



Silicide-based Josephson field effect transistors for superconducting qubits

Tom Vethaak

► To cite this version:

Tom Vethaak. Silicide-based Josephson field effect transistors for superconducting qubits. Mesoscopic Systems and Quantum Hall Effect [cond-mat.mes-hall]. Université Grenoble Alpes [2020-..], 2021. English. NNT : 2021GRALY017 . tel-03445492

HAL Id: tel-03445492

<https://theses.hal.science/tel-03445492>

Submitted on 24 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Physique de la Matière Condensée et du
Rayonnement

Arrêté ministériel : 25 mai 2016

Présentée par

Tom VETHAAK

Thèse dirigée par **François LEFLOCH** et
codirigée par **Fabrice Nemouchi**, CEA-Grenoble

préparée au sein du **Laboratoire PHotonique, ELectronique
et Ingénierie QuantiqueS**
dans l'**École Doctorale Physique**

Circuits supraconducteurs en technologie silicium

Silicide-based Josephson field effect transistors for superconducting qubits

Thèse soutenue publiquement le **7 juillet 2021**,
devant le jury composé de :

Monsieur FRANCK BALESTRO

PROFESSEUR DES UNIVERSITÉS, Université Grenoble Alpes,
Président

Monsieur FRANCESCO GIAZOTTO

DIRECTEUR DE RECHERCHE, Istituto Nanoscienze CNR,
Rapporteur

Monsieur CHRISTOPHE DETAVERNIER

PROFESSEUR, Universiteit Gent, Rapporteur

Monsieur SHI-LI ZHANG

PROFESSEUR, Uppsala Universitet, Examineur

Monsieur DOMINIC HORSMAN

DOCTEUR EN SCIENCES, Université Grenoble Alpes, Examineur



Abstract

Scalability in the fabrication and operation of quantum computers is key to move beyond the NISQ era. So far, superconducting transmon qubits based on aluminum Josephson tunnel junctions have demonstrated the most advanced results, though this technology is difficult to implement with large-scale facilities. Recently, the alternative “gatemon” qubit has appeared, which uses hybrid superconducting/semiconducting (S/Sm) devices as gate-tuned Josephson junctions. Current implementations of these use nanowires, of which the large-scale fabrication has not yet matured. Alternatively, CMOS Josephson Field-Effect Transistors may be used as tunable weak link in a scalable gatemon design, where an ideal device has leads with a large superconducting gap that contact a short channel through high-transparency interfaces. High transparency, or low contact resistance, is achieved in the microelectronics industry with silicides, of which some turn out to be superconducting. The first part of the experimental work in this thesis covers material studies on two such materials: V_3Si and $PtSi$, which are interesting for their high T_c , and mature integration, respectively. The second part covers experimental results on 50 nm gate length $PtSi$ transistors, where the transparency of the S/Sm interfaces is modulated by the gate voltage. At low voltages, the transport shows no conductance at low energy, and well-defined features at the superconducting gap. The barrier height at the S/Sm interface is reduced by increasing the gate voltage, until a zero-bias peak appears around zero drain voltage, which reveals the appearance of an Andreev current. The successful gate modulation of Andreev current in a silicon-based transistor suggests the feasibility of fully CMOS-integrated superconducting quantum computers.

Contents

1	Quantum computing	7
1.1	Introduction	7
1.2	Quantum circuits and gates	9
1.2.1	Single-qubit gates	9
1.2.2	Two-qubit gates	12
1.3	Algorithms and computational complexity	16
1.3.1	Types of quantum algorithms	17
1.3.2	Deutsch's problem	18
1.3.3	The Quantum Fourier Transform and Shor's algorithm	21
1.3.4	Grover's algorithm	25
1.3.5	Quantum annealing (QA) and the quantum approximate optimization algorithm (QAOA)	27
1.4	Physical implementations of quantum computers	28
1.4.1	Ion traps	28
1.4.2	Trapped spins in semiconductors	30
1.4.3	Superconducting qubits	30
2	Complementary Metal Oxide Silicon gatemons	41
2.1	Introduction	41
2.2	The quantum harmonic oscillator	42
2.2.1	Qubit from it	43
2.2.2	The transmon Hamiltonian	44
2.2.3	Gatemons	47
2.3	Superconducting transport across semiconducting junctions	48
2.3.1	Andreev reflection and the proximity effect	50
2.3.2	Diffusion	53
2.3.3	Decoherence	55
2.3.4	Energy, length and time scales	57
2.4	Transistors	58
2.4.1	The field effect	59
2.4.2	Schottky barriers	61
2.4.3	JoFET considerations	62
3	Silicides	71
3.1	Introduction	71
3.2	V3Si formation by pure metal deposition	72
3.2.1	Early encouragement	73
3.2.2	Signs of trouble	76

3.2.3	Lessons learned the hard way	79
3.3	V ₃ Si formation by deposition from a compound target	82
3.3.1	Strain, sound velocity and a drop in critical temperature	82
3.3.2	Substrate-induced thermal stress	83
3.3.3	Strain development during thermal processing	89
3.3.4	Stability of V ₃ Si thin films during thermal processing	94
3.3.5	Conclusion	100
3.4	PtSi	101
3.4.1	Split 1: platinum deposition	102
3.4.2	Split 2: rapid thermal processing (RTP)	106
3.4.3	Split 3: A look under the hood	111
3.4.4	Conclusion	116
4	Josephson Field Effect Transistors	123
4.1	Introduction	123
4.2	Test structures	123
4.2.1	Academic mask set: Python code for adjustable patterns	123
4.2.2	Industrial mask set: TASP	128
4.3	Proximity effect in a PtSi SBMOSFET	131
4.3.1	Description of the samples	131
4.3.2	Room-temperature tests	132
4.3.3	Cryogenic measurement setup	134
4.3.4	Demonstration of the field effect at low temperature	136
4.3.5	Evidence of superconductivity	138
4.3.6	Gate modulation of the proximity effect	142
4.3.7	Estimation of the Schottky barrier height	149
4.3.8	Conclusion	152
	Appendices	I
A	Samples for material studies	III
A.1	Samples prepared by pure V sputtering	III
A.2	Samples prepared by compound V ₃ Si sputtering	VI
B	Devices measured at low temperature	XI

Introduction

“You have nothing to do but mention the quantum theory, and people will take your voice for the voice of science, and believe anything.”

– George Bernard Shaw

In front of you are 10^7 bits of information. That sounds like a lot, and it certainly took me a while to put together, but it isn't nearly enough to describe what is going on inside a chip like Sycamore at any point in time [1], even after accounting for redundancies due to error rates and such [2]. What's remarkable is that all this complexity is accounted for by only 53 physical qubits. This exponential scaling of the Hilbert space with the number of qubits is the basis for the general excitement about the promise of quantum computing. Possible applications range far beyond the familiar breaking of RSA cryptography, from optimization problems and quantum machine learning, to a complete overhaul of chemistry, materials science and biology by the simulation of nano-scale electronic interactions. Broadening the range of computable problems is of course an exciting, rewarding and even economically sensible goal, and justifies the recent surge in activity. On the other hand, the cautious skepticism towards claims of any imminent *useful* advantage is equally justified. But even if quantum computers only become competitive with supercomputer clusters on actual applications after our lifespans, or if there turn out to be practical limits to the scale at which coherence can be maintained, this field still has an additional appeal on a more basic level.

What gets me excited is the realization that at its most fundamental, existence itself is computation [3–5], and the laws of physics don't deal with classical bits, they deal with qubits [6–8]. No matter what underlying structure there turns out to be to the universe, whether it contains supersymmetric siblings of every elementary particle, or if they are all excitations of strings vibrating in 10 or 11 dimensions [9, 10], in the end it is all described by *information*; a vast register of qubits in a cosmic quantum computer. The quantum nature of reality has been tested (and always verified) since Einstein's interpretation of the photoelectric effect, and improvements in the understanding of its principles has led to revolutions in science and technology alike [11]. But only recently have we begun to isolate, couple and probe individual degrees of freedom in physical qubits, such that we can now set, flip or measure the state of a single fundamental unit of information. Now that we have broken it all down to this core component, we can start to build something of our own. I do not much care what the appropriate level of skepticism is towards any predictions of near-term realizations of large-scale quantum computers, for to me *this*, the gaining of control over qubits, is already a turning point in science.

