THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : PHYS - Physique
Spécialité : Physique de la Matière Condensée et du Rayonnement
Unité de recherche : Laboratoire d'Électronique et de Technologie de l'Information (LETI - CEA)

Électronique cryogénique pour l'ingénierie quantique

Cryogenic electronics for quantum engineering

Présentée par :

Loïck Le Guevel

Direction de thèse :

Gaël Pillonnet DIRECTEUR DE RECHERCHE, CEA-LETI	Directeur de thèse
Louis Jansen INGENIEUR DE RECHERCHE, CEA-IRIG	Co-Directeur de thèse

Rapporteurs :

Joseph Bardin
PROFESSEUR, University of Massachusetts Amherst

Pietro Maris Ferreira
PROFESSEUR ASSOCIE, Université Paris-Saclay

Thèse soutenue publiquement le **13 janvier 2023**, devant le jury composé de :

Joseph Bardin PROFESSEUR, University of Massachusetts Amherst	Rapporteur
Sylvain Bourdel PROFESSEUR DES UNIVERSITES, Université Grenoble-Alpes	Examinateur Président du jury
Andreia Cathelin INGENIEUR, STMicroelectronics	Examinatrice
Pietro Maris Ferreira PROFESSEUR ASSOCIE, Université Paris-Saclay	Rapporteur
Salvador Mir DIRECTEUR DE RECHERCHE, CNRS	Examinateur
Fabio Sebastiano PROFESSEUR ASSOCIE, Delft University of Technology	Examinateur

Invités :

Gérard Billiot
INGENIEUR DE RECHERCHE, CEA-LETI

Abstract

Thousands to millions of sensitive signals will need to be conveyed through all the temperature stages of a dilution fridge to operate future large-scale quantum processors made of many quantum bits. The exploding number of heat-conductive coaxial cables will overwhelm the fridge cooling capabilities, detrimental to the quantum core. Moving the control electronics down to cryogenic temperatures allows the use of already-available superconducting cables, alleviating heat conduction between low-temperature stages, and appears as a clear path towards scalability in the number of operating qubits.

This Ph.D. work aimed to explore the use of the industrial CMOS 28nm Fully-Depleted Silicon-On-Insulator (FD-SOI) technology at cryogenic temperatures for quantum computing applications. Our first objective is to extend the sparse existing knowledge about the FD-SOI 28nm transistors at cryogenic temperatures for practical aspects of circuit design and later for developments of compact models.

To speed-up the characterization of single devices with the inherent hour-long cooling cycles, we designed an integrated circuit multiplexing a thousand of transistors for different geometries and gate-stack flavors for low-frequency measurement of the current-voltage characteristics and pair-matching analysis from 300 down to 0.1K. We discuss and analyze the evolution trends with varying geometry at different temperatures for important quantities in circuit design, such as the transconductance, the conductance, and the transconductance over drain current ratio of individual transistors.

Secondly, we explore the low-temperature co-integration and the full on-chip integration of semiconductor quantum devices with classical electronics aimed at specific measurements down to the millikelvin range.

We first focus on the sub-nanoampere current measurement of quantum dot devices by designing and characterizing a low-power transimpedance amplifier (TIA). The high-gain amplifier is successfully applied to measure the current across single- and double quantum-dot devices, respectively wire-bonded a few mm away or on-chip integrated a few micrometers-away. To further leverage the integration into the same substrate, we connected GHz-range voltage-controlled oscillators to one of the gates of the double dot in an attempt to observe discrete charge pumping in a fully-integrated device.

Lastly, we tackled the measurement of the gate capacitance of single quantum devices by proposing a new measurement scheme making use of the cryogenic electronics capabilities as an alternative to the well-known reflectometry. By integrating a voltage-controlled current excitation and a voltage-sensing amplifier in the 200 MHz range, both in close proximity to the quantum device connected to an LC tank, the read-out circuitry of a variation in the device capacitance becomes a purely lumped-element system with impedance measurement of the resonant circuit without any wave propagation like in reflectometry. This approach increases simplicity and compactness of the measurement set-up. We even replace the bulky passive inductor used in reflectometry by an active inductor made of transistors and capacitors, offering improved scalability with a 3-orders of magnitude lower area for the same inductance. The resulting circuit successfully measured aF-capacitance variations of nanometric transistors at 4.2K revealing oscillatory quantum effects in the gate capacitance as a function of the gate and back-gate voltages.

At the end of this dissertation, a picture is given with the challenges laying ahead related to circuit architecture and design for the ultimate goal of entering into the era of large-scale quantum processing.

Résumé

Des milliers, voire des millions de signaux sensibles devront être acheminés à travers tous les étages de température d'un réfrigérateur à dilution pour faire fonctionner les futurs processeurs quantiques à grande échelle composés de nombreux bits quantiques (qubits). L'explosion du nombre de câbles coaxiaux thermo-conducteurs va submerger les capacités de refroidissement du réfrigérateur, au détriment du noyau quantique. Refroidir l'électronique de contrôle, aujourd'hui à températures ambiantes, à des températures cryogéniques pourrait permettre l'utilisation de câbles supraconducteurs, thermiquement isolant, et ouvrir la voie vers l'augmentation du nombre de qubits.

Ces travaux de doctorat visent à explorer l'utilisation de la technologie CMOS FD-SOI 28nm à des températures cryogéniques pour des applications dans le calcul quantique. Notre premier objectif fut d'étendre les connaissances existantes sur les transistors FD-SOI 28 nm à des températures cryogéniques pour les aspects pratiques de la conception de circuits et plus tard pour le développement de modèles compactes.

Pour accélérer la caractérisation de dispositifs CMOS, longue à cause de la durée inhérente des cycles de refroidissement, nous avons conçu un circuit intégré multiplexant un millier de transistors aux différentes géométries et empilements de grilles pour la mesure des caractéristiques courant-tension basses-fréquences et des disparités entre dispositifs identiques de 300 à 0.1K. Nous discutons et analysons le comportement en fonction de la température et de la géométrie des transistors en se concentrant sur des quantités importantes lors de la conception de circuits intégrés, telles que la transconductance, la conductance et le rapport de la transconductance sur courant de drain des transistors.

Deuxièmement, nous explorons la co-intégration à basse température et l'intégration complète sur un même substrat silicium de dispositifs quantiques semi-conducteurs avec leur électronique classique visant à les mesurer jusqu'à des températures de l'ordre du millikelvin.

Nous nous concentrons d'abord sur la mesure du courant à travers des puits quantiques en concevant et en caractérisant un amplificateur à transimpédance de faible puissance (TIA). L'amplificateur de fort gain permet de mesurer avec succès le courant à travers des dispositifs à un et deux puits quantiques, respectivement placés à quelques mm ou intégrés sur puce à quelques micromètres. Afin de tirer davantage parti de l'intégration sur le même substrat, nous avons connecté des oscillateurs commandés en tension dans la gamme GHz à l'une des portes du double puit quantique dans le but d'observer le pompage de charge discret au sein d'une unique puce.

Finalement, nous avons abordé la mesure de la capacité de grille de dispositifs nanométriques en proposant un nouveau mode de mesure qui tire parti de l'électronique cryogénique et qui se substitue à la célèbre réflectométrie. En intégrant une source de courant commandée en tension et un amplificateur de tension opérant à des fréquences de l'ordre de 200 MHz à proximité immédiate du dispositif quantique et du résonateur, le circuit de lecture de la capacité de grille devient un circuit électronique discret sans ondes propagatrices, contrairement à la réflectométrie, pour une plus grande simplicité et meilleure compacité. Nous remplaçons même l'encombrante inductance passive utilisée en réflectométrie par une inductance active composée de transistors et de condensateurs, offrant une surface active réduite. Le circuit conçu mesure avec succès les variations de capacité de grille des transistors nanométriques à 4,2 K, ainsi révélant des effets quantiques oscillatoires dans la capacité de grille en fonction des tensions de grille.

À la fin de cette thèse, une image est dépeinte concernant les futurs défis et obstacles liés à l'architecture et à la conception des circuits dans le but ultime d'entrer dans l'ère du calcul quantique à grande échelle.

Contents

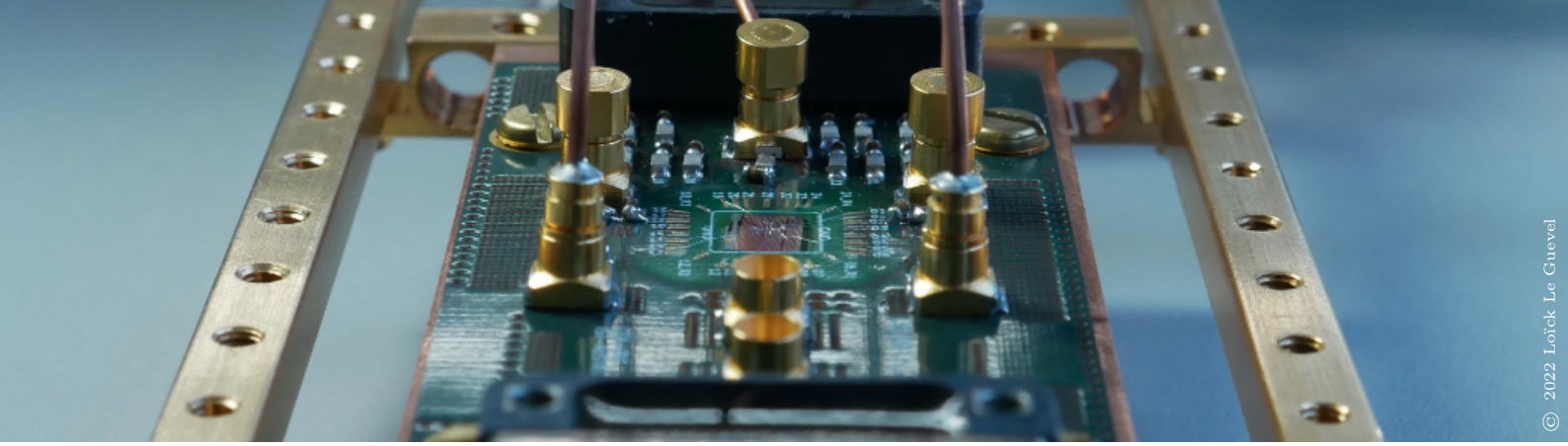
List of Figures

1 Introduction

2 Deep-Cryogenic Massive DC Characterization of FD-SOI MOSFETs

E Quantum Full Integration with the Impedancemetry Chip

List of Tables



CHAPTER 1
Introduction

The development of computing hardware is undoubtedly one of the most impactful inventions of the last century. The wide range of applications of programmable computing quickly intruded many aspects of our modern society from fundamental research to every-day problem solving.

Digital computing, that we will refer to as classical computing in this thesis, is the dominant type of computation. Classical computing has been using increasingly small switch-type devices, evolving from decimetric vacuum tubes to the modern nanometric transistors. The ability of switch-type devices to toggle other switch-type devices allows to chain tens of them to realize simple or more complex digital operations, commonly known as digital gates (AND, NOT, NAND,...). Assembling digital gates into higher abstraction levels made of 10^3 to 10^{10} transistors, e.g. building arithmetic logic units, central processing units, graphics processing units, or tensor processing units, leads to the modern computation units, able to self-drive a car, to make better medical visual diagnostic than humans, and to beat any human player in extremely complex games. Although classical computing has shown to be able to resolve more and more complex problems by improving hardware, some particular problems remain extremely hard to solve exactly, requiring an incommensurable amount of time or resources.

I Classical computing and its limits

One well known example of classical computing limit is the factorization of a large number into a (unique) multiplication of 2 or more prime numbers. As much as it is easy for off-the-shelf computers to decompose rather small number such as 3233 into the product of prime numbers 61×53, it becomes exponentially more complex as the factorized number increases, even for a classical super-computer. This out-of-reach complexity is at the heart of today's encryption routines such as the well-known Rivest-Shamir-Adleman (RSA) code. The use of large numbers typically between 3×10^{616} (2048 bits) and 1×10^{1233} (4096 bits) prevents any near-future classical computer to find the prime number factorization.

The complexity of solving a problem is often expressed with the $O(\cdot)$ operator describing how the number of operations scales with the problem size n. The faster the number of operations grows with the problem size, the harder the problem is. Problems with a polynomial complexity expressed as $O(n^k)$ with $k \geq 1$ are considered possible to solve in a reasonable time frame. Logarithmic scaling $O(\log n)$ such as search algorithms in a large database is extremely efficient in dealing with larger problems, as the number of operations only doubles when going from a size n to n^2. On the opposite side, an exponential complexity $O(e^n)$ leads to problems impossible to solve in a reasonable time for sufficiently large system size n. Prime

number factorization complexity grows much faster than polynomial, almost exponential [1]. Algorithms can be classified in a simplified complexity picture shown in Figure 1.1 for the context of classical computing. Complexity classes are represented as expanding circles as $O(\log n) \Rightarrow O(n^k) \Rightarrow O(e^n)$. Solvable large problems are usually considered to be polynomial complexities or below.

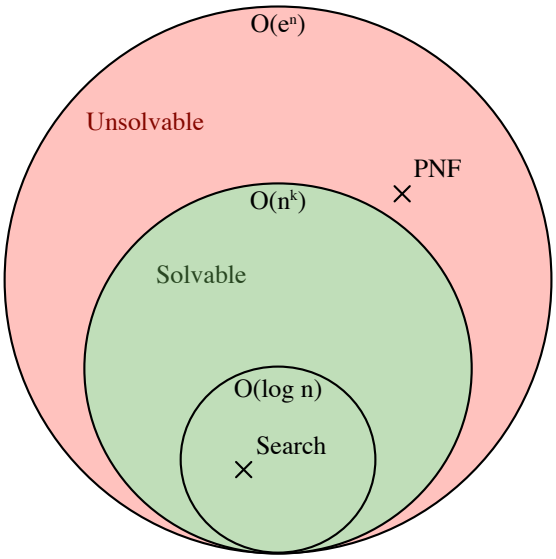


Figure 1.1 | Complexity-based algorithm classification
Classification of algorithms based on the complexity to solve the problem for a classical computer. Green area represents problems solvable in a reasonable time while the red section represents harder problems, with the crosses for typical examples of the search and Prime Number Factorization algorithms.

Figure 1.1 suggests that some problems are intrinsically hard to solve on a classical computer. To this day, it is not known whether such hard-to-solve problems exist. The answer to this question is related to the famous P versus NP problem, one of the four Millenium Prize Problems. Although not proven, it is widely believed that $P \neq NP$ for the classes P and NP of problems that can be solved in, respectively, polynomial and non-polynomial time, meaning that there are problems intrinsically hard to solve with a classical computer. Prime Number Factorization (PNF) is conjectured to be one of them, so much so that most cybersecurity systems rely on this problem being hard to solve on existing computing platforms. Many ongoing research tries to break the complexity floor of PNF, but without any success within the classical computing framework.

PNF with polynomial complexity was demonstrated by P. Shor in 1994 by switching the computing paradigm[1]. Shor's algorithm leverages the laws of quantum physics using the concept of quantum computing. Entanglement and superposition between quantum states brings an intrinsic parallelism that can be advantageously exploited to speed up some classical algorithms, known as quantum acceleration. For the first time, a new computing hardware paradigm offered the possibility to solve some of the hard problems.

II Quantum computing and its undertakings

Similarly to classical computing complexity, one can define a quantum complexity as the number of quantum operations needed to solve a problem. The same problems can be tackled by both classical and quantum computers, but might involve different complexities (as seen for prime number factorization). This is represented in Figure 1.2 where the complexity classes of classical computing (as seen in Figure 1.1) are combined with the complexity classes for quantum computing. In all generality, every complexity class

1. Strictly speaking, PNF complexity is sub-exponential, i.e. $O(k^n)$ with $k > 1$, in between polynomial and exponential complexities.

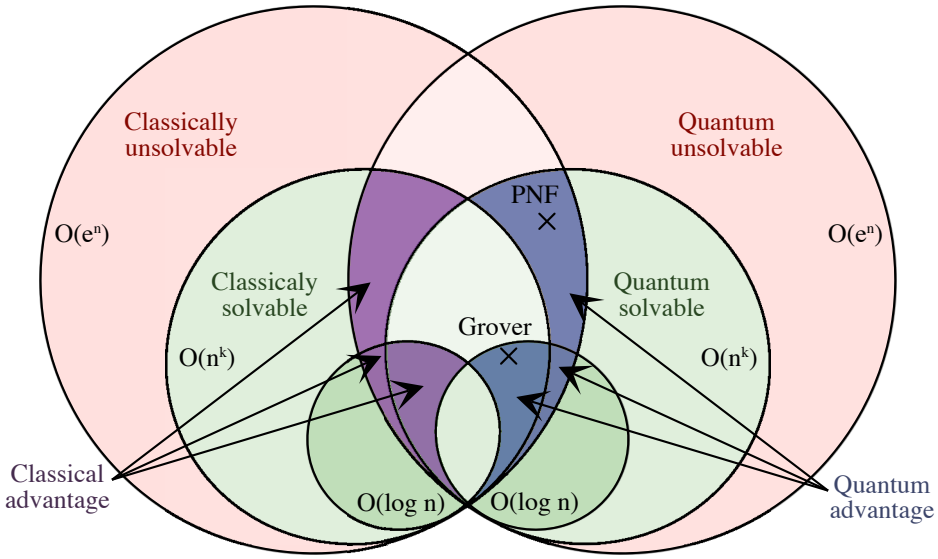


Figure 1.2 | Classical vs quantum complexity classes

Schematic intersection between classical and quantum computational complexities to solve problems, with the indicative positions of the PNF and Grover algorithms (crosses).

might intersect with some problems easier to solve with a classical computer than with a quantum one, and inversely. Shor's algorithm[1] shows quantum advantage from sub-exponential classical complexity to polynomial quantum complexity. Other quantum algorithms have been demonstrated to show a quantum advantage such as the Grover algorithm[2] for database search (see Figure 1.2).

Although classical algorithms can be implemented on a quantum processor, it is more favorable to run part of the algorithm on a classical computing unit due to the associated cost and technical difficulties in the early development of quantum computation. For this reason, quantum computing is very likely to co-exist with classical computing to efficiently solve problems. For this reason, we'll prefer talking about quantum processing unit or quantum processor instead of quantum computer in this thesis. A near to long term vision of computers is pictured in Figure 1.3a where the quantum Processing Unit is used as an accelerator to the central PU, at the same level of the graphical and tensor PUs.

III Anatomy of a quantum computer

A quantum processing unit is made of several functional layers to interface the quantum part to the classical world (see Figure 1.3b). Top layers handle the algorithmic part with e.g. compilation and error correction, while the bottom half concerns the hardware side. In this section, we describe what makes a quantum computer following a bottom-up approach. The fundamental layer of a quantum processor is the quantum chip, made of interacting quantum objects, named quantum bits (or qubits), as an analogy with the classical digital bits.

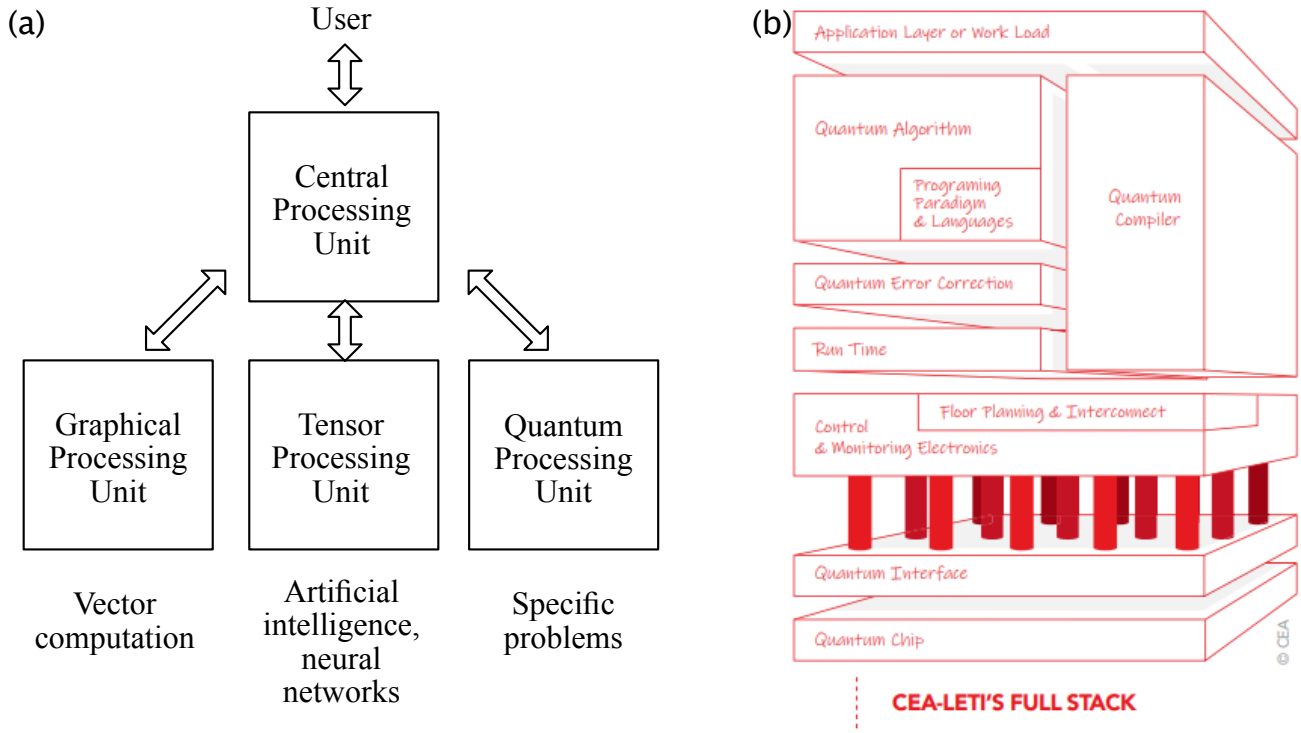


Figure 1.3 | Computing architecture

(a) Schematic representation of the distribution of tasks for the different kinds of processing units depending on the problem. The Quantum Processing Unit is likely to be used for problems requiring a high level of parallelism. (b) Proposition of the different functional layers found in a quantum processing unit (© CEA-LETI).

a) Elementary building blocks: quantum bits

The concept of quantum bits

Quantum bits, also known as qubits for short, are the smallest quantum object able to reveal superposition, a fundamental property of quantum mechanics allowing an object to be simultaneously in two different states. Qubits are made of only two quantum states, conventionally noted $|0\rangle$ and $|1\rangle$. In all generality, a quantum state, noted $|\Psi\rangle$, is defined by two parameters: the probability p_0 to be in state $|0\rangle$ (equivalently the probability p_1 to be in $|1\rangle$) and an angle ϕ.

$$|\Psi\rangle = \sqrt{p_0}\,|0\rangle + \sqrt{1-p_0}e^{i\phi}\,|1\rangle \tag{1.1}$$

The squared absolute value of the coefficient in front of $|0\rangle$ (respectively $|1\rangle$) gives the probability to find the qubit in $|0\rangle$ (resp. $|1\rangle$) when measured. Qubit states are represented as a point on a 3D sphere named the Bloch sphere with the pure states $|0\rangle$ and $|1\rangle$ at the poles. The angle ϕ defined in the xy-plane sets the longitude while p_0 appear as the projection onto the z-axis and sets the latitude.

Real-world hardware implementation of qubits

Pure two-level systems barely exist in experimental realizations. The quantum bit behavior is experimentally approximated from a more complicated quantum system with a large number of accessible quantum states. This is usually achieved by ensuring a sufficiently large energy separation between the two states chosen to implement the qubit state and the remaining higher-energy states.

Several hardware platforms exist for the implementation of qubits which can be divided in two main categories: optically-controlled qubits and electrically-controlled qubits. Optically controlled qubits require the use of laser beams to manipulate the quantum state such as trapped ions, atoms, or photons. This type of qubits achieves an astonishing level of accuracy, and is often used for sensing applications. On the other side, electrically-controlled qubits require the use of electrical signals (voltages, currents, magnetic or electric fields) such as superconducting qubits and charge/spin qubits. In this PhD thesis, we focus on the electrically-controlled qubits.

Superconducting qubits (shown in Figure 1.4a) are made of two superconducting islands linked by a very thin layer of dielectric, forming a Josephson junction through which discrete quantum Cooper-pair are tunneling. The application of a microwave voltage at the qubit frequency on one of the superconducting islands via capacitive coupling is used to manipulate the qubit quantum state. Replacing the single junction by two parallel junctions allows to tune the superconducting qubit properties such as its state frequencies by threading a magnetic field in the Josephson loop.

Charge and spin qubits (shown in Figure 1.4b) are implemented by trapping one or a few charges on a nanometric metallic-like island, named quantum dot. The island can be made of a highly-doped semiconductor or a metal but is more commonly implemented in a transistor-like structure. A voltage applied on a metallic gate deposited on an undoped semiconductor layer, traps charges at the surface of the semiconductor layer (shown in Figure 1.4b). The application of a magnetic field reveals the spin of the charge, used to realize spin qubits. The application of a voltage on the gate, or of a magnetic field allow to control the qubit parameters and affect its quantum state.

These two platforms of superconducting and spin qubits have similar control requirements with qubit frequencies corresponding to the energy separation of the two-level system typically in the range from 1 to 10 GHz. These qubits are highly sensitive to their environment and require to be cooled down at low temperatures below 1 K to avoid excessive thermal excitation from the ground state compromising the quantum state integrity. Quantum chips are placed at the coldest stage in dilution fridges to maintain such a low temperature environment, typically down to 10 mK.

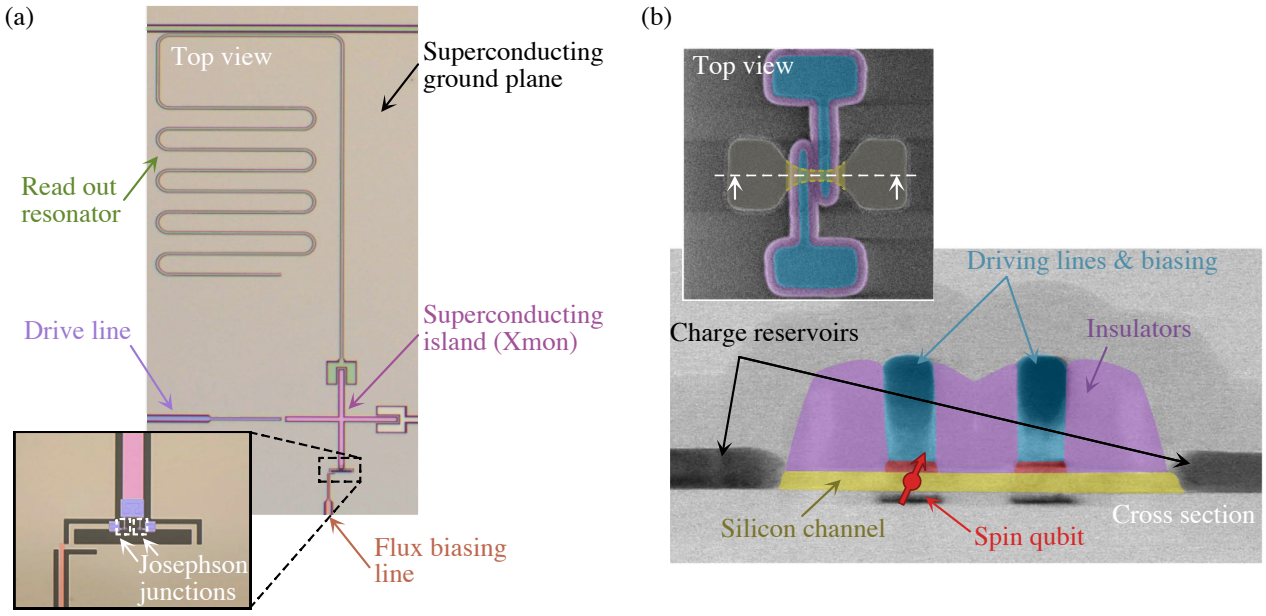


Figure 1.4 | Electrically-controlled superconducting and spin qubits

(a) Top view of a superconducting-metal chip containing a typical cross-shaped transmon device, also known as Xmon, with its capacitively-coupled drive and readout lines (blue and green) (adapted from[3]). The qubit is made of a superconducting island (pink) shorted to the superconducting ground plane with one or two Josephson junctions (white). In this example, two Josephson junctions are used to form a loop in which magnetic field is threaded via the flux line (red) to tune the qubit parameters. (b) Top view and cross section of a typical spin qubit embedded in a CMOS transistor-like structure (adapted from [4]). A single spin, coming from the charge reservoirs, is confined under the oxide of a gate by applying DC biasing voltages. A magnetic field is applied via an external coil or nearby flux line (not shown). Spin qubit state is read via a nearby resonator made of a normal or superconducting inductor (not shown).

In the next sections we describe the control and read-out of the quantum states considering both platforms as examples.

Quantum state determination

The measurement of the quantum state is generally accomplished by measuring the properties of an LC tank coupled to the qubit. Measuring a qubit refers to the action transforming the quantum state into a pure (classical) state (being either $|0\rangle$ or $|1\rangle$) following the probability distribution set by the qubit-state parameter p_0 (respectively $p_1 = 1 - p_0$) to get $|0\rangle$ (resp. $|1\rangle$).

Read-out of the quantum state is usually done at frequencies far-detuned from the qubit state transition frequency, typically 6-10 GHz or 0.1-1 GHz, and at very low amplitudes (typically -110 dBm) to prevent accidental alteration of the qubit while determining its state. Combined with the requirement of a high signal to noise ratio and fast read-out to accurately measure the qubit, read-out requires an ultra-low noise amplification chain with an input noise temperature below 1 K. Such low level of input noise is commonly obtained with quantum-limited amplifiers with an input noise temperature below 200 mK such as Josephson Parametric Amplifiers (JPA), or Traveling-Wave Parametric Amplifiers (TWPA), followed by cryogenic low-noise amplifiers with a few Kelvin input noise.

Signal requirements to tune and manipulate quantum bits

Qubit properties are usually tuned in-situ to the desired value for the point of operation by the application of DC signals to compensate for non-ideal fabrication. For example, a DC magnetic field will modify the qubit frequency, or a DC voltage can change the confinement potential in quantum dots.

Qubit manipulation refers to the ability to move a quantum state anywhere in the Bloch sphere by the application of tailored high-frequency signals. Conventionally, qubits are manipulated using harmonic pulses tuned to be on-resonant with the qubit frequency f_q (defined by the energy spacing between $|0\rangle$ and $|1\rangle$). Pulse length is typically between 10 to 100 times $1/f_q$ (see Figure 1.5), usually set by a trade-off between coherence loss and non-coherent errors. The phase of the signal determines the axis of rotation in the equator while its amplitude determines the angle of rotation. Length of the pulse could be used to modify the angle of rotation, however the sampling rate of off-the-shelf AWGs is usually insufficient to accurately set the angle. Only a subset of all possible rotations are sufficient to move a quantum state everywhere on the Bloch sphere. Rotations are usually denoted as X_θ or Y_θ where X and Y denotes the rotation axis while θ is the rotation angle. Typical gate operations comprise $X_\pi, Y_\pi, X_{\pi/2}, Y_{\pi/2}, X_{-\pi/2}, Y_{\pi/8}, \cdots$.

The accuracy of quantum gates is usually evaluated with a quantity named fidelity F. Fidelity is usually expressed in % and evaluates how close the implemented rotation is to the ideal rotation, with a fidelity of 100% corresponding to a perfect gate. Two main definitions of the fidelity exist: the Pauli fidelity (also known as process or entanglement fidelity) and the average fidelity. The former measures how well entanglement with other systems is preserved by the quantum gate[5] and is experimentally measured through Quantum Process Tomography (QPT)[6]. The latter measures how well the gate performs as expected throughout the entire initial state space[7] and is experimentally measured with randomized benchmarking[8] or cross-entropy benchmarking[9]. The two obtained fidelity numbers differ only slighlty e.g. 99.90% process fidelity becomes 99.93% average fidelity[10]. Measuring the average gate fidelity has become the main tool to evaluate quantum gates as the benchmarking methods require a lower number of measurements than QPT and are insensitive to errors in initial state preparation and final state readout. Fidelity is an important metrics to implement quantum algorithms and typical aimed levels are above 99.9%.

Single qubit gates are insufficient to implement quantum algorithms and multi-qubit gates are required. Multi-qubit gates generate entanglement between qubits, a key ingredient for quantum acceleration.

b) Interacting qubits

Qubits in a quantum chip are usually placed close to each other. Qubits are coupled to their neighbors to allow multi-qubit gates and qubit-qubit entanglement. Different strategies are used with fixed or tunable coupling.

Coupling quantum bits

Coupling between qubits is usually achieved via electric (capacitive) or magnetic interaction from e.g. capacitive coupling of superconducting qubits or spin-spin interaction of spin qubits. A mediator can be placed in between the coupled qubits to enhance or deactivate coupling set by a voltage or current (magnetic flux). Spin qubits confined in quantum dots are separated by a gate, named plunger, acting as the interaction mediator between the two qubits. The plunger gate determines the electrostatic potential that screens or enhances the interaction. For superconducting qubits, tunable couplers are made with flux-tunable Josephson junctions to allow resonant coupling.

Two-qubit gates rely on qubit-qubit coupling to generate entanglement and are a crucial element of quantum computing.

Two-qubit quantum gates

Multi-qubit gates generate entanglement between several qubits, one of the key components for exponential acceleration. Two-qubit gates are most commonly used for their simplicity and are sufficient to implement all quantum algorithms. In a multiqubit configuration, the gates usually perfom rotations of the quantum state of the target qubit depending on the state of a control qubit, such as control-Z (C-Z) or control-X (C-X, or CNOT) gates performing a rotation along Z or X of the target qubit. As for single qubit control, different implementations of two qubit gates exist and can be classified in these two categories: baseband pulses, and RF pulses.

Baseband pulses turn on and off the interaction between two qubits for a given time by e.g. tuning one qubit frequency to be resonant with a second one, or via tunable couplers. RF pulses can be used to drive a two qubit transition (such as $|10\rangle \rightarrow |01\rangle$) thanks to a non-zero qubit-qubit coupling. Alternatively, RF pulses have been widely used in cross-resonance (CR) scheme, where the first qubit is driven at the second qubit frequency with a constant qubit-qubit coupling. Effectively, the second qubit (usually named target) will feel a different-amplitude drive depending on the state of the first qubit (control), resulting in a conditional quantum gate.

c) The challenge of up-scaling the qubit number

As shown in Figure 1.5, the quantum chip requires many signals to perform computation: from DC electric signals to baseband and radio-frequency shaped pulses.

The need to control and measure hundreds to thousands qubits requires a careful system-level engineering of both the quantum part (qubit type, number of qubits, qubit connectivity, qubit-qubit couplings,...) and the control-electronics part (e.g. AWG, mixing, amplification, real-time feedback,...). Recent achievements with 50 qubits such as [11] require custom-made room-temperature electronics with state-of-the-art performance to run experiments occupying several large racks.

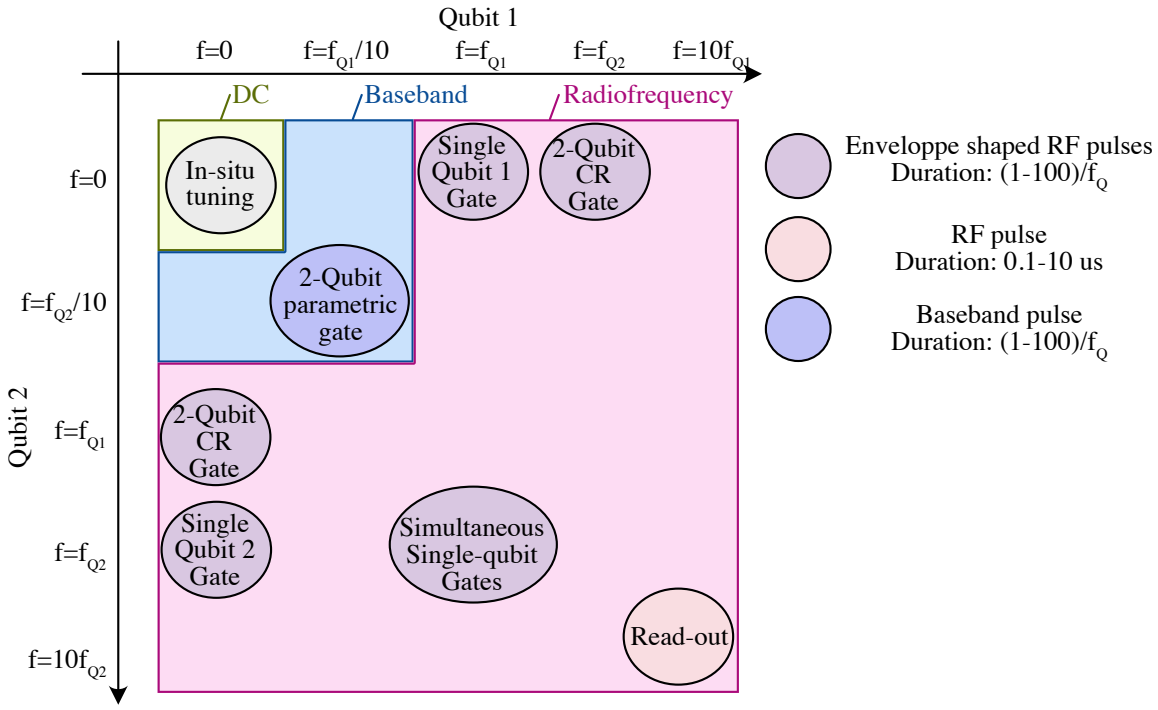


Figure 1.5 | Signal requirements for quantum computing

Representation of the signal requirements to perform universal quantum computation based on two qubits Q1 and Q2, showing the frequency regions for DC, baseband, and RF signals. The qubit frequency f_Q typically ranges from 1 to 20 GHz, leading to a few tens of nanoseconds pulses for a single operation. Read-out duration only depends on the resonator linewidth and integration time, not on the qubit frequency.

Signals are carried to and from the quantum chip resting at the colder stage of a dilution fridge through coaxial cables. Large attenuation on each cable, distributed at different fridge stage, is primordial to lower the input noise and thermal radiation. Moreover, 20cm-long resistive cables are required between different temperature stages above 10 K to limit the heat conduction. Below 10 K, superconducting coaxial cables are ideal with low electrical loss, low heat conduction, and recently-demonstrated dense assembly. The current approach of linearly scaling the electronics with the number of qubits (required today to compensate for imperfect qubit fabrication) approaches a ceiling with fridges full of cables for only one experiment. A strategy change for control electronics of a large number of qubits is necessary to continue the up-scaling.

Control electrical requirements can be lowered by improving the quantum part. Such methods includes the use of frequency multiplexing for control and read-out, large-scale architecture with repeated pattern, lower-frequency qubits, autonomously-corrected qubits, or bosonic encoded states. Although these are very interesting to reduce the number of control signals, it significantly adds complexity on the quantum system, and requires long development cycles.

A different approach started to appear in 2017, and consists in moving part of the electronics to the cryogenic stage of the dilution fridge below 10 K to reduce the number of connections from room temperature. A large number of connections to the qubit stage can be handled by short superconducting cables, opening the prospect of a large qubit array with linear scaling of the electronics with the number of qubits for mid-term applications.

IV Cryogenic electronics

Placing part of the control electronics at a lower temperature stage solves the non-superconducting cable crowding issue at temperature above 10 K. Cryogenic electronics has long been used and developed for space applications with the development of low-noise amplifiers down to 4.2 K. However, the space-focusing cryogenic electronics mostly aim low-frequency (sub-GHz) circuits and GHz-range Low-Noise Amplifier. The requirement of quantum computing for high-frequency circuits other than LNAs such as radiofrequency pulse generators and demodulators called for a new field of research.

The semiconductor community have been the first one to explore the idea of cryogenic integrated circuits for quantum computing. The first demonstrations of integrated circuits operating at 4.2 K and below for quantum computing applications used the CMOS Silicon-On-Sapphire (SOS) technology, known for its robustness in harsh environments such as radiations and high temperatures. A fast pulse generator in 2008[12] and a Digital-to-Analog Converter in 2016[13], targeting dynamic voltage biasing of a semiconductor qubit, were shown to operate at 4.2 K and below, opening the prospect for the co-integration of semiconductor electronics with quantum devices. In 2016, the first co-integrations of a quantum device with low-frequency and high-frequency CMOS ICs were successfully demonstrated. A cryogenic Transimpedance Amplifier[14] made of a commercial CMOS bulk 0.35 µm were used to measure charge stability diagrams, the "identity card" of a quantum dot. A 0.05-1.2 GHz voltage-controlled ring oscillator[15] made of a CMOS SOI nanowire technology was hooked up to a double quantum dot to induce a discrete charge flow through the dots at each clock cycle, known as quantized charge pumping.

The year 2017 marked a turning point in the involvement of the electrical engineering community in cryogenic electronics for quantum applications. IEEE conference papers[16, 17] advocating in favor of cryogenic electronics to enable fault-tolerant quantum computing led to a fertile ground for funding and new research in electrical engineering. Shortly, the first demonstration of qubit control with cryogenic circuits[18] was showcased at the newly-created quantum computing session at the 2019 International Solid-State Circuits Conference (ISSCC). Accompanied by Google's demonstration of quantum supremacy[11]

opening the path towards large-scale quantum computing, the field boomed with the presence of large companies such as Google, IBM, and Intel and academia research centers such as TU Delft, and CEA-LETI Grenoble, UNSW, University of Toronto or University of Dublin, and others.

New technologies had to be explored to realize high-frequency System-on-Chip (SoC) operating at cryogenic temperatures. The initial lack of transistor modeling at deep cryogenic temperatures (4.2 K and below) impacts the design choices and many groups rely on room temperature compact models. Significant efforts have been devoted to the characterization of transistors at 4.2 K and below for bulk technologies (180[19, 20], 160[21], and 40[21, 22, 23], and 28 nm[24]), FDSOI technologies (28[25, 26, 27, 28, 29, 30], and 22 nm[31, 32]), and FinFET technologies (64[33], 35[34], 16[35] and 14 nm[33]). Typical CMOS technology nodes below 40 nm show interesting transistor characteristics with minimal faulty cryogenic effects such as kink effects or hysteresis while remaining commercial technologies with quick fabrication cycles.

Shortly, large high-frequency SoC showcased the success of modern CMOS technologies to be used at cryogenic temperature for quantum computing applications. Successive yearly ISSCCs have been the place for new key demonstrations: from sub-Kelvin integration[36], to qubit control[18, 37, 38, 39], to qubit read-out[38, 40, 41]. Recently, the IBM teams demonstrated the first cryogenic qubit controller to achieve a state-of-the-art single-transmon gate fidelity of 99.9 % (presentation slides of [39]).

Despite all these milestones, the prospect for large-scale quantum computing processor is still far away. Focusing on the cryogenic IC aspect, none of the demonstrated ICs qualifies to drive more than a few hundreds qubits at best with present cooling capabilities. Lowering the power consumption per driven qubit in cryogenic electronics while maintaining the stringent requirements for quantum computing is a big challenge and will require new innovations from circuit architecture to the use of novel technologies.

V Context of this Ph.D. work

Rewinding back to 2017, the year this Ph.D. work started, the development of cryogenic electronics for quantum computing applications was in its infancy. To tackle the fundamentally interdisciplinarity of this research field, involving electrical engineering, quantum mechanics, and cryogenics, the French Alternative Energies and Atomic Energy Commission (CEA) in Grenoble united the efforts of the fundamental research institute - the Interdisciplinary Research Institute of Grenoble (IRIG) - and of the technology research institute - the Laboratoire d'Électronique et de Technologie de l'Information (LETI). The history, knowledge, and experience of both institutes influenced the overall direction and target of this Ph.D. work.

We naturally decided to study the STMicroelectronics CMOS FD-SOI technology as CEA-LETI has been a pioneer in developing this technology over the last 20 years with both Soitec and STMicroelectronics being part of the Grenoble-Crolles ecosystem. The acquired in-house knowledge on this technology from fabrication to modeling and circuit design at room temperatures will accelerate the development of cryogenic electronics. Moreover, the access to the SOI back gate was promising to compensate the increase of threshold voltages at cryogenic temperatures, unlike any other technology, making it particularly relevant for high-frequency circuits that could be affected by high threshold voltages.

On the fundamental side, the CEA-IRIG has a long-standing experience in a particular type of qubit: the spin qubit in CMOS quantum dots. The team is familiar with radio-frequency off-the-shelf equipment as well as with the use of cryogenic instrumentation from helium dip-stick to custom-made dilution fridges. In 2016, the CEA-IRIG demonstrated the first spin qubit made on an 300mm wafer from a slightly modified CMOS commercial technology[4], opening the path for large-scale uniform arrays.

VI Resume

This Ph.D. dissertation explores the use of advanced CMOS technologies for quantum computing applications. The first two chapters focus on the changes in the electrical characteristics of basic CMOS elements when cooling down to sub-Kelvin temperatures.

Chapter 2 focuses on important single transistor quantities for circuit design such as the threshold voltage, the transconductance, or the ON current. Transistors with vastly varying lengths and widths and under various biasing regime (including forward body biasing) are studied from room to cryogenic temperatures thanks to a novel approach which drastically speeds up the measurement time. To this end, single transistors are embedded in a multiplexing matrix made of on-chip switches and decoding circuitry leading to a single cooldown to characterize more than 1,000 devices.

Chapter 3 concentrates on how up-to-10's transistors assembled in larger circuits behave at cryogenic temperatures compared to room temperature. We approach both digital and analog circuits in a unified way focusing on particular quantities defined across several circuits and extracted at both temperatures. For example, the evolution of the input voltage range in which the circuit is functional is compared between voltage-controlled oscillators, level-shifters, and operational amplifiers. Similarly, digital circuits such as ring oscillators are set side by side with analog circuits such as amplifiers to track the energy efficiency with lowering temperature. Although, circuit performance mostly degrades at low temperatures, we show that the high Forward Body-Biasing (FBB) permitted by the FD-SOI technology allows to recover better performance.

The last two main chapters center on the application of cryogenic circuits for the study of quantum objects. Following the long history of the CEA Grenoble with the development and study of quantum dots in CMOS technology, we applied the developed cryogenic circuits to measurements on single and double quantum-dot structures.

Chapter 4 presents two applications of cryogenic electronics to quantum dots based on current read-out with a low-power cryogenic transimpedance amplifier. First, a single quantum dot at is co-integrated via wire-bonding to the cryogenic TIA. We demonstrate the successful assembly down to 4.2 K by measuring typical charge stability diagrams, commonly used to derive important dot energy scales such as the charging energy. Then, we integrate a double-quantum dot on the same FD-SOI chip with the TIA allowing to place nearby an excitation circuit generating tunable GHz square voltage signals. The cryogenic TIA probes the dot response to the excitation. The entire circuit operates at the low temperature of 110 mK in a dilution fridge thanks to its low power dissipation.

Chapter 5 addresses the capacitive read-out of quantum dots, recently favored to the current read-out for its increased speed and sensitivity. We replaced the typically used passive inductor to form the read-out LC resonator with an active inductor made of transistors, offering in-situ tunability. The much smaller footprint of the active inductor would be ideal for scalability. Having all the read-out circuitry close to the measured quantum dot, we replace the measurement of the scattering parameters (usually S_{12} in transmission or S_{11} in an RF reflectometry setup) by a measurement of the LC tank impedance to have a sensitive probe to capacitive changes. We validate all the functionalities of the read-out circuit with the active inductor at 4.2 K, such as in-situ tuning of resonance frequency and detection sensitivity or device multiplexing. As a demonstration, capacitive signatures of quantum effects in nanometric CMOS transistors are revealed by the read-out circuitry at cryogenic temperatures as a demonstration.

[FR] Introduction

Le développement des processeurs numériques est sans aucun doute l'une des avancées technologiques les plus marquantes du siècle dernier. Le large éventail d'applications du numérique a rapidement envahi de nombreux aspects de notre société moderne, de la recherche fondamentale à la résolution de problèmes quotidiens.

L'informatique numérique, que nous appellerons informatique classique dans cette thèse, est le type de calcul dominant. L'informatique classique utilise des dispositifs au comportement d'interrupteur de plus en plus petits, évoluant des anciens tubes à vide décimétriques aux modernes transistors nanométriques. La capacité qu'ont les dispositifs de type interrupteur à contrôler d'autres dispositifs de même nature permet leur assemblage a différents niveaux d'abstractions : des portes numériques faites de quelques-uns jusqu'aux unités de calcul contenant jusqu'à 10^{10} éléments. Ces complexes unités de calcul allant du processeur graphique jusqu'au processeur tensoriel font le succès de l'informatique d'aujourd'hui, capable de conduire une voiture, de faire un meilleur diagnostic visuel médical que les humains et de vaincre n'importe quel joueur à des jeux extrêmement complexes. Bien que l'informatique classique soit capable de résoudre des problèmes de plus en plus complexes en améliorant le matériel, certains problèmes restent extrêmement difficiles à résoudre, nécessitant un temps et/ou des ressources incommensurables.

Un exemple bien connu des limites du calcul classique concerne la factorisation d'un grand nombre en une multiplication (unique) de 2 nombres premiers ou plus. Bien qu'il soit trivial pour un ordinateur de décomposer un nombre relativement petit tel que 3233 en produit de nombres premiers 61×53, cette tache devient exponentiellement plus couteuse à mesure que le nombre factorisé augmente, et va même au-delà des capacités des meilleurs super-ordinateurs classiques. Cette complexité hors-de-portée des ordinateurs d'aujourd'hui est au cœur des routines de chiffrement modernes telles que le code de Rivest-Shamir-Adleman (RSA). L'utilisation de grands nombres généralement compris entre 3×10^{616} (2048 bits) et 1×10^{1233} (4096 bits) empêche tout ordinateur classique de trouver la factorisation en nombres premiers. De ce fait, la factorisation des nombres premiers (FNP) est largement utilisée pour des applications de cybersécurité et d'encryptions des donnes. De nombreuses recherches tentent de résoudre ce problème de manière plus optimale mais sans succès dans le cadre informatique classique. Cependant, en 1994, P. Shor démontre que la factorisation en nombres premiers peut être faite de manière extrêmement efficace en changeant le paradigme de calcul. L'algorithme de Shor exploite les lois de la physique quantique dans le cadre de ce qui est appelé aujourd'hui l'informatique quantique. L'intrication et la superposition entre états quantiques apportent un parallélisme intrinsèque qui peut être avantageusement exploité pour accélérer certains calculs classiques, connu sous le nom d'accélération quantique. Pour la première fois, une nouvelle plateforme offre la possibilité de résoudre certains des problèmes les plus difficiles. Bien qu'à ce jour aucunes des unités de calcul quantique existantes ne soit capable de telles prouesses, plusieurs démonstrations clefs ont entre-ouvert une telle possibilité.

a) L'informatique quantique

Une unité de traitement quantique est composée de plusieurs couches fonctionnelles pour interfacer la partie quantique avec le monde classique (voir Figure 1.3b). Les couches supérieures gèrent la partie algorithmique avec, par exemple, la compilation et la correction d'erreurs, tandis que la moitié inférieure concerne le côté matériel. Dans cette section, on aborde la partie matérielle, en lien avec les travaux effectués lors de cette thèse.

Le concept de bits quantiques

Au cœur d'un ordinateur quantique se trouve la puce quantique, constituée d'objets quantiques en interaction, appelés bits quantiques (ou qubits), par analogie avec les bits numériques de l'informatique classique. Ces bits quantiques peuvent prendre une infinité de valeurs entre les deux états d'énergie fondamentaux $|0\rangle$ et $|1\rangle$.

Les systèmes à seulement deux niveaux existent très rarement dans la nature. Le comportement des bits quantiques est approximé expérimentalement à partir d'un système quantique plus compliqué avec un grand nombre d'états quantiques accessibles. Ceci est généralement réalisé en assurant une séparation énergétique suffisante entre les deux états choisis pour implémenter le qubit et les états restants à plus hautes énergies.

Les qubits supraconducteurs (illustrés à la Figure 1.4a) sont constitués de deux îlots supraconducteurs reliés par une très fine couche de diélectrique, formant une jonction Josephson à travers laquelle des paires de Cooper quantiques discrètes tunnelles. L'application d'une tension radiofréquence résonante avec le qubit sur l'un des îlots supraconducteurs via un couplage capacitif permet de manipuler l'état quantique du qubit. Le remplacement de la jonction par deux jonctions en parallèles permet d'ajuster les propriétés du qubit telles que les fréquences de transitions grâce à l'application d'un champ magnétique.

Les qubits de charge et de spin (illustrés à la Figure 1.4b) sont implémentés en piégeant une ou plusieurs charges sur un îlot nanométrique, appelé puit quantique. L'îlot peut être constitué d'un semi-conducteur fortement dopé ou d'un métal, mais est plus couramment réalisé au sein d'une structure de type transistor. Une tension appliquée sur la grille métallique déposée sur une couche semi-conductrice non dopée permet de piéger des charges à la surface du semi-conducteur (représentée sur la Figure 1.4b). L'application d'un champ magnétique révèle le spin de la charge, utilisé pour réaliser des qubits de spin. L'application d'une tension sur la grille, ou d'un champ magnétique permet de contrôler les paramètres du qubit et de manipuler son état quantique.

Ces deux plates-formes de qubits, supraconducteurs et de spin, ont des exigences de contrôle similaires avec des fréquences de qubit généralement dans la plage de 1 à 10 GHz. Ces qubits sont très sensibles à leur environnement et nécessitent d'être refroidis à des températures inférieures à 1 K afin de conserver l'intégrité de l'état quantique. Les puces quantiques sont placées à l'étage le plus froid dans des réfrigérateurs à dilution généralement allant jusqu'à 10 mK.

Le défi de la mise à l'échelle du nombre de qubits

La puce quantique nécessite de nombreux signaux électriques, continus et radiofréquences, qui doivent être acheminés à chaque qubit afin d'effectuer des calculs quantiques. La nécessité de contrôler et de mesurer des centaines voire des milliers de qubits requiert une attention particulière lors de la conception d'un processeur quantique. En effet, les réalisations récentes avec seulement 50 qubits telles que [11] nécessitent une électronique sur mesure occupant déjà plusieurs grands racks à température ambiante et requiert des centaines de câbles coaxiaux entrant le réfrigérateur a dilution, posant de fortes contraintes thermiques. L'approche actuelle consistant à mettre à l'échelle linéairement l'électronique avec le nombre de qubits (nécessaire aujourd'hui pour compenser la fabrication imparfaite de qubits) s'approche d'un plafond avec des réfrigérateurs remplis de câbles, affectant fortement les capacités de refroidissement. Un changement de stratégie pour l'électronique de contrôle d'un grand nombre de qubits est nécessaire pour poursuivre l'amélioration des plateformes de calculs quantiques existantes.

Une nouvelle approche a commencé à apparaître en 2017, et consiste à placer une partie de l'électronique de control à un étage cryogénique du réfrigérateur à dilution en dessous de 10 K afin de réduire le nombre de câbles thermo-conducteurs entre les températures ambiantes et celles cryogéniques. Un grand nombre

de connexion entre l'électronique cryogénique et la puce quantique peut être réalisé avec des câbles supraconducteurs courts, thermiquement isolants, ouvrant la voie vers le control d'un grand nombre de qubits visant des applications à moyens termes.

b) Électronique cryogénique

Placer une partie de l'électronique de commande à un étage de température plus bas résout le problème d'encombrement des câbles non-supraconducteurs aux températures supérieures à 10 K. L'électronique cryogénique a longtemps été utilisée et développée pour des applications spatiales avec le développement d'amplificateurs à bas bruit jusqu'à 4.2 K. Cependant, l'électronique cryogénique à but spatial vise principalement des circuits basses fréquences et l'amplification de signaux à bas bruit dans la gamme GHz. Le besoin de l'informatique quantique pour des circuits hautes fréquence autres que les amplificateurs bas bruits tels que les générateurs d'impulsions radiofréquences et les démodulateurs requiert de nouveaux développements.

La communauté des qubits semi-conducteurs a été la première à explorer l'idée de circuits intégrés cryogéniques pour l'informatique quantique. Les premières démonstrations de circuits intégrés fonctionnant à 4.2 K pour des applications d'informatique quantique utilisaient les technologies CMOS Silicium sur saphir (SOS), connues pour leur robustesse dans des environnements difficiles tels que les radiations et les hautes températures. Un générateur d'impulsions rapides en 2008[12] et un convertisseur numérique-analogique en 2016[13] ciblant la polarisation dynamique en tension des qubits semi-conducteurs ont été démontrés à 4.2 K, ouvrant la voie vers la cointégration de l'électronique avec des dispositifs quantiques. En 2016, les premières cointégrations d'un dispositif quantique avec des circuits intégrés CMOS basses et hautes fréquences ont été démontrées avec succès. Un amplificateur cryogénique à transimpédance[14] réalisé avec une technologie commerciale CMOS 0.35 µm a permis de mesurer les diagrammes de stabilité de charge, la « carte d'identité »des puits quantiques. Un oscillateur en anneau commandé en tension de 0,05 à 1,2 GHz[15] fait dans une technologie de nanofils CMOS SOI a été connecté à un double puit quantique afin de stimuler le transport de quelques charges à chaque oscillation, connu sous le nom de pompage de charges quantifiées.

L'année 2017 a marqué un tournant dans l'implication de la communauté des ingénieurs dans l'électronique cryogénique pour les applications quantiques. Des articles de conférence publiés au sein de l'IEEE[16, 17] plaidant en faveur de l'électronique cryogénique pour les calculs quantiques ont conduit à un terrain fertile pour des nouveaux financements et des nouvelles recherches en génie électronique. Peu de temps après, la première démonstration du contrôle d'un qubit avec un circuit cryogénique[18] a été présentée lors de la session d'informatique quantique nouvellement créé au sein de l'International Solid-State Circuits Conference (ISSCC) en 2019. Accompagné par la démonstration de la suprématie quantique de Google[11] ouvrant la voie à l'informatique quantique à grande échelle, le domaine a explosé avec la présence de grandes entreprises telles que Google, IBM et Intel et de centres de recherche universitaires tels que TU Delft, CEA-LETI Grenoble, UNSW, Université de Toronto ou l'université de Dublin, et autres.

De nouvelles technologies ont dû être explorées pour réaliser des systèmes sur puce (SoC) hautes fréquences fonctionnant à des températures cryogéniques. Cependant, le manque de modélisation des transistors à des températures cryogéniques (4.2 K et en dessous) impacte les choix de conception. De fait, des efforts importants ont été consacrés à la caractérisation des transistors à des températures en deçà de 4.2 K réalisés avec les technologies bulk (180[19, 20], 160[21] et 40[21, 22, 23] et 28 nm[24]), les technologies FD-SOI (28 [25, 26, 27, 28, 29, 30] et 22 nm [31, 32]), et les technologies FinFET (64 [33], 35 [34], 16 [35] et 14 nm [33]). Les nœuds de technologie CMOS en dessous de 40 nm présentent des caractéristiques de transistor intéressantes, sans non-idéalités dues à des effets cryogéniques tels que des effets de kink ou d'hystérésis tout en restant des technologies commerciales avec des cycles de fabrication rapides.

Peu de temps après, de nombreux SoC hautes fréquences ont vu le jour et ont fait le succès des technologies CMOS modernes pour des applications en informatique quantique. Chaque années, ISSCC a été le théâtre de nouvelles démonstrations clefs : de l'intégration avec des dispositifs quantique en dessous de 1 K[36], au contrôle des qubits [18, 37, 38, 39] jusqu'à la lecture des qubits[38, 40, 41].

Malgré tous ces jalons, la perspective d'un processeur informatique quantique à grande échelle est encore loin. Aucuns de ces circuits intégrés n'est adapté pour le contrôle de plus de quelques centaines de qubits en vue des capacités actuelles de refroidissement des cryostats. Réduire la dissipation d'énergie par qubit des circuits cryogéniques tout en maintenant la précision des signaux générés est un défi de taille et nécessitera de nouvelles innovations, de l'architecture des circuits à l'utilisation de nouvelles technologies.

c) Contexte de ce doctorat

En 2017, les travaux de ce doctorat ont commencé lorsque l'électronique cryogénique pour les applications en informatique quantique en était à ses balbutiements. Afin d'aborder l'interdisciplinarité intrinsèque de ce domaine de recherche, associant génie électronique, mécanique quantique, et la cryogénie, le Commissariat aux énergies alternatives et à l'énergie atomique (CEA) de Grenoble a uni les efforts de l'institut de recherche fondamentale - l'Institut de recherche interdisciplinaire de Grenoble (IRIG) - et de l'institut de recherche technologique - le Laboratoire d'Électronique et de Technologie de l'Information (LETI). L'histoire, les connaissances et l'expérience des deux instituts ont influencé les décisions et la direction de ce doctorat.

Naturellement, nous avons décidé d'étudier la technologie CMOS FD-SOI de STMicroelectronics puisque le CEA-LETI a été pionnier dans le développement de cette technologie au cours des 20 dernières années, Soitec et STMicroelectronics faisant partie de l'écosystème Grenoble-Crolles. Les connaissances acquises sur cette technologie, de la fabrication à la modélisation et à la conception de circuits à température ambiante promettaient d'accélérer le développement de l'électronique cryogénique. De plus, l'accès à la grille arrière SOI se présentait comme une opportunité unique de compenser l'augmentation des tensions de seuil à froid, particulièrement pertinent pour les circuits hautes fréquences qui sont affectés par des tensions de seuil élevées.

Sur le plan fondamental, le CEA-IRIG a une longue expérience sur un type particulier de qubit : les spin qubits contenus dans des puits quantiques réalisés en technologie CMOS. L'équipe est familière avec les équipements standards de mesure radiofréquence ainsi qu'avec l'utilisation des cryostats allant du dip-stick plongé dans l'hélium liquide aux réfrigérateurs à dilution faits sur mesure. En 2016, le CEA-IRIG a démontré le premier qubit de spin réalisé sur un wafer de 300 mm à partir d'une technologie commerciale CMOS légèrement modifiée [4], ouvrant la voie à la réalisation d'un grand nombre de qubits identiques.

d) Résumé

Ce doctorat explore l'utilisation des technologies CMOS avancées pour les applications en informatique quantique. Les deux premiers chapitres se concentrent sur les changements des caractéristiques électriques des éléments CMOS de base lors du refroidissement à des températures cryogénique jusqu'à 100 mK.

Le chapitre 2 se concentre sur les quantités importantes des transistors individuels pour la conception de circuits, telles que la tension de seuil, la transconductance ou le courant drain-source. Des transistors de longueurs et de largeurs variables et sous divers régimes de conductions (y compris avec la tension de grille arrière) sont étudiés des températures ambiantes jusqu'aux températures cryogéniques grâce à une nouvelle approche qui accélère considérablement le temps de mesure. À cette fin, des transistors de test sont intégrés au sein d'une matrice de multiplexage avec un circuit de décodage afin de caractériser plus de 1000 dispositifs en un seul cycle de refroidissement.

Le chapitre 3 se concentre sur le comportement de petits circuits constitués de quelques transistors à des températures cryogéniques par rapport au fonctionnement à température ambiante. Nous abordons les circuits numériques et analogiques sous une même approche en nous concentrant sur des quantités communes définies pour plusieurs circuits et extraites à plusieurs températures. Par exemple, l'évolution de la plage de tension d'entrée pour laquelle le circuit est fonctionnel est comparée entre des oscillateurs commandés en tension, des convertisseurs de niveau et des amplificateurs opérationnels. De même, l'efficacité énergétique des circuits numériques tels que des oscillateurs en anneau sont comparés avec celle des circuits analogiques tels que des amplificateurs. Bien que les performances des circuits se dégradent à basse température, nous montrons que l'utilisation de la grille arrière permis par la technologie FD-SOI permet de retrouver des performances comparables, si ce n'est meilleures, qu'aux températures ambiantes.

Les deux derniers chapitres portent sur l'application des circuits cryogéniques à l'étude des dispositifs quantiques. Suite à la longue histoire du CEA Grenoble avec le développement et l'étude des puits quantiques en technologie CMOS, nous avons mesuré des structures à simple et double puits quantiques avec les circuits cryogéniques développés.

Le chapitre 4 présente deux applications de l'électronique cryogénique à la mesure du transport au sein de puits quantiques avec un amplificateur à transimpédance (TIA) cryogénique de faible puissance. Tout d'abord, un unique puit quantique est cointégré par wire-bonds au TIA cryogénique. Nous démontrons le succès de l'assemblage jusqu'à 4.2 K en mesurant les diagrammes de stabilité de charge, couramment utilisés pour déduire les différentes énergies caractéristiques des puits quantiques. Ensuite, nous intégrons un dispositif à double puits quantiques sur la même puce FD-SOI avec le TIA, permettant aussi de placer un circuit d'excitation générant des signaux en tension de l'ordre du GHz à proximité. Le TIA cryogénique est utilisé pour mesurer la réponse du dispositif quantique à l'excitation haute-fréquence. L'ensemble du circuit fonctionne à la basse température de 110 mK dans un réfrigérateur à dilution grâce à sa faible dissipation de puissance.

Le chapitre 5 aborde la lecture capacitive des puits quantiques, récemment préférée à la lecture en courant pour sa vitesse et sa sensibilité accrues. Nous avons remplacé l'encombrante inductance passive utilisée pour former le résonateur de lecture par une inductance active constituée de transistors et de condensateurs, offrant la possibilité de varier la valeur d'inductance en situation pour une occupation en silicium réduite. En tirant profit des circuits de lecture au plus proche du puit quantique mesuré, nous remplaçons la mesure des paramètres de diffusion (généralement S_{12} en transmission ou S_{11} dans une configuration de réflectométrie radiofréquence) par une mesure de l'impédance du résonateur de lecture afin de détecter des petits changements capacitifs. Nous validons toutes les fonctionnalités du circuit de lecture avec l'inductance active à 4.2 K, telles que le réglage in-situ de la fréquence de résonance et de la sensibilité de détection ou le multiplexage des dispositifs mesurés. À titre de démonstration, des signes d'effets quantiques dans les transistors CMOS nanométriques sont révélées par le circuit de lecture à des températures cryogéniques.

Bibliography

[1] Peter W. Shor. **Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer**. *arXiv:quant-ph/9508027*, January 1996. Available from: http://arxiv.org/abs/quant-ph/9508027.

[2] Lov K. Grover. **Quantum Mechanics Helps in Searching for a Needle in a Haystack**. *Physical Review Letters*, **79**(2):325–328, July 1997. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.79.325.

[3] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O'Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and John M. Martinis. **Coherent Josephson Qubit Suitable for Scalable Quantum Integrated Circuits**. *Physical Review Letters*, **111**(8):080502, August 2013. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.111.080502.

[4] R. Maurand, X. Jehl, D. Kotekar-Patil, A. Corna, H. Bohuslavskyi, R. Laviéville, L. Hutin, S. Barraud, M. Vinet, M. Sanquer, and S. De Franceschi. **A CMOS silicon spin qubit**. *Nature Communications*, **7**:13575, November 2016. Available from: https://www.nature.com/articles/ncomms13575.

[5] Evan M. Fortunato, Marco A. Pravia, Nicolas Boulant, Grum Teklemariam, Timothy F. Havel, and David G. Cory. **Design of strongly modulating pulses to implement precise effective Hamiltonians for quantum information processing**. *The Journal of Chemical Physics*, **116**(17):7599–7606, May 2002. Available from: https://aip.scitation.org/doi/10.1063/1.1465412.

[6] Yu-xi Liu, L. F. Wei, and Franco Nori. **Tomographic measurements on superconducting qubit states**. *Physical Review B*, **72**(1):014547, July 2005. Available from: http://arxiv.org/abs/quant-ph/0407197.

[7] Line Hjortshøj Pedersen, Niels Martin Møller, and Klaus Mølmer. **Fidelity of quantum operations**. *Physics Letters A*, **367**(1-2):47–51, July 2007. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0375960107003271.

[8] Easwar Magesan, Jay M. Gambetta, B. R. Johnson, Colm A. Ryan, Jerry M. Chow, Seth T. Merkel, Marcus P. da Silva, George A. Keefe, Mary B. Rothwell, Thomas A. Ohki, Mark B. Ketchen, and M. Steffen. **Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking**. *Physical Review Letters*, **109**(8):080505, August 2012. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.109.080505.

[9] Sergio Boixo, Sergei V. Isakov, Vadim N. Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J. Bremner, John M. Martinis, and Hartmut Neven. **Characterizing**

quantum supremacy in near-term devices. *Nature Physics*, **14**(6):595–600, June 2018. Available from: https://www.nature.com/articles/s41567-018-0124-x.

[10] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. **General teleportation channel, singlet fraction, and quasidistillation**. *Physical Review A*, **60**(3):1888–1898, September 1999. Available from: https://link.aps.org/doi/10.1103/PhysRevA.60.1888.

[11] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. **Quantum supremacy using a programmable superconducting processor**. *Nature*, **574**(7779):505–510, October 2019. Available from: http://www.nature.com/articles/s41586-019-1666-5.

[12] S. Ramesh Ekanayake, Torsten Lehmann, Andrew S. Dzurak, and Robert G. Clark. **Qubit Control-Pulse Generator Circuits for Operation at Cryogenic Temperatures**. In *2008 8th IEEE Conference on Nanotechnology*, pages 472–475, August 2008.

[13] M.T. Rahman and T. Lehmann. **A cryogenic DAC operating down to 4.2 K**. *Cryogenics*, **75**:47–55, April 2016. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0011227516000278.

[14] M. L. V. Tagliaferri, A. Crippa, S. Cocco, M. De Michielis, M. Fanciulli, G. Ferrari, and E. Prati. **Modular Printed Circuit Boards for Broadband Characterization of Nanoelectronic Quantum Devices**. *IEEE Transactions on Instrumentation and Measurement*, **65**(8):1827–1835, August 2016.

[15] P. Clapera, S. Ray, X. Jehl, M. Sanquer, A. Valentian, and S. Barraud. **A quantum device driven by an on-chip CMOS ring oscillator**. In *2014 11th International Workshop on Low Temperature Electronics (WOLTE)*, pages 73–76, July 2014.

[16] E. Charbon, F. Sebastiano, A. Vladimirescu, H. Homulle, S. Visser, L. Song, and R. M. Incandela. **Cryo-CMOS for quantum computing**. In *2016 IEEE International Electron Devices Meeting (IEDM)*, pages 13.5.1–13.5.4, December 2016.

[17] E. Charbon, F. Sebastiano, M. Babaie, A. Vladimirescu, M. Shahmohammadi, R. B. Staszewski, H. A. R. Homulle, B. Patra, J. P. G. van Dijk, R. M. Incandela, L. Song, and B. Valizadehpasha. **Cryo-CMOS circuits and systems for scalable quantum computing**. In *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, pages 264–265, February 2017.

[18] J. C. Bardin, E. Jeffrey, E. Lucero, T. Huang, O. Naaman, R. Barends, T. White, M. Giustina, D. Sank, P. Roushan, K. Arya, B. Chiaro, J. Kelly, J. Chen, B. Burkett,

Y. Chen, A. Dunsworth, A. Fowler, B. Foxen, C. Gidney, R. Graff, P. Klimov, J. Mutus, M. McEwen, A. Megrant, M. Neeley, C. Neill, C. Quintana, A. Vainsencher, H. Neven, and J. Martinis. **A 28nm Bulk-CMOS 4-to-8GHz <2mW Cryogenic Pulse Modulator for Scalable Quantum Computing**. In *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, pages 456–458, February 2019.

[19] Teng-Teng Lu, Zhen Li, Chao Luo, Jun Xu, Weicheng Kong, and Guoping Guo. **Characterization and Modeling of 0.18μm Bulk CMOS Technology at Sub-Kelvin Temperature**. *IEEE Journal of the Electron Devices Society*, **8**:897–904, 2020.

[20] Yuanke Zhang, Tengteng Lu, Wenjie Wang, Yujing Zhang, Jun Xu, Chao Luo, and Guoping Guo. **Characterization and Modeling of Native MOSFETs Down to 4.2 K**. *IEEE Transactions on Electron Devices*, **68**(9):4267–4273, September 2021.

[21] R. M. Incandela, L. Song, H. A. R. Homulle, F. Sebastiano, E. Charbon, and A. Vladimirescu. **Nanometer CMOS characterization and compact modeling at deep-cryogenic temperatures**. In *2017 47th European Solid-State Device Research Conference (ESSDERC)*, pages 58–61, September 2017.

[22] P. A. T Hart, M. Babaie, Edoardo Charbon, Andrei Vladimirescu, and Fabio Sebastiano. **Characterization and Modeling of Mismatch in Cryo-CMOS**. *IEEE Journal of the Electron Devices Society*, **8**:263–273, 2020. Available from: `https://ieeexplore.ieee.org/document/9015956/`.

[23] P. A. t Hart, M. Babaie, E. Charbon, A. Vladimirescu, and F. Sebastiano. **Subthreshold Mismatch in Nanometer CMOS at Cryogenic Temperatures**. In *ESSDERC 2019 - 49th European Solid-State Device Research Conference (ESSDERC)*, pages 98–101, September 2019.

[24] Arnout Beckers, Farzan Jazaeri, and Christian Enz. **Characterization and Modeling of 28-nm Bulk CMOS Technology Down to 4.2 K**. *IEEE Journal of the Electron Devices Society*, **6**:1007–1018, 2018.

[25] S. Bonen, U. Alakusu, Y. Duan, M. J. Gong, M. S. Dadash, L. Lucci, D. R. Daughton, G. C. Adam, S. Iordănescu, M. Păşteanu, I. Giangu, H. Jia, L. E. Gutierrez, W. T. Chen, N. Messaoudi, D. Harame, A. Müller, R. R. Mansour, P. Asbeck, and S. P. Voinigescu. **Cryogenic Characterization of 22-nm FDSOI CMOS Technology for Quantum Computing ICs**. *IEEE Electron Device Letters*, **40**(1):127–130, January 2019.

[26] H. Bohuslavskyi, S. Barraud, M. Cassé, V. Barrai, B. Bertrand, L. Hutin, F. Arnaud, P. Galy, M. Sanquer, S. De Franceschi, and M. Vinet. **28nm Fully-depleted SOI technology: Cryogenic control electronics for quantum computing**. In *2017 Silicon Nanoelectronics Workshop (SNW)*, pages 143–144, June 2017.

[27] Bruna Cardoso Paz, Loïck Le Guevel, Mikaël Cassé, Gérard Billiot, Gaël Pillonnet, Aloysius Jansen, Sebastien Haendler, Andre Juge, Emmanuel Vincent, Philippe Galy, Gérard Ghibaudo, Maud Vinet, Silvano de Franceschi, Tristan Meunier, and Fred Gaillard. **Integrated Variability Measurements of 28 nm FDSOI MOSFETs down to 4.2 K for Cryogenic CMOS Applications**. In *2020 IEEE 33rd International Conference on Microelectronic Test Structures (ICMTS)*, pages 1–5, May 2020.

[28] B. Cardoso Paz, L. Le Guevel, M. Cassé, G. Billiot, G. Pillonnet, A. G. M. Jansen, R. Maurand, S. Haendler, A. Juge, E. Vincent, P. Galy, G. Ghibaudo, M. Vinet, S. de Franceschi, T. Meunier, and F. Gaillard. **Variability Evaluation of 28nm FD-SOI Technology at Cryogenic Temperatures down to 100mK for Quantum Computing**. In *2020 IEEE Symposium on VLSI Technology*, pages 1–2, June 2020.

[29] Lucas Nyssens, Arka Halder, Babak Kazemi Esfeh, Nicolas Planes, Denis Flandre, Valeriya Kilchytska, and Jean-Pierre Raskin. **28-nm FD-SOI CMOS RF Figures of Merit Down to 4.2 K**. *IEEE Journal of the Electron Devices Society*, **8**:646–654, 2020.

[30] P. Galy, J. Camirand Lemyre, P. Lemieux, F. Arnaud, D. Drouin, and M. Pioro-Ladrière. **Cryogenic Temperature Characterization of a 28-nm FD-SOI Dedicated Structure for Advanced CMOS and Quantum Technologies Co-Integration**. *IEEE Journal of the Electron Devices Society*, **6**:594–600, 2018.

[31] Wriddhi Chakraborty, Khandker Akif Aabrar, Jorge Gomez, Rakshith Saligram, Arijit Raychowdhury, Patrick Fay, and Suman Datta. **Characterization and Modeling of 22 nm FDSOI Cryogenic RF CMOS**. *IEEE Journal on Exploratory Solid-State Computational Devices and Circuits*, **7**(2):184–192, December 2021.

[32] Hung-Chi Han, Farzan Jazaeri, Antonio D'Amico, Zhixing Zhao, Steffen Lehmann, Claudia Kretzschmar, Edoardo Charbon, and Christian Enz. **In-depth Cryogenic Characterization of 22 nm FDSOI Technology for Quantum Computation**. In *2021 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EuroSOI-ULIS)*, pages 1–4, September 2021.

[33] Omar López-L, I. Martínez, D. Durini, E. A. Gutiérrez-D, D. Ferrusca, M. Velázquez, F. J. De la Hidalga, and V. Gómez. **A Performance Comparative at Low Temperatures of Two FET Technologies: 65 nm and 14 nm**. In *2020 17th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, pages 1–4, November 2020.

[34] Lin-Jie Fan, Jin-Shun Bi, Xue Fan, Gao-Bo Xu, Yan-Nan Xu, Kai Xi, and Zhan-Gang Zhang. **Cryogenic Characterization of Nano-scale Bulk FinFETs**. In *2020 IEEE 15th International Conference on Solid-State & Integrated Circuit Technology (ICSICT)*, pages 1–3, November 2020.

[35] Hung-Chi Han, Farzan Jazaeri, Antonio D'Amico, Andrea Baschirotto, Edoardo Charbon, and Christian Enz. **Cryogenic Characterization of 16 nm FinFET Technology for Quantum Computing**. In *ESSCIRC 2021 - IEEE 47th European Solid State Circuits Conference (ESSCIRC)*, pages 71–74, September 2021.

[36] Loick Le Guevel, Gerard Billiot, Xavier Jehl, Silvano De Franceschi, Marcos Zurita, Yvain Thonnart, Maud Vinet, Marc Sanquer, Romain Maurand, Aloysius G. M. Jansen, and Gael Pillonnet. **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**. In *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, pages 306–308, San Francisco, CA, USA, February 2020. IEEE. Available from: `https://ieeexplore.ieee.org/document/9063090/`.

[37] B. Patra, J. P. G. van Dijk, S. Subramanian, A. Corna, X. Xue, C. Jeon, F. Sheikh, E. Juarez-Hernandez, B. P. Esparza, H. Rampurawala, B. Carlton, N. Samkharadze, S. Ravikumar, C. Nieva, S. Kim, H. Lee, A. Sammak, G. Scappucci, M. Veldhorst, L. M. K. Vandersypen, M. Babaie, F. Sebastiano, E. Charbon, and S. Pellerano. **19.1 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**. In *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, pages 304–306, 2020.

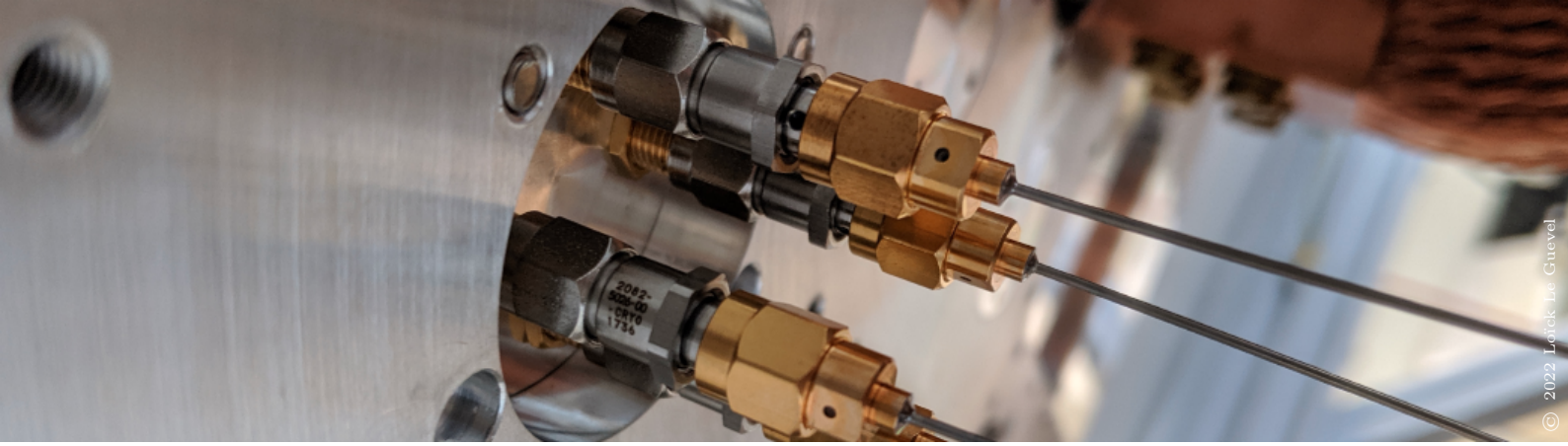
[38] J.-S. Park, S. Subramanian, L. Lampert, T. Mladenov, I. Klotchkov, D. J. Kurian, E. Juarez-Hernandez, B. Perez-Esparza, S. R. Kale, K. T. Asma Beevi, S. Premaratne, T. Watson, S. Suzuki, M. Rahman, J. B. Timbadiya, S. Soni, and S. Pellerano. **A Fully**

Integrated Cryo-CMOS SoC for Qubit Control in Quantum Computers Capable of State Manipulation, Readout and High-Speed Gate Pulsing of Spin Qubits in Intel 22nm FFL FinFET Technology. In *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, **64**, pages 208–210, February 2021.

[39] David J. Frank, Sudipto Chakraborty, Kevin Tien, Pat Rosno, Thomas Fox, Mark Yeck, Joseph A. Glick, Raphael Robertazzi, Ray Richetta, John F. Bulzacchelli, Daniel Ramirez, Dereje Yilma, Andrew Davies, Rajiv V. Joshi, Shawn D. Chambers, Scott Lekuch, Ken Inoue, Devin Underwood, Dorothy Wisnieff, Chris Baks, Donald Bethune, John Timmerwilke, Blake R. Johnson, Brian P. Gaucher, and Daniel J. Friedman. **A Cryo-CMOS Low-Power Semi-Autonomous Qubit State Controller in 14nm FinFET Technology**. In *2022 IEEE International Solid- State Circuits Conference (ISSCC)*, **65**, pages 360–362, February 2022.

[40] G. Kiene, A. Catania, R. Overwater, P. Bruschi, E. Charbon, M. Babaie, and F. Sebastiano. **13.4 A 1GS/s 6-to-8b 0.5mW/Qubit Cryo-CMOS SAR ADC for Quantum Computing in 40nm CMOS**. In *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, **64**, pages 214–216, February 2021.

