

The Potential and Challenges of Two-Dimensional Materials in Sub-5 nm CMOS Devices

Tingan Wang

WLSA Shanghai Academy,
Shanghai, China
Corresponding author:
tinganwang618@outlook.com

Abstract:

The relentless scaling of CMOS transistors below 5 nm has pushed silicon to its fundamental limits, resulting in severe short-channel effects, degraded carrier mobility, and excessive leakage. Two-dimensional (2D) materials—graphene, molybdenum disulfide (MoS₂), tungsten disulfide (WS₂), and black phosphorus (BP)—have emerged as attractive alternatives to silicon channel materials due to their atomically thin bodies, favorable carrier mobility, and excellent electrostatic controllability. In this paper, we carefully review the capability of these 2D materials to serve as channel materials for sub-5 nm CMOS devices in terms of transport characteristics, bandgap, scalability, and integration. We critically compare the distinct advantages and associated challenges of each material, including contact resistance, fabrication scalability, environmental stability, and gate dielectric compatibility. Our analysis concludes that although 2D materials offer clear promise for extending Moore's Law beyond the silicon era, extensive research and engineering are still required to address the remaining issues before they can be feasibly adopted in future CMOS technology.

Keywords: 2D materials; CMOS; sub-5 nm node.

1. Introduction

For half a century, Moore's Law has been leading to unprecedented scaled-down transistor dimensions and density enhancements in each generation of integrated circuits [1, 2]. With the upcoming sub-5 nm technology node, however, the scaling of silicon-based CMOS is facing serious fundamental limits. Under extreme dimensions, traditional silicon channel conductors are greatly deteriorated by short-channel effects (SCEs), diminishing the effec-

tiveness of the gate control and results in high off-state leakage and sharp subthreshold swing. Quantum mechanical tunneling from source to drain becomes a significant concern at short gate lengths, i.e. below ~10 nm, leading to high offstate currents in silicon transistors. In addition, lowering the Si body thickness (in ultra-thin body, or nanowire devices) leads to decreased carrier mobility, e.g. due to surface roughness, as well as quantum confinement effects, leading to more difficult drive currents as a function of size.

These problems suggest that single silicon may not be able to satisfy power and performance of sub-5 nm nodes. As such researchers are investigating on new channel materials and device architecture to achieve scaling. This has led to increased attention on two-dimensional (2D) materials, as promising substitute to silicon for nanoscale transistor [1,2]. 2D materials are crystals of single-or few-atom-thick crystal layers with strong in-plane bonding but weak van der Waals interlayers interaction. These are graphene, based on carbon, transition metal dichalcogenides (TMDs). This family includes carbon-based graphene, transition metal dichalcogenides (TMDs) like MoS₂ and WS₂, and black phosphorus (a layered allotrope of phosphorus). With their atomic thinness, they afford the ultimate scaling advantage in that a 2D monolayer can operate as a channel that is only 0.3–0.7 nm thick, leading to good electrostatics gate control and effective short-channel effect suppression [2]. In fact, both theory and simulation suggest that a sub-nm 2D channel can achieve sub-70 mV/decade subthreshold swing even at extremely short length scales compared to bulk silicon which would lead to bad leakage [3]. Additionally, charge transport in multiple 2D materials actually does not decline when the thickness of the materials decrease— 2D semiconductors in fact maintain relatively high carrier mobilities as their thicknesses are reduced to monolayer due to the defect free surface without dangling bonds [2]. This is in clear contrast to silicon where mobility decreases drastically when the channel is confined to several nanometers due to surface scattering.

Of equal significance, 2D materials provide the possibility of having diverse intrinsic electronic properties that would be potentially valuable in advanced CMOS. As graphene is a semimetal with ultrahigh carrier mobility (more than 10^5 cm²/V·s at room temperature for encapsulated material [4]), it offers fast switching and high operating frequency. The TMDs (MoS₂ and WS₂) are intrinsic semiconductors with 1– 2 eV bandgaps [4], or larger than silicon's,

that can be used in low-leakage digital transistors. Black phosphorus (in the few-layer form, aka. phosphorene) is a semiconductor with thickness dependent bandgap (~0.3 eV in bulk increasing to ~1.5–2.0 eV in monolayer) and high hole mobility up to ~1000 cm²/V·s for ~10 nm thin films [5]. These various properties provide opportunities for 2D materials not just to substitute silicon in logic circuits, but also to provide new functionalities (such as anisotropic conduction of BP or high-frequency capabilities of graphene) in future chips.

In this work, we compare the four prominent 2D material candidates; graphene, MoS₂, WS₂, and black phosphorus to be suitable channel materials of sub-5 nm CMOS devices. We first compare major material properties related to the transistors performance, which are carrier mobility, the bandgap, electrostatic gate control, scalability, and the environmental stability. We next present the potential opportunities of each material towards enhancing/continuing CMOS as well as a detailed review of the challenges (one per material, and one per 2D material) that need to be met to integrate the materials into devices that can be used in practice. Important concepts, such as better electrostatics of the 2D channel and device designs utilizing 2D materials, will be illustrated through figures and tables. In this review, we hope that readers will understand the potential of 2D material to make up the limits of silicon for devices nanometer scale and know research direction to realize 2D material-based CMOS technology.

