INVITED PAPER

Tunneling Transistors Based on Graphene and 2-D Crystals

Graphene-based tunneling transistors and how these compare to 2-D transistors made from the GaAs/AlGaAs materials systems is the topic of discussion in this paper.

By Debdeep Jena, *Member IEEE*

ABSTRACT | As conventional transistors become smaller and thinner in the quest for higher performance, a number of hurdles are encountered. The discovery of electronic-grade 2-D crystals has added a new "layer" to the list of conventional semiconductors used for transistors. This paper discusses the properties of 2-D crystals by comparing them with their 3-D counterparts. Their suitability for electronic devices is discussed. In particular, the use of graphene and other 2-D crystals for interband tunneling transistors is discussed for low-power logic applications. Since tunneling phenomenon in reduced dimensions is not conventionally covered in texts, the physics is developed explicitly before applying it to transistors. Though we are in an early stage of learning to design devices with 2-D crystals, they have already been the motivation behind a list of truly novel ideas. This paper reviews a number of such ideas.

KEYWORDS | Graphene; semiconductors; transistor; tunneling

I. INTRODUCTION

Semiconductors come in many crystal forms. Since their discovery in the early 20th century, the semiconductors used in electronic and optical devices are of the 3-D crystal form. Three-dimensional crystal semiconductors have remained at the heart of such devices from the earliest "cat's whisker" detectors [1] to the latest billion-transistor silicon complementary metal–oxide–semiconductor (CMOS) [2], [3] and quantum-well (QW) lasers [4]. As the understanding of the physics of electron transport and electron–photon coupling sharpened, it became clear that controlling the potential energy landscape of electrons could lead to massive boosts in device functionality and performance.

The first level of direct control of the "energy-band diagrams" was by chemical doping, which involved replacing a small number of atoms of the 3-D semiconductor by those with higher or lower valence. The next advance involved varying the chemical nature of the crystal along specific directions, which marked the birth of semiconductor heterostructures [5]. These advances taught electrons "new tricks," and made possible the smallest and fastest electronic switches [6], high-density memories, and the most efficient light-emitting diodes (LEDs) and lasers [7]. These devices form the bedrock of computation, data storage, solid-state lighting, and communication in today's information age.

At this time, in the early part of the 21st century, these building blocks based on traditional device concepts are approaching their performance limits. Therefore, new ideas and new materials are necessary. For example, photonic crystal, metamaterial, and plasmonic concepts are advancing the area of optoelectronic devices beyond what was thought possible before [8], [9]. Strong light–matter interaction has been exploited to demonstrate polariton lasers that take advantage of Bose–Einstein condensation at room temperature for ultralow threshold lasing [10].

Similarly, for electronic switching devices, a number of approaches are being taken to address the future beyond scaling. Conventional field-effect and bipolar transistors operate on the basis of energy filtering of electrons (or holes) flowing over a barrier. The barrier is electrostatically controlled with a voltage. In an electrostatically well-designed device, all of the control voltage is spent in

Manuscript received July 21, 2012; revised December 12, 2012 and February 9, 2013; accepted February 13, 2013. Date of publication May 3, 2013; date of current version June 14, 2013. This work was supported in part by the National Science Foundation under Grant 0802125, the Semiconductor Research Corporation (SRC) through the Nanoelectronics Research Initiative (NRI) program at the Midwest Institute of Nanoelectronics Descovery (MIND) center, and the U.S. Air Force Office of Scientific Research (AFOSR).

The author is with the University of Notre Dame, Notre Dame, IN 46556 USA (e-mail: djena@nd.edu).

Digital Object Identifier: 10.1109/JPROC.2013.2253435

moving the barrier. The electrons carrying the current are spread in a band according to the Fermi–Dirac distribution, with a Boltzmann tail in energy. The energy filtering thus leads to a current dependence of the form $I \sim \exp[qV/k_BT]$, where q is the electron charge, V is the voltage, T is the temperature, and k_B is the Boltzmann constant. When operated in this fashion, the current cannot be changed any steeper than $S \sim (k_BT/q)\ln 10 \sim$ 60 mV/decade. This subthreshold swing (SS) "limit" is often referred to as the "thermal limit" or the "Boltzmann limit" (though Boltzmann did not set this limit). We refer to this condition as the SS limit to avoid confusion.

An electronic switch must have its ON- and OFF-states clearly demarcated for performing digital (Boolean) logic. Let us say this demarcation is set to $I_{ON}/I_{OFF} = 10^4$. To achieve it, a voltage supply of at least 4×60 mV $= 0.24$ V is necessary. Since the speed of switching and the dynamic and static power dissipation of transistors are strong functions of the supply voltage, the SS limit sets a floor of minimum power dissipation. This issue is described in sufficient detail in a number of recent articles that motivate the search for new materials and ideas for going beyond the SS limit [11]–[13].

Now there is nothing particularly fundamental about the SS limit. Devices that do not operate on the traditional transistor mechanism exist today and operate below the SS limit. An example is a nanoelectromechanical system (NEMS), which is the analog of a mechanical relay. Substantial progress has been made in this area [14]. Due to mechanical moving parts, these devices are currently slow, but are expected to improve with scaling.

A number of relatively new ideas are being explored at this time for switching devices beyond the SS limit. Some exploit impact ionization to obtain sub-SS limit operation [15], [16]. Other devices aim to use correlated electron effects; for example, if electrons can be made to "pair up" similar to Cooper pairs in superconductors, but at room temperature, the SS limit would be cut in half. If the control voltage could be internally "stepped up" through novel ferroelectric gates, sub-SS limit devices can be realized [17]. Other routes involve the internal transduction of the voltage into other state variables such as strain, spin, or electron localization [18]. Among these strategies, a transistor concept based on interband tunneling transport has emerged as an attractive candidate for switching. This paper will focus on this device. The tunneling field-effect transistor (TFET) can be realized in traditional 3-D crystal semiconductors and their heterostructures.

However, since the discovery of graphene in 2004, device engineers have a new class of materials in 2-D crystals at their disposal. In this paper, we discuss possible realizations of TFETs with 2-D crystals, and compare them with 3-D crystal counterparts. In the process of this discussion, a number of novel features of 2-D crystals will emerge that distinguish them from traditional 3-D crystal semiconductors. These novel features of the new material family offer a compelling case for investigating them further. To motivate their suitability for electronic devices, we first discuss the various 2-D crystal materials and their properties. We do so against the backdrop of their ubiquitous 3-D crystal semiconductor counterparts.

II. TWO-DIMENSIONAL CRYSTALS

Two-dimensional crystals exploded into the limelight in 2004 with the remarkable reports of the isolation of atomically thin graphene [19]–[21]. What is often overlooked is that the early reports [22] also presented evidence of the isolation of single-layers of BN—an insulator or a wide-bandgap semiconductor, MoS_2—a traditional semiconductor, and $NbSe_2$—a superconductor with possible charge-density wave electronic phases. Single layers of the cuprate high-T_c superconductors were also isolated. It is interesting to note that the voltage "scaling" of silicon CMOS processors stalled around the same time, marking the move toward multicore processors [23]. One of the reasons for paradigm shift was the unsustainable increase in dynamic and OFF-state power dissipation due to the SS limit and high-frequency operation. Whether 2-D crystals can help in this arena remains to be seen. We first discuss a few properties of 2-D crystals and their suitability for electronic devices.

Fig. 1 is a schematic representation of the structure of crystals of various dimensions. The bottom row shows the atomic building blocks. The first column shows the ubiquitous 3-D crystal semiconductors. The second column shows the emerging family of 2-D crystals and their many variants. The third and fourth columns indicate ideal 1-D and 0-D structures. Atomic chains have been investigated for their transport properties [24], and a benzene ring can be considered either as an atomic "ring," or even a basic 2-D crystal unit. An atom is a perfect 0-D structure in which electrons are localized in all three dimensions. We note that the electrons in an atom still move in 3-D, but their energy spectra are discrete and gapped; they do not form bands that are necessary for transport. It is in this sense that they are 0-D. We focus our attention on 2-D crystals, and their differences from 3-D crystal semiconductors.

The building blocks for 3-D semiconductors are typically tetrahedrally bonded atoms. The lattice is 3-D, and the basis typically consists of two atoms. For example, electrons in 3-D crystals from group IV elements (Si, Ge, etc.) occupy [core] ms^2mp^2 orbitals, where m is the row number in the periodic table, and [core] represents the core electrons that do not participate in chemical bonding directly. Electrons from the outermost s and p orbitals of nearest neighbor atoms pair up to form sp^3 bonds. An sp^3 bond is inherently 3-D, and so is the

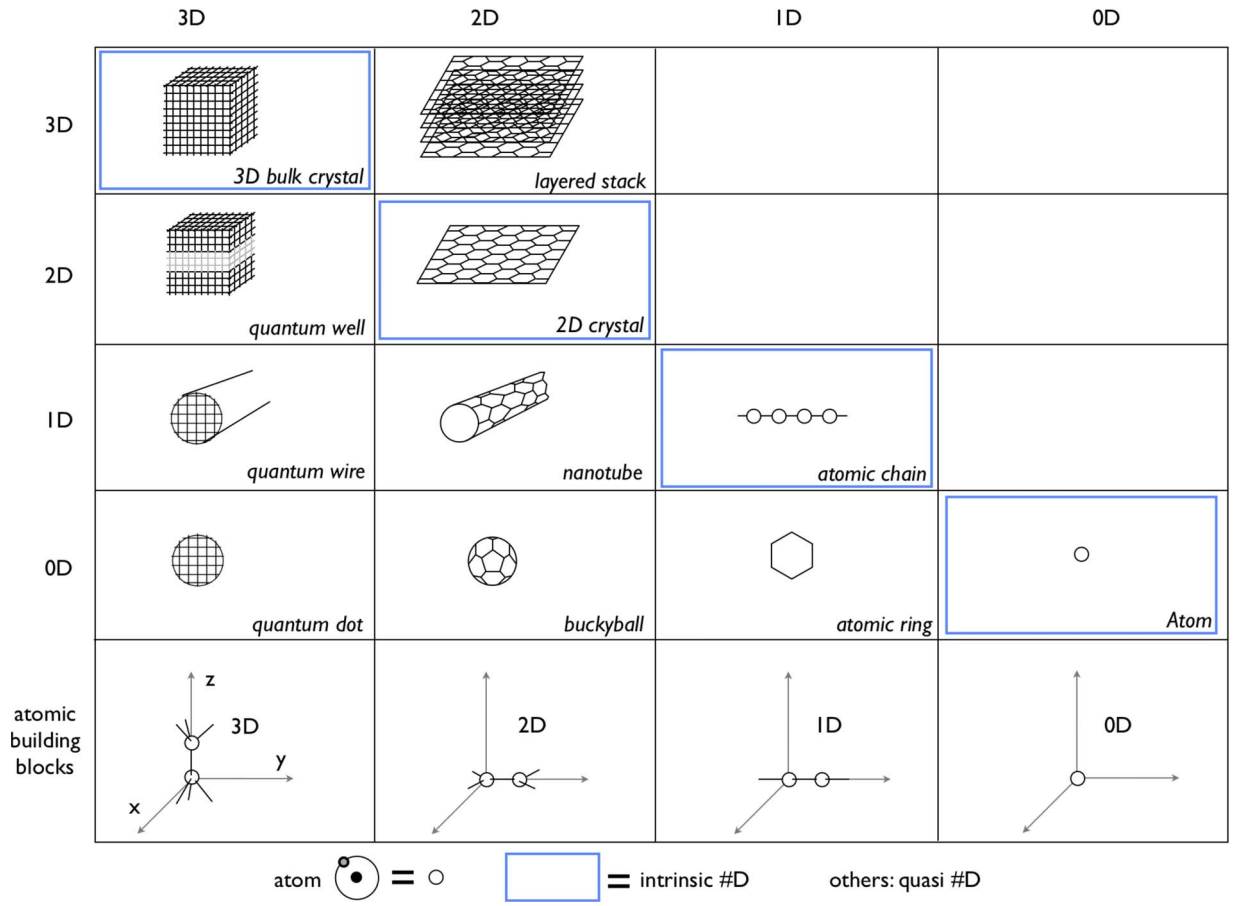

Fig. 1. *A schematic representation of "crystals" of the many spatial dimensions that result from various building blocks. The building blocks contain atomic bases that form 3-D bonds in the first column, 2-D planar bonds in the second column, and 1-D linear bonds in the third column. The ideal 0-D structure is an atom in the fourth column.*

resulting semiconductor crystal. The natural crystal is thus a bulk 3-D semiconductor. A termination such as a surface results in dangling bonds, a fraction of which might reconstruct.