[41] A. Ruffino, Y. Peng, T.-Y. Yang, J. Michniewicz, M. F. Gonzalez-Zalba, and E. Charbon. **A Fully-Integrated 40-nm 5-6.5 GHz Cryo-CMOS System-on-Chip with I/Q Receiver and Frequency Synthesizer for Scalable Multiplexed Readout of Quantum Dots**. In *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, **64**, pages 210–212, February 2021.



CHAPTER 2
Deep-Cryogenic Massive DC Characterization of FD-SOI MOSFETs

I Introduction

Circuit design requires accurate and computationally efficient models of both active and passive elements within a chosen technology to keep up with increasingly complex systems. Industry-standard models are developed and optimized for operation at temperatures ranging from -40 to 85 °C, going as far as from -55 to 125 °C for military-grade models (equivalently from 218 to 398 K). These so-called compact models consider hundreds of intricated phenomena: from the well-known semiconductor physics to e.g. quantification phenomena in thin-films. As the transistor's size shrinks with Moore's law down to a few nanometers, the number of parameters in these models exploded to keep-up with the rich physics.

To exploit the compact models, its is necessary to determine the model parameters from accurate electrical measurements of transistor characteristics. Multiple iterations through an optimized procedure are required to converge to a physically reasonable and accurate set of parameters. Low-frequency models form the basis to more complicated models with additional parasitic elements for high-frequency simulations. A large set of transistor DC curves are fitted to the model description following a procedure that allows to isolate different contributions (short vs long channels, wide vs narrow,...) and extract physically-meaningful parameters.

Measuring the intrinsic properties of single-devices can be challenging as the physical access to the device port require additional fabrication stages (e.g. metal routing and silicidation of contacts) that adds parasitic access resistance. To fit a physical model on measurements, the Back-End Of Line (BEOL) effect on the device characteristics has to be taken into account or removed with e.g. the four-point method[1] (also known as the Kelvin method). Typical devices in modern CMOS silicon technologies require measuring sub-mA currents with multiple volt-level biasing voltages (e.g. source, drain, gate, substrate,...).

To characterize plenty of devices at room temperature, an automated wafer probe station browses all individual devices and measures them one-by-one, moving the probe along each device. This method allows to precisely characterize thousands of devices across a single wafer in a few hours. Although cryogenic wafer probe stations have been developed[2], they remain expensive and limited to temperatures not lower than a few Kelvins. The requirement of having an entire wafer with multiple area-consuming wire-bonding pads per device also makes this solution not cost-effective.

For the characterization of thousands of devices at cryogenic temperatures, cooling down and warming up a cryogenic setup for a few devices per run is not an option as a full temperature cycle can take from 2 hours with a 4.2 K dip-stick to more than 24 hours with a dilution fridge. Inspired by local-mismatch studies at room temperature that leverage on-chip multiplexing to characterize many closely-packed identical transistors[3, 4, 5], several groups developed multiplexing matrices to measure current-voltage characteristics of a wide variety of transistors down to cryogenic temperatures with e.g. 66 devices per matrix in 2018[6, 7] and, later, our realization with 1024 devices per matrix published in 2020[8, 9].

In this chapter, we describe and discuss the implementation of our multiplexed matrix composed of 1024 devices that operates down to 0.1 K[8, 9]. From the obtained DC characteristics, we derive and study the dependence on temperature and voltage biasing at 300, 4.2, and 0.1 K of single-transistor quantities which will be of importance for integrated circuit design.

II Cryogenic Addressable Matrix

The addressable matrix for DC current-voltage (I-V) characterizations of individual devices is composed of multiple unit cells each containing a switch structure for multiplexing and the Device Under Test (DUT). Every unit cell is activated or deactivated depending on the programmed input sent to an on-chip digital circuit following a communication/addressing protocol. All structures switchable via the digital command have common analog inputs: current forcing ports (Force-I or FI), voltage sensing ports (Sense-V or SV), and voltage inputs for biasing (Force-V or FV) as shown in Figure 2.1.

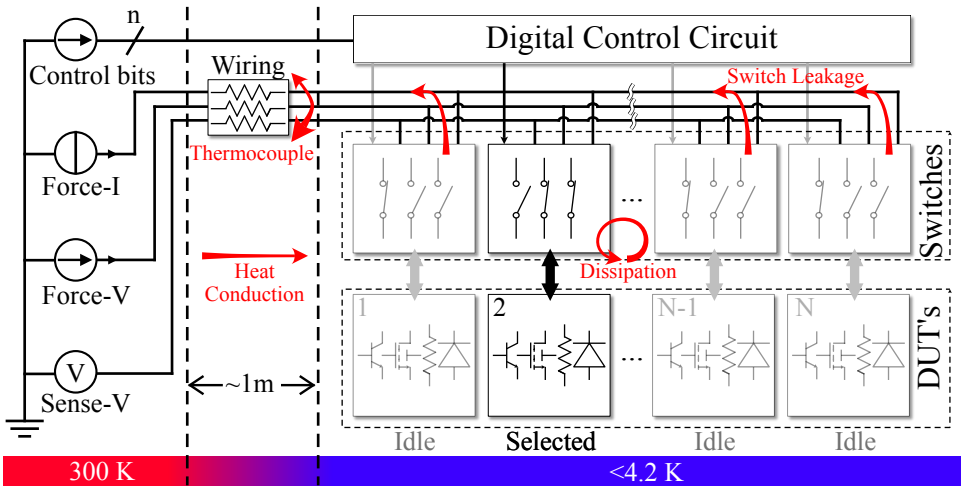


Figure 2.1 | System-view of a multiplexing matrix.

Functional view of an addressable matrix for current-voltage measurements at cryogenic temperatures of individual DUT's such as bipolar transistors, diodes, resistors, or MOSFETs. The structure of addressable switches reduces the number of output lines from N$\sim 10^3$ (one line per DUT) to a few lines ~ 10. The on-chip digital control circuit decodes the user inputs and controls the switches. Signal addressing and sensing with room-temperature electronics is performed through meter-scale wiring with non-negligible resistance. The items in red indicate important issues for cryogenic operation such as the dissipation at low-temperatures, the heat-load from room-temperature, the leakage current through open switches, and the thermoelectric voltage across the wiring.

The multiplexing of a matrix of Devices Under Test (DUT's) requires to satisfy the following criteria to accurately measure individual DC I-V curves:

i. Measurement of each DUT individually with proper DC biasing;

ii. Scalability to thousands of devices compatible with a wide range of geometry or device type;

iii. Negligible impact of access resistances on measurements;

iv. Negligible impact of idle devices on measurements.

The switch structures break the current path from the current port (FI) to the unselected DUT's to isolate the drain-source current I_{ds} of the selected DUT (criterion i.). Generally, switches are the highest constraint when designing a multiplexing matrix as the total switch area scales linearly with N the total number of devices. The switches need to be small enough to allow scalability (criterion ii.) but wide enough to reduce access resistance (criterion iii.). An increased switch resistance results in an increased voltage drop, reducing the maximum DUT source-drain voltage for high currents. The switch resistance sets the maximum W/L dimensional ratio of devices that we can measure at high V_{ds}, with W (respectively L) the transistor gate width (respectively the gate length).

Switch nonidealities become important when scaling N to thousands of devices. Individual leakage current from the switch bit-line command to the current path (FI) results in a $N-1 \sim 10^3$ times higher leakage current measured in parallel with the DUT. This leakage limits the minimum measurable current and can even prevent the measurement of devices with a sufficiently small W/L ratio.

Digital DUT selection

The digital control circuit receives from the user the addressing signal enabling the measurements of only one DUT by toggling selected switches.

The "cumulative" method of addressing uses only two user inputs. Every unit cell control signal is connected to the next cell by a flip-flop (see Figure 2.2a). The first unit cell signal is initialized to a "1" and the first DUT can be measured. Flip-flops are then triggered by a common signal that allow the "1" to transit to the next cell and the second DUT can be measured. By resetting flip-flops and sending p pulses that triggers flip-flops p times, the p-th DUT is selected for measurement. This method is very suitable for up-scaling with only O(1) input lines and is very powerful when $> 10^5$ identical devices are connected to perform large-scale mismatch studies. However, this method has its limitations for the measurements of transistors with varying properties (e.g. width, length, flavor, type) as the addressing depends on all previous errors in triggering. A slight shift in the addressing of different transistors makes the entire measurement useless. This phenomenon becomes even more important as in harsh environment transistors are more faulty. However, this method has been used at cryogenic temperatures[6, 7] without major reported failures for 66 addressable devices.

With no prior knowledge on the eventual faultiness of transistors at cryogenic temperatures, we decided to implement the "parallel" method that uses an unique DUT address based on a $\log_2(N)$-bit word sent by the user (see Figure 2.2b). This solution with O($\log(N)$) input lines is less scalable than the cumulative method. However, unique addressing is more robust to faulty gates and leads to more reliable measurements for a matrix containing differently designed DUT's.

Current measurement at cryogenic temperatures

To measure the current, we use the Kelvin method shown in Figure 2.2d that avoids the impact of access resistance on the measured characteristics[4] accounting for switch, cable, and metal-routing access resistances. To reduce thermal conduction from room-temperature to cryogenic temperatures,

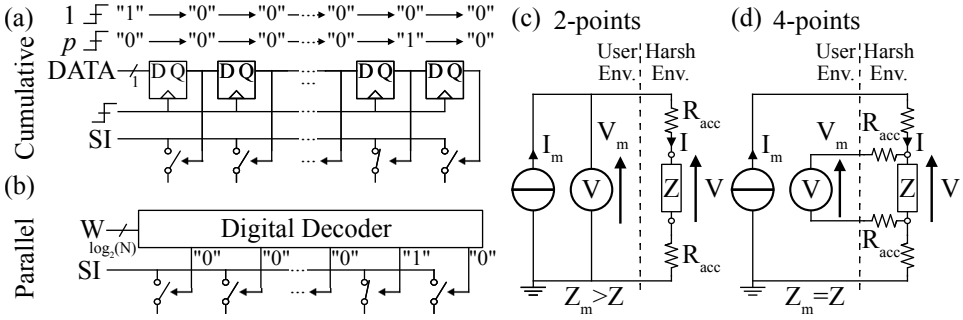


Figure 2.2 | Addressing and measurement methods.

(a) Cumulative control circuit using flip-flops. Only two inputs (1-bit DATA and trigger lines) are required independently from the number of devices N. Each clock triggering shifts registers. (b) Parallel control circuit where the input word W of $n = \log_2(N)$ bits represents the unique address of one DUT. (c) Two-points method using only two ports but adding the access resistance R_{acc} in series with the true impedance Z to the measured device impedance Z_m. (d) Four-points Kelvin method using four ports to be insensitive to access resistance R_{acc}.

resistive wiring is used, even further justifying the use of the Kelvin method in all realizations at cryogenic temperatures[6, 7, 8, 9].

The Kelvin method alleviates the access resistance but other phenomena in cryogenic setups might become problematic. The matrix chip is placed at cryogenic temperatures while the voltage and current measurement apparatus are at room-temperature. The wiring linking the chip to the room-temperature electronics experiences a large temperature gradient, conducive to thermocouple effects. The Seebeck effect is the build-up of an electric potential ΔV across a wire under a temperature gradient ΔT[10]. The thermocouple is the difference in potential between two dissimilar materials due to a temperature gradient. Different materials have different Seebeck coefficients $S(T) = -\Delta V/\Delta T$ that depend on temperature, typically with a linear dependence at high temperature and saturating at low temperatures. Let's look at the case where we are using low-resistance copper wiring for V_s carrying I_{ds} to reduce dissipation and higher-resistance constantan wiring for the gate-source voltage V_{gs} not carrying any current. If we apply $V'_{gs} = V'_g - V'_s$ at room-temperature on the hot end of the cables linking to V_g and V_s at cold, thermocoupling add a contribution to the applied potential which is: $\delta V_s = V'_s - V_s = S_{cu}\Delta T$ and $\delta V_g = V'_g - V_g = S_{Const}\Delta T$. As $S_{cu} \neq S_{Const}$, the gate-source voltage V_{gs} applied on the device at cryogenic temperatures becomes temperature-dependent as $V_{gs} = V'_{gs} + (S_{Const} - S_{Cu})\Delta T$. $S_{Const} - S_{Cu}$ represents the thermocoupling between constantan and copper, roughly equal to $-29\,\mu V/K$. The thermocouple results in a lower V_{gs} voltage on the DUT by 8.6 mV for a voltage gradient $\Delta T = 300\,K - 4.2\,K$. Thermocoupling on the gate-source voltage is only a fraction of V_{gs} and shouldn't be too problematic. However, if we consider constantan wiring of Sense-V and copper wiring for Force-I in the source-drain transport measurement, then the thermocouple voltage appears non-negligible in the measured mV-level of V_{ds}. For example, applying a zero voltage at room temperature actually generates a non-zero V_{ds} at low-temperature thus generating current which might seem perturbing. To avoid thermocoupling, we use copper wiring for S-V, F-I and the transistor source and constantan wiring for all other connections.

a) Implementation & Design

The implemented multiplexed array consists of 1024 transistors arranged in 512 pairs (all NMOS or PMOS) for DC I-V characterizations and DC mismatch studies. The 512 pairs are organized as described in Table 2.1 with the indicated command bits for device selection. The entire data set $I_{ds}(V_{gs}, V_{ds}, V_{bs})$ can be sensed for every transistor for $|V_{gs}|$ and $|V_{ds}|$ between 0 and 1.8 V. $|V_{ds}|$ and $|V_{gs}|$ are kept below the recommended supply voltage of 1 V for thin-oxide (GO1) and 1.8 V for thick-oxide (GO2) devices. The gate connection of thin-oxide and thick-oxide uses different voltage lines to avoid stressing thin-oxide devices when measuring thick-oxide devices. At room-temperature, it is recommended to keep $|V_{bs}|$ below 2 V to avoid potential oxide damage. At low-temperature, we were able to increase the back-gate voltage up to 7 V without any noticed stress. However, we keep a reasonable maximum $|V_{bs}|$ of 4 V for the systematic characterizations.

Table 2.1 | Organization of the addressing of 1024 transistors in one matrix.

The addressing control bits determines the flavor (RVT or LVT) and dimension (W,L, oxide thickness) of the selected DUT among the 512 pairs of two identical devices.

Command bits	b_9	$b_8b_7b_6b_5b_4b_3b_2b_1$	b_0
Description	V_{th}-flavor	Pair Number	Element of pair
Composition	512 pairs (256 LVT, 256 RVT)	256 pairs (128 GO1, 128 GO2)	2 devices (identical)

DUTs properties for characterization and mismatch studies

The transistor channel lengths have been chosen to obtain a representative set of the technology for high-speed circuit design with minimal length 28 nm for thin-oxide devices (resp. 130 nm for thick-oxide devices) but also higher length up to 25 µm for precise analog functions and transistor modeling. Widths vary from the minimal width of 80 nm for thin-oxide devices (respectively 150 nm for thick-oxide devices) to 25 µm allowing spanning W/L from 10^{-3} to 10^{3}. The investigated range of transistor dimensions offers a broad experimental data set for the development of cryogenic models of transistors. Up to 26 pairs of the same geometry are introduced to perform a DC mismatch study at several WL products. The number of pairs per geometry is a trade-off between the number of measured geometries versus the accuracy of the mismatch extraction. With 30 pairs, we obtain with a 96 % confidence a typical accuracy on the mismatch parameters (standard deviation) of ±28 %[11]. Measuring multiple chips accounting for a total of 100 pairs reduce the error interval to 15 %. The obtained uncertainty is higher than typical room-temperature mismatch studies performed on a few thousand pairs, leading to a <5 % uncertainty with 99.6 % fidelity. However, the obtained level of accuracy at cryogenic temperature is sufficient to extract the general trend of mismatch parameters for preliminary results while allowing a wide range of geometries inside the matrix. Details on the total number of devices and their respective dimensions are shown in Figure A.2 and A.3 for, respectively, thin-oxide and thick-oxide devices.

Leakage-canceling switch structure

The unit cells in the matrix are composed of 12 analog switches, 1 DUT pair, and 1 NAND gate (see Figure 2.3a, NAND gate not shown). The NAND gate is connected to a common row and common column line and enables the row-column addressing of unit cells. The DUTs in each unit-cell pair are placed as

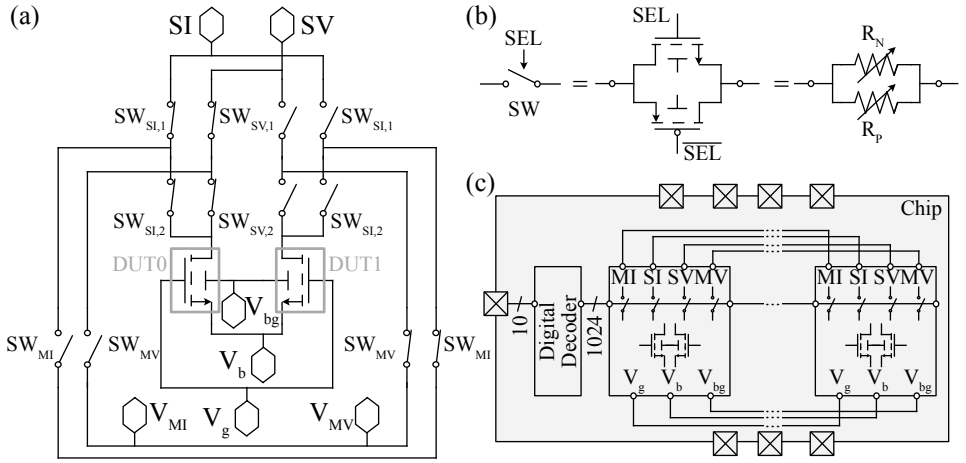


Figure 2.3 | Matrix unit-cell implementation.

Implementation of the addressable matrix. (a) Unit cell with two DUT's for pair matching. Only one of the two DUT's is activated according to the selection bits. (b) Switch implementation with pass-gates and equivalent circuit. Each transistor acts as a variable resistor that depends on the switch voltage drop. (c) Implementation of the multiplexed array with only 8 inputs/outputs.

close as allowed by the Design Rule Manual (DRM) for pair mismatch studies. Dummies are added on each side to limit systematic stress-induced mismatch when sufficient room is available. Each DUT is attached to 6 switches in a structure very similar to prior art[3] for reduced off-switch leakage impact.

For the selected DUT in the unit cell, $SW_{SV,1}$, $SW_{SV,2}$, $SW_{SI,1}$, and $SW_{SI,2}$ are closed and enable a 3-point measurement of the DUT between SI, SV, and V_b. SW_{MI} and SW_{MV} are used to reduce current leakage from idle DUT's to SI. These two switches are left open for the measured DUT. V_b is set to 0 V (resp. 1.8 V) for NMOS (resp. PMOS) and V_{ds} is defined by the measured voltage V_{SV} (resp. $V_{SV} - 1.8\,\mathrm{V}$).

Drain-source current leakage from idle DUT switches is suppressed by zeroing the voltage drop across SW_{SI} and SW_{SV}. $SW_{SI,1}$ (resp. $SW_{SV,1}$) is open and SW_{MI} (resp. SW_{MV}) is closed and apply a voltage identical to the SI (resp. SV) voltage. $SW_{SI,2}$ and $SW_{SV,2}$ break the potentially damaging current path between MI or MV to V_b. Removing currents flowing through idle devices is of the higher importance as it would result in additional dissipation of the matrix, potentially heating the fridge or the chip itself, leading to less accurate measurements. In this matrix, power is dissipated in the SI switch and DUT as the currents are flowing. For the highest measured V_{ds} of 0.9 V, the voltage drop in the current path from V_{FI} to V_b is equally distributed on the SI switch and DUT. Hence, in the worst case, the same power dissipated in the DUT will be dissipated by the switches, resulting in similar heating effects.

Switches are implemented via thick oxide pass-gates to benefit from higher oxide-breakdown voltage (equivalently higher power-supply voltage), giving more voltage room for applied drain-source voltage V_{ds}. Pass-gates (see Figure 2.3b) are composed of a NMOS transistor in parallel with a PMOS to ensure low access resistance for all V_{ds} values. The PMOS conducts when voltage signals are close to 0 V and NMOS conducts when they are close to 1.8 V. Preliminary measurements of single MOSFETs indicate that despite the V_{th} increase at cryogenic temperatures, thick oxide transistors keep a sufficiently high gate overdrive $|V_{DD} - V_{gs}|$ thanks to a higher supply voltage of 1.8 V (compared to thin-oxide transistors with $V_{DD} = 1\,\mathrm{V}$). Thick-oxide transistors are used as the V_{th} increase does not significantly alter the pass-gate resistance (see chapter 3). We expect the switch resistance and leakage to decrease at cryogenic

Table 2.2 | Switch transistor dimensions of length L and width W.

		$SW_{SI,1}$	$SW_{SI,2}$	SW_{MI}	$SW_{SV,1}$	$SW_{SV,2}$	SW_{MV}
NMOS	L (μm)	0.15	0.15	1	0.15	0.15	1
	W (μm)	12	12	0.5	1	1	0.5
	W/L	80	80	0.5	6.7	6.7	0.5
	W×L (μm^2)	1.8	1.8	0.5	0.15	0.15	0.5
PMOS	L (μm)	0.15	0.15	1	0.15	0.15	1
	W (μm)	36	36	0.5	3	3	0.5
	W/L	240	240	0.5	20	20	0.5
	W×L (μm^2)	5.4	5.4	0.5	3	3	0.5

temperature, thus making switches better at low temperatures than at room temperature. For this reason, the switch transistor dimensions (shown in Table 2.2) have been optimized via simulation with foundry models at room temperature.

Assembly of the unit cells

All unit cells are arranged in a 32×16 matrix shown in Figure 2.4b. The unit cells shown in Figure 2.4a have fixed width of 12 µm and variable height from 24 to 39 µm to accommodate long transistors. Unit cells are sorted by height needed to fit transistors to gain vertical space by affording variable height of rows. The final matrix of 1024 transistors occupies 0.216 mm^2. Each DUT is selected with a line-column addressing forming a 10-bit word whose most significant bit represents the V_{th}-flavor (left half-part: RVT and right half-part: LVT) and whose least significant bit represents one of the two element DUT0 and DUT1 of each pair inside the unit cells. A digital circuit decodes the 10-bit address and sends 0's and 1's to the corresponding row and column NAND gates, toggling switches inside the unit cells (see Figure 2.3c). The triple well structure is shared on each row with a back-gate voltage V_{bg} tap every 12 µm. SI, SV, MI, MV, V_b, V_{bg}, and V_g are common to every unit cell. In total, 17 wires are necessary to measure all 1024 DUT's.

b) Experimental Setup & Validation

The addressable matrix fits in a 1 mm^2 silicon substrate consisting of the matrix with a surrounding 48-port I/O ring. The wire-bonding pads are routed to the I/O ring in which ElectroStatic Discharge (ESD) diodes protect the circuit from external discharge (by e.g. the human body or the wiring-bonding tool). The die is wire-bonded to a 48-pin Dual-In Line (DIL) package mounted in a socket soldered onto a 4-layer Printed Circuit Board (PCB). The PCB routes the DIL pins to the setup connector, universal across all our fridges.

Validation of the experimental setup

We used many connection layers in our setup for an increased modularity and adaptability thanks to the package, the socket, and the PCB. We started our first cryogenic tests by ensuring that the chip embedding is fully functional at cryogenic temperature. The procedure described below is performed as a sanity-check for all tested circuits presented in this thesis at cryogenic temperatures.

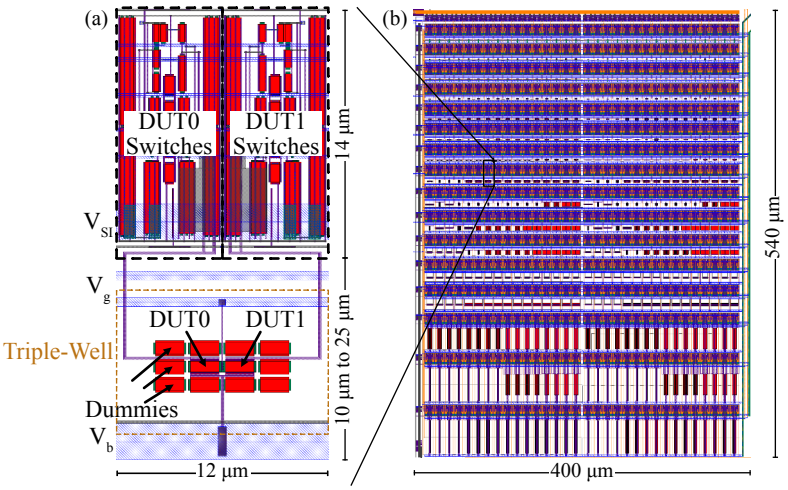


Figure 2.4 | Matrix layout.

(a) Unit cell layout with dummies and triple-well structure. (b) Matrix layout with 512 unit cells and line-column addressing.

We first characterized the ESD diodes by applying a voltage on each I/O ring pad until significant current flows through the ESD diodes validating both the connection from room-temperature electronics to the chip. The required voltage to turn on the diodes is in the typical 1 V range and should not damage the devices. This suggest that ESD diodes continue to offer protection against high-voltages (mistakenly applied by the user or via discharges). Although rigorous dynamic ESD tests are required to judge the protection at cryogenic temperatures, the circuits seemed well protected from experience (compared to our previous experience with unprotected wire-bonded single-devices).

With all working connections from room to cryogenic temperatures, we then proceeded to the validation of the matrix inside the I/O ring.

Validation of the matrix

A semiconductor parameter analyzer performs the 3-points measurement and apply $V_{MV} = V_{SI}$. A 1 TΩ input-impedance off-the-shelf buffer copies V_{SI} to V_{MI} without affecting the measured current on SI. An Arduino board controls the digital bits. Before data collection of the individual transistors in the matrix, the functioning of the circuit has been validated via a few preliminary experiments.

The first test consists in biasing DUT's at the point of maximum current to observe the voltage drop over a SI switch element, i.e. the voltage between SV and SI. For NMOS (resp. PMOS) devices, V_{gs} is set to 0.9 V (resp. -0.9 V) and $V_{ds} = 0.9$ V by adjusting V_{SI} to get $V_{SV} = 0.9\,\text{V} - V_b$ with V_b a bias voltage set to 0 V for NMOS and 1.8 V for PMOS. The highest voltage drop ΔV_{max} over the switch is measured and the equivalent worst-case switch resistance R_{SW} is extracted by dividing with the corresponding current $I(\Delta V_{max})$. Results are presented in Table 2.3. At 4.2 K, the worst-case R_{SW} is reduced by a factor of 1.3 for NMOS and a factor of 1.2 for PMOS. ΔV_{max} slightly increases at 4.2 K as the corresponding measured current I_{ds} increases by 1.4 for NMOS and 1.3 for PMOS. At 4.2 K, with respect to room temperature, the matrix performance is improved with measuring currents as high as 3.3 mA.

The leakage current I_{leak} corresponding to the 1023 unselected matrix elements in parallel is measured when all DUT's are open with $V_{gs} = 0$ V at V_{ds} equal to 0.05 and 0.9 V. At weak V_{ds}, the total leakage current is reduced by a factor of 11 at 4.2 K compared to measurements at 300 K (see Table 2.3). The

Table 2.3 | Switching resistance and leakage currents of the multiplexing matrix at 300 and 4.2 K.

	ΔV_{max} (V)	$I(\Delta V_{max})$ (mA)	R_{SW} (Ω)	$I_{leak}(0.05\,\mathrm{V})$ (nA)	$I_{leak}(0.9\,\mathrm{V})$ (nA)
300 K	0.81/-0.41	2.4/-1.3	340/320	0.21	1.1
4.2 K	0.90/-0.43	3.3/-1.7	270/260	0.020	0.91

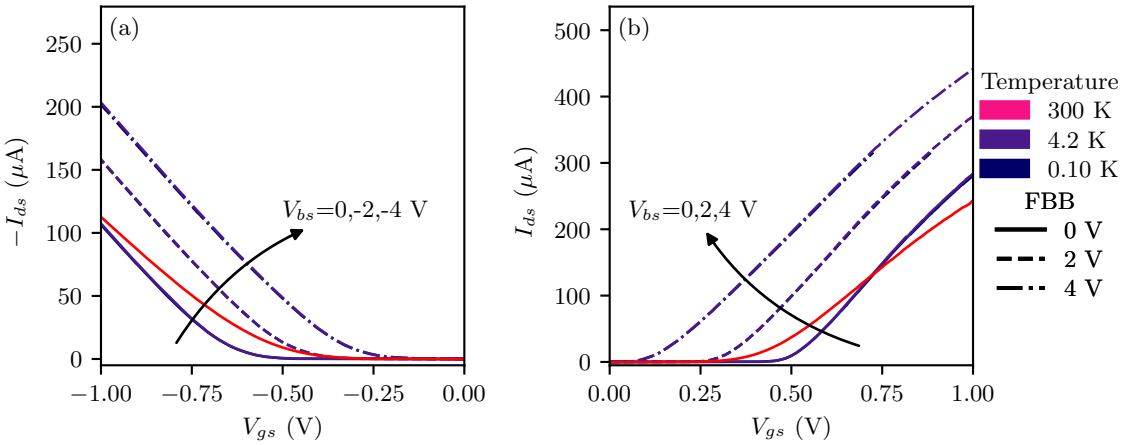


Figure 2.5 | Typical I-V characteristics vs temperature and back-gate voltage.

$I_{ds} - V_{gs}$ current-voltage characteristics at $|V_{ds}| = 0.9$ V as a function of V_{bs} at different temperatures T for a thin-oxide (a) PMOS and (b) NMOS with $L = 28$ nm and $W = 320$ nm. The red curve at 300 K and $V_{bs} = 0$ V is shown for comparison.

leakage is most likely due to Gate-Induced Drain Leakage (GIDL) as the switch V_{gs} is high. At higher V_{ds}, a higher leakage of 0.91 nA is measured with only a 17% decrease in temperature. The minimum measurable current becomes 20 pA at low temperature.

The measured switch resistance and lower leakage at cryogenic temperature validate our procedure of designing the matrix at room-temperature with accurate models. Finally, the multiplexed array has been proven to operate from 300 down to 0.1 K for the measurement of all DUTs individually for the full bias range. An example of such measurement of $I_{ds}(V_{gs})$ is given for a transistor with $L = 28$ nm and $W = 320$ nm in Figure 2.5 at 300, 4.2, and 0.1 K and for a back-gate voltage V_{bg} equal to 0, 2, and 4 V (corresponding to V_bs of 0,2, and 4 V for N-type devices and 0,-2, and -4 V for P-type devices).

The characterization of the matrix consists in measuring the following curves of individual transistors at a given temperature and back-gate voltage V_{bs}:

— $I_{ds} - V_{gs}$ for V_{gs} from 0 to V_{DD} with a step of 3 mV for $V_{ds} = 50$ mV (linear regime);

— $I_{ds} - V_{gs}$ for V_{gs} from 0 to V_{DD} with a step of 3 mV for $V_{ds} = 0.9$ V (saturation regime);

— $I_{ds} - V_{ds}$ for V_{ds} from 0 to V_{DD} with a step of 10 mV for V_{gs} equal to 0.7, 0.8, and 0.9 V (linear and saturation).

Two chips (PMOS and NMOS) each with a matrix of 1024 transistors have been entirely characterized

at 300, 100, 77, 40, 30, 20, 4.2, and 0.1 K. Thin oxide LVT transistors from 5 NMOS chips (resp. 1 PMOS chip) have been characterized at 300 and 4.2 K for a mismatch study leading to a total of 3968 transistors with 1 to 60 identical pairs per geometry[8, 9].

In this chapter, we only present treated data for thin-oxide LVT devices in a flip-well structures that are the devices more likely to be used for high-speed electronics with reasonable V_{th} values and for more compact layout by allowing sharing the triple well while being forward body-biased.

c) Perspectives in view of possible improvements

Voltage limitation for thick oxide devices

The designed matrix allows measuring the drain-source current I_{ds} of thin-oxide transistors at high source-drain bias of 0.9 V from 0.9 nA to $\sim$ 3.3 mA at cryogenic temperatures. The current range is sufficient to study the I-V characteristics in inversion for all thin-oxide devices.

Thick-oxide devices supporting voltages up to 1.8 V could only be measured at a moderate source-drain bias $V_{ds} = 0.9$ V as increasing V_{ds} lowers the voltage room for a voltage drop in the switch resulting in a lower upper limit on the measured current. If measuring these devices at high V_{ds} is a priority for e.g. studying the transport under high electrical fields, the switch resistance has to be reduced to limit the voltage drop. As in most analog circuits V_{ds} is lower than half the power supply voltage, the absence of data at higher V_{ds} is not an immediate roadblock for circuit design.

Reduction of the minimal measurable current

The lower limit in the measured current of 0.9 nA prevents the accurate extraction of e.g. the subthreshold swing or the Gate-Induced Drain Leakage (GIDL) as it requires measuring currents down to $\sim$ pA. These low-current parameters are important in some specific applications as for holding an analog voltage on a capacitor with sample and hold circuits, or for low-power digital applications aiming for the lowest subthreshold swing. As the minimum measurable current comes from leakage in the switches, the number of parallel switches on the F-I line would have to be reduced, resulting in fewer devices in the matrix. Also, the switch width could be reduced to lower the leakage because the current is so small that it should not create a significant voltage drop at the device source-drain despite the increased switch resistance. However, reducing the width reduces the maximum measurable current and will eventually prevent the measurement of devices in inversion.

A better solution to decrease the minimal measurable current consists in using a switch network with multiple stages to limit the total switch leakage in the Force-I port. The idea is to split the matrix of N devices in m sub-matrices of N/m devices. Each sub-matrix is addressable with additional switches as shown in Figure 2.6b, resulting in a doubling of the access resistance for two switches in series. If we consider that the matrix is divided m times, then the total number of switches leaking in F-I becomes equal to $N/m + m - 2$. If $m = 1$ or $m = N$, the architecture is equivalent to the original matrix architecture with doubling of the switch resistance. $m = 2$ corresponds to the matrix split in two parts with a gain in leakage by a factor of 2 with doubling of the switch resistance. Now, for $m = 4$, the switch resistance is still increased by a factor of 2 but the leakage is reduced by a factor of 4 resulting in an overall improvement of the matrix for smaller current with limited impact on the higher measurable currents. In fact, there is an optimal value of m for a given N which is $m = \sqrt{N}$. In this case, the switch resistance doubles while the leakage is divided by 16 for $N = 1024$. And actually, the higher the number of devices N, the higher is the gain in leakage compared to the original approach. For 1024 devices, we can expect a smaller measurable current of 62 pA under high source-drain voltage V_{ds} and 1.3 pA at low V_{ds}.

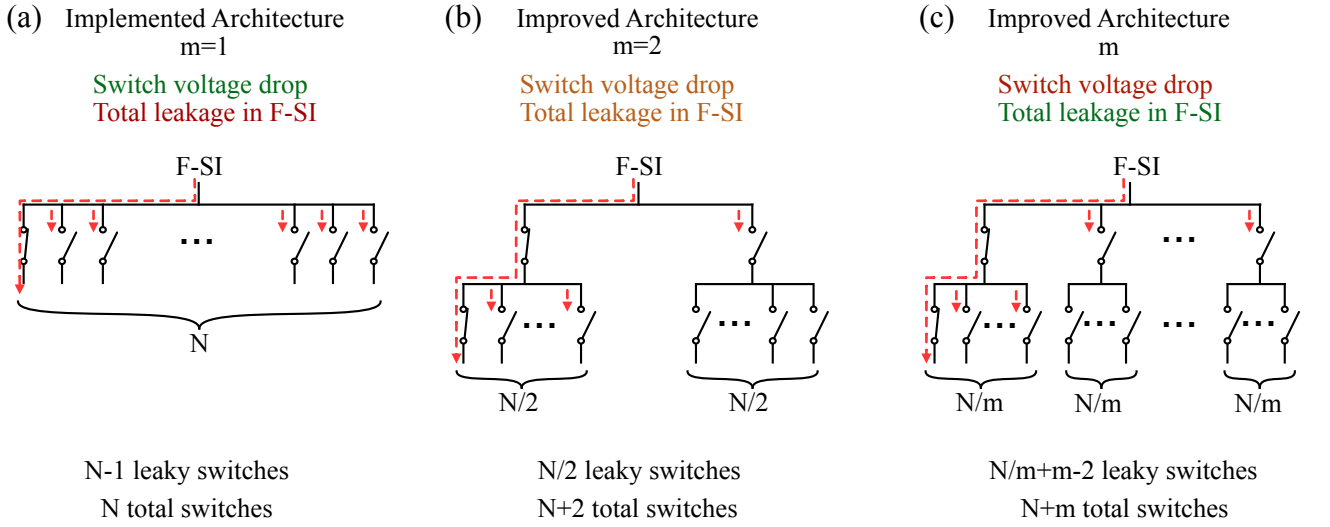


Figure 2.6 | Alternative matrix architecture for switch leakage reduction.

(a) Switch structure implemented in the presented addressable matrix with one switch per device under test. The Force current port F-SI is connected to every switch resulting in a leakage scaling with the number of DUTs N. (b) Splitting the matrix in two halves reduces the leakage by 2 at the cost of doubling of the total switch resistance. (c) Generalization of (b) for a matrix split m times. Compared to (a), the leakage is reduced by a factor $\propto 1/\sqrt{N}$ for the optimal $m = \sqrt{N}$.

Perspectives

The integration of a high-impedance follower to reduce source-drain leakage from switches and the use of a transimpedance amplifier to measure the current could be envisioned in a future version of the matrix design to ease the testing process and to eventually speed up the measurements.

The application of the original design can be applied to passive devices as e.g. resistors, or diodes or even to accurately measure a large number of temperature sensor (made of resistors or diodes with identical geometry) scattered across an active circuit to study chip cooling and temperature gradients.

The characterization of leakage in CMOS transistors might not require a dedicated addressable matrix for a large-scale measurement. The leakage generally follows simple scaling rules and is more accurately measured in single-device setups, optionally with triaxial cables. However, the proposed multiplexing matrix to measure low-frequency low-current I-V characteristics is of interest to massively characterize quantum-dot structures with currents in the pA-nA range. The variability study in the quantum transport of nominally identical quantum-dot devices will be important to optimize the fabrication process in view of the impact of device geometry and materials on the quantum properties. The proposed matrix architecture will be an efficient tool to reach the goal of low qubit variability, important to successfully implement large-scale control electronics for quantum processors.

III Subthreshold regime & subthreshold swing

The typical cross-sectional view of a MOSFET in the CMOS FD-SOI technology is shown in Figure 2.7a with its four terminals: the source, the drain, the gate, and the back-gate. We first consider the case where all applied voltages are close to 0 V. In such a configuration, a strong energy barrier separates

the source from the drain as shown in the band diagram along the channel axis (see Figure 2.7e). Some thermally excited charges cross the barrier from source to drain and drain to source, generating currents in opposite directions. As the source-drain voltage V_{ds} is equal to zero, the probability that a charge goes from source to drain $p_{s\to d}$ is equal to the probability of a charge going from drain to source $p_{d\to s}$, resulting in no effective current flowing in the device. Slightly increasing V_{ds} disturbs this equilibrium by decreasing the probability for a charge to go from source to drain leading to $p_{d\to s} > p_{d\to s}$ that effectively generates a current flowing from drain to source I_{ds}. As the barrier height at low gate-source voltage V_{gs} is rather large with respect to the thermal energy k_BT for $T \leq 300\,\mathrm{K}$, thermal charge excitation are well described by Boltzmann's statistics leading to $p \propto \exp\frac{\Delta V}{V_T}$ with ΔV the barrier height, V_T the thermal voltage k_BT/e, and e the electron's charge. The current flowing from drain to source is expressed as $I_{ds} = \mu eN$ with μ the charge mobility and N the number of effective charges going from drain to source. Hence, the flowing current becomes[12]:

$$I_{ds} \propto e(p_{d\to s} - p_{s\to d}) = e\left(\exp\frac{V_d - \Delta V}{V_T} - \exp\frac{V_s - \Delta V}{V_T}\right) \tag{2.1}$$

$$= e\left(\exp\frac{V_{ds} - \Delta V}{V_T} - \exp\frac{-\Delta V}{V_T}\right) \qquad \text{assuming} \quad V_s = 0 \tag{2.2}$$

$$= e\exp\frac{-\Delta V}{V_T}\left(\exp\frac{V_{ds}}{V_T} - 1\right) \tag{2.3}$$

with the thermal voltage $V_T = k_BT/e$ varying from 0.36 mV to 26 mV from 4.2 to 300 K. In most cases, V_{ds} is much larger than V_T which simplifies equation 2.1 to:

$$I_{ds} \propto e\exp\frac{-\Delta V}{V_T} \tag{2.4}$$

In this regime, the device acts as a current source whose current is exponentially dependent on the barrier height. The barrier height ΔV is actually controlled with the gate-source voltage V_{gs}. We can consider the gate to be capacitively coupled to the channel voltage with a capacitance C_{g-ch} while the channel is capacitively coupled to the well with C_{g-bg}. The channel can also be capacitively connected to source and drain, with respectively C_{s-ch} and C_{ch-d}, but we neglect this contribution for now. Now, it becomes clear that ΔV is controlled by V_{gs} via the capacitive divider that sets the level-arm of the gate on the channel $\alpha = \Delta V/V_{gs} = C_{g-bg}/(C_{g-ch} + C_{g-bg})$. In FD-SOI 28nm, the gate oxide has an equivalent oxide thickness (EOT) of 1.1 nm while the buried oxide (BOX) has a thickness of 25 nm[8] leading to a level arm $\alpha = 0.95$ that we consider equal to 1. Increasing V_{gs} by $2.3V_T$ decreases ΔV by $2.3V_T$ and leads to an increase of I_{ds} by a factor of 10. Each time V_{gs} increases by $2.3V_T$, the current gains one order of magnitude leading to the exponential dependence of I_{ds}. $2.3V_T$ is called the subthreshold swing SS in mV/dec. At room-temperature, devices are optimized to get the lowest SS[13], defined by the ambient temperature T for conventional CMOS Bulk and FD-SOI technologies, as higher SS can lead to increased power consumption of digital circuits. The measured SS at room temperature is close to 60 mV/dec. However, at cryogenic temperatures, the measured SS differs from the expected linear temperature dependence and saturates at sufficiently low temperatures. As an example, the theoretical SS at 4.2 K is 0.83 mV/dec while the measured SS are higher with saturating values of $3 - 10\,\mathrm{mV/dec}$ below approximately 40 K[14, 15]. This saturation can be explained by an exponential tail of states near the band edges within the band-gap that contributes to conduction and whose typical width sets SS[16]. As the temperature is reduced, the Boltzmann distribution becomes narrower than the band tail and transport becomes dominated by the available states in the band gap leading to a saturation of SS at a higher value (see Figure 2.7e).

The lower SS at cryogenic temperatures leads to a steeper reduction of current with applied source-gate voltage V_{gs}. As the current measurement with the matrix is limited at 20 pA, we could not extract the subthreshold swing at cryogenic temperatures with sufficient precision.

As V_{gs} increases, the current sharply increases by decreasing the barrier height ΔV with V_{gs} to the point where ΔV becomes of the order of $V_{T0} = k_B T_0$ with T_0 the effective temperature (highest characteristic temperature between the thermodynamic temperature and the band-tail width). From this point, charges from source and drain start to accumulate under the gate as there is no barrier anymore (see Figure 2.7f). They accumulate so much that they form a thin metallic-layer under the gate, named inversion layer (see Figure 2.7b). This term comes from older CMOS technologies with a doped silicon channel in which the effective doping was reversed by the high concentration of minority charge carriers in the inversion region. The presence of the inversion layer at high V_{gs} changes the current dependence with V_{gs} and V_{ds}. The two regimes can be separated at a characteristic gate-source voltage V_{gs} named the threshold voltage V_{th}. For $V_{gs} < V_{th}$, the device operates in subthreshold regime, considered to be in the off-state, with an exponential dependence of I_{ds} with V_{gs}. The subthreshold regime is valid for all V_{ds} as the gate-channel voltage remains below V_{th} (see Figure 2.7c). For $V_{gs} > V_{th}$, the device operates in inversion regime, considered to be in the on-state, with a different I_{ds} dependence with V_{gs} as explained in the following.

IV Linear regime at low source-drain voltage

For gate-source voltages V_{gs} above the threshold voltage V_{th}, charges accumulate under the gate and form a very thin inversion layer (see Figure 2.7b). As long as V_{ds} is kept below the gate-overdrive $V_{gs} - V_{th}$, the local gate-channel potential $V_{g-ch}(x)$, defined as the voltage drop at the top oxide along the channel position $0 \leq x \leq L$ with L the gate length, stays above V_{th} (see Figure 2.7d). In this condition, the inversion layer takes up the entire channel and electrically links the source and the drain and acts as a metallic electrode with the gate. Applying a small V_{ds} bias creates an electrical field from source to drain that forces a current to flow through the inversion layer. The inversion layer acts as a gate-tunable resistor at $V_{gs} > V_{th}$ and $V_{ds} < V_{gs} - V_{th}$. The transistor operates in the linear regime and is described by the threshold voltage V_{th} and channel resistance R_{ch} described in this section.

a) Inversion layer & threshold voltage

The threshold voltage V_{th} separates the off and on state of the transistor depending on V_{gs}. For gate-source voltages $|V_{gs}|$ below $|V_{th}|$, little current flows through the device. However once $|V_{gs}|$ becomes greater than $|V_{th}|$, the device starts conducting.

Extraction method of threshold voltage

Different definitions of V_{th} exist and often vary with the considered simplified model of the transistor[17]. We extract V_{th} from 300 to 0.1 K with the constant current method (CC) for its simplicity. The CC method is widely used for modern nodes for which the definition of V_{th} is not straightforward as many intricated phenomena play a role. This method consists in extracting the source-gate voltage V_{gs} at a given current density $I_{\square} = I_{ds} L/W$ with L the device gate-length and W the device gate-width. At room-temperature, $I_{\square}$ is chosen to remain close to more complicated extraction strategies[18]. A typical value of $I_{\square}$ is 10^{-7} A/$\square$ for N-type devices and 0.8×10^{-7} A/$\square$ for P-type devices. A slightly lower value for PMOS is chosen to account for a generally lower mobility, translating in lower source-drain currents (see Figure 2.9, A.4 and A.5). We use the same definition at cryogenic temperature and the extraction

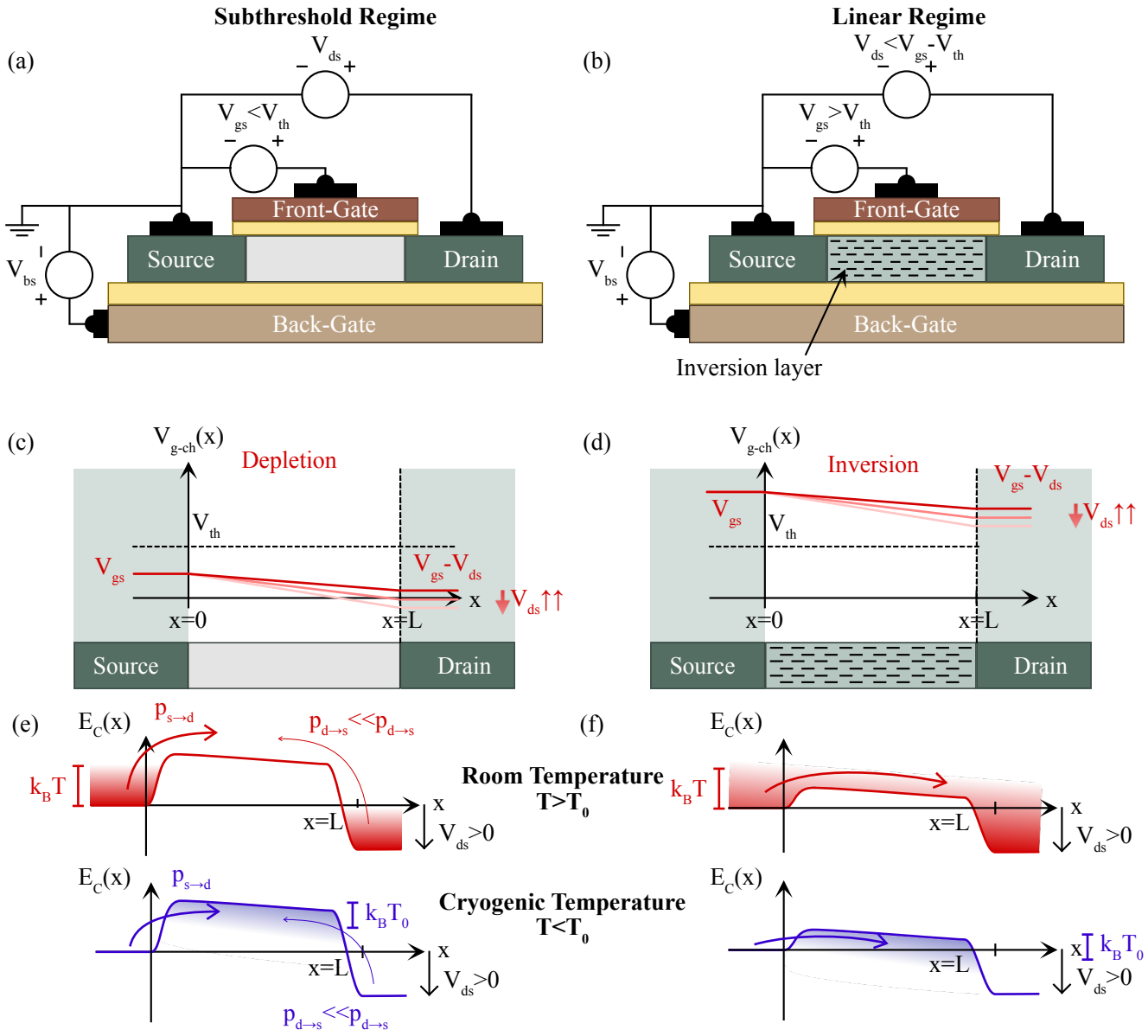


Figure 2.7 | Subthreshold and linear regime of MOSFETs.

Cross-sectional view of an FD-SOI transistor in (a) subthreshold regime with $V_{gs} < V_{th}$ and (b) in linear regime with $V_{gs} > V_{th}$ and $V_{ds} < V_{gs} - V_{th}$. Band-diagram (c) in subthreshold regime and (d) in linear regime. At room-temperature, the conduction is dominated by thermal excitation of charges in source and drain that crosses the potential barrier. (e-f) At cryogenic temperatures T, the conduction is dominated by band-tail states with a typical width of $k_B T_0$ greater than the thermal energy $k_B T$.

process is shown in Figure A.6. This definition leads to reasonable values of V_{th} at cryogenic temperatures and does not seem to require a redefinition.

Temperature evolution of threshold voltage

The threshold voltage V_{th} of devices with different lengths extracted from the addressable matrix data is shown in Figure 2.8 from 300 to 0.1 K for (a) P-type and (b) N-type devices. From 300 down to 100 K, V_{th} decreases linearly with increasing temperature T with a typical slope $\alpha_{th} = dV_{th}/dT$ of 0.73 mV/K for NMOS and $0.8 - 1$ mV/K for PMOS (see Figure 2.8cd) similar to the obtained values of 0.7 mV/K in [14] for the same technology and of 0.6 for the FD-SOI 22nm in [19]. Below 100 K, V_{th} deviates from the linear behavior and saturates to its cryogenic value $V_{th}(4.2\ K)$ below ∼70 K. The saturation of V_{th} below temperatures of a few tens of Kelvin is a general feature of CMOS technologies (bulk[20], and FD-SOI[14]). This saturation is reminiscent to the subthreshold swing saturation (see section III), also observed in [16, 14]. Knowing α_{th} and $V_{th}(4.2\,\mathrm{K})$, we can estimate the associated equivalent temperature T_0 defined as $V_{th}(4.2K) = V_{th}(300K) + \alpha_{th}(300K - T_0)$. This phenomenon is attributed to non-uniformly distributed states in the silicon band-gap as seen for V_{th} and heavily discussed in [21, 22, 16].

The threshold voltage of thin-oxide devices (GO1) for N- and P-type devices as a function of gate-length L is presented in Figure 2.9 at 300, 4.2, and 0.1 K for (a) PMOS and (b) NMOS. The increase of V_{th} from room temperature to cryogenic temperature $\Delta_T V_{th} = V_{th}(4.2K) - V_{th}(300K)$ depends on the length as already seen in FD-SOI[14] and bulk[20]. The increase $\Delta_T V_{th}$ is 156 mV for lengths below 65 nm. For longer transistors, $\Delta_T V_{th}$ continuously increases up to 178 mV at $L = 1\,\mu$m. These values are higher than in bulk[23] where it ranges from 80 to 130 mV depending on the substrate structure (well doping type & doping levels) for the same 28 nm node.

At a given temperature, $|V_{th}|$ decreases with smaller gate-length L below 100 nm due to Short-Channel Effects (SCE). For long-channel, the gate level-arm on the channel potential is close to unity and fully controls the channel potential. The electrical field from source and drain does not penetrate on a significant portion of the channel. However, as the gate-length is reduced, the field penetrates on a non-negligible portion of the channel. Previously we considered the level-arm to be solely defined by the gate-channel and channel-BOX capacitance. As the gate-length is reduced, the associated capacitance channel-source and channel-drain, respectively C_{ch-s} and C_{ch-d}, increases. The source and drain voltages now contributes to the channel potential. As a consequence, V_{th} decreases with smaller lengths L as observed in Figure 2.9.

Back-gate impact on threshold voltage

The BOX and doped well under the channel, the unique feature of SOI technologies, acts as a second gate. For N-type devices, a positive back-gate voltage V_{bg} acts as an increase of V_{gs} and lowers the threshold voltage. For P-type devices, a negative V_{bg} reduces the threshold voltage. When the applied back-gate voltage V_{bg} lowers the threshold voltage, the device is forward body biased. In all the following, we only talk of forward body bias (FBB) to avoid unnecessary distinction between N- and P-type devices. Although not desirable, the device can also be reverse body biased (RBB) with positive (respectively negative) V_{bg} for P-type (resp. N-type) devices, resulting in an increased $|V_{th}|$. As a consequence, applying a non-zero FBB lowers $|V_{th}|$. This can be seen with the simplified view in Figure 2.10b. In all generality, ϕ_b is capacitively coupled to the gate, the drain, and also the back-gate. C_{g-ch} is greater than C_{bg-ch} as the oxide is thinner at the front interface than at the back interface. V_{th} can be lowered or increased by applying a back-gate voltage. The threshold voltage change ΔV_{th} is linear with V_{bg} and linked by the back-gate level arm $\delta_{th} = C_{bg-ch}/C_{ch-ch}$ with $C_{ch-ch} = C_{g-ch} + C_{d-ch} + C_{bg-ch}$ the total capacitance seen by ϕ_b.

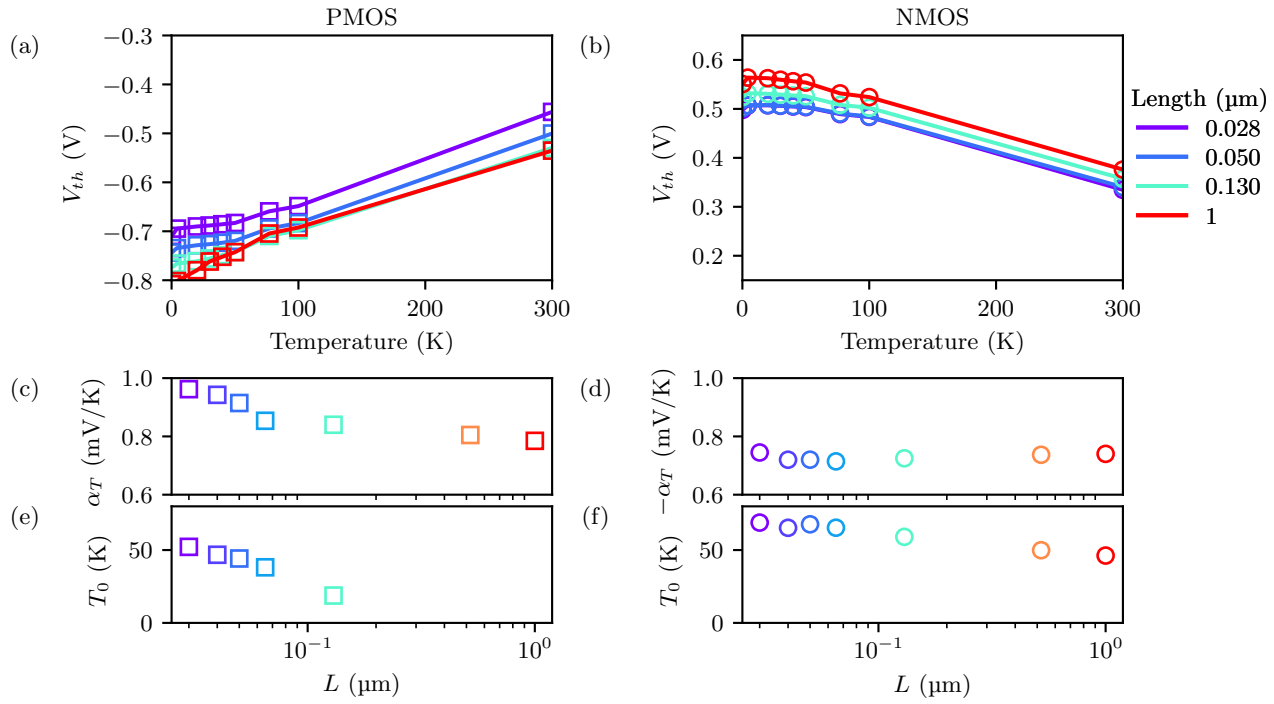


Figure 2.8 | Threshold voltage evolution with temperature.

Threshold voltage as a function of temperature for (a) PMOS and (b) NMOS from 300 to 0.25 K for different transistors with lengths of 0.028, 0.05, 0.13, and 1 µm. Linear coefficient $\alpha_T = \Delta_T V_{th}/\Delta T$ in the linear region at high temperatures $T \geq 100\,\mathrm{K}$ for (c) PMOS and (d) NMOS as a function of transistor gate-length L. Estimate of the equivalent temperature of saturation of V_{th} with temperature $T_0 = 300\,\mathrm{K} - (V_{th}(300\,\mathrm{K}) - V_{th}(4.2\,\mathrm{K}))/\alpha_T$ for (e) PMOS and (f) NMOS as a function of the length L of the transistor.

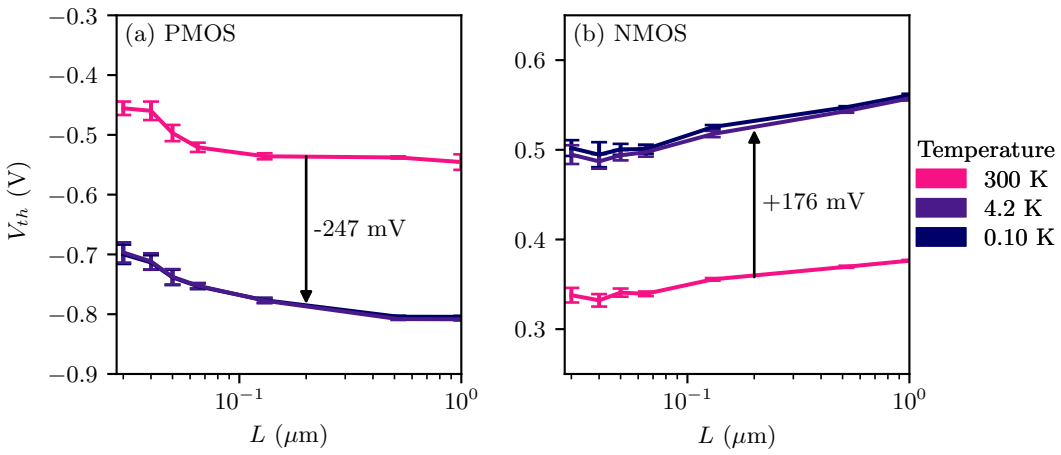


Figure 2.9 | Threshold voltages at 300, 4.2, and 0.1 K.

Threshold voltages V_{th} in linear regime for thin-oxide (a) PMOS and (b) NMOS as a function of gate length at 300, 4.2, and 0.1 K. Strong short-channel effects appear below 100 nm and result in a decrease of $|V_{th}|$.

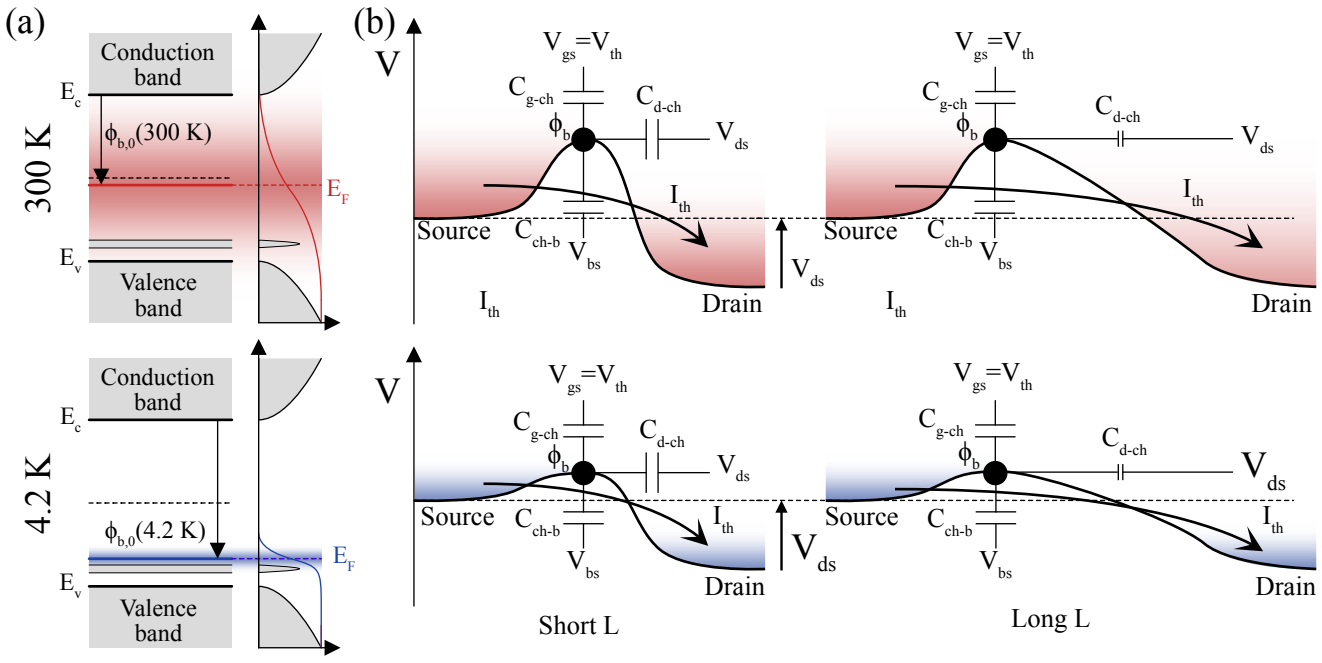


Figure 2.10 | Band diagram and threshold voltages.

(a) Band diagram (left) and state density (right) of N-type devices at 300 and 4.2 K. The Fermi level E_F is lowered at low temperature and results in a potential barrier of height $\phi_{b,0}$ at $V_{gs} = 0$ V. (b) Cross-section of the conduction band along short and long channel devices at 300 and 4.2 K. C_{d-ch} decreases with increasing length, thus reducing DIBL effects.

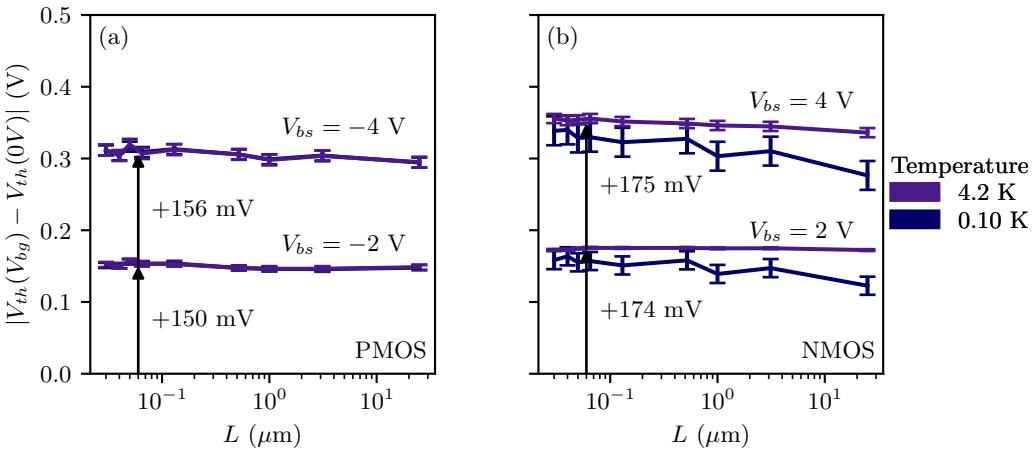


Figure 2.11 | Back-gate effect on threshold voltages.

Absolute change in threshold voltage $|V_{th}(V_{bg}) - V_{th}(V_{bg} = 0)|$ at FBB 2 and 4 V for thin-oxide (a) P-type and (b) N-type devices at 4.2 and 0.1 K.

V_{th} is measured for an FBB of 2 and 4 V. The difference of $V_{th}(FBB)$ with respect to $V_{th}(FBB = 0\,\mathrm{V})$ is shown in Figure 2.11. The back-gate effect remains even at 0.1 K, showing the absence of charge freeze-out in the underneath well. In modern technologies, the high doping dose generally prevents charge freeze-out in doped regions. The back-gate effect is independent of the device geometry, type (N or P), or flavor (LVT or RVT). Only the oxide thicknesses of the top and bottom interface have a significant impact. The back-gate level-arm δ_{th} is extracted to be 81 mV/V at 4.2 and 0.1 K valid for V_{bg} from -4 to 4 V. These values are consistent with [8, 9] and close to the obtained values at 4.2 K for the FD-SOI 22 nm[19]. As ϕ_b is linear with the voltages V_{ds} and V_{bg}, the total V_{th} variation ΔV_{th} is the sum of the respective contributions of ΔV_{ds} and ΔV_{bg}. The FD-SOI back-gate compensates for the V_{th} increase at lower temperatures.

b) Channel resistance

Once V_{gs} becomes higher than V_{th}, an inversion layer forms under the gate and acts as a resistor R_{ch} under a small $V_{ds} < V_{gs} - V_{th}$.

Extraction of channel resistance

In Figure A.4 and A.5, we show the dependence of I_{ds} with V_{ds} and V_{gs}. The current follows a linear relationship with V_{ds} in linear regime, therefore acting as a resistor. The channel resistance R_{on} is defined as V_{ds}/I_{ds} extracted from DC curves $I_{ds} - V_{gs}$ at $V_{ds} = 0.05$ V and $V_{gs} = 0.9$ V in, respectively, linear and strong inversion regimes. More details on the channel resistance extraction are shown in Figure A.7.

Temperature dependence of channel resistance

The normalized channel resistance R_{ch}/W at $FBB = 0$ V averaged on all devices of constant length L is shown in Figure 2.12. Long-channel devices exhibit the usual W/L dependence shown by the unity slope in Figure 2.12. The on-resistance per square $R_{ch}^{\square} = R_{ch} \times L/W$ of long-channel NMOS devices decreases by $\times$2.9 from 300 to 4.2 and 0.1 K (see Table 2.4). Phonon freeze-out at cryogenic temperature increases the

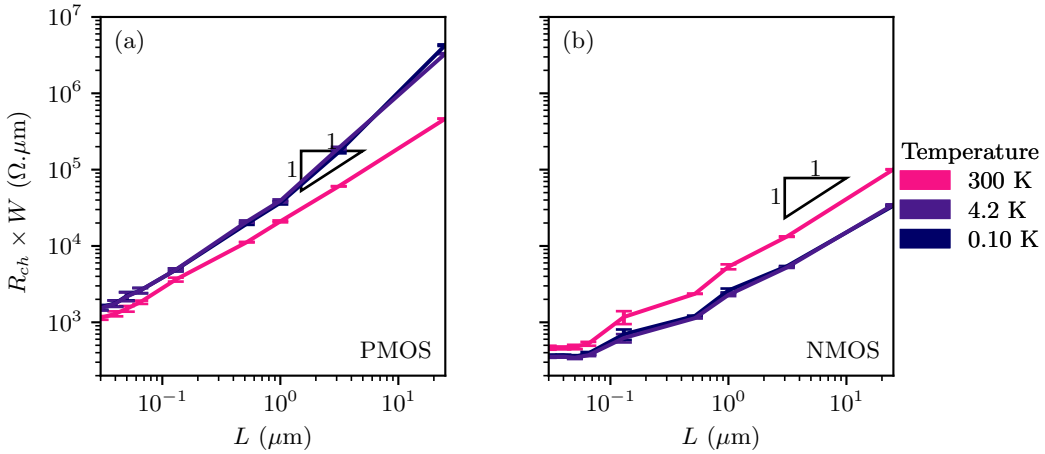


Figure 2.12 | Channel resistance of transistors down to 0.1 K.

Channel resistance R_{ch} of (a) PMOS and (b) NMOS at FBB= 0 V and at 300, 4.2, and 0.1 K. The triangles gives the slope for the expected dimensional variation $R_{ch} \propto W/L$.

Table 2.4 | On-resistance values for NMOS and PMOS at 300, 4.2, and 0.25 K.

	V_{bg} (V)	$R_{ch}^{\square}$ (kΩ/□)			$R_{ch}(30\,\text{nm})/W$ (kΩ/µm)		
		300 K	4.2 K	0.25 K	300 K	4.2 K	0.25 K
NMOS	0	4.0	1.4	1.3	0.42	0.35	0.35
	2	-	0.94	1.0	-	0.29	0.30
	4	-	0.62	0.67	-	0.26	0.26
PMOS	0	19	140	190	1.0	1.4	1.4
	-2	-	11	9.9	-	0.80	0.81
	-4	-	5.4	5.3	-	0.61	0.61

mobility of carriers leading to lower R_{ch}. However, $R_{ch}^{\square}$ of long-channel PMOS devices increases by ×7.4 from 300 to 4.2 and 0.1 K. Despite the increased mobility, this opposite phenomenon in the temperature dependence of the channel resistance comes from the increased mobility with increased threshold voltage at low temperature. PMOS devices have significantly higher V_{th} than NMOS (see Figure 2.9), reducing the gate overdrive to about only 100 mV at $V_{gs} = 0.9$ V with $V_{th} \simeq 0.8$ V. The gate overdrive is not sufficient to ensure the strong inversion and leads to higher R_{ch}. Forward body biasing of devices lowers V_{th} and increases the gate-overdrive $V_{gs} - V_{th}$, resulting in lower R_{ch} as discussed later.

Short-channel devices with $L \leq 100$ nm deviate from the usual W/L dependence by becoming independent of the length L. As the length decreases below the mean free path of charge carriers, conduction goes from a diffusive regime to a ballistic regime. In the ballistic regime, the conduction no longer depends on channel length but only on the source barrier that depends on V_{gs} and V_{bg}. Once electrons have passed the barrier, they almost instantly arrive in the drain and have no time to interact with the crystal. R_{ch} at minimal gate length of 30 nm given in Table 2.4 decreases by ∼ 20 % for NMOS and increases by ∼ 40 % for PMOS. The discrepancies of the linear resistance between long-channel NMOS and PMOS devices increase at low temperature. $R_{ch}^{\square}$ (resp. $R_{ch}(30\,\text{nm})/W$) for NMOS is ×4.8 (resp. ×2.4) higher than PMOS at 300 K and ×10 (resp. ×4) at 4.2 K.

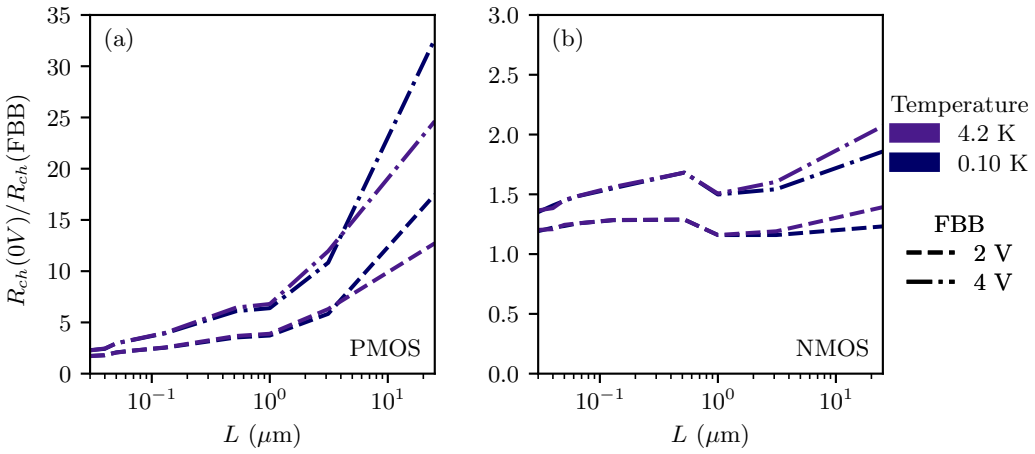


Figure 2.13 | Channel resistance reduction with back-gating.

Ratio value between R_{ch} at FBB= 0 V and R_{ch} at FBB equal to 2 and 4 V for (a) PMOS and (b) NMOS at 4.2, and 0.1 K.

Back-gate effect on channel resistance

PMOS suffers from increased V_{th} at low temperature (see Figure 2.9a). The high PMOS V_{th} of about -0.75 V at 4.2 K (with respect to the power supply of 1 V for thin-oxide devices) leads to an insufficient gate overdrive $V_{gs} - V_{th}$ that degrades the channel resistance R_{ch} at $V_{gs} = 0.9$ V. The ratio between the on-resistance at FBB$\neq$ 0 V and at FBB= 0 V is plotted in Figure 2.13 for NMOS and PMOS. The back-gate reduces V_{th} and drastically improves the PMOS specifications. With $FBB = 4$ V, $R_{ch}^{\square}$ (resp. $R_{ch}(30\,\text{nm})/W$) of PMOS devices is divided by $\times$13 (resp. $\times$1.8) reaching a value $\times$3.5 (resp.$\times$1.6) higher than the room-temperature value (see Table 2.4). NMOS also benefits from back-gating with a further decrease by $\times$2.3 for $R_{ch}^{\square}$ and by $\times$1.4 for $R_{ch}(30\,\text{nm})/W$.

c) Cryogenic irregularities for large devices under high back-gate biasing

Irregularities in the $I_{ds} - V_{gs}$ characteristics of large-area N-type devices were observed at cryogenic temperatures at strong FBB. A typical curve obtained for a NMOS with gate-length of 25 µm and gate-width of 1 µm is shown in Figure 2.14a as a function of the gate-source voltage V_{gs} and forward-body biasing (FBB) at 4.2 K. A non-monotonous hump of increasing magnitude appears in the $I_{ds} - V_{gs}$ characteristics at $V_{ds} = 50$ mV with the back-gate biasing FBB increasing from 0 to 4 V. At an FBB of 4 V, the current hump induces a decrease of current for an increase of V_{gs} in a small window around 0.6 V resulting in a negative transconductance dI_{ds}/dV_{gs} as shown in Figure 2.14b. The hump is reduced at higher source-drain voltage V_{ds} as shown in Figure 2.14c for $V_{ds} = 0.9$ V.

To understand which devices are affected, we ran tests on all acquired data at 4.2 K for FBB of 4 V. All measured sizes are presented in Figure A.10b in which the circled points of the N-type devices show significant amount of negative transconductance, at least greater than a tenth of the maximum positive transconductance. The measured curves, $I_{ds}(V_{gs})$ and its derivative, are shown in Figure A.10a and b. Devices with large area $> 0.2\,\mu\text{m}^2$ showed significant negative transconductance while devices with smaller area (squared points) exhibits a much smaller current hump that results in no negative transconductance. The negative-transconductance phenomena only appears at cryogenic temperatures and is only observed for N-type devices although P-type devices also showed a small hump.

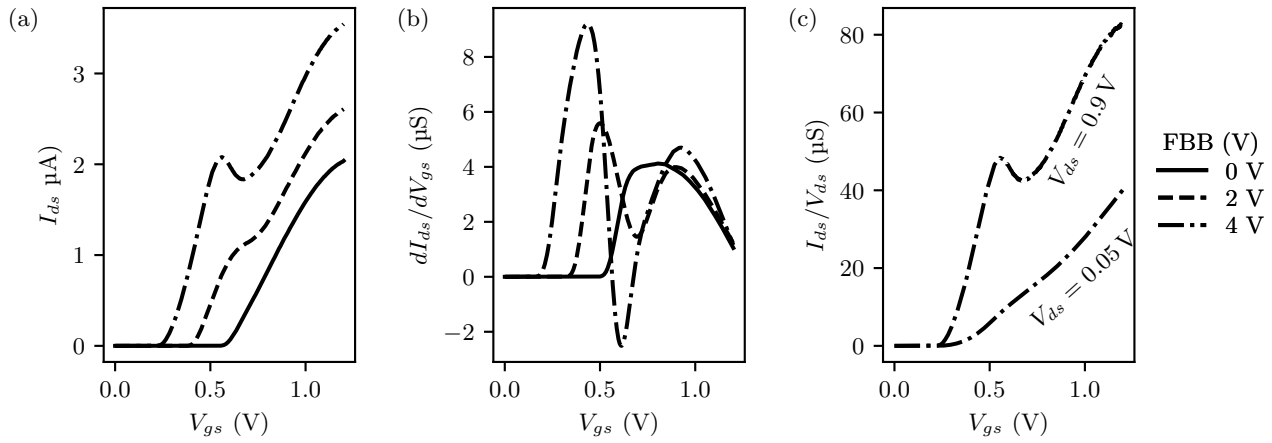


Figure 2.14 | Interface-coupling in long channel devices.

(a) Source-drain current of an N-type device with gate-length $L = 25\,\mu m$ and gate-width $W = 1\,\mu m$ for increasing FBB from 0 to 4 V. A hump in the current appears at FBB$> 2\,V$. (b) Derivative of the source-drain current with respect to the gate source voltage dI_{ds}/dV_{gs}. The current hump magnitude increases with increasing FBB and results in a negative derivative at FBB$= 4\,V$. (c) Channel conductance I_{ds}/V_{ds} at low source-drain bias $V_{ds} = 0.05\,V$ and high source-drain bias $V_{ds} = 0.9\,V$. The current hump disappears at high V_{ds}.

This feature in the source-drain current at low source-drain bias was later explained in [24]. Electrons in the channel are confined in depth thanks to the smallness of the 7 nm thick silicon channel. From 1D Poisson-Schrödinger numerical computations, the authors of [24] extracted the first three quantized energy levels with the associated electron wave functions for various gate-source and back-gate voltages as shown in Figure 2.15. With no back-gate voltage FBB$= 0\,V$, the charge centroid of the electron wave function is close to the front-gate interface and the inversion layer is located near the top oxide interface (see Figure 2.15a-c). Increasing V_{gs} increases the number of electrons in the inversion layer and brings the centroid even closer to the interface. In this regime, only the first energy level contributes to the inversion layer.

For a higher FBB of 5 V, the charge centroid is closer to the bottom interface at the onset of conduction with only the first band occupied (see Figure 2.15d). Increasing V_{gs} pulls electrons closer to the front-gate interface and results in an occupation increase of the second band whose wave function centroid is closer to the front-interface (see Figure 2.15e). The two wave functions associated with the two populated bands have non-negligible overlap that results in increased Coulomb interaction. The increased Coulomb interaction reduces the mobility of electrons as soon as the second band is populated, leading to a decrease of the device transconductance with increasing gate-source voltage at high FBB. This phenomenon is named intersubband scattering as mobility is reduced due to scattering between two conduction subbands. The two bands continues to interact with each other even at higher V_{gs} (see Figure 2.15).

The difference in energy between the two bands ΔE is computed to be 30 meV. Hence, as V_{ds} increases such that $qV_{ds} \gg \Delta E$, the hump disappears as the two sub-bands are populated at the onset of conduction and this is observed in Figure 2.14c and in [24]. For the same reason, the hump disappears at temperatures larger than 100 K as the associated thermal energy $k_B T$ exceeding ΔE forces the two first sub-bands to be simultaneously populated at higher temperatures.

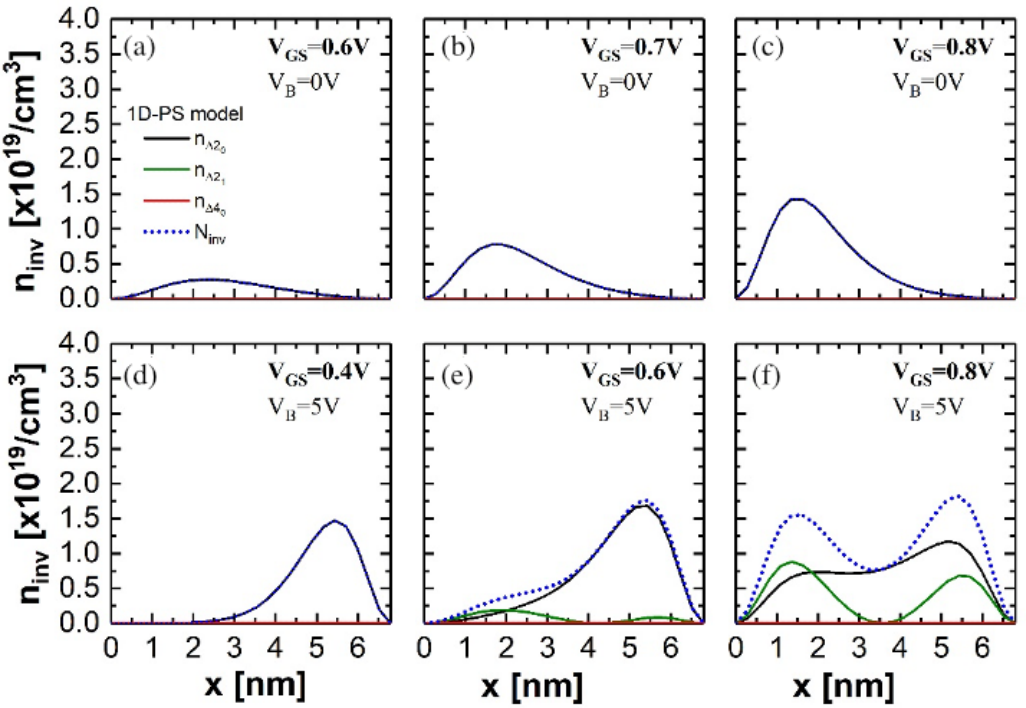


Figure 2.15 | Electron wave functions in the 7 nm-thick silicon channel.

