A 6-to-8GHz 0.17mW/Qubit Cryo-CMOS Receiver for Multiple Spin Qubit Readout in 40nm CMOS Technology

Prabowo, Bagas; Zheng, Guoji; Mehrpoo, Mohammadreza; Patra, Bishnu; Harvey-Collard, Patrick; Dijkema, Jurgen; Sammak, Amir; Scappucci, Giordano; Charbon, Edoardo; Sebastiano, Fabio

DOI
10.1109/ISSCC42613.2021.9365848

Publication date
2021

Document Version
Final published version

Published in
2021 IEEE International Solid-State Circuits Conference, ISSCC 2021 - Digest of Technical Papers

13.3 A 6-to-8GHz 0.17mW/Qubit Cryo-CMOS Receiver for Multiple Spin Qubit Readout in 40nm CMOS Technology

Bagas Prabowo[1,2], Guoji Zheng[1,2], Mohammadreza Mehrpoo[1,3], Bishnu Patra[1,2], Patrick Harvey-Collard[1,2], Jurgen Dijkema[1,2], Amir Sammak[4], Giordano Scappucci[1,2], Edoardo Charbon[1,2,5], Fabio Sebastiano[1,2], Lieven M. K. Vandersypen[1,2], Masoud Babaie[1,2]

[1]Delft University of Technology, Delft, The Netherlands
[2]QuTech, Delft, The Netherlands
[3]now with Broadcom Netherlands, Bunnik, The Netherlands
[4]TNO, Delft, The Netherlands
[5]EPFL, Neuchatel, Switzerland

2021 IEEE International Solid- State Circuits Conference (ISSCC) | 978-1-7281-9549-0/20/$31.00 ©2021 IEEE | DOI: 10.1109/ISSCC42613.2021.9365848

Quantum computers (QC) promise to solve certain computational problems exponentially faster than a classical computer due to the superposition and entanglement properties of quantum bits (qubits). Among several qubit technologies, spin qubits are a promising candidate for large-scale QC, since (1) they have a small footprint allowing them to be densely integrated and (2) they can operate at relatively high temperatures (>1K) [1], potentially reducing system cost and complexity.

Executing practical quantum algorithms with real-time quantum error correction will require thousands of qubits to be controlled and read out simultaneously faster than the qubits' decoherence time (<<T_2^* ~120µs). The scalability of the control/readout electronics to support thousands of qubits will be limited by following the current approach due to the use of thousands of cables between the qubits and the bulky instruments operating at room temperature (RT). To address this issue, the electronics interface should be placed in close proximity or ideally co-integrated with the qubits at a close temperature [1]. Previous works have shown qubit control [2,3] and DC readout of a double quantum dot (DQD) [4] with ICs operating at 3K and 110mK, respectively. This work presents the first cryogenic CMOS (Cryo-CMOS) IC for gate-based RF readout of spin qubits.

Figure 13.3.1 shows a DQD with two single-electron spin qubits in which the left qubit (green) is used as data qubit (LQ) and the right qubit (purple) is used for readout (RQ). The SEM image of the DQD realized in a ^{28}Si/SiGe heterostructure and its equivalent schematic are shown at the bottom right of Fig. 13.3.1. By controlling the electrochemical potential difference between the two dots, i.e. the detuning (ε), using left and right plunger gate voltages (V_{LP}, V_{RP}), the qubits can be decoupled at ε<0 or confined in the right dot at ε>0. In the former regime, LQ is manipulated by applying a microwave pulse with a frequency corresponding to its Zeeman energy splitting (E_z) [5], as shown in the energy level diagram. To read the LQ's state, RQ is initialized to state $|0\rangle$ (spin down) and ε is set to 0. An opposite spin of the LQ relative to the RQ allows it to tunnel between the two different regimes of the DQD. These regimes are shown in the charge stability diagram in the red and blue regions representing one electron in each dot (1,1) and two electrons on the right dot (0,2), respectively. Tunneling alters the total capacitance of the DQD, which can be detected by probing the shift of the resonance frequency of a high-impedance (~1KΩ) NbTiN nanowire resonator connected to the readout gate of the DQD. The high-impedance NbTiN nanowire enables a higher sensitivity in detecting the DQD's capacitance variation.