2. Comparative Material Properties

In this section, we contrast key graphene, MoS₂, WS₂, and black phosphorous properties relevant to sub-5 nm CMOS applications. Material parameters that play major roles in determining the feasibility as a transistor channel of each of the above materials are carrier mobility, bandgap, electrostatic gate control, scalability, and environmental stability and are highlighted in Table 1.

Table 1. Comparative properties of candidate 2D channel materials (graphene, MoS₂, WS₂, black phosphorus) relevant to sub-5 nm CMOS devices.

Material	Carrier Mobility (cm ² /V·s)	Bandgap (eV)	Electrostatic Control	Scalability (Synthesis)	Environmental Stability
Graphene	~200,000 (intrinsic mobility in ideal samples) [4]; typically 10,000–40,000 on SiO ₂ substrates at RT	0 (semimetal)	Excellent gate control due to atomic thinness; however, zero bandgap causes poor turn-off (leakage current remains high) [4]	Large-area CVD growth demonstrated (transfer to wafers required); difficult to pattern at <5 nm dimensions without inducing a bandgap	Chemically inert and stable in ambient conditions (graphitic structure does not oxidize easily)

MoS ₂	~200–500 (monolayer, with high-κ dielectric engineering) [2]; up to ~700 in multilayer with optimized contacts [6]	~1.8 (monolayer direct gap) [2]; ~1.2 (bulk indirect gap)	Strong gate electrostatic control even at monolayer thickness (suppresses SCEs); intrinsic bandgap allows transistor to turn off, yielding high on/off ratios ~10 ⁸ [2]	Wafer-scale growth possible by CVD/MOCVD (demonstrated on up to 8" wafers) but with grain boundaries and defects; still maturing for uniform device fabrication	Stable in ambient (chemically robust sulfur surface); minor oxidation at high temperature or harsh environment, but generally not prone to rapid degradation
WS ₂	~50–200 (monolayer values in literature; lower than MoS ₂ due to heavier effective mass) [7]	~2.1 (monolayer direct gap)	Comparable electrostatic control to MoS ₂ (atomically thin with sizable bandgap); robust gate control in scaled transistors demonstrated in simulations [9]	Wafer-scale synthesis similar to MoS ₂ (CVD growth reported), but large-area quality and uniformity need improvement; less studied than MoS ₂ for mass production	Stable in ambient (similar to MoS ₂ , no rapid ambient degradation); good chemical stability of the WS ₂ layer in air
Black Phosphorus (BP)	~1000 (for ~10 nm thick flakes at RT) [5]; can exceed 200–500 in few-layer (~5 nm) devices; highly anisotropic (higher mobility in one in-plane direction)	~0.3 (bulk); 0.5 – 0.8 (few-layer); up to ~1.5–2.0 (monolayer) [8] (direct gap, thickness-dependent)	Good electrostatic control when thinned to few-layer (~1 nm) thickness; supports high on/off ratios (~10 ⁵) in <10 nm thick FETs [5]. However, bandgap decreases with thickness, so thicker flakes trade off electrostatics for mobility.	No mature method for wafer-scale BP film growth (flakes obtained by exfoliation or chemical synthesis in small batches); scalability is poor – integrating BP into large-scale manufacturing remains challenging	Poor stability: Degrades rapidly in ambient air due to oxidation [8] (monolayers oxidize within minutes, ~10 nm flakes within hours). Requires encapsulation (e.g. with hBN or Al ₂ O ₃) to stabilize in devices.

Figure 1 below illustrates one of the key advantages of 2D semiconductor channels – their superior electrostatic control at extremely scaled lengths – by comparing the subthreshold swing (SS) behavior of various transistor geometries.

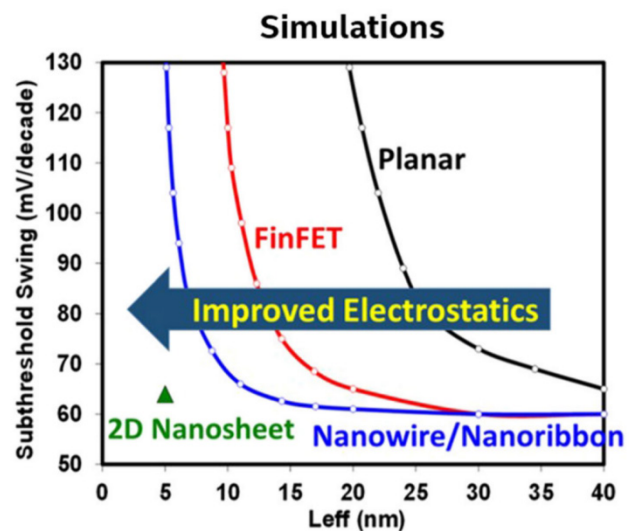


Fig. 1 Comparison of Subthreshold Swing Versus Channel Length for Various Transistor Architectures[3]

Figure 1 shows the simulated subthreshold swing (SS) as a

function of channel length for several transistor structures, including a planar bulk transistor, a FinFET, a nanoribbon (nanowire) FET, and a 2D material monolayer nanosheet FET [3]. At very short gate lengths (far right of the plot), the 2D nanosheet device maintains a much lower SS than the silicon-based devices, indicating better control of the channel and lower leakage [3]. In essence, as gate length scales into the sub-10 nm regime, traditional 3D structures suffer from increased leakage (SS rises sharply) due to direct source–drain tunneling, whereas a monolayer TMD channel can sustain subthreshold swings closer to the ideal 60 mV/decade limit. This highlights the fundamental promise of 2D materials in enabling effective transistor switching at dimensions where silicon would cease to turn off properly.

