Meeting the challenge of integrating III-Vs with deep submicron silicon

High-mobility devices based on indium gallium arsenide (InGaAs) channels could benefit the performance of mainstream silicon integrated circuits. Researchers are working to meet the challenges of bringing the different device traditions together. Mike Cooke reports on contributions at December's IEEE International Electron Devices Meeting (IEDM) in Washington DC.

In principle, the high mobilities of InGaAs and related III-V compound semiconductors should lead to faster switching speeds with larger drive currents. These benefits could be used to replace silicon channels altogether or the critical parts of a complementary metal-oxide-semiconductor (CMOS) integrated circuit (IC).

Such developments are not simple drop-in replacements, however. While the cut-off frequencies can be much higher than silicon, other factors such as high on/off ratios and device dimension scaling (shrinks) are more difficult to achieve with III-Vs. Many III-V production processes are incompatible with silicon processing.

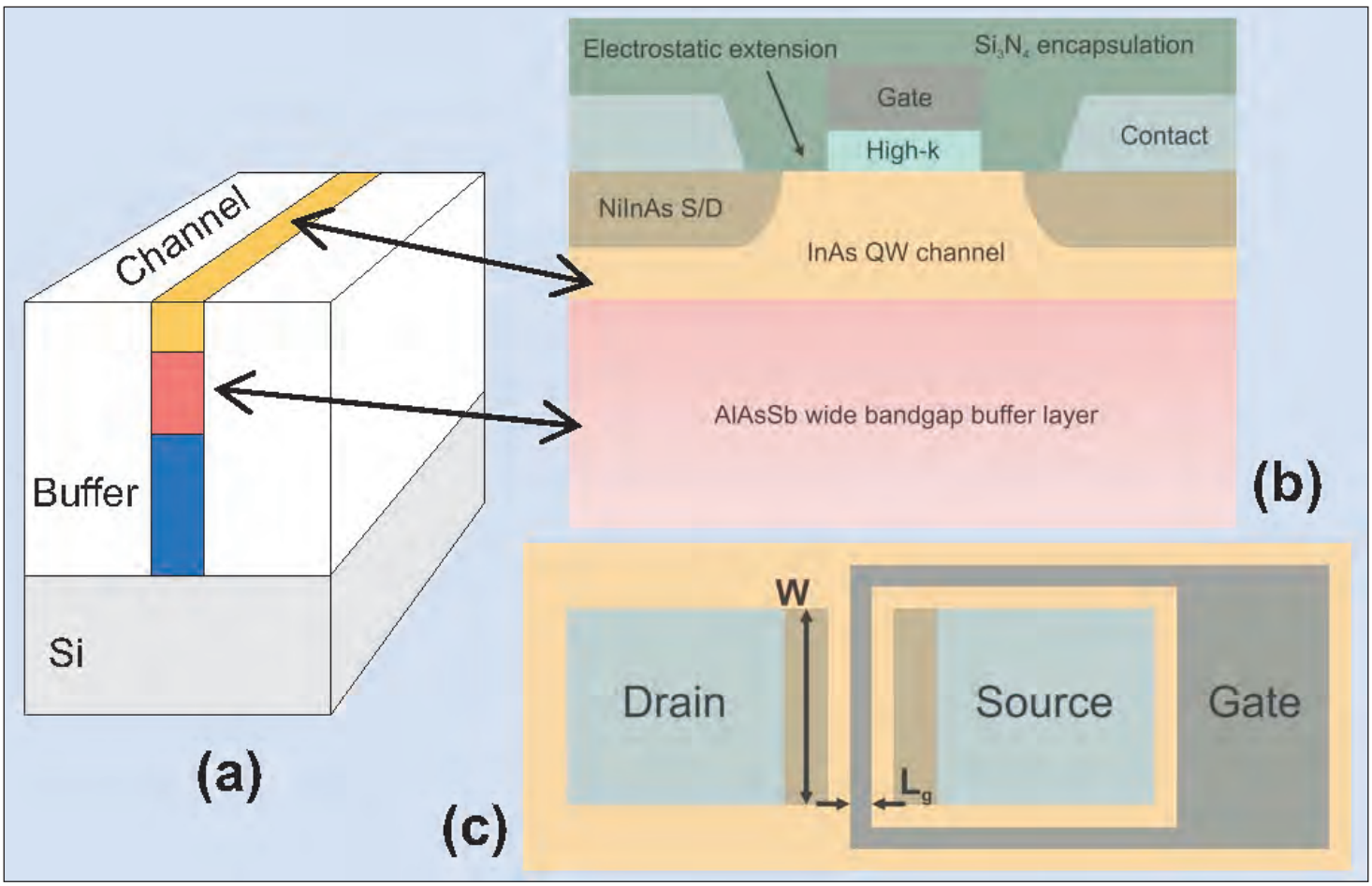

Figure 1. (a) The STI ART technique to introduce high-mobility channels into a silicon CMOS platform. (b) Representative 6.1Å III-V quantum well (QW) n-MOSFET architecture for implementation in CMOS. (c) Schematic of device layout used in S.W. Chang et al, showing active device width W and gate length L_g.

Here we look at some attempts to tackle these issues and others as reported at IEDM.

Beating HEMTs

Record-setting III-V n-type MOS field-effect transistors (FETs) were claimed by TSMC (Belgium and Taiwan), University of Glasgow (UK), Texas State University (USA), and Lund University (Sweden) [S.W. Chang et al, 16.1]. The team comments: "For the first time performance better than state-of-the-art HEMTs is demonstrated."