The corresponding building block of a 2-D crystal consists of a *planar* 2-D lattice. For graphene and BN, the basis consists of two atoms attached to a hexagonal planar lattice. These chemical bonds in the two-atom basis for graphene and BN are of the sp^2 type. So the chemical bonds of their basis are *also* planar. In the second column of Fig. 1, the underlying planar structure of 2-D crystals is shown. Attached to each point of intersection is one carbon atom for graphene, alternating B and N atoms for BN, and a basis of X-M-X for transition metal dichalcogenides (TMDs). TMD 2-D crystals share the same planar lattice geometry of graphene and BN. But the basis of TMD 2-D crystals consists of three atoms of the form MX_2, where M is the transition metal chemically bonded to two chalcogenide atoms X. The chemical bonds in TMD 2-D crystals (e.g., MoS_2, WSe_2, WS_2, etc.) involve s-, p-, and d-orbitals, and the two M-X bonds stick out of the center 2-D plane containing the transition metal atom M [25]. Thus, unlike its lattice, the basis of TMD 2-D crystals is not perfectly planar. Recent reports also indicate the possible existence of 2-D forms of Si (silicene), Ge (germanene), and possibly AlN and GaN [26]–[28], [102]. Single layers of 2-D crystals are typically less than 1 nm in thickness. An exotic form of a 2-D crystal semiconductor may also exist when two surfaces of topological insulators come close to each other [29]. These materials have been less explored than the others discussed here.

Unlike a perfect 3-D crystal, a perfect 2-D crystal has no broken/dangling bonds on its surface. The quasi-low-dimensional structures formed from 3-D crystals such as 2-D nanomembranes, 1-D nanowires, and 0-D nanocrystals are still volume elements deriving from 3-D bonding, and necessarily have dangling bonds on their surfaces. These broken bonds may be passivated by either dielectrics, or by lattice-matched or strained heterostructures. In contrast, the various dimensional structures deriving from 2-D crystals are "hollow" and are "all-surface."

Two-dimensional crystal sheets may be stacked to form 3-D structures with weak van-der-Waal's interlayer bonding. They can be rolled up into quasi-1-D nanotubes, or into 0-D buckyballs (C_{60}). The symmetry of a 2-D crystal is broken at its *edge*. Similar to the surface state reconstruction or passivation of the surfaces of 3-D crystals, the edge states can reconstruct and tie up the dangling bonds. For special cases, such as in buckyballs, the chemical bonding is seamless and there are no broken bonds. Indeed, the icosahedral geometry of the buckyball belongs to one of the five platonic solids, which have mathematically represented "perfection" in shape since the earliest times [30], [31].

For electronic devices using field effect, the *absence of dangling bonds* is a major advantage for planar 2-D crystals, since electrons trapped in them serve to shield electric field lines from entering the bulk of the corresponding 3-D semiconductors. We now discuss the electronic properties of 2-D crystals and compare them to those of 3-D crystal semiconductors.

A. Electronic Properties of 2-D Crystals

The electronic orbitals that form the family of 2-D crystals are shown in Fig. 2. Electron states at the conduction and valence band edges of 3-D semiconductors derive from various admixtures of sp^3 bonds. For direct-gap semiconductors such as GaAs and GaN, the conduction band edge is mostly s-like. The spherical symmetry of the s-orbitals imparts electrons in the conduction band their isotropic nature. The electronic states at the valence band edge on the other hand are more p-like. Because p-orbitals are directional, the hole effective mass is anisotropic. The imbalance of the nature of chemical bonding in 3-D crystals semiconductors thus also results in an *asymmetry* in the curvature or the effective mass of the conduction and valence band states. In modern complementary logic devices, symmetry is a highly desirable characteristic. The degree of asymmetry between, for example, nMOS and pMOS devices dictates the geometry and layout of circuits that could be considerably simplified by symmetry.

The covalent bonds in graphene and BN are of the sp^2 kind. They are responsible for the structural properties of the crystal. The leftout p_z orbital sticks out of the 2-D plane. The electrons in these orbitals can hop between nearest neighbors, leading to the electronic conductivity and optical properties of such crystals. In graphene and BN, the structural properties such as thermal conductivity and mechanical stability derive from the covalent sp^2 bonds. But the electronic and optical properties derive from the delocalized p_z orbitals. There is a wide energy separation between the sp^2 and p_z energy bands. In this sense, the electronic properties of such 2-D crystals have a different origin than their structural properties. This is in contrast to 3-D semiconductors, where the structural and electronic properties derive from the same sp^3 electronic band states.

Electrons in 3-D crystals can be quantum-mechanically confined to move in 2-D and 1-D, or localized in 0-D by chemical and geometrical constraints in heterostructures, as shown in the first column in Fig. 1. This is achieved by taking advantage of energy band offsets around the bandgap. Conduction band offsets ΔE_C confine electrons, and valence band offsets ΔE_V confine holes. We note here

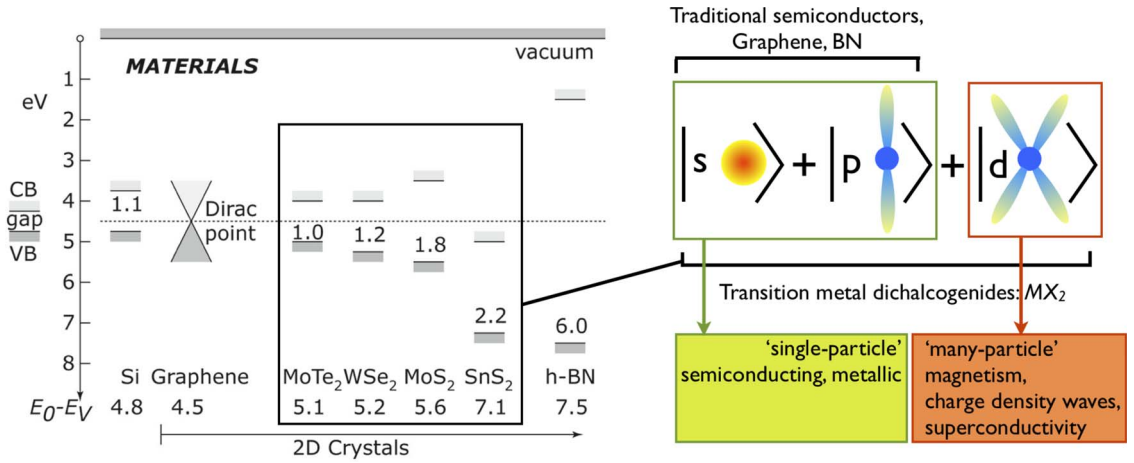

Fig. 2. *Energy band alignments of various 2-D crystals compared to silicon. The relative energy band offsets of graphene, BN, and transition-metal dichalcogenides are shown. The numbers at the center indicate the respective bandgaps reported at this time, but are subject to refinement with further experiments. An energy scale from the vacuum level is also indicated, showing a work function (or electron affinity) of intrinsic zero-gap 2-D graphene to be ~4.5 eV. The conduction and valence band edge states of Si, graphene, and BN are formed of linear combinations of* $|s>$ *- and* $|p>$ *-orbitals, whereas those of the transition-metal dichalcogenide 2-D crystals involve* $|d>$ *-orbital states at the band edges. The presence of* d*-orbital states near the Fermi level implies that some of these 2-D crystals can exhibit electronic phenomena that require many-particle effects such as magnetism and superconductivity.*

that the 2-D confinement of electrons in a quantum well in a 3-D crystal leads to a quasi-2-D electron gas (2-DEG). This means there are multiple 2-D electronic subbands whose spacing in energy grows as the inverse square of the spatial confinement. In sharp contrast, there is just *one* band for 2-D electron systems in single-layer 2-D crystals, since the electron wave function cannot spread sufficiently out of the plane in equilibrium.

The energy bandgaps and the band lineups of a few 2-D crystals are shown in Fig. 2. The figure also indicates the chemical bonding schemes that characterize them, along with their relative positions with respect to the vacuum energy level [32]–[34]. A distinctive feature of the 2-D crystals is that their energy gap windows are not populated by surface states in sufficiently crystalline sheets, as is necessarily the case for 3-D crystals. Thus, the measurements of their band alignments are relatively simpler, as described in [33].

Graphene is a zero-bandgap semiconductor, with the energy dispersion $E(k_x, k_y) = \pm\hbar v_F|\mathbf{k}|$, where $\hbar$ is the reduced Planck's constant, $v_F = 10^8$ cm/s is called the Fermi velocity, and $|\mathbf{k}| = \sqrt{k_x^2 + k_y^2} = 2\pi/\lambda$ is the electron wave vector. The dispersion is taken around the Dirac points in the band structure, which are located at the twofold degenerate K-points in the k-space, as shown in Fig. 3. These states are similar to the conduction band edge or valence band edge states of 3-D semiconductors. The positive branch is the conduction band, and the negative branch is the valence band. We note the perfect symmetry of the bands, which is quite distinct from traditional 3-D semiconductors. This symmetry is special, and has an important bearing on tunneling transistors discussed later. The energy bandgap is zero. The density of states (DOS) of 2-D graphene is given by $\rho_{gr}^{2\text{-D}}(E) = [g_s g_v/2\pi(\hbar v_F)^2] \times |E|$, where $g_s = 2$ is the spin degeneracy and $g_v = 2$ is the valley degeneracy [35].

Two-dimensional BN has an energy bandgap of ~6.0 eV as a consequence of the broken crystal symmetry in the basis, but its band extrema also occur at the K-points in the Brillouin zone. Thus, it has the same valley degeneracy as graphene. The effective mass characterizing the symmetric conduction and valence bands of 2-D BN is $m^* \sim 0.6m_0$, where m_0 is the free-electron mass [36]. The DOS of 2-D BN looks like those of conventional 2-DEGs, $\rho_{\text{BN}}^{2\text{-D}}(E) = (g_s g_v m^*/\pi\hbar^2) \times \theta[E - E_C]$. The major difference is the absence of higher subbands owing to the absence of atoms out of the plane.

The bandstructures of 2-D crystal semiconductors of the TMD family are being evaluated at this time [37], [38]. Initial experiments and theoretical models point out that they too have their band extrema at the K-points. The conduction and valence bands in single-layer TMDs appear less symmetric than graphene and BN, but much more symmetric than traditional 3-D semiconductor crystals. Effective masses ranging from $m^* \sim 0.34m_0 - 0.76m_0$ have been calculated, and are expected to undergo refinement through experimental measurements [39].

Two-dimensional crystal sheets typically occur in nature in their stacked layered forms. The band structures of the multilayer variants of graphene, BN, and TMDs are distinct from the single-layer counterparts. The bandgap of a stacked 2-D crystal is smaller than the single layer [40], [41]. For example, graphite becomes a semimetal with a negative bandgap. Similarly, when 2-D crystals are used to form 1-D nanotubes (Fig. 1), quasi-1-D subbands appear, and the bandgaps increase due to additional quantum confinement. The DOS then acquires van Hove singularities in a manner similar to quasi-1-D nanowires or quantum wires formed of 3-D semiconductor crystals. In this paper, we maintain focus on single-layer 2-D crystals and occasionally mention their quasi-1-D and quasi-3-D variants when they appear in context. The discussion of the electronic band structures of 2-D crystals leads us naturally to a point where we can gauge their suitability for electronic devices. We start by discussing their suitability for traditional field-effect transistors (FETs).

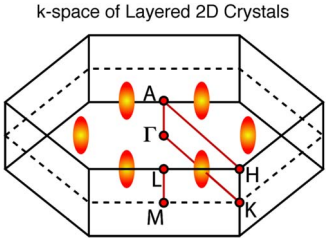

Fig. 3. *The k-space picture of 2-D crystals such as graphene, BN, and the transition-metal dichalcogenide MX_2 compounds. A good understanding of the k-space picture is important for choosing the right materials for device applications, and especially important for tunneling transistors. Since the real-space lattice is hexagonal in the 2-D plane, so is the k-space lattice. Since the interlayer separation is larger than the in-plane lattice constant, the hexagonal Brillouin zone is shorter in the vertical direction. The important high-symmetry points are labeled. Graphene, BN, and single-layer MoS_2 have their conduction band edge and valence band edges at the K-points, which leads to twofold degeneracy. The conduction band edge of multilayer MoS_2 at this point is believed to be along the $\Gamma - K$ minimum as shown, which makes it an indirect-bandgap semiconductor, and imparts to it a valley degeneracy of 6 by symmetry, similar to silicon.*

B. Suitability of 2-D Crystals for Traditional Transistors

The operation of a FET hinges on electrostatics and transport of charge carriers. FETs based on 3-D crystal semiconductors have been scaled to ~10 s of nanometer channel lengths in the quest to achieve higher performance. As the source/drain separations have been scaled, it has become necessary to reduce the channel thickness. This requirement is driven by the need for a gate metal to exercise electrostatic control over mobile electrons and holes. If the gate is farther away from the carriers than the S/D distance, it loses control over them. The device then

cannot be switched on or off as effectively as is needed for the transistor to operate in a circuit. This necessity is at the root of the reason for the move to silicon-on-insulator (SOI) and FinFET type of topologies [42]. The silicon channels have thus become more 2-D in SOI structures [43], and closer to 1-D in FinFETs and nanowire geometries.

