Non-volatile memories based on graphene and related two-dimensional materials

Simone Bertolazzi, Paolo Bondavalli, Stephan Roche, Tamer San, Sung-Yool Choi, Luigi Colombo, *Francesco Bonaccorso, *Paolo Samori*

Dr. S. Bertolazzi, Prof. P. Samori
Université de Strasbourg, CNRS, ISIS UMR 7006, 8 allée Gaspard Monge, 67000 Strasbourg, France
E-mail: samori@unistra.fr

Dr. P. Bondavalli
Chemical and Multifunctional Materials lab, Thales Research and Technology, 91767 Palaiseau, France

Prof. S. Roche
Catalan Institute of Nanoscience and Nanotechnology (ICN²), CSIC and The Barcelona Institute of Science and Technology, Campus UAB, Bellaterra, 08193 Barcelona, Spain
&
ICREA - Institució Catalana de Recerca i Estudis Avançats, 08070 Barcelona, Spain

Dr. T. San
Texas Instruments, Dallas, TX 75243, USA

Prof. S. –Y. Choi
School of Electrical Engineering, Graphene/2D Materials Research Center, KAIST, 34141, Daejeon, Korea

Prof. L. Colombo
Department of Materials Science and Engineering, University of Texas at Dallas, Richardson, TX 75080, USA
E-mail: colombo@utdallas.edu

Dr. F. Bonaccorso
Istituto Italiano di Tecnologia, Graphene Labs, Via Morego 30, I-16163 Genova, Italy
&
BeDimensional Srl, Via Albisola 121, 16163 Genova, Italy
E-mail: Francesco.Bonaccorso@iit.it

Keywords: non-volatile memories, two-dimensional materials, graphene, transition metal dichalcogenides, black phosphorous
Abstract

The pervasiveness of information technologies in all aspects of our daily lives has brought us to an impressive generation of data, which need to be stored and accessed very quickly. Non-volatile memories (NVMs), with their ever-growing capacity and speed, are making inroads into high-capacity storage to replace hard disk drives, fuelling the rapid expansion of the global storage-class memory market. As silicon-based flash memories are approaching their fundamental limit, vertical stacking of multiple memory-cell layers (*i.e.* 3D integration), as well as innovative device concepts and novel materials are being intensively investigated for the development of new NVMs. In this context, emerging two-dimensional (2D) materials, such as graphene, transition metal dichalcogenides (TMDs) and black phosphorous (BP), offer a host of outstanding physical and chemical properties, which could both improve existing memory technologies and enable the next-generation of low-cost, flexible and wearable information-storage devices. In this review article, we provide an overview on the exploitation of graphene and related 2D materials (GRMs) in different types of NVM cells, including resistive random access, flash, magnetic and phase-change memories. We discuss in depth the physical and chemical mechanisms underlying the non-volatile switching of GRM-based memory devices developed at the laboratory scale in the last decade. Although at this stage most of the proof-of-concept devices developed in academia do not compete with state-of-the-art market products, a number of promising technological advancements have emerged, particularly within the area of low-cost and flexible electronics, which could become the focus of considerable development efforts in the forthcoming years. Here, the most relevant material properties and device structures are analysed, emphasizing both opportunities and challenges towards the realization of practical NVM devices that exploit the unique properties of GRMs.

1. Introduction

The digital universe, quantified by the number of bits generated annually by human kind, is undergoing a relentless expansion and is expected to reach 44 zettabytes (*i.e.*, 44 trillion gigabytes)
by 2020. To deal with such a large amount of data, the next generation of non-volatile memory (NVM) technologies must offer ever-growing performance in terms of speed, capacity, endurance and retention, as well as energy cost. Nowadays, novel device architectures, alternatives to the widespread silicon flash memory, are being intensively investigated for future NVMs with the objective of scaling the feature size, increasing the write/erase speed and reducing the data access time. The 2017 technology and market report by Yole Développement (ref. 6) highlights the growing importance of emerging/prototypical NVM technologies (see Figure 1a), such as resistive random access memories (ReRAMs), phase change memories (PCMs) and magnetic random access memories (MRAMs). These technologies are foreseen to play a leading role in the rising market of storage-class memory (SCM), which represents an intermediate memory hierarchy between cost-effective data storage, e.g. flash NOT-AND (NAND) and disks, and high-performance working memories, e.g. static/dynamic random access memories (S/DRAMs), providing simultaneously permanent storage and fast processing of large volumes of data. Today, several companies are developing and/or introducing new NVM products based on emerging technologies, such as for instance carbon-nanotube RAMs (NRAM) and spin-transfer torque (STT) MRAMs. These NVM devices can provide operation speeds comparable to DRAMs and are expected to replace SRAMs in future mobile and high-performance computing; this would represent a major step forward towards the so-called universal memory, that is a single information-storage technology that combines together the best properties of data storage and working memories, eliminating the need for multiple memory hierarchies within the same computing system. The development of such universal memory can lead to cost and complexity reduction and is expected to improve the overall system speed by suppressing the time required for transferring data among different memory levels. Universal memories would allow sustaining the ever-growing demand for cheaper and smaller NVMs with higher density storage, greater endurance and higher speed, with respect to current technologies.

Since the first prototypes of solid-state memories were developed about five decades ago, impressive progress has been made thanks to continuous advancements in fabrication processes, as
well as through the introduction of novel device concepts stimulated by the discovery of new materials and phenomena at the nanoscale. Today, a number of different (nano)materials are being explored for improving the figures of merit (FoM) of NVM devices, including graphene\textsuperscript{[27-28]} and related two-dimensional materials (GRMs),\textsuperscript{[29-33]} which have received tremendous attention over the past decade. In addition to their atomic-scale thickness, these materials have unique chemical and physical properties, including flexibility and transparency, which are highly desirable for the development of information-storage devices to be integrated in wearable systems and smart objects.\textsuperscript{[34]} As consumer electronics moves towards pervasive connectivity (e.g., Internet of Things, IoT), as well as mobile and data-centric applications,\textsuperscript{[5]} the market size of low-cost, lightweight, portable/wearable NVM devices is expected to grow steadily in the next years. In this context, the large family of GRMs can offer a wealth of opportunities. Materials of interest includes 2D semiconducting sheets of transition metal dichalcogenides (TMDs)\textsuperscript{[35-37]}, e.g., MoS\textsubscript{2}, WS\textsubscript{2}, MoSe\textsubscript{2} and WSe\textsubscript{2}, and black phosphorous (BP),\textsuperscript{[38-40]} topological insulators (e.g. silicene and other buckled 2D Xenes),\textsuperscript{[41-42]} insulators such as hexagonal boron nitride (h-BN),\textsuperscript{[43-44]} as well as highly conducting layers such as semimetallic graphene\textsuperscript{[27-28, 31]} and superconducting TMDs (e.g. NbSe\textsubscript{2}).\textsuperscript{[45-46]} Moreover, the possibility to assemble artificial van der Waals heterostructures composed of multiple GRMs has paved the way towards novel nanomaterials with optical and electronic properties \textit{ad-hoc} for various technological applications.\textsuperscript{[47-49]} Such a broad spectrum of materials/properties makes GRMs appealing for use in a large number of technologies, including different types of NVM devices (e.g. Figure 1b-e). Since 2008, being just four years after the isolation of graphene, numerous academic research groups have been exploring the use of GRMs in NVM technologies, starting from graphene-based ReRAMs,\textsuperscript{[50-52]} ferroelectric random access memories (FeRAMs)\textsuperscript{[53]} and flash memories.\textsuperscript{[54]} A few reports appeared also on the use of graphene in magnetic tunnel junctions (MTJs) for MRAMs\textsuperscript{[55]} and in PCMs.\textsuperscript{[56]} More recently, new types of NVM cells enabled by the unique properties of GRMs were demonstrated, such as two-terminal tunnelling memories\textsuperscript{[57]} and programmable p-n junctions,\textsuperscript{[58]} both based on artificially stacked van der Waals heterostructures of 2D crystals.
The large variety of GRMs has sparked the creativity of scientists and engineers to improve the performance of NVMs and to develop novel device structures based on 2D materials (e.g., refs [57-59]). In most cases, GRMs have been introduced in existing NVM technologies in order to improve the figures of merit (FoM) of scaled devices. For example, few-layer graphene has been investigated as a potential floating gate (FG) material in future flash memories\(^\text{[60-61]}\) with the aim of reducing leakage currents through the gate stack and minimize capacitive coupling interferences among neighbouring cells (see Section 6). The insertion of single-layer graphene in PCMs as a thermal resistance layer between the Ge-Sb-Te (GST) phase-change material and the tungsten heater electrode has resulted in memory cells with improved energy efficiency with respect to the baseline devices.\(^\text{[62]}\)

Graphene has also been widely used as ultrathin flexible/transparent electrode or as an interfacial layer in ReRAMs for lowering power consumption and for suppressing detrimental surface effects.\(^\text{[63-67]}\) Alongside graphene, other members of the GRM family, in particular TMDs, h-BN and BP, have been used in ReRAMs\(^\text{[64, 68]}\) and flash memories (e.g., refs [61, 69-71]). A few reports also suggest that GRMs might be introduced in current NVM technologies through the development of novel device concepts enabled by 2D van der Waals heterostructures.\(^\text{[57-59]}\) However, it should be noticed that the development of high-performance memory devices incorporating GRMs requires significant resources before the latter can be integrated into a “conventional Si device flow”.\(^\text{[34]}\) On the contrary, it is more likely that GRMs could find application in low-cost portable/wearable information-storage devices,\(^\text{[72-73]}\) thanks to the availability of cost-effective solution-processing techniques, such as spray coating and ink-jet printing,\(^\text{[74-75]}\) which are particularly suitable for the production of memory devices on flexible plastic substrates.\(^\text{[76-78]}\)

Here, we provide a comprehensive overview of the most significant advancements in the field of GRM-based NVMs, from the first graphene resistive memories demonstrated in 2008\(^\text{[50-52]}\) to the latest memory cells based on 2D crystals-based heterostructures.\(^\text{[57-58, 79]}\) For each device structure, we describe the physical and chemical mechanisms that allow for writing, reading and storing digital information. In addition, we present a comprehensive analysis/comparison of the FoM of the memory cells based on GRMs, reported by academic research groups in the last decade, highlighting the
opportunities and the challenges associated with the introduction of such materials in future NVM technologies.

2. 2D materials: production and processing

Development and production of electronic devices inherently depends on the properties of available materials. Although some proof-of-concept 2D material-devices have been demonstrated exploiting micromechanically cleaved samples, the development of scalable processes with “on-demand” tuning of structural and electronic properties is a “must” for the practical realization of this technology. The growth of large area high quality single crystal 2D materials (Figure 2 a-c) is perhaps one of the most challenging aspects of this research area, which is especially true for the multicomponent 2D materials. The requirement to have control at the monolayer level needs surface physics and chemistry understanding that hitherto has not yet been demonstrated on a large scale. Graphene can be “easily” grown on some metal substrates, and progress is being made towards large area single crystals, a viable process that might yield high quality thin films. Growth techniques reported in the literature such as chemical vapour deposition (CVD), see Figure 2a, atomic layer deposition (ALD), molecular beam epitaxy (MBE) etc., which have been traditionally used to grow multicomponent heterostructures of II-VI, III-V and oxide materials, are also promising for the large scale production of hexagonal boron nitride (h-BN) and TMDs. At this time, high quality large area polycrystalline and single crystal graphene has only been grown by CVD and thermal desorption of Si from SiC single crystals. Notable steps forward have been made in the growth of graphene on metals and on silicon carbide (SiC), although it is more difficult to exploit the latter process for NVM applications due to its ultimate size limitations and inability to be integrated in a Si-flow process. Progress is also reported on the growth of h-BN and TMD materials. However, the growth of large area monolayer or few-layer single crystals of h-BN and TMDs is still a major challenge that will require continued significant research efforts. Two-dimensional growth has already been clearly demonstrated for graphene on copper. Deterministic
nucleation and growth of graphene on copper (see Figure 2b) has also been demonstrated to yield hexagonal graphene single crystals.\cite{94-95} This process may require edge functionalization for graphene growth on non-catalytic surfaces. In general, 2D materials have a higher edge growth rate compared to cubic systems as clearly demonstrated in graphene\cite{83, 96} and more recently for TMD materials.\cite{97}

By carefully choosing precursors of a particular compound and growth conditions, one can tune the growth parameters, creating “materials on-demand” for the design and realization of heterostructures based on 2D materials.\cite{98} The functionality of such heterostructures, that is not simply given by the combined properties of the individual layers but how they interact, can give rise to transport properties that enable the creation of a truly “quantum leap” in the functionality of electronic devices.\cite{48}

Two-dimensional materials can be integrated in device flows in several ways, including via 1) the in-situ thin film (selective) growth on pre-patterned substrates, and 2) the direct transfer of individual layers or stacks from any substrate to the desired support for device fabrication. In-situ and selective growth of any of the 2D material family will require continued basic understanding of nucleation and growth on dissimilar surfaces or epitaxial growth. These technologies are being investigated but are still in the embryonic stage. The advantage of selective growth relies on the fact that if a single crystal is needed, the area requirement for its growth will be much lower in comparison to global single crystal growth. Direct transfer of individual films has the advantage of growing the films on an optimum substrate but has the big disadvantage that other than cost, the substrate with the device needs to be planarized, thus potentially limiting the usefulness of this approach. The availability of large area high quality synthetic 2D films will enable the development of equipment for transfer and alignment of 2D materials for the fabrication of various types of devices including NVMs. The development of such new equipment (e.g. ref. \cite{99}) will open new opportunities for the integration of these material structures in current NVM device flows. Given the early stage of development for selective growth, it is foreseen that, at least for the next few years, in order to achieve high quality stacked films, the transfer processes will be the most feasible route.
Hitherto several transfer processes, classified as wet- or dry-transfer, have been developed. In the case of the wet-transfer process, the as-grown 2D materials are in contact during at least one step of the procedure with a liquid.\cite{74} This promotes the presence of trapped adsorbates onto the 2D materials surface, critically affecting the quality of the interfaces. In order to overcome such problem, dry transfer protocols, where 2D materials are protected against the contact with any liquid, have been developed to obtain cleaner interfaces.\cite{100} Cleaner surfaces will enable the achievement of the ultimate fundamental properties of 2D materials, namely extremely low interface trap density or dangling bonds.\cite{101,102} The transfer of graphene using pick-and-place techniques\cite{43,101,103} enabled the demonstration of extremely high charge-carrier mobility ($\approx 140,000$ cm$^2$V$^{-1}$s$^{-1}$ at room temperature) in graphene transistors using $h$-BN as the gate insulator.

The direct exfoliation of bulk layered crystals by liquid-phase exfoliation (LPE),\cite{104-107} see Figure 2d, is another industrially relevant strategy for the scalable production of 2D materials. The LPE process enables the formulation of inks of 2D materials in different solvents.\cite{108-110} This is the starting point for reliable production of devices based on printed technology,\cite{108-110} and thus for the development of 2D-materials-based flexible devices.\cite{75}

Liquid-phase exfoliation (LPE) is a versatile technique that can be exploited for the exfoliation of layered materials\cite{104-107} such as graphite, TMDs, BP and $h$-BN, just to cite a few. The LPE process of bulk crystals (see Figure 2d) generally involves three steps: (i) dispersion in a solvent; (ii) exfoliation; (iii) “sorting”.\cite{74,108} The LPE process starts with the dispersion of bulk crystals in an appropriate medium, which can be either organic solvent\cite{104-107} or aqueous solution, in the latter case with the aid of surfactants\cite{106,111-113} or polymers.\cite{114-115} The exfoliation process is commonly carried out by means of ultrasonication of bulk crystals. However, while this approach allows the production of low-defect flakes (\textit{i.e.}, no significant additional defects are introduced during the exfoliation) with concentrations of only several g/l,\cite{116} it is not easily scalable to large volumes.\cite{108} To overcome this issue, other approaches have been proposed such as ball milling,\cite{117-119} shear exfoliation,\cite{120-121} and micro-fluidization,\cite{122-125} each having its own advantages and disadvantages,\cite{108} especially in terms of quality, yield of exfoliation, cost, scalability and defect density.\cite{108} Feng, Müllen and co-workers
have demonstrated that electrochemical exfoliation of graphite (see Figure 2e) provides graphene
flake \(^{[126]}\) from one- to three-layers with a high yield of greater than 80\% and a high C/O ratio of
(~12), a sheet resistance value of 4.8 kΩ/□ and hole mobility of 233 cm\(^2\)/Vs for a single sheet. \(^{[126]}\)
These features are key for further development of NVM technology based on GRM. Another very
promising approach is the use of high-pressure wet-jet-milling (WJM).\(^{[127]}\) In fact, the WJM process
will facilitate the production of defect-free and high quality 2D-crystal (single- and few-layer)
dispersions on a large scale, i.e., 2 L hr\(^{-1}\) at a concentration of 10 gL\(^{-1}\).

