Frank Schwierz, Jörg Pezoldt and Ralf Granzner

Two-dimensional materials and their prospects in transistor electronics

Original published in:

Original published: 12 April 2015
ISSN (online): 2040-3372
ISSN (print): 2040-3364
DOI: 10.1039/C5NR01052G
URL: http://dx.doi.org/10.1039/C5NR01052G
[Visited: 2015-07-17]

This article is licensed under a Creative Commons Attribution 3.0 Unported Licence. [ http://creativecommons.org/licenses/by/3.0/ ]
Two-dimensional materials and their prospects in transistor electronics

F. Schwierz,* J. Pezoldt and R. Granzner

During the past decade, two-dimensional materials have attracted incredible interest from the electronic device community. The first two-dimensional material studied in detail was graphene, and, since 2007, it has intensively been explored as a material for electronic devices, in particular, transistors. While graphene transistors are still on the agenda, researchers have extended their work to two-dimensional materials beyond graphene and the number of two-dimensional materials under examination has literally exploded recently. Meanwhile several hundreds of different two-dimensional materials are known, a substantial part of them is considered useful for transistors, and experimental transistors with channels of different two-dimensional materials have been demonstrated. In spite of the rapid progress in the field, the prospects of two-dimensional transistors still remain vague and optimistic opinions face rather reserved assessments. The intention of the present paper is to shed more light on the merits and drawbacks of two-dimensional materials for transistor electronics and to add a few more facets to the ongoing discussion on the prospects of two-dimensional transistors. To this end, we compose a wish list of properties for a good transistor channel material and examine to what extent the two-dimensional materials fulfill the criteria of the list. The state-of-the-art two-dimensional transistors are reviewed and a balanced view of both the pros and cons of these devices is provided.

1. Introduction

The FET (field-effect transistor) is the backbone of today’s semiconductor electronics. It represents the basic building block of the systems of modern information and communication technology and progress in this important field critically depends on rapid improvements of FET performance. An efficient option to achieve this goal is the introduction of novel channel materials into FET technology. In this regard, 2D (two-dimensional) materials have attracted considerable attention from the transistor community. The rise of 2D materials began in 2004 with the successful preparation of graphene samples.\(^1,2\) Particularly the observed high carrier mobilities raised expectations that graphene could be the perfect channel material for FETs and will become the successor of conventional semiconductors. After the enthusiastic early days of graphene transistor research it became clear, however, that graphene would not be able to fulfill these high expectations since it does not possess a bandgap which is mandatorily needed for proper FET operation.\(^3\)

Just when the interest of the transistor community in graphene began to subside,\(^4,5\) a paper on the fabrication of single-layer MoS\(_2\) FETs\(^6\) gave new momentum to the research on 2D materials. Over a surprisingly short period of time, entire classes of new 2D materials have been discovered. A steadily increasing number of groups worldwide are now working intensively on 2D FETs, chipmakers pay attention to the progress in the field, and since 2011 the ITRS (International Technology Roadmap for Semiconductors), the strategic planning document of the semiconductor industry, has mentioned 2D materials beyond graphene as candidates for future electronics.\(^7\) Nevertheless, the prospects of the 2D materials in electronics are controversially debated. While part of the community is very optimistic, another fraction believes that 2D transistors are of academic interest only and are useful for niche applications at best.

The strong interest of the community in 2D transistors is manifested in the remarkable number of recent review papers on this topic, each with a specific focus and flavor. Particular mention should be made of the thorough overviews on 2D materials by Butler et al.\(^8\) and on 2D transistors by Lemme et al.\(^9\) and of the survey on the status of TMD (transition metal dichalcogenide) as a prominent group of 2D materials beyond graphene.\(^10\) The reader will also appreciate the excellent discussions on 2D transistors for digital logic by Fiori et al.\(^11\) and for flexible electronics by Akinwande et al.\(^12\)

The present paper focuses on the properties and physics of 2D materials and 2D FETs and is organized as follows. Section
2 reviews important trends in semiconductor electronics and introduces the figures of merit we later use to assess the performance of transistors. Section 3 presents an overview of the different classes of 2D materials. In section 4 we compose a wish list of properties desirable for FET channel materials and examine to what extent the 2D materials meet the requirements of our list. To keep the discussion manageable, neither the entries of the wish list nor the set of transistor figures of merit are intended to be exhaustive. Instead, we limit the discussion to a selected set of material parameters as entries for the wish list and a few key transistor figures of merit. Section 5 provides a discussion of the current status, prospects, and problems of 2D transistors, and finally section 6 concludes the paper.


2.1 More Moore trends

The overall semiconductor chip market has an annual volume of about $300 billion and can, as shown in Fig. 1, broadly be divided into the two main domains More Moore and More Than Moore.

More Moore encompasses digital ICs (integrated circuits) such as microprocessors and memories and covers around 70% of the overall chip market. Interestingly, today’s digital ICs are based on single semiconducting material, Si, one single transistor type, the Si MOSFET, and one single circuit technique, Si CMOS (complementary MOS) employing both n-channel (electrons constitute the transistor’s output current) and p-channel MOSFETs (holes carry the current). For decades, chipmakers have introduced CMOS ICs with an exponentially increasing number of individual MOSFETs and, simultaneously, an exponentially decreasing price per transistor, to the market. Key to this evolution is the continuous shrinking of the MOSFET size, for short scaling. As shown in Fig. 2, since 2014, processors containing five billion MOSFETs with gate lengths around 20 nm have been commercially available and the introduction of a 10-billion processor has been announced for 2015.13 This trend of continuous device miniaturization and increasing circuit complexity is called Moore’s Law.14

It is important to recognize that, to follow Moore’s Law, increasing circuit complexity by scaling alone is not sufficient. Instead, the electrical performance of the scaled MOSFETs must be improved as well. In particular, transistor switching speed should increase exponentially and the energy needed for a single switching event should decrease exponentially. Only if these “triple dividends of MOSFET scaling”15 – smaller, faster, more power-efficient – can be achieved in concert, Moore’s Law will survive.

In the recent past it has become more and more difficult to achieve the required performance improvements when scaling the Si MOSFET. A material property closely related to transistor performance is the carrier mobility, and a high mobility is always beneficial for transistor speed and power efficiency. Since Si, the backbone of today’s More Moore circuits, offers only moderate mobilities, rigorous efforts are made to implement alternative high-mobility MOSFET channel materials into Si CMOS. A first step was the introduction of strain into the Si MOSFET channels leading to enhanced mobilities. Meanwhile, strained Si is widely used in CMOS production16 and research on high-mobility III–V and Ge channels for future MOSFET generations is aggressively pushed forward.15,17

Beyond the horizon of the current ITRS edition, when gate lengths below 5 nm will be required, MOSFETs with high-mobility (and thus light carrier effective mass) channels will encounter a fundamental problem. At this stage of scaling, the distance between the MOSFET’s source and drain regions

Fig. 1 The More Moore and More Than Moore domains of semiconductor electronics, together with important trends and the semiconductors used in these domains. Note that the year 2028 indicates the end of the time horizon of the 2013 ITRS edition.7

Fig. 2 Evolution of the MOSFET gate length and the number of transistors integrated on a single microprocessor chip. The numbers above the gate length curve indicate the processor supply voltage $V_{DD}$. Note the continuous decrease of $V_{DD}$ in the past and the required continuation of this trend in the future. The ITRS targets refer to MOSFETs for high-performance logic as specified in the 2013 ITRS edition.7
becomes so short that quantum-mechanical source–drain tunneling impedes proper FET operation. It may turn out that for sub-5 nm MOSFETs heavy effective mass materials with lower mobilities will become preferable since a heavy carrier effective mass suppresses source–drain tunneling.18,19

2.2 More Than Moore trends

More Than Moore is not primarily focused on increasing circuit complexity but rather on enhancing the functionality of ICs and electronic systems by combining digital electronics with components such as analog/RF (radio frequency) and high-voltage circuitry, sensors, actuators, etc. In the More Than Moore domain, traditionally the material basis has been much broader compared to Moore and, apart from Si, a variety of alternative semiconductors is used to complement Si.

More Than Moore covers a wide and heterogeneous field, see Fig. 1. Here we focus solely on RF electronics where at present, in addition to Si, the compound semiconductors based on GaAs, InP, and GaN are very popular. It should be noted that in RF electronics only n-channel FETs are used since in most materials electrons are faster than holes. A major trend in RF electronics is the striving for higher transistor operating frequencies and for extending transistor operation into the THz range. The evolution of the frequency performance of RF FETs in terms of the characteristic frequencies $f_T$ (cutoff frequency) and $f_{\text{max}}$ (maximum frequency of oscillation) is shown in Fig. 3.

As can be seen, the current record $f_{\text{max}}$ of RF FETs is slightly above 1 THz, which means that FETs capable of THz amplification with reasonable power gain are still missing. On the other hand, the frequency range around and above 1 THz is attractive for applications in a variety of fields, such as security, medicine, and ultra-high-speed communications, to name just a few.40,41 Therefore, RF device engineers are looking for alternative FET channel materials offering improved mobility and closely follow the research on 2D materials.

2.3 Transistor figures of merit

A FET consists of a channel region connecting two reservoirs of mobile charges called source and drain. The third electrode (gate) is separated from the channel by a thin barrier. The applied gate–source voltage $V_{GS}$ controls the conductivity of the channel and the applied drain–source voltage $V_{DS}$ drives a drain current $I_D$ through the transistor. Fig. 4 shows a generic FET, together with the structures of Si MOSFETs, III–V HEMTs, and 2D MOSFETs. FETs are devices with the ability (i) to switch and (ii) to amplify signals and provide gain. For digital electronics, i.e., More Moore, switching is relevant while amplifying signals is important for RF applications.

In digital logic, the FET is supposed to switch between on and off. In the on-state, see Fig. 4(a), the channel has a low resistance and a large on-current $I_{on}$ can flow through it. As per definition, $I_{on}$ is the transistor current for the bias condition $V_{GS} = V_{DS} = V_{DD}$. In the off-state, see Fig. 4(b), on the other hand, the FET should block the current, the channel resistance should be high, and only a very small off-current $I_{off}$ is allowed to flow. The off-current is defined as the current flowing under the condition $V_{GS} = 0$ and $V_{DS} = V_{DD}$. The gate voltage at which the transistor is just at the verge of switching on is the threshold voltage $V_{Th}$. The transfer characteristics of a MOSFET shown in Fig. 5 indicate that in the subthreshold region ($V_{GS} < V_{Th}$) the drain current depends exponentially on $V_{GS}$, followed by a transition region around $V_{Th}$ and finally by the superthreshold region where the drain current is essentially linearly dependent on $V_{GS}$. As we have shown in Fig. 2, the supply voltage $V_{DD}$ of logic circuits has been decreased continuously over the years and a further reduction is required for the future. This means that logic transistors must switch from off to on within a very limited range of $V_{GS}$.

