Science

AAAS

Heterogeneous Three-Dimensional Electronics by Use of Printed Semiconductor Nanomaterials
Jong-Hyun Ahn, *et al.*
Science **314**, 1754 (2006);
DOI: 10.1126/science.1132394

9. M. Schwarzschild, R. Härm, *Astrophys. J.* **150**, 961 (1967).
10. R. H. Sanders, *Astrophys. J.* **150**, 971 (1967).
11. M. Busso, R. Gallino, G. J. Wasserburg, *Annu. Rev. Astron. Astrophys.* **37**, 239 (1999).
12. O. Straniero *et al.*, *Astrophys. J.* **440**, L85 (1995).
13. J. W. Truran, I. Iben Jr., *Astrophys. J.* **216**, 797 (1977).
14. H. Beer, R. L. Macklin, *Astrophys. J.* **339**, 962 (1989).
15. C. Abia *et al.*, *Astrophys. J.* **559**, 1117 (2001).
16. D. L. Lambert *et al.*, *Astrophys. J.* **450**, 302 (1995).
17. J. Tomkin, D. L. Lambert, *Astrophys. J.* **523**, 234 (1999).
18. A. Banerjee, D. Das, V. Natarajan, *Europhys. Lett.* **65**, 172 (2004).
19. N. Grevesse, A. J. Sauval, *Space Sci. Rev.* **85**, 161 (1998).
20. P. S. Chen, R. Szczerba, S. Kwok, K. Volk, *Astron. Astrophys.* **368**, 1006 (2001).
21. F. M. Jiménez-Esteban, P. García-Lario, D. Engels, in *Planetary Nebulae as Astronomical Tools*, R. Szczerba, G. Stasinska, S. K. Gorny, Eds. (AIP Conference Proceedings, vol. 804, American Institute of Physics, Melville, NY, 2005), pp. 141–144.
22. P. Ventura *et al.*, *Astrophys. J.* **550**, L65 (2001).
23. C. M. O'D. Alexander, *Philos. Trans. R. Soc. London Ser. A* **359**, 1973 (2001).
24. E. Zinner *et al.*, *Geochim. Cosmochim. Acta* **69**, 4149 (2005).
25. K. D. McKeegan, A. M. Davis, in *Meteorites, Planets, and Comets*, A. Davis, Ed., vol. 1 of *Treatise on Geochemistry* (Elsevier-Pergamon, Oxford, 2003), pp. 431–460.
26. F. A. Podosek *et al.*, *Geochim. Cosmochim. Acta* **55**, 1083 (1991).
27. Supported by the Spanish Ministerio de Educación y Ciencia (MEC) grants AYA 2004-3136 and AYA 2003-9499 (A.M. and P.G.L.) and a MEC JdC grant (J.M.T.R.). This work is based on observations obtained at the 4.2-m William Herschel Telescope, operated on the island of La Palma by the Isaac Newton Group in the Observatorio del Roque de Los Muchachos of the Instituto de Astrofisica de Canarias, and on observations obtained with the 3.6-m telescope at ESO–La Silla Observatory (Chile).

Heterogeneous Three-Dimensional Electronics by Use of Printed Semiconductor Nanomaterials

Jong-Hyun Ahn,[1,2,3] Hoon-Sik Kim,[5] Keon Jae Lee,[1,3] Seokwoo Jeon,[1,2,3] Seong Jun Kang,[1,2,3] Yugang Sun,[1,2,3] Ralph G. Nuzzo,[1,2,3,4] John A. Rogers[1,2,3,4,5]*

We developed a simple approach to combine broad classes of dissimilar materials into heterogeneously integrated electronic systems with two- or three-dimensional layouts. The process begins with the synthesis of different semiconductor nanomaterials, such as single-walled carbon nanotubes and single-crystal micro- and nanoscale wires and ribbons of gallium nitride, silicon, and gallium arsenide on separate substrates. Repeated application of an additive, transfer printing process that uses soft stamps with these substrates as donors, followed by device and interconnect formation, yields high-performance heterogeneously integrated electronics that incorporate any combination of semiconductor nanomaterials on rigid or flexible device substrates. This versatile methodology can produce a wide range of unusual electronic systems that would be impossible to achieve with other techniques.

