PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Beyond CMOS: heterogeneous integration of III–V devices, RF MEMS and other dissimilar materials/devices with Si CMOS to create intelligent microsystems

Thomas E. Kazior

Phil. Trans. R. Soc. A 2014 **372**, 20130105, published 24 February 2014

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

rsta.royalsocietypublishing.org

Review

Cite this article: Kazior TE. 2014 Beyond CMOS: heterogeneous integration of III–V devices, RF MEMS and other dissimilar materials/devices with Si CMOS to create intelligent microsystems. *Phil. Trans. R. Soc. A* **372**: 20130105.
http://dx.doi.org/10.1098/rsta.2013.0105

One contribution of 11 to a Discussion Meeting Issue 'Beyond Moore's law'.

Subject Areas:
electrical engineering, nanotechnology, microsystems

Keywords:
Si CMOS, III–V devices, heterogeneous integration, 'system-on-a-chip'

Author for correspondence:
Thomas E. Kazior
e-mail: thomas_e_kazior@raytheon.com

Royal Society Publishing

Beyond CMOS: heterogeneous integration of III–V devices, RF MEMS and other dissimilar materials/devices with Si CMOS to create intelligent microsystems

Thomas E. Kazior

Raytheon Integrated Defense Systems, 362 Lowell Street, Andover, MA 01810, USA

Advances in silicon technology continue to revolutionize micro-/nano-electronics. However, Si cannot do everything, and devices/components based on other materials systems are required. What is the best way to integrate these dissimilar materials and to enhance the capabilities of Si, thereby continuing the micro-/nano-electronics revolution? In this paper, I review different approaches to heterogeneously integrate dissimilar materials with Si complementary metal oxide semiconductor (CMOS) technology. In particular, I summarize results on the successful integration of III–V electronic devices (InP heterojunction bipolar transistors (HBTs) and GaN high-electron-mobility transistors (HEMTs)) with Si CMOS on a common silicon-based wafer using an integration/fabrication process similar to a SiGe BiCMOS process (BiCMOS integrates bipolar junction and CMOS transistors). Our III–V BiCMOS process has been scaled to 200 mm diameter wafers for integration with scaled CMOS and used to fabricate radio-frequency (RF) and mixed signals circuits with on-chip digital control/calibration. I also show that RF microelectromechanical systems (MEMS) can be integrated onto this platform to create tunable or reconfigurable circuits. Thus, heterogeneous integration of III–V devices, MEMS and other dissimilar materials with Si CMOS enables a new class of high-performance integrated circuits that enhance the capabilities of existing systems, enable new circuit architectures and facilitate the continued proliferation of low-cost micro-/nano-electronics for a wide range of applications.

1. Introduction

Advances in silicon technology continue to revolutionize micro-/nano-electronics. The functionality and complexity of complementary metal oxide semiconductor (CMOS) circuitry has increased and the cost decreased to the point where Si microelectronics has become ubiquitous and touches every aspect of our lives, including communication, transportation, healthcare, commerce and leisure/entertainment. In fact, just about everything we touch contains at least one silicon chip.

Will this continue indefinitely into the future? In general, yes. Is silicon the best technology for every function and application? This answer is not black and white. For example, from the 2011 Radio Frequency and Analog/Mixed Signal chapter of the International Technology Roadmap for Semiconductors, ITRS (http://www.itrs.net/Links/2011ITRS/ 2011Chapters/2011RFAMS.pdf): *For RF, microwave, millimeter wave and mixed signal applications, if the performance requirements can be satisfied by silicon they will be, primarily for cost and integration density reasons*. In some cases, new devices (for example SiGe heterojunction bipolar transistors (HBTs)) or components (silicon microelectromechanical systems (MEMS)) based on similar or compatible materials were integrated onto the CMOS platform to 'enhance' performance and/or provide new capabilities/functions. However, there are cases where continued scaling of CMOS or integration with other Si-based devices/components does not result in increased product performance. In addition, there are cases where non-Si device and component technologies provide superior performance and are required for applications where performance requirements cannot be met with Si. As examples, III–V devices provide superior frequency, gain, noise and power performance, especially for emerging millimetre wave and terahertz applications, and radio-frequency (RF) MEMS provide superior switching characteristics. As a result, *III–V devices will continue to serve niche or performance-driven applications where silicon performance is not adequate* (http://www.itrs.net/Links/2011ITRS/2011Chapters/2011RFAMS.pdf). However, III–V devices and MEMS have traditionally been limited in integration density.

(a) More than Moore

So how can we take advantage of the integration density of Si CMOS and the superior performance of non-Si devices or components? More importantly, what is the best (and most cost-effective) way to integrate dissimilar material and device technologies to create integrated micro- and nano-systems? Traditionally, multichip assemblies (or MCMs) have been used to integrate the 'best' devices, components or technologies into microsystems that meet the performance requirements. An example is the mobile phone, which contains Si and GaAs components. Over time, the complexity, integration density and performance of these multichip assemblies has increased dramatically with ever decreasing footprints. The evolution in the size and functionality of smart phones and tablets is an example. This has been enabled, in part, by advanced chip-level packaging and board-level integration techniques such as ball grid arrays and flip chip assembly (bump bonding). More recently, silicon back-end-of-line (BEOL) micro-fabrication techniques and through-substrate via holes have been applied to packaging and the manufacture of multichip assemblies. Examples include Freescale's RCP [1], DRAPER's iUHD [2], IBM's system-on-package technology based on silicon carriers [3], and RTI's Si interposers (http://www.rti.org/brochures/rti_3d_integration.pdf) [4,5]. However, there are still size constraints and integration and performance limitations for microsystems realized with these approaches. So what is next?

Recent advances in the heterogeneous integration of III–V devices with Si CMOS on silicon substrates will enable the realization of RF and mixed signal circuits that take advantage of the superior performance of III–V devices and the high integration density of Si CMOS (http://www.itrs.net/Links/2011ITRS/2011Chapters/2011RFAMS.pdf). In this review article, I present several approaches for heterogeneously integrating high-performance III–V devices, such as InP HBTs and GaN high-electron-mobility transistors (HEMTs) with Si CMOS on a common substrate.

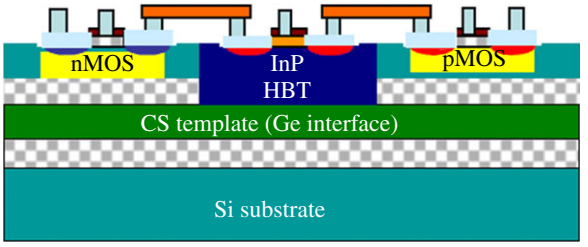

Figure 1. Schematic cross section of Raytheon's COSMOS approach. InP HBTs are integrated with Si CMOS using a III–V BiCMOS process.