The LPE of layered materials is also a valuable approach to produce TMDs with different phases. For
example MoS\(_2\) can be formed in both the semiconducting 2H (trigonal prismatic) \(^{[128-129]}\) and the
metallic 1T (octahedral),\(^{[130]}\) by LPE thus permitting resistive switching. However, whatever
exfoliation process is used, the key issue of LPE is that the samples are always highly polydispersed
with broad flake size and thickness distributions.\(^{[108]}\) Therefore, it is necessary to fine tune the
morphological properties i.e., the separation of large from small\(^{[113]}\) and thick from thin\(^{[112]}\) flakes.
This final step is usually carried out by using ultracentrifugation processes.\(^{[112, 131]}\) The exfoliated 2D
crystals will then have to be sorted both by lateral size and thickness by following different strategies
based on ultracentrifugation in a uniform\(^{[132]}\) (sedimentation based-separation -SBS-) or density
gradient\(^{[132]}\) medium (density gradient ultracentrifugation -DGU-).

A key issue of 2D flakes produced by LPE is the agglomeration following the deposition/coating
process and how this affects the electronic, i.e., charge carrier mobility, contact resistance, as well as
the physical properties such as the roughness of the as-deposited film. These aforementioned
problems will have to be solved for a successful integration in NVM devices. The addition of
stabilizing agents,\(^{[111-113, 131]}\) physically hindering the flakes from contacting each other, could be an
option to overcome flake agglomeration. While the stabilizing agents can minimize agglomeration,
they could also have a negative effect by diluting/lowering electrical performance of the assembled
films. Some of the layered materials, such as BP, are unstable in ambient conditions or in the presence
of water. Some of the these stability issues are also valid for other 2D materials grown by bottom-up
approach (e.g., CVD)\(^{[74, 83, 93, 133]}\) or produced by micromechanical cleavage,\(^{[27, 31, 134-135]}\) and can be
overcome by the introduction of a protective solvent shell. This solvent shell, and the residual surfactants/polymers adsorbed onto the flake surface, introduces an intrinsic doping of the flakes,[108] which can be later exploited to control the transport properties of the deposited films. The same is also valid for transferred 2D materials grown by bottom-up techniques.[74, 83, 86, 93, 133, 136]

The LPE process can be exploited not only for the exfoliation of pristine bulk layered materials but also for the exfoliation of graphite oxide[137] to produce GO, largely used in NVMs technology.[74, 107]

In particular, graphite oxide is prepared by various methodologies (e.g., the modified Hummer’s method)[138] which involves aggressive chemical processes that introduce functional groups both at the edges (e.g., carboxylic and carbonyl groups, as well as phenol, lactone and quinone) and on the basal plane (hydroxyl or epoxide groups).[139-140] The presence of these functional groups is fundamental for the GO production by thermal expansion,[141] ultrasonication,[142] stirring[143] of graphite oxide followed by liquid dispersion, which can be carried out in aqueous solutions,[142, 144] since GO flakes are strongly hydrophilic. Graphene oxide provides a unique platform for reversible and non-volatile chemical switching being a wide bandgap material (as high as 6 eV), whose electronic properties can be tuned by the amount, nature and position on the GO flakes of the functional groups.[145-146] However, although GO flakes can have lateral size up to several microns,[147] they are defective,[140] because the aforementioned chemical treatments disrupt the sp²-bonded network compromising their structural and electronic integrity.[74] In order to restore, although only partially, the electrical and thermal conductivity of pristine graphene flakes, several procedures have been devised to chemically reduce the GO flakes, by both chemical[139, 144] and physical[140-141, 148] processes. These reduction processes have recently been optimized yielding electrical properties truly approaching those of pristine graphene, with room temperature in field-effect transistor mobility values exceeding 1000 cm²V⁻¹s⁻¹ for microwave-reduced GO.[149]

The key features of the LPE processes are its scalability and versatility, which, associated with the low-cost production technique, can provide 2D materials in bulk quantities. Moreover, the possibility of having a large class of solution-processed 2D materials enables their integration with polymeric materials or deposition/coating on different substrates. In this context, progress on large-scale
placement of 2D materials-based inks by the various deposition/coating techniques such as Langmuir–Blodgett,[150] spin-, [151] spray-, [152] and rod-coating,[153] and inkjet printing,[109-110] is enabling printing of 2D materials-based films and heterostructures on a large scale.[108] Nonetheless, apart from the uniformity of large area films, the roughness of the deposited film is still an issue for both optical and electronic properties of the deposited films and worse than those obtained by micromechanical cleavage or the direct growth. However, unlike the transfer approach, drop-on-demand printing[108] could meet the high-volume-manufacturing (HVM) requirements of 2D-materials-based devices. A key advantage of this approach could be represented by the possibility to integrate/complement other production approaches, for example for the realization of contacts. This was recently demonstrated for a programmable logic memory device (i.e., graphene/WS2/graphene) realized by inkjet printing technology.[75] Drop-on-demand ink-jet printing has been demonstrated in an all-printed, vertically stacked transistor device flow with graphene-based source, drain, and gate electrodes, a TMD channel, and a h-BN dielectric[154] having a charge carrier mobility of ~0.22 cm²/Vs.[154] However, the contacts are ~ 400 nm thick, the roughness is extremely high (>50 nm) and the charge carrier mobility is rather low. [154] Moreover, it is still unclear if the h-BN acts as dielectric, because ionic liquid is still needed. [154] Considering the aforementioned issues, new insights are needed to further improve the performance of the printed electronic devices. The challenges here are two-fold: first, the ink formulation determination/selection that can fulfill the requirements of morphological (i.e., lateral size and thickness of the dispersed 2D crystal flakes) and rheological (i.e., surface tension and viscosity of the inks) properties; second, the printing parameters for the deposition of homogeneous 2D crystal-based layers with clean interfaces will have to be optimized. In particular, the interfaces are strongly affected by the solvent and additives (i.e., surfactants/stabilizers) residuals,[108] which need to be minimized.

Notwithstanding the production method, understanding the precise layering and interface structure of the various 2D films is of utmost importance. The local mapping of strain and/or variations in lattice parameters, chemical composition, defect types, surface chemistry/composition, and resulting interface band structure are also critical in the design of electronic devices. Further, chemical doping
and functionalization are important in tuning optical and electronic properties of the devices. However, a detailed understanding of the charge transfer and transport properties, defects (dopants, grain boundaries -GBs-, point defects, etc.) or ambient contaminants (such as adsorbates, e.g., surfactants), chemical reactivity and edge terminations is still missing for the evaluation of 2D material (opto)electronic properties.

3. Figures of merit

The integration of GRMs in NVM cells aims at improving the figures of merit (FoM) of information-storage devices towards faster, smaller and cheaper memories. In assessing the performance of new technology approaches based on GRMs, it is important to benchmark the devices under development against state-of-the-art market products (Table 1), and compare their FoM with the corresponding requirements/projections of semiconductor industry roadmaps[^155-158] (e.g. Table 2). NVMs are commonly evaluated in terms of (1) speed, (2) scalability, (3) power consumption, (4) reliability, and (5) cost.

1. **Speed.** The speed of a memory device depends on the random access time to individual memory cells and on the effective time required to perform the write/erase operations (latency).[^159-160] Nowadays, there is still a large discrepancy between the data rate of processors — typically of the order of nanoseconds — and that of flash memories, which is limited by a long write/erase time of the order of hundreds of microseconds and by the slow serial access in NAND structures.[^3, 160] Processors are normally interfaced with high-performance yet volatile working memories, namely SRAMs and DRAMs, which are characterized by short latencies (1-100 ns), but are more expensive and occupy larger chip areas as compared to silicon flash memories.[^2, 5] It is worth noting that emerging NVMs based on the STT technology have already shown random access and write/erase time comparable to that of SRAMs/DRAMs,[^161] though at this stage they cannot compete with NAND chips as far as cost and bit density (e.g. 4 Gb/chip for the best STT-RAM[^162] vs 256 Gb/chip for 3D NAND[^163]). On the other hand, PCMs offer an interesting trade-off between speed and capacity. Recently, Micron
and Intel introduced in the market the novel 3D XPoint memory chips (128 Gb per die fabricated with the 20 nm node process)\cite{14-15, 164}, based on GST phase-change materials\cite{165} – that were claimed to have latencies of ~0.1–1 µs, i.e. 1000 times shorter than flash NAND\cite{6, 164}. These are the first example of SCM devices filling the speed-gap between working memories and storage\cite{6}.

2. **Scalability.** To further increase the density of memory devices, the memory cell has to be scaled. However, the scaling not only introduces challenges in processing but also makes the crosstalk (fringing fields) between neighbouring cells a limiting factor\cite{3, 160}. Today, the semiconductor industry is introducing products with novel 3D integration schemes and stacking beyond 64 layers of memory on one chip, achieving a record-high bit density, close to ≈0.5 GB∙mm$^{-2}$\cite{166}. Typical values of bit density for state-of-the-art market products are reported in Table 1 (expressed as the number of GB per chip) together with the corresponding cell size (expressed as multiples of F$^2$, where F is the technology feature size). For comparison, Table 2 displays the cell-size requirements for NVMs according to semiconductor industry roadmaps\cite{156}.

3. **Power consumption.** It refers to both the dynamic power consumption, quantified by the energy required for memory transitions (e.g. program energy per bit), and the static power dissipation, which stems from leakages during the storage time. In the case of flash memories, both static and dynamic power consumption increase upon scaling. Emerging NVM technologies, such as Spin-Transfer Torque Random Access Memories (STT-MRAM), Ferroelectric Random Access Memories (FeRAM), Phase-Change-memories (PCMs), and Resistance Random Access Memories (ReRAMs) consume significantly less power than silicon flash memories (see Table 1) and are more robust against power-consumption degradation upon miniaturization. For this reason, such technologies are expected to take over flash NANDs in SCM applications in helping data centres with their ever-increasing energy needs. Two-terminal memory cells, e.g. ReRAMs and PMCs, are often assessed in terms of their switching voltage $V_{\text{SET/RESET}}$ or switching current $I_{\text{SET/RESET}}$. This FoM is defined as the voltage/current that has to be applied to the device for inducing SET (high-to-low resistance) and RESET (low-to-high resistance) transitions and should be minimized to limit the dynamic power consumption.
4. **Reliability.** Proof-of-concept GRM-based NVM devices reported in the literature are evaluated mostly in terms of reliability, which comprises data retention and write endurance. The former refers to the amount of time for which the information can be retained within the cell and must be at least 10 years for any practical application (>3×10^8 sec). The latter, instead, quantifies the resistance to fatigue degradation, being defined as the highest number of write/erase (or program/erase) cycles that can be performed before the memory cell becomes unreliable. For example, flash NAND can withstand up to 10^5 write/erase cycles, whereas emerging NVM technologies such as STT-MRAM offer high endurance over 10^15 cycles (see Table 1 and 2). High endurance and long retention are essential to avoid bit errors and maintain good readability at any stage of the device lifetime. In this context, two FoM are frequently encountered in the literature of GRM-based NVMs. The first is the \( I_{on}/I_{off} \) switching ratio (or equivalently \( R_{on}/R_{off} \)), which is intended as the highest possible ratio between the current in the program (bit “0”) and erase (bit “1”) states. The second is the memory window \( \Delta V \) – commonly used for flash memories – that is the difference between the threshold voltages for the program and erase states of the transistor in the memory cell. Though there are no specific requirements on these FoM, both \( I_{on}/I_{off} \) and \( \Delta V \) should be maximized for improving the readability, as well as for enabling multilevel operation, i.e. the capability of storing multiple bits of information in a single cell.

5. **Cost.** Though strictly speaking cost is not a FoM, the cost of materials, processes and systems involved in the manufacturing of a memory chip must be carefully analysed to establish the viability of any new NVM technology. It is worth noting that the maximum acceptable cost for future successful market products strictly depends on the targeted devices application (e.g. embedded, stand-alone, high-speed, high-capacity, low-power memories, etc.). Table 1 displays typical price ranges for NVM chips currently available on the market. Besides magnetic hard-disk drives, which represent the cheapest (~0.1 $/GB) yet slowest (3-10 ms latency) storage option nowadays, NAND flash technology provides the smallest price per unit GB, though it is not fast enough to be used as a working memory. It is worth mentioning that bit-cost scalable (BiCS) 3D NANDs, originally introduced by Toshiba in 2007 and mass-produced since 2015 with a 48-layer stacking
process\textsuperscript{[160]}, can be fabricated with the same number of lithography steps regardless of the number of vertically-stacked layers, which allows for a continuous reduction of bit cost.\textsuperscript{[170]} Indeed, 3D NAND is the most promising and mature NVM technology offering ever-growing capacitance for low-cost massive data storage.

Table 1. Typical FoM values and market readiness for established and emerging memory technologies. The table is based on the data reported in refs \textsuperscript{[6-7, 155-157, 165, 167, 171-182]} and shows representative values, which may vary significantly in specific products.

<table>
<thead>
<tr>
<th>FoM</th>
<th>SRAM</th>
<th>DRAM</th>
<th>Flash NAND (planar)</th>
<th>ReRAM</th>
<th>FeRAM</th>
<th>PCM</th>
<th>STT-MRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density (bytes per chip)</td>
<td>≈10 MB</td>
<td>1-10 GB</td>
<td>≈10 GB</td>
<td>≈1 GB</td>
<td>≈1 MB</td>
<td>1-10 GB</td>
<td>10-100 MB</td>
</tr>
<tr>
<td>Cell size in F\textsuperscript{2}</td>
<td>&gt;100</td>
<td>6-10</td>
<td>4.5</td>
<td>6-20</td>
<td>15-40</td>
<td>6-20</td>
<td>35-40</td>
</tr>
<tr>
<td>Write time</td>
<td>&lt;10 ns</td>
<td>≈10 ns</td>
<td>≈100 µs</td>
<td>10-100 ns</td>
<td>≈100 ns</td>
<td>10-100 ns</td>
<td>≈10 ns</td>
</tr>
<tr>
<td>Program energy per bit</td>
<td>1-10 pJ</td>
<td>1-10 pJ</td>
<td>≈10 nJ</td>
<td>≈10 pJ</td>
<td>1 pJ</td>
<td>0.1-1 nJ</td>
<td>&lt;1 pJ</td>
</tr>
<tr>
<td>Retention</td>
<td>Volatile</td>
<td>Volatile</td>
<td>Non-volatile</td>
<td>Non-volatile</td>
<td>Non-volatile</td>
<td>Non-volatile</td>
<td>Non-volatile</td>
</tr>
<tr>
<td>Endurance (cycles)</td>
<td>&gt;10\textsuperscript{14}</td>
<td>&gt;10\textsuperscript{15}</td>
<td>10\textsuperscript{9}-10\textsuperscript{10}</td>
<td>10\textsuperscript{9}</td>
<td>&gt;10\textsuperscript{15}</td>
<td>10\textsuperscript{9}-10\textsuperscript{10}</td>
<td>&gt;10\textsuperscript{15}</td>
</tr>
<tr>
<td>Maturity</td>
<td>Product</td>
<td>Product</td>
<td>Product</td>
<td>Early product</td>
<td>Product</td>
<td>Early product</td>
<td>Early product</td>
</tr>
<tr>
<td>Market price ($/GB)</td>
<td>10k-100k</td>
<td>≈10</td>
<td>≈1</td>
<td>≈1k</td>
<td>10k-100k</td>
<td>10-100</td>
<td>1k-10k</td>
</tr>
</tbody>
</table>

Table 2. Technology requirements for NVM devices according to semiconductor industry roadmaps (refs \textsuperscript{[155-156]}).
4. Resistive NVMs based on graphene and its derivatives

Resistive random access memories are a class of memory devices that exploit the resistance switching (RS) of a material, commonly an insulator, to store digital information.\cite{5, 159} The prototypical ReRAM cell consists of an insulating layer, such as a binary metal oxide film (e.g. NiO\textsubscript{x}, TiO\textsubscript{x}, HfO\textsubscript{x}, etc.), sandwiched between two metal electrodes.\cite{5, 159, 183-185} Upon application of a SET bias voltage (V\textsubscript{SET}), the cell switches from a high resistance state (HRS, OFF state) to a low resistance state (LRS, ON state). The opposite transition (i.e. LRS→HRS) is obtained through the application of the RESET voltage (V\textsubscript{RESET}), which has the same (opposite) polarity of V\textsubscript{SET} in unipolar (bipolar) memory cells. A preliminary “forming” process, carried out at bias voltages higher than the SET and RESET voltages, occurring at the soft breakdown of the film, is required in order to achieve an effective HRS →LRS switching.\cite{186}

The first observation of hysteretic current-voltage characteristics in metal-insulator-metal (MIM) structures dates back to 1962 with the pioneering work of Hickmott,\cite{187} who was exploring the electrical properties of thin anodic oxide films sandwiched between different metal electrodes (e.g. Al-SiO-Au). In 1967, Simmons and Verderber\cite{188} reported on the voltage-controlled resistivity in Al-SiO-Au structures, a phenomenon that was attributed to the electrolytic injection of Au ions in the SiO insulator, resulting in a broad band of impurity levels that allows the electrons to move across the insulating barrier via tunneling between adjacent localized states.\cite{188} Since then, numerous RS phenomena have been reported and the corresponding charge-carrier conduction behaviours have been rationalized with a variety of models, such as for instance space-charge limited conductivity (SCLC), trap charging/discharging, Schottky barrier and Poole-Frenkel emission, as well as filamentary mechanisms that involve the formation and rupture of conducting filaments.\cite{5, 159, 189-190}

Compared to conventional flash memories, ReRAMs are advantageous in terms of power consumption and speed.\cite{5} Different from DRAM and FeRAM, the read operation in ReRAMs is non-destructive; moreover, these NVMs display excellent scalability and compatibility with mainstream CMOS technology.\cite{5, 159} Thanks to the high R\textsubscript{on}/R\textsubscript{off} ratio, ReRAM cells can in principle be
implemented without switching (or selector) device,\cite{183,191} as in the case of flash NAND,\cite{192-196} thus enabling the realization of high-density crossbar arrays with minimal footprint of $4F^2$.\cite{159} However, selector devices (e.g. transistors or rectifying diodes) are in practice required to reduce “sneak” leakage paths from unselected cells during the read operation.\cite{197} We refer the interested reader to the review article by A. Chen for more details on recent research on this topic (ref. \cite{197}). Hereafter, we will present and discuss the research efforts carried out in the last decade to integrate GRMs in ReRAM cells with the aim of enabling a new generation of low-cost flexible/transparent NVMs by making use of the unique physical and chemical properties of such promising nanomaterials. It is worth noting that different device architectures have been investigated, with both planar and vertical structures, and a variety of GRMs prepared with both top-down or bottom-up approaches (see Section 2) have been explored for applications in ReRAM devices. We will start with the description of the first planar graphene-based devices\cite{50-52} and we will then survey the most significant results on the use of graphene derivatives (e.g. graphene oxide -GO-), graphene-based composites, as well as 2D materials beyond graphene in selector-less ReRAM cells. Graphene and related 2D material-based selector devices are not discussed in this review as only few papers have been published so far on this subject, \cite{198-199} and further investigations are required towards high-density and integrated 2D ReRAMs.