Many existing and emerging electronic devices benefit from the heterogeneous integration of dissimilar classes of semiconductors into single systems, in either two-dimensional (2D) or 3D layouts (*1*, *2*). Examples include multifunctional radio-frequency communication devices, infrared imaging cameras, addressable sensor arrays, and hybrid silicon complementary metal oxide semiconductor (CMOS) circuits and nanowire devices (*3*–*7*). In some representative systems, compound semiconductors or other materials provide high-speed operation, efficient photodetection, or sensing capabilities; the silicon CMOS provides digital readout and signal processing in circuits that often involve stacked 3D configurations. Wafer- or chip-scale bonding (*1*, *2*, *6*, *8*–*10*) and epitaxial growth (*3*, *11*, *12*) represent the two most widely used methods for achieving these types of integrated systems.

The bonding processes use fusion processes (*8*, *9*) or layers of adhesives (*6*, *10*) to combine integrated circuits, photodiodes, or sensors formed separately on different semiconductor wafers. This approach works well in many cases, but it has important drawbacks (*1*, *9*), including (i) limited ability to scale to large areas (i.e., larger than the wafers) or to more than a few layers in the third (i.e., stacking) dimension; (ii) incompatibility with unusual materials (such as nanostructured materials) and/or low-temperature materials and substrates; (iii) challenging fabrication and alignment for the through-wafer electrical interconnects; (iv) demanding requirements for planar bonding surfaces; and (v) bowing and cracking that can occur from mechanical strains generated by differential thermal expansion and contraction of disparate materials. Epitaxial growth provides a different approach, which uses molecular beam epitaxy or other means to form thin layers of semiconductor materials directly on the surfaces of wafers of other materials. Although this method avoids some of the aforementioned problems, the requirements for epitaxy place severe restrictions on the quality and type of materials that can be grown, even when buffer layers and other advanced techniques are used (*1*, *13*).

By contrast, nanoscale wires, ribbons, membranes, or particles of inorganic materials, or carbon-based systems such as single-walled carbon nanotubes (SWNTs) or graphene sheets (*14*–*17*), can be grown and then suspended in solvents or transferred onto substrates in a manner that bypasses the need for epitaxial growth or wafer bonding. Recent work has shown the integration of isolated crossed nanowire diodes in 2D layouts formed by solution casting (*18*). Our results show how dissimilar single-crystal inorganic semiconductors (such as micro- and nanoscale wires and ribbons of GaN, Si, and GaAs) can be combined with one another and also with other classes of nanomaterials (such as SWNTs) with the use of a scalable and deterministic printing method to yield complex, heterogeneously integrated electronic systems in 2D or 3D layouts. The capabilities of this process are demonstrated by ultrathin multilayer stacks of high-performance metal oxide semiconductor field-effect transistors (MOSFETs), high electron mobility transistors (HEMTs), thin-film transistors (TFTs), photodiodes, and other components that are integrated into device arrays, logic gates, and actively addressable photodetectors on rigid inorganic and flexible plastic substrates.

Figure 1 illustrates representative steps for producing these types of systems. The process begins with the synthesis of the semiconductor nanomaterials, each on their own source substrate. The devices shown in Fig. 1 allow micro- and nanoscale wires and ribbons of single-crystalline GaN, GaAs, and Si that were formed with the use of wafer-based source materials and lithographic etching procedures (*19*–*23*) to be integrated with each other or with networks of SWNTs that were grown by chemical vapor deposition (*16*, *23*). Scanning electron micrographs at the top of Fig. 1 show these semiconductor nanomaterials after their removal from the source substrates. For circuit fabrica-

[1]Department of Materials Science and Engineering, University of Illinois, Urbana-Champaign, IL 61801, USA. [2]Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign, IL 61801, USA. [3]Frederick Seitz Materials Research Laboratory, University of Illinois, Urbana-Champaign, IL 61801, USA. [4]Department of Chemistry, University of Illinois, Urbana-Champaign, IL 61801, USA. [5]Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, IL 61801, USA.