I focus on an approach that is analogous to a SiGe BiCMOS process (BiCMOS integrates bipolar junction and CMOS transistors) in which the III–V device is integrated into the CMOS process flow between the CMOS front-end-of-line (FEOL) and BEOL processes. We also present three-dimensional approaches for integration of MEMS switches with III–V devices and/or Si CMOS to create 'intelligent' or 'reconfigurable' circuits. Thus, three-dimensional and heterogeneous integration will be used to create 'intelligent' micro- and nano-systems or sensors on a chip.

2. Chip-scale heterogeneous integration

(a) III–V devices and Si CMOS

Under DARPA funding, several different approaches for integrating Si CMOS with III–V devices have been demonstrated [6]. Each approach has benefits, limitations and challenges. All of the approaches are capable of supporting interconnect lengths and pitches of the order of micrometres (rather than tens of micrometres for traditional multichip assembly integration approaches). In one approach developed at Northrup Grumman [7], a wafer-scale assembly approach is used to integrate completed III–V 'chiplets' on top of a completed CMOS wafer. This approach uses 'microbumps' to interconnect two or more different device and component types with the Si CMOS. To preserve device performance, the different devices are fabricated on their native substrates. In another approach (HRL Laboratories, LLC) [8], III–V epitaxy is transferred to the top of a completed CMOS wafer. The III–V devices are then fabricated *in situ* on top of the CMOS wafer to provide precision placement of the III–V device relative to the Si transistors. A similar approach has been developed at the University of California at Santa Barbara to integrate and precision align III–V lasers and detectors to optical waveguides formed on Si wafers to create optoelectronic integrated circuits [9]. MIT Lincoln Laboratory has adapted three-dimensional integration approaches originally developed for stacking of Si CMOS wafers [10]. In their approach, partially processed CMOS wafers are 'wafer-bonded' face-to-face to partially processed III–V wafers. The Si substrate is removed and through-dielectric via holes are formed to interconnect the CMOS and III–V devices. Additional CMOS layers can be added to create higher levels of three-dimensional integration.

(i) III–V BiCMOS (InP and Si CMOS)

Under the DARPA COSMOS Program, the Raytheon team [11] integrated high-performance InP HBTs with Si CMOS on a common silicon substrate in a process similar to a SiGe BiCMOS process (figure 1). That is, the InP HBT was integrated into the Si CMOS process between the Si CMOS FEOL process and BEOL interconnect process. As device-quality (low defect/dislocation density) III–V epi-layers cannot easily be grown directly on Si substrates, our direct integration approach is based on 'engineered' silicon substrates, which are a variant of standard silicon-on-insulator (SOI) wafers. The starting substrate is known as silicon on lattice-engineered substrate (SOLES). SOLES contains a buried III–V template layer that enables the direct growth of high-quality III–V

epitaxial material in windows directly on the silicon substrate. During the COSMOS Program, the buried III–V template layer was Ge, although the substrate fabrication process is compatible with GaAs or InP template layers as well. Ge was selected for the initial demonstrations because (i) Ge is already used in Si foundries, removing the barrier of introducing a non-standard substrate into a Si foundry, and (ii) high-quality GaAs buffer layers can easily be grown on Ge (near-perfect lattice match). Device-quality InP epitaxial films can easily be grown on the GaAs buffer using metamorphic growth techniques. The SOLES wafers were manufactured by Soitec using their SmartCut^{TM} process, and were successfully scaled to 200 mm diameter and processed in a silicon CMOS foundry. Details of the SOLES have been previously published [12–14].

The III–V BiCMOS process flow consists of: (i) starting substrate (SOLES); (ii) Si CMOS device fabrication; (iii) III–V 'window' formation; (iv) III–V epitaxial growth; (v) III–V device fabrication; and (vi) multilayer interconnect fabrication. After completion of CMOS device fabrication, windows are lithographically defined and etched to reveal the III–V template layer. As the III–V growth windows are defined as part of the CMOS fabrication process, the III–V epitaxial material can be grown selectively and arbitrarily across the substrate as required for the particular circuit or application.

For InP HBTs, a GaAs nucleation layer on the Ge template layer is grown, followed by a metamorphic buffer layer and then the InP HBT epi-layers. One of the biggest challenges of this approach is the growth of high-quality III–V epitaxial material in windows on the Ge template layer. (Note: all of the III–V epitaxial material reported in this work is grown by molecular beam epitaxy (MBE).) This requires pristine, damage-/defect-free surfaces in the bottom and on the edges of the windows in the SOLES wafers containing CMOS transistors. Surface defects could nucleate dislocations, which propagate through the buffer layer and into the device active layer, compromising device performance and reliability. Most of the surface defects/dislocations could be annihilated by increasing the thickness of the metamorphic buffer layer. However, owing to the poor thermal conductivity of the metamorphic buffer layer, the buffer layer thickness must be minimized to prevent the InP HBT performance from being thermally limited. Thus, a compromise must be made between thermal resistance and defect density. Edge defects could nucleate the growth of nanowires that penetrate into the device active layer. Additional challenges include: optimization of growth methods that support the nucleation of low-defect-density, anti-phase domain-free GaAs on the Ge surface in windows in SOLES; and prevention of Ge out-diffusion. Details of the InP epi growth process and challenges for the growth in windows were previously published [11,15]. It should be noted that some of these growth challenges (e.g. elimination of metamorphic buffer layer and GaAs–Ge interface) would be eliminated by the use of an InP template layer in the SOLES wafer.

After extensive windows etch and epi growth process optimization, device-quality InP HBT epi was successfully and repeatably grown in windows on 100 and 200 mm diameter SOLES wafers containing Si CMOS. An example of InP HBT adjacent to Si CMOS is shown in figure 2. InP HBTs with a $0.5 \times 5\,\mu m^2$ emitter grown in windows on SOLES as small as $15 \times 15\,\mu m^2$ exhibit gain (β), f_t and f_{max} of 40, more than 200 GHz and more than 200 GHz, respectively, performance similar to the InP HBTs on a native InP substrate [16].

For the initial demonstrations, a standard, Au-based multilayer interconnect process, commonly used in III–V foundries, was adapted to interconnect the InP HBTs and Si CMOS. To facilitate interconnecting of the III–V devices and CMOS transistors, the thickness of the III–V epitaxial layers and depth of the windows are optimized such that the III–V devices and CMOS transistors are coplanar. This approach provides the maximum flexibility and intimacy (shortest interconnect length) for placement of III–V devices relative to Si CMOS. Heterogeneous interconnect yields on daisy chain test structures with up to 4000 interconnects exceeded 99.9% with CMOS to III–V device spacing (interconnect length) as small as 2.5 μm and interconnect pitch as small as 5 μm [17]. In the future, standard Cu and Al CMOS multilayer interconnect metallurgy will be adopted. This will allow the integration process to take advantage of the high yield of the CMOS BEOL.