Data for the materials of table 1 illustrates that each 2D material has different trade-off of properties. Graphene provides extraordinary carrier mobility with ultra-atom thickness, but its zero bandgap precludes the possibility of switching current off except bandgap engineering. MoS₂ and WS₂ have adequate bandgaps to give decent on/off ratios and low offstate leakage. In practice MoS₂ shows higher mobility than WS₂ in practice (due in part to more mature research and possibly lighter effective carrier mass), but both TMDs have considerably lower intrinsic mobilities than graphene. Black phosphorus is an intermediate-bandgap material with mobility values between graphene and MoS₂ and has the notable advantage of being able to support greater drive currents than TMDs on account of its larger mobility, yet the problem of stability is more serious here. The four materials are all atomically thin, providing exquisite electrostatic gate modulation; each would eliminate short-channel effects in a sub-5 nm transistor geometry. Their scalability and integration challenges differ significantly. Graphene and TMDs can already be produced at large scale, though further improvement is needed, whereas black phosphorus remains difficult to prepare at the wafer level and is further limited by environmental instability. The following sections analyze in detail the opportunities that these 2D materials bring to advanced CMOS technology and the specific challenges that each material must overcome.

3. Opportunities of 2D Materials

There are two- opportunities where 2D materials offer significant potential to push CMOS performance and scaling beyond silicon. In this section, we highlight each of the 4 candidate materials – graphene, MoS₂, WS₂, and black phosphorus – as being suitable for sub- 5 nm node transistors and advanced CMOS architectures. We describe how each of the 4 materials present opportunities in areas

and specific advantages that could be exploited in future devices.

3.1 Graphene

Although graphene is not a semiconductor in the strict sense (due to its gapless band structure), it exhibits extraordinary carrier transport properties that would have been desirable in certain application scenarios. The carrier mobility in graphene can be as high as $10^5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ (orders of magnitude higher than Si's mobility) in high-quality samples [4], which allows high-density drive current to flow through the graphene channels and high-frequency operation is also achieved. Even though graphene FET's cutoff frequency has been pushed into the hundreds of GHz even at short gate length [4], it points toward the potential of RF and analog applications. For logical computation, we require a bandgap but the ultra-thin thickness of graphene allows for near-ideal electrostatic control, i.e., if a bandgap can be engineered (e.g. by patterning graphene into nanoribbons or by applying a bilayer perpendicular electric field), the device could be switched with negligible short-channel effects. Moreover, graphene's toughness and chemical stability are appealing for integration – it can handle high current densities ($>10^8 \text{ A/cm}^2$) and it could be used as a future interconnect material or the channel of a tunneling device. Its two-dimensional nature also allows for stacking several layers of graphene, with insulators, for three-dimensional integration. Finally, in a nutshell, the opportunity of graphene is its speed and its conductance – it may realize transistors with remarkably fast switching or be used in applications where the carrier velocity is very high (e.g. for THz electronics), provided that the absence of bandgap is solved in the device.

3.2 Molybdenum Disulfide (MoS₂)

MoS₂ is amongst the most used 2D semiconductors for electronics and was the first TMD shown to demonstrate a transistor effect at monolayer thickness [2]. Its major benefit for CMOS is that monolayer MoS₂ has a direct bandgap $\sim 1.8 \text{ eV}$ [2], which is sufficiently large to ensure low off-state leakage and high on/off ratios in transistors. In fact, MoS₂ FETs have reported on/off current ratios of the order of 10^8 at room temperature [2] that is several orders of magnitude higher than achievable by graphene FETs without a bandgap. Thus, MoS₂ is suitable for digital logic switches. Moreover, monolayer MoS₂ ($\sim 0.65 \text{ nm}$ thick) can be scaled with reasonable carrier mobility. Early devices exhibit low mobility ($\sim 1\text{--}10 \text{ cm}^2/\text{V}\cdot\text{s}$), which was improved by engineering the gate dielectric (e.g., using high-k oxides such as HfO₂) and substrate optimization to get the room temperature mobility of monolayer MoS₂

up to $\sim 200 \text{ cm}^2/\text{V}\cdot\text{s}$ [2]. Mobilities significantly higher ($\sim 700 \text{ cm}^2/\text{V}\cdot\text{s}$) were obtained for multi-layer MoS_2 (\sim few nanometers thick), e.g., by applying a 10 nm MoS_2 flake with optimized Scandium contact [3]. These mobility values, although lower compared to silicon at high carrier densities, are adequate for CMOS applications given the material's ultra-thin geometry and electrical strength. Another potential benefit of MoS_2 is that it can interface with current CMOS processes – i.e., that it can survive the integration of high- κ dielectric and that it has been made into 300 mm wafer devices (via transfer or by direct growth) in experiment. Additionally, MoS_2 transistors have been demonstrated functioning to gate length down to the ~ 5 nm gate length with sufficient performance, which is not reachable with Si due to excessive leakage [3]. Therefore, MoS_2 provides a silicon-like semiconducting channel in atomically thin manner with capability of continuous downscaling of gate length and device footprint while short-channel effects under control. The monolayer material is thus seen as one of the top candidates for a 2D-material CMOS node which can be further extended to a “monolayer electronics” era.

