Cryo-CMOS Circuits and Systems for Quantum Computing Applications

Bishnu Patra, Student Member, IEEE, Rosario M. Incandela, Student Member, IEEE, Jeroen P. G. van Dijk, Harald A. R. Homulle, Lin Song, Mina Shahmohammadi, Member, IEEE, Robert Bogdan Staszewski, Fellow, IEEE, Andrei Vladimirescu, Fellow, IEEE, Masoud Babaie, Member, IEEE, Fabio Sebastiano, Senior Member, IEEE, and Edoardo Charbon, Fellow, IEEE

Abstract—A fault-tolerant quantum computer with millions of quantum bits (qubits) requires massive yet very precise control electronics for the manipulation and readout of individual qubits. CMOS operating at cryogenic temperatures down to 4 K (cryo-CMOS) allows for closer system integration, thus promising a scalable solution to enable future quantum computers. In this paper, a cryogenic control system is proposed, along with the required specifications for the interface of the classical electronics with the quantum processor. To prove the advantages of such a system, the functionality of key circuit blocks is experimentally demonstrated. The characteristic properties of cryo-CMOS are exploited to design a noise-canceling low-noise amplifier for spin-qubit RF-reflectometry readout and a class-F2.3 digitally controlled oscillator required to manipulate the state of qubits.

Index Terms—Class-F oscillator, CMOS characterization, cryo-CMOS, low-noise amplifier (LNA), noise canceling, phase noise (PN), quantum bit (qubit), quantum computing, qubit control, single-photon avalanche diode (SPAD).

I. INTRODUCTION

QUANTUM computing is a new paradigm that exploits basic principles of quantum mechanics, such as entanglement and superposition [1], potentially enabling unprecedented speedups in solving intractable problems [2]. The new computing opportunities include prime factorization, quantum simulations for synthesis of drugs and materials, and complex optimizations [3].

In its fundamental embodiment, a quantum computer comprises a quantum processor and a classical electronic controller. The quantum processor consists of a set of quantum bits (qubits) operating at extremely low temperatures, typically a few tens of mK, while the classical electronic controller is used to read out and control the quantum processor, as shown in Fig. 1. Although the classical controller is implemented today with room-temperature laboratory instruments, this approach becomes increasingly challenging and less cost-effective as the number of qubits grows toward the thousands and millions, as required by practical quantum algorithms [2].

This paper proposes the monolithic integration of the readout and control circuitry in a standard CMOS technology operating at cryogenic temperatures (cryo-CMOS). Although other specialized electronic technologies can handle cryogenic temperatures, only CMOS can work down to at least 30 mK while providing complex system-on-chip integration capable of handling thousands or millions of qubits [4]–[8]. A drastic reduction of the complex interconnections between the cryogenic chamber and the room-temperature electronics will result in enhanced compactness and reliability, thus paving the way to the creation of practical quantum computers. Moreover, the cryo-CMOS circuits and systems could prove useful in other domains, for example, in applications that require cryogenic environments as an integral part of their operation, such as space and high-energy-physics experiments, or whenever extremely low noise is essential, such as in metrology, imaging, and instrumentation.

Cryogenic CMOS circuits have been proposed before for applications ranging from space missions to low-noise amplifiers (LNAs) [9]–[11]. However, quantum processors require extremely high performance from the classical electronic controller in terms of bandwidth and noise, so as to ensure accuracy and speed in the control and readout of the qubits. In this paper, we demonstrate two critical sub-systems of the electronic controller: an LNA and an RF oscillator, which are optimized for cryogenic operation [12].

An overview of the classical electronic controller and its specifications is presented in Section II, while the main characteristics of the cryo-CMOS devices are described in
Section III. The design and experimental validation of the LNA and the RF oscillator are presented in Sections IV and V, respectively. Conclusions are drawn in Section VI.

II. ARCHITECTURE OF THE CLASSICAL CONTROLLER

Qubit control and readout require generation and acquisition of specific electronic and/or optical signals. Typically, for the manipulation of a single qubit, microwave bursts of a short duration (<1 μs) must be applied, e.g., 4-8-GHz bursts for transmons [13], [14] and >12-GHz bursts for spin qubits [15]–[18]. Two-qubit operations typically require current (for transmons [19]) or voltage pulses (for spin qubits [17]) with a bandwidth of tens of MHz. The readout is often performed fully electrically, e.g., by measuring the resonance frequency of a microwave resonator in the case of transmons [20], or by measuring the impedance of a charge detector, such as a quantum point contact, in the case of spin qubits [21].

In order not to disturb the quantum state of the qubit, the controller must generate accurate and extremely low-noise signals. However, even if an ideal control signal is applied, the state of the qubit is destined to be lost in a relatively short time, usually characterized by the dephasing time $T_\phi^*$ [22], which is much shorter than the duration of practical quantum algorithms. Therefore, the classical controller is also responsible for implementing qubit corrections, such as echo techniques [16], [22]–[24] and quantum error correction (QEC) schemes [25], designed to maintain the quantum state over longer periods. QEC algorithms achieve fault tolerance by redundancy, i.e., by encoding the state of a single logical qubit on many physical qubits. Consequently, even for non-trivial problems that require only 100 logical qubits, such as quantum chemistry algorithms [26], millions of physical qubits could be required, which must all be served by the classical controller.