This thesis explores a path towards large-scale quantum computers. The central idea is to leverage the existing fabrication processes already in use for classical computer chips,

and adapt them to mass-produce quantum computers made of superconducting circuits. We will first discuss the broader context of quantum computing in chapter 1, before delving into the details of designing a superconducting qubit in chapter 2. Here we will make the argument that, like with so many instances of technological innovation, the key is in the development of the right materials. In our case, this means studying silicides, intermetallic compounds that are part silicon and part metal, that are traditionally used in the contacts to the semiconducting silicon channel of transistors. Chapter 3 then elaborates on experiments with two such silicides that superconduct, V_3Si and $PtSi$. This last material was used in the late 90's in a special kind of transistor where, unusually, the contact silicides extended to underneath the gate electrode. As we will discuss in chapter 4, the particular combination of design parameters made it hard to switch this transistor off at room temperature, rendering it unsuitable for its original purpose. Fortunately, these very same properties allowed for current to flow even when the device was cooled to a few tens of millikelvin, leading to the observation of induced superconductivity inside the silicon channel. Although work remains to be done, we hope to convince you that this is a promising result that hints at the possibility of large-scale fabrication of silicon-based superconducting qubits.

Bibliography

- [1] Frank Arute et al. “Quantum supremacy using a programmable superconducting processor”. In: *Nature* 574.7779 (2019), pp. 505–510.
- [2] Yiqing Zhou, E Myles Stoudenmire, and Xavier Waintal. “What limits the simulation of quantum computers?” In: *Physical Review X* 10.4 (2020), p. 041038.
- [3] Richard P Feynman. “Simulating physics with computers”. In: *International Journal of Theoretical Physics* 21.6/7 (1982).
- [4] David Deutsch. “Quantum theory, the Church-Turing principle and the universal quantum computer”. In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 400.1818 (1985), pp. 97–117.
- [5] Seth Lloyd. “Computational capacity of the universe”. In: *Physical Review Letters* 88.23 (2002), p. 237901.
- [6] David Deutsch. *The fabric of reality*. Penguin, 1998. ISBN: 9780140275414.
- [7] Scott Aaronson. *Quantum computing since Democritus*. Cambridge University Press, 2013. ISBN: 9780521199568.
- [8] Jonathan P Dowling. *Schrödinger’s Killer App: Race to Build the World’s First Quantum Computer*. CRC Press, 2013. ISBN: 978-1439896730.
- [9] Leonard Susskind. *The black hole war: My battle with Stephen Hawking to make the world safe for quantum mechanics*. Little, Brown & Company, 2009. ISBN: 9780316016414.
- [10] Sabine Hossenfelder. *Lost in math: How beauty leads physics astray*. Hachette UK, 2018. ISBN: 978-0465094257.
- [11] Jonathan P Dowling and Gerard J Milburn. “Quantum technology: the second quantum revolution”. In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 361.1809 (2003), pp. 1655–1674.

Acknowledgements

“Teamwork makes the dream work.”

– Scott Kelly

De tous ceux que je dois énumérer, sans doute premier parmi ses pairs est Frederic Gustavo. Fred, merci de m’avoir appris les bonnes manières de la salle blanche, et de m’avoir aidé à supporter la charge des expériences. C’est un fait indiscutable que pas un seul des résultats discutés dans cette thèse n’a été obtenu sans ton aide. Mais surtout, je veux te remercier pour les nombreuses heures pendant lesquelles nous avons parlé de tout *sauf* de science. Je suis sûr que je me souviendrai avec émotion de nos discussions et dîners longtemps après avoir oublié la plupart des résultats scientifiques dans ce manuscrit.

As with Fred, I would like to thank also François Lefloch in the first place for the human component that he brings to his research. His concern for the well-being of those around him is of a level that is rare not just in scientific environments. To name just one instance, I’m glad that I could take off an entire month when my grandmother was about to pass away, without any questioning of my commitment, nor any push to work remotely. This freedom extended to the work itself, and has helped me enormously to gain the confidence to do research independently. As co-supervisor, Fabrice “the magician” Nemouchi has been absolutely irreplaceable to help me navigate the LETI, and has taught by example the importance of all the soft skills involved in any collaborative effort. The sense of humor that he shares with his taller alter ego Philippe Rodriquez has brought color to even the most mundane discussions.

I’m greatly indebted to Valérie Lapras, who has introduced me to all the principles of the cleanroom, and who has sacrificed many weeks of her time to cook up the process flows for nearly all the lots that were prepared. For the materials part of this thesis I have relied heavily on Shi-Li Zhang, who became my mentor on all things related to silicides, and who together with Tomas Kubart took good care of me in Uppsala. Last, I am grateful to Sylvan Brocard, who has been my soundboard and intellectual anchor throughout the past three years.

Chapter 1

Quantum computing

1.1 Introduction

“It is difficult to make predictions, especially about the future.”¹

– Danish proverb

Quantum computers will change our world.

To see why, first consider how we became aware of the power of *classical* computation. Around the time that Alan Turing wrote his prescient paper “*Computing machinery and intelligence*” [1], where he introduced the concept of machine intelligence, the general public at large had not yet had any interaction with an artificial computer². It then took 40 years from the invention of the integrated circuit [2] to the explosion in software applications that shaped our modern world [3], and only this last decade have we seen convincing implementations akin to Turing’s vision. What Turing had touched on that allowed him to preconceive the future that we are only now embarking on, was a concept called *universality*³ [4, 5], the idea that there is a small set of operations that all logical processes can be reduced to⁴. Since there was (is?) no reason to assume that human thought is not a logical process⁵, he applied this principle to reason that computing machines, once they would be fast enough, could emulate us as well. The impact of computation on the world was predicted not by extrapolating from contemporary capabilities or applications, but by taking a fundamental statement about computability to its logical conclusion.

¹“*Det er vanskeligt at spaa, især naar det gælder Fremtiden.*”

²“Artificial”, to distinguish the machines from the profession.

³While it is now known as a “universal Turing machine” due to Church, he originally named his invention more humbly a “universal computing machine” in the paper cited above. He starts section 6 with the phrase “*It is possible to invent a single machine which can be used to compute any computable sequence*”, which is now known as the Church-Turing thesis.

⁴The original example of a universal machine was a tape reader with a set of internal states, that moves along a tape with bits of 0 and 1. It would move, read, write, or change its configuration depending on the combination of its current configuration and the value of the bit it was on.

⁵At the time that Turing, Shannon and von Neumann were making these comparisons, McCulloch and Pitts’ Boolean framework [6] was still current. Cobb details how we now know that the brain is not just a network of logical gates (signals are non-binary, bi-directional, and can be modulated by semi-global variables) [7], but far as I understand computational complexity theory, there is still a chance that it can be simulated efficiently on a Turing machine.

Now, how does this relate to the current situation around quantum computers, and what kind of fundamental statement can we make about them? After Turing had established that all naturally computable functions can be computed by a universal machine *in principle*, the field started wondering whether such a machine could always do this *efficiently*. And so Church and Turing’s original thesis was extended, to state that for any problem, if it is possible to make a special dedicated machine tailored to that problem that can solve it efficiently, then a Turing machine can solve it efficiently as well (with at most polynomial overhead). This qualifier about efficiency makes an important difference in practical terms. For example, if a traveling salesperson wanted to minimize the distance she would have to travel to visit each of a set of cities, she would need an amount of time exponential in the number of cities when consulting a Turing machine⁶. Given the extended thesis, she would concede that there is not much point in building a specialized tool either, since whatever speedup it would give, it could still never do better than exponential time⁷. The extended thesis puts a cap on the potential of computers: a Turing machine is the most powerful thing that you could possibly build, and those problems that are out of its reach will remain forever so. But here’s the crux: it turns out to be *wrong* [8].