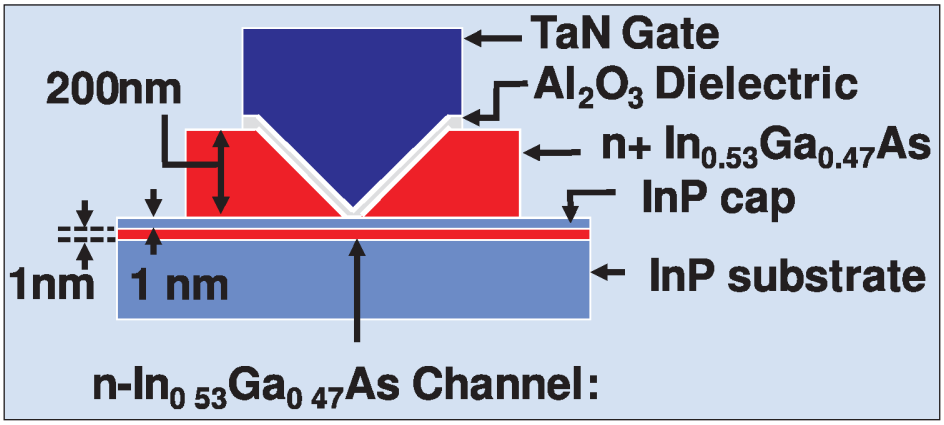

Figure 2. Structure of National University of Singapore device with ultra-thin (1nm) and ultra-short (~6nm) channel.

The process used is compatible with shallow-trench isolation (STI) aspect-ratio trapping (ART) growth techniques (Figure 1) and with antimony (Sb)-based p-FETs for a full III-V CMOS structure.

The MOSFET consisted of a 10nm surface channel of unstrained InAs with a 130nm gate length. The channel layer was molecular beam epitaxy (MBE) grown on a lattice-matched aluminium arsenide antimonide (AlAsSb) barrier on InAs substrate. Metal-organic chemical vapor deposition (MOCVD) growth has also recently been demonstrated.

The device structure had a gate that surrounds the gate contact. "This alleviates the need for isolation by, for example, mesa etch, which is not straightforward to implement due to the conductive nature of exposed InAs surfaces," the researchers explain. The gate stack consisted of ALD high-k dielectric and palladium metal electrode.

The source/drain regions consisted of titanium-nickel InAs, followed by titanium/gold contacts. Low-sheet-resistance electrostatic extensions and encapsulation were achieved using silicon nitride. The extensions induce carriers by pinning the Fermi surface above the conduction band minimum.

At 0.5V drain bias, the on current was 601μA/μm and the off current was 100nA/μm. The extrinsic transconductance was 2.72mS/μm (5.5mS/μm, intrinsic), exceeding the best reported data for state-of-the-art planar III-V n-FETs. The researchers point out that this performance is higher than aggressively scaled devices, despite the relatively long gate length.

Subthreshold swing (SS) and drain-induced barrier lowering (DIBL) were 85mV/dec and 40mV/V, respectively. The mobility was 7100cm^2/V-s and the carrier density was 6.7x10^{12}/cm^2.

Shrink to 6nm

National University of Singapore (NUS) reported the realization of $In_{0.53}Ga_{0.47}As$ junctionless FETs (JLFETs) with the shortest channel length for any III-V transistor of 6nm [K.H. Goh et al, 16.5]. The 1nm-thick channel was sandwiched between a 1nm InP cap/etch stop and the InP substrate. Etched raised source/drain structures consisted of n^+- $In_{0.53}Ga_{0.47}As$, giving a contact resistance of 165Ω-μm. The gate stack consisted of tantalum nitride electrode on aluminium oxide (Al_2O_3) dielectric with an equivalent oxide thickness (EOT) of 2.5nm (Figure 2).

The peak transconductance of 1480μS/μm was achieved at 0.7V drain bias. The device achieved a ballistic transport efficiency of 0.82. The on–off current ratio was ~10^4 at 0.1V drain. A drive current of 480μA/μm was achieved at 0.7V drain and 0.7V gate potential over threshold. Short-channel effects were quite severe, however, although these were considered "well controlled down to 16nm" channel length.

NUS has also developed a passivation process for germanium FETs using high-quality $In_{0.48}Al_{0.52}P$. The large band offset between the materials enabled record high mobility values in electron transport in n-FET inversion layers of ~958cm^2/V-s. The carrier density was 6x10^{11}/cm^2. Hole transport mobility in p-FETs was also high, at ~390cm^2/V-s.

The researchers comment: "The InAlP-capped Ge CMOS technology could enable a common gate stack, common channel material option for sub-10nm technology nodes."

Results and prospects

The Massachusetts Institute of Technology (MIT) group led by Jesús del Alamo contributed three papers on InGaAs transistor technology [2.1, 16.2, 28.4] (along with one on GaN MISHEMTs [6.2], see Mike Cooke, Semiconducor Today 2014 February, page 83).

One presentation [J. A. del Alamo et al, 2.1] gave a progress report comparing MOSFET technology with the more established HEMT technology (Figure 3). Global Foundries and SEMATACH were also involved in the work. Recent MOSFET work has managed to beat HEMTs in terms of low on-resistance and has come close to HEMT transconductance performance.