*To whom correspondence should be addressed. E-mail: jrogers@uiuc.edu

tion, these elements remain in the configurations defined on the wafers during the fabrication or growth stage: aligned arrays in the case of the GaN, GaAs, and Si materials, and submonolayer random networks for the SWNTs. High-temperature doping and annealing procedures for ohmic contacts to the GaN, GaAs, and Si can be performed on the source substrates.

The next step involves transferring these processed elements, with the use of an elastomeric stamp–based printing technique (*19*), from the source substrates to a device substrate, such as a sheet of polyimide (PI) (Fig. 1). In particular, laminating a stamp of polydimethylsiloxane against the source substrate establishes soft, van der Waals adhesion contacts to the semiconductor nanomaterial elements. We contacted the "inked" stamp to a device substrate (such as a PI sheet) with a thin, spin-cast layer of a liquid prepolymer (such as polyamic acid) on its surface and then cured the polymer, which left these semiconductor materials embedded on and well adhered to this layer (*19–22*) when the stamp was removed. Similar procedures work well with a range of substrates (i.e., rigid or flexible and organic or inorganic) and semiconductor nanomaterials. We used a slightly modified version of this process for the SWNT devices (*23*). The thickness of the interlayer (PI in this case) can be as small as 500 nm and was typically 1 to 1.5 μm for the systems we describe.

After some additional processing—including deposition and patterning of gate dielectrics, electrodes, and interconnects—the transfer printing and device fabrication steps can be repeated, beginning with spin-coating a new prepolymer interlayer on top of the previously completed circuit level. Automated stages specially designed for transfer printing or conventional mask aligners enable overlay registration accuracy of ~2 μm over several square centimeters (fig. S1). The spatial distortions associated with the printing had a mean value of ~0.5 μm (fig. S2). The yields for printing of Si, SWNT, GaAs, and GaN were >99%, >99%, >95%, and >85%, respectively. Defects in these last two cases were associated with fracture and impartial transfer for the relatively wide GaAs ribbons and relatively thick GaN bars, respectively (fig. S3). Layer-to-layer interconnects (*24*) were formed simply by evaporating metal lines over and into openings in the interlayers defined by photopatterning and/or dry etching.

This fabrication approach has several important features. First, all of the processing on the device substrate occurs at low temperatures, thereby avoiding differential thermal expansion and shrinkage effects that can result in unwanted deformations in multilayer stacked systems. This operation also enables the use of low-temperature plastic substrates and interlayer materials, and it helps to ensure that underlying circuit layers are not thermally degraded by the processing of overlying devices.

Second, the method is applicable to broad classes of semiconductor nanomaterials, including emerging materials such as SWNTs. Third, the soft stamps enable nondestructive contacts with underlying device layers; these stamps, together with the ultrathin semiconductor materials, can also tolerate surfaces that have some topography. Fourth, the ultrathin device geometries and interlayers allow easy formation of layer-to-layer electrical interconnects by direct metallization over the device structure. These features overcome many of the disadvantages of conventional approaches.

Figure 2 presents three-layer, 3D stacks of arrays of Si MOSFETs fabricated by the general process flow shown in Fig. 1. We used single-crystalline silicon nanoribbons with doped contacts (formed on the source wafer), SiO_2 dielectrics formed by plasma-enhanced chemical vapor deposition, and Cr/Au metallization for the source, drain, and gate electrodes (*25*). Each device uses three aligned nanoribbons, each with length $L = 250$ μm, width $W = 87$ μm, and thickness = 290 nm. Figure 2A shows an