Since contemporary quantum computers operate at deep cryogenic temperatures, well below 1 K [13]–[19], individually interconnecting each of potentially millions of qubits to a room-temperature controller becomes infeasible due to the sheer interconnect complexity, cost, and poor system reliability. To address those issues, we propose a cryo-CMOS controller [5]–[7], [12], whose block diagram is shown in Fig. 2. It comprises both the control and readout sections, in addition to service blocks, such as current, frequency, and voltage references [27], and a digital controller [28]–[30]. While the quantum processor and classical controller ideally reside at the same temperature, not enough cooling power is available in existing refrigerators to extract heat from the full controller at the temperature of the quantum processor (<1 mW below 100 mK [31]). Only some (de)multiplexing, such as frequency [17], [32] or time [33] multiplexing, is envisioned at the quantum-processor temperature to reduce the number of interconnects. The main part of the controller is then operated at 4 K, where a few watts of cooling power are available. Moreover, the possibility of bringing the qubits to a higher temperature, e.g., 4 K, is currently being investigated [33]. A short-term target for the proposed controller is to serve over 1000 qubits, leading to a target power consumption of only $\sim$1 mW/qubit.

For the spin-qubit readout, an LNA can be employed to sense the reflection coefficient at the end of a 50-Ω line connected to the charge sensor in a so-called reflectometry readout [32], [34]. Due to the limited sensor sensitivity ($\Delta R/R \approx 1\%$), as well as the fact that qubits are sensitive to electric field variations, only a very weak reflected power ($-135$ dBm in 1-MHz bandwidth per qubit) is thus realistically available. It means that the readout circuitry must be operating at very low-noise levels (at least below 40 pV/$\sqrt{\text{Hz}}$ for an SNR $> 0$ dB). However, by taking advantage of the low thermal noise at cryogenic temperatures, a good power efficiency can be achieved using a wideband LNA offering frequency multiplexing of many qubit channels, as described in Section IV.

The lower thermal noise at cryogenic temperatures also benefits the photon detector employed for the readout of qubits implemented using nitrogen-vacancy centers in diamond (operating from 500 to 1550 nm [35]). As described in Section III, a single-photon avalanche diode (SPAD) has a highly improved dark count rate (DCR) at 77 K over a wavelength range of 350–800 nm.

On the control side, the application of microwave bursts requires the generation of a carrier using an oscillator in a phase-locked loop (PLL). In order not to reduce the dephasing
Fig. 3. Output characteristics of (a) thick-oxide and (b) thin-oxide NMOS and PMOS in 160-nm CMOS technology.

Fig. 4. Output characteristics of (a) thick-oxide and (b) thin-oxide NMOS and PMOS in 40-nm CMOS technology.

time ($T_{2}^{*} = 120$ $\mu$s for the state-of-the-art spin qubit in purified silicon [18]), the frequency noise (FN) should be lower than 1.9 kHz$_{\text{rms}}$ over a bandwidth with the upper bound set by the qubit operation speed (currently around 1 MHz and extending in the future to 10 MHz) and the lower bound set by the echo-period [24], [36]. A class-$F_{2,3}$ oscillator optimized for cryogenic operation to achieve such performance is described in Section V.

III. CRYO-CMOS CHARACTERIZATION

The first challenge to address when designing CMOS circuits at cryogenic temperatures is the availability of device models. Figs. 3 and 4 show the measured $I_D$-$V_{DS}$ characteristics for both NMOS and PMOS transistors at 300 and 4 K in 160- and 40-nm CMOS technologies, respectively. As expected, the drain current at 4 K is higher than that measured at 300 K, mainly due to increased carrier mobility. The mobility-induced current increase is partially mitigated by the increase in threshold voltage that is shown in Fig. 5 for both thin-oxide 160-nm NMOS and PMOS.

In addition to a large variation of the transistor parameters, specific cryogenic non-idealities can be present, such as a kink and hysteresis. For example, the thick-oxide transistors in 160-nm CMOS show a clear kink at higher $V_{DS}$ [Fig. 3(a)], due to the bulk current generated by impact ionization at the drain combined with increased resistivity of the frozen-out substrate, leading to a decrease in the threshold voltage [37]. As an additional challenge, prior work suggests that mismatch deteriorates below 70 K due to freeze-out in the base [27].

Among devices that are feasible in a CMOS process technology, a SPAD implemented in 180-nm CMOS (Fig. 8) was characterized over temperature down to 77 K. A SPAD is sensitive to single photons due to an extremely high electric field in the reverse biased p-n junction, so as to operate above breakdown. The DCR, i.e., the occurrence of randomly
triggered spurious events in the junction, is heavily dependent on temperature [39]. In general, trap-assisted mechanisms dominate at or near room temperature, while tunneling dominates at lower temperatures.

Significant changes are also observed in digital logic. Despite the drawbacks related to the higher threshold voltage, digital logic benefits from the higher \( I_{ON}/I_{OFF} \) ratio due to the steeper sub-threshold slope. The measured delay of an inverter in 160-nm CMOS at a nominal supply voltage improves by 20%, from 38.3 ps at 300 K to 30.6 ps at 4 K. In 40-nm CMOS, the improvement is 36%, indicating that significant speedups can be achieved at cryogenic temperatures.

While cryogenic behavior of CMOS devices has been explained and explored in depth [37], no standard cryogenic models exist for commercial technologies and simulators. Therefore, to enable the design of cryogenic circuits, a modified MOS11 model [40] for transistors operating at 4 K has been developed [41] and extensively used in the simulations shown in Section IV.