A further FOM (figure of merit) related to switching is the on–off ratio $I_{on}/I_{off}$. For FETs to be used in digital logic, on–off ratios in the range of $\approx 10^4–5 \times 10^7$ are required.7 As a basic rule, $I_{off}$ should be as low as possible and both $I_{on}$ and the on–off ratio should be as high as possible. A low $I_{off}$ is needed for a low static power consumption of logic circuits while a high $I_{on}$ is relevant for a high transistor switching speed.

When operated as an amplifier, on the other hand, the FET does not necessarily need to switch off. Instead, in most RF amplifier configurations the FET is permanently operated in the on-state and small signals applied to its input, i.e., the gate, appear amplified at the output. The extent to which the input signal is amplified is called gain. The small-signal current gain, for example, is defined as the RF output current of the transistor divided by the RF input current. Gains are frequency dependent and decrease with increasing frequency. Two important FOMs of RF transistors are the characteristic frequencies $f_T$ and $f_{\text{max}}$. The cutoff frequency $f_T$ is the frequency at which the small-signal current gain $h_{\text{f1}}$ of the transistor has dropped to unity (i.e., 0 dB) and the maximum frequency of oscillation $f_{\text{max}}$ is the frequency where the unilateral power gain $U$ becomes unity. It should be noted that for most RF applications, power gain and $f_{\text{max}}$ are more important.
than current gain and $f_T$. Moreover, as a rule of thumb, the operating frequency should be lower than 20% of the used transistors’ $f_{\text{max}}$ to guarantee sufficient power gain. Fig. 6 shows the small-signal current and power gains of an RF FET as a function of frequency, together with $f_T$ and $f_{\text{max}}$.

Commonly the voltage gain of RF FETs is not discussed explicitly since, if both current and power gain are reported, this data pair contains information on the voltage gain. Because FETs with gapless channels, e.g. graphene MOSFETs, suffer from poor power gain, an inspection of the voltage gain is advisable, however. The frequency-dependent voltage gain $A_V$ is defined as

$$A_V = \frac{z_{21}}{z_{11}}$$  \hspace{1cm} (1)

where $z_{21}$ and $z_{11}$ are ac impedance parameters. At low frequencies, $A_V$ approaches the so-called intrinsic gain $G_{\text{int}}$ given as

$$G_{\text{int}} = \frac{g_m}{g_d}$$  \hspace{1cm} (2)

where $g_m$ is the transconductance (i.e., the slope of the transfer characteristics, see Fig. 5, at the dc operating point) and $g_d$ is the drain conductance (i.e., the slope of the $I_D-V_{DS}$ output characteristics). A transistor with low intrinsic gain, e.g., caused by a large drain conductance, will always suffer from a low power gain.

All semiconductor devices generate fluctuations of voltage and current called noise. Noise is always undesirable and particularly critical for the amplification of small RF signals.

Fig. 4  Basic FET structures. Generic structure of a FET (a) in the on-state and (b) in the off-state. (c) Conventional Si n-channel MOSFET. (d) HEMT. (e) 2D MOSFET. (f) Back-gate 2D MOSFET frequently used for proof-of-concept purposes. Note that HEMTs and 2D MOSFETs do not possess pn junctions as present in conventional Si MOSFETs shown in (c) but rather resemble the junctionless MOSFET which is intensively investigated at present.

Fig. 5  Transfer characteristics of an n-channel FET showing the drain current $I_D$ as a function of the gate-source voltage $V_{GS}$ together with the on and off operating points for CMOS logic. Note that the scale of the left current axis is logarithmic and that of the right current axis is linear.

Fig. 6  Small-signal current gain $h_{21}$ and unilateral power gain $U$ of a RF FET as a function of frequency, after ref. 43 and 44. The characteristic frequencies $f_T$ and $f_{\text{max}}$ are obtained by extrapolating the measured $h_{21}$ and $U$ with the characteristic slope of $-20$ dB per dec.
A measure for the noise generated in a transistor is the noise figure $NF$, usually given in units of dB and defined as

$$NF \text{ [dB]} = 10 \times \log \frac{P_{Si}/P_{NI}}{P_{SO}/P_{NO}} \quad (3)$$

where $P_{Si}$ and $P_{SO}$ are the signal powers at the input and output, and $P_{NI}$ and $P_{NO}$ are the noise powers at the input and output, respectively. Under optimum matching and bias conditions, the noise figure reaches a minimum called minimum noise figure $NF_{\text{min}}$. It is the relevant FOM to characterize the noise performance of RF transistors. For a good RF FET the characteristic frequencies $f_r$ and $f_{\text{max}}$ should be high and the minimum noise figure $NF_{\text{min}}$ should be low.

### 3. Overview of 2D materials

Inspired by the successful preparation of graphene, researchers have intensively examined options to obtain stable 2D materials beyond graphene. These efforts include real-world experiments where 2D materials have been prepared and analyzed, as well as computational experiments where the thermodynamic stability and the band structure of existing and hypothetical 2D materials have been computed. For the latter, particularly the Atlas of 2D Materials reporting the properties of more than 140 different 2D materials is worth mentioning.

Recent experimental and theoretical efforts have shown that a variety of 2D materials beyond graphene do exist and that their electronic properties span the full range from metallic to insulating. Of particular interest for FET channels are the semiconducting and, to a lesser extent, the gapless semi-metallic 2D materials. Fig. 7 shows schematically the relevant parts of the band structure of different classes of these 2D materials.

#### 3.1 X-enes

Single-layer materials consisting of atoms of one single element arranged in a hexagonal lattice are designated as X-enes. So far, graphene as well as its Si-, Ge-, and P-based counterparts silicene, germanene, and phosphorene have been experimentally realized, and the band structure of stanene, the X-ene based on tin (Sn), has been calculated.

Compared to graphene, much less is known on the other X-enes. As of the end of 2014, the Web of Science database listed almost 91,000 entries under the search term graphene compared to 654 entries for silicene, 133 for germanene, 81 for phosphorene, and 8 for stanene. As shown in Fig. 7, graphene, silicene, germanene, and stanene are gapless and have cone-shaped conduction and valence bands. The cones are frequently called Dirac cones and, correspondingly, graphene, silicene, germanene, and stanene are designated as Dirac materials. Phosphorene, on the other hand, is a semiconductor with a sizeable gap.

#### 3.2 X-anes

The crystallographic structure of the X-anes is closely related to that of the X-enes. They also possess a hexagonal lattice of carbon (graphene), silicon (silicene), germanium (germanane), or tin (stanane) atoms. However, their lattice atoms are additionally out of plane bonded to hydrogen atoms – this is frequently called hydrogenated. Graphane was predicted to exist in 2005 and was produced experimentally shortly afterwards. Recently, germanane could also be realized experimentally.

Most relevant for electronic applications is the fact that graphene, silicane, and germanane possess sizeable bandgaps.

#### 3.3 Fluoro-X-enes

The structure of these materials is very similar to that of the X-enes. Here, the lattice atoms are bonded out of plane to F (fluorine) atoms. Fluorographene has already been produced experimentally and shown to have a wide gap of around, possibly even exceeding, 3 eV. Theory has confirmed these results for fluorographene and predicted a gap around 1 eV for fluorosilicene while fluorogermanene seems to be gapless.

#### 3.4 TMDs

The TMDs constitute a group of materials consisting of a transition metal $M$ (elements of groups 4, 5, and 6 of the periodic table of elements) and a chalcogen Q, i.e., sulfur (S), selenium (Se), or tellurium (Te). These M and Q elements form covalently bonded 2D layers of the MQ$_2$ type (e.g., MoS$_2$) with a hexagonal lattice. Single-layer TMDs consist of three atomic layers where a layer of M atoms is sandwiched between two layers of Q atoms. For example, single-layer MoS$_2$ is composed of one layer of molybdenum atoms and two layers of sulfur atoms. Today, more than 40 different types of TMDs are known. While many of them are metallic, those containing Mo and W (i.e., MoS$_2$, WS$_2$, etc.) as well as several of the Hf, Pd, Pt, and Zr-based TMDs are semiconductors with bandgaps of the order of 1–2 eV. It should be noted that TMDs occur in different polytypes 1T, 1T’, 2H, and 3R, where T means trigonal, T’ distorted trigonal, H hexagonal, and R rhombohedral, and 1, 2, or 3 indicates the number of TMD layers in the unit cell, which have different properties. For

![Fig. 7 Schematic band structure of 2D materials relevant for transistors. BLG: bilayer graphene; GNR: graphene nanoribbon; TMD: transition metal dichalcogenide; SMC: semimetal chalcogenide.](image-url)
example, the common 2H polytypes of the Mo- and W-based TMDs are semiconducting while their metastable counterparts of the 1T type are metallic.\textsuperscript{58}

3.5 SMCs

SMCs consist of a semimetal M (Ga or In) and a chalcogen (S or Se). In contrast to the TMDs, they are expected to occur in M\textsubscript{2}X\textsubscript{2} stoichiometry in a four-layer X–M–M–X configuration and to be semiconducting.\textsuperscript{45}

3.6 MX-enes

There is a material class called the MAX phase family comprising more than 60 individual ternary layered materials.\textsuperscript{59} These materials have a hexagonal lattice and the composition M\textsubscript{n+1}AX\textsubscript{n} where M is an early transition metal, A is a group 13 or 14 element, X is either carbon or nitrogen, and n is an integer equal to 1, 2, or 3. The bonds between the M and X atoms are much stronger than the M–A bonds so that the A atoms can easily be removed, e.g., by an acid treatment. By a subsequent sonication, single M\textsubscript{n+1}X\textsubscript{n} layers, the so-called MX-enes, can be obtained. As the first MX-ene, Ti\textsubscript{3}C\textsubscript{2} has been successfully prepared from the MAX material Ti\textsubscript{3}AlC\textsubscript{2},\textsuperscript{60} soon followed by the experimental verification of five further MX-enes,\textsuperscript{61} and the existence of even more MX-enes has been predicted. Furthermore, F\textsubscript{2}, (OH)\textsubscript{2}, and O\textsubscript{2} groups can be attached to the pure MX-enes of the M\textsubscript{2}X configuration, resulting in the formation of the modified MX-enes M\textsubscript{2}XF\textsubscript{2}, M\textsubscript{2}X(OH)\textsubscript{2}, and MXO\textsubscript{2}.\textsuperscript{62} Several modified MX-enes have been predicted to be semiconductors with sizeable bandgaps.\textsuperscript{62}

3.7 Further 2D materials

By first-principles calculations, the phonon properties and band structures of entire classes of 2D IV–IV and III–V compounds and of many other 2D materials have been investigated. It has been shown that a large body of these 2D materials with both hexagonal and tetragonal lattice structures should be stable and show bandgaps between 0.2 and 5 eV.\textsuperscript{45,63,64} Although it is uncertain whether all these 2D materials can be synthesized, at least part of them may become available for experiments in the future.