4.1. Lateral ReRAMs based on graphene and its derivatives

Lateral ReRAMs consist in devices structured in a metal-RS-metal architecture with a standard planar geometry, in a sort of transistor configuration. Graphene and its derivatives have been investigated as the RS medium of the lateral ReRAM. For example, graphene, produced by micromechanical cleavage of pristine graphite,\cite{31} was the first 2D material to be investigated as a potential candidate for use in NVMs. In 2008, Echtermeyer et al.\cite{50} reported on RS in double-gated field-effect transistors (FETs) based on mechanically-exfoliated graphene sheets. The switching was attributed to chemical modifications of the graphene’s crystal structure, into GO or graphane (i.e. hydrogenated graphene), induced by the electrostatic field of the gate electrode in the presence of OH$^-$ and H$^+$.
species unintentionally adsorbed on the surface of the SiO$_2$/Si substrates in ambient air. This represents the first observation of reversible RS in graphene, which required the use of three-terminal device architecture. In the same year, Standley et al.\textsuperscript{[52]} presented a two-terminal metal-graphene-metal device with planar geometry displaying $R_{\text{on}}/R_{\text{off}} \approx 100$ and switching time $\tau_S \approx 100 \, \mu$s. Soon after, Li et al.\textsuperscript{[51, 200]} developed planar ReRAM cells using thin graphitic stripes (5-10 nm thick, 0.2-5.0 $\mu$m wide) as channel layers, achieving $R_{\text{on}}/R_{\text{off}}$ up to $10^7$ and $\tau_S \approx 1 \, \mu$s. When compared to graphene the main advantages are the simplicity of the production and the much smaller cost of the material. Upon application of a sufficiently large voltage, namely ca. 6 V (ref. \textsuperscript{[52]}) and 3-6 V \textsuperscript{[51, 200]}, the channel layer breaks down leading to a significant conductance drop, as illustrated in Figure 3 a-c. The mechanism responsible for the non-volatile RS relies on the bridging (SET) and gaping (RESET) of the graphene/graphite sheets occurring as the current density surpasses a critical value.\textsuperscript{[52]} In particular, the SET transition was attributed to the formation of conducting chains of carbon atoms moving under the driving force of the electric field and bridging the gap.\textsuperscript{[52]} Noticeably, the voltage required to break the junction to perform the RESET operation was found to be proportional to the channel length, which implies that the energy consumption diminishes upon miniaturization. However, Yao et al.\textsuperscript{[201-204]} observed RS in devices solely consisting of nano-gapped electrodes on SiO$_2$, raising doubts on the effective RS mechanism, which could eventually be ascribed to the reversible formation and modification of conducting silicon nanocrystals within the oxide substrate.\textsuperscript{[201-204]} New evidence in favour of the gaping/bridging hypothesis was enforced in 2012 by Zhang et al.,\textsuperscript{[205]} who reported on the direct visualization of the electrical breakdown of suspended graphene sheets via scanning electron microscopy (SEM), see Figure 3d-e. The observation of RS in suspended graphene layers allowed ruling the SiO$_2$ substrate as the main reason for the non-volatile RS occurring in the lateral device structures implemented by Standley et al.\textsuperscript{[52]} and the Tour group.\textsuperscript{[51, 200]}

The majority of reports dealing with GRM-based ReRAMs have been focused on chemical switching, which stems from electric-field induced redox reactions occurring within the 2D sheets, especially in graphene and its derivatives (e.g., changes between sp$^2$ and sp$^3$ carbon). In this context, GO
and reduced graphene oxide (RGO, conducting)\textsuperscript{215-218} have been widely investigated thanks to the possibility to tune their electronic, mechanical and optical properties by oxidation/reduction processes,\textsuperscript{146} see section 2.

By making use of the same planar configuration described previously for graphene/graphite layers, Panin et al.\textsuperscript{219} fabricated ReRAM cells by making use of thin films of GO flakes (few-layer thick) deposited via spray-coating techniques on SiO\textsubscript{2}/Si substrates pre-patterned with Al electrodes (channel length L \approx 20-25 \, \mu m). Upon application of high forming voltage (\approx 5 \, V), a well-defined unipolar RS was observed (R\text{on}/R\text{off} \approx 10^3). The authors explained the reversible voltage-driven RS with a “cluster structure” model, in which the GO flakes were assumed to consist of sp\textsuperscript{2} clusters within a sp\textsuperscript{3} matrix enriched with oxygen.\textsuperscript{219} Such sp\textsuperscript{2} graphene clusters are formed at high electric fields near the Al/GO interface through the electro-diffusion of oxygen (pre-forming process). During the SET/RESET operations, the clusters undergo a reversible sp\textsuperscript{2} \rightleftharpoons sp\textsuperscript{3} structure reconstruction, resulting in significant local resistance changes, which are at the origin of the observed RS behaviour.

Wu et al. \textsuperscript{220} investigated the electrical properties of ReRAM cells based on an individual graphene sheet suspended on pre-patterned ITO electrodes. Current-voltage (I-V) measurements carried out on the planar device revealed a large I\text{on}/I\text{off} ratio (\approx 10^6) and promising retention capability under ambient conditions (>10\textsuperscript{4} \, s). Data programming could be performed by applying a large bias voltage (\approx 7 \, V) between the two ITO electrodes, inducing a significant conductance drop.\textsuperscript{220} The latter was attributed to the local oxidation of the graphene sheet at the graphene/ITO interface, a mechanism that is believed to be responsible for the write-once read-many (WORM) behaviour of the device.\textsuperscript{220}

Though the memory cell could not be re-written electrically, the high-conductance state was recovered through a simple heat treatment (i.e. 10 mins at 200\textdegree C in both argon and air atmosphere). Even though lateral ReRAMs show intriguing results such as the bridging (SET) and gaping (RESET) of the graphene/graphite sheets and the electric-field induced redox reactions occurring within the 2D sheets, there are several issues regarding power-consumption and complexity of integration due to relatively large lateral dimension (i.e. \textit{> 50 \times 50 \, \mu m}^2), thus being much greater compared to state-of-
the-art. Because of these reasons, vertical ReRAMs may be more easily fabricated for practical applications and potential for 3D integration in multi-layered structures.

4.2. Vertical ReRAMs based on graphene derivatives

ReRAM devices based on solution-processed GO or RGO nanosheets have been often fabricated with vertical geometry by stacking metal-insulator-metal (MIM) layers in a crossbar array geometry (e.g. Figure 4a). When compared to lateral ReRAMs, the vertical ReRAMs displays various advantages including the low-driving voltage, simple fabrication process, the potential integration in multi-layered structures and which could yield in a dramatic increase of the storage density. In such a way, He et al. developed ReRAMs by making use of GO thin films prepared by vacuum filtration method. The Cu/GO/Pt memory cells displayed an appreciable non-volatile RS effect – characterized by $R_{on}/R_{off} \approx 20$, retention time $\tau_r \geq 10^4$ seconds, and switching voltage $|V_{SET/RESET}| \leq 1$ V – that was ascribed to the desorption/absorption of oxygen-containing groups in the GO sheets, as well as to the formation/rupture of conducting filaments diffused from the top Cu electrodes.

In 2010, Jeong et al. demonstrated flexible ReRAMs based on Al/GO/Al crossbar arrays fabricated on plastic substrates (i.e. polyethersulfone, PEI), as shown in Figure 4a. The insulating GO films, consisting of multiple overlapped or stacked GO flakes, were prepared via a simple spin-casting method. The resulting memories could maintain $R_{on}/R_{off}$ up to $10^3$ during 1000 mechanical bending cycles. By using transmission electron microscopy (TEM) technique, Kim et al. revealed that the switching mechanism of Al/GO/Al and Au/GO/Al devices is attributed to the formation of conducting filament induced by the migration of oxygen functional groups under the electric field. Hong et al. carried out a detailed analysis of the RS failure mechanisms revealing the critical role of the top electrode (Au or Al). In the case of Au, no oxygen migration occurs, whereas the Al electrode tends to be oxidized upon application of a high voltage bias due to the oxygen present in the GO sheets. The surface roughness of the bottom electrode can also play an important role. In fact, if the roughness is high, the top metal electrode could easily penetrate into the insulating GO
multilayer film, composed by more GO flakes stacked in a disordered way, and thus leading to filament formation, commonly called dendrites,\[^{226}\] hindering the switching operations and degrading the device endurance (≈ 100 cycles). X-ray photoelectron spectroscopy (XPS) measurements revealed the presence of metal atoms near the bottom electrode, which were considered to be responsible for the device failure. According to Hong et al.,\[^{54, 208}\] the formation of conducting filaments induced by Al diffusion is at the origin of device failure, whereas oxygen migration is the main mechanism responsible for the RS effect. Therefore, the choice of metals for top and bottom electrodes is critical for the reliable operation of GO-based ReRAM.

In order to avoid device failure due to the formation of permanent conductive paths, it is necessary to improve the GO deposition techniques to obtain continuous films with uniform coverage and quite low surface roughness (≤ 1 nm). One possibility consists in spray-coating while heating the substrate to avoid the so-called “coffee-ring effect”,\[^{227-228}\] which is known to degrade the quality of the deposited films.\[^{229-230}\] This method is compatible with large surface areas (e.g. > 10 × 10 cm\(^2\)) and it allows reducing the roughness of the GO film and preventing short-circuits due to dendrites from the top contact. Considering that GO is stable in water, a green-type low-temperature process to deposit GO films on large-area plastic substrates could be easily implemented (e.g. roll-to-roll).

Thin films of RGO have also been investigated as RS layers. Vasu et al.\[^{231}\] reported unipolar ReRAM cells based on thin RGO films (≈20 nm thick) displaying \(R_{\text{on}}/R_{\text{off}}\) up to \(10^5\). Such films were prepared via a room temperature drop-casting process onto ITO substrates, followed by Al or Au deposition. The ITO/RGO/Al and ITO/RGO/Au have shown similar current-voltage (I-V) characteristics, as in this case the unipolar RS effect could be attributed to the formation of nano-filaments of carbon atoms. Interestingly, the ReRAM cells could be programmed 10 µs voltage pulses, resulting in with \(R_{\text{on}}/R_{\text{off}}\approx 100\). However, RGO-based ReRAMs have relatively low \(R_{\text{on}}/R_{\text{off}}\) ratio due to their high off-current. Because of its high degree solubility in aqueous solutions, GO represents a better choice over RGO for application in ReRAM devices on flexible substrates, where solution-processing techniques for large-area coatings are required. However, promising strategies for the preparation of ReRAMs based on spin-coated RGO thin films have also been demonstrated. Figure 5a shows the
approach adopted by Zhao et al.\textsuperscript{[232]} which relies on the reduction of GO via a chemoselective photo-deoxidization process based on the ultraviolet irradiation of GO in the presence of 2,2,6,6-tetramethyl-4-piperidinol (TMP). This green organo-catalytic method allows for selective removal of carbonyl groups while leaving hydroxyl and epoxy groups on the surface of the sheets, thereby providing a good solubility in aqueous solutions. As a proof-of-concept, thin films ($\approx$30 nm thick) of chemo-selectively reduced GO (CRGO) were prepared via spin-coating from ethanolic solutions and were subsequently used as resistive-switching layers in MIM structures with ITO/CrGO/Au configurations. As-fabricated memory cells displayed a WORM-type switching behaviour (Figure 5b) with $I_{on}/I_{off} \approx 10^3$, which was ascribed to the irreversible degradation of oxygen-containing functional groups at bias voltages of $\approx$5 V. The retention time of the CRGO-based ReRAMs ($\approx 10^5$ s) was higher than the one achieved by equivalent devices based on GO. This difference is ascribed to the poor stability of the carbonyl and carboxyl groups, which are abundant on the GO surface but not on that of its decarbonylated counterpart, namely CRGO.

4.3. ReRAMs based on graphene composites

Various research groups have investigated graphene-based composite materials, such as combinations of GO/RGO with polymers,\textsuperscript{[207, 210-211, 233-235]} small molecules (e.g. ferrocene),\textsuperscript{[236]} metal nanoparticles (NPs)\textsuperscript{[216]} and organic nanocrystals (e.g. cellulose)\textsuperscript{[237]}, for application in ReRAMs. For instance, Zhuang et al.\textsuperscript{[210]} developed a soluble GO-polymer nanocomposite by grafting triphenylamine-based polyazomethine (TPAPAM) with terminal NH$_2$ groups to GO sheets, as illustrated in Figure 5c. Thin films ($\approx$50 nm thick) of such nanocomposites were sandwiched between ITO (bottom) and Al (top) electrodes, and served as RS layers in NVM cells (Figure 5d), which showed rewritable memory characteristics with stable current states for more than $10^8$ cycles and $I_{on}/I_{off}$ ratio exceeding $10^3$. Here, the switching mechanism was attributed to the reversible electrochemical reduction of GO enabled by charge-transfer interactions between the TPAPAM polymer and the GO sheets driven by the electric field. Nanocomposite materials based on combinations of Au NP and nanosheets of RGO were also developed and tested for application in ReRAM devices (ref. \textsuperscript{[216]}). A $\pi$-conjugated bifunctional molecular linker, namely 4-mercapto-
benzenediazonium tetrafluoroborate (MBDT) salt, was used to covalently attach Au NPs to RGO sheets obtained via chemical reduction of GO with hydrazine vapours.\textsuperscript{[216]} The resulting hybrid material was then used as channel layer in planar back-gated FETs (see Figure 5e), as well as in vertical two-terminal NVM cells consisting of \(\approx 50\) nm thick films embedded within ITO and Al electrodes, as portrayed in Figure 5f. The former devices displayed a significant nonlinear hysteresis with low \(I_{on}/I_{off}\) ratio (\(\approx 2\)), whereas the latter has shown stable ON/OFF current states (\(\geq 10^3\) s) and \(I_{on}/I_{off}\) up to \(\approx 100\). Control experiments carried out in the absence of Au NPs and/or of the MBDT linker, revealed that the memory switching arises from electron trapping/de-trapping at Au NPs covalently bound to the RGO sheets.\textsuperscript{[216]} Due to the large potential barrier between RGO and the Au NPs connected through MBDT, charge transfer between the two materials forming the nanocomposite occurs for bias voltages \(\geq 3\) V. After removal of the bias voltage, the trapped charge can be effectively stored within the Au NPs, enabling stable current states and relatively long retention time (\(\geq 10^3\) s). It is expected that the switching/retention characteristics of the hybrid RGO-NP layers can be further optimized via engineering of its material components, such as for instance by tuning the length of the molecular linker or the size/density of the NPs covalently bound to the RGO sheets.

Table 3 shows the FoM of the first resistive-switching memory cells based on graphene sheets, as well as those of the following ReRAM devices based on (functionalized-) GO, RGO and their related nanocomposites. It should be noticed that data on program/erase speeds and power consumption have not been included in the table, since they are rarely discussed and reported in the literature. Moreover, endurance and retention are often investigated for limited number of cycles and time, much smaller than the minimal requirements for NVMs (see Table 2). It is worth noting that the experimental
studies conducted by academic research laboratories in this research area give more emphasis to the preparation/synthesis of novel RS materials, accompanied by the demonstration of proof-of-concept memory cells. Hence, the majority of the publications cited so far do not provide a comprehensive benchmarking of the ReRAM devices, therefore a systematic analysis and comparison of their FoM is not always possible. At this stage, the graphene-based ReRAMs still require significant improvements, particularly in terms of data retention and cyclability, and intensive research and development efforts are necessary towards practical NVM technologies.

