source and drain alignment at room temperature remains a substantial challenge preventing these devices from entering the market in the near future.

In contrast, a magnetic tunnel junction is an excellent candidate for realizing power-reducing approaches, as it possesses a simple structure, long retention time, high endurance, fast operation speed, and yields high integration density. Magnetic tunnel junctions with large magnetoresistance ratio are perfectly suited as key elements of non-volatile magnetoresistive memory compatible with the complementary metal-oxide-semiconductor technology and capable to replace dynamic and potentially static random access memories. We review the present status of the technology, remaining challenges, as well as approaches to resolve the remaining problems.

Regarding active power reduction, delegating data processing capabilities into the non-volatile segment and combining non-volatile elements with CMOS allows for efficient power gating. It also paves the way for a new low-power and high-performance computing paradigm-based on an intrinsic logic-in-memory architecture, where the same non-volatile elements are used to store and to process the information.

I. Introduction

The breathtaking increase in performance and speed of integrated circuits has been enabled by continuous miniaturization of complementary metal-oxide semiconductor (CMOS) devices. On this exciting path numerous outstanding technological challenges have been resolved. Among the most crucial technological changes recently adopted by the semiconductor industry to boost CMOS performance while maintaining gate control over the semiconductor channel are the introduction of strain [1], high-k gate dielectrics and metal gates [2], and a three-dimensional (3D) tri-gate transistor architecture [3-5]. The successes and innovative solutions developed for the microelectronics technology have been always supported by sophisticated simulation tools which allow reducing the research and development costs by 35-40% [6].

Although transistor sizes are scaled down, the on-currents cannot be further decreased due to the need to charge/discharge the load capacitances and to maintain the clock frequency which is saturated at approximately 4.0 GHz. Increasing the clock frequency results in an active power penalty, while continuous transistor scaling results in growing leakages and stand-by power. Novel revolutionary approaches are desperately needed in the long run to sustain the vital societal and industrial progress in computing performance whilst simultaneously reducing power consumption.

The ultimate solution to one of the primary issues – the power reduction – is to introduce non-volatility into the circuits. Non-volatility is crucial for eliminating the leakage power dissipation and data refreshment cycles. Apart from stand-alone applications, it is particularly promising to use non-volatility in the main computer memory as a replacement of conventional volatile CMOS-based dynamic random-access memory (DRAM) [7]. This will help bridging the speed gap between the last level cashes and main memory, as CMOS SRAM is much faster compared to CMOS DRAM. In modern multicore processors, much of the energy consumption appears in the hierarchical multi-level cache memory structure. To reduce this energy consumption, a viable approach is to replace SRAM with non-volatile elements, which will also support a reduced memory cell size [7].

To be competitive with traditional volatile memories, emerging non-volatile memories must offer a fast switching time, a high integration density supported with good scalability, a long retention time, a high endurance, and a low power consumption. At the same time, they must possess a simple structure to reduce fabrication costs and the new non-volatile circuit elements must be compatible with CMOS technology in order to benefit from the advantages provided by the well-developed CMOS fabrication technology.

A spin field-effect transistor (SpinFET) is a future semiconductor device employing the electron spin. The

V. Sverdlov, A. Makarov, and S. Selberherr are with the Institute for Microelectronics, TU Wien

J. Weinbub is with the Christian Doppler Laboratory for HP TCAD at the Institute for Microelectronics, TU Wien Gußhausstraße 27-29, 1040 Wien, Austria
E-mail: {last name}@iue.tuwien.ac.at

978-1-5386-2562-0/17/$31.00 ©2017 IEEE

non-volatility in SpinFETs is added by replacing the non-magnetic source and drain in a FET by its ferromagnetic counterparts. Metallic ferromagnetic contacts serve not only as an injector/detector of the spin-polarized electron current in the channel, they also provide an additional current modulation due to their capabilities to inject/detect spins [8]. The electron current gets enhanced in the case of parallel alignment between the source/drain electrodes, while the current is suppressed for anti-parallel magnetization alignment [8]. As the magnetization of the source/drain can be manipulated by means of an external magnetic field and/or current (by means of the spin-transfer torque), the two on-current states for parallel/anti-parallel magnetization alignment potentially enable reprogrammable logic [9]. The relative magnetization orientation between source and drain is preserved without external power, which makes reprogrammable logic (partly) non-volatile. We briefly discuss recent advances and remaining challenges to realize SpinFET-based logic.

Magnetoresistive random access memory (MRAM) and in particular spin transfer torque (STT) MRAM is a perfect candidate for future universal memory applications. MRAM is CMOS-compatible which increases its potential to replace processor-embedded SRAM and DRAM. With STT MRAM currently emerging as a commercial product for stand-alone applications, it will be critically important to introduce STT MRAM in the main computer memory. This would create a new innovative multi-billion dollar industry and will sustain the breathtaking path of electronics by delivering cheaper, faster, and environmentally friendlier compact and mobile devices. Bringing STT MRAM into the vast computer memory market for embedded and stand-alone applications in traditional high-performance and low-power mobile platforms will result in an exponential growth of the non-volatile memory market share in the near future.

A critical aspect of the currently used STT MRAM technology is a relatively high switching current. This might prevent the MRAM from successfully entering the computer memory market. The problem of high switching current and large active writing energy may jeopardize the advantages provided by non-volatility. Several plausible approaches to reduce the current were suggested including the replacement of the in-plane magnetization orientation in magnetic tunnel junctions (MTJs) with perpendicular magnetization, the use of composite free recording layers, decoupling the write and read current paths, controlling magnetization by voltage, and employing new materials with improved properties and characteristics.

