Cryogenic Control Architecture for Large-Scale Quantum Computing

J. M. Hornibrook,1 J. I. Colless,1 I. D. Conway Lamb,1 S. J. Pauka,1 H. Lu,2 A. C. Gossard,2 J. D. Watson,3,4 G. C. Gardner,3,4 S. Fallahi,5,4 M. J. Manfra,3,4,5,6 and D. J. Reilly1,*

1ARC Centre of Excellence for Engineered Quantum Systems, School of Physics, The University of Sydney, Sydney, New South Wales 2006, Australia
2Materials Department, University of California, Santa Barbara, California 93106, USA
3Department of Physics, Purdue University, West Lafayette, Indiana 47907, USA
4Birck Nanotechnology Center, Purdue University, West Lafayette, Indiana 47907, USA
5School of Materials Engineering, Purdue University, West Lafayette, Indiana 47907, USA
6School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907, USA

(Received 4 October 2014; revised manuscript received 21 December 2014; published 23 February 2015)

Solid-state qubits have recently advanced to the level that enables them, in principle, to be scaled up into fault-tolerant quantum computers. As these physical qubits continue to advance, meeting the challenge of realizing a quantum machine will also require the development of new supporting devices and control architectures with complexity far beyond the systems used in today’s few-qubit experiments. Here, we report a microarchitecture for controlling and reading out qubits during the execution of a quantum algorithm such as an error-correcting code. We demonstrate the basic principles of this architecture using a cryogenic switch matrix implemented via high-electron-mobility transistors and a new kind of semiconductor device based on gate-switchable capacitance. The switch matrix is used to route microwave waveforms to qubits under the control of a field-programmable gate array, also operating at cryogenic temperatures. Taken together, these results suggest a viable approach for controlling large-scale quantum systems using semiconductor technology.

DOI: 10.1103/PhysRevApplied.3.024010

I. INTRODUCTION

Realizing the classical control and readout system of a quantum computer is a formidable scientific and engineering challenge in its own right, likely requiring the invention of a suite of new devices with tailored physical properties. Already under way for this purpose is the development of near-quantum-limited amplifiers [1–4], small-footprint circulators [5], ultralow-loss resonators [6,7], cryogenic filters [8], and interconnect solutions [9–11]. The hardware for classical data conversion and processing, however, has yet to be tightly integrated with the quantum technology. Such a classical control interface must be fast, relative to the time scales of qubit decoherence low noise, so not to disturb the fragile operation of qubits, and scalable with respect to physical resources [12–14]. A particular challenge is ensuring that the footprint for routing signal lines or the operating power does not grow rapidly as the number of qubits increases [15,16]. As solid-state quantum processors will likely operate below 1 K [17–21], there are advantages to also locating components of the control system in a cryogenic environment, adding further constraints.

Similar challenges have long been addressed in the satellite and space exploration community [22], where the need for high-frequency electronic systems operating reliably in extreme environments has driven the development of new circuits and devices [23]. Quantum-computing systems, on the other hand, have to date largely relied on brute-force approaches, controlling a few qubits directly via room-temperature electronics that is hardwired to the quantum device at cryogenic temperatures.

Here we present a control architecture for operating a cryogenic quantum processor autonomously and demonstrate its basic building blocks using a semiconductor qubit. This architecture addresses many aspects relating to scalability of the control interface by embedding multiplexing subsystems at cryogenic temperatures and separating the high-bandwidth analog-control waveforms from the digital addressing needed to select qubits for manipulation. Our demonstration makes use of a semiconductor switch matrix constructed using high-electron-mobility transistors and a new type of microwave switch element based on the gate-tuneable capacitive response of a heterostructure device. Under the control of a commercial field-programmable gate array (FPGA) made to operate at 4 K, the switch matrix is used to route microwave signals to selected quantum-dot qubits at 20 mK. Bringing these subsystems together in the context of our control architecture suggests a path for scale up of control hardware needed to manipulate the large numbers of qubits in a useful quantum machine.