3.8 Production of 2D materials

To fabricate 2D transistors, first the 2D starting material must be produced, preferably in the form of large-area sheets with uniform thickness and high crystallographic quality. Layered van der Waals materials, where the stacked layers are bound by weak van der Waals forces, can be exfoliated. Mechanical exfoliation is widely used to produce graphene,\textsuperscript{1} TMDs,\textsuperscript{6} and phosphorene\textsuperscript{48} layers. While this approach is simple and does not require expensive equipment, it is time consuming and provides flakes of limited size only. Liquid exfoliation as a second exfoliation method delivers 2D flakes dispersed in a liquid.\textsuperscript{65} This method is very effective, but the dispersed flakes are small, which makes transistor and circuit processing challenging.

As an alternative, wafer-scale 2D materials can be grown on substrates. Epitaxial graphene on SiC has successfully been formed and used for transistor processing.\textsuperscript{66} Graphene can also be grown by CVD (chemical vapor deposition) on metals and subsequently be transferred to insulating substrates such as oxidized Si wafers.\textsuperscript{67} Furthermore, different TMDs such as MoS\textsubscript{2}, WS\textsubscript{2}, and WSe\textsubscript{2} have been grown by CVD directly on oxidized Si wafers.\textsuperscript{68,69} Finally, silicene has been grown on metals and transferred to SiO\textsubscript{2}/Si substrates for device processing.\textsuperscript{70}

4. Properties of 2D materials relevant for transistors

4.1 The ideal material for a FET channel – a wish list†

To assess the potential of novel materials for FET channels it is not sufficient to consider only one single material property such as the carrier mobility. Instead, a set of properties should be examined to get a realistic impression on the suitability of the new material for transistors. Therefore, in the following we compose a (certainly not complete) wish list for the material properties of a FET channel, particularly for use in digital logic and RF FETs, and examine the extent to which the 2D materials meet our wishes. In the discussion it should always be kept in mind that research on most of the 2D materials has just begun and is in an embryonic stage compared to the conventional 3D bulk semiconductors such as Si, Ge, and the III–V compounds. This leads to the situation that, while our wish list itself is well established and relies on the experiences of decades of transistor research, the available data for the relevant properties of the 2D materials are rather fragmentary.

4.1.1 Wish list entry #1: bandgap

The bandgap is a key property of semiconductors. It decisively affects their applicability to electronic devices and the existence of a gap is essential for proper FET operation. As discussed in section 2.3, FETs for digital logic need a high on–off ratio. While the on-current of a FET is not directly related to the bandgap $E_G$, the off-current strongly depends on $E_G$ according to

$$I_{\text{off}} \propto \exp \left( -\frac{E_G}{mk_BT} \right)$$

where $m$ is a factor of 2 (ref. 72) or larger (depending on the specific FET design), $k_B$ is the Boltzmann constant, and $T$ is the temperature. Thus, the on–off ratio follows

$$\frac{I_{\text{on}}}{I_{\text{off}}} \propto \exp \left( \frac{E_G}{mk_BT} \right)$$

Estimations suggest that at room temperature a gap of the order of 400 meV is needed to achieve a sufficiently good switch-off and the required on–off ratios.\textsuperscript{72–74}

Since RF FETs can permanently be operated in the on-state and do not need to switch off, one might conclude that here a gap is not needed. Unfortunately, the situation is more

†The idea for the title of this section originates from ref. 71.
complex. It has been discussed in detail that to achieve a high power gain and high $f_{\text{max}}$, the FET needs to show a sufficiently good saturation of the drain current and for that, in turn, a gap is needed. Current gain and $f_T$, on the other hand, are less affected by a missing or weak current saturation. Thus, transistors with gapless channels can provide high current gain and $f_T$ but suffer from poor power gain and $f_{\text{max}}$. It is not exactly known how wide the gap of the channel of a good RF FET should be. Most likely, the requirements are less stringent than for logic FETs. The fastest and least noisy RF FETs are InP HEMTs and GaAs mHEMTs having In$_x$Ga$_{1-x}$As channels with high In contents $x$ ranging from 0.7 ($E_G \approx 0.65$ eV) to 1 (i.e., InAs, $E_G = 0.35$ eV). Experiments with InSb channel HEMTs ($E_G$ InSb = 0.17 eV), on the other hand, revealed that these transistors, despite being fast, show lower $f_T$ and $f_{\text{max}}$ compared to InP HEMTs and GaAs mHEMTs with a similar gate length. This suggests that the 0.17 eV gap of InSb might already be too narrow for ultra-high-performance RF FETs.

4.1.2 Entry #2: carrier transport and effective mass. Logic and RF FETs should be fast, i.e., they should react quickly on the variations of their input signals, and show a large on-state current. For this, fast carriers are needed. Measures for the speed of carrier transport are the mobility $\mu$, the peak velocity $v_{\text{peak}}$, and the saturation velocity $v_{\text{sat}}$. When a low electric field $E$ acts on a carrier, its drift velocity $v$ is given by $v = \mu \times E$. The velocity is inversely proportional to the carrier effective mass $m_{\text{eff}}$ and high $m_{\text{eff}}$ is a precondition for a high $\mu$. Under high field conditions, the carrier velocity no longer follows the field linearly. The high-field velocity for holes shows a soft saturation and approaches $v_{\text{sat}}$ at high fields. For electrons, the situation is more diverse. In some semiconductors (e.g., GaAs), the electron velocity shows a pronounced peak ($v_{\text{peak}}$) at a certain field, then decreases at higher fields, and eventually approaches the saturation velocity. In other semiconductors, e.g., Si, soft saturation without a velocity peak occurs. Fig. 8 shows exemplarily the $v$–$E$ (velocity–electric field) characteristics for electrons and holes in Si and GaAs. The channel material of a fast FET should show a high mobility (and thus a light $m_{\text{eff}}$), preferably combined with a high $v_{\text{peak}}$ and/or $v_{\text{sat}}$.

For logic FETs, a high mobility and a high saturation velocity are always desirable. It should be mentioned, however, that Si MOSFETs show surprisingly good switching characteristics in spite of the rather moderate $\mu$ (200–500 cm$^2$ V$^{-1}$ s$^{-1}$) and $v_{\text{sat}}$ (10$^7$ cm s$^{-1}$). The In$_x$Ga$_{1-x}$As channels of the best RF FETs, i.e., InP HEMTs and GaAs mHEMTs, show mobilities of 10 000–15 000 cm$^2$ V$^{-1}$ s$^{-1}$ and peak velocities of 3–4 × 10$^7$ cm s$^{-1}$, and alternative channel materials that are to compete with In$_x$Ga$_{1-x}$As for high-performance RF FETs should show at least similar transport characteristics, combined with an appropriate bandgap. Reasonably good RF performance, however, can be achieved already with lower mobility channels, e.g., Si channels of RF Si MOSFETs, provided the contact resistance (see entry #4) is low and the scale length (entry #5) is short.

4.1.3 Entry #3: heat transport. If a voltage is applied to a FET and a current is flowing through its channel, electrical energy is converted into heat and this heat must be removed to avoid unacceptable self-heating. To accomplish an effective heat removal, the channel material must show a high thermal conductivity $\kappa$.

The thermal conductivity of the channel material, however, does not tell the whole story of heat transport, in particular if the transistor is located on a substrate different from the channel material. 2D FETs, such as graphene or TMD FETs (see Fig. 4e showing a MoS$_2$ MOSFET), are frequently realized on the surface of oxidized Si wafers. Here, the heat generated in the channel first has to cross the channel–SiO$_2$ interface acting as a thermal boundary resistance $R_{\text{TB}}$, move through the SiO$_2$ layer, and then cross the SiO$_2$–substrate interface forming a second $R_{\text{TB}}$ before it can spread across the Si substrate. To limit self-heating, the thermal conductivity of the FET channel material and of all materials underneath should be high and the thermal boundary resistances of all interfaces between the channel and the back side of the chip should be low.

4.1.4 Entry #4: contact resistance. In a FET, the gate–source voltage controls the channel conductivity and thus the drain current $I_D$. A closer inspection shows that not the gate–source voltage applied between the gate and source terminals, but rather the intrinsic gate–source voltage $V_{\text{GS-int}}$, i.e., the potential difference between the gate and the source-side end of the channel, is controlling the current. Carriers coming from the source terminal first cross the metal–semiconductor interface which hinders the carrier flow and acts as a resistance called source contact resistance $R_{\text{co-S}}$. Next, the carriers flow through an ungated semiconductor region until they reach the channel underneath the gate. This ungated region represents the source series resistance $R_{\text{ser-S}}$. The overall parasitic resistance at the source side, $R_S$ is the sum $R_{\text{co-S}} + R_{\text{ser-S}}$. The same situation occurs at the drain side of the FET leading to the parasitic drain resistance $R_D$.

Due to the voltage drop across $R_S$, part of the applied $V_{\text{GS}}$ is lost for the control of the current as can be seen from

$$V_{\text{GS}} = V_{\text{GS-int}} + I_D(R_{\text{co-S}} + R_{\text{ser-S}}) = V_{\text{GS-int}} + I_DR_S$$ (6)
Similarly, the applied drain–source voltage \( V_{DS} \) is different from the intrinsic drain–source voltage \( V_{DS\text{-int}} \) according to
\[
V_{DS} = V_{DS\text{-int}} + I_D(R_S + R_D) \tag{7}
\]

The parasitic source and drain resistances \( R_S \) and \( R_D \), and thus the contact resistance \( R_{co} \), deteriorate transistor performance. Therefore, the \( R_{co} \) should be as low as possible.