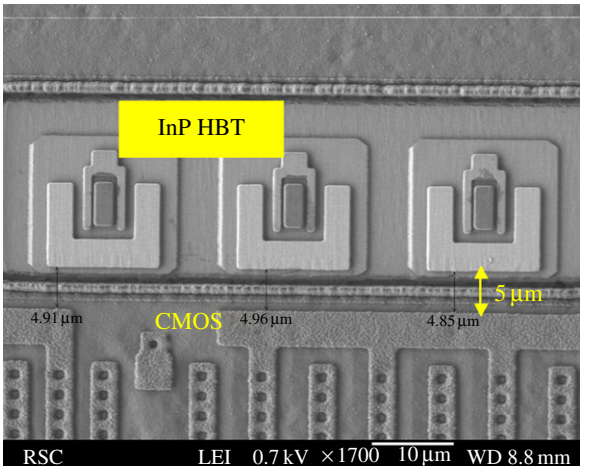

Figure 2. SEM image showing InP HBTs in windows adjacent to Si CMOS. (Online version in colour.)

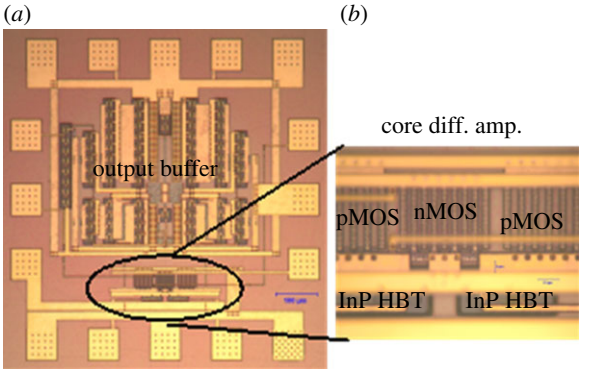

Figure 3. Optical image of InP–Si CMOS differential amplifier with output buffer and bias circuit (*a*) and differential amplifier core (*b*). (Online version in colour.)

Examples of completed demonstration circuits are shown in figure 3 (a high-speed InP–Si CMOS differential amplifier [18]—the COSMOS Phase 1 demonstration vehicle containing high-speed InP HBT differential pairs and CMOS current sources and an InP HBT buffer amplifier) and 4 (a high-dynamic-range digital-to-analogue converter (DAC) with over 50 000 CMOS transistors and 1500 InP HBTs) [19]. The DAC was implemented with 180 nm CMOS. Fully functional DACs were realized. Also shown in figure 4 is a schematic showing the intimate placement and intermixing of InP HBTs relative to the Si CMOS that is enabled by our integration approach. A key feature of our III–V BiCMOS integration process is the ability to integrate control and calibration circuitry directly on the chip. (All CMOS calibration circuits are shown in the upper left-hand corner of the DAC in figure 4.) This intimate integration of feedback circuitry enables the circuit performance to be optimized to eliminate as-fabricated performance variation, to adapt the chip to its operational environment (e.g. *in situ* correction for ambient temperature variation) and to extend chip lifetime by *in situ* calibration to counteract effects of drift and aging.

(b) GaN and Si CMOS

Using a similar process, we also integrated GaN HEMTs with Si CMOS [20] (figure 5). As device-quality (low defect/dislocation density) III–N material can be grown directly on Si ⟨111⟩ substrates, SOLES wafers are not required. Rather, GaN HEMT epi was grown on Si surfaces in

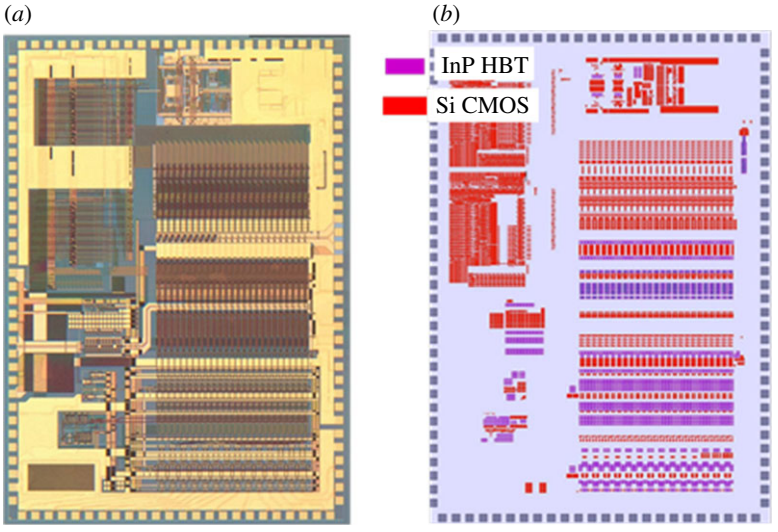

Figure 4. Optical image of completed COSMOS DAC-containing on-chip calibration circuitry (*a*) and CAD drawing showing intimate intermixing of InP HBTs and Si CMOS enabled by our integration approach (*b*). The circuitry in the upper left-hand corner is the on-chip calibration circuitry. (Online version in colour.)

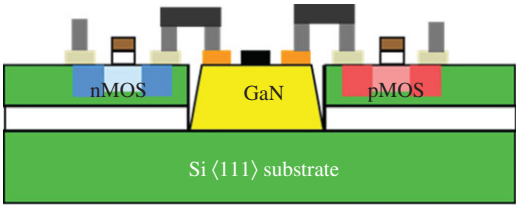

Figure 5. Schematic cross section of Raytheon's GaN–Si CMOS integration process. GaN HEMTs are fabricated in windows on Si SOI wafer. (Online version in colour.)

windows on commercially available SOI wafers, where the Si ⟨100⟩ handle wafer of traditional SOI is replaced by a high-resistivity Si ⟨111⟩ substrate. We incorporated lessons learned from the growth on InP HBT epi in windows to optimize the windows etch process to eliminate surface and edge defects. Nevertheless, several challenges remained, including optimization of the nucleation and buffer layer to simultaneously minimize stress (usually achieved by thicker strain relief buffer layers) and wafer bow and minimize buffer thermal resistance (achieved by thinner buffer layers) as well as eliminate auto-doping at the Si–GaN interface. For the purposes of this initial demonstration, 1 µm CMOS was fabricated in a silicon fab on 100 mm diameter wafers and the GaN transistors and interconnects were fabricated in a III–V fab. As this approach uses a variation of standard SOI wafers, 200 mm (or greater) diameter wafers are readily available providing a path to integrating advanced Si CMOS (180 nm or smaller) with GaN as well as performing the entire fabrication process in a silicon fab.

MBE is the preferred 'GaN in windows' growth technique owing to the lower growth temperature when compared with more standard metal organic chemical vapour deposition (MOCVD). The MOCVD growth temperature (approx. 1000°C) leads to degradation of Si CMOS device characteristics. Details of the growth of GaN HEMT epitaxy in windows on Si by MBE have been reported previously [21]. To facilitate heterogeneous interconnect formation, the thickness of the GaN HEMT epi is chosen such that the top of the GaN transistor is coplanar with the CMOS transistors.