Electron wave functions in inversion for different back-gate voltages computed from 1D Poisson-Schrödinger simulations at 4.2 K adapted from [24]. The amplitude of the electron wave function is expressed in terms of the inversion charge density along the channel height where $x = 0\,\text{nm}$ (respectively $x = 7\,\text{nm}$) corresponds to the top-gate interface (resp. back-gate interface). (a-c) At zero FBB, the electrons are localized near the top-interface only populating the first conduction band for all V_{gs}. (d-e) At higher FBB of 5 V, the electrons are localized at the bottom interface. At the onset of conduction, only the first band is populated (d). Electrons are pulled closer to the top-interface with increasing gate-source voltage resulting in the population of higher levels (e-f). Non-zero overlapping between the wave function of the first two populated stated in (e-f) results in increased Coulomb interaction that reduces the electron mobility. This translates to the hump in current observed in Figure 2.14 due to intersubband scattering.

This irregularity with negative transconductance for large-area devices under high FBB disappears for high source-drain voltage and should not be an issue for circuit operation. This phenomenon has to be taken into account in transistor modeling to ensure good predictions, even at smaller lengths.

Whether this range of geometry, prone to this phenomenon, has to be avoided is unclear. In one hand, this phenomenon allows to reach lower channel resistivity as Coulomb scattering is removed at the onset of conduction. In the other hand, the dynamic behavior and potential variability is unknown. In either case, more studies are required to fully understand it, and eventually leverage it in circuits.

d) Beyond the linear regime

In the linear regime with $V_{gs} > V_{th}$ and $V_{ds} < V_{gs} - V_{th}$, transistors act as gate-controlled resistors. The increase in V_{th} at cryogenic temperatures limits the channel resistance values R_{ch} of P-type devices. N-type devices benefit from a lower R_{ch} thanks to intrinsically lower V_{th}. Applying a back-gate voltage alleviates the overdrive-limited R_{ch} by reducing V_{th} to its room-temperature value. Important resistance reduction up to a factor of 10 are obtained at cryogenic temperatures leading to improved performance. This resistance reduction allows to reduce the typical pass-gate area by 10 at cryogenic temperatures for a given aimed value compared to room temperature.

Increasing the source-drain voltage V_{ds} beyond the gate overdrive $V_{gs} - V_{th}$ forces the inversion layer to recede depending on V_{ds}, thus entering the saturation regime.

V Saturation regime at high source-drain bias

At the onset of saturation with $V_{ds} = V_{gs} - V_{th}$, the local gate-channel potential at the drain $V(L)$ is equal to V_{th}. As V_{ds} further increases, the local gate-channel potential near the drain falls below V_{th} and forces the inversion layer to move back towards the source, creating a depletion region around the drain (see Figure 2.16). The inversion layer is stopped at this point $V(x) = V_{th}$ and create what is called a pinched-off region (see Figure 2.16a). The term pinched-off originates from bulk technologies in which the inversion layer is literally pinched off near the drain with an inversion layer depth that linearly decreases from Debye length near the source to 0 at the pinched-off region. In FD-SOI technologies, the inversion layer occupies the entire silicon channel as its depth is much smaller than Debye length. The FD-SOI inversion layer is not exactly pinched off near the drain, but we'll keep this conventional appellation. As V_{ds} increases, the pinched-off point moves back thus reducing the current. Interestingly, R_{ch} increases proportionally to V_{ds} leading to a plateau of current as a function of V_{ds}. This typical behavior is similar to an ideal current source. In modern technologies with smaller gate lengths, the receding of the pinch-off region decreases the effective length of the transistor with V_{ds}. Increasing V_{ds} reduces therefore the effective channel length and leads to a slightly higher current I_{ds}. This dependence is similar to a non-ideal current source with a parallel conductance g_{ds} (see typical characteristics in Figure A.4 and A.5).

Keeping V_{ds} constant, increasing the gate-source voltage V_{gs} increases the number of mobile charges in the inversion proportional to $V_{gs} - V_{th}$. We assumed that the charges just punch through the depletion region as they acquired sufficient speed. In the inversion layer, the lateral electrical field is equal to $V_{gs} - V_{th}$ which then gives a square dependence of I_{ds} with $V_{gs} - V_{th}$. With the high electrical fields in sub-100nm transistors, the mobility reduces due to Coulomb scattering which generally leads to a linear dependence of I_{ds} with $V_{gs} - V_{th}$. The transconductance $g_m = dI_{ds}/dV_{gs}$ of a transistor determines the voltage-to-current gain under high electrical field and depends on the mobility and V_{gs}.

In saturation, transistors act as voltage-controlled non-ideal current sources with a gate-control of the current via g_m and a parallel conductance g_{ds}.

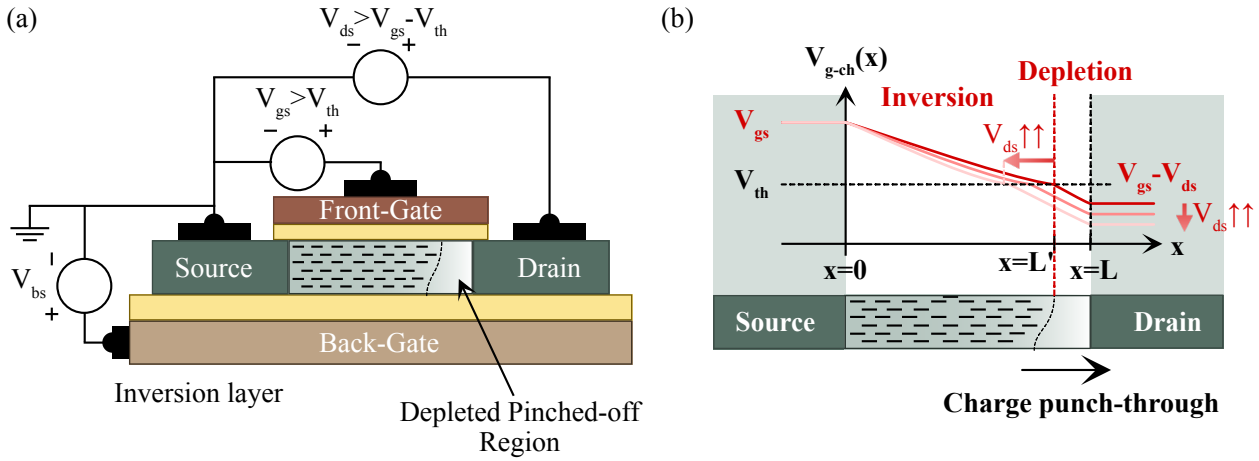


Figure 2.16 | Saturation regime in MOSFETs.

(a) Cross-view of a MOSFET in saturation with $V_{gs} > V_{th}$ and $V_{ds} > V_{gs} - V_{th}$. The high drain voltage creates a depletion region in place of a portion of inversion layer. (b) Gate-channel potential V_{g-ch} as a function of the channel position $0 \leq x \leq L$ with L the gate length. Near the drain, V_{g-ch} falls below V_{th} inducing a depletion region. Charges punch through the thin depletion region as they acquired sufficient speed from the source to the depletion region.

a) Drain-induced barrier lowering

In Figure 2.10, we saw that the channel potential could be partly controlled by the source-drain voltage V_{ds}. The contribution of the drain becomes non-negligible at high V_{ds} resulting in a lowering of the gate-source voltage required to induce the inversion layer. This effect is known as the Drain-Induced Barrier Lowering (DIBL) and becomes more important as the channel length decreases. In the previous section, we defined the threshold voltage in the linear regime V_{th}. Similarly, we define the threshold voltage in the saturation regime $V_{th,sat}$ at $V_{ds} = 0.9\,\mathrm{V}$ extracted with the Constant-Current method at the same current density $I_{\square}$ (see section IV). We define the DIBL coefficient $\delta_{th}(V_{ds})$ in mV/V as $(V_{th} - V_{th,sat})/(V_{ds,lin} - V_{ds,sat})$ with $V_{ds,lin}$ (respectively $V_{ds,sat}$) the source-drain voltage at which we measured V_{th} (resp. $V_{th,sat}$). Now, V_{th} depends on V_{ds} and is equal to $V_{th}(V_{ds}) = V_{th} - \delta_{th} V_{ds}$.

The DIBL coefficient δ_{th} is presented in Figure 2.17 as a function of device gate-length L for (a) PMOS and (b) NMOS. For long-channel devices $L \geq 1\,\mu\mathrm{m}$, the DIBL coefficient δ_{th} is rather low. For a high source-drain voltage V_{ds} of 0.9 V, the shift of threshold voltage is less than 10 % with a decrease of $20 - 40\,\mathrm{mV}$. For shorter channels, δ_{th} becomes non-negligible and results in a shift of V_{th} up to 180 mV for PMOS devices and 100 mV for NMOS devices.

Applying a non-zero back-gate voltage does not significantly affect the DIBL effect. If we go back to the capacitive picture in Figure 2.10, we see that the back-gate effect is superimposed to the impact of the source-drain. The increase in V_{th} happens for all V_{ds}, hence the DIBL coefficient stays unchanged with FBB.

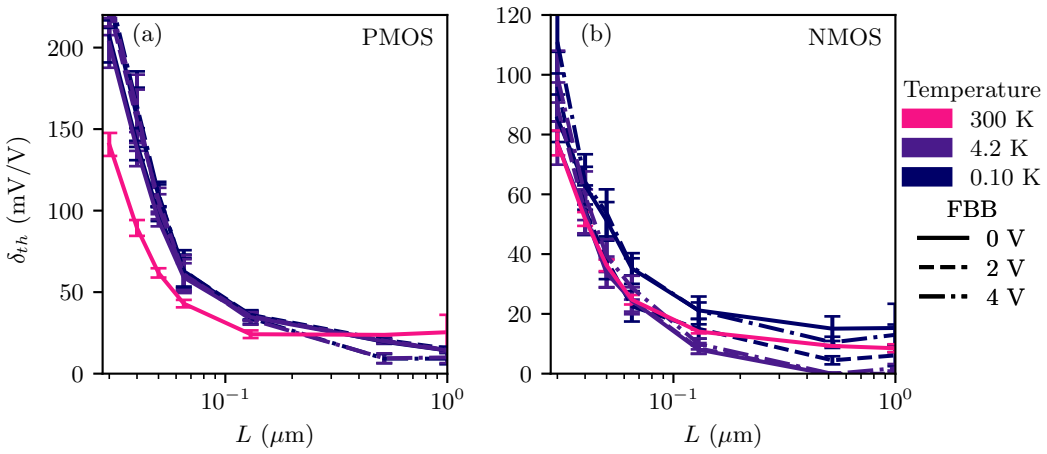


Figure 2.17 | Drain-Induced Barrier Lowering.

DIBL effect on V_{th} measured with δ_{th} for thin-oxide (a) PMOS and (b) NMOS at 4.2 and 0.1 K for FBB= $0, 2, 4$ V as a function of length. The red curve at 300 K and FBB= 0 V is given for comparison.

b) Transconductance

In saturation regime, the transistor acts as a voltage-controlled current source of current to voltage gain $g_m = dI_{ds}/dV_{gs}$, named the transconductance.

Extraction of transconductance

We extract the maximum transconductance at high source-drain voltage $V_{ds} = 0.9\,\text{V}$ by differentiating the drain-source current I_{ds} with respect to the gate-source voltage V_{gs}. To remove excessive noise inherent to the derivation of signals, we use the Savitzky-Golay method[25] of first order. Typical analysis curves are shown in Figure A.8.

Temperature variation of transconductance

The extracted maximum transconductance per unit width averaged for all widths at 300, 4.2, and 0.1 K is shown in Figure 2.18 as a function of channel length L. Long-channel devices exhibit the W/L dependence at all temperatures as emphasized with the indicated slopes in Figure 2.18ab. The transconductance per square $g_m^{\square}$ is extracted by linear fitting $1/g_m$ as a function of transistor length $L > 1\,\mu\text{m}$. $g_m^{\square}$ increases by $\times 3.3$ (resp. $\times 1.8$) for NMOS (resp. PMOS) from 300 to 4.2 and 0.1 K (see Table 2.5). Less phonon interactions at cryogenic temperatures increase the mobility and leading to higher transconductance.

Short-channel devices have a transconductance lower than $g_m^{\square} \times W/L$ due to short-channel effects. The transconductance per unit width at $L = 28\,\text{nm}$ is shown in Table 2.5 for all temperatures. $g_m(28\,\text{nm})/W$ increases by $\times 1.4$ (resp. $\times 1.2$) for NMOS (resp. PMOS) from 300 to 4.2 and 0.25 K.

Back-gate effect on transconductance

As the back-gate voltage increases, conduction takes place in different portions of the conduction channel leading to different mobility parameters. By applying the back-gate at cryogenic temperature,

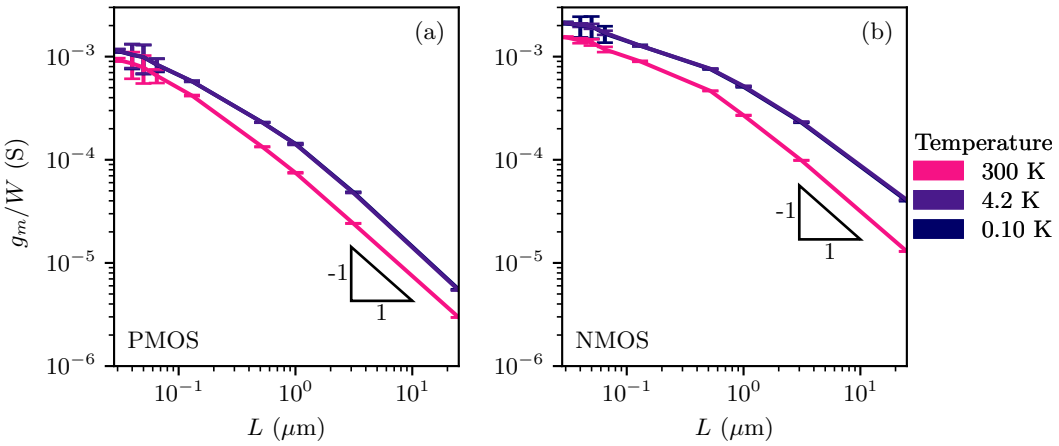


Figure 2.18 | Transconductance down to 0.1 K.

Extracted transconductance g_m at 300, 4.2, and 0.1 K as a function of transistor length L. The transconductance g_m is taken as the maximum of the dI_{ds}/dV_{gs} curve at high drain-source voltage $V_{ds} = 0.9$ V. Error bars represent the standard deviation among all widths.

Table 2.5 | Maximum transconductance of N- and P-type devices at 300, 4.2, and 0.1 K.

Extracted values for the transconductance at 300, 4.2, and 0.1 K at various forward body biasing (FBB) for NMOS and PMOS.

	FBB (V)	$g_m^{\square}$ (mS/□)			g_m(30 nm)/W (mS/µm)		
		300 K	4.2 K	0.1 K	300 K	4.2 K	0.1 K
NMOS	0	0.33	1.1	1.1	1.5	2.1	2.1
	2	-	1.2	1.2	-	2.1	2.1
	4	-	1.6	1.6	-	2.0	1.9
PMOS	0	0.074	0.13	0.14	0.93	1.1	1.1
	2	-	0.21	0.21	-	1.2	1.2
	4	-	0.29	0.30	-	1.2	1.2

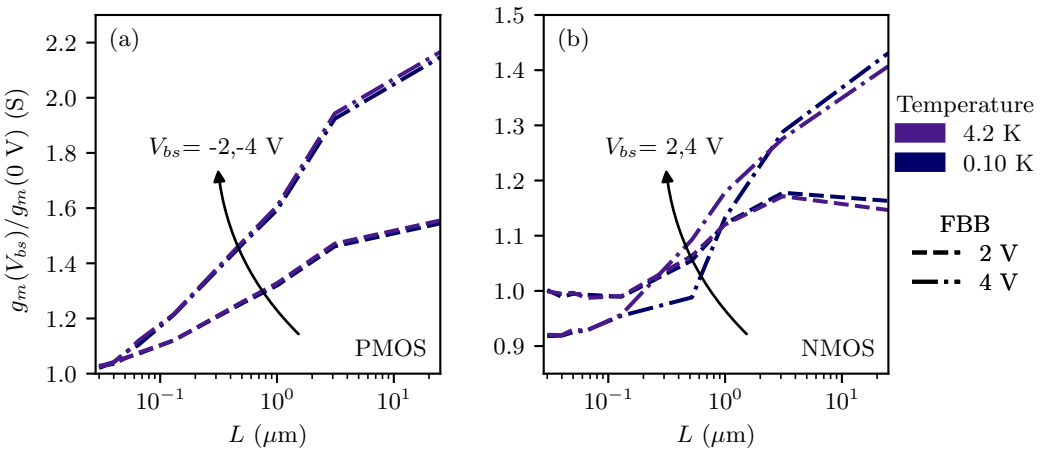


Figure 2.19 | Transconductance improvement with forward body biasing at 4.2 and 0.1 K.
Ratio of the transconductance with FBB on the one with FBB= 0 V at 4.2 and 0.1 K for (a) PMOS and (b) NMOS. The transconductance improves with FBB except for sub-1 µm N-type devices at a strong FBB of 4 V.

$g_m^{\square}$ increases by ×1.6 (resp. 2.2) for NMOS (resp. PMOS) for FBB= 4 V (see Figure 2.19). g_m/W for short-channel N-type devices actually decreases by less than 10% at FBB=4 V (see Figure A.8).

c) Conductance

The transistor is a non-ideal current source defined by a parallel conductance $g_{ds} = dI_{ds}/dV_{ds}$. The conductance measures the ability of the transistor to impose a current through a load of impedance R. A lower conductance means that the transistor behaves as a better current source as it can imposes a current through higher impedance load as long as $R < 1/g_{ds}$. Inversely, if the conductance is high, the transistor can only impose a current through low-impedance loads.

Extraction of conductance

The conductance is extracted at $V_{gs} = 0.9$ V by linear fitting the $I_{ds} - V_{ds}$ curve for $V_{DD} - V_{gs} > 100$ mV as shown in Figure A.9.

Temperature variation of conductance

The conductance is extracted for all devices and averaged over the width for each length at 300, 4.2, and 0.1 K (see Figure 2.20). The conductance g_{ds} follows a power law of the form $S_0 \times (L/L_0)^{\alpha}$ with α the exponent of the power law, and S_0 the conductance at the arbitrary normalization length L_0 chosen equal to 1 µm (see Table 2.6). α undergoes very slight changes in temperatures from 300 to 0.1 K suggesting the same physics. α is determined by the dependence of the mobility μ and the λ parameter for the reduced effective channel length. Mobility Degradation (MD) leads to a term in $(1 + A/L)^{-1}$ with A a constant, and λ leads to a L^{-1} term due to Channel-Length Modulation (CLM). From this simple model, α is expected to have a value between -1 if MD is negligible in front of CLM and -2 if they are of equal importance. α of -1.3 for PMOS suggests a domination of Channel-Length Modulation (CLM) with still influence of mobility degradation while α of -2.1 for NMOS suggests that CLM and MD are of equal

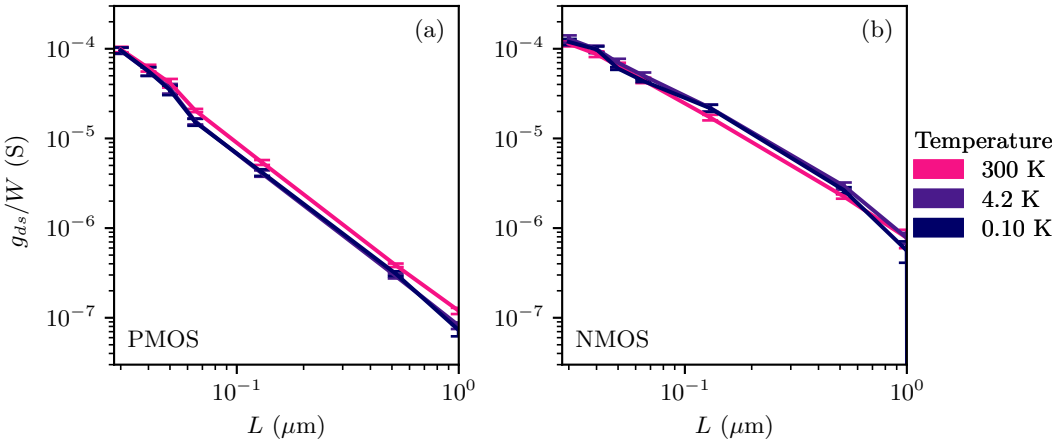


Figure 2.20 | Conductance down to 0.1 K.

Extracted conductance averaged on all width as a function of length L at 300, 4.2, and 0.1 K. Error bars represent the standard deviation among all widths.

Table 2.6 | Power-law fit results of the conductance at 300, 4.2, and 0.1 K.

Extracted values of the least-square fitting of $g_{ds}(L) = S_0 \times (L/1\,\mu\text{m})^{\alpha}$ for N- and P-type devices without FBB. S_0 is the conductancee at $L = 1\,\mu$m in µS and α is the power-law exponent.

	FBB	Temperature		
		300 K	4.2 K	0.1 K
NMOS	0 V	$0.94 \times (L/1\,\mu\text{m})^{-1.4}$	$1.3 \times (L/1\,\mu\text{m})^{-1.3}$	$1.2 \times (L/1\,\mu\text{m})^{-1.3}$
PMOS	0 V	$0.10 \times (L/1\,\mu\text{m})^{-2.0}$	$0.070 \times (L/1\,\mu\text{m})^{-2.1}$	$0.075 \times (L/1\,\mu\text{m})^{-2.0}$

importance. The conductance of NMOS devices decreases by ×2.1 from 300 to 4.2 and 0.25 K while the conductance of PMOS increase by ×2.1 from 300 to 4.2 and 0.1 K. A higher g_{ds} of NMOS despite the lower gate-source overdrive $V_{gt} = V_{gs} - V_{th}$ suggests than the conductance tends to increase at cryogenic temperature for a constant V_{gt}. Likewise, the measured slight improvement of g_{ds} for PMOS is certainly a combination of the highly-reduced V_{gt} to only −100 mV with an increased conductance at constant overdrive.

VI Overview of single-transistor behavior at cryogenic temperatures

We realized an addressable multiplexing matrix operating from 300 to 0.1 K to characterize thousands of devices in one cool-down, drastically reducing the measurement time per device, especially at sub-Kelvin temperatures. The circuit allows to measure the DC current-voltage characteristics of 1024 DUT for currents as high as 2 mA under strong source-drain bias V_{ds} of 0.9 V. Because of leakage through open switches, the minimum measurable current is limited to 26 pA at low source-drain bias and at 0.9 nA at higher bias. The minimum current prevented the accurate extraction of the subthreshold swing but allowed to measure all other parameters in inversion. For future design improvement, a revised architecture has been described to reduce the minimum measurable current which would be particularly relevant for a

Table 2.7 | Evolution of single-device quantities by cooling from 300 to 4.2 K.

The arrows indicate the evolution of each quantities from room-temperature with no FBB to 4.2 K with the indicated FBB. ⇈ (respectively ⇊) indicate an increase (resp. a decrease) when lowering the temperature and applying FBB. The arrows colors suggest whether the evolution is beneficial (green) or disadvantageous (red).

	FBB	$\lvert V_{th}\rvert$	$\lvert V_{th,sat}\rvert$	R_{ch}	g_m	g_{ds}
NMOS	0 V	⇈	⇈	⇊	⇈	⇈
	2 V	⇊	⇊	⇊	⇈	
	4 V	⇊	⇊	⇊	⇈	
PMOS	0 V	⇈	⇈	⇈	⇈	⇊
	2 V	⇈	⇈	⇊	⇈	
	4 V	⇊	⇊	⇊	⇈	

massive characterization of quantum-dot devices.

We extracted the important single-device parameters in linear and saturation regime from 300 to 0.1 K. As a summary in Figure 2.21 for NMOS and Figure 2.22 for PMOS, we plot the 2D histograms of all measured quantities at 300 K at zero forward-body bias versus the same quantity at 4.2 K for increasing forward-body bias from 0 to 4 V. The values at 300 and 4.2 K are well correlated as shown with the linear fits. At 4.2 K, the threshold voltages in saturation $V_{th,sat}$ and in linear regime V_{th} increases by $\times 1.5$, reaching levels above 0.45 V for NMOS and 0.65 V for PMOS. The significant increase of the threshold voltages above half of the power-supply voltage $V_{DD} = 1$ V is problematic for the operation of cryogenic electronics. To maintain sufficient gain, the gate-source voltages of all transistors in a circuit have to be kept above V_{th}. As circuits generally consist of multiple transistors, stacked between the ground and V_{DD}, high V_{th} limits the usable architectures. Thanks to the unique feature of SOI, V_{th} is reduced with forward body-biasing. V_{th} at 4.2 K with comparable values at 300 K are retrieved by applying an FBB greater than 2 V for NMOS and 4 V for PMOS. Ultra-low threshold voltages are reached at 4 V with V_{th} below 100 mV. Such low voltages combined with increasing subthreshold slope are interesting for low-power digital circuit operation[26].

The higher threshold voltage levels of PMOS devices also affect the channel resistance in linear regime for high source-drain bias, leading to a $-25\,\%$ evolution of R_{ch} from 300 to 4.2 K. Reducing V_{th} with FBB quickly allows recovering better values than at room-temperature with up to a $+130\,\%$ lower resistance for both N- and P-type devices. Higher mobility of the charge carriers at cryogenic temperatures reduces R_{ch} with less phonon interactions. The increased mobility leads to an improved maximum transconductance g_m increasing by $\times 1.6 - 1.8$ up to a factor of $\times 2$ for PMOS with FBB of 4 V.

The device conductance g_{ds} at high gate-source voltage does not significantly change at lower temperatures with a change of only $\simeq 10\,\%$. For N-type devices, a higher g_{ds} despite the lower gate-source overdrive $V_{gt} = V_{gs} - V_{th}$ suggests than the conductance tends to increase at cryogenic temperature for a constant V_{gt}. For P-type devices, the measured slight improvement of g_{ds} is certainly a combination of the highly-reduced V_{gt} to only -100 mV with an increased conductance at constant overdrive.

Overall, the specifications of single devices reveal a better performance at cryogenic temperatures, especially with forward body-biasing. To fully understand the effect of forward body-biasing on single devices, we look at the correlations between each quantity at 0 and 4 V as shown with the 2D histograms in Figure 2.23 and 2.24 for N-type and P-type devices. FBB of 4 V mainly offsets the threshold voltages V_{th} and $V_{th,sat}$ by 465 mV as seen in Figures 2.21ae and 2.22ae. The effect of drain-induced barrier lowering is

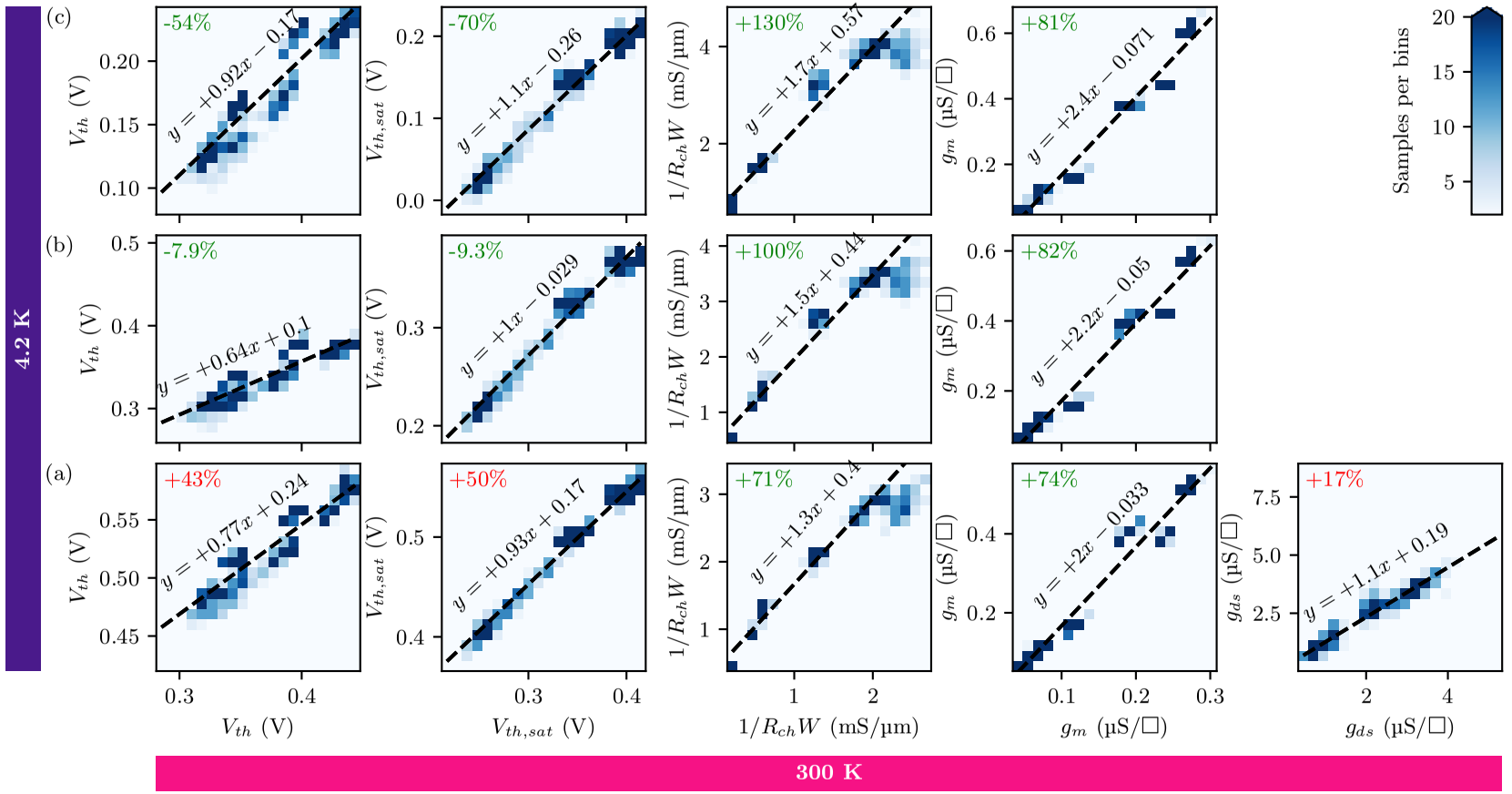


Figure 2.21 | Impact of the temperature and back-gate on the extracted N-type single-device parameters.

The extracted quantities of V_{th}, $V_{th,sat}$, $1/R_{ch}$, g_m, and g_{ds} presented in this chapter for 1000 thin-oxide N-type devices are represented as 2D histograms with column values at 300 K and row values at 4.2 K for increasing forward body-biasing at (a) 0 V, (b) 2 V, and (c) 4 V. The color-scale shows bins with a high number of sample points in darker colors. Above 20 samples the color is unchanged. For all subplots, a general trend of the correlations between data at 300 and 4.2 K is extracted by fitting a linear dependence (black dashed lines) with the given equation where x (respectively y) is the value at 300 K (resp. 4.2 K). In the upper left corner of the graphs, the relative evolution of the parameters from 300 to 4.2 K averaged over all data points is shown in green for improvement and in red for worsening in percentage.

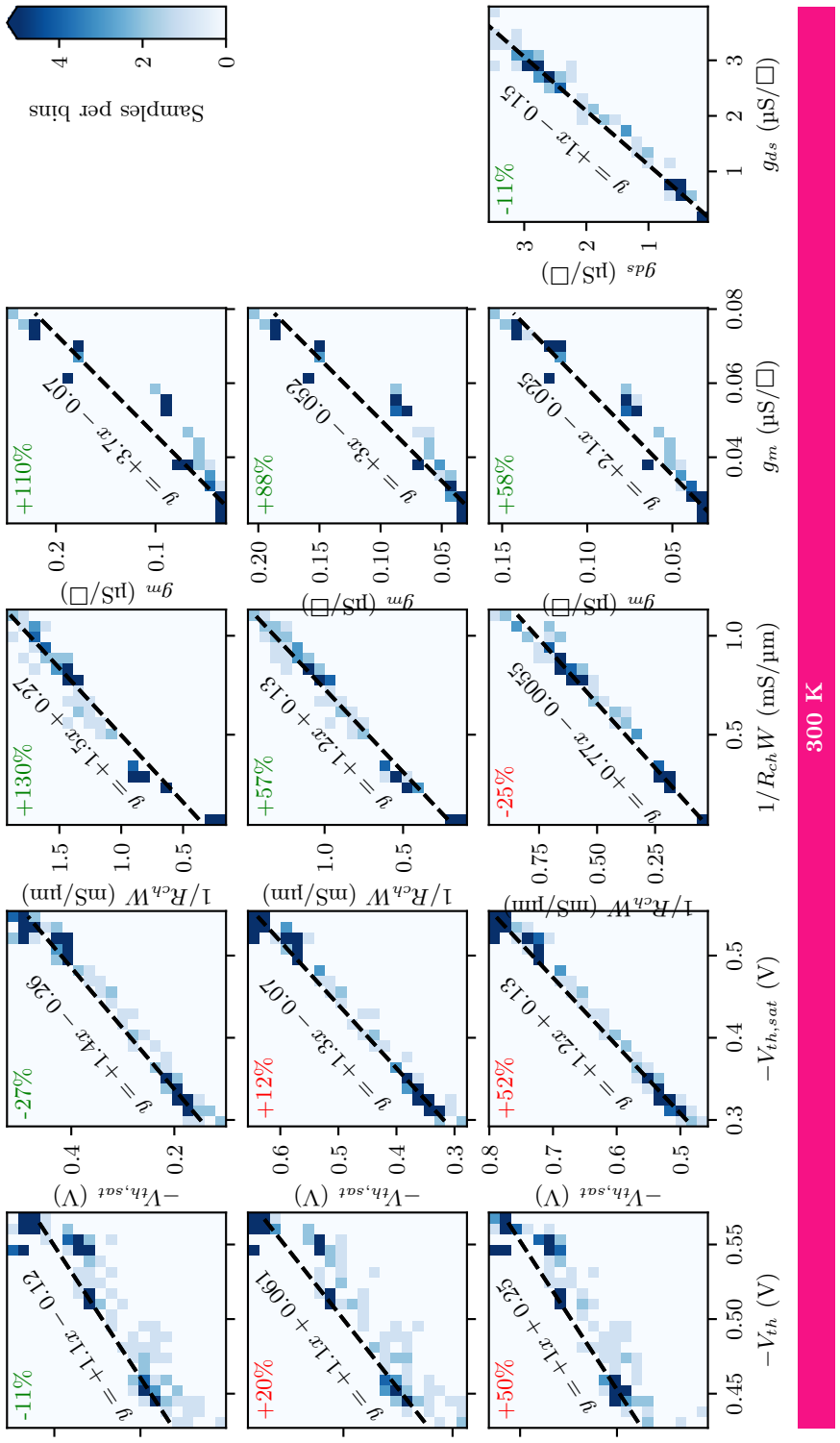


Figure 2.22 | Impact of the temperature and back-gate on the extracted P-type single-device parameters.
The extracted quantities of V_{th}, $V_{th,sat}$, $1/R_{ch}$, g_m, and g_{ds} presented in this chapter for 200 thin-oxide P-type devices are represented as 2D histograms with column values at 300 K and row values at 4.2 K for increasing forward body-biasing at (a) 0 V, (b) 2 V, and (c) 4 V. The color-scale shows bins with a high number of sample points in darker colors. Above 5 samples the color is unchanged. For all subplots, a general trend of the correlations between data at 300 and 4.2 K is extracted by fitting a linear dependence (black dashed lines) with the given equation where x (respectively y) is the value at 300 K (resp. 4.2 K). In the upper left corner of the graphs, the relative evolution of the parameters from 300 to 4.2 K averaged over all data points is shown in green for improvement and in red for worsening in percentage.

observed in Figures 2.21b and 2.22b by the banana-shaped trace slightly offset to the reported fit of V_{th}. The channel conductance $1/R_{ch}$ is nicely linear with V_{th} in Figures 2.21c and 2.22c. The V_{th} reduction with FBB leads to an increase of $1/R_{ch}$ with FBB as observed in Figures 2.21h and 2.22h. As R_{ch} is extracted in linear regime and linear to V_{th}, we can also observe the DIBL effect on R_{ch} in Figures 2.21f and 2.22f. The transconductance g_m follows the classical relation $g_m \propto 1/(V_{gs} - V_{th})$ as shown in Figures 2.21g and 2.22g. From these plots, we observe that the FBB only results in a shift of V_{th} that impacts subsequently the other quantities. No strong effect of the FBB was observed on e.g. the mobility.

No major changes are observed upon reducing the temperature from 4.2 to 0.1 K.Only long N-type devices at high FBB exhibits higher V_{th} at 0.1 K that might come from the N-type substrate becoming less conductive combined with potentially too distant metallic taps into the well.

VII From single devices to circuits

In circuits, several transistors are assembled in different topologies depending on the circuit task like e.g. signal amplification (see Figure 2.25a for two-transistor stages). Transistors are often stacked between the two power-supply voltages V_{DD} and V_{SS}. To maintain the device biasing in saturation, the gate-source and source-drain voltages have to respect the saturation inequalities $V_{gs} > V_{th}$ and $V_{ds} > V_{gs} - V_{th}$. As the transistors bias voltages and currents become interdependent in circuits, it is more complicated to predict the temperature dependence of circuit performance from single device characteristics only. In this section, we study meaningful quantities for circuit design as a function of temperature and forward body biasing.

a) Intrinsic voltage gain

Transistors in saturation behave as a current source of current $g_m V_{gs}$ with a parallel conductance g_{ds}. From the Norton-Thévenin equivalence, the transistor can be seen as a non-ideal voltage source of voltage $A_V V_{gs}$ and output impedance of $1/g_{ds}$ when biased with an ideal current source. In this case, the transistor behaves as a voltage amplifier of gain A_V with $A_V = g_m/g_{ds}$ the intrinsic voltage gain. A_V is an interesting quantity for transistors as it sets an upper bound on the voltage amplification of one single transistor in perfect biasing condition with ideal current-source. A_V is the voltage gain of a common-source amplifier biased with a perfect current source (see Figure 2.25b). In reality, non-ideal current sources made of transistors lowers the gain A_V by adding a contribution to the amplifying transistor conductance g_{ds}.

Extraction of intrinsic voltage gain

The intrinsic voltage gain $A_V = g_m/g_{ds}$ is computed from the extracted data of g_{ds} and g_m at the bias point $V_{ds} = V_{gs} = 0.9\,\mathrm{V}$. g_{ds} is the same quantity as in section c) but g_m is extracted at constant V_{gs} contrary to the presented $g_{m,max}$ in section b). A_V depends on the biasing condition and generally is maximized at lower voltage bias as g_m peaks at lower V_{gs} voltages while g_{ds} increases at lower V_{gs}. A_V could be characterized as a function of the bias point with the designed matrix but this would require new measurements. From the acquired data, we extract $A_V = g_m/g_{ds}$ at high bias $V_{ds} = V_{gs} = 0.9\,\mathrm{V}$ to get insights on the amplification properties of transistors at cryogenic temperatures.

Temperature dependence of intrinsic voltage gain

In Figure 2.26, we present the intrinsic voltage gain A_V for (a) PMOS and (b) NMOS as a function of the transistor gate-length L for temperatures of 300, 4.2, and 0.1 K. The intrinsic voltage gain increases with increasing length due to the stronger dependence of $g_{ds} \propto L^{-2}$ on L than $g_m \propto L^{-1}$. A_V at 4.2 K

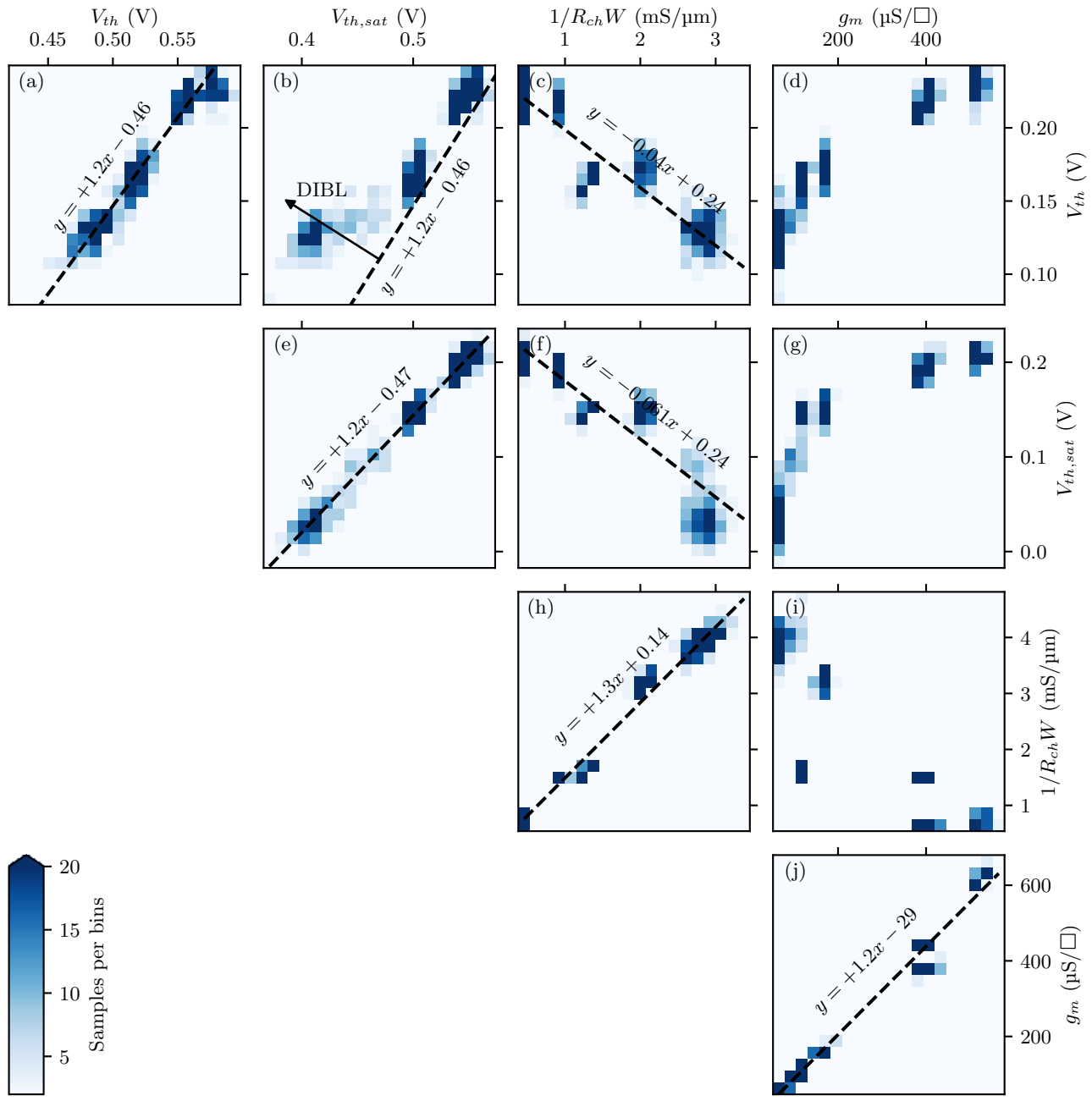


Figure 2.23 | Correlations between different body bias voltage for the extracted quantities at 4.2 K.

The subplots correspond to the 2D histograms of all extracted quantities of 1000 N-type thin-oxide devices. The x-axis corresponds to values with no FBB while the y-axis are the values at a high FBB of 4 V. General trend curves are represented as dashed lines further discussed in the text. From this graph, we can deduce that positive FBB mainly contributes to lower V_{th} with no major change on e.g. the mobility.

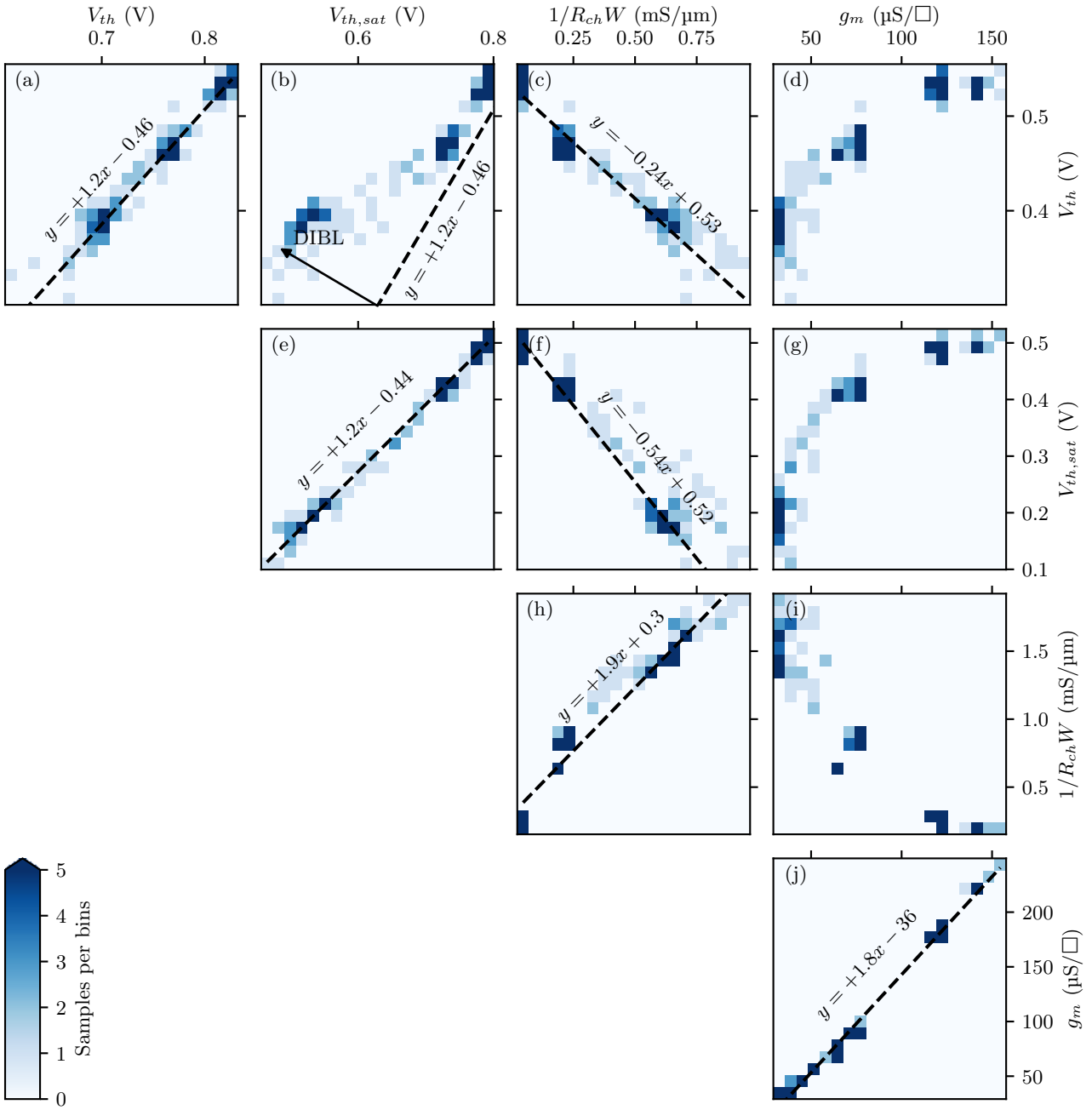


Figure 2.24 | Correlations between different body bias voltage for the extracted quantities at 4.2 K.
The subplots correspond to the 2D histograms of all extracted quantities of 200 P-type thin-oxide devices. The x-axis corresponds to values with no FBB while the y-axis are the values at a high FBB of 4 V. General trend curves are represented as dashed lines further discussed in the text. From this graph, we can deduce that positive FBB mainly contributes to lower V_{th} with no major change on e.g. the mobility.

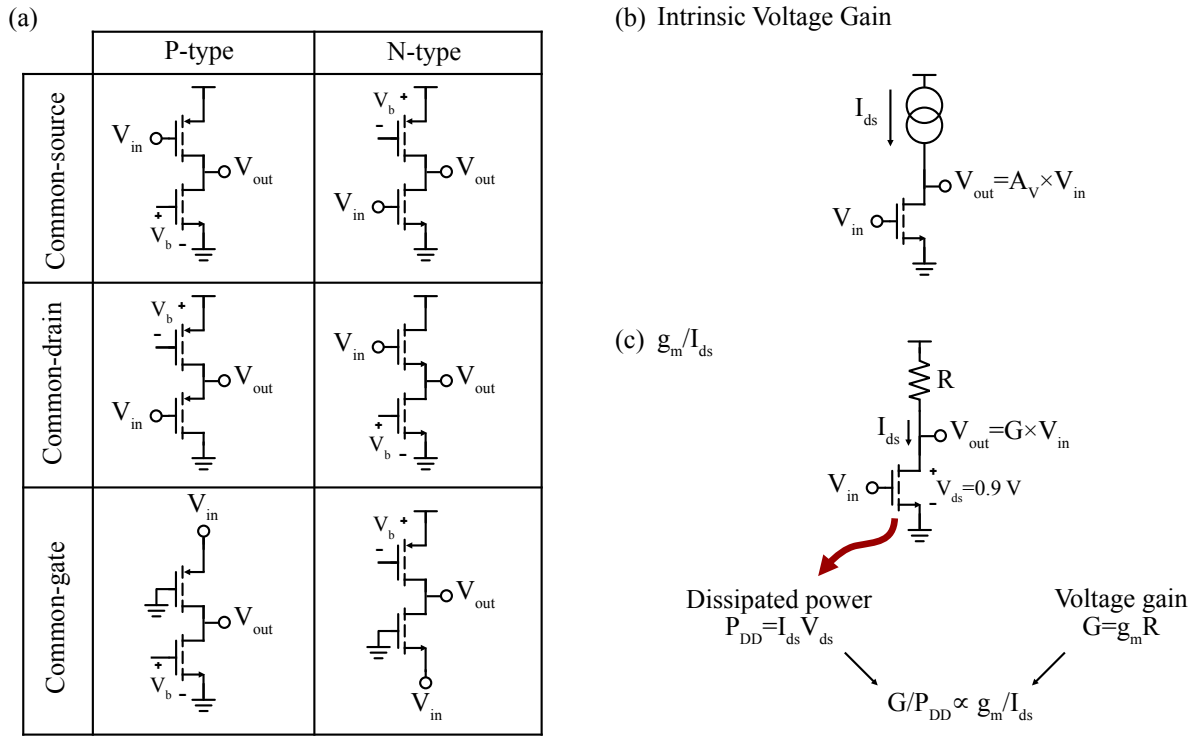


Figure 2.25 | From transistor to circuits.

(a) The 6 two-port two-transistor assemblies made of N- and P-type transistors. V_{in} is the input and V_{out} is the output voltage of the stage. V_b is a DC voltage ensuring that the devices operate in saturation. The common-source architecture is often use for signal amplification while the common-drain is mostly used for signal following and buffering. More architectures with >2 transistor are also possible. (b) The intrinsic voltage gain A_V corresponds to the voltage gain of a common-source (N-type shown) stage with ideal current source. Real-world current source are non-ideal and lowers the reachable voltage gain, thus A_V is the maximum voltage gain reachable for a given device. (b) The energy efficiency of a simple one transistor stage with resistive load defined as the ratio of the voltage gain G over the dissipated power P_{DD} is proportional to the g_m/I_{ds} ratio. The higher g_m/I_{ds} is synonym of a higher energy efficiency.

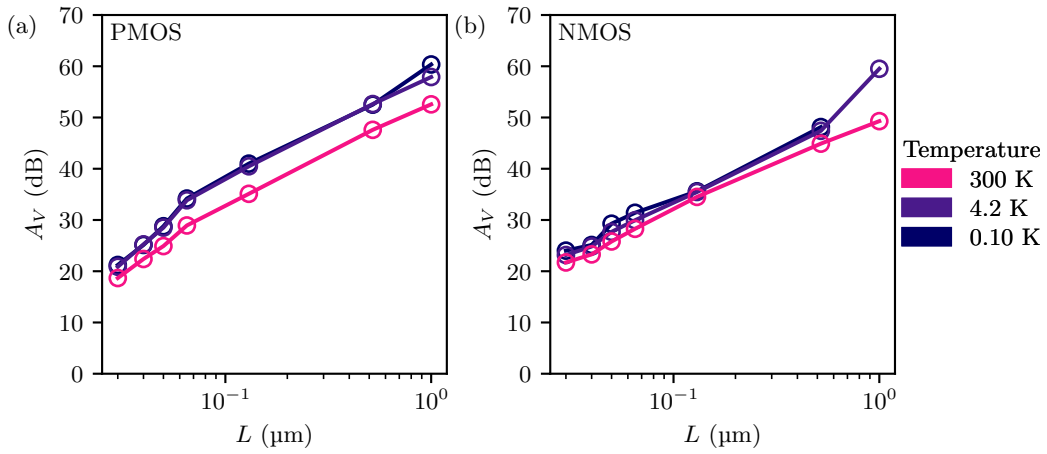


Figure 2.26 | Temperature dependence of the intrinsic voltage gain.

Intrinsic voltage gain $A_V = g_m/g_{ds}$ at strong bias $V_{ds} = V_{gs} = 0.9\,\mathrm{V}$ of (a) P-type and (b) N-type single-transistor devices as a function of gate-length L at temperatures ranging from 0.1 to 300 K.

and 0.1 K is significantly higher than at room-temperature thanks to the g_m increase (see section b)) with an increase of the gain from 18 to 33 % for N-type devices and from 28 to 84 % for P-type devices. The increase of A_V represents a gain of 1.5 to 5.3 dB leading to a gain at 4.2 K of 21 dB (respectively 23 dB) for $L = 28\,\mathrm{nm}$ to 53 dB (resp. 47 dB) at $L = 600\,\mathrm{nm}$ for P-type (resp. N-type) devices. As already observed for all other quantities, the voltage gain remains identical at 4.2 and 0.1 K.

As we already observed that the g_m peaks at higher values at cryogenic temperatures for lower V_{gs} voltages, we expect A_V to further increase with respect to room-temperature values at lower V_{gs}. The higher intrinsic voltage gain of transistor will improve the amplifiers gain at cryogenic temperature.

Circuits, such as voltage amplifiers, placed at cryogenic temperature must not dissipate more power than the cooling capacity of the fridge. It is instructive to look at the energy efficiency of single devices at cryogenic temperatures to deduce if circuits might consume higher power at low temperatures than at room-temperature. A good quantity to measure energy efficiency of single devices is the g_m/I_{ds} ratio that tells how much current-to-voltage gain we can expect for a given source-drain current. Higher source-drain currents often leads to higher power consumption and higher g_m such that g_m/I_{ds} depends only weakly on the device geometry but could strongly depend on temperature. At low-temperatures, we know that g_m increases but so do I_{ds}. In the next section, we look at the g_m/I_{ds} as a function of temperature to get an idea of the power-efficiency of single transistors at cryogenic temperatures.

b) Design considerations with $\mathbf{g_m/I_{ds}}$

Power consumption $P_{DD} = I_0(V_{DD} - V_{SS})$ of circuits is easily computed from the sum of all bias currents I_0 between V_{DD} and V_{SS}. Power dissipated by the cryogenic circuits must be compensated by the fridge cooling power to avoid heating above base temperature. In circuits, P_{DD} is proportional to the drain-source current I_{ds} that is imposed by the circuit designer given the aimed circuit specifications. For circuit design, it is interesting to focus on the dependence of transistor parameters on the source-drain current I_{ds}. Circuit design methodologies exploiting the current-dependence of quantities allow to easily design circuits using simple calculations despite increasing non-ideal transistor behavior present in modern

technologies. Initially, the methodology was derived for the prediction of g_m for a given bias current I_{ds}, hence the name "GM/ID methodology"[27]. This methodology has been developed to take into account other important parameters such as parasitic capacitance or conductance[28].

The GM/ID methodology is based on the quantity g_m/I_{ds} that has only power-law dependence on bias voltages V_{gs} and V_{ds} and weak dependence on device geometry L and W, even in subthreshold regime as both for g_m and I_{ds} the exponential dependence cancels out[27]. g_m/I_{ds} is independent of the width W and only second-order dependent on L due to short-channel effects (as both g_m and I_{ds} are proportional to first order to W/L). Moreover, the g_m/I_{ds}methodology can be extended to a wide temperature range[29], making it usable in harsh environments. This capacity to use one quantity in all regimes and for all widths allow designing circuit with transistors even in subthreshold that leads to lower current I_{ds} hence lower power consumption P_{DD}. g_m/I_{ds} can also be seen as the transistor gain for a given power consumption P_{DD}, hence a kind of power efficiency of transistors (see Figure 2.25c).

To give an example, we assume that we aim a power consumption P_{DD} of 100 µW and transconductance g_m of 1 mS for some circuit. We know that I_{ds} has to be equal to 100 µA with $V_{DD} = 1\,\mathrm{V}$ hence g_m/I_{ds} is equal to $10\,\mathrm{V}^{-1}$. Selecting a g_m/I_{ds} of 10 gives different values of L/W that all meet the previous criteria of transconductance and power. Generally, we also want a given conductance g_{ds} that defines the transistor length L. We can then deduce the required transistor width W from the known g_m/I_{ds} curves. Different design strategies depending on the design goal (power first, gain first,...) can be implemented.

g_m/I_{ds} curves at 4.2 K

In Figure 2.27, we show the dependence of g_m/I_{ds} averaged across all widths W as a function of the current density $I_{ds} \times L/W$ for multiple lengths L for for some transistors studied at 4.2 K. Long-channel devices with $L \geq 1\,\mu\mathrm{m}$ exhibit the typical inverse square-root dependence with I_{ds} at strong gate-source voltage $V_{gs} \gg V_{th}$ as seen in Figure 2.27. Indeed, long-channel transistors often exhibit a square-law dependence with $V_{gt} = V_{gs} - V_{th}$ according to the simplest model of I_{ds}:

$$I_{ds} \propto W/L\mu(V_{gs} - V_{th})^2 \tag{2.5}$$

with μ the constant effective mobility of charge carriers for long-channel devices. Given equation 2.5, g_m/I_{ds} is proportional to $I_{ds}^{-1/2}$ using $g_m \propto I_{ds}^{1/2}$.

With smaller length, g_m/I_{ds} decreases more rapidly with I_{ds} as the mobility of charge carrier is no longer independent of V_{gt} but decreases with increasing V_{gt} (equivalently higher I_{ds}). The power-law exponent of g_m/I_{ds} with I_{ds} changes from -0.5 for long devices down to -1.6 (respectively -0.9) for P-type (resp. N-type) 28nm gate-length devices. The increased access resistance and surface roughness in shorter gate transistors accelerate the decrease of μ with V_{gt} (equivalently I_{ds}), leading to the observed power-law exponents < -0.5.

Strong oscillations in g_m/I_{ds} appear in the subthreshold regime for $V_{gs} < V_{th,sat}$ at cryogenic temperature, reminiscent of the subthreshold swing oscillations. With I_{ds} following the typical exponential behavior as in equation 2.4, g_m/I_{ds} becomes equal to the inverse of the thermal voltage V_T^{-1} (assuming a gate level-arm of 1), thus inversely proportional to the subthreshold swing.

Temperature & back-gate impact on the g_m/I_{ds} ratio

The value of g_m/I_{ds} in Moderate Inversion (MI), noted $(g_m/I_{ds})_{MI}$, is extracted at a constant current density $I_{ds}^{\square}$ of $1 \times 10^{-6}\,\mathrm{A}/\square$ (resp. $3 \times 10^{-6}\,\mathrm{A}/\square$) for PMOS (resp. NMOS) and are shown in Figure 2.28ab. The maximum $(g_m/I_{ds})_{MI}$ is $27\,\mathrm{V}^{-1}$ (resp. $16\,\mathrm{V}^{-1}$) for NMOS (resp. PMOS) at 300 K for the longer

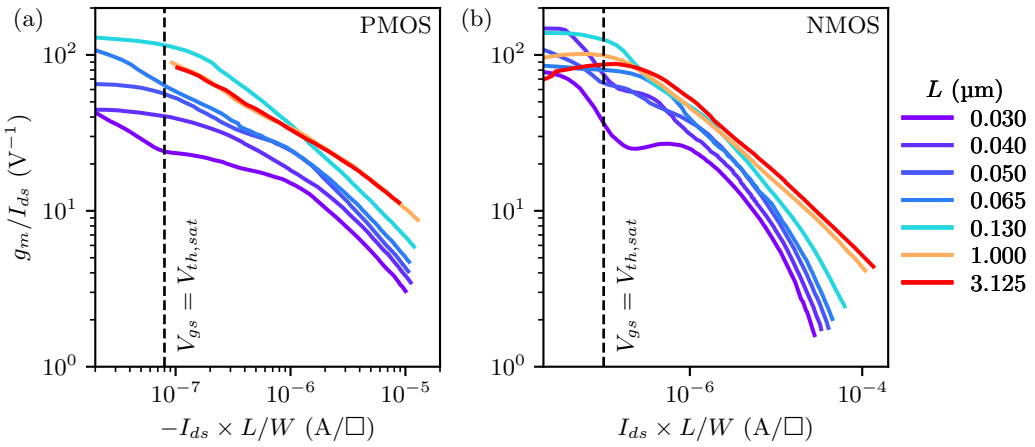


Figure 2.27 | g_m/I_{ds} with varying transistor length at 4.2 K.

g_m/I_{ds} at $|V_{ds}| = 0.9\,\mathrm{V}$ as a function of the current density $I_{ds}^{\square} = I_{ds} \times L/W$ for (a) PMOS and (b) NMOS at 4.2 K with width of 0.32 µm and lengths from 30 nm to 3.125 µm.

device with $L = 1\,\mu\mathrm{m}$. This value increases by ×1.8 (resp. ×2.2) from 300 to 4.2 K. The back-gate further increases $(g_m/I_{ds})_{MI}$ up to $74\,\mathrm{V}^{-1}$ (resp. $36\,\mathrm{V}^{-1}$) for NMOS (resp. PMOS).

VIII Conclusion

In this chapter, we acquired knowledge on single devices of the CMOS FD-SOI 28nm technology at cryogenic temperatures. We extracted meaningful quantities of transistors for the design of circuits at a few special bias points. Circuits are composed of multiple transistors assembled in a certain topology to achieve a special task like e.g. signal amplification and signal following. Each transistor relies on the other to keep a sufficient source-drain and gate-source voltage that might strongly deviate from the special bias points used for single-device parameter extraction.

To gain insights on the operation of circuits at cryogenic temperatures, we characterized and analyzed circuits made of 4 to 100's of transistors from room to cryogenic temperature in the following chapter. We highlight the change of circuit specifications with temperature with respect to the single-device quantities extracted here.

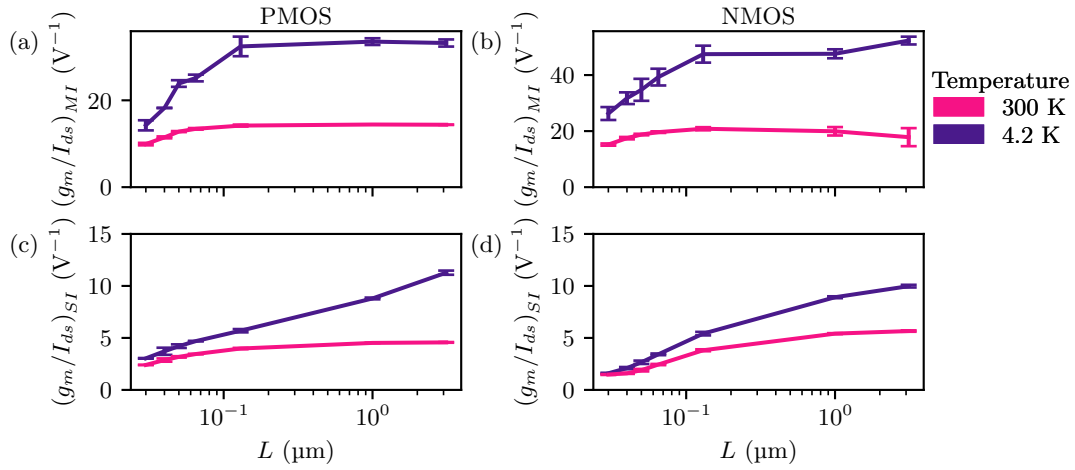


Figure 2.28 | g_m/I_{ds} in moderate and strong inversion at 300 and 4.2 K for various FBB conditions.

g_m/I_{ds} in moderate (respectively strong) inversion as a function of transistor gate length L for (a) PMOS and (b) NMOS (resp. (c) PMOS and (d) NMOS) devices with 0.32 µm width for temperatures of 300 and 4.2 K at 0 V FBB.

[FR] Caractérisation basse-fréquence massive des transistors FD-SOI

La conception de circuits nécessite des modèles précis et rapides des éléments actifs et passifs au sein de la technologie choisie afin de comprendre et de concevoir des systèmes complexes. Ces modèles prennent en compte des centaines de phénomènes physiques : de la physique bien connue des semi-conducteurs à par exemple les phénomènes de quantification au sein des matériaux de taille nanométrique. Alors que les transistors continuent à rapetisser suivant la loi de Moore jusqu'à quelques nanomètres, le nombre de paramètres de ces modèles a explosé pour modéliser la physique riche.

Pour exploiter ces modèles à des fin de conception de circuits intégrés, il est nécessaire de déterminer les paramètres du modèle à partir de mesures électriques précises des caractéristiques des transistors individuels. Les modèles basses fréquences constituent la base de modèles plus complexes prenant en compte les condensateurs et inductances parasites supplémentaires pour les simulations hautes fréquences. Plusieurs itérations à travers une procédure optimisée sont nécessaires pour converger vers un ensemble de paramètres physiquement raisonnables et précis. Un grand nombre de mesures des transistors est requis afin d'isoler les différentes contributions (canaux courts vs longs, larges vs étroits,...) et d'extraire des bons paramètres.

Mesurer les propriétés intrinsèques des dispositifs CMOS peut être difficile car l'accès physique aux terminaux nécessite des étapes de fabrication supplémentaires (par exemple, acheminement du métal et siliciuration des contacts) qui ajoutent des composantes extrinsèques à la mesure, tels que les résistances d'accès. Pour ajuster un modèle physique sur les mesures, l'effet du Back-End Of Line (BEOL) sur les caractéristiques des dispositifs doit être pris en compte ou compensé grâce à des calibrations. Les dispositifs typiques des technologies de silicium CMOS modernes nécessitent de mesurer des courants inférieurs au mA avec plusieurs tensions de polarisation de l'ordre du volt (par exemple source, drain, grille et substrat).

Pour caractériser de nombreux transistors à température ambiante, une station sous points automatisée mesure rapidement tous les éléments un par un en déplaçant les pointes, souvent sur un wafer entier. Cette méthode permet de caractériser précisément des milliers de dispositifs sur un même wafer en quelques heures. Bien que des stations sous points cryogéniques aient été récemment développées[2], elles restent coûteuses et limitées à des températures au-dessus de quelques Kelvins.

Le cycle de refroidissement inhéremment long des systèmes cryogéniques de l'ordre de quelques heures pour des dip-stick jusqu'à plus de 24h pour des réfrigérateurs a dilution rend la caractérisation de milliers d'éléments CMOS à des températures cryogéniques quasiment irréalisable. Inspiré par des études de désappareillement local entre transistors identiques à température ambiante qui multiplexent un grand nombre de transistors proches les uns des autres[3, 4, 5], plusieurs groupes ont développé des matrices de multiplexage adaptées aux températures cryogéniques pour la caractérisation basse-fréquence de 66 dispositifs par matrice en 2018 [6, 7] et, plus tard, notre réalisation avec 1024 dispositifs par matrice publiée en 2020 [8, 9]. Dans ce chapitre, nous décrivons et discutons l'implémentation de notre matrice de multiplexage composée de 1024 dispositifs qui fonctionne jusqu'à 0.1 K faites avec la technologie CMOS FD-SOI 28 nm[8, 9]. À partir des caractéristiques DC obtenues, nous dérivons et étudions des quantités clefs des transistors en fonction de la température et du régime de conduction, telles que la transconductance, la conductance et le ratio de transconductance sur le courant drain-source. La réalisation et la mesure de cette matrice de multiplexage nous as permis d'acquérir énormément de connaissances sur les dispositifs CMOS individuels de la technologie CMOS FD-SOI 28nm à des températures cryogéniques.

Les circuits sont composés de plusieurs transistors assemblés dans une certaine topologie pour réaliser une tâche spéciale comme par ex. amplifier ou suivre un signal. Chaque transistor devient dépendant des autres au sein d'un même circuit, influençant les points de polarisation comme les tensions de source, de drain, et de grille. Pour mieux comprendre le fonctionnement des circuits à des températures cryogéniques, nous avons caractérisé et analysé des circuits constitués de 4 à 100 transistors des températures ambiantes

jusqu'aux températures cryogéniques dans le chapitre suivant. Nous mettons en relation les changements de performance des circuits étudiés en fonction de la température avec l'évolution des caractéristiques des transistors individuels extraites dans ce chapitre.

Bibliography

[1] Michael B. Heaney. **Electrical Conductivity and Resistivity**. *Electrical measurement, signal processing, and displays*, January 2003. Available from: `https://www.academia.edu/29112469/Electrical_Conductivity_and_Resistivity`.

[2] R. Pillarisetty, H. C. George, T. F. Watson, L. Lampert, N. Thomas, S. Bojarski, P. Amin, R. Caudillo, E. Henry, N. Kashani, P. Keys, R. Kotlyar, F. Luthi, D. Michalak, K. Millard, J. Roberts, J. Torres, O. Zietz, T. Krähenmann, A. M. Zwerver, M. Veldhorst, G. Scappucci, L. M. K. Vandersypen, and J.S. Clarke. **High Volume Electrical Characterization of Semiconductor Qubits**. In *2019 IEEE International Electron Devices Meeting (IEDM)*, pages 31.5.1–31.5.4, December 2019.

[3] S.B. Yeo, J. Bordelon, S. Chu, M.F. Li, B.A. Tranchina, M. Harward, L.H. Chan, and A. See. **A robust and production worthy addressable array architecture for deep sub-micron MOSFET's matching characterization**. In *Proceedings of the 2002 International Conference on Microelectronic Test Structures, 2002. ICMTS 2002.*, pages 229–234, April 2002.

[4] Y. Shimizu, M. Nakamura, T. Matsuoka, and K. Taniguchi. **Test structure for precise statistical characteristics measurement of MOSFETs**. In *Proceedings of the 2002 International Conference on Microelectronic Test Structures, 2002. ICMTS 2002.*, pages 49–54, April 2002.

[5] M. Quarantelli, S. Saxena, N. Dragone, J.A. Babcock, C. Hess, S. Minehane, S. Winters, Jianjun Chen, H. Karbasi, and C. Guardiani. **Characterization and modeling of MOSFET mismatch of a deep submicron technology**. In *International Conference on Microelectronic Test Structures, 2003.*, pages 238–243, March 2003.

[6] P. A. 't Hart, J. P. G. van Dijk, M. Babaie, E. Charbon, A. Vladimircscu, and F. Sebastiano. **Characterization and Model Validation of Mismatch in Nanometer CMOS at Cryogenic Temperatures**. In *2018 48th European Solid-State Device Research Conference (ESSDERC)*, pages 246–249, September 2018.

[7] P. A. t Hart, M. Babaie, E. Charbon, A. Vladimirescu, and F. Sebastiano. **Subthreshold Mismatch in Nanometer CMOS at Cryogenic Temperatures**. In *ESSDERC 2019 - 49th European Solid-State Device Research Conference (ESSDERC)*, pages 98–101, September 2019.

[8] Bruna Cardoso Paz, Loïck Le Guevel, Mikaël Cassé, Gérard Billiot, Gaël Pillonnet, Aloysius Jansen, Sebastien Haendler, Andre Juge, Emmanuel Vincent, Philippe Galy, Gérard Ghibaudo, Maud Vinet, Silvano de Franceschi, Tristan Meunier, and Fred Gaillard. **Integrated Variability Measurements of 28 nm FDSOI MOSFETs down to 4.2 K for Cryogenic CMOS Applications**. In *2020 IEEE 33rd International Conference on Microelectronic Test Structures (ICMTS)*, pages 1–5, May 2020.

[9] B. Cardoso Paz, L. Le Guevel, M. Cassé, G. Billiot, G. Pillonnet, A. G. M. Jansen, R. Maurand, S. Haendler, A. Juge, E. Vincent, P. Galy, G. Ghibaudo, M. Vinet, S. de Franceschi, T. Meunier, and F. Gaillard. **Variability Evaluation of 28nm FD-SOI Technology at Cryogenic Temperatures down to 100mK for Quantum Computing**. In *2020 IEEE Symposium on VLSI Technology*, pages 1–2, June 2020.

[10] A.W. Van Herwaarden and P.M. Sarro. **Thermal sensors based on the seebeck effect**. *Sensors and Actuators*, **10**(3-4):321–346, November 1986. Available from: `https://linkinghub.elsevier.com/retrieve/pii/0250687486800531`.

[11] A. Pergoot, B. Graindourze, E. Janssens, J. Bastos, M. Steyaert, P. Kinget, R. Roovers, and W. Sansen. **Statistics for matching**. In *Proceedings International Conference on Microelectronic Test Structures*, pages 193–197, March 1995.

[12] T. Grotjohn and B. Hoefflinger. **A parametric short-channel MOS transistor model for subthreshold and strong inversion current**. *IEEE Transactions on Electron Devices*, **31**(2):234–246, February 1984.

[13] Y. Khatami and K. Banerjee. **Steep Subthreshold Slope n- and p-Type Tunnel-FET Devices for Low-Power and Energy-Efficient Digital Circuits**. *IEEE Transactions on Electron Devices*, **56**(11):2752–2761, November 2009.

[14] A. Beckers, F. Jazaeri, H. Bohuslavskyi, L. Hutin, S. De Franceschi, and C. Enz. **Design-oriented modeling of 28 nm FDSOI CMOS technology down to 4.2 K for quantum computing**. In *2018 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, pages 1–4, March 2018.

[15] H. Bohuslavskyi, S. Barraud, M. Cassé, V. Barrai, B. Bertrand, L. Hutin, F. Arnaud, P. Galy, M. Sanquer, S. De Franceschi, and M. Vinet. **28nm Fully-depleted SOI technology: Cryogenic control electronics for quantum computing**. In *2017 Silicon Nanoelectronics Workshop (SNW)*, pages 143–144, June 2017.

[16] H. Bohuslavskyi, A. G. M. Jansen, S. Barraud, V. Barral, M. Cassé, L. Le Guevel, X. Jehl, L. Hutin, B. Bertrand, G. Billiot, G. Pillonnet, F. Arnaud, P. Galy, S. De Franceschi, M. Vinet, and M. Sanquer. **Cryogenic Subthreshold Swing Saturation in FD-SOI MOSFETs Described With Band Broadening**. *IEEE Electron Device Letters*, **40**(5):784–787, May 2019. Available from: `https://ieeexplore.ieee.org/abstract/document/8660508`.

[17] A. Ortiz-Conde, F.J. García Sánchez, J.J. Liou, A. Cerdeira, M. Estrada, and Y. Yue. **A review of recent MOSFET threshold voltage extraction methods**. *Microelectronics Reliability*, **42**(4-5):583–596, April 2002. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S0026271402000276`.

[18] Kazuo Terada, Katsuhiko Nishiyama, and Kei-Ichi Hatanaka. **Comparison of MOSFET-threshold-voltage extraction methods**. page 6, 2001.

[19] Bruna Cardoso Paz, Mikaël Cassé, Christoforos Theodorou, Gérard Ghibaudo, Thorsten Kammler, Luca Pirro, Maud Vinet, Silvano de Franceschi, Tristan Meunier, and Fred Gaillard. **Performance and Low-Frequency Noise of 22-nm FDSOI Down to 4.2 K for Cryogenic Applications**. *IEEE Transactions on Electron Devices*, pages 1–5, 2020.

[20] A. Beckers, F. Jazaeri, A. Ruffino, C. Bruschini, A. Baschirotto, and C. Enz. **Cryogenic characterization of 28 nm bulk CMOS technology for quantum computing**. In *2017 47th European Solid-State Device Research Conference (ESSDERC)*, pages 62–65, September 2017.

[21] Arnout Beckers, Farzan Jazaeri, Alexander Grill, Subramanian Narasimhamoorthy, Bertrand Parvais, and Christian Enz. **Physical Model of Low-Temperature to Cryogenic Threshold Voltage in MOSFETs**. *IEEE Journal of the Electron Devices Society*, **8**:780–788, 2020.

[22] Arnout Beckers, Farzan Jazaeri, and Christian Enz. **Revised theoretical limit of the subthreshold swing in field-effect transistors**. *arXiv:1811.09146 [cond-mat]*, November 2018. Available from: `http://arxiv.org/abs/1811.09146`.

[23] Joseph C. Bardin, Evan Jeffrey, Erik Lucero, Trent Huang, Sayan Das, Daniel Thomas Sank, Ofer Naaman, Anthony Edward Megrant, Rami Barends, Ted White, Marissa Giustina, Kevin J. Satzinger, Kunal Arya, Pedram Roushan, Benjamin Chiaro, Julian Kelly, Zijun Chen, Brian Burkett, Yu Chen, Andrew Dunsworth, Austin Fowler, Brooks Foxen, Craig Gidney, Rob Graff, Paul Klimov, Josh Mutus, Matthew J. McEwen, Matthew Neeley, Charles J. Neill, Chris Quintana, Amit Vainsencher, Hartmut Neven, and John Martinis. **Design and Characterization of a 28-nm Bulk-CMOS Cryogenic Quantum Controller Dissipating Less Than 2 mW at 3 K**. *IEEE Journal of Solid-State Circuits*, **54**(11):3043–3060, November 2019.

[24] Mikaël Cassé, Bruna Cardoso Paz, Gérard Ghibaudo, Thierry Poiroux, Emmanuel Vincent, Philippe Galy, André Juge, Fred Gaillard, Silvano de Franceschi, Tristan Meunier, and Maud Vinet. **Evidence of 2D intersubband scattering in thin film fully depleted silicon-on-insulator transistors operating at 4.2 K**. *Applied Physics Letters*, **116**(24):243502, June 2020. Available from: `http://aip.scitation.org/doi/10.1063/5.0007100`.

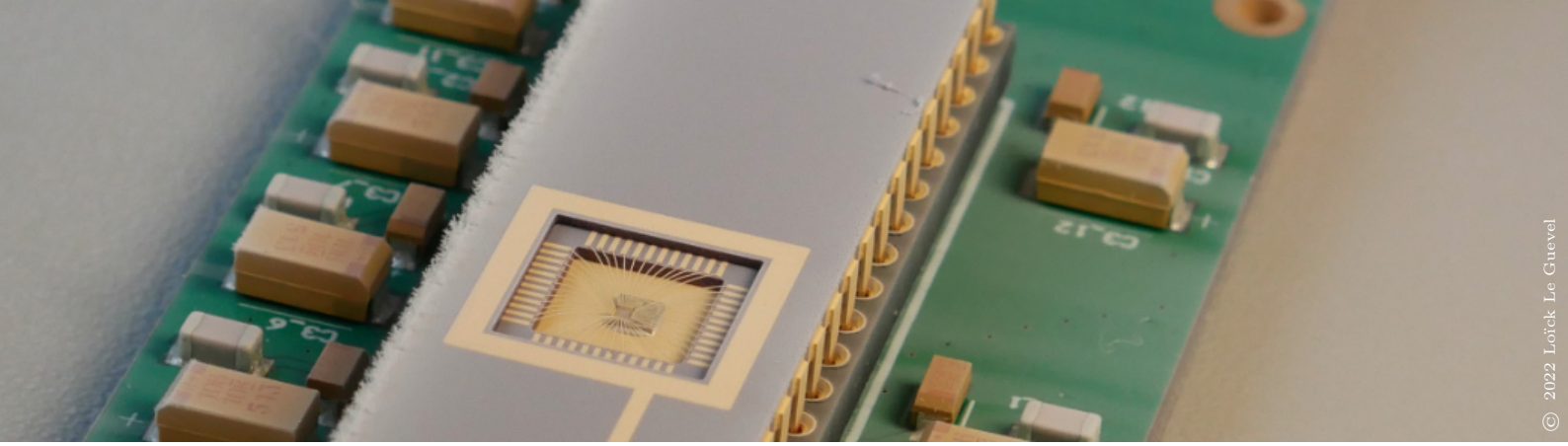
[25] Abraham. Savitzky and M. J. E. Golay. **Smoothing and Differentiation of Data by Simplified Least Squares Procedures.** *Analytical Chemistry*, **36**(8):1627–1639, July 1964. Available from: `https://pubs.acs.org/doi/abs/10.1021/ac60214a047`.

[26] H. Bohuslavskyi, S. Barraud, V. Barral, M. Cassé, L. Le Guevel, L. Hutin, B. Bertrand, A. Crippa, X. Jehl, G. Pillonnet, A. G. M. Jansen, F. Arnaud, P. Galy, R. Maurand, S. De Franceschi, M. Sanquer, and M. Vinet. **Cryogenic Characterization of 28-nm FD-SOI Ring Oscillators With Energy Efficiency Optimization**. *IEEE Transactions on Electron Devices*, **65**(9):3682–3688, September 2018.

[27] F. Silveira, D. Flandre, and P. G. A. Jespers. **A g/sub m//I/sub D/ based methodology for the design of CMOS analog circuits and its application to the synthesis of a silicon-on-insulator micropower OTA**. *IEEE Journal of Solid-State Circuits*, **31**(9):1314–1319, September 1996.

[28] Mostafa N. Sabry, Hesham Omran, and Mohamed Dessouky. **Systematic design and optimization of operational transconductance amplifier using gm/ID design methodology**. *Microelectronics Journal*, **75**:87–96, May 2018. Available from: `https://linkinghub.elsevier.com/retrieve/pii/S0026269217307905`.

[29] João Roberto Raposo De Oliveira Martins, Ali Mostafa, Jérome Juillard, Rachid Hamani, Francisco De Oliveira Alves, and Pietro Maris Ferreira. **A Temperature-Aware Framework on gm/ID-Based Methodology Using 180 nm SOI From -40°C to 200°C**. *IEEE Open Journal of Circuits and Systems*, **2**:311–322, 2021.



CHAPTER 3
Low-temperature Impact on Circuits

I Introduction

In an electrical circuit, the transistors sharing voltage nodes and currents allow to implement various functionalities (e.g. amplification) depending on the topology. The circuit topology induces interdependence between transistor characteristics that leads to complicated analysis asking for an accurate model of transistor behavior. To further understand the impact of cryogenic temperatures for the FD-SOI technology on top of the single-device measurements presented in chapter 2, we studied low-frequency and high-frequency circuits composed of a few 10's to 100's of transistors from 300 to 4.2 K.

a) Circuits under test

We realized and tested basic transistor circuits to study both the analog behavior and the digital operation. The circuits are labeled with the format $\mathrm{NAME}_{L}^{W/L}$ with NAME the circuit name, L the transistor gate-length and W/L the gate-width to -length ratio of the employed N-type devices. The P-type devices are often sized with a two-times larger width to account for the typical difference in transistor currents between N- and P-type long-channel devices. The investigated circuits span a wide range of transistor types and geometries with minimal gate-length of 28 nm up to a relaxed gate-length of 130 nm with W/L values of 5 or 10 for N-type devices (see appendices B and C for dimension details).