The contact resistance is proportional to the inverse of the contact width, \( R_{co} \propto 1/W \). To compare the contact resistances obtained from structures with different contact widths, \( R_{co} \) is commonly normalized with respect to the contact width and is given in units of \( \Omega \) mm.

### 4.1.5 Entry #5: scale length and channel thickness

The scale length \( \lambda \) is not a material property in the strict sense. It provides, however, valuable information on the electrostatic integrity, the ability to suppress undesirable short-channel effects, and the scaling limits of a certain FET design. In ref. 80, the scale length expression
\[
\lambda = \sqrt{\varepsilon_{\text{ch}}t_{\text{bar}}/\varepsilon_{\text{bar}}} \tag{8}
\]
has been derived, where \( \varepsilon_{\text{ch}} \) and \( \varepsilon_{\text{bar}} \) are the dielectric constants of the channel and the barrier separating gate and channel (e.g., the gate oxide in MOSFETs), \( t_{\text{ch}} \) is the thickness of the channel region, and \( t_{\text{bar}} \) is the barrier thickness. If a transistor with a gate length \( L \) fulfills the condition
\[
L \geq a \times \lambda \tag{9}
\]
where \( a \) is a constant, short-channel effects, most notably the undesirable degradation of the subthreshold behavior and the increase of the off-current, are sufficiently suppressed. Therefore, a short scale length and thus a combination of a thin channel region and a thin barrier are always desirable.

Three issues related to the scale length should be borne in mind. First, eqn (8) has been derived for fully depleted SOI (silicon on insulator) MOSFETs with relatively thick Si channels. Thus, strictly speaking, eqn (8) does not apply to FETs with extremely thin channels, e.g., channels of 2D materials. Second, the effect of the insulating layer underneath the channel has not been taken into account. Third, the scale length concept is valid only for semiconducting channels and does not apply to gapless channels. The main message of eqn (8) and (9) is that a thin channel region is beneficial for suppressing short-channel effects and FET scaling. If not only the trend is of interest, but actual numbers for the scale length of MOSFETs with 2D channels are needed, more elaborate scale length expression developed for atomically thin channels should be used.\(^{81,82}\)

### 4.1.6 Addendum to entry #2

We have seen that FET channel materials having high mobility and light effective mass are desirable, and indeed, chipmakers put significant effort in increasing the channel mobility in CMOS logic FETs. It should be noted, however, that a light effective mass may also cause two problems. First, a material with a light effective mass \( m_{\text{eff}} \) has a low DOS (density of states), and a low DOS means that, to achieve a certain variation of the channel charge \( \Delta Q_{\text{ch}} \) (and thus of the drain current), a larger variation of the gate-source voltage \( \Delta V_{GS} \) is needed.\(^{82}\) The consequence of this effect called the DOS bottleneck is that FETs with light-\( m_{\text{eff}} \) channels may suffer from a depressed control effect of the gate and thus a degraded transconductance \( g_m \) that obeys the relation
\[
g_m \propto \frac{\Delta Q_{\text{ch}}}{\Delta V_{GS}} \tag{10}
\]
where \( v_{\text{eff}} \) is the effective carrier velocity in the channel, which is related to \( \mu, v_{\text{peaks}} \), and \( v_{\text{sat}} \). While a light \( m_{\text{eff}} \) is beneficial for a high \( v_{\text{eff}} \), the opposite is the case for \( \Delta Q_{\text{ch}}/\Delta V_{GS} \). Thus, a light \( m_{\text{eff}} \) is preferable only as long as its positive effect on \( v_{\text{eff}} \) overcompensates its deteriorating influence on \( \Delta Q_{\text{ch}}/\Delta V_{GS} \).

Second, beyond the horizon of the 2013 ITRS edition, when gate lengths of 5 nm and below will be needed and source-to-drain tunneling will become an issue, FETs with high-mobility light-\( m_{\text{eff}} \) channels may fail and channel materials with wider bandgap, heavier carrier effective mass (and, consequently, lower mobility) may be preferred to suppress source-to-drain tunneling.\(^{18,19}\)

Table 1 summarizes our discussion on the wish list for the properties of FET channel materials.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Property</th>
<th>Desirable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bandgap</td>
<td>Logic: ≥0.4 eV hp: high-performance; L: gate length. Note that for logic (L &gt; 5 nm) a carrier effective mass &lt; 0.1 is desirable in general while for holes ( m_{\text{eff}} \leq 0.2 ) is probably more realistic since for many semiconductors the effective mass for holes is noticeably heavier than for holes. The effective mass is given in units of the electron rest mass ( m_0 )</td>
</tr>
<tr>
<td>2</td>
<td>Carrier transport</td>
<td>Carrier effective mass: Logic: light (L &gt; 5 nm), ( m_{\text{eff}} &lt; 0.1 ) heavy (L ≤ 5 nm), ( m_{\text{eff}} \geq 0.5 ) hp RF: very light, ( m_{\text{eff}} &lt; 0.05 ) hp Logic: high, &gt; 500 cm² V⁻¹ s⁻¹ hp RF: very high, &gt; 10000 cm² V⁻¹ s⁻¹</td>
</tr>
<tr>
<td>3</td>
<td>Mobility</td>
<td>Logic: high, &gt; 100 cm s⁻¹ hp RF: very high, ≥3 × 10⁷ cm s⁻¹</td>
</tr>
<tr>
<td>4</td>
<td>Peak/saturation velocity</td>
<td>High</td>
</tr>
<tr>
<td>5</td>
<td>Heat transport</td>
<td>Thermal conductivity: Low</td>
</tr>
<tr>
<td>6</td>
<td>Contact resistance</td>
<td>Low, ≤0.03 Ω mm</td>
</tr>
<tr>
<td>7</td>
<td>Scale length</td>
<td>Small</td>
</tr>
<tr>
<td>8</td>
<td>Channel thickness</td>
<td></td>
</tr>
</tbody>
</table>
wide body of results obtained from first-principle (ab initio) calculations. Calculations provide not only the bandgap itself, but the entire band structure, from which the carrier effective masses can be extracted. Thus, first-principle calculations are a very useful means of assessing the suitability of a semiconductor for FET channels. It should be mentioned, however, that different methods of first-principle calculations can lead to quite different bandgaps for one and the same material. The popular DFT (density functional theory) systematically underestimates the bandgap, while computationally more demanding approaches, such as GW, predict wider gaps which, however, in some cases exceed the measured gaps.45,84

Two examples for the considerable differences between the bandgaps obtained from DFT calculations and those calculated using more rigorous methods are given below. For germanane, the bandgap predicted by DFT ranges from 0.95 eV85 to 1.53 eV83 compared to values in the range of 1.84 eV55 to 3.6 eV86 obtained by other methods and a measured gap of 1.59 eV.53 As a second example we consider GNRs. For N = 7 armchair GNRs (N is the number of carbon atoms along the GNR width and N = 7 corresponds to a width of about 0.74 nm), a gap of 3.8 eV has been calculated by the GW method87 compared to 1.5–1.6 eV obtained by DFT87,88 and measured gaps of 2.3–2.8 eV.89,90 Thus, calculated bandgaps should be treated as estimates rather than accurate predictions, and those obtained by DFT calculations as a lower limit. On the other hand, the shape of the individual bands calculated by DFT and the extracted carrier effective masses can be considered as reasonable guides.45

Large-area graphene is gapless and the same holds for silicene and germanene.45 There are, however, options to open a gap in these materials. For graphene, the first approach is to form narrow GNRs by either chemical synthesis91,92 or lithographic patterning93,94 Chemically synthesized GNRs with atomically precise edges have been reported while patterned GNRs typically show non-ideal edges that degrade carrier transport. Fig. 9 compiles theoretical bandgap data (calculated by the GW method) together with experimental bandgap data for GNRs and shows the general trend of an increasing gap for decreasing width. As can be seen, narrow ribbons having a width of 10 nm or less are needed to achieve the 0.4 eV gap required for digital logic.

The second option to open a gap in graphene is by applying a perpendicular electric field to BLG, i.e., two graphene layers, one located directly on top of the other.96,97 As shown in Fig. 7, the conduction and valence bands of biased BLG are not parabolic as for most other semiconductors, but Mexican-hat-shaped.96 For FET-relevant BLG structures, realistically gaps up to 130 meV can be expected.97 This might be helpful for RF FETs but is not sufficient for digital logic FETs.

Recent calculations have predicted a gap opening in silicene and germanene nanoribbons98 as well as in single-layer silicene and germanene when a vertical electric field is applied.99 For a given field, however, the gap in silicene and germanene is smaller than that in biased BLG. Thus, regarding the bandgap, large-area silicene and germanene are not suited for logic FETs and their potential for RF FETs is unclear.

The bandgaps of the 2D materials (except Dirac materials, nanoribbons, BLG, and biased silicene and germanene) are summarized in Fig. 10. To guarantee consistency, all gaps shown are calculated by DFT. True semiconductors (Ec0 = 0.5–2 eV) are phosphorene, germanane, fluorosilicene, the Mo- and...
W-based TMDs, as well as HfS₂, ZrS₂, and TiS₂, GaSe, InSe, and several MXenes. Graphene, silicane, fluorographene, GaS, InS, and SiC can be considered as wide bandgap semiconductors while BN is an insulator. For comparison, bandgap data for conventional 3D semiconductors widely used for FET channels are also shown in Fig. 10.

Due to its wide gap, BN is not really a candidate channel material. It has, however, successfully been used as a gate dielectric for graphene MOSFETs and moreover shown to have a beneficial effect on the mobility in graphene channels underneath or above it.

The bandgap of TMDs and phosphorene depends on the layer number. It is widest for single layers and gradually decreases with increasing layer number toward the bulk value. This is important since the channels of experimental TMD and phosphorene MOSFETs frequently consist of few-layer material instead of single layers. The thickness dependence of the gap is particularly pronounced for phosphorene where the gap decreases from 1 eV for single layers (see Fig. 10) down to 0.67 and 0.52 eV for bilayer and trilayer materials.