II. SpinFET and SpinMOSFET

In a SpinFET [8] the electrons with spin 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 and the electron spin polarization at the end of the semiconductor

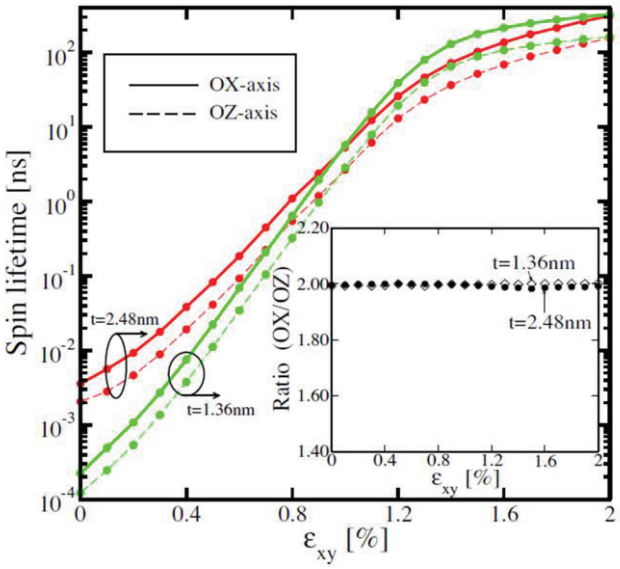

Fig. 1. Spin lifetime in (001) Si film as a function of shear strain, for two film thicknesses. The spin lifetime is a factor of two longer for spins injected in-plane as compared to that for spins injected perpendicular to the film (Inset).

channel. The electron spin orientation at the end of the channel is determined by the source magnetization and can be additionally manipulated by tuning the strength of the effective spin-orbit interaction in the channel by the gate voltage. In contrast to the electron charge, the injected spin relaxes to its equilibrium zero value while propagating through the channel. Spin relaxation is an important characteristic as it reduces the current modulation and affects the SpinFET functionality. It is governed by the spin-orbit interaction (SOI) and scattering, both spin-dependent and spin-independent, and manifests itself differently in semiconductor crystals with or without the inversion symmetry [10].

In crystals obeying the inversion symmetry (silicon, germanium) the spin relaxation is governed be the Elliott-Yafet mechanism [10]. The wave function with a fixed spin projection is not an eigenstate of the microscopic Hamiltonian because of the electron momentum-dependent SOI. The SOI forces the eigenstate wave function to possess a small but finite contribution with an opposite spin projection. Therefore, the finite amplitude to flip the electron spin appears at every spin-independent scattering event – the Elliott process [11]. This is complemented by the Yafet spin-flip events due to SOI-dependent electron-phonon scattering. In silicon the electron spin relaxation is determined by the inter-valley transitions [11] and can be efficiently controlled by stress [12]. In silicon channels, uniaxial [110] stress generates shear strain, which suppresses the spin relaxation [13] by lifting the valley degeneracy [14]. In silicon films the spin lifetime depends on the spin injection direction [15] shown in Fig.1.

In III-V materials the crystal inversion symmetry is broken. This results in the up and down spin states with the same electron momentum having different energies, and

the spin relaxation is governed by the Dyakonov-Perel mechanism [10]. However, the SOI does not always play a detrimental role. The inversion symmetry in the channel can be violated by applying the gate voltage. In this case the strength of the effective SOI depends on the effective electric field perpendicular to the channel [16]. The strength of the gate voltage-dependent SOI can be used to modulate the current between the ferromagnetic source and drain by means of an additional spin modulation in the channel resulting in a different spin orientation relative to the drain as compared to the case without SOI [8]. Importantly, as the strength of the SOI in the channel depends on the effective electric field, the suggested method provides purely electrical means of manipulating the electron spins and thus the current in the channel.

The voltage-induced SOI in III-V materials can also be used for an efficient spin injection into the channel from point contacts [17]. Additional gates create the point contacts to the two-dimensional (2D) electron gas by confining the gas in the III-V channel under the gates. Applying a voltage between the two gates generates the spin-orbit Rashba field perpendicular to the point contact. By properly tuning the chemical potential in the contact one can achieve that, due to this SOI, all electrons moving to the right are spin-polarized (while electrons moving to the left are polarized in the opposite direction). Thereby an efficient and purely electrical spin injection/detection is achieved. Using this injection scheme, the ever first reliable demonstration [17] of a working SpinFET [8] suggested in 1990 was achieved.

New 2D materials (graphene, transition metal dichalcogenides) are attractive for future microelectronics applications. Recently, a new concept for realizing a spin switch with a graphene channel was demonstrated [18]. Spin-polarized electrons are injected into the graphene, a good spin conductor due to low SOI, and reach the drain electrode, if the electrochemical potential in MoS_2 is tuned into the forbidden energy gap. In this case the electrons do not enter MoS_2. The situation is completely changed, if the electrochemical potential is tuned by the gate into the MoS_2 conduction band. In this case a parallel path for electrons through MoS_2 with high SOI is open, which results in strong spin relaxation, so that the current reaching the drain is not spin-polarized.

Regardless of the successful demonstration of the SpinFET, the conductance modulations were only resolved at temperatures far below 300K. Recently, the first successful demonstration of a silicon spin metal-oxide-semiconductor field-effect transistor (SpinMOSFET) at room temperature [19] was presented. In silicon the strength of the Rashba SOI is much smaller compared to that in III-V semiconductors. Therefore, the SOI cannot be used for spin manipulation. Thus, the current modulation in the SpinMOSFET is only achieved by altering the relative magnetization between the ferromagnetic source and drain. This way, a high difference between the on-currents in parallel and anti-parallel source/drain configuration was demonstrated. However, the relative ratio of the currents, a characteristic similar to the tunnel magnetoresistance (TMR) ratio, is still several orders of magnitude lower [19] than the TMR in MTJs.