In the next paragraphs, we quickly introduce the following investigated circuits: level-shifters, ring oscillators, analog pass-gates (also known as transmission gates or switches), amplifiers, and digital-to-analog converters.

Level shifters

Level shifters are often used when transferring data between chip in- and outputs operating at a high supply voltage $V_{DD} > 1\,\mathrm{V}$ and a low-power digital processor with a lower supply voltage $V_{DDL} < 1\,\mathrm{V}$. The designed level shifters, named LS_{28}^{10} and LS_{130}^{10}, transform a digital input with logic-0 defined by V_{SS} and logic-1 defined by V_{DD} to a digital output with logic-1 defined by $V_{DDL} < V_{DD}$. In figure 3.1a, we show the details of the electrical scheme of the investigated level shifters creating a digital output V_{out} at V_{DDL} level depending on the digital input V_{in} at V_{DD} level. The two upper transistors (P1, P2) are cross-coupled to maintain a complementary stable digital output at V_{DDL} while the bottom transistors (N1, N2) are activated or deactivated depending on the digital input at V_{DD} using an inverter scheme.

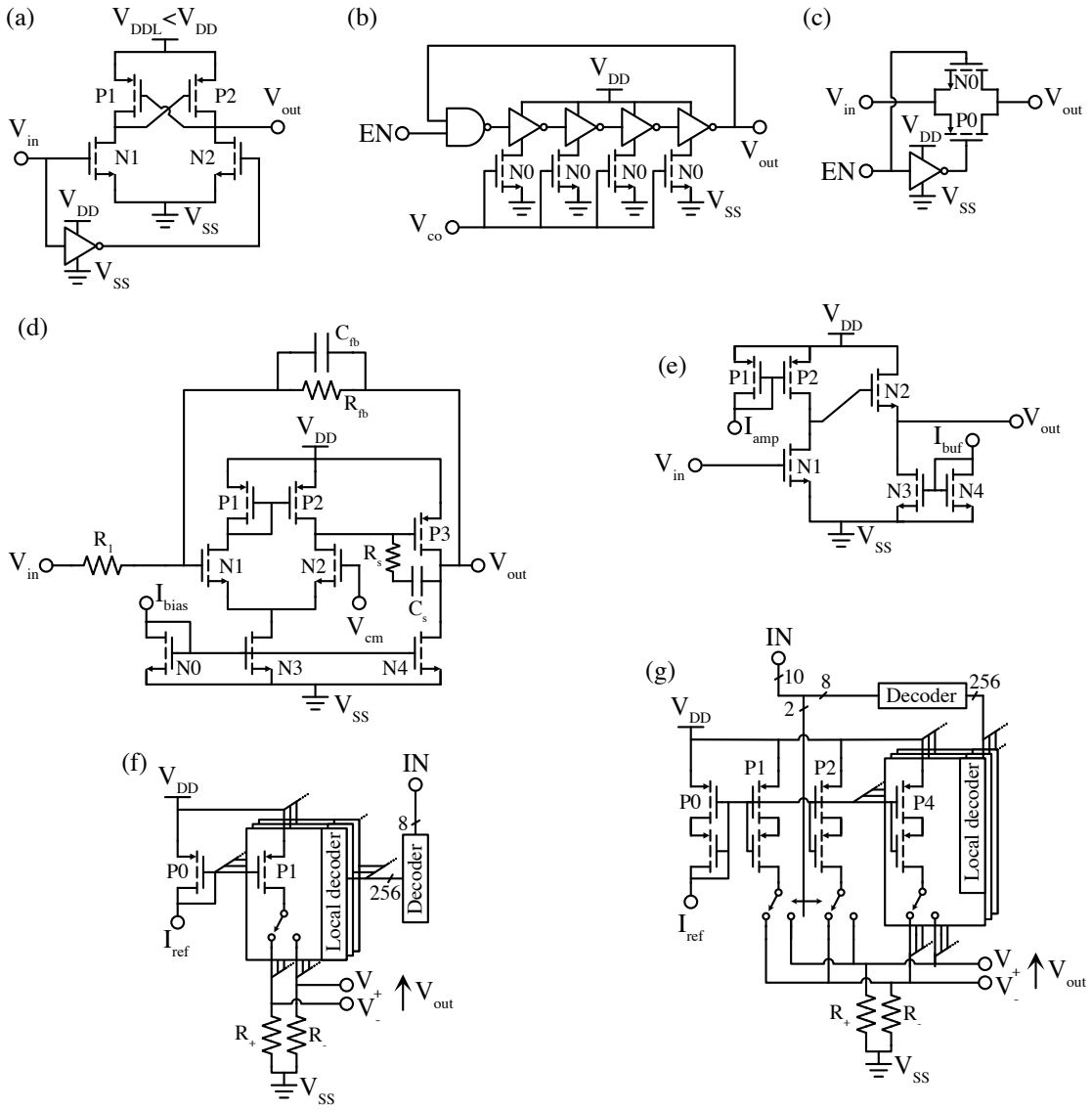


Figure 3.1 | Schematics of the circuits under test.

(a) Level-shifter schematics to reduce the power supply voltage V_{DD} of the input bit V_{in} to a lower voltage $V_{DDL} < V_{DD}$ of the output bit V_{out}. (b) Ring-oscillator implementation with 5 inverting stages. One inverting stages is a NAND gate to enable oscillation by setting EN to 1. Each inverting stage is current-limited with an additional N-type transistor N0. (c) Implementation of the pass-gate with a P- and N-type transistor in parallel. The pass-gate conducts for EN corresponding to a digital 1. (d) Voltage amplifier of gain $V_{out}/V_{in} = R_{fb}/R_1 = 10$ based on a two-stage miller-compensated operational amplifier with a resistive feedback loop. (e) Voltage amplifier made of an N-type common-source stage followed by common-drain stage. (f) Non-segmented 8-bit current steering DAC made of 256 identical P-type current sources. The current of each source is conveyed to either the + or − port depending on the input code IN. The total added current flowing through a resistor in each output generate the $\mathrm{DAC_{fine}}$output differential voltage $V_{out} = V_+ - V_-$. (g) Segmented 10-bit current steering DAC made of an 8-bit section similar to (f) and two binary-weighted P-type current sources for the two least significant bits. Cascoded current sources increase the output impedance to reduce non-linearities associated with the wide output voltage range of $\mathrm{DAC_{coarse}}$.

LS^{10}_{28}, respectively LS^{10}_{130}, is composed of 28 nm (resp. 130 nm) gate-length transistors and W/L of 5 (resp. 10) (see Table B.1).

We investigated the DC behavior of the two level-shifters at 300 and 4.2 K with and without applying a forward back-bias voltage FBB by setting the input bit at $V_{DD} = 1\,\mathrm{V}$ and measuring the output-bit voltage upon varying V_{DDL} from 0 to 1 V. For high V_{DDL} close to 1 V, the level-shifter outputs a value close to V_{DDL}. As V_{DDL} is decreased, the output follows V_{DDL} down to a certain value from which the level shifters stop following V_{DDL} (see Figure B.1). Interestingly, the value at which the level-shifters stop operating is below the V_{th} value. In fact, the implemented architecture allows transistors to operate even in subthreshold regime as long as the transistors source-drain voltage is sufficiently large to remain in saturation-like regime.

In this chapter, we study the operating range of level-shifters with respect to V_{DDL}.

Voltage-controlled ring-oscillators

Ring oscillators (RO) generate a square wave between the two power rail voltages V_{SS} and V_{DD}. An odd number of inverting gates looped together in a feedback as shown in Figure 3.1b becomes unstable and starts oscillating. We designed three ring oscillators, named RO^{5}_{28}, RO^{10}_{28}, and RO^{10}_{130}, made of 5 inverting stages of which one NAND gate for enabling oscillations. RO^{5}_{28} and RO^{10}_{28} have minimal gate-length L of 28 nm and different aspect ratio W/L of 5 and 10 for N-type devices (and 10 and 20 for P-type devices). RO^{10}_{130} has a longer gate of 130 nm and W/L of 10. The oscillation frequency is defined by the charging time of the inverting transistors and can be controlled with an additional transistor N0 cascoded with the inverters that modifies the discharging time depending on its gate-source voltage V_{co}, thus reducing the signal frequency by reducing the control voltage V_{co}.

From room-temperature simulations with foundry models, the ring oscillators generate square waves from 100 MHz to a few GHz. Ten dynamic flip-flops divide the output frequency of each RO by 1024 such that a digital buffer is able to send the MHz-frequency signal to room-temperature through meter-long coaxial cables, required during cryogenic measurements. We designed two frequency dividers: one with minimal gate-length of 28 nm and a second one with relaxed gate-length of 130 nm. RO^{10}_{28} and RO^{10}_{130} output frequencies are measured with the 28 nm frequency divider as they generate GHz signals while the slower RO^{5}_{130} uses the 130 nm frequency divider. More details on the dynamic flip-flops for the frequency divider can be found in appendix B. We measured the output frequency of the ring-oscillators with V_{co} varied from 0 to 1 V at 300 and 4.2 K with and without FBB. The obtained frequency characteristics $f(V_{co})$ are shown in Figure B.4. RO^{10}_{28} is the fastest oscillator thanks to its minimal gate-length, followed by RO^{10}_{130} then RO^{5}_{130} (see Table 3.1 for a summary of the maximally attained oscillator frequencies).

The maximal output frequency of the ROs increases by 7 to 12 % by cooling the circuit from 300 to 4.2 K with no FBB. Applying a FBB significantly increases the RO speed by a factor $2.9 - 3.7$ culminating to a whopping frequency of 9.47 GHz for RO^{10}_{28} at 4.2 K and FBB of 4 V.

In this chapter, we study the operating voltage range, the static leakage, and the dynamic power as a function of temperature and FBB.

Analog pass-gates

Pass-gates are switches for analog signals that depending on the activation voltage pass all signals or none. The design shown in Figure 3.1c has one P-type and one N-type transistor (N0, P0) in parallel ensuring a good conduction upon activation over the entire rail-to-rail voltage range. Once activated, the pass-gates behave as a input-dependent resistor and once deactivated, is left at high-impedance. Pass-gates were extensively used in the addressable matrix presented in chapter 2 to isolate the DC signals of the

Table 3.1 | Maximal output frequency of the investigated ring-oscillators at 300 and 4.2 K for increasing FBB.

The maximal output frequency of RO^{10}_{28}, RO^{5}_{130}, and RO^{10}_{130} is obtained from the $f(V_{co})$ curves shown in Figure B.4.

T		300 K	4.2 K		
FBB		0 V	0 V	2 V	4 V
f (GHz)	RO^{10}_{28}	3.24	3.47	5.79	9.47
	RO^{10}_{130}	1.43	1.53	2.78	
	RO^{5}_{130}	1.08	1.21	2.37	3.97

Table 3.2 | Equivalent resistance of pass-gates at 300 and 4.2 K with FBB.

The equivalent resistance of pass-gates is extracted from $R(V_{cm})$ curves shown in Figure C.1. R_N (respectively R_P) is the equivalent resistance extracted at $V_{cm} = 100\,\text{mV}$ (resp. $V_{cm} = V_{DD} - 100\,\text{mV}$) corresponding to the N-type (resp. P-type) conducting path.

T		300 K	4.2 K		
FBB		0 V	0 V	2 V	4 V
PG^{10}_{130}	R_N (Ω)	263	174	142	125
	R_P (Ω)	297	151	148	135
PG^{13}_{150}	R_N (Ω)	286	144	134	94
	R_P (Ω)	319	139	144	102

measured devices under test from all other devices. We designed pass-gates with thin (respectively thick) oxide named PG^{10}_{130} (resp. PG^{13}_{150}) with gate-length of 130 nm (resp. with minimal gate-length of 150 nm) (see Table C.1 for more details on transistor properties).

We measure the small-signal equivalent resistance R_{PG} of the pass-gates as a function of the DC common-mode voltage of the input signal $V_{in} = V_{cm}$ with a lock-in amplifier (see measurement details in appendix C, and the obtained $R_{PG}(V_{in})$ curves in Figure C.1). We extract the equivalent resistance in the two regimes: where the PMOS conducts for $V_{cm} = 0\,\text{V}$ and where the NMOS conducts for $V_{cm} = V_{DD}$ (see Table 3.2) with V_{DD} the power supply voltage equal to 1 V (respectively 1.8 V) for thin-oxide (resp. thick-oxide) devices. The pass-gate resistance decreases by a factor $1.4 - 2.0$ with cooling from 300 to 4.2 K. However, an anomalous effect in the pass-gate resistance arises for thin-oxide devices at 4.2 K without FBB: the equivalent resistance at $V_{cm} = 0.5\,\text{V}$ culminates at 138 kΩ, which is almost a factor 10^3 greater value than the pass-gate resistance at $V_{cm} = 0\,\text{V}$ (see Figure C.1a). Increasing FBB to 2 V reduces the resistance increase at mid-V_{DD} to 1.3 kΩ. At FBB of 4 V, the pass-gate retrieve an almost constant resistance, similar to its room-temperature behavior. This effect is also observed in the thick-oxide device for no FBB but only with a factor-3 increase (see Figure C.1b). This anomaly for the thin-oxide devices comes from the fact that V_{th} increases to about 0.6 V at cryogenic temperatures, i.e. above half V_{DD}, such that for $V_{cm} = V_{DD}/2$ neither the PMOS nor the NMOS are conducting, leading to a high equivalent resistance. Decreasing the V_{th} below $V_{DD}/2$ with FBB or using thick-oxide devices that benefits from a higher V_{DD} solves the problem by ensuring $V_{th} \ll V_{DD}/2$. The multiplexing matrix presented in chapter 2 is made of thick-oxide pass-gates, hence not affected by this effect.

In this chapter, we study the operating range of pass-gates.

Amplifiers

We implemented two amplifiers of gain $\sim 10\,\mathrm{V/V}$, named AMP_{28} and AMP_{130}. AMP_{130} is a two-stage Operational Amplifier with resistive feedback network made of an on-chip polysilicon resistor $R_{fb} = 11.6\,\mathrm{M\Omega}$ and an off-chip Metal Electrode Leadless Face (MELF) resistor $R_1 = 1\,\mathrm{M\Omega}$ (see Figure 3.1d). The ratio of resistance R_{fb}/R_1 fixes the voltage gain $|V_{out}/V_{in}|$ to ~ 10. The operational amplifier follows a two-stage Miller-compensated architecture. The first stage is a N-type differential stage with P-type active load and the second stage is a P-type common-source stage. All transistors of the operational amplifier have relaxed gate-length above 130 nm to increase its open-loop gain (see chapter 2, section a). The operational amplifier with feedback resistor R_{fb} is later used to form a transimpedance amplifier for the current read-out of quantum dots, described in more details in chapter 4.

AMP_{28} is composed of an input amplifying stage (N-type common-source) and an output buffer stage (N-type common-drain) as shown in Figure 3.1e. The two stages are composed of transistors with minimal gate length of 28 nm to reach GHz-bandwidths. AMP_{28} is used to extract a small-signal voltage from the capacitance read-out circuit presented in more details in chapter 5.

In this chapter, we study the noise and energy efficiency of amplifiers. The amplifier measurements are further detailed in chapters 4 and 5.

Digital-to-analog converters

Digital-to-analog converters (DAC) generate a voltage depending on a digital input code w. DACs are often used to generate arbitrarily-shaped time-signals from high-speed digital data stream imposing the output voltage values. Marcos Zurita (postdoc, CEA-LETI) implemented two current-steering DACs: a fine-grained 8-bit DAC ($\mathrm{DAC}_{\mathrm{fine}}$) and a coarse-grained 10-bit DAC ($\mathrm{DAC}_{\mathrm{coarse}}$) that could be assembled together to create a >16-bit DAC.

The $\mathrm{DAC}_{\mathrm{fine}}$[1] follows a typical 8-bit non-segmented thermometric architecture operating under 1.8 V power supply voltage (see Figure 3.1f). A reference current I_{ref} is mirrored to 256 identical current sources made of a single 160nm-long P-type LVT thick-oxide (GO2) device, providing 12 µA each when biased. The individual currents are steered to one of the two polysilicon 1 kΩ resistors to generate the differential output voltages V_+ and V_-, depending on the input 8-bit code w. The steering switch structure includes charge-balancing dummies to prevent voltage glitches at high speed. The fine DAC covers the ±1.54 mV differential output range for a total static power of 7.3 µW.

The $\mathrm{DAC}_{\mathrm{coarse}}$adopts a segmented 10-bit architecture under 1.8 V supply voltage, with the 8 Most-Significant Bits (respectively 2 Least-Significant Bits) following a thermometric encoding (resp. binary encoding) (see Figure 3.1g). A diode-mounted P-type device (P0) generates the gate-source voltage fed to the DAC output current sources (P1, P2, and P4), corresponding to the reference current I_{ref}. The thermometric part associated with the 8 MSB is made of 256 cascoded 1.1 µA current sources made of long thick-oxide P-type transistors (0.625 and 5 µm) to reduce the non-linearity caused by the wide output range and non-zero transistor conductances. Two binary-weighted current sources made of two cascoded P-type thick-oxide (GO2) devices respectively generate 0.275 and 0.55 µA output currents. The current of every individual source is steered to one of the two 1 kΩ polysilicon resistors to generate the differential output voltages V_+ and V_-. The coarse DAC spans the wide 1 V differential output voltage range for a total power of $\sim 2\,\mathrm{mW}$.

b) Extracted circuit quantities

From systematic measurements of the investigated circuits at 300 and 4.2 K, we are able to compare the circuit behavior with and without the use of forward body-biasing (FBB). Instead of analyzing each circuit

Table 3.3 | Investigated circuits from 300 to 4.2 K and extracted circuit-level quantities.

We investigated multiple version with different transistor dimensions of each circuit accounting in total for: 2 level-shifters, 3 ring-oscillators, 2 pass-gates, 2 amplifiers, and 2 DACs.

Circuits	Performance			Energy		System
	Operating Range	Inversion Point	Noise	Leakage	Efficiency	Variability
Level Shifters	✗					
Ring Oscillators	✗	✗			✗	
Analog pass-gates	✗	✗		✗		
Amplifiers	✗		✗		✗	
DACs						✗

separately, we studied universal quantities of interest divided in three categories: performance-related, energy-related, and system-related quantities presented in Table 3.3.

We first present the performance of the circuits by analyzing the voltage range in which the circuit operates as intended, the typical voltage level at which cryogenic circuits surpasses room-temperature circuits, and the output noise. Then, the energetic capabilities of circuits are treated by looking at the leakage current, and at the energy efficiency defined as the ratio of power and bandwidth. Finally, we focus on system-level properties that emerge from the source-drain current variability considering thousands devices.

II Circuit performance

In this section, we investigate quantities related to the circuit operation and performance with varying temperature and FBB. We first look at the operating voltage range in which circuits operate. In this operating voltage range, we extract a typical voltage separating the voltage range in two: one part in which the circuits show improved performance at cryogenic temperatures and the other part in which the circuits operate better at room temperature. We finally look at the circuit output noise.

a) Operating Voltage Range

Circuits generally operate with voltages comprised between the two power-supply rail voltages V_{SS} and V_{DD} to avoid damaging devices with, e.g. oxide breakdown. In the FD-SOI 28 nm technology, thin-oxide (respectively thick-oxide) devices support a maximum rail-to-rail voltage $\Delta V_{rr} = V_{DD} - V_{SS}$ of 1 V (resp. 1.8 V). In the following, we consider V_{SS} tied to the ground at 0 V.

Circuits employ transistors in specific regimes, e.g. in saturation or in triode regime with imposed conditions between source-drain and gate-source voltages. These conditions can't be fulfilled for all input voltages and the circuit stops working as intended outside a given operating voltage range (OVR) that can be significantly reduced compared to the full voltage range $\Delta V_{rr} = V_{DD} - V_{SS}$. The OVR can be limited by a large voltage excursion at any node of the circuit that modifies the transistor operating regime. As an example, the OVR of an amplifier is likely to be limited by its output stage as the highest voltage change is expected at the output upon varying the input voltage, leading to transistors entering the triode regime with greatly reduced the gain. Inversely, a voltage-controlled oscillator might be limited by the input voltage that controls the frequency preventing oscillations if too low or too high.

Although no general definition of the operating voltage range is applicable to all circuits, we are looking

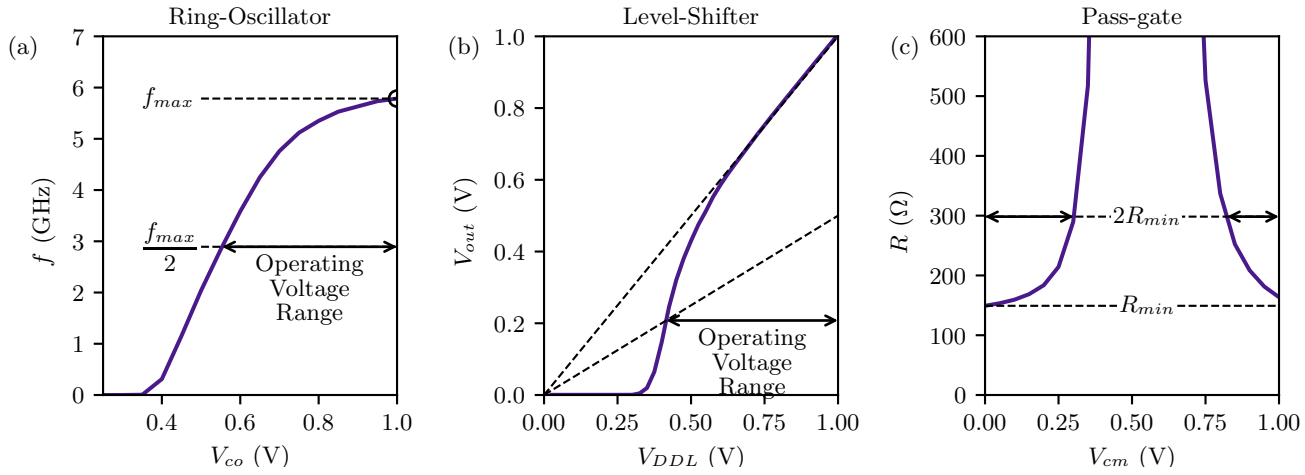


Figure 3.2 | Extraction of the operating voltage range for each circuit.

Criteria for the extraction of the operating voltage range for (a) ring oscillator, (b) level-shifter, and (c) pass-gate. All presented data are extracted at 4.2 K for a FBB of 2 V.

for a general trend of the operating range from room to cryogenic temperatures. To this goal, we give a slightly different definition of the OVR for each circuit, all answering the same question: "What is the input-voltage window in which the circuit operates as intended ?". It might be tempting to resume the OVR to a change in threshold voltage, as if voltages fall below V_{th} we expect transistors to be turned off resulting in dysfunctionning. However, circuits might operate as intended even with transistors in the subthreshold regime leveraging only the diffusion of electrons linked to the thermal voltage V_t given by k_BT/e (see chapter 2).

Level-shifters

The OVR of the level shifters is extracted from the DC characteritics $V_{out}(V_{DDL})$ (shown in Figure B.1) as the voltage range where the output voltage stays above half the requested V_{DDL} value, avoiding the poor operation of the level shifter (see Figure 3.2b). At room-temperature, the level-shifters operate in a rather large operating voltage range of 0.49 and 0.65 V for respectively LS^{10}_{130} and LS^{10}_{28} (see Table 3.4), roughly corresponding to the difference between V_{DD} and V_{th} of P1 and P2 (see Figure 3.1a). The higher value obtained for LS^{10}_{28} comes from short-channel effects that reduces V_{th} for 28 nm gate-length devices compared to 130-nm gate-length devices (see chapter 2, section IV). Cooling down the circuits with FBB at 0 V leads to a reduced OVR by 0.22 and 0.26 V mostly linked to the V_{th} increase at cryogenic temperatures (see Figure 2.8). The obtained OVR of only 0.23 V obtained for LS^{10}_{130} greatly limits the use of the level-shifter as V_{DDL} can only be reduced to 0.77 V, which is generally insufficient for a low-power core (with typical supply of 0.6-0.7 V). Turning on FBB of 2 V significantly increases the OVR to usuable values as V_{th} is reduced. The OVR only retrieves its room-temperature value for all investigated circuits under a high FBB of 4 V.

Ring-oscillators

The OVR of ring oscillators is extracted as the frequency control voltage V_{co} above which the observed oscillation frequency f is greater than the half maximum frequency $f_{max}/2$ of the ring oscillator (reached at the maximal V_{co} of 1 V, see Figure 3.2a and the $f(V_{co})$ characteristics in Figure B.4). For the three

Table 3.4 | Extracted operating voltage range for the circuits at 300 and 4.2 K for different FBB.
Operating voltage range in Volt according to the criteria of extraction shown in Figure 3.2. The numbers are colored in green if the OVR at 4.2 K is greater than at 300 K and in red otherwise. Typically, FBB of 2 V is sufficient to recover room-temperature OVR.

T		300 K	4.2 K		
FBB		0 V	0 V	2 V	4 V
OVR (V)	LS^{10}_{28}	0.65	0.43	0.55	0.69
	LS^{10}_{130}	0.49	0.23	0.39	0.56
	RO^{10}_{28}	0.33	0.26	0.36	0.48
	RO^{5}_{130}	0.34	0.25	0.35	0.50
	RO^{10}_{130}	0.33	0.27	0.38	
	PG^{10}_{130}	0.91	0.43	0.69	1.00†
	PG^{13}_{150}	1.80†	1.38	1.80†	1.80†

† Maximal OVR reaching the power supply ΔV_{rr}

oscillators with different gate-length and W/L ratio, the OVR is similar for all the three oscillators at respective temperature and FBB conditions (see Table 3.4). The OVR decreases from 0.33-0.34 V to 0.25-0.27 V from 300 to 4.2 K. The room-temperature values of OVR are retrieved with only a 2 V FBB. Higher FBB significantly increases the OVR by 46 % compared to its room-temperature value, offering a wider range of frequencies with FBB at cryogenic temperatures.

Pass-gates

For the level-shifters and ring-oscillators, the operating voltage range (OVR) was easily derived from the characteristics as these circuits operate at high voltages (V_{DDL} and V_{co}) and stop working below a given voltage. However, pass-gates operate well at low common-mode V_{cm} (PMOS conduction) and high V_{cm} (NMOS conduction) but stop working well around mid-V_{DD} (especially for PG^{10}_{130} made of thin-oxide devices with low V_{DD}). We compute the OVR by adding the voltage ranges in which the PMOS conducts and in which the NMOS conducts with an equivalent resistance below $2 \times R_{min}$ with R_{min} the minimum pass-gate resistance (achieved at V_{cm} equal to 0 or V_{DD}, see Figure C.1). At room-temperature, the OVR corresponds to a large portion of the available voltage range ΔV_{rr} ($\Delta V_{rr} \equiv V_{DD} - V_{SS} = 1\,\text{V}$ for thin-oxide LS^{10}_{130} and 1.8 V for thick-oxide LS^{13}_{150}). The pass-gates have OVR above 91 % of ΔV_{rr} at 300 K that largely decreases to only 43 % (respectively 77 %) for LS^{10}_{130} (resp. LS^{13}_{150}) at 4.2 K with no FBB. A 2 V FBB is sufficient to reach a maximal OVR equal to ΔV_{rr} for LS^{13}_{150} made of thick-oxide devices. However, a 4 V FBB is required by the LS^{10}_{130} made of thin-oxide devices to counteract the lower ΔV_{rr} and reach a maximal OVR.

General trend

We gathered in Figure 3.3 the extracted relative variation of the operating voltage range from 300 to 4.2 K for all investigated circuits at various FBB. It appears clearly that cooling down circuits with no FBB decreases by $20 - 52$ % the voltage range in which circuits operate properly. The reduced OVR limits the use of the circuit with e.g. level-shifters that can only shift the input to a minimum V_{DDL} of 0.6-0.9 V. With a higher supply voltage of 1.8 V, circuits made of thick-oxide devices such as PG^{13}_{150} are

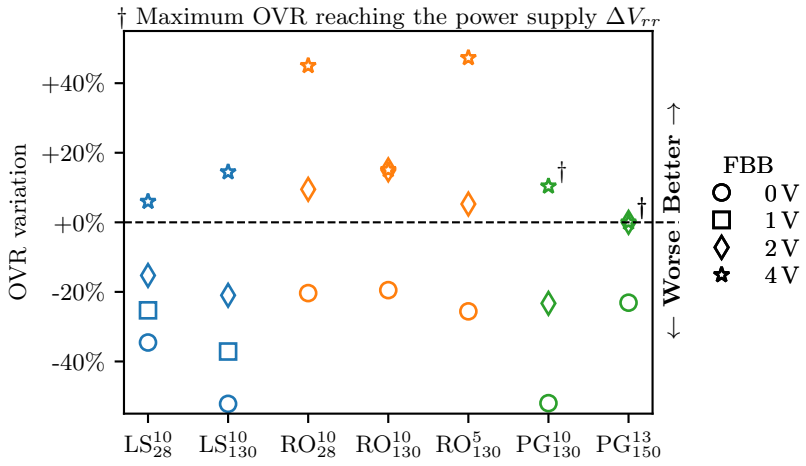


Figure 3.3 | Relative operating voltage range evolution from 300 to 4.2 K with FBB from 0 to 4 V.
All data presented in this graph are gathered from Table 3.4. Overall, a FBB of 2 and 4 V allow to recover standard room-temperature values.

less sensitive to the increased V_{th} at cryogenic temperatures, thus only exhibits a -20 % reduction of the OVR. Ring oscillators and pass-gates surpasses room-temperature values of the OVR for a FBB of 2 V, while level shifters require a high FBB of 4 V.

In this section, we investigated the voltage range in which circuits operate properly and in which they can be tuned. We observed that using FBB we are able to retrieve room-temperature values giving us the same flexibility at cryogenic temperature than at room-temperature. It becomes natural to ask what happens in this operating voltage region: are circuits operating better at room or at cryogenic temperatures ? We investigate the circuit performance evolution from 300 to 4.2 K inside the OVR in the following section.

b) Performance inversion-voltage

As the temperature decreases, the freeze-out of lattice vibrations (phonons) reduces the scattering of majority carriers improving the mobility while the reduction of thermally excited charge carriers increases the threshold voltage V_{th}. The concurrence of both phenomena generally leads to a particular gate-bias point in the MOSFET drain current I_{ds} versus the gate-source voltage V_{gs} called the Zero Temperature Coefficient point at which the drain current is almost independent of temperature T[2] (see Figure 2.5). V_{th} has a roughly linear T dependence from 300 K down to about ~ 50 K after what it saturates (see chapter 2). This increase of V_{th} with decreasing temperatures leads to a T^2 (for square-law devices $I_{ds} \propto (V_{gs} - V_{th})^2$) or T (for linear-law devices $I_{ds} \propto (V_{gs} - V_{th})$) dependence of the source-drain current I_{ds} in inversion. The decrease of I_{ds} with decreasing temperature T can be counteracted by the mobility temperature dependence T^{-1} or T^{-2}. As the mobility dependence on temperature depends on the scattering type (phonon, surface roughness, or coulomb interaction) that changes with increasing V_{gs}, we often observe a ZTC point in MOSFET characteristics. Despite the saturation of V_{th} at the lowest temperatures, the point of ZTC remains, suggesting that the mobility saturates as well at about the same temperature which is coherent with the influence of band-edge broadening at low temperatures[3].

In analogous way, we define a gate voltage for circuits made of multiple transistors where the circuit output characteristics at 300 K equals that at 4.2 K. To avoid confusion, we define this point as the

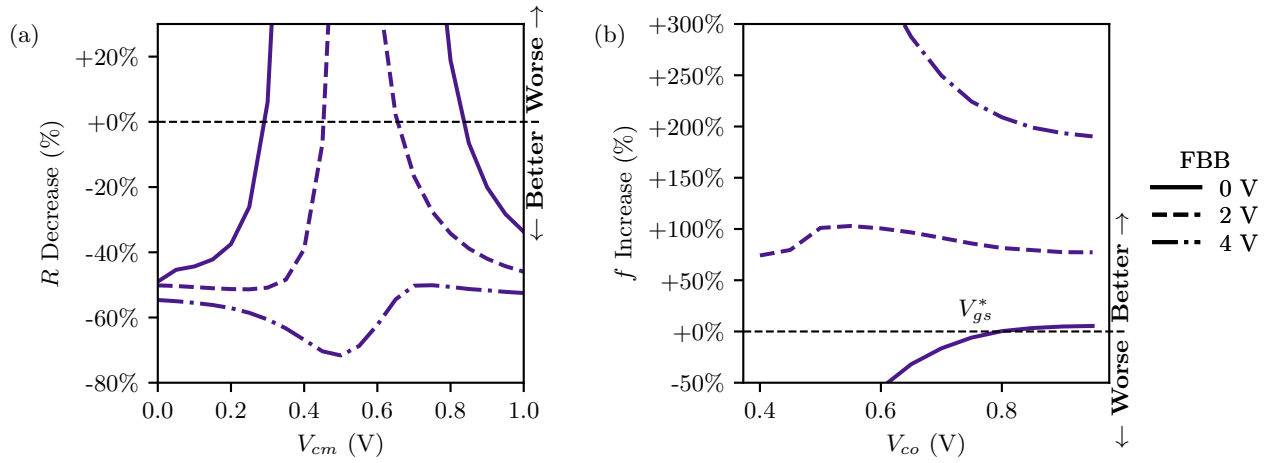


Figure 3.4 | Performance evolution from 300 to 4.2 K for FBB from 0 to 4 V.

(a) Pass-gate resistance R evolution from 300 to 4.2 K for FBB from 0 to 4 V. (b) Ring-oscillator output frequency f evolution from 300 to 4.2 K for FBB from 0 to 4 V. The performance inversion point V_{gs}^* is extracted at each FBB when the measured characteristic at 300 K equals the one at 4.2 K, i.e. for a 0 % change.

performance inversion point V_{gs}^* as we only studied the two extreme temperature points instead of the entire temperature range. As the studied circuits are made of a few transistors, we define V_{gs}^* as the gate-source voltage at which the circuit performance is identical at 300 and 4.2 K. The V_{gs}^* voltage splits the operating voltage range in two: one side $V_{gs} > V_{gs}^*$ for which the circuit behaves better at cryogenic temperature because the mobility increases enough to counteract the V_{th} increase, and the second one $V_{gs} < V_{gs}^*$ for which the circuit behaves better at room-temperature. By convention, we define V_{gs}^* equal to V_{DD} if the circuit behaves worse at cryogenic temperatures for all applied gate voltages and, inversely, $V_{gs}^* = 0\,\text{V}$ if the circuit performance are higher at cryogenic temperatures for all gate voltages. We performed this analysis in terms of the inversion point V_{gs}^* for the ring oscillators and for the pass gates.

Ring-oscillators

For ring oscillators, the output frequency f curves as a function of the gate voltage V_{co} at 300 and 4.2 K cross at a particular V_{co} of about of $0.77 - 0.80\,\text{V}$ at zero FBB, which we define as the inversion point V_{gs}^* (see Figure B.4). V_{co} corresponds to the gate-source voltage of the transistor N0 controlling the discharging time (see Figure 3.1b). For $V_{co} > V_{gs}^*$, the output frequency at cryogenic temperatures is higher than at room temperatures with a maximum increase of +12% for RO_{130}^{10}. In this regime, the decrease in channel resistance at cryogenic temperatures totally counteracts the V_{th} increase. Lowering V_{th} with FBB of 2 V leads to an improvement in most of the operating voltage range with V_{gs}^* of 0.35 for RO_{28}^{10} and 0.45 for RO_{130}^{5} and RO_{130}^{10}. The output frequency for all ring oscillators at 4.2 K under strong FBB of 4 V surpasses the frequency at 300 K for all V_{co}, leading to $V_{gs}^* = 0\,\text{V}$. The output frequency of ring-oscillators at 4.2 K with a FBB of 2 V (respectively 4 V) increases by a factor of $\times 1.8 - 2.2$ (resp. $\times 2.9 - 3.6$) with respect to the room-temperature specifications.

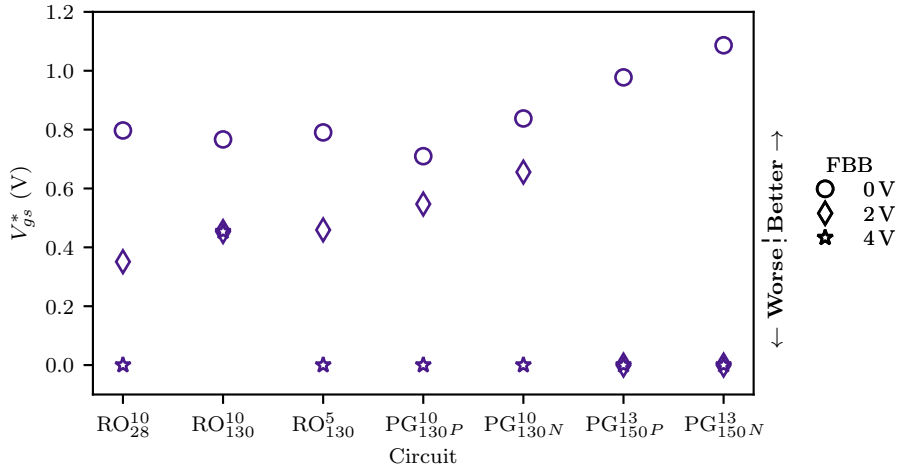


Figure 3.5 | Performance inversion point at 4.2 K for FBB from 0 to 4 V.

Performance inversion point for every studied circuit. A FBB of 4 V ensures that $V_{gs}^* = 0\,\mathrm{V}$, i.e. the circuit shows improvement at 4.2 K compared to 300 K over the entire voltage range.

Pass-gates

Similarly to the OVR definition for pass-gates, we separate the voltage space V_{cm} into two intervals: one interval (respectively the second interval) for V_{cm} close to 0 V (resp. V_{DD}) corresponding to PMOS (resp. NMOS) conduction. In Figure 3.4a, we show the equivalent resistance decrease from 300 to 4.2 K (positive percentage corresponds to a decrease in resistance, thus an improvement of the circuit). For FBB of 0 and 2 V, there is always a region mid-ΔV_{rr} in which the pass-gate resistance is higher at low temperatures due to V_{th} crossing as detailed in the previous section. There are two crossings of the 0 % level corresponding to the definition of two V_{gs}^*: the left crossing corresponding to $V_{DD} - V_{gs,P}^*$ of the PMOS and the right for the corresponding $V_{gs,N}^*$ of the NMOS. The thick-oxide devices of PG_{150}^{13} show a higher performance inversion point than thin-oxide devices of PG_{130}^{10}. However, the higher V_{gs}^* of thick-oxide devices have less impact on the circuit performance thanks to the higher power supply voltage of thick-oxide devices. PMOS devices exhibit a higher performance inversion point than NMOS devices by $110 - 130\,\mathrm{mV}$. A FBB of 2 V is sufficient to obtain better performance of PG_{150}^{13} at 4.2 K while a 4 V FBB is required for PG_{130}^{10}.

General trend and design implications

We observe a typical V_{gs}^* of 0.8 V (respectively 1 V) for all circuits made of thin-oxide (resp. thick-oxide) devices as shown in Figure 3.5. The almost constant value of V_{gs}^* extracted from all tested circuits is an interesting quantity for circuit designers. When designing circuits with no accurate models available at cryogenic temperatures, circuit designers often prefer to design circuits with accurate room-temperature models to encompass all transistor non-idealities (e.g. finite conductance and parasitic capacitance) and to extrapolate the circuit behavior at lower temperatures. With the definition of V_{gs}^* it becomes easier to predict that a simple circuit will keep about the same properties at low temperatures if V_{in} is kept close to V_{gs}^*, be better if $V_{in} > V_{gs}^*$, or worse if $V_{in} < V_{gs}^*$. In this last case, greater margins are required to meet the circuit requirements. FBB drastically boosts the circuit performance at 4.2 K as seen in Figure 3.5. By ensuring that gate-voltages are above V_{gs}^*, the IC designer is certain that the circuit will behave the

same or better at low temperatures, thereby mitigating the need of back-gating in design.

Circuits at cryogenic temperature with FBB surpass the performance of the same circuit at room-temperature with respect to, e.g., higher speed and lower resistance. Another key quantity related to circuit performance is the output noise of circuits that depends on individual transistor contribution within the complete circuit architecture.

c) Noise

Sensitive electronics, as required for quantum computing applications, require a careful design of circuits with respect to noise from transistors (see chapter 4 and 5 for a system-level analysis of noise). Broadband thermal noise decreases at cryogenic temperature with a theoretical factor of 10 from 300 to 4.2 K for resistors and transistors. However, as we approach nanometric gate-lengths, the 1/f-flicker noise becomes preponderant in a large range of low frequencies with a typical corner frequency up to a few tens of MHz at room temperature. The main cause of flicker noise comes from traps at the oxide interface between the channel and the front-gate high-k dielectric used in recent technologies. Despite having a thermally activated behavior at room-temperature, the flicker noise usually increases up to a factor of 2 and rarely decreases at cryogenic temperatures.

We measure the output noise of the amplifiers AMP_{28} and AMP_{300}, and the gyrator with a lock-in amplifier. The details on operation and noise values are presented in chapter 4 for AMP_{300} and in chapter 5 for AMP_{28} and GYR_{28}in the context of complete circuits with a dedicated function. A typical noise-spectrum graph with frequency is shown in Figure 3.6a for AMP_{28} at 300 and 4.2 K. For all circuits, we observe a typical noise spectrum with a flicker component at low frequency, and a broadband thermal noise at higher frequency if the circuit bandwidth is high enough. The corner frequency defined as the intersection between the $1/f$ and thermal noise component is around 250 kHz for AMP_{28}.

We extract the noise characteristics from the obtained output noise spectrum of the three circuits with least-square fits to the $1/f^n$ frequency dependence. The obtained flicker noise at 1 kHz, the flicker exponent n, and the thermal noise level are shown in Table 3.5 at 300 and 4.2 K with and without FBB. Without FBB, the flicker noise amplitude increases by 43 % at low temperatures while the flicker exponent keeps the same value of $0.47-0.48$. Surprisingly, the flicker noise extracted at 1 kHz decreases by $53-58\%$ when applying a 2 V FBB. With the decrease of flicker noise with FBB, the flicker exponent also changes from $0.46-0.47$ at 300 K to $0.56-0.57$ at 4.2 K suggesting a slight change of the noise origin. The origin of the decrease of flicker noise with FBB is not yet known and needs to be investigated further. One possible explanation could be that the traps in the top-gate interface causing the flicker noise get less coupled to the conduction channels as the inversion layer approaches the much cleaner bottom interface with FBB (see the charge carrier wavefunction with FBB in Figure 2.15). This phenomenon has been observed on both thin-oxide and thick-oxide devices. Further investigations are on their way to better understand this noise reduction.

The decrease of flicker noise at low temperatures with FBB is an excellent discovery for circuit design. However, the flicker noise remains at a high level compared to other bipolar technologies and might significantly affect the performance of circuits. Drift effects related to flicker-noise could induce significant systematic errors in e.g. qubit operations. In the long run, feedback circuits and modulation techniques (e.g. chopping and double correlated sampling) must be considered with CMOS nanotechnologies.

In this section, we investigated the performance of circuits when cooled down from room to cryogenic temperatures. We showed that the operation at cryogenic temperatures leads to a smaller voltage range for circuit operation, mainly due to the increased threshold voltage at cryogenic temperatures. The increased threshold voltage competes with the increased mobility at lower temperatures, sometimes strongly altering

Table 3.5 | Output noise characteristics of all circuits at 300 and 4.2 K with and without FBB.
Extracted values of the linear fit for the flicker noise and thermal noise. The flicker noise is presented as $A \times (f/1\,\mathrm{kHz})^{-n}$ with A the noise amplitude at 1 kHz, and n the flicker exponent. The flicker noise amplitude increases at 4.2 K without FBB but is reduced by applying FBB. Thermal noise decreases by a factor of 20.

	Flicker Noise ($\mu V/\sqrt{Hz}$)			Thermal Noise ($\mu V/\sqrt{Hz}$)	
T	300 K	4.2 K		300 K	4.2 K
FBB	0 V	0 V	2 V	0 V	2 V
AMP_{300}	$0.21 \times (f/1\,\mathrm{kHz})^{-0.47}$	$0.30 \times (f/1\,\mathrm{kHz})^{-0.48}$	-	-	-
AMP_{28}	$4.5 \times (f/1\,\mathrm{kHz})^{-0.46}$	-	$2.1 \times (f/1\,\mathrm{kHz})^{-0.57}$	0.21	0.095
GYR_{28}	$1.2 \times (f/1\,\mathrm{kHz})^{-0.45}$	-	$0.50 \times (f/1\,\mathrm{kHz})^{-0.56}$	-	-

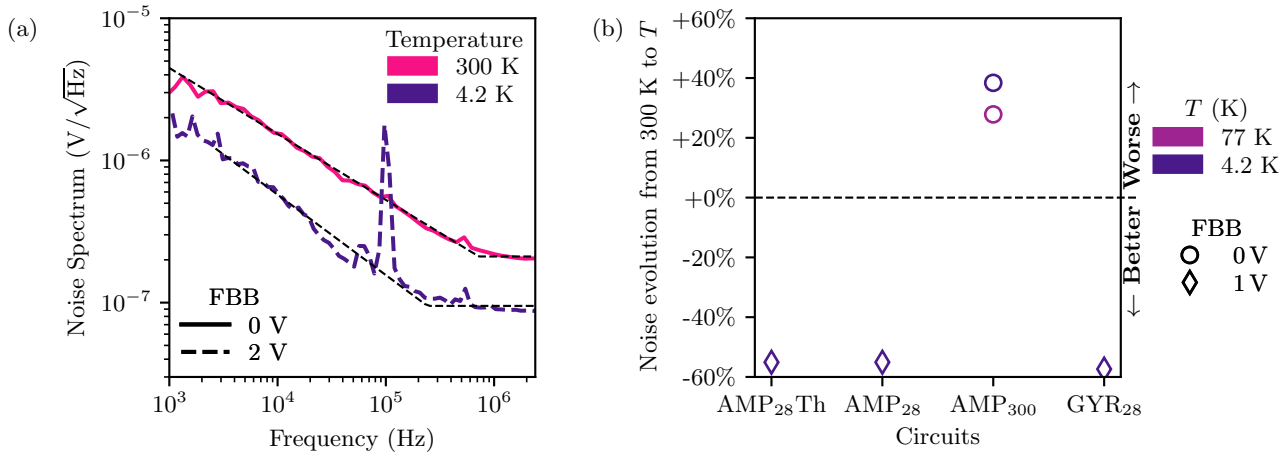


Figure 3.6 | Evolution of noise from 300 to 4.2 K.
(a) Example of noise spectrum of the voltage amplifier AMP_{28} at 300 and 4.2 K. For all circuits, we extract the thermal noise floor at high frequency and the $1/f^n$ flicker noise with the flicker exponent n, as shown by the black dashed lines. (b) Evolution of noise from 300 to 4.2 K extracted from the voltage amplifier, the OP-AMP, and the active inductance.

the circuit performance. Forward body biasing of FD-SOI circuits removes the negative effects from increased V_{th} and leads to circuits operating with better specifications at cryogenic temperatures, fully leveraging the gain in mobility with standard-level V_{th}. FBB also significantly reduces the flicker-noise level at cryogenic temperatures by more than ~ 50 %. Circuits operating at cryogenic temperatures under FBB behave better than at room-temperature. However, faster circuit operation leads to a higher power consumption which has to be controlled at cryogenic temperature to remain within the limits of the fridge cooling power. In the next section, we investigate the power efficiency of the CMOS FD-SOI technology at cryogenic temperatures.

III Energy-aware quantities

Power consumption of circuits operating at cryogenic temperatures has to be reduced to stay within the fridge cooling capabilities. We investigate the power efficiency of the CMOS FD-SOI technology at 4.2 K compared to 300 K. The comparison is interesting as it allows us to get a predictive sense of the power cost we can expect from circuits at low temperatures based on the existing room temperature realizations.

We analyzed the transistor leakage extracted from multiple devices connected to the same node and the energy cost per unit of bandwidth for digital and analog circuits.

a) Reduced Leakage

Device leakage has been increasingly important with node down-scaling as thinner oxides increase current leakage by thermal or even quantum effects. Current leakage increases the power consumption in large digital systems accounting for about 60 % of the total power consumption[4]. Leakage also reduces the retention time of voltage on high-impedance nodes and could prevent DRAM-like architecture to bias sequentially thousands of qubits[5, 6]. In a scalable architecture of a 2D matrix of qubits, high retention times are the key to be able to bias thousands qubit gates allowing the use of only one voltage source refreshing periodically every DC voltage bias. Low temperature greatly reduces leakage by suppressing thermal effects. However, quantum-based tunneling leakage from gate to source for thin-oxide devices might not be significantly affected due to a weak dependence on temperature. In this section, we extract the typical evolution of leakage current from 300 to 4.2 K.

Typical leakage currents in single devices are of the order of a few fA, which is difficult to measure. We extract the typical leakage in the FD-SOI technology thanks to the shared measurement lines with 1023 devices in the multiplexing matrix presented in chapter 2.

Top-gate leakage

We measure the top-gate leakage I_{gg} of all thin-oxide devices sharing their gate connection by measuring the total current as a function of the gate-source voltage V_{gs} of all DUTs inside the multiplexing matrix from chapter 2 at 300 and 4.2 K (see Figure 3.7a). All DUTs account for a total gate area of $5152\,\mu m^2$. The maximum measured current corresponding to the leakage in strong inversion $V_{gs} \gg V_{th}$ decreases by a factor 3.9 from 300 to 4.2 K for a gate-source leakage of only $260\,\mu A/mm^2$ at 4.2 K. For gate voltages V_{gs} above 0.5 V (typical value of V_{th} at low temperatures), the leakage reduction by a factor 3.9 remains the same. In subthreshold $V_{gs} < V_{th}$, the leakage rapidly falls below the background leakage associated with meter-long cabling corresponding to a $\sim 50\,pA$ leakage current. However, an estimated $>10^2$-fold decrease of the leakage is estimated at $V_{gs} \simeq 0.45\,V$, likely to be maintained all throughout the subthreshold region. This rapid decrease below 0.5 V at 4.2 K suggests a high reduction of thermally activated leakage at low temperatures.

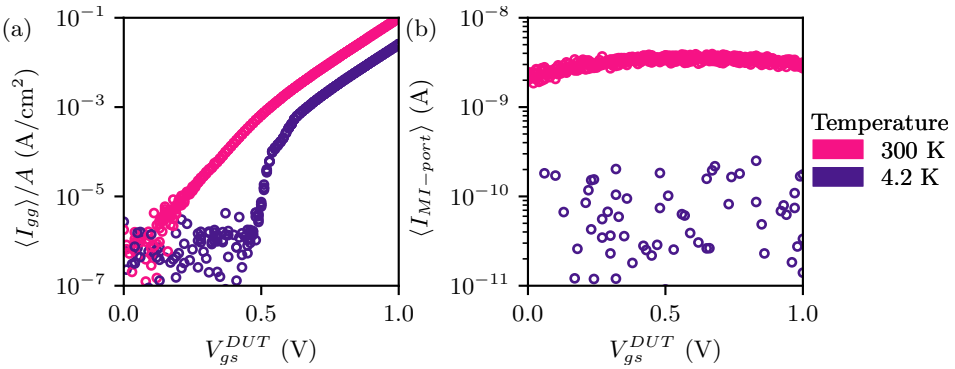


Figure 3.7 | Gate and source-drain current leakage per transistor channel area at 300 and 4.2 K.
(a) Gate-source leakage per device area for thin-oxide N-type devices as a function of gate voltage extracted from a shared connection on all gates of devices accounting for a total area of 5152 µm^2. (b) Total current measured at the MI port of the multiplexing matrix as function of the selected DUT gate-source voltage V_{gs}^{DUT} (see Figure 2.3a). This leakage current corresponds to the combined leakage of the 1023 off-state switches with contributions from both the source and the gate leakage, independently of the selected DUT gate-voltage.

Drain leakage

As an additional data point, we investigate the total leakage $I_{MI-port}$ measured at the MI port of the multiplexing matrix (see Figure 2.3a) The current at the MI-port is a combination of gate-source leakage and drain-source leakage from the 1023 off-state switches, and is consequently independent of the DUT gate-source voltage V_{gs}^{DUT}. The measured leakage shown in Figure 3.7b decreases by more than 2 orders of magnitude from 300 to 4.2 K going from 4 nA to <30 pA at 4.2 K limited by the measurement setup background leakage due to meter-long cabling.

General trend

The overall trend of leakage-current reduction upon reducing the temperature from 300 to 4.2 K is shown in Figure 3.8. Leakage always decreases at cryogenic temperature by a factor of 3.9 to 100 depending on the relative importance of leaking mechanism (tunneling vs thermal). MOSFETs become less leaky at cryogenic temperature allowing to increase the retention times up to a factor 100. The leakage reduction allows to imagine a voltage-biasing architecture of thousands of quantum devices by exploiting the high cryogrenic retention time as proposed and shown in [6]. The reduction at cryogenic temperature of static power dissipation from leakage will lead to a higher efficiency of cryogenic digital processors. For the RISC-V processor in 40 nm bulk, a static power consumption reduction by a factor 10 was observed when cooling the processor from 300 to 4.2 K[7].

In the mentioned RISC-V processor, the static power reduction does not affect the total power consumption. Digital processors are limited by the power consumption associated with static leakage but also by the dynamical power consumption due to the charging and discharging of gate-capacitors. In the next section, we look at the energy efficiency of circuits from 300 to 4.2 K.

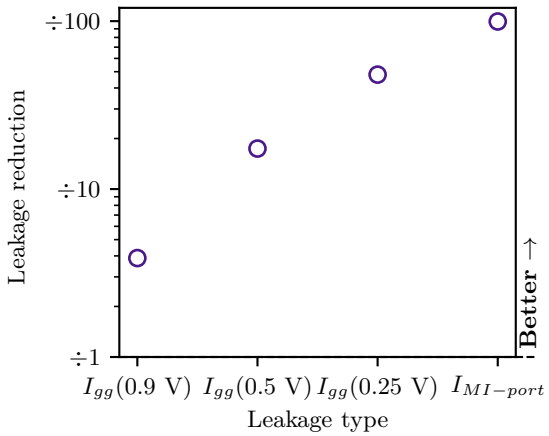


Figure 3.8 | Relative leakage evolution from 300 to 4.2 K.

Evolution of leakage currents extracted from Figure 3.7 for the gate leakage I_{gg} of 512 thin-oxide transistors in parallel at several gate-source voltages corresponding to subthreshold, moderate inversion, and strong inversion, and the average drain leaking measured at the MI-port of the matrix. All leakage currents reduce at cryogenic temperatures due to less thermally-activated conduction, up to more than two orders of magnitude in weak inversion.

b) Higher energy efficiency

Experimental apparatus to reach cryogenic temperatures can only sustain limited dissipation called the heat load. Dilution fridges typically can sustain 300 µW at 100 mK and a few W at a few Kelvin while a dip-stick in liquid helium sustains a few tens of W at the cost of more helium consumption. To implement bigger systems, it is required to limit the dissipation of electronics placed at the cryogenic stages. In this section we look at the energy efficiency of electronics at cryogenic temperatures.

Digital dynamic power

We first investigate the power consumption P_{DD} of the three ring oscillators by measuring the ring oscillator power consumption as a function of the output frequency f. The $P_{DD}(f)$ characteristics of all the ring oscillators from 300 to 4.2 K exhibit a linear relationship. The power increasing linearly with the frequency, named dynamic power P_{dyn}, comes from the incessant charging and discharging of the transistor gate capacitances at a rate equal to the oscillation frequency. P_{dyn} is generally expressed as $nC_gV_{DD}^2f$ with nC_g the total capacitance charged and discharged at every period with n the RO stage number ($n = 5$), and C_g the RO stage capacitance. The stage capacitance C_g depends on the transistor geometry with width W and length L and can be simplified to a component proportional to the device area WL corresponding to the parallel plate capacitor made of the inversion channel and the gate, and a component proportional to the device width W corresponding to stray capacitance.

We extract the dynamic power P_{dyn} by fitting $P_{DD}(f)$ with a linear function of f which is plotted in Figure 3.9a as a function of the stage area. P_{dyn} remains roughly constant with cooling from 300 to 4.2 K, explained by the weak dependence of gate capacitance on temperature.

With hardly any dependence in temperature of P_{dyn}, the energy efficiency remains the same from 300 to 4.2 K at constant power supply voltage V_{DD}. The dissipated power by the RO can be reduced

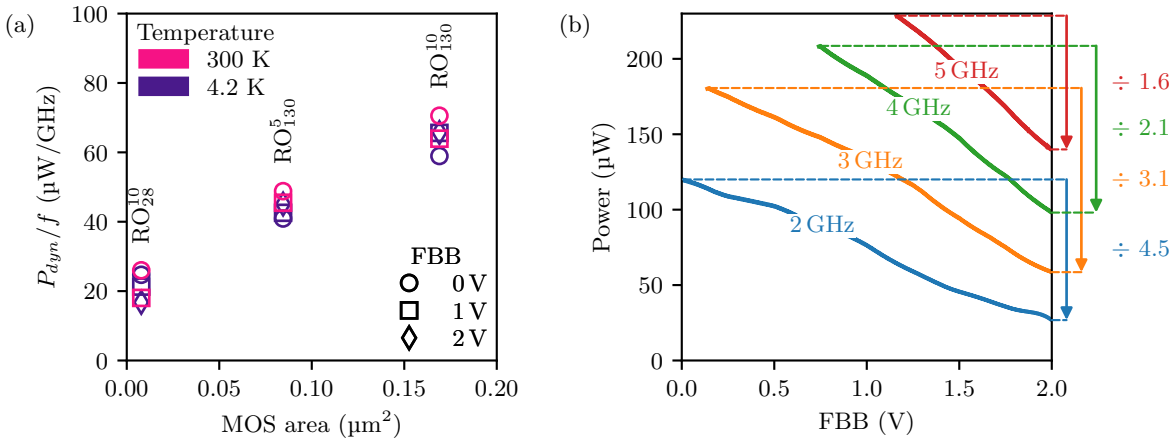


Figure 3.9 | Energy efficiency of ring oscillators at 300 and 4.2 K.

(a) Dynamic power per frequency $P_{dyn}/f = nCV_{DD}^2$ of each ring oscillator at 300 and 4.2 K for FBB from 0 to 4 V. P_{dyn} is extracted by a linear fit of the power of the ring oscillator as a function of the output frequency f. The dynamic power is not significantly dependent on the temperature nor the FBB, and only depends on the device geometry. The constant dynamic power implies that the gate-capacitance of each transistor does not significantly vary with temperatures from 300 to 4.2 K. (b) Ring oscillator power reduction of RO^{10}_{28} by simultaneously lowering V_{DD} while increasing FBB to maintain a constant frequency. Arrows indicate the power reduction at each frequency.

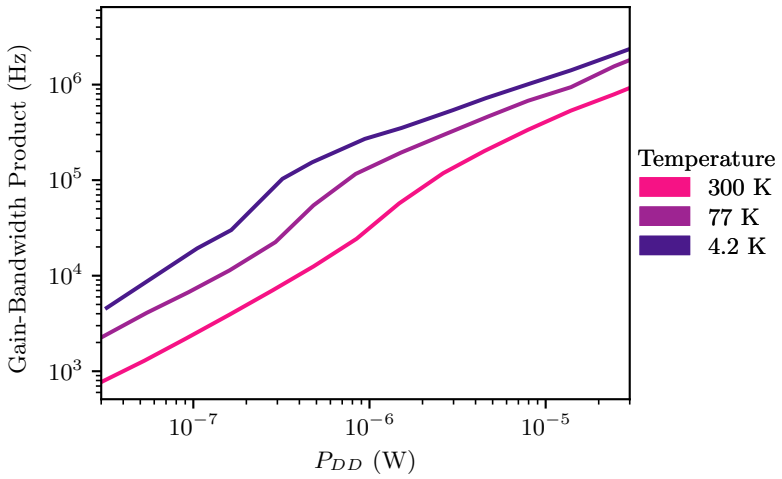


Figure 3.10 | Gain-bandwidth product of AMP_{300} versus power consumption from 300 to 4.2 K.
The gain-bandwidth product corresponds to the amplifier bandwidth if its gain equals unity. The *GBW* is indicative of the energy efficiency of the OP-AMP to provide a given voltage gain. The bandwidth for a given power consumption P_{DD} increases with decreasing temperature leading to a better energy efficiency at cryogenic temperatures.

by lowering the supply voltage while maintaining the gate overdrive by applying FBB to maintain the same oscillation speed. Applying FBB while reducing the power supply to maintain a constant frequency drastically reduces the total dissipated power by up to a factor 4.5 for back-gate voltages of 2 V at cryogenic temperatures (see Figure 3.9b).

Analog dynamic power

Power consumption of analog circuits is less straightforward than for digital circuitry with many interdependence between transistor geometry of transistors and circuit topology. Analog circuit power can largely vary depending on e.g. the required speed, noise level, and output impedance. We take the example of the Operational Amplifier (OP-AMP) and measure its gain-bandwidth product (GBW), indicative of how energy efficient the amplifier is in delivering a given voltage gain. We refer the reader to chapter 4 for an extensive description of the OP-AMP. We measure the gain-bandwidth product at 300, 77, and 4.2 K as a function of the OP-AMP power consumption P_{DD} as the bias current is varied from 0.01 to 10 µA (see Figure 3.10). The GBW expression is proportional to the device transconductance g_m. At a constant P_{DD}, the OP-AMP GBW increases with decreasing temperature as g_m at constant bias current (i.e. constant P_{DD}) increases at low temperature (see the data for g_m/I_{ds} in chapter 2). The gain in GBW is greater when transistors operate in subthreshold $P_{DD} < 1$ µW as g_m/I_{ds} drastically increases in this regime from 300 to 4.2 K. We define the energy efficiency of the OP-AMP as the power divided by the gain-bandwidth product. This quantity corresponds to the energy required by the OP-AMP mounted as a follower to follow a signal at a frequency of GBW. The energy efficiency increases by a factor of ~ 5 from 300 to 4.2 K is attributed to the improvement of g_m/I_{ds} which forms a key quantity for low-power analog electronics.

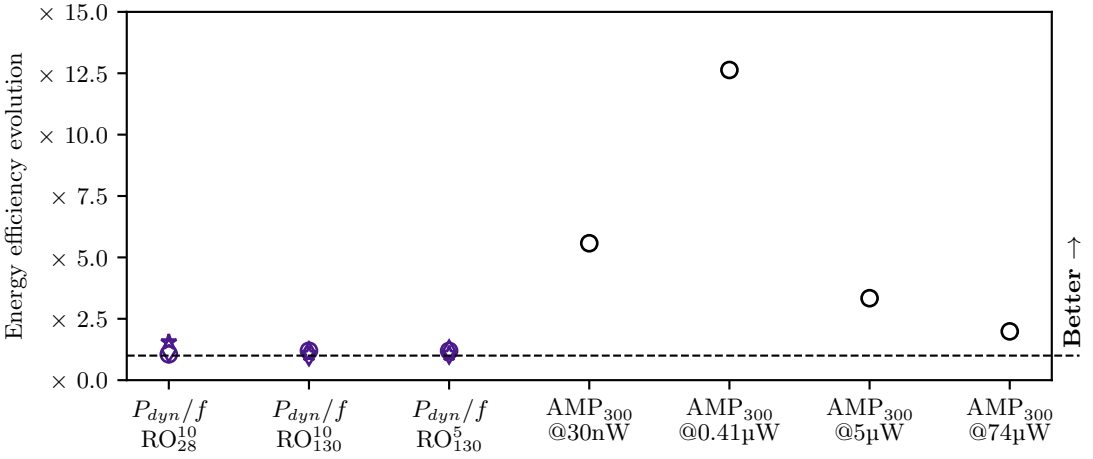


Figure 3.11 | Energy efficiency evolution from 300 to 4.2 K.

Evolution of the energy efficiency when cooling down the ring oscillators and the operational amplifier. The RO dynamic power stays unchanged as the MOSFET capacitance is only weakly dependent on temperature. The OP-AMP power efficiency defined as the gain-bandwidth divided by the power shows consequent improvements at cryogenic temperatures thanks to the increase of the g_m/I_{ds} ratio, more pronounced in weak and moderate inversion at low power.

General trend

We gathered all the presented data for the energy efficiency in Figure 3.11. The dynamic power does not significantly change with temperature or FBB as the transistor capacitance in inversion remains a constant given by geometry. The only lever arm to reduce the circuit power consumption concerns the more analog-based power. Indeed, the dynamic power can be reduced by increasing FBB while reducing the supply voltage, effectively reducing the number of charge transferred per cycle while keeping the same oscillation frequency thanks to FBB. On the pure analog side for AMP$_{300}$, the energy efficiency is improved by a factor of up to 10 at low temperatures thanks to the large increase of g_m/I_{ds} from 300 to 4.2 K (see chapter 2).

In this section, we found that the power required by circuits at cryogenic temperature tends to reach lower values than at room-temperature. Leakage reduction combined with FBB and reduced V_{th} can lead to drastic reduction of the power consumption of cryogenic digital processors[7, 8]. Analog circuits also benefit from higher g_m/I_{ds} at cryogenic temperatures.

IV System-level considerations

In this last section, we focus on the variability effects arising when considering thousands of transistors.

a) Increased Source-Drain Current Variability

Despite identical laid-out dimensions, transistor characteristics will vary from device to device due to fabrication imperfections or proximity effects. This variation, also named mismatch, generally varies by a few % but becomes of major importance in certain circuits that require almost identical operation of the

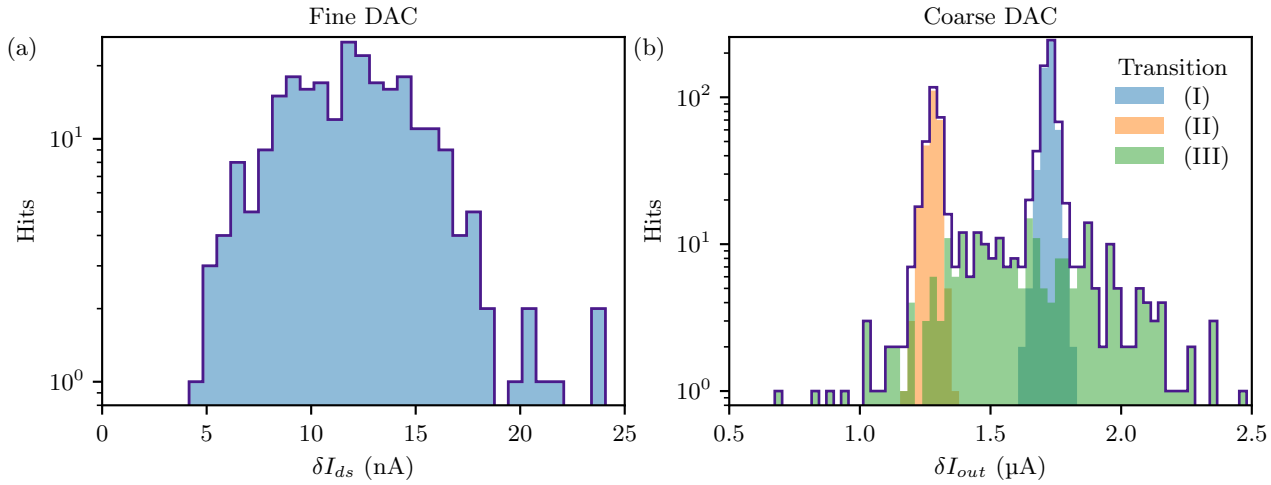


Figure 3.12 | Fine and coarse DACs mismatch at 4.2 K.

Histogram of $\delta I_{out}(w)$ at 4.2 K for the fine DAC (left) and coarse DAC (right) on a logarithmic y-axis to improve readability. The step-like solid line corresponds to the histogram data extracted from all words w. For the coarse DAC, the filled areas outline the three histograms generated from subsets of δI_{out} corresponding to the three identified transitions.

transistors. Differential amplifiers and analog-to-digital or digital-to-analog converters are among the most well-known examples sensitive to component mismatch. Here, we investigate the mismatch behavior of the two DACs from 300 to 4.2 K made with thick-oxide devices.

Mismatch of the non-segmented thermometric fine DAC

Due to the non-segmented thermometric architecture of the DAC_{fine}, incrementing the input digital word w results in only one additional current source being switched on (see Figure 3.13a). Hence, the voltage step δV_{out} when incrementing the input word directly relates to the transistor current mismatch $\delta I_{out} = \delta I_{ds} = \delta V_{out} R$ with $R = 1\,\text{k}\Omega$ corresponding to the output resistor value. The histogram of δI_{ds} exhibits a single gaussian, expected for this architecture, with a mean value of 11.8 nA and standard deviation of 3.4 nA at 4.2 K (see Figure 3.12a). The current mismatch of 29% at 4.2 K is significantly higher than the expected value of 3% at 300 K obtained with Monte-Carlo simulations. The unusually high mismatch is attributed to the combination of cryogenic temperatures, layout effects, and transistors operating in weak to moderate inversion.

Mismatch of the segmented coarse DAC

In a similar fashion, we obtain the output current variation δI_{out} for $\text{DAC}_{\text{coarse}}$and plot the corresponding histogram in Figure 3.12b. The distribution of δI_{out} is made of three gaussian contributions due to the segmented architecture with two binary-weighted LSB. The different contributions correspond to three different switching scenarios resulting from the DAC architecture as shown in Figure 3.13b. For simplicity, we decompose the 10-bit input digital word w into a 8-bit word w_{th} corresponding to the thermometric part and a 2-bit word w_{bin} corresponding to the binary LSBs. The switch sequence follows a 4-periodic pattern, starting from an arbitrary w_{th}, w_{bin} goes from 00, to 01, to 10, to 11, and go back to 00 while incrementing w_{th} by 1 (respectively named (Ia), (II), (Ib), and (III)). Among those four transitions, 00

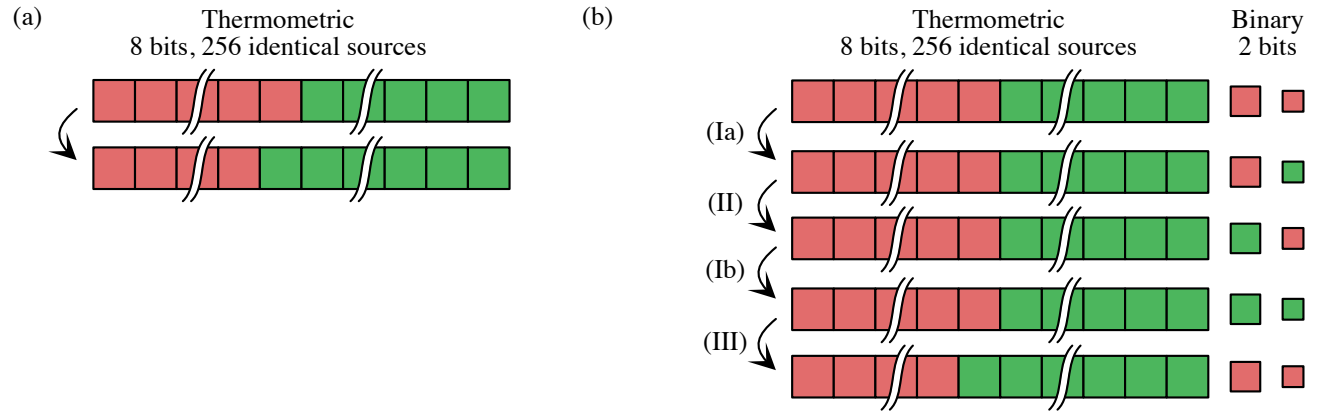


Figure 3.13 | Illustration of the different switching scenarios in both DACs

Squares represent individual current sources in the two DACs with their state indicated with green if switched on or red if switched off. In the thermometric 8-bit part of both DACs, 256 identical current sources are switched one by one as the input digital word is incremented by one, resulting in a single type of transition for DAC_{fine}, shown in (a). DAC_{coarse}has three more transition types due to the 2-bit binary part added combined with the 8-bit thermometric part as shown in (b). Transitions (Ia) and (Ib) both correspond to the same device being switched on, hence are indistinguishable and adds up under the same gaussian distribution.

to 01 (Ia) and 10 to 11 (Ib) are identical from a mismatch point of view as the same device is switched while all other devices remain unchanged (shown in Figure 3.13b). Hence, it becomes straightforward to identify the different contributions in Figure 3.12b as we expect twice as many hits for (I)-type transitions compared to (II), the remaining one being of the type (III). The transitions (I) and (II) exhibit low mismatch values of 1.4 and 2.1% at cryogenic temperatures, only increasing by $\times$1.4 and $\times$1.6 compared to their room-temperature values. However, type-(III) transitions have a larger mismatch of 19%, $\times$4.1 higher than at room temperatures.

General trend

The evolution of mismatch with cooling from 300 to 4.2 K is shown in Figure 3.14. Depending on the biasing regime and geometry of the transistor, the current mismatch at constant V_{gs} and resistive load R measured for thick-oxide devices in the DAC_{coarse}increases up to $\times$4. Overall, the extracted mismatch of both thick- and thin-oxide devices inevitably increase at low temperatures as the impurities are no longer smoothed out by the higher temperatures in both bulk[9, 10] and FD-SOI[11, 12] technologies. Compared to bulk technologies, FD-SOI is showing the lowest mismatch[12], an asset for reliable and precise large-scale systems.

V Conclusion

We investigated in this chapter the behavior of circuits at cryogenic temperature compared to room-temperature. The overall conclusions are presented in Table 3.6 and are further discussed in the next

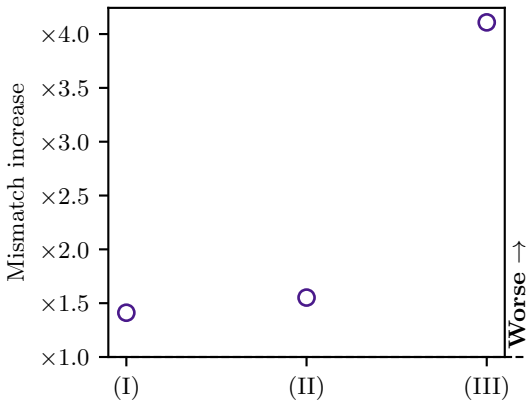


Figure 3.14 | Evolution of mismatch from 300 to 4.2 K.

Mismatch data gathered from the three transition types within the coarse DAC. Overall, the mismatch increases at cryogenic temperatures.

Table 3.6 | Evolution of circuit-related quantities from 300 to 4.2 K with and without FBB.

Forward-body biasing of devices in circuits results in an overall improvement of the studied quantities except for mismatch. Green-colored (respectively red-colored) arrows represent improvement (resp. deterioration) compared to 300 K.

	Performance			Energy		System
	Operating Range	Inversion Point	Noise	Leakage	Efficiency	Variability
No FBB	⇊	⇈	⇈	⇊	⇈	⇊
With FBB	⇈	⇊	⇊	⇊	⇈	⇊

paragraphs.

The circuit performance deteriorates from 300 to 4.2 K without FBB. The higher threshold voltage V_{th}, close to the supply voltage V_{DD}, at cryogenic temperature reduces the voltage range in which transistors remain in saturation. Insufficient gate-overdrive $V_{gt} = V_{gs} - V_{th}$ from high V_{th} prevents to benefit from mobility improvement at cryogenic temperatures, resulting in lower speed circuits. The low-frequency noise increases at cryogenic temperature while the thermal noise limit, at higher frequencies, significantly decreases as expected. With FBB, reduced V_{th} back to their room-temperature values leads to an increased operating voltage range and sufficient gate-overdrive to benefit from the higher mobility at 4.2 K with less phonon interactions. The FBB pulls the inversion layer farther away from the less clean top-interface and leads to a significant reduction of the low-frequency noise by 60 %.

Circuits also tend to consume less energy at cryogenic temperatures. Digital circuits highly benefits from a large reduction of leakage that only leaves the dynamic power in the digital processor's power bill. The dynamic power can be further reduced at cryogenic temperature by operating at a lower supply voltage V_{DD} thanks to the steeper subthreshold slope. Analog circuits greatly benefits from the higher charge-carrier mobility at 4.2 K that leads to higher g_m/I_{ds} ratio. At constant dissipated power, the analog circuits see their bandwidth extended with the increased g_m at fixed I_{ds}, or inversely, at constant bandwidth, the analog circuits can operate at a lower power budget.

We investigated the mismatch evolution at cryogenic temperature in real-world applications such as current-steering DACs. The mismatch inevitably increases at cryogenic temperatures but remains at a reasonable level for circuit design. To compensate for the increase of mismatch for precise applications, it becomes more important to consider larger devices or more complex circuit architectures, less sensitive to device mismatch.

Across the board, FBB drastically improves circuit performance at 4.2 K by compensating for the increased threshold voltage. The availability of the back-gate makes the SOI technologies attractive for high-speed cryogenic electronics.

In the last chapters, we explore in-depth more complex circuits co-integrated with quantum devices for specific measurements at low temperatures. We, first, designed a cryogenic transimpedance amplifier for current spectroscopy of quantum-dot structures with respect to applied quantum-dot gate and source voltages.

[FR] Comportement des circuits aux basses températures

Le partage des tensions et des courants entre les transistors d'un même circuit électronique permet d'implémenter différentes fonctions en fonction de la topologie, telles que l'amplification, la dérivation ou le suivi en tension. La topologie du circuit induit une interdépendance entre les caractéristiques des transistors qui conduit à une analyse compliquée, grandement simplifiée par l'utilisation de modèles numériques précis du comportement des transistors. Pour mieux comprendre l'impact des températures cryogéniques sur les circuits réalisés avec la technologie FD-SOI au-delà des mesures de transistors individuels présentées au chapitre 2, nous avons étudié des circuits basses- et hautes-fréquences composés de quelques dizaines à quelques centaines de transistors de 300 à 4.2 K tels que des convertisseurs de niveaux, des oscillateurs en anneaux, des amplificateurs ou encore des convertisseurs numérique-analogique.

Nous avons comparé dans ce chapitre le comportement de plusieurs circuits aux températures cryogéniques par rapport aux températures ambiantes. Les performances des circuits se détériorent de 300 à 4.2 K sans utiliser la grille arrière. Les tensions de seuil V_{th} plus élevées, proches de la tension d'alimentation V_{DD}, réduit la plage de tension dans laquelle les transistors restent en saturation aux températures cryogéniques. Une tension de grille insuffisante $V_{gs} \sim V_{th}$ ne permet pas une amélioration des transistors à froid malgré l'augmentation de la mobilité des porteurs de charges, rendant les circuits moins rapides à basse température. Le bruit basse fréquence tel que le bruit de flicker augmente à plus basse température tandis que le bruit thermique à des fréquences plus élevées diminue de manière significative. La réduction des tensions de seuil à leurs valeurs usuelles à température ambiante grâce à la polarisation de la grille arrière conduit à une plus grande plage de fonctionnement en tension et a une vitesse accrue à 4.2 K. Appliquer une tension de grille élevée éloigne la couche d'inversion de la grille avant et conduit même à une réduction significative du bruit basse-fréquence de 60 %.

Les circuits ont également tendance à consommer moins d'énergie aux températures cryogéniques. Les circuits numériques bénéficient d'une forte réduction des courants de fuites qui ne laisse que la puissance dynamique dans la consommation totale des processeurs numériques. La puissance dynamique peut être fortement réduite aux températures cryogéniques en fonctionnant à une tension d'alimentation V_{DD} plus faible grâce à la plus forte pente sous le seuil et grâce aux tensions de seuil réduites avec l'utilisation de la grille arrière. Les circuits analogiques bénéficient de la mobilité plus élevée des porteurs de charge à 4.2 K qui conduit à un rapport de transconductance sur courant drain-source plus élevé. A puissance dissipée constante, les circuits analogiques voient leur bande passante étendue, ou inversement, à bande passante constante, les circuits analogiques peuvent fonctionner avec une consommation réduite.

Nous avons étudié l'évolution du dépareillement entre transistors à température cryogénique au sein de convertisseurs numérique-analogique. Les disparités entre transistors de même taille augmentent inévitablement aux températures cryogéniques mais reste à un niveau raisonnable pour la conception de circuits. Pour compenser l'augmentation du dépareillement pour des applications cryogéniques de précision, il est primordial de considérer des transistors plus larges ou des architectures de circuits plus complexes, moins sensibles au dépareillement.

Dans l'ensemble, l'utilisation de la grille arrière améliore considérablement les performances du circuit à 4.2 K en compensant l'augmentation de la tension de seuil à froid. La disponibilité de la grille arrière rend les technologies SOI attrayantes pour l'électronique cryogénique à grande vitesse.

Dans les derniers chapitres, nous explorons des circuits plus complexes cointégrés avec des dispositifs quantiques. Nous avons conçu un amplificateur cryogénique à transimpédance pour la spectroscopie en courant des structures de puits quantiques.

Bibliography

[1] Marcos E. P. V. Zurita, Loïck Le Guevel, Gérard Billiot, Adrien Morel, Xavier Jehl, Aloysius G. M. Jansen, and Gaël Pillonnet. **Cryogenic Current Steering DAC With Mitigated Variability**. *IEEE Solid-State Circuits Letters*, **3**:254–257, 2020.

[2] I. M. Filanovsky and A. Allam. **Mutual compensation of mobility and threshold voltage temperature effects with applications in CMOS circuits**. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, **48**(7):876–884, July 2001.

[3] H. Bohuslavskyi, A. G. M. Jansen, S. Barraud, V. Barral, M. Cassé, L. Le Guevel, X. Jehl, L. Hutin, B. Bertrand, G. Billiot, G. Pillonnet, F. Arnaud, P. Galy, S. De Franceschi, M. Vinet, and M. Sanquer. **Cryogenic Subthreshold Swing Saturation in FD-SOI MOSFETs Described With Band Broadening**. *IEEE Electron Device Letters*, **40**(5):784–787, May 2019. Available from: `https://ieeexplore.ieee.org/abstract/document/8660508`.

[4] Rohit Lorenzo and Saurabh Chaudhury. **Review of Circuit Level Leakage Minimization Techniques in CMOS VLSI Circuits**. *IETE Technical Review*, **34**(2):165–187, March 2017. Available from: `https://www.tandfonline.com/doi/full/10.1080/02564602.2016.1162116`.

[5] M. Veldhorst, H. G. J. Eenink, C. H. Yang, and A. S. Dzurak. **Silicon CMOS architecture for a spin-based quantum computer**. *Nature Communications*, **8**(1):1766, December 2017. Available from: `http://www.nature.com/articles/s41467-017-01905-6`.

[6] Simon Schaal, Alessandro Rossi, Virginia N. Ciriano-Tejel, Tsung-Yeh Yang, Sylvain Barraud, John J. L. Morton, and M. Fernando Gonzalez-Zalba. **A CMOS dynamic random access architecture for radio-frequency readout of quantum devices**. *Nature Electronics*, **2**(6):236–242, June 2019. Available from: `http://www.nature.com/articles/s41928-019-0259-5`.