*Corresponding author.
david.reilly@sydney.edu.au
II. CONTROL MICROARCHITECTURE

Our control microarchitecture executes a quantum algorithm by decomposing it into a sequence of universal quantum gates, allowing for arbitrary logic operations to be realized using a small set of repeated single- and two-qubit unitaries applied in sequence. At the level of physical qubits in the solid state, whether they are spins [24], transmons [17,19], or quasiparticles [25], these elemental gate operations amount to applying calibrated electrical waveforms to a particular set of qubits or pairs of qubits each clock cycle as determined by a quantum algorithm.

A key aspect of our control architecture is the separation of these analog “prime waveforms,” which are typically pulses at microwave frequencies, from the digital qubit addressing information that determines which waveform is directed to which qubit, at a particular point in the code. In comparison to brute-force approaches, this scheme lifts the need of having a separate waveform generator and transmission line for each qubit, taking advantage of a small universal gate set that uses the same analog waveforms over and over throughout the algorithm. As realistic qubits will inevitably include variations in their physical parameters, the control architecture must also incorporate means of calibrating and adjusting the response of the qubit to the control waveforms, as described below.

Our prime-line bus (PL) and address-line bus (AL) architecture is shown schematically in Fig. 1, where we draw part of a circuit for implementing a quantum error-correcting surface code [26,27]. Precisely timed analog prime waveforms generated at cryogenic or room temperature propagate cyclicly on a high-bandwidth prime-line bus that is terminated with a matched impedance at a location in the system where heat can be dissipated. The quantum algorithm is then executed exclusively via the digital address-line bus selecting qubits and qubit pairs to receive the appropriate prime waveform at the correct clock cycle in the circuit. Readout proceeds in a similar way, with the digital address bus selecting a particular qubit (or readout device) for interfacing with multiplexing devices [28,29] and analog readout circuitry such as a chain of amplifiers and data converters.

III. IMPLEMENTATION OF THE CONTROL ARCHITECTURE

Realizing our PL-AL architecture requires integrating multicomponent control and readout hardware with the quantum system of qubits fabricated on a chip. Owing to the large number of qubits that are likely to be needed for quantum computation and the time scales involved in their control, there are advantages to locating subsystems of the control architecture at cryogenic temperatures, either on chip with the physical qubits or in close proximity and connected via integrated multichip modules [30] and compact transmission lines. Aspects of the control system will, however, generate significant heat or fail to function at the millikelvin temperatures needed for qubit operation. The competing constraints of interconnect density, heat generation, signal latency, footprint, and noise performance suggest a control architecture that is distributed across a cryostat, taking advantage of the significantly different thermal budgets available at each temperature stage. This distributed architecture is illustrated in Fig. 2, where control subsystems are positioned at different temperature stages of a cryogen-free dilution refrigerator. Below we describe and provide a basic demonstration of these subsystems.

A. Switch matrix

The key subsystem underpinning the control microarchitecture is a switch matrix or routing technology that steers the prime waveforms to particular qubits based on a
digital address. This technology is ideally located in close proximity to the qubits to avoid latency and synchronization challenges that arise when signals propagate over length scales comparable to the electromagnetic wavelength (typically centimeters for quantum-control waveforms). Physically integrating the switch matrix and qubit system has the further advantage of significantly reducing the wiring and interconnect density by making use of silicon lithography (or multichip module packaging) to provide the wiring and interconnect density by making use of silicon lithography (or multichip module packaging) to provide the wiring and interconnect density by making use of silicon lithography (or multichip module packaging) to provide.

FIG. 2. Schematic of a control microarchitecture that distributes subsystems across the various temperature stages of a dilution refrigerator, depending on the available cooling power (image is of a Leiden Cryogenics CF450). A millikelvin switch matrix on the same chip as the qubit device or close to it steers a small number of control pulses to qubits using addressing information from cryogenic logic at 4 K. The matrix will incorporate a level of digital decoding to enable switch addresses to be transmitted on a relatively small number of serial lines. The cryogenic logic also interfaces with multiplexed readout and digital-to-analog converters (DAC). The 4-K stage typically has a cooling power of approximately 1 W, with the 20-mK stage having less than 10 μW.

be used to account for the variation in physical parameters that will inevitably occur with systems comprising large numbers of qubits.