3.3 Tungsten Disulfide (WS_2)

WS_2 is a TMD close cousin of MoS_2 and also has a direct bandgap in monolayer (~ 2.1 eV) that is favorable for transistor operation. Although the work on WS_2 slightly trails MoS_2 , there is some opportunity that is not present in MoS_2 . For instance, one study reported that monolayer WS_2 FETs can show superior performance than MoS_2 at comparable channel length [9]. WS_2 has comparable ultra-thin structure and therefore shares the good gate control benefit with MoS_2 . Its slightly lower bandgap can be used to further lower off-state leakage current, which is advantageous for low-power electronics. While measured mobilities in WS_2 devices have been modest, usually tens of $\text{cm}^2/\text{V}\cdot\text{s}$ [7], there is opportunity for improvement by optimizing contact engineering and optimizing the dielectric environment, as for MoS_2 . WS_2 is also found to form high-quality interfaces with high- κ dielectrics and to grow by CVD and there have been some initial results indicating large-area WS_2 can be generated and this opens the possibility of wafer-scale devices. WS_2 may present an opportunity unique to it, and possibly to some W-based TMDs, because these materials can be operated in the p-type regime, since the valence band of WS_2 is at a similar energy as some metal work functions, potentially allowing for hole injection with proper contact metal. Although MoS_2 transistors are typically n-channel, p-type transistors of WS_2 (as well as the related material WSe_2) can also be fabricated. Overall, the electrostatics of WS_2

is strong and the band structure of the material can be very flexible, and it could complement the MoS_2 material for building all-2D material logic that uses both n- and p-channel devices. Its challenge is to approach the same level of performance of MoS_2 with ongoing investigations for enlarging the set of 2D semiconductors for device applications.

3.4 Black Phosphorous (BP)

Black phosphorous is distinguishable from graphene and TMDs by its large intrinsic carrier mobility and tunable bandgap. Few-layer BP, a direct-bandgap semiconductor, exhibited mobilities of $\sim 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ on ~ 10 nm thick films at room temperature [5], which is two orders of magnitude higher than those seen in most monolayer TMDs. Even monolayer or bilayer phosphorene can achieve mobilities of several hundred $\text{cm}^2/\text{V}\cdot\text{s}$ and BP's mobility is strongly anisotropic (carriers travel more easily in the armchair direction of its crystal lattice) – its anisotropy could be exploited in device engineering (e.g. alignment of transistors to achieve optimum performance) [8]. BP's bandgap, ≈ 0.3 eV in bulk with increasing to ~ 2 eV in monolayer, means one can select the layer number to achieve a desired bandgap. In practice, few-layer thickness (5–10 nm) induces a moderate bandgap (~ 0.5 eV) that balances ON current and OFF leakage in FETs. FETs fabricated with ~ 5 – 6 nm BP exhibit on/off ratio 10^5 and good current saturation. The strength with BP is being able to offer higher drive currents than MoS_2/WS_2 at comparable channel length due to its larger mobility and reasonable bandgap. This makes BP very appealing in high-performance logic transistors and possibly also analog devices. Further, because BP's spectral bandgap tunability (through thickness or strain) allows it to serve useful optoelectronic applications (photodetectors, etc.) integrated on chip, spanning the visible to mid-infrared range. Overall, BP can provide the combination of high mobility and electrically controlled bandgap, making it have high potential to be used in high-speed and high-drive current devices in scaled CMOS technologies. Provided stability issues can be addressed, BP can beat TMDs in some performance characteristics and add functionalities (e.g., anisotropic transport, broadband optoelectronics) to advanced electronics.

Beyond the individual materials, 2D materials collectively offer some general opportunities for CMOS integration. One such opportunity is the ability to stack multiple channels vertically in a gate-all-around geometry. Since 2D materials can be layered without covalent bonds (using van der Waals stacking), it is feasible to create a multilayer channel structure where several monolayer ribbons are

placed on top of each other, separated by thin dielectrics. This is effectively a nanosheet GAA transistor with multiple 2D sheets. Such a structure can increase drive current per footprint area. In fact, a recent study illustrated that in a given vertical space, one can stack more 2D material sheets than silicon nanowires (for example, six monolayer

ribbons versus four Si nanoribbons in the same height), meaning a 2D-channel transistor could achieve equal or greater total drive current even if each individual 2D sheet carries somewhat less current than a thicker Si nanowire [10]. Figure 2 demonstrates this concept of vertical stacking advantage for 2D materials.