Although there is a gap in current-gain cut-off frequency (f_T) performance, the researchers comment: "It is only

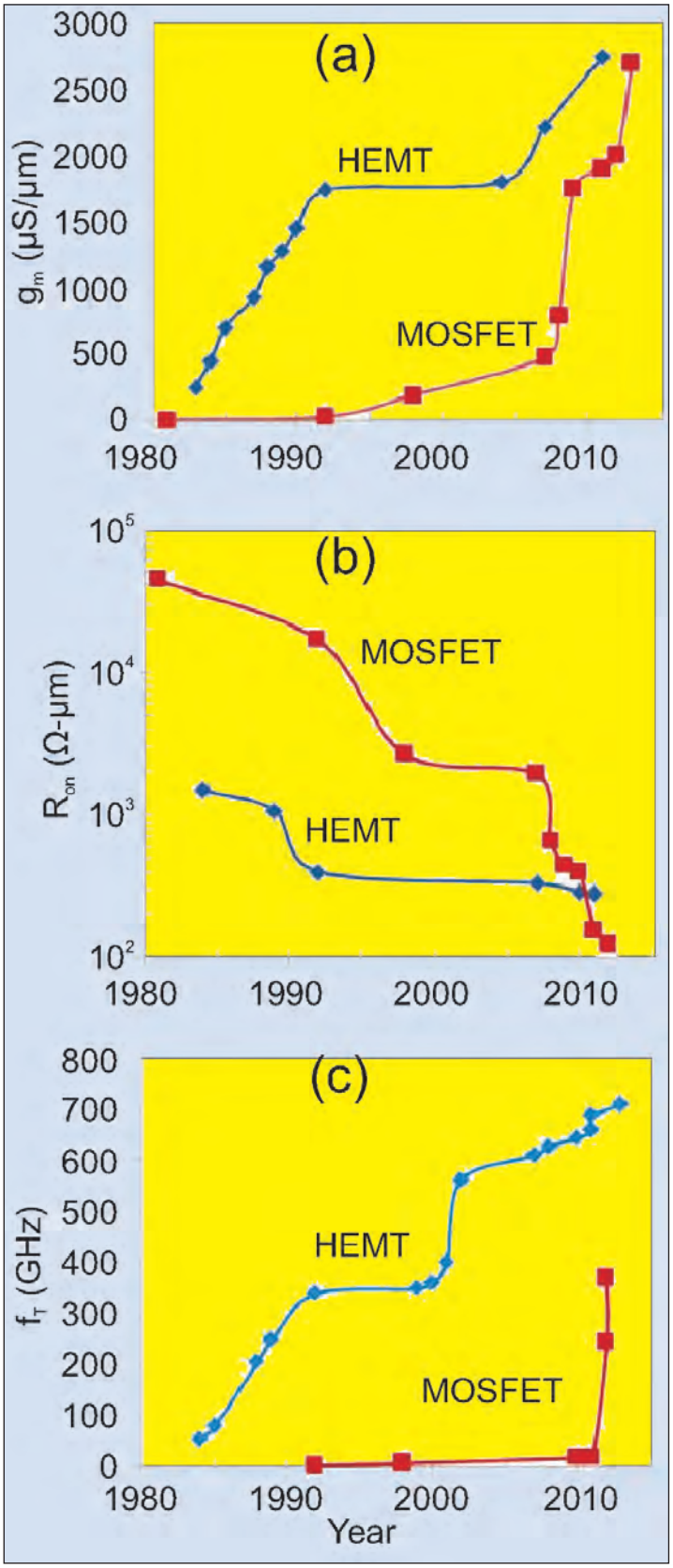

Figure 3. Performance comparison of inversion-type InGaAs MOSFETs and HEMTs (with InAs composition between 0 and 1) vs year: (a) transconductance, (b) ON resistance, (c) current-gain cut-off frequency, f_T.

a matter of time before low-parasitic-capacitance MOSFET designs are developed and improved high-frequency characteristics are demonstrated."

However, for integration into 10nm-scale mainstream manufacture these devices need to be shrunk to an ultra-small footprint. Further, the process needs to be 'self-aligned' to lower complexity and cost. Present research devices tend to have micron-scale contacts that need to be reduced by orders of magnitude. Finally, the transistors need to be inserted into a production environment where three-dimensional (3D) stacking of devices is an increasing trend.

Also, for consistency with mainstream silicon device processing, traditional III-V methodologies such as gold-based metallization, wet etch, and lift-off must be replaced with their silicon-compatible counterparts.

The team has developed molybdenum contacts that give a low contact resistivity of 0.67Ω-μm^2, but this value is not maintained below 110nm contact lengths. Alternative materials such as nickel, cobalt or palladium give inferior results.

The team has also worked on the gate dielectric, seeking EOTs smaller than 1nm for adequate electrostatic control of channel current by the gate. Sub-1nm EOTs have been achieved with a buried channel under an InP layer on which hafnium dioxide (HfO_2) is deposited. The direct deposition of HfO_2 on InP has been found to give lower interface trap densities, compared with Al_2O_3.

MOSFETs with 50nm gates and HfO_2/InP gate barriers have been produced with SS of 95mV/dec in 0.5V operation. "This nearly matches the characteristics of InGaAs trigate FETs of identical gate length, even though the trigate MOSFET has intrinsically better short-channel effects," the team writes.