**IV. LOW-NOISE AMPLIFIER FOR SPIN-QUBIT REFLECTOMETRY**

As stated in Section II, an LNA for spin-qubit reflectometry readout must demonstrate extremely low noise, i.e., an input-referred noise power spectral density (PSD) \( S_{tot} = (40 \text{ pV}/\sqrt{\text{Hz}})^2 \), while amplifying signals from a 50-\( \Omega \) input line. Such requirements greatly differ from those of prior cryogenic CMOS LNAs designed for different applications [9]. The proposed LNA exploits a noise-canceling technique to satisfy both noise and input impedance matching requirements [42]. As shown in the simplified schematic of Fig. 9, this technique enables canceling the noise of the input device (\( M_1 \)) by feeding it to the output through two paths (\( M_5 \) and \( M_3 \)) with the same gain magnitude but opposite phase. The input device can then be optimized solely for input matching, since \( Z_{in} \approx 1/gm_1 \), where here and thereafter \( gm_1 \) and \( ID_i \) are the transconductance and drain current of transistor \( M_1 \), respectively. Noise cancellation is achieved when

\[
\frac{gm_3}{gm_5} = 1 + \frac{R_f}{R_s} = 1 + |G_1| \tag{1}
\]

which links the transconductance of \( M_5 \) (\( gm_3 \)) and \( M_3 \) (\( gm_5 \)) to the feedback resistance (\( R_f \)) and the source impedance (\( R_s \)), or to the voltage gain of the amplifier (\( G_1 \)) [42]. In the full schematic of Fig. 10 (device sizes are shown in Table I), transistors \( M_1 \) and \( M_3 \) are cascoded by \( M_2 \) and \( M_4 \), respectively, to enhance their output impedance. Furthermore, the currents through \( M_1-M_3 \) are made tunable, either externally or by two current digital-to-analog converters (IDACs) referenced to an on-chip constant-\( gm \) bias circuit (\( I_{BIAS1} \) and \( I_{BIAS2} \)). This enables an independent calibration of input matching and noise cancellation, both at room and cryogenic temperatures, against expected worse matching and process spread [38]. Three additional gain stages and a 50-\( \Omega \) output driver follow the LNA to boost the gain of the full amplifier and to facilitate testing.

The amplifier is optimized for the minimum dissipated power normalized to the number of qubit channels processed by the LNA, i.e., for the minimum power per qubit (\( P_{qubit} \)). The following assumptions are made in the optimization.

1) The LNA is followed by a second single-transistor amplifying stage contributing to the overall noise, with only thermal noise being considered. Any other following stage would only contribute negligible noise.

2) Transistors are assumed in strong inversion with \( \beta = gm/ID = 10 \text{ V}^{-1} \).

3) The noise-canceling condition [see (1)] is met.

4) The LNA power is a sum of a noise-dependent contribution \( P_{LNA,noise} \) and a constant contribution \( P_{LNA,\text{const}} \). The former is related to the transistors contributing to the overall noise (\( M_3 \) and \( M_5 \)); i.e., \( P_{LNA,noise} = V_{DD}ID_3 \), while the latter refers to the power required for \( M_1 \) to provide the input matching, i.e., \( P_{LNA,\text{const}} = V_{DD}ID_1 = V_{DD}(gm_1/\beta) = (V_{DD}/\beta Z_{in}) \), with \( Z_{in} = 50 \text{ } \Omega \) and supply voltage \( V_{DD} \). The second-stage power \( P_{2,noise} \) is assumed to be fully noise-dependent.

5) The bandwidth of the system is assumed to be limited by the LNA.

\[1\]The other transistors in Fig. 10 are considered negligible in terms of noise because of proper sizing or because of noise cancellation (\( M_1 \)).
Fig. 8. (a) SPAD micrograph. The device was fabricated in 180-nm standard CMOS technology and has a diameter of 12 μm. (b) DCR, normalized by the active area of the SPAD, for two types of SPADs, which have different multiplication regions for optimized sensitivity in different regions of the spectrum.

With $S_1$ and $S_2 = (S_{\text{tot}} - S_1) \cdot G_1^2$ being the input-referred noise PSD of the LNA and the second amplifier, respectively, the noise-dependent power of the second stage can be derived as

$$P_{2,\text{noise}} = V_{\text{DD}} \cdot I_{\text{stage2}} = V_{\text{DD}} \cdot \frac{g_m \cdot I_{\text{stage2}}}{\beta} = V_{\text{DD}} \cdot \frac{4k_B T \gamma}{\beta S_2} = \frac{4k_B T \gamma V_{\text{DD}}}{\beta(S_{\text{tot}} - S_1) \cdot G_1^2},$$

(2)

where $I_{\text{stage2}}$ and $g_m \cdot I_{\text{stage2}}$ are the drain current and transconductance of the second-stage transistor, respectively, $k_B$ is the Boltzmann’s constant, $T$ the absolute temperature, and $\gamma$ the transistor excess-noise coefficient. By using (1), the LNA input-referred noise, which includes contributions from $R_f$, $M_3$, and $M_5$ in Fig. 9, can be expressed as

$$S_1 = \frac{4k_B T}{|G_1|^2} \left[ R_f + \gamma \frac{(1 + |G_1|)(2 + |G_1|)}{\beta G_{1,\text{noise}}} \right].$$

(3)