It should be mentioned that graphene has also been used as the electrode material in flexible/transparent ReRAMs (e.g. refs \cite{27, 117, 238}) or as an interface layer between the RS material and the electrodes leading to a number of benefits to the device properties, including transparency, high chemical stability, high thermal heat dissipation, low-power consumption,\cite{239} integration of built-in selector,\cite{240} and suppression of programming failure\cite{241} as recently reviewed by Hui \textit{et al.}\cite{64}.

\begin{table}
\centering
\begin{tabular}{|l|l|l|l|l|l|l|}
\hline
Active layer (thickness) & Electrodes (structure) & Flexible (substrate) & Proposed switching mechanism(s) & Current switching ratio & | Set voltage | Retention & Endurance [cycles] & Ref. \\
\hline
Graphite around SiO₂/Si NWs (5-10 nm) & Pt/Pt (planar) & No & Formation and breaking of atomic chains in nanogaps & $1.5 \times 10^7$ & 4-6 & 2 weeks & $>10^3$ & \cite{51} \\
Graphene (1-2L) & metal/metal* (planar) & No & Break junction and filament formation & $10^2$ & $\approx$6 & 24 h & $>10^3$ & \cite{52} \\
\hline
\end{tabular}
\caption{Summary of the main results obtained on resistive-switching NVMs based on graphene, GO, RGO, as well as graphene-based composites. Notes: The Al (or Au) electrode is commonly employed as the cathode, the ITO electrode as the anode. (*) Contact metal not specified. (**) CRGO: GO reduced \textit{via} a chemo-selective photodeoxidization process that preferentially removes carbonyl groups.}
\end{table}
### 5. 2D materials ‘beyond’ graphene for resistive NVMs

In addition to graphene and its derivatives/composites, a number of promising proof-of-concept devices have been implemented by making use of nanosheets of TMDs, in particular few, and single-layers of MoS$_2$, as well as insulating 2D materials (e.g. h-BN); more recently also BP has been explored for application in ReRAMs.\[64, 68, 189\] These 2D materials offer a wealth of properties,
complementary to those of graphene, such as on-demand energy bandgaps together with tuneable oxidation states and surface chemistry.

5.1. MoS$_2$-based nanomaterials via solution processing

MoS$_2$ is the most studied semiconductor among the family of layered TMDs. In the monolayer form, it has an optical bandgap of $\sim$1.9 eV\cite{242-243} combined with excellent mechanical flexibility\cite{244} and high charge-carrier mobility (> 20 cm$^2$/Vs for $N \sim 10^{11}$ cm$^{-2}$),\cite{245} which make it a prime candidate for next-generation flexible (opto)electronic devices, including memories.\cite{246-247} In comparison to graphene and its derivatives/composites, pristine nanosheets of MoS$_2$ do not display significant RS behaviour.\cite{64} However, the latter can be introduced via chemical functionalization methods or by mixing solution-processed MoS$_2$ nanosheets with other materials, such as dielectric polymers.\cite{248} In 2012, the Zhang’s group developed blends of 2D MoS$_2$ and polyvinylpyrrolidone (PVP),\cite{249} proving the potential of such hybrid materials for applications in rewritable ReRAMs. Their approach consists in sonicating the MoS$_2$ powder in ethanol in the presence of PVP, as illustrated in Figure 6a. The addition of the PVP was crucial, since MoS$_2$ does not possess suitable physical-chemical properties, e.g., surface tension, Hansen and Hildenrand parameters, for its exfoliation and dispersion in ethanol.\cite{108} Thin films of MoS$_2$/PVP blends were deposited by spin-coating on solution-processed RGO electrodes transferred on polyethylene terephthalate (PET) substrates.\cite{249} The fabrication process was completed by thermal evaporation of Al top electrodes, resulting in flexible rewritable memory devices with stable resistance states (see Figure 6b).\cite{249} The switching mechanism was deduced by fitting the I-V curves to power-law functions ($I \sim V^m$), which revealed the occurrence of space-charge limited conduction (SCLC) within the voltage range from 0.5 to +3.5 V ($m \approx 2$), and Ohmic conduction in the LRS ($m \approx 1$).\cite{249} The abrupt change in electrical resistivity upon application of a sufficiently high voltage – i.e., $\approx$ +3.5 V (SET) and -4.5 V (RESET) – across a $\approx$70 nm thick active layer, was ascribed to a possible trapping and de-trapping of charge carriers within the MoS$_2$ sheets of the composite material.\cite{249} A $\approx$10$^2$ $I_{on}/I_{off}$ ratio, maintained also during bending tests, revealed the potential of TMD materials for use in flexible NVMs.\cite{249} However, control experiments
with devices based on pure PVP films sandwiched between RGO and Al, necessary to confirm the role of MoS$_2$, were not reported by Zhang et al.,$^{[249]}$ and doubts remain on possible secondary effects, such as migration of oxygen species from the RGO surface or Al diffusion from the top electrode. It is worth noting that the reports cited by the authors to support their conclusions refer to MIM devices based on polyvinylphenol (PVPh) films sandwiched between Al and p-Si$^{[250]}$ or between two Al electrodes$^{[251]}$, but not to PVP films between RGO and Al.

More recently, the same group developed NVM cells based on hybrid films of MoS$_2$ nanobelts decorated with PtAg NPs and dispersed in a PVP polymer matrix.$^{[252]}$ The I-V characteristics display a marked hysteresis with negative differential resistance (NDR) occurring at high-voltage biases ($\pm 5$ V), which was attributed to charge trapping/de-trapping within the hybrid active layer.$^{[252]}$ However, the rapid discharging of the PtAg-MoS$_2$ nanobelts resulted in short-time data storage ($\ll 10$ years), which is not suitable for NVM applications.

Combinations of GO and MoS$_2$ via solution processing techniques have been investigated for use in flexible/translucent GRM-based NVMs.$^{[235, 253]}$ Two main approaches have been explored, namely (i) the deposition of MoS$_2$-GO mixtures from aqueous solutions by means of spraying methods,$^{[235]}$ and (ii) the sequential deposition of GO-MoS$_2$-GO stacks by spin-casting.$^{[253]}$ In the first case, the MoS$_2$ nanosheets, prepared by lithium-ion intercalation$^{[254]}$, were used to increase the electrical conductivity of the active layer with the aim to promote the migration of oxygen species from/to the GO sheets. The multicomponent MoS$_2$-GO films possess promising characteristics with low switching voltage ($\leq 1.5$ V) and appreciable $I_{on}/I_{off}$ ($\approx 10^2$).$^{[253]}$ In the GO-MoS$_2$-GO stacks deposited by sequential spin-casting, the disconnected metallic 1T-MoS$_2$ are embedded between two GO layers (see Figure 6d), acting as charge trapping centres. The amount of charge carriers trapped within the potential well of the GO-MoS$_2$-GO heterostructure (band diagram in Figure 6d) can be modulated by applying a voltage between the two electrodes of the memory cell.$^{[253]}$ The I-V characteristics are reported in Figure 6c and show a pronounced hysteresis, which stems from the trapping/de-trapping of charges in the MoS$_2$ nanosheets during the voltage sweep. ReRAM devices based on such multicomponent films, sandwiched between Al (bottom) and Au (top) electrodes, display $I_{on}/I_{off}$ as
high as $10^4$, which is about two orders of magnitude greater than in equivalent devices based on GO. Such remarkable memory effect enabled the realization of multilevel memory cells with at least four distinct resistance states, which could be systematically programmed by controlling the magnitude of the RESET voltage.[255]

In the last four years, a number of different strategies have been explored to improve the switching/retention capability of MoS$_2$-based ReRAMs, e.g. the preparation of core-shell structures consisting of MoS$_2$ thin layers (core) and metal-organic frameworks (shell, e.g. zeolitic imidazolate frameworks, ZIF-8),[256] the synthesis of hybrid nanofibers based on MoS$_2$ nanosheets and achiral copolymers — such as Pluronic P123 (PEO$_{20}$PPO$_{70}$PEO$_{20}$)[257] — as well as the combination of MoS$_2$ with different dielectric polymers, including polymethyl methacrylate (PMMA)[258-260] and polyvinyl alcohol (PVA).[261]

In early 2015, Bessonov et al.[78] reported breakthrough experiments on two-terminal memory cells based on vertical heterostructures of MoS$_2$ (or WS$_2$) and MoO$_x$ (or WO$_x$) encapsulated between Ag electrodes. The fabrication process of the ReRAM device is shown in Figure 6e and is briefly described in the following. Liquid-phase exfoliated semiconducting MoS$_2$ nanosheets are deposited by a modified Langmuir-Blodgett spreading technique[262] in the form of thin films – with thickness varying between 50 and 600 nm – on flexible polyethylene naphthalate (PEN) substrates pre-patterned with an array of silver electrodes obtained by screen-printing.[78] Prior to the deposition of the Ag top electrode, a thermal annealing step in ambient air (180-200 °C, 3 hours) is performed in order to oxidize the top MoS$_2$ surface, into a thin layer (< 3 nm) of MoO$_x$. The resulting devices display record-high $I_{on}/I_{off}$ ($\approx 10^6$) and remarkably low programming voltages, namely between 0.1 and 0.2 V (see Figure 6f), likely thanks to minimal Schottky barriers between the (doped) transition metal oxide and the Ag electrode. Moreover, bending tests confirmed the excellent mechanical strength and flexibility of the memory devices (> $10^4$ cycles). However, the exact mechanism underlying the memory switching is not yet fully understood. It might include a combination of phenomena, such as ion vacancy migration and trapping/de-trapping of charge carriers at the MoO$_x$/Ag interface.[78]
Non-volatile memory cells based on solution-processed 1T-MoS\(_2\) nanosheets sandwiched between Ag electrodes have been recently reported (ref. \([263]\)). Here, the choice of the 1T phase was found to be critical for the reliable operation of the device. In fact, resistive switching was observed only in the case of the 1T-MoS\(_2\) polytype, and no memory effect was observed for pristine (i.e., not oxidized) semiconducting 2H MoS\(_2\) using the same device configuration.\([263]\) Such a phase-dependent memristive behaviour was attributed to the hybridization of atomic orbitals leading to electron delocalization in the distorted 1T phase.\([263]\) Upon application of an external electric field, the displacement of Mo and S ions results in a lattice distortion, which enhances the electron delocalization and increases the conductivity of the active layer enabling the switching from HRS to LRS.\([263]\) The proof-of-concept devices show record-low switching voltages (< 100 mV), as well as appreciable endurance (> 1000 cycles) at the early stage of development.\([263]\) However, a systematic study of the stability/retention of the memory states has not been reported, so that the real potential of 1T-MoS\(_2\) nanosheets for NVM technologies remains to be explored.

Table 4 shows the FoM of the memory cells that exploit the properties of MoS\(_2\) nanosheets. It is worth noting that critical information such as time retention, endurance, active-layer thickness, are not always reported, hampering the systematic assessment and comparison of memory performance. Moreover, data on programming/erasing speeds are missing in all such publications, likely due to lack of appropriate device structures and electrical characterization equipment. At this stage, it appears that all the proof-of-concept devices, though revealing some intriguing characteristics, do not possess satisfactory FoM, especially in terms of time retention (<< 10 years), to compete with state-of-the-art ReRAM devices, where the HfO\(_2\)/TaO\(_x\) bilayer ReRAM shows the read/write latency of 300/100 ns, endurance of 10\(^7\) cycles and the retention time of 3 years.\([264]\)

Table 4. Comparison among the FoM of resistive NVMs (vertical sandwich structure) that incorporate solution-processed MoS\(_2\)-based nanomaterials. Notes: The Al (or Au) electrode is commonly employed as the cathode, the ITO electrode as the anode. If available, the thickness of the RS layer is reported in parenthesis. (*) FTO: fluorine doped tin oxide. (**) WORM: write once read many.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MoS2-PVP (≈70 nm)</td>
<td>Al/RGO</td>
<td>Yes (PET)</td>
<td>Charge trapping</td>
<td>10^3</td>
<td>3.5</td>
<td>-</td>
<td>-</td>
<td>[249]</td>
</tr>
<tr>
<td>PtAg-MoS2 nanobelts in PVP</td>
<td>Al/ITO</td>
<td>No</td>
<td>Charge trapping</td>
<td>-</td>
<td>-5</td>
<td>Volatile (DRAM)</td>
<td>-</td>
<td>[250]</td>
</tr>
<tr>
<td>MoS2/GO hybrid film (≈100 nm)</td>
<td>Al/ITO</td>
<td>No</td>
<td>Oxygen migration, Charge trapping</td>
<td>10^2</td>
<td>-2</td>
<td>-</td>
<td>-</td>
<td>[251]</td>
</tr>
<tr>
<td>GO/MoS2/GO heterostructure (total ≈ 20 nm)</td>
<td>Al/Au</td>
<td>No</td>
<td>Charge trapping</td>
<td>10^4</td>
<td>-4</td>
<td>10^4 (multilevel)</td>
<td>10^2</td>
<td>[252]</td>
</tr>
<tr>
<td>Pt-MoS2@ZIF-8</td>
<td>RGO/RGO</td>
<td>Yes (PET)</td>
<td>Charge trapping</td>
<td>10^3-10^4</td>
<td>3.3</td>
<td>1.5×10^3</td>
<td>WORM(□)</td>
<td>[253]</td>
</tr>
<tr>
<td>MoS2/P123 nanofibers</td>
<td>RGO/RGO</td>
<td>Yes (PET)</td>
<td>Charge trapping</td>
<td>10^3-10^3</td>
<td>4</td>
<td>4×10^3</td>
<td>-</td>
<td>[254]</td>
</tr>
<tr>
<td>CVD MoS2 coated with PMMA</td>
<td>Au/graphene</td>
<td>NO</td>
<td>Charge trapping</td>
<td>10^3-10^3</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>[255]</td>
</tr>
<tr>
<td>MoS2-QD-PMMA (total ≈ 200 nm)</td>
<td>FTO(□)/Au</td>
<td>NO</td>
<td>Charge trapping Quantum tunneling</td>
<td>10^2</td>
<td>0.5</td>
<td>10^4</td>
<td>10^2</td>
<td>[256]</td>
</tr>
<tr>
<td>MoS2 nanostructures in PMMA</td>
<td>Cu/ITO</td>
<td>Yes (PET)</td>
<td>Charge trapping</td>
<td>10^3-10^3</td>
<td>-2</td>
<td>10^5</td>
<td>10^3</td>
<td>[257]</td>
</tr>
<tr>
<td>MoS2-PVA</td>
<td>Ag/Ag</td>
<td>Yes (PET)</td>
<td>Charge trapping</td>
<td>10^2</td>
<td>3</td>
<td>10^4</td>
<td>10^2</td>
<td>[258]</td>
</tr>
<tr>
<td>MoOx/MoS2 (total 50-600 nm)</td>
<td>Ag/Ag</td>
<td>Yes (PEN)</td>
<td>Diffusion of oxygen vacancies</td>
<td>10^6</td>
<td>0.1-0.2</td>
<td>10^4</td>
<td>10^4</td>
<td>[71]</td>
</tr>
<tr>
<td>1T-MoS2 film (≈ 550 nm)</td>
<td>Ag/Ag</td>
<td>No</td>
<td>MoS displacement, Lattice distortion</td>
<td>10^2</td>
<td>&lt; 0.1</td>
<td>-</td>
<td>10^2</td>
<td>[259]</td>
</tr>
</tbody>
</table>

5.2. CVD-grown MoS2 and h-BN: the role of grain boundaries

A different type of NVM cell, with lateral structure (see Section 4.2), was demonstrated by Sangwan et al.,[265] who investigated memristive phenomena in monolayer MoS2 films grown by CVD. The authors reported on the electrical characterization of back-gated FETs in which grain boundaries (GBs) were intentionally included within the semiconducting channel connecting the source and drain electrodes (Au). Different electrode-GB geometries have been explored, and the best performing devices were found to be those with one GB connected to a single electrode, as illustrated in Figure 7a.[265] The resistance of such device could be modulated by means of relatively small in-plane electric
fields ($\approx 10^4$ V cm$^{-1}$) enabling $R_{\text{on}}/R_{\text{off}}$ up to $10^3$.\textsuperscript{[265]} The voltage-controlled RS was ascribed to the migration of dopant species, such as sulphur vacancies, between the GBs and their adjacent channel regions.\textsuperscript{[265]} Noteworthy, the switching voltage $V_{\text{SET}}$ could be controlled through the gate electric field, opening intriguing opportunities for developing electronic devices with both data-storage and computing capabilities. However, at this stage a few critical challenges have to be addressed, among them are the control the GBs’ alignment, orientation, chemical composition, and stoichiometry. We point out that the MoS$_2$ NVM based on GBs has not been termed ReRAM but rather as a memristor. The latter is the fourth fundamental passive circuit element, joining the resistor, capacitor and inductor, and was predicted theoretically by L. Chua in 1971\textsuperscript{[266]} and demonstrated experimentally by Strukov et al. in 2008.\textsuperscript{[267]} Whereas ReRAMs commonly display two distinct resistance states (i.e., HRS and LRS) with ohmic conduction,\textsuperscript{[268]} the typical memristor presents non-linear hysteretic I-V characteristics arising from the motion of charged atoms or molecular species.\textsuperscript{[267-268]} The difference between ReRAMs and memristors is actually rather subtle and the two terms have been often used interchangeably. It is worth noting that the borderline between the two fields may depend on the adopted definitions;\textsuperscript{[269-270]} for instance, according to the nomenclature given in ref. \textsuperscript{[270]} ReRAMs are a particular class of memristors. For further details on this topic we refer the reader to the review paper by E. Galle \textsuperscript{[268]}, which presents typical aspects of ReRAMs and memristors focusing on the prototypical case of TiO$_2$-based MIM devices.