For tackling the problems of dry/plasma etch of indium-containing materials, the team has developed an inductively coupled plasma reactive-ion process using boron trichloride, silicon tetrachloride and argon. The researchers report: "When combined with digital etching, we have realized 20nm fins and pillars on InGaAs/InAlAs/InP heterostructures without notching and minimum footing and trenching."

Some of these techniques were demonstrated in the two device presentations from MIT. Jianqiang Lin et al [16.2] reported on a "new self-aligned quantum-well MOSFET architecture fabricated by a scalable tight-pitch process". The device was produced using a three-step gate recess process (Figure 4).

A 70nm gate length InAs MOSFET with 5nm ledge length achieved a record transconductance of 2.7mS/μm. A different device with a 70nm ledge achieved a record 410μA/μm on-current. The latter MOSFET also demonstrated SS of 90mV/dec at 50mV and 94mV/dec at 0.5V. "This is the lowest S demonstrated to date at this gate length among III-V MOSFETs," the researchers claim.

The short-ledge device also has low source–drain resistance. To combine the positive features of the two types of device would require improvements in access resistance.

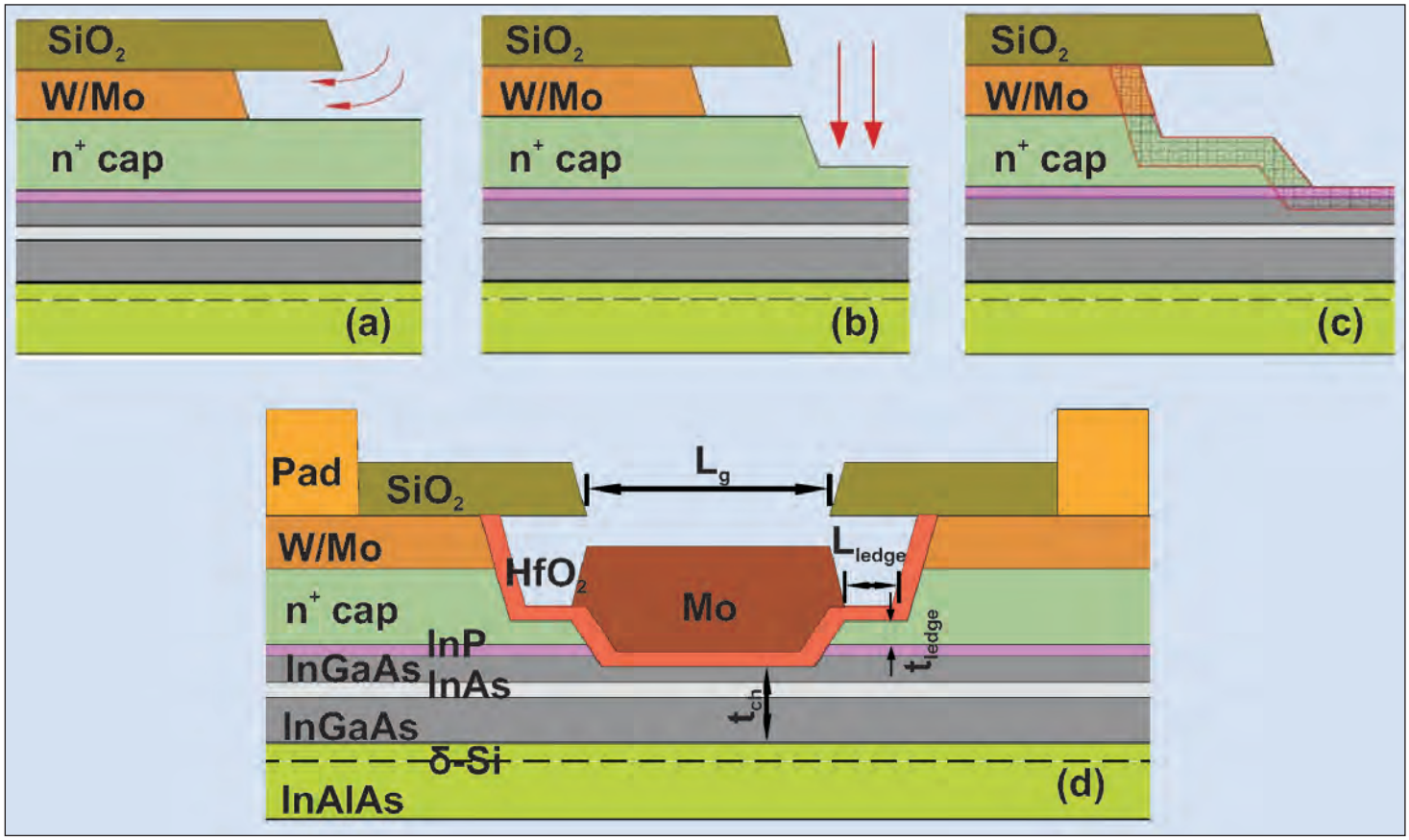

Figure 4. MIT's 3-step gate recess process: (a) W/Mo pull in, (b) time-controlled Cl_2 dry etch, (c) cap and barrier digital etch. (d) Cross section schematic of complete device structure.

The process was also used to create working 20nm gate-length devices with "very tight metal contact spacing". The researchers comment: "To our knowledge, this is the smallest III-V MOSFET demonstrated so far."