The SpinMOSFET can function only, if the electron spins are efficiently injected/extracted in/from the channel. As there is no known semiconductor ferromagnet at room temperature, to achieve the efficient injection/detection from a metal ferromagnets in the semiconductor channel and vice versa, a thin tunneling barrier must be placed between the electrodes and the channels [20] to mitigate the spin impedance mismatch. However, the signal attributed to the spin injection [21] appears to be much weaker as compared to the large effect [22] currently attributed to the spin-dependent resonant tunneling [23-25], and the development of efficient ways to electrically inject spins from a ferromagnetic metal in a semiconductor has become an area of active research

Although many fundamental challenges have been resolved and both a SpinFET and a SpinMOSFET have been successfully demonstrated, an enhancement of the on-current ratio between the parallel and anti-parallel source/drain magnetization alignment at room temperature remains one of the main challenges. In addition, both SpinFET and SpinMOSFET still rely on the charge current to transfer the spin, which may set some limitations for the applicability of such devices in main-stream microelectronics in the future. Non-volatile devices based on MTJs possess the TMR suitable for practical applications and are reviewed below.

III. Magnetoresistive Memories

Applications driven by magnetic moments and induced magnetic fields have a large share in typical information technology products. An efficient coupling between the electrical and the magnetic degree of freedom is achieved on a quantum mechanical level and is known as the giant magneto-resistance (GMR) effect. Based on this principle hard drive storage devices with extremely high density appeared on the market. The enormous impact of this discovery on the development of information technology was recognized by awarding the inventors the Nobel Prize in 2007 [26]. A next generation of storage devices with higher density is based on the unique properties of the MTJ that the tunneling current through the structure strongly depends on the relative polarization of the ferromagnetic contacts. The difference in the MTJ resistivity can reach several hundred percent at room temperature [27]. Memory cells are arranged in a matrix connected with bit and world lines. A cell in this cross-point architecture is written by simultaneously selecting the cell with current pulses applied to the corresponding world and bit lines. The problem of half-selected cells [28] is solved by the application of a certain pulse sequence to the lines supplemented with a special design of the free layer arranged as a synthetic anti-ferromagnet [29]. This results in a deterministic, toggle-like fast switching of the free layer.

In order to be used in computer memories, MTJs must be complemented with the ability to efficiently convert charge information into magnetic moment orientation as writing the state by the magnetic field is not scalable [28]. The STT effect [30] has been proven to be a perfect alternative to the magnetic field for magnetization switching. The STT is used for purely electrical data writing by passing the current through the MTJ. The memory technology based on MTJs and the STT effect has resulted in the development of STT MRAM characterized by lower power-consumption and better scalability than conventional MRAM, where the switching is performed by the magnetic field [31].

A. STT MRAM

The magnetic field employed for switching prevents MRAM from scaling down beyond 90nm [32] as the current needed for the field generation rapidly increases with scaling. STT [30] opened a new way of manipulating magnetization dynamics by using spin-polarized currents instead of magnetic fields. The spin-polarized current allows writing the information into the memory cell purely electrically by spin transfer torques. The interest in STT MRAM has increased significantly after the observation of spin torque induced switching in AlO_x-based [33] and MgO-based [34] STT MRAM cells. Depending on the orientation of the layer magnetizations the magnetic pillars can be divided into two categories: (1) perpendicular with out-of-plane magnetization direction and (2) in-plane with the magnetization in the plane of the magnetic layer. The introduction of STT MRAM with in-plane magnetization orientation to the market has already begun with 64Mb chips by Everspin Technologies [35].

Switching the magnetization can occur spontaneously due to thermal fluctuations. This is an undesired event, which leads to the loss of the stored information. For gigabit applications the thermal stability barrier should be at least 80kT to guarantee the required retention time of 10 years. Achieving large thermal stability and a low switching current simultaneously represents one of the main challenges to engineer a good MRAM cell. Perpendicular MTJs (p-MTJs) with the thermal barrier equal to the switching barrier are preferred for applications, since they operate with a lower switching current. In addition, p-MTJs are better suited for high-density memory [36].

Both field-induced and STT switching can be complemented with heat assisted switching [28]. In addition to assisting switching, heat can facilitate new unique functionalities, for instance, using an MRAM cell with a soft reference layer as a magnetic logic unit [37]. This extends the MRAM research and development area towards logic-in-memory architectures and non-volatile computing.

Regardless of the undisputable success of the first MRAM products on the market, several important challenges remain. The general requirements for any memory type including MRAM are: (i) ability to write the data with low energy without damaging the device; (ii) data retention within a given long time interval; (iii) ability to read the recorded data without destroying the data.

Improving one or two aspects usually leads to a degradation of the remaining functionality [28]. Therefore, a careful parameter optimization specific to a particular technology is the main subject which must be addressed in order to facilitate production of high density memory arrays suitable for replacing SRAM caches and DRAM-embedded main computer memory. A careful and innovative design yielded recently a successful implementation of 8Mb 1T-1MTJ STT MRAM embedded in a 28nm CMOS logic platform [38].