The use of insulating 2D materials for application in ReRAM technologies has been explored by various research groups.\textsuperscript{[271-274]} In 2016, Qian and co-workers reported on flexible memory devices based on vertical MIM structures consisting of a $\sim 3$ nm thick CVD-grown h-BN film sandwiched between Cu (cathode) and Ag (anode) electrodes (Figure 7 e-f).\textsuperscript{[273]} The memory switching was found to originate from the electrochemical formation/disruption of conducting filaments bridging the two electrodes.\textsuperscript{[273]} Such mechanism enabled the fabrication of ReRAMs on bendable PET substrates with switching ratio of $\approx 100$, which was preserved during more than 700 bending cycles.\textsuperscript{[273]} High-resolution TEM investigations revealed that the Ag filaments grow from the anode to the cathode
with the narrowest part of the filament lying at the interface between h-BN and the cathode (Figure 7f)[273]

More recently, Lanza’s group[271] carried out a comprehensive experimental investigation of ReRAM cells based on multilayer h-BN grown by CVD on Cu and Ni-doped Cu foils. More than 300 memory cells were fabricated directly on the growth substrate using various combinations of electrode materials and active-layer thicknesses.[271] A promising combination of $R_{on}/R_{off}$ (up to $10^6$), low switching voltage ($\approx 0.4$ V) and endurance (> 600 cycles) was achieved, though data retention ($\approx 10$ hours) needs to be significantly enhanced for practical applications.[271] Two typical sets of FoM for thin (5-7 layers) and thick (15-20 layers) h-BN films are reported in Table 5. The RS was attributed to the formation of conductive filaments as a result of the diffusion of metallic ions from the electrodes through the GBs of CVD-grown multilayer polycrystalline h-BN.[271]

5.3. Solution-processed black phosphorous nanosheets

After the successful demonstration in 2014 of FETs based on atomically-thin BP,[38] research efforts has been devoted to the study and optimization of the physical properties and device applications of BP.[275-276] The main challenge when dealing with BP is its instability, i.e. degradation, under ambient conditions that hinders the use of LPE methods for large-scale production of BP nanosheets, although progress has recently been made with the use of anhydrous solvents, both high and low boiling ones.[127, 276-278] Hao et al.[279] succeeded in developing an effective drop-casting technique for the fabrication of flexible ReRAMs based on BP thin films sandwiched between ITO (bottom) and Al (top) electrodes. Their approach exploits the solvent-dependent degradation of BP upon exposure to ambient air, leading to the formation of an amorphous BP layer on the BP crystal, which is referred to as top degraded layer (TDL).[279] Besides preventing further degradation of the BP film, the TDL serves as thin insulating barrier below the top Al electrode and enables excellent memory switching characteristics, such as high $I_{on}/I_{off}$ ratio (up to $3\times10^5$) and retention time > $10^5$ sec.[279] The transition between HRS and LRS for this structure was ascribed to multiple mechanisms, including the diffusion oxygen ions deep into the BP film, charge trapping/de-trapping processes, as well as the formation
and disruption of conducting filaments inside the TDL. In 2015, Zhang’s group reported a solution-based approach for the preparation of quantum dots (QD) of BP (2-6 layers thick) with average lateral size ≈5 nm.[280] The QDs were found to be stable in N-methyl-2-pyrrolidone (NMP) and could be used in combination with the PVP polymer for the realization of flexible resistive memory cells. The latter displayed good stability and high $I_{on}/I_{off}$ ratio ($> 6 \times 10^4$), which is more than two orders magnitude larger than in ReRAM devices based on MoS$_2$-PVP films.[249]

Table 5. Comparison among the FoM of resistive memory cells based on CVD-grown MoS$_2$ and h-BN, as well as solution-processed BP. If available, the thickness of the active layer is reported in parenthesis.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CVD MoS$_2$ with GB (1L)</td>
<td>Ti-Au/Ti-Au (planar)</td>
<td>No</td>
<td>Migration of S vacancies at GBs</td>
<td>$\approx 10^3$</td>
<td>8.3 (V$_g$ tunable)</td>
<td>$&gt; 120$</td>
<td>-</td>
<td>[265]</td>
</tr>
<tr>
<td>CVD h-BN ($\approx 3$ nm)</td>
<td>Cu/Ag (vertical)</td>
<td>Yes (PET)</td>
<td>CF growth</td>
<td>$10^2$</td>
<td>0.72</td>
<td>$3 \times 10^3$</td>
<td>550</td>
<td>[273]</td>
</tr>
<tr>
<td>CVD h-BN with GB (5-7 L)</td>
<td>Cu/Ti (vertical)</td>
<td>No</td>
<td>CF growth at GBs with B vacancies</td>
<td>$10^6$</td>
<td>2.2</td>
<td>-</td>
<td>$&gt;350$</td>
<td>[271]</td>
</tr>
<tr>
<td>CVD h-BN with GB (15-20 L)</td>
<td>CuNi/Ti (vertical)</td>
<td>No</td>
<td>CF growth at GBs with B vacancies</td>
<td>$10^6$</td>
<td>2.2</td>
<td>-</td>
<td>-</td>
<td>[271]</td>
</tr>
<tr>
<td>BP film by solution processing ($\approx 3 \mu$m)</td>
<td>ITO/Al (vertical)</td>
<td>Yes (PET)</td>
<td>Oxygen diffusion Charge trapping CF growth in TDL</td>
<td>$3 \times 10^3$</td>
<td>1.5-2</td>
<td>$10^5$</td>
<td>-</td>
<td>[279]</td>
</tr>
<tr>
<td>BP-QD/PVP</td>
<td>Au/Ag (vertical)</td>
<td>Yes (PET)</td>
<td>Charge trapping</td>
<td>$6 \times 10^4$</td>
<td>-1.2</td>
<td>$1.1 \times 10^3$</td>
<td>-</td>
<td>[280]</td>
</tr>
</tbody>
</table>

33
6. Flash memories based on GRMs

Flash memory is an electrically erasable programmable read only memory (EEPROM) in which individual blocks (i.e. sub-arrays of memory cells) can be erased at one time.\textsuperscript{[160, 192]} Since the demonstration of the first prototype by Masuoka et al.\textsuperscript{[281]} (Toshiba) in 1984, flash memories have been the workhorse NVM technology; nowadays they are making inroads in high-capacity solid state drives (SSDs) that complement/replace magnetic hard-disk drives.

During the last three decades, a large variety of flash-memory devices have been developed and manufactured to satisfy the requirements of numerous technological applications. They can be classified according to multiple criteria, such as for instance the programming/erasing mechanisms – e.g. Fowler-Nordheim tunneling or channel hot electron (CHE) injection – or the access type to the memory cells, e.g. parallel or serial.\textsuperscript{[192]} The market is currently dominated by two types of flash memories, namely NOR and NAND. The former provides high speed, random access to data and is commonly used for storing code in embedded systems, whereas the latter allows for high-capacity data storage though with longer read latencies.\textsuperscript{[3, 160, 192]} Vertical stacking of multiple NAND layers in the so called 3D NAND technologies, e.g., BiCS (bit-cost scalable, Toshiba)\textsuperscript{[168, 282]} and TCAT (terabit cell array transistor, Samsung)\textsuperscript{[283]}, is nowadays the most mature industrial approach to maintain the pace with the ever-growing NVM market, which requires continuous reduction of bit cost as well as relentless improvements in capacity and speed of data-storage devices.

The basic building block of flash memory is the floating-gate field-effect transistor (FG-FET), see Figure 8. The FG electrode is embedded within the gate insulator stack and is separated from the transistor channel by the tunnelling oxide and from the control gate by the blocking oxide.\textsuperscript{[3]} The information stored within the cell (bit “0” or “1”) depends on the amount of charge accumulated in the FG layer, which determines the threshold voltage of the transistor and thus also the magnitude of the current that can be read upon application of a bias voltage between the source and drain electrodes. As the FG is entirely surrounded by an insulating material, the charge can be maintained for long time (> 10 years) also in the absence of a supply power. In conventional flash memories polysilicon
is used as FG-electrode material for storing charges, whereas in charge-trap flash memories\textsuperscript{[284]} a nitride (Si$_3$N$_4$) trapping layer is used, such as for example in SONOS (Si-Oxide-Nitride-Oxide-Si) or TANOS (TaN-Al$_2$O$_3$-Nitride-Oxide-Si) flash technologies.\textsuperscript{[160, 284]}

Figure 8a portrays a schematic view of three adjacent FG-FETs connected in series, as in the case of Si NAND flash memories. Due to parasitic fringing fields, the FG of one cell can interfere with the FGs of the neighbouring cells, resulting in undesired threshold-voltage variations via capacitive coupling. Such interferences can be limited by scaling the thickness of the FG (t$_{FG}$) or by making use of insulating charge-trapping layers, such as nitrides.\textsuperscript{[284]} On the one hand, the FG-thickness scaling is hampered by the onset of leakage currents, as schematically illustrated in Figure 8b, which severely degrades the operation and reliability of the memory cell for t$_{FG}$ ≤ 10 nm.\textsuperscript{[285]} On the other hand, the use of charge-trapping layers poses new challenges for achieving both good data retention and fast erase speed via Fowler-Nordheim tunneling.\textsuperscript{[284]} In this context, high-κ dielectric materials (e.g., HfO$_2$, Al$_2$O$_3$, ZrO$_2$, Ta$_2$O$_5$, etc.) are being investigated as tunnelling/blocking layers in charge-trapping flash memories, as they enable further scaling the gate stack along the vertical direction.\textsuperscript{[286]}

However, the use of metal control-gate electrodes in combination with high-κ oxides results in significant reliability issues, as described by Degraeve et al.\textsuperscript{[287]}. In fact, a number of critical hurdles have to be faced to continue the miniaturization of flash memories, and memory manufacturers are making use of vertical integration/stacking strategies to manufacture 3D flash memories consisting of multi-stacked arrays of memory cells.\textsuperscript{[288]} We refer the reader to the review articles by K. Takeuchi \textsuperscript{[289]} and S. Lee\textsuperscript{[290]} for a detailed analysis of the scaling challenges and their potential solutions.

Graphene and related 2D materials, in particular graphene and 2D semiconductors (e.g., monolayers of TMDs or BP), have been investigated for application in flash devices since 2010. The first reports on this subject focused on the use of GO,\textsuperscript{[291]} graphene,\textsuperscript{[292]} multilayer graphene (≤ 5 nm)\textsuperscript{[60, 293]} as charge-storage layers in FG-FETs at the place of polysilicon. Besides their minimal thickness which is ideal for minimizing cell-to-cell interferences, thin graphitic layers may help suppress ballistic leakage currents thanks to their relatively high resistivity along the out-of-plane c-axis ($\rho_\perp$), approximately 100 times larger than its in-plane counterpart ($\rho_\parallel$).\textsuperscript{[294]} Moreover, the breadth of
electron properties available in the class of GRMs has been explored for improving various FoM, e.g., retention and memory window, through a careful choice of the material building blocks as well as through engineering of the work function and density of electronic states.\cite{295}

In 2010, Wang et al.\cite{291} replaced the nitride layer of a TANOS charge-trap flash memory with monolayer GO sheets to obtain the following stack: TaN-gate/Al$_2$O$_3$/GO/SiO$_2$/p-Si. In such device architecture, electrons can be stored within deep-level traps in the GO sheets, resulting in a memory window $\Delta V$ as high as $\approx$ 7.5 V (see Table 6). Soon after, Hong et al.\cite{292} explored the use of single-layer graphene (SLG) and multilayer graphene (MLG) as charge-trapping layers in two-terminal flash memories, as illustrated in Figure 9a-c. Such devices displayed a wide memory window, up to $\approx$ 7 V for MLG devices and $\approx$ 2 V for SLG, as well as good retention time ($\approx$ 8% loss in 10 years). Mishra and co-workers\cite{60,293} proposed to use MLG as the FG material rather than SLG, due to its higher density of states ($\approx$ 4.4x10$^{13}$ cm$^{-2}$eV$^{-1}$ vs $\approx$ 8x10$^{12}$ cm$^{-2}$eV$^{-1}$)\cite{296} and higher work function ($\approx$ 4.7 eV for MLG vs $\approx$ 4.2 eV for SLG).\cite{297} Moreover, the high electrical resistivity of MLG along the out-of-plane ($\rho_\perp/\rho_\parallel \approx 100$)\cite{294} would allow suppressing detrimental leakage currents hampering the vertical scaling of flash memories.\cite{60} Single layer graphene was also exploited as the control-gate electrode in charge-trap flash memories\cite{298} with the aim of overcoming reliability issues associated with the high-$\kappa$/metal-gate stack.\cite{287} It was shown that replacing the metal control gate with an ultrathin graphene sheet allows minimizing the detrimental mechanical stress that affects the endurance and retention of the memory cell.\cite{299}

After the demonstration of the first top-gated monolayer MoS$_2$ FET in 2011\cite{300} the potential of 2D semiconductors for pursuing aggressive miniaturization of electronic devices started to be widely recognized,\cite{301-302} and MoS$_2$ and 2D semiconductor-based transistors became one of the top-ten “hottest research front” in Physics according to a citation-based study released in 2015.\cite{303}

Atomically-thin semiconducting materials rapidly attracted attention also for use in flash memories. In early 2013, Bertolazzi et al.\cite{61} implemented a proof-of-concept FG-FET (see Figure 9d and Figure 1c) that combined the unique properties of 2D MoS$_2$, used as the transistor channel material, with those of MLG (4-5 layers thick), which served as charge-trapping layer. High-$\kappa$ HfO$_2$ films on the
other hand were used as both tunnelling and blocking oxide layers (≈ 6 nm and 30 nm thick, respectively), whereas CVD-grown SLG sheets were used as source and drain electrodes resulting in linear current-voltage output characteristics at room temperature. Following this approach, a large memory window (≈ 8 V) and program/erase current ratio of ≈10^4 were demonstrated (Figure 9e). Soon after, Choi et al. presented FG-FETs based on heterostructures of graphene, h-BN, and MoS2, which were electrostatically gated via a SiO2/Si gate stack. Thin flakes of h-BN (5-15 nm) served as tunnelling barriers, while graphene and MoS2 sheets, obtained via micromechanical cleavage, were alternatively used as the charge-storage and channel layer. The two configurations displayed remarkably different characteristics due to the semiconducting and semi-metallic nature of MoS2 and graphene, respectively. For instance, when graphene was used as the channel (Figure 9f) a low program/erase current ratio was obtained (I_{on}/I_{off} ≈2, Figure 9g), whereas in the opposite configuration I_{on}/I_{off} ratio as high as ≈10^4 was achieved, comparable to the results obtained by Bertolazzi et al. After these early works, a number of follow-up studies were conducted to identify the best GRMs for use as charge-storage layers in flash memories (see Table 6). Various research groups focused their attention on the combination of metal NPs, polymers, high-κ dielectrics and 2D TMDs for improving the memory window, the charge retention and the switching ratio of flash-memory cells. The group of S.-Y. Choi developed a low-power flash-memory cell consisting of a few-layer MoS2 channel, a polymer tunnelling layer, namely a ≈10 nm thick film of poly(1,3,5-trimethyl-1,3,5-trivinyl cyclotrisiloxane) (pV3D3) formed via a solvent-free initiated chemical vapor deposition (iCVD) process, as well as an Al2O3 blocking oxide (≈20 nm) deposited by ALD onto a Au-nanoparticle charge-storage layer (≈4 nm). The memory cells displayed promising characteristics, namely endurance >10^3 cycles, state-of-the-art retention time (>10 years) and high switching ratio (I_{on}/I_{off} ≈ 10^6), being the highest obtained so far in GRM-based flash devices (see Table 6).

Given the potential of BP as a 2D semiconductor, Dong et al. built a flash memory cell based on heterostructures of BP (channel, ≈ 8 nm), h-BN (tunneling layer, ≈ 25 nm), MoS2 (charge-
storage layer, ≈ 8 nm), SiO$_2$ (blocking layer, 300 nm) on heavily-doped Si (control gate, substrate), exhibiting a moderate program/erase ratio (≈50) due to the small bandgap of BP (< 0.4 eV for ≥ 5 layers)$^{[312]}$ but with a remarkable memory window (i.e., ≈ 60 V for a ±40 V gate-voltage range). The latter was attributed to the ambipolar nature of BP that allows for injection and storage of both electrons and holes within the MoS$_2$ charge-storage layer. Subsequent experimental$^{[70, 313]}$ and theoretical$^{[314]}$ studies were focused on heterostructures of high-κ dielectrics (Al$_2$O$_3$)$^{[70, 313]}$ and mechanically-exfoliated multilayer BP sheets (≥ 5 nm). The latter were used both as channel and charge-storage layers,$^{[313]}$ while Al$_2$O$_3$ served as tunneling and blocking layer.$^{[70, 313]}$ The results are reported in Table 6, which compares the materials, the structures and the most important FoM of the proof-of-concept flash-memory cells discussed in this section. At this stage, it is clear that a significant amount of work needs to be carried out to improve the retention and the endurance of most of the GRM-based flash-memory cells, similar to the case of the ReRAM devices discussed in Section 4. Indeed, the most important challenge consists in developing suitable integration strategies to develop high-quality heterostructures of multiple GRMs (TMDs, BP, h-BN, graphene, etc.), see Section 2 for further details, as well as stacks of GRMs and high-κ dielectrics with a reduced number of defects and charge traps.$^{[315]}$ These challenges have to be overcome to make the lab-to-fab transition a realistic path for GRMs towards flash-memory devices that exploit their unique properties.
Table 6. Comparison among the FoM of flash memories that exploit the properties of GRMs.