The second MIT device report was "Vertical nanowire InGaAs MOSFETs fabricated by a top-down approach" from Xin Zhao et al [28.4]. The nanowire (NW) device fabrication included digital etch — effectively atomic-layer deposition in reverse — that increased the transconductance by 20%.

The researchers involved in this project comment: "A vertical NW device structure uncouples gate length scaling and footprint scaling. As a result, device density goals can be reached with far better short-channel effects and performance than in planar MOSFETs, FinFETs or lateral NW-FETs."

The vertical nanowires were carved out of InGaAs on InP substrates with the ICP-RIE described above and digital etch to smooth out surface roughness. The use of 'top-down' etching, rather than a 'bottom-up' process, to create nanowires is rare because of the difficulty in obtaining smooth sidewalls. Bottom-up processes use a gold seed, which is not compatible with mainstream silicon processing.

The InGaAs layers consisted of an undoped intrinsic region sandwiched between heavily doped n-type material. A device with 80nm channel (the intrinsic InGaAs layer) length, 50nm diameter wires and EOT of 2.2nm for the gate dielectric (spin-on glass) achieved 720μS/μm transconductance at 0.5V.

Subtheshold swing and drain-induced barrier lowering characteristics are improved by using smaller-diameter NWs, at the cost of reduced transconductance. The researchers add: "Our devices demonstrate a performance in terms of the balance between short-channel effects and transport that matches that of the best vertical NW III-V MOSFETs fabricated by bottom-up techniques."

Self-alignment and co-integration

IBM researchers at the T. J. Watson Research Center have developed "self-aligned fully-depleted III-V MOSFETs using CMOS-compatible device structures and manufacturable process flows" [Y. Sun et al, 2.7]. Devices with gate lengths as short as 30nm were produced. The peak saturation transconductance was 1140μS/μm for a 60nm gate-length transistor at 0.5V operation.

The process involved "gate definition and spacer formation using RIE, and formation of self-aligned source/drain extensions (SDE) and self-aligned raised source/drain (R_{SD})".

The IBM team chose to work with $In_{0.53}Ga_{0.47}As$ channels since it expected this to lead to lower leakage at 10nm gate length due to "reduced direct source-to-drain tunneling and band-to-band tunneling, even when quantization is taken into account". Also, the material

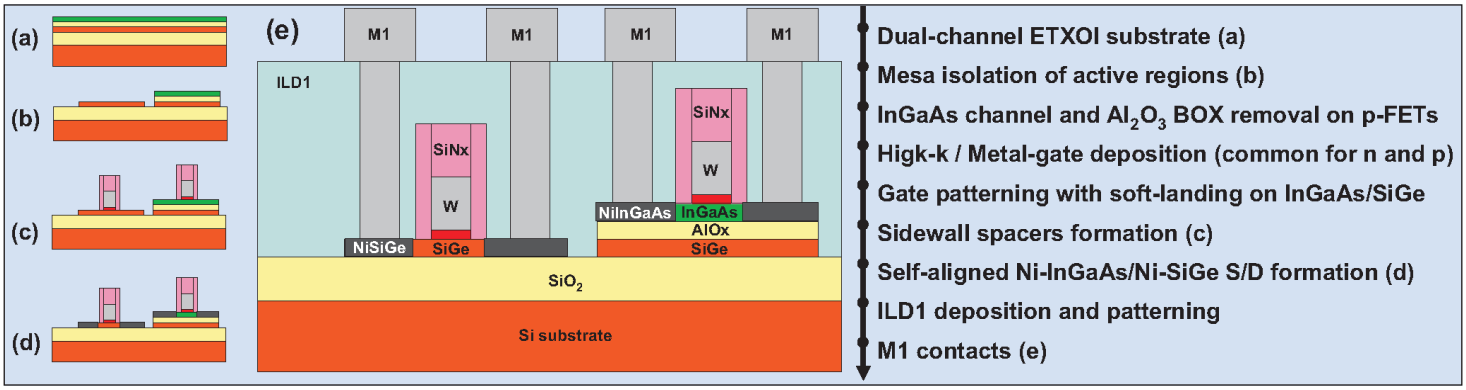

Figure 5. Process flow description for the co-planar co-integration of InGaAs n-FETs and SiGe p-FETs with a common front-end (a, b, c, d, e).

is likely to optimize on-current at shorter gate lengths.

The device layers were grown on p-type InP substrates using metal-organic chemical vapor deposition. An InAlAs buffer layer with wider bandgap than the channel provided back-barrier insulation similar to that achieved with silicon-on-insulator technologies. The researchers believe that replacing the InAlAs with a true insulator such as silicon dioxide (SiO_2) or Al_2O_3 could help them to better tackle short-channel effects. Channels thinner than 20nm could also be beneficial.

A peak mobility at 1550cm^2/V-s for the device was 4–5x that achieved for silicon-based NFETs. The researchers believe that "exploring methods to impart uniaxial tensile stress is more fruitful than trying to simply integrate high-In-content channels". They project that with gate lengths of less than 20nm, such MOSFETs would operate within 5% of the ballistic limit (i.e. where carrier transport occurs without scattering losses).