A quantitative statement of the importance of electrostatics is obtained from a solution of the Poisson equation for a FET. For a FET with a semiconductor layer of thickness t_s of dielectric constant ε_s gated through an insulator of thickness t_{ox} and dielectric constant ε_{ox}, the Poisson equation for the electric potential V takes the form $\partial_x^2 V \sim V/l^2$, where $l = \sqrt{t_s t_{ox}(\varepsilon_s/\varepsilon_{ox})}$ is the characteristic "scaling length" [44]. This length determines the smallest distances over which electric potential may be dropped. Therefore, for scaling to the smallest lengths, high-K insulators and ultrathin channels are desirable. This argument, in conjunction with the absence of dangling bonds and the associated interface traps highlights the attractive feature of 2-D crystals for ultrascaled FETs based on electrostatics arguments alone. Furthermore, 2-D crystal insulators such as BN can eliminate dangling bonds altogether in planar FET geometries.

As the channels of 3-D semiconductors are thinned down, the roughness of the surfaces causes degradation of the carrier transport due to surface-roughness scattering. The root of this form of scattering is the effect of the roughness on the quantization of energy levels. For example, in a SOI structure of thickness t, the quantization energy of subbands varies as $E \sim \hbar^2/m^* t^2$. Variation of the layer thickness by Δt leads to a perturbation of the subband edge by $\Delta E \sim (2\hbar^2/m^* t^3)\Delta t$. Since the scattering rate is proportional to the square of the perturbation, the mobility degrades as $\mu \sim t^6$, i.e., roughly as the sixth power of the width [45]. Thus, for very thin layers of a 3-D semiconductor, such as those used in ultrathin body (UTB) transistors, the transport properties suffer from the surface roughness. Two-dimensional crystals offer an ideal solution to this problem. Two-dimensional crystals are intrinsically of an atomically thin body (ATB) nature. When sufficiently pure, they do not have surface roughness. The attractiveness of TMD 2-D crystal semiconductors was brought to sharp focus with the demonstration of single-layer MoS_2 FETs [46]. A FET with 10^8 ON/OFF ratio at room temperature and electron mobility of $\sim$200 cm^2/Vs was achieved with a single layer of MoS_2 2-D crystal of thickness < 1 nm. The SS was close to ideal, thanks to the absence of broken bonds and associated interface traps. Such performance has never been measured in devices made from 3-D crystals of the same thickness. Though the initial results look promising, the dynamic range and reliability of the performance metrics will be assessed carefully in the next few years.

Additional novel features of charge transport in 2-D crystals that have been predicted and recently observed include dielectric-mediated carrier mobilities. The basic premise is that the Coulomb interaction $V \sim q/4\pi\varepsilon r$ between charged impurities and mobile channel carriers is mediated by the dielectric constant ε of the space separating them. In 3-D semiconductors, the Coulomb interaction is dominated by the bulk dielectric constant of the semiconductor itself (i.e., $\varepsilon = \varepsilon_s$) since the charged impurity and the charge carrier are effectively buried inside and in close proximity. On the other hand, in 2-D crystals, most of the electric field lines connecting the charged impurity to the mobile carrier actually lie *outside* the 2-D crystal itself, in the surrounding dielectric. This effectively provides an external knob to damp Coulomb scattering and improve carrier mobilities, since $\varepsilon \sim \varepsilon_{ox}$ for this interaction [47]. Use of high-K dielectrics has been observed to damp scattering and improves charge mobility in 2-D crystals such as graphene [48], [49] and MoS_2 [46]. The exact mechanisms likely also include phonons.

At this time, the understanding of transport in 2-D crystals is evolving. It is clear that the interactions that limit charge transport in 3-D semiconductors and heterostructures were intrinsic to the 3-D crystal itself. But for 2-D crystals, these interactions can be tuned based on what we put around them. This is because in 2-D crystals we have direct access to the electrons, their spins, and atomic vibrations to an unprecedented degree. As our understanding of these mechanisms evolves, the level of direct access to the physical properties may well prove to be *the* defining factor that differentiates 2-D crystal devices from their 3-D counterparts. This feature is simultaneously an advantage *and* a challenge, since noise and reliability of the desired nanoscale devices must be robust for usability.

C. Possibility of 2-D Crystal Heterostructures

Heterostructures based on 3-D crystals take advantage of energy band offsets that originate from differences in chemical composition. The concept of quasi-electric fields in heterostructures breaks the symmetry of electrical forces acting on electrons and holes. In a semiconductor of constant chemical composition (uniformly doped, or p-n homojunctions), the electric force acting on electrons and holes is the same. This is not true in a heterostructure [5]. This broken symmetry is central to quantum confinement and high oscillator strengths that have led to high-efficiency LEDs and lasers. QW FETs and even the MOSFET gain from the concept of quantum confinement. In graded-base heterostructure bipolar transistors (HBTs), the broken symmetry is central in speeding up electrons with a quasi-electric field in the same region in space where there is no field acting on holes [50]. Examples of such heterostructures based on 3-D crystal semiconductors include SiGe/Si, AlGaAs/GaAs/InGaAs, and AlGaN/GaN/InGaN material systems. Except in special cases, most of such 3-D crystal heterostructures have strain due to the lattice mismatch. Strain can be desirable for affecting the carrier

transport or as the driving force for the formation of quantum dots by the Stranksi–Krastanov mechanism during epitaxy. Strain can often be undesirable, since it can lead to relaxation and defect formation beyond certain critical thicknesses.

Heterostructures based on 2-D crystals are at their infancy. However, a number of interesting features are likely to emerge in them. Initial demonstrations of in-plane 2-D crystal heterostructures such as graphene seamlessly connected to BN have been experimentally observed, and provide exciting opportunities in device design [51]. Hybrid heterostructures composed of 2-D crystals such as graphene placed on 3-D semiconductors such as silicon have been used to demonstrate new device concepts. One recent example is a graphene–Si Schottky diode where graphene may be thought of as the Schottky "metal" contact. However, unlike a typical metal, the Fermi level of graphene can be tuned with a third gate electrode, which leads to a variable Schottky-barrier height [52]. This idea was used to demonstrate a variable-barrier transistor (or the so-called "Barristor").

Out-of-plane or vertical heterostructures are also realized when 2-D crystals are stacked on each other. Such heterostructures do not suffer from lattice mismatch requirements, since there are no interlayer covalent bonds. The weak van der Waal's interlayer bonding in principle allows unstrained integration of 2-D crystal layers of different material properties. One may envision vertical heterostructures of 2-D crystal metals, semiconductors, insulators, and perhaps a wider range of materials. Due to the absence of broken bonds, the interfaces are expected to be pristine and devoid of electronic trap states. Interlayer transport of electrons would involve tunneling. The rotational alignment of the 2-D crystal layers might play an important role in such heterostructures. These features are currently under investigation, and are certain to lead to a range of new applications. Initial demonstrations of a graphene–BN–graphene and graphene–MoS_2–graphene heterostructures tunneling transistors have been recently reported [53]. A proposed device called the bilayer pseudospin FET (BiSFET) is based on many-body excitonic condensation of electron–hole pairs in closely spaced layers of graphene. It falls under the category of vertical 2-D crystal heterostructures [54]. Its single-particle counterpart, a tunneling transistor that takes advantage of the symmetry of the bandstructure of some 2-D crystals, is called the "SymFET" [55]. These tunneling devices that are rooted in 2-D crystals are described in Section V.

D. Maturity of 2-D Crystals and Material Challenges

Since the field of 2-D crystal semiconductors is relatively young, a short discussion of the material challenges is necessary. Since the initial demonstrations in 2004, the large-area growth capability of single-layer graphene has expanded rapidly [21]. At this time, epitaxial single-layer graphene on several-inch-diameter SiC wafers are available [56], [57]. Chemical vapor deposition (CVD)-grown graphene has been realized on metals, and transferred to other substrates [58]. Nanoribbons have been fabricated on CVD-grown graphene [59]. CVD-grown graphene has shown promise for larger area crystals than epitaxial graphene, which is limited to the size of the starting 3-D crystal substrate. The crystal quality is not perfect yet, but as was the case in the development of 3-D crystals, there is reason to believe it will undergo drastic improvements in the near future.

Similarly, BN 2-D crystals have been grown by CVD, as have electronic-grade MoS_2 and WS_2 layered materials [60]–[62]. However, it is also important to realize that most forms of 2-D crystals have been produced in large volumes in their layered forms [63]. They have already found industrial applications in chemical catalysis (MoS_2, graphite), lithium–ion batteries (lithium cobaltate and layered carbon), lubricants (MoS_2), neutron moderation in nuclear reactors (graphite), and thermally and mechanically refractory crucibles used in much of electronic material and device processing (BN and graphite). The development of electronic grade counterparts thus is expected to heavily leverage the considerable prior existing knowledge and industrial base for these materials.

A major immediate challenge is to develop methods of doping and controlling the Fermi level in 2-D crystals. Possible methods with TMD 2-D crystals include chemical substitutional doping, and/or modulation doping by taking advantage of the rich intercalation chemistry of such layered materials. Since doping control is intimately connected to the ability to form low-resistance contacts, this challenge assumes increased importance.

The development of electronic grade 2-D crystals is expected to be rapid. The first active device applications are expected to be in traditional FETs. For example, TMD-based transistors offer attractive routes to large-area thin-film transistors (TFTs) by virtue of low SS values and respectable mobilities when compared to organic semiconductors and 3-D oxide materials [64]. But can they offer new functionalities for high-performance devices beyond what is being envisioned with 3-D crystal semiconductors? To address that question, we focus the rest of the paper on one of the possible candidates for high-performance and low-power energy-efficient logic devices: the tunnel FET (TFET).

III. TUNNELING TRANSPORT IN SEMICONDUCTORS

Following the motivation provided earlier, we start with a short introduction to tunneling transport and its incorporation into the heart of the transistor operation. The discussion starts with an evaluation of the effect of dimensionality on interband Zener tunneling [65], [66].

Consider the $p - i - n$ junction shown in Fig. 4. We make some simplifying assumptions that allow us to zone

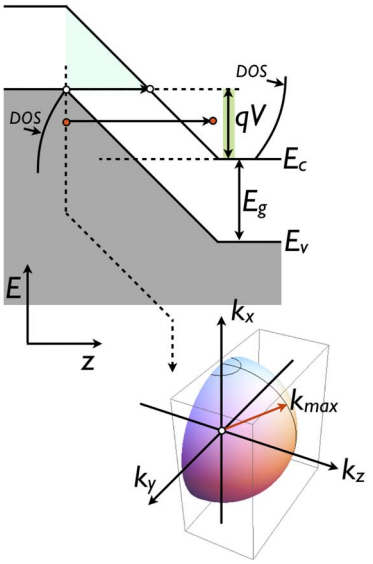

Fig. 4. *Interband tunneling in a reverse-biased $p-i-n$ junction diode. Most TFETs use the reverse-bias Zener tunneling as the mechanism of current conduction in their ON-states. The current may be calculated by integrating over the k-states at the injection point as outlined in the text.*

into the relevant physics immediately. Assume the doping in the p- and n-sides are just enough to align the Fermi levels at the respective band edges. Then, under no bias, $E_V^p = E_c^n$ and no net current flows across the junction. Under the application of a reverse bias voltage V, a finite energy window is created for electrons since $E_V^p - E_C^n = qV$. Within this energy window, electrons from the valence band can tunnel into the conduction band on the other side, as indicated.

The current is calculated by summing the individual contributions by each k-state electron. There are many approaches to evaluate currents, but none is as transparent as the formalism in the k-space. To illustrate, we write the tunneling current as

$$I_T = q\frac{g_s g_v}{L}\sum_k v_g(k)(f_v - f_c)T \tag{1}$$

where $g_s = 2$ is the spin degeneracy and g_v is the valley degeneracy. L is the macroscopic length along the electric field (which will cancel out), $v_g(k) = \hbar^{-1}\nabla E(k)$ is the group velocity of carriers in the band $E(k)$, f_v, f_c are the Fermi–Dirac occupation factors of the valence and conduction bands, and T is the tunneling probability. The sum is over k-states for electrons that are allowed to tunnel. We illustrate the clarity of this approach by using the same expression for evaluating Zener tunneling currents for $p-i-n$ junctions made of 3-D, 2-D, and 1-D crystals. We first consider semiconducting crystals that have a bandgap. Then, we remove the bandgap criteria to allow for special cases such as graphene.