[7] E. Schriek, F. Sebastiano, and E. Charbon. **A Cryo-CMOS Digital Cell Library for Quantum Computing Applications**. *IEEE Solid-State Circuits Letters*, pages 1–1, 2020.

[8] H. Bohuslavskyi, S. Barraud, V. Barral, M. Cassé, L. Le Guevel, L. Hutin, B. Bertrand, A. Crippa, X. Jehl, G. Pillonnet, A. G. M. Jansen, F. Arnaud, P. Galy, R. Maurand, S. De Franceschi, M. Sanquer, and M. Vinet. **Cryogenic Characterization of 28-nm FD-SOI Ring Oscillators With Energy Efficiency Optimization**. *IEEE Transactions on Electron Devices*, **65**(9):3682–3688, September 2018.

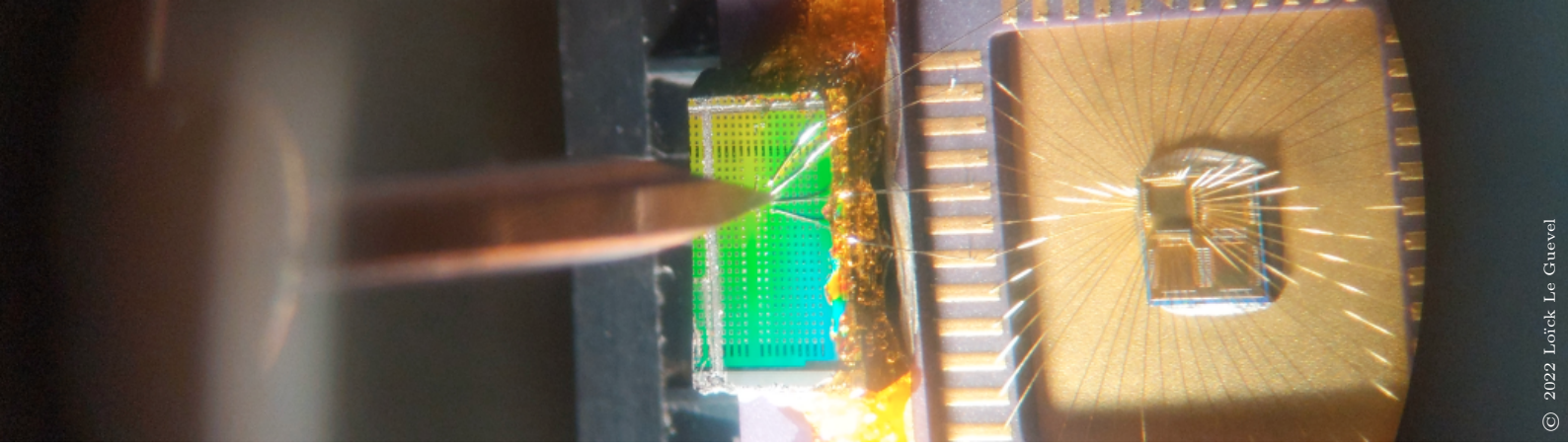
[9] P. A. 't Hart, M. Babaie, E. Charbon, A. Vladimirescu, and F. Sebastiano. **Subthreshold Mismatch in Nanometer CMOS at Cryogenic Temperatures**. *IEEE Journal of the Electron Devices Society*, pages 1–1, 2020.

[10] P. A. T Hart, M. Babaie, Edoardo Charbon, Andrei Vladimirescu, and Fabio Sebastiano. **Characterization and Modeling of Mismatch in Cryo-CMOS**. *IEEE Journal*

of the Electron Devices Society, **8**:263–273, 2020. Available from: `https://ieeexplore.ieee.org/document/9015956/`.

[11] Bruna Cardoso Paz, Loïck Le Guevel, Mikaël Cassé, Gérard Billiot, Gaël Pillonnet, Aloysius Jansen, Sebastien Haendler, Andre Juge, Emmanuel Vincent, Philippe Galy, Gérard Ghibaudo, Maud Vinet, Silvano de Franceschi, Tristan Meunier, and Fred Gaillard. **Integrated Variability Measurements of 28 nm FDSOI MOSFETs down to 4.2 K for Cryogenic CMOS Applications**. In *2020 IEEE 33rd International Conference on Microelectronic Test Structures (ICMTS)*, pages 1–5, May 2020.

[12] B. Cardoso Paz, L. Le Guevel, M. Cassé, G. Billiot, G. Pillonnet, A. G. M. Jansen, R. Maurand, S. Haendler, A. Juge, E. Vincent, P. Galy, G. Ghibaudo, M. Vinet, S. de Franceschi, T. Meunier, and F. Gaillard. **Variability Evaluation of 28nm FD-SOI Technology at Cryogenic Temperatures down to 100mK for Quantum Computing**. In *2020 IEEE Symposium on VLSI Technology*, pages 1–2, June 2020.



CHAPTER 4
Current read-out of quantum dots

I Introduction

Current sensing of high-impedance quantum-dot structures is done using a current-to-voltage amplifier, also known as a Transimpedance Amplifier (TIA). At low temperatures, electrical transport measurements of quantum-dot systems give access to their charge and spin properties[1]. The application of TIAs in these quantum systems can be divided in two main categories: high-speed detection of moving charges and more time-consuming accurate characterization of the quantized states.

High-speed detection of charge movements in a quantum-dot array allows fast spin-qubit read-out by e.g. probing a nearby capacitively-coupled Single-Electron Transistor (SET) as shown in Figure 4.1a. Biasing the SET voltages (e.g. the gate and drain) near a charge-state transition results in a current step of typical amplitude from 0.01 to 1 nA when a charge moves between nearby dots thanks to the capacitive coupling with the SET current-voltage characteristic (see Figure 4.1b). For quantum computing applications, we wish to detect this current step in a few $0.1 - 10\,\mu s$, comparable to the quantum gate duration, corresponding to bandwidths of $0.1 - 10\,\text{MHz}$ for high gains of $10^9\,\text{V/A}$. In typical setups with electronics at room-temperature, the amplification is split between a TIA of gain $Z_T = 10^7\,\text{V/A}$ followed by a low-noise amplifier (LNA) of gain $20 - 40\,\text{dB}$ to achieve the relatively high bandwidth given the high gain. As an example from [2, 3], the TIA[4] dominates the amplification noise with an input-referred noise of $43\,\text{fA}/\sqrt{\text{Hz}}$ accounting for 98 % of the total amplification noise ($1\,\text{fA}/\sqrt{\text{Hz}}$ contribution from the LNA[5]). Here, the TIA noise is near the theoretical limit of noise for resistive TIA at room-temperature due to the feedback resistor noise ($\sqrt{4k_BT/Z_T} = 41\,\text{fA}/\sqrt{\text{Hz}}$). The read-out fidelity depends on the signal-to-noise ratio (SNR) defined as the square of the ratio between the measured current-step amplitude and the TIA noise computed over the integration time[6]. With this level of noise, we can achieve SNR of 38.2 for an integration time of 3.21 µs and signal of 100 pA that results in fidelities above 99.9% (assuming perfect detection capability of the SET[6]). Usual specifications of TIAs at room-temperature are sufficient to perform single-shot read-out at high fidelities if the read-out current is ∼100 pA. However, if the measured current is reduced to 10 pA, the integration time becomes 321 µs to achieve the 99.9% fidelity. To compensate for the lower signal, the noise can be reduced by a factor of 10 to obtain a read-out time in the µs-range by placing the TIA at cryogenic temperatures.

In Figure 4.2, we schematically show the drawbacks and advantages of placing the TIA at low temperature close to the Device Under Test (DUT) by comparing a measuring configuration with the TIA at room temperature to that at cryogenic temperatures. The short-distance connection between TIA and

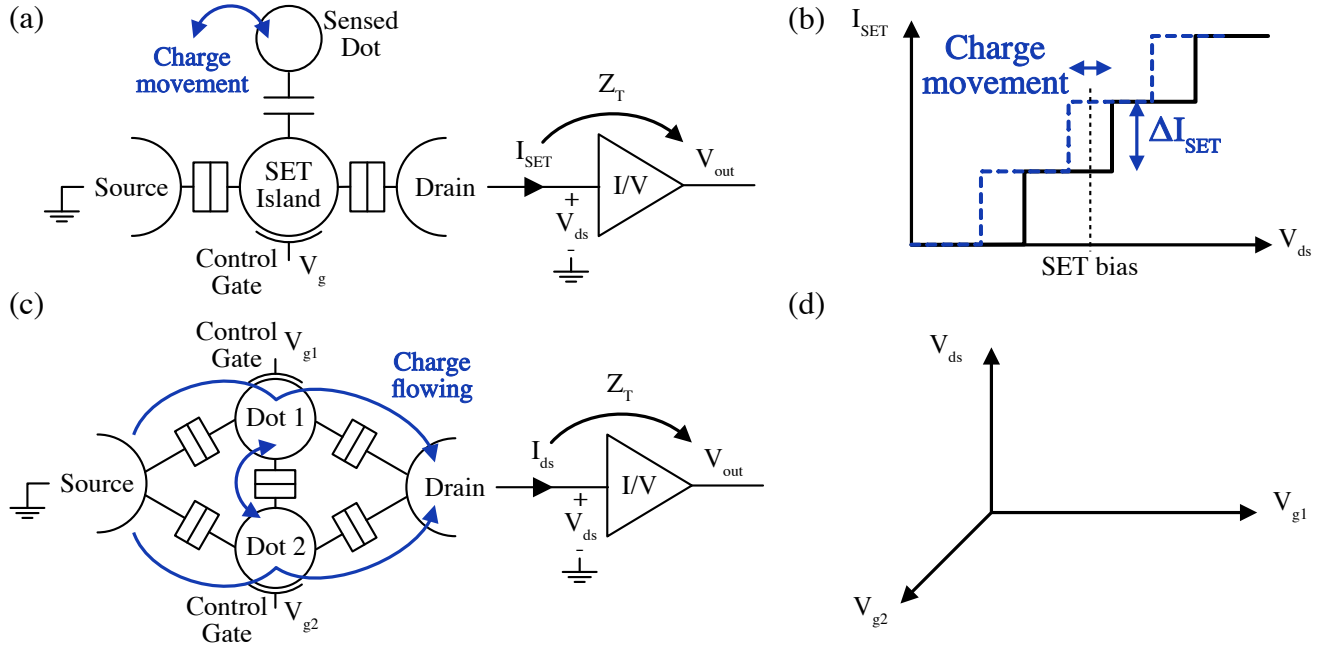


Figure 4.1 | Two applications of transimpedance amplifiers with quantum-dot systems.

(a) Single-electron transistor (SET) made of a single quantum dot with tunneling junctions with source and drain used as an electrometer. The SET current I_{SET} at a specific bias point (V_g, V_{ds}) is measured with a transimpedance amplifier of gain Z_T. A sensed dot with varying charge occupation induces a capacitive shift in the SET characteristics as shown in (b). The capacitive coupling induces a measured current step ΔI_{SET} if the number of charges in the sensed dot changes, leading to the read-out signal. (c) Fine measurement of the quantum transport properties of a more complex quantum-dot system consisting as an example of two coupled quantum dot in parallel. The current through the quantum structure I_{ds} is measured with a TIA with varying V_{ds}, V_{g1}, V_{g1} to probe the entire charge stability diagram. These measurement are time-consuming due to the high dimensionality of the parameter space shown in (d) for a double-dot structure to the low level of measured currents.

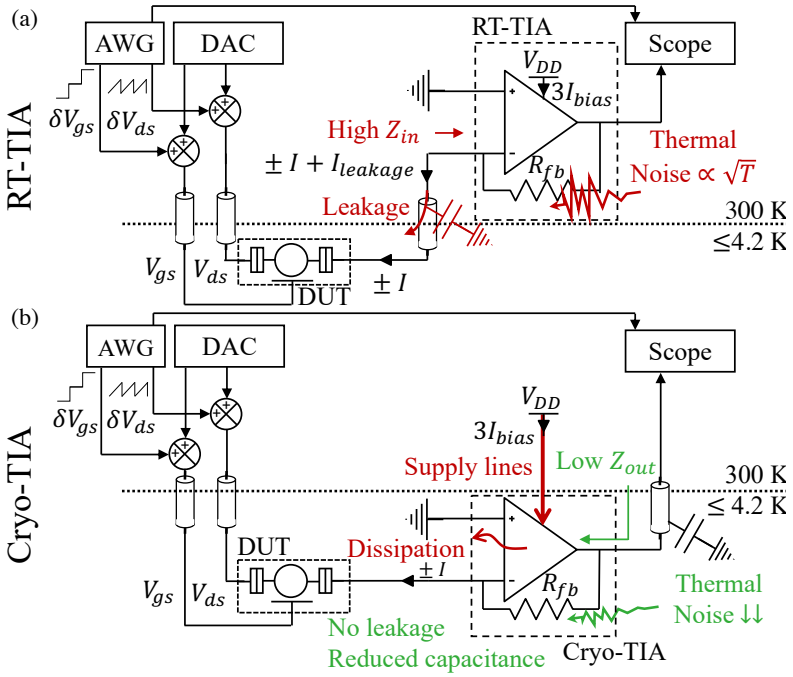


Figure 4.2 | Comparison of quantum-dot measurements between a room-temperature and a cryogenic transimpedance amplifier.
Schematics of the TIA amplifier for high-speed current measurement of a high-impedance quantum dot (Device Under Test, DUT). The comparison between 300 (a) and ≤4.2 K (b) highlights the issues of leakage current, thermal noise, dissipation, and trace impedance. It is clearly visible that the short-distance connection between TIA and DUT will reduce leakage currents because of reduced cable length and allow the detection of smaller current signals. The TIA at low temperatures has intrinsically a lower thermal noise.

DUT increases the maximum achievable bandwidth and reduces leakage currents because of reduced cable length thus combining faster detection of smaller features in the measured current. The dissipation of the TIA at low temperatures must be compatible with the cooling capacity of the low-temperature stage to keep the fridge temperature stable, about 1 W at 4.2 K and 300 µW at 0.1 K. Even with a few tens of µW, the TIA can generate hot-spots in the circuits (due to e.g. self-heating) leading to temperature gradients affecting the quantum device. Accidental heating of the quantum device by the TIA would result in a degradation of quantum properties. This subject highly depends on the rather unknown thermalization of the chips with the fridge base-temperature plate. The second type of heating, easier to quantify, is the effective increase of the charge carrier's temperature[7] due to excessive voltage noise. For a quantum device at a typical temperature of 300 mK, the associated thermal voltage $v_T = k_b T/q \simeq 26\,\mu\text{V}$ requires limiting injection of voltage noise from the TIA to the measured device in order to prevent the increase of the charge carrier's temperature. The TIA at low temperatures benefits from the intrinsically lower thermal noise compared to room-temperature with a theoretical reduction by a factor of more than 9 for temperatures below 4.2 K.

Prior to this work, the realized cryogenic TIA[8] in the literature dissipated mW power. The high power consumption forces the TIA to be placed at the 4.2 K stage at which sufficient cooling power is available. TIAs with the same resistive-feedback architecture as the room-temperature TIA[4] achieve

a 4-times lower noise and a similar bandwidth for the same gain as shown in [2, 3]. The obtained lower noise at 4.2 K allows gaining a factor of 16 in the read-out times, leading to a µs read-out at 99.9% fidelity even for 10 pA.

To further improve the read-out of SETs by e.g. extending the bandwidth to allow frequency multiplexing, new architectures are envisioned. Adding a cryogenic current preamplification based on bipolar technologies such as heterojunctions (HBT) results in a bandwidth improvement up to 650 kHz[9, 10, 11]. Taking a different direction, researchers leveraged the binary origin of the read-out output (either $|0\rangle$ or $|1\rangle$) and proposed a current comparator that only detects if the SET current is above or below a current threshold[12]. This new method achieves a detection time of a few µs, yet only for currents of the order of nA. Combining the comparator with a bipolar current amplifier will enhance the detection capabilities for a fast read-out time and reduced power consumption of a few µW. This new architecture leverages the binary-origin of the read-out output to greatly improve the SNR for a higher speed and lower power consumption, and certainly constitutes the best actual solution for single-SET state detection. However, this scheme is incompatible with usual frequency-multiplexing techniques and no clear path towards scalability has yet been presented.

TIAs are also useful for the more time-consuming energy-level characterization of quantum states in the quantum-dot devices (see Figure 4.1c). The experimental setup for current measurement is far less involved than other characterization methods such as reflectometry (see chapter 5 for more details) which require additional elements (bias-tees, directional couplers, inductors, ...). However, current TIAs are unable to detect transport in the regime of a few excess charges[13] that would result in very low currents. The first charge-level with tunneling rates t of about 10 MHz[13] results in currents $e \times t$ of only 2 pA. In the already presented cryogenic TIA[8], such arrangements still use cables of tens of cm for thermal isolation between the sub-K and 4.2 K stages at which the TIA dissipates a few mW. This cable lengths results in additional measured leakage of a few 10's pF limiting the minimum measured current. Leakage can be suppressed by placing the TIA closer to the quantum device at the sub-Kelvin stage.This implementation requires amplifiers to operate at sub-K temperatures[11] assembled with a quantum device a few mm-away. In such a scheme, it becomes primordial to operate at low power consumption (typically smaller than 10's µW) to avoid thermal destruction of quantum effects in quantum dots[7].

Complete integration at µm distance between amplifier and quantum device can be reached by producing the classical electrical circuit and the quantum device in the same semiconductor technology.The recently developed CMOS spin qubit made on an industrial foundry for 300 mm wafers[14] would open this route towards on-chip integrated electronics for qubit control and read-out. The measurement throughput can be increased by parallel processing requiring low-footprint identical amplifiers which motivated us to use industrial state-of-the-art silicon-based CMOS technologies for the development of cryogenic electronic circuits.

In this chapter, we present our first realization of a cryogenic TIA for the fine measurement of quantized transport in quantum dots. The TIA[15, 16] is designed to achieve a high gain for low power dissipation to measure the current through a nearby-placed high-impedance quantum-dot device as a function of varying gate and bias voltages at sub-Kelvin temperatures. The cryo-TIA is based on an Operational Amplifier (OP-AMP) with on-chip resistive feedback for a high gain of 10^7 V/A at cryogenic temperatures. The OP-AMP is made of a two-stage Miller-compensated operational amplifier dissipating only 1 µW, resulting in a negligible heat load compared to the usually available cooling power of a ^{3}He/^{4}He dilution refrigerator (typical cooling power $\gtrsim$100 µW at 100 mK). We systematically investigate the temperature dependence of gain, gain-bandwidth product, and noise between 0.25 and 300 K with respect to results presented in chapter 3. To demonstrate the low-temperature operation of a quantum device with classical electronics, a nanometer-sized quantum-dot device was wired-bonded to the TIA after positioning the respective chips

at a few-mm distance. The configuration with the TIA at cryogenic temperatures is compared with the configuration with the TIA outside the cryostat at room temperature linked to the quantum device by meter-scale wiring. We show that the cryogenic configuration is less prone to current leakage and achieves a better bandwidth. Further improvements are identified to increase the performance of the cryogenic TIA with respect to bandwidth and noise for the same power consumption of a few µW in future designs.

Finally, to reach complete integration, we validate the use of the cryo-TIA down to 10 mK by characterizing the low-frequency behavior of a quantum device made in the same silicon substrate following the foundry design rules. The cryo-TIA and quantum device are incorporated in a larger circuit with GHz voltage excitation made with circuits presented in chapter 3 to probe the high-frequency behavior of the quantum device operating as a charge pump. We discuss the impact of the high-frequency dissipating electronics on the nearby quantum device. The on-chip realization of a circuit combining classical and quantum devices demonstrates the full capabilities of FD-SOI CMOS technologies down to 10 mK for quantum computing applications.

II Design of the cryogenic transimpedance amplifier

The required high gain of the TIA prevents the use of capacitive TIAs that would benefit from a noiseless feedback as the transimpedance gain Z_T, defined as a ratio of capacitances, results in an unreasonable die area. As an example, considering a unit 5-metal-level Metal-Oxide-Metal capacitor of 10 fF for an area of 1.5 µm^2, a gain of $Z_T = 10^7$ V/A requires a second capacitor of 100 nF that would occupy 15 mm^2. The use of resistive-feedback is mandatory and mitigation of the associated thermal noise is primordial. At room-temperature, the few examples that reach lower input-referred noise[17] use pre-amplification stages before the resistive TIA with e.g. bipolar current pre-amplification or with a noiseless capacitive-feedback pre-amplification. These architectures require the design of multiple amplifiers and lead to a complicated and intricate design. As a first step towards the direction of low-noise TIA, we designed a TIA based on an operational amplifier (OP-AMP) and a resistive feedback. The OP-AMP architecture consists of an N-type differential stage with active load followed by a P-type common-source amplification stage. This architecture is well-known and allows reaching high open-loop gain for low power[18].

a) Implementation details

Using the characterizations of the 28nm FD-SOI thin-oxide devices presented in chapter 2, we detail the operation of the TIA composed of an Operational Amplifier (OP-AMP) with a resistive feedback (see Figure 4.3). The feedback polysilicon resistor R_{fb} (150 nm width, 31.5 µm length) sets the desired current-to-voltage closed-loop gain $Z_T \simeq 11 \times 10^6$ V/A of the TIA. The power consumption P_{DD} of the TIA that needs to be compensated by the cooling capacity of the fridge consists of the sum of the joule heating in R_{fb} and the OP-AMP dissipation. For the transimpedance gain Z_T and maximum output voltage of 1 V set by V_{DD}, the maximum measurable current in absolute value is equal to $I_{max} = V_{DD}/(2Z_T) = 46$ nA assuming a symmetric transfer function. In the resistive-feedback architecture, the measured current passes through R_{fb} resulting in a heat dissipation of only $R_{fb}I_{max}^2 = 23$ nW. For such low values of measured currents, the TIA power consumption is dominated by the OP-AMP. Moreover, the TIA performance is closely related to the OP-AMP specifications on which we focus in the next paragraphs.

The OP-AMP is composed of two amplifying stages: an N-type input differential stage with P-type active load, and a P-type output stage, made of thin-oxide Regular Threshold Voltage (regular well structure) devices. The OP-AMP total open-loop gain $G_{ol} = A_1A_2$ where A_1 (resp. A_2) is given by the voltage gain A_V of the first (resp. second) stage corresponding to the intrinsic voltage gain of transistors

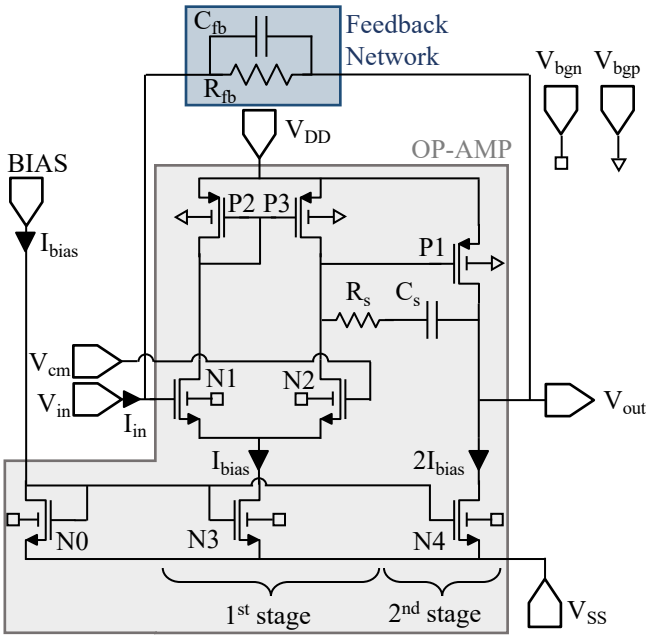


Figure 4.3 | Transistor implementation of the transimpedance amplifier.
The Miller-compensated two-stage topology of the OP-AMP mounted as TIA with the feedback network. The expected TIA gain is $G = V_{out}/I_{in} = R_{fb}$. The common-mode V_{cm} is set to 0.65 V. R_s and C_s ensure the stability of the two stages. In this work, a polysilicon resistor R_s in series with a MOM capacitor C_s ensures the two-stage stability with a phase margin above 75° from 30 nW to 30 µW. The MOM capacitor C_{fb} in parallel with R_{fb} ensures the stability in closed-loop mode considering the interconnection capacitance $C_{in} \simeq 3\,\text{pF}$ present at the input V_{in}. Transistor sizes are given in Table 4.1.

N1 and N2 (resp. P1) indicated in Table. 4.1. The high G_{ol} of 110 dB ensures that the closed-loop gain remains constant as a function of the device impedance by keeping a good virtual ground at the V_{in} input. The NMOS current sources shown (Figure 4.3, bottom MOSFETs) bias the two stages by copying the reference bias current I_{bias}. The first stage (resp. second stage with wider transistors) is biased with I_{bias} (resp. $2I_{bias}$). The OP-AMP power consumption $P_{DD} = 3V_{DD}I_{bias}$ is given by the product of the power-supply voltage $V_{DD} = 1\,\text{V}$ and the added currents of the two branches $3I_{bias}$.

b) Power dissipation and self-heating within the amplifier

Reducing the OP-AMP power to a few µW to reduce its impact on the measured quantum device requires to decrease the biasing current I_{bias}. We decide to maximize the power-efficiency of the OP-AMP by sizing all transistors to operate in weak inversion with g_m/I_{bias} of about $\sim 20\,\text{V}^{-1}$ at room-temperature. From room-temperature simulations, the OP-AMP operates for I_{bias} from 10 nA to 10 µA leading to a power consumption from 0.03 to 30 µW. For the obtained P_{DD}, we can estimate the MOSFET local temperature increase with respect to the fridge temperature due to self-heating effects. From self-heating studies at 4.2 K[19], the computed temperature increase would be 2 mK at 0.1 µW (respectively 1.8 K at 100 µW) if all power was consumed in only one transistor. However, only N1 is directly connected to the quantum device consuming only one ninth of the power dissipated within the TIA (assuming a typical value of $V_{ds} = V_{DD}/3$). In the mid-range at 1 µW, heating would then lead to only a 2 mK higher temperature of N1. At sub-Kelvin temperatures, assuming the same saturating trend below 10 K as observed in [19], the temperature increase will remain identical with a negligible heating of the Device Under Test (DUT).

c) Bandwidth and stability analysis

The dominant pole in the frequency response for V_{out}/V_{in} of the OP-AMP corresponds to the Miller capacitance of the second stage $(1 + A_2)C_s$ and the output impedance of the first stage $R_{out,1}$ leading

Table 4.1 | Transistor properties at 300 and 4.2 K.

Single MOSFET properties from foundry models at 300 K and custom models at 4.2 K for a power consumption $P_{DD} = 1\,\mu W$. Cryogenic models are obtained by fitting a UTSOI model[20] on cryogenic transistor data from [21] in saturation.

MOSFETs	L (µm)	W (µm)	T (K)	A_V (dB)	g_m (µS)	g_{ds} (nS)	g_m/I_{ds} (V^{-1})
N1, N2	0.6	20	300	37	5.8	1.6	35
			4.2	47	12	-	72
N0, N3	0.3	20	300		12	13	35
			4.2		21	-	61
N4	0.3	40	300		23	19	35
			4.2		21	-	32
P1	0.4	40	300	43	22	16	33
			4.2	59	15	-	23
P2, P3	0.4	20	300		5.6	4.4	34
			4.2		8.2	-	50

to a cut-off frequency $f_{op} = 1/\left[2\pi R_{out,1} C_s \left(1 + A_2\right)\right]$ where C_s is the added compensation capacitance, and $R_{out,1}$ is the first-stage output impedance (cf. Table 4.2). This leads to a simulated gain-bandwidth product GBW$= f_{op} G_{ol}$ at 1 µW of 32 kHz at 300 K (see Table 4.2) which would be the maximal TIA bandwidth when measuring an infinite impedance current source (without C_{fb}). Assuming that N1 and N2 mainly contribute to A_1 and $A_2 \gg 1$, the approximated expression GBW$= g_m^{N1,N2}/2\pi C_s$ differs less than 1 % from the predictions at room-temperature. This expression for GBW can then be used at 4.2 K where the conductance g_{ds} is not reliably known giving a GBW of 66 kHz. The second pole corresponds to the TIA output impedance R_{out} in parallel with the large access capacitance $C_{cables} \simeq 100\,pF$. The associated cut-off is shifted to high frequency (>1 MHz) thanks to the low output impedance $R_{out} \simeq 1/A_1 g_m^{P1}$ induced by the TIA feedback (see Table 4.2).

The AC stability analysis of the OP-AMP was performed with foundry models by extrapolating results down to cryogenic temperatures. A polysilicon resistor $R_s = 10.6\,k\Omega$ in series with a MOM capacitor $C_s = 29.1\,pF$ (see Figure 4.3) ensures the two-stage stability with a phase margin above 75° from 30 nW to 30 µW. The MOM capacitor $C_{fb} = 3.2\,pF$ in parallel with R_{fb} ensures the stability in closed-loop mode with an interconnection capacitance $C_{in} \simeq 3\,pF$ present at the input V_{in}.

The added capacitors C_s and C_{fb} for the two-stage stability have an impact on the TIA bandwidth. The RC feedback network sets the maximal TIA bandwidth $f_{fb} = 1/2\pi R_{fb} C_{fb} = 4.29$ kHz (see Table 4.2) regardless of the measured device impedance R_m as long as $R_m > R_{fb} \times f_{fb}/GBW \simeq 715\,k\Omega$.

d) Evaluation of charge-carrier heating by the amplifier

In absence of prior knowledge of noise at cryogenic temperature for the used technology, no further optimization has been performed in the design procedure. With room-temperature models, the TIA output-noise is limited by the flicker noise in the first stage that is amplified by the second. P2 and P3 represent 64 % of the output noise, contributing 98 % together with N1 and N2. The output noise generates a voltage noise directly on the measured device at the input. This kick-back voltage noise is independent of the DUT impedance and is estimated to be equal to $350\,\mu V_{rms}$ at 300 K integrated from 1 Hz to 1 GHz.

Table 4.2 | Operational amplifier and transimpedance amplifier specifications at 300 and 4.2 K.
Predicted OP-AMP and TIA specifications at $P_{DD} = 1\,\mu W$ obtained from the MOSFETs properties (in Table 4.1) and passive elements at 4.2 and 300 K (see Figure A.1). G_{ol}, $R_{out,i}$, GBW are respectively the open-loop gain, the output impedance of the i-th stage and the gain-bandwidth product of the OP-AMP. R_{out} is the output impedance of the TIA. Gain and range of linear response are estimated for the TIA.

Circuit	Property	Temperature 4.2	Temperature 300	Unit K
OP-AMP	G_{ol}	106*	114	dB
	$R_{out,1}$	-	167	MΩ
	$R_{out,2}$	-	29	MΩ
	GBW	66	32	kHz
TIA	Gain	10.6	10.6	10^6 V/A
	Range	±47	±47	nA
	R_{out}	298*	642	Ω
	$R_{fb} /\!/ C_{fb}$	4.29	4.29	kHz

*Estimated from A_V of N1, N2, and P1 (Table 4.1)

These random voltage fluctuations are smaller than the usually applied V_{ds} in the observed Coulomb blockade features but remain non-negligible and could result in potential electron heating. However, due to the presence of capacitance at the TIA input from parasitics (bonding pads, interconnections,...) and at the TIA output from cables, the kick-back noise reduces to $32\,\mu V_{rms}$ with 3 pF at the input and 100 pF at the output from cables (see Table D.1). The frequency-integrated kick-back noise is dominated by high-frequency thermal noise expected to decrease at cryogenic temperatures. This reduced voltage noise V at 300 K is equivalent to a charge-carrier temperature of $eV/k_B = 370$ mK and might decrease at cryogenic temperature. As the charge-carrier temperature is about a few 100's mK in usual setups, the noise seen by the quantum device at the TIA input stays at reasonable values.

All transistors are placed in a triple well structure (N-well for PMOS and P-well for NMOS, the latter being isolated from the substrate by a surrounding N-well, see chapter 1). This well structure allows for independent back-gate biasing of NMOS (resp. PMOS) at voltages V_{bgn} (resp. V_{bgp}). In this chapter, back-gate biasing voltages are kept at 0 V.

III Cryogenic characterization of the current-to-voltage amplifier

As a first step towards using the TIA to measure a nearby quantum device, we characterized the electrical properties of the TIA and OP-AMP from 0.25 to 300 K. The IC chip is wire-bonded to a package placed on a PCB with routing and decoupling capacitors (see pictures of the PCB and IC chip in Figure 4.4). Testing from 300 to 4.2 K is done with a dip-stick in liquid helium, liquid nitrogen, and ambient air. Down to 50 mK, the tests are done in a wet dilution fridge.

The best low-power/high-bandwidth trade-off in analog circuits usually coincides with transistors operating in moderate inversion. We study the TIA over a wide range of bias currents I_{bias} (setting the power consumption) to find the best bias with regard to power consumption and bandwidth.

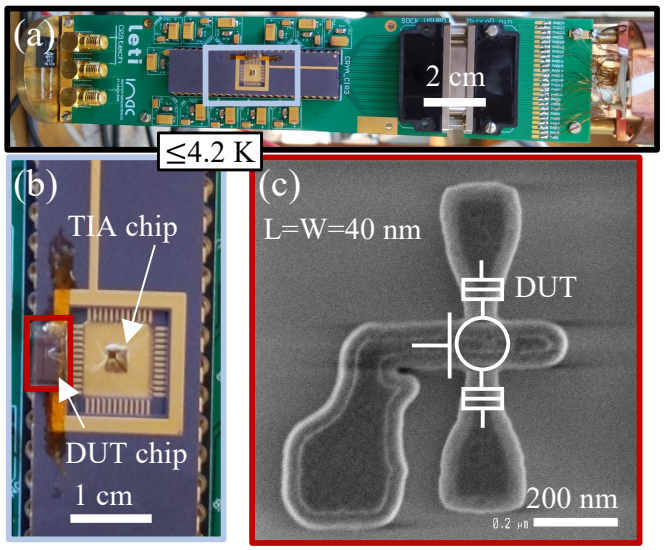


Figure 4.4 | Picture of the experimental devices.
(a) PCB with the co-integrated TIA and quantum dot device to be placed on a dip-stick ($\geq$ 4.2 K) or inside a dilution fridge (down to 10 mK). (b) The FD-SOI chip with the TIA circuit is bonded on a 48-pin DIL package. The quantum dot chip (DUT) is glued nearby and wire-bonded to the CMOS chip. (c) SEM image of the FD-SOI quantum dot with source and drain contacts and gate dimensions $L = W = 40$ nm.

a) DC transfer function of the transimpedance amplifier

The TIA output V_{out} as a function of the input current I_{in}, shown in Figure 4.5a, exhibits a rail-to-rail linear behavior with a transimpedance gain $Z_T = 11.6 \times 10^6$ V/A from -35 to 55 nA both at 50 mK for a 1 μW power consumption and at 250 mK for a larger range of P_{DD} allowed by the higher cooling power. An identical transfer function is observed at higher temperatures 77 and 300 K as the TIA gain is set by the temperature-independent feedback resistor R_{fb} (see Figure A.1 and D.2). The transimpedance gain Z_T and the supply voltage $V_{DD} = 1$ V set the nominal current span $\Delta I = V_{DD}/Z_T = 86$ nA while the common-mode voltage $V_{cm} = 0.65$ V sets the center of the span $I_0 = (V_{cm} - 0.5)/Z_T = 13$ nA (see Figure 4.5a). Due to increased threshold voltage $V_{th} \simeq 0.6$ V at 4.2 K and the need for $V_{cm} > V_{th}$ (saturation region), V_{cm} cannot be chosen equal to the ideal value $V_{DD}/2 = 0.5$ V thus leading to a slight asymmetry in the measured current-range for linear response.

b) AC measurements

To measure AC properties of the TIA and OP-AMP, a 1 MΩ resistor R_1 is connected to the V_{in} input at a close distance of 5 cm on the PCB at the cryogenic stage. Placing the resistor close to the input reduces the interconnection capacitance to avoid an additional cut-off in the measurement chain (see Figure D.1a). The TIA with the input resistor can be equivalently seen as a TIA measuring a 1 MΩ DUT (see Figure D.1c) as well as a voltage-amplifier of gain $R_{fb}/R_{in} = 10$ V/V (see Figure D.1b). Adding the resistor allows extracting electrical quantities (e.g. bandwidth and noise) of the TIA and OP-AMP at the same time in relevant experimental condition from 300 to 0.25 K. The low-frequency voltage noise of the TIA is measured with a lock-in amplifier by averaging the demodulated signal from the IC output. The input current noise of the TIA is obtained by dividing the output voltage noise by the transimpedance Z_T (see Figure 4.5b). The noise as a function of the lock-in reference frequency f, plotted in Figure 4.5b, shows the characteristic $f^{-1/2}$ dependence for flicker noise. The flicker noise increases by only 30 % from 300 K to 250 mK which is ten times less than what has been observed in bulk CMOS[22]. This difference in flicker noise between SOI and bulk technologies is known at room-temperature[23]. The SOI flicker noise has been reported to even decrease at 77 K for some PMOS devices[24]. The minimal achievable current noise at 4.2 K, given by the broad-band thermal noise of the 11.6 MΩ feedback resistor, is 5 fA/$\sqrt{\text{Hz}}$. The TIA has an enhanced noise of 300 fA/$\sqrt{\text{Hz}}$ at 1 kHz resulting from flicker noise (see Figure 4.5b). As

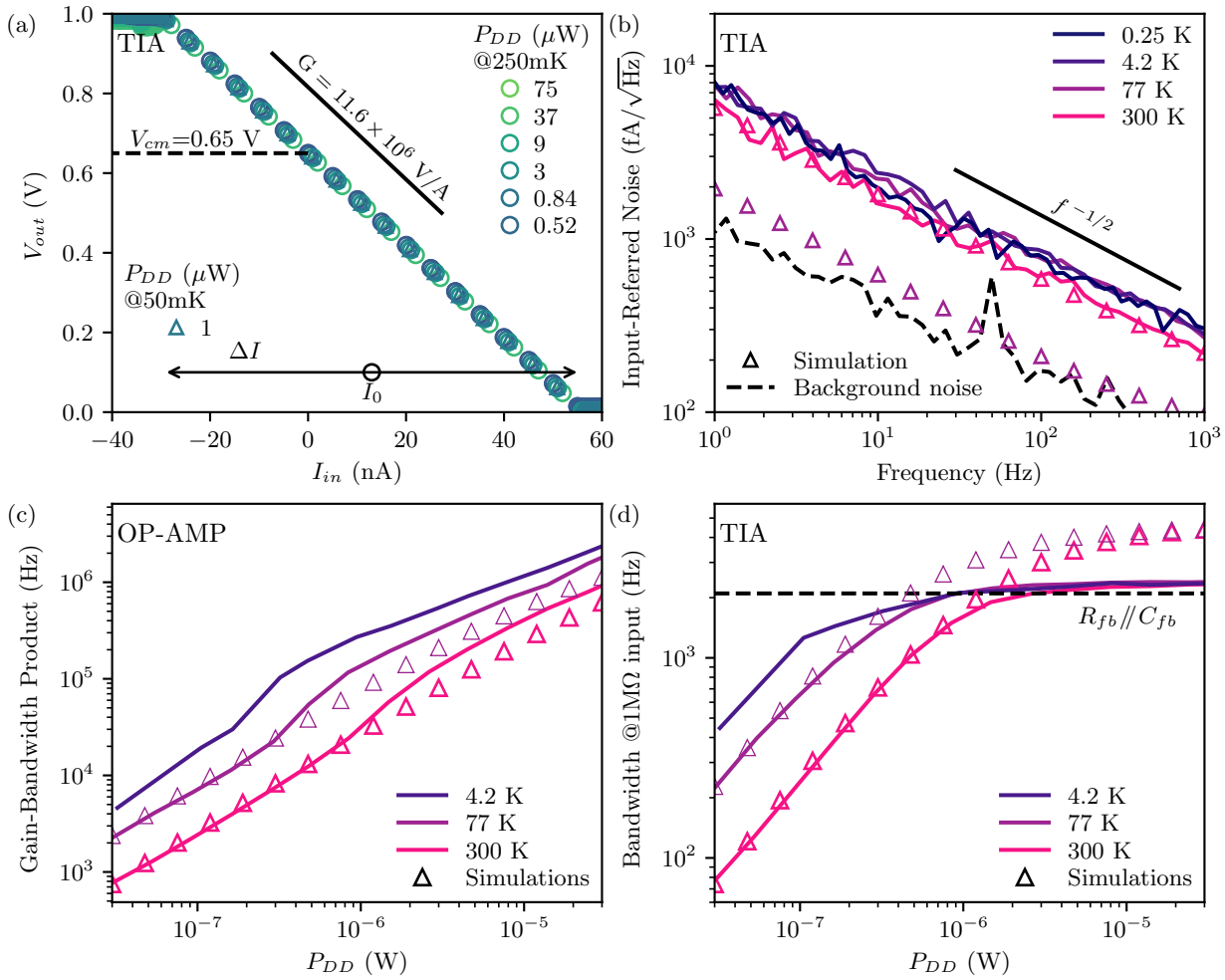


Figure 4.5 | Characterization of the operational and transimpedance amplifiers from 300 to 0.25 K.
(a) TIA output voltage V_{out} as a function of the input current I_{in} for several values of power consumption $P_{DD} = 3I_{bias}V_{DD}$ at 0.25 K and for $P_{DD} = 1$ μW at 50 mK. The impedance gain $Z_T = R_{fb}$ is extracted from the slope. (b) Input-referred current noise of the TIA as a function of frequency for different temperatures, with the background noise of our measuring instrument. The straight line indicates the expected slope for the $1/f$ flicker noise. (c) OP-AMP Gain-Bandwidth product and (d) TIA bandwidth when measuring a 1 MΩ impedance as a function of the power consumption P_{DD} for the indicated temperatures.

the overall noise changes only slightly with temperature, we expect that transistors P2 and P3 stay the major contributors at cryogenic temperatures as confirmed with room-temperature models. For a typical measurement of quantum dots with currents of 100 pA, the cryo-TIA achieves a signal-to-noise ratio of 38.2 (fidelity of 99.9%) for an integration time of 172 µs (neglecting bandwidth limitations for now). The noise has to be reduced to reach high fidelities for shorter integration times (see section IV for future improvements).

The frequency response of the TIA measuring a 1 MΩ DUT is equal to the bandwidth of the equivalent voltage amplifier of gain $\simeq 10$ (see appendix D, section II). From the same measurement, we can also extract the OP-AMP gain-bandwidth by multiplying the cut-off by the voltage gain (see Figure D.1b). The bandwidth is measured with the lock-in amplifier from 10 Hz to 50 MHz. Two poles can be identified: one that depends on the bias current, attributed to the Miller capacitance inside the OP-AMP, and one that does not depend on the bias current, attributed to the $R_{fb}C_{fb}$ feedback network (see Figure D.3). In Figure 4.5c, the gain-bandwidth product GBW for the bias-current dependent bandwidth of the OP-AMP is shown as computed from the measured voltage gain and bandwidth at 4.2, 77, and 300 K. The increase of the OP-AMP bandwidth by a factor of 5 at low temperatures corresponds to the transconductance at constant current (similarly, g_m/I_{ds}) largely increasing at cryogenic temperatures in moderate inversion[25]. The second cut-off, independent on the bias current, is shown in Figure 4.5d for the bandwidth data as a function of $P_{DD} \propto I_{bias}$ when measuring a 1 MΩ device. At low power consumption, the frequency limitation is due to the finite gain-bandwidth product of the op-amp while at higher power-consumption, the measured bandwidth is limited by the feedback network $R_{fb}C_{fb}$ at 2.6 kHz, independently of the temperature.

c) Comparison of results with simulations

Within the typical accuracy range for temperature from -40 to 185°C, the foundry model predicts well the noise and bandwidth of the OP-AMP and TIA at 300 K (see Figure 4.5bcd). Simulation results at 77 K still predict the TIA and OP-AMP bandwidth as seen in Figure 4.5cd. The observed difference between simulations and experiment in the TIA bandwidth at high P_{DD} (Figure 4.5d) comes from the variability with respect to the nominal resistance value of the highly doped resistor R_{fb}. Models predict a decreased flicker noise at 77 K not observed experimentally (see Figure 4.5b). Below 77 K, the model stops converging and is unable to predict the change in circuit specifications from 77 and 4.2 K.

IV Integrated circuit improvement in next design iteration

The use of back-gate voltage, as can be done with the FD-SOI technology, gives possibilities to improve the performance of the cryogenic TIA. Firstly, the threshold gate-voltage of the transistors can be lowered at the usual level of about 0.4 V by tuning the back-gate voltage thus allowing the common-mode voltage to lie at $V_{DD}/2$, resulting in a symmetric range of ±43 nA for the measurable currents. Secondly, the flicker noise, presumably mainly coming from the top gate-oxide/silicon interface, will be reduced for applied back-gate voltage as the conduction is displaced towards the much cleaner bottom-gate interface.

The bandwidth of the cryogenic TIA is limited at 2.6 kHz by the RC-network in the feedback loop. While the resistor R_{fb} sets the necessary gain, C_{fb} was added to ensure stability with the large cable capacitance ~ 100 pF present at the V_{cm} pin. C_{fb} can be greatly reduced when measuring a device at mm distance with reduced interconnection capacitance ~ 3 pF compared to the cable capacitance. A capacitance C_{fb} of about 10 fF could be sufficient to ensure stability and would lead to a bandwidth of 1 MHz, compatible with the 1 MHz GBW product of the OP-AMP at 10 µW. At such high frequency,

the $1/f$ flicker noise contribution will be highly reduced, reaching the broad-band thermal noise that is expected to even decrease at cryogenic temperature.

Adding a pre-amplification stage with noiseless feedback[17] or with bipolar current amplification[9, 10] will improve the signal-to-noise ratio of the measurement chain with cryo-TIA to meet the requirement for fast quantum dot characterization and read-out. As an example, a current pre-amplification stage of gain 200 A/A with input-referred noise of $19\,\mathrm{fF}/\sqrt{\mathrm{Hz}}$ for a power of 0.8 µW as in [11] is sufficient to reach an input-referred noise of $\sim 20\,\mathrm{fF}/\sqrt{\mathrm{Hz}}$ for a bandwidth of 2.6 kHz, a gain of 2×10^9 V/A, and a total power consumption of only 1.8 µW with the presented cryo-TIA.

V Assembly of quantum-dots devices with transimpedance amplifiers down to 4.2 K

a) Assembly and Measurement with the cryogenic transimpedance amplifier

We have employed the cryo-TIA to measure the current through high-impedance quantum dot devices. We present a control experiment to validate the operation of the cryo-TIA for Quantum Dot measurements and compare with a room-temperature TIA. This experiment is realized at 4.2 K to benefit from the quick cool-down and warm-up to easily switch between RT-TIA and cryo-TIA on daily basis with the dip-stick in liquid helium. At 4.2 K, the thermal energy k_bT of 0.36 meV has to be smaller than the typical energy level spacing of the quantum-dot structure to measure Coulomb blockade conduction. The energy level spacing of a single dot relates to the charging energy to add one electron in the dot as $E_C = e^2/C$ (typically equal to a few meV) with $C = \epsilon_0\epsilon_{SiO2}A/EOT$ the dot capacitance, ϵ_0 the permittivity, A the dot area, and EOT the SiO_2-equivalent oxide thickness. Assuming an EOT of 1 nm with modern high-k dielectrics[21], the relation $E_C > 10 \times k_bT$ ensuring clear Coulomb blockade features gives a typical area of $A \simeq (40\,\mathrm{nm})^2$. Already for a single dot, the condition to observe Coulomb blockade effects requires small dots. As for double- or triple-dot structures, the considered overlap between single-dot energy levels requires an even smaller dot dimension (or lower temperature) to resolve Coulomb blockade effects. For this reason, we chose a state-of-the-art FD-SOI single quantum dot fabricated at CEA-Leti with nanowires such as [14] as presented next.

We mounted the semiconductor chip containing the quantum dot at a few-mm distance of the chip embedding the cryo-TIA (see Figure 4.4). Electrical connections between the two chips were made using standard bonding wires. The quantum dot device consists of an n-type SOI transistor with a 40-nm-wide silicon channel and a 40-nm-long gate electrode (see Figure 4.4 for a SEM image). Because of the intentional implementation of wider insulating spacer layers, the gated region is well separated from the heavily doped source/drain contact regions. As a consequence, the electron accumulation induced by a sufficiently positive gate voltage results in the formation of an electron quantum dot under the gate [26]. The dot is weakly coupled to the source and drain contacts via tunnel barriers naturally formed under the gate spacers. Thanks to the small quantum-dot size, electron transport through the quantum dot is governed by the Coulomb blockade effect[26], even at 4.2 K.

The right panel of Figure 4.6b shows a color-scale plot of the source-drain current I_{ds} as a function of the voltages V_{gs} and V_{ds} applied to the gate and to the drain electrodes. This measurement of the stability diagram $I_{ds}(V_{gs}, V_{ds})$ was taken with the quantum dot and the cryo-TIA at 4.2 K. The cryo-TIA was set to dissipate 1 µW in order to reach a 2.6 kHz bandwidth for a measured impedance $\geq$1 MΩ. The two-dimensional plot shows the characteristic diamond structures where current transport is suppressed due to Coulomb blockade. Each of these regions is associated with a well-defined integer number, N, of excess electrons in the dot, increasing progressively from left to right (we estimate $N \sim 100$ in the

displayed V_{gs} range).

b) Comparison between room-temperature and cryogenic-temperature amplifiers

To confirm the faithful operation of our quantum-classical circuit, we carried out a control experiment by replacing the cryo-TIA with a room-temperature transimpedance amplifier (RT-TIA) with comparable performance. The quantum dot can be measured by either the cryo-TIA or a usual TIA at room-temperature by respectively disconnecting or reconnecting a pin of the DIL package. More specifically, the RT-TIA consisted of an off-the-shelf operational amplifier with the same circuit architecture as in the cryo-TIA and a feedback resistor of 10 MΩ. It exhibits a 18 MHz gain-bandwidth product and a 2 mW power consumption[27]. In Figure 4.6a, we compare the frequency response of the cryo-TIA and the RT-TIA when measuring a resistance of 14 MΩ. While the cryo-TIA reaches the expected 2.6 kHz bandwidth, the response of the RT-TIA is cut at 600 Hz. We ascribe this to the RC time associated with pi-filters with nF-capacitance and the input-impedance of the RT-TIA. The pi-filters for moderate- to high-frequencies filtering were built-in in the matrix box used to dispatch cryogenic measurement cables to room-temperature apparatus.

The stability diagram $I_{ds}(V_{gs}, V_{ds})$ measured with the RT-TIA and with the quantum dot device at 4.2 K is shown in the left panel of Figure 4.6b. We find no significant differences with respect to data from the cryo-TIA (right panel). We emphasize that for a fair comparison the two data sets were acquired by spanning the same V_{gs} and V_{ds} ranges in the same amount of time (about 2 min). For further clarity, we compare in Figure 4.6c representative horizontal line cuts, $I_{ds}(V_{gs})$, at $V_{ds} = \pm 1$ mV. The Coulomb blockade oscillations are clearly visible with comparable quality. The cryo-TIA is much less power consuming (1 µW against 2 mW). Moreover, it shows a lower current-offset (0.3 pA against 26 pA for the RT-TIA), which can be attributed to the higher leakage current associated with the long cable connecting the quantum dot to the RT-TIA. We note that for both measurement configurations, the TIA offset was zeroed by adjusting the common-mode voltage V_{cm} when the device was open, i.e. no current flowing.

The co-integration experiment at 4.2 K validates the use of the cryo-TIA to characterize quantum-dot devices. At 4.2 K, the surrounding exchange gas with the liquid helium bath and the out-of-chip wire-bonding help to thermalize the quantum-dot and the cryo-TIA chips to base temperature. The low power of the cryo-TIA allows to place it at sub-Kelvin temperature in the vicinity of the DUTs. At sub-Kelvin in dilution fridges, thermalization only happens by thermal conduction via the mounted PCB which does not prevent an eventual temperature gradient between the two chips. Thermalization becomes even harder when both the cryo-TIA and the quantum device are in the same silicon chip at a few µm-distance. In the next section, we demonstrate and study such fully on-chip integration of quantum device and cryo-TIA in the same silicon substrate down to 10 mK. We also added power-consuming circuits for high-frequency excitation to study its impact on the quantum device.

VI Full integration of a double quantum dot with a transimpedance amplifier at sub-Kelvin temperatures with GHz excitation

The small gate lengths of modern industrial CMOS technologies enable the fabrication of nm-sized MOS quantum dots on a 300mm wafer. Tiny MOSFETs turn into quantum devices such as single electron transistors or Double Quantum Dot (DQD) at cryogenic temperature (4.2 K and below) under the right gate-voltage biasing. More complex quantum-dot structures are then built such as in [16, 28].

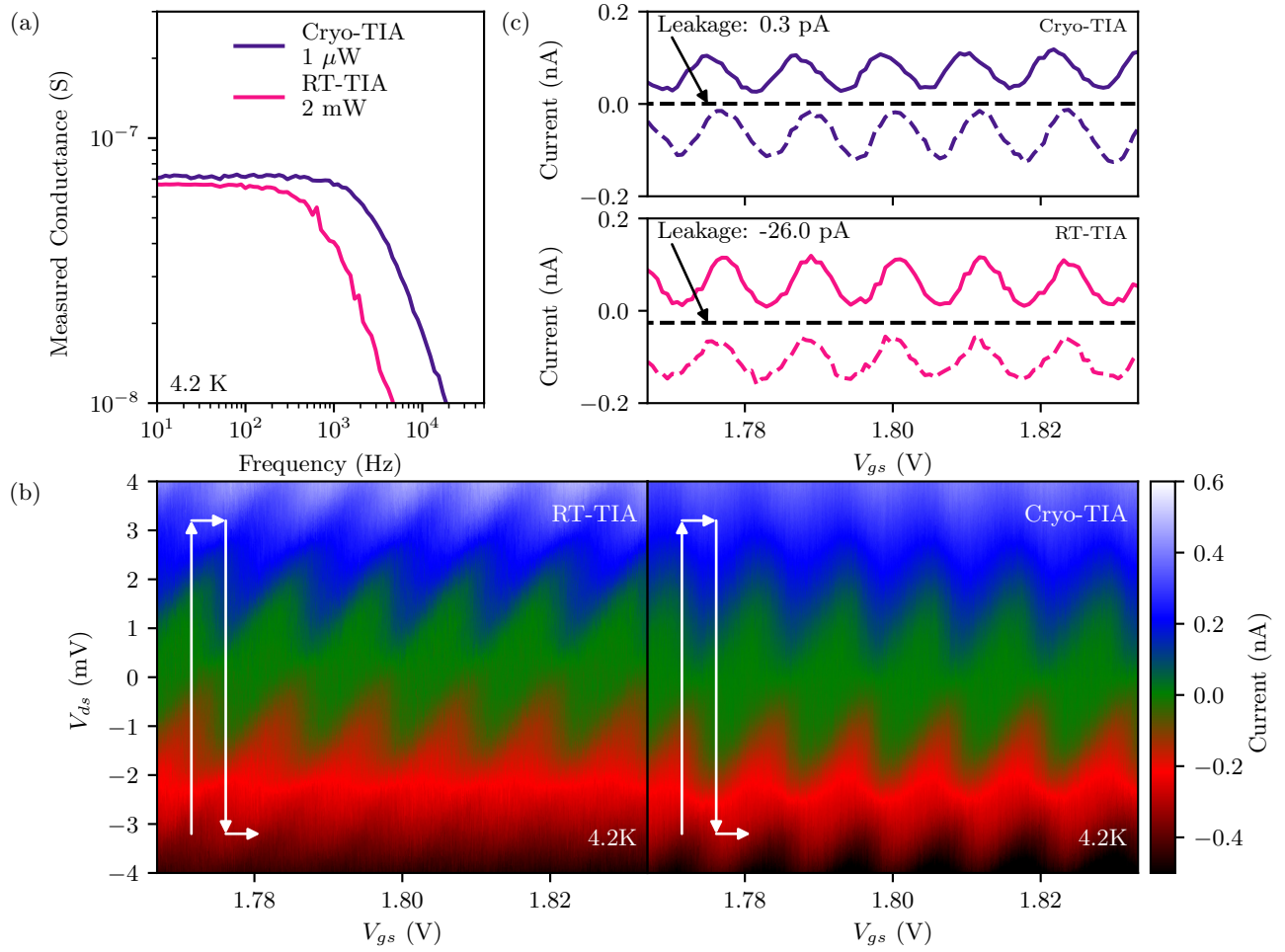


Figure 4.6 | Measurement of a co-integrated single quantum dot with the transimpedance amplifier at 4.2 K.

(a) The frequency response of the RT-TIA at 300 K and the cryo-TIA at 4.2 K when measuring the quantum-dot placed at 4.2 K for a resistance of about 14 MΩ at the same (V_{ds}, V_{gs}) bias point. Extracted bandwidths of 0.6, and 2.6 kHz for, respectively, RT-TIA, and cryo-TIA. (b) Coulomb-diamond structure in the measured current $I_{ds}(V_{ds}, V_{gs})$ for the same single quantum dot at 4.2 K measured with the cryo-TIA and the RT-TIA for the same acquisition time (2 min). White arrows illustrate how voltages V_{ds} and V_{gs} are swept. V_{ds} is a triangular signal of period $2t$ and V_{gs} is stepped each $t = 260$ ms by 100 μV. Coulomb oscillations in (c) are extracted from (b) at $V_{ds} = \pm 1$ mV. The leakage currents extracted by averaging the signals for positive and negative V_{ds} are indicated in the top-left corner.

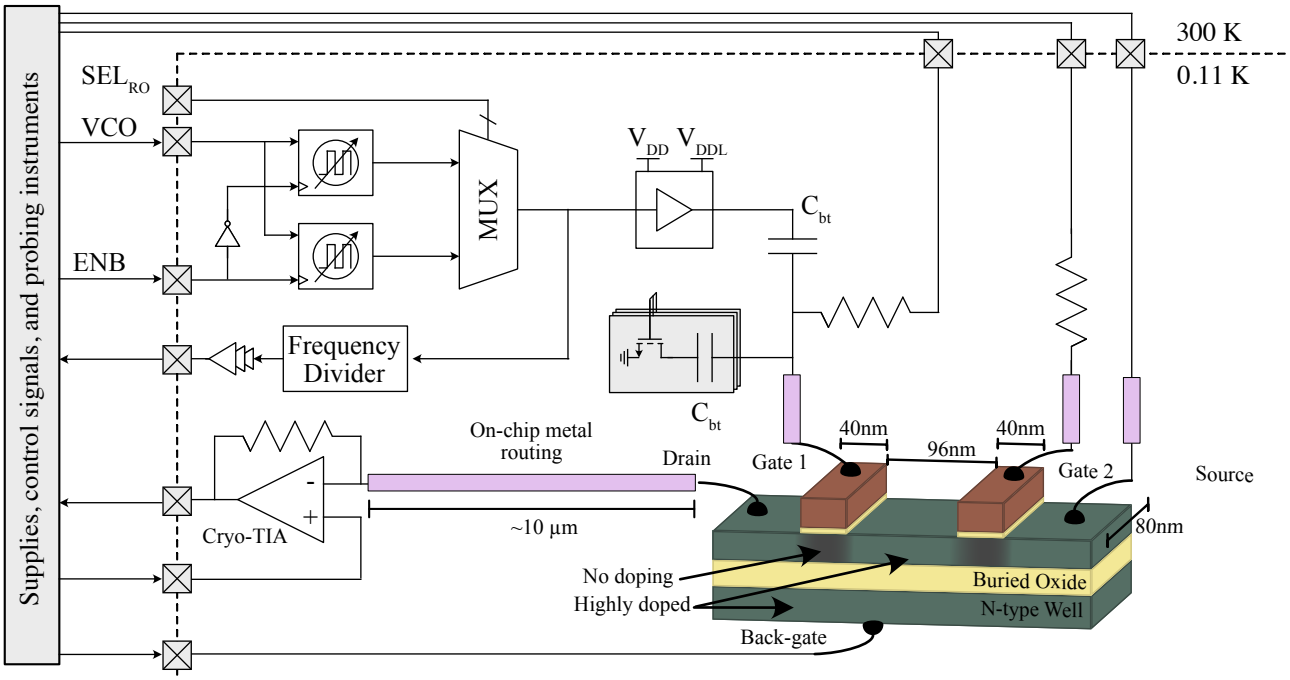


Figure 4.7 | Schematics of the complete on-chip integration of the transimpedance amplifier with the double quantum dot-device.
Quantum-classical chip with the GHz excitation circuit, the cryo-TIA, and the integrated quantum device. We laid out a quantum device as shown in the bottom right corner with two minimally spaced gates on top of a narrow FD-SOI channel. The two 40 nm-long polysilicon gates are deposited on top of a 80 nm-wide FD-SOI channel. The two gates are placed in agreement with the founder design rules, being spaced by 96 nm. High doping dose is injected in the source, drain, and inter-gate areas. The source is directly routed to a pad and linked to room-temperature via constantan wiring. The drain is routed to the on-chip cryo-TIA with 10 µm-long first level metal routing. DC gate-voltages are applied from room temperature through an on-chip bias tee made of a 1 MΩ polysilicon resistor and metal-oxide-metal capacitors routed to high-frequency circuitry. The AC excitation from the high-frequency circuit is superimposed to the DC bias on the first gate.

We combined circuit blocks presented in chapter 3 to form a monolithic circuit containing GHz-voltage excitation of a quantum-dot device with current reading at sub-Kelvin temperature, respecting the maximum power dissipation allowed by the typical cooling power of the dilution fridges. This circuit acts as a proof-of-concept of what could one-day evolve into a silicon quantum processor with dissipating high-speed electronics and sensitive quantum devices. The idea of the realized circuit is to induce charge pumping in a Double Quantum Dot (DQD) with GHz excitation while only having low-frequency room-temperature electronics linked to the sub-Kelvin stage.

a) Circuit architecture

We laid out a circuit realized in the CMOS FD-SOI 28 nm technology as shown in Figure 4.7. The quantum device consists of two 40nm-long polysilicon gates deposited on top of a 80nm-wide FD-SOI channel. The two gates are placed as close as allowed by the foundry design rules and spaced by 96 nm.

Compared to the single dot measured in section V which is optimized to operate as a quantum device with wider spacers, the realized quantum-dot structure consists of two standard MOSFETs in series, initially optimized by the foundry for the realization of classical circuits. High doping dose is injected in the source, drain, and inter-gate areas. At cryogenic temperatures and under mV-level drain-source voltages, this device consists of three quantum dots in series. The highly-doped inter-gate region forms a metallic quantum dot with equally-spaced size-quantized states related to the island area 96 nm × 80 nm. Under each gate, a quantum dot can be created by applying the appropriate gate-voltage around the threshold voltage of the associated FET-behavior. Each quantum dot is linked to its neighboring regions by tunnel junctions. Applying a voltage difference between the source and drain V_{ds} forces a current to flow through the quantum dots, depending on the impedance. Transport through the quantum structure is governed by the lowest tunneling rate of the barriers and by the alignment of quantized states in the V_{ds} window. By applying a strong positive back-gate voltage on the N-type device, eventually the quantum dots under each gate are suppressed[29] leaving only the metallic dot with tunneling junctions on each side to source and drain. In this regime the device resembles an electron pump formed of a single quantum dot with tunable tunneling junctions to contact reservoirs. By applying a square voltage wave on one gate, we will be able to generate non-adiabatic charge pumping which is the generation of a DC current proportional to the excitation frequency[30]. To observe charge pumping, it is required to have full control on the DC biasing of the quantum device (source, drain, and both gates), and on the AC wave amplitude and frequency.

For the control of DC biasing of the quantum device, the source is directly routed to a pad and linked to room-temperature via constantan wiring. The drain is routed to the on-chip cryo-TIA with 10 µm-long first level metal routing (see Figure 4.7). DC gate-voltages are applied from the room-temperature DAC channels through on-chip 1 MΩ polysilicon resistors R_{bt} in order to isolate the DC lines from high-frequency on-chip gate voltages.

High-frequency signals are generated with two ring oscillators (RO_{28}^{10} and RO_{130}^{10} in chapter 3) selectable with a digital multiplexer generating a $1\,V_{pp}$ square wave at frequencies continuously tunable from a few 100's MHz to 7 GHz (see Figure D.7). The excitation frequency is measured at room-temperature after dividing its frequency by 1024 with a chain of 10 dynamic flip-flops ended with a digital buffer. The power consuming frequency-extraction circuitry is only punctually activated to reduce the average heat load at the bottom plate of the dilution fridge. The AC wave amplitude is coarsely tunable with a capacitive divider made of C_{bt} and a digitally-controlled capacitor bank of MOM capacitors to achieve a nominal output amplitude of 167, 91, 50, 24, 14, 7.3, and 4.9 mV depending on the digital selection bits $SEL[2:0]$ (see Figure D.8). A level shifter before the capacitive divider allows continuous fine-tuning of the excitation amplitude from 0.5 to 1 V with a maximum bandwidth of 2.8 GHz (see Figure D.7). The combination of the capacitive divider and level-shifter allows to cover the continuous range of excitation amplitude from 2.5 to 167 mV.

The chip containing the entire circuit mounted on the PCB is placed at the mK temperature stage of a dry dilution fridge (see Figure D.4).

b) On-Chip DC measurements of the double quantum-dot device at 10 mK

In a first step to analyze the circuit, we characterize the DC behavior of the quantum device by deactivating all power-consuming high-frequency circuitry. The TIA dissipates only 1 µW leading to a fridge temperature of 10 mK (see Figure D.10).

The DC current through the double-dot structure is measured by applying $V_{ds} = 1$ mV as a function of the gate voltages with a voltmeter connected to the TIA output (see Figure 4.8ab). Under zero back-gate voltage, transport is dominated by the under-gate dots as attested by the 2D pattern of current triangles as a function of the two gate voltages shown in Figure 4.8c. The observed trianges are typical for transport

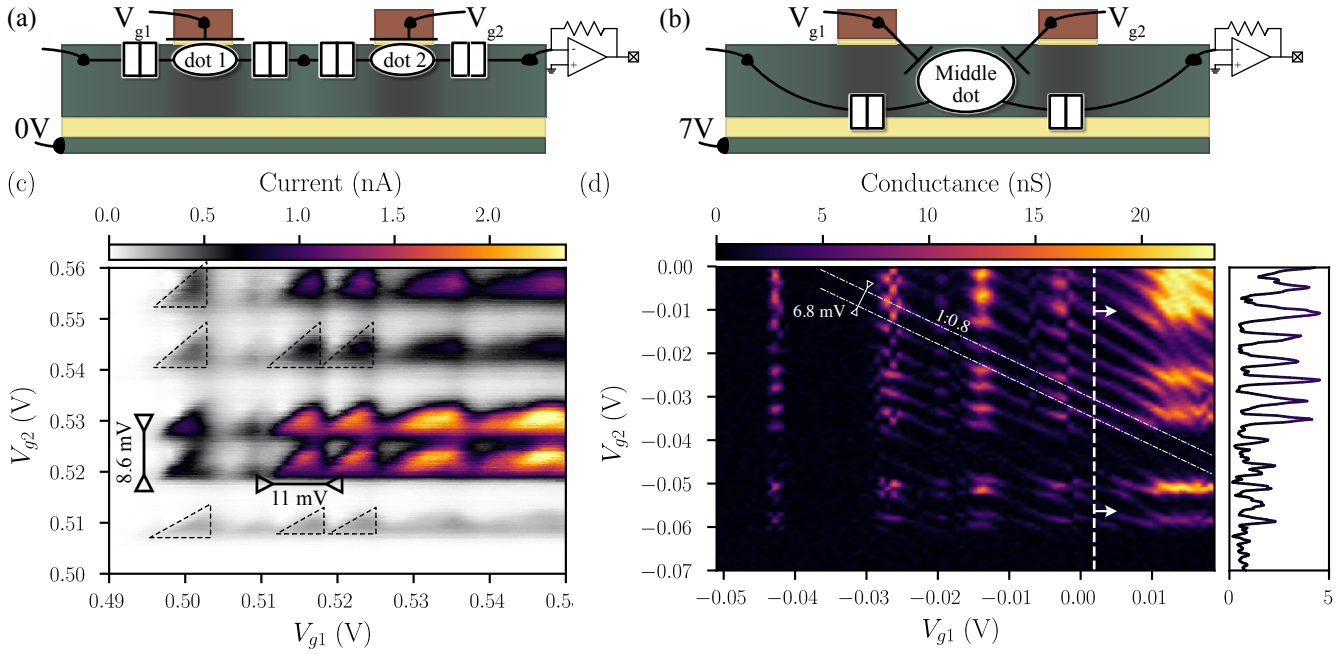


Figure 4.8 | On-chip integration of a double quantum-dot device with the transimpedance amplifier at 10 mK.

Schematic of the on-chip quantum device and equivalent quantum-dot structure for V_{bs} equals (a) 0 V and (b) 7 V coupled with the cryo-TIA in the same silicon substrate to measure stability diagrams of the transport current across the quantum-dot structure at a fridge temperature of 10 mK. (a) Double quantum dot structure under zero back-gate voltage. (b) Single dot situated on the highly-doped metallic island equally coupled to both gates under high $V_{bs} = 7$ V. (c) Current as a function of DC gate voltages V_{g1} and V_{g2} at $V_{bs,dot} = 0$ V revealing triangular structures. (d) Lock-in conductance measurement of the quantum device with $V_{bs} = 7$ V revealing parallel lines of non-zero conductance, signature of a single dot equally coupled to both gates.

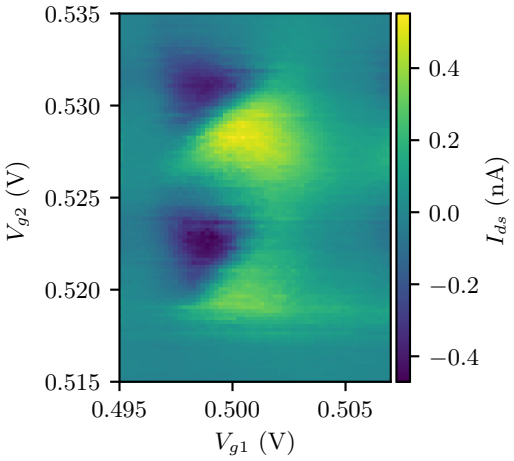


Figure 4.9 | Conduction triangles with on-chip dissipation.
Conduction triangles of the DQD with a small source-drain bias estimated to be $\simeq 100\,\mu V$ at $V_{bg} = 0\,V$ and $T = 110\,mK$. The opposite triangles corresponding to opposite drain-source voltage biases appear due to crosstalk with high-frequency signals generating a voltage noise greater than the imposed DC bias.

across two quantum dots in series. The operating regime of the double quantum-dot device is depicted in Figure 4.8a.

As the back-gate voltage increases, enhancing back-channel conduction, the under-gate dots are bypassed leading to a direct tunneling junction between the metallic dot and the source and drain. As a result, the single dot behavior emerges in the current signal as a function of the two gate voltages, as shown in Figure 4.8b. To measure the current through the single dot configuration with better sensitivity, we measure the conductance $G_{ds} = dI_{ds}/dV_{ds}$ by applying an $5\,mV_{pp}$ AC excitation on V_{ds} and demodulating the cryo-TIA output with a large integration time of 10 ms. Single dot characteristics of Coulomb peaks, equally coupled to each gate, starts to dominate transport as seen in Figure 4.8d with the tilted lines. A cut view of $G_{ds}(V_{g2})$ at $V_{g1} = 2\,mV$ demonstrates clear regular Coulomb peaks spaced by a charging energy of $\sim$7 mV. The peak width is limited by the AC source-drain voltage of 5 mV.

c) GHz-excited double quantum-dot device at 110 mK

With the high-frequency circuitry turned on, the bottom plate temperature increases due to the additional power dissipation. At the nominal operating point with a full-scale 2.8 GHz excitation, the circuit dissipates 295 µW leading to a fridge temperature of 110 mK (see Figure D.10). Although the bottom plate is at 110 mK as indicated by the plate temperature sensor, it is harder to define the local chip temperature. Non-negligible self-heating of the circuit transistors as shown at 4.2 K[19] could induce temperature increase of the quantum-dot device, eventually killing the quantum effects.

By leaving the high-frequency circuitry running in background and deactivating the high-frequency signal path by setting the level shifter V_{DDL} to 0 V, we still are able to observe quantum effects as attested by the conduction triangles observed at zero back-gate voltage and small source-drain bias $V_{ds} \simeq 100\,\mu V$ (see Figure 4.9). Interestingly, we observe simultaneously two opposite-oriented triangles corresponding to both positive and negative polarization. We interpret this behavior as an interference effect via capacitive coupling from the nearby oscillator generating an alternating voltage excitation V_{ds} greater than the applied DC value. Despite this anomalous interference effect, we conclude from the observed triangular features that the device temperature remains under a few Kelvins with double-dot quantum characteristics.

The DC current as a function of both gate voltage is impacted by the high-frequency excitation (see Figure D.11 on the DQD regime). At high back-gate voltage, an interesting feature with triangle structure appears in the single dot regime. Changing the excitation frequency f from 0.4 to 2.1 GHz for an estimated voltage excitation of 5 mV, we obtain the 2D maps presented in Figure 4.10. The reduction in size of the

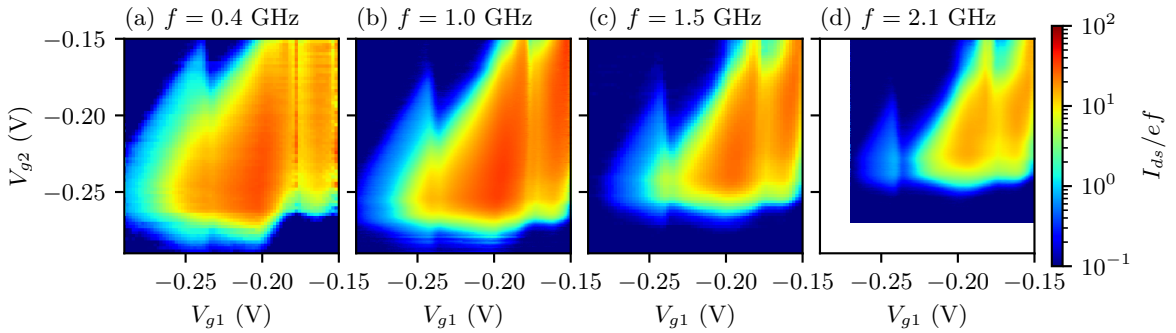


Figure 4.10 | GHz-voltage excitation in the single-dot regime at 110 mK.

Triangle structure in the single dot regime at 110 mK with $V_{bg} = 8\,\mathrm{V}$ appearing when excited with a 5 mV excitation at the indicated frequency f in (a-d). The reduction in size of the triangular features with increasing frequency f is attributed to the decreasing RF amplitude because of the high-frequency cut-off of the level shifter.

triangle features with higher f might come from the level-shifter attenuating the AC gate-signal amplitude at high frequencies (see Figure D.7c). In these plots, we normalize the current by $e \times f$ with e the electron charge which is the expected value of quantized conductance in the presence of charge pumping. The current amplitude remains in the $1 - 30\,ef$ order of magnitude, however no clear quantization plateaus have been observed in any of the acquired data.

VII Conclusion

The important transistor parameters studied in chapter 2 at 300, 4.2, and 0.25 K allowed us to detail the TIA operation down to 10 mK for a 1 µW power consumption to measure high-impedance quantum-dot devices. The TIA is composed of an OP-AMP with an integrated resistive feedback. The two-stage Miller-compensated operational amplifier whose power consumption can be varied from 300 nW to 30 µW via the bias current achieves a 250 kHz Gain-Bandwidth product at 1 µW that rises to 1 MHz at 10 µW. The obtained 1.9 kΩ output impedance combined with the 100 pF cable capacitance placed at the OP-AMP output leads to a higher RC frequency cut-off (5.3 MHz). As a result, the OP-AMP bandwidth is only limited by its intrinsic gain-bandwidth product.

The TIA uses a resistive feedback made of a practically temperature-independent polysilicon resistor R_{fb} to achieve a high transimpedance gain of 11.6×10^6 V/A from room to sub-Kelvin temperatures with a linear response for ±40 nA. When electrically contacted for characterization, the transimpedance amplifier has a large capacitance placed at the input V_{in} from meter-length cables that require the use of the capacitor C_{fb} placed in parallel with R_{fb} to ensure the stability of the TIA circuit. The resulting $R_{fb}C_{fb}$ time limits the maximum bandwidth of the TIA to 2.6 kHz for power consumption above 1 µW. For lower power consumption, the bandwidth is limited by the GBW product of the op-amp and varies with power. The input-referred noise within the TIA bandwidth is limited by the $1/f$ flicker noise with a current noise given by $300\,\mathrm{fA}/\sqrt{\mathrm{Hz}}$ at 1 kHz which shows an increase of only 30 % from 300 down to 0.25 K. The TIA bandwidth can be improved in future iterations to 220 kHz by reducing the parallel capacitance C_{fb}, resulting in a ×9.2 lower $1/f$ noise level at higher frequencies. The cryo-TIA operates from 300 to 0.01 K for a low power-consumption of 1 µW.

At 4.2 K, the cryogenic amplifier was assembled with an FD-SOI single quantum dot positioned a

few mm away to assess its performance. The results where compared to those obtained with a more conventional setup employing a room-temperature TIA connected to the quantum chip via a meter-long cable. For the same acquisition time, both TIAs were able to measure the usual Coulomb-diamond structure in the transport current across the dot. Thanks to reduced interconnection length between cryo-TIA and quantum dot, the cryo-TIA exhibits a low 300 fF leakage current while the RT-TIA shows a leakage of 26 pA.

At a fridge temperature of 10 mK, the TIA measurements on a double quantum dot indicate that the quantum properties due to the confined geometries can be observed down to the lowest temperatures for the fully on-chip integration of measuring electronics and quantum device. Even with a high on-chip dissipation of 295 µW from high-frequency circuitry, quantum effects were preserved in the transport current measurements of the double quantum dot. Although we did not observe charge pumping upon GHz excitation, we have been able to measure the FD-SOI commercially-fabricated quantum dot structure in the double and single quantum-dot regime excited with a GHz voltage tone and measured with the µm-away cryo-TIA. These results suggest that dissipating electronics is not a showstopper for quantum effects and pave the way towards integrated near-qubit electronics for faster and better measurements.

Besides their use for the characterization of quantum-dot stability diagrams, TIAs are implemented to perform single-shot spin-state read-out. By measuring the current of a quantum dot near a current step results in high-speed and low-noise response to a capacitive change occurring due to e.g. moving electrons in nearby dots. Spin-state read-out requires only a binary output and not the continuous analog output of usual amplifiers. With the availability of electronics at low-temperature, one way towards scalability in the number of qubits is to enhance the amplifier specificity in order to exactly match the needs. This allows to reduce the power consumption and increase the read-out speed. This approach of tailored integrated electronics attracts much attention, and examples have already been presented[12]. However, the scalability of such solutions has yet to be proposed as frequency-multiplexing is not always applicable.

Following this idea of tailored electronics, in the next chapter, we tackle the gate-based radio-frequency read-out of quantum devices by proposing a new read-out architecture.

[FR] Mesure en transport de puits quantiques

La mesure du courant au travers de structures à puits quantiques est effectuée à l'aide d'un amplificateur courant-tension, également appelé amplificateur à transimpédance (TIA). À basse température, les mesures de transport électrique des systèmes de puits quantiques donnent accès à leurs propriétés de charge et de spin[1]. L'application des TIA dans ces systèmes quantiques peut être divisée en deux catégories : la détection à grande vitesse des mouvements de charge et la caractérisation précise plus longue des états quantifiés.

Dans ce chapitre, nous présentons notre première réalisation d'un TIA cryogénique pour la caractérisation de puits quantiques. Le TIA[15, 16] est conçu pour obtenir un gain élevé pour une faible puissance dissipée afin de mesurer le courant à travers un dispositif à puits quantiques placé à proximité en fonction des tensions de grille et de drain à des températures au-dessous de quelques Kelvins. Le cryo-TIA est basé sur un amplificateur opérationnel (OP-AMP) avec un feedback résistif intégré sur puce afin d'obtenir un gain élevé de 10^7 V/A à des températures cryogéniques. L'OP-AMP est constitué de deux étages amplificateurs avec compensation de Miller ne dissipant que 1 µW, résultant en une puissance dissipée très faible par rapport à la puissance de refroidissement disponible d'un réfrigérateur à dilution ^{3}He/^{4}He de l'ordre de 100 µW à 100 mK. Nous étudions systématiquement la dépendance en température du gain, du produit gain-bande-passante et du bruit entre 0.25 et 300 K, et les mettons en regard des résultats présentés au chapitre 3.

Le TIA est cointégré avec un puit quantique à 4.2 K afin d'en démontrer le bon fonctionnement à basse température. Le dispositif quantique est relié par wire-bond au TIA après avoir positionné les deux puces à quelques mm de distance. Le mode de mesure du puit quantique avec le TIA à température cryogénique est comparé à la configuration habituelle avec le TIA à température ambiante à l'extérieur du cryostat et relié au dispositif quantique par un câble de plus d'un mètre de long. Nous montrons que la configuration cryogénique est moins sujette aux fuites de courant due à la taille réduite des interconnexions et permet d'obtenir une meilleure bande passante. Des futures améliorations sont identifiées afin d'augmenter les performances du TIA cryogénique en terme de bande passante et de bruit tout en conservant la même consommation électrique de quelques µW.

Finalement, pour démontrer la possibilité d'une intégration complète, nous validons l'utilisation du cryo-TIA jusqu'à 10 mK en caractérisant le comportement basse-fréquence d'un dispositif quantique réalisé dans le même substrat de silicium suivant les règles de conception de la technologie FD-SOI 28 nm. Pour aller plus loin, le cryo-TIA et le dispositif quantique sont incorporés dans un circuit plus conséquent composé d'une excitation de tension GHz réalisée avec des circuits présentés au chapitre 3 afin d'observer le comportement à hautes fréquences du dispositif quantique dans un régime de pompage de charges. Nous discutons de l'impact de la puissance dissipée par l'électronique haute-fréquence sur le dispositif quantique à proximité. Cette réalisation sur une même puce de dispositifs classiques et quantiques démontre le potentiel des technologies CMOS FD-SOI pour les applications d'informatique quantique jusqu'à 10 mK.

Outre leur utilisation pour la caractérisation des diagrammes de stabilité des puits quantiques, les TIA sont aussi utilisés afin d'effectuer une lecture de l'état de spin en une seule mesure, appelé « single-shot ». En mesurant le courant d'un puit quantique proche d'une transition de charge, il est possible d'obtenir une réponse rapide et claire à un changement capacitif se produisant dans l'environnement immédiat dû à par ex. un déplacement de charges dans des puits quantiques voisins. La lecture de l'état de spin est une sortie binaire, et non une sortie analogique continue des amplificateurs habituels. Avec la disponibilité de l'électronique à basse température, une voie vers la lecture d'un grand nombre de qubits consiste à améliorer la spécificité de l'amplificateur à transimpedance afin de n'en obtenir qu'une sortie binaire. Cela permet de réduire la consommation d'énergie et d'augmenter la vitesse de lecture comme déjà présenté[12].