Combining (1)–(3), an expression for the total power of the LNA and of the second stage can be derived as

$$P = P_{\text{LNA, constellation}} + P_{\text{LNA, noise}} + P_{2, \text{noise}}$$

$$= \frac{V_{\text{DD}}}{\beta Z_{\text{in}}} \cdot \frac{4k_B T \gamma V_{\text{DD}}}{|G_1|^2} \cdot \frac{1 + |G_1|}{S_1 G_1^2 - 4k_B T R_f}$$

$$+ 4k_B T \gamma V_{\text{DD}} \left( \frac{2}{|G_1|^2} + 1 \right).$$

(4)

(5)

It can be shown that $P$ has a minimum with respect to the input-referred PSD of the LNA, $S_1$, and that such a minimum $P_{\text{min}}(G_1)$ is a function of the LNA gain $G_1$. In fact, if $S_1$ is small, the power is dominated by the LNA, while if $S_1$ approaches $S_{\text{tot}}$, the second stage needs to consume more power because of its tighter noise budget. Furthermore, it can be shown that $P_{\text{min}}(G_1)$ increases for lower $|G_1|$, since, for a low gain, the noise budget is completely dominated by $R_f$, thereby resulting in a larger power in the active components ($M_3$ and $M_5$).

The other parameter that plays a crucial role in $P_{\text{qubit}}$ optimization is the number of qubit channels ($n_{\text{qubit}}$), which is directly proportional to the LNA bandwidth. By assuming $C_a$, $C_b$, and $C_{\text{load}}$ in Fig. 9 to be the dominant capacitors and extracting their values from circuit simulations, the frequency of the zero ($f_z$) and of the poles ($f_a$, $f_b$, and $f_{\text{load}}$ related to $C_a$, $C_b$, and $C_{\text{load}}$, respectively) of the LNA can be readily calculated. An approximated expression for the bandwidth $BW_{\text{LNA}}$ is then derived and shown in Fig. 11, which also shows that the speed of the amplifier decreases with $G_1$. In fact, a high gain requires a large $M_3$ [from (1)] and, therefore, a large capacitance $C_a$.

The total power is then calculated as

$$P_{\text{qubit}} = P \cdot \frac{n_{\text{qubit}}}{n_{\text{qubit}}} = P \cdot \left[ \frac{BW_{\text{LNA}} - f_{\text{start}}}{BW_{\text{qubit}}} \right]^{-1}$$

(6)

where $BW_{\text{qubit}} = 2$ MHz is the bandwidth per qubit that includes both the signal bandwidth (1 MHz) and the same amount for channel spacing, and $f_{\text{start}} = 200$ MHz is the minimum allowed channel frequency to avoid large passive components in the matching networks close to the qubits.

According to the analysis, at $|G_1| \approx 21$ V/N, an overall bandwidth of the LNA of $(BW_{\text{LNA}})$ of 850 MHz, a total power of 50 mW (48 mW for the LNA), and $P_{\text{qubit}} = 150$ μW are anticipated.

Three diode-loaded common-source stages are cascaded to boost the total gain to 60 dB while an output driver matches the output impedance of the amplifier to 50 Ω.

The pole at the lowest frequency ($f_z$) is assumed to be compensated by the zero, thus resulting in the LNA bandwidth being approximated by the frequency of the pole at the immediately higher frequency ($f_a$ or $f_{\text{load}}$).
TABLE I

<table>
<thead>
<tr>
<th>Device</th>
<th>W/L [μm/μm]</th>
<th>Device</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>58/0.16</td>
<td>$R_{bias1}$</td>
<td>200 kΩ</td>
</tr>
<tr>
<td>$M_2$</td>
<td>30.16/0.2</td>
<td>$R_{bias2}$</td>
<td>100 kΩ</td>
</tr>
<tr>
<td>$M_3$</td>
<td>1276/0.2</td>
<td>$C_{bias1}$</td>
<td>12 pF</td>
</tr>
<tr>
<td>$M_4$</td>
<td>324.8/0.2</td>
<td>$C_{bias2}$</td>
<td>51 pF</td>
</tr>
<tr>
<td>$M_5$</td>
<td>30.16/0.16</td>
<td>$C_{AC}$</td>
<td>0.6 pF</td>
</tr>
<tr>
<td>$M_6$</td>
<td>122.96/0.2</td>
<td>$M_7$</td>
<td>157.76/0.2</td>
</tr>
<tr>
<td>$M_8$</td>
<td>30.16/0.2</td>
<td>$M_9$</td>
<td>39.44/0.2</td>
</tr>
<tr>
<td>$M_{10}$</td>
<td>928/1</td>
<td>$M_{11}$</td>
<td>92.8/1</td>
</tr>
</tbody>
</table>

A test chip was fabricated in the SSMC 160-nm CMOS technology (Fig. 12), comprising the proposed LNA, the three cascaded stages, and the output driver. The circuit comprises only thin-oxide devices, which do not suffer from the kink effect, as shown in Section III. The bias circuit includes a constant-$g_m$ block that generates the necessary bias current for the amplifier. This circuit proved to be functional both at 300 K, generating approximately 103 μA, and at 4 K, generating 123 μA. Fig. 13 shows the measured S-parameters representing gain ($S_{21}$) and input matching ($S_{11}$) both at room temperature and 4 K. The amplifier has a gain of 40 dB and a −3-dB bandwidth of 400 MHz at room temperature, which improve to 57 dB and 500 MHz at 4 K, respectively, enabling the allocation of 150 qubit channels. Input matching improves by 3 dB by going to lower temperatures, reaching a value of −8 dB at 4 K. These parameters are measured with IDAC settings for the lowest noise figure (NF) at room temperature and 4 K.