Various technologies appear suitable for constructing such a switch matrix, including semiconducting devices [31–33], mechanical systems [34,35], and superconducting logic [36,37]. For qubit technologies built from semiconductors [21,38], field-effect-based devices are ideally suited owing to their subnanosecond switching speed, gigahertz transmission bandwidth, low dissipation, small footprint, cryogenic compatibility, and opportunity for integration with qubits. Below we demonstrate the operation of such devices using GaAs high-electron-mobility transistor (HEMT) circuits configured as a switch matrix with variable amplitude and phase response. We note that complex circuits constructed from HEMTs demonstrate that these devices are well suited to extensive fan-out [39].

B. HEMT switching elements

A prototype HEMT-style microwave switch based on a GaAs/Al0.3Ga0.7As heterostructure is shown in Figs. 3(a) and 3(b). Fabrication of these switching elements follows a similar procedure to quantum-dot qubit devices (allowing easy integration). The mesa is wet etched using sulphuric acid before Au/Ge/Ni Ohmic contacts are thermally...
evaporated and annealed at 470 °C for 100 sec. The final metal layer is thermally evaporated TiAu (10 nm/100 nm). In the on state, the switch is configured to have a characteristic impedance of approximately 50 Ω owing to its coplanar waveguide (CPW) geometry. Prime waveforms are fed to and from the HEMT two-dimensional electron gas (2DEG) via eutectic Ohmic contacts and TiAu planar transmission lines. In the off state, a negative voltage applied to the TiAu top gate pinches off the electron gas channel, reflecting the prime waveform signal due to the large impedance of the HEMT relative to the characteristic impedance of approximately 50-Ω feed line. The transmission response of the switch is shown in Fig. 3(c), with an on-off ratio (OOR) above 40 dB in the frequency range 0–2.5 GHz, suitable for control of spin qubits [40]. For these prototype devices, a large insertion loss of 10–20 dB is observed, owing mostly to the resistance of the Ohmic contacts, which is currently a few hundred Ω in our process. Precise control of the contact resistance and capacitance using ion implantation can overcome this limitation and also dramatically shrink the footprint of these devices [41,42].

The time-domain response of the switch is demonstrated by amplitude modulating an applied 120-MHz constant wave tone, as shown in Fig. 3(d). To determine the maximum switching time of the HEMT, we modulate a 5-GHz carrier tone with a sinusoidal waveform applied to the gate and measure the depth of modulation as a function of gate frequency, as indicated in Fig. 3(e). For these prototype devices, the switching time is of order 1 ns.

FIG. 4. A switch design that produces an impedance mismatch by depleting the transmission-line ground plane. Shown is an image (a) and cross sections (b) of the device. The input is a coplanar line [i] which transitions to a microstrip using the 2DEG as the ground plane [ii,iii]. The ground plane is tapered so that it lies beneath the signal track [iv, v]. When a negative voltage is applied to the signal track, the ground plane is depleted [v], and the impedance mismatch reflects the input signal. (c) Transmission through the switch in the on (blue) and off (red) states. (d) Frequency response of capacitors formed using surface gates and 2DEG as a parallel plate (inset).

C. Capacitive switching elements

Microwave switching devices based on the depletion of an electron gas also enable a new capacitive mode of operation. In this configuration, the CPW feed line transitions to a microstrip geometry by contacting the electron gas to the planar ground planes using Ohmic contacts, as illustrated in Figs. 4(a) and 4(b). The two conductors in the microstrip transmission line are, thus, constructed using the top gate and electron gas as ground. This device can act as a reflective switch by depleting the effective ground plane using a negative bias on the gate. Depletion reduces the capacitance between the conductors of the microstrip and modulates the device impedance. Transmission through the switch is shown in Fig. 4(c) in the on (blue) and off (red) state, with an OOR greater than 25 dB for 0–8 GHz.

The switch is capacitively coupled to the input and output ports, with a planar spiral inductor at one port forming a bias tee to provide the dc gate voltage needed to deplete the electron gas. In place of a planar interdigitated capacitor, we make use of the GaAs heterostructure to provide a low-footprint parallel-plate capacitor formed between the CPW central track and the electron gas, as shown in the inset of Fig. 4(d). The frequency response of this capacitor is shown in Fig. 4(d).