Suivant cette idée d'électronique sur mesure, dans le chapitre suivant, nous abordons la lecture

radiofréquence basée sur la mesure de la capacité de grille des dispositifs quantiques en proposant une nouvelle architecture de lecture.

Bibliography

[1] W. G. van der Wiel, S. De Franceschi, J. M. Elzerman, T. Fujisawa, S. Tarucha, and L. P. Kouwenhoven. **Electron transport through double quantum dots**. *Reviews of Modern Physics*, **75**(1):1–22, December 2002. Available from: https://link.aps.org/doi/10.1103/RevModPhys.75.1.

[2] Serwan Asaad, Vincent Mourik, Benjamin Joecker, Mark A. I. Johnson, Andrew D. Baczewski, Hannes R. Firgau, Mateusz T. Mądzik, Vivien Schmitt, Jarryd J. Pla, Fay E. Hudson, Kohei M. Itoh, Jeffrey C. McCallum, Andrew S. Dzurak, Arne Laucht, and Andrea Morello. **Coherent electrical control of a single high-spin nucleus in silicon**. *Nature*, **579**(7798):205–209, March 2020. Available from: http://arxiv.org/abs/1906.01086.

[3] R. Zhao, T. Tanttu, K. Y. Tan, B. Hensen, K. W. Chan, J. C. C. Hwang, R. C. C. Leon, C. H. Yang, W. Gilbert, F. E. Hudson, K. M. Itoh, A. A. Kiselev, T. D. Ladd, A. Morello, A. Laucht, and A. S. Dzurak. **Single-spin qubits in isotopically enriched silicon at low magnetic field**. *Nature Communications*, **10**(1):5500, December 2019. Available from: http://www.nature.com/articles/s41467-019-13416-7.

[4] *Variable Gain Low Noise Current Amplifier (FEMTO DPLCA-200)*. Available from: https://www.femto.de/en/products/current-amplifiers/variable-gain-up-to-500-khz-dlpca.html.

[5] *Small Instrumentation Modules: Low noise voltage preamplifiers (SIM910 and SIM911)*. Available from: https://www.thinksrs.com/products/sim910911.html.

[6] J.P.G. van Dijk, E. Kawakami, R.N. Schouten, M. Veldhorst, L.M.K. Vandersypen, M. Babaie, E. Charbon, and F. Sebastiano. **Impact of Classical Control Electronics on Qubit Fidelity**. *Physical Review Applied*, **12**(4):044054, October 2019. Available from: https://link.aps.org/doi/10.1103/PhysRevApplied.12.044054.

[7] J. Pettersson, P. Wahlgren, P. Delsing, D. B. Haviland, T. Claeson, N. Rorsman, and H. Zirath. **Extending the high-frequency limit of a single-electron transistor by on-chip impedance transformation**. *Physical Review B*, **53**(20):R13272–R13274, May 1996. Available from: https://link.aps.org/doi/10.1103/PhysRevB.53.R13272.

[8] M. L. V. Tagliaferri, A. Crippa, S. Cocco, M. De Michielis, M. Fanciulli, G. Ferrari, and E. Prati. **Modular Printed Circuit Boards for Broadband Characterization of Nanoelectronic Quantum Devices**. *IEEE Transactions on Instrumentation and Measurement*, **65**(8):1827–1835, August 2016.

[9] H. Birk, K. Oostveen, and C. Schönenberger. **Preamplifier for electric-current noise measurements at low temperatures**. *Review of Scientific Instruments*, **67**(8):2977–2980, August 1996. Available from: https://aip.scitation.org/doi/10.1063/1.1147083.

[10] M J Curry, T D England, N C Bishop, G Ten-Eyck, J R Wendt, T Pluym, M P Lilly, S M Carr, and M S Carroll. **Cryogenic preamplification of a single-electron-transistor using a silicon-germanium heterojunction-bipolar-transistor**. *Appl. Phys. Lett.*, page 6, 2015.

[11] M. J. Curry, M. Rudolph, T. D. England, A. M. Mounce, R. M. Jock, C. Bureau-Oxton, P. Harvey-Collard, P. A. Sharma, J. M. Anderson, D. M. Campbell, J. R. Wendt, D. R. Ward, S. M. Carr, M. P. Lilly, and M. S. Carroll. **Single-Shot Readout Performance of Two Heterojunction-Bipolar-Transistor Amplification Circuits at Millikelvin Temperatures**. *Scientific Reports*, **9**(1):16976, December 2019. Available from: `http://www.nature.com/articles/s41598-019-52868-1`.

[12] Kushal Das, Torsten Lehmann, and Andrew S. Dzurak. **Sub-Nanoampere One-Shot Single Electron Transistor Readout Electrometry Below 10 Kelvin**. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **61**(10):2816–2824, October 2014.

[13] M. G. House, T. Kobayashi, B. Weber, S. J. Hile, T. F. Watson, J. van der Heijden, S. Rogge, and M. Y. Simmons. **Radio frequency measurements of tunnel couplings and singlet–triplet spin states in Si:P quantum dots**. *Nature Communications*, **6**(1):8848, December 2015. Available from: `http://www.nature.com/articles/ncomms9848`.

[14] R. Maurand, X. Jehl, D. Kotekar-Patil, A. Corna, H. Bohuslavskyi, R. Laviéville, L. Hutin, S. Barraud, M. Vinet, M. Sanquer, and S. De Franceschi. **A CMOS silicon spin qubit**. *Nature Communications*, **7**:13575, November 2016. Available from: `https://www.nature.com/articles/ncomms13575`.

[15] L. Le Guevel, G. Billiot, B. Cardoso Paz, M. L. V. Tagliaferri, S. De Franceschi, R. Maurand, M. Cassé, M. Zurita, M. Sanquer, M. Vinet, X. Jehl, A. G. M. Jansen, and G. Pillonnet. **Low-power transimpedance amplifier for cryogenic integration with quantum devices**. *Applied Physics Reviews*, **7**(4):041407, December 2020. Available from: `http://aip.scitation.org/doi/10.1063/5.0007119`.

[16] Loick Le Guevel, Gerard Billiot, Xavier Jehl, Silvano De Franceschi, Marcos Zurita, Yvain Thonnart, Maud Vinet, Marc Sanquer, Romain Maurand, Aloysius G. M. Jansen, and Gael Pillonnet. **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**. In *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, pages 306–308, San Francisco, CA, USA, February 2020. IEEE. Available from: `https://ieeexplore.ieee.org/document/9063090/`.

[17] M. A. Ghanad and C. Dehollain. **A sub 1 dB NF high dynamic range low-input impedance CMOS amplifier**. In *2016 IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pages 1–4, October 2016.

[18] W.C. Black, D.J. Allstot, and R.A. Reed. **A high performance low power CMOS channel filter**. *IEEE Journal of Solid-State Circuits*, **15**(6):929–938, December 1980.

[19] K. Triantopoulos, M. Cassé, S. Barraud, S. Haendler, E. Vincent, M. Vinet, F. Gaillard, and G. Ghibaudo. **Self-Heating Effect in FDSOI Transistors Down to Cryogenic Operation at 4.2 K**. *IEEE Transactions on Electron Devices*, pages 1–8, 2019.

[20] T. Poiroux, O. Rozeau, P. Scheer, S. Martinie, M. Jaud, M. Minondo, A. Juge, J. C. Barbé, and M. Vinet. **Leti-UTSOI2.1: A Compact Model for UTBB-FDSOI Technologies—Part II: DC and AC Model Description**. *IEEE Transactions on Electron Devices*, **62**(9):2760–2768, September 2015.

[21] Bruna Cardoso Paz, Loïck Le Guevel, Mikaël Cassé, Gérard Billiot, Gaël Pillonnet, Aloysius Jansen, Sebastien Haendler, Andre Juge, Emmanuel Vincent, Philippe Galy, Gérard Ghibaudo, Maud Vinet, Silvano de Franceschi, Tristan Meunier, and Fred Gaillard. **Integrated Variability Measurements of 28 nm FDSOI MOSFETs down to 4.2 K for Cryogenic CMOS Applications**. In *2020 IEEE 33rd International Conference on Microelectronic Test Structures (ICMTS)*, pages 1–5, May 2020.

[22] C. R. Viswanathan. **Low-temperature characterization of CMOS devices**. In *International Symposium on VLSI Technology, Systems and Applications,*, pages 207–212, May 1989.

[23] L. Pirro, O. Zimmerhackl, A. Zaka, L. Miiller-Meskamp, R. Nelluri, T. Hermann, I. Cortes-Mayol, A. Huschka, M. Otto, E. Nowak, A. Mittal, and J. Hoentschel. **RTN and LFN Noise Performance in Advanced FDSOI Technology**. In *2018 48th European Solid-State Device Research Conference (ESSDERC)*, pages 254–257, September 2018.

[24] D. M. Binkley, C. E. Hopper, B. J. Blalock, M. M. Mojarradi, J. D. Cressler, and L. K. Yong. **Noise performance of 0.35-µm SOI CMOS devices and micropower preamplifier from 77-400 K**. In *2004 IEEE Aerospace Conference Proceedings (IEEE Cat. No.04TH8720)*, **4**, pages 2495–2506 Vol.4, March 2004.

[25] H. Bohuslavskyi, S. Barraud, M. Cassé, V. Barrai, B. Bertrand, L. Hutin, F. Arnaud, P. Galy, M. Sanquer, S. De Franceschi, and M. Vinet. **28nm Fully-depleted SOI technology: Cryogenic control electronics for quantum computing**. In *2017 Silicon Nanoelectronics Workshop (SNW)*, pages 143–144, June 2017.

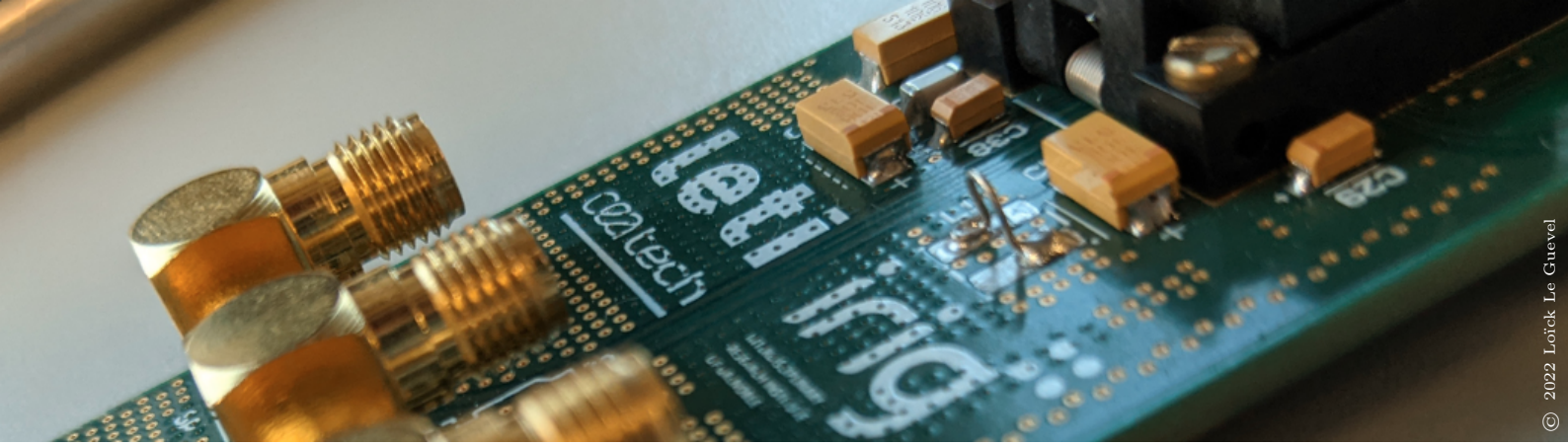
[26] M. Hofheinz, X. Jehl, M. Sanquer, G. Molas, M. Vinet, and S. Deleonibus. **Individual charge traps in silicon nanowires: Measurements of location, spin and occupation number by Coulomb blockade spectroscopy**. *The European Physical Journal B*, **54**(3):299–307, December 2006. Available from: `http://link.springer.com/10.1140/epjb/e2006-00452-x`.

[27] *LTC6240/LTC6241/LTC6242, Single/Dual/Quad 18MHz, Low Noise, Rail-to-Rail Output, CMOS Op Amps*. Linear Technology.

[28] I. Bashir, E. Blokhina, A. Esmailiyan, D. Leipold, M. Asker, E. Koskin, P. Giounanlis, H. Wang, D. Andrade-Miceli, A. Sokolov, A. Koziol, T. Siriburanon, and R. B. Staszewski. **A Single-Electron Injection Device for CMOS Charge Qubits Implemented in 22 nm FD-SOI**. *IEEE Solid-State Circuits Letters*, pages 1–1, 2020.

[29] P. Clapera, S. Ray, X. Jehl, M. Sanquer, A. Valentian, and S. Barraud. **Design and Cryogenic Operation of a Hybrid Quantum-CMOS Circuit**. *Physical Review Applied*, **4**(4):044009, October 2015. Available from: `https://link.aps.org/doi/10.1103/PhysRevApplied.4.044009`.

[30] Bernd Kaestner and Vyacheslavs Kashcheyevs. **Non-adiabatic quantized charge pumping with tunable-barrier quantum dots: a review of current progress**. *Reports on Progress in Physics*, **78**(10):103901, September 2015. Available from: `https://doi.org/10.1088%2F0034-4885%2F78%2F10%2F103901`.



CHAPTER 5
Capacitive read-out of quantum dots

I Introduction

The electronic quantum states within a quantum-dot structure can be advantageously probed with a gate-only measurement. Gate-based measurements rely on the availability of a nearby metallic electrode capacitively coupled to the quantum dot. The electrode is often placed above the dot, reminiscent of the transistor gate in conventional CMOS transistors, hence its name. This type of measurement is advantageous as it is a single-port measurement and is generally non-invasive compared to transport measurements (as no additional contacts in the substrate are needed, see Chapter 4). The out-of-plane probe relaxes the constraints on the quantum-dot array topology leading to a higher qubit-qubit connectivity and better regularity.

a) The concept of quantum capacitance in quantum-dot devices

For a capacitor made of two metallic electrodes, the accumulated charges on each plate is considered continuous as a large number of charges are available. A change in the potential V across the capacitor results in a continuous rearrangement of charges q in the electrodes following the well-known law:

$$q = C \times V \tag{5.1}$$

In gate-based measurement, one electrode is the quantum dot holding a low number of electrons, and most importantly a discrete number of them. A change of the gate-potential of a quantum dot does not always lead to a change of the number of charges due to Coulomb blockade. In such a regime, the potential across the capacitor changes but the number of charges in the dot remains the same, requiring a change of capacitance C_q to accommodate equation 5.1. In the most general case, C_q is named the parametric capacitance and appears in capacitors in which one the two plates is made of a material with a low electron density. In some cases, C_q is named the quantum capacitance.

Parametric capacitance

The parametric capacitance is a correction to the metal-like capacitance of a system resulting from a low density of states in part of the capacitor. It appears in parallel to the geometric capacitance C_{geom} and can be negative (keeping in mind that $C_q \ll C_{geom}$).

The parametric capacitance in the adiabatic regime at zero temperature is proportional to the energy-band curvature with respect to the gate voltage:

$$C_p = -\frac{\partial^2 E}{\partial V_g^2} \tag{5.2}$$

with $E(V_g)$ the state energy as a function of the applied DC gate voltage V_g. Given two states with very distinct curvatures, it is possible to distinguish the two states by measuring the capacitance, thus resulting in a quantum state measurement. In the next section, we compute and evaluate equation 5.2 in the context of a double quantum-dot system, widely used to implement qubits.

A charge qubit: Double quantum dot

A double quantum-dot system consists in two dots, close to each other such that electrons can tunnel back and forth, depending on the electrostatic environment. Each quantum dot is capacitively coupled to a gate, such that a positive (respectively negative) gate voltage attracts (resp. repels) electrons in (resp. out) the dot. We note V_g the differential voltage $V_R - V_L$ that controls the energy detuning $\epsilon = e\alpha V_g$ between the two quantum dots with e the electron charge, and α the gate-level arm depending on all capacitance. The gate-level arm α in V/V is defined by the ratio of the capacitance between the dot and the gate and the total dot capacitance. By definition α is equal or lower than 1, typical values of the level-arm are 0.6 for MOS-based devices with nanometric oxides and 0.01-0.1 for thicker oxides in III-V materials. The two quantum dots are tunnel-coupled with a typical energy t, named the tunnel coupling at the interdot transition. Energies, usually in µeV, in quantum mechanics are often given in frequency units by dividing the energy by planck's constant $h = 4.14\,\mu\text{eV/GHz}$.

With one moving electron in our system, the two lowest-energy levels E_- and E_+ of a double quantum dot represented in Figure 5.1a are given by:

$$E_\pm(V_g) = \pm\frac{1}{2}\sqrt{(e\alpha V_g)^2 + (2t)^2} \tag{5.3}$$

Injecting equation 5.3 into 5.2 leads to the following capacitive contribution as a function of the gate voltage:

$$C_p^\pm(V_g) = \mp\frac{(e\alpha)^2}{4}\frac{(2t)^2}{\left((e\alpha V_g)^2 + (2t)^2\right)^{3/2}} \tag{5.4}$$

The parametric capacitance is maximal at zero detuning $V_g = 0$ (equivalently $\epsilon = 0$) where the curvature is maximal (see Figure 5.1a), reaching an absolute value of $(e\alpha)^2/8t$.

We consider typical values for a silicon quantum dot with $\alpha = 0.5$ and $t = 1\,\text{GHz}$, such that the parametric capacitance of a charge in a quantum dot near the interdot transition has a typical amplitude of 1.2 fF. The parametric capacitance of more complicated system such as the spin-qubit in a double quantum dot at non-zero magnetic field tends to exhibit a smaller parametric capacitance by a factor of 10-100, down to a few aF due to weaker band curvatures.

In real systems, the extrinsic capacitance from strain capacitance, and interconnections, largely overpowers the intrinsic geometric capacitance (and parametric capacitance), by orders of magnitude. The fast µs detection of the tiny parametric capacitance diluted in parasitics requires dynamic ranges up to 10^6 in current systems. The low value of the parametric capacitance and required high-speed measurement for quantum computation are challenging and require carefully-designed experimental setups to reach high-fidelity read-out of the dot quantum state.

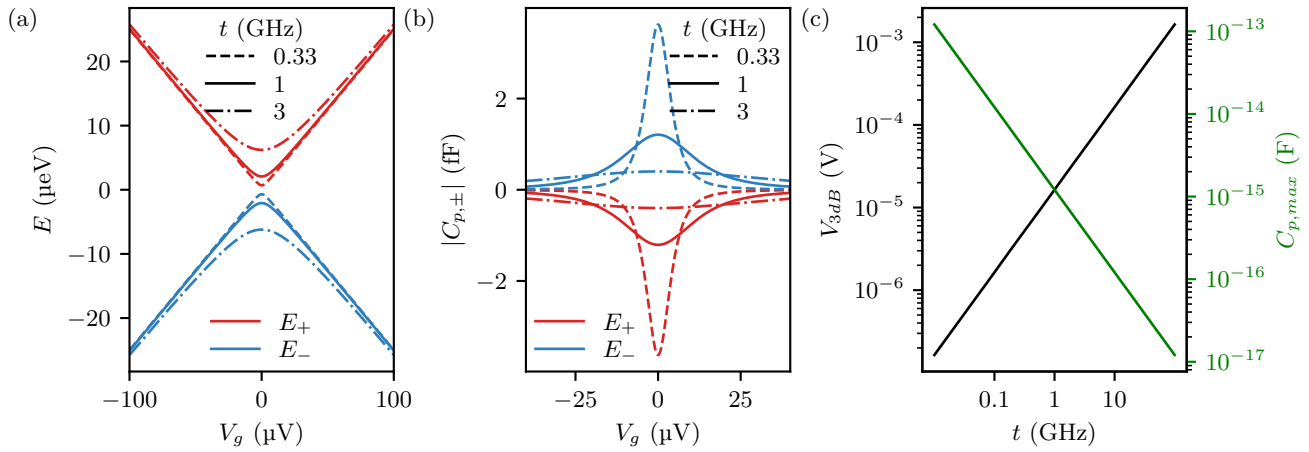


Figure 5.1 | Parametric capacitance of a double quantum dot.

(a) Two lowest energy bands of the double quantum dot as a function of the gate voltage V_g for three different tunnel-coupling values of 0.33, 1, and 3 GHz. (b) Parametric capacitance $C_{p,\pm}$ as a function of V_g for the ground (blue) and excited state (red) for different values of tunnel coupling. (c) Gate voltage amplitude at which the measured capacitance is reduced due to compression effects as a function of tunneling coupling (left y-axis). Reduction of the measured parametric capacitance with stronger tunnel coupling (right y-axis). A tunnel-coupling sweet-spot exists for optimal read-out that depends on the electronics specifications.

To improve the read-out speed, the parametric capacitance C_q needs to be as large as possible. The decrease of the tunneling coupling t increases C_q but also reduces the required range in gate-voltage V_g to avoid compression effects (see Figure 5.1c). Moreover, the adiabatic regime hypothesis used to derive equation 5.2 breaks if the tunneling rate $\propto t$ is slower than the excitation frequency. Lowering the excitation frequency slows the detection hence t can't be decreased too much. C_q can be optimized by adjusting t and the gate level-arm α but little improvement is expected due to the interplay of the different read-out components defining the fidelity.

The treatment of the read-out by considering the adiabatic regime allows to get a feeling of how quantum states can be measured. We've seen that the strong hypothesis leads to a constrained problem in which the optimization of the total read-out signal-to-noise ratio relies on the electronics. Fortunately, the notion of parametric capacitance in the adiabatic regime is the tip of the iceberg of quantum-state detection. More complicated analysis and modeling of the read-out process considering more "quantumness" (with e.g. a quantum resonator) offers new trade-offs in the optimization of the read-out.

In the following sections, we always consider the adiabatic regime to derive circuit specifications for its simplicity.

b) Probing the quantum capacitance

Measuring the tiny capacitance associated with the electronic quantum states is a challenge as it is diluted in the much larger geometric capacitance. Moreover, the experimental conditions related to the cryogenic temperatures required to witness quantum effects leads to an even more challenging task. The capacitance has to be probed with μV-level signals conveyed through often-lossy meter-long wiring from

room to cryogenic temperature. A widely-anchored solution across all solid-state qubit platforms has been to measure the frequency-shift of a radio-frequency/microwave resonator coupled to the quantum device.

Coupling a passive resonator to a quantum device

The resonance frequency $f_r = 1/2\pi\sqrt{LC_p}$ of an LC-tank depends on the tank capacitance C_p parallel to the inductance L. Adding the quantum device in parallel gives the total shunt capacitance $C_p + C_{DUT}$ with C_{DUT} the gate-capacitance of the quantum device accounting for both the geometric capacitance and quantum capacitance. Monitoring the LC-tank resonance frequency deviation from the bare resonator frequency gives access to the quantum capacitance of a DUT. More generally, the resonator frequency shift resulting from the occupation of a quantum state is a general feature of a quantum system coupled to a resonator and is a valid picture, even in the strongly-coupled or diabatic regime.

Experimentally, a shift δf of the bare resonance frequency f_r can be accurately detected by extracting the signal of the phase change at f_r with homodyne detection. Around f_r, the phase signal θ is linked to the frequency shift δf by the measurement-setup quality factor Q. The type of signal used in the homodyne detection depends on the implementation of the measurement system. In the next section, we detail the reflectometry implementation that measures the resonator phase shift by probing the reflection coefficient of a wave against the tank.

c) The well-established reflectometry

The fast detection of the parametric capacitance starts to be routinely done with the well-established reflectometry method using harmonic propagating waves. In this implementation, a microwave signal near the resonant frequency f_r that can range from ~ 0.3 up to $\sim 3\,\mathrm{GHz}$ is sent to the resonator on which it gets reflected. The frequency range of f_r combined with meter-long wiring from cryogenic to room-temperature requires to study the system in the microwave framework with propagating waves in $50\,\Omega$-matched lines. The phase shift between the incoming and reflected waves becomes an image of the resonator phase shift δf and allows the detection of the quantum states. Effectively, reflectometry measures the phase of the S_{11} scattering parameters.

The typical experimental setup to measure quantum states is shown in Figure 5.2a. The quantum device is anchored at the lowest-temperature stage and linked to room-temperature wiring with meter-long cables traversing each temperature stage within the dilution fridge with distributed attenuation to reduce thermal noise along the way. At the low-temperature stage, the reflected wave is isolated from the incoming wave with a directional coupler, then amplified with cryogenic LNAs. At room-temperature, a lock-in amplifier or any equivalent measurement unit demodulates the reflected wave using the excitation wave to recover the phase shift.

Alternatively, S_{12} can be measured in transmission instead of S_{11} in reflection, often with GHz-frequency resonators, reminiscent of the superconducting qubit read-out measurements. This solution is often implemented with custom-made inductors, such as superconducting inductors, that can be easily provided with a built-in capacitive coupling to the transmission line. In this section, we focus our analysis on the reflectometry measurement, knowing that the same reasoning can be applied to the S_{12} measurement.

The measured scattering parameter S_{11} is given by the following formula:

$$S_{11} = \frac{Z - Z_0}{Z + Z_0} \tag{5.5}$$

with $Z_0 = 50\,\Omega$ the characteristic line impedance and Z the RLC tank impedance. The results of this complex transformation is shown in Figure 5.2b with the measured amplitude and phase of S_{11}. S_{11} dips

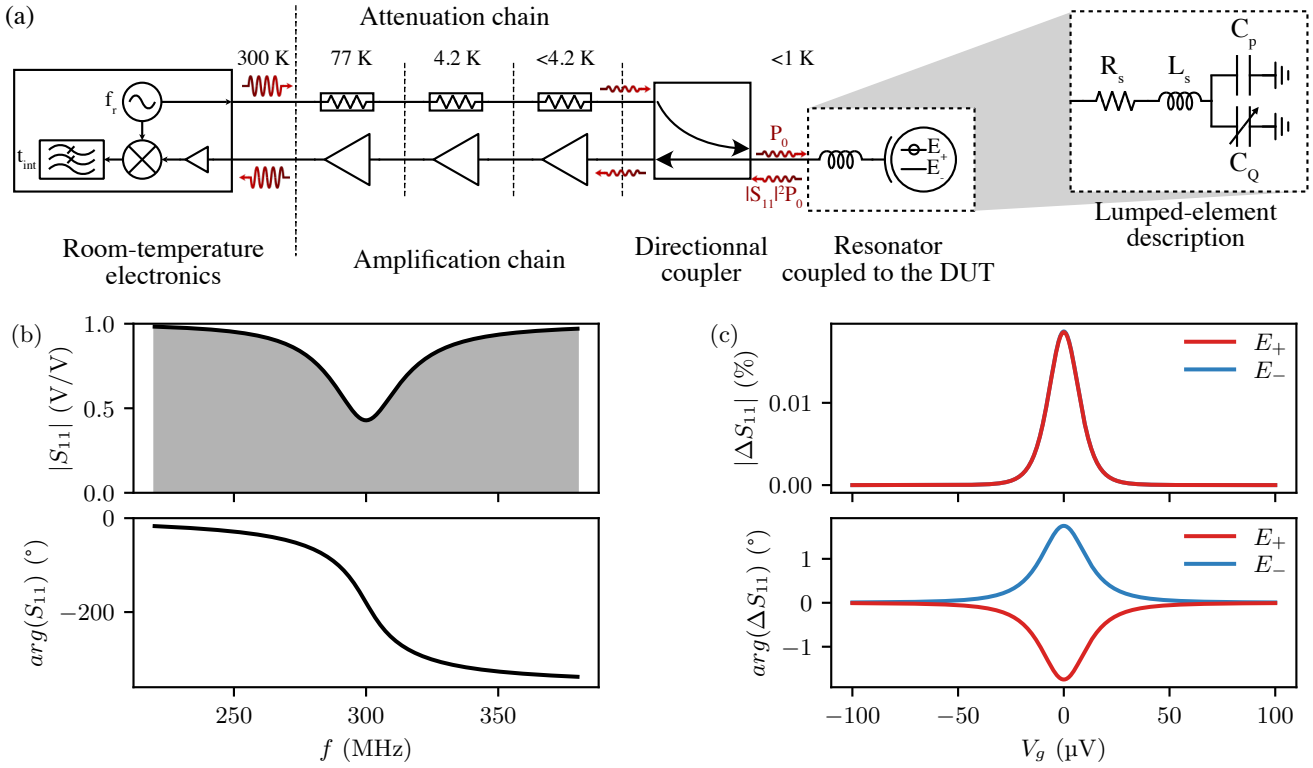


Figure 5.2 | Reflectometry for the detection of parametric capacitance.

(a) Reflectometry setup to probe the gate capacitance of a quantum device at sub-K temperature in a cryogenic refrigerator. The room-temperature electronics sends an harmonic signal at f_r inside the fridge. The incoming wave is successively attenuated at different temperature stages to reduce the signal floor noise from thermal radiations. The wave passes through a directional coupler and is reflected on the series combination of the quantum device and a passive inductor depending on the scattering parameter S_{11}. The reflected wave passes through the lossless port of the directional coupler, is amplified with cryogenic low-noise amplifiers and further amplified at room-temperature. The lock-in amplifier performs the homodyne detection at f_r, the resonator frequency, to obtain the phase difference between the outgoing and incoming signal that depends on the device parametric capacitance. (b) Scattering parameter S_{11} of the series combination of the quantum device gate capacitance and passive inductor as a function of frequency. In this configuration, the signal at f_r is attenuated by the resonance while all other frequency components remain intact, increasing the impact of out-of-resonance components (e.g. noise and parasitics tones). (c) Amplitude and phase difference between the incoming and reflected waves from homodyne detection for a varying gate voltage crossing the interdot region. A clear signature appears in the phase signal at the interdot transition and its value depends on the quantum state of the quantum device ($|0\rangle$ in red and $|1\rangle$ in blue).

at the resonant frequency f_r at which Z is purely real and equal to R_s. If $R_s = Z_0$, S_{11} is zero and no signal is returned due to a perfect load matching between the signal generator and resonator. Generally, S_{11} is of the order of 0.5 to get a decent reflected signal.

An example of the scattering parameter variation $\Delta S_{11} = S_{11}(\epsilon) - S_{11}(0)$ as a function of the device gate biasing V_g near an inter-dot transition is shown in Figure 5.2c. The measured phase after demodulation $\arg(\Delta S_{11})$ is highly sensitive to the operation frequency f_r, itself sensitive to a change in the quantum capacitance $C_p^{\pm}$. The measured change of the phase signal $\Delta\phi$ at fixed frequency f_r becomes an image of the small capacitance variation due to the parametric capacitance $C_p^{\pm} \ll C_{DUT}$ according to:

$$\Delta\phi = QC_p^{\pm}/C_{DUT} \quad \text{with} \quad C_{DUT} \gg C_P^{\pm} \tag{5.6}$$

Equation 5.4 and 5.6 allow to compute the expected phase signal given the circuit and quantum device properties for a charge-like device.

Spin qubit have a more complicated energy diagram with the splitting of singlet and triplet states with magnetic field lifting the degeneracy. Due to smaller difference in energy curvature between a singlet and triplet state, the resulting quantum capacitance is significantly lower by a factor of about 10-100 compared to charge qubits.

We revisited the reflectometry measurement setup to take advantage of the availability of cryogenic electronics. We proposed the impedancemetry measurement setup, replacing the measurement of the scattering parameters S by the resonator impedance Z by adding buffering circuits at cryogenic temperatures as presented next.

II A different approach: the impedancemetry

In this section, we introduce a different technique to probe the resonance of the LC tank coupled to the quantum device that simplifies the system analysis by getting rid of propagating waves. The LC tank response is no longer probed via the scattering parameters, subject to propagating waves physics, but instead via the impedance or admittance parameters Z or $Y = 1/Z$, giving it the name of impedancemetry. The use of impedance parameters instead of the scattering parameters simplifies the system-level analysis as no $50\,\Omega$ impedance matching, reflections, or standing-wave have to be taken into account. However, propagating waves can only be neglected if all wiring has a length l below the signal typical wavelength λ by a few orders of magnitude (typically $l \leq \lambda/20$). For typical frequencies of a few 100's MHz to 1 GHz, the wiring length has to be below $l \leq \lambda/20 = 1.5 - 15\,\text{cm}$ (assuming $c_{wave} = c_{light}$). The suppression of propagating-wave analysis comes at the price of a strong reduction in wire length which with the LC tank at dilution-fridge temperatures $< 1\,\text{K}$ imposes the use of cryogenic electronics. With the acquired knowledge from all previous chapters, we tackled this challenge in the next section.

a) Description

As for reflectometry, the gate of the quantum device under test (DUT) is connected to an inductor to form an LC tank (see Figure 5.3a). The impedance Z (respectively admittance Y) of the LC tank can be probed by exciting the tank with a current I_{in} (resp. voltage V_{in}) and measuring the voltage $V_{out} = ZI_{in}$ (resp. current $I_{out} = YV_{in}$). In this chapter, we focus on the impedance measurement.

Impedancemetry uses currents to excite and probe the resonator via the impedance parameters without the need of bulky coupling elements. The incoming signal V_{in} at the resonant frequency f_r, generated at room temperature, is converted in a current $I_{in} = g_m V_{in}$ with a voltage-controlled current source of transimpedance g_m at the base-temperature stage. The input current I_{in} creates a voltage $V_{out} = ZI_{in}$

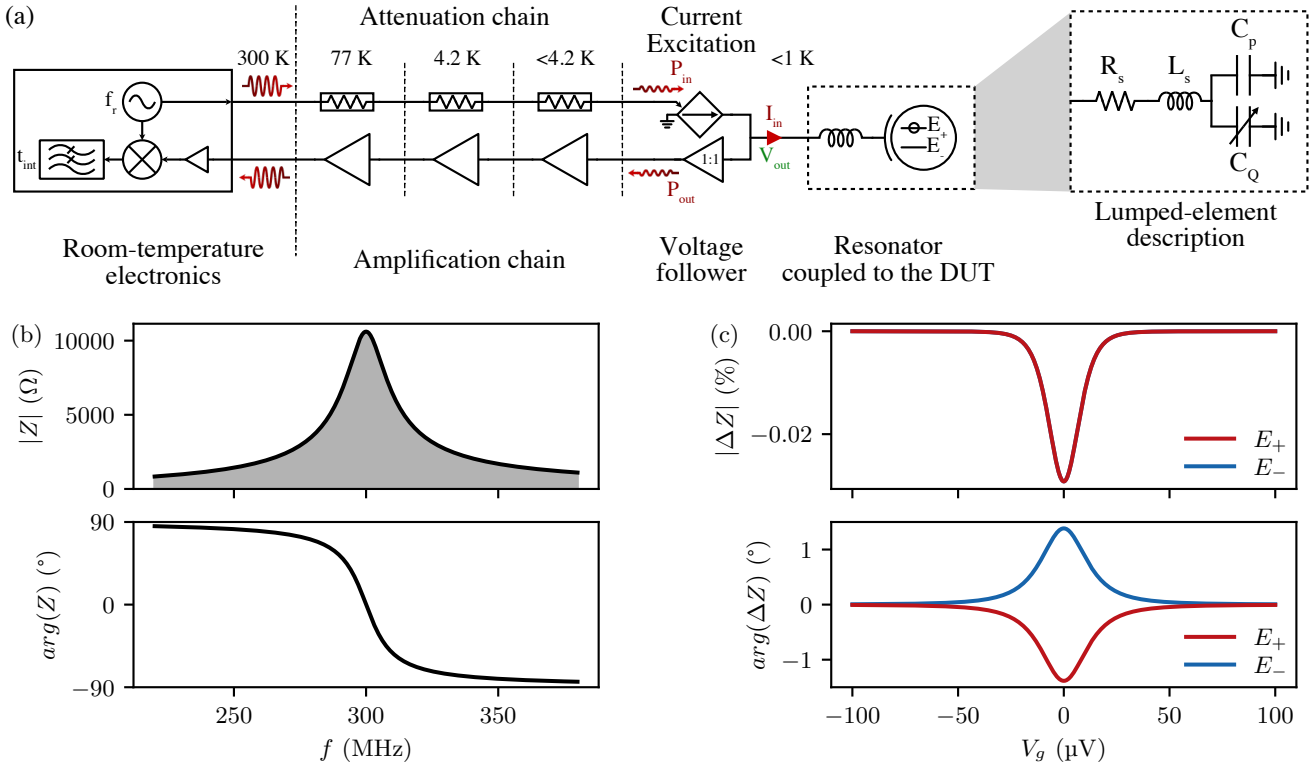


Figure 5.3 | Impedancemetry for the detection of parametric capacitance.

(a) Impedancemetry experimental setup to probe the gate capacitance of a quantum device at sub-K temperature in a cryogenic refrigerator. The room-temperature electronics sends an harmonic signal at f_r inside the fridge. The incoming wave is successively attenuated at different temperature stage to reduce the signal floor noise. The wave is converted to a current thanks to a voltage-controlled current source, exciting the nearby series combination of the quantum device and a passive inductor. The returned signals that only depends on the impedance Z of the LC tank thanks to the short interconnection length is buffered by a voltage-follower and is amplified with cryogenic low-noise amplifier and further amplified at room-temperature. The lock-in amplifier performs the homodyne detection at f_r, the resonator frequency, to obtain the phase difference between the outgoing and incoming signal that depends on the device parametric capacitance. (b) Impedance $Z = V_{out}/I_{in}$ of the series combination of the quantum device gate capacitance and passive inductor as a function of frequency. The input noise from the excitation current source is filtered by the LC tank impedance as illustrated by the grey area in the amplitude response. (c) Amplitude and phase difference of the measured impedance Z from homodyne detection for a varying gate voltage crossing the interdot region. A clear signature appears in the phase signal at the interdot and its value depends on the quantum state of the quantum device ($|0\rangle$ in red and $|1\rangle$ in blue).

through the tank impedance Z that carries the information about the DUT capacitance. V_{out} is conveyed to a low-power unity-gain amplifier (follower) placed nearby the DUT to reduce parasitic capacitance C_{DUT}. The main amplification is placed at a higher temperature (typically 4.2 K) to benefit from higher cooling power.

b) Advantages & challenges of impedancemetry

Impedancemetry has the advantage over reflectometry that the 50 Ω impedance matching plays no role in the optimization of the resonant circuit depending on the inductor and the parasitic capacitors. However, the cryogenic circuitry required by impedancemetry generates extra noise compared to reflectometry, which needs to be minimized. The impedance of the resonator naturally filters out-of-resonance components such as noise and spurious tones as illustrated with the grey area of the impedance $|Z|$ in Figure 5.3a compared to the bigger grey area of the scattering parameter $|S_{11}|$ in the reflectometry measurement in Figure 5.2a. In the perspective of quantum computing involving a qubit matrix, V_{in} could contain a comb of excitation frequencies to excite a set of frequency-selective resonators.

In the case of impedancemetry, without the need of directional couplers, the footprint of the read-out circuitry is reduced. Using modern CMOS technologies with sub-100nm nodes, the additional circuitry of current sources and followers easily fits on a chip with size comparable to the hundreds of qubits chip ($<$mm^2) such that the total footprint is limited by the size of passive µH inductance occupying a few mm^2. Despite the gain in area by switching to the impedancemetry, the resulting area remains orders of magnitude bigger than quantum devices and remains an issue for scaling up the qubit number in quantum processors.

III Shrinking the inductors to improve scaling of the read-out with the qubit number

In perspective of a large number of qubits, frequency-multiplexing is often an important argument to justify the scalabilty of a method. A general estimate for frequency-multiplexing considers a frequency separation of 10 MHz to reach 1 µs read-out. Scaling will always achieve the point at which the available frequency spectrum is filled, crowded by all read-out harmonic tones. At this particular point, the read-out circuitry will have to be duplicated or the read-out will have to be sequentially performed. The latter solution is the easiest method but would limit the speed of execution of quantum algorithms that require read-out at mid-operation such as error correction codes. Likely both solutions will have to be co-jointly used but we consider the case where we duplicate the read-out circuitry for scaling purposes.

The relatively low operating frequency below 1 GHz of the very-first gate-capacitance measurements of quantum devices required the use of large discrete passive inductors with no ferrite materials to sustain the magnetic field required for spin qubit operation. The inductor size in typical gate-capacitance measurement setups of $\sim$ 1 mm^2 severely hinders the scalability to many qubits. More recently, superconducting resonators at a few GHz, smaller than their coil-like counterparts and with higher quality factors, have been used, similarly to the ones used for superconducting qubits. The integration of superconducting elements with CMOS devices requires new fabrication methods but leads to much denser inductance values, improving scalability.

More innovative solutions could also be used such as Josephson junction chains or active inductors to further improve inductance density. We chose to explore the latter, leveraging the availability of cryogenic electronics to emulate the inductance behavior, known as active inductors.

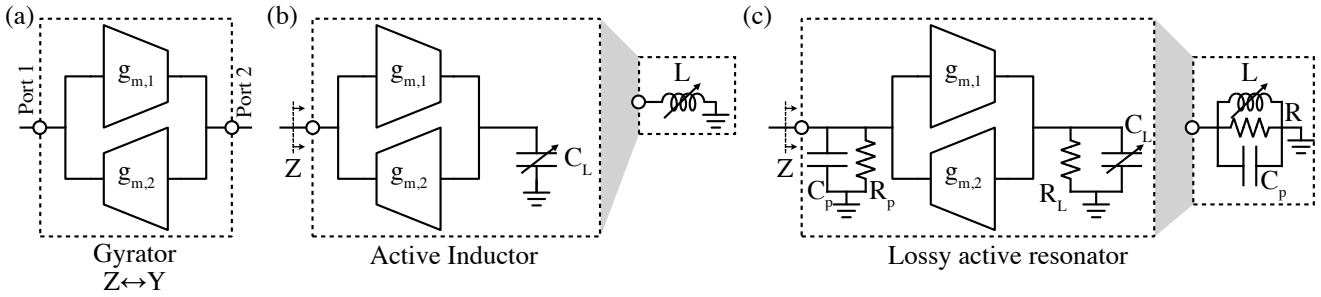


Figure 5.4 | Active inductor and resonator

(a) Gyrator made of two head-to-tail transconductance stages. Port 1 and 2 are symmetrical. Any impedance on one port is perceived as admittance on the other port modulo the product of transconductance $g_{m,1}g_{m,2}$. (b) Active inductor made of a gyrator coupled to a capacitor C_L. The inductance value $L = C_L/g_{m,1}g_{m,2}$ can be easily made tunable by using varactors or capacitor banks. (c) Lossy active resonator made of an active inductor with losses R_p, R_L and parasitics capacitance C_p.

IV Active inductors for high compacity

Ideal inductors follows the well-known impedance law $Z = V/I = j2\pi fL$ with $j = \sqrt{-1}$ the imaginary number, L the inductance value, and f the operation frequency. If we were able to find a circuit capable of inverting the roles of current I and voltage V, we theoretically would be able to transform an impedance Z to an admittance Y with the same frequency dependence. This trick would allow to transform dense MOM capacitors into compact inductors.

Gyrators

Such a circuit, named a gyrator, is able to transform a capacitance C to an inductance L, and vice versa. The simplest gyrator consists of two head-to-tail transconductance stages g_m and $-g_m$ (shown in Figure 5.4a). When an impedance Z_1 (respectively admittance Y_1) is present at port 1, the admittance Y_2 (resp. impedance Z_2) seen from port 2 is $1/(g_m^2 Z_1)$ (resp. Y_1/g_m^2). The gyrator is symmetrical and the same transformation is valid from port 2 to 1.

Placing a capacitor C_L with the ideal gyrator results in the emulation of an ideal inductor of inductance $L = C_L/g_{m,1}g_{m,2}$ (see Figure 5.4b). The higher capacitance density in recent CMOS technologies results in higher inductance density than passive inductors. Moreover, the inductance value of the active inductor can be made tunable with the use of varactors or capacitor banks (tunable C_L in Figure 5.4b).

Considering a second capacitor C_0 placed in parallel due to e.g. interconnexions transforms the emulated inductor into a lossless active resonator of resonant frequency f_r given by:

$$f_r = \frac{1}{2\pi\sqrt{LC_0}} = \frac{1}{2\pi}\sqrt{\frac{g_{m,1}g_{m,2}}{C_L C_0}} \tag{5.7}$$

Without any losses, the active resonator has infinite quality factor Q. Losses in the gyrator due to non-ideal transistors and passive components set Q as discussed below.

Losses

Gyrators are made of active transistors playing the role of transconductance $g_m = dI_{ds}/dV_{gs}$ with I_{ds} the drain-source current and V_{gs} the gate-source voltage. Transistors have losses represented by the conductance g_{ds} that limits the highest achievable Q factor of the active resonator.

Interestingly, losses in the system can be balanced by injecting energy in the resonator thanks to the used active elements, resulting in higher Q factors. An equivalent point of view is to place a negative resistance (made of active elements) in parallel with the active resonator to virtually reduce losses.

A negative resistance can be realized by leveraging second-order effects $(j\omega)^2 = -\omega^2$ between C_L and a second capacitor C_R to compensate the intrinsic losses in the active resonator. To understand the mechanism better, we introduce the concept of current conveyors of type II (CCII) for a generalization of the transistor behavior. CCII are 3-port devices (as transistors, forgetting about substrate or back-gate): two high-impedance ports Y and Z, and one low-impedance X (see Figure 5.5a). The voltage V_Y at the port Y is copied to the low-impedance port X. The current flowing through X is then output at the high-impedance port Z. Depending on the sign $\pm$ of the output current at Z we define CCII+ and CCII-. CCIIs can be seen as idealized transistors where the Y port would be the gate, X the source, and Z the drain. Combined with passive elements, CCII devices can represent different transistor behavior and topologies such as a positive or negative transconductance or a current mirror/conveyor (see Figure 5.5b).

The gyrator, made of C_L and two transconductances, is shown with two CCII in Figure 5.5c as an equivalent representation of the two head-to-tail transconductance stages shown in Figure 5.4. In this picture, the negative resistance can be understood with a CCII+ placed in a current conveyor/mirror configuration combined with a high-pass filter made of a capacitor C_Q and a resistor $1/g_{m,2}$, representing a common-gate transistor (shown later in the chosen implementation of the active inductance in Figure 5.6b). The added capacitor and resistor allows to inprint part of the current with a phase shift of 90°, transformed into a phase shift of 0° at C_L, converted to a phase shit of 180° by the negative transconductance stage, leading to a negative real term in the impedance as shown in Figure 5.5d.

Tunability

Active inductance architectures incorporating a negative resistor lead to quality factor Q up to a few hundred with independent tuning of the inductance value L and the quality factor Q[1, 2]. Fine calibration of the tunable inductance value using variable capacitors allows a precise in-situ definition of the resonant frequency value, ideal for optimal frequency-multiplexing of large qubit matrices. The tunability of the Q-factor enables different modes of read-out: high-Q gives a precise measurement of quantum capacitance to calibrate qubit matrices, while lower-Q is more suitable for fast read-out during quantum computation.

In this chapter, we implement the proposed impedancemetry as an alternative read-out technique to the reflectometry, leveraging cryogenic electronics to remove propagative waves. We design the on-chip integrated circuit implementing the impedancemetry setup involving a cryogenic CMOS-based active inductor for improved scalability. The read-out circuit is composed of a current source exciting the active LC tank, an amplifier to read the voltage response, and a multiplexed capacitor bank to select different devices under test (DUT). Besides its reduced footprint favoring scalability, this CMOS inductor offers the possibility to tune the characteristic frequency and the quality factor of the resonator, which is instrumental in optimizing measurement sensitivity. We characterize the circuit sensitivity and tunability at 4.2 K demonstrating its capability to measure capacitances as low as 10 aF. By applying our technique to a gate-coupled MOSFET co-integrated on the same chip we reveal typical signatures of quantized electronic states at the onset of conduction.

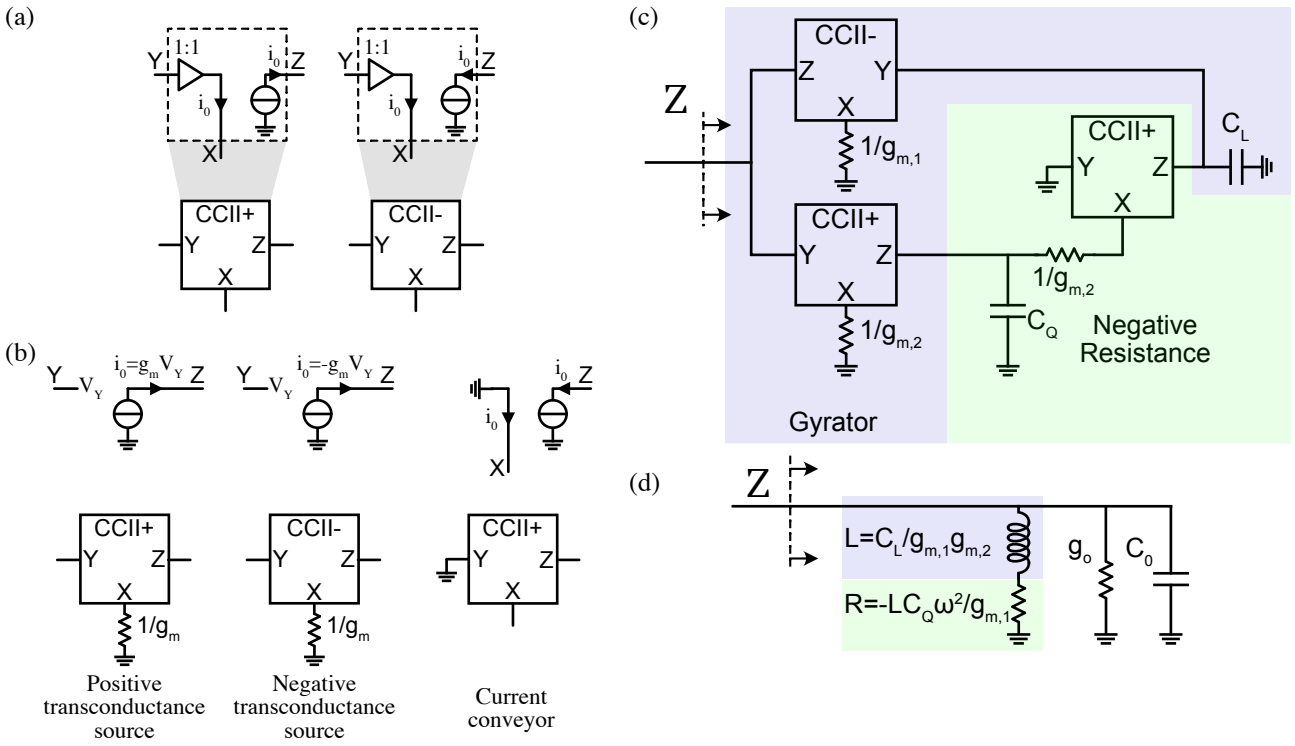


Figure 5.5 | Current conveyors to explain the negative resistance in gyrators

(a) Introduction of the concept of current conveyors of type II (CCII), an idealization of the transistor behavior. The only difference between CCII+ and CCII- is the output current sign. (b) Representation with CCIIs of different transistor topologies such as positive and negative transconductance as well as the current conveyor/mirror. (c) Representation with CCIIs of the active inductor made of the gyrator and capacitor C_L emulating the inductance in purple and the negative resistance made of a current conveyor and high-pass filter highlighted in green. (d) Lumped element representation of the circuit active inductor circuit shown in (c). The negative resistance R appears in series with the inductance to compensate for the intrinsic losses g_0.

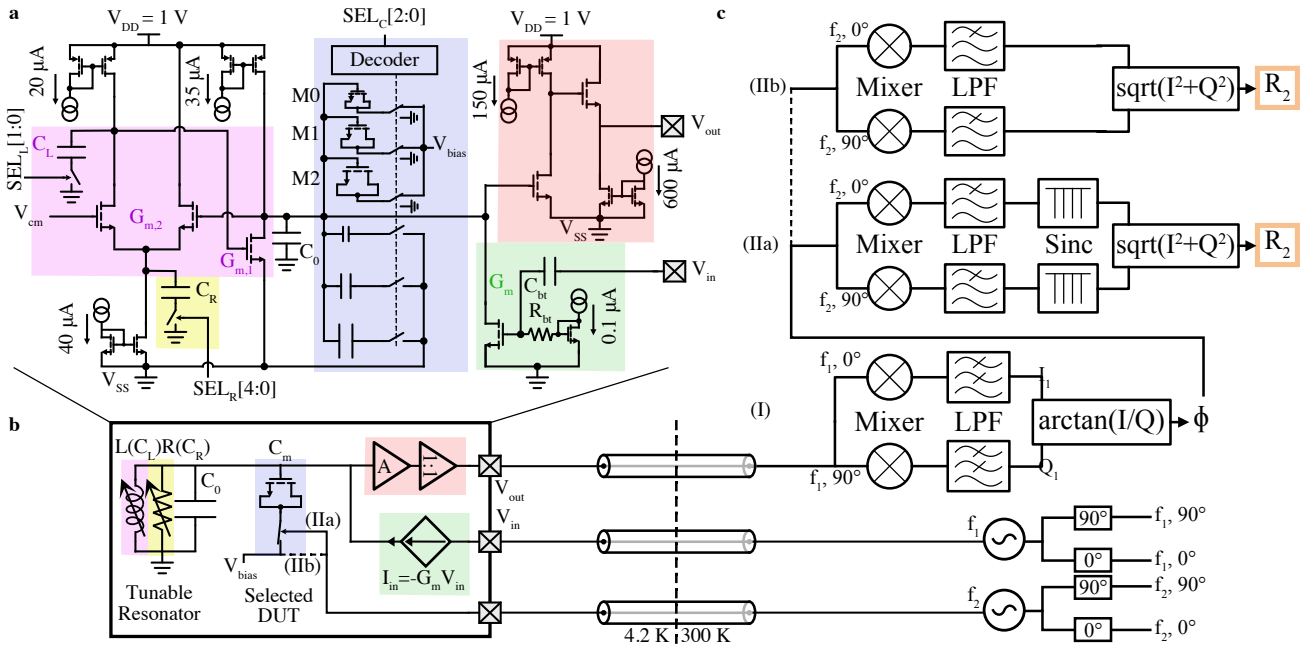


Figure 5.6 | Setup with on-chip electronics.

(a) On-chip circuit implementation of the active inductance (pink), current excitation (green), test capacitor bank (blue), and amplification stage (red). For clarity, the bias MOSFETs operating in DC are drawn of smaller size than MOSFETs in the high-frequency signal chain. (b) Simplified view of the on-chip resonant circuit placed at 4.2 K with tunable resonator, DUT, current excitation, and voltage amplification, linked to room-temperature phase-sensitive electronics via meter-long cables. (c) Room-temperature homodyne detection with single (I) and double (II) demodulation of the circuit output V_{out} and generation of voltage excitation V_{in} at modulation frequencies f_1 (150-200 MHz) and f_2 (1 kHz).

V On-chip impedancemetry as a proof-of-concept

We designed an integrated circuit (IC) to demonstrate the viability of the impedancemetry combined with active inductors to significantly increase the read-out scalability[3]. The designed IC incorporates the amplification chain and quantum device to ease the experimental demonstration without the need of co-integration.

The impedancemetry experiment was integrated on a single chip with multiplexed quantum devices using Low Threshold Voltage (flip-well structure) thin-oxide (GO1) devices within the Fully-Depleted Silicon-On-Insulator (FD-SOI) 28nm CMOS technology. The realized integrated circuit contains the current source, the active inductance with addressable capacitor banks for tunability, the multiplexed DUTs, and the amplification stage (Figure 5.6a, b).

a) Design and simulations

During design, we focused on bringing down the footprint and power consumption of the active inductance being the main original component of our circuit. The complete amplification and current generation was added on-chip to facilitate testing the concept of impedancemetry at 4.2 K. In the absence of high-frequency models of FD-SOI transistors at cryogenic temperatures, we designed the integrated

circuit with accurate room-temperature models supplied by the foundry[4]. The evolution of transistor characteristics towards the lowest temperatures was extrapolated from the temperature variation in foundry models but also from acquired 4.2 K data of single transistors[5, 6].

Active resonator

The active inductance follows a known NMOS-based Karsilayan-Schaumann architecture[7, 8]. The gyrator is made of a single-ended negative transconductance $-G_{m,1}$ and a differential transconductance stage $G_{m,2}$. The gyrator transforms a tunable capacitance C_L into an inductance $L(C_L) = C_L/G_{m,1}G_{m,2}$. An added metal-oxide-metal (MOM) capacitor C_0 of 136 fF parallel to L controls the resonant frequency $f_r = 1/2\pi\sqrt{L(C_L)C_0}$. No dependence in temperature is expected for MOM capacitors[9]. Adding C_0 makes the measuring circuit less sensitive to the DUT-capacitance with the increased tank capacitance but avoids the influence of unknown parasitic capacitances at cryogenic temperatures (e.g. substrate parasitics). Hence, the resonant frequency f_r is set by C_0, C_L, and $G_{m,i=1,2}$. C_L is implemented with one main metal-oxide-metal (MOM) capacitor of 362 fF in parallel with two digitally-controlled binary-weighted MOM capacitors of 68 and 136 fF. At room temperature, the emulated L ranges from 5.3 to 8.4 µH to reach f_r from 128 to 165 MHz. The estimated power consumption of the resonator is 85 µW and corresponds to a footprint of 8.5 µW/qubit assuming a reasonable frequency multiplexing of 10 quantum devices.

Adding a capacitance C_R at the foot of the differential transconductance stage allows to introduce a negative resistance in series with the active inductance, leading to a higher Q-factor with an increased parallel effective resistance $R(C_R, C_L)$. The Q-factor of the active inductance defined as $Q = R(C_L, C_R)\sqrt{L(C_L)/C_0}$ depends on C_R and C_L. By tuning C_L, then C_R, L and Q can be adjusted to any desired value apart from possible instabilities. To tune the Q factor, we choose to cover a wide range of C_R values in steps of 23 fF by selecting 4 binary-weighted MOM capacitors of 23, 46, 92, and 184 fF. From room-temperature simulations, these settings allow to cover a wide range of quality factors Q from 7 to 300, including the unstable states with negative Q.

Current excitation

The voltage-controlled current source exciting the resonator is made of a current mirror combined with an RC bias tee. The bias tee superimposes DC signals from the diode transistor to set the DC operating point of the current source and AC signals from the excitation input V_{in} to generate the AC current I_{in}. The RC filter of the bias tee consists of R_{bt} (polysilicon resistor of 10 MΩ) and C_{bt} (MOM capacitor of 406 fF) to reach a characteristic frequency of 39 kHz. As no large signals V_{in} are required, the current source operates in subthreshold regime with a bias current of only 0.1 µA to minimize its conductance for a negligible impact on the resonator and obtain a desirable low transconductance for nA-current excitation. From foundry models, we get for the current generating transistor a transconductance G_m of 3.4 nA/mV and bandwidth of 3.5 GHz (see Appendix E).

Multiplexing of devices under test

To investigate the active inductance circuit with different DUTs, we added an addressable bank of 6 capacitors. Three MOM capacitors of 2, 4, and 8 fF have the purpose of calibrating the active inductance on known values. Three additional MOSFETs (M0, M1, M2) of width 80 nm and length 28, 60, and 120 nm are used as test-bench for the investigation of quantum properties. The source and drain voltage of the quantum MOSFETs are grounded when unselected and polarized at V_{bias} when selected. The differential transconductance stage of the active inductance copies the DC common-mode voltage V_{cm} to the DUT gate potential, such that the DC gate voltage $V_{gs} = V_{cm} - V_{bias}$ can be varied via V_{bias} (see Figure 5.6).

Amplification and transmission

Once excited by I_{in}, the tank voltage is amplified, then sent through a unit-gain buffer for detection at room temperature via meter-long cable. The amplifier is a common-source N-type single-stage and the 1:1 buffer is a common-drain N-type single stage (see Appendix E). Based on room-temperature simulations, the amplifier has a gain A of 15 dB and a bandwidth of 1.8 GHz for a power consumption of 150 µW. The buffer reaches a bandwidth of 92 MHz for a cable capacitance of 50 pF and a power consumption of 2.4 mW. The net amplification at 165 MHz becomes 8 dB.

b) Noise analysis

Transistor noise translates into transconductance noise that generates perturbing variations into the parameters determining the active inductance. A varying L modulates f_r and generates phase noise in V_{out}. The phase noise spectrum of V_{out} around the carrier frequency f_r extracted from room-temperature steady-state simulations (SST) exhibits a flicker component on time-scale >10 ms, induced by a noisy modulated L. For a typical Q of 81 with sufficiently fast measurements to avoid $1/f$ noise, we get a phase noise of $0.002\,^\circ/\sqrt{\text{Hz}}$ that gives an input-referred noise of $3.2\,\text{aF}/\sqrt{\text{Hz}}$.

VI Cryogenic demonstration

a) Impedancemetry circuit characterization

Without the assistance of low-temperature models, the operating point of the circuit had to be determined experimentally starting from room-temperature settings of bias voltages and currents. The increase in threshold voltage of NMOS (resp. PMOS) transistors at 4.2 K is compensated by applying a back-gate voltage of 1.2 V (resp. −2 V). The optimal cryogenic common-mode voltage $V_{cm} = 0.48$ V was obtained while monitoring the tank impedance via repeated frequency sweeps until a typical resonance behavior up to 200 MHz emerges for the lowest values of C_L and C_R. The gain of the low-temperature amplification stage at f_r is optimized with respect to the current bias of amplifier and buffer. The main results of the impedancemetry with respect to tunability and detection sensitivity are shown in Figure 5.7.

The amplitude and phase of V_{out} using single homodyne detection (I) without any connected DUT are shown in Figure 5.7a for the 4 C_L values from 362 to 566 fF and a few C_R values ranging from 0 to 322 fF depending on C_L. V_{out} at maximal amplitude was kept equal to 1.8 mV by adjusting V_{in} to avoid non-linearities coming from non-linear MOSFETs behavior. The resonance frequency f_r varies by 5.1 % from 189.1 to 199.0 MHz by tuning C_L. The quality factors Q extracted from a linear fit of the phase around f_r are shown in Figure 5.7b. The Q values range from 80 to 250, and can be tuned by a factor > 2 for every C_L by adjusting C_R. These data demonstrate that Q can be tuned almost independently of the resonance frequency with a frequency variation of less than 0.22 % across the entire C_R range (see Figure 5.7b).

For the minimum value of C_L with the highest resonance frequency, we calibrate the capacitance sensitivity of the circuit for each Q by switching on and off the DUT MOM capacitors C_m=2, 3, and 8 fF and using double homodyne detection (II a) (see Figure 5.7c). The capacitance sensitivity α is extracted from a least-square linear fit of the phase change $\Delta\phi = QC_m/C_0 \equiv \alpha C_m$ for a given Q as shown in Figure 5.7d. The sensitivity α increases linearly with Q from 0.76 to $1.9\,^\circ$/fF. From the linear fit in Figure 5.7d, we obtain $C_0 = 137$ fF, in good agreement with the designed value (136 fF). In usual circuits without an additional input capacitance[7], the parasitic capacitance of the MOSFETs determines the

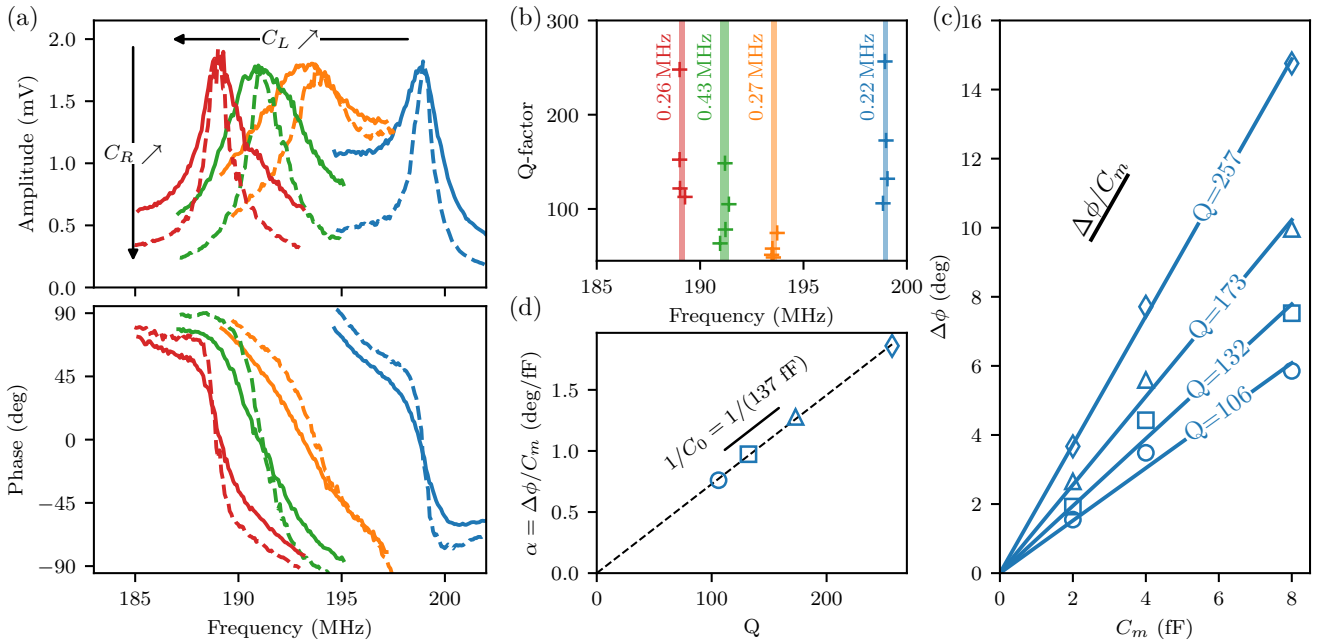


Figure 5.7 | Characterization of the resonant circuit at 4.2 K for capacitance detection.

(a) Amplitude and phase of the demodulated circuit output V_{out} for several active inductance settings. The resonance frequency shifts to lower frequency as the inductance value increases with increasing C_L (different colors). The continuous (low-Q) and dashed line (high-Q) show the signals for different values of C_R. (b), Data points for the resonance frequency f_r when the extracted Q is tuned with C_R. The colored bars of width given by the written maximal deviation indicate the low dispersion of f_r for fixed C_L when varying Q with C_R. (c), Measured phase shift for MOM capacitor C_m of 2, 4, and 8 fF in several Q-factor settings. The capacitance sensitivity $\Delta\phi/C_m$ of the circuit is extracted from the slope with a least square fit at given Q. (d), Capacitance sensitivity extracted from (c) as a function of the Q factor. A least square linear fit of $\Delta\phi/C_m(Q)$ allows to extract the capacitance C_0 parallel with the active inductance.

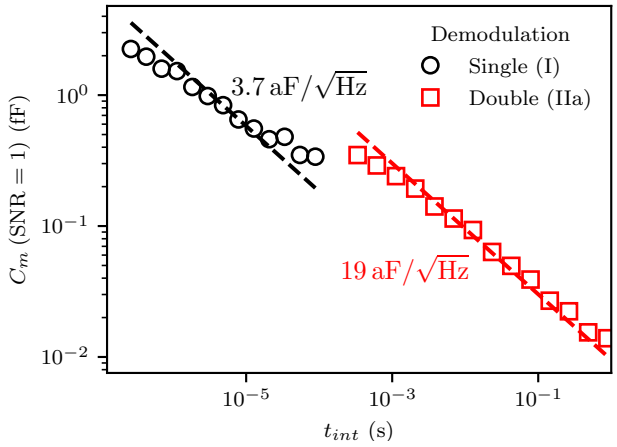


Figure 5.8 | Capacitance resolution of the measurement set-up.

Extrapolated capacitance C_m at signal-to-noise ratio equal to 1 for single (I) (black circles) and double (IIa) (red squares) homodyne detection of the capacitance measurement as a function of the integration time t_{int}. Dashed lines are least-square fits $C_m = a\, t_{int}^{-1/2}$ with $a = \sqrt{0.250}\, S_c$ and S_c the equivalent noise spectral density in aF/$\sqrt{\text{Hz}}$ of the capacitance measurement.

resonance frequency. In future design with accurate cryogenic compact models, this capacitance can be reduced significantly leading to higher resonance frequency and improved sensitivity.

From C_0 and f_r, we are now able to deduce the inductance value L. By adjusting C_L, L varies from 2.42 to 5.18 µH. For a total footprint of 60 µm × 50 µm, the active inductance density of 1.73 mH/mm^2 is five orders of magnitude higher than previously used passive inductors (55 nH/mm^2)[10] and three orders of magnitude higher than superconding inductors (1.6 µH/mm^2)[11].

b) Capacitance resolution

We now turn to the resolution in capacitance of the set-up, we derive the input-referred noise in aF/$\sqrt{\text{Hz}}$ from the signal-to-noise ratio (SNR) as a function of the integration time t_{int}.

For this, we generate a capacitance change by continuously connecting and disconnecting $C_m = 2$ fF at a rate of 1 kHz. Using the demodulation method (I), the square-wave of the phase ϕ at f_r with a rise time given by the integration time is used to extract the signal power P_{sig} and noise power P_{noise} by separating the corresponding frequency components in the power spectrum. The resulting SNR= P_{sig}/P_{noise} is used to extract the capacitance resolution given by the equivalent $C_m(SNR = 1) = C_m/SNR$ shown in Figure 5.8 as a function of t_{int} from 100 ns to 100 µs. A capacitance of 1 fF can be detected with an integration time of 1 µs with $SNR = 1$. The capacitance resolution follows a square-root law with t_{int} from which we extract the equivalent input-referred noise of 3.7 aF/$\sqrt{\text{Hz}}$, two orders of magnitude higher than the best reported sensitivity using an ultra-low noise SQUID amplifier[12]. The increased noise of the measurement circuit would prevent fast single-shot read-out of quantum states in a few 100 ns but remain reasonable for device characterization with integration times on the order of 1 ms or more.

As correlated noise appears on time scales longer than 1 ms originating probably from the $1/f$ flicker noise of the transistors, we add a second demodulation (IIa) (see Figure 5.6) at the capacitance switching frequency of 1 kHz to remove phase noise originating from a varying L. The 1 kHz square-wave ϕ from (I) with an integration time of 100 µs is demodulated by (IIa) at 1 kHz to obtain its amplitude $|\phi|$. The

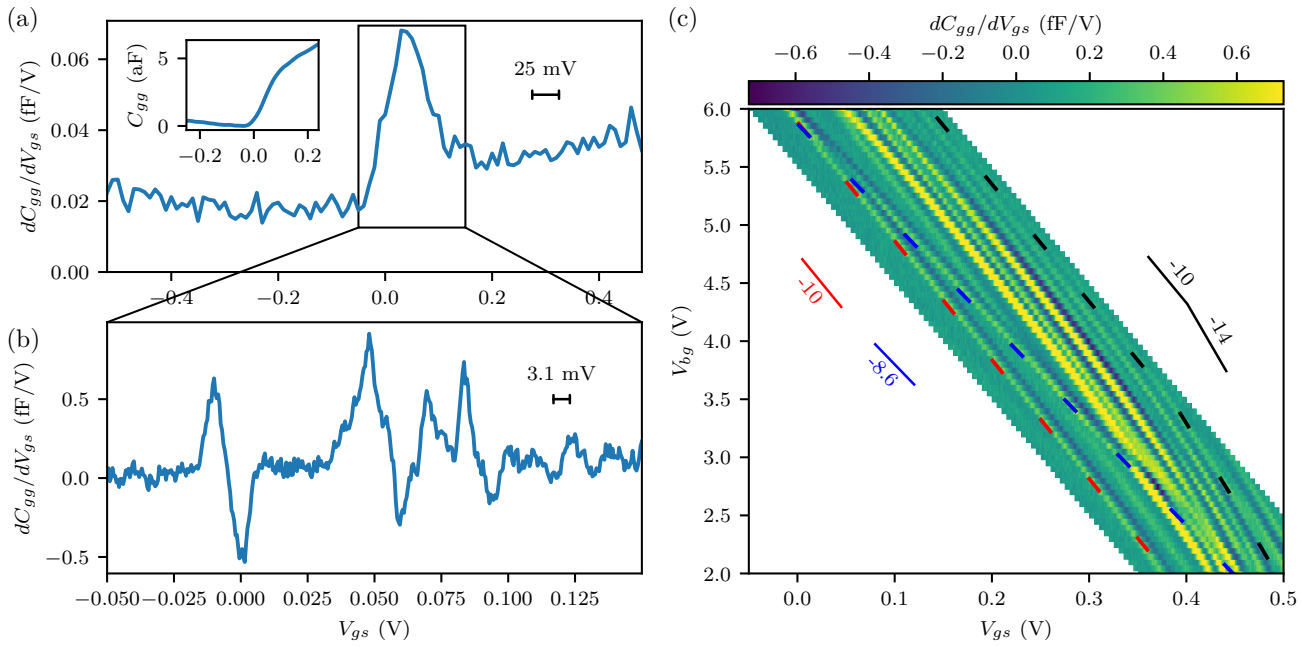


Figure 5.9 | Quantum capacitance measurement of an integrated MOSFET with channel length 120 nm and width 80 nm.

(a) Measurement of the first derivative of the gate capacitance C_{gg} with respect to V_{gs} by applying a gate-source AC excitation of 25 mV. The inset shows the capacitance $C_{gg}(V_{gs})$ computed from the integrated signal of the derivative. (b) Expanded view of dC_{gg}/dV_{gs} around the off-on transition of the MOSFET measured with a smaller excitation of 3.1 mV. The resolved features are signatures of quantized electronic states in the measured capacitance of the MOSFET channel. (c) Evolution of dC_{gg}/dV_{gs} with the back-gate voltage V_{bg} and the gate-source voltage V_{gs}. The indicated slopes $\beta = dV_{bg}/dV_{gs} \simeq C_{g-ch}/C_{bg-ch}$ represent the relative coupling strength of the detected quantized states with respect to back- and front-gate.

capacitance resolution as a function of the second integration time for the 1 kHz demodulation is extracted by taking the ratio of the average and the standard deviation of the $|\phi|$ signal and is shown in Figure 5.8. With an integration time of 1 s, the resolution becomes as low as 10 aF.

VII Quantum capacitance measurements

With the calibrated impedancemetry circuit, we are able to detect the gate quantum capacitance C_{gg} of the multiplexed tiny MOSFETs (M0, M1, M2 in Figure 5.6a) similar to the ones used to implement spin qubits[13] or single-electron transistors for read-out with CMOS technology. Measurements will be presented for M2 with a gate length of 120 nm and a gate width of 80 nm.

The total gate capacitance C_{gg} of such devices corresponds to the sum of the capacitance to drain, source, back-gate, and MOSFET channel of which the gate to channel capacitance depends highly on the gate-source voltage V_{gs} controlled by the DC component of V_{bias} (see Figure 5.6a). As C_{gg} of nanometric devices is extremely small compared to C_0, we don't expect to have sufficient SNR for small capacitance variations at reasonable integration times. Better sensitivity can be obtained by modulating V_{gs} (method

IIb in Figure 5.6c) to measure after demodulation the first derivative dC_{gg}/dV_{gs} as C_{gg} varies a lot in a small V_{gs} window.

While the resonator impedance is probed at 199 MHz, V_{gs} is modulated at 1 kHz with mV-range excitation on V_{bias} (see Figure 5.6bc). The obtained result with a relatively large 25 mV modulation of V_{gs} (shown in Figure 5.9a) is reminiscent of the typical gate capacitance variation around threshold voltage $V_{th} \simeq 0\,\mathrm{V}$ at $V_{bg} = 6\,\mathrm{V}$. C_{gg} (see inset of Figure 5.9a) reflects the typical behavior for a FET capacitance from the subthreshold regime $V_{gs} \ll V_{th}$ to the strong inversion regime $V_{gs} \gg V_{th}$. Upon decreasing the amplitude of the V_{gs} modulation to only 3.1 mV, the observed dC_{gg}/dV_{gs} signal in Figure 5.9b reveals a fine structure around V_{th} consisting of successive peak-dip oscillations. Following numerical integration, these features result in a series of peaks in C_{gg}, which we interpret as quantum contributions to the capacitance coming from electrons tunneling in and out of localized quantum states within the transistor channel.

To further identify these quantum states, we acquire dC_{gg}/dV_{gs} for different back-gate voltage V_{bg} from 2 to 6 V as shown in Figure 5.9c. As V_{bg} increases, all observed features shift to lower V_{gs} with a slope close to the ratio β of gate-channel capacitance C_{g-ch} over the backgate-channel capacitance C_{bg-ch} alike the V_{th}-shift with back-gate for similar FET devices[6]. For $V_{bg} > 4\,\mathrm{V}$, all features have a coupling ratio of 10 except for one with a lower coupling 8.6 attributed to an impurity closer to the back-gate interface. No anomalous impurity structure is detected in the smaller 60 nm × 80 nm device. At lower V_{bg}, the coupling increases with V_{gs} from 10 to 14 as the electron-filled inversion layer is brought back to the top-interface.

These measurements of integrated quantum devices demonstrate that the capacitive signature of structure in the electronic density of states of quantum dots can be probed via impedancemetry.

VIII Conclusions

We reported an integrated circuit that performs impedancemetry of a resonator coupled to a quantum dot at cryogenic temperatures. The active inductance of the resonator allowed the controlled tuning of the resonance frequency and quality factor, which will be of importance for optimal frequency-spectrum crowding in multiplexed read-out schemes. The employed multiplexing of nanometric quantum devices with on-chip switches could be beneficial for reduced power per qubit in a scalable multi-qubit architecture. Novel read-out architectures with cryogenic electronics, such as the active inductance, have the potential to increase scalability and flexibility in the design and exploitation of quantum processors.

Further work towards lower noise and lower power design with more accurate high-frequency models at cryogenic temperatures will improve the final performance. Measuring multiplexed out-of-chip capacitances of quantum devices will be also promising for the screening of quantum devices with a simpler experimental setup than reflectometry. In the long run, the realization of tailored high-end analog electronics at cryogenic temperatures will improve and accelerate the up-scaling of quantum processors.

[FR] Mesure capacitive de puits quantiques

Les états quantiques électroniques au sein des structures à puits quantiques peuvent être efficacement détectés grâce à l'accès à une unique électrode couplé au dispositif. L'électrode est souvent placée au-dessus du puit, similaire à la grille des transistors CMOS usuels. Ce type de mesure est avantageux puisque l'accès a la grille est souvent disponible au sein de telles structures, faisant de cette méthode une mesure non-invasive contrairement aux mesures en transport qui requièrent des contacts supplémentaires au sein du substrat. L'unique accès à la grille requis assouplit les contraintes sur la topologie du réseau de puits quantiques dans la couche substrat, conduisant à une meilleure connectivité entre qubits et à une meilleure régularité.

La capacité de grille des dispositifs quantiques est composée d'une composante géométrique C_{DUT} due principalement aux interconnections et d'une correction d'origine quantique $C_p \ll C_{DUT}$, appelée capacité paramétrique ou quantique, due au blocage de Coulomb au sein du puit. La capacité quantique est de l'ordre du femtofarad pour des états de charge et de l'ordre quelques attofarads pour des états de spin pour des dispositifs quantiques typiques.

La mesure rapide en quelques µs de la petite capacité paramétrique diluée au sein des capacités parasites nécessite des plages dynamiques lors de la mesure allant jusqu'à 10^6 dans les systèmes actuels. La faible valeur de la capacité paramétrique et la mesure à grande vitesse requise pour le calcul quantique rendent la détection difficile et nécessitent des conditions expérimentales soigneusement choisies pour atteindre une lecture haute-fidélité. De plus, la capacité doit être sondée avec des tensions de l'ordre du µV à travers un câblage souvent long de plus d'un mètre reliant les températures ambiantes aux températures cryogéniques. Une solution largement ancrée au travers de toutes les plates-formes de qubits à l'état solide consiste à mesurer le décalage de la fréquence d'un résonateur radiofréquence/micro-onde couplé à la grille du dispositif quantique.

La fréquence de résonance $f_r = 1/2\pi LC_0$ d'un résonateur LC dépend de la capacité totale $C_0 = C_{DUT} + C_p$ en parallèle de l'inductance L. Ainsi, toute variation de la capacité quantique C_p est répercutée sur la fréquence de résonance, souvent observée au travers de la phase lors d'une détection homodyne à la fréquence du résonateur.

La détection rapide de la capacité paramétrique commence à être effectuée de manière routinière grâce à la méthode de réflectométrie utilisant des ondes propagatrices harmoniques. Dans cette implémentation, une onde radiofréquence à des fréquences allant de 0.3 à 3 GHz est réfléchie par le résonateur, avec une phase ajoutée dépendante de C_p. La gamme de fréquences de f_r combinée à un câblage d'un mètre de long nécessitent d'étudier le système dans le cadre des micro-ondes avec des ondes se propageant dans des lignes adaptées de $50\,\Omega$. De ce fait, la réflectométrie mesure la phase des paramètres de diffusion S_{11}.

Dans ce chapitre, nous avons revisité la configuration de mesure de réflectométrie pour tirer parti de la disponibilité de l'électronique cryogénique. Nous avons proposé la configuration de mesure d'impédancemétrie, en remplaçant la mesure des paramètres de diffusion par l'impédance du réservoir grâce à l'ajout des circuits d'excitation et de lecture du résonateur proche du dispositif quantique.

Nous présentons un circuit intégré qui implémente l'impédancemétrie d'un résonateur couplé à des puits quantiques à des températures cryogéniques. De plus, le remplacement de l'inductance passive par une inductance active au sein du résonateur permet le réglage in-situ de la fréquence de résonance et du facteur de qualité, important lorsque l'on considère l'encombrement optimal du spectre en fréquences dans les schémas de lecture multiplexée. Le multiplexage de dispositifs quantiques avec des commutateurs sur puce, démontré aux températures cryogéniques, est intéressant afin de réduire la puissance dissipée par qubit dans une architecture multi-qubit. Les nouvelles architectures de lecture impliquant de l'électronique classique cryogénique, telles que l'inductance active, ont le potentiel d'améliorer et de simplifier la conception et l'exploitation des futurs processeurs quantiques.

Des travaux supplémentaires vers une conception à faible bruit et à faible puissance avec des modèles hautes-fréquences plus précis aux températures cryogéniques amélioreront certainement les performances finales. À long terme, la réalisation d'électronique analogique sur-mesure à des températures cryogéniques améliorera et accélérera la mise à l'échelle des processeurs quantiques.

Bibliography

[1] Fei Yuan. *CMOS Active Inductors and Transformers*. Springer US, Boston, MA, 2008. Available from: http://link.springer.com/10.1007/978-0-387-76479-5.

[2] A. Ilker Karsilayan and R. Schaumann. **A high-frequency high-Q CMOS active inductor with DC bias control**. In *Proceedings of the 43rd IEEE Midwest Symposium on Circuits and Systems (Cat.No.CH37144)*, **1**, pages 486–489 vol.1, August 2000.