As already noted, for about 10 years data retention the thermal stability barrier must be at least 80kT for gigabit MRAM arrays. However, increasing the barrier also results in an increase of the switching current density, which is proportional to the thermal barrier for p-MTJs. In order to reduce the switching current density and preserve the large thermal barrier at the same time one has to reduce the Gilbert damping and increase the spin current polarization. An interface-induced p-MTJ structure with a composite free layer CoFeB/Ta/CoFeB with two MgO interfaces [39] allows simultaneously boosting the thermal barriers and reducing damping.

For in-plane MTJs the faster switching can also be achieved, when the composite free layer is made of two half-ellipses separated by a narrow gap. The peculiarities of the magnetization dynamics of the two parts of the composite free layer [40,41], which occur in opposite senses to each other, lead to the magnetization switching in-plane. This way the large demagnetization penalty of the magnetization getting out of plane is avoided, and the switching barrier becomes equal to the thermal barrier. Because the thermal barrier depends on the free layer volume, the required large thermal stability factors of ~80kT are easily achieved in this structure.

A large TMR ratio is needed for reliably reading the information in MRAM. Indeed, the middle reference resistance to which the low and high resistance MTJ states are compared must be well separated from either of them. However, since a bit-to-bit resistance variation within a memory array is increasingly difficult to control with device sizes scaling down, the dispersion increases and so must the TMR. Obtaining a large TMR is more difficult in interface-induced p-MTJs, because the layer width must be reduced in order to boost the magnetic anisotropy; however, a TMR ratio as large as 350% has been demonstrated [42].

With growing data services such as Big Data analysis the need for additional memory capacity and speed as well as in-memory computing has increased dramatically. The last level cache memory must be increased [43-45] to bridge the memory-bandwidth gap between central processing units (CPUs) and the main memory. The CPU performance can be significantly boosted by using fast non-volatile memories in cache for data storing without the need to address the main memory. STT MRAM is characterized by high-speed access with less than 10ns and

is thus suitable for the last level three cache, where it guarantees an approximately ten times reduction in power consumption [46-48], while other types of non-volatile memories cannot provide the required high-speed access.

B. Advanced MRAM

Although STT MRAM is competitive with DRAM and can also be used in level two and level three caches in CPUs, increasing write currents for faster switching prevents it from being used in level one caches, where very fast switching is required. An ultimate swap to p-MTJs and Gilbert damping reduction are two common paths to reduce the switching current; however, these efforts are counteracted by the necessity to maintain high thermal stability which requires high perpendicular magnetic anisotropy [28].

There are indications that by downscaling the p-MTJ diameter the switching current decreases faster than the thermal stability factor, which has been shown to be as high as 120 in p-MTJs with a diameter of 30nm [49]. Nevertheless, it is preferred to have an alternative way to switch the free layer.

Interface-induced perpendicular magnetic anisotropy materials provide a sufficiently large thermal stability factor for free layers with diameters down to 12-14nm. Since the anisotropy is determined by the interface properties, it can be altered by applying an electric field. The electric field polarizes the charge densities of the interfacial atoms, thereby modifying overlap integrals and exchange interactions. This may soften the perpendicular magnetic anisotropy thus reducing the switching energy barrier and even changing it to in-plane. The magnetization can easily be pushed over the barrier and stabilized in the state with an opposite magnetization after the voltage is removed.

An MRAM controlled by voltage [50-53] is a viable option for last level cache applications. The voltage-controlled MRAM switching principle is based on voltage-mediated removal of the potential barrier separating the two stable magnetization orientation states. Without the barrier the magnetization precesses around the effective magnetic field and can be put into the alternative magnetization state, when the potential barrier is re-introduced at the end of the voltage pulse [54].

Because the voltage-induced switching is unipolar, the voltage controlled MRAM is free from the read disturb which is characteristic to STT MRAM. Although voltage controlled MRAM is a two-terminal device, the separation between read and write is performed by alternating the polarities during these two operations.

Voltage controlled MRAM has a few unsolved issues so far preventing it from being broadly used in applications. One of the problems originates in the precession at switching and thus depends on the initial state determined by the fluctuating thermal and unwanted variability. This variability results in write errors and must be suppressed.

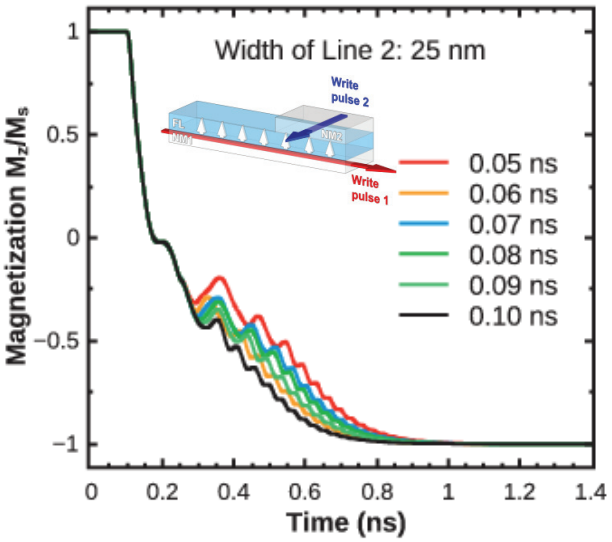

Fig. 2. Magnetization switching as a function of the second pulse duration when the two consecutive orthogonal pulses generating are applied. The first pulse duration 0.1ns is fixed.

The second problem is a larger resistance of the memory cell compared to STT MRAM, which results in smaller currents. Small currents lead to a longer delay while reading the state by a sense amplifier. As it was shown recently [53], both problems can be solved by carefully tailoring and optimizing the entire circuit.