The capacitance-based switch has improved performance at higher frequencies than the HEMT-based switch, although it has a larger footprint due to both the length of line needed for adiabatic tapering from 50 to 200 Ω and for the coplanar-to-microstrip transition. Working with a characteristic impedance of approximately 200 Ω minimizes the area of electron gas and reduces Ohmic loss. The improved frequency performance stems from the absence of a gate structure, which in the HEMT switch capacitively couples the source and drain contacts, even in the off state. The required footprint is reduced significantly in an all-microstrip circuit that is designed to operate at a characteristic impedance close to 200 Ω. In their current form, the performance of both kinds of switches is better suited to controlling spin qubits, where the frequency of signals is of order 1 GHz. For superconducting qubits, we envisage extending the operation of these switches to frequencies in the 4–12 GHz range by shrinking their footprint to suppress parasitic capacitances and inductances that lead to resonances in the present design.

D. 2:2 switch matrix

We demonstrate cryogenic operation of a prototype routing matrix based on HEMT switches with two input and two output ports. A magnified image of the device is shown in Fig. 5(a) with an associated schematic in Fig. 5(b). Each input port is split and connected to each output port via a switch so that the transmission parameters $S_{ij}$ of the device are controlled by the respective gate voltages $V_{i,j}$. The output ports include bias tees, which are needed for use with qubits based on semiconductor quantum dots.
We also demonstrate basic quadrature (IQ) modulation using our switch matrix by applying rf tones at both inputs with a 90° phase offset between them. The 90° shift can be produced by a length of transmission line (with narrowband response) or as a separate quadrature prime waveform. The output waveform at angular frequency $\omega$ is $A \sin(\omega t) + B \cos(\omega t) = R \sin(\omega t + \phi)$, where the magnitude $R$ and phase $\phi$ are determined by the amplitudes $A$ and $B$ controlled by the gate voltages $V_{i,j}$. After the calibration function $R, \phi = F(V_{i,j})$ is generated once, we can select the appropriate $V_{i,j}$ to produce a tone with arbitrary phase and amplitude in the first quadrant of the complex plane, as shown in Fig. 5(e). The corresponding voltage output along a quarter circle of constant amplitude is shown in Fig. 5(f).

By controlling the amplitude and phase shift using the integrated switch matrix, the connection between each qubit and the prime-line bus can be specifically adjusted to compensate for the inevitable variation in parameters between physical qubits [43].

### E. Cryogenic logic

For controlling and programming the switch matrix via the address bus, we envisage a layer of fast classical logic that serves as an interface between the physical qubits and the compiled quantum algorithm (which will likely comprise mostly an error-correcting code). This layer of classical logic is also needed for executing various automatic sequences associated with fast feedback for qubit stabilization, readout signal conditioning, or open-loop error suppression [44,45]. For controlling a large-scale quantum computer, there are many advantages to locating this classical logic and associated data converters close to the qubits inside a dilution refrigerator. In comparison to room-temperature-based control systems, cryogenic operation results in an enhanced clock speed, improved noise performance, reduced signal latency and timing errors, and larger bandwidth. Some of these aspects stem from the ability to make use of compact superconducting transmission lines and interconnects at cryogenic temperatures.

We note that locating control electronics inside the vacuum space of the refrigerator allows it to be positioned physically close to the qubit device, even if the qubits and control systems are at moderately different temperatures.