Turing’s universal machine appears not to be universally efficient after all: there are problems that can be solved in polynomial time on a quantum computer, that are exponentially hard on a classical one. Though not proven rigorously before 1993 [9], the difficulty of simulating quantum systems and therefore the relative advantage of computers based on them was already realized by Feynman in 1982 [10]⁸. We do not yet exactly know, however, how many such problems there are. For example, whether the traveling salesperson mentioned above may one day find her shortest route by sending a query to a quantum server is unknown: no-one has yet proven whether NP-complete problems fall within BQP (Bounded-error Quantum Polynomial, loosely speaking the set of problems that can be solved in polynomial time on a quantum computer), or any other complexity class that takes advantage of quantum mechanics [12]. Neither do we know whether the exponential speedups of some of our best quantum algorithms are truly insurmountable by any classical means.

Nobody has yet found proof that Shor’s prime factorization algorithm [13], arguably the best-known example of the potential that quantum computers hold, cannot be equaled in speed by something run on a classical Turing machine. As far as we know, finding

⁶Though “determining” the solution is exponentially hard, checking it can be done in polynomial time, which means that its complexity is *non-deterministically polynomial* (NP). What’s more, once you have found the shortest route, you can use that solution to efficiently solve a range of other NP problems, such as finding the lowest-energy configuration of an Ising spin glass. This equivalence groups these problems together in what is called NP-complete.

⁷Since e^x is an infinite series of all powers of x , division by any finite polynomial still returns an exponential.

⁸The paper contains many profound insights, such as the idea that the correctness of a law depends on its computability, since the things that that law describes could themselves be used as the computer. Once you realize that quantum mechanics cannot be efficiently simulated by a classical computer, two options remain: either you assume the extended Church-Turing thesis, taking computability to mean *computable by a Turing machine*, and conclude that our physical laws must be wrong, or you take quantum physics to be true and conclude that the extended thesis is wrong. Luckily the Bell test was done a decade earlier by Freedman and Clauser, helping him conclude the latter [11]. He also posited the idea of a probabilistic computer, which are now known to be faster in some situations (e.g. they can test whether a number is prime in polynomial time [12]), giving rise to the complexity class BPP. It was relative to BPP that Bernstein and Vazirani proved quantum computers to be more powerful, $BPP \subseteq BQP$ and $BPP \neq BQP$ [8] (here BPP stands for Bounded-error Probabilistic Polynomial time).

the prime factors of large numbers is exponentially hard on a classical computer. This is so hard that even if every particle in the entire universe were to be used as a logical component in an enormous computer, which would be left to calculate prime factors for the entire age of the universe, it could at most have factored only a million-bit number (less than a megabyte) [14]⁹. But this is only true if our best classical algorithms really are the best ones *possible*. In July 2018, 18-year old Ewin Tang posted a preprint [16] describing a classical algorithm that could solve the problem of sampling sparse matrices in polylogarithmic time. This was a shock to the computational complexity community, who had considered the quantum algorithm for this problem [17] a promising indication that quantum machine learning (one of three loosely defined branches of quantum algorithms, along optimization and simulation) could offer exponential speedups. To further illustrate the state of the field, not only are we unsure of BQP’s relationship to NP, we cannot even say with certainty that $P \neq NP$!

So here is where we are today: we do not yet really know what limits there are to the computational power of a classical computer, but we are sure that whatever it *can* do, a quantum computer can do more. We already have many quantum algorithms for more or less useless problems that are certainly faster than anything that could run on a classical computer, and a healthy number of quantum algorithms for actually useful problems that are probably faster than classical alternatives. Though currently more or less limited to the simulation of quantum systems, finding prime factors and quantum walk graph traversal, large efforts are under way to add to the list of useful problems with certain, exponential quantum advantage.

1.2 Quantum circuits and gates

Feynman ended the first half of his 1982 talk by saying “*I therefore believe it’s true that with a suitable class of quantum machines you could imitate any quantum system, including the physical world.*”, and asked “*What, in other words, is the universal quantum simulator?*” This question was answered by David Deutsch three years later [18], who gave a proof of existence of a universal quantum Turing machine¹⁰. In short, Deutsch showed that there exist initialization programs ρ for each L -qubit state $|\psi\rangle$ such that $\rho|\psi\rangle = |0_L\rangle$. These programs together with their inverses, combined with the ability to multiply the ground state by a phase factor, can then perform arbitrary unitary transformations on any L -qubit state. Though this was an important result as it proved that we can construct a universal quantum Turing machine in principle, the proposed operations are not very practical. Much simpler and more convenient gate sets have since been developed, and we will see below that any universal set of single-qubit gates plus a single entangling two-qubit gate is sufficient.

1.2.1 Single-qubit gates

When discussing single-qubit gates, the word “universal” refers to the possibility of obtaining any (normalized) complex combination of $|0\rangle$ and $|1\rangle$ in a unitary manner¹¹.

⁹The hardness of this problem is the basis for RSA encryption [15].

¹⁰That quantum Turing machines can simulate each other efficiently was only proven in 1993 [19].

¹¹Unitarity implies reversibility, which means that information is always preserved. This fundamental principle of quantum mechanics led Don Page to conclude that black holes must preserve information, contradicting his PhD supervisors Stephen Hawking and Kip Thorne. It became a productive controversy, leading to the holographic principle [20] (which in turn is one of the pillars of Erik Verlinde’s entropic

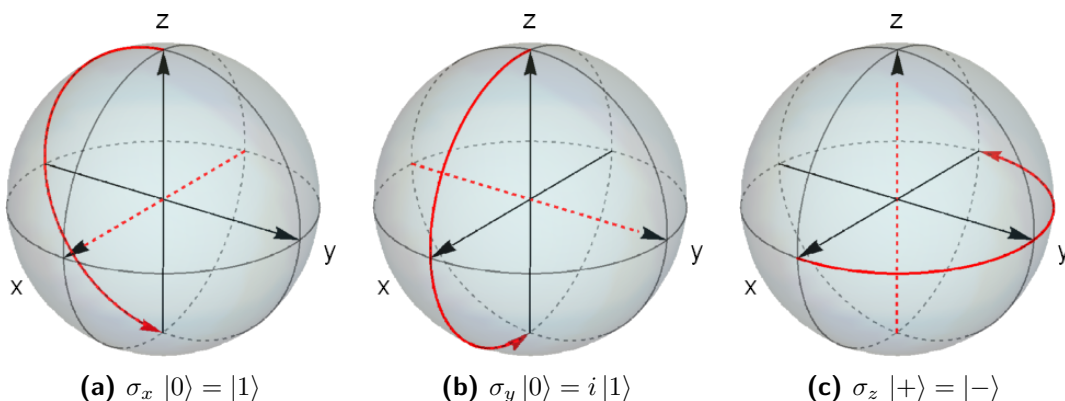


Figure 1.1: Each of the Pauli matrices is a π rotation around the relevant axis. Since σ_x turns $|0\rangle$ into $|1\rangle$ and vice versa, it is also known as the “bit-flip operator”. Rotations around the y -axis do the same, except that they also introduce a phase shift. This distinction between x and y is of course just a matter of choice in the case of single qubits, where we can forget about the global phase. Once we entangle multiple qubits, however, the phase becomes relative and really does matter (hence the name of the i SWAP operation that we will meet in section 1.2.2). The identity matrix, σ_0 , preserves the state.