The tunneling probability is obtained by the Wentzel–Kramers–Brillouin (WKB) approximation [67]. For electrons in the valence band of the p-side with transverse kinetic energy $E_\perp = \hbar^2 k_\perp^2/2m_v^*$, the WKB tunneling probability is given by [68]

$$T_{\mathrm{WKB}} = \exp\left[-\frac{4\sqrt{2m_R^*}(E_g + E_\perp)^{\frac{3}{2}}}{3q\hbar F}\right] \approx T_0 \exp\left[-\frac{E_\perp}{\overline{E}}\right] \tag{2}$$

where $T_0 = \exp[-4\sqrt{2m_R^*}E_g^{3/2}/3q\hbar F]$, $\overline{E} = q\hbar F/2\sqrt{2m_R^* E_g}$, F is the (constant) electric field in the junction, and m_R^* is the reduced effective mass given by $m_R^* = m_c^* m_v^*/(m_c^* + m_v^*)$. This expression is found to be consistent with experimental results [69]. Note that the tunneling probability of electrons is lowered exponentially with their transverse kinetic energy. To evaluate the tunneling current, we attach this tunneling probability to each electronic k-state, and sum it over all electrons incident on the tunneling barrier.

Three-dimensional semiconductors: Consider the case when the $p-i-n$ junction is made of 3-D crystal semiconductors. In Fig. 4, we concentrate on a particular 2-D plane as shown by the dashed line, at the $p-i$ junction. Half of the electrons in the valence band in that plane move to the right in the $+k_z$ direction, as indicated in the hemisphere in the k-space. Since there are negligible electrons in the conduction band in that plane, the current there must be carried by electrons in the valence band. But which of these right-going electrons are allowed to tunnel through the gap? In the absence of phonon scattering, tunneling is an elastic process. This enforces the energy requirement

$$E_v^p - \frac{\hbar^2}{2m_v^*}\left(k_{xp}^2 + k_{yp}^2 + k_{zp}^2\right) = E_c^n + \frac{\hbar^2}{2m_c^*}\left(k_{xn}^2 + k_{yn}^2 + k_{zn}^2\right) \tag{3}$$

with the additional requirement that the lateral momentum be conserved. To simplify the analytical treatment, and in preparation for 2-D crystals, we further assume that the bands are symmetric, i.e., $m_c^* \approx m_v^* = 2m_R^*$. The energy and momentum conservation requirements thus lead to the relation

$$2k_\perp^2 + k_{zp}^2 = \frac{4m_R^* qV}{\hbar^2} - k_{zn}^2 \tag{4}$$

where $k_\perp^2 = k_{xp}^2 + k_{yp}^2$. Let us define $k_{\max}^2 = 4m_R^* qV/\hbar^2$. Since there is an electric field in the z-direction, momentum in that direction will not be conserved. For the

electron to emerge on the right (n-)side, k_{zn} must be nonzero, and thus $k_{zn}^2 \geq 0$, which implies

$$2k_{\perp}^2 + k_{zp}^2 \leq k_{\max}^2. \tag{5}$$

The above condition defines a restricted volume Ω_T of the k-space hemisphere for electron states that are allowed to tunnel. We are now in a position to evaluate the tunneling current for 3-D semiconductor $p-i-n$ junctions. In the expression for the tunneling current [see (1)], the group velocity term is that of the valence band k-state $v_g(k) = \hbar k_z/m_v^*$. We skip the p- or n-subscripts, since it is clear that the electrons tunnel from the valence band. The expression for the tunneling current is then

$$I_T = q\frac{g_s g_v}{L_z}\sum_{(k_x,k_y,k_z)\in\Omega_T}\frac{\hbar k_z}{m_v^*}(f_v - f_c)T_0\exp\left[-\frac{\hbar k_{\perp}^2}{2m_v^*\overline{E}}\right]. \tag{6}$$

The sum over k-states is converted into an integral via the recipe $\sum_k(\ldots) \rightarrow L_xL_yL_z/(2\pi)^3 \times \int dk_x dk_y dk_z(\ldots)$. To evaluate the tunneling current in the restricted volume, we use spherical coordinates $(k_x, k_y, k_z) = (k\sin\theta\cos\phi, k\sin\theta\sin\phi, k\cos\theta)$ to obtain the restricted k-space volume $k^2 \leq k_{\max}^2/(1+\sin^2\theta)$. This relation is representative of the "filtering" brought about by the requirements of energy and momentum conservation. Electrons incident normal to the junction have no transverse momentum. For them $\theta = 0$, and they are allowed to tunnel. The number of electron states allowed to tunnel reduces as their transverse directed momentum increases. The current carried by these states with transverse momentum is further damped by the $\exp[-E_{\perp}/\overline{E}]$ factor, leading to further filtering and momentum collimation.

To evaluate the current, the integral in k-space should be evaluated. To simplify the evaluation in 3-D without losing much accuracy, we assume $f_v - f_c \approx 1$ for the energy window of current-carrying electrons. This relation is exact at 0 K, and remains an excellent approximation even at room temperature. The tunneling current *density* is then given by

$$\begin{aligned} J_T^{3\text{-D}} &= \frac{I_T^{3\text{-D}}}{L_xL_y} \\ &= q\frac{g_s g_v \hbar}{(2\pi)^3 m_v^*}T_0 \times \int_{\phi=0}^{2\pi} d\phi \int_{\theta=0}^{\frac{\pi}{2}} d\theta \sin\theta\cos\theta \\ &\quad \times \int_{k=0}^{\frac{k_{\max}}{\sqrt{1+\sin^2\theta}}} dk\cdot k^3 \exp\left[-\frac{\hbar^2k^2}{2m_v^*\overline{E}}\sin^2\theta\right] \end{aligned} \tag{7}$$

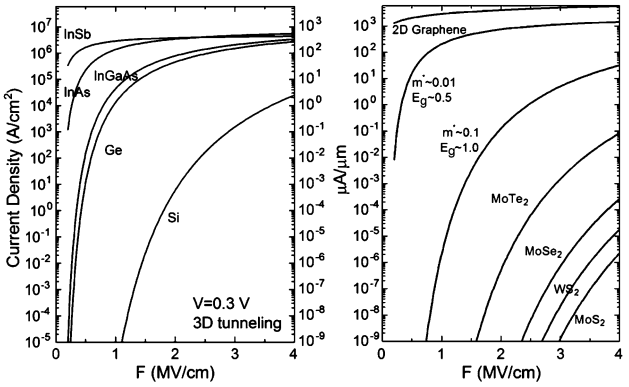

Fig. 5. *Calculated interband tunneling current densities in a few 3-D and 2-D semiconductor crystal $p-n$ junctions. The left figure shows the calculated tunneling current densities in reverse-biased $p-n$ homojunctions. If the current per unit area is assumed constant for a layer thickness of 10 nm, then the effective current per unit width is shown in the right axis of the left plot. This estimation neglects quantization. The right figure shows the calculated tunneling current per unit widths of some 2-D crystals. The transition metal dichalcogenides have low current densities due to high bandgaps, whereas 2-D graphene has the highest current density. Two-dimensional tunneling currents for two small bandgap and effective masses are also shown.*

where the k-space integral is evaluated over the restricted volume Ω_T. The units are in current per unit area (A/cm^2), as it should be. The integral yields an analytical result. Using the symmetric band approximation $m_c^* \approx m_v^* = 2m_R^*$, we get

$$J_T^{3\text{-D}} = \frac{q^2 g_s g_v\sqrt{2m_R^*}F}{8\pi^2\hbar^2\sqrt{E_g}}T_0\left[qV - 2\overline{E}\left\{1-\exp\left(-\frac{qV}{2\overline{E}}\right)\right\}\right] \tag{8}$$

where the symbols have been defined earlier. For extremely small reverse bias voltages $qV \ll 2\overline{E}$, the tunneling current varies as $J_T^{3\text{-D}} \sim V^2$ to leading order. For larger voltages when $qV \gg 2\overline{E}$, $J_T^{3\text{-D}} \sim V$ and this is the condition used in most TFETs. The expression for the tunneling current shows the dependences on various band structure and junction parameters explicitly.

The calculated interband tunneling current densities for 3-D semiconductors are shown in Fig. 5(left) for a reverse bias voltage of 0.3 V. As is evident, the smaller bandgaps of InSb and InAs favor high tunneling current densities that approach $\sim 10^6$ A/cm^2. If we assume that the body thickness of the $p-i-n$ junction is 10 nm, the effective current per unit width is also shown in the right axis of Fig. 5(left). However, this value of the current does not account for the increase in the bandgap due to quantization, which we address shortly. We now apply the same technique for calculating tunneling currents in 2-D crystal semiconductors.

Two-dimensional semiconductors: The same recipe is repeated for 2-D crystals. If the transport is along the x-direction, the transverse momentum component consists of one component k_y, and the restricted k-space volume is given by $2k_y^2 + k_x^2 \leq k_{\max}^2$. The interband tunneling current *per unit width* in a 2-D crystal $p-i-n$ junction then evaluates to

$$J_T^{2\text{-D}} = \frac{qg_s g_v \sqrt{2m_R^* \overline{E}}}{2\pi^2\hbar^2} T_0 \times \left[(qV - \overline{E})\sqrt{\pi}\,\mathrm{Erf}\left[\sqrt{\frac{qV}{2\overline{E}}}\right] + \sqrt{qV \cdot 2\overline{E}}\exp\left[-\frac{qV}{2\overline{E}}\right]\right] \quad (9)$$

where $\mathrm{Erf}[\ldots]$ stands for the error function, and $\overline{E} = q\hbar F/2\sqrt{2m_R^* E_g}$ as before. For extremely small reverse bias voltages $qV \ll 2\overline{E}$, the tunneling current varies as $J_T^{2\text{-D}} \sim V^{3/2}$ to leading order. For larger voltages when $qV \gg 2\overline{E}$, $\mathrm{Erf}[\ldots] \to 1$, and we get a linear dependence of the tunneling current on the reverse-bias voltage $J_T^{2\text{-D}} \approx (q^2 g_s g_v \sqrt{2\pi m_R^* \overline{E}}/2\pi^2\hbar^2) T_0 V$. We note that the units are in current per unit width (mA/μm), as should be the case for 2-D crystals. In quasi-2-D systems, multiple subbands may be involved in transport. Then, we sum the current from each subband with the respective band parameters.

For the special case of 2-D graphene, the band structure is conical, and the bandgap is zero. The interband tunneling probability for a graphene in-plane $p-n$ junction is given by $T(E,\theta) = \exp[-\pi E^2 \sin^2\theta/q\hbar v_F F]$, where θ is the angle between the incident electron momentum and the junction electric field F, and E is the electron energy [70]. The requirement of lateral momentum conservation effectively opens a bandgap proportional to the lateral momentum of electrons. The doping in the p- and n-graphene regions are such that the Fermi level to Dirac point energies are E_{Fp} and E_{Fn}, respectively, and the junction "depletion width" is L_{pn}. The reverse-bias tunneling current in the 2-D graphene $p-n$ junction is then given by [71]

$$J_T^{Gr} = \frac{q^2 V}{\pi^2\hbar}\sqrt{\frac{E_{Fp} + E_{Fn} - qV}{\hbar v_F L_{pn}}}. \quad (10)$$

Let us assume that the applied reverse bias voltage is small compared to the degeneracy energies, and approximate the junction field by $qF \sim (E_{Fp} + E_{Fn})/L_{pn}$. Then, we obtain an approximate expression for the interband reverse-bias tunneling current per unit width in 2-D graphene $p-n$ junctions to be $J_T^{Gr} \sim (q^2\sqrt{qF}/\pi^2\hbar\sqrt{\hbar v_F})V$.

The interband tunneling current densities of various 2-D crystals are plotted in Fig. 5(right). The material constants (bandgaps and effective masses) are obtained from [39]. The values of tunneling currents for transition-metal dichalcogenides are low owing to their large bandgaps. For example, the current density approaches $\sim$0.1 μA/μm for $MoTe_2$ at a high field of 4 MV/cm. The tunneling current density of 2-D graphene is the highest ($\sim$several mA/μm), but it lacks a bandgap. As new 2-D crystals come to the fore, it is desirable to have smaller bandgaps for boosting the current, as indicated by the two curves corresponding to hypothetical 2-D crystals with bandgaps of 0.5 and 1.0 eV, respectively. Such small-bandgap materials could be intrinsic 2-D crystals, or derived from interaction-induced bandgap of Dirac-cone surface states in thin topological insulator materials [29]. Another possibility is in bilayer graphene, where breaking the layer symmetry by vertical electric fields opens a small bandgap [72]–[74]. It is clear that at this stage the currently available TMD family of 2-D crystal semiconductors can enable tunneling transistors. But for in-plane tunneling geometries, the current densities will be low. This feature can be effectively addressed by either narrower gap 2-D crystal semiconductors, or by interlayer tunneling device geometries. We address interlayer tunneling devices in Section V, after discussing the treatment of tunneling in 1-D semiconductors.