Fig. 14 shows the gain ($S_{21}$) of the full chain comparing measurements with simulations of the post-layout netlist including parasitic inductance of the bondwires connected to input–output and supply pads at 300 K and 4 K. The SPICE models adapted for cryogenic temperatures (see Section III) have been employed in the simulations. The comparison shows a fairly good match with the simulations at both temperatures. Noise measurements reveal an in-band NF of 0.1 dB±0.05 dB at 4 K and 0.8 dB±0.3 dB at 300 K, as shown in Fig. 15, which is equivalent to a noise temperature of 7 K and 60 K, respectively. The degradation of the NF is due to the AC-coupling ($C_{AC}R_{bias1}$ in Fig. 10) and the flicker noise of $M_1$ and $M_3$ at low frequency, in combination with the decrease of gain at high frequencies. Large peaks are visible in the NF plots, but they are associated with external electromagnetic interference (EMI). Such interference is not expected to be an issue in the final quantum computer application thanks to the EMI shielding in the refrigerator. The measured NF translates to an improvement of a ∼10× noise factor at 4 K with respect to 300 K that can be attributed to the large decrease in thermal noise [43].
of which 54.9 mW (45.9 mW) is for the LNA, and the power, which is below the state-of-the-art LNAs, high linearity is not required in the Although the linearity performance is not comparable with the amplifying transistors caused by the increase of mobility. at lower temperature and the reduced overdrive voltage in expected non-linearity at 4 K is mainly due to the larger gain higher gain and the reduced overdrive voltage.

Fig. 15. Measured NF at 300 K and 4 K. The peaks come from external electromagnetic interference.

Fig. 16. Measured IIP3 (top) and IIP2 (bottom) at 300 K and 4 K, respectively. At 4 K, the linearity is around 10 dB worse because of the higher gain and the reduced overdrive voltage.

Fig. 14. Comparison of $S_{21}$ between measurements and simulations both at 300 K and 4 K with the use of the developed SPICE models for low temperatures.

V. DIGITALLY CONTROLLED OSCILLATOR FOR QUBIT CONTROL

As mentioned in Section II, the tolerable FN of a PLL-stabilized oscillator for qubit control should be lower than 1.9 kHz$_{rms}$ for a 6-GHz carrier. Since the total integrated phase noise (PN) is crucial for this application, the PLL’s bandwidth ($f_{PLL}$) should be optimized for the in-band and out-of-band PN performance. Beyond the PLL’s bandwidth, the PN is dominated by the oscillator, and can be expressed as $\mathcal{L}_{osc}(\Delta f) = \alpha/\Delta f^2$, where $\Delta f$ is the frequency offset from the carrier and $\alpha$ is a constant coefficient$^3$ in Hz$^2$/Hz. The PLL’s in-band PN is constrained by the PN of phase detector and reference clock, which should be lower than the oscillator’s PN at $f_{PLL}$. Consequently, the PLL’s PN can be expressed roughly by

$$\mathcal{L}_{PLL}(\Delta f) = \begin{cases} \alpha/f_{PLL}^2 & \Delta f \leq f_{PLL} \\ \alpha/\Delta f^2 & \Delta f \geq f_{PLL}. \end{cases}$$

Furthermore, the integrated FN of the PLL can be estimated by [44]

$$FN_{PLL} = \sqrt{2 \cdot \int_{f_{PLL}}^{f_a} (\Delta f)^2 \cdot \mathcal{L}_{PLL}(\Delta f) \cdot d(\Delta f).}$$

The lower integration bound $f_a$ is set by the quantum operation cycle (worst case: $1/T_q^3 \approx 8.3$ kHz) and the higher limit is determined by the qubit operation speed (e.g., 10 MHz). By inserting (7) into (8)

$$FN_{PLL} = \alpha \left( 2f_b - f_{PLL} \left[ \frac{4}{3} + \frac{2}{3} \left( \frac{f_a}{f_{PLL}} \right)^3 \right] \right)^{1/2} < 1.9 \text{ kHz} \quad (f_a < f_{PLL} < f_b). \quad (9)$$

By considering a typical PLL bandwidth of 300 kHz and exploiting (9), $\alpha$ is calculated to be lower than 0.2 Hz$^2$/Hz. This translates to the in-band PN of$^4$ $-115$ dBc/Hz and oscillator’s PN of $-147$ dBc/Hz at 10-MHz offset from the carrier. Even by considering the thermal noise reduction of passive and active devices at cryogenic temperatures, achieving such PN performance is challenging with traditional MOS cross-coupled $LC$ oscillators [46]. Furthermore, as mentioned before, the power consumption of the control electronics is severely limited by the cooling power of the dilution refrigerator. Hence, a power-efficient oscillator topology with high figure-of-merit (FoM)$^5$ is essential. Moreover, it has been observed that the flicker noise of MOS transistors can significantly worsen at lower temperatures [47]. This may result in a much higher $1/f^3$ PN corner that can exceed the PLL’s bandwidth, thus degrading the integrated FN. Consequently, an oscillator topology with low $1/f^3$ PN corner is desired.

$^3$The oscillator’s $1/f^3$ corner is assumed to be below the PLL’s bandwidth.