Extending the ideas of voltage-controlled magnetic anisotropy, the voltage pulse can be applied not only to reduce the potential barrier between the two magnetization states but also to boost it to make the switching harder [55]. The switching is mediated by a spin-orbit and/or spin Hall torque generated by the current flowing through a conductive line made of a heavy metal underneath the magnetic MTJ, ensuring the write operation without an external magnetic field [56]. If reading is performed by applying the voltage pulse with its polarity opposite to that used for writing, the potential barrier is increased, hardening the cell immunity against read disturb errors.

Among the newly discovered physical phenomena suitable for next-generation MRAM are the spin Hall effect and the spin-orbit torque (SOT) switching [57-61]. Current passing in a material with a high spin Hall angle/SOI results in spin-orbit torques capable to switch the free layer of an MTJ. This way the read and write currents are decoupled, which prevents the tunnel barrier from damage and improves device reliability.

The spin Hall effect and/or SOT alone do not provide switching in devices with perpendicular magnetization. To guarantee switching, it is required to apply an external magnetic field. Switching can be reliably achieved by complementing the SOT with the voltage pulse, which reduces the magnetic anisotropy [56]. Alternatively, to accomplish reliable switching without external field, we suggest to employ two orthogonal SOT generating pulses. The second pulse must run through the line which

underlaps with the structure. The dependence of the switching on the second pulse duration is shown in Fig.2. Innovative materials are required to increase the torques and to boost the switching efficiency. New materials with a strong SOI, e.g., topological insulators, allow the current to flow only at their interface states [62]. Due to the spin-momentum locking characteristic to these states the passing current results in a large spin accumulation at the interface [63,64] and the SOT aids the magnetization to switch.

A potential disadvantage of the write and read current paths' separation is that these devices appear in a three-terminal cell configuration [65]. Therefore, they can be used only for applications in which the density is not the top priority and are thus suitable for high-speed reliable operation needed for SRAM. The need to decrease the critical current in spin orbit MRAM accelerates the search for new materials with large SOI.

IV. Non-Volatile Logic

MRAM is attractive for use with CMOS-based logic applications. Fast non-volatile memory combined with non-volatile processing elements is a fertile ground for realizing the first microprocessors with reduced power consumption working on an entirely new principle. MRAM arrays can be embedded directly on top of CMOS logic [66] to reduce the length of interconnects and the corresponding delay time. This architecture is traditionally called logic-in-memory, although as of yet the information is not processed in the MRAM. Power-efficient MRAM-based logic-in-memory circuits have already been demonstrated [67]. They include field-programmable gate arrays and ternary content addressable memory as well as other variants. These CMOS/spintronic hybrid solutions are already competitive in comparison to the conventional CMOS technology with respect to power consumption and speed.

The power consumption problem in modern integrated circuits with ultra-scaled CMOS devices is becoming critical, which prompts various power reduction technologies to be used for keeping the heat dissipation under control. The techniques based on reduced voltage operation, clock gating, and power gating modes allow to address the problem to a certain extent, however, they also result in an increase of the time delay to get into or out from these modes. The use of non-volatile MRAM-based devices [68-70] with fast access to the stored data allows cutting out the penalty of stand-by power and eliminates the delays when using energy saving modes. The first microcontroller unit with zero standby power featuring non-volatile elements was shown to operate at 8MHz [71]. In order to boost the operating frequency, spin-based non-volatile flip-flops were recently used to demonstrate a power-gating microcontroller unit [67] fabricated with 90nm CMOS technology with an additional MTJ process. The chip features a very short delay in entering/exiting power-on/power-off with the potential to be further reduced by optimizing parasitic capacitances.

Another new circuit example is a field programmable gate array built with non-volatile devices. Here, temporal data is quickly saved in magnetic tunnel junctions before the power is turned off. This has a great potential to reduce the power consumption, which becomes a critical issue in conventional SRAM-based gate arrays [72-74]. By using a logic-in-memory structure [75,76], replacing SRAM cells with non-volatile flip-flops [77] and smartly connected redundant MTJs to avoid resistance variations [78], the area of a six-input look-up table is shown to be reduced by about 50% [67].

Ternary content-addressable memory (TCAM) is able to perform a very high-speed search to match an input [79]. CMOS-based TCAM suffers from standby power losses and relatively high due to its complex structure [79]. Employing a 2T-2MTJ structure for the equality search logic part reduces the TCAM cell area [80,81]. A 1Mb non-volatile TCAM chip with a 6T-2MTJ cell structure fabricated in 90nm CMOS and perpendicular MTJ technologies has been demonstrated [82], with 9T-2MTJ [83], 7T-2MTJ [84], 4T-2MTJ [85], and 5T-4MTJ [86] modifications for high-speed accessibility and reduced variation effects being also reported. Currently, the TCAM cell structure design as well as the word segmentation algorithm optimization is under intense investigation [67] in order to increase the speed and reduce the area.

A motion-vector prediction circuit is critical for performing mobile video compression by finding motion vectors between two adjacent frames. It has been demonstrated that the introduction of non-volatile elements to implement a full adder helps making the circuit compact, fast, and stable [87-91]. The introduction of non-volatility and a logic-in-memory architecture helps reducing power consumption by 45% [92]. With the activation ratio of embedded clusters decreased, a reduction of 97% is possible [67]. Another example of an application specific circuit currently under thorough investigation is a brain-inspired computing network with non-volatile elements, which also demonstrates a large, i.e., more than 90%, power reduction on average when compared to its CMOS based counterparts [67].