One-dimensional semiconductors: For 1-D tunneling, we obtain an exact analytical result even when we include the Fermi–Dirac occupation factors in the source and the grain sides of the $p-i-n$ junction. In the ideal 1-D case, $E_\perp = 0$ since electrons cannot have transverse momentum. When a voltage V is applied, $f_v = 1/(1+\exp[(E-qV)/kT])$ and $f_c = 1/(1+\exp[E/kT])$ are the occupation functions of the source and drain sides. The interband tunneling current is evaluated by the same prescription followed for the 3-D and 2-D cases to be [75]

$$I_T^{1\text{-D}} = \frac{q^2}{h} g_s g_v T_0 \times \frac{kT}{q}\ln\left[\frac{1}{2}\left\{1 + \cosh\left(\frac{qV}{kT}\right)\right\}\right]. \quad (11)$$

Note the explicit appearance of the Landauer conductance in the expression. This expression for tunneling current holds for quasi-1-D semiconductors such as semiconducting nanowires, carbon nanotubes, or semiconducting graphene nanoribbons (GNRs). The appropriate WKB tunneling probability should be used. For nanowires made from conventional 3-D semiconductor crystals, the probability is $T_0 = \exp[-4\sqrt{2m_R^*}E_g^{3/2}/3q\hbar F]$ as before. For CNTs and GNRs, the unconventional band structure is captured in a modified WKB tunneling probability, which is given by $T_0 = \exp[-\pi E_g^2/4q\hbar v_F F]$, where v_F is the Fermi velocity [75]. If there are multiple subbands involved in the transport, we add the currents from each subband with the right bandgap.

Fig. 6 shows the effect of quantization on bandgaps of 3-D crystals on the left, and the calculated 1-D

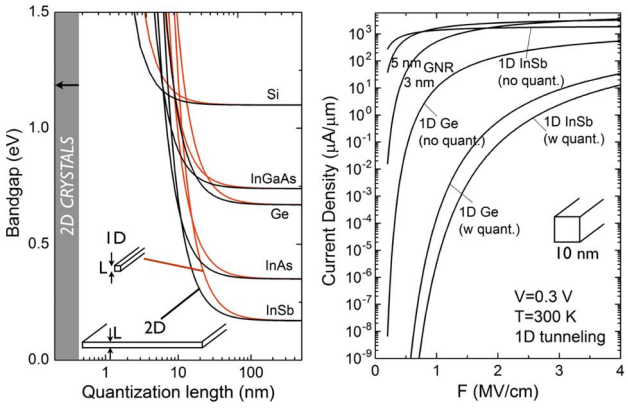

Fig. 6. *The effect of quantization on the bandgap of some 3-D crystals. The plot is generated assuming a particle-in-a-box quantization, and is meant to illustrate the approximate trends. The effect of quantization is the most severe for narrow bandgap semicondctors. The increase in bandgap will reduce interband tunneling currents. The right figure shows the 1-D tunneling currents for GNRs, and InSb and Ge. Note the large reduction of current due to quantization effects in Ge and especially in InSb. A major advantage of 2-D crystals is their inherently thin nature. In addition, their large effective masses make them robust to quantization effects when rendered 1-D.*

semiconductor tunneling current densities on the right. The effect of quantization is to increase the bandgap, in turn reducing the interband tunneling current. The right figure shows the current densities of "1-D" semiconductors such as graphene nanoribbons (GNRs) and Ge and InSb nanowires. The tunneling current densities of GNRs are the highest of all materials calculated that possess bandgaps. The effect of quantization on Ge and InSb nanowire 1-D $p-n$ junction structures is evident from the precipitous drop in the interband tunneling currents in them.

The increase in the bandgap for both 2-D and 1-D confinement is calculated using a simple particle-in-a-box model with the band-edge effective masses of the 3-D semiconductors. The values are meant to be representative of the trends; more accurate electronic structure calculations should be used for direct validation. However, it is clear that as 3-D crystals are scaled in thickness (for making them 2-D) or in diameter (for making them 1-D), the corresponding increase in bandgap is rapid. Large bandgap semiconductors are more robust to quantization since they possess heavier effective masses. This is a dilemma for the scaling of tunneling transistors. As shown in the shaded region in the left of Fig. 6, 2-D crystals are typically of ~nanometer thicknesses and span bandgaps from 0 eV (graphene) to several eVs (BN). This regime remains inaccessible to 3-D crystal semiconductors due to quantization. It is possible to access this regime with 3-D semiconductors only if the band structure allows for extreme anisotropies [76], but such highly desirable properties are yet to be demonstrated in 3-D crystal semiconductors.

IV. TUNNELING TRANSISTORS WITH 3-D CRYSTALS

The unified view of tunneling transport discussed in the last section provides a framework for comparative studies of the effect of dimensionality on tunneling transistors. Based on the discussion of transport in two-terminal tunnel junctions, we now discuss the electrostatics and device embodiments of the corresponding three-terminal TFETs.

In a TFET, a gate terminal electrostatically controls the energy-band alignment of the $p-n$ junction, as indicated in Fig. 7. In the OFF-state of the device, electrons in the valence band of the source are energetically forbidden to tunnel to the drain since the channel length exponentially damps the direct source-to-drain tunneling probability. To turn the device on, the gate pushes the channel bands to align the conduction band edge of the channel region with the valence band of the source. Electrons can now tunnel through the tunneling barrier, which is much smaller than the OFF-state. The goal therefore is to allow a large current to flow in the ON-state, while cutting the current off as much as possible in the OFF-state.

The performance requirements of a TFET are indicated schematically in Fig. 7. Compared to a MOSFET, the steeper SS slope of a TFET enables a higher ON-current at a smaller gate overdrive voltage. This feature is expected to enable scaling of the voltage supply V_{DD} to lower values while maintaining a substantial ON/OFF ratio. The issues of electrostatics and transport have been discussed at length in various articles [12], [69]. We refer the reader to these articles for detailed historical perspectives and further technical details. Here, we qualitatively discuss a few embodiments and issues with TFETs realized with 3-D crystal semiconductors. The discussion naturally motivates the case for 2-D crystal realizations of the device.

The electric field lines emanating from the gate metal of a TFET need to access the $p-n$ junction. Therein lies a dilemma for TFETs based on 3-D crystal semiconductors. As shown in Fig. 8, if the tunneling current flows in the lateral direction and the gate field is vertical, the channel needs to be thinned down to exercise substantial electrostatic control over the entire junction thickness. As the

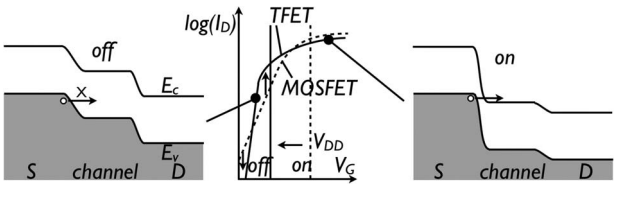

Fig. 7. *TFET operation and requirements. The left figure shows the OFF-state energy-band diagram of the TFET along the tunneling direction. The right figure shows the ON-state. The channel band is controlled by the gate. TFETs are expected to lower the supply voltage V_{DD} since a steeper SS swing leads to a higher ON-current at a smaller voltage, as shown in the middle.*

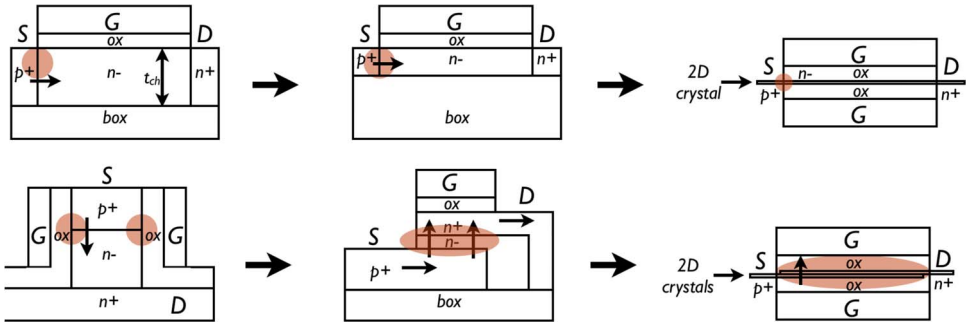

Fig. 8. *Schematic representation of various topologies of TFETs. The top row shows TFETs where the tunneling current flows laterally in the $p^+ - n^- - n^+$ junction. The circles are indications of the region in space where most of the interband tunneling current flows. Since the gate is on the top, parts of the junction farther away from it are not effectively gated in the top left TFET. The top middle geometry is the same as the left, but with a thinner channel for more uniform electrostatic gate control of tunneling current. The right figure on the top row is the 2-D crystal realization of the lateral TFET. As the channel thickness is reduced in 3-D semiconductors, quantum confinement increases the bandgap and reduces the tunneling current. This is avoided in 2-D crystals. To increase the net current, vertical TFETs are being considered. The bottom row indicates some realizations of TFETs in which the tunneling current flows vertically. The left figure shows a side-gate geometry, and the middle figure is a geometry in which the current flow is not over a "line," but an "area," as shown by the shaded ellipse. The right figure shows the realization of a vertical double-gate TFET with p^+ and n^+ 2-D crystal layers. It highlights the electrostatic advantage and simplicity.*

channel thickness is scaled down, quantum confinement increases the bandgap, and thus the interband tunneling current reduces (see Fig. 6). A 2-D crystal does not suffer from such a problem, and thus offers a way to fight quantization effects. In addition, it offers a solution to surface state related trap states, and simpler integration of double-gate geometries, as indicated in Fig. 8.

A number of TFETs with subthreshold slopes less than the SS limit of 60 mV/decade have been demonstrated, proving the feasibility of the concept. Such devices have been made with 3-D crystal semiconductors (Si, Ge, etc.) as well as with carbon nanotubes [77]–[80]. However, for most realizations, the ON-state current falls below the $\sim$1 mA/μm range necessary for high-performance operation. Low ON-current TFETs can enable various new applications where performance (speed) requirements are not as critical as the requirement of low power consumption. For high-performance TFETs, various approaches are being pursued to increase the ON-current. These approaches involve using heterojunctions that have staggered or broken-gap band alignments, or through changes in the device topology. An approach based on the device topology is indicated in Fig. 8.

The shaded regions in Fig. 8 indicate the location of current flow. To increase the tunneling current per unit width, it is necessary to increase the net *area* of tunneling current flow. The vertical geometries shown in Fig. 8 allow this change [81], [82]. The gate field effect is in the *same* direction as the tunneling current flow in such devices. The tunneling current follows a nonlinear path (shaped like an "S") laterally from the source, vertically into the drain, and then out laterally into the drain. The device geometry requires careful processing. For this geometry, two layers of 2-D crystals, one doped *p*-type and the other *n*-type, promise efficient vertical scaling and electrostatic control as shown in the figure. It may also enable a simplification of the processing requirements.

V. TUNNELING TRANSISTORS WITH 2-D CRYSTALS

The 2-D crystal realizations of TFETs discussed here involve *in-plane* tunneling for the lateral device and *interlayer* tunneling in the vertical TFET. We discuss them in greater detail here. Note that due to the relatively early phase of material development, we estimate and project the performance advantages in cases where experimental results are not available yet.

A. In-Plane Tunneling: 2-D Crystal Semiconductors

The in-plane interband tunneling currents calculated in Fig. 5 show that smaller bandgap semiconductor 2-D crystals are required for boosting the ON-state current of the devices. Low-power TFETs are realizable with the transition-metal dichalcogenide semiconductors. The effective masses of the conduction and valence band edges of TMD 2-D crystals have been calculated to be rather symmetric. For example, the electron effective mass of MoS_2 is $\sim$0.57, and the hole effective mass is $\sim$0.66 [39]. The symmetry in the band structure is expected to lead to symmetric performance of nTFETs and pTFETs, which would be essential for complementary logic circuits.

It has been found that multilayer versions of TMD 2-D crystals have smaller bandgaps than the single-layer counterpart, and are generally of indirect bandgap nature [40], [41]. This is also true when one considers single-layer graphene (direct bandgap) and graphite (which is a semimetal). Furthermore, it has recently been reported that

carrier inversion can be achieved in multilayer TMD crystals by the field effect. A hole channel was observed in a nominally *n*-type layered semiconductor [64]. Consider a few-layer stack of 2-D TMD crystals. By using two gates, it is possible to create an electron channel at one interface and a hole channel in the other. These channels can be placed several nanometers apart by controlling the number of layers. The wave function overlap between these states is small at no bias owing to the high effective mass for carrier motion between planes. The geometry then allows for a TFET similar to the vertical structure shown in Fig. 8, but without the need to chemically dope the individual layers. A major challenge in such structures is in the formation of ohmic contacts to the individual layers. Note that such a device has also been recently proposed for thin layer Si [83]. The realization with multilayer version of 2-D crystals can be an alternative approach that can leverage the robustness against quantization effects, and relative insensitivity to surface and interface trap effects.

B. In-Plane Tunneling: 2-D Graphene

As shown in Fig. 5, the ON-state interband tunneling current density in 2-D graphene is the highest due to the absence of a bandgap. For the same reason, it is difficult to obtain the OFF-state condition using monolayer 2-D graphene. Field-tunable bandgaps in bilayer graphene have been proposed as a possible approach to achieving ON/OFF ratios in TFETs [84]. There have also been recent reports of the observation of negative differential resistance in monolayer 2-D graphene FETs [85]. The proposed mechanism responsible for such behavior relies entirely on gate electrostatics and the unique band structure with the zero-gap nature of 2-D graphene. More experimental work and understanding of NDR mechanisms in 2-D graphene can lead to useful device applications in the analog arena to complement TFETs. To decrease the OFF-state current for in-plane tunneling devices, it is necessary to create bandgaps in graphene. One approach is to use CNTs or lithographically patterned GNRs, which is discussed next.