One might think that the familiar Pauli matrices,

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (1.1)$$

which form a complete basis for the 2×2 space of Hermitian matrices, should be able to rotate the qubit in any direction. In fact these matrices all cause rotations of the basis states by multiples of π , so you cannot, for example, use them to go from a z -eigenstate to an x -eigenstate. One popular way of visualizing this is the Bloch sphere, shown in Fig. 1.1, in which the ground state points up along the z -axis, the excited state points down, and the equator represents equal superpositions of these two with varying phases.

To create a superposition, the Hadamard gate can be used,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = |+\rangle \langle 0| + |-\rangle \langle 1|, \quad (1.2)$$

which is a rotation around a tilted axis that maps $|0\rangle$ to $|+\rangle = \sqrt{1/2}(|0\rangle + |1\rangle)$ and $|1\rangle$ to $|-\rangle = \sqrt{1/2}(|0\rangle - |1\rangle)$ (and vice versa), as illustrated in Fig. 1.2. Though superposition and entanglement are purely quantum-mechanical phenomena that cannot appear in classical systems, they still are not sufficient to give a computational advantage. Gottesman and Knill showed [27] that any combination of Pauli and Hadamard gates, together with an entangling operation, can be simulated efficiently (i.e. in polynomial time) on a probabilistic classical computer. We therefore need to add one more gate, which gives a smaller rotation around the z -axis: the T or $\pi/8$ gate,

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} = e^{i\pi/8} \begin{pmatrix} e^{-i\pi/8} & 0 \\ 0 & e^{i\pi/8} \end{pmatrix}. \quad (1.3)$$

gravity theory, that does away with the need for dark matter [21]), and fruitful links between general relativity and quantum information theory [22, 23], two opposite ends of the field of Physics that famously appear incompatible (e.g., their estimates of the vacuum energy density are off by 121 orders of magnitude, predicting 10^{-9} and 10^{113} J/m³, respectively [24]). To bring us full circle, these links can actually be studied with quantum computers! [25, 26]

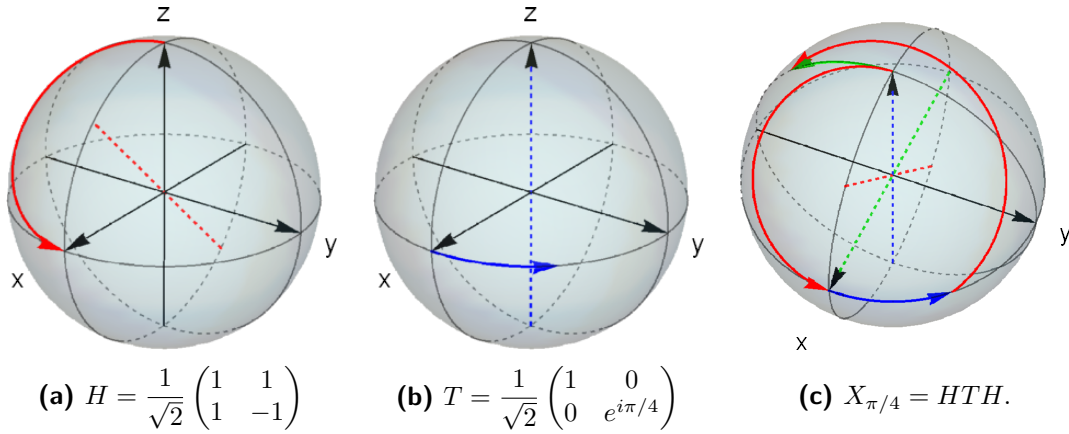


Figure 1.2: (a) The Hadamard gate H rotates by π around a diagonal axis, allowing for the creation of a superposition. (b) The T gate introduces a small rotation of only $\pi/4$ about the z -axis. (c) By combining the two, smaller rotations about other axes can be generated, which in turn can be combined to reach any point on the Bloch sphere.

Luckily, adding this final gate is enough to access any point on the Bloch sphere, and we do not need to include gates with infinitesimal rotations¹². It was proven by Kitaev and Solovay [29, 30] that as long as we have a gate that can perform a rotation by $\pi/4$ around one of the three axes, in addition to the Hadamard, we can approximate any other gate to arbitrary precision ϵ in logarithmic time $\mathcal{O}(\log(1/\epsilon))$.

Though a qubit state can be prepared at any point on the Bloch sphere, such detail can only ever be known to us in a statistical sense¹³. Every time a measurement is performed, all that we will get out is either a $|0\rangle$ or a $|1\rangle$ in the measurement basis, with the chance of either described by Born’s rule. This rule is usually stated as follows: when given a state Ψ , the probability of measuring outcome A is proportional to [24]

$$P(A) \propto |\langle A|\Psi \rangle|^2, \quad (1.4)$$

i.e. the absolute square of Ψ ’s amplitude over that state. But since physicists don’t quite agree what it means to *measure* something¹⁴, I prefer to follow Hossenfelder [34] and phrase it just as the transition probability from state Ψ_0 to state Ψ_1 :

$$P(|\Psi_0\rangle \rightarrow |\Psi_1\rangle) \propto |\langle \Psi_1|\Psi_0 \rangle|^2, \quad (1.5)$$

where Ψ_1 can then be some basis state aligned with a “measurement outcome” A (e.g. $|0\rangle$ or $|1\rangle$), and leave it at that. This relation, at least, is free of controversy, and can even be derived from first principles [34].

¹²These T gates are so essential to quantum computers’ advantage over classical computers, that some researchers estimate that they will make up around 90% of the gates in a typical quantum algorithm. Since they are costly to implement, efforts are underway to bring this number down [28].

¹³The state can also be *inside* the Bloch sphere, which means that there is less information about its state. This happens for example when two qubits are entangled, at which point the individual states no longer exist. Once you learn something about one of the qubits, the other will purify and move towards the surface of the Bloch sphere.

¹⁴Does the universe split into as many branches as there are basis states with nonzero amplitude [31], does the wavefunction “collapse” to classical certainty (Copenhagen interpretation), do we become entangled with the measured object and remain forever in superposition [32, 33], or should we just not be allowed to even ask this kind of question?

1.2.2 Two-qubit gates

Algorithms with exponential speedups are the holy grail of quantum computing, hence the fame of Deutsch’s black box problem [18], Shor’s algorithm for finding the prime factors of large integers [13], and Childs’ procedure for quantum random walks on a graph [35] (each forms the basis of their own class of algorithms, where they are used as subroutines). This impressive advantage relative to classical computers has its origin in the amount of information that can be stored in a set of entangled qubits. To see this, consider first the four possible ways that we can combine two classical bits:

$$00, \quad 01, \quad 10, \quad 11. \quad (1.6)$$

In general, though we can generate 2^n numbers with n bits, we only need n fundamental “units” of information [36] to describe their state. Two quantum bits can also be in these four configurations¹⁵,

$$|00\rangle, \quad |01\rangle, \quad |10\rangle, \quad |11\rangle, \quad (1.8)$$

but what sets them apart from classical bits is that they can be in any linear combination or *superposition* of them,

$$|\Psi\rangle = \alpha |00\rangle + \beta |01\rangle + \gamma |10\rangle + \delta |11\rangle. \quad (1.9)$$

So now, to fully describe the state, two numbers are no longer enough: we cannot just use one number to express the state of the first qubit, and then a second one to indicate that of the other. Neither can we extract such individual information: if we were to measure the first bit and got 0, and then the second one and got 0 again, then all that we would know is that $\alpha \neq 0$. Once two qubits become entangled, they give up their individual identity, and it no longer makes sense to talk about what state each of them is in. In turn, this loss of information is more than made up for by the exponential amount of it contained in the correlations between the two qubits. In general, we need $2^n - 1$ complex numbers to describe an n -qubit state, where the -1 comes from a global phase that can be ignored¹⁶.