The idea of combining MTJs with a common free layer enables the realization of an efficient nano-oscillator [93] and a non-volatile magnetic flip-flop [94]. Complementing flip-flops with a spin torque majority gate enables the realization of a 1-bit full adder [95]. It is important that computation is done in magnetic domain. Placing the actual computation into the magnetic domain reduces the need of converting magnetically stored information into currents and voltages for processing and helps not only to simplify the circuit layout but also increases the integration density. The idea is to use MTJs as elementary blocks for non-conventional logic-in-memory architectures. Our invention shows that on an MRAM array any two of the coupled 1T-1MTJ cells can serve simultaneously as non-volatile memory and computing units by performing a logical implication operation [96]. These structures inherently

realize non-volatile logic-in-memory circuits with zero-standby power, where the same elements are used for storing and also processing information. They have a great potential for Big Data storing and computing, as they are also opening a path for developing computing architectures conceptually different from the still standard Von Neumann architecture. A new design of an implication-based full adder involves only six 1T-1MTJ cells with 27 subsequent FALSE and material implication operations [97].

V. Conclusion

Spin transistors have been recently successfully demonstrated, however, an enhancement of the on-current ratio between the parallel and anti-parallel source/drain magnetization configuration at room temperature remains one of the main challenges. As both SpinFET and SpinMOSFET still rely on the charge current to transfer the spin, it sets limitations for the applicability of such devices in main-stream microelectronics, and new ideas are needed for the future.

Non-volatile devices based on MTJs possess a TMR suitable for practical applications. Although STT MRAM is currently hitting the market, the switching current reduction and increasing speed and the thermal stability may prevent their use in first-level caches. Thus, novel innovative non-volatile devices with improved switching characteristics and low power consumption are required on a longer run.

Finally, the successful adoption of non-volatility by the microelectronics industry by developing various logic-in-memory architectures will inevitably result in increasing disseminations of this technology for other applications such as ultralow-power electronics, high-performance computing, the Internet of Things, and Big Data analysis.

Acknowledgement

Fruitful discussions with Dr. T. Windbacher and the financial support by the Austrian Federal Ministry of Science, Research and Economy and the National Foundation for Research, Technology and Development are gratefully acknowledged.

References

[1] S.-E. Thompson, M. Armstrong, C. Auth *et al., IEEE T. Electron Dev.*, vol.**51**, pp. 1790-1797, 2004.
[2] K. Mistry, C. Allen, C. Auth *et al., IEDM Techn. Digest*, pp. 247-250, 2007.
[3] S. Natarajan, M. Armstrong, M. Bost *et al., IEDM Techn. Digest*, pp. 941-943, 2008.
[4] R. Xie, P.Montanini, K.Akarvardar *et al., IEDM Techn. Digest*, pp. 47-50, 2016.
[5] S.-Y. Wu, C.Y. Lin, M.C. Chiang *et al., IEDM Techn. Digest*, pp. 43-46, 2016.
[6] International Technology Roadmap for Semiconductors (2016).
[7] H. Ohno, M. Stiles, B. Dieny, *Proc. of the IEEE*, vol.**104**, pp. 1782-1786, 2016.
[8] S. Datta, B. Das, *Appl. Phys. Lett.*, vol. **56** (7), pp. 665–667, 1990.
[9] S. Sugahara, J. Nitta, *Proc. of the IEEE*, vol. **98**, pp. 2124-2154, 2010.
[10] I. Zutic, J. Fabian, S. Das Sarma, *Rev. Mod. Phys*., vol. **76**, pp. 323–410, 2004; J. Fabian, A. Matos-Abiaguea, C. Ertler *et al., Acta Phys. Slovaca*, vol. **5**, pp. 565–907, 2007.
[11] P. Li, H. Dery, *Phys. Rev. Lett*. vol. **107**, 107203, 2011.
[12] O. Chalaev, Y. Song, H. Dery, *Phys.Rev.B*, vol. **95**, 035204, 2017.
[13] V. Sverdlov, S. Selberherr, *Phys. Rep.*, vol. **585**, pp. 1 – 40, 2015.
[14] V. Sverdlov, Strain-induced Effects in Advanced MOSFETs, Springer, 2011.
[15] V. Sverdlov, J. Ghosh, S. Selberherr, in *Proc. of WINDS*, p.7, 2016.
[16] Y. Bychkov, E. Rashba, *JETP Lett*. vol. **39**, pp. 78–81, 1984.
[17] P. Chuang, S.-C. Ho, L. W. Smith *et al.*, *Nature Nanotechnol*., vol. **10**, pp. 35-39, 2015.
[18] W. Yan, O. Txoperena, R. Llopis *et al.*, *Nature Communications*, vol. **7**, 13372, 2016.
[19] T. Tahara, H. Koike, M. Kameno *et al*., *Appl. Phys. Express*, vol. **8,** 11304, 2015.
[20] E. I. Rashba, *Phys. Rev. B*, vol. **62**, R16267–R16270, 2000.
[21] T. Tahara, Y. Ando, M. Kameno *et al.*, *Phys. Rev. B*, vol. **93**, 214406, 2016.
[22] R. Jansen, *Nat. Mater.*, vol. **11**, pp. 400–408, 2012.
[23] Y. Song and H. Dery, *Phys. Rev. Let.,* vol. **113**, 047205, 2014.
[24] Z. Yue, M. C. Prestgard, A. Tiwari, M. E. Raikh, *Phys. Rev. B,* vol. **91**, 195316 2015.
[25] V. Sverdlov, J. Weinbub, S. Selberherr, in *Proc. of IWCN,* pp. 88 – 90, 2017
[26] A. Fert, Nobel Lecture: Origin, Development, and Future of Spintronics, *Rev. Modern Phys*., vol. **80**, pp. 1517–1530, 2008.
[27] S. Ikeda, J. Hayakawa, Y. Ashizawa *et al*., *Appl. Phys. Lett.*, vol. **93**, 082508, 2008.
[28] D. Apalkov, B. Dieny, and J.M. Slaughter, *Proc. of the IEEE*, vol. **104**, pp. 1796-1830, 2016.
[29] L. Savtchenko, B. Engel, N. Rizzo, M. Deherrera, J. Janesky, U.S. Patent 6545906 B1, 2003.
[30] J. Slonczewski, *J. Magn. Mater*., vol. **159**, pp. L1-L7, 1996;
[31] Electronicdesign.com, 4-Mbit Device is First Commercially Available MRAM, Electronic Design, 2006.
[32] R. Sbiaa, H. Meng, and S.N. Piramanayagam, *Phys. Stat. Solidi (RRL),* vol. **5**, pp. 413–419, 2011.
[33] Y. Huai, F. Albert, P. Nguyen *et al., Appl. Phys. Lett.,* vol. **84**, pp. 3118–3120, 2004.
[34] Z. Diao, D. Apalkov, M. Pakala *et al*., *Appl. Phys. Lett.,* vol. **87**, 232502, 2005.
[35] N.D. Rizzo, D. Houssameddine, J. Janesky *et al., IEEE Trans. Magn*., vol. **49**, pp. 4441–4446, 2013.
[36] S.-W. Chung, T. Kishi, J.W. Park *et al., IEDM Techn. Digest*, pp. 659-662, 2016.
[37] B. Dieny, R. Sousa, S. Bandiera *et al., IEDM Techn.Digest*, pp. 1.3.1–1.3.4, 2011.
[38] Y.J. Song, J.H. Lee, H.C. Shin *et al., IEDM Techn. Digest*, pp. 663-666, 2016.
[39] H. Sato, M. Yamanouchi, S. Ikeda *et al.*, *IEEE Trans. Magn*., vol. **49**, pp. 4437–4440, 2013.
[40] A. Makarov, V. Sverdlov, D. Osintsev, S. Selberherr, *Phys.Stat.Solidi (RRL)*, vol. **5**, pp. 420–422, 2011.