GaN HEMTs have been successfully fabricated in windows on SOI wafers containing Si CMOS transistors. The GaN HEMT fabrication steps are identical to the process steps used for fabricating

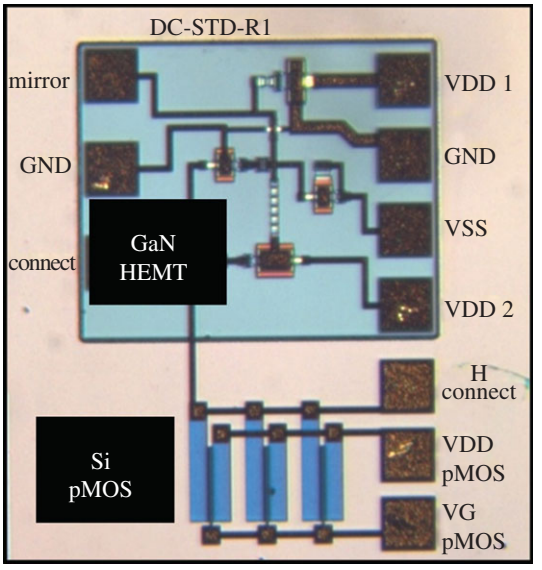

Figure 6. Optical micrograph of the first GaN–Si CMOS heterogeneously integrated MMIC: GaN–Si pMOS current mirror for gate bias control of a GaN amplifier. (Online version in colour.)

GaN HEMTs on SiC substrates. The GaN HEMT in windows achieved an open channel current (I_{max}) of 1.1 A mm^{-1}, a transconductance (g_m) of 270 mS mm^{-1} and good pinch-off characteristics. RF power measurements revealed similar performance as GaN on SiC substrates [20].

As a demonstration vehicle, we designed and fabricated a GaN amplifier heterogeneously integrated with pMOS gate bias control circuitry (a D-mode field effect transistor (FET) current mirror with RF output stage). This simple building block circuit, which consists of four GaN transistors, several GaN diodes for level shifting and a pMOS transistor as a constant-current source, takes advantage of superior V_t control in CMOS to set GaN FET quiescent bias. An optical image of the heterogeneously integrated circuit is shown in figure 6. This was the first successful demonstration of a fully functional heterogeneously integrated GaN HEMT–Si CMOS circuit. Using this approach, we are currently fabricating amplifier circuits with *in situ* gate and drain bias control. In addition, the process is being scaled to 200 mm wafers and 180 nm CMOS under the DARPA DAHI Program.

3. Integration with MEMS

(a) MEMS and GaAs

Under the DARPA IRFFE Program, we successfully integrated RF MEMS capacitive membrane switches with GaAs to create a tunable broadband RF amplifier [22,23]. The GaAs chip included a single stage amplifier with MEMS tunable input and output matching networks and on-chip power and thermal sensors for providing feedback to an off-chip controller (figure 7). (Note: for this demonstration, no effort was made to create a compact design.) When compared to semiconductor-based switches, the Raytheon capacitance MEMS switch shows a 20–50 times improvement in switching figure of merit (on-resistance × off-capacitance) [24]. The RF performance of the broadband, MEMS-tuned amplifier could be optimized (tuned) for narrow instantaneous performance over a more than 8 to 1 frequency band in a matter of tens of milliseconds convergence or settling time. More importantly, RF output power and efficiency approached the power and efficiency of fixed tuned narrowband amplifiers and was significantly better than what could be achieved with a fixed tuned broadband amplifier.

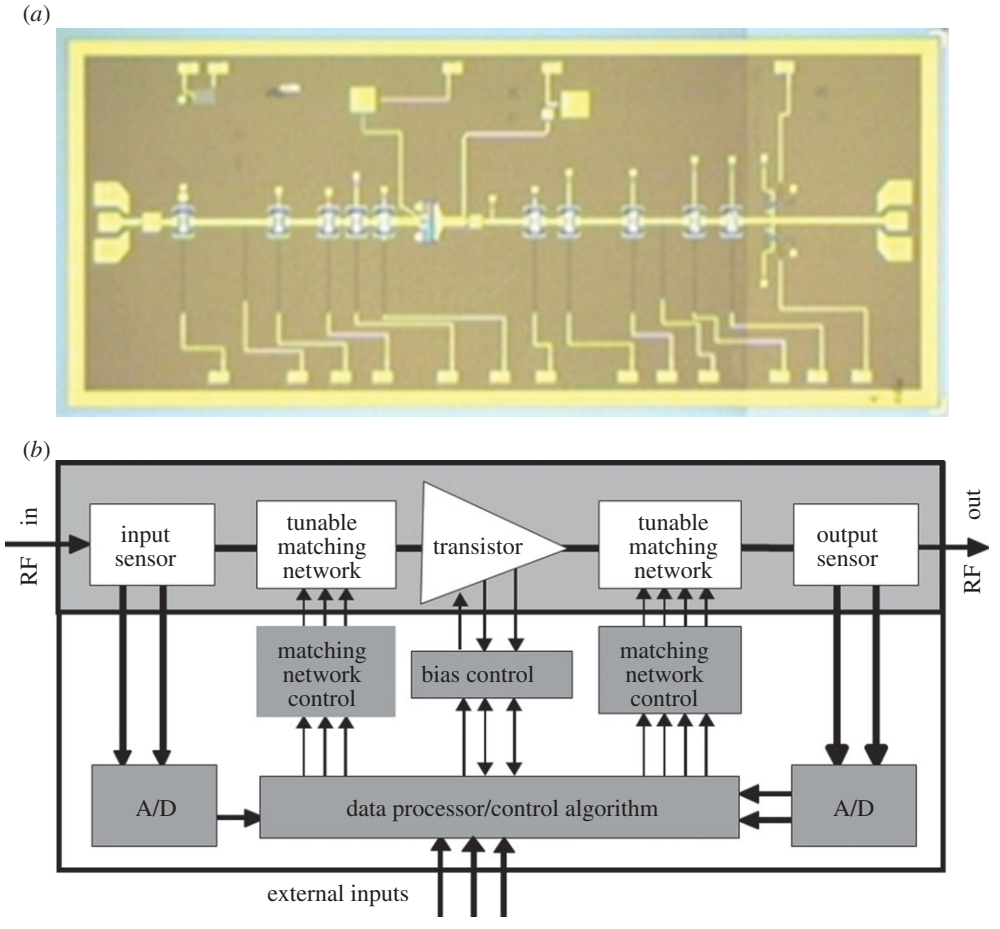

Figure 7. (*a*) Optical image and (*b*) block diagram of RF MEMS integrated with a GaAs amplifier to create a tunable broadband RF amplifier. In this iteration, the control circuitry (white box) was off-chip. (Online version in colour.)