[3] L. Le Guevel, G. Billiot, S. De Franceschi, A. Morel, X. Jehl, A. G. M. Jansen, and G. Pillonnet. **Compact gate-based read-out of multiplexed quantum devices with a cryogenic CMOS active inductor**, February 2021. Available from: http://arxiv.org/abs/2102.04364.

[4] T. Poiroux, O. Rozeau, P. Scheer, S. Martinie, M. Jaud, M. Minondo, A. Juge, J. C. Barbé, and M. Vinet. **Leti-UTSOI2.1: A Compact Model for UTBB-FDSOI Technologies—Part II: DC and AC Model Description**. *IEEE Transactions on Electron Devices*, **62**(9):2760–2768, September 2015.

[5] Bruna Cardoso Paz, Loïck Le Guevel, Mikaël Cassé, Gérard Billiot, Gaël Pillonnet, Aloysius Jansen, Sebastien Haendler, Andre Juge, Emmanuel Vincent, Philippe Galy, Gérard Ghibaudo, Maud Vinet, Silvano de Franceschi, Tristan Meunier, and Fred Gaillard. **Integrated Variability Measurements of 28 nm FDSOI MOSFETs down to 4.2 K for Cryogenic CMOS Applications**. In *2020 IEEE 33rd International Conference on Microelectronic Test Structures (ICMTS)*, pages 1–5, May 2020.

[6] B. Cardoso Paz, L. Le Guevel, M. Cassé, G. Billiot, G. Pillonnet, A. G. M. Jansen, R. Maurand, S. Haendler, A. Juge, E. Vincent, P. Galy, G. Ghibaudo, M. Vinet, S. de Franceschi, T. Meunier, and F. Gaillard. **Variability Evaluation of 28nm FD-SOI Technology at Cryogenic Temperatures down to 100mK for Quantum Computing**. In *2020 IEEE Symposium on VLSI Technology*, pages 1–2, June 2020.

[7] Haiqiao Xiao and Rolf Schaumann. **A 5.4-GHz high-Q tunable active-inductor bandpass filter in standard digital CMOS technology**. *Analog Integrated Circuits and Signal Processing*, **51**(1):1–9, April 2007. Available from: https://doi.org/10.1007/s10470-007-9040-1.

[8] H. Barthelemy and W. Rahajandraibe. **NMOS transistors based Karsilayan Schaumann gyrator :'lowpass and bandpass filter applications'**. In *2003 46th Midwest Symposium on Circuits and Systems*, **1**, pages 97–100 Vol. 1, December 2003.

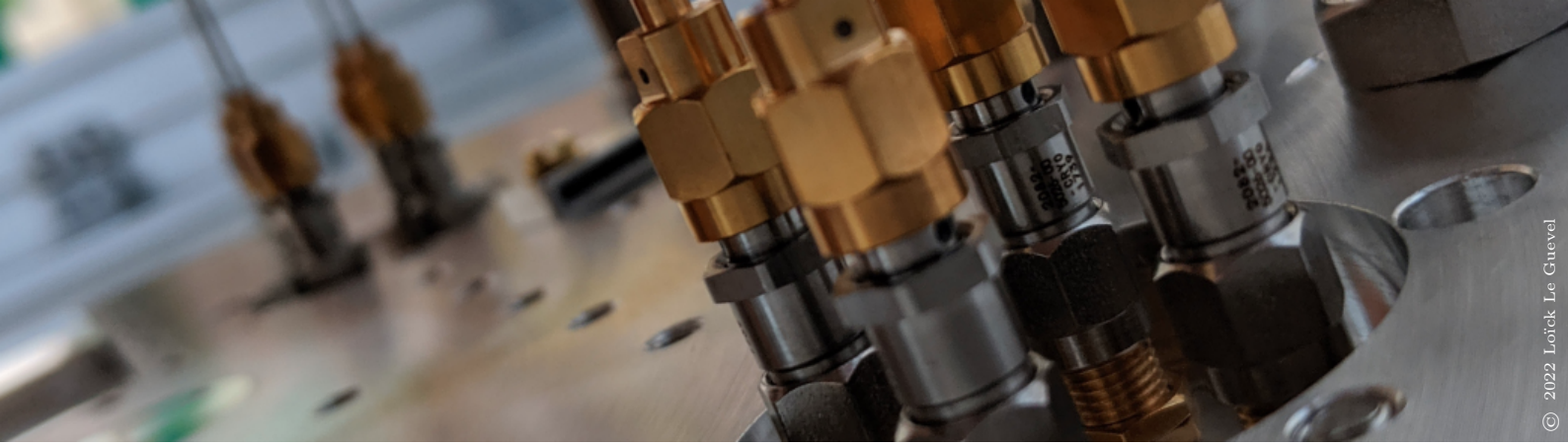
[9] Bishnu Patra, Mohammadreza Mehrpoo, Andrea Ruffino, Fabio Sebastiano, Edoardo Charbon, and Masoud Babaie. **Characterization and Analysis of On-Chip Microwave Passive Components at Cryogenic Temperatures**. *IEEE Journal of the Electron Devices Society*, **8**:448–456, 2020.

[10] A. Crippa, R. Ezzouch, A. Aprá, A. Amisse, R. Laviéville, L. Hutin, B. Bertrand, M. Vinet, M. Urdampilleta, T. Meunier, M. Sanquer, X. Jehl, R. Maurand, and S. De Franceschi. **Gate-reflectometry dispersive readout and coherent control of a spin qubit in silicon**. *Nature Communications*, **10**(1):2776, December 2019. Available from: http://www.nature.com/articles/s41467-019-10848-z.

[11] J. M. Hornibrook, J. I. Colless, A. C. Mahoney, X. G. Croot, S. Blanvillain, H. Lu, A. C. Gossard, and D. J. Reilly. **Frequency multiplexing for readout of spin qubits**. *Applied Physics Letters*, **104**(10):103108, March 2014. Available from: http://aip.scitation.org/doi/10.1063/1.4868107.

[12] F. J. Schupp, F. Vigneau, Y. Wen, A. Mavalankar, J. Griffiths, G. A. C. Jones, I. Farrer, D. A. Ritchie, C. G. Smith, L. C. Camenzind, L. Yu, D. M. Zumbühl, G. A. D. Briggs, N. Ares, and E. A. Laird. **Radio-frequency reflectometry of a quantum dot using an ultra-low-noise SQUID amplifier**. *arXiv:1810.05767 [cond-mat, physics:physics, physics:quant-ph]*, June 2020. Available from: http://arxiv.org/abs/1810.05767.

[13] R. Maurand, X. Jehl, D. Kotekar-Patil, A. Corna, H. Bohuslavskyi, R. Laviéville, L. Hutin, S. Barraud, M. Vinet, M. Sanquer, and S. De Franceschi. **A CMOS silicon spin qubit**. *Nature Communications*, **7**:13575, November 2016. Available from: https://www.nature.com/articles/ncomms13575.



CHAPTER 6
Conclusion & Perspectives

The research described in this thesis aimed to explore and evaluate the industry-standard CMOS FD-SOI 28 nm technology for use at cryogenic temperatures in quantum computing applications. The control and probing of quantum devices require accurate and low-noise high-frequency circuitry for high fidelity operations. Moreover, the ability of CMOS control electronics to favorably scale with the qubit number is primordial to unleash the full capabilities of quantum computing.

This concluding chapter gives a preliminary answer on how the FD-SOI technology could be a good contender for cryogenic operation. We highlight the lines of research in this Ph.D. work from systematic characterization of individual transistors at cryogenic temperatures to the implementation of on-chip circuitry for specific measurements of quantum devices in the millikelvin range. The FD-SOI technology could evolve into a turnkey solution for quantum computing applications based on a complete design kit for circuit design at cryogenic temperatures. We end this dissertation by picturing the challenges ahead related to circuit architecture and design for ultimately entering into the large-scale quantum processing era.

I CMOS FDSOI 28nm for cryogenic circuits

a) A prototypical technology down to sub-Kelvin temperatures

A CMOS technology aimed for general-purpose integrated circuits must contain at least on-chip transistors, resistors, capacitors, and inductors to reach the full high-speed potential via monolithic integration. This work demonstrates that transistors, resistors, and capacitors all work perfectly from room down to sub-Kelvin temperatures at low and moderate frequencies ($\simeq 1\,\mathrm{GHz}$).

Metal-Oxide-Metal capacitors are used in all our cryogenic circuits to ensure feedback stability (Chapter 4) or to tune circuit parameters with e.g. switchable banks (Chapter 5). This common type of capacitor offers high density of capacitance with a weak temperature dependence down to the lowest temperatures, making it ideal for cryogenic circuit design.

Common types of resistors, such as silicided and unsilicided resistors (both active and polysilicon) were separately measured down to cryogenic temperatures. None of these resistors transitioned from metal to insulator at cryogenic temperature thanks to their high doping dose. Silicided active, silicided polysilicon and unsilicided active resistors all show a metal-like behavior with a linear decrease of resistance down to a few tens of Kelvin followed by a resistivity saturation at sufficiently low temperature. These resistor flavors lose between 25 to 37 % of their value from 300 to 4.2 K. On the other end, unsilicided polysilicon resistors show a very low dependence on temperature with a variation below 5 % on the entire temperature

range, thanks to an intrinsic high level of impurities. This last type of resistor has been used in some of our cryogenic circuits to set the transimpedance gain (Chapter 4), or to realize filter functions in the bias-tee used to make the current excitation of the active inductance (Chapter 5).

The transistors made with the FD-SOI 28nm technology have been extensively studied from room to sub-Kelvin temperatures as they are the center piece for any circuit. A good knowledge of their behavior is primordial to make optimal circuit-design choices. The current-voltage characteristics of thousands of transistors were efficiently acquired by integrating them on-chip inside a large multiplexing matrix made of digitally-controlled switches (Chapter 2). This novel method to characterize a large number of devices down to sub-Kelvin temperatures offered the first extensive data set about a CMOS technology at these temperatures. The multiplexing matrix of transistors contains a representative set of different transistors for circuit design and transistor modeling, including variations in geometry (length and width), doping type, oxide thickness, and back-gate doping (threshold flavors). Transistors were organized in matching pairs with added redundancy for some chosen geometries to study the mismatch properties down to cryogenic temperatures. To remain concise, this dissertation focused on the transistor properties of low-threshold voltage (LVT) thin-oxide transistors such as the threshold voltage or the transconductance. This device flavor is likely to be favored in cryogenic applications for the already-lower threshold voltage and the small geometries. Despite the threshold voltage increase at cryogenic temperatures, the transistor transconductance and conductance remain of the same order as at room temperatures, making the transistors comparably good and usable. These two properties increase almost linearly with smaller length without any surprising dependence.

To account for the transistor-transistor biasing interdependence in real circuit implementation, we also measured digital and analog circuits made of a few transistors to investigate the performance of the technology at 4.2 K against 300 K. These basic circuit elements included level-shifters, ring-oscillators, DACs, TIA, and pass-gates (Chapter 3). We defined circuit level quantities assessing their performance such as how large is the input voltage range, how fast the circuit is, or how energy efficient it is. These quantities tend to worsen at cryogenic temperatures, mainly due to the higher threshold voltage, severely impacting the transistor biasing in all the tested topologies.

Luckily, the unique feature of the back-gate in the FD-SOI technology provides a flexible solution to compensate for the temperature dependence of the threshold voltage.

b) Forward body biasing as a cryogenic circuit performance booster

A positive voltage for NMOS or a negative voltage for PMOS on the back-gate effectively lowers the threshold voltage for the front gate. The increase of threshold voltage at cryogenic temperatures can be compensated by applying a suitable back-gate voltage, leaving the threshold voltage at a typical value with an increased transconductance and lower thermal noise of the transistors at lower temperatures. The enhanced transistor performance at cryogenic temperature with forward back bias translates into a boost of circuit performance (Chapter 3). Most circuits tested at 4.2 K with FBB surpass their room-temperature counterparts for the same range of inputs. Combined with the already lower thermal noise at cryogenic temperatures, the flicker noise is reduced with FBB by pulling the conduction channel away from the irregular high-k top-gate oxide. Additionally, analog circuits become more energy efficient with an increased ratio of transconductance and drain-source current while digital circuits still benefit from the reduced leakage.

The application of FBB at cryogenic temperatures with the gain in performance could be the key for energy-efficient low-noise circuitry as required for quantum computing applications. The capability of the FD-SOI technology to properly operate down to the qubit temperature stage opens the way for electronics to co-exist in the closest proximity to the quantum devices. The already demonstrated FD-SOI quantum

devices could ultimately share the same die, reducing the interconnections to sub-micron lengths.

c) A single technology for quantum and classical electronics

We explored the co- and full-integration of quantum devices with classical electronics at cryogenic temperatures. Co-integration refers to the assembly of several separate dies with short external Input/Output connections (I/Os) such as wire-bonding while integration refers to the use of one unique die with internal I/Os.

The co-integration of quantum and classical electronics allows the combination of different technologies involving separate optimization for better overall performance. Moreover, it allows separate temperatures for both dies, or separate thermalization links via clever engineering of the environment to loosen the power dissipation requirement for the control electronics, and offer a better environment for the quantum devices. We co-integrated a low-power transimpedance amplifier (TIA) made of the industry-standard FD-SOI 28nm technology to perform pA-resolved current measurements through a single quantum dot made with a similar FD-SOI process but slightly modified in the LETI clean room for better quantum performance (Chapter 4). The quantified current through the dot at 4.2 K is measured by the TIA wire-bonded millimeters-away. The measured Coulomb diamonds with the cryogenic TIA demonstrate the successful co-integration. The preserved quantum characteristics confirm the absence of significant heating of the nearby quantum dot, attributed to the low power dissipation of the TIA. While the dissipation from the classical electronics can be alleviated by good thermal engineering, external I/Os limit the number of measured devices as wire-bonding pads are $\sim 10^3$ larger than the typical quantum devices. This I/O crowding can become a bottleneck for quantum computing applications as thousands of devices need to be simultaneously measured and controlled. A drastic solution to the I/O bottleneck could be to integrate quantum and classical devices within a single silicon die allowing I/O pitches of the order of the quantum device size.

We integrated quantum-dot devices with classical electronics within the same substrate of the CMOS FD-SOI 28nm technology from STMicroelectronics (Chapter 4). Single and double quantum dots are made of one or two minimum size transistors. Despite having less fine features (about twice the possible width with the Leti fabrication), clear quantum behavior appears at temperatures of 4.2 K or below. We first integrated on-chip a double quantum dot with the same TIA as for the co-integration assembly and cooled the single die to 100 mK in a dilution fridge. The observation of clear Coulomb triangles in the measured current at 100 mK validates that the heating of the nearby electronics does not disturb the quantum behavior of the quantum dot device micrometers away. For even further integration with high-frequency excitation of the quantum-dot system, an on-chip GHz variable oscillator was connected to one of the dot gates. The high-frequency excitation alters the DC current flowing through the two dots, measured with the integrated TIA. Even with the high-frequency generator and current read-out amplifier, the chip dissipation remained within the cooling capabilities of the fridge, thus showcasing the energy-efficiency of the FD-SOI technology at low temperatures.

The second demonstration of integration consisted in measuring the gate capacitance of nanometric transistors to detect signs of quantum-dot effects (Chapter 5). The detection of the so-called quantum capacitance is at the heart of spin and charge qubit read-out schemes, essential for quantum algorithms on a semiconductor platform. Conventionally, the nanometric transistor-like device is embedded in an LC tank made of a mm^2 nH surface-mounted inductor to resonate at frequencies around 0.5 MHz. Keeping track of the resonator frequency change gives direct information on the quantum capacitance of the device, used to infer the qubit state. While this solution is widely used in semiconductor quantum experiments with a few devices, the large size of the discrete inductor (10^5 times bigger than a single qubit device), raises questions about the possibility of scaling this read-out setup to tens to thousands of devices. We decided to approach

this issue with an innovative look allowed by the use of cryogenic integrated circuits. The ability to place electronics close to the device allowed us to ditch the typical 50 Ω matching environment, required for remote probing in e.g. transmission and reflection measurements done so far. The impedancemetry scheme directly measures the impedance of the resonator by applying an current and reading the voltage at the resonator node. Moreover, the use of an active inductance occupying on the chip, effectively emulating the behavior of a passive inductor has the potential to increase the inductance density by a factor of 10^3. The interconnction were kept to a minimum by placing the excitation current source and read-out voltage buffer immediately next to the active inductance. The integration of a buffering circuit with the quantum device in the same substrate really minimized the interconnection parasitic capacitance, further reducing the required inductance value, hence its size. The obtained circuit for the LC tank has a surface of only $50\,\mu m \times 60\,\mu m$, only 10^2 times bigger than a single qubit device. Finally, the cryogenic circuit revealed changes in the gate capacitance in time-multiplexed nanometric transistors due to quantum effects near the inversion onset of the transistor channel.

II Perspectives

By validating the cryogenic functionality of the CMOS FD-SOI technology from basic building blocks to circuits, this thesis work opens the path towards the design of a control and read-out system meeting the stringent requirements of quantum computing. However, crucial challenges remain ahead, ranging from the understanding of cryogenic temperature effects on CMOS technologies to the conception of a scalable low-noise circuit architecture.

a) Towards a cryogenic design kit for circuit designers

Accurate modeling and simulation of CMOS technologies is the key to realize intricated systems meeting the specifications of the targeted application. While technologies at room temperatures benefit from decades of development, the lack of a complete design kit at cryogenic temperatures is a serious impediment. Different approaches towards a complete design kit are taken by different teams. The long-term approach extends physics-based models to the low-temperature regime while the more pragmatic short-term approach fits a crude model of the transistor characteristics onto experimental data taken at cryogenic temperatures. Both approaches for the development of a design kit require more measurements of devices and circuits compared to the set of the low-frequency data presented in this work.

High-frequency data along with noise analysis of the transistors, the passive elements, and the interconnects at cryogenic temperature are primordial for circuit design operating at frequencies in the GHz range. Cryogenic probe stations are largely used for careful electrical characterization of single devices up to 10's GHz for temperatures above a few Kelvins which would be sufficient to capture most cryogenic effects on transistors. Subsequently, the model parameters adapted for cryogenic effects can be extracted from the acquired data. Models are the key for fast simulation procedures to describe a wider continuous range of device geometries as would be required to properly design circuits.

b) Quantum architecture targeted integrated circuits

With a design kit accurately describing circuit behavior at cryogenic temperatures, the next challenge for quantum computing applications is to design an architecture able to run quantum algorithms outperforming classical solvers. The required simultaneous control and read-out of thousands to more than millions of qubits pose serious constraints and trade-offs on the integrated-circuit side of quantum processing units.

The widely-adopted current approach to the architecture design follows a quantum-device first approach, in which the quantum part dictates the requirements for the rest of the system. This approach is the result of the early pioneering development of qubits lacking design flexibility compared to the electronics component.

Recent work demonstrated high-fidelity control and read-out of one and two qubit devices with cryogenic integrated circuits operating within the fridge power budget[1, 2, 3, 4, 5]. However, the high power consumption per qubit of a few 10's mW strongly limits the number of simultaneously-controllable qubits, thus hindering the algorithmic capacity. The major foreseen challenge in scaling up the qubit number is the development of ICs meeting the accuracy and noise requirement for high-fidelity operations at a reduced fraction of the fridge cooling power.

Several axes of research are taken to tackle this problem: from reducing the power consumption with clever circuit design, to temperature-staggered electronic architecture, or to cryogenic fridge improvements. The first approach is a purely IC-focused improvement in which the power consumption is minimized by a careful and meticulous system design. Likely, the use of e.g. bipolar technologies for high-speed and low-noise parts will be mixed with modern nanometric technologies such as FinFET for low-power digital processing power. Some proposed electronics architectures leverage the higher cooling power of fridges at higher temperature stages such as 77 K or even at 300 K. The total power consumption is separated between different cooling power budgets to reduce the heat load on the critical 4.2 K stage. This approach requires careful design and consideration for I/Os communication bottlenecks between each temperature stage due to the limited number of connections. A few dilution-fridge companies are already working on the increase of the cooling power at 4.2 K with e.g. stronger pulse tubes. This cooling power boost will allow to fit more electronics, hence increasing the number of active qubits.

Temperature-staggered architectures are based on the co-optimization of very different hardware parts: the cooling system and the control system. These two parts are no longer separately designed but considered as one big engineering challenge to reach new performance heights. This simultaneous inter-disciplinary optimization might be endemic of future developments of quantum processing units in the next decades when quantum devices will be better understood and qubit design will be an extensive and prolific research field. Transitioning from a quantum-first to a system-first in which every part is equally important will enable the control of large number of qubits for the given available technology at the given design time. Already today, we can see new qubit types with high-temperature spin qubits operating above 1 K or new low-frequency superconducting qubits quickly approaching the transmon quality. Soon enough, the qubit type and various properties will enter the optimization loop to allow large-scale quantum computation. Tomorrow's quantum computer architects will have to co-optimize all components of the quantum processing unit: from the cryogenics, to the electronics, to the algorithms, and to the quantum devices.

[FR] Conclusion & perspectives

La recherche décrite dans cette thèse vise à explorer et à évaluer la technologie commerciale CMOS FD-SOI 28 nm pour une utilisation à des températures cryogéniques pour des applications en calcul quantique. Le contrôle et la mesure des dispositifs quantiques nécessitent des circuits hautes-fréquence précis et à bas bruit pour des opérations sans fautes. Ces mêmes circuits doivent aussi pouvoir suivre l'augmentation du nombre de qubit avec des faibles consommations pour débloquer les avantages du calcul quantique à long terme.

Ce dernier chapitre donne une réponse préliminaire sur la façon dont la technologie FD-SOI pourrait être un bon candidat pour la réalisation de circuits cryogéniques. Nous mettons en évidence les différents axes de recherche de ce doctorat : de la caractérisation systématique de transistors individuels à des températures cryogéniques à la mise en œuvre de circuits intégrés pour des mesures spécifiques de dispositifs quantiques jusqu'aux températures de l'ordre de quelques millikelvins. Il est envisageable que la technologie FD-SOI puisse évoluer vers une solution clé-en-main pour les applications d'informatique quantique basée sur un kit de conception complet pour la conception de circuits aux températures cryogéniques. Nous terminons cette thèse en décrivant les défis à venir liés à l'architecture et à la conception des circuits afin d'entrer, un jour, dans l'ère du calcul quantique à grande échelle.

c) La technologie CMOS FDSOI 28nm pour des circuits cryogéniques

Une technologie complète jusqu'à des températures inférieures au Kelvin

Une technologie CMOS destinée à la réalisation de circuits intégrés doit contenir à minima : des transistors, des résistances, des condensateurs et des inductances sur puce pour atteindre le plein potentiel à grande vitesse via une intégration monolithique. Ce travail démontre que les transistors, les résistances et les condensateurs fonctionnent tous parfaitement des températures ambiantes jusqu'aux températures cryogéniques bien inférieures au Kelvin à des fréquences basses et modérées ($< 1\,\mathrm{GHz}$).

Les condensateurs métal-oxyde-métal (MOM) sont utilisés dans tous nos circuits cryogéniques pour assurer la stabilité de la boucle de rétroaction (Chapitre 4) ou pour régler les paramètres des circuits avec, par exemple, des banques commutables (Chapitre 5). Ce type de condensateur très utilisé offre une haute densité de capacité avec une faible dépendance en température, et ce, jusqu'aux plus basses températures, les rendant idéales pour la conception de circuits cryogéniques.

Les résistances siliciurées et non siliciurées (à la fois « active » et « polysilicium ») ont été mesurées individuellement jusqu'aux températures cryogéniques. Aucune de ces résistances n'est passée de métal à isolant à température cryogénique grâce à leur forte dose de dopage. Les résistances « active » siliciurées, « polysilicium » siliciurées et « active » non siliciurées présentent toutes un comportement de type métal avec une diminution linéaire de la résistance jusqu'à quelques dizaines de Kelvin suivie d'une saturation de résistivité à des températures suffisamment basses. Ces types de résistance perdent entre 25 et 37 % de leur valeur de 300 à 4.2 K. D'un autre côté, les résistances en « polysilicium » montrent une très faible dépendance à la température avec une variation inférieure à 5 % sur toute la plage de température grâce à un nombre élevé d'impuretés. Ce dernier type de résistance a été utilisé dans certains de nos circuits cryogéniques pour régler le gain de transimpédance (Chapitre 4) ou pour réaliser des filtres tels que dans le « bias-tee » utilisé pour faire l'excitation en courant de l'inductance active (Chapitre 5).

Les transistors fabriqués avec la technologie FD-SOI 28 nm ont été largement étudiés des températures ambiantes aux températures inférieures au Kelvin, puisqu'ils sont les pièces maîtresses de tout circuit intégré. Une bonne connaissance de leur comportement est primordiale pour faire des choix optimaux lors de la conception de circuits. Les caractéristiques courant-tension de milliers de transistors ont été

rapidement acquises en intégrant 1024 transistors sur une même puce au sein d'une grande matrice de multiplexage composée de commutateurs à commande numérique (Chapitre 2). Cette nouvelle méthode capable de caractériser un grand nombre de dispositifs jusqu'à des températures inférieures au Kelvin a offert le premier ensemble de données basses-fréquences complet sur une technologie CMOS à ces températures. La matrice de multiplexage contient un ensemble représentatif de différents transistors pour la conception de circuits et la modélisation des transistors, avec des variations de géométrie (longueur et largeur), du type de dopage, de l'épaisseur d'oxyde et du dopage de la grille arrière. Les transistors ont été organisés en paires assorties avec une redondance supplémentaire pour certaines géométries, dans le d'étudier le dépareillement entre transistors de même géométrie jusqu'aux températures cryogéniques. Pour rester concis, cette thèse s'est concentrée sur les propriétés des transistors à oxyde fin (GO1) à faibles tensions de seuil (LVT) telles que les valeurs de tension de seuil, la transconductance ou bien encore la conductance. Ce type de transistor est susceptible d'être favorisée pour les applications cryogéniques avec des tension de seuil réduites et des petites géométries. Malgré l'augmentation des tensions de seuil aux températures cryogéniques, la transconductance et la conductance des transistors restent du même ordre qu'aux températures ambiantes, rendant les transistors relativement bons et utilisables. Ces deux propriétés augmentent presque linéairement avec des longueurs plus petites sans aucune dépendance anormale jusqu'aux plus basses températures.

Pour tenir compte de l'interdépendance entre transistors dans la mise en œuvre réelle de circuits, nous avons également étudié des circuits numériques et analogiques constitués de quelques transistors afin de comparer les performances de la technologie entre 4.2 et 300 K. Ces circuits de base comprennent des translateurs de niveau, des oscillateurs en anneaux, des convertisseurs digital vers analogue, un amplificateur opérationnel, et des « pass-gates » (Chapitre 3). Nous avons défini des grandeurs évaluant leurs perfonces telles que la plage de bon fonctionnement en tension d'entrée, la vitesse du circuit ou bien son efficacité énergétique. Ces quantités ont tendance à diminuer aux températures cryogéniques, principalement due aux tensions de seuil plus élevées, impactant sévèrement la polarisation des transistors dans toutes les topologies testées.

Cependant, l'accès à la grille arrière, unique aux technologies FD-SOI, offre une solution flexible afin de compenser l'augmentation des tensions de seuils aux basses températures.

Polarisation de grille arrière pour booster les performances des circuits cryogéniques

Une tension positive pour les NMOS ou une tension négative pour les PMOS sur la grille arrière abaisse la tension de seuil de la grille avant. L'augmentation de la tension de seuil aux températures cryogéniques peut être compensée en appliquant une tension de grille arrière appropriée, laissant la tension de seuil à une valeur typique avec une transconductance accrue et un bruit thermique réduit des transistors aux températures cryogéniques. Les performances améliorées des transistors à température cryogénique avec une polarisation de la grille arrière se traduisent directement par une amélioration des performances des circuits intégrés (Chapitre 3). La plupart des circuits testés à 4.2 K avec l'utilisation de la grille arrière surpassent leurs homologues à température ambiante. Combiné avec le bruit thermique déjà plus faible aux températures cryogéniques, le bruit de « flicker » est aussi réduit avec l'application d'une tension de grille arrière qui éloigne le canal de conduction de l'oxyde de grille avant, fortement irrégulier. Finalement, les circuits analogiques deviennent plus économes en énergie avec un rapport plus élevé de transconductance sur le courant drain-source g_m/I_{ds}, tandis que les circuits numériques bénéficient des fuites de courant fortement réduites aux basses températures.

La polarisation de la tension de grille arrière aux températures cryogéniques pourrait être la clé pour la réalisation de circuits à bas bruit et économes en énergie comme l'exigent les applications de calcul quantique. La capacité qu'a la technologie FD-SOI à fonctionner correctement jusqu'aux températures les

plus basses, auxquelles les qubit opèrent, ouvre aussi la voie à la coexistence de l'électronique classique au plus près des dispositifs quantiques. Les dispositifs quantiques FD-SOI déjà démontrés pourraient finalement partager le même substrat, réduisant la longueur des interconnexions au-dessous du micromètre.

d) Une technologie unique pour l'électronique quantique et classique

Nous avons exploré la cointégration et l'intégration de dispositifs quantiques avec de l'électronique classique à des températures cryogéniques. La cointégration fait référence à l'assemblage de plusieurs puces séparées par de courtes connexions externes telles que des fils de « wirebond », tandis que l'intégration fait référence à l'utilisation d'un seul et même substrat.

La cointégration de l'électronique quantique et classique permet l'utilisation de différents procédés de fabrication afin d'optimiser séparément chaque technologie. De plus, il est même possible d'avoir des températures différentes entre chaque puce grâce à une conception soigneuse de l'environnement thermique, pouvant améliorer les conditions d'opération des deux puces. Nous avons cointégré un amplificateur à transimpédance (TIA) de faible puissance réalisé avec la technologie commerciale CMOS FD-SOI 28 nm afin d'effectuer des mesures de courant de l'ordre du picoampère à travers un puit quantique fabriqué avec un processus FD-SOI similaire dans la salle blanche du CEA-LETI, légèrement modifié pour de meilleures performances quantiques (Chapitre 4). Le courant quantifié à travers le puit à 4.2 K est mesuré par le TIA placé à quelques millimètres et relié par un fil de « wirebond ». Les diamants de Coulomb mesurés avec le TIA cryogénique démontrent le succès de la cointégration aux basses températures. La conservation des caractéristiques quantiques confirme l'absence d'échauffement significatif du puit quantique voisin, attribué à la faible puissance dissipée du TIA. Alors que la dissipation de l'électronique classique peut être absorbée grâce l'ingénierie thermique du système, les entrées et sorties (E/S) externes limitent le nombre de dispositifs mesurés car les plots de connexion par fil de « wirebond » sont $\sim 10^3$ plus encombrant que les dispositifs quantiques étudiés. Cet encombrement des E/S peut devenir un frein pour les applications de calcul quantique, car des milliers de dispositifs doivent être simultanément mesurés et contrôlés. Une solution à ce problème pourrait consister à intégrer des dispositifs quantiques et classiques dans une seule et même puce en silicium permettant des E/S sur puce du même ordre de grandeur que les dispositifs quantiques.

Nous avons intégré des dispositifs à puits quantiques avec de l'électronique classique dans le même substrat réalisés avec la technologie commerciale CMOS FD-SOI 28 nm (Chapitre 4). Les puits quantiques simples et doubles sont constitués d'un ou deux transistors de taille minimale. Malgré des caractéristiques géométriques moins fines (environ deux fois la largeur possible avec la fabrication du CEA-LETI), un comportement quantique clair apparaît à des températures de 4.2 K ou moins. Nous avons d'abord intégré sur puce deux puits quantiques avec le TIA cryogénique, refroidi dans un réfrigérateur à dilution. L'observation des triangles de Coulomb dans le courant mesuré à 100 mK confirme que l'échauffement de l'électronique à proximité ne perturbe pas le comportement quantique du dispositif mesuré, même à quelques micromètres de distance. Pour une intégration encore plus poussée, un oscillateur sur puce couvrant des fréquences de l'ordre du GHz a été connecté à l'une des grilles du dispositif quantique. Même avec le générateur haute-fréquence et l'amplificateur de lecture en courant, la dissipation de la puce reste dans les limites des capacités de refroidissement du réfrigérateur, démontrant ainsi l'efficacité énergétique de la technologie FD-SOI à basse température.

La seconde démonstration d'intégration a consisté à mesurer la capacité de grille de transistors nanométriques pour détecter des signes d'effets quantiques (Chapitre 5). La détection de la capacité de grille des dispositifs quantiques dite quantique est au cœur des schémas de lecture des qubits de spin et de charge, essentiels pour l'exécution d'algorithmes quantiques. Le dispositif de type transistor nanométrique est généralement connecté à un résonateur LC résonnant à des fréquences autour de 0.3 à 3 GHz. Le

suivi de la fréquence du résonateur donne une information directe sur la capacité quantique du dispositif, utilisée pour déduire l'état du qubit. Bien que cette solution soit largement utilisée avec quelques dispositifs quantiques, la grande taille de l'inductance discrète (10^5 fois plus grande qu'un seul dispositif quantique), soulève des questions quant à la possibilité de mettre à l'échelle cette configuration de lecture à des dizaines voire des milliers de dispositifs mesurés. Nous avons décidé d'aborder cette problématique avec un regard innovant, uniquement permis par l'utilisation de circuits intégrés cryogéniques. La possibilité de placer l'électronique à proximité du dispositif quantique nous a permis d'abandonner l'environnement d'adaptation d'impédance typique de $50\,\Omega$, requis lors mesures en transmission et en réflexion effectuées jusqu'à présent. Le schéma d'impédancemétrie mesure directement l'impédance du résonateur en appliquant un courant et en mesurant la tension. De plus, l'utilisation d'une inductance active émulant le comportement d'une inductance passive a le potentiel d'augmenter la densité d'inductance d'un facteur 10^3. Les interconnexions ont été réduites au minimum en plaçant la source de courant d'excitation et l'amplificateur de lecture en tension immédiatement à proximité de l'inductance active. L'intégration de l'électronique de lecture avec le dispositif quantique au sein du même substrat silicium a permis de minimiser les capacités parasites, réduisant la valeur d'inductance requise et donc sa taille. Le circuit obtenu pour le réservoir LC a une surface de seulement $50\,\mu m \times 60\,\mu m$, seulement 10^2 fois plus grande qu'un seul dispositif qubit. Enfin, le circuit cryogénique a permis de révéler des changements dans la capacité de grille des transistors nanométriques multiplexés dues à des effets quantiques au début de l'inversion du canal du transistor.

e) Perspectives

En validant le bon fonctionnement de la technologie CMOS FD-SOI a des températures cryogéniques pour des circuits de taille modérée, ce travail de thèse ouvre la voie vers la conception de systèmes de contrôle et de lecture répondant aux exigences strictes du calcul quantique. Cependant, des défis majeurs restent à relever, allant de la compréhension des effets des températures cryogéniques sur les technologies CMOS à la conception d'une architecture de circuit de contrôle à bas bruit et haute efficacité énergétique.

Vers un kit de conception cryogénique pour la conception de circuits intégrés

La modélisation et la simulation des technologies CMOS sont la clé pour réaliser des systèmes complexes répondant aux spécifications de l'application ciblée. Alors que les technologies à température ambiante bénéficient de décennies de développement, l'absence d'un kit de conception complet à des températures cryogéniques est un sérieux obstacle. Différentes approches vers un kit de conception complet sont prises par différentes équipes. L'approche à long terme étend les modèles existants à la physique des basses températures tandis que l'approche à court terme plus pragmatique adapte un modèle brut des caractéristiques du transistor aux données expérimentales prises à des températures cryogéniques. Les deux approches pour le développement d'un kit de conception nécessitent plus de mesures des transistors et des circuits par rapport à l'ensemble des données basses fréquences présentées dans ce travail.

Les données hautes-fréquences ainsi que l'analyse du bruit des transistors, des éléments passifs et des interconnexions à température cryogénique sont primordiales pour la conception de circuits fonctionnant à des fréquences de l'ordre du GHz. Les stations sous pointes cryogéniques sont largement utilisées pour la caractérisation électrique minutieuse des éléments actifs et passifs jusqu'à $\sim 10\,GHz$ à des températures supérieures à quelques Kelvins, ce qui est suffisant pour capturer la plupart des effets cryogéniques. Par la suite, les paramètres des modèles adaptés aux effets cryogéniques peuvent être extraits. Les modèles sont la clé des procédures de simulation en décrivant de manière précise le comportement des dispositifs pour une large gamme continue de géométries, nécessaire pour concevoir des circuits intégrés.

Circuits intégrés et architecture pour le calcul quantique

Avec un kit de conception décrivant avec précision le comportement des circuits à des températures cryogéniques, le prochain défi pour les applications en calcul quantique est de concevoir une architecture capable d'exécuter des algorithmes quantiques surpassant les algorithmes classiques. Le contrôle et la lecture simultanés de milliers voir plus de millions de qubits posent de sérieuses contraintes sur les circuits intégrés de contrôle des unités de traitement quantique.

L'approche actuelle largement adoptée lors de la conception des architectures de contrôle suit une approche « dispositif quantique *first* », dans laquelle la partie quantique dicte les exigences pour le reste du système. Cette approche est le résultat du développement précoce des qubits qui manquent de flexibilité par rapport aux composants électroniques.

Des travaux récents ont démontré un contrôle et une lecture haute-fidélité de dispositifs à un et deux qubits avec des circuits intégrés cryogéniques fonctionnant dans les limites du budget de refroidissement des réfrigérateurs. Cependant, la forte consommation d'énergie par qubit de quelques mW limite fortement le nombre de qubits contrôlables simultanément, affectant négativement les capacités algorithmiques. Le défi majeur lors de l'augmentation du nombre de qubits est le développement de circuits intégrés répondant aux exigences de précision et de bas bruit à une fraction de la puissance de refroidissement du réfrigérateur.

Plusieurs axes de recherche sont suivis pour résoudre ce problème : de la réduction de la consommation d'énergie grâce à une conception intelligente des circuits, à l'architecture électronique échelonnée en température ou aux améliorations des réfrigérateurs cryogéniques. La première approche est une amélioration purement axée sur les circuits intégrés dans laquelle la consommation d'énergie est minimisée par une conception de système soignée et méticuleuse. Par exemple, l'utilisation de technologies bipolaires pour les circuits à haute vitesse et à bas bruit pourraient être intégrés avec des technologies nanométriques modernes telles que le FD-SOI ou le FinFET pour une puissance de traitement numérique à faible puissance. D'autres architectures électroniques proposées tirent parti de la puissance de refroidissement plus élevée des réfrigérateurs à des étages en température plus élevés tels qu'a 77 K ou même à 300 K. La consommation électrique totale est séparée entre différents budgets de puissance de refroidissement afin de réduire la charge thermique sur les étages cryogéniques, plus critiques. Cette approche nécessite une conception soignée et une prise en compte des effets d'entonnoir des E/S entre chaque étage de température en raison du nombre limité de connexions. Finalement, quelques fabricants de réfrigérateurs à dilution travaillent déjà sur l'augmentation de la puissance de refroidissement à 4.2 K avec, par exemple, des tubes à impulsions plus puissants. Cette augmentation de la puissance de refroidissement permettra d'installer plus d'électronique, augmentant ainsi le nombre de qubits actifs.

Les architectures étagées en température reposent sur la co-optimisation de parties « hardware » très différentes : le système de refroidissement et le système de contrôle. Ces deux pièces ne sont plus conçues séparément mais considérées comme un grand problème d'ingénierie afin d'atteindre de nouveaux sommets de performance. Cette optimisation interdisciplinaire simultanée pourrait être endémique des développements futurs des unités de traitement quantique au cours des prochaines décennies, lorsque les dispositifs quantiques seront mieux compris et que la conception de qubits sera un domaine de recherche étendu et prolifique. La transition d'une approche quantique à une approche système dans laquelle chaque composant est d'égale importance permettra de contrôler un grand nombre de qubits avec les technologies disponibles au moment de la conception. Déjà aujourd'hui, nous pouvons voir de nouveaux types de qubits avec des qubits de spin fonctionnant à hautes températures au-dessus de 1 K ou de nouveaux qubits supraconducteurs à basses fréquences se rapprochant rapidement de la qualité des célèbres qubits « transmon ». Bientôt, le type de qubit et leurs diverses propriétés entreront dans la boucle de conception pour permettre le calcul quantique à grande échelle. Les futurs architectes d'ordinateurs quantiques devront co-optimiser tous les composants de l'unité de calcul quantique : de la cryogénie à l'électronique, en passant

par les algorithmes et les dispositifs quantiques.

Bibliography

[1] J. C. Bardin, E. Jeffrey, E. Lucero, T. Huang, O. Naaman, R. Barends, T. White, M. Giustina, D. Sank, P. Roushan, K. Arya, B. Chiaro, J. Kelly, J. Chen, B. Burkett, Y. Chen, A. Dunsworth, A. Fowler, B. Foxen, C. Gidney, R. Graff, P. Klimov, J. Mutus, M. McEwen, A. Megrant, M. Neeley, C. Neill, C. Quintana, A. Vainsencher, H. Neven, and J. Martinis. **A 28nm Bulk-CMOS 4-to-8GHz <2mW Cryogenic Pulse Modulator for Scalable Quantum Computing**. In *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, pages 456–458, February 2019.

[2] Bishnu Patra, Jeroen P. G. van Dijk, Sushil Subramanian, Andrea Corna, Xiao Xue, Charles Jeon, Farhana Sheikh, Esdras Juarez-Hernandez, Brando Perez Esparza, Huzaifa Rampurawala, Brent Carlton, Nodar Samkharadze, Surej Ravikumar, Carlos Nieva, Sungwon Kim, Hyung-Jin Lee, Amir Sammak, Giordano Scappucci, Menno Veldhorst, Lieven M. K. Vandersypen, Masoud Babaie, Fabio Sebastiano, Edoardo Charbon, and Stefano Pellerano. **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**. In *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, pages 304–306, 2020.

[3] J.-S. Park, S. Subramanian, L. Lampert, T. Mladenov, I. Klotchkov, D. J. Kurian, E. Juarez-Hernandez, B. Perez-Esparza, S. R. Kale, K. T. Asma Beevi, S. Premaratne, T. Watson, S. Suzuki, M. Rahman, J. B. Timbadiya, S. Soni, and S. Pellerano. **A Fully Integrated Cryo-CMOS SoC for Qubit Control in Quantum Computers Capable of State Manipulation, Readout and High-Speed Gate Pulsing of Spin Qubits in Intel 22nm FFL FinFET Technology**. In *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, **64**, pages 208–210, February 2021.

[4] A. Ruffino, Y. Peng, T.-Y. Yang, J. Michniewicz, M. F. Gonzalez-Zalba, and E. Charbon. **A Fully-Integrated 40-nm 5-6.5 GHz Cryo-CMOS System-on-Chip with I/Q Receiver and Frequency Synthesizer for Scalable Multiplexed Readout of Quantum Dots**. In *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, **64**, pages 210–212, February 2021.

[5] David J. Frank, Sudipto Chakraborty, Kevin Tien, Pat Rosno, Thomas Fox, Mark Yeck, Joseph A. Glick, Raphael Robertazzi, Ray Richetta, John F. Bulzacchelli, Daniel Ramirez, Dereje Yilma, Andrew Davies, Rajiv V. Joshi, Shawn D. Chambers, Scott Lekuch, Ken Inoue, Devin Underwood, Dorothy Wisnieff, Chris Baks, Donald Bethune, John Timmerwilke, Blake R. Johnson, Brian P. Gaucher, and Daniel J. Friedman. **A Cryo-CMOS Low-Power Semi-Autonomous Qubit State Controller in 14nm FinFET Technology**. In *2022 IEEE International Solid- State Circuits Conference (ISSCC)*, **65**, pages 360–362, February 2022.

APPENDIX A

FD-SOI Devices

I Integrated Passive Resistors down to 4.2 K

We independently measured passive resistors within the CMOS FD-SOI 28 nm technology from 300 to 4.2 K. Each resistor is wire-bonded to a sample holder placed inside a Physical Property Measurement System, made to measure device resistivity as a function of temperature and magnetic field. Results for different types of resistors are shown in Figure A.1.

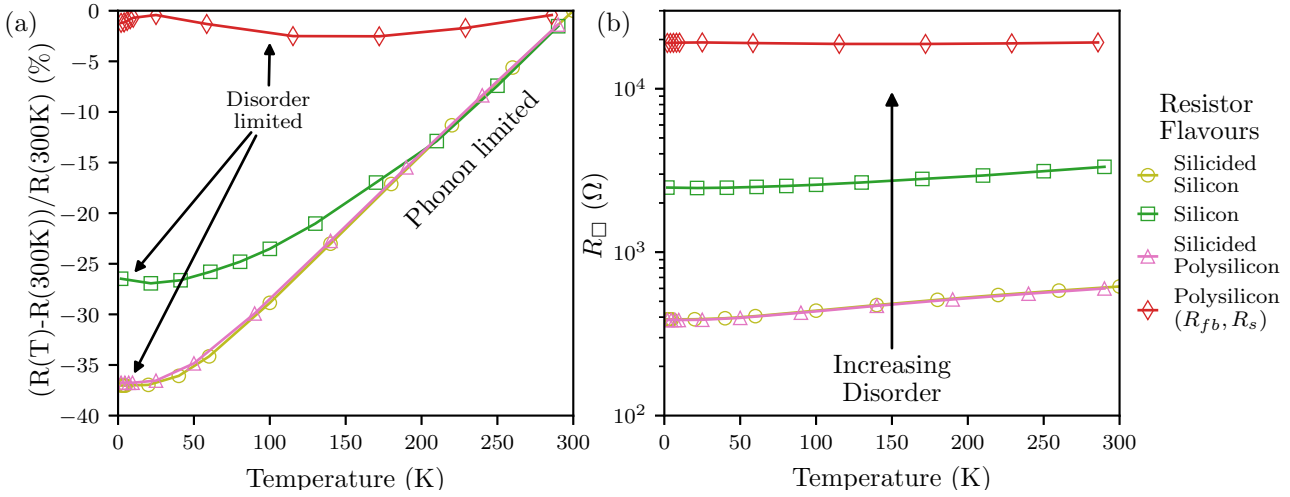


Figure A.1 | Behavior of integrated resistors of the FDSOI 28nm technology from 300 to 4.2 K.
(a) Relative resistance variation and (b) sheet resistance variation from 300 to 4.2 K for silicon, silicided silicon, polysilicon, and silicided polysilicon. Silicon, silicided silicon, and silicided polysilicon show a strong dependence on temperature related to the suppression of phonon scattering with a low residual sheet resistance. Polysilicon resistance only varies by 4% and has the largest sheet resistance due to its high disorder.

II Transistors

Transistors within the CMOS FD-SOI 28 nm technology are studied at cryogenic temperatures by embedding thousands of them in a single multiplexing matrix for quick low-frequency characterizations (see Chapter 2). Several matrices have been measured and a total count accounting for all the measured thin- (respectively thick-) oxide devices is shown in Figure A.2 (resp. in Figure A.3). We measured I-V curves as a function of all bias voltages including gate, source, drain, and back-gate as shown in Figure A.4 (respectively in Figure A.5) for N-type (resp. P-type) devices.

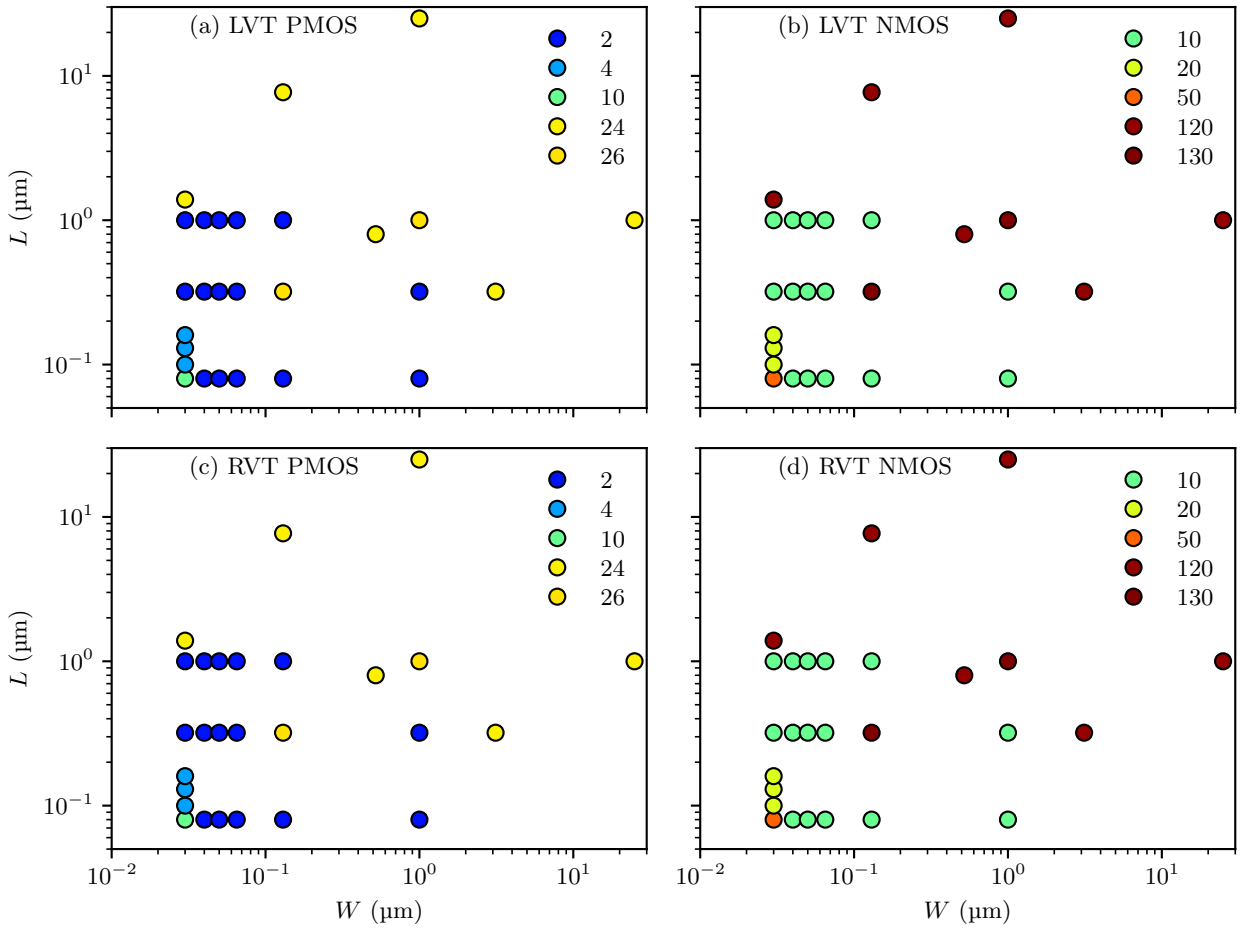


Figure A.2 | Measured thin-oxide transistors at 4.2 K.

Representation of the total number of measured thin-oxide (GO1) devices at 4.2 K for each length L and width W. The number of devices n for a given geometry is made of $n/2$ mismatch pairs. All flavors are represented with (a) LVT P-type, (b) LVT N-type, (c) RVT P-type, and (d) RVT N-type. In chapter 2, we focus on the LVT devices.

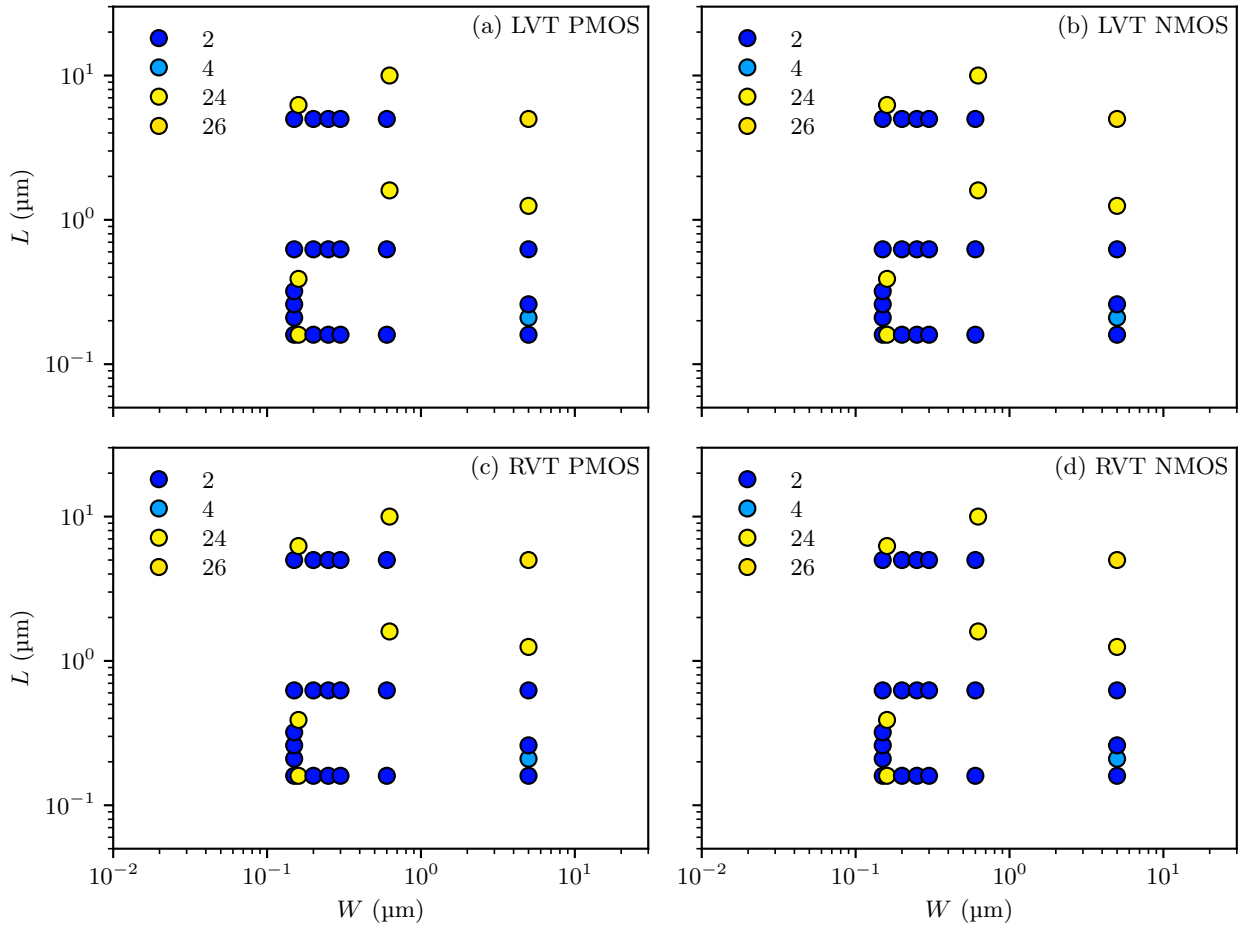


Figure A.3 | Measured thick-oxide transistors at 4.2 K.

Representation of the total number of measured thick-oxide (GO2) devices at 4.2 K for each length L and width W. The number of devices n for a given geometry is made of $n/2$ mismatch pairs. All flavors are represented with (a) LVT P-type, (b) LVT N-type, (c) RVT P-type, and (d) RVT N-type. Fewer thick-oxide devices were measured as we envision to use thin-oxide devices for most of the low-power and high-speed circuits.

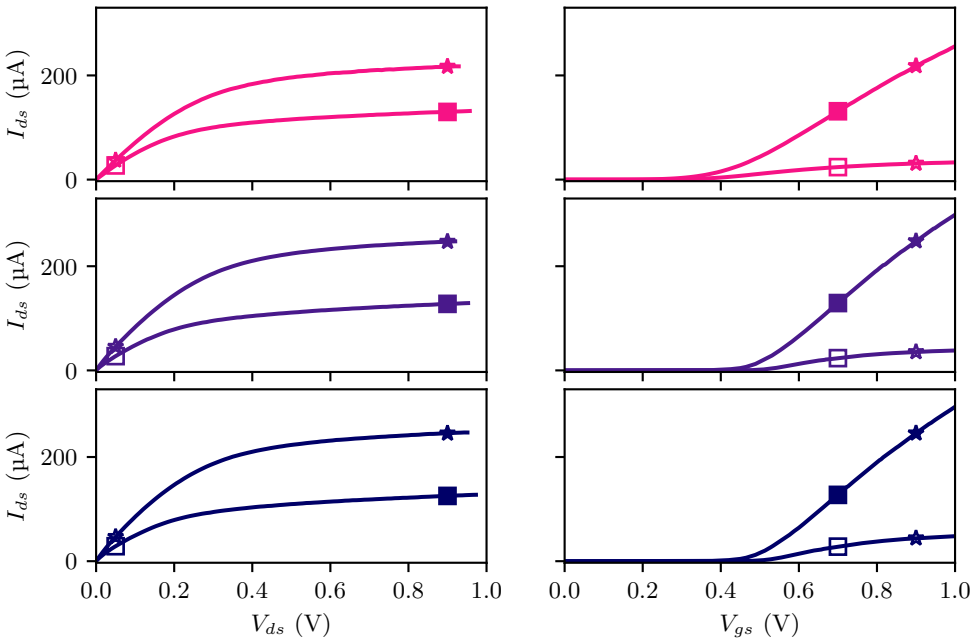


Figure A.4 | Typical current-voltage characteristics of N-type devices at 300, 4.2 and 0.1 K. (a-c) Source-drain current I_{ds} increase with the source-drain voltage V_{ds} for V_{gs} of 0.7 and 0.9 V at the indicated temperatures. (d-f) Source-drain current I_{ds} increase with the gate-source voltage V_{gs} for V_{ds} of 50 mV and 0.9 V at the indicated temperatures. The correspondence between (a-c) and (d-f) is shown with markers: empty markers for the linear regime $V_{ds} < V_{gs} - V_{th}$ and full markers for the saturation regime $V_{ds} > V_{gs} - V_{th}$. Square markers (respectively star markers) corresponds to V_{gs} of 0.7 V (resp. 0.9 V).

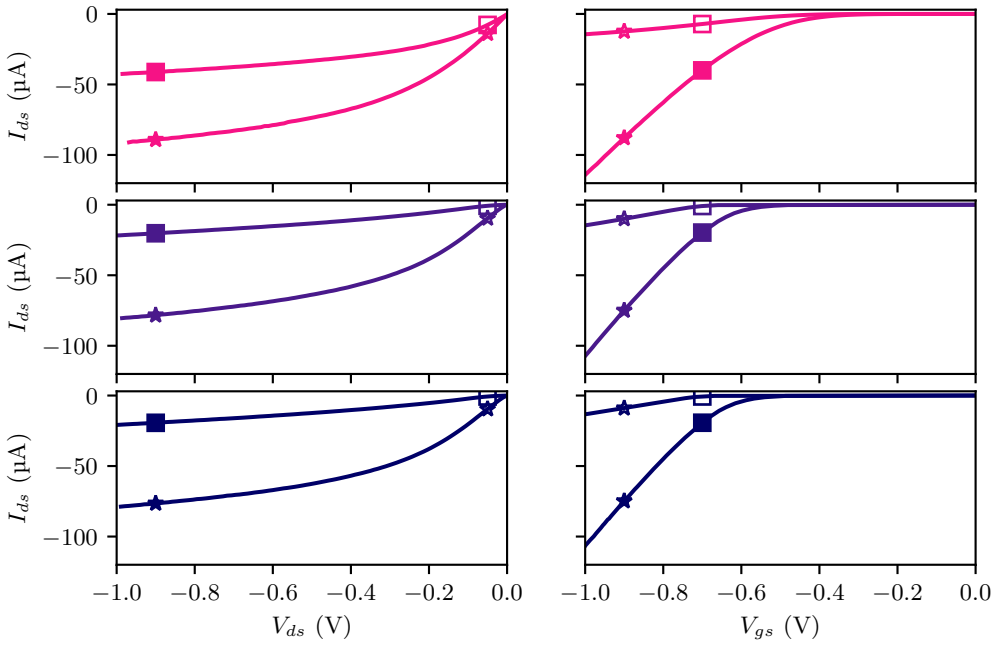


Figure A.5 | Typical current-voltage characteristics of P-type devices at 300, 4.2, and 0.1 K. (a-c) Source-drain current I_{ds} with the source-drain voltage V_{ds} for V_{gs} of -0.7 and -0.9 V at the indicated temperatures. (d-f) Source-drain current I_{ds} increase with the gate-source voltage V_{gs} for V_{ds} of -50 mV and -0.9 V at the indicated temperatures. The correspondence between (a-c) and (d-f) is shown with markers: empty markers for the linear regime $V_{ds} < V_{gs} - V_{th}$ and full markers for the saturation regime $V_{ds} > V_{gs} - V_{th}$. Square markers (respectively star markers) corresponds to V_{gs} of $-0.7\,\mathrm{V}$ (resp. $-0.9\,\mathrm{V}$).

In Chapter 2, we extracted single transistor quantities from 300 to 4.2 K such as the threshold voltage (Figure A.6), the channel resistance (Figure A.7), the transconductance (Figure A.8), and the conductance (see Figure A.9). Typical curves and the extraction method are presented the associated figures.

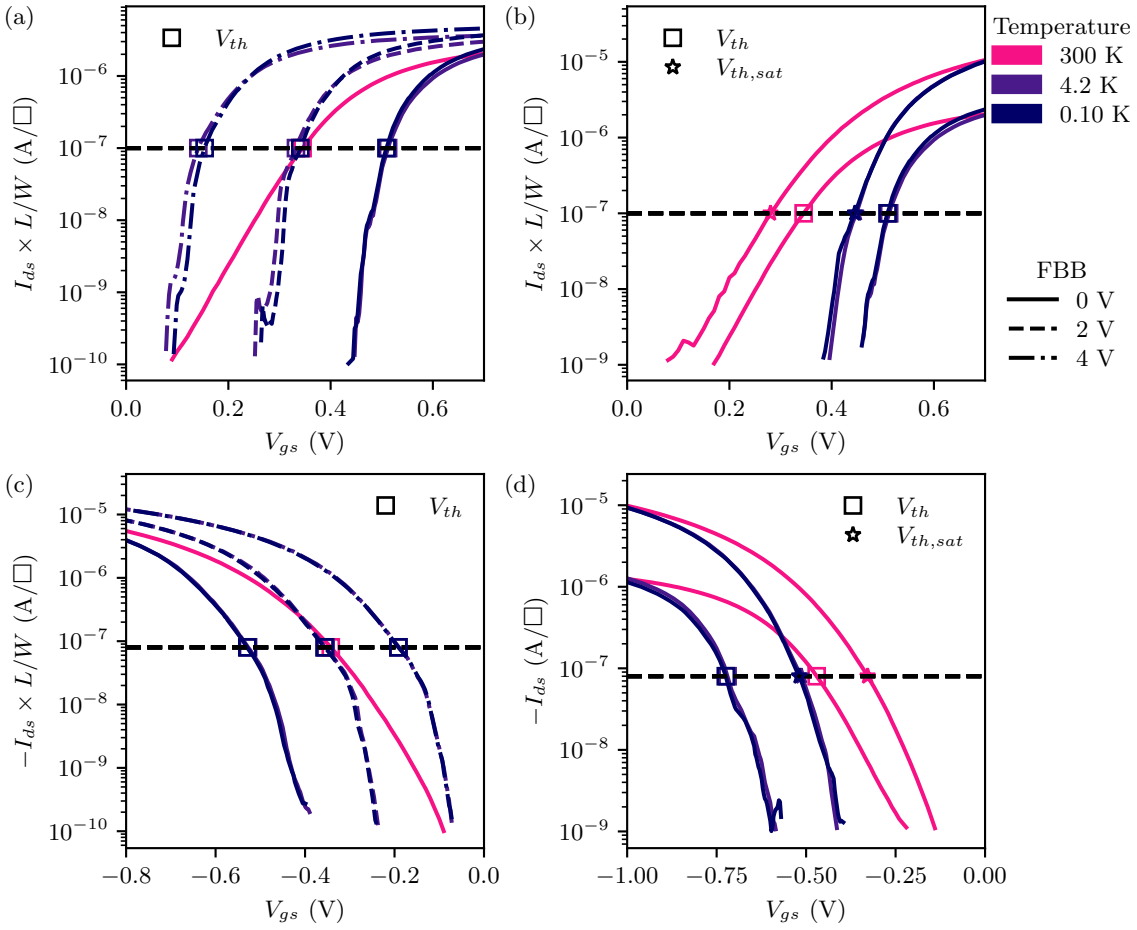


Figure A.6 | Threshold voltage extraction.

Log-plot of the drain-source current $|I_{ds}| \times L/W$ in linear regime $|V_{ds}| = 50\,\text{mV}$ at 300 and 4.2 K for several FBB voltages for (a) N-type and (b) P-type devices. The threshold voltage V_{th} in linear regime is extracted at constant current density: $10^{-7}\,\text{A}/\square$ for NMOS and $0.8 \times 10^{-7}\,\text{A}\square$ for PMOS. Log-plot of the drain-source current $|I_{ds}| \times L/W$ for FBB of 0 V in linear regime $|V_{ds}| = 50\,\text{mV}$ and in saturation $|V_{ds}| = 0.9\,\text{V}$ at 4.2 and 300 K for (b) NMOS and (d) PMOS. The threshold voltage in saturation $V_{th,sat}$ is significantly lower than V_{th} in linear regime due to drain-induced barrier lowering effects in short-channel devices.

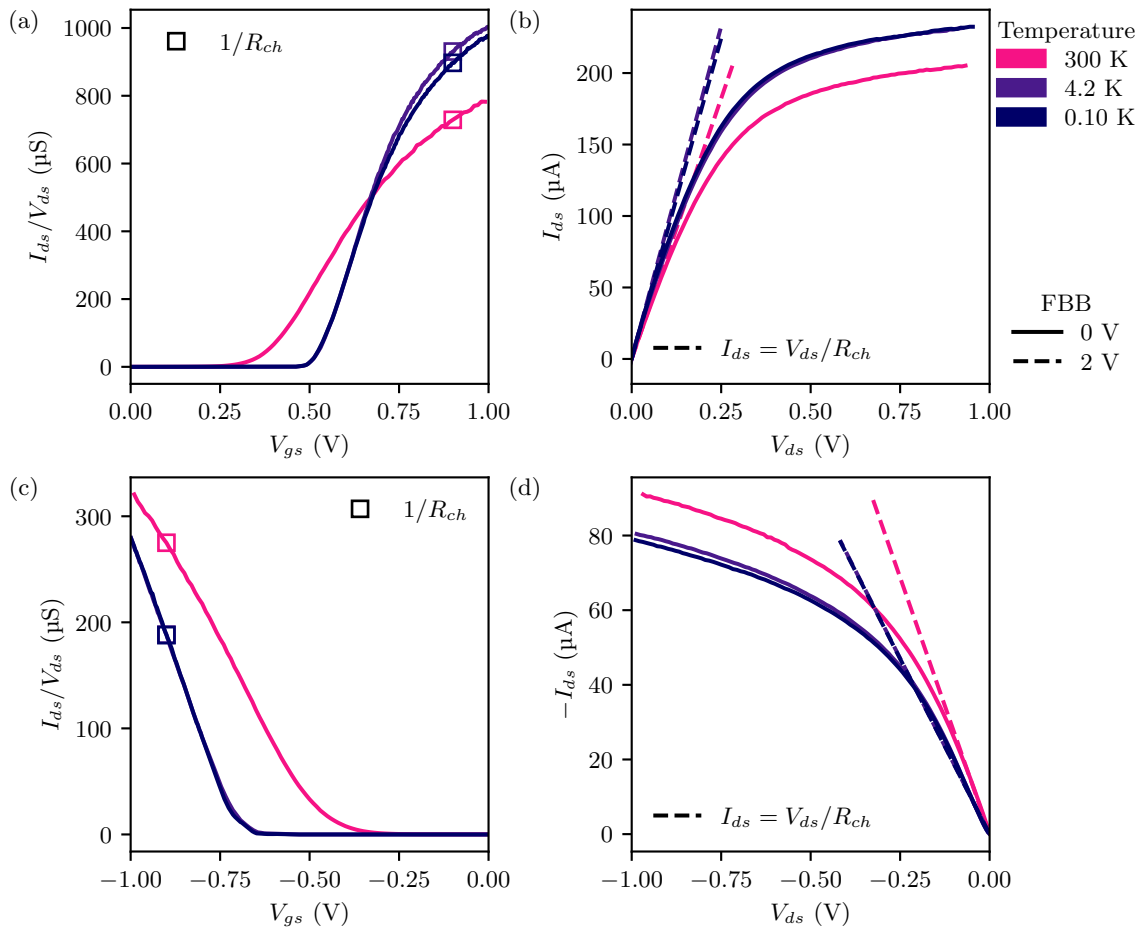


Figure A.7 | Channel resistance extraction.

Channel conductance I_{ds}/V_{ds} evolution with the gate-source voltage V_{gs} for (a) NMOS and (b) PMOS in linear regime $|V_{ds}| = 50\,\text{mV}$ at 300, 4.2, and 0.1 K. The channel resistance $R_{ch} = V_{ds}/I_{ds}$ is computed from the channel conductance in strong inversion at $|V_{gs}| = 0.9\,\text{V}$ indicated by the square markers. The extracted R_{ch} gives the slope in linear regime of the source-drain current I_{ds} as a function of drain-source voltage V_{ds} as shown by the dashed lines for (b) NMOS and (c) PMOS.

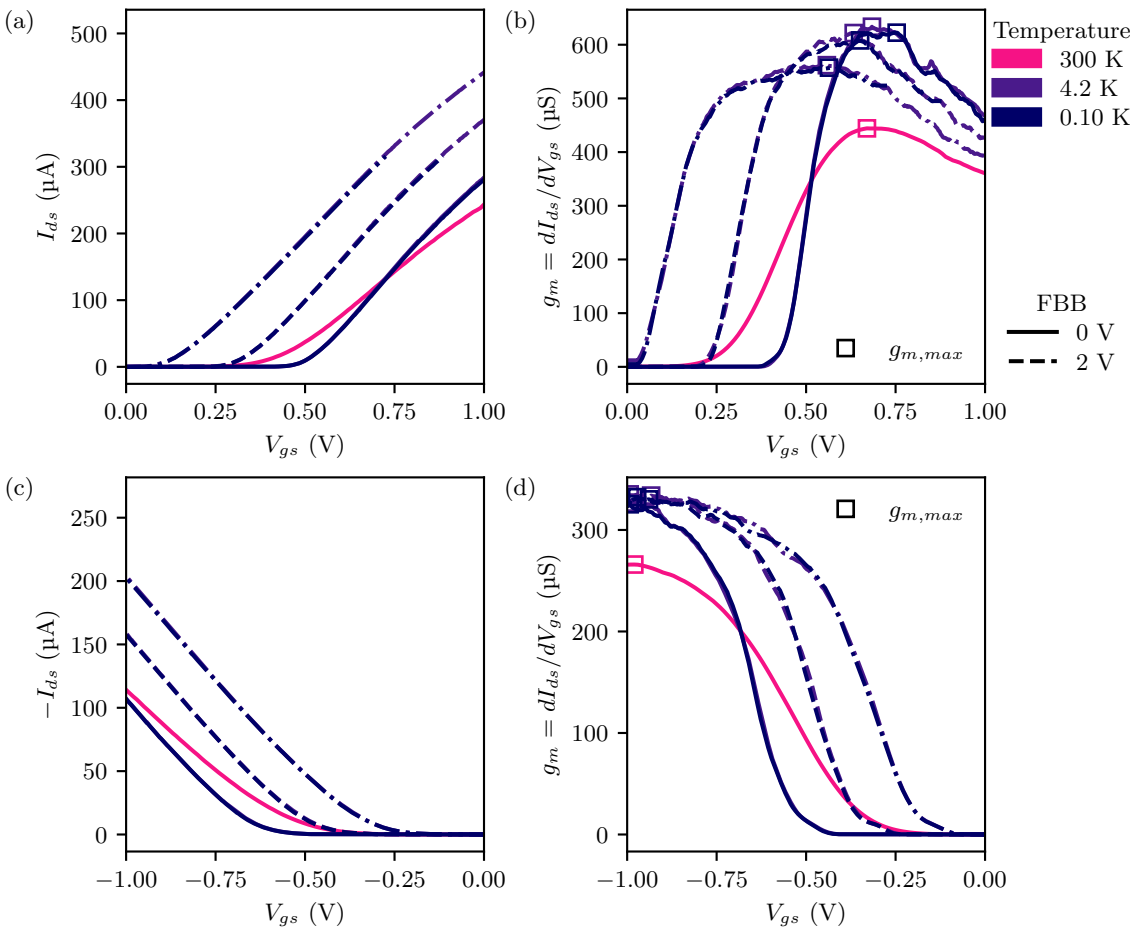


Figure A.8 | Transconductance extraction.

Linear plot of the drain-source current I_{ds} versus the gate-source voltage V_{gs} in saturation with $|V_{ds}| = 0.9\,\mathrm{V}$ at 300 and 4.2 K and several FBB voltages for (a) NMOS and (c) PMOS. The transconductance $g_m = dI_{ds}/dV_{gs}$ is computed from the curves in (a) and (c) with the Savitzky-Golay differentiation and is shown as a function of V_{gs} for (b) NMOS and (d) PMOS. The transconductance peaks at a given V_{gs} at which we extract the maximum transconductance $g_{m,max}$ shown with the square markers.

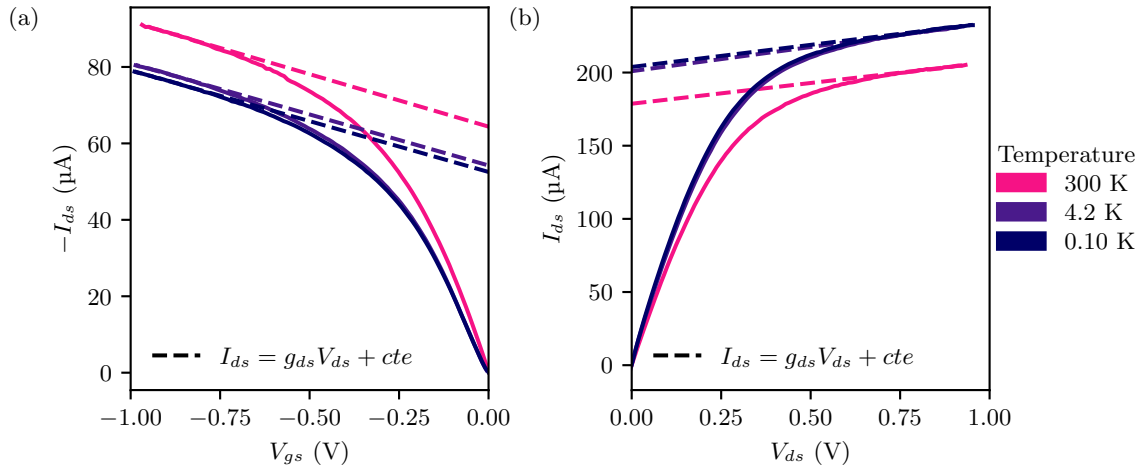


Figure A.9 | Conductance extraction.

Drain-source current I_{ds} versus the drain-source voltage V_{ds} for temperature from 300 to 0.1 K at FBB of 0 V for (a) PMOS and (b) NMOS. The conductance in saturation g_{ds} is extracted by fitting $I_{ds}(V_{ds})$ with a linear function using least-squares optimization for $|V_{gs} > 0.8\,\mathrm{V}$. The linear fit is shown with the dashed lines.

III Details on subband scattering

During the measurement of single devices at cryogenic temperatures under a strong FBB, we noticed the appearance of negative transconductance in some large-area devices, predominantly N-type devices. Typical $I_{ds}(V_{gs})$ curves at low and strong source-drain voltage exhibiting a clear negative transconductance at FBB of 4 V are shown in Figure A.10a. All these devices have an area greater than $0.5\,\mu m^2$ as illustrated by the circle marker in Figure A.10b representing the device width and length showing negative transconductance. Little humps in the transconductance appear for smaller devices (square markers in Figure A.10) as shown in the $I_{ds}(V_{gs})$ curves at 4.2 K and FBB of 4 V plotted in Figure A.10c. As the devices get smaller, the hump reduces and disappears.

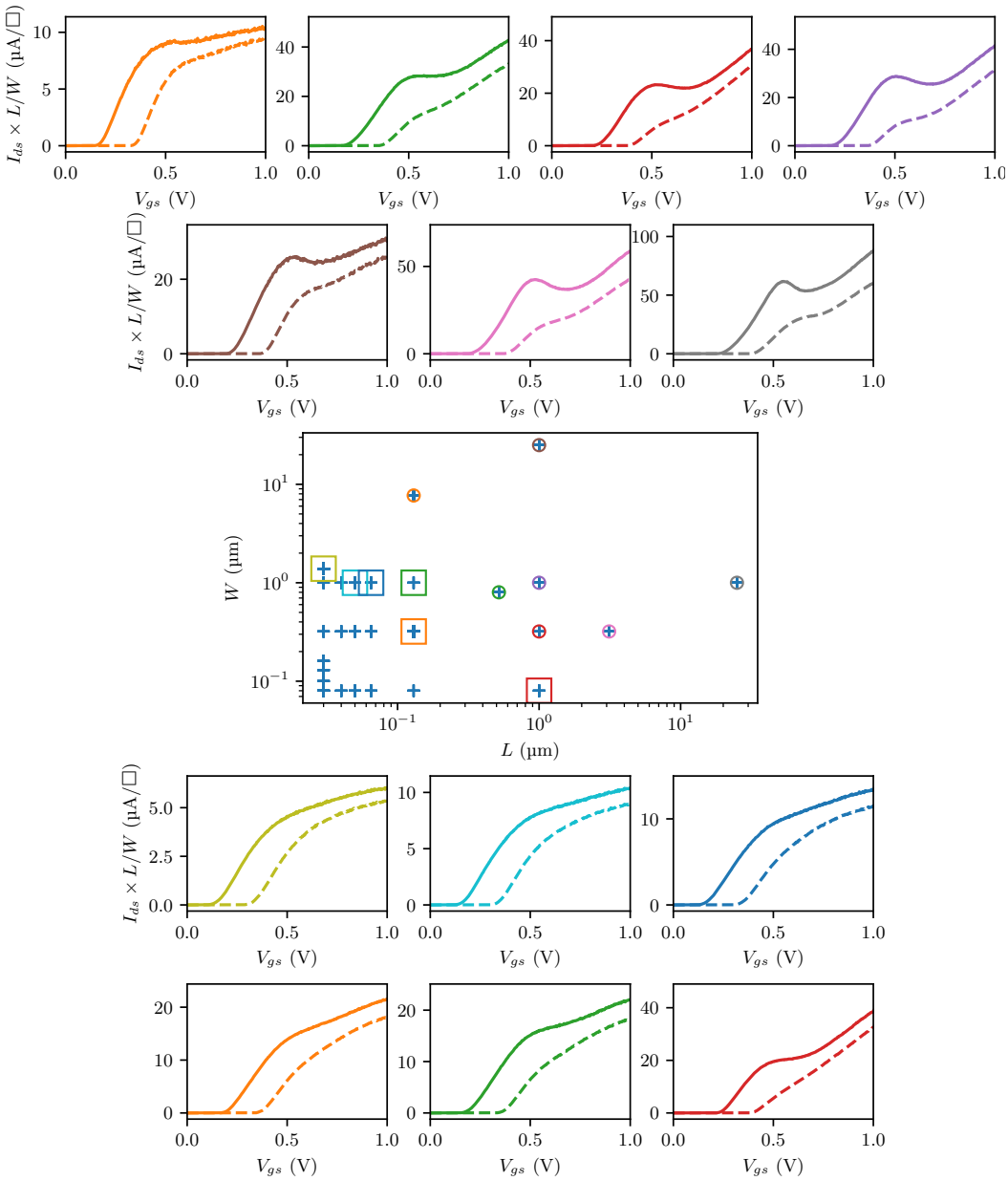


Figure A.10 | Negative transconductance at 4.2 K and high FBB.

(a) $I_{ds}(V_{gs})$ curves normalized to the geometry at $V_{ds} = 0.05\,\mathrm{mV}$ (solid line) and $0.9\,\mathrm{V}$ (dashed line) for all measured devices at 4.2 K and FBB of 4 V showing a clear negative transconductance. The colors correspond to the length and width shown in circle markers in (b). (b) All measured thin-oxide device lengths and widths at 4.2 K and high FBB of 4 V. Circle markers show devices showing a clear negative transconductance. The associated colors correspond to the color in (a). Square markers show devices showing a hump but no clear negative transconductance. The associated colors correspond to the color in (c). (c) $I_{ds}(V_{gs})$ curves normalized to the geometry at $V_{ds} = 0.05\,\mathrm{mV}$ (solid line) and $0.9\,\mathrm{V}$ (dashed line) for all measured devices at 4.2 K and FBB of 4 V showing a hump but no clear negative values in the transconductance.

APPENDIX B
Complementary Data: Digital Circuits

I Level-Shifters

We designed level-shifters in the CMOS FD-SOI 28 nm technology with thin-oxide devices with the geometries shown in Table B.1. Typical curves of the output voltage V_{out} as a function of the low output supply voltage V_{DDL} when the input is high are shown in Figure B.1 at 300 and 4.2 K for various FBB voltages.

Table B.1 | Sizing of the level-shifters.

Device geometries and flavors used in the realization of the two level shifters LS^{10}_{28} and LS^{10}_{130} studied in Chapter 3.

	LS^{10}_{28}		LS^{10}_{130}	
	N1,N2	P1, P2	N1, N2	P1, P2
Type	NMOS	PMOS	NMOS	PMOS
Oxide	Thin		Thin	
V_{th} flavor	RVT		RVT	
Length (µm)	0.028		0.13	
Width (µm)	0.3	0.6	1.3	1.3

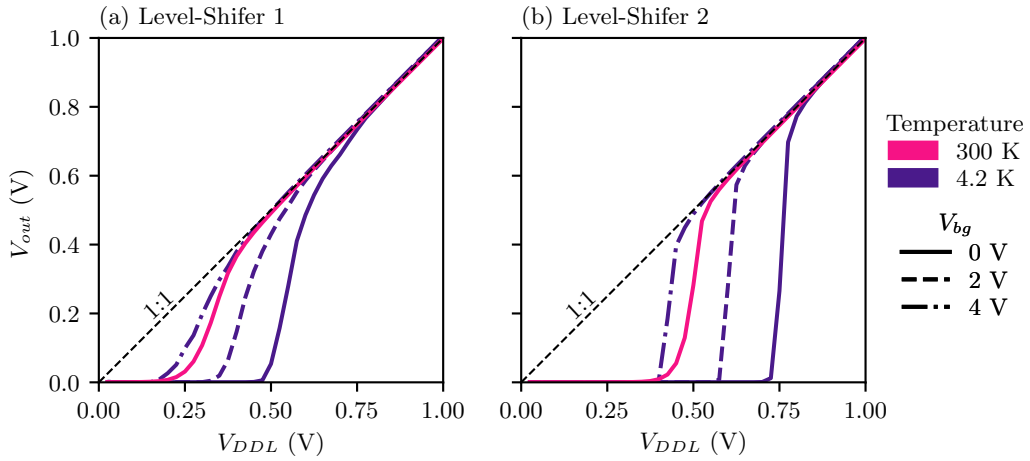


Figure B.1 | Output of level-shifters to a logic-1 at 300 and 4.2 K with positive FBB. Characteristics curves $V_{out}(V_{DDL})$ of the two level-shifters.