C. In-Plane Tunneling: CNTs and GNRs

One of the early reports of sub-60-mV/decade SS slope TFET behavior was observed in semiconducting carbon nanotubes at room temperature [80]. Analysis of the device performance for CNT TFETs [86] and GNR TFETs [87] shows that they are attractive for desirable ON-currents, ON/OFF ratios, and sub-60-mV/decade SS slopes. CNTs do not have edge states, and are the most attractive from a performance viewpoint. Bandgap control, chemical doping, and patterned assembly on large wafers still remain challenging for CNTs, though rapid progress is being made [88].

Their close cousins, GNRs are also highly attractive candidates for TFETs. For example, Fig. 9 shows the device structure, energy band diagrams, and the projected device characteristics of complementary GNR TFETs. The inclusion of parasitic elements to the intrinsic model still maintains a high performance. GNRs can be integrated on planar surfaces, and can be made lithographically in

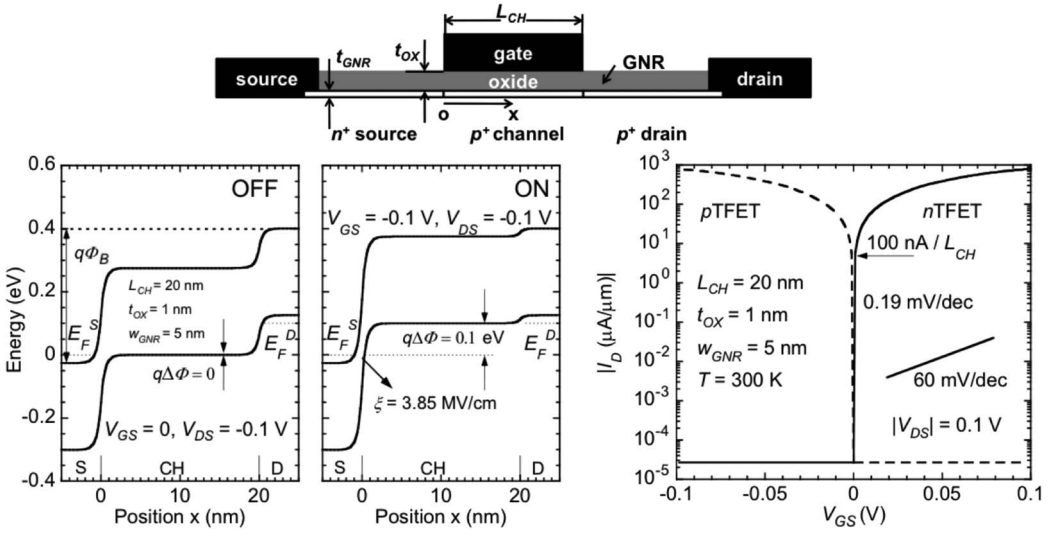

Fig. 9. *A proposed GNR TFET geometry, energy band diagram, and the calculated transistor transfer curves. The device structure consists of a GNR p – n junction that is gated through an insulator from the top gate. The energy band diagrams are for a 20-nm-long channel device with a 5-nm-wide GNR. The energy band diagrams indicate the OFF- and ON-states of the device, where the channel potential is moved with the gate voltage. The resulting transfer curve shows a high ON-current, a low OFF-current, and a low SS slope, below the 60-mV/decade limit. Though the calculations are for an ideal case, they represent the attractiveness of GNRs as possible candidates for TFETs. The figure has been adapted from [87].*

parallel arrays to boost the net current in a realistic TFET device geometry. The available dangling bonds at the edges can be used to chemically dope them; initial reports indicate this possibility [89]. The major challenges at this stage for the realization of GNR TFETs lie in the narrowness of the GNR widths necessary to avail high-performance levels. The energy bandgap of a semiconducting GNR of width W is $E_g \sim 1.4/W$ eV, where W is in nanometers. Based on theoretical estimates, GNRs of widths ≤ 10 nm are necessary. The line-edge roughness that might result from process variations for the thinnest GNRs can degrade the performance of GNR TFETs, as has been analyzed in [90]. On the other hand, advances in process control in the fabrication of thin films in Si FinFETs can be effectively leveraged for fabrication of wafer-scale GNRs. A number of variants of the GNR TFETs have also been proposed to improve the device performance [91]–[95].

The symmetry of the band structure of CNTs and GNRs is a major advantage that allows for the realization of nTFETs and pTFETs on equal footing. Combined with the scaling advantages that stem from their atomically thin body nature, they are highly desirable for nanoscale TFETs. The approach to high-performance TFETs using 3-D crystal semiconductors is taking the path toward materials with successively smaller bandgaps to increase the ON-current. The approach with graphene, CNTs, and GNRs is from the other extreme, where we start from zero bandgap and very high ON-currents, and now need to open bandgaps controllably to lower the OFF-current. While this is an attractive and complementary approach, 2-D crystals also offer the possibility of *interlayer* tunneling transistors, which we discuss now.

D. Interlayer Tunneling Devices, BiSFETs, and SymFETs

Electron tunneling out of the plane of a 2-D crystal is under intense scrutiny at this time [96]. The electronic band structure of the 2-D crystal is defined in the plane but not out of it. The conventional approach to tunneling calculations requires the knowledge of band parameters such as the effective mass of the evanescent band structure in the direction of the tunneling. Since this feature is not well defined for 2-D crystals, it is more feasible to use scattering rate formalisms for quantitative calculations of interlayer tunneling. The Bardeen transfer-Hamiltonian approach, used in scanning tunneling microscopy [97], [98] and in superconducting Josephson junctions [99] allows such evaluation. We do not derive the quantitative results here, but refer the reader to recent articles that approach the subject of interlayer tunneling using the Bardeen method.

A prototype interlayer-tunneling device is a graphene–insulator–graphene (GIG) junction. In a recent work [100], the interlayer tunneling current in such a GIG junction was explicitly evaluated using the Bardeen method. The predicted I–V characteristics are rather remarkable, and highlight the strong role of the symmetry of the band structure of graphene.

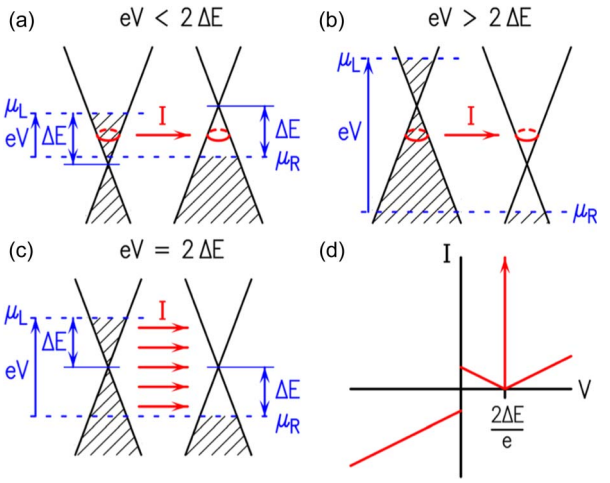

Fig. 10. *Band alignments of GIG interlayer tunnel junctions under various bias conditions from [100]. The graphene layers are doped to form a $p-n$ junction. In (a) and (b), the symmetry of the band structure restricts electrons at only one energy to carry interlayer current due to the requirement of transverse momentum conservation. A special case occurs when the Dirac points align: electrons at all energies are now allowed to tunnel, leading to a spike in the current, as shown schematically in (d).*

Fig. 10 shows the energy band alignments and projected device performance of a GIG interlayer tunnel junction device. The two graphene layers are "independent" in the sense that they do not form a bilayer, and they are doped p- and n-type as captured by their Fermi level degeneracies. Ohmic contacts are made to the two layers independently. A voltage is applied across the junction. When the Dirac points of the two layers are misaligned, a small interlayer tunneling current flows. The circles indicated on the Dirac cones in Fig. 10(a) and (b) show the states that participate in the interlayer tunneling process. Electrons that have energy halfway between the Dirac points carry the current. This is because transverse momentum conservation requires the radii of the iso-energy circles to be the same in both layers.

However, at the particular voltage when the Dirac points align, as shown in Fig. 10(c), electrons at *all* energies are now allowed to tunnel, leading to a large spike in the current. This is schematically shown in Fig. 10(d) as a Dirac-delta function. A quantitative evaluation leads to broadening, but with a very large NDR effect. Note that the peak would be much smaller if the band structure was not symmetric. Since then the requirement of transverse momentum conservation would restrict the current to flow at a particular energy, and a collective tunneling condition as in Fig. 10(c) cannot be achieved. The large tunneling current peak is a direct consequence of the symmetric band structure of 2-D graphene.

The GIG $p-n$ junction structure can be connected to gates to realize an interlayer-tunneling transistor. Such a device, called the symmetric-FET (SymFET) has been recently proposed [55]. In addition to performing logic operations, the inherently fast tunneling feature and large NDR promises to also enable analog applications such as high-harmonic generation, and high-speed oscillator design. Note that the SymFET device structure is similar to gated RTD structures [101] realized in 3-D semiconductor heterostructures, but takes advantage of the band structure symmetry of graphene to deliver a stronger NDR behavior. The first experimental report of such a structure did not show NDR, but exhibited TFET-like behavior with a few orders ON/OFF ratio at room temperature. The structure used consisted of graphene–BN–graphene and graphene–MoS_2–graphene heterostructures [53].

The SymFET structure is based on single-particle tunneling. Realistic fabrication of the device calls for rotational alignment of the graphene layers. By adjusting the interlayer distance and the carrier densities, the Coulombic forces between the electrons and the holes in the two graphene layers can be made strong enough to form excitonic quasi-particles. Under suitable bias conditions, the interlayer current flow can take a collective many-body form triggered by a Bose–Einstein condensation of the excitons. The condensate can boost the interlayer current significantly. The proposed device, called the bilayer pseudospin FET (BiSFET) is insensitive to the rotational alignment of the two graphene layers. It has been shown that if the BISFET can be realized, it can perform digital logic by consuming many orders of magnitude lower energy than conventional MOSFETs [54]. The SymFET and the BiSFET are fundamentally new types of devices with no direct analogs to conventional semiconductors. This is because of their unique band structures and their 2-D crystal nature. Their discussion is an ideal point to end this review paper and to wrap up with a few concluding remarks.

VI. FUTURE PERSPECTIVES AND CONCLUSION

The emergence of 2-D crystal materials has marked a new phase for the development of semiconductor devices. It may rank at the same level as the origin and proliferation of heterostructures in 3-D semiconductors. The materials and the resulting devices are at their infancy, as are many device ideas based on tunneling that are at proposal stages. But the novelty the family of 2-D crystal has brought to the field becomes evident by the string of new device concepts based on tunneling. The addition of graphene with its unique band structure, BN as a 2-D crystal insulator, and transition-metal dichalcogenides with material properties ranging from semiconducting to metallic and superconducting casts a much wider net than has been possible with conventional materials. The possibility of integration of diverse material properties in 2-D crystal heterostructures has breathed new life into existing paradigms of electronic device technologies. This is an exciting time when creative ideas are needed to exploit the power of this new material system. Though it is impossible to predict the exact path forward, we can be sure that electronic devices that go far beyond the current state of the art will result from the new material family. ■

Acknowledgment

The author acknowledges discussions and exchange of ideas with collaborators over the past few years on the topic of 2-D crystals and tunneling devices.

REFERENCES

[1] M. Riordan and L. Hoddeson, *Crystal Fire*. New York, NY, USA: Norton, 1998, pp. 88–90.

[2] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed. New York, NY, USA: Wiley-Interscience, 2006.

[3] J. S. Kilby, "Turning potential into realities: The invention of the integrated circuit," *Nobel Lecture*, 2000.

[4] Z. I. Alferov, "Nobel lecture: The double heterostructure concept and its applications in physics, electronics, and technology," *Rev. Mod. Phys.*, vol. 73, pp. 767–782, Oct. 2001.

[5] H. Kroemer, "Nobel lecture: Quasielectric fields and band offsets: Teaching electrons new tricks," *Rev. Mod. Phys.*, vol. 73, pp. 783–793, Oct. 2001.

[6] D. H. Kim, B. Brar, and J. A. del Alamo, "$f_T = 688$ GHz and $f_{max} = 800$ GHz in $L_g = 40$ nm $In_{0.7}Ga_{0.3}As$ MHEMTs with $g_m(max) > 2.7$ mS/μm," in *IEEE Int. Electron Devices Meeting Tech. Dig.*, 2011, pp. 319–322.