Figure 13.3.2 depicts the proposed single-shot multi-qubit readout architecture. Each qubit is connected to an individual resonator with a unique resonant frequency ($\omega_{r,i}$) that is also capacitively coupled to a common transmission line (TL), allowing simultaneous readout of all qubits. During the readout operation, a short (<50µs) weak (~-110dBm) RF pulse composed of tones at the qubits' resonant frequencies is applied to the TL to observe the state-dependent amplitude and phase response of individual resonators. In order not to degrade the inherent SNR of the qubit signals, a superconducting parametric amplifier with a near *quantum-limit* noise temperature (Tn≈0.4K) or a HEMT LNA (Tn≈2.2K, BW=4-to-8GHz) must be placed as the first stage to amplify the signal >20dB. The rest of the readout chain, including RF amplifiers, I/Q downconversion mixers, and baseband amplifiers/filters, is implemented in cryo-CMOS chip operating at 4K, replacing the RT lab instrument with noise figure (NF)=1.9dB, Gain=37dB, as used in [6]. After the digitization by an off-the-shelf ADC, a digital processor selects the corresponding signal from each qubit and determines its state based on the observed amplitude/phase information.

The transistor 1/f noise corner is expected to be much higher at cryogenic temperatures. Thus, as depicted in Fig. 13.3.3, a high-IF receiver (RX) is designed to maintain the qubits' SNR over the entire bandwidth (BW). However, a ~2× higher BW baseband circuitry is required, slightly increasing the system's power consumption (P_{DC}). Moreover, the noise from the image band folds on top of the desired band, degrading the NF. Hence, three doubly-tuned transformers are employed to provide in-band (6-to-8GHz) input and interstage matching while effectively attenuating the image band (3.8-to-5.8GHz) before the frequency downconversion. The transistor channel noise does not scale linearly with temperature [5]. Consequently, a cascoded common-source LNA with inductive degeneration (L_s) is employed due to its operation close to the minimum NF along with its high gain and low P_{DC}. A large capacitor (C_S) is placed between L_s and the secondary winding of the transformer to provide a short impedance path for RF current, thus suppressing the effect of the parasitic inductance of the ground bondwires on the input matching. The second stage provides a single-ended to differential conversion and amplification before downconversion by a quadrature double-balanced active mixer driven by on-chip I/Q LO drivers. Current bleeding and shunt peaking techniques are also used to reduce the 1/f noise contribution of the switching quad and enhance the conversion gain-bandwidth product of the mixer, respectively. Finally, two IF amplifiers are cascaded to provide the necessary in-band voltage gain and out-of-band filtering while driving 50Ω cables.

Fabricated in 40nm CMOS technology, the RX chip with a core area of 0.68mm^2 consumes 66/70mW at 4K/RT. Figure 13.3.4 shows the RX electrical performance measured in a cryogenic probe station both at RT and 4K. Benefitting from the higher transistor mobility and better quality factor (Q) of passives at 4K, a ~5dB higher gain and a 25% increase in BW results in 58dB gain and 2GHz BW, respectively. A double sideband (DSB) NF of ~0.6dB shows only a 4× reduction compared to RT, possibly limited by temperature-independent shot noise of active devices. The in-band IIP3 and IP1dB are -50.8/-55dBm and -58.4/-44.9dBm at 4K/RT, respectively. A 200MHz, -70dBm, 16-QAM input signal results in an EVM of -19dB, demonstrating the overall RX performance at 4K.

Figure 13.3.7 shows the dilution refrigerator setup for qubit readout with the RX chip and the qubit sample, which are mounted at the 4k plate and 10mK mixing chamber, respectively. For comparison, the signals from the DQD are read out using the RT rackmount setup, RX chip operating at RT and RX chip installed at 4K plate of the dilution refrigerator. The corresponding DQD charge stability diagrams are measured by monitoring the RX output power (S_{21}), as shown in Fig. 13.3.5. It can be observed that the three configurations show similar performance, proving that the qubits can be readout by the RX chip fully enclosed in a dilution refrigerator. The slight shift in (V_{LP}, V_{RP}) is due to drift in DQD over time. With the RX chip at 4K, an RF tone (6.9GHz) located at the DQD resonant frequency is applied while sweeping V_{LP} and V_{RP}. The measured S_{21} around the interdot transition (red dashed line) indicates ~9dB dip at the interdot crossing (at V_{RP}≈-210mV) due to the frequency shift of the resonator. In this system, SNR = $(A/B)^2$, where A is the dip amplitude and B is the RMS noise amplitude measured at the Coulomb blockade region (V_{RP}≈–230mV). The achieved SNR versus integration time using the RX chip operating at 4K is fitted linearly and extrapolated to extract t_{min} of 1.96µs where SNR=1. Due to time constraints, the results obtained using the RX chip inside the dilution refrigerator at 4K is provisional and sub-optimal.