optical micrograph of an edge of the system; the layout of the system was designed to reveal separately the parts of the substrate that support one, two, and three layers of MOSFETs. A 90° rotation of the device geometry for the second layer, relative to the first and third, helps to clarify the layout of the system. Schematic cross-sectional and angled views of the stacked structure are shown in Fig. 2B. The sample can be viewed in 3D using confocal optical microscopy. Figure 2C shows top and angled views of such images. (The image quality degrades somewhat with depth because of scattering and absorption from the upper layers). Figure 2D presents electrical measurements of representative devices in each layer. Devices on
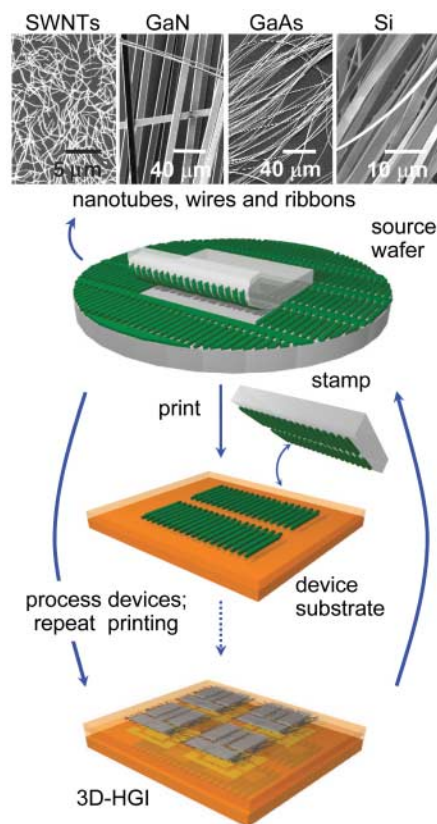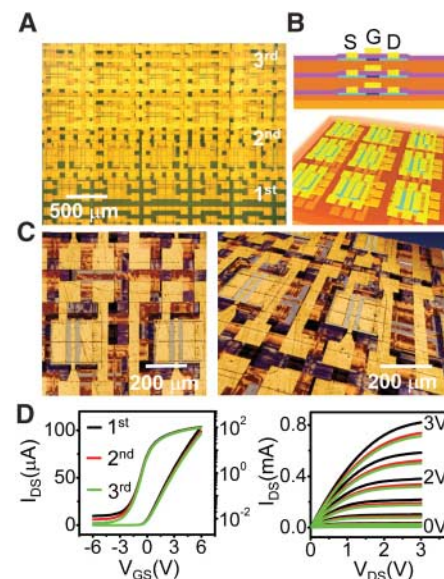
Fig. 1. Schematic illustration of a printed semiconductor nanomaterials–based approach to heterogeneous 3D electronics. The process involves the repetitive transfer printing of collections of nanotubes, nanowires, nanoribbons, or other active nanomaterials, separately formed on source substrates, to a common device substrate to generate interconnected electronics in ultrathin, multilayer stack geometries. HGI, heterogeneous integration.

Fig. 2. (**A**) Optical micrograph view of the top of a 3D multilayer stack of arrays of single-crystal silicon MOSFETs that use printed silicon nanoribbons for the semiconductor. The bottom (1st), middle (2nd), and top (3rd) parts of this image correspond to regions with one, two, and three layers of devices, respectively. (**B**) Schematic cross-sectional (top) and angled (bottom) views. S, D, and G refer to source, drain, and gate electrodes (all shown in gold), respectively. The light and dark blue regions correspond to doped and undoped regions of the silicon ribbons; the purple layer is the SiO_2 gate dielectric. (**C**) 3D images (left, top view; right, angled view) collected by confocal microscopy on a device substrate similar to that shown in (A) and (B). The layers are colorized (gold, top layer; red, middle layer; blue, bottom layer; silicon, gray) for ease of viewing. (**D**) Current-voltage characteristics of Si MOSFETs in each of the layers, showing excellent performance (mobilities of 470 ± 30 cm^2/Vs) and good uniformity in the properties. The channel lengths and widths are 19 and 200 μm, respectively. The overlap lengths, as defined by distance that the gate electrode extends over the doped source and drain regions, are 5.5 μm. I_{DS}, drain current; V_{GS}, bias voltage; V_{DS}, drain voltage.

each of the three layers, which are formed on a PI substrate, show excellent properties with linear mobilities of 470 ± 30 cm^2/Vs (where the error is SD), on/off ratios greater than 10^4, and threshold voltages of -0.1 ± 0.2 V; there are no systematic differences between devices in different layers. Additional layers can be added to this system by repeating the same procedures. To investigate issues related to mismatches in coefficients of thermal expansion in these systems, we evaluated the behavior of the devices under thermal cycling (60 times) between room temperature and 90°C. Small changes were observed for the first few cycles followed by stable behavior (fig. S9).