[7] S. Pimputkar, J. S. Speck, S. P. DenBaars, and S. Nakamura, "Prospects for LED lighting," *Nature Photon.*, vol. 3, pp. 180–182, 2009.

[8] J. D. Joannopoulus, S. G. Johnson, and J. N. Winn, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. Princeton, NJ, USA: Princeton Univ. Press, 2008.

[9] W. Cai and V. Shalaev, *Optical Metamaterials: Fundamentals and Applications*, 1st ed. New York, NY, USA: Springer-Verlag, 2009.

[10] S. Christopoulos, G. Baldassarri Hoger von Hogersthal, A. J. D. Grundy, P. G. Lagoudakis, A. V. Kavokin, J. J. Baumberg, G. Christman, R. Butte, E. Feltin, J.-F. Carlin, and N. Grandjean, "Room-temperature polariton lasing in semiconductor microcavities," *Phys. Rev. Lett.*, vol. 98, pp. 126405-1–126405-4, Mar. 2007.

[11] T. N. Theis and P. M. Solomon, "It's time to reinvent the transistor!" *Science*, vol. 327, pp. 1600–1601, Mar. 2010.

[12] A. M. Ionescu and H. Riel, "Tunnel field-effect transistors as energy-efficient electronic switches," *Nature*, vol. 479, pp. 329–337, Nov. 2011.

[13] T. N. Theis and P. M. Solomon, "In quest of the "next switch": Prospects for greatly reduced power dissipation in a successor to the silicon field-effect transistor," *Proc. IEEE*, vol. 98, no. 12, pp. 2005–2014, Dec. 2010.

[14] T.-J. King, D. Markovic, V. Stojanovic, and E. Elon, "MEMS switches for low-power logic," *IEEE Spectrum*, pp. 38–43, Apr. 2012.

[15] K. Gopalakrishnan, P. B. Griffin, and J. D. Plummer, "I-MOS: A novel semiconductor device with a subthreshold slope lower than kT/q," in *IEEE Int. Electron Devices Meeting Tech. Dig.*, 2002, pp. 289–292.

[16] M. T. Bjork, O. Hayden, H. Schmid, H. Riel, and W. Reiss, "Vertical surround-gated silicon nanowire impact ionization field-effect transistors," *Appl. Phys. Lett.*, pp. 142 110–142 292, 2007.

[17] S. Salahuddin and S. Datta, "Use of negative capacitance to provide voltage amplification for low power nanoscale devices," *Nano. Lett.*, vol. 8, pp. 405–410, 2008.

[18] D. Newns, B. Elmegreen, X. H. Liu, and G. Martyna, "A low-voltage high-speed electronic switch based on piezoelectric

transduction," *J. Appl. Phys.*, vol. 111, pp. 084509-1–084509-18, 2012.

[19] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric field effect in atomically thin carbon films," *Science*, vol. 306, pp. 666–669, 2004.

[20] Y. Zhang, Y. W. Tan, H. Stormer, and P. Kim, "Experimental observation of the quantum Hall effect and Berry's phase in graphene," *Nature*, vol. 438, pp. 201–204, 2005.

[21] C. Berger, Z. Song, X. Li, X. Wu, N. Brown, C. Naud, D. Mayou, T. Li, J. Hass, A. N. Marchenkov, E. H. Conrad, P. N. First, and W. A. de Heer, "Electronic confinement and coherence length in patterned epitaxial graphene," *Science*, vol. 312, pp. 1191–1196, 2006.

[22] K. S. Novoselov, D. Jiang, F. Schedin, T. J. Booth, V. V. Khotkevich, S. V. Morozov, and A. K. Geim, "Two-dimensional atomic crystals," *Proc. Nat. Acad. Sci.*, vol. 102, pp. 10451–10453, 2005.

[23] D. J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur, and H.-S. P. Wong, "Device scaling limits of Si MOSFETs and their application dependencies," *Proc. IEEE*, vol. 89, no. 3, pp. 259–288, Mar. 2001.

[24] A. I. Yanson, G. R. Bollinger, H. E. van den Brom, N. Agrait, and J. M. van Ruitenbeek, "Formation and manipulation of a metallic wire of single gold atoms," *Nature*, vol. 395, pp. 783–785, 1998.

[25] R. H. Friend and A. D. Yoffe, "Electronic properties of intercalation complexes of the transition metal dichalcogenides," *Adv. Phys.*, vol. 36, pp. 1–94, 1987.

[26] P. Vogt, P. D. Padova, C. Quaresima, J. Avila, E. Frantzeskakis, M. C. Asensio, A. Resta, B. Ealet, and G. Le Lay, "Silicene: Compelling experimental evidence for graphene-like two-dimensional silicon," *Phys. Rev. Lett.*, vol. 108, pp. 1055501-1–1055501-5, 2012.

[27] Z. Ni, Q. Liu, K. Tang, J. Zheng, J. Rhou, R. Qin, Z. Gao, D. Yu, and J. Lu, "Tunable bandgap in silicene and germanene," *Nano Lett.*, vol. 12, pp. 113–118, 2012.

[28] C. L. Freeman, F. Claeyssens, N. L. Allan, and J. H. Harding, "Graphitic nanofilms as precursors to Wurtzite films: Theory," *Phys. Rev. Lett.*, vol. 96, pp. 066102-1–066102-4, 2006.

[29] J. Chang, L. F. Register, and S. K. Banerjee, "Possible applications of topological insulator thin films for tunnel FETs," in *IEEE Device Res. Conf. Tech. Dig.*, 2012, pp. 31–32.

[30] H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley, "C_{60}: Buckminsterfullerene," *Nature*, vol. 318, pp. 162–163, 1985.

[31] M. Atiyah and P. Sutcliffe, "Polyhedra in physics, chemistry and geometry," *Milan J. Math.*, vol. 71, pp. 33–58, 2003.

[32] R. Wang, Q. Zhang, W. Li, I. Calizo, T. Shen, C. A. Richter, A. R. Hight-Walker, X. Liang, A. Seabaugh, D. Jena, H. G. Xing, D. J. Gundalch, and N. V. Nguyen, "Determination of graphene work function and graphene-insulator-semiconductor band alignment by internal photoemission spectroscopy," *Appl. Phys. Lett.*, vol. 101, pp. 022105-1–022105-4, 2012.

[33] W. Monch, "Valence-band offsets and Schottky barrier heights of layered semiconductors explained by interface-induced gap states," *Appl. Phys. Lett.*, vol. 72, pp. 1899–1901, 1998.

[34] R. Schlaf, O. Lang, C. Pettenkofer, and W. Jaegermann, "Band lineup of layered semiconductor heterointerfaces prepared by van der Waals epitaxy: Charge transfer correction term for the electron affinity rule," *J. Appl. Phys.*, vol. 85, pp. 2732–2753, 1999.

[35] T. Fang, A. Konar, H. Xing, and D. Jena, "Carrier statistics and quantum capacitance in graphene sheets and ribbons," *Appl. Phys. Lett.*, vol. 91, pp. 092109-1–092109-3, 2007.

[36] K. Watanabe, T. Taniguchi, and H. Kanda, "Direct-bandgap properties and evidence for ultraviolet lasing of hexagonal boron nitride single crystal," *Nature Mater.*, vol. 3, pp. 404–409, 2004.

[37] K. Wood and J. B. Pendry, "Layer method for band structure of layer compounds," *Phys. Rev. Lett.*, vol. 31, pp. 1400–1403, 1973.

[38] K. Kaasbjerg, K. S. Thygesen, and K. W. Jacobsen, "Phonon-limited mobility in n-type single-layer MoS_2 from first principles," *Phys. Rev. B*, vol. 85, pp. 115317-1–115317-16, 2012.

[39] L. Liu, S. B. Kumar, and J. Guo, "Performance limits of monolayer transition metal dichalcogenide transistors," *IEEE Trans. Electron Devices*, vol. 58, no. 9, pp. 3042–3047, Sep. 2011.

[40] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, "Atomically thin MoS_2: A new direct-gap semiconductor," *Phys. Rev. Lett.*, vol. 105, pp. 136805-1–136805-4, 2010.

[41] A. Splendiani, L. Sun, Y. Zhang, T. Li, J. Kim, C.-Y. Chim, G. Galli, and F. Wang, "Emerging photoluminescence in monolayer MoS_2," *Nano Lett.*, vol. 10, pp. 1271–1275, 2010.

[42] D. Hisamoto, W. C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T. J. King, J. Bokor, and C. Hu, "FinFET-A self-aligned double-gate MOSFET scalable to 20 nm," *IEEE Trans. Electron Devices*, vol. 47, no. 12, pp. 2320–2325, Dec. 2000.

[43] G. Tsutsui, M. Saitoh, and T. Hiramoto, "Experimental study on superior mobility in [110]-oriented UTB SOI pMOSFETs," *IEEE Electron Device Lett.*, vol. 26, no. 11, pp. 836–838, Nov. 2005.

[44] R. S. Muller, T. I. Kamins, and M. Chan, *Device Electronics for Integrated Circuits*, 3rd ed. New York, NY, USA: Wiley, 2003, pp. 485–487.

[45] H. Sakaki, T. Noda, K. Hirakawa, and T. Matsusue, "Interface roughness scattering in GaAs/AlAs quantum wells," *Appl. Phys. Lett.*, vol. 51, pp. 1934–1936, 1987.

[46] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, "Single-layer MoS_2 transistors," *Nature Nanotechnol.*, vol. 6, pp. 147–150, 2011.

[47] D. Jena and A. Konar, "Enhancement of carrier mobility in semiconductor nanostructures by dielectric engineering," *Phys. Rev. Lett.*, vol. 98, pp. 136805-1–136805-4, 2007.

[48] C. Jang, S. Adam, J. H. Chen, E. D. Williams, S. Das Sarma, and M. S. Fuhrer, "Tuning the effective fine structure constant in graphene: Opposing effects of dielectric screening on short- and long-range potential screening," *Phys. Rev. Lett.*, vol. 101, pp. 146805-1–146805-4, 2008.

[49] A. K. M. Newaz, Y. S. Puzyrev, B. Wang, S. Pantelides, and K. I. Bolotin, "Probing charge scattering mechanisms in suspended graphene by varying its dielectric environment," *Nature Commun.*, vol. 3, pp. 7341–7346, 2012.

[50] H. Kroemer, "Two integral relations pertaining to the electron transport through a bipolar transistor with a nonuniform energy gap in the base region," *Solid State Elecron.*, vol. 28, pp. 1101–1103, 1985.

[51] L. Ci, L. Song, C. Jin, D. Jariwala, D. Wu, Y. Li, A. Srivastava, Z. F. Wang, K. Storr, L. Balicas, F. Liu, and P. M. Ajayan, "Atomic layers of hybridized boron nitride and graphene domains," *Nature Mater.*, vol. 9, pp. 430–435, 2010.

[52] H. Yang, J. Heo, S. Park, H. J. Song, D. H. Seo, K. E. Byun, P. Kim, I. Yoo, H. J. Chung, and K. Kim, "Graphene barristor, a triode device with a gate-controlled Schottky barrier," *Science*, vol. 336, pp. 1140–1143, 2012.

[53] L. Britnell, R. V. Gorbachev, R. Jalil, R. D. Belle, F. Schedin, A. Mischenko, T. Georgiou, M. I. Katsnelson, L. Eaves, S. V. Morozov, N. M. R. Peres, J. Leist, A. K. Geim, K. S. Novoselov, and L. A. Ponomarenko, "Field-effect tunneling transistor based on vertical graphene heterostructures," *Science*, vol. 335, pp. 947–950, 2012.

[54] S. K. Banerjee, L. F. Register, E. Tutuc, D. Reddy, and A. H. MacDonald, "Bilayer pseudospin field-effect transistor (BiSFET): A proposed new logic device," *IEEE Electron Device Lett.*, vol. 30, no. 2, pp. 158–160, Feb. 2009.

[55] P. Zhao, R. M. Feenstra, G. Gu, and D. Jena, "SymFET: A proposed symmetric graphene tunneling field-effect transistor," in *IEEE Device Res. Conf. Tech. Dig.*, 2012, pp. 33–34.

[56] G. Gu, S. Nie, R. M. Feenstra, R. P. Devaty, W. J. Choyke, W. K. Chan, and M. G. Kane, "Field effect in epitaxial graphene on a silicon carbide substrate," *Appl. Phys. Lett.*, vol. 90, pp. 253507-1–253507-3, 2007.

[57] J. S. Moon, D. Curtis, M. Hu, D. Wong, C. McGuire, P. M. Campbell, G. Jernigan, J. L. Tedesco, B. VanMil, R. Myers-Ward, C. Eddy, and D. K. Gaskill, "Epitaxial-graphene RF field-effect transistors on Si-face 6H-SiC substrates," *IEEE Electron Device Lett.*, vol. 30, no. 6, pp. 650–652, Jun. 2009.