Compared to state-of-the-art RT rackmount setup (Fig. 13.3.6), the cryo-CMOS RX achieves similar performance, while significantly improving the form factor and P_{DC}. The use of high Q (~2500) resonators allows a 5MHz frequency spacing between the FDMA qubits. Thus the RX chip can simultaneously read out ~400 qubits, resulting in a P_{DC} of 170µW/qubit. Compared to the prior-art cryogenic DQD DC readout IC [4], our chip supports FDMA for scalability, 500× faster readout essential for error-correction, with 3× lower energy consumption per readout cycle for each qubit. This is a critical step towards the integration of cryogenic readout systems, facilitating the development of scalable quantum computers.

Acknowledgement:
The authors would like to acknowledge Intel Corp. for funding the project and Z. Y. Chang, O. Benningshof, N. Alberts from TU Delft for measurement support.

References:
[1] L. M. K. Vandersypen et al., "Interfacing Spin Qubits in Quantum Dots and Donors—Hot, Dense, and Coherent," *npj Quantum Information*, vol. 3, no. 34, 2017.
[2] J. C. Bardin et al., "A 28nm Bulk-CMOS 4-to-8GHz 2mW Cryogenic Pulse Modulator for Scalable Quantum Computing," *ISSCC*, pp. 456-458, Feb. 2019.
[3] B. Patra et al., "A Scalable Cryo-CMOS 2-to-20GHz Digitally Intensive Controller for 4×32 Frequency Multiplexed Spin Qubits/Transmons in 22nm FinFET Technology for Quantum Computers," *ISSCC*, pp. 304-306, Feb. 2020.
[4] L. Le Guevel et al., "A 110mK 295µW 28nm FDSOI CMOS Quantum Integrated Circuit with a 2.8GHz Excitation and nA Current Sensing of an On-Chip Double Quantum Dot," *ISSCC*, pp. 306-308, Feb. 2020.
[5] A. H. Coskun and J. C. Bardin, "Cryogenic Small-Signal and Noise Performance of 32nm SOI CMOS," IEEE MTT-S International Microwave Symposium (IMS2014), pp. 1-4, June 2014.
[6] G. Zheng et al., "Rapid Gate-Based Spin Read-Out in Silicon Using an On-Chip Resonator", *Nature Nanotechnology*, vol. 14, no. 8, pp. 742-746, 2019.

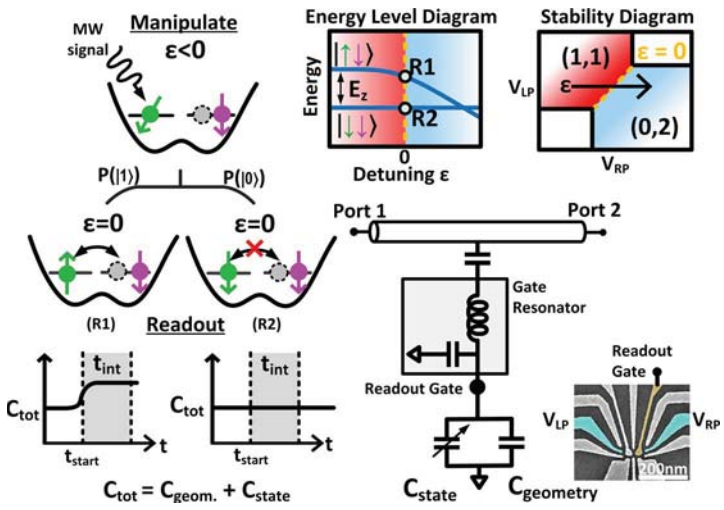

Figure 13.3.1: Schematic representation of the operation of a DQD and the equivalent capacitance in different regimes. Energy level diagram and charge stability diagram highlighting the two regimes.