In addition to 3D circuits with a single semiconductor, Fig. 3 illustrates that the capability to combine various semiconductors can be used in multiple layers. We fabricated arrays of HEMTs, MOSFETs, and TFTs—with the use of GaN bars, Si nanoribbons, and SWNT films, respectively, on PI substrates. Figure 3, A and B, shows high-magnification optical and confocal images, respectively, of the resulting devices. The GaN HEMTs on the first layer use ohmic contacts (Ti/Al/Mo/Au, annealed on the source wafer) for the source and drain, and Schottky (Ni/Au) contacts for the gates. Each device uses GaN ribbons (composed of multilayer stacks of AlGaN/GaN/AlN) interconnected electrically by processing on the device substrate. The SWNT TFTs on the second layer use SiO$_2$ and epoxy for the gate dielectric and Cr/Au for the source, drain, and gate. The Si MOSFETs use the same design as those shown in Fig. 2. Various other devices can be constructed with different combinations of Si, SWNT, and GaN (figs. S4 and S5). Figure 3C presents the current-voltage characteristics of typical devices in the systems of Fig. 3, A and B. In all cases, the properties are similar to those fabricated on the source wafers: The GaN HEMTs have threshold voltages (V_{th}) of -2.4 ± 0.2 V, on/off ratios greater than 10^6, and transconductances of 0.6 ± 0.5 mS; the SWNT TFTs have $V_{th} = -5.3 \pm 1.5$ V, on/off ratios greater than 10^5, and linear mobilities of 5.9 ± 2.0 cm^2/Vs; and the Si MOSFETs have $V_{th} = 0.2 \pm 0.3$ V, on/off ratios greater than 10^4, and linear mobilities of 500 ± 30 cm^2/Vs.

Another interesting aspect of these devices, which follows from the use of thin PI substrates (25 μm), devices (<1.7 μm), and PI interlayers (1.5 μm), is their mechanical bendability. This characteristic is important for applications in flexible electronics, for which these systems might provide attractive alternatives because of their enhanced capabilities compared with those of conventional organic-based devices. We evaluated the effective transconductance (g_{eff}) for the Si, SWNT, and GaN devices in the system of Fig. 3A as a function of bend radius. Figure 3D, which shows these data normalized to the transconductance in the unbent state (g_{0eff}), demonstrates the stable performance for

bend radii down to 3.7 mm. To explore the response of devices to operation under various conditions, such as repeated bending and electrical testing, we carried out two sets of experiments. Repeated bending (up to 2000 cycles) resulted in no substantial change in the properties of the devices (fig. S8). Repeated electrical testing showed stable responses (~10% changes in properties, or less) (fig. S11). Figures
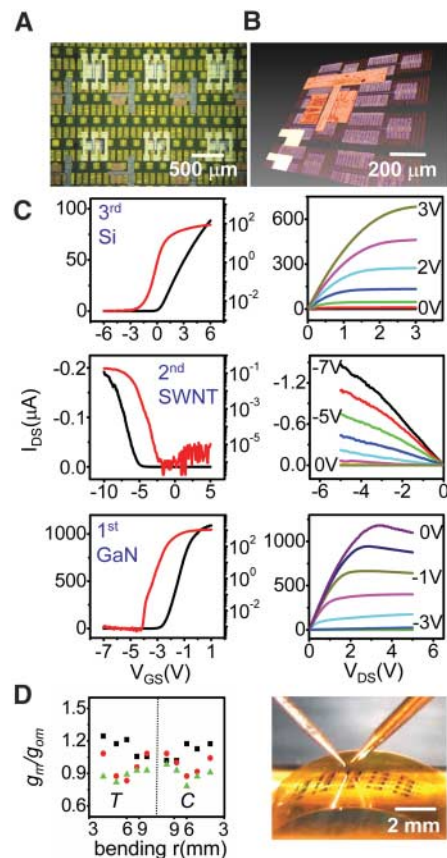
S12 to S15 present information on variation in device properties.