For this demonstration, the chip was mounted onto a separate controller board. Using the III–V BiCMOS process described in §2, the CMOS controller can be directly integrated with the MEMS-tuned III–V amplifier to create an 'intelligent', high-performance reconfigurable amplifier on a chip.

(b) MEMS and GaN

Under the AFRL ARCHIPELA Program, we extended this work and integrated the RF MEMS capacitive membrane switches with GaN transistors to create a reconfigurable-on-the-fly GaN RF amplifier (figure 8) [25]. The design also supports integration of 'on-chip' bias control and the ability to tune the amplifier performance to compensate for impedance variations.

There were two key new features in this approach. (i) To create a more compact design, the MEMS switches were (three dimensionally) integrated on top of the GaN devices (figure 8) and connected with RF through dielectric vias. (ii) The GaN transistor was fabricated on Si substrates. This provides a path to heterogeneous integration of the GaN amplifiers with Si CMOS control circuitry using the approach described in §2b. Future integration with Si CMOS will allow *in situ* control (adaptive bias, modulation, linearization and self-healing) as well as fully reconfigurable amplifiers or RF field programmable arrays and ultimately intelligent, reconfigurable microsystems.

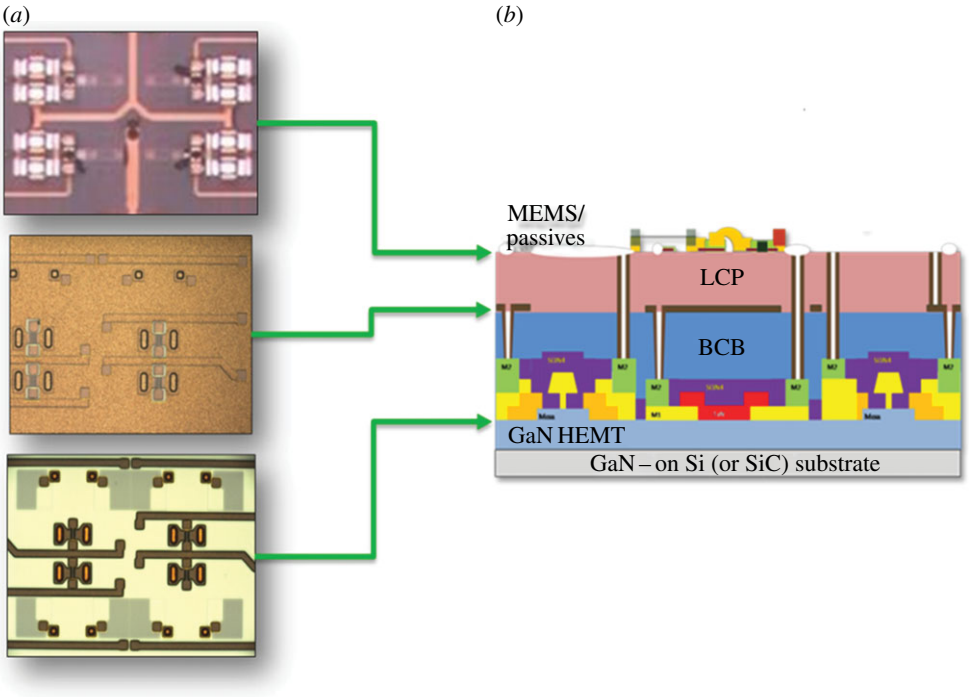

Figure 8. Schematic cross section (*b*) and optical images (*a*) of different layers of three-dimensional, integrated RF MEMS on GaN on silicon HEMT. (Online version in colour.)

(c) *In situ* performance sensing and autonomous recalibration

The IRFFE and ARCHIPELA Programs described above employed closed-loop sensing and healing to increase RF amplifier performance. In each of these cases, on-chip sensors serve as micro-instruments to provide the feedback of system response for automatic gain equalization to allow for agile self-reconfiguration. This concept of automatic gain control for an individual amplifier can be extended to multiple amplifiers chained together in serial and parallel combinations, reconfiguring the system for changes in frequency, bandwidth and dynamic range. In this case, automatic gain control in the receiver can adjust to the new nonlinearities presented by a reconfigured transmitter to reduce dynamic range requirements on the digital back end.

For example, next-generation communication systems will require the capability to sense their electromagnetic environment, and adapt component configurations to maintain system performance or to reconfigure for an alternative mode or frequency of operation based on available spectrum. Raytheon has been investigating techniques for sensing and responding to changes in circuit performance for both yield enhancement and autonomous recalibration of RF integrated circuits (ICs) in the field. Silicon sensors and actuators were integrated with an on-chip digital processor to perform calibration of gain flatness, phase error and linearity/dynamic range of a silicon microwave receiver [26]. The on-chip sensors, actuators and control algorithms can also be used for self-healing, tuning or calibrating for circuit drift due to aging or temperature fluctuation. The *in situ* control circuitry can also be used to improve circuit yield (to performance), by compensating for (calibrating out) inherent non-uniformities or variability associated with scaled advanced semiconductor technologies and helping to further drive down the cost of microsystems on a chip.

An important aspect of healing architectures is the ability to accomplish healing with minimal overhead in circuit area and power. The Raytheon team employs surrogate modelling and

efficient healing algorithms to minimize the computational load, reducing both of these important parameters [27].

4. Heterogeneous/three-dimensional integration with other components

While the MEMS capacitance membrane switches, described in §3, have been used to realize tunable matching networks and high-Q tunable filters, and can be integrated in multichip assemblies or on top of CMOS circuitry and III–V-based RF amplifiers, the resulting tunable components are quite large (occupy a significant fraction of the physical size of any transceiver), require hermetic packaging to prevent MEMS reliability problems, such as stiction, and are relatively slow (reconfigurability of the order of tens to hundreds of microseconds). For chip-scale microsystems, the ideal situation would be to heterogeneously integrate a MEMS structure or other type of high-speed switch material directly into the CMOS or III–V wafer. The intimate on-chip connection would further reduce parasitics, improving insertion loss (IL), bandwidth and tuning speed.

(a) Si and GaN chip-scale resonators

To this end, several investigators have demonstrated Si-based resonators that are fabricated as part of the standard Si FEOL process (e.g. Weinstein & Bhave [28]). One such device, known as a resonant body transistor, has been demonstrated with resonant frequencies up to 37 GHz and an fQ product of 2.2×10^{13} and a footprint of the order of a CMOS transistor cell [29]. These novel, high-Q, semiconductor-based bulk acoustic wave (BAW) resonators are 1000 times smaller than traditional EM resonators, and have been demonstrated using standard Si CMOS. Resonator figures of merit could be improved dramatically if these resonators could be realized with piezoelectric material. With the emergence of GaN on Si technology, several investigators have demonstrated GaN-based BAW resonators (figure 9) [30–32]. The strong piezoelectric characteristics and high breakdown fields of the GaN material system are ideal for realizing high-power handling BAW resonators with frequency × quality factor products $fQ > 10^{13}$. GaN HEMT BAW resonators provide three attractive functionalities:

- active HEMT sensing enabling operation at orders of magnitude higher frequency than passive sensing;
- out-of-line switching to provide 10^5 times signal suppression and low IL; and
- low parasitic capacitance.