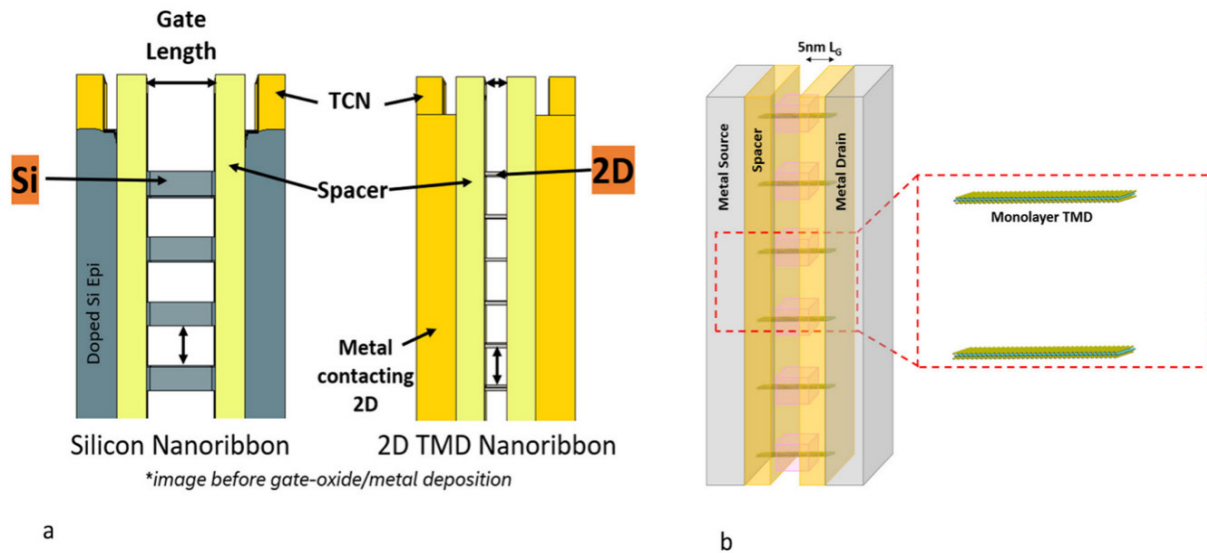


Fig. 2 (a) Silicon GAA nanoribbon transistor structure. (b) 2D TMD GAA nanoribbon transistor with higher stacking density [10]

Figure 2 demonstrates schematic cross-sections of stacked gate-all-around (GAA) nanoribbon transistors of (a) silicon and (b) 2D material (TMD) channels at the same overall stack height. In (a) four Si nanoribbons (gray) are vertically stacked, delimited by dielectric spacers (yellow) and gates oxide (pink). In (b) due to the atomic level thickness of 2D TMD sheets six nanoribbon channels fit in the same height as (a) [10]. Consequently, while each 2D ribbon may require a lower drive current than a Si nanoribbon, overall drive current of the six layer 2D stack can reach or surpass the drive current of the four layer Si stack [10]. This demonstrates the unique potential of 2D materials. They can make possible new device architectures (e.g., densely stacked nanosheets) that increase performance density greater than what is realized using thicker semiconductor channels. [10]

Another general opportunity is the incorporation of 2D materials in tunneling field-effect transistors (TFETs) and other steep-slope devices for ultra-low-power logic. The direct and moderate bandgaps of TMDs and BP are suitable for band-to-band tunneling; indeed, monolayer MoS₂ has been proposed for TFETs to achieve sub-60 mV/dec swings [2]. 2D materials' atomically sharp junctions (when stacked heterojunctions are made) and their ability to tolerate high doping without much short-channel tunneling

make them promising for novel device concepts that could drastically reduce power consumption.

To conclude, while each 2D material offers its own benefits, e.g., graphene's highspeed mobility, atomic-scale-well-behaved MoS₂ switches, good options for digital complementary devices and low leakages in WS₂, or high current drive and greater versatility in black phosphorus, combined they provide routes to keep extending CMOS scaling (e.g., ultra-thin channel and stacked architectures) and to enhance performance (in terms of higher mobilities, or new device physics) in the sub-5 nm regime. The second equally important aspect is introduced in the next section, which examines the obstacles that must be addressed for each material to realize its potential in advanced CMOS-grade technologies.

4. Challenges of 2D Materials

4.1 Major Challenges

While there is large potential for 2D materials in nanoelectronics, there are still major challenges ahead if they are to replace or complement silicon in mainstream CMOS. Here, we analyze the materials-specific challenges of each of the four materials - graphene, MoS₂,

WS₂ and black phosphorous - as well as more general integration challenges. Among the challenges is intrinsic material limitation (for example graphene doesn't have a bandgap), device level issue (making low resistance contacts to TMDs), or manufacturing issues (for example, wafer scale growth and stability of these materials). Table 2 summarizes the outstanding challenges for each material and techniques for overcoming them.

4.1.1 Graphene

Due to a lack of bandgap, graphene is unable to achieve an off state, causing extremely low on/off ratios in FETs [4]. Bandgap engineering, such as nanoribbons, bilayer bias, or chemical functionalization, allows for digital logic, but often degrades the mobility [4]. Contacting and doping become complicated since the Dirac point results in Fermi level that cannot pin contacts, and controlling the threshold voltage will not be easy without a bandgap. Integrating with high- κ dielectrics is limited due a lack of dangling bonds on the inert surface of graphene, while ALD of a gate oxide usually uses seeding techniques that introduce defects. Nanoribbon or quantum confinement patterns graphene into <10 nm ribbons to create a bandgap (~0.1-0.2 eV) between the conduction and valence bands, however this comes with a tradeoff that drops mobility down to only 200 cm²/V·s, which competitive with silicon [4]. Bilayer graphene with a vertical field applies a perpendicular electric field to bilayer graphene to create a tunable bandgap, including bilayer graphene that requires applying a high bias and with precise control done [4]. Chemical functionalization is the addition of adatoms or covalent bonds that break symmetry and open a gap, but at the possible cost of degrading carrier transport. Contact engineering uses 2D metals or edge contacts to graphene to reduce access resistance and can dope the contact regions with charge transfer (for example, depositing metals like Ti that donate electrons). Oxide deposition involves using seeding layers (e.g., ~1 nm of Al to facilitate Al₂O₃ growth) or remote plasma ALD to add high- κ materials to graphene without altering it significantly.