To create these correlations, we need two-qubit entangling gates. The simplest of

¹⁵This choice of basis vectors is common in quantum computation. For some platforms it may be useful to instead choose a basis where three of the states have a total spin of 1 (the “triplet”), and one with zero spin (the “singlet”):

$$s = 1 : \begin{cases} |\uparrow\uparrow\rangle & (m = 1) \\ \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) & (m = 0) \\ |\downarrow\downarrow\rangle & (m = -1) \end{cases}, \quad s = 0 : \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \quad (m = 0), \quad (1.7)$$

where s is the amplitude of the spin, and m its projection along the z -axis. In platform-agnostic notation, $|0\rangle = |\uparrow\rangle$ and $|1\rangle = |\downarrow\rangle$.

¹⁶Of course the global phase of an entangled state *exists*, but in order to estimate it with the quantum phase estimation algorithm (QPE, based on the QFT discussed in section 1.3.3) [37], you need to introduce an extra qubit that you will measure, and then you still don’t know the global phase of *the whole system* including that extra qubit! Hence there is no useful information in the global phase.

these is the controlled-NOT (CNOT for short), which we draw as follows¹⁷.

$$\begin{array}{c}
 |\psi_1\rangle \text{---} \bullet \\
 | \\
 |\psi_2\rangle \text{---} \oplus
 \end{array}
 \quad (1.10)$$

The horizontal lines are the qubits, labeled $|\psi_1\rangle$ and $|\psi_2\rangle$, which are connected through the CNOT gate in such a way that the state of the first “controls” the second. If $|\psi_1\rangle = |0\rangle$, then nothing happens to $|\psi_2\rangle$, but if $|\psi_1\rangle = |1\rangle$, $|\psi_2\rangle$ will be flipped, which can be conveniently summarized in a truth table:

	input	output
CNOT :	$ 00\rangle$	$ 00\rangle$
	$ 01\rangle$	$ 01\rangle$
	$ 10\rangle$	$ 11\rangle$
	$ 11\rangle$	$ 10\rangle$

(1.11)

This CNOT is universal on classical machines, so we already see that we can always simulate a classical Turing machine on a quantum one. This gate by itself is not enough to generate entanglement though, which has the requirement that the states cannot be separated. If we really do have as input one of the two-qubit basis states, e.g. $|11\rangle$, such that we end up with $|10\rangle$ after the operation, this state can still be written as

$$|10\rangle = \underbrace{|1\rangle}_{\psi_1} \otimes \underbrace{|0\rangle}_{\psi_2}, \quad (1.12)$$

which means that we can get information about the first qubit without learning anything about the second: we can still talk about the two states separately. To properly entangle the two states such that their individual identity disappears, we need to combine the CNOT with a gate that creates a superposition, such as the Hadamard introduced earlier (Fig. 1.2):

$$\begin{array}{c}
 \psi_1 = |+\rangle \\
 | \\
 |\psi_1\rangle = |0\rangle \text{---} \boxed{H} \text{---} \bullet \\
 | \\
 |\psi_2\rangle = |0\rangle \text{---} \oplus
 \end{array}
 \quad (1.13)$$

Here we start with both qubits in the ground state $|0\rangle$, rotate the first by π around an axis tilted by $\pi/4$ from the z -direction (see Fig. 1.2) to create the superposition $|\psi_1\rangle = \sqrt{1/2}(|0\rangle + |1\rangle) = |+\rangle$ in the top line, and *then* apply the CNOT. Since the input state right before this last operation is now a superposition of $|00\rangle$ and $|10\rangle$ (which is still separable), this finally gives us the inseparable state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) \neq |\psi_1\rangle \otimes |\psi_2\rangle, \quad (1.14)$$

which cannot be rewritten as a single product of two single-qubit states.

¹⁷The \oplus symbol is used in math for “modulo plus”, which means that you add two numbers *modulo* the base. So if you use base 10, then $5 \oplus 7 = 12 \bmod 10 = 2$. In binary this is the same as the XOR (eXclusive OR) gate, and you can see in truth table (1.11) that the output of the target qubit is just the XOR of the two input states.

We can make other two-qubit gates using the same idea of controlling one with the other, by placing other single-qubit gates on the target timeline, e.g.

$$\begin{array}{c} |\psi_1\rangle \\ \text{---} \bullet \text{---} \\ | \\ |\psi_2\rangle \text{---} \boxed{Y} \text{---} \end{array}, \quad \begin{array}{c} |\psi_1\rangle \\ \text{---} \bullet \text{---} \\ | \\ |\psi_2\rangle \text{---} \boxed{Z} \text{---} \end{array}, \quad \begin{array}{c} |\psi_1\rangle \\ \text{---} \bullet \text{---} \\ | \\ |\psi_2\rangle \text{---} \boxed{H} \text{---} \end{array}, \text{ etc.} \quad (1.15)$$

The matrix form of these gates is easy to find when the first qubit is the control. In this case the matrix simply has the identity in the top-left, and the gate in the bottom-right:

$$\begin{array}{c} \text{---} \bullet \text{---} \\ | \\ \text{---} \oplus \text{---} \end{array} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \boxed{0} & \boxed{1} \\ 0 & 0 & \boxed{1} & \boxed{0} \end{pmatrix}, \quad \begin{array}{c} \text{---} \bullet \text{---} \\ | \\ \text{---} \boxed{Y} \text{---} \end{array} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \boxed{0} & \boxed{i} \\ 0 & 0 & \boxed{-i} & \boxed{0} \end{pmatrix}. \quad (1.16)$$

σ_x σ_y

When the second qubit is used as control, matrices can be found using truth tables like the one in eq. (1.11),

$$\begin{array}{c} \text{---} \boxed{Y} \text{---} \\ | \\ \text{---} \bullet \text{---} \end{array} = \begin{array}{cc} \text{input} & \text{output} \\ \hline |00\rangle & |00\rangle \\ |01\rangle & -i |11\rangle \\ |10\rangle & |10\rangle \\ |11\rangle & i |01\rangle \end{array} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{0} & 0 & \boxed{-i} \\ 0 & 0 & 1 & 0 \\ 0 & \boxed{i} & 0 & \boxed{0} \end{pmatrix}, \quad (1.17)$$

σ_y^T

where in general you'll find that you need to switch the middle rows and columns and transpose the target matrix¹⁸. Luckily we don't need all of these, since it turns out that the combination of single-qubit H and T , and the two-qubit CNOT (known as the "Clifford+T" gate set) is universal [38].

However, since this is a thesis about superconducting qubits, there is one more important gate that we need to discuss, that is especially relevant to transmon qubits: the i SWAP,

$$\begin{array}{c} \text{---} \boxed{i\text{SWAP}} \text{---} \\ | \\ \text{---} \end{array} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (1.18)$$

As will be discussed in more detail in section 1.4.3, in the transmon limit of large Josephson coupling and small charging energy (large capacitance), $E_J \gg E_C$, little energy is associated with the number of charges on either side of the Josephson junction, and the qubit states are instead mostly defined by the phase fluctuations across the junction (though they are not phase states [39]!). This means that (in the limit where the capacitance *between* the qubits is much smaller than the capacitances of the qubits themselves) a capacitive coupling between two transmons is *transverse* (orthogonal) to the quantization axis of the system [40], and the off-diagonal interaction causes an exchange of energy. The simplest capacitive interaction occurs when the two qubits have their resonance frequencies tuned to the same energy, such that there is no energy difference

¹⁸The logic here is that you could just list the basis states in a different order, such that the first qubit becomes the second and vice versa.

between $|01\rangle$ and $|10\rangle$, and the two-qubit spectrum hybridizes at a so-called avoided crossing [41]. At this point an excitation in one qubit can be transferred to the other without having to either release energy to, or absorb energy from the environment, and a lossless “swap” can be performed.