$^4$This can be easily achieved by employing an integer-N sub-sampling PLL architecture [45] with 50 MHz reference clock ($-155$ dBc/Hz noise floor).

$^5$FoM = $|PN| + 20 \cdot \log_{10} \left( \frac{f_{ PLL}}{f_{qubit}} \right) - 10 \cdot \log_{10} \left( \frac{P_{PLL}}{1\text{mW}} \right)$.
To address the aforementioned issues, a transformer-based class $F_{2,3}$ oscillator was designed [48]. It is an extension of a class-$F_3$ oscillator [49], which introduces an auxiliary impedance peak $\omega_{1,DM}$ around the third harmonic of the fundamental differential mode (DM) oscillation frequency ($\omega_{1,DM} = 3\omega_0,DM$) in order to convert the third harmonic of the active device current into voltage, thus creating a pseudo-square oscillation waveform. As a result, the oscillator’s impulse sensitivity function is reduced, especially when the $g_{m}$-devices enter the triode region and inject large noise into the tank. Consequently, thermal-to-PN upconversion is lower in this operation. On the other hand, to reduce the flicker noise upconversion, the oscillator tank should exhibit an auxiliary common-mode (CM) resonance at the second harmonic of the fundamental DM oscillation frequency $\omega_{CM} = 2\omega_0,DM$ [48], [50]. As discussed in [51], this class-$F_2$ operation can be realized by exploiting the different behavior of the transformer at CM and DM excitation. To exploit the advantages of both operations, a class-$F_{2,3}$ oscillator exhibiting two DM resonances and one CM resonance was thus designed. This was achieved by careful design of the inductance values of the primary ($L_p$) and secondary ($L_s$) windings of a step-up (1:2) transformer, along with single-ended primary ($C_p$) and differential secondary capacitor banks ($C_s$).

In the proposed transformer, the CM input signal can neither see the secondary winding of the transformer nor the $C_s$ tuning capacitors. Consequently, single-ended switched capacitors are required to obtain the desired CM resonance. However, differential capacitors are employed at the secondary, due to their higher $Q$-factor in the ON-state. For this particular design, the class-$F_{2,3}$ operation is satisfied with $L_s C_s = 3.8 L_p, d C_p$ for a coupling factor ($k_m$) of 0.67.

The channel resistance of long-channel-length MOS devices significantly reduces in the kink region [see Fig. 3(a)]. If the $g_{m}$-devices are biased in this region, they significantly load the tank at the oscillation zero-crossings, degrading the effective $Q$-factor and PN of the oscillator. Even though the nominal $V_{DD}$ of thick-oxide devices is 2.5 V in this technology, the oscillator’s $V_{DD}$ is chosen as 1 V to limit the oscillation swing at the drain nodes in order to avoid this probable kink region. However, a 1:2 step-up transformer is employed to boost the oscillation voltage at the gate nodes where most of the tank capacitance is located, consequently improving the oscillator’s PN and its startup margin. The schematic and the oscillation waveforms are shown in Fig. 17. At the bottom of the core $g_{m}$ transistors, a 5-bit binary-weighted switchable resistor with an LSB of 5 $\Omega$ is implemented to roughly control the oscillator current. The single-ended primary and differential secondary capacitor banks are realized using two 6-bit switchable metal–oxide–metal (MoM) capacitors with an LSB of 5 and 10 $fF$, respectively. Due to the class-F operation, a pseudo-square waveform is realized at the drain of the $g_{m}$-devices. However, the transformer filters out the harmonic components of the drain oscillation voltage and hence, a sinusoidal waveform is restored at the gate.

The oscillator was prototyped in a 40-nm 1P7M CMOS process with an ultra-thick metal layer. The chip micrograph is shown in Fig. 18. The oscillator performance was characterized across a wide temperature range from 300 K to 4 K. As shown in Fig. 19, the measurement setup consists of a hollow steel pipe, where the device under test is mounted at one end, while the other end features the connectors to interface with the instruments at room temperature. This pipe
Fig. 19. Oscillator measurement setup for temperature sweep.

Fig. 20. Oscillator current consumption versus temperature.

is dipped into a helium barrel and its position inside the barrel is changed to adjust the temperature. The oscillator’s temperature is monitored using a sensor mounted on the back side of the PCB, right under the chip. A comprehensive investigation of discrete components on the PCB was done for proper measurement of the chip at cryogenic temperatures. For example, capacitance of widely used X5R and X7R capacitors would reduce by almost 90% and their effective series resistance would increase at least $10^5$. Hence, NP0- and COG-type capacitors are employed on this PCB [30] (see Fig. 19) with values ranging from 47 pF to 100 μF, apart from an on-chip decoupling capacitor of 27 pF. NP0/COG capacitors are available up to a value of 1 μF and are much larger in dimension than commercially available X5R or X7R capacitors, due to lower dielectric constant. Moreover, due to internal space constraints of the dilution refrigerator, capacitors above 1 μF have to be electrolytic, although they work rather poorly at cryogenic temperatures [30].

Fig. 20 reveals that, for any fixed switched-resistor configuration, the oscillator dissipates more power at lower temperatures. One of the reasons is the increase in electron mobility at colder temperatures, as justified in Section III. Another reason is a reduction in resistance of the poly resistors (10% from 300 K to 4 K) that are used to control the oscillator’s current. It is also observed that the oscillation frequency shifts up ∼7% from 300 K to 4 K, although the frequency span is largely maintained, as shown in Fig. 21 for different values of the coarse and fine-tuning capacitance. The inductance change is on the order of ppm/°C [52]. Hence, it becomes evident that the frequency shift is caused by a reduction in capacitance of the MoM capacitors used for frequency tuning and parasitics of the $g_m$-devices, as well as the transistors in the switched capacitor array.