4.1.2 MoS₂

Carrier mobility and drive current limitations are significant hurdles as mobility for the monolayer (~100-200 cm²/V·s) is lower, as compared to silicon, which limits the on-current [2]. While multilayer structures will improve mobility, the trade-off is worsened electrostatics. The contact resistance is rooted from Schottky barriers at the metal/MoS₂ interface attributed to the Fermi level pinning toward the conduction band, and result in large contact resistances and lower drive current [6]. Growth scalability is still an issue, as difficulty arises when growing uniform

monolayer MoS₂ over full wafers. CVD grown MoS₂, provides difficulty with not only having grain boundaries, but also thickness variations. Interface traps at the dielectric/MoS₂ interface will create localized trap states that will scatter carriers and provide instability in threshold voltage. Finally, the absence of p-type devices is still an obstacle since MoS₂ is naturally n-type, stable p-channel behavior, for specification of CMOS duality, would require extensive doping, supplementation with complementary materials such as WSe₂ are used to achieve this. Using low-workfunction metals (Sc, Ti) is an example of contact engineering where both yielded demonstrated ohmic-like contacts achieved mobility ~700 cm²/V·s [6]. Contacts that are phase engineered to convert the 2H phase into 1T phase locally can also reduce barriers. Two-dimensional metallic contacts can be achieved by using graphene or a metallic 1T'-phase TMD interlayer at contacts to avoid the Schottky barrier. Mobility improvements can be achieved through a high- κ dielectric environment like HfO₂ top gates to screen impurities [2] and dual-gate or local gating configurations for improved electrostatics and carrier injection. Small increases in mobility in MoS₂ from strain engineering are also possible. Further improves in atomic layer or molecular beam epitaxy for MoS₂ synthesis are yielding even larger grain sizes, while seeding techniques and more qualified precursors are leading to lower defect density. Post-growth grain stitching or selective area growth is also being considered to create uniform films. In CMOS logic some p-type paths continue to explore electrostatic doping through a back-gate as a global p-bias, while using another 2D material like WSe₂ for a p-MOSFET with an n-MOSFET MoS₂. Integration of two TMDs remains a frontier in study.

4.1.3 WS₂

A key limitation is lower mobility because the mobility of WS₂ monolayer, usually a few tens of cm²/V·s, is lower than that of MoS₂ and could limit drive current [7]. The contact issues are the same as those previously described with MoS₂, where contacts can form Schottky barriers. Also, there is little data on the optimized contacts for WS₂, so the effect of contact resistance has hardly been explored yet. Another concern is material uniformity; while WS₂ films can be uniform, thickness non-uniformity often occurs. Since WS₂ is less studied than MoS₂, many methods for defects in WS₂ are still basic and still being evaluated. P-type operation control is difficult; even though WS₂ can be made p-type in principle, it is not a straightforward task to make reliable p-channel WS₂ devices without suitable contacts and doping. Methods similar to those for MoS₂ are used, specifically, either lower-workfunction metal contacts or 2D edge contacts to optimize and reduce con-

tact resistance. For example, early experiments with WS₂ transistors used Ti/Au contact metal combinations and had moderate success, and current work is continuing to look to metals like Scandium following the MoS₂ example. Enhancing mobility relies on ensuring high crystal quality by utilizing better CVD or epitaxial techniques on the oriented substrate. Since encapsulation of 2D materials, such as WS₂ with hBN, can lessen scattering and enhance mobility similar to other 2D materials, there is justifiable hope that this filtration technology may help WS₂ to eventually achieve better mobility. Alloys of WS₂ with MoS₂ creating Mo_{1-x}W_xS₂ may also allow for the combination of benefits of each material and better mobility overall. In terms of CVD growth of WS₂, wafer-scale growth efforts are increasing, with some promising routes to achieve monolayers uniformly across growth substrates, which represent a more involved sulfurization approach or utilizing seeding promoters to enhance uniformities. There has not yet been the same level of maturity around growth with WS₂ as seen with MoS₂. Studies of device architectures suggest that WS₂ may demonstrate more advantageous performance in much shorter channels, as some simulations appear to suggest [9]. Therefore, if WS₂ is going to be useful in a device, it may benefit utilization in fewer layers in a multilayer stacked-channel GAA structures, such as viewed in Figure 2, to compensate for the per layer reduced drive current capability. Using MoS₂ and WS₂ matched together as p-type and n-type channels, respectively, could also allow for fully 2D CMOS circuits.

4.1.4 Black Phosphorus

Environmental instability presents challenges for black phosphorus (BP) since it is very sensitive to oxygen, moisture, and light, with even thin layers degrading in air from minutes to hours [8]. Device performance degrades rapidly unless steps are taken toward encapsulation. Scalability is poor because there is no reliable growth method for large areas; devices rely on exfoliated crystals, which is incompatible with VLSI. BP generally exhibits p-type character at contacts of Au or Pd because of work-function matching, while n-type operation requires either low-workfunction metals or electrostatic doping, both of which are challenging to obtain. Making an Ohmic contact for both transport is currently problematic. Thickness is also a very important variable as BP's bandgap is dependent on the number of layers. Bulk BP (>10 nm) with a small bandgap (~0.3–0.4 eV) has high off-current states, while thin BP has a good off-state but reduced mobility, relying on balance per device. Integration with dielectrics is an additional problem, since atomic layer deposition methods can oxidize BP; deposition must take place in an inert or benign manner, as interface traps and charges