[41] A. Makarov, T. Windbacher, V. Sverdlov, S. Selberherr, S*emicond. Sci. and Tech.*, vol. **31**, 113006, 2016.
[42] M. Krounbi, V. Nikitin, D. Apalkov *et al*., in *Proc. of ECS Meeting*, 2015.
[43] M.T. Chang, P. Rosenfeld, S.L. Lu B. Jacob, in *Proc. of HPCA*, pp. 143–154, 2013.
[44] T. Endoh, T. Ohsawa, H. Koike *et al.,* in *Symp. VLSI Technol.*, pp. 89–90, 2012.
[45] T. Endoh, in *Proc. of IEEE ISCAS*., pp. 13–16, 2015.
[46] H. Noguchi, K. Ikegami, N. Shimomura *et al., in Symp. VLSI Circuits*, pp. 97-98, 2014.
[47] G. Jan, L. Thomas, S. Le *et al*., *in Symp. VLSI Technol.*., pp. 18-19, 2016.
[48] H. Noguchi, K. Ikegami, S. Takaya *et al*., *in ISSCC Dig. Tech. Papers*, pp. 132-133, 2016.
[49] L. Thomas, G. Jan, J. Zhu *et al., J. Appl. Phys.,* vol. **115**, 172615, 2014.
[50] T. Nozaki, Y. Shiota, M. Shiraishi *et al*., *Appl. Phys. Lett*., vol. **96**, 022506, 2010.
[51] Y. Shiota, T. Nozaki, F. Bonell *et al., Nat. Mater.*, vol. **11**, pp. 39-43, 2012.
[52] Y. Shiota, T. Nozaki, S. Tamaru *et al., Appl. Phys. Express*, vol. **9**, 013001, 2015.
[53] K.L. Wang, X. Kou, P. Upadhyaya *et al.*, *Proc. of the IEEE*, vol. **104**, pp. 1974-2008, 2016.
[54] M.K. Niranjan, CG. Duan, S.S. Jaswal, E.Tsymbal, *Appl. Phys. Lett*., vol. **96**, 222504, 2010.
[55] H. Noguchi, K. Ikegami, K. Abe *et al.,* in *IEDM Techn. Digest*, pp.675-678, 2016.
[56] H. Yoda, N. Shimomura, Y. Ohsawa *et al., in IEDM Techn. Digest*, pp.679-682, 2016.
[57] I.M. Miron, G. Gaudin, S. Auffret *et al., Nat. Mater.*, vol. **9**, pp. 230–234, 2010.
[58] I.M. Miron, K. Garello, G. Gaudin *et al., Nature*, vol. **476**, pp. 189–193, 2011.
[59] L. Liu, O.J. Lee, T.J. Gudmundsen *et al*., *Phys. Rev. Lett*., vol. **109**, 096602, 2012.
[60] A. Brataas and K. M.D. Hals, *Nat. Nanotechnol*., vol. **9**, pp. 86–88, 2014.
[61] M. Cubukcu, O. Boulle, M. Drouard *et al., Appl. Phys. Lett*., vol. **104**, 042406, 2014.
[62] Y. Fan, P. Upadhyaya, X. Kou *et al., Nat. Mater.*, vol. **13**, pp. 699–704, 2014.
[63] C.H. Li, O.M.J. Van't Erve, J.T. Robinson *et al., Nat. Nanotechnol*., vol. **9**, pp. 218–224, 2014.
[64] C.H. Li, O.M.J. Van't Erve, S. Rajput *et al., Nature Communications*, vol. **7**, 13518, 2016.
[65] S.-W. Lee, K.-J. Lee, *Proc. of the IEEE*, vol. **104**, pp. 1831-1843, 2016.
[66] J.-G. Zhu, *Proc. of the IEEE*, vol. **96**, pp. 1786–1798, 2008.
[67] T. Hany, T. Endoh, D. Suzuki *et al., Proc. of the IEEE*, vol. **104**, pp. 1844-1863, 2016.
[68] T. Endoh, S. Togashi, F. Iga *et al*., *IEDM Techn. Digest*, pp.75–78, 2011.
[69] N. Sakimura, Y. Tsuji, R. Nebashi *et al., in Proc. of IEEE ISSCC,* pp. 184–185, 2014.
[70] H. Koike, T. Ohsawa, S. Ikeda *et al., in Proc. of IEEE A-SSCC.*, pp. 317–320, 2013.