The measured PN plot is shown in Fig. 22 for different temperatures at the same power consumption, $P_{dc}$. The level of the spurs, especially at low offset frequencies, increases at cryogenic temperatures, indicating the inability of on-board decoupling capacitors to filter supply noise [30]. Fig. 23 shows the oscillator PN in the thermal noise region at $Δf = 30$ MHz offset frequency versus temperature, again at a constant $P_{dc}$. To better understand this plot, it is instructive to analyze the variation of temperature-dependent terms in the PN equation [53]

$$L(Δω) = 10 \cdot \log_{10} \left( \frac{k_B T}{2} \cdot \frac{1}{Q^2 \cdot αV \cdot αI \cdot P_{dc}} \cdot \left( \frac{f_0}{Δf} \right)^2 \cdot (1 + γ) \right).$$

(10)
where $\alpha_I$ and $\alpha_V$ are the oscillator’s current and voltage efficiency, respectively. Note that these parameters are mainly dependent on the oscillator topology; they are only weakly dependent on temperature. From 300 to 170 K, the PN improves by 30 dB/decade, out of which 10 dB/decade is attributed to the temperature reduction, while the tank’s $Q$-factor enhancement realizes the remaining 20 dB/decade. At room temperature, magnetically induced image currents from the tank flow in the low-resistive substrate, reducing the quality factor. As mentioned in Section III, due to the carrier freeze-out at cryogenic temperatures, the substrate becomes highly resistive, hence lowering substrate losses considerably. Moreover, the inductor’s series resistance also decreases at lower temperatures, thus further improving the tank’s $Q$-factor. Note that the transistor’s ON-resistance does not dramatically reduce at cryogenic temperatures. Consequently, below a certain temperature, the tank’s $Q$-factor is dominated by the switched capacitors. Hence, the PN improvement as a function of temperature reduces below 170 K. Moreover, it can be observed that the PN at 30 MHz offset does not improve linearly with the temperature below 70 K. This could be due to an increase of the channel noise factor, $\gamma$ [54].

Even though the flicker noise upconversion is significantly suppressed by virtue of the chosen topology, the oscillator’s $1/f^3$ PN corner improves dramatically at lower temperatures, as shown in Fig. 24, thus hinting at a larger $1/f$ noise corner for MOS transistors at cryogenic temperatures.\(^6\) Another reason could be the worsening of mismatch [38] between the two core transistors at cryogenic temperatures, resulting in an asymmetric rise/fall times of oscillation waveforms and, consequently, resulting in a larger dc value of the impulse sensitivity function and higher flicker noise upconversion [55]. For completeness, the PN values at 4 K at different offset frequencies across the tuning range are shown in Fig. 25.

Even without a PLL, the FN of the proposed oscillator is as low as 3.4 kHz\(_{\text{rms}}\) over a worst-case integration bandwidth of 8.3 kHz to 10 MHz. Hence, by using, for example, a sub-sampling (digital or analog) PLL with a bandwidth of $\leq 1$ MHz, the desired integrated FN of 1.9 kHz\(_{\text{rms}}\) can easily be achieved.

### VI. Conclusion

This paper describes the design and experimental validation of several major circuit blocks critical for the implementation of a CMOS classical electronic controller to operate at cryogenic temperatures (i.e., cryo-CMOS) in order to interface with a practical quantum processor. We have demonstrated a 4-K 160-nm LNA capable of reading out 150 1-MHz qubit channels with a power efficiency better than 700 $\mu$W/qubit and a 4-K 40-nm 6-GHz class-$F_{2.3}$ oscillator with an integrated FN of 3.4 kHz\(_{\text{rms}}\) over $\sim 10$ MHz bandwidth, which is low enough to drive the state-of-the-art qubits without limiting their performance. Such performance is achieved by carefully employing standard circuit design techniques while exploiting specific characteristics of the adopted cryo-CMOS devices, such as increased speed and low thermal noise. Although further effort is required to develop a full cryogenic controller demonstrating the required performance in the very tight power budget set by existing dilution refrigerators, the proposed circuits show that cryo-CMOS is a viable technology for the implementation of such classical electronic controllers, thus establishing cryo-CMOS circuits and systems as an enabling technology for the fabrication of practical quantum computers with thousands or even millions of qubits.

### Acknowledgment

The authors would like to thank Dr. M. Aminian for SPAD characterizations, NXP Semiconductors for 0.16-$\mu$m CMOS foundry services and A. Akhnoukh, W. Straver, Z. Chang and R. Puffelen for measurement support.
Mina Shahmohammadi (’S12–M’17) received the B.Sc. degree in communication systems from the Amirkabir University of Technology, Tehran, Iran, in 2005, the M.Sc. degree in electronics from the University of Tehran, Tehran, in 2007, and the Ph.D. degree from Electronics Research Laboratory, Delft University of Technology (TU Delft), Delft, The Netherlands, in 2010, with a focus on wide tuning range and low flicker noise RF-CMOS oscillators. From 2007 to 2011, she was with Rezvan Engineering Company, Tehran, as an Analog Designer. She was a Research Assistant with the Electronic Instrumentation Laboratory, TU Delft, from 2011 to 2013, where she was involved in resistor-based temperature sensors. She is currently an Analog Designer with Catena B.V., Delft. Her current research interests include analog and RF integrated circuits design.