<table>
<thead>
<tr>
<th>Active layer (thickness)</th>
<th>Charge storage layer (thickness)</th>
<th>Tunnel/control layer (thickness)</th>
<th>Control gate electrode</th>
<th>Memory window [V] (V_{ox} range)</th>
<th>Current switching ratio</th>
<th>Retention</th>
<th>Endurance [cycles]</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-type Si (bulk)</td>
<td>Isolated GO sheets (1L)</td>
<td>SiO_{2}/Al_{2}O_{3} (5/15 nm)</td>
<td>TaN</td>
<td>7.5 (from -5 to 14 V)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>[21]</td>
</tr>
<tr>
<td>p-type Si (bulk)</td>
<td>CVD graphene (SLG or MLG)</td>
<td>SiO_{2}/Al_{2}O_{3} (5/35 nm)</td>
<td>Ti/AlAu</td>
<td>≈2, SLG (±7 V) ≈7, MLG (±7 V)</td>
<td>-</td>
<td>8% loss in 10 years</td>
<td>-</td>
<td>[24]</td>
</tr>
<tr>
<td>p-type Si (bulk)</td>
<td>RGO sheets (2-3 nm thick)</td>
<td>SiO_{2}/Al_{2}O_{3} (EOT: 13.5 nm)</td>
<td>TiN</td>
<td>2.61 (±10 V) 6.6 (±18 V)</td>
<td>-</td>
<td>N.A.</td>
<td>-</td>
<td>[30]</td>
</tr>
<tr>
<td>p-type Si (bulk) with n++ source/drain</td>
<td>RGO sheets (&lt;5 nm)</td>
<td>SiO_{2}/Al_{2}O_{3} (8/22 nm)</td>
<td>TiN</td>
<td>9.4 (±20 V)</td>
<td>-</td>
<td>26% loss in 10 years</td>
<td>≈10^3</td>
<td>[20]</td>
</tr>
<tr>
<td>p-type Si (bulk)</td>
<td>SiN4 (6 nm)</td>
<td>SiO_{2}/Al_{2}O_{3} (4/12 nm)</td>
<td>CVD graphene</td>
<td>16.8 (E = 1.6 V cm^{-1})</td>
<td>-</td>
<td>&gt;10 years</td>
<td>-</td>
<td>[20]</td>
</tr>
<tr>
<td>MoS_{2} (1L)</td>
<td>MLG (4-5 layers)</td>
<td>HfO_{2}/HfO_{2} (6/30 nm)</td>
<td>Cr/Au</td>
<td>&gt;8 (±15 V)</td>
<td>≈10^4</td>
<td>70% loss in 10 years</td>
<td>&gt;120</td>
<td>[31]</td>
</tr>
<tr>
<td>Graphene (1L)</td>
<td>MoS_{2} (5 nm)</td>
<td>h-BN/SiO_{2} (6/280 nm)</td>
<td>Si substrate</td>
<td>&gt;20 (±40 V)</td>
<td>≈2</td>
<td>&gt;1200 s</td>
<td>&gt;100</td>
<td>[30]</td>
</tr>
<tr>
<td>MoS_{2} (3L)</td>
<td>Graphene (2L)</td>
<td>h-BN/SiO_{2} (12/280 nm)</td>
<td>Si substrate</td>
<td>&gt;15 (±15 V)</td>
<td>≈10^4</td>
<td>&gt;1400 s</td>
<td>&gt;100</td>
<td>[30]</td>
</tr>
<tr>
<td>MoS_{2} (few-layer)</td>
<td>Au NP (&gt;4 nm)</td>
<td>PV3D3/Al_{2}O_{3} (10/20 nm)</td>
<td>Cr/Au</td>
<td>5.2 (±13 V)</td>
<td>≈10^5</td>
<td>&gt;10 years</td>
<td>&gt;10^5</td>
<td>[31]</td>
</tr>
<tr>
<td>p-type Si (bulk)</td>
<td>Isolated nanographene sheets (1L)</td>
<td>SiO_{2}/Al_{2}O_{3} (4/15 nm)</td>
<td>Al</td>
<td>4.5 (±8 V)</td>
<td>-</td>
<td>44% loss in 10 years</td>
<td>&gt;10^5</td>
<td>[32]</td>
</tr>
<tr>
<td>Pentacene (40 nm)</td>
<td>RGO (±1 nm) on Au NPs (±15 nm)</td>
<td>Al_{2}O_{3}/Al_{2}O_{3} (10/30 nm)</td>
<td>Ag on PET substrate</td>
<td>1.95 (-5-0 V)</td>
<td>≈10^4</td>
<td>&gt;10^5 s</td>
<td>&gt;10^5</td>
<td>[32]</td>
</tr>
<tr>
<td>Pentacene (40 nm)</td>
<td>MoS_{2} (1L) on Au NPs (±5 nm)</td>
<td>Al_{2}O_{3}/SiO_{2} (5/100 nm)</td>
<td>Si substrate</td>
<td>19 (±50 V)</td>
<td>≈10^3</td>
<td>&gt;10^5 s</td>
<td>&gt;100  (8 levels)</td>
<td>[32]</td>
</tr>
<tr>
<td>MoS_{2} (1L)</td>
<td>Au NPs (3-5 nm)</td>
<td>HfO_{2}/HfO_{2} (6.5/20 nm)</td>
<td>Cr/Au</td>
<td>&gt;10^4</td>
<td>≈10^3</td>
<td>40% loss in 10 years</td>
<td>-</td>
<td>[32]</td>
</tr>
<tr>
<td>MoS_{2} (3-4L)</td>
<td>HfO_{2} (8 nm)</td>
<td>Al_{2}O_{3}/Al_{2}O_{3} (7/30 nm)</td>
<td>Cr/Au</td>
<td>&gt;20 (±26 V)</td>
<td>≈10^4</td>
<td>28% loss in 10 years</td>
<td>&gt;120</td>
<td>[31]</td>
</tr>
<tr>
<td>BP (8 nm)</td>
<td>MoS_{2} (6 nm)</td>
<td>h-BN/SiO_{2} (25/300 nm)</td>
<td>Si substrate</td>
<td>&gt;60 (±40 V)</td>
<td>≈10^3</td>
<td>&gt;10^7</td>
<td>40</td>
<td>[71]</td>
</tr>
<tr>
<td>BP (+11L)</td>
<td>BP (±10L)</td>
<td>Al_{2}O_{3}/Al_{2}O_{3} (5/35 nm)</td>
<td>Au</td>
<td>&gt;22 (±25 V)</td>
<td>≈10^3</td>
<td>&gt;10^7</td>
<td>-</td>
<td>[31]</td>
</tr>
<tr>
<td>BP (5nm)</td>
<td>control oxide acts as CSL</td>
<td>-Al_{2}O_{3} (-/30 nm)</td>
<td>Cr/Au</td>
<td>23-27 (±20 V tunable)</td>
<td>40×10^{4} tunable</td>
<td>100% loss in 2×10^7 s</td>
<td>&gt;100</td>
<td>[30]</td>
</tr>
</tbody>
</table>
7. Other emerging NVM cells based on GRMs

Resistive random access memories and flash memories are seemingly the most investigated devices architecture in the context of GRM research. However, several reports also exist on ferroelectric random-access memories (FeRAMs), PCMs, as well as tunneling random access memories (TRAMs), which we briefly review in the next sub-sections.

7.1. Ferroelectric memories

The high polarization field and dielectric constants of ferroelectric materials produce very efficient gating effects in 2D materials, modulating charge density and all physical properties over wide range of applied voltages. Based on this, many novel device functionalities and architectures have been recently proposed including ferroelectric 2D memory devices,[85, 316-320] highly sensitive phototransistors,[317, 321-322] low-power field-effect transistors enabled by high-κ ferroelectrics,[321, 323-325]. Most of the reports in this research area have focused on ferroelectric FET structures (FeFET) where a ferroelectric layer serves as gate dielectric material.[326] Ferroelectric devices incorporating a graphene layer have been investigated since 2009.[53, 318-319, 327] Zheng et al.[53] were the first to demonstrate a ferroelectric-gated graphene FET with double-gate structure consisting of a thin film of poly(vinylidene fluoride/trifluoroethylene), P(VDF/TrFE), see Error! Reference source not found. a, b, deposited by spin-coating on the graphene surface. The memory cell exhibited two bistable current states (program/erase current ratio of ≈ 2), whose origin was identified in the electrostatic doping of graphene due to gate-tunable electric dipoles at the P(VDF/TrFE)/graphene interface.[53] Noticeably, these novel graphene-polymer memory cells can be easily integrated onto flexible substrates with high mechanical flexibility and transparency.[328]

A similar approach was adopted with monolayer MoS2,[329] (Figure 10 c, d) and BP[316] (Figure 10 e, f) resulting in high-mobility transistors, up to 220 cm²V⁻¹s⁻¹ ([329]) and ≈ 1160 cm²V⁻¹s⁻¹ ([316]), and NVMs with program/erase ratio exceeding 10³ in both cases. Thin films of ferroelectric lead-zirconate-titanate (PZT) were investigated by Song et al.,[318] who used PZT as substrate and dielectric material in back-gated FETs based on graphene grown by CVD. The authors attributed the large
memory window ($\Delta V_M \approx 4.2$ for a gate-voltage range of $\pm 6$ V) observed in their devices not only to interfacial electric dipoles, but also to a significant trapping/de-trapping of charge carriers at the PZT/graphene interface.[318] Recently, Ko et al.[330] reported on ferroelectric transistors based on mechanically-exfoliated nanosheets of TMDs, among which MoS$_2$ and WSe$_2$, deposited with the scotch-tape method[27] on highly-crystalline PZT thin films prepared by pulsed laser deposition on SrTiO$_3$ substrates. These devices have shown very promising results, such as current switching exceeding $10^4$ and low programming/erasing voltages (<2.5 V).[330]

Recently, Wang et al.[331] developed a top-gate NVM FeFETs with MoSe$_2$ nanosheets of varying thickness as the channel gated by a ferroelectric film. The best performant device, i.e., in term of retention time and endurance, was obtained for the monolayer thick MoSe$_2$, exhibiting the largest hysteresis with $I_{on}/I_{off} \approx 10^5$.[85] The increase of the MoSe$_2$ thickness (i.e., from single-layer to $\approx 10$ nm) determined a narrowing of the memory window and a decrease of the write/erase ratio by two orders of magnitudes. In Figure 11, the best performances of the FeFET devices (with a single-layer MoSe$_2$ channel) are presented through the transfer curves acquired at different source/drain voltages. Two memory states (write and erase) are clearly seen with an $I_{on}/I_{off}$ greater than $10^5$ at $V_G = 0$ V and $V_{SD} = 1$ V. The bottom left panel shows the data write and erase speed tests at various voltage pulse widths and for a pulse height of 40 V. With increasing channel thickness, degraded memory window, retention as well as endurance characteristics are reported; however, the transistors still exhibit reasonable memory performances.[85]

Recent progress in FeFET-based devices combining GRMs and organic ferroelectric gate dielectrics brings novel prospects for future applications in non-volatile ferroelectric memories. A comparison among the FoM of the proof-of-concept FeRAM devices is provided in Table 7.
Table 7. Comparison among the FoM of FeFETs that exploit the properties of GRMs.

<table>
<thead>
<tr>
<th>Active layer (thickness)</th>
<th>Ferroelectric layer (thickness)</th>
<th>Control gate (position)</th>
<th>Current Switching Ratio</th>
<th>Current Switching Ratio (Δn or ΔV [V])</th>
<th>V_{sw} range [V]</th>
<th>Retention [sec]</th>
<th>Endurance [cycles]</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphene (1L)</td>
<td>P(VDF-TrFE) (~700 nm)</td>
<td>Au (top)</td>
<td>4.5</td>
<td>Δn=6</td>
<td>±85</td>
<td>-</td>
<td>-</td>
<td>[32]</td>
</tr>
<tr>
<td>Graphene (1L)</td>
<td>P(VDF-TrFE) (~500 nm)</td>
<td>Au (top)</td>
<td>6</td>
<td>Δn=2</td>
<td>±30</td>
<td>-</td>
<td>10^3</td>
<td>[233]</td>
</tr>
<tr>
<td>CVD graphene (1L)</td>
<td>P(VDF-TrFE) (220 nm)</td>
<td>Pt (top)</td>
<td>7.75</td>
<td>Δn=3</td>
<td>±30</td>
<td>10^3</td>
<td>10^3</td>
<td>[333]</td>
</tr>
<tr>
<td>Graphene (1L)</td>
<td>P(VDF-TrFE) (200 nm)</td>
<td>Au (top)</td>
<td>2.5</td>
<td>Δn=15</td>
<td>±20</td>
<td>10^5</td>
<td>-</td>
<td>[334]</td>
</tr>
<tr>
<td>CVD graphene (1L)</td>
<td>PZT (150 nm)</td>
<td>Pt (bottom)</td>
<td>4.2</td>
<td>ΔV=4.2</td>
<td>±6</td>
<td>10^3</td>
<td>-</td>
<td>[318]</td>
</tr>
<tr>
<td>CVD graphene (1L)</td>
<td>Epitaxial PZT (140 nm)</td>
<td>SrRuO3 (bottom)</td>
<td>-</td>
<td>Δn=10</td>
<td>±5</td>
<td>3×10^4</td>
<td>10^3</td>
<td>[335]</td>
</tr>
<tr>
<td>CVD graphene (1-2L)</td>
<td>PZT (360 nm)</td>
<td>Pt (bottom)</td>
<td>&gt;10</td>
<td>Δn&gt;10</td>
<td>±7</td>
<td>-</td>
<td>6×10^4</td>
<td>[337]</td>
</tr>
<tr>
<td>GNR, ≈37 nm wide (1L)</td>
<td>Epitaxial BiFeO3 (100 nm)</td>
<td>PtNb-SrTiO3 (bottom)</td>
<td>≈5</td>
<td>-</td>
<td>±5</td>
<td>10^6</td>
<td>10^10</td>
<td>[336]</td>
</tr>
<tr>
<td>CVD graphene (1L)</td>
<td>PMN-PT substrate (~0.5 nm)</td>
<td>Au (bottom)</td>
<td>5.5</td>
<td>Δn=6.7</td>
<td>±300</td>
<td>-</td>
<td>-</td>
<td>[337]</td>
</tr>
<tr>
<td>MoS2 (1L)</td>
<td>P(VDF-TrFE) (200 nm)</td>
<td>Al (top)</td>
<td>5×10^3</td>
<td>ΔV=14</td>
<td>±20</td>
<td>10^3</td>
<td>-</td>
<td>[322]</td>
</tr>
<tr>
<td>MoS2 (3L)</td>
<td>P(VDF-TrFE) (~300 nm)</td>
<td>Al (top)</td>
<td>&gt;10^4</td>
<td>ΔV=25</td>
<td>±300</td>
<td>-</td>
<td>-</td>
<td>[322]</td>
</tr>
<tr>
<td>MoS2 (several layers)</td>
<td>P(VDF-TrFE) (~150 nm)</td>
<td>Pt (bottom)</td>
<td>8×10^5</td>
<td>ΔV=16</td>
<td>±26</td>
<td>3×10^4</td>
<td>10^3</td>
<td>[338]</td>
</tr>
<tr>
<td>MoS2 (few layers)</td>
<td>PZT (260 nm)</td>
<td>Pt (bottom)</td>
<td>10^10</td>
<td>ΔV=6</td>
<td>±8</td>
<td>10^4</td>
<td>100</td>
<td>[324]</td>
</tr>
<tr>
<td>MoS2 (1-3L)</td>
<td>PZT (100 nm)</td>
<td>TiO2/Ir (bottom)</td>
<td>10-10^4 variable due to traps</td>
<td>ΔV=4</td>
<td>±6</td>
<td>10^4</td>
<td>500</td>
<td>[317]</td>
</tr>
<tr>
<td>WS2 (3L)</td>
<td>Epitaxial PZT (500 nm)</td>
<td>SrRuO3 (bottom)</td>
<td>10^10</td>
<td>-</td>
<td>-4.3</td>
<td>2×10^4</td>
<td>400</td>
<td>[333]</td>
</tr>
<tr>
<td>MoSe2 (1L)</td>
<td>P(VDF-TrFE) (300 nm)</td>
<td>Al (top)</td>
<td>10^3</td>
<td>ΔV=30</td>
<td>±35</td>
<td>2×10^3</td>
<td>10^4</td>
<td>[88]</td>
</tr>
<tr>
<td>BP (~1.8 nm)</td>
<td>P(VDF-TrFE) (220 nm)</td>
<td>Al (top)</td>
<td>10^4</td>
<td>ΔV=14.6</td>
<td>±20</td>
<td>10^3</td>
<td>-</td>
<td>[316]</td>
</tr>
</tbody>
</table>
7.2. Phase change memories

Phase change memories (PCM) have been extensively investigated over the past decades because of their scalability to the nanometer regime and several semiconductor companies have transferred this technology into manufacturing. The principle of the PCM based on GeSbTe (GST) for NVM has been presented and described by Takaura et al.\cite{339} Briefly, when heated to a high temperature, the phase-change chalcogenides, such as GST, exhibit a reversible structural phase transition between amorphous and crystalline phases. This phase change modulates the resistivity of the material allowing electrical detection of the phase as memory state.\cite{84,340-341} The RESET operation (i.e., from crystalline to amorphous) is accomplished by applying large current amplitude with a short pulse width and a sharp falling edge as described by Wuttig and Yamada.\cite{342} A large current heats up the GST above the melting temperature and a quenching process freezes the softened, almost melted, GST in a disordered amorphous state. The opposite SET transition (i.e., from amorphous to crystalline), on the other hand, requires lower current amplitude with a longer pulse width, compared to what needed for the RESET operation.\cite{159} The time to complete the phase transition depends on the melting/crystallization temperature. Usually, a SET time on the order of 100 ns is achieved (limited by the crystallization process), which is longer than the conventional volatile memory devices like DRAM but faster than the storage devices like NAND (see Table 1 for comparison). The operating speed of PCM cells is primarily related to the time required to transform the GST into the crystalline phase and the time to quench the crystallized GST back into the glassy state. The reliability, on the other hand, is set by the repeatability of the meta-stable amorphous state and the changes in the integrated storage element components from cycle to cycle. This type of device also exhibits a thermal proximity disturb, effects or heat dissipation from nearby memory cells. In other words, a reset bit may be affected by the heat generated by adjacent bits under programming potentially leading to an accelerated crystallization in the reset bit.\cite{343} In the last two years, GRMs have also been investigated to form part of the PCM memory cell.\cite{56,62,344-346} In order to improve the energy-efficiency of the PCM Ahn et al.\cite{62} have used polycrystalline graphene as a thermal barrier
between the PCM and a tungsten bottom electrode heater so as to decrease the reset current as shown in Figure 12. Owing to its ultrathin aspect ratio and anisotropic thermal conductivity, graphene confined the heat within the active volume of the cell, enabling a 40% reduction in the RESET current.\cite{62} Recent investigations also focused on the phase-changing properties of TMDs, which can be transformed among different structural polytypes with distinct electronic properties.\cite{345-347} However, there is still more work to be done before graphene or TMDs are inserted into the conventional memory device flow.