In summary we state that in terms of bandgap many 2D materials fulfill the requirements of our wish list and therefore are suitable for FET channels.

### 4.3 Carrier transport and effective mass

It is well established that the carrier mobility of semiconductors tends to decrease with increasing bandgap. From Fig. 11 showing this trend for the electron mobility \( \mu_n \), we see that for the III-V compounds (black solid circles), the electron mobility reduces from 77 000 cm² V⁻¹ s⁻¹ for the narrow-bandgap InSb to 3000 cm² V⁻¹ s⁻¹ for the wider-bandgap Ga₀.₅₃In₀.₄₇P (\( E_g = 1.85 \) eV). Si (\( E_g = 1.12 \) eV) and Ge (\( E_g = 0.66 \) eV) follow this tendency although for these materials the electron mobility is lower than that of the III-V compounds with a comparable gap. Also graphene cannot escape from this mobility-bandgap trend. Suspended gapless graphene shows very high electron mobilities of up to 200 000 cm² V⁻¹ s⁻¹. The electron mobility in gapless graphene on insulating substrates is lower, but still impressively high as shown in Table 2. The gap opening in GNRs and biased BLG results, however, in a dramatic mobility reduction.

The electron mobility data reported by early 2015 for 2D materials beyond graphene are also included in Fig. 11. For MoS₂, phonon limited mobilities, i.e., the upper mobility limit for defect-free undoped material, ranging from 130 cm² V⁻¹ s⁻¹ (ref. 104) to 410 cm² V⁻¹ s⁻¹ (ref. 107), have been calculated and mobilities between 1 and 300 cm² V⁻¹ s⁻¹ have been extracted from experimental MoS₂ MOSFET structures, see, e.g., ref. 6, 109–112. Note that (i) the mobility reported for back-gated MoS₂ MOSFETs (a few to a few tens of cm² V⁻¹ s⁻¹) is lower than the mobility observed in top-gated MoS₂ FETs (several tens to a few hundreds of cm² V⁻¹ s⁻¹), and (ii) for top-gated MoS₂ FETs occasionally electron mobilities of up to 1000 cm² V⁻¹ s⁻¹ (ref. 121) have been reported. These data are not included in Fig. 11 since it most likely results from an irregular extraction approach.

For single-layer WSe₂, a hole mobility of 250 cm² V⁻¹ s⁻¹ has been reported that compares favorably with the best electron mobilities around 200 cm² V⁻¹ s⁻¹ observed in multilayer WSe₂ and in ref. 127 it has been shown that the hole mobility in WSe₂ can indeed be higher than the electron mobility. In few-layer phosphorene, a remarkable hole mobility of 1000 cm² V⁻¹ s⁻¹ has been measured.

Velocity-field characteristics for gapless graphene have been simulated and measured, and for BLG, GNRs, silicene, and MoS₂ high-field transport data have been simulated. The reported \( v-E \) characteristics of 2D materials show a soft saturation with a slight decline of the velocity at high fields. Table 3 summarizes the available saturation and peak velocity data for 2D materials. These high-field transport data can be compared with the \( v-E \) characteristics for Si and GaAs in Fig. 8.

It should be mentioned that for future generations of Si-based logic ICs, ultra-thin-body SOI MOSFETs and possibly Si nanowire MOSFETs with ultra-small body cross-section will be considering the electric field dependence of the mobility.

### Table 2 Measured electron mobility in gapless graphene produced by different methods. Note that the predicted upper limit for the mobility in gapless graphene on SiO₂/Si is 40 000 cm² V⁻¹ s⁻¹ (ref. 120)

<table>
<thead>
<tr>
<th>Graphene type</th>
<th>Supporting layer/substrate</th>
<th>Mobility (cm² V⁻¹ s⁻¹)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exfoliated</td>
<td>SiO₂/Si</td>
<td>24 000</td>
<td>115</td>
</tr>
<tr>
<td>Exfoliated</td>
<td>WS₂/SiO₂/Si</td>
<td>38 000</td>
<td>116</td>
</tr>
<tr>
<td>CVD</td>
<td>SiO₂/Si</td>
<td>16 000</td>
<td>117</td>
</tr>
<tr>
<td>Epitaxial</td>
<td>Si-face SiC</td>
<td>2000</td>
<td>118</td>
</tr>
<tr>
<td>Epitaxial</td>
<td>C-face SiC</td>
<td>8700</td>
<td>119</td>
</tr>
</tbody>
</table>
needed and that in these structures, the mobility is degraded and much smaller compared to bulk Si. For example, in sub-5 nm diameter Si nanowire MOS structures, the electron mobility at low perpendicular fields can be below 100 cm$^2$ V$^{-1}$ s$^{-1}$ and drop further at higher effective fields down to a few tens of cm$^2$ V$^{-1}$ s$^{-1}$. Thus, several 2D semiconductors can be considered as a viable alternative.

If carrier transport data for a new material are not available, an examination of the material's band structure and carrier effective mass can be helpful. Although the effective mass is not the only quantity influencing the mobility, a light effective mass is always an indication for a high mobility. In Fig. 12, the electron effective mass of 2D materials and of conventional semiconductors is plotted as a function of the bandgap.

For conventional semiconductors (Si, Ge, III–V compounds), the hole mobility $\mu_p$ is always lower than the electron mobility. This is particularly true for the III–V compounds with very high electron mobility. The ratio $\mu_p/\mu_n$ is around 0.5 for Ge, 0.3 for Si, 0.05 for GaAs, and approaches 0.01 for the narrow bandgap compounds InAs and InSb. For the Mo- and W-based TMDs and for germanane, as well as for phosphorene for transport in the armchair direction, $\mu_p/\mu_n$ ratios ($m_{\text{eff}}$ and $m_{\text{eff}}$ are the electron and hole effective masses) of 0.8–1 have been calculated and band structure calculations for BLG and GNRs also reveal a high degree of symmetry of the top of the valence band and the bottom of the conduction band. This suggests $\mu_p/\mu_n$ ratios close to unity for these materials. For CMOS logic, $\mu_n \approx \mu_p$ is highly desirable since under these conditions a symmetric design of n- and p-channel MOSFETs is possible.

When comparing the electron mobilities of the 2D materials in Fig. 11 with those of the conventional 3D semiconductors it seems that the semiconducting 2D materials show rather moderate mobilities and, with the exception of germanane, do not compete well. This leads us to the conclusion that the 2D materials cannot compete with the high-mobility III–V compounds and will not be suitable for ultra-fast high-performance transistors. On the other hand, their mobilities are appropriate for many other applications where high speed is not of primary importance. The calculated electron mobility of more than 18 000 cm$^2$ V$^{-1}$ s$^{-1}$ for germanane, on the other hand, is very promising. However, this result needs to be confirmed by experiments and to be reproduced by calculations of other groups.

### 4.4 Heat transport

Table 4 summarizes the currently available data for the thermal conductivity of 2D materials in comparison with that of 3D bulk materials. The thermal conductivity of suspended graphene exceeds that of metals (e.g., $\kappa_{\text{Cu}} = 400$ W m$^{-1}$ K$^{-1}$) and the maximum reported record value from ref. 138 is even above that of bulk graphite. The thermal conductivity of graphene on SiO$_2$ is still high but lower compared to suspended graphene due to interface interactions. GNRs as well as the Mo- and W-based TMDs show lower thermal conductivities.

Table 5 shows the thermal boundary resistance $R_{\text{th}}$ of graphene/SiO$_2$, graphene/SiC, and graphene/BN junctions and for comparison that of silicon on SiO$_2$ as used in modern SOI structures. Unfortunately, information on the thermal boundary resistance for the 2D materials beyond graphene is not yet

---

**Table 3** Peak and saturation velocities for 2D materials and the corresponding fields (in kV cm$^{-1}$) and carrier sheet densities

<table>
<thead>
<tr>
<th>Material</th>
<th>$v_{\text{peak}}$ (10$^7$ cm s$^{-1}$)</th>
<th>$v_{\text{sat}}$ (10$^7$ cm s$^{-1}$)</th>
<th>Comment</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphene</td>
<td>—</td>
<td>1–3 @ 20</td>
<td>Exp. + fit, different $n_s$</td>
<td>129</td>
</tr>
<tr>
<td>Graphene</td>
<td>5.8 @ 15</td>
<td>5.2 @ 50</td>
<td>Sim., $n_s = 10^{11}$ cm$^{-2}$</td>
<td>130</td>
</tr>
<tr>
<td>BLG</td>
<td>3.4 @ 5</td>
<td>2.9 @ 20</td>
<td>Sim., $n_s = 5 \times 10^{11}$ cm$^{-2}$</td>
<td>131</td>
</tr>
<tr>
<td>BLG</td>
<td>—</td>
<td>3.0 @ 25</td>
<td>Sim., $n_s = 10^{12}$ cm$^{-2}$, $E_G = 0.1$ eV</td>
<td>130</td>
</tr>
<tr>
<td>GNR</td>
<td>3.7 @ 10</td>
<td>$&lt;3$ @ &gt;30</td>
<td>Sim., $n_s = 10^{12}$ cm$^{-2}$, $w = 10.1$ nm</td>
<td>132</td>
</tr>
<tr>
<td>GNR</td>
<td>$&gt;3.1$ @ &gt;100</td>
<td>—</td>
<td>Sim., $n_s = 10^{12}$ cm$^{-2}$, $w = 2.62$ nm</td>
<td>130</td>
</tr>
<tr>
<td>Silicene</td>
<td>0.5 @ 30</td>
<td>0.46 @ 50</td>
<td>Sim.</td>
<td>104</td>
</tr>
<tr>
<td>Si</td>
<td>—</td>
<td>0.34 @ 100</td>
<td>Sim.</td>
<td>104</td>
</tr>
<tr>
<td>MoS$_2$</td>
<td>—</td>
<td>1.5 @ 100</td>
<td></td>
<td>108</td>
</tr>
<tr>
<td>MoS$_2$</td>
<td>—</td>
<td>0.34 @ 100</td>
<td></td>
<td>108</td>
</tr>
</tbody>
</table>

---

Fig. 12 Electron effective mass of 2D and conventional 3D semiconductors vs. bandgap. III–V compounds (black solid circles), from left to right InSb, InAs, In$_{0.53}$Ga$_{0.47}$As, InP, GaAs, Al$_{0.5}$Ga$_{0.5}$As: Data for TMDs and graphene from ref. 45, for germanane from ref. 53, for GNRs from ref. 135, and for phosphorene from ref. 136 and 137. The numbers at the GNR data points indicate the ribbon width according to the width-gap relationship from ref. 87.
available. The measured $R_{\text{TB}}$ of single-layer graphene on SiO$_2$ is more than a factor of 10 larger than that calculated for Si on SiO$_2$. This factor of 10, however, is less critical than it seems on first sight. Usually the thermal resistance encountered by the heat flow from the chip surface (where the devices are located and the heat is generated) to the heat sink at the back side of the chip is much larger than the thermal boundary resistance the heat has to surmount when flowing from the 2D channel to the underlying substrate. To illustrate this issue, let us consider a graphene device layer located on a Si wafer (thickness $t_{\text{Si}} = 360 \mu$m) covered with SiO$_2$ (thickness $t_{\text{SiO}_2} = 90 \text{ nm}$) and assume the simplified case of a homogeneous heat flow from the graphene layer downward to the bottom of the substrate. The heat generated in the graphene layer first crosses the graphene/SiO$_2$ interface having a thermal boundary resistance $R_{\text{TB1}}$, flows through the SiO$_2$ layer with a thermal resistance $R_{\text{th-SiO}_2}$, then crosses the SiO$_2$/Si interface ($R_{\text{TB2}}$), and finally flows through the Si substrate ($R_{\text{th-Si}}$).