Electrical interconnections formed between different levels in these devices can create
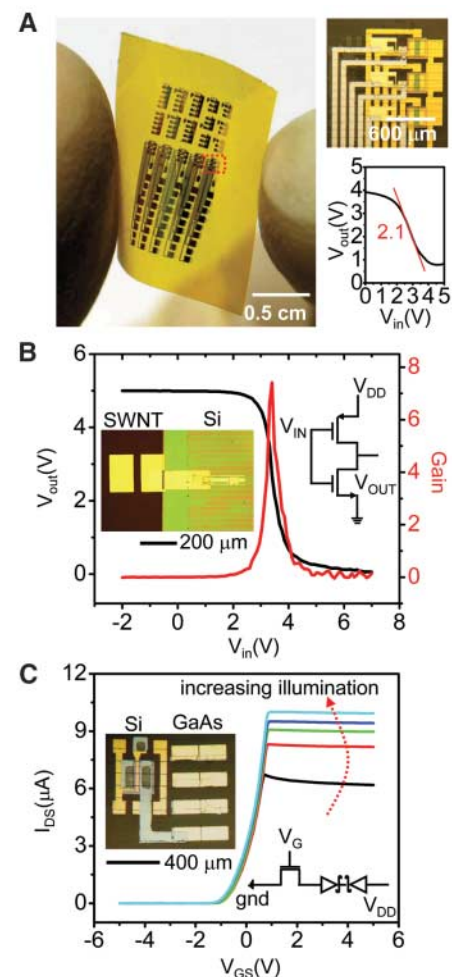

Fig. 3. (**A**) Optical micrograph of 3D heterogeneously integrated electronic devices, including GaN nanoribbon HEMTs, Si nanoribbon MOSFETs, and SWNT network TFTs, in a three-layer stack. (**B**) 3D image collected by confocal microscopy. The layers are colorized (gold: top layer, Si MOSFETs; red: middle layer, SWNT TFTs; pink: bottom layer) for ease of viewing. (**C**) Electrical characteristics of GaN devices on the first layer (channel lengths, widths and gate widths of 20, 170, and 5 μm, respectively, and ribbon thicknesses, widths, and lengths of 1.2, 10, and 150 μm, respectively), SWNT devices on the second layer (channel lengths and widths of 50 and 200 μm, respectively, and average tube diameters and lengths of ~1.5 nm and ~10 μm, respectively), and Si devices on the third layer (channel lengths and widths of 19 and 200 μm, respectively). (**D**) (Left) Normalized transconductances (g_m/g_{0m}) of devices in each layer (black squares, Si MOSFETs; red circles, SWNT TFTs; green triangles, GaN HEMTs) as a function of bending radius of the plastic substrate. *T*, tension; *C*, compression. (Right) Image of the bent system and probing apparatus.

Fig. 4. (**A**) Image of a printed array of 3D silicon n-channel metal oxide semiconductor inverters on a PI substrate. The inverters consist of MOSFETs (channel lengths of 4 μm, load-to-driver width ratio of 6.7, and a driver width of 200 μm) on two different levels, interconnected by electrical via structures. The image on the top right provides a magnified view of the region indicated by the red box in the left frame. The graph on the bottom right shows transfer characteristics of a typical inverter. (**B**) Transfer characteristics of a printed complementary inverter that uses a p-channel SWNT TFT (channel length and width of 30 and 200 μm, respectively) and an n-channel Si MOSFET (channel length and width of 75 and 50 μm, respectively). The insets provide an optical micrograph of an inverter (left) and a circuit schematic (right). (**C**) Current-voltage response of a GaAs MSM (channel length and width of 10 and 100 μm, respectively) integrated with a Si MOSFET (channel length and width of 9 and 200 μm, respectively) at different levels of illumination from dark to 11 μW with an infrared light source at 850 nm. The insets show an optical image (left) and a circuit diagram (right). gnd, ground.