7.3. Spintronic memories

Spintronics is a key enabling technology for “beyond-CMOS” era, in which information is carried by spin instead of charge, and which allows not only efficient data storage but also possible new paradigms in spin manipulation and active spin logic device and architecture design. Beyond existing needs of efficient memory storage technologies in the field of personal computers, powerful workstations or computing centers, the world of Internet of Things (IoTs) also relies on highly demanding node memories to process data among sensors, cloud and RF front-end. Both mainstream and emerging memories are intensively investigated for achieving highly energy efficiency. In that context, the development of spin transfer torque magnetic tunnel junction (STT-MTJ)-based non-volatile memory has demonstrated great performance in terms of zero standby power, switching power efficiency, infinite endurance and high density, but high dynamic write energy, large latency, yield and reliability need to be improved for future market impact. The development of voltage-controlled magnetic anisotropy has appeared as a promising solution for reaching for high-performance and ultra-low power consumption targets of the IoT node memory\cite{348} so that any solution to improve and control the magnetic anisotropy would bring a clear competitive advantage. Two-dimensional materials-based spintronic devices are expected to address some of the main challenges in spintronic technology (such as ultralow power consumption and versatile integration in I technologies) but could also enable new applications in the field of spin-based computing
technologies, especially thanks to the wealth of opportunities offered through harvesting proximity effects between 2D materials. In fact, despite the weak inter-layer interactions in van der Waals heterostructures, a plethora of new properties can be imprinted to graphene by proximity effects\cite{341}. Recent advances on magnetic proximity effects have been made using both low and high Curie temperature materials such as EuO, EuS and yttrium iron garnet or Y$_3$Fe$_5$O$_{12}$ (YIG)\cite{349,351}, or strong spin-orbit coupling materials\cite{352-353} which evidence the richness and diversification of opportunities.

In 2017, spin transport in graphene was found to be strongly influenced by proximity effects in heterostructures of graphene and TMDCs, such as molybdenum disulphide and tungsten disulphide. Through inter-layer coupling, the behaviors of the parallel and perpendicular spin orientations in graphene is dramatically altered, leading to an anisotropic spin relaxation from one to several orders of magnitude, opening the route to designing spin filters and spin switches\cite{354-355}

After one decade of intense research efforts to achieve long spin lifetimes in graphene\cite{356-357}, the potential of GRMs for spintronic applications is now attracting the attention of large companies in the sectors of advanced memories and future spin logic (such as INTEL, SAMSUNG, GLOBAL FOUNDRY and TSMC). Interestingly, first proofs of principle of graphene-based spin logic devices including spin-based field-effect transistors\cite{358} or the design of XOR logic functions using magnetologic gates have already been reported, paving the way towards all spin-based information processing circuits. MRAM and STT-MRAM are reaching a level of maturity that makes them leading prospects in the market of emerging memories\cite{359}. These approaches currently use conventional insulating materials within the tunnel barriers, such as MgO. Unfortunately, further reduction of the cross-sectional area of the junction, as density increases, is expected to lead to prohibitively high junction resistances largely exceeding than 100 kΩ, which will result in too large energy consumption. The industry already envisions the replacement of the insulating materials, or its complete elimination, using standard giant magnetoresistance structures. The latter, however, leads to a very significant reduction of the magnetoresistance, and therefore graphene and/or other 2D materials could help overcome the present and future challenges.
In fact, graphene and h-BN have proven to be useful materials to unlock low-cost processes such as ALD for the fabrication of large-scale spintronic devices, which represents an essential step towards the realization of practical novel memory concepts. Graphene-based materials (including multilayers and graphite) as well as h-BN and their heterostructures, have definitely a strong potential for spintronic devices based on vertical geometry. It has been predicted\cite{360} and demonstrated\cite{361} that graphene/ferromagnets interfaces can efficiently filter spin channels. This could significantly improve interfacial spin polarization giving rise to high tunnel magneto-resistance values, similarly to the case of epitaxial MgO\cite{362}. Experimental demonstration of this effect has been obtained using CMOS-compatible processes ($<450^\circ$C large scale CVD processes) for the direct integration of graphene in devices\cite{363}. Already these novel electrodes have been shown to be oxidation-resistant and allowed to unlock low-cost processes for the fabrication of MTJ such as ALD\cite{364}. A recent report has clarified the huge impact of h-BN as a tunnel barrier between ferromagnetic materials (Co/hBN/Fe interfaces), with tunneling magneto-resistance reaching values above 50\% at room temperature\cite{365} and resulting from a change in the nature of hybridization from physiorption to chemisorption depending on the interface magnetic material.

An important next step is to explore the behavior of non-collinear spintronic phenomena such as spin transfer torque, important for the development of STT-MRAM\cite{366} and spin torque nano-oscillators\cite{367}, which can play a significant role in next generation high speed communications. STT-MRAM are mainly based on MTJs (Figure 13a), and have two operational functions\cite{22, 161, 182}. The write operation uses the spin-transfer torque mechanism, which is to pilot the magnetization reversal by driving a spin polarized (or pure spin current) through the junction, while the Read operation is based on the tunneling magneto-resistance across the junction. One direction of development is to use graphene coating to enhance the efficiency of the writing process, by increasing the perpendicular magneto-crystalline anisotropy (PMA), thus reducing the driving current intensity (lower energy consumption). The other direction of work is to harvest the spin Hall Effect mechanism to create a pure spin current perpendicular to the charge flow, which will act on the magnetization of the ferromagnetic material\cite{159, 161} (illustrated in Figure 13b). One denotes such memory as spin-orbit
torque (SOT-MRAM). This could further downsize the energy consumption cost during magnetization switching operations. In this perspective, the understanding and optimization of upper spin Hall angle (SHA) values achievable with two-dimensional materials is of great concern and is currently under fierce debate, given promising but contradictory experimental results.\textsuperscript{[368-369]} The optimized use of TMDC as a substrate for graphene could achieve large enough signal and SHA to become technology relevant as discussed by Torres et al.\textsuperscript{[370]} and Garcia et al.\textsuperscript{[371]}. The improvement of magnetic properties of FM/graphene interfaces such as PMA is a key direction to pursue for the downscaling of spintronic devices to improve energy efficiency, and this has sparked large efforts in both theoretical and experimental investigation. Recently, Yang and co-workers reported that the coating of Co films with graphene could result in strong enhancement of the PMA of the magnetic materials.\textsuperscript{[372]} The ab-initio simulations show that the surface anisotropy of Co films can double the value of its pristine counterpart by graphene coating, with a possible extension of the out-of-plane effective anisotropy up to a thickness of 25 Å.\textsuperscript{[372]} Such results are fully consistent with experimental data, using graphene coating of Co films which are grown on iridium substrate. Besides, a theoretical scenario of superexchange stabilized Co–graphene heterostructures with a robust constant effective PMA and linearly increasing interfacial anisotropy as a function of film thickness was proposed, pointing to possible engineered graphene/ferromagnetic metal heterostructures with giant PMA more than 20-times larger compared to conventional multilayers, a true hallmark for future graphene and traditional spintronic technologies. In 2013, a first experimental measurement of spin transfer torque in graphene lateral nonlocal spin valve devices was reported.\textsuperscript{[373]} The experiment consists of activating an input magnet from which pure spin currents are driven to a receiving magnet, of which the magnetization is controllable by the input currents, while the tunneling magnetoresistance is the relative change of the resistance of the junction when magnets are polarized parallel or antiparallel to each other’s. In reference \textsuperscript{[373]}, such phenomenon was assisted by an external magnetic field, and the magnetization switching is reversible between parallel and antiparallel configurations, depending on the polarity of the applied charged current. The presented results stand as an important step forward towards the development of graphene-based spin logic and engineering
of spin-transfer torque technologies, but more practical realizations are needed, especially in a fab environment. On the theoretical side, current-induced spin polarization of weakly magnetized graphene with Rashba spin-orbit interaction has been investigated\cite{374}. The authors found that for such material all components of the current-induced spin polarization are nonzero, which contrasts to the nonmagnetic case, where the only non-vanishing component of spin polarization is the one in the graphene plane and normal to the electric field. Kawakami\cite{357} recently proposed a new concept for spin amplification for monitoring a ferromagnet attached to a non-magnetic spin channel, with small spin currents.\cite{357} The proposed idea is to drive a ferromagnet into an unstable symmetric state, in which a weak spin current will be enough to drive the magnetization direction of the ferromagnet, by spin transfer torque.

Finally, it is worth noting that thanks to the development of large scale CVD processes, other 2D materials beyond graphene have to be investigated for magnetic tunnel junctions.\cite{365} In graphene-based magnetic tunnel junctions, the contact resistance can be tailored to desired regime by introducing heterostructures with h-BN tunnel barriers without compromising spin polarization, which will be useful for high-density MRAM.\cite{375} However, more remain to be done, and in particular the exploration of STT-related phenomena for lateral and more complex geometries using hybrid 2D materials-based heterostructures. We note that the use of such 2D materials heterostructures in magnetic tunnel junctions could also provide flexible spintronic devices. It is finally worth to mention that IMEC in Belgium has very recently reported the first full-scale integration of top-pinned perpendicular MTJ on 300 mm wafer using CMOS-compatible processes for SOT-MRAM architectures.\cite{376} Ultrathin SOT layers were fabricated to operate at low power (about 300pJ), sub-ns switching (210 ps) and with excellent endurance (> 5x1010) and all in the absence of electromigration effect. A successful integration of 2D materials in such technology will therefore potentially further improve the data storage performance of an emerging technology on its way to the market.
7.4. Tunneling memories

Finally, a new type of memory cell, named Tunneling Random Access Memory (TRAM), was revealed last year by the group of Y. H. Lee,[57, 377] who assembled a van der Waals (vdW) heterostructures of multiple GMRs, namely graphene, h-BN and MoS$_2$, see Figure 14 a, b, in a device architecture similar to that developed by Choi et al. (ref. [69], Section 6). However, the device was operated using only two terminals, and the program/erase operations were achieved by means of drain-field induced tunneling of charge carriers between the MoS$_2$ channel and the graphene FG. The two-terminal geometry might enable further scaling of charge-based memory devices, for instance through the implementation of cross-point arrays, as in the case of ReRAMs and memristors.[377] Noticeably, the TRAM cells have been integrated on stretchable substrates suitable for wearable electronics (Figure 14 c-e).[57, 377]

8. GRM-based NVMs: challenges and outlook

Today’s portable electronics, such as smart phone and smart watch, have been transforming the paradigm of electronics industry from performance-oriented electronics to human-friendly electronics. Recent advances of artificial intelligence (AI) and Internet of Things (IoT) technologies have accelerating the paradigm shift with breakthroughs in novel high-tech products performing intelligent tasks such as real-time big data analytics, self-driving automobile navigation, and smart home appliances. In particular, flexible and wearable electronics will be an ideal platform to provide handily the powerful services with users, because they provide human-friendly interfaces, outstanding portability, and convenience. To efficiently process the growing flood of data arising from these technologies, it is inevitable to develop flexible, low-power nonvolatile memory with high-density due to not only its fundamental roles in electronic system such as data storage, processing, and communication with external device. A number of research groups have explored a variety of organic thin film-based (flexible) memories[378-382] due to their well-known inherent flexibility and cost-effective synthesis over large area. However, there are still big hurdles for
developing high-performance flexible memory with high-density, including how to improve insufficient performance stemming from inherent material properties and noncompatibility with photolithography process.

In this review, we have provided an overview on the use of graphene and related 2D materials (GRMs) in different types of non-volatile memory (NVM) cells. In particular, we focused our attention on the physical and chemical mechanisms underlying the non-volatile switching of GRM-based NVM devices, addressing the key material properties and device structures. We emphasized both opportunities and challenges towards the realization of practical NVM devices that exploit the unique properties of GRMs. Based on the results and analysis at the laboratory scale over the past decade, two types of technologies are envisioned for GRMs: (i) low-cost flexible/transparent information-storage devices to be integrated in wearable systems or smart objects, and (ii) high-speed and high-capacity NVMs. The former has less stringent materials requirements, are easier to develop and are expected to enter the market of portable/wearable electronics sooner. In this context, the availability of large amounts of high-quality GRMs produced either via solution processing or CVD-growth techniques is of paramount importance. On the other hand, developing high-performing memory technologies based on GRMs, requires solving integration issues associated with the introduction of 2D materials in existing conventional device flows. Currently, the most challenging aspects of 2D materials integration are direct growth on the surface of traditional dielectrics in a conventional Si flow, as well as transfer processes compatible with the Si flow while keeping the properties of the 2D materials unchanged. The challenges associated to the growth of graphene, h-BN or TMDs are related to thermal budget, composition control, and epitaxial growth compatible with the memory flow. For example, it is possible to grow graphene on a dielectric surface by diffusing C through a metal but controlling the thickness or the grain size has been difficult to date. The same is true for h-BN; it is possible to grow it under a metal but grain size, orientation, and surface roughness are difficult to achieve and thus it is important to identify and understand the material requirements for the final NVM cell. Growth under a metal for both h-BN and graphene could simplify the fabrication process of NVM cells. In the case of TMDs, it is likely that it is not necessary to grow large area single crystal;
if this is so then thermal budget, selectivity, and composition control remain to be addressed. While there has not been much development yet in selective growth processes, this process could simplify the integration of TMDs in the Si flow for both single films as well as heterostructures. In addition, however, it is extremely important to ensure the stability of TMDs in contact with dielectrics and metals when exposed to a Si thermal budget. In the case of graphene in PCM devices, there is a clear benefit and the requirements are not as stringent as for active components. For other technologies, unless there is a significant driver for the use of these new materials, it is unlikely that they can be integrated in state of art silicon fab within the next few years. However, it should be noted that game-changing technologies, such as chalcogenide-based PCMs, as well as many new materials currently in use in integrated circuits required one to two decades of intense research and development activities by both industry and academia before finally entering the market place. In the long term, the advantages of 2D materials cannot be underestimated. In the case of GRMs-based NVMs we are still in the early stages of development, and in order to take full advantage of the outstanding properties of 2DM collaboration between industrial and academic laboratories is necessary to ensure the timely introduction of these new materials in future products.

Acknowledgements

We acknowledge funding from the European Commission through the Graphene Flagship (GA-696656) and the Marie-Curie IEF MULTI2DSWITCH (GA-700802), the MAECI (Ministero degli Affari Esteri e della Cooperazione Internazionale) for the bilateral Italy–China GINSENG projects, as well as the Agence Nationale de la Recherche through the Labex projects CSC (ANR-10-LABX-0026 CSC) and NIE (ANR-11-LABX-0058 NIE) within the Investissement d’Avenir program (ANR-10-120 IDEX-0002-02), and the International Center for Frontier Research in Chemistry (icFRC).