A team mainly based at IBM Zurich Research Laboratory presented the first demonstration of "dense co-integration of co-planar nano-scaled SiGe p-FETs and InGaAs n-FETs" [L. Czornomaz et al, 2.8]. Direct wafer bonding techniques were used to create "hybrid substrates containing extremely-thin SiGe and InGaAs layers" on an insulating layer of SiO_2 on silicon. The stacked high-mobility III-V semiconductor layers were about 6nm thick. The silicon germanium (SiGe) layer was 8nm. A 10nm aluminium oxide layer was used for bonding the SiGe-on-insulator and III-V layers.

The hybrid substrates enabled the creation of n- and p-channel FETs with ultra-thin bodies on a buried oxide, similar in design to extremely thin body silicon-on-insulator technology (Figure 5). The researchers comment: "Working CMOS inverters are obtained using a common front-end which confirms the viability of this integration scheme for hybrid high-mobility dual-channel CMOS." However, it was not possible to produce ring oscillators since the output from one inverter was not sufficient to trigger another.

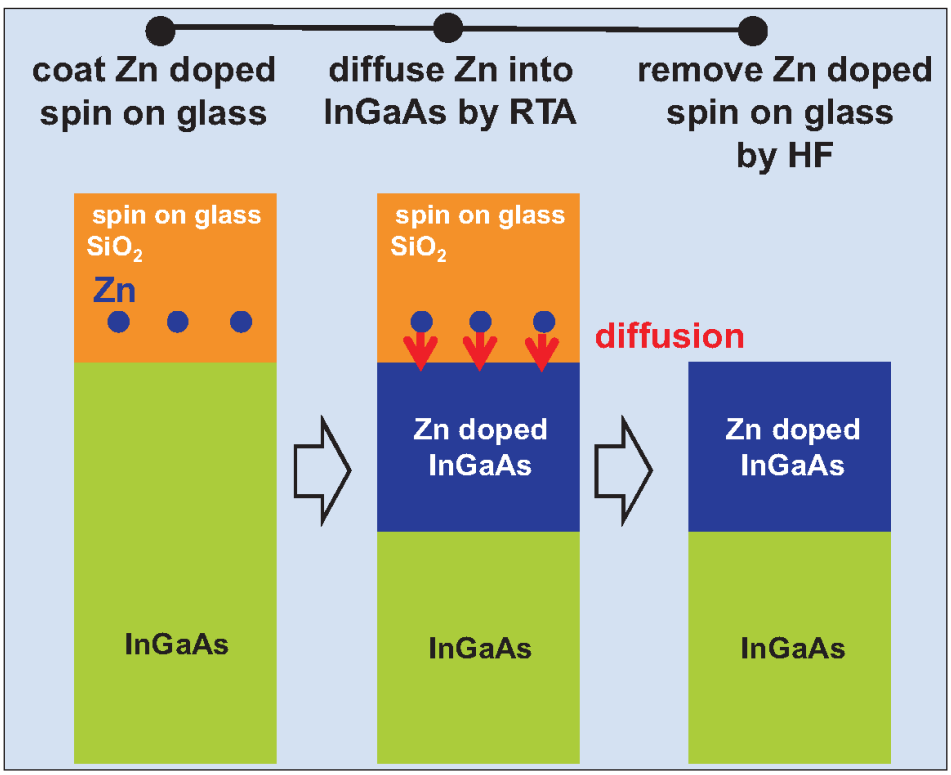

Figure 6. Schematic process flow of Zn diffusion into InGaAs by using Zn-doped spin-on-glass.

Tunneling

University of Tokyo and Sumitomo have developed planar InGaAs tunneling FETs (TFETs) using solid-phase zinc diffusion (Figure 6) to achieve steep p^+/n source junctions without defects [Munetaka Noguchi et al, 28.1]. The structure achieved a low SS of 64mV/dec and high on/off current ratio of 10^6. The SS and on/ratio values are records for planar-type III-V TFETs.

TFETs achieve a steep switching slope by using bandgap filtering of carriers in the source from the tail of the Fermi distribution.

Vertical TFETs have achieved 60mV/dec, the theoretical limit for traditional

planar MOSFETs. However, planar TFETs would be preferable in terms of integration with mainstream silicon processing technologies. Planar InGaAs TFETs with beryllium doped source junctions have only achieved 230mV/dec SS. Besides the poorer performance, one should probably avoid exposure to beryllium since it is poisonous when inhaled and classified as a group 1 carcinogen by the International Agency for Research on Cancer.

Penn State University, US National Institute of Standards and Technology, and IQE Inc. reported on high switching speeds attained by $In_{0.9}Ga_{0.1}As/GaAs_{0.18}Sb_{0.82}$ tunneling FETs with 200nm channel length. The devices use a near-broken gap structure to achieve record drive current of 740μA/μm at 0.5V drain bias. The peak intrinsic RF transconductance was 700μS/μm. These values were, respectively, 2.3x and 2.6x the characteristics reported by Penn State in 2012. The f_T cut-off frequency was 19GHz. At 0.3V drain, f_T was 10GHz.

The device used heterojunction technology to boost on-currents without increasing off-currents at the same time. This was achieved by decreasing the effective barrier height to 0.02eV, compared with ~0.25eV in 2012. The researchers used an internal photoemission spectroscopy (IPE) technique involving a graphene electrode to determine the heterojunction band alignments in the process of tuning the epitaxial metamorphic growth on lattice mismatched InP substrate using solid source MBE.