will shift the threshold voltage. Encapsulating with hBN or by using Al₂O₃ self assembled by atomic layer deposition immediately after exfoliation is always the preferred approach to protect BP [8]. Encapsulated devices operated stably for months. Other coatings or functionalization slow down degradation [11], but h-BN performs the best [8]. BP can only be fabricated in a glovebox or vacuum environment, which makes processing be cumbersome, although vacuum transfer systems may be developed in the future. Early synthesis attempts include gas-phase transport, epitaxy, or red-phosphorus conversion [12] as the basis for synthesis, but, again, processing has not yet reached wafer scale. For p-type material, Au contacts have been reported with hole mobility close to 1000 cm²/V·s [5], and for n-type, Mg or Ti contacts and top or dual gates can induce conduction. Hybrid 2D/3D contacts such as graphene electrodes potentially lower resistance without damaging BP. Utilizing ~5 nm BP material with ~0.6–0.7 eV bandgap helps balance mobility and leakage, while multigate designs can further limit off-current. BP is a good candidate for future analog or mixed-signal circuits due to its anisotropy and moderate bandgap.

4.2 General Challenges

Beyond the above-mentioned material-specific problems listed above, there are general challenges that would face all 2D material for integration into the sub-5 nm CMOS process.

4.2.1 Wafer-Scale integration

In contrast to silicon, which can be grown into arbitrarily large single crystals and oxidized to obtain an atomically clean interface, 2D materials have to be put or grown onto substrates to be compatible with fabrication. Flakes cannot be transferred for fabrication purposes, and one has to grow them on a wafer via chemical vapor deposition (CVD) or some other approach. Consistently obtaining monolayer coverage with controlled layer number and low defect density over a whole wafer is a persistent challenge. As an instance, for CVD-grown MoS₂, the inclusion of grain boundaries can still degrade consistency amongst devices. Ability to align crystallographic direction is important if multiple 2D layers are to be stacked, or if anisotropy (as with BP) is to be aligned. Note that researchers are also working towards developing methods for growing 2D materials directly on dielectric substrates or metal foils and then transfer [7, 8], wherein the defects and contamination of transfer can erode the performance. A critical hurdle to be crossed for realizing a scalable synthesis of high-quality 2D semiconductors is an effective, facile exfoliation [9, 10].

4.2.2 Contacts and interconnects

Contact resistance is an ever-present concern for all 2D materials (certainly not except graphene itself, that is semi-metallic). At channel-lengths below 5 nm, contact resistances of a few hundred ohms· μm become dominant for device behavior. Novel contact schemes are under exploration, e.g., one-dimensional (edge) contacts using a metal spanning the open edge of a 2D sheet to contact all layers or employing conductive 2D material as the interfacial layer to reduce disruption of the channel. Another consideration is contact compatibility with semiconductor processing. For example, low-WF metal Sc is highly reactive, and unlikely to be used for CMOS fabs, necessitating alternatives or encapsulation. However, besides the source/drain contacts, how to connect the 2D materials with existing BEOL (back-end-of-line) interconnects is also a problem – although it has been proposed that graphene may be used as a substitute for Cu interconnect due to the conductivity, making reliable via contact for graphene and metal is still a problem.

4.2.3 Gate Dielectrics and Interface Quality

Creating a gate oxide of high quality remains challenging on air-stable 2D materials such as MoS₂ or WS₂, although there has been some advancement in this area. ALD HfO₂ or Al₂O₃ will also deposit with proper seeding, such as atomic layer deposition following ozone pre-treatment to facilitate nucleation on MoS₂, though such steps lead to interfacial traps that affect threshold voltage and sub-threshold swing, even when they may not be apparent for the first stages of atomic layer deposition. For both graphene and BP, the problem of oxide growth is even more difficult, as oxide growth on top of graphene could damage graphene and BP oxidizes with low bias. Another route is to grow a thin insulating 2D layer, such as hBN, as gate dielectric, but hBN has a low dielectric constant of ~ 4 and is also limited in scalable area. This makes integration more complex. Therefore high- κ /2D interfaces are actively studied, for example deposition in low temperature, plasma-free ALD or even formation of an oxide in situ (i.e. oxidize a layer of metal evaporated on top of the 2D material). A scalable and low defect gate stack is essential for 2D FET performance and reliability.

4.2.4 Device Architecture and Variability

To maximally utilize 2D materials, new device architectures are desired (e.g., GAA nanosheets, see Figure 2). The architectures bring new hurdles to manufacturing themselves – e.g., how to etch and pattern multiple stacked monolayers homogeneously, how to isolate the devices, etc. The device parameter variability introduced by material defects/non-uniformities could also be a con-

cern. Flake size and thickness, and even grain boundaries and strain can introduce threshold voltage variations and mobility variation. For use in digital circuits, yield demands minimization of variability which may necessitate approaches like doping or stitching together, statistically speaking, flake-to-flake. The variability problem is exacerbated by relative immaturity of 2D material growth in contrast with silicon. It has been observed that currently the device-to-device variations are large for 2D transistors, but this should decrease once as the quality of the material increases.

4.2.5 Thermal management

With aggressively scaled devices, power density may be high. 2D materials, being atomically thin, have alternative heat dissipation paths. Graphene has good in-plane thermal conductivity, but when used as a channel under a gate stack heat must flow out-of-plane into the substrate or along the channel to contacts. Compared to graphene, thermal conductivity of TMDs are lower, and layers such as MoS₂ may be at risk of self-heating if the thermal management is not adequate, in addition the presence of poor thermal contact with underlying substrates is due to the van der Waals gap blocking the heat transfer. This phenomenon may need to be taken into consideration by device design through the addition of heat-spreading layers [2] or a more optimal geometry for ensuring heat removal [3].