The MIT Hybrid MEMS Group is taking this one step further. The GaN on Si resonators described above are being integrated into the GaN–Si CMOS platform described in §2b. Rather than building RF transceivers with traditional large off-chip filter banks, the chip-scale resonators can be integrated with FET switches directly into the transceiver chip with *in situ* CMOS control to create compact, real-time-configurable, chip-scale filter banks. The high Q of the constituent resonators ($Q > 500$ at 18 GHz) provides low IL (IL < 4 dB) and excellent frequency selectivity. The mechanical coupling, isolating the input and output electrical signals, generates superior stop-band rejection (more than 60 dB). Strong mechanical coupling ensures low IL, large power handling (more than 20 dBm) and insensitivity to fabrication variations. Out-of-line switching using active elements embedded inside the acoustic resonators provides 50 dB isolation between the input and output and introduces no additional IL into the filter network. The intimate integration of this filter with the GaN amplifiers and CMOS control circuitry ensures fast response time (critical for 'on-the-fly configuration'), reduced loss and interference due to off-chip parasitics and the smallest possible form-factor transceiver.

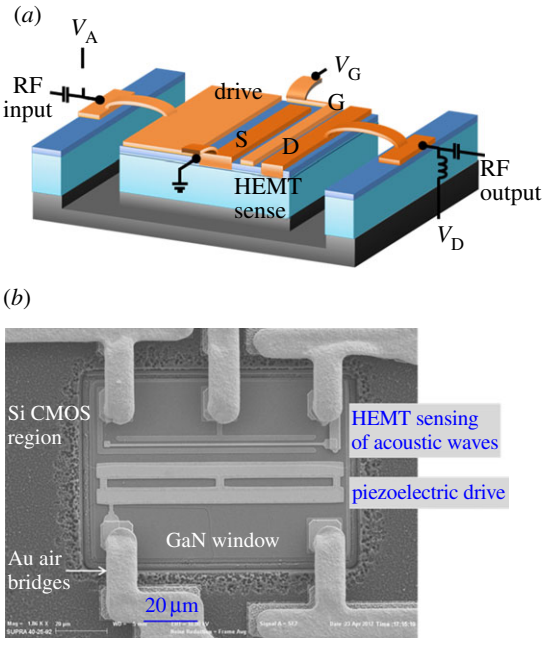

Figure 9. Schematic of a GaN HEMT BAW resonator (*a*). SEM image of a GaN HEMT BAW resonator in 'windows' in SOI wafers (*b*). (Online version in colour.)

(b) Phase-change materials as switches

Another attractive switch material is phase-change materials, for example the chalcogenides. These materials undergo reversible, temperature-dependent changes in resistivity between conducting and insulating states in the tens of nanoseconds time scale (more than 100 times faster than MEMS). These materials can be deposited at relatively low temperatures, are compatible with standard CMOS fabrication deposition and patterning processes and, therefore, can be integrated directly into a standard CMOS BEOL process flow. As an example, the phase-change material germanium antimonide telluride (GST) has been developed and implemented for non-volatile memory applications [33]. The chalcogenides (GST, germanium telluride, germanium antimonide) also show promise as low-impedance (low-loss), high-speed (tens of nanoseconds) switches for RF circuits with off/on resistance ratios as high as 10^5 [34]. Another phase-change material that has received attention as a low-loss (1 dB at 40 GHz), fast-response-time (30 ns) switch, compatible with IC technology, is vanadium oxide (VO_x) [35]. Vanadium oxide has been used to fabricate low-loss, shunt and series, broadband (50 MHz–35 GHz) switches and tunable band-stop filters operating around 10 GHz [36].

5. Future prospects

For size- and performance-driven applications in which high levels of integration are needed, there will be a continued drive towards heterogeneous integration of III–V devices and/or devices/components based on other dissimilar materials with silicon technologies. In the future, in addition to the MEMS and phase-change materials discussed above, these other 'dissimilar' materials will include: carbon-based materials (carbon nanotubes, graphene and diamond), semiconducting and metallic nanowires, two-dimensional semiconductors, complex oxides, etc. The integration of these dissimilar materials will require novel Si fab compatible fabrication techniques. One approach showing significant promise is the electrostatic assisted self-assembly technique being developed at Penn State [37]. The approach is being used to integrate 'chemically

functionalized' metallic and semiconducting nanowires and 'chiplets' onto Si CMOS wafers to create chemical and biological sensors.

The heterogeneous integration examples presented in this paper were realized with hybrid or split manufacturing approaches where Si CMOS was manufactured in CMOS foundries and the non-Si devices/components were integrated in III–V or boutique foundries. While this is useful for demonstrating the viability of these integration approaches, there are severe limitations in cost and yield. The ultimate cost-effective solution will come from fabricating the non-Si devices on large-diameter wafers in the same silicon foundries as the Si-based devices. To achieve this vision for the III–V devices, significant development is required to create Si fab compatible III–V process modules. As a step in this direction, several investigators have reported Au-free contacts to III–V devices with similar electrical performance as traditional Au-based contacts typically used for III–V devices [38–40]. Recent advances in the epitaxial growth of GaN on silicon has shown that the growth of device-quality GaN HEMTs epitaxy on 200 mm diameter Si substrates is feasible [40,41]. Both of these results open up the possibility of fabricating GaN HEMTs in existing 200 mm Si foundries, thereby providing a cost-effective solution for heterogeneous integration of GaN and Si CMOS. Some of these integration challenges will be addressed under the DARPA DAHI Foundry Program.

6. Summary

As presented in this paper, there has been significant advancement in the heterogeneous integration of III–V and other non-Si materials and devices on a Si CMOS platform. Heterogeneous integration will enable designers to take advantage of the high-frequency performance of the III–V devices, the unique performance characteristics of other dissimilar devices and materials systems, and the high-density integration and processing capabilities of silicon, to create a new class of intelligent, 'mostly digital' RF and mixed signal circuits whose performance cannot be achieved with either Si, III–V or other non-Si technology alone nor with traditional multichip assemblies.

Some of the technology benefits include:

— flexible functionality enabling a system to reconfigure itself to perform multiple missions;
— optimization of performance including maximizing gain and efficiency under conditions of varying centre frequency, varying output power, varying output load impedance and varying requirements for linearity;
— eliminating performance variations caused by temperature drift and aging; and
— increasing manufacturing tolerances and reducing the need for tuning of individual amplifiers to meet system specifications.