The overall thermal resistance is the sum of the individual contributions and reads as

$$R_{\text{th}} = \left( R_{\text{TB1}} + R_{\text{th-SiO}_2} + R_{\text{TB2}} + R_{\text{th-Si}} \right) \times \frac{1}{A},$$

where $A$ is the area of the device layer. Using the data from Tables 4 and 5, one obtains $R_{\text{TB1}} = 10^{-8} \text{ m}^2 \text{ K W}^{-1}$, $R_{\text{th-SiO}_2} = 6 \times 10^{-8} \text{ m}^2 \text{ K W}^{-1}$, $R_{\text{TB2}} = 10^{-9} \text{ m}^2 \text{ K W}^{-1}$, and $R_{\text{th-Si}} = 2.8 \times 10^{-6} \text{ m}^2 \text{ K W}^{-1}$, which clearly shows the dominating role of the substrate for heat removal.

Note that the heat transport parameters from Tables 4 and 5 relate to room temperature. When the temperature rises beyond room temperature, the thermal conductance decreases (undesirable) and the thermal boundary resistance decreases (desirable) as well.

4.5 Contact resistance

The first contact experiments for 2D materials have been focused on graphene and revealed that achieving low-resistance contacts is challenging. In late 2010, the best metal–graphene contact resistances have been in the range $0.5–10 \Omega \text{ mm}$, i.e., orders of magnitude higher than for contacts to Si and III–V compounds. Therefore, a lot of effort has been invested in reducing the metal–graphene contact resistance that finally led to significant improvements. Today, metal–graphene contact resistances as low as $0.01–0.2 \Omega \text{ mm}$ have been achieved. This already comes close to the resistance of state-of-the-art metal contacts on Si and III–V semiconductors.

Meanwhile, data for metal–TMD and metal–phosphorene contacts have been reported as well. Here, however, the contact resistance is still too high and typically exceeds $0.5 \Omega \text{ mm}$. Table 6 summarizes the state-of-the-art metal contacts on 2D materials and on conventional semiconductors.

From the ITRS one can get an impression about the contact requirements related to logic MOSFETs. During the entire horizon of the 2013 ITRS edition, the maximum allowed parasitic resistance (which includes both the contact resistance itself and the semiconductor series resistance, see eqn (4)) is around $0.065 \Omega \text{ mm}$. Thus, the contact resistance must be significantly below $0.065 \Omega \text{ mm}$. As Table 6 indicates, so far only

<table>
<thead>
<tr>
<th>Transistor type</th>
<th>$R_{\text{co}}$ (Ω mm)</th>
<th>Metal</th>
<th>Comment</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphene MOSFET</td>
<td>0.011–0.08</td>
<td>Ni, Ti</td>
<td>BL</td>
<td>155</td>
</tr>
<tr>
<td>MoS$_2$ MOSFET</td>
<td>0.1–0.2</td>
<td>Ti, Ni, Pd/Au, Cr/Au</td>
<td>SL</td>
<td>155–158</td>
</tr>
<tr>
<td>WSe$_2$ MOSFET</td>
<td>0.2–1.6</td>
<td>Ni/Au, Ti/Au, Au</td>
<td>FL</td>
<td>159–162</td>
</tr>
<tr>
<td>Phosphorene MOSFET</td>
<td>1.75</td>
<td>Ni/Au, Pd/Au</td>
<td>FL</td>
<td>166</td>
</tr>
</tbody>
</table>

Table 6 Contact resistance of different FET structures. SL: single layer; BL: bilayer; FL: few-layer.

Table 4 Room temperature thermal conductivity $\kappa$ of 3D bulk and single-layer 2D materials (unit W m$^{-1}$ K$^{-1}$). e: Experiment; s: simulated.

<table>
<thead>
<tr>
<th>3D bulk materials</th>
<th>2D materials</th>
<th>Substrate</th>
<th>$\kappa$ (W m$^{-1}$ K$^{-1}$)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>Graphene</td>
<td>SiO$_2$</td>
<td>e, 600</td>
<td>141</td>
</tr>
<tr>
<td>SiC</td>
<td>GNr</td>
<td>SiO$_2$</td>
<td>e, 80</td>
<td>142</td>
</tr>
<tr>
<td>SiO$_2$</td>
<td>MoSe$_2$</td>
<td>No substrate</td>
<td>s, 103</td>
<td>144</td>
</tr>
<tr>
<td>Graphite</td>
<td>MoSe$_2$</td>
<td>No substrate</td>
<td>s, 54</td>
<td>144</td>
</tr>
<tr>
<td>SiC</td>
<td>WS$_2$</td>
<td>No substrate</td>
<td>s, 142</td>
<td>144</td>
</tr>
<tr>
<td>SiO$_2$</td>
<td>WSe$_2$</td>
<td>No substrate</td>
<td>s, 53</td>
<td>144</td>
</tr>
</tbody>
</table>

Table 5 Thermal boundary resistance of graphene and Si on different substrates in units of $10^{-8}$ m$^2$ K W$^{-1}$. xL: $x$ layers; e: experiment; s: simulated.
Table 7 Scale lengths of 2D MOSFETs and competing conventional FET types, all with one single top-gate

<table>
<thead>
<tr>
<th>Transistor class</th>
<th>ε₁(ch)</th>
<th>ε₁(bar)</th>
<th>t(Å)</th>
<th>f₁bar (THz)</th>
<th>λ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si SOI MOSFET</td>
<td>11.9</td>
<td>3.9</td>
<td>5</td>
<td>0.6</td>
<td>3</td>
</tr>
<tr>
<td>InP HEMT &amp; GaAs mHEMT</td>
<td>14.1</td>
<td>12.7</td>
<td>15</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>GNR MOSFET</td>
<td>1.8</td>
<td>3.9</td>
<td>0.35</td>
<td>0.6</td>
<td>&lt;1</td>
</tr>
<tr>
<td>MoS₂ MOSFET</td>
<td>2.8</td>
<td>3.9</td>
<td>0.72</td>
<td>0.6</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

Graphene has fulfilled this requirement. Given the successful reduction of $R_{on}$ for metal–graphene contacts by more than an order of magnitude within a few years, we expect sizeable improvements for TMD and phosphorene contacts in the near future as well.

4.6 Scale length

Table 7 shows the scale length $\lambda$ of GNR, MoS₂, and Si MOSFETs as well as of InP HEMTs and GaAs mHEMTs obtained using eqn (8). The results indicate that, in terms of scale length and thus regarding the suppression of short-channel effects and scaling limits, the FETs with 2D channels behave much better than those with conventional channel materials.

5. State-of-the-art 2D FETs

In the following, the status of research on 2D transistors is reviewed by presenting experimental transistor data collected from the literature. In addition, transistor performance trends are discussed on the basis of theoretical considerations.

5.1 X-ene FETs

5.1.1 Graphene FETs. Recently, several comprehensive review papers on the state-of-the-art graphene FETs have been published. Therefore, we only summarize the main messages of these papers and refer the reader to ref. 3, 102, 170 and 171 for more details.

5.1.2 Graphene MOSFETs for digital logic. MOSFETs with gapless large-area graphene channels do not switch off and show on–off ratios of only 2–10. Thus, large-area graphene is not a suitable channel material for logic MOSFETs. Back-gate GNR MOSFETs with on–off ratios of $10^{5}$ to $10^{6}$ have been successfully fabricated. Due to the thick back-gate dielectrics, however, these transistors need unacceptably large gate voltage swings of 3–20 V for switching. As we have shown in Fig. 2, the supply voltage for logic circuits (and thus the maximum available gate voltage swing) is currently below 1 V and is required to decrease further in the future. Therefore, top-gate GNR MOSFETs with very thin gate dielectrics are needed to achieve good switching behavior with a sub-1 V gate voltage swing. Moreover, as has been shown in Fig. 11, the gap opening in GNRs is accompanied by a dramatic mobility reduction. For these reasons, in the short to medium term the application of GNR MOSFETs in digital logic is rather unlikely. Things may change at 5 nm and below gate length levels when direct source–drain tunneling becomes an issue.

To circumvent the problems of gapless large-area graphene, vertical non-FET transistor concepts that do not require a gap for switch-off have been elaborated. These concepts rely either on graphene–Si Schottky barriers as described in ref. 174, where on–off ratios of up to $10^{5}$ have been demonstrated, on tunneling through an insulator between two graphene layers, or on using graphene as the base of a hot electron transistor. These devices represent interesting options to exploit gapless graphene in logic transistors and can in general be realized also using 2D materials beyond graphene. It is difficult, however, to assess their true potential for future logic at the moment.

5.1.3 Graphene MOSFETs for RF. So far, only experimental graphene RF FETs with gapless channels have been reported. Soon after the demonstration of the first graphene MOSFET with a gapless channel, the RF capabilities of such transistors have been investigated, and meanwhile many groups have realized RF MOSFETs with gapless graphene channels. Table 8 summarizes the best reported cutoff frequencies $f_{T}$ and maximum frequencies $f_{max}$ of such graphene MOSFETs, together with the $f_{T} - f_{max}$ performance of competing RF FET types. More details can be found in the $f_{T}$ vs. $L$ and $f_{max}$ vs. $L$ plots of ref. 3.