In summary, while 2D materials face significant challenges – including the lack of a bandgap in graphene, relatively low mobility and contact issues in TMDs, and severe instability and scalability problems for BP – substantial research efforts are addressing these issues. Several solutions are starting to show promise or have even shown promise on a small scale, including bandgap tuning of graphene, better contacts to significantly improve MoS₂ performance [6], encapsulation methods for the stabilization of BP, and new 2D heterostructures to enable functional devices. However, the realization of these 2D materials in commercial technology remains uncertain, as the challenges facing wafer scale growth and device uniformity will take time and perhaps new manufacturing methodologies to overcome. The next section concludes the paper by reflecting on the outlook for 2D materials in future nanoelectronics.

5. Conclusion

Two-dimensional materials provide a potential avenue for continuing to scale CMOS technology to and beyond the 5 nm node. Their atomic thickness allows for advantageous electrostatic-control sensitivity that suppresses

short-channel effects that are a constraint for traditional silicon devices. Each of the two-dimensional materials—graphene, molybdenum disulfide (MoS_2), tungsten disulfide (WS_2), and black phosphorus—has its own set of benefits: graphene has astonishingly high carrier mobility; MoS_2 and WS_2 provide intrinsic bandgaps that can be manipulated for logic applications; and black phosphorus has high mobility and a tunable bandgap. These characteristics allow for the eventual realization of transistors with low leakage currents, subthreshold slopes as steep as theoretical limits permit, and possibly, new 3D integration schemes enabled by densely stacked nanosheet channels that are the nontrivial extension of traditional silicon.

Nevertheless, using these materials for sub-5 nm CMOS offers basic challenges. Graphene will require artificial bandgap schemes developed for these materials to switch effectively for the sake of logic presentation. Transition metal dichalcogenides (TMDs) will require improved contact resistance and charge transport schemes. Black phosphorus has limitations in stability and production schemes for large areas and scales, similarly to TMDs. Finally, creating these materials at scale requires a method to produce large, uniform 2D films with minimal point defects, and then a fabrication method that is reliable and repeatable for the least amount of variability.

Despite a large amount of research over the past ten years improving TMD transistors and even building demonstrable prototype 2D CMOS circuits, this is still some distance from becoming marketable. Continued improvement in several areas is expected over the next several years, including synthesis, hybrid integration with silicon, and understanding the atomic scale physics when you create chargers and devices. Ultimately, if these technical barriers are overcome, two-dimensional materials would be a mechanism for advancing Moore's Law by creating smaller, faster, and more energy-efficient electronic devices. Whether they remain a research field away from becoming a more mainstream component of CMOS technology will depend on continuing to work on materials engineering and the fabrication of 2D devices.

References

- [1] Schwierz, F. (2013). Graphene transistors: Status, prospects, and problems. *Proceedings of the IEEE*, 101(7), 567-1584.
- [2] Radisavljevic, B., Radenović, A., Brivio, J., et al. (2011). Single-layer MoS_2 transistors. *Nature Nanotechnology*, 6(3), 147-150.
- [3] O'Brien, K. P., Dorow, C. J., Naylor, C. H., et al. (2023). Process integration and future outlook of 2D transistors. *Nature Communications*, 14, 6400.
- [4] Kalantar-Zadeh, K., Kis, A., Coleman, J. N., et al. (2012). Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. *Nature Nanotechnology*, 7(11), 699-712.
- [5] Li, L., Yu, Y., Ye, G. J., et al. (2014). Black phosphorus field-effect transistors. *Nature Nanotechnology*, 9(5), 372-377.
- [6] Das, S., Chen, H-Y., Penumatcha, A. V., & Appenzeller, J. (2013). High performance multilayer MoS_2 transistors with scandium contacts. *Nano Letters*, 13(1), 100-105.
- [7] Sebastian, A., Pendurthi, R., Choudhury, T. H., Redwing, J. M., & Das, S. (2021). Benchmarking monolayer MoS_2 and WS_2 field-effect transistors. *Nature Communications*, 12, 693.
- [8] Clark, N., Nguyen, L., Hamer, M. J., et al. (2018). Scalable patterning of encapsulated black phosphorus. *Nano Letters*, 18(9), 5373-5381.
- [9] Islam, M. S., Hasan, M. S., Rahman, M. Z., et al. (2023). Quantum transport properties of monolayer MoS_2 , WS_2 , and black phosphorus: A comparative study. *Journal of Condensed Matter*, 1(02).
- [10] O'Brien, K. P., Dorow, C. J., Penumatcha, A., et al. (2021, December). Advancing 2D monolayer CMOS through contact, channel and interface engineering. In *IEDM* (pp. 7.1.1-7.1.4). San Francisco: IEEE.
- [11] Abate, Y., Alessandra, A., et al. (2018). Recent progress on stability and passivation of black phosphorus. *Advanced Materials*, 30(29), 1704749.
- [12] Ahmed, T., Balendhran, S., & Walia, S. (2017). Degradation of black phosphorus is contingent on UV-blue light exposure. *npj 2D Materials and Applications*, 1, 18.