The matrix form of this interaction can be derived by first recognizing that there are two terms: a lowering operator on the excited qubit times a raising operator on the other, for each the two relevant input states $|01\rangle$ and $|10\rangle$ that can be swapped [42],

$$H_{\text{transv. int.}} = g \left(\underbrace{\sigma_+^{(1)}}_{\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}} \otimes \underbrace{\sigma_-^{(2)}}_{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}} + \sigma_-^{(1)} \otimes \sigma_+^{(2)} \right) = g \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (1.19)$$

$|10\rangle \rightarrow |01\rangle$

$|01\rangle \rightarrow |10\rangle$

This then gives rise to the unitary evolution [43]

$$U_{\text{transv. int.}} = e^{iHt} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(gt) & -i \sin(gt) & 0 \\ 0 & -i \sin(gt) & \cos(gt) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (1.20)$$

which causes the complex amplitudes associated with $|01\rangle$ and $|10\rangle$ to oscillate out of phase with a period¹⁹ of $\tau = 2\pi/g$. One might expect a swap to occur after half a period, but that would only flip the signs of each component. Instead, when the interaction is turned on for just a *quarter* oscillation, we get

$$U_{\text{transv. int.}} \left(t = \frac{\pi}{2g} \right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (1.21)$$

revealing the i in the name “ i SWAP”. Coupling the two qubits for an even shorter amount of time of only an eighth of a period, will perform half a swap or \sqrt{i} SWAP, turning either of the swappable two-qubit basis states into entangled pairs,

$$\sqrt{i}\text{SWAP} |01\rangle = \frac{1}{\sqrt{2}} (|01\rangle - i |10\rangle), \quad \sqrt{i}\text{SWAP} |10\rangle = \frac{1}{\sqrt{2}} (-i |01\rangle + |10\rangle). \quad (1.22)$$

This gate is thus a very efficient “native” or “primitive” means of creating large multi-qubit entangled states in circuits of tunable-frequency transmons, able to generate 2^n -qubit

¹⁹To see why we get 1s in the top-left and bottom-right corners of this 4×4 unitary time evolution matrix, while the Hamiltonian matrix itself only has nonzero entries in the central 2×2 block, expand the exponential $\exp(iHt)$:

$$\exp \left(\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \right) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \right)^n = I + \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} + \frac{1}{2} \left(\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \right)^2 + \dots,$$

where the identity gives us those corner 1s, while the higher-order terms will only change the central block. We then get sines and cosines by summing the odd (off-diagonal) and even (diagonal) matrix products, respectively, keeping the diagonal real since it only has even factors of the imaginary i .

entanglement after a circuit depth²⁰ n .

Though such entanglement will allow us to access an exponentially large Hilbert space, unfortunately this does not mean that we can also get exponentially much information out of the computation. The Holevo bound simply states that since we will always have to perform a measurement at the end, all we can ever expect is an output of as many classical bits as we use qubits [45]. This means that to have a speedup, interference of all these exponentially many quantum states needs to be orchestrated between the initialization and readout steps [18]. It is this design of interference patterns, in a way that depends on the specifics of the problem at hand, that we are concerned with when developing quantum algorithms.

1.3 Algorithms and computational complexity

“Shut up and calculate!”

– David Mermin

“OK, but what?”

– Scott Aaronson

Consciously or not, we use algorithms every day, and have in fact been using them for much longer than we have had digital computers. An algorithm is any set of logical operations that solves a problem, and can be as mundane as the following thought process that you may go through when choosing a pack of toilet paper in the supermarket:

Algorithm 1.1: Choose a product in the supermarket.

```

1 PriceWeight ← −2
2 QualityWeight ← 1
3 foreach brand in tpbrands:
4     brand.score ← PriceWeight * brand.Price
5     brand.score += QualityWeight * brand.PerceivedQuality
6 tpbrands ← sort(tpbrands, score)
7 buy(tpbrands[0])

```

More complex routines like baking a cake or writing a thesis can in principle also be reduced to similar sets of operations, as prescribed by the Church-Turing thesis discussed earlier [1].

One important property of an algorithm is its complexity in terms of time, energy, memory and other resources that it requires. These are in turn linked to the complexity of the problem that it is designed to solve, and are bound by the physical characteristics of the hardware it is running on. The extended Church-Turing thesis states that any efficiently computable function is also efficiently computable on a universal Turing machine, which means that there is at most a polynomial overhead when an algorithm is compiled to a form that a Turing machine can interpret. The following sections will pick apart how the existence of quantum algorithms have shattered this bedrock of classical computer science.

²⁰In qubit-speak, the “volume” is the square of the number of qubits that can effectively be used on your physical chip (or 2 to the power of that number if you work at IonQ [44]), the “size” of a logical circuit is the total number of gates, and the “depth” is the maximum number of sequential operations on a single qubit in terms of the gates native to the physical system.

1.3.1 Types of quantum algorithms

There are different ways that we can group quantum algorithms that have an advantage relative to classical ones. One popular way is to sort them into three broad groups by the type of problem that they solve [46].

Optimization: This can generally be rephrased as minimizing a cost function, such as the one shown in algorithm 1.1.

Problems: finding the ground state of an Ising Hamiltonian, compiling an investment portfolio or picking locations for vaccination centres.

Algorithms: Quantum annealing (QA), quantum approximate optimization algorithm (QAOA), and combinations thereof [47].

Quantum machine learning (QML): Reducible to sampling probability distributions, as problems can be reformulated to finding the probability that a function accepts (i.e. returns 1 or TRUE) over many degrees of freedom.

Problems: analyzing data from quantum circuits, possibly using quantum circuits to analyze classical data (image processing, unsupervised learning). Not yet clear what kind of problems can have quantum speedups [46].

Algorithms: quantum neural networks, Gaussian boson sampling with gradient descent [48], quantum circuit Born machines (QCBM) [49], variational quantum algorithms (VQA) [50].

Simulation: Quantum systems are intrinsically hard to simulate with classical machines [10, 51]; calculating the energy spectrum of a single large atom takes more computing power than all supercomputers in the world combined can deliver in any reasonable time, while all it takes on a quantum computer is a few hundred qubits [52–54].

Problems: quantum chemistry, condensed matter physics, calculating the ground state or the dynamics of any quantum system.

Algorithms: Hamiltonian simulation [52, 55], variational quantum eigensolver (VQE) [56].

At the time of this writing, the first category may seem to be off to a head start, as commercial superconducting quantum annealing chips are running optimization problems with thousands of qubits [57]. This is in part because it is perceived to have direct applications to lucrative disciplines [58], leading analysts to predict that it will continue to be an important part of the end-user market [59]. It should be noted however, that known quantum optimization algorithms give at most a polynomial speedup relative to classical alternatives, with practical examples typically solving problems only quadratically faster.

In these discussions, we generally only care about the *asymptotic* behavior of an algorithm; the way that the demand on some resource scales with the input size of the problem or accuracy of the output without caring about constant prefactors. We can describe the complexity of an algorithm with at least one of the following:

- Ω : the lower bound, best-case complexity. See e.g. section 1.3.2.
- \mathcal{O} (“big-oh”): the upper bound, worst-case complexity. See sections 1.3.3 and 1.3.5.

- Θ : both a lower and upper bound: the best and worst case scenarios scale in the same way. See section 1.3.4.

Figuring out the *type* of speedup that can be achieved is essential, since it is the way that the computation time scales with the size of the problem that give quantum computers their edge. Performing a single two-qubit operation takes on the order of 10^3 to 10^6 times longer than a simple switch of a transistor, even without taking into account the overhead introduced by error correction²¹. Fully fault-tolerant two-qubit operations may even be up to 10^7 times slower than classical two-bit operations, which means that a quadratic speedup with $n_Q = \sqrt{n_C}$ operations and a fault-tolerant gate time of $\tau_Q = 10^{-2}$ s still gives cross-over times on the order of

$$t = \tau_Q \frac{\tau_Q}{\tau_C} = 10^{-2} \text{ s} \times 10^7 = 10^5 \text{ s} \approx 1 \text{ day}. \quad (1.23)$$

More comprehensive estimates that also take into account classical parallelization and typical circuit implementations range from hundreds of days (10^3 classical cores) to thousands of years (10^6 cores) [65]. Improving the order of the polynomial speedup to cubic or quartic can bring these cross-over times down to mere hours or minutes. Exponential speedups such as those seen for finding the prime factors of large numbers [13] or traversing graphs [35] would give a more fundamental quantum advantage.