interesting circuit capabilities. The thin polymer interlayers allow robust interconnects to be formed easily by evaporating metal lines over lithographically defined openings. Thermal cycling tests showed no changes in their properties (fig. S10). Figure 4A shows a 3D n-channel metal oxide semiconductor inverter (logic gate) in which the drive ($L = 4$ μm, $W = 200$ μm) and load ($L = 4$ μm, $W = 30$ μm) Si MOSFETs are on different levels. With a supply voltage of 5 V, this double-layer inverter exhibits well-defined transfer characteristics with gains of ~2, comparable to the performance of conventional planar inverters that use similar transistors. Figure 4B shows an inverter with a complementary design (CMOS) with the use of integrated n-channel Si MOSFETs and p-channel SWNT TFTs, designed to equalize the current-driving capability in both pull-up and pull-down directions. Transfer curves collected with a supply voltage (V_{DD}) of 5 V and gate voltage (input) swept from 0 to 5 V appear in Fig. 4B. The curve shapes and gains (as high as ~7) are qualitatively consistent with numerical circuit simulations (fig. S6). As a third example, we built GaAs metal-semiconductor-metal (MSM) infrared detectors (26), integrated with Si MOSFETs on flexible PI substrates, to demonstrate a capability for fabricating unit cells that could be used in active infrared imagers. In this case, printed nanoribbons of GaAs ($L = 400$ μm, $W = 100$ μm, and thickness = 270 nm) transferred onto a substrate with a printed array of Si nanoribbon MOSFETs form the basis of the MSMs. Electrodes (Ti/Au) deposited on the ends of these GaAs nanoribbons form back-to-back Schottky diodes with separations of 10 μm. The resulting detector cells exhibit current enhancement as the intensity of infrared illumination increases (Fig. 4C), con-

sistent with circuit simulation (fig. S7). A responsivity of about 0.30 A/W at the 850-nm wavelength is observed from 1 to 5 V. (This value underestimates the true responsivity because it ignores optical reflection). The bendability of this system, which is comparable to that of the devices in Fig. 3, could be useful for advanced systems such as curved focal plane arrays for wide-angle infrared night vision imagers.

Printed semiconductor nanomaterials provide new approaches to 3D heterogeneously integrated systems that could be important in various fields of application, including not only those suggested by the systems reported here but also others such as microfluidic devices with integrated electronics, chemical and biological sensor systems that incorporate unusual materials with conventional silicon-based electronics, and photonic and optoelectronic systems that combine light emitters and detectors of compound semiconductor with silicon drive electronics or microelectromechanical structures. Furthermore, the compatibility of this approach with thin, lightweight plastic substrates may create additional opportunities for devices that have unusual form factors or mechanical flexibility as key features.