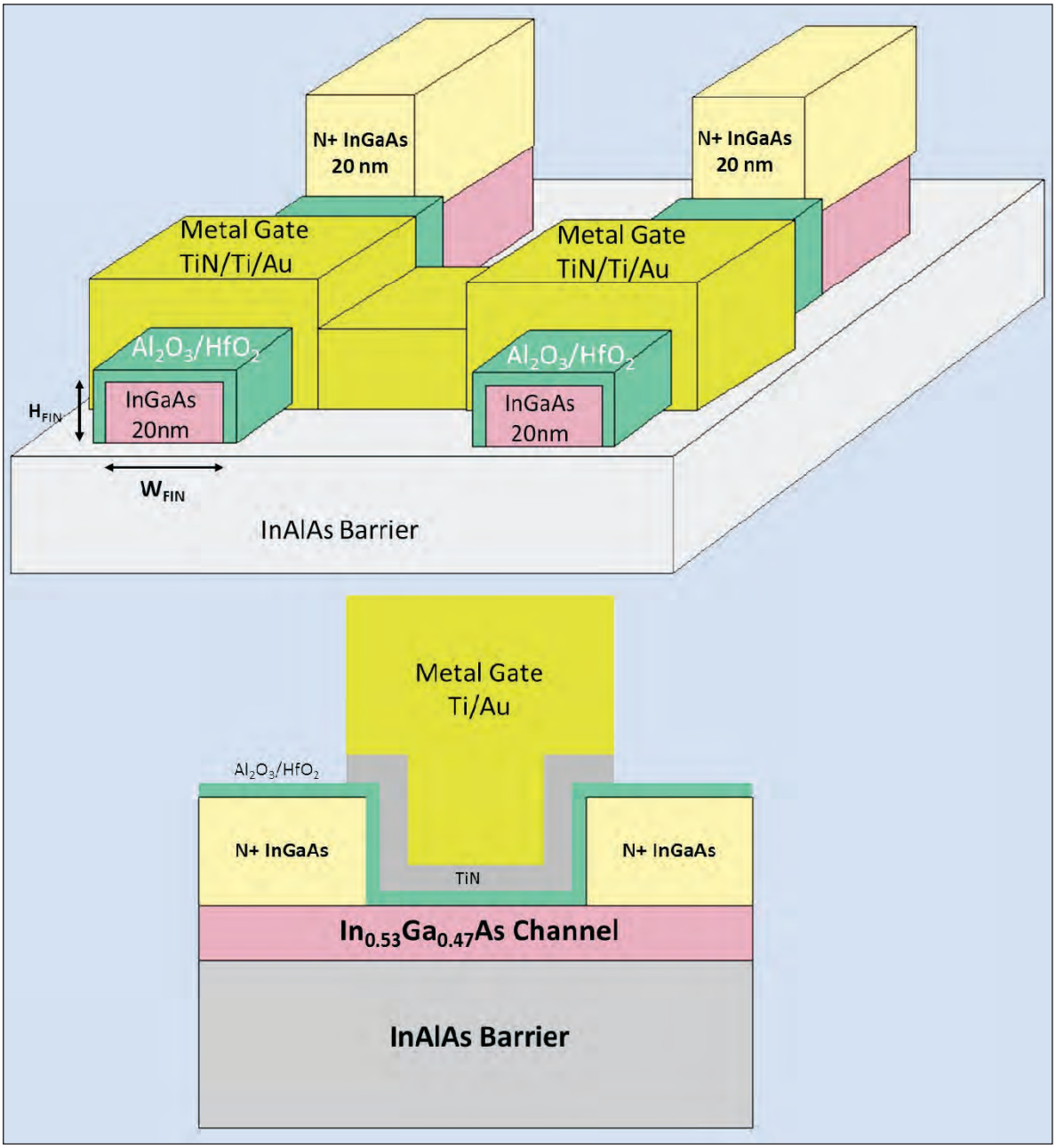

Figure 7. Top: InGaAs quantum-well (QW) tri-gate MOSFET architecture. Bottom: cross-sectional schematic of each fin along gate-length direction.

Contact and gate stack developments

SEMATECH was involved in two pieces of research towards promoting InGaAs technology: in one, a self-aligned nickel source/drain contact process module was developed using wet etch [Rinus T.P. Lee et al, 2.6]; in the other, trigate MOSFETs were produced with a gate stack EOT of less than 1nm aimed at low-power logic [T.-W. Kim et al, 16.9].

The nickel process module work also involved New York State's College of Nanoscale Science and Engineering (CNSE), along with assignees from TSMC and GLOBALFOUNDRIES. The contacts helped the researchers achieve a 25% increase in saturated drain current over non-self-aligned molybdenum contact devices. The reduction in parasitic resistance was 27%.

The thermal stability of the nickel contacts was also

increased over previous reports to 500ºC by the use of a 'novel interlayer', enabling compatibility with standard very large scale integration (VLSI) back-end-of-line metallization processes.

The researchers comment: "Reduced parasitic resistance with a thermally stable, single contact scheme is significant to achieve a manufacturable path for III-V on Si."

The trigate MOSFET development involved University of Texas Austin, CNSE, Tokyo Electron Ltd, Yonsei University, and GLOBALFOUNDRIES. The transistor achieved SS as low as 77mV/dec, and DIBL of 10mV/V. The peak transconductance was 1.5mS/μm at 0.5V drain. The researchers write: "This result is the best balance of $g_{m,max}$ and SS in any reported III-V MOSFETs."