In terms of $f_{max}$, graphene MOSFETs perform competitively. They outperform Si MOSFETs with comparable size and compete well with InP HEMTs and GaAs mHEMTs (which are the fastest RF FETs of all) down to gate lengths of about 60 nm. Regarding the more important FOM $f_{max}$, however, the picture looks less promising for graphene MOSFETs. While the record $f_{max}$ of InP HEMTs and GaAs mHEMTs exceeds 1 THz and Si MOSFETs with an $f_{max}$ of 420 GHz have been reported, the best graphene RF FETs show an $f_{max}$ of about 100 GHz only. It has been discussed in detail that the main reason for this poor $f_{max}$ performance is the missing gap in large-area graphene. Thus, no matter how carefully the design of large-area graphene FETs is optimized, these transistors will

Table 8 The best $f_{T}$ and $f_{max}$ data for graphene RF MOSFETs with gapless channels and for competing RF FETs. $L$ is the gate length

<table>
<thead>
<tr>
<th>FET type</th>
<th>$L$ (nm)</th>
<th>$f_{T}$ (GHz)</th>
<th>$f_{max}$ (GHz)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphene MOSFET</td>
<td>67</td>
<td>427</td>
<td>—</td>
<td>179</td>
</tr>
<tr>
<td>(gapless channel)</td>
<td>40</td>
<td>350</td>
<td>22</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>144</td>
<td>300</td>
<td>—</td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>93</td>
<td>105</td>
<td>182</td>
</tr>
<tr>
<td>InP HEMT</td>
<td>30</td>
<td>644</td>
<td>681</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>600</td>
<td>1200</td>
<td>36</td>
</tr>
<tr>
<td>GaAs mHEMT</td>
<td>40</td>
<td>688</td>
<td>800</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>515</td>
<td>&gt;1000</td>
<td>76</td>
</tr>
<tr>
<td>Si MOSFET</td>
<td>29</td>
<td>485</td>
<td>420</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>360</td>
<td>—</td>
<td>185</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>395</td>
<td>410</td>
<td>186</td>
</tr>
</tbody>
</table>
never compete well with III–V HEMTs and Si MOSFETs in terms of power gain and $f_{\text{max}}$.

When comparing experimental $f_T$ and $f_{\text{max}}$ data of graphene MOSFETs with that of other RF FETs, one should take a closer look at the applied de-embedding procedure. De-embedding is a common practice in RF electronics to eliminate the effect of the parasitics of the measurement environment from the measured RF data. Usually all parasitics down to the large pads (needed for the RF probes) are de-embedded while the metal lines from the pads to the transistor are not de-embedded. In the RF characterization of graphene MOSFETs, however, frequently these metal lines are de-embedded as well. This full de-embedding procedure provides the RF parameters of the intrinsic device which are difficult to compare with those obtained by the common pad de-embedding approach and leads to a very optimistic picture of the transistor’s RF performance. This issue has nicely been discussed in ref. 187 and its relevance becomes evident from Table 9 comparing the $f_T$ and $f_{\text{max}}$ data of a 260 nm gate graphene MOSFET obtained by different de-embedding procedures.

So far, only a little amount of data is available on the RF noise for graphene MOSFETs. In ref. 188 and 189 the noise performance of graphene MOSFETs with gapless large-area channels has been measured up to 8 GHz. To enable a comparison of competing FET technologies with different channel materials and gate lengths in terms of noise, we define the noise-related FOM $M_N$ as

$$M_N = \frac{T_{N_{\text{min}}}}{L}$$

where $L$ is the gate length in µm and $T_{N_{\text{min}}}$ is the minimum noise temperature at a given frequency defined as

$$T_{N_{\text{min}}} = T_0 \left(10^{NF_{\text{min}}/10} - 1 \right)$$

where $T_0$ is the ambient temperature during measurement. It should be noted sometimes that other expressions such as $M_N = T_{N_{\text{min}}}/(f \times L)$ or $NF_{\text{min}}/(f \times L)$ are used as noise-related FOM.188,190 These figures, however, may be misleading since they result in noticeably different numbers of $M_N$ for one and the same transistor at different frequencies. Table 10 shows $M_N$ after eqn (12) for graphene MOSFETs and competing state-of-the-art RF FETs based on noise figures consistently measured at a frequency of 8 GHz. The choice of this frequency is a compromise since it marks the upper bound up to which experimental noise data for graphene MOSFETs are available at present and the lower bound of published experimental noise data for III–V HEMTs which are usually characterized at much higher frequencies.

### Table 9 $f_T$ and $f_{\text{max}}$ of 260 nm gate graphene MOSFETs obtained by different de-embedding procedures

<table>
<thead>
<tr>
<th>Parameter</th>
<th>As measured</th>
<th>Pad de-embedding</th>
<th>Full de-embedding</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_T$ (GHz)</td>
<td>23.6</td>
<td>38.7</td>
<td>198</td>
</tr>
<tr>
<td>$f_{\text{max}}$ (GHz)</td>
<td>6.5</td>
<td>7.6</td>
<td>28.2</td>
</tr>
</tbody>
</table>

### Table 10 Noise performance of graphene MOSFETs and competing RF FET types at a frequency of 8 GHz

<table>
<thead>
<tr>
<th>FET type</th>
<th>$L$ (nm)</th>
<th>$NF_{\text{min}}$ (dB)</th>
<th>$T_{\text{min}}$ (K)</th>
<th>$M_N$ (K µm$^{-1}$)</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphene MOSFET</td>
<td>1000</td>
<td>4.27$^a$</td>
<td>502</td>
<td>502</td>
<td>188</td>
</tr>
<tr>
<td>(gapless channel)</td>
<td>150</td>
<td>4.85$^b$</td>
<td>616</td>
<td>616</td>
<td></td>
</tr>
<tr>
<td>InP HEMT</td>
<td>150</td>
<td>0.14</td>
<td>9.83</td>
<td>7573</td>
<td>189</td>
</tr>
<tr>
<td>GaAs mHEMT</td>
<td>250</td>
<td>0.21</td>
<td>14.9</td>
<td>65.5</td>
<td>191</td>
</tr>
<tr>
<td>GaAs pHEMT</td>
<td>150</td>
<td>0.16</td>
<td>11.26</td>
<td>75.1</td>
<td>193</td>
</tr>
<tr>
<td>Si MOSFET</td>
<td>180</td>
<td>0.96</td>
<td>74.2</td>
<td>412</td>
<td>197</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>0.6</td>
<td>44.4</td>
<td>556</td>
<td>198</td>
</tr>
</tbody>
</table>

$^a$ De-embedded. $^b$ As measured.

As Table 10 shows, the noise performance of graphene transistors significantly lags behind that of InP and GaAs HEMTs, but their $M_N$ is close to that of RF Si MOSFETs.

While so far experimental RF data for GNR MOSFETs have not been reported, their $f_T$ performance has been studied using simulations. An overview of recent simulation activities for graphene RF MOSFETs with both gapless large-area and GNR channels can be found in ref. 199. Taking carrier scattering into account, a cutoff frequency of 5 THz has been simulated for a 10 nm gate transistor with a 10 nm wide GNR channel.200 Without experimental results for RF GNR MOSFETs it is hard to judge this result, particularly since, in simulations, usually idealized device structures and conditions are assumed. Although $f_{\text{max}}$ simulations for GNR MOSFETs are still missing, we expect that due to the gap opening in narrow GNRs a notable improvement of $f_{\text{max}}$ compared to gapless graphene MOSFETs should be possible. On the other hand, achieving an $f_{\text{max}}$ performance of GNR MOSFETs better than that of InP HEMTs and GaAs mHEMTs is rather unlikely owing to the much lower mobility of GNR channels.

An interesting direction in 2D transistor research that concerns transistors with gapless graphene and TMD channels and that recently has attracted considerable interest is the development of MOSFETs on flexible substrates. Graphene and the 2D materials beyond graphene are bendable and can easily be transferred to flexible substrates without seriously affecting the carrier mobility. For example, electron and hole mobilities of 8000 cm$^2$ V$^{-1}$ s$^{-1}$ and 6000 cm$^2$ V$^{-1}$ s$^{-1}$, respectively, have been reported for CVD-grown graphene transferred to polyimide.201 This is orders of magnitude more than the mobility of organic semiconductors which are commonly used for flexible electronics. Although flexible graphene transistors suffer from the missing switch-off as do their counterparts on rigid substrates, they show promise for flexible RF electronics. Fig. 13 summarizes the state-of-the-art RF performance of flexible 2D transistors. Particularly remarkable are the results for a 260 nm gate flexible graphene MOSFET showing an $f_T$ of 198 GHz and an $f_{\text{max}}$ of 28 GHz.187 For comparison, the fastest
from black phosphorus crystals show p-type conductivity and hole mobilities ranging from 286 cm$^2$ V$^{-1}$ s$^{-1}$ (ref. 100) to almost 1000 cm$^2$ V$^{-1}$ s$^{-1}$ (ref. 100). As can be expected from the reasonably wide gap of phosphorene, high on-off ratios of up to $10^8$ have been reported. Particularly interesting are a 300 nm RF phosphorene MOSFET showing an $f_T$ of 12 GHz and an $f_{max}$ of 20 GHz, see Fig. 13, and the first flexible phosphorene MOSFET circuits. Phosphorene transistors have (even for 2D transistor standards) a short history. Nevertheless, remarkable results have been obtained in such a short period of time and further progress in phosphorene transistor processing and performance is expected in the near future. More work is needed, however in order to assess the true potential of these transistors.

In 2015, the first silicene MOSFET was demonstrated. This back-gate device shows an on-off ratio of around 10 only due to the gapless silicene channel.

5.2 X-ane FETs

The current–voltage characteristics of graphene n- and p-channel MOSFETs have been simulated and, assuming ballistic carrier transport, on-off ratios of $10^5$ to $10^6$ have been obtained for a gate voltage swing of 0.8 V. Due to the fact that in graphite (in contrast to other bulk and 2D semiconductors) the electron effective mass is heavier than the hole effective mass, a higher on-current has been predicted for the p-channel MOSFET. Recently the first experimental graphene-like MOSFET has been reported. For this back-gate transistor with a hydrogenated graphene channel having a hydrogen coverage of 25% (compared to 100% for true graphite) and a gap of 4 eV, an on-off ratio of $3 \times 10^8$ has been measured for a gate voltage swing of more than 40 V.