References and Notes

1. K. Banerjee, S. J. Souri, P. Kapur, K. C. Saraswat, *Proc. IEEE* **89**, 602 (2001).
2. S. F. Al-Sarawi, D. Abbott, P. D. Franzon, *IEEE Trans. Components Packaging Manufacturing Technol. Part B* **21**, 2 (1998).
3. A. S. Brown *et al.*, *Mater. Sci. Eng. B* **87**, 317 (2001).
4. Y.-C. Tseng *et al.*, *Nano Lett.* **4**, 123 (2004).
5. C. Joachim, J. K. Gimzewski, A. Aviram, *Nature* **408**, 541 (2000).
6. G. Roelkens *et al.*, *Opt. Express* **13**, 10102 (2005).
7. D. B. Strukov, K. K. Likharev, *Nanotechnology* **16**, 888 (2005).
8. Q. Y. Tong, U. Gosele, *Semiconductor Wafer Bonding: Science and Technology* (John Wiley, New York, 1999).
9. M. A. Schmidt, *Proc. IEEE* **86**, 1575 (1998).
10. P. Garrou, *Semicond. Int.* **28**, SP10 (February, 2005).
11. H. Amano, N. Sawaki, I. Akasaki, Y. Toyoda, *Appl. Phys. Lett.* **48**, 353 (1986).
12. T. Kuykendall *et al.*, *Nat. Mater.* **3**, 524 (2004).
13. J. C. Bean, *Proc. IEEE* **80**, 571 (1992).
14. A. M. Morales, C. M. Lieber, *Science* **279**, 208 (1998).
15. M. Law *et al.*, *Science* **305**, 1269 (2004).
16. J. Kong, H. T. Soh, A. M. Cassell, C. F. Quate, H. Dai, *Nature* **395**, 878 (1998).
17. K. S. Novoselov *et al.*, *Science* **306**, 666 (2004).
18. Y. Huang, X. Duan, C. M. Lieber, *Small* **1**, 142 (2005).
19. M. A. Meitl *et al.*, *Nat. Mater.* **5**, 33 (2006).
20. E. Menard, K. J. Lee, D. Y. Khang, R. G. Nuzzo, J. A. Rogers, *Appl. Phys. Lett.* **84**, 5398 (2004).
21. Y. Sun, S. Kim, I. Adesida, J. A. Rogers, *Appl. Phys. Lett.* **87**, 083501 (2005).
22. S.-H. Hur, D.-Y. Khang, C. Kocabas, J. A. Rogers, *Appl. Phys. Lett.* **85**, 5730 (2004).
23. Materials and methods are available as supporting material on *Science* Online.
24. S. Linder, H. Baltes, F. Gnaedinger, E. Doering, in *Proceedings of the 7th IEEE International Workshop on Micro Electro Mechanical Systems*, Oiso, Japan, 25 to 28 January 1994 (IEEE, Piscataway, NJ, 1994), pp. 349–354.
25. J.-H. Ahn *et al.*, *IEEE Electron Devices Lett* **27**, 460 (2006).
26. J. B. D. Soole, H. Schumacher, *IEEE J. Quantum Electron.* **27**, 737 (1991).
27. The research was supported by the U.S. Department of Energy, Division of Materials Sciences under award no. DEFG02-91ER45439, through the Frederick Seitz Materials Research Laboratory (FS-MRL). We thank T. Banks and K. Colravy for help with cleanroom and other facilities at the Frederick Seitz Materials Research Laboratory and H. C. Ko, Q. Cao, P. Ferreira, J. Dong, and E. Menard for help with printing and distortion measurements using facilities and manufacturing approaches developed at the Center for Nanoscale Chemical Electrical Mechanical Manufacturing Systems at the University of Illinois (funded by the NSF under grant DMI-0328162). All imaging and surface analysis was performed at the FS-MRL Center for Microanalysis of Materials at the University of Illinois at Urbana-Champaign, supported by award no. DEFG02-91ER45439.

Quantum Spin Hall Effect and Topological Phase Transition in HgTe Quantum Wells

B. Andrei Bernevig,[1,2] Taylor L. Hughes,[1] Shou-Cheng Zhang[1]*

We show that the quantum spin Hall (QSH) effect, a state of matter with topological properties distinct from those of conventional insulators, can be realized in mercury telluride–cadmium telluride semiconductor quantum wells. When the thickness of the quantum well is varied, the electronic state changes from a normal to an "inverted" type at a critical thickness d_c. We show that this transition is a topological quantum phase transition between a conventional insulating phase and a phase exhibiting the QSH effect with a single pair of helical edge states. We also discuss methods for experimental detection of the QSH effect.

The spin Hall effect (1–5) has recently attracted great attention in condensed matter physics, not only for its fundamental scientific importance but also because of its potential application in semiconductor spintronics. In particular, the intrinsic spin Hall effect promises the possibility of designing the intrinsic electronic properties of materials so that the effect can be maximized. On the basis of this line of reasoning, it was shown (6) that the intrinsic spin Hall effect can in principle exist in band insulators, where the spin current can flow without dissipation. Motivated by this suggestion, researchers have proposed the quantum spin Hall (QSH) effect for graphene (7) as well as for semiconductors (8, 9), where the spin current is carried entirely by the helical edge states in two-dimensional samples.

Time-reversal symmetry plays an important role in the dynamics of the helical edge states (10–12). When there is an even number of pairs of helical states at each edge, impurity scattering or many-body interactions can open a gap at the edge and render the system topologically trivial. However, when there is an odd number of pairs of helical states at each edge, these effects cannot open a gap unless time-reversal symmetry is

[1]Department of Physics, Stanford University, Stanford, CA 94305, USA. [2]Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA.

*To whom correspondence should be addressed. E-mail: sczhang@stanford.edu