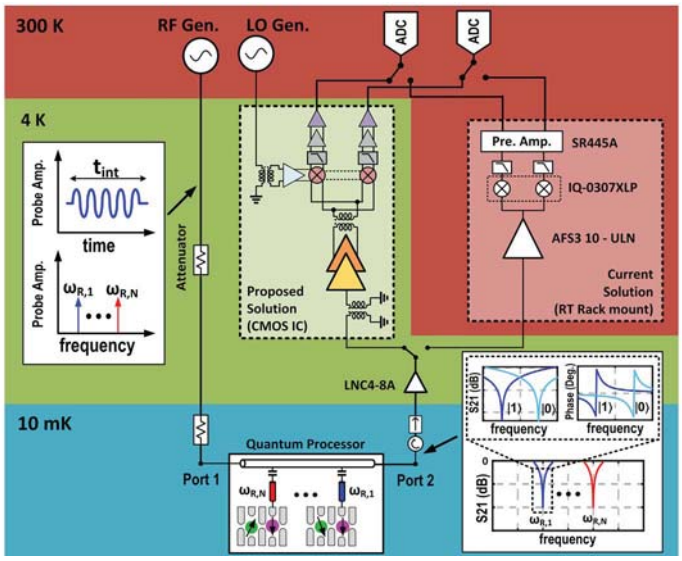

Figure 13.3.2: System level schematic of the proposed frequency multiplexed spin qubit readout chain. A room temperature readout path is added for comparison.

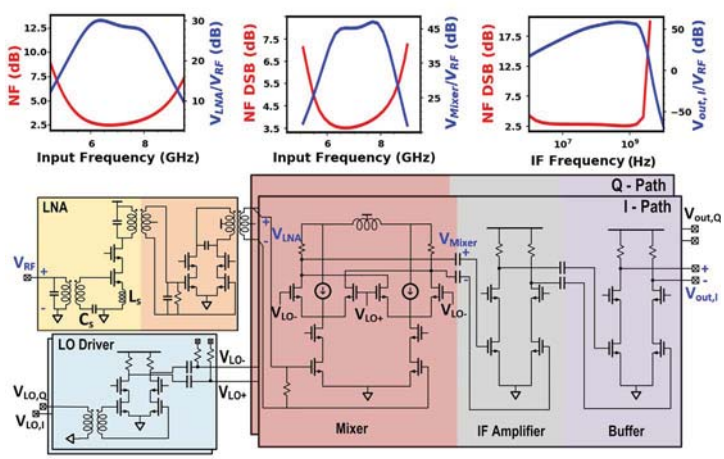

Figure 13.3.3: Transistor level schematic of the cryo-CMOS receiver with the simulated transfer function at various stages.

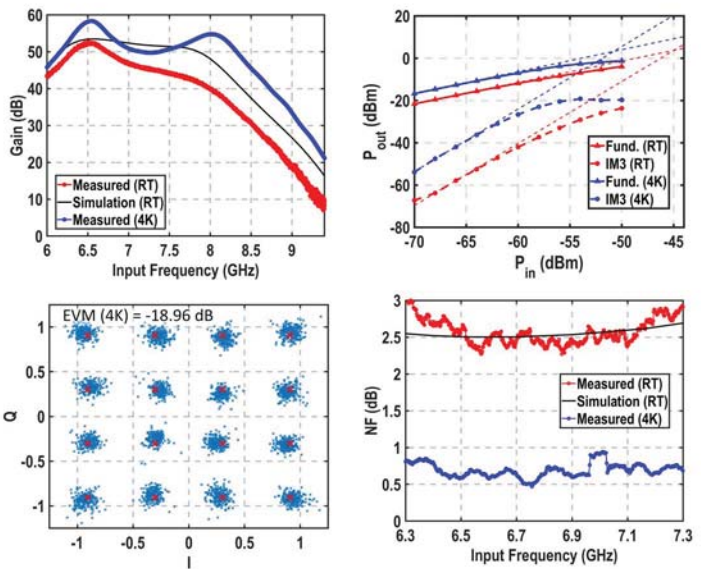

Figure 13.3.4: Measured gain, linearity, EVM, and NF of the receiver at RT and 4K.