The availability of intelligent circuits (microsystems- and sensors-on-a-chip) will have a revolutionary impact on a large number of systems and continue the revolution of micro- and nano-electronics.

Funding statement. This work is supported in part under the DARPA DAHI Program (Dr Sanjay Raman), including COSMOS (contract no. N00014-07-C-0629 monitored by Dr Harry Dietrich and Dr Paul Maki–ONR) and GaN–Si CMOS Integration (contract no. FA86500-10-C-7055 monitored by Mr Jim Sewell–AFRL), the DARPA Intelligent RF Front End (IRFFE), monitored by the Air Force Research Laboratory, Sensors Directorate, and the United States Air Force Research Laboratory ARCHIPELA Program (contract no. FA8650-09-C-1504) monitored by Mr Kelson Chabak.

References

1. Keser B, Amrine C, Trung D, Hayes S, Leal G, Lytle W, Mitchell D, Wenzel R. 2008 Advanced packaging: the redistributed chip package. *IEEE Trans. Adv. Packag.* **31**, 39–43. (doi:10.1109/TADVP.2007.909456)

2. Thompson J, Tepolt G, Mueller LRA, Langdo T, Gauthier D, Smith B. 2011 Embedded package wafer bow elimination techniques. In *IEEE 61st Electronic Components and Technology Conference* (*ECTC*), *Lake Buena Vista, FL, 31 May–3 June*, pp. 55–58. Piscataway, NJ: IEEE. (doi:10.1109/ECTC.2011.5898491)
3. Knickerbocker J *et al.* 2005 Development of next-generation system-on-package (SOP) technology based on silicon carriers with fine-pitch chip interconnection. *IBM J. Res. Dev.* **49**, 725–753. (doi:10.1147/rd.494.0725)
4. Temple D, Malta D, Lannon JM, Lueck M, Huffman A, Gregory C. 2008 Bonding for 3-D integration of heterogeneous technologies and materials. *ECS Trans.* **16**, 3–13. (doi:10.1149/1.2982849)
5. Temple D, Lannon JM, Malta D, Robinson JE, Coffman PR, Welch TB, Skokan MR, Moll AJ, Knowlton WB. 2007 Advances in 3D integration of heterogeneous materials and technologies. *Proc. SPIE* **6544**, 654401. (doi:10.1117/12.722502)
6. Raman S, Chang T-H, Dohrman CL, Rosker MJ. 2010 The DARPA COSMOS Program: the convergence of InP and silicon CMOS technologies for high-performance mixed-signal. In *Int. Conf. on Indium Phosphide and Related Materials* (*22nd IPRM*), *Kagawa, Japan, 31 May–4 June*, PL2, 5 pp. Piscataway, NJ: IEEE. (doi:10.1109/ICIPRM.2010.5516241)
7. Gutierrez-Aitken A *et al.* 2009 Advanced heterogeneous integration of InP HBT and CMOS Si technologies for high performance mixed signal applications. In *IEEE MTT-S Int. Microwave Symposium Digest* (*MTT'09*), *Boston, MA, 7–12 June*, pp. 1109–1112. Piscataway, NJ: IEEE. (doi:10.1109/MWSYM.2009.5165895)
8. Li JC *et al.* 2008 Heterogeneous wafer-scale integration of 250 nm, 300 GHz InP DHBTs with a 130 nm RF-CMOS technology. In *IEEE Int. Electron Devices Meeting* (*IEDM 2008*), *San Francisco, CA, 15–17 December 2008*, pp. 944–948. Piscataway, NJ: IEEE. (doi:10.1109/IEDM.2008.4796802)
9. Liang D, Bowers JE. 2010 Optoelectronic device integrated on Si. In *III–V compound semiconductors: integration with silicon-based microelectronics* (eds T Li, M Mastro, A Dadgar), pp. 281–334. Boca Raton, FL: CRC Press.
10. Chen CL *et al.* 2009 Wafer-scale 3D integration of InGaAs image sensors with Si readout circuits. In *IEEE Int. Conf. on 3D System Integration* (*3DIC 2009*), *San Francisco, CA, 28–30 September*, 4 pp. Piscataway, NJ: IEEE. (doi:10.1109/3DIC.2009.5306556)
11. Kazior TE, Herrick K, LaRoche JR. 2010 Direct growth of III–V devices on silicon. In *III–V compound semiconductor: integration with silicon-based microelectronics* (eds T Li, M Mastro, A Dadgar), pp. 257–280. Boca Raton, FL: CRC Press.
12. Dohrman CL, Chilukuri K, Isaacson DM, Lee ML, Fitzgerald EA. 2006 Fabrication of silicon on lattice-engineered substrate (SOLES) as a platform for monolithic integration of CMOS and optoelectronic devices. *Mater. Sci. Eng. B* **135**, 235–237. (doi:10.1016/jmseb.2006.08)
13. Chilukuri K, Mori MJ, Dohrman CL, Fitzgerald EA. 2007 Monolithic CMOS-compatible AlGaInP visible LED arrays on silicon on lattice-engineered substrates (SOLES). *Semicond. Sci. Tech.* **22**, 29–34. (doi:10.1088/0268-1242/22/2/006)
14. Maleville C, Mazuré C. 2004 Smart-Cut® technology: from 300 mm ultrathin SOI production to advanced engineered. *Solid State Electron.* **48**, 1055–1063. (doi:10.1016/j.sse.2003.12.029)
15. Liu WK *et al.* 2009 Monolithic integration of InP-based transistors on Si substrates using MBE. *J. Crystal Growth* **311**, 1979–1983. (doi:10.1016/j.jcrysgro.2008.10.061)
16. Ha W *et al.* 2008 Small-area InP DHBTs grown on patterned lattice-engineered silicon substrates. In *Proc. 66th Device Research Conf., Santa Barbara, CA, 23–25 June*, pp. 311–312. Piscataway, NJ: IEEE.
17. Kazior TE, LaRoche JR, Hoke W. 2012 Heterogeneous integration of III–V devices and Si CMOS on a Si substrate. *ECS Trans.* **50**, 1039–1045. (doi:10.1149/05009.1039ecst)
18. Kazior TE *et al.* 2009 A high performance differential amplifier through the direct monolithic integration of InP HBTs and Si CMOS on silicon substrates. In *IEEE MTT-S Int. Microwave Symposium Digest* (*MTT'09*), *Boston, MA, 7–12 June*, pp. 1113–1116. Piscataway, NJ: IEEE. (doi:10.1109/MWSYM.2009.5165896)
19. Kazior TE *et al.* 2010 High performance mixed signal circuits enabled by the direct monolithic heterogeneous integration of InP HBT and Si CMOS on a silicon substrate. In *IEEE Compound Semiconductor Integrated Circuit Symp., Monterey, CA, 3–6 October 2010*, H.1, 4 pp. Piscataway, NJ: IEEE. (doi:10.1109/CSICS.2010.5619670)