Robert Bogdan Staszewski (M’97–SM’05–F’09) was born in Białystok, Poland. He received the B.Sc. (summa cum laude), M.Sc., and Ph.D. degrees from The University of Texas at Dallas, Richardson, TX, USA, in 1991, 1992, and 2002, respectively, all in electrical engineering. From 1991 to 1995, he was with Alcatel Network Systems, Richardson, TX, USA, where he was involved in SONET cross-connect systems for fiber optics communications. In 1995, he joined Texas Instruments Inc., Dallas, TX, USA, where he was elected as a Distinguished Member of Technical Staff (limited to 2% of technical staff). From 1995 to 1999, he was involved in advanced CMOS read channel development for hard disk drives. In 1999, he co-started the Digital RF Processor (DRP) Group within Texas Instruments Inc., with a mission to invent new digitally intensive approaches to traditional RF functions for integrated radios in deeply scaled CMOS technology. From 2007 to 2009, he was a CTO of the DRP Group. In 2009, he joined the Delft University of Technology, Delft, The Netherlands, where he currently holds a guest appointment of Full Professor (Antoni van Leeuwenhoek Hogereleraar). Since 2014, he has been a Full Professor with the University College Dublin, Dublin, Ireland. He has authored or co-authored four books, five book chapters, 210 journal and conference publications, and holds 160 issued U.S. patents. His current research interests include nanoscale CMOS architectures and circuits for frequency synthesizers, transmitters, and receivers.


Masoud Babaie (S’12–M’16) received the B.Sc. (Hons.) degree in electrical engineering from the Amirkabir University of Technology, Tehran, Iran, in 2004, the M.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, in 2006, and the Ph.D. (cum laude) degree from the Delft University of Technology, Delft, The Netherlands, in 2016.

In 2006, he joined the Kavoshcom Research and Development Group, Tehran, where he was involved in designing wireless communication systems. From 2009 to 2011, he was a CTO of the company. From 2013 to 2015, he was consulting for the RF Group of Taiwan Semiconductor Manufacturing Co., Ltd., Hsinchu, Taiwan, where he was involved in designing 28-nm all-digital phase-locked loop and Bluetooth low-energy transceiver chips. From 2014 to 2015, he was a Visiting Scholar Researcher with the Berkeley Wireless Research Center, Berkeley, CA, USA, with the Group of Prof. A. Niknejad. In 2016, he joined the Delft University of Technology as an Assistant Professor. His current research interests include analog and RF/millimeter-wave integrated circuits and systems for wireless communications, and cryogenic electronics for quantum computation.

Dr. Babaie was a recipient of the 2015–2016 IEEE Solid-State Circuits Society Pre-Doctoral Achievement Award. He serves as a Reviewer of the IEEE JOURNAL OF SOLID-STATE CIRCUITS.

Fabio Sebastiano (S’09–M’10–SM’17) was born in Teramo, Italy, in 1981. He received the B.Sc. (cum laude) and M.Sc. (cum laude) degrees in electrical engineering from the University of Pisa, Italy, in 2003 and 2005, respectively, the M.Sc. degree (cum laude) from Sant’Anna School of Advanced Studies, Pisa, Italy, in 2006, and the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands, in 2011.

In 2006 to 2013, he was with NXP Semiconductors Research, Eindhoven, The Netherlands, where he was involved in conducting research on fully integrated CMOS frequency references, deep-submicron temperature sensors, and area-efficient interfaces for magnetic sensors. In 2013, he joined the Delft University of Technology, where he is currently an Assistant Professor. He has authored or co-authored one book, nine patents, and over 40 technical publications. His current research interests include sensor read-outs, fully-integrated frequency references, and cryogenic electronics for quantum applications.

Edoardo Charbon (SM’00–F’17) received the Diploma degree from ETH Zurich, the M.S. degree from the University of California at San Diego, and the Ph.D. degree from the University of California at Berkeley, Berkeley, CA, USA, in 1988, 1991, and 1995, respectively, all in electrical engineering and EECS.

He has consulted with numerous organizations, including Bosch, X-Fabs, Texas Instruments Inc., Dallas, TX, USA, Maxim, Sony, Agilent, and the Carlyle Group. He was with Cadence Design Systems from 1995 to 2000, where he was the Architect of the company’s initiative on information hiding for intellectual property protection. In 2000, he joined Canesta Inc., as the Chief Architect, where he led the development of wireless 3-D CMOS image sensors. Since 2002, he has been a member of the faculty of Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, where he has been a Full Professor since 2015. From 2008 to 2016, he was with Delft University of Technology, Delft, The Netherlands, as the Chair of VLSI design. He has been the driving force behind the creation of deep-submicron CMOS single-photon avalanche diode technology, which has been mass-produced since 2015 and is present in telemeters, proximity sensors, and medical diagnostics tools. He is currently a Distinguished Visiting Scholar with the W. M. Keck Institute for Space at Caltech and a fellow of the Kavli Institute of Nanoscience Delft. He has authored or co-authored over 250 papers and two books, and he holds 20 patents. His current research interests include 3-D vision, FLIM, FCS, NIROT to super-resolution microscopy, time-resolved Raman spectroscopy, and cryo-CMOS circuits, and systems for quantum computing.