Abbreviations
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>Two-dimensional</td>
</tr>
<tr>
<td>ALD</td>
<td>Atomic layer deposition</td>
</tr>
<tr>
<td>BiCS</td>
<td>Bit-cost scalable</td>
</tr>
<tr>
<td>CF</td>
<td>Conducting filaments</td>
</tr>
<tr>
<td>CVD</td>
<td>Chemical vapor deposition</td>
</tr>
<tr>
<td>CrGO</td>
<td>Chemoselectively reduced graphene oxide</td>
</tr>
<tr>
<td>CSL</td>
<td>Charge storage layer</td>
</tr>
<tr>
<td>DRAM</td>
<td>Dynamic random access memory</td>
</tr>
<tr>
<td>EOT</td>
<td>Equivalent oxide thickness</td>
</tr>
<tr>
<td>FeRAM</td>
<td>Ferroelectric random access memory</td>
</tr>
<tr>
<td>FET</td>
<td>Field-effect transistor</td>
</tr>
<tr>
<td>FG</td>
<td>Floating gate</td>
</tr>
<tr>
<td>FG-FET</td>
<td>Floating-gate field-effect transistor</td>
</tr>
<tr>
<td>FTO</td>
<td>Fluorine-doped tin oxide</td>
</tr>
<tr>
<td>GB</td>
<td>Grain boundaries</td>
</tr>
<tr>
<td>GNR</td>
<td>Graphene nanoribbon</td>
</tr>
<tr>
<td>GO</td>
<td>Graphene oxide</td>
</tr>
<tr>
<td>GRM</td>
<td>Graphene and related materials</td>
</tr>
<tr>
<td>h-BN</td>
<td>Hexagonal boron nitride</td>
</tr>
<tr>
<td>ITRS</td>
<td>International Technology Roadmap for Semiconductors</td>
</tr>
<tr>
<td>MBDT</td>
<td>4-mercaptobenzenediazonium tetrafluoroborate</td>
</tr>
<tr>
<td>MLG</td>
<td>Multilayer graphene</td>
</tr>
<tr>
<td>MRAM</td>
<td>Magnetic random access memory</td>
</tr>
<tr>
<td>MIM</td>
<td>Metal insulator metal</td>
</tr>
<tr>
<td>MW</td>
<td>Memory window</td>
</tr>
<tr>
<td>NP</td>
<td>Nanoparticle</td>
</tr>
<tr>
<td>NMP</td>
<td>N-methyl-2-pyrrolidone</td>
</tr>
<tr>
<td>NVM</td>
<td>Non-volatile Memory</td>
</tr>
<tr>
<td>NW</td>
<td>Nanowire</td>
</tr>
<tr>
<td>P123</td>
<td>PEO_{20}PPO_{70}PEO_{20}</td>
</tr>
<tr>
<td>PCM</td>
<td>Phase change memory</td>
</tr>
<tr>
<td>PEN</td>
<td>Polyethylene naphthalate</td>
</tr>
<tr>
<td>PES</td>
<td>Polyethersulfone</td>
</tr>
<tr>
<td>PMMA</td>
<td>Polymethyl methacrylate</td>
</tr>
</tbody>
</table>
| PMN-PT       | \((1-x)[Pb(Mg_{1/3}Nb_{2/3})O_3]-x[PbTiO_3]_0.3\)
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PV3D3</td>
<td>Poly(1,3,5-trimethyl-1,3,5-trivinyl cyclotrisiloxane)</td>
</tr>
<tr>
<td>PVA</td>
<td>Polyvinyl alcohol</td>
</tr>
<tr>
<td>PVP</td>
<td>Polyvinylpyrrolidone</td>
</tr>
<tr>
<td>PVPh</td>
<td>Polyvinylphenol</td>
</tr>
<tr>
<td>P(VDF-TrFE)</td>
<td>Poly(vinylidene fluoride-trifluoroethylene)</td>
</tr>
<tr>
<td>PVK</td>
<td>poly(N-vinylcarbazole)</td>
</tr>
<tr>
<td>PZT</td>
<td>Pb(Zr0.5Ti0.5)O3 (PZT)</td>
</tr>
<tr>
<td>QD</td>
<td>Quantum dot</td>
</tr>
<tr>
<td>RAM</td>
<td>Random access memory</td>
</tr>
<tr>
<td>ReRAM</td>
<td>Resistive random access memory</td>
</tr>
<tr>
<td>SCM</td>
<td>Storage class memory</td>
</tr>
<tr>
<td>SLG</td>
<td>Single layer graphene</td>
</tr>
<tr>
<td>STT</td>
<td>Spin transfer torque</td>
</tr>
<tr>
<td>STT-MRAM</td>
<td>Spin transfer torque magnetic random access memory</td>
</tr>
<tr>
<td>TMD</td>
<td>Transition metal dichalcogenide</td>
</tr>
<tr>
<td>TMP</td>
<td>2,2,6,6-tetramethyl-4-piperidinol</td>
</tr>
<tr>
<td>TPAPAM</td>
<td>triphenylamine-based polyazomethine</td>
</tr>
<tr>
<td>TRAM</td>
<td>Tunneling Random Access Memory</td>
</tr>
<tr>
<td>WORM</td>
<td>Write only read many</td>
</tr>
<tr>
<td>YIG</td>
<td>Yttrium Iron Garnet</td>
</tr>
<tr>
<td>ZIF-8</td>
<td>Zeolitic imidazolate frameworks</td>
</tr>
</tbody>
</table>
References


61


**Figures**

**Figure 1.** (a) Diagram of established and emerging/prototypical memory technologies (non-exhaustive list). A more detailed taxonomy can be found in semiconductor industry roadmaps (e.g. ref. [158]). (b-e) Examples of proof-of-concept NVM cells that incorporate 2D materials: flash memory based on graphene/MoS$_2$ heterostructures (b), ferroelectric transistor with graphene channel (c), two-terminal resistive memory cell based on GO and MoS$_2$ nanosheets (d), and TRAM cell based on MoS$_2$/h-BN/graphene heterostructures. (b) Reproduced with permission.[61] Copyright 2013, American Chemical Society. (c) Reproduced with permission.[53] Copyright 2009, AIP Publishing. (d) Reproduced with permission.[253] Copyright 2016, IOP Publishing. (e) Copyright 2016, Nature Publishing Group.
Figure 2. Schematic illustration of the GRM production techniques and the respective materials that they can used to prepare and memory cells they can be used to fabricate: (a) chemical vapour deposition (CVD), (b) nucleation and growth by chemical vapour deposition, (c) precipitation in metals, (d) liquid phase exfoliation (LPE), and (e) electrochemical exfoliation. Each one of these techniques has already achieved a level of maturity where the processes have been adopted in pilot production lines.

TMDs = transition metal dichalcogenides; MLG = multi-layer graphene; SLG = single layer graphene; FLG = few-layer graphene; FeRAM = Ferroelectric random access memory; MRAM = Magnetic random access memory; STT-RAM = Spin transfer torque random access memory; TRAM = Tunneling Random Access Memory; PCM = Phase change memory; ReRam = Resistive random access memory.
Figure 3. Memory switches based on graphene/graphitic nanosheets. (a) Tilted-view SEM image of a few-layer graphene memory device after formation of the nanogap with a voltage $V > V_{\text{break}}$. Adapted with permission.[200] Copyright 2009, American Chemical Society. (b) Schematics of the device operating in the ON (top) and OFF (bottom) states. (c) Time-dependent measurement of the current, $I$, flowing through the gapped nanosheet upon application of a voltage ramp. Adapted with permission.[52] Copyright 2008, American Chemical Society. (d) SEM image of a suspended-graphene device after formation of a nanogap (OFF state). (e) Time dependent measurements of current $I$ and conductance $G$ upon application of voltage pulses. Reproduced with permission.[205] Copyright 2012, American Chemical Society.
Figure 4. Resistive NVM devices based on GO nanosheets. (a) A schematic illustration of a GO based flexible crossbar memory device. (b) Typical $I-V$ curve of a Al/GO/Al/PES device plotted on a semilogarithmic scale. (c) Schematic of the proposed bipolar resistive switching (BRS) model for Al/GO/Al electrode crossbar memory device. Top: the pristine device is in the OFF state due to the (relatively) thick insulating top interface layer formed by a redox reaction between Al and the GO film. Bottom: the ON state is induced by the formation of local filaments in the top interface layer due to oxygen ion diffusion back into the GO film by an external negative bias on the top electrode. Adapted with permission.[76] Copyright 2010, American Chemical Society.
**Figure 5.** Resistive NVM devices based on (functionalized-) RGO and GO. (a) Schematic illustration of the chemoselective reduction of GO via the ultraviolet irradiation catalyzed with 2,2,6,6-tetramethyl-4-piperidinol (TMP). (b) Current-voltage ($I$-$V$) characteristics of Au/RGO/ITO memory cells. Inset: schematic illustration of the device structure. (c) $I$-$V$ characteristics of a memory device based on RGO functionalized with Au NPs. The active material is sandwiched between the ITO substrate and the Al top electrodes (inset). (d) Schematic drawing of the GO layer functionalized with Au NPs by making use of a bifunctional molecular linker (MBDT salt). The functionalized GO was employed also in ReRAM devices with horizontal FET structure. (e) Schematics of the functionalization of GO with the conjugated polymer TPAPAM. (f) Typical $I$-$V$ curves of a memory device based on a TPAPAM-GO active layer sandwiched between ITO and Al electrodes (structure in the inset). (a, b) Adapted with permission.\cite{232} Copyright 2012, American Chemical Society. (c, d) Adapted with permission.\cite{216} Copyright 2011, American Chemical Society. (e, f) Adapted with permission.\cite{210} Copyright 2010, John Wiley & Sons.
Figure 6. Resistive NVMs based on solution-processed MoS$_2$ nanosheets. (a) Illustration of the polymer-assisted liquid-phase exfoliation of MoS$_2$ via ultrasonication. (b) Current-voltage ($I$-$V$) characteristics of the flexible ReRAM fabricated in ref. [249]. Inset: schematics of the RGO/MoS$_2$-PVP/Al device. (c) Multilevel $I$-$V$ curves of memory cells based on multicomponent GO-MoS$_2$-GO films, which were obtained by making use of different RESET voltages. (d) Device structure (left) and band diagram (right) of the GO-MoS$_2$-GO heterostructure. (e) Left: photograph of a cross-point memristor array on PEN (scale bar, 10 mm). Right: cross-section SEM image of the device based on MoO$_x$/MoS$_2$ heterostructures (scale bar, 100 nm). The inset at the top shows the schematic structure of the memory cell. (f) $I$-$V$ characteristics of the ReRAM device shown in panel e. (a, b) Reproduced with permission. [249] Copyright 2013, John Wiley & Sons. (c, d) Reproduced with permission. [253] Copyright 2016, IOP Publishing. (e, f) Reproduced with permission. [78] Copyright 2015, Nature Publishing Group.
Figure 7. Resistive memory cells based on CVD-grown MoS$_2$ and $h$-BN. (a) Schematic illustration of an MoS$_2$ memristor where two GBs, intersecting each other at a vertex within the channel, are connected to one of the electrodes.$^{[265]}$ (b) Partial $I-V$ characteristics of an intersecting-GB memristor. The resistance switch occurs at a voltage $V_{\text{SET}} \approx 8.3$ V. Inset: Full $I-V$ characteristics. (c) Atomic force microscopy (AFM) phase image of an MoS$_2$ flake with multiple GBs. The dashed white lines indicate the location of the electrode edges. Color scale bars show the phase angles in degrees. (d) Schematic drawing of the ReRAM device based on CVD-grown $h$-BN (ref. $^{[273]}$) and (e) corresponding $I-V$ characteristics after an electroforming process. (f) Top: high-resolution TEM image of incomplete Ag conducting filament (highlighted by the closed region) formed within a $h$-BN ReRAM. Bottom: schematic illustration of the filament growth: the $\text{Ag}^+$ ions are reduced inside $h$-BN by capturing free electrons. (a-c) Reproduced with permission.$^{[265]}$ Copyright 2015, Nature Publishing Group. (d-f) Adapted with permission.$^{[273]}$ Copyright 2016, John Wiley & Sons.
**Figure 8.** Scaling challenges in silicon-based flash memories. (a) Schematic drawing of three adjacent FG-FETs within a typical bit-line array. The parasitic capacitive coupling between neighboring FGs gives rise to detrimental cell-to-cell interferences. (b) Band diagram of the gate stack in a typical silicon FG-FET, showing the occurrence of ballistic leakage currents for the channel to the control gate.
Figure 9. Flash memory cells based on GRMs. (a) Schematic illustration of the flash memory device based on MLG (or SLG) charge-trapping layer. The silicon substrate acts as bottom electrode, SiO$_2$ as tunneling oxide and Al$_2$O$_3$ as control oxide. The top electrode consists in a metal stack of Ti/Al/Au. (b) HR-TEM cross sectional image of the MLG device. (c) Typical capacitance-voltage ($C$-$V$) measurements on MLG flash memories, showing a memory window of $\approx$ 6 V. (d, f) Schematic drawings of the memory devices based on heterostructures of graphene and MoS$_2$. In (d), HfO$_2$ serves as tunnel/control oxide, while monolayer MoS$_2$ and MLG act as transistor channel and FG, respectively (ref. $^{[61]}$). In (f), h-BN serves as tunneling layer, graphene as transistor channel and MoS$_2$ as charge-trapping layer (ref. $^{[69]}$). (e, g) Temporal evolution of the drain-source current in the erase (release) and program (trap) states for the device structures shown in panel (d) and (e), respectively. (a-c) Adapted with permission.$^{[292]}$ Copyright 2011, American Chemical Society. (d, e) Reproduced with permission.$^{[61]}$ Copyright 2013, American Chemical Society. (f, g) Reproduced with permission.$^{[69]}$ Copyright 2013, Nature Publishing Group.
Figure 10. Ferroelectric memory devices based on GRMs. (a) Electric displacement vs top-gate voltage ($D$ vs $V_{TG}$) characteristics of a ferroelectric memory based on graphene/P(VDF-TrFE). Inset a: schematics of the device and electric displacement continuity equation at ferroelectric/graphene interface. Inset b: schematic drawing of a polarized P(VDF-TrFE) molecule. (b) Resistance $R$ as a function of $V_{TG}$. The plot shows the electric hysteresis loop of the graphene-ferroelectric FET. (c) Schematic of a ferroelectric memory cell built on a SiO$_2$/Si substrate. The cell includes a semiconducting MoS$_2$ channel, a P(VDF-TrFE) insulating polymer and an Al top gate. (d) Typical drain current vs top-gate voltage curves of the ferroelectric device schematized in panel c. (e) Cross-sectional schematic views of the BP ferroelectric FET (ref. $^{[316]}$), showing the mechanisms for programming/erasing: the p-type channel is turned ON after a $-20$ V pulse (left) and turned OFF after $+20$ V pulse (right). (f) Black line: drain current $I_D$ vs top-gate voltage $V_G$ hysteresis loop of a BP FeFET. Red and blue lines: $I_D$ vs $V_G$ sweeps acquired in the small range $-0.5 \leq V_G \leq 0.5$ V after programming (red) and erasing (blue) pulses. (a, b). Reproduced with permission.$^{[53]}$ Copyright 2009, AIP Publishing. (c, d) Reproduced with permission.$^{[329]}$ Copyright 2012, John Wiley & Sons. (e, f) Adapted with permission.$^{[316]}$ Copyright 2015, American Chemical Society.
Figure 11. FeRAMs based on MoSe$_2$ nanosheets. (a) Optical images of a single-layer MoSe$_2$ transistor before coating with P(VDF-TrFE) and after device fabrication (inset). Scale bar, 20 µm. (b) Schematic diagram of the device configuration. (c) Transfer curves at different source/drain voltages. (d) Data of write and erase speed tests at different voltage pulse widths and with a pulse height of 40 V. (a-d) Adapted with permission.$^{[85]}$ Copyright 2017, IOP Publishing.
Figure 12. Phase-change memories with graphene confinement layer. (a) Cross-sectional HR-TEM image of the PCM device incorporating a graphene thermal barrier. (b) Reduction of the RESET current $I_{\text{RESET}}$ in patterned graphene PCMs (red, G-PCM) compared to devices without graphene (black, PCM) and with large unpatterned graphene (blue, control sample). (c) Schematic representation of the graphene-based PCM. The inset shows two different types of devices, i.e. control sample and G-PCM, fabricated with different graphene lateral sizes $D_G$. (d) Cross-sectional TEM image of a PCM device, where graphene is embedded between the bottom (10 nm thick) and top (30 nm thick) GST layers. (e) Zoom-in image corresponding to the rectangle area in (a), showing the amorphous and crystalline GST layers on both sides of graphene. (f) Simulation of the RESET temperature distributions in PCM cells with (left) and without (right) graphene. The initial temperature is 300 K. (a-c) Reproduced with permission.[62] Copyright 2015, American Chemical Society. (d-f) Reproduced with permission.[344] Copyright 2016, AIP Publishing.
Figure 13. MRAMs based on spin-transfer and spin-orbit torque effects. (a) Schematic of the basic STT-MRAM unit showing the local magnetic element, whose state is controlled by a spin-polarized charge current flow. (b) Typical junction between a strong SOC-material giving rise to the spin Hall effect and a ferromagnet on which a pure spin current act to modulate/switch the magnetization (so-called spin-orbit torque effect).
Figure 14. Two-terminal floating-gate memory. (a) Schematic of the TRAM cell based on MoS$_2$/h-BN/graphene heterostructures. (b) Typical current-voltage ($I_{ds}$ vs $V_{ds}$) characteristics for a device incorporating a 5.5-nm thick $h$-BN layer. The four stages are: (i) Programme, (ii) Read, (iii) Erase and (iv) Read. (c, d) Electrical measurements of flexible TRAMs upon stretching: $I_{ds}$ vs $V_{ds}$ curves acquired for different values of strain (c), and strain-dependent ON and OFF currents at fixed $V_{ds} = 1$V. The maximum strain before memory failure is 20%. (e) Photograph of the flexible TRAM device on a PET substrate. (a-e) Reproduced with permission.[57] Copyright 2016, Nature Publishing Group.