The measured device (Figure 7) had a 60nm gate length with a fin width x height cross-section of 30nm x 20nm. Simulations suggest that the technology could be used for sub-10nm devices.

The sub-1nm EOT was achieved by using two layers of high-k dielectric: 0.7nm of Al_2O_3 and 1.6nm of HfO_2. The InGaAs channel was 20nm thick. The gate electrode metal was titanium nitride applied using atomic-layer deposition. In addition, on-resistance was 253Ω-μm, "the lowest reported in any non-planar III-V MOSFET", according to the team.

Channel thickness and nanowires

The IMEC research center in Belgium has been studying the effect of InGaAs channel thickness on transistor performance [Alireza Alian, 16.6]. The team found that thinner channels improved electrostatic control, but at the cost of degraded mobility. A 3nm channel had a low mobility of 110cm^2/V-s, resulting in loss of drive current. A 10nm channel was able to maintain mobility above 1000cm^2/V-s while still delivering reasonable electrostatic control, as indicated by the SS of 77mV/dec with 10μm gate length. The on/off current ratio was 50,000.

The mobility droop in thin channels is attributed to increased scattering effects from oxide interface roughness/defects, border traps, and InP/InGaAs interface dipoles. A further problem suggested by simulations is that the carriers are effectively heavier. Heavy carriers tend to have lower mobility. IMEC also sees its work as having implications for nanowire devices with small cross-section parameters.

Purdue University reported "for the first time" on the variability and reliability of gate-all-around transistors (Figure 8) using multiple InGaAs nanowires [S. H. Shin et al, 7.5]. Two particular problems are: self-heating with high densities of nanowires to achieve high on-current that also increases the off-state leakage current, degrading the on/off ratio; and, variability in threshold voltage and substhreshold swing of individual nanowires negatively impacts the overall substhreshold performance. With 19 parallel nanowires, the researchers found the self-heating led to a temperature of 420K (~150ºC).

The researchers conclude: "Such variability and heat dissipation must be carefully optimized to fully realize the dramatic scaling potential promised by surrounding-gate transistors."

Finally, researchers in Japan (AIST, Sumitomo Chemical Ltd., and Tokyo Institute of Technology) successfully fabricated triangular InGaAs-on-insulator (InGaAs-OI) n-type MOSFETs with smooth (111)B side surfaces on Si [E Moriyama et al, 2.2]. The triangular-shaped channels were formed by MOVPE growth on narrow InGaAs-OI fins. The bottom widths of the triangles were as small as 30nm. The use of (111)B surfaces for the transistor channel improved device mobility over reference InGaAs-OI tri-gate (by 1.9x) as well as bulk (100) InGaAs nMOSFETs (1.6x). The researchers suggest the new devices have lower interface trap density in the conduction band, suppressing carrier trapping at the MOS interface. The current reached 930μA/μm with 0.3μm gate length. ■

The author Mike Cooke is a freelance technology journalist who has worked in the semiconductor and advanced technology sectors since 1997.

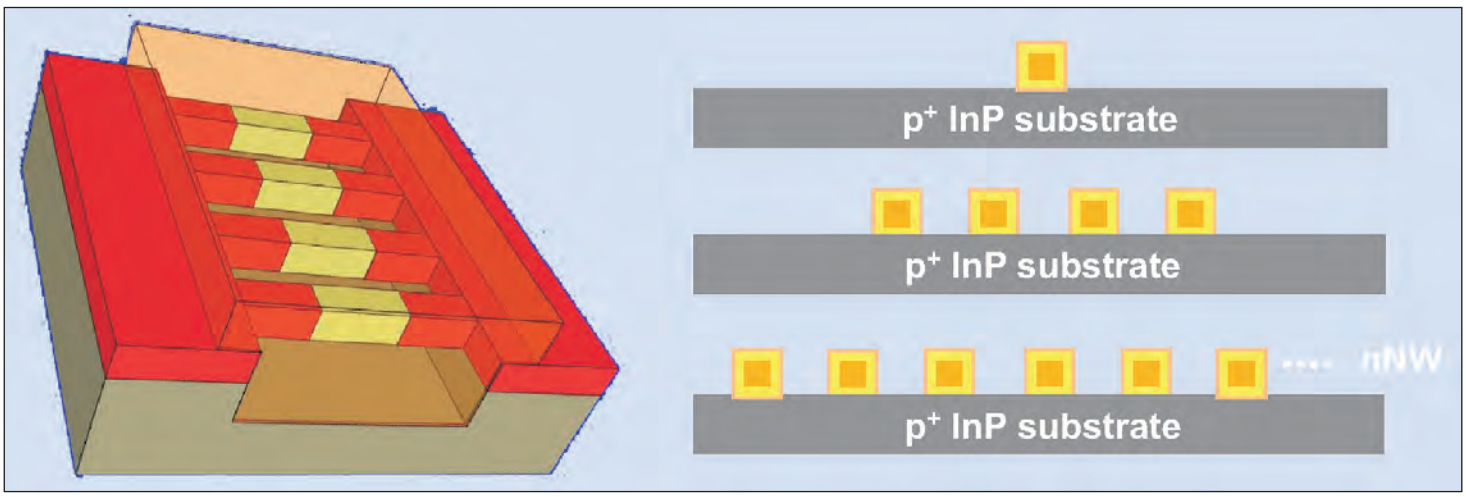

Figure 8. InGaAs GAA NW MOSFET.