The choice of technology for constructing this layer of classical control is largely dictated by the qubit coherence times, control-signal bandwidth, and the number of simultaneous qubits under control. With a convergence of solid-state qubit coherence times now approaching 1 ms [46–48], present-day CMOS-based FPGAs or application-specific integrated circuits (ASICs) operating at 4 K are a viable control platform. Higher-performance control systems that are likely to be realized in the longer term include technologies based on InP devices [49], SiGe BiCMOS [50,51], and superconducting flux logic [36,37].
For the basic demonstration of the PL-AL scheme considered here, the classical logic is implemented using a commercial FPGA (Xilinx Spartan-3A) that we make operational at the 4-K stage of a dilution refrigerator. To achieve cryogenic operation, the FPGA chip is mounted on a custom cryogenic printed circuit board that includes components which vary little in their parameters at cryogenic temperatures [10,11]. Power and clock signals to the FPGA are adjusted for cryogenic operation using room-temperature sources, and a semirigid coax line is configured for sending serial commands, with reprogramming of the low-temperature array occurring via a dedicated ribbon cable. With the FPGA mounted at the 4-K stage, we measure an idle power dissipation of approximately 30 mW, with negligible increase during dynamic logic operations for the simple code executed here. We estimate a dynamic power dissipation of approximately 100 mW for computational operations that use most of the gates in the Spartan-3 array (further details of cryogenic operation of FPGAs are given elsewhere [52]). The FPGA is programmed to interpret serial communication and output a 3.3-V signal on selected pins to activate prime waveform routing in the switch matrix. These outputs are combined with a negative voltage provided from room temperature via a cold resistive adder so that the switch matrix gates receive $-50$ mV for the on-state and $-380$ mV for the off-state voltage.

IV. SEMICONDUCTOR QUBIT CONTROL

We combine the building blocks of our microarchitecture described above to demonstrate that a semiconductor qubit can feasibly be controlled autonomously without introducing additional noise or heating to the quantum system. The qubit is a GaAs double quantum dot configured as a charge or spin qubit in the few-electron regime (the heterostructure has a carrier density $2.4 \times 10^{-15}$ m$^{-2}$ and mobility 44 m$^2$/V s at 20 K). These qubits are commonly controlled using dc-pulse waveforms on the gates to rapidly manipulate the energy levels of the quantum dots [53]. A typical setup connects a waveform generator to each gate using a separate high-bandwidth coaxial cable and bias tee.

For this demonstration, we connect a single coaxial cable from a waveform generator at room temperature to the input of the 2:2 switch matrix, with the two output ports connected to the two plunger gates LP and RP of the double dot, as shown schematically in Fig. 6(a). The waveform generator produces a prime waveform consisting of a 100-kHz square wave [shown in Fig. 6(b)], which is then steered by the 4-K FPGA by opening and closing switches in the matrix depending on commands sent from room temperature.

The charge state of the double dot is sensed using a rf quantum point contact [29,54], which provides a readout signal $V_{rf}$ as a function of the gate voltages $V_L$ and $V_R$ indicated in Fig. 6(c). With both switches of the matrix set to the off state, a standard charge-stability diagram is detected indicating that the off state provides sufficiently high isolation between input and output ports, as shown in Fig. 6(d) [55]. In contrast to using the qubit decoherence time to detect additional noise sources from the control circuits, we note that the width and jitter of a quantum-dot charge transition provides a broadband probe of electrical noise, including fluctuations that occur on time scales much longer than the qubit coherence.

FIG. 6. Simple implementation of the microarchitecture introduced in Fig. 2. (a) Experimental setup for measuring a double quantum dot using a cryogenic FPGA to steer pulses via a millikelvin switch matrix. Charge-state readout is performed using a rf quantum point contact (QPC). (b) Switch matrix output showing a 100-kHz square wave directed to plunger gates of the quantum dot. (c) Micrograph of the quantum-dot device. The shaded gates labeled LP and RP are connected to the switch matrix output. (d)–(g) Charge sensing of the double quantum dot in the few-electron regime, with electron occupancy indicated by the labels (m,n). The color axis is the derivative of the sensing signal $V_{rf}$ with respect to $V_R$. When the FPGA-controlled switch matrix blocks waveforms (d), a standard double-dot stability diagram is detected. When the square wave is directed to either LP (e), RP (f), or both (g), copies of the stability diagram appear (see text). These measurements demonstrate that the double-dot potential can be controlled autonomously by the switch matrix and cold FPGA.
Sending a command to the cold FPGA allows the prime waveform to be directed to the left, right, or both plunger gates, producing two copies of the charge-stability diagram. These copies appear because, on the time scale of the readout, a square wave with 50% duty cycle configures the double dot in two distinct charge states that are offset from one another by the voltage $\Delta V_R$ or $\Delta V_L$, as shown in Figs. 6(e)–6(g). We note that the shift measured in Fig. 6(g) is the vector sum of the shifts in Figs. 6(e) and 6(f) and accounts for the cross capacitance between each gate and each dot [56]. In comparison to data taken on the bare quantum dot, we are unable to detect any additional noise or an increase in the electron temperature (which is of order 100 mK) when configuring the charge state using the cryogenic FPGA and switch matrix.