It may therefore be more productive to have a bottom-up classification of quantum algorithms, looking first at the kinds of asymptotic speedups that we know of for existing problems, focusing on those that are better than quadratic, and then finding real-world examples that they can be applied to. In this spirit, the following sections will briefly review the black-box problems of Deutsch and Simon, the period-finding Quantum Fourier Transform (QFT), the unsorted database search by Grover, and the quantum annealing and quantum approximate optimization algorithms.

1.3.2 Deutsch’s problem

Deutsch’s problem deals with what we call an “oracle”: some black-box function that we take to exist, and that we want to find information about. Though this specific problem was clearly contrived for the express purpose of demonstrating a quantum advantage, oracles are actually found in many real-life applications. Historically, Deutsch’s problem as it was posed in 1985 [18] was the first to show a concrete example of a situation (besides simulation) in which a quantum computer would give a speedup. Given a function $f(x)$ with $x, f(x) \in \{0, 1\}$, the problem is to find out whether $f(0) \oplus f(1) = 0$ or $f(0) \oplus f(1) = 1$. In the first case, the function is said to be “constant” since it gives the same output no matter the input, while in the second the function is “balanced”. To

²¹It is estimated that error correction requires anywhere from a 13 to a 10^4 physical to logical qubit ratio, depending on the two-qubit gate fidelity [60]. Ion traps have the advantage of higher fidelity, which also means that they can run larger circuits before even needing error correction at all, but a single gate operation takes around a thousand times longer than on a superconducting system [61, 62]. By the way, the existence of error correction [63] is guaranteed by the linearity of quantum mechanics, preventing errors from cascading into noise [12, 64].

illustrate, consider the four possible functions.

$f(0)$	$f(1)$	$f(0) \oplus f(1)$	
0	0	0	constant
0	1	1	balanced
1	0	1	balanced
1	1	0	constant

(1.24)

Any classical algorithm deciding whether the function is balanced will need to query f twice: once to extract $f(0)$, and then again for $f(1)$. Deutsch imagined a Turing-like machine where a quantum program $\pi(f, a, b)$ applies a function f to slot a and stores the result in slot b , thus performing the calculation²²

$$|\pi(f, 2, 3), i, j\rangle \mapsto |\pi(f, 2, 3), i, j \oplus f(i)\rangle. \quad (1.25)$$

If a superposition of $|0\rangle$ and $|1\rangle$ is placed in register i while register j is initialized in the ground state, then this program evaluates to

$$\frac{1}{\sqrt{2}} \sum_{i \in \{0,1\}} |\pi(f, 2, 3), i, 0\rangle \mapsto \frac{1}{\sqrt{2}} \sum_{i \in \{0,1\}} |\pi(f, 2, 3), i, f(i)\rangle. \quad (1.26)$$

We can then perform a simultaneous measurement on the last two qubits in the four-eigenstate basis

$$\left\{ \begin{array}{l} |\text{zero}\rangle \equiv \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle), \\ |\text{one}\rangle \equiv \frac{1}{2}(|00\rangle - |01\rangle - |10\rangle + |11\rangle), \\ |\text{fail}\rangle \equiv \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle), \\ |\text{error}\rangle \equiv \frac{1}{2}(|00\rangle + |01\rangle - |10\rangle - |11\rangle), \end{array} \right. \quad (1.27)$$

where the amplitudes for the outcomes of “constant” functions represented by the first and last rows in eq. (1.24) add up in the eigenstates $|\text{zero}\rangle$ and $|\text{fail}\rangle$, while they cancel out in $|\text{one}\rangle$ and $|\text{error}\rangle$. Similarly, the “balanced” outcomes add up in $|\text{one}\rangle$ and $|\text{fail}\rangle$, while they cancel out in $|\text{zero}\rangle$ and $|\text{error}\rangle$. Constant functions thus have a 50% chance of outputting $|\text{zero}\rangle$, while balanced functions have a 50% chance of outputting $|\text{one}\rangle$, and both also have a 50% chance of $|\text{fail}\rangle$, while $|\text{error}\rangle$ should never be observed. This means that if we are allowed only a single query of f , Deutsch’s algorithm has a 50% chance of answering the question of whether it is constant, while a classical algorithm would have 0% chance of solving it, giving it a complexity²³ $\Omega(1/2)$.

The algorithm can be improved to have a 100% chance of solving the problem by placing also the target qubit j in superposition [66]. This is most easily visualized using our familiar quantum circuit notation [37, 67]:

$$\left. \begin{array}{l} |0\rangle \xrightarrow{H} \text{---} i \text{---} \boxed{U_f} \text{---} i \xrightarrow{H} |f(0) \oplus f(1)\rangle \\ |1\rangle \xrightarrow{H} \text{---} j \text{---} \boxed{U_f} \text{---} j \oplus f(i) \text{---} \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \end{array} \right\} \times \text{global phase}. \quad (1.28)$$

²²The storage in slot b needs to be additive to ensure unitarity.

²³In this case, no asymptote can be calculated, since the problem only exists for a single n .

Here the unitary U_f is the circuit version of the program π described above, and performs a rotation in the two-qubit Hilbert space. The perhaps counter-intuitive result that the solution is output on the top channel i , while the function $f(i)$ is added to the bottom qubit, is due to a phenomenon called “phase kickback”, where the global phase is effectively *kicked back* to the control qubit (i in this case). This is seen more clearly when writing down the U_f operation explicitly:

$$\begin{aligned}
\frac{U_f}{2} [(|0\rangle + |1\rangle) \otimes (|0\rangle - |1\rangle)] &= (|0\rangle + |1\rangle) \otimes \frac{e^{i\pi(f(0)+f(1))}}{2} (|0\rangle - |1\rangle) \\
&= \frac{e^{i\pi(f(0)+f(1))}}{2} (|0\rangle + |1\rangle) \otimes (|0\rangle - |1\rangle) \\
&= \begin{cases} (i) (|0\rangle + |1\rangle) \otimes (|0\rangle - |1\rangle)/2, \\ (ii) (|0\rangle - |1\rangle) \otimes (|0\rangle - |1\rangle)/2, \end{cases}
\end{aligned} \tag{1.29}$$

where options (i) and (ii) are:

$$\begin{aligned}
(i) \quad f(0) \oplus f(1) &= 0 \quad (\text{constant}), \\
(ii) \quad f(0) \oplus f(1) &= 1 \quad (\text{balanced}).
\end{aligned} \tag{1.30}$$

The final Hadamard on the first qubit ensures that we can measure the output in the z -basis.

While the above version of the problem with only a single control qubit is important for having established that quantum computers can be faster in principle, the n -qubit version proves that even exponential speedups can be achieved [9, 66, 68]. In this more general problem, the input for function f is no longer just 0 or 1, but an n -bit string of 0s and 1s, and f is said to be balanced iff the single-bit output is 0 for exactly half the inputs. Since there are 2^n possible bit strings, this problem would require $2^n/2 = 2^{n-1}$ queries of the oracle in a classical setting. A quantum computer could solve this with a single query and by measuring only n qubits, using a modified version of eq. (1.28) [37, 66, 69]:

$$\begin{aligned}
&|0\rangle^{\otimes n} \xrightarrow{H} \text{---} i \text{---} \boxed{U_f} \text{---} i \xrightarrow{H} 2^{-n} \sum_{x,y \in \{0,1\}^n} e^{i\pi x \cdot z + f(x)} |y\rangle \\
&|1\rangle \xrightarrow{H} \text{---} j \text{---} \boxed{U_f} \text{---} j \oplus f(i) \text{---} \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle)
\end{aligned} \tag{1.31}$$

In this larger circuit, if the function is balanced, then there is at least one bit in the input string for which flipping it changes the outcome,

$$\exists m \in \{0, \dots, n-1\} : f(|\dots 0_m \dots\rangle) \oplus f(|\dots 1_m \dots\rangle) = 1. \tag{1.32}$$

For example, if all odd inputs x give $f(x) = 0$ while all even ones give $f(x) = 1$, such that the function only depends on the last qubit in the register ($m = n-1$), then that qubit will get a phase kickback. In general, any balanced function will necessarily have at least one (and possibly all) of the control qubits flip to $|1\rangle$ in eq. (1.31), and finding a single nonzero entry is sufficient to determine that the function is balanced, giving us $\Omega(1)$ and $\mathcal{O}(n)$. In contrast, there will be no phase kickback at all to any of the n control

qubits if the function is constant, and the output will be $|0\rangle^{\otimes n}$. It takes at most one query of the oracle and n measurements to analyze the function, an exponential speedup relative to the 2^n queries and measurements on a classical machine.