5.3 TMD FETs

The first true 2D TMD transistor has been a 500 nm top-gate MOSFET with an exfoliated single-layer MoS$_2$ channel reported in 2011. Due to the $\approx 2$ eV bandgap of 2D MoS$_2$, this device showed good switch-off and an on-off ratio of $10^8$ for a gate voltage swing of 4 V. Subsequently other groups followed and fabricated MOSFETs with single- and multi-layer MoS$_2$ channels. In most cases exfoliated MoS$_2$ has been used, e.g. with hydrogenation, although also CVD-grown MoS$_2$ has been reported. In addition to single transistors, recently simple circuits with MoS$_2$ MOSFETs have been realized.

Since MoS$_2$ is bendable, it is attractive for flexible electronics. Indeed, MoS$_2$ MOSFETs on flexible substrates have been demonstrated. Due to their relatively low channel mobility, MoS$_2$ MOSFETs show worse $f_T$ performance compared to graphene MOSFETs, see Fig. 13(a). They have, however, the big advantage of a semiconducting channel with a sufficiently wide gap leading to excellent switch-off as well as good drain current saturation and thus reasonable power gain and $f_{max}$. The current record RF performance for MoS$_2$ MOSFETs is defined by a 68 nm gate transistor on a rigid SiO$_2$/Si substrate showing an $f_T$ of 42 GHz and an $f_{max}$ of 50 GHz and its counterpart on a flexible substrate with $f_T = 13.5$ GHz.
and \( f_{\text{max}} = 10.5 \text{ GHz} \). Thus, MoS\(_2\) MOSFETs are very promising for flexible digital and RF applications.

While experimental work on TMD transistors has so far focused on MoS\(_2\) MOSFETs, recently the first single- and/or multi-layer WSe\(_2\), WS\(_2\), MoSe\(_2\), and MoTe\(_2\) MOSFETs have been demonstrated.

On the theoretical side, a lot of work has been done too, particularly to investigate the advantages and drawbacks of TMD MOSFETs for future logic transistor generations and to compare their performance with that of Si MOSFETs. In ref. 220 and 221 it has been shown that in the ballistic limit TMD MOSFETs show slightly higher on-currents for a given off-current, i.e., better on-off ratios, than Si MOSFETs. It has also been shown that among the TMD MOSFETs those with tungsten-based channels show the highest on-currents due to the lower carrier effective mass compared to the Mo-based TMDs.

As we have seen in Fig. 2, logic MOSFETs with gate lengths around 5 nm will be needed in 2028. This target was the motivation for numerous theoretical studies on the behavior of 5 nm gate MOSFETs. It has been shown that in such short channels direct source–drain tunneling becomes a serious issue, and that particularly high-mobility (and thus low-effective-mass) channels will suffer from unacceptably large tunneling currents. This suggests that channel materials with heavier carrier effective mass such as the TMDs would be a viable option. While the tunneling currents in 5 nm TMD channels have been analyzed by simulations, a study comparing the tunneling tendency of 5 nm Si and III–V channels with that of TMD channels is still missing. In the following we provide such a comparison for n-channel MOSFETs employing a simple first-order approximation.

In ref. 19 the source–drain tunneling current in III–V, Ge, and Si NW (nanowire) MOSFETs has been investigated. First, the band structure of the NWs has been calculated, and the bandgap and the carrier effective masses have been extracted (note that the band structure of small diameter NWs is different from that of the corresponding bulk material). These data and a simplified rectangular potential profile in the MOSFET channel have then been used to calculate the source–drain tunneling current in 5 nm NW n-channel MOSFETs. In the present study, we take both the electron effective masses for NWs with 6 nm diameter and the potential profile from ref. 19 and calculate the transmission coefficient \( T_t \) through the barrier according to ref. 225

\[
T_t = \left[ 1 + \frac{E_0^2 \sinh^2(|k|)W}{4E(E_0 - E)} \right]^{-1}
\]

Here \( E_0 \) is the barrier height, \( W \) is the 5 nm barrier width, \( E \) is the electron energy at source, and \( k \) is given by \( \sqrt{2m_{\text{eff}}(E - E_0)}/\hbar \), where \( m_{\text{eff}} \) is the electron tunneling effective mass and \( \hbar \) is the reduced Planck constant. The tunneling current \( I_t \) is proportional to the transmission coefficient and can be estimated using

\[
I_t \approx CM_T^T
\]

where \( c \) is a constant and \( M \) is the number of propagating modes (4 for the Si and Ge NWs and 1 for the III–V NWs). As can be seen from Fig. 14(a), our simple approach reproduces the tunneling currents and the order of the materials from ref. 19 properly. This shows that the transmission coefficient according to eqn (14) is a reasonable measure to assess and compare the source-to-drain tunneling tendency in different materials.

Fig. 14(b) shows the transmission coefficient through a 5 nm wide barrier as a function of the carrier effective mass of the channel material. The transmission coefficients have been calculated using eqn (14) and assuming the potential profile from ref. 19. Indicated by symbols are the transmission coeffi-
Fig. 15 Possible scenario for the selection of the channel materials for logic MOSFETs. sSi means strained Si, \( \mu \) is the carrier mobility, \( m_{\text{eff}} \) is the carrier effective mass, and 2Ds means semiconducting 2D materials.

Table 11 Main advantages (pros), drawbacks (cons), and potential applications of selected 2D channel materials

<table>
<thead>
<tr>
<th>Material class</th>
<th>Material</th>
<th>Main pros</th>
<th>Main cons</th>
<th>Potential for</th>
<th>Not suited for</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-enes</td>
<td>LA graphene</td>
<td>High ( \mu )</td>
<td>Zero gap</td>
<td>flex. mp RF</td>
<td>Logic, hp RF</td>
</tr>
<tr>
<td></td>
<td>GNRs</td>
<td>?</td>
<td>Low ( \mu )</td>
<td>Logic at ( L \leq 5 ) nm</td>
<td>hp RF</td>
</tr>
<tr>
<td></td>
<td>BLG</td>
<td>?</td>
<td>Low ( \mu ), narrow gap</td>
<td>?</td>
<td>Logic</td>
</tr>
<tr>
<td></td>
<td>Silicene</td>
<td>?</td>
<td>Zero gap</td>
<td>?</td>
<td>Logic, hp RF</td>
</tr>
<tr>
<td></td>
<td>Germanene</td>
<td>Reasonable ( \mu )</td>
<td>Zero gap</td>
<td>?</td>
<td>Logic at ( L \leq 5 ) nm</td>
</tr>
<tr>
<td></td>
<td>Phosphorene</td>
<td>Reasonable ( \mu )</td>
<td>?</td>
<td>flex, mp RF &amp; logic</td>
<td>hp RF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hp logic ( L &gt; 5 ) nm</td>
<td>Log</td>
</tr>
<tr>
<td>X-anes</td>
<td>Germanane</td>
<td>High ( \mu ) ( ^* )</td>
<td>?</td>
<td>hp RF</td>
<td>Logic at ( L \leq 5 ) nm</td>
</tr>
<tr>
<td></td>
<td>TMDs</td>
<td>Mo-based</td>
<td>Reasonable ( E_g )</td>
<td>Moderate ( \mu )</td>
<td>flex, mp RF &amp; logic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>W-based</td>
<td>Reasonable ( E_g )</td>
<td>Moderate ( \mu )</td>
<td>hp logic ( L \leq 5 ) nm</td>
</tr>
</tbody>
</table>

\( ^* \) The high mobility of germanene is yet to be confirmed. hp: high performance; mp: medium performance; flex: flexible. A question mark indicates that the advantages, drawbacks, and potential applications are not clear yet.

6. Outlook

Research on 2D materials for electronic applications is a new field as can be seen from the fact that the first graphene, MoS\(_2\), and phosphorene MOSFETs have been reported in 2007, 2011, and 2014, respectively. Given this short history, the achievements made so far are remarkable and further progress is expected. On the other hand, it is extremely difficult to assess the real potential of the 2D materials in electronics. Table 11 lists, based on our current state of knowledge and provided that high-quality 2D layers can be realized in an industrial environment, our expectations on future applications of several 2D materials for MOSFET channels.

Due to their zero gap, the gapless X-enes are not suitable for logic and high-performance RF FETs. The bendability and high mobility make graphene, however, a candidate for flexible medium-performance RF FETs. The semiconducting 2D materials with the exception of germanane suffer from low mobilities and therefore show only a little promise for high-performance logic (as long as source–drain tunneling is not an issue) and RF transistors, while they may find applications in flexible medium-performance logic and RF circuits. Due to their thinness (resulting in short scale lengths) and relatively heavy carrier effective mass (leading to suppressed tunneling)
drain tunneling) they may become an option for sub-5 nm logic FETs. It should be recognized, however, that multiple-gate Si NW transistors will be strong competitors in this field. In case germanane actually shows such a high mobility as predicted in ref. 53, it could become a very interesting material for high-performance logic and RF FETs.

We note that the fascinating variety of 2D materials is both a blessing and a curse – a blessing for researchers for whom the 2D materials are a new and wide field for exciting science, and a curse since the available funding for research spreads across many materials. Thus, sooner or later funding and research activities need to be focused on a limited number of promising 2D materials.

In conclusion we believe that research on 2D materials is just at the beginning and will stay a very exciting field. We are convinced that eventually certain 2D materials will find their applications in electronics, particularly since they can be used not only in transistors but also for other purposes, such as transparent electrodes, sensors, touch screen displays, etc., which are possibly closer to industrial fabrication than 2D transistors.

Acknowledgements

This work was supported by the Excellence Research and the Intra-Faculty Research Grants of Technische Universität Ilmenau, Germany, and by DFG under contract numbers SCHW 729/16-1 and PE 624/11-1. The authors would like to thank A. Alsioufy and M. Lootze for processing MoS2 and WS2 devices as well as F. Bechstedt, T. Heine and A. MacDonald for fruitful discussions and D. Akinwande for providing experimental results prior to its publication.

References

14 G. E. Moore, Dig. ISSCC, 2003, 20–23.