[71] S.C. Bartling, S. Khanna, M.P. Clinto *et al.*, *in Proc. of IEEE ISSCC,* pp. 432–433, 2013.
[72] Y. Guillemenet, L. Torres, G. Sassatelli *et al.*, *in Proc. of FPL*, pp. 421–426, 2008.
[73] Y.Y. Liauw, Z. Zhnag, W. Kim *et al., in Proc. of IEEE ISSCC,* pp. 406–408, 2012.
[74] Z. Zhang, Y.Y. Liauw, C. Chen, S.S. Wong, *Proc. of the IEEE*, vol. **103**, pp. 1197–1210, 2015.
[75] D. Suzuki, M. Natsui, S. Ikeda *et al., in Symp. VLSI Circuits*, pp. 80–81, 2009.
[76] D. Suzuki, M. Natsui, T. Endoh *et al., J. Appl. Phys*., vol. **111**, 07E318, 2012.
[77] S. Yamamoto, Y. Shuto, and S. Sugahara, *Jpn. J. Appl. Phys.*, vol. **51**, pp. 11PB021–11PB025, 2012.
[78] D. Suzuki, M. Natsui, A. Mochizuki *et al., in Symp. VLSI Circuits*, pp. 172–173, 2015.
[79] K. Pagiamtzis and A. Sheikholeslami, *IEEE J. Solid-St. Circ.*, vol. **41**, pp. 712–727, 2006.
[80] S. Matsunaga, K. Hiyama, A. Matsumoto *et al., Appl. Phys. Express*, vol. **2**, pp. 0230041–0230043, 2009.
[81] S. Matsunaga, M. Natsui, S. Ikeda *et al., in Proc. of ASP-DAC,* pp. 475–476, 2012.
[82] S. Matsunaga, A. Katsumata, M. Natsui *et al., in Symp. VLSI Circuits*, pp. 106–107, 2013.
[83] S. Matsunaga, A. Katsumata, M. Natsui *et al., Jpn. J. Appl. Phys.*, vol. **51**, pp. 02BM061–02BM066, 2012.
[84] S. Matsunaga, A. Katsumata, M. Natsui *et al., J. Appl. Phys*., vol. **111**, pp. 07E3361–63, 2012.
[85] S. Matsunaga, S. Miura, H. Honjou *et al.,* in *Symp. VLSI Circuits*, pp. 44–45, 2012.
[86] S. Matsunaga, A. Mochizuki, N. Sakimura *et al., IEICE Electron. Express*, vol. **11**, pp. 02971–02977, 2014.
[87] S. Matsunaga, J. Hayakawa, S. Ikeda *et al., Appl. Phys. Express*, vol. **1**, pp. 0913011–0913013, 2008.
[88] H.-P. Trinh, W. Zhao, J.-O. Klein *et al., IEEE Trans. Circuits Syst.*, vol. **60**, pp. 1469–1477, 2013.
[89] E. Deng, Y. Zhang, W. Kang *et al., IEEE Trans. Circuits Syst.*, vol. **62**, pp. 1757–1765, 2015.
[90] M. Natsui, D. Suzuki, N. Sakimura *et al., in Proc. of IEEE ISSCC.*, pp. 194–195, 2013.
[91] M. Natsui, D. Suzuki, N. Sakimura *et al., IEEE J. Solid-St. Circ.*, vol. **50**, pp. 476–489, 2015
[92] Y. Ma, T. Shibata, T.Endoh, i*n Proc. of IEEE ISCAS,* pp. 1248–1251, 2013.
[93] A. Makarov, V. Sverdlov, S. Selberherr, in *Proc. of the ICNM*, p.69, 2013.
[94] T. Windbacher, H. Mahmoudi, V. Sverdlov, S. Selberherr, *Patent*: International, No. Wo 2014/154497 A1; EP 13161375.4.
[95] T. Windbacher, A. Makarov, V. Sverdlov, S. Selberherr, in "Future Trends in Microelectronics - Journey into the Unknown, S.Luryi, J.Xu, A.Zaslavsky (ed); J.Wiley&Sons, 2016.
[96] H. Mahmoudi, T. Windbacher, V. Sverdlov, S. Selberherr, *Patent:* International, No. Wo 2014/079747 A1; EP 12193826.0.
[97] H. Mahmoudi, V. Sverdlov, S. Selberherr, in *Proc. of ESSDERC*, pp. 254 – 257, 2012.