20. Kazior TE, Chelakara R, Hoke W, JBettencourt J, Palacios T, Lee HS. 2011 High performance mixed signal and RF circuits enabled by the direct monolithic heterogeneous integration of GaN HEMTs and Si CMOS on a silicon substrate. In *IEEE Compound Semiconductor Integrated Circuit Symp., Waikoloa, HI, 16–19 October 2011*, E.1, 4 pp. Piscataway, NJ: IEEE. (doi:10.1109/CSICS.2011.6062443)
21. Hoke W *et al.* 2012 Monolithic integration of silicon CMOS and GaN transistors in a current mirror circuit. *J. Vac. Sci. Technol. B* **30**, 02B101. (doi:10.1116/1.3665220)
22. Lardizabal S, Pillans B, Molfino R, Bettencourt J, Malczewski A, Herrick K, O'Connor F, Jerinic G. 2005 Wideband and multi-octave MEMS-reconfigurable RF power amplifiers. In *GOMACTech Digest of Papers, Las Vegas, NV, April 2005*.
23. Herrick KJ, Jerinic G, Molfino RP, Lardizabal SM, Pillans B. 2006 S-Ku band intelligent amplifier microsystem. *Proc. SPIE* **6232**, 62320W. (doi:10.1117/12.669718)
24. Malczewski A, Pillans B, Morris F, Newstrom R. 2011 A family of MEMS tunable filters for advanced RF applications. In *IEEE MTT-S Int. Microwave Symposium Digest* (*MTT 2011*), *Baltimore, MD, 5–10 June*, 4 pp. Piscataway, NJ: IEEE. (doi:10.1109/MWSYM.2011.5972889)
25. Lardizabal S, Molfino R, Wang C, Moody C, Pillans B. 2011 Component and technology development toward an adaptive RF system. In *GOMACTech Digest of Papers, Orlando, FL, 21–24 March*, pp. 491–494.
26. Sollner G *et al.* 2011 Tunable receiver for 6–18 GHz with autonomous self-healing. In *GOMACTech Digest of Papers, Orlando, FL, 21–24 March*.
27. Yelten MB, Zhu T, Koziel S, Franzon PD, Steer MB. 2011 Demystifying surrogate modeling for circuits and systems. *IEEE Circuits Syst. Mag.* **12**(1), 45–63. (doi:10.1109/MCAS.2011.2181095)
28. Weinstein D, Bhave SA. 2010 The resonant body transistor. *Nano Lett.* **10**, 1234–1237. (doi:10.1021/nl9037517)
29. Weinstein D, Bhave SA. 2010 Acoustic resonance in an independent-gate FinFET. In *Solid State Sensor, Actuator and Microsystems Workshop, Technical Digest, Hilton Head Island, SC, 6–10 June*, pp. 459–462. Transducer Research Foundation.
30. Popa LC, Weinstein D. 2013 Switchable piezoelectric transduction in AlGaN/GaN MEMS resonators. In *Transducers and Eurosensors XXVII, at 17th Int. Conf. on Solid-State Sensors, Actuators and Microsystems, Barcelona, Spain, 16–20 June*, pp. 2461–2464. Piscataway, NJ: IEEE. (doi:10.1109/Transducers.2013.6627304)
31. Ansari A, Gokhale V, Thakar VA, Roberts J, Rais-Zadeh M. 2011 Gallium nitride-on-silicon micromechanical overtone resonators and filters. In *IEEE Int. Electron Devices Meeting* (*IEDM*), *Washington, DC, 5–7 December*, pp. 485–488. Piscataway, NJ: IEEE. (doi:10.1109/IEDM.2011.6131590)
32. Müller A *et al.* 2009 6.3-GHz film bulk acoustic resonator structures based on a gallium nitride/silicon thin membrane. *IEEE Electron Device Lett.* **30**, 799–801. (doi:10.1109/LED.2009.2023538)
33. Wong H-SP, Raoux S, Kim SB, Liang J, Reifenberg JP, Rajendran B, Asheghi M, Goodson KE. 2010 Phase change memory. *Proc. IEEE* **98**, 2201–2227. (doi:10.1109/JPROC.2010.2070050)
34. Lo H *et al.* 2010 Three-terminal probe reconfigurable phase-change material switches. *IEEE Trans. Electron Devices* **57**, 313–319. (doi:10.1109/TED.2009.2035533)
35. Sovero EA, Deakin Jr DS, Higgins JA, DeNatale JF, Pittman S. 1990 Fast thin film vanadium dioxide microwave switches. In *Gallium Arsenide Integrated Circuit* (*GaAs IC*) *Symp., Technical Digest, New Orleans, LA, 7–10 October*, pp. 101–103. Piscataway, NJ: IEEE. (doi:10.1109/GAAS.1990.175460)
36. Crunteanu A, Givernaud J, Blondy P, Orlianges J-C, Champeaux C, Catherinot A. 2010 Exploiting the semiconductor–metal phase transition of VO_2 materials: a novel direction towards tuneable devices and systems for RF–microwave applications. In *Advanced microwave and millimeter wave technologies: semiconductor devices, circuits and systems* (ed. M Mukherjee), pp. 35–56. Rijeka, Croatia: InTech. (doi:10.5772/8768)
37. Smith BD, Mayer TS, Keating CD. 2102 Deterministic assembly of functional nanostructures using nonuniform electric fields. *Annu. Rev. Phys. Chem.* **63**, 241–263. (doi:10.1146/annurev-physchem-032210-103346)

38. Malmros A, Blanck H, Rorsman N. 2011 Electrical properties, microstructure, and thermal stability of Ta-based ohmic contacts annealed at low temperature for GaN HEMTs. *Semicond. Sci. Technol.* **26**, 075006. (doi:10.1088/0268-1242/26/7/075006)
39. Lee H-S, Lee DS, Palacios T. 2011 AlGaN/GaN high-electron-mobility transistors fabricated through a Au-Free technology. *IEEE Electron Device Lett.* **32**, 623–625. (doi:10.1109/LED.2011.2114322)
40. De Jaeger B, Van Hove M, Wellekens D, Kang X, Liang H, Mannaert G, Decoutere S. 2012 Au-free CMOS-compatible AlGaN/GaN HEMT processing on 200 mm Si substrates. In *24th Int. Symp. on Power Semiconductor Devices and ICs, Bruges, Belgium, 3–7 June*, pp. 49–52. Piscataway, NJ: IEEE. (doi:10.1109/ISPSD.2012.6229020)
41. Tripathy S *et al.* 2012 AlGaN/GaN two-dimensional-electron gas heterostructures on 200 mm diameter Si(111). *Appl. Phys. Lett.* **101**, 082110. (doi:10.1063/1.4746751)