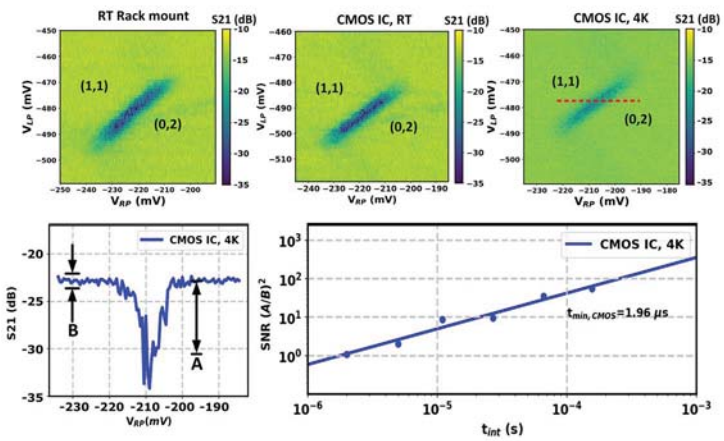

Figure 13.3.5: The measured DQD charge stability diagrams using three different setups (top). Red dashed line indicates the V_{RP} sweep across the interdot crossing. The measured S_{21} and integrated SNR using the cryo-CMOS RX operating at 4K plate (bottom).

	This work		RT Rack mount
Operating Temp.	300 K	4.2 K	300 K
Technology	Bulk CMOS 40nm		-
Operating Freq.	6 - 8 GHz		4 - 8 GHz
Gain	52 dB	58 dB	36.5 dB
NF (DSB)	2.5 dB	0.6 dB	1.9 dB
IIP3	-44.9 dBm (Δ=50 MHz)	-50.8 dBm (Δ=50 MHz)	-9.2 dBm*
IP1dB	-55 dBm	-58.4 dBm	-19.2 dBm*
Power Dissipation	70 mW	66 mW	3 W#
Area	0.68 mm²		Rack mount

	This work	L. Le Guevel, ISSCC'20 [4]
Readout scheme	Gate-based RF readout	DC readout
FDMA	Yes	No
Components	LNA, Mixer, baseband amplifier	TIA
Operating temperature	4.2 K	110 mK
Pdc/qubit	170 μW	1 μW
Readout BW	2 MHz	1.1 kHz
Min. readout time (t_{min})	1.96 μs	1 ms
Energy/Qubit↓	333 pJ	1 nJ

↓ Energy/qubit = (P_{DC} / qubit) × t_{min}
* Linearity performance only considering the commercial LNA and Mixer from the datasheet
# Power excluding the SR445 Preamplifier

Figure 13.3.6: Comparison with RT rackmount setup for gate-based RF readout scheme (left). Comparison with prior-art cryogenic DQD DC readout (right).

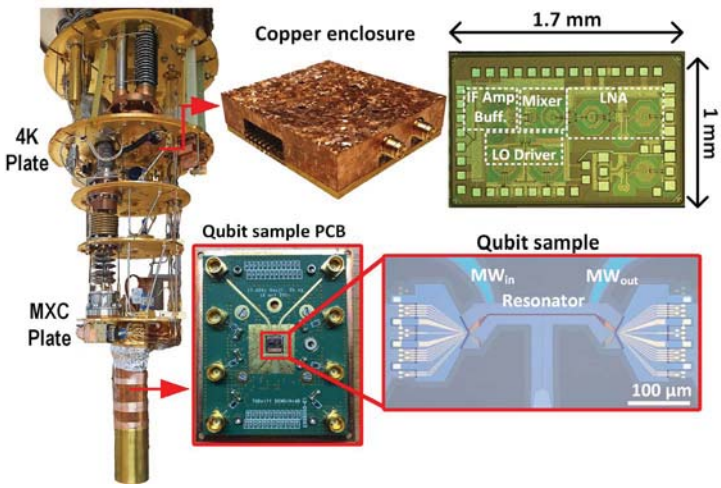

Figure 13.3.7: Top to bottom, left to right: dilution refrigerator setup, gold-plated copper enclosure, receiver chip micrograph, qubit sample PCB and qubit sample micrograph.

978-1-7281-9549-0/21/$31.00 ©2021 IEEE