[58] X. Li, W. Cai, S. Kim, J. Nah, D. Yang, R. Piner, A. Velamakanni, I. Jung, E. Tutuc, S. K. Banerjee, L. Colombo, and R. S. Ruoff, "Large-area synthesis of high-quality and uniform graphene films on copper foils," *Science*, vol. 324, pp. 1312–1314, 2009.

[59] W. S. Hwang, K. Tahy, X. Li, H. G. Xing, A. C. Seabaugh, C. Y. Sung, and D. Jena, "Transport properties of graphene nanoribbon transistors on chemical-vapor-deposition grown wafer-scale graphene," *Appl. Phys. Lett.*, vol. 100, pp. 203107-1–203107-3, 2012.

[60] L. Song, L. Ci, H. Lu, P. Sorokin, C. Jin, J. Ni, A. G. Kvashnin, D. G. Kvashnin, J. Lou, B. I. Yakobson, and P. M. Ajayan, "Large scale growth and characterization of atomic hexagonal boron nitride layers," *Nano Lett.*, vol. 10, pp. 3209–3215, 2010.

[61] Y. H. Lee, X. Q. Zhang, W. Zhang, M. T. Chang, C. T. Lin, K. D. Chang, Y. C. Yu, J. T. W. Wang, C. S. Chang, L. J. Li, and T. W. Lin, "Synthesis of large-area MoS_2 atomic layers with chemical vapor deposition," *Adv. Mater.*, vol. 24, pp. 2320–2325, 2012.

[62] W. S. Hwang, M. Remskar, R. Yan, V. Protasenko, K. Tahy, S. D. Chae, P. Zhao, A. Konar, H. G. Xing, A. C. Seabaugh, and D. Jena, "Transistors with chemically synthesized layered semiconductor WS_2 exhibiting 10^5 room temperature modulation and ambipolar behavior," *Appl. Phys. Lett.*, vol. 101, pp. 013107-1–013107-4, 2012.

[63] J. M. Martin, C. Donnet, T. Le Mogne, and T. Epicier, "Superlubricity of molybdenum disulphide," *Phys. Rev. B.*, vol. 48, pp. 10583–10586, 1993.

[64] S. Kim, A. Konar, W. S. Hwang, J. H. Lee, J. Lee, J. Yang, C. Jung, H. Kim, J. B. Yoo, J. Y. Choi, Y. W. Jin, S. Y. Lee, D. Jena, W. Choi, and K. Kim, "High-mobility, low-power thin-film transistors based on multilayer MoS_2 crystals *Nature Commun.*, article 1011, DOI: 10.1038/ncomms2018.

[65] C. Zener, "A theory of the electrical breakdown of solid dielectrics," *Proc. R. Soc. Lond. A.*, vol. 145, pp. 523–529, 1934.

[66] C. Wittig, "The Landau-Zener formula," *J. Phys. Chem. B*, vol. 109, pp. 8428–8430, 2005.

[67] R. Shankar, *Principles of Quantum Mechanics*, 2nd ed. New York, NY, USA: Plenum, 1994, pp. 435–449.

[68] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed. New York, NY, USA: Wiley-Interscience, 2006.

[69] A. C. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond CMOS logic," *Proc. IEEE*, vol. 98, no. 12, pp. 2095–2110, Dec. 2010.

[70] V. Cheianov and V. I. Fal'ko, "Selective transmission of Dirac electrons and ballistic magnetoresistance of n-p junctins in graphene," *Phys. Rev. B.*, vol. 74, pp. 041403(R)-1–041403(R)-4, 2006.

[71] T. Fang, "Carrier transport in graphene, graphene nanoribbons, and GaN HEMTs," Ph.D. dissertation, Dept. Electr. Eng., Univ. Notre Dame, Notre Dame, IN, USA, 2012.

[72] E. McCann, "Asymmetry gap in the electronic band structure of bilayer graphene," *Phys. Rev. B*, vol. 74, pp. 161403(R)-1–161403(R)-4, 2006.

[73] Y. Zhang, T. T. Tang, C. Girit, Z. Hao, M. C. Martin, A. Zettl, M. F. Crommie, Y. R. Shen, and F. Wang, "Direct observation of a widely tunable bandgap in bilayer graphene," *Nature*, vol. 459, pp. 820–823, 2009.

[74] F. Xia, D. B. Farmer, Y. M. Lin, and P. Avouris, "Graphene field-effect transistors with high on/off current ratio and large transport band gap at room temperature," *Nano Lett.*, vol. 10, pp. 715–718, 2010.

[75] D. Jena, T. Fang, Q. Zhang, and H. Xing, "Zener tunneling in semionducting nanotube and graphene nanoribbon p-n junctions," *Appl. Phys. Lett.*, vol. 93, pp. 112106-1–112106-3, 2008.

[76] M. Rodwell, W. Frensley, S. Steiger, E. Chagarov, S. Lee, H. Ryu, Y. Tan, G. Hegde, L. Wang, J. Law, T. Boykin, G. Klimek, P. Asbeck, A. Kummel, and J. N. Schulman, "III-V FET channel designs for high current densities and thin inversion layers," in *IEEE Device Res. Conf. Tech. Dig.*, 2010, pp. 149–152.

[77] W. Y. Choi, B. G. Park, J. D. Lee, and T. K. K. Liu, "Tunneling field-effect transistors (TFETs) with subthreshold swing (SS) less than 60 mV/dec," *IEEE Electron Device Lett.*, vol. 28, no. 8, pp. 743–745, Aug. 2007.

[78] S. H. Kim, H. Kam, C. Hu, and T. K. K. Liu, "Germanium-source tunnel field-effect transistors with record high Ion/Ioff," in *Proc. Very Large Scale Integr. (VLSI) Technol. Symp.*, 2009, pp. 178–179.

[79] F. Mayer, C. L. Oryer, J. F. Damlencourt, K. Romanjek, F. Andrieu, C. Tabone, B. Previtali, and S. Deleonibus, "Impact of SOI,$Si_{1-x}Ge_x$OI and GeOI substrates on CMOS compatible tunnel FET performance," in *IEEE Int. Electron Device Meeting Tech. Dig.*, 2008, pp. 163–166.

[80] J. Appenzeller, Y. M. Lin, J. Knoch, and P. Avouris, "Band-to-band tunneling in carbon nanotube field-effect transistors," *Phys. Rev. Lett.*, vol. 93, pp. 196805-1–196805-4, 2004.

[81] C. Hu, "Reduce IC power consumption by $> 10\times$ with a green transistor?" in *IEEE Device Res. Conf. Tech. Dig.*, 2009, pp. 9–10.

[82] R. Li, Y. Lu, G. Zhou, Q. Liu, S. D. Chae, T. Vasen, W. S. Hwang, Q. Zhang, P. Fay, T. Kosel, M. Wistey, H. Xing, and A. Seabaugh, "AlGaSb/InAs tunnel field-effect transistor with on-current of 78 μA/μm at 0.5 V," *IEEE Electron Device Lett.*, vol. 33, no. 3, pp. 363–365, Mar. 2012.

[83] L. Lattanzio, L. De Michielis, and A. M. Ionescu, "Complementary germanium electron-hole bilayer tunnel FET for sub-0.3 V operation," *IEEE Electron Device Lett.*, vol. 33, no. 2, pp. 167–169, Feb. 2012.

[84] G. Fiori and G. Iannaccone, "Ultralow-voltage bilayer graphene tunnel FET," *IEEE Electron Device Lett.*, vol. 30, no. 10, pp. 1096–1098, Oct. 2009.

[85] Y. Wu, D. B. Farmer, W. Zhu, S. J. Han, C. D. Dimitrakopoulos, A. A. Bol, P. Avouris, and Y. M. Lin, "Three-terminal graphene negative differential resistance devices," *ACS Nano.*, vol. 6, pp. 2610–2616, 2012.

[86] J. Knoch and J. Appenzeller, "Tunneling phenomena in carbon nanotube field-effect transistors," *Phys. Stat. Solidi (A)*, vol. 205, pp. 679–694, 2008.

[87] Q. Zhang, T. Fang, H. Xing, A. Seabaugh, and D. Jena, "Graphene nanoribbon tunnel transistors," *IEEE Electron Device Lett.*, vol. 29, no. 12, pp. 1344–1346, Dec. 2008.

[88] M. S. Arnold, A. A. Green, J. F. Hulvat, S. I. Stupp, and M. C. Hersam, "Sorting carbon nanotubes by electronic structure using density differentiation," *Nature Nanotechnol.*, vol. 1, pp. 60–65, 2006.

[89] X. Wang, X. Li, L. Zhang, Y. Yoon, P. K. Weber, H. Wang, J. Guo, and H. Dai, "N-doping of graphene through electrothermal reactions with ammonia," *Science*, vol. 324, pp. 768–771, 2009.

[90] M. Luisier and G. Klimeck, "Performance analysis of statistical samples of graphene nanoribbon tunneling transistors with line edge roughness," *Appl. Phys. Lett.*, vol. 94, pp. 223505-1–223505-3, 2009.

[91] S. O. Koswatta, S. J. Koester, and W. Haensch, "On the possibility of obtaining MOSFET-like performance and sub-60-mV/dec swing in 1D broken-gap tunnel transistors," in *IEEE Int. Electron Device Meeting Tech. Dig.*, 2010, pp. 3222–3230.

[92] K. T. Lam, D. Seah, S. K. Chin, S. B. Kumar, G. Samudra, Y. C. Yeo, and G. Liang, "A simulation study of graphene-nanoribbon tunneling FET with heterojunction channel," *IEEE Electron Device Lett.*, vol. 31, no. 6, pp. 555–557, Jun. 2010.

[93] H. Da, K. T. Lam, G. Samudra, S. K. Chin, and G. Liang, "Graphene nanoribbon tunneling field-effect transistors with a semiconducting and a semimetallic heterojunction channel," *IEEE Trans. Electron Devices*, vol. 59, no. 5, pp. 1454–1461, May 2012.

[94] Y. Yoon and S. Salahuddin, "Barrier-free tunneling in a carbon heterojunctions transistor," *Appl. Phys. Lett.*, vol. 97, pp. 033102-1–033102-3, 2010.

[95] D. Sarkar, M. Krall, and K. Banerjee, "Electron-hole duality during band-to-band tunneling process in graphene-nanoribbon tunnel-field-effect-transistors," *Appl. Phys. Lett.*, vol. 97, pp. 236109-1–236109-3, 2010.

[96] S. Agarwal and E. Yablonovitch, "Pronounced effect of pn-junction dimensionality on tunnel switch sharpness." [Online]. Available: arXiv:1109.0096

[97] J. Bardeen, "Tunneling from a many-particle point of view," *Phys. Rev. Lett.*, vol. 6, pp. 57–59, 1961.

[98] R. M. Feenstra, Y. Dong, M. P. Semstiv, and W. T. Masselink, "Influence of tip-induced band bending on tunneling spectra of semiconductor surfaces," *Nanotechnology*, vol. 18, pp. 044105-1–044105-7, 2007.

[99] V. Ambegaokar and A. Baratoff, "Tunneling between superconductors," *Phys. Rev. Lett.*, vol. 10, pp. 486–489, 1963.

[100] R. M. Feenstra, D. Jena, and G. Gu, "Single-particle tunneling in doped graphene-insulator-graphene junctions," *J. Appl. Phys.*, vol. 111, pp. 043711-1–043711-10, 2012.

[101] J. A. Simmons, M. A. Blount, J. S. Moon, S. K. Lyo, W. E. Baca, J. R. Wendt, J. L. Reno, and M. J. Hafich, "Planar quantum transistor based on 2D-2D tunneling in double quantum well heterostructures," *J. Appl. Phys.*, vol. 84, pp. 5626–5634, 1998.

[102] H. Sahin, S. Cahangirov, M. Topsakal, E. Bekaroglu, E. Akturk, R. T. Senger, and S. Ciraci, "Monolayer honeycomb structures of group-IV elements and III-V binary compounds: First-principles calculations," *Phys. Rev. B.*, vol. 80, pp. 155453-1–155453-12, 2009.

ABOUT THE AUTHOR

Debdeep Jena (Member, IEEE) received the B.Tech. degree in electrical engineering and a minor in physics from the Indian Institute of Technology (IIT), Kanpur, India, in 1998 and the Ph.D. degree in electrical and computer engineering from the University of California Santa Barbara (UCSB), Santa Barbara, CA, USA, in 2003.

He joined the faculty of the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN, USA, in 2003. His research and teaching interests are in the MBE growth and device applications of quantum semiconductor heterostructures; investigation of charge transport in nanostructured semiconducting materials such as graphene, nanowires, and nanocrystals, and their device applications; and in the theory of charge, heat, and spin transport in nanomaterials. He has authored/coauthored more than 100 journal articles.

Prof. Jena received two best student paper awards in 2000 and 2002 for his Ph.D. dissertation research, the National Science Foundation (NSF) CAREER award in 2007, the Joyce award for excellence in undergraduate teaching in 2010, the International Symposium of Compound Semiconductors Young Scientist award in 2012, and the IBM Faculty award also in 2012.