II Ring Oscillators & Frequency Dividers

We designed 3 ring-oscillators (RO) RO^{10}_{28}, RO^{10}_{130}, and RO^{5}_{130} made of 5 inverting stages whom one is a NAND gate for enabling oscillations (see Figure B.2ab for the implementation and Table B.2 for the device geometries). We measured the 3 ROs outputting GHz-speed signals by appending a frequency divider made of 10 dynamic flip-flop to reduce the frequency by 1024 (see Figure B.3ab). Typical curves $f(V_{co})$ of the output frequency as a function of the control voltage are shown in Figure B.4.

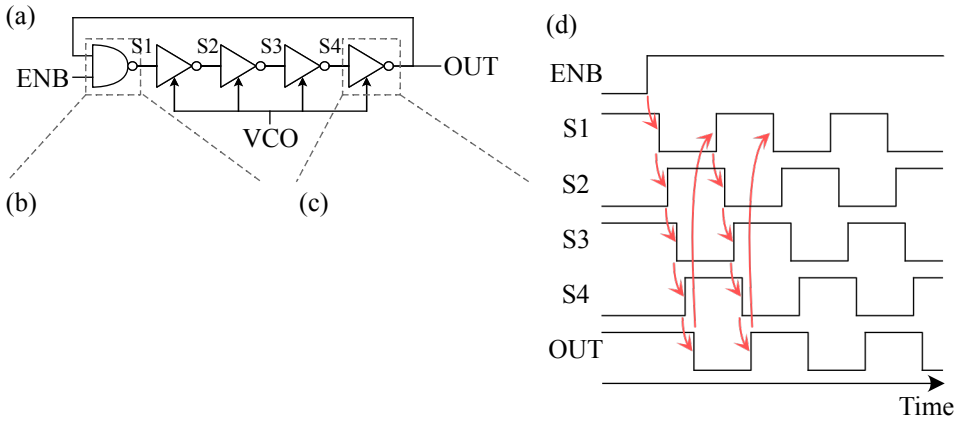


Figure B.2 | Ring-oscillators.

(a) 5-stage voltage-controlled ring oscillator with a NAND gate for enable. (b) NAND gate with equivalent capacitive load than inverting stages. (c) Current-limited inverters to control the RO frequency.

Table B.2 | Sizing of the ring-oscillators.

		Inverter		NAND gate		
		N1	P1	Nx	Px	N0
RO^{10}_{28}	Type	NMOS	PMOS	NMOS	PMOS	NMOS
	Oxide	Thin		Thin		Thin
	V_{th} flavor	RVT		RVT		RVT
	Length (µm)	0.028		0.028		0.028
	Width (µm)	0.3	0.6	0.3	0.6	xx
RO^{10}_{130}	Type	NMOS	PMOS	NMOS	PMOS	NMOS
	Oxide	Thin		Thin		Thin
	V_{th} flavor	RVT		RVT		RVT
	Length (µm)	0.13		0.13		0.13
	Width (µm)	1.3	1.3	1.3	1.3	xx
RO^{5}_{130}	Type	NMOS	PMOS	NMOS	PMOS	NMOS
	Oxide	Thin		Thin		Thin
	V_{th} flavor	RVT		RVT		RVT
	Length (µm)	0.13		0.13		0.13
	Width (µm)	0.65	0.65	0.65	0.65	xx

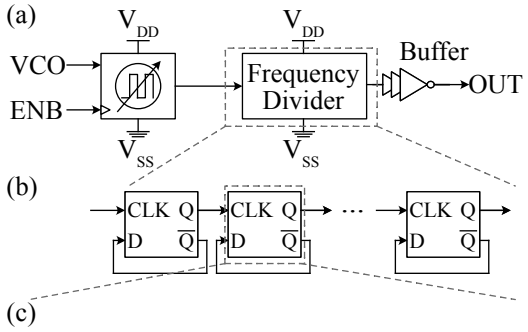


Figure B.3 | Frequency divider.

(a) Frequency division by 1024 and buffering of signals to load the cable capacitance. (b) Frequency divider made of 10 dynamic flip-flops and its implementation (c). (d) Buffer output received at room-temperature for a RO frequency of 10 GHz.

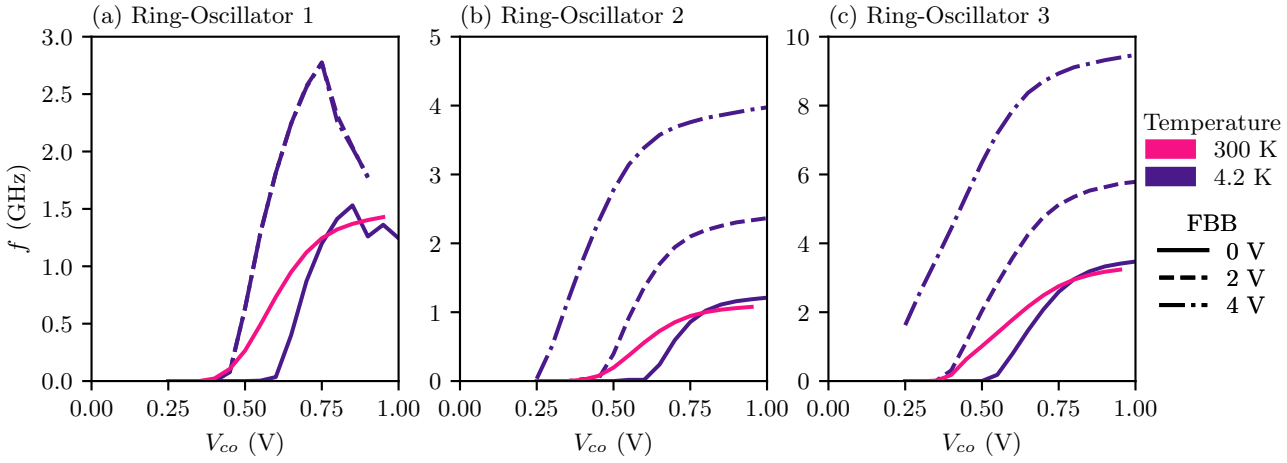


Figure B.4 | Output frequency of ring oscillators at 300 and 4.2 K with positive FBB.

Output frequency f of (a) RO^{10}_{130}, (b) RO^{5}_{130}, and c RO^{10}_{28} as a function of the control voltage V_{co} at 300 and 4.2 K for FBB of 0, 2, and 4 V. The frequency decrease seen in (a) corresponds to the frequency divider made of long transistors failing at high frequency. (b) and (c) use minimal 28-nm long transistors in the frequency divider.

APPENDIX C
Complementary Data: Analog Circuits

I Pass-gates

We designed 2 pass-gate circuits PG^{10}_{130} and PG^{13}_{150} (also known as transmission gates or switches) with both thin and thick- oxide devices and with geometries shown in Table C.1. We measured the small-signal resistance of the pass-gate when enabled as a function of the DC voltage V_{cm} shown in Figure C.1 at 300 and 4.2 K at different values of FBB.

Table C.1 | MOSFETs in the analog pass-gates.

Transistor sizes and types in the two measured pass-gates PG^{130}_{10} and PG^{150}_{13}.

	PG^{10}_{130}		PG^{13}_{150}	
	N0	P0	N0	P0
Type	NMOS	PMOS	NMOS	PMOS
Oxide	Thin		Thick	
V_{th} flavor	RVT		RVT	
Length (µm)	0.13		0.15	
Width (µm)	3.6	9	3.8	11.4

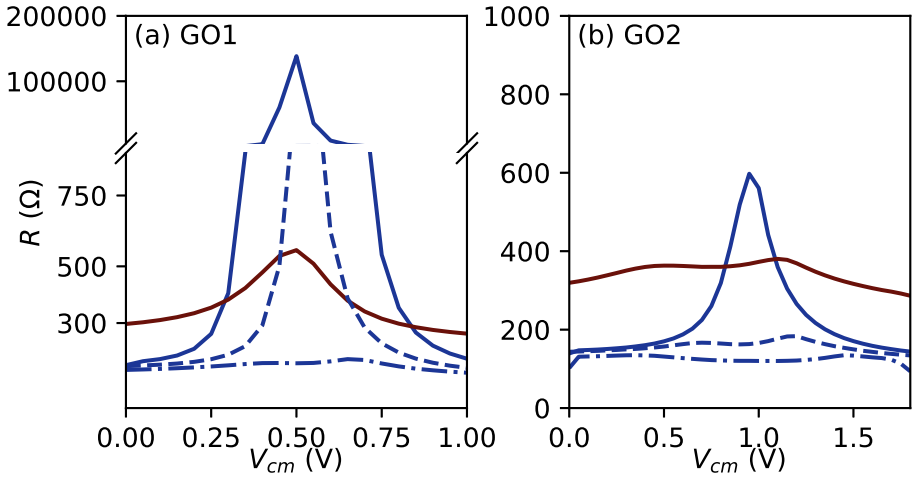


Figure C.1 | Small-signal pass-gate resistance at 300 and 4.2 K at different values of FBB.

Small-signal resistance of (a) PG^{10}_{130} and (b) PG^{13}_{150} as a function of input voltage $V_{in} = V_{cm}$ from 0 V to $V_{DD} = 1$ V at FBB of 0, 2, and 4 V.

II Digital-to-Analog Converters

Marcos Zurita (postdoc at LETI) designed two digital-to-analog converters $DAC_{fine}$and DAC_{coarse}within the CMOS FD-SOI 28 nm technology with thick-oxide devices. The output voltage as a function the input code is shown in Figure C.2ab at 300 and 4.2 K without any FBB. The step size between consecutive codes is plotted in Figure C.2.

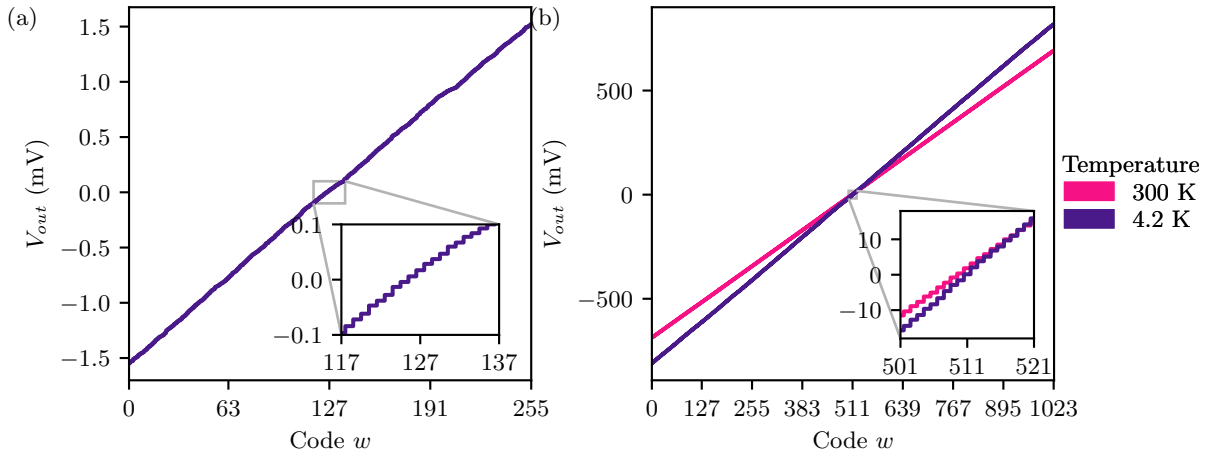


Figure C.2 | Output voltage of the fine and coarse DACs at 300 and 4.2 K.

Output voltage of (a) DAC_{fine} and (b) DAC_{coarse} as a function of the input digital code measured at 300 and 4.2 K.

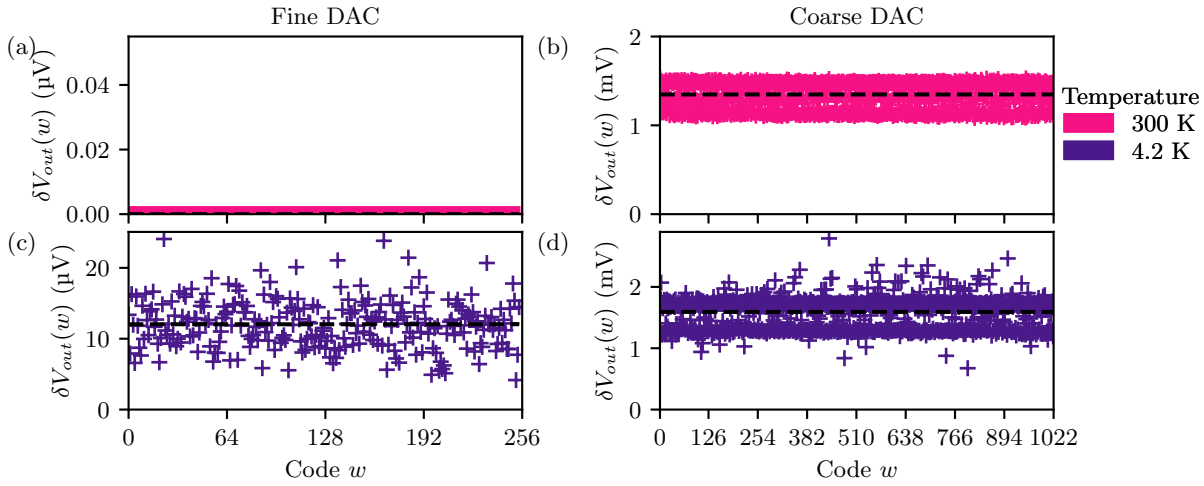


Figure C.3 | Output voltage step of the fine and coarse DACs at 300 and 4.2 K.

Voltage step between successive input code for respectively the $\mathrm{DAC_{fine}}$ at (c) 4.2 K (resp. $\mathrm{DAC_{coarse}}$ at (b) 300 and (d) 4.2 K). Data for the $\mathrm{DAC_{fine}}$ at 300 K in (a) are unavailable.

APPENDIX D
Quantum Co- and Full Integration with the Cryogenic Transimpedance Amplifier

I Simulations from Foundry Models at 300 K

To maintain a low impact of the measuring TIA on the nearby quantum-dot device, we estimate the voltage noise seen by the DUT by integrating the noise at the TIA input on a large frequency range for different scenarios of parasitic capacitance from interconnections (see Table D.1).

Table D.1 | Kick-back voltage noise.

Evaluated kick-back noise on the measured quantum device evaluated from foundry models noise data at 300 K. The noise density is integrated from 10 Hz to 1 GHz.

$C_{in} = 3$ pF	$C_{out} = 100$ pF	Kick-back noise (μV_{rms})
no	no	350
no	yes	58
yes	no	36
yes	yes	32

II Experimental Setup & Measurements

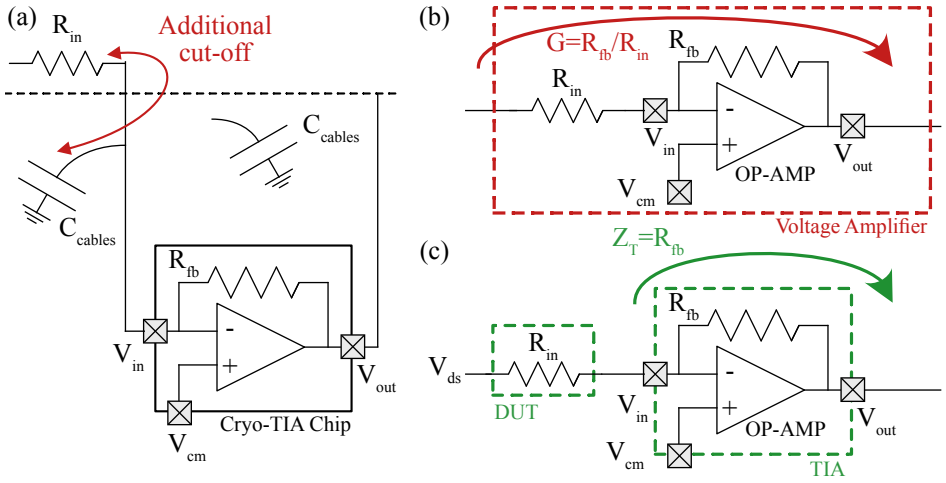


Figure D.1 | Circuit schematics of different measuring configurations of the TIA and OP-AMP.
(a) Measurement configuration of the TIA and OP-AMP placed at cryogenic temperatures with a room-temperature resistor R_{in}. For high values of R_{in} of a few MΩ, the cable capacitance from R_{in} to the TIA adds a low-frequency cut-off, preventing the bandwidth measurement of the cryogenic circuit. Placing R_{in} at cryogenic temperatures drastically reduces C_{cables}, moving the cut-off to higher frequencies, allowing the bandwidth measurement. (b) Voltage amplifier configuration by placing the TIA chip in series with an off-chip resistor R_{in}. (c) TIA configuration when R_{in} is used to generate a current to the TIA. In this configuration, the off-chip resistor R_{in} acts as the Device Under Test.

III Measurement data of the TIA and OP-AMP down to cryogenic temperatures

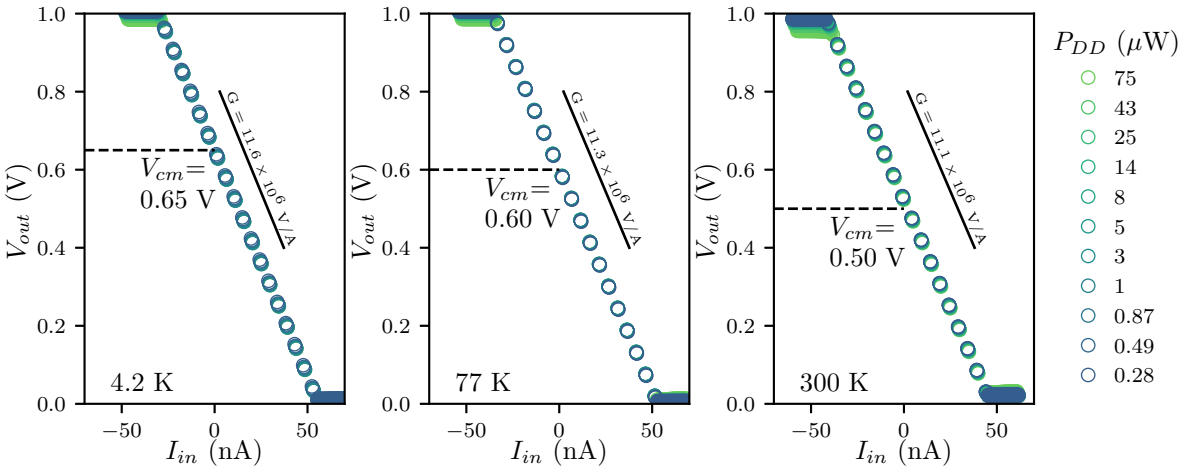


Figure D.2 | Cryo-TIA output characteristics from 300 to 4.2 K.

Linear response of the TIA as a function of input current I_{in}. Almost-constant transimpedance gain is conserved at 300, 77, and 4.2 K for power varying from 0.28 to 75 μW. V_{cm} is increased at low temperatures to compensate for increasing threshold voltages thus increasing the asymmetry in the measured current range.

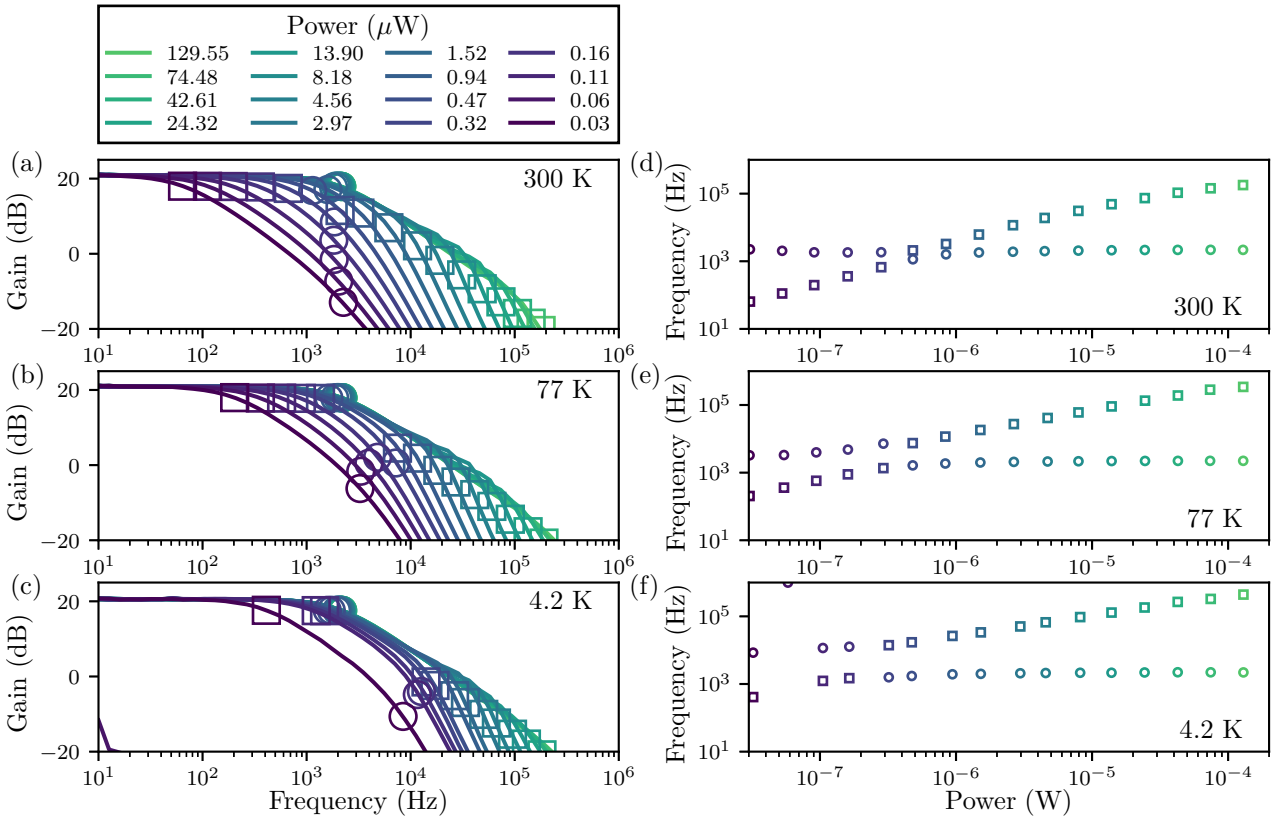


Figure D.3 | Frequency response of the cryogenic TIA mounted as a gain-10 voltage amplifier.

(a,b,c) Frequency response of the voltage gain at 300, 77, and 4.2 K for different power consumption P_{DD} of the OP-AMP mounted as a voltage amplifier of gain 10 with 1 MΩ at the input. Two cut-off frequencies are identified by fitting the frequency response to a third order system, highlighted for each P_{DD} as circles for the first order cut-off and as squares for the second order cut-off. (d,e,f) Extracted cut-off frequencies at 300, 77, and 4.2 K as a function of power consumption P_{DD}. The first order cut-off (circles) is independent of P_{DD} and corresponds to $R_{fb}C_{fb}$. The second order cut-off (squared symbols) increases with increasing power and is attributed to the increase of the internal OP-AMP Miller cut-off frequency. The gain-bandwidth product is extracted by multiplying the latter cut-off by the voltage gain obtained from (a,b,c) at low frequencies.

IV Full integration of the TIA with a double-dot device

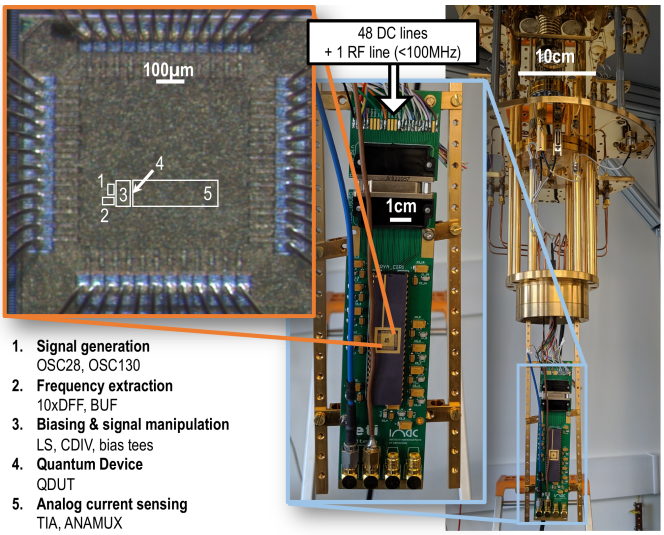


Figure D.4 | Experimental setup of the full integration experiment with the TIA.
Die micrograph, packaged chip mounting on a PCB with decoupling capacitors, and dilution-fridge mounting.

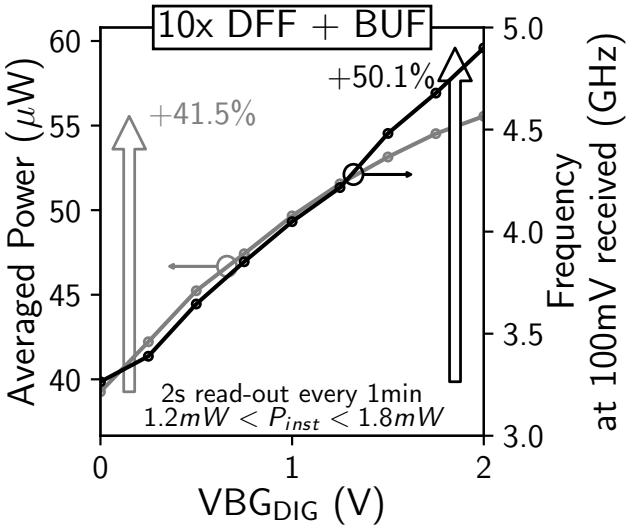


Figure D.5 | Power consumption of the frequency-measuring circuit.
Average power of the frequency divider by 1024 made of 10 dynamic flip flops followed by a digital buffer at 100 mK as a function of the digital back-gate voltage.

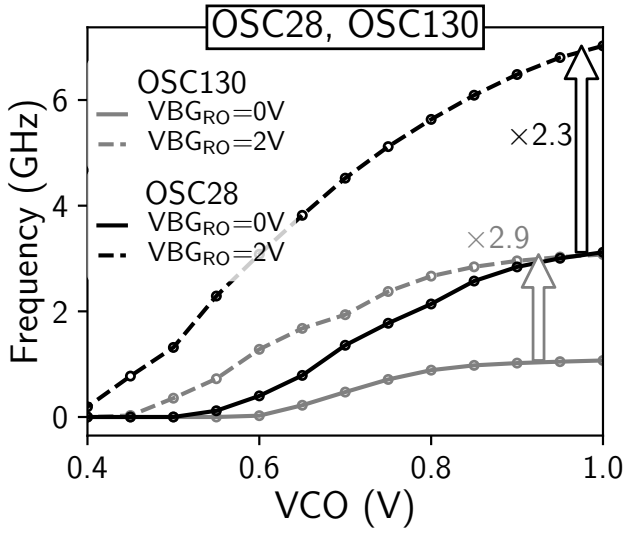


Figure D.6 | Ring-oscillator output frequency at 100 mK.
Output frequency of the ring oscillators RO^{10}_{28} and RO^{10}_{130} at 100 mK for FBB of 0 and 2 V.

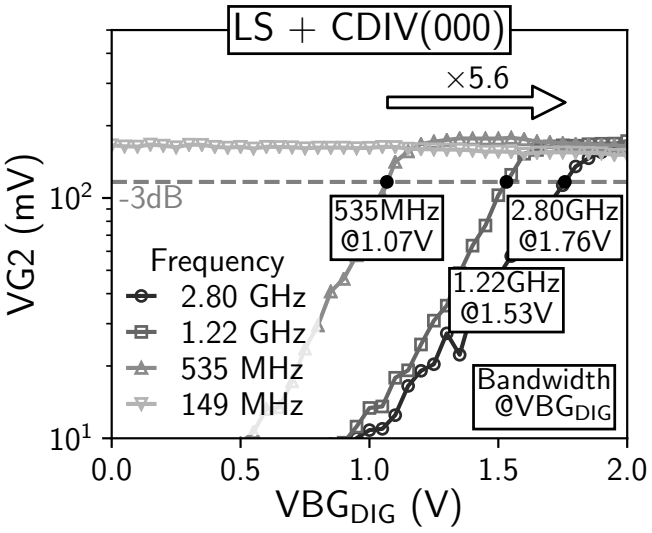


Figure D.7 | Quantum-dot GHz voltage excitation amplitude.
Voltage GHz excitation amplitude of the quantum dot gate as different excitation frequency and FBB.

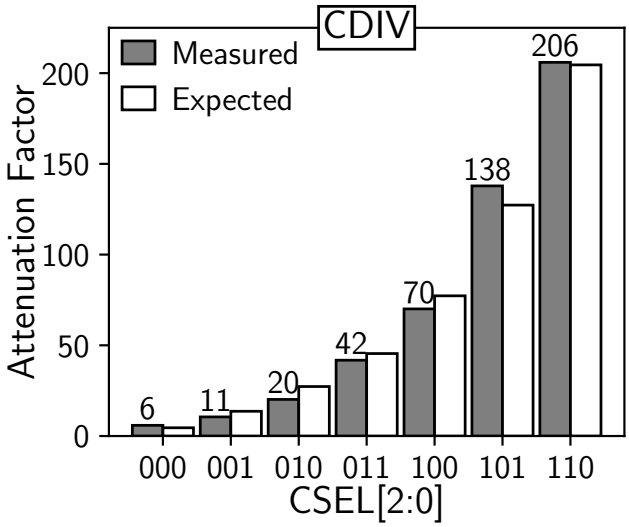


Figure D.8 | Attenuation factor of the capacitive divider at 100 mK.
Attenuation factor simulated and measured values at 100 mK used to tune the excitation amplitude applied on one of the quantum dot gate.

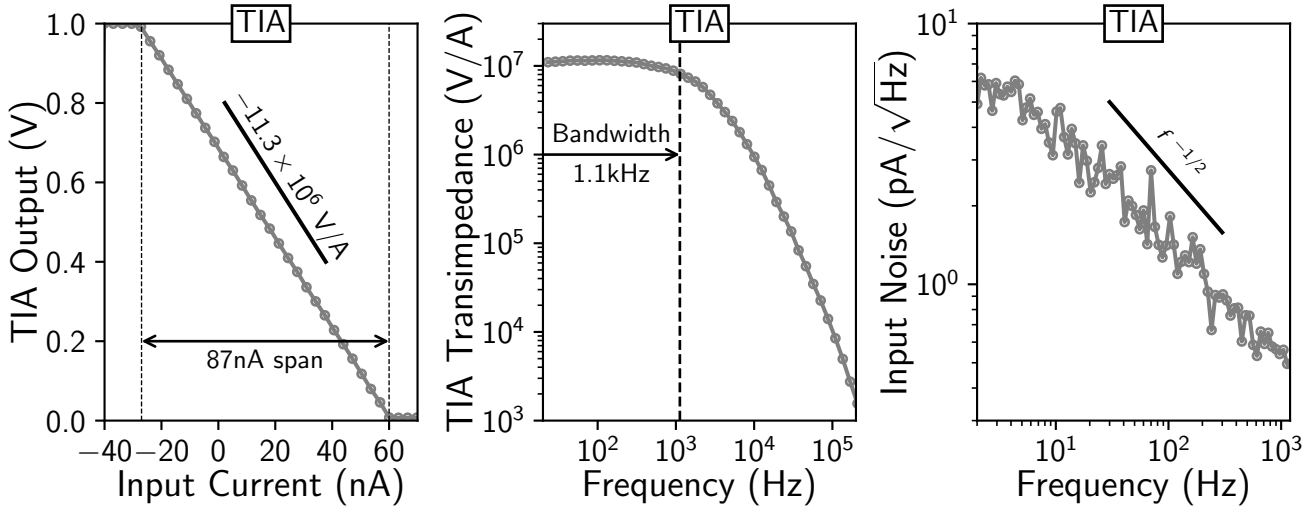


Figure D.9 | TIA performance at 100 mK.

From left to right: transimpedance gain, bandwidth, and noise of the TIA measured at 100 mK.

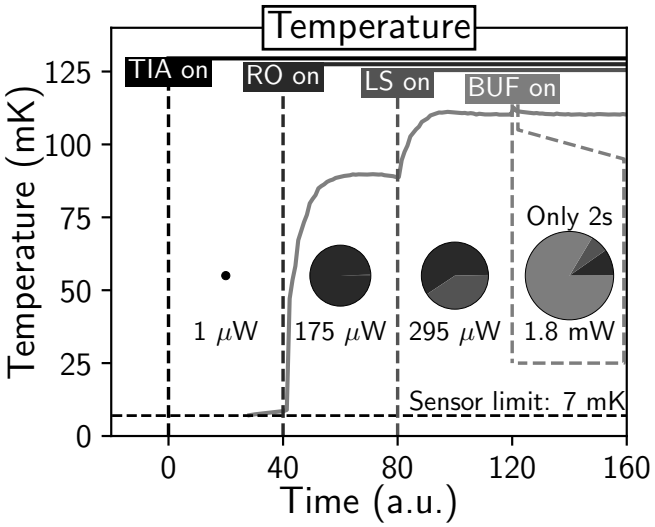


Figure D.10 | Fridge temperature rise from the IC power consumption.

Fridge bottom plate temperature due to the cryogenic circuit dissipation power.

V GHz excitation and DC read-out of the integrated double quantum dot

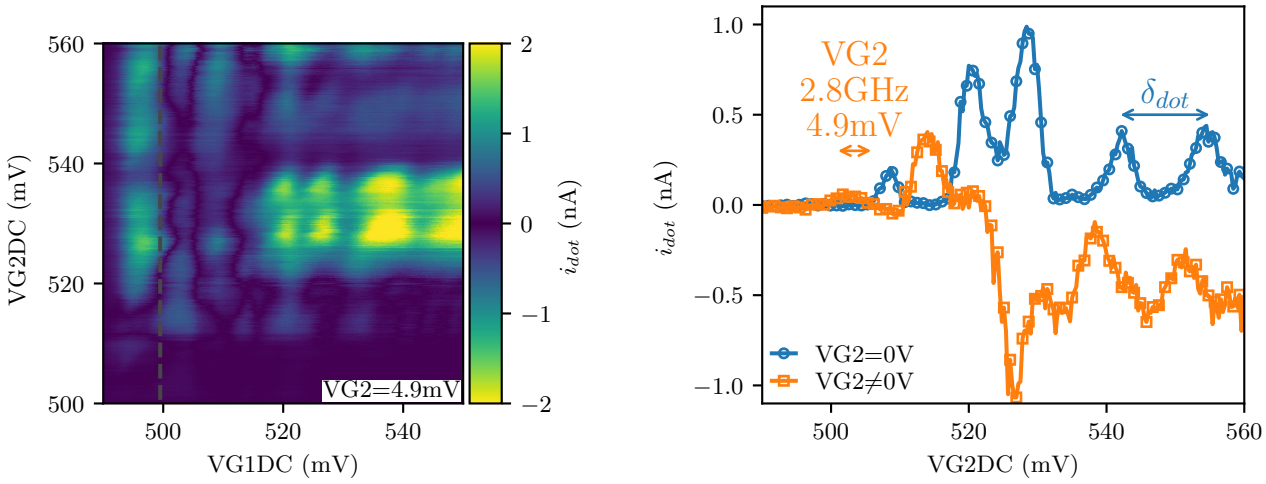


Figure D.11 | Quantum-dot measurement with a GHz voltage excitation at 100 mK.

Left: Quantum dot DC current measured with the micrometer-away TIA at 100 mK under a 2.8 GHz 4.9 mV excitation. Right: Quantum dot DC current at fixed $VG1DC \sim 500\,\mathrm{mV}$ with and without the high-frequency excitation $VG2$.

APPENDIX E
Quantum Full Integration with the Impedancemetry Chip

I Design of the integrated circuit

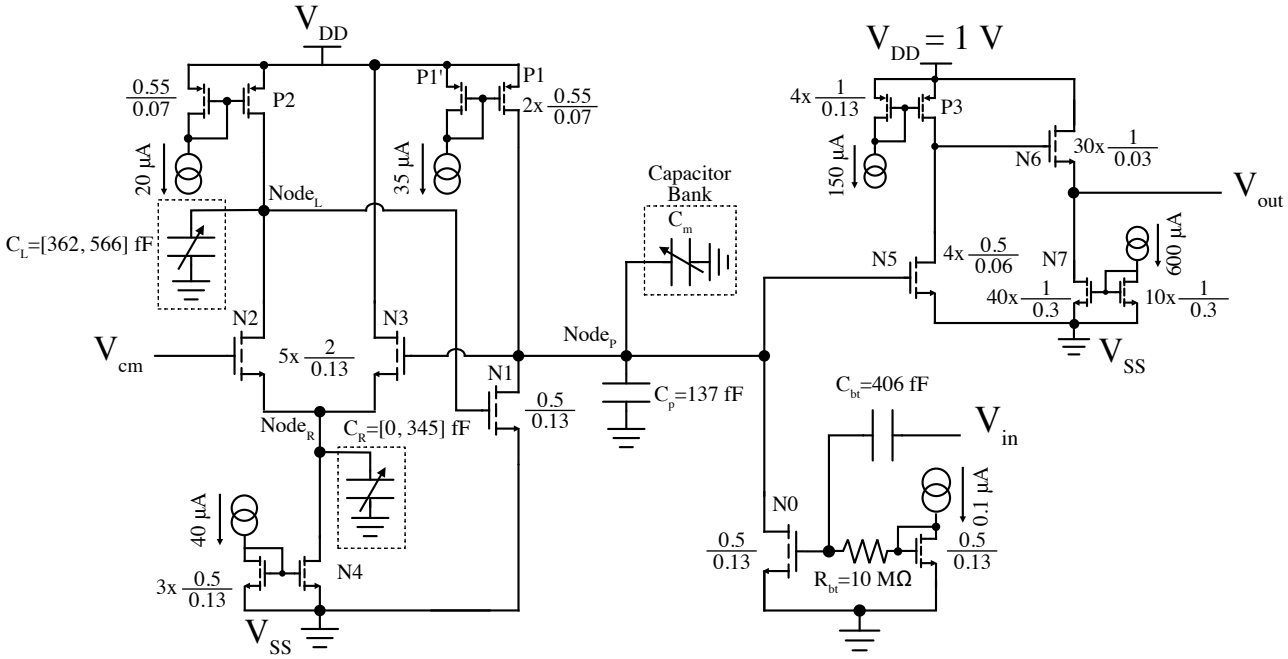


Figure E.1 | Design implementation of the impedancemetry chip.

Details of the components with transistor dimensions in the integrated circuit. The capacitor banks used as variable capacitors (C_L, C_R, and C_m) are detailed in Figure E.2 and E.3. Each transistor dimension is indicated as $m \times \frac{W}{L}$ with W (resp. L) the gate-finger width (resp. length) of the gate in µm and m is the number of fingers. The indicated current references of the diode-mounted transistors are generated at room temperature (see Figure E.14).

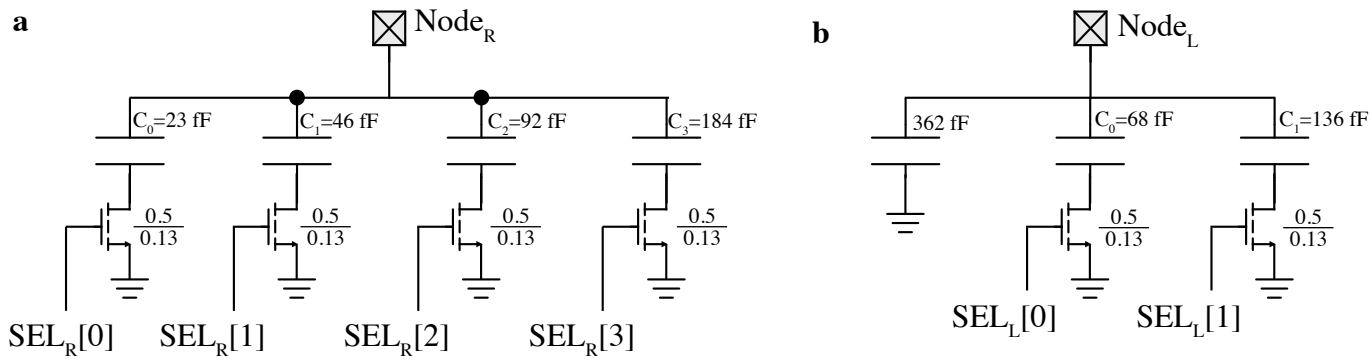


Figure E.2 | Design implementation of the variable capacitors.

a, Implementation of the variable capacitor C_R at Node$_R$ in Figure E.1. The 4 binary-weighted MOM capacitors are selected with NMOS switches activated by the selection bits $SEL_R[3:0]$ for C_R variation from 0 to 345 fF. **b**, Similar implementation of C_L with 3 binary-weighted MOM capacitors in parallel from 362 to 566 fF.

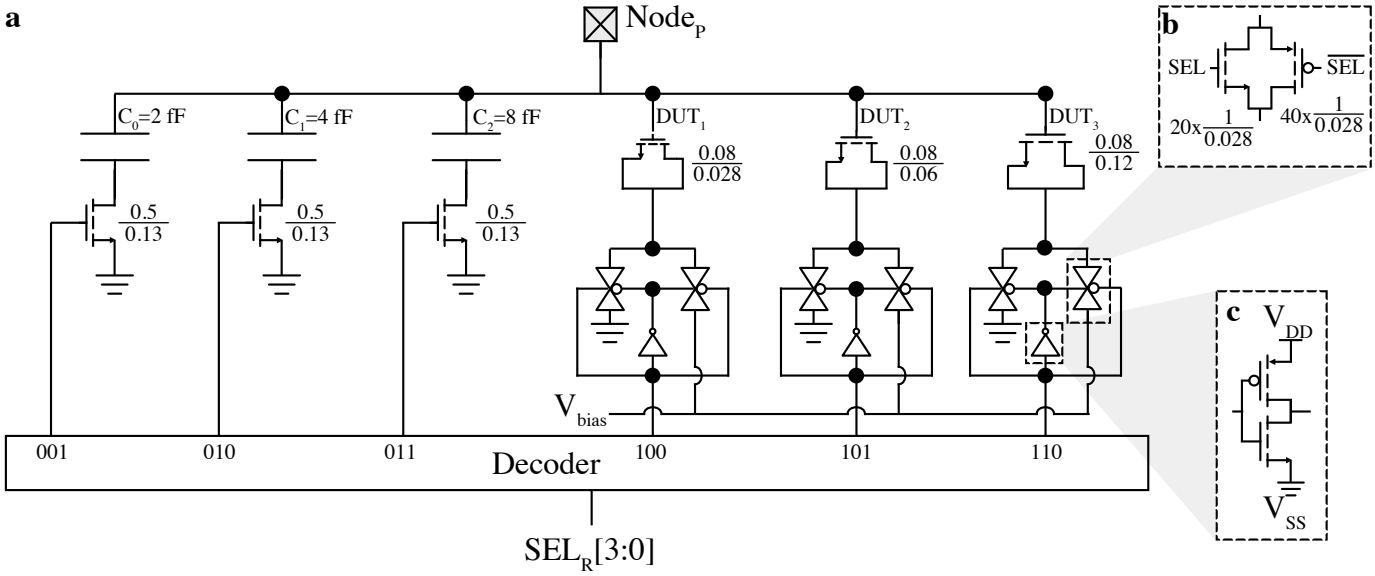
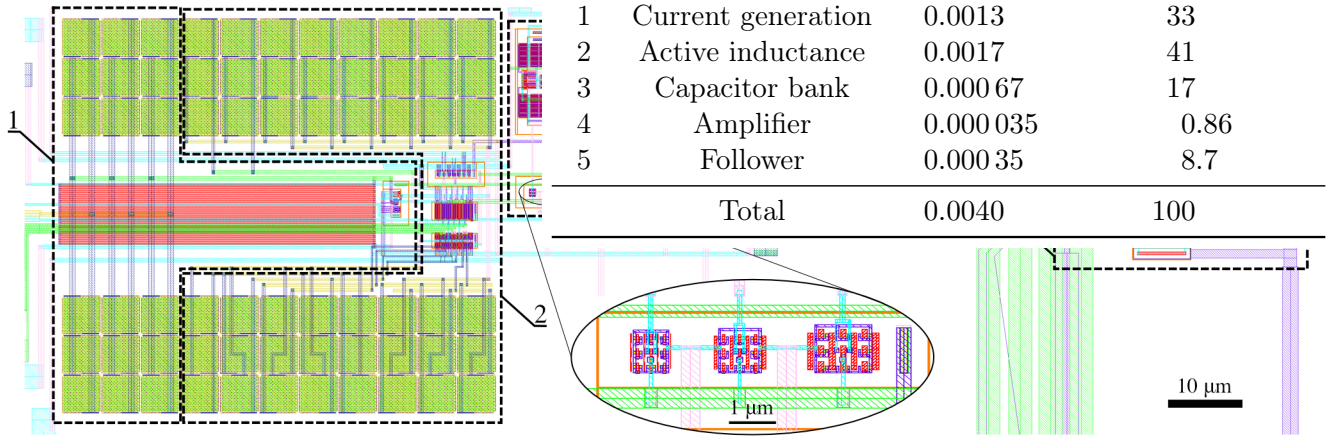


Figure E.3 | Multiplexing of the Device Under Test (DUT).

a, Design implementation of the capacitor bank with 3 MOM capacitors for capacitive calibration and 3 nanometer-sized MOSFETs for quantum capacitance measurements. The MOM capacitors are selected with NMOS switches. The MOSFETs are selected with a pair of pass-gates as detailed in **b**. When the MOSFETs are unselected, the drain and source are set to ground while the selected-MOSFET drain and source are linked to V_{bias} to change V_{gs}. Pairs of pass-gate are made complementary (when one is OFF, the other one is ON) with one inverter as shown in **c**.

a



b

ID	Block name	Area (mm^2)	Contribution (%)
1	Current generation	0.0013	33
2	Active inductance	0.0017	41
3	Capacitor bank	0.000 67	17
4	Amplifier	0.000 035	0.86
5	Follower	0.000 35	8.7
	Total	0.0040	100

Figure E.4 | Layout and footprint of the impedancemetry circuit.

a, Layout view of the impedancemetry circuit. The labeled areas correspond to the circuit block names of table **b**. The bottom inset shows a zoomed window for the measured quantum devices. Each device is surrounded by dummies to improve the fabrication quality of the nanometer-sized devices. **b**, Table of the occupied area of each block for a total footprint of 0.004 mm^2.

II Simulation Results at 300 K

Table E.1 | Noise contribution from linear AC simulations at 300 K.
Listed noise contributions of the 7 transistors generating 84 % of the total output noise of the impedancemetry chip extracted from foundry models at 300 K. The listed transistor devices belong to the active inductance which constitutes the main source of noise in the circuit.

Rank	Device	Contribution
1	P1	25%
2	N1	23%
3	P1’	14%
4	P2	10%
5	N4	5%
6	N2	4%
7	N3	3%

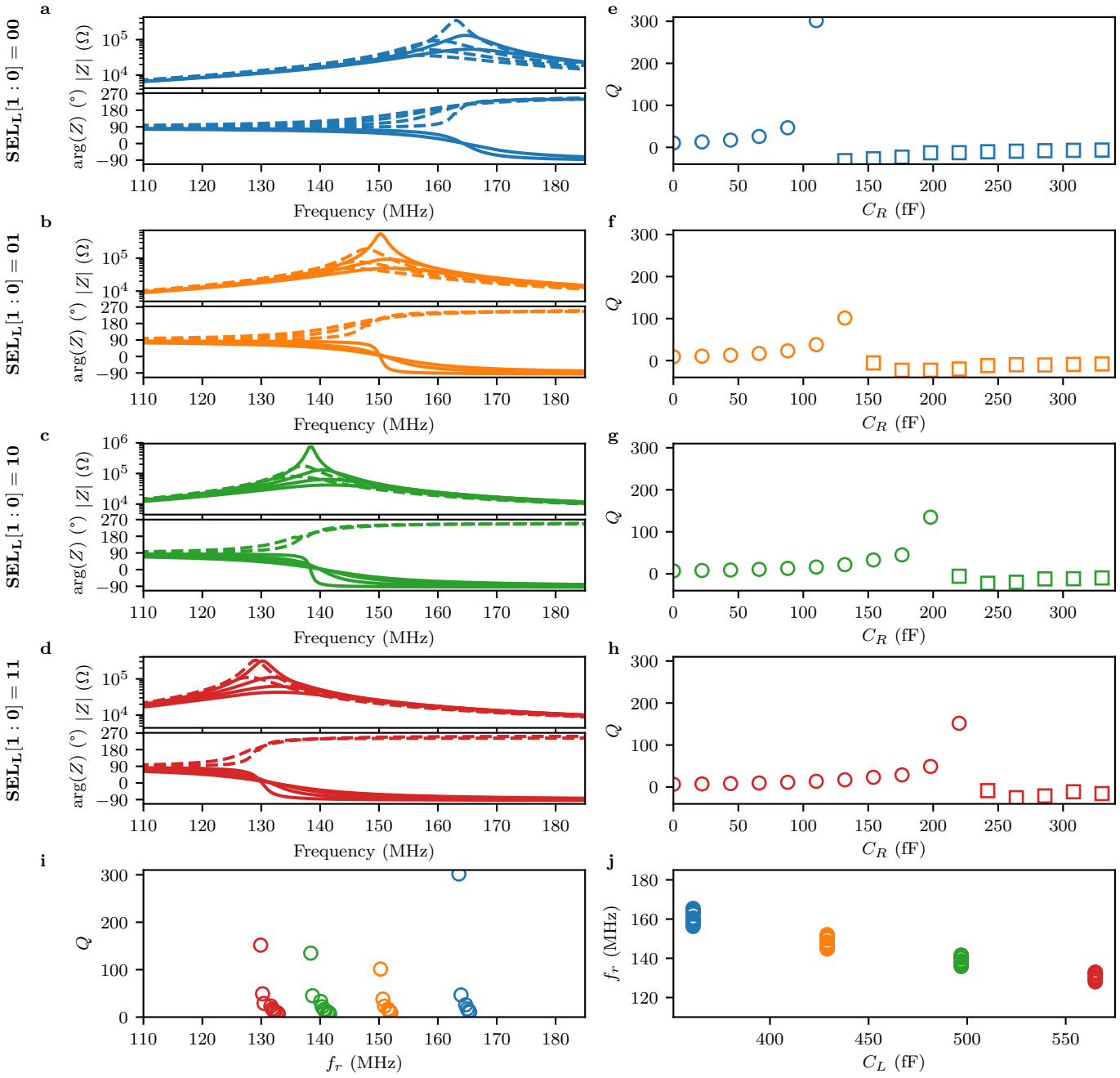


Figure E.5 | Impedance of the active inductance from simulations at 300 K.

a-d, Complex impedance of the active inductance for a few C_R values with C_L equal to: **a** 362, **b** 420, **c** 498, and **d** 566 fF. **e-h**, Quality factor Q of the active inductance as a function of C_R extracted from the tank impedance. In **a-d** (respectively **e-h**), stable resonance data with $Q \geq 0$ are shown in continuous lines (resp. round markers) while unstable resonance data with $Q < 0$ are shown as dashed lines (resp. square markers). **i**, Q as a function of the resonant frequency f_r showing small dispersion. **j**, Evolution of f_r with C_L at all C_R values.

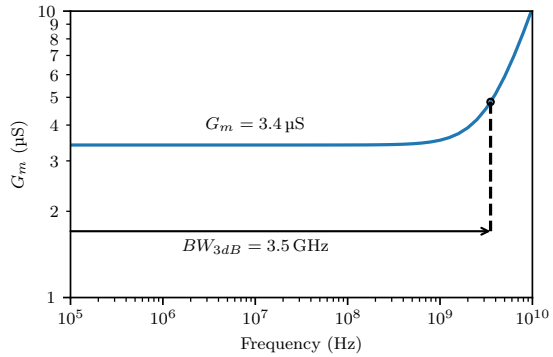


Figure E.6 | Voltage-to-current conversion for the current excitation of the tank from simulations at 300 K.
Transimpedance G_m and bandwidth of the current generation as a function of frequency extracted from foundry models at 300 K.

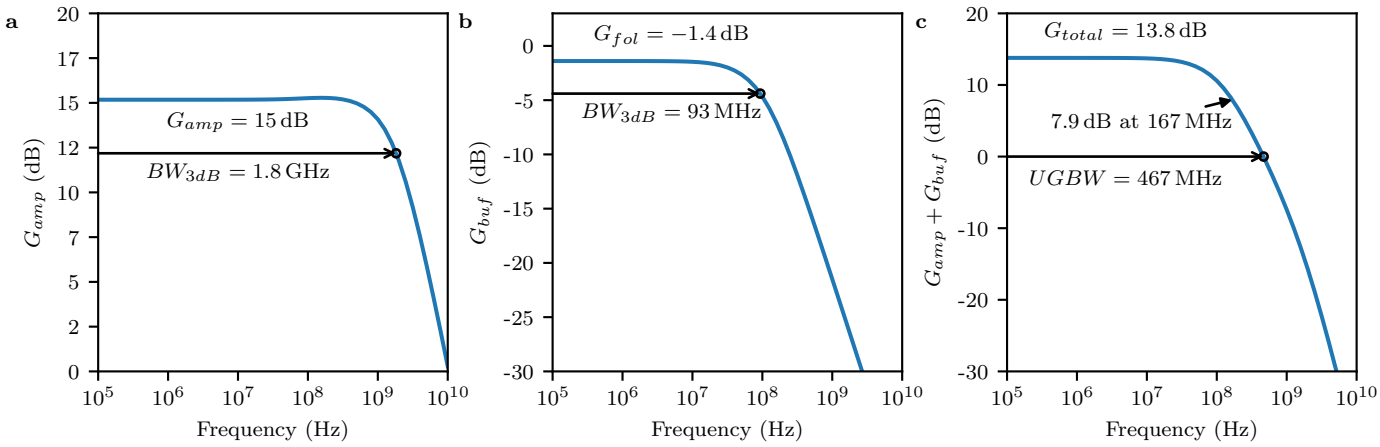


Figure E.7 | Amplifier and follower characteristics from simulations at 300 K.

a, Gain of the amplifier as a function of frequency extracted from AC simulations with foundry models at 300 K. **b**, Gain of the follower loading a 50 pF cable capacitance at the chip output. **c**, Total amplification of amplifier and follower as a function of frequency. Despite large cable capacitance from 4.2 K to 300 K stage, the unity-gain bandwidth of 467 MHz allows to keep the gain above 1 at the resonant frequencies f_r.

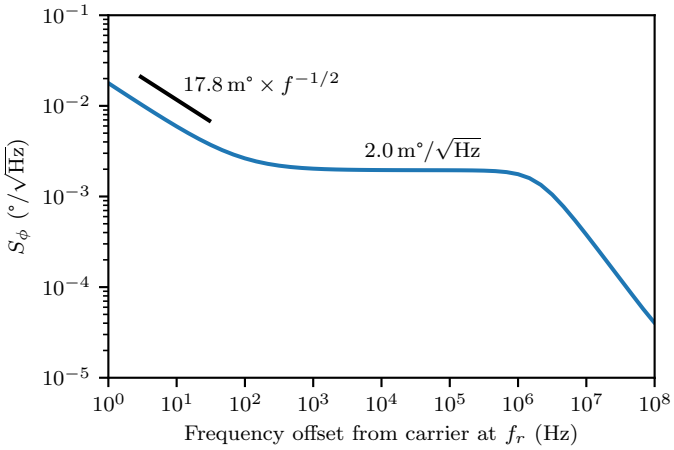


Figure E.8 | Phase noise of the impedancemetry setup from 300 K simulations.

Phase noise as a function of the frequency offset with respect to the resonance frequency of the tank $f_r = 167\,\text{MHz}$ extracted at the chip output from Steady STate (SST) simulations at 300 K with foundry models. The plateau above about 100 Hz with the cut-off around $\sim 2\,\text{MHz}$ is the phase noise behavior we could expect from a linear AC noise analysis. The cut-off corresponds to the width of the resonant peak $\kappa \sim f_r/Q \simeq 2.1\,\text{MHz}$ with the quality factor $Q = 81$. From the plateau value of $2\,\text{m°}/\sqrt{\text{Hz}}$, we estimate the input-referred noise to $3.7\,\text{aF}/\sqrt{\text{Hz}}$. The appearance of a flicker component in the phase noise is the evidence of the mixing between $1/f$ low-frequency noise and the tank high-frequency signal. This additional non-linear noise is attributed to transistor noise that translates into noise contribution to the inductance L and quality factor Q resulting in low-frequency noise up-mixing.

III Complementary data of the resonant circuit at 4.2 K

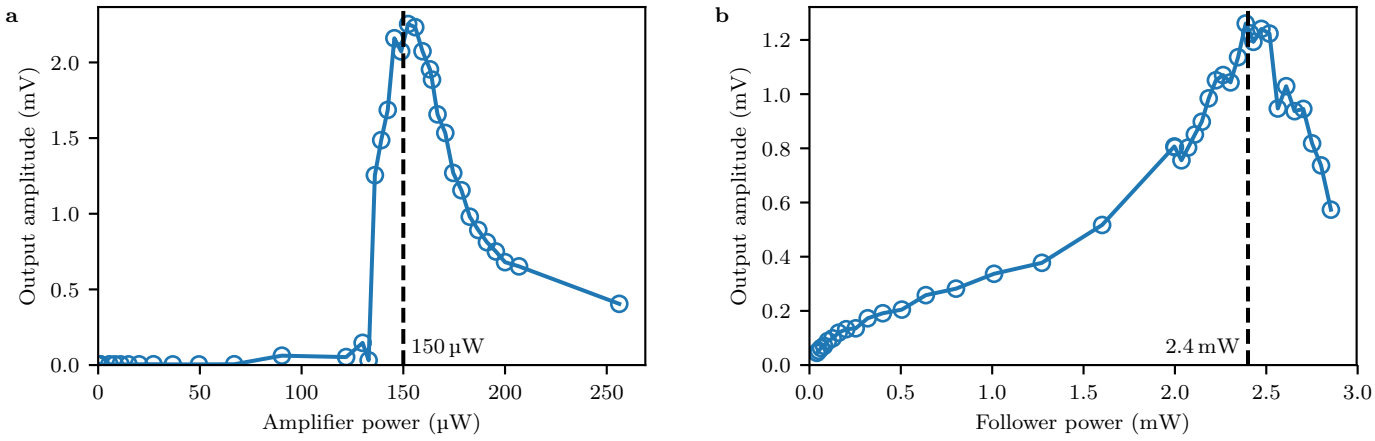


Figure E.9 | Amplifier and follower optimization at 4.2 K.

a, Output voltage V_{out} at the tank resonant frequency $f_r = 199\,\text{MHz}$ as a function of the amplifier power at 4.2 K. To maximize the gain, we choose the operating power of the amplifier at $150\,\mu\text{W}$. **b**, Output voltage V_{out} at the tank resonant frequency $f_r = 199\,\text{MHz}$ as a function of the follower power at 4.2 K. To maximize the gain, we choose the operating power of the follower at $2.4\,\text{mW}$.

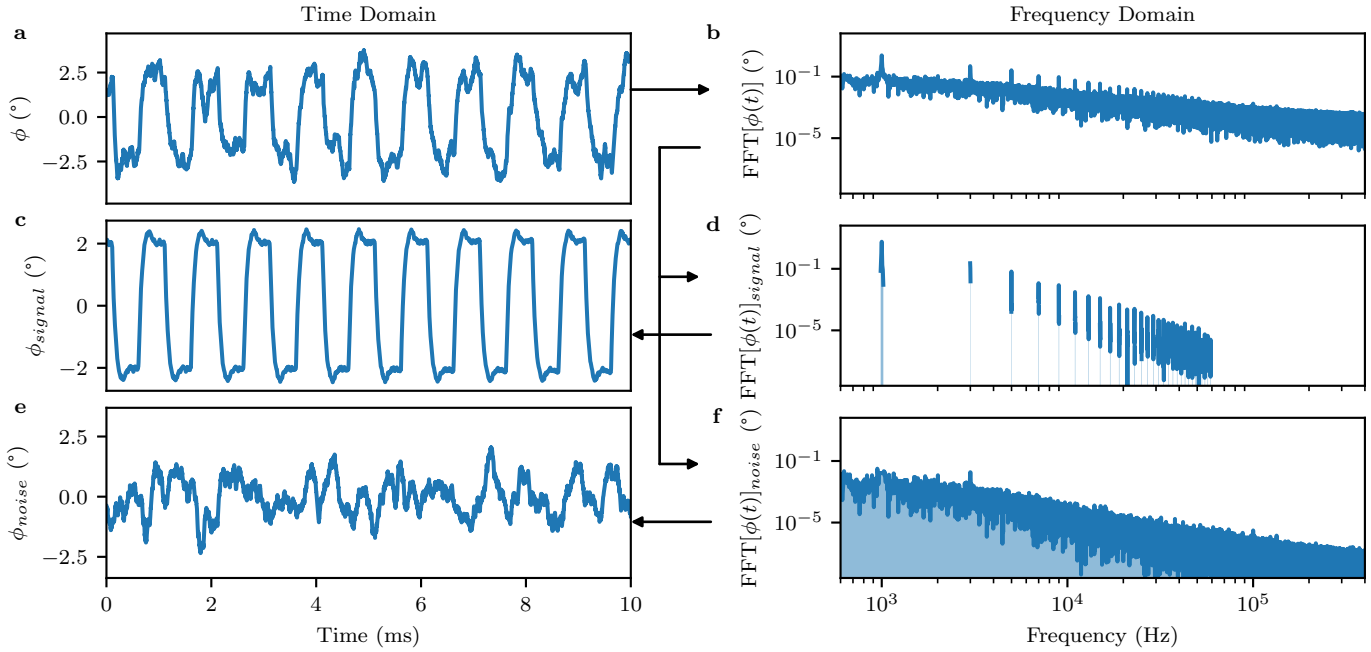


Figure E.10 | Extraction of the signal-to-noise ratio.

a, Signal of the phase output after single demodulation (I) at the tank frequency f_r when the DUT capacitance $C_m = 2\,\text{fF}$ is continuously connected and disconnected at a 1 kHz rate. **b**, Fourier power spectrum of ϕ with the typical signature of a square wave with exponential transients with the harmonics $2n + 1$. The signal power spectrum related to the square wave is isolated in **d** and the time-domain signal trace is recovered in **c** by inverted Fourier transform. The signal power P_{sig} is computed by integrating the power spectrum of the signal. **f**, Noise power spectrum in ϕ extracted as the complementary power spectrum to the signal spectrum shown in **b**. The noise power P_{noise} is computed by integrating the power spectrum of the noise. **e**, Time-domain noise extracted from **f**. The signal-to-noise ratio is extracted as the ratio of P_{sig}/P_{noise} and is equal to 5.7 for an integration time t_c of 55 µs in this example.

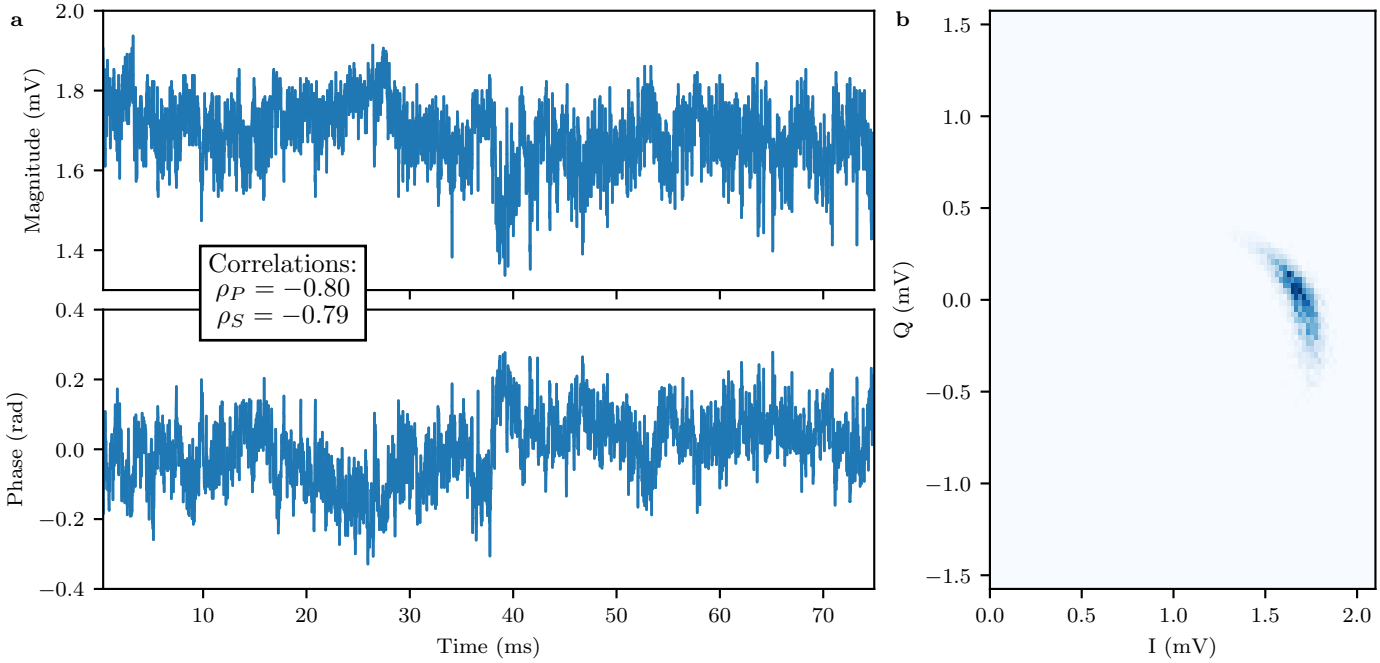


Figure E.11 | Correlated noise in the resonator output signal at 4.2 K.

(a) Magnitude and phase of the output signal after demodulation at the resonance frequency (method I). Clear correlations between magnitude and phase are observed as a function of time. Both the Spearman ρ_S and Pearson ρ_P computed on the entire time-trace exhibit strong anti-correlation between phase and magnitude with values of about -0.8. These correlations appear roughly after a 1 ms time-scale. **b**, 2D histogram in the I-Q plane of the data in **a**. The non-gaussian banana-shaped spot distribution reflects the evidence for correlated noise in the I-Q plane. This noise is attributed to transistor noise that leads to a noisy inductance value L and quality factor Q.

IV Measurement of the gate capacitance of a different DUT

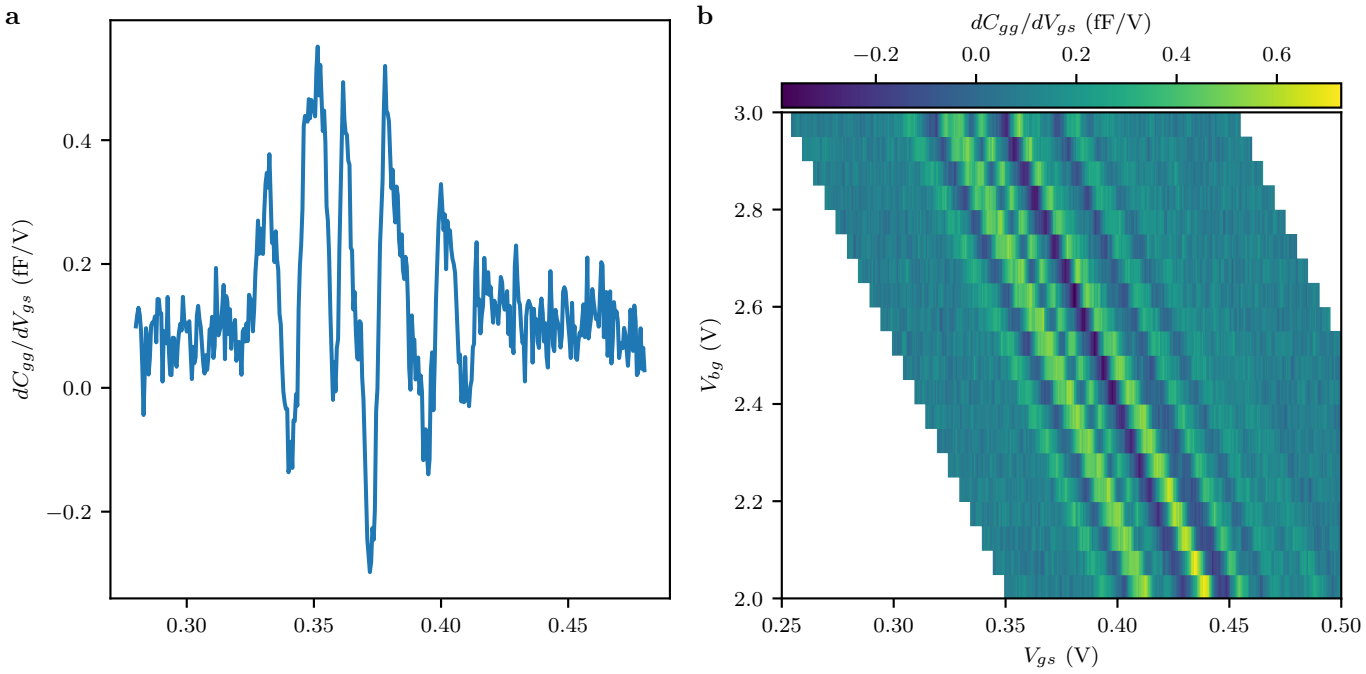


Figure E.12 | Measurement of the N-type DUT transistor with $L = 60\,\text{nm}$ and $W = 80\,\text{nm}$.

(a) First derivative of the gate capacitance dC_{gg}/dV_{gs} of a N-type MOSFET gate-capacitance as a function of the gate-source voltage V_{gs} measured with the impedancemetry setup at 4.2 K. Oscillations in the derivative capacitance is a sign of quantum capacitance from confined electronic state in the MOSFET channel. **b**, dC_{gg}/dV_{gs} evolution with front-gate V_{gs} and back-gate V_{bg} voltages. All features shift to lower V_{gs} for increasing V_{bg}. No traces attributed to an impurity state with an anomalous slope as seen in Figure 5 in the main text are detected for this device. These measurements are the same as presented in the main paper (Figure 5) for a longer device with $L = 120\,\text{nm}$ and $W = 80\,\text{nm}$.

V Experimental setup

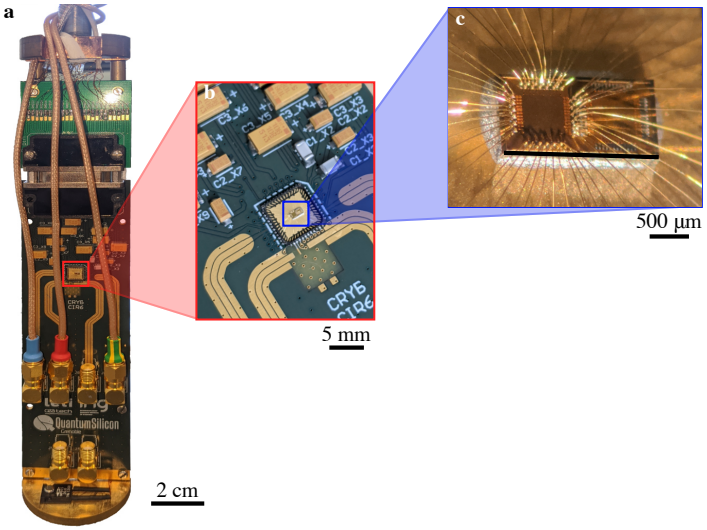


Figure E.13 | Experimental setup of the cryogenic sample holder.

Picture of the PCB (**a**) with soldered QFN48 (**b**), and wire-bonded integrated circuit (**c**). The PCB mounted at the end of a dip-stick is enclosed in a metallic tube filled with He gas for thermal exchange with a liquid helium bath at 4.2 K. High frequency signals are routed to SMA connectors at the end of the PCB with grounded coplanar waveguides and conveyed to room temperature with coaxial cables. SMD capacitors close to the chip reject the noise at sensitive voltage nodes (V_{DD}, V_{SS},...).

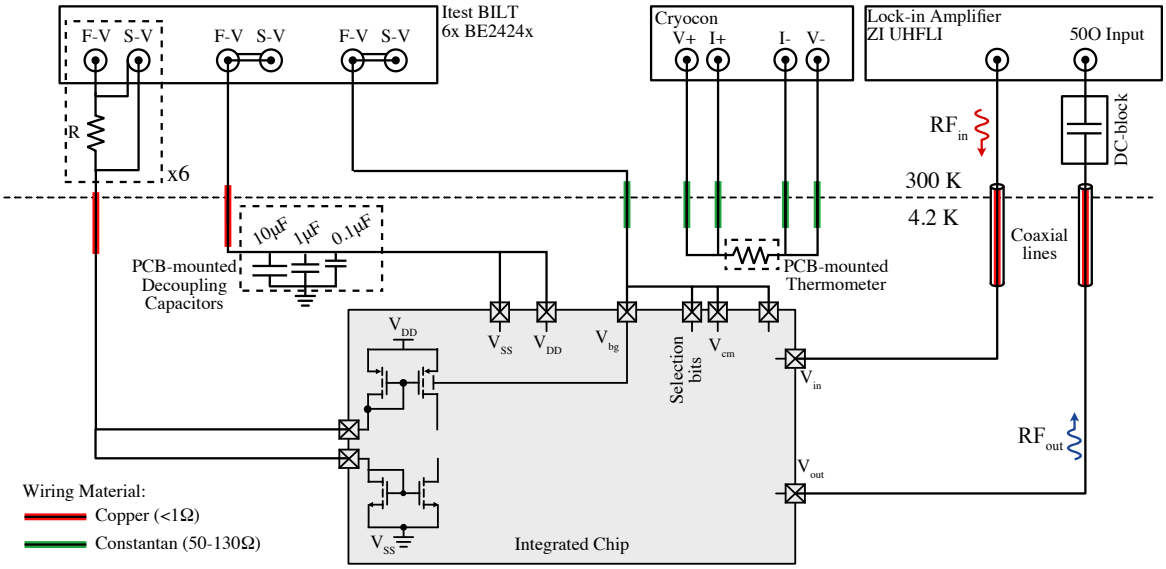


Figure E.14 | Instrumentation with connections to the chip.

The chip is anchored at 4.2 K while all the equipment for voltage and current references, excitation, and signal detection are placed at room-temperature. The chip power supply lines V_{DD} and V_{SS} are generated from a low-noise voltage source and conveyed with copper wring to avoid voltage drop along the line. Supply voltages are stabilized at cryogenic temperature with 3 discrete SMD capacitors (2 ceramic capacitors of 0.1 and 1 µF and 1 tantalum capacitor of 10 µF) placed on the PCB close to the chip pins. Current references are generated by applying a command voltage across a resistor R using the Sense (S-V) and Force (F-V) of the low-noise voltage sources. The 6 resistors R of 0.20, 0.40, 0.47, 0.90, 1.4, and 200 kΩ are placed in a grounded metallic enclosure for shielding from environmental noise. DC voltages on high-Z inputs (arriving on transistor gates, or back-gate wells) are applied with a low-noise voltage source with constantan wring to reduce heat conduction from 300 to 4.2 K. The chip excitation V_{in} and the output voltage V_{out} conveyed with coaxial SMA lines are generated and treated by a lockin-amplifier with 50 Ω-matched ports. A DC block at V_{out} prevents DC currents from the buffer output to flow in the 50 Ω port. A thermometer anchored at the PCB ground back-plane is used to monitor the PCB temperature to detect eventual heating above base temperature.

APPENDIX F
Ph.D. publications

SINGLE-TRANSISTOR CHARACTERIZATION AT CRYOGENIC TEMPERATURES

Cryogenic Subthreshold Swing Saturation in FD-SOI MOSFETs Described With Band Broadening
H. Bohuslavskyi, A. G. M. Jansen, S. Barraud, V. Barral, M. Cassé, **L. Le Guevel**, X. Jehl, L. Hutin, B. Bertrand, G. Billiot, G. Pillonnet, F. Arnaud, P. Galy, S. D. Franceschi, M. Vinet, and M. Sanquer
IEEE Electron Device Letters, vol. 40, no. 5, pp. 784–787, May 2019 (10.1109/LED.2019.2903111)

Integrated Variability Measurements of 28 nm FDSOI MOSFETs down to 4.2 K for Cryogenic CMOS Applications
B. C. Paz, **L. Le Guevel**, M. Cassé, G. Billiot, G. Pillonnet, A. Jansen, S. Haendler, A. Juge, E. Vincent, P. Galy, G. Ghibaudo, M. Vinet, S. de Franceschi, T. Meunier, and F. Gaillard
2020 IEEE 33rd International Conference on Microelectronic Test Structures (ICMTS), pp. 1–5, May 2020 (10.1109/ICMTS48187.2020.9107906)

Variability Evaluation of 28nm FD-SOI Technology at Cryogenic Temperatures down to 100mK for Quantum Computing
B. C. Paz, **L. Le Guevel**, M. Cassé, G. Billiot, G. Pillonnet, A. G. M. Jansen, R. Maurand, S. Haendler, A. Juge, E. Vincent, P. Galy, G. Ghibaudo, M. Vinet, S. de Franceschi, T. Meunier, and F. Gaillard
2020 IEEE Symposium on VLSI Technology, pp. 1–2, Jun. 2020 (10.1109/VLSITechnology18217.2020.9265034)

CIRCUIT TESTING AND ANALYSIS AT CRYOGENIC TEMPERATURES

Cryogenic Characterization of 28-nm FD-SOI Ring Oscillators With Energy Efficiency Optimization
H. Bohuslavskyi, S. Barraud, V. Barral, M. Cassé, **L. Le Guevel**, L. Hutin, B. Bertrand, A. Crippa, X. Jehl, G. Pillonnet, A. G. M. Jansen, F. Arnaud, P. Galy, R. Maurand, S. D. Franceschi, M. Sanquer, and M. Vinet
IEEE Transactions on Electron Devices, vol. 65, no. 9, pp. 3682–3688, Sep. 2018 (10.1109/TED.2018.2859636)

Cryogenic Current Steering DAC With Mitigated Variability
M. E. P. V. Zurita, **L. Le Guevel**, G. Billiot, A. Morel, X. Jehl, A. G. M. Jansen, and G. Pillonnet
IEEE Solid-State Circuits Letters, vol. 3, pp. 254–257, 2020 (10.1109/LSSC.2020.3013443)

INTEGRATION OF CRYOGENIC CLASSICAL ELECTRONICS WITH QUANTUM DEVICES

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
L. Le Guevel, G. Billiot, X. Jehl, S. De Franceschi, M. Zurita, Y. Thonnart, M. Vinet, M. Sanquer, R. Maurand, A. G. M. Jansen, and G. Pillonnet
2020 IEEE International Solid- State Circuits Conference - (ISSCC), San Francisco, CA, USA, Feb. 2020, pp. 306–308 (10.1109/ISSCC19947.2020.9063090)

Low-power transimpedance amplifier for cryogenic integration with quantum devices
L. Le Guevel, G. Billiot, B. Cardoso Paz, M. L. V. Tagliaferri, S. De Franceschi, R. Maurand, M. Cassé, M. Zurita, M. Sanquer, M. Vinet, X. Jehl, A. G. M. Jansen, and G. Pillonnet *Applied Physics Reviews*, vol. 7, no. 4, p.041407, Dec. 2020 (10.1063/5.0007119)

Compact gate-based read-out of multiplexed quantum devices with a cryogenic CMOS active inductor
L. Le Guevel, G. Billiot, S. De Franceschi, A. Morel, X. Jehl, A. G. M. Jansen, and G. Pillonnet
arXiv, Feb. 2021 (arXiv:2102.04364)

PATENT

Spin qubit quantum device read by impedance measurement
L. Le Guevel, G. Billiot, A. Jansen, and G. Pillonnet
US Patent 11,387,828, Jul. 2022

Acronyms

BEOL Back-End Of Line. 25
BOX Buried OXide. 36

CC Constant-Current. 37, 48
CLM Channel-Length Modulation. 51
CMOS Complementary Metal Oxide Semiconductor. 10, 104, 105, 109, 113, 115
CR Cross-resonance. 8

DAC Digital-to-Analog Converter. xiv, 10, 75, 90, 93
DIBL Drain-Induced Barrier Lowering. 41, 48, 49
DQD Double Quantum Dot. 113, 115
DRM Design Rule Manual. 30
DUT Device Under Test. 26, 27, 84, 101, 103, 106, 107, 109, 113, 132, 187, 188

EOT Equivalent Oxide Thickness. 36, 112
ESD Electrostatic Discharge. 31

FBB Forward Body-Biasing. 12, 39, 44, 73, 75, 78, 80, 87, 88, 93, 174, 177, 183, 184
FD-SOI Fully-Depleted Silicon-On-Insulator. 76, 84, 105, 109, 111, 115, 119, 120

GBW Gain-Bandwidth. 88, 107
GIDL Gate-Induced Drain Leakage. 33, 34
GO1 Thin-oxide. 29, 138, 158
GO2 Thick-oxide. 29, 75

I-V current-voltage. 26
IC Integrated Circuit. 11, 138
ISSCC International Solid-State Circuits Conference. 11

LNA Low-Noise Amplifier. 10, 101, 130
LSB Least-Significant Bits. 75, 90
LVT Low Threshold Voltage (flip-well structure). 29, 75, 138, 158

MD Mobility Degradation. 51
MELF Metal Electrode Leadless Face. 75
MI Moderate Inversion. 61
MOM Metal-Oxide-Metal. 106, 107, 116, 135, 157
MSB Most-Significant Bits. 75

OP-AMP Operational Amplifier. 75, 83, 88, 89, 104–106, 108, 110, 111, 119
OVR Operating Voltage Range. 77

PCB Printed Circuit Board. 31, 109, 113, 116
PNF Prime Number Factorization. 2
PU Processing Unit. 3

QD Quantum Dot. 112
QPU Quantum Processing Unit. 4

RBB Reverse Body-Biasing. 39
RO Ring Oscillator. 73, 86, 89, 178
RVT Regular Threshold Voltage (regular well structure). 29, 105

SCE Short-Channel Effects. 39
SEM Scanning Electron Microscope. 112
SET Single-Electron Transistor. 101
SNR Signal-to-noise ratio. 101
SoC System-on-Chip. 11
SOI Silicon-On-Insulator. 10, 93, 109, 112
SOS Silicon-On-Sapphire. 10

TIA Transimpedance Amplifier. 10, 101, 103–120, 159

Last update on April 12, 2023