V. DISCUSSION

Our simple demonstration of a multicomponent control architecture provides a path for scaling up the classical support system needed for operating a large-scale quantum computer. Aspects of this demonstration will also likely find immediate use in improving the performance of few-qubit experiments using electron spins in quantum dots. For example, in using the switch matrix to produce multiple out-of-phase copies of control waveforms, cross-talk can be suppressed by canceling the voltage that is capacitively coupled to neighboring surface gates [57]. Using the switch matrix as a high-frequency cryogenic multiplexer will also enable the automated testing and characterization of many devices in a given cooldown experiment. In the longer term, our microarchitecture can be extended to allow additional functionality of the switch matrix, providing qubit control frequency correction by using the HEMTs as mixers or as cryogenic adder circuits that reduce the complexity or resolution needed for biasing surface gates that define quantum dots.

We comment here also on the possibility of implementing our PL-AL control architecture using a switch matrix based on single-flux-quanta superconducting logic [37]. Such logic already appears well suited to control-flux-based qubits at high speed and with low dissipation. To what extent these devices are compatible with magnetic fields and the need to generate and steer large-voltage waveforms required in the operation of semiconductor qubits remains an open question.

At the layer of classical logic, our demonstration shows that commercial FPGA devices can be configured to work at cryogenic temperatures and are compatible with controlling qubits in close proximity. Beyond the control architecture presented here, the use of cold low-latency classical logic will likely improve the performance of feedback systems generally needed for adaptive measurement, quantum-state distillation, and error-correction protocols. A further consideration is the heat generated by the switch matrix, which must operate at the mixing chamber stage of a dilution refrigerator. Given that these switches are reflective rather than dissipative at microwave frequencies, heat generation will be dominated by the charging of the gate capacitance with each switch, as is the case for today's room-temperature CMOS technology. For 1000 HEMT switches of the kind shown here operating at a clock frequency of 1 GHz, we estimate a total power dissipation of 100 $\mu$W. Straightforward improvements in switch design, such as a reduction in subthreshold voltage swing at low temperature, can likely reduce dissipation by a factor of 100. Even so, improvements in cryogenic refrigeration technology, both at the chip level [58] and cryostat, similar to what has been achieved in rare-event physics [59], will likely be needed to enable large-scale physics [59], will likely be needed to enable large-scale physics.

VI. CONCLUSION

We propose a microarchitecture for the control of a large-scale quantum processor at cryogenic temperatures. The separation of analog-control prime waveforms from the digital addressing needed to select qubits offers a means of scaling this approach to the numbers of qubits needed for a computation. To demonstrate the feasibility of our scheme, we show that a semiconductor qubit can be controlled using a cryogenic FPGA system and custom switch matrix for steering analog waveforms at low temperature. We anticipate that integrated autonomous control systems of this kind will be increasingly important in the development and demonstration of fault-tolerant quantum machines.

ACKNOWLEDGMENTS

We thank B. Smith, D. Tuckerman, D. Wecker, K. Svore, C. M. Marcus, L. DiCarlo, L. P. Kouwenhoven, and M. Freedman for useful conversations. Devices are fabricated at the New South Wales Node of the Australian National Fabrication Facility (ANFF). This research is supported by the Office of the Director of National Intelligence, Intelligence Advanced Research Projects Activity (IARPA), through the Army Research Office Grant No. W911NF-12-1-0354, the Australian Research Council Centre of Excellence Scheme (Grant No. EQuS CE110001013), and Microsoft Research.


[43] Calibration of the switch response and qubit can be performed at the same time, measuring the qubit evolution as a function of switch gate voltage.


[52] I. D. Conway Lamb and D. J. Reilly, Field programmable gate arrays (FPGA) in cryogenic environments (to be published).


[55] A very small amount of jitter in the charge transitions can be seen due to coupling of the rf QPC carrier to the gates.