This algorithm is an example of “quantum parallelism”, and Deutsch famously remarked that the speedup due to it implies that Everett’s multiverse interpretation of quantum mechanics²⁴ [32] must be correct, for else [18, 31], *where was it computed?* The term has unfortunately also given rise to one of quantum computer scientists’ pet peeves in popular communications: that quantum computers somehow allow for “massive parallel computation” of any classically parallelizable circuit, simply performing all calculations in parallel in the exponentially large Hilbert space. Parallel operations can indeed be performed on all basis states in a superposition, but the idea of unbridled parallelism falls apart at the actual measurement of the output: while the information stored inside the register can in the case of maximum entanglement reach 2^n bits, only n bits of classical information can ever be read out. This limit is known as the Holevo bound [45], and suggests a requirement for achieving a speedup: interference between the possible outcomes must be orchestrated, such that the information about the quadratic or exponential number of parallel computations can be read out from only a linear number of qubits.

Deutsch’s efforts inspired Simon to construct a similar problem [70, 71], where instead of a single bit of information (whether a function is balanced or not), an entire bit string can be extracted by using n target qubits on the bottom line of eq. (1.31). It was upon seeing this result that Shor realized a method for finding the prime factors of integers [13], which we will discuss next.

1.3.3 The Quantum Fourier Transform and Shor’s algorithm

Any periodic function can be expressed as a (possibly infinite) series of sines and cosines, written most compactly using Euler’s identity:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/L}, \quad c_n = \frac{1}{2L} \int_{-L}^L f(x) e^{-in\pi x/L} dx. \quad (1.33)$$

As an extension of this idea to non-periodic functions by letting $L \rightarrow \infty$, the Fourier transform takes as input such a function $f(x)$ and outputs its conjugate $\hat{f}(w)$, a continuous equivalent of the coefficients²⁵ c_n [72]:

$$\begin{aligned} \mathcal{F}(f(x)) &= \hat{f}(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{iwx} dx, \\ \mathcal{F}^{-1}(\hat{f}(w)) &= f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(w) e^{-iwx} dw. \end{aligned} \quad (1.34)$$

The relevance of the Fourier transform to physics is perhaps clearest from the concept of conjugate variables, the transformation between which in the form of series expansions not only provides access to mathematical techniques [72, 73], but also gives more direct

²⁴Everett never called it that, but used the name “relative state interpretation”, insisting that the observed outcome of a measurement necessarily depends on the final state of the observer, just like the velocity of an object can only be measured relative to a rest frame.

²⁵Perhaps you prefer to normalize these transforms asymmetrically by placing a factor $1/2\pi$ only in front of the inverse Fourier transform, but then I suppose you also follow the mostly minuses convention of the metric tensor.

insights from quantum mechanical uncertainty and commutation relations [74]. The importance to classical physics was stressed by Noether [75], who pointed out that the symmetry of a system under differentiation to any variable implies a conservation of its conjugate (e.g. time invariance implies energy conservation, translational invariance implies momentum conservation etc). The transform also plays a role in computer science, where its application to signal processing is ubiquitous, and the bit-wise product representation that we will see below is a prime example of the divide and conquer strategy to solving recursive problems. Finding a faster way of performing this operation could thus have consequences for a broad range of studies and applications.

Analytical methods to perform the Fourier transform are often unavailable, and we have to resort to numerics. In this case we replace the infinite integrals in eq. (1.34) by finite sums,

$$\begin{aligned}\mathcal{F}(\mathbf{f}) &= \hat{\mathbf{f}}, & \hat{f}_k &= \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} f_j e^{ikx_j}, \\ \mathcal{F}^{-1}(\hat{\mathbf{f}}) &= \mathbf{f}, & f_j &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{f}_k e^{-ikx_j},\end{aligned}\tag{1.35}$$

where \mathbf{f} and $\hat{\mathbf{f}}$ are vectors of size N , and x_j are N points at which the function f is evaluated (often equally spaced at $x_j = 2\pi j/N$). Calculating all N components of the transformed vector thus amounts to computing the multiplication by an $N \times N$ matrix, which naively takes $\mathcal{O}(N^2)$ logical operations, or $\mathcal{O}(N \log N)$ when the matrix is recursively divided (divide and conquer).

Now imagine that we store the input vector \mathbf{f} of size N in a $\log_2 N = n$ qubit string using the exponential scaling of the storage capacity as discussed around eq. (1.9). When the j th component of the vector $\hat{\mathbf{f}}$ in eq. (1.35) is calculated, we don't store it in the j th register as we would classically, but assign it instead as an amplitude to the j th term in the n -qubit superposition. For example, if we would have 2 qubits, we could store a four-component input vector \mathbf{f} like so:

$$\mathbf{f} = f_0 \underbrace{|00\rangle}_{j=0} + f_1 \underbrace{|01\rangle}_{j=1} + f_2 \underbrace{|10\rangle}_{j=2} + f_3 \underbrace{|11\rangle}_{j=3}.\tag{1.36}$$

We can then write the definitions of the discrete Fourier transform in eq. (1.35) in the appropriate bracket notation [13],

$$\hat{\mathbf{f}} = \sum_{k=0}^{N-1} \hat{f}_k |k\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} f_j e^{ikx_j} |k\rangle,\tag{1.37}$$

where each of the basis states $|j\rangle = |j_0 \cdots j_{N-1}\rangle$, such as the four in eq. (1.36), contributes

$$\mathcal{F}(|j\rangle) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{ikx_j} |k\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{2\pi i j k / N} |k\rangle.\tag{1.38}$$

Here we assumed in the last step that the function is restricted to²⁶ $[0, 2\pi)$ and that x_j represents it optimally. Taking again the 2-qubit basis of eq. (1.36) as example, this

²⁶That would be $[0, 2\pi[$ in French notation.

would result in

$$\begin{aligned}
\mathcal{F}(|j_0 j_1\rangle) &= \frac{1}{\sqrt{4}} \sum_{k=0}^3 e^{2\pi i [j_0 j_1] k/4} |k\rangle \\
&= \frac{1}{2} \left(|00\rangle + e^{\pi i [j_0 j_1]/2} |01\rangle + e^{\pi i [j_0 j_1]} |10\rangle + e^{\pi i [j_0 j_1] 3/2} |11\rangle \right) \\
&= \frac{1}{2} \left(|0\rangle + e^{i\pi \cancel{j_1} j_2} |1\rangle \right) \otimes \left(|0\rangle + e^{i\pi [j_0 j_1]/2} |1\rangle \right),
\end{aligned} \tag{1.39}$$

where $[j_0 j_1]$ is the binary number represented by the state $|j\rangle$, and we can drop any full rotations by 2π (hence the $\cancel{j_1} j_2 = [j_1 j_2] \bmod 2$). In general, for an n -qubit register we find [76],

$$\hat{f}_k = 2^{-n/2} \left(|0\rangle + e^{2\pi i 0 \cdot j_{n-1}} |1\rangle \right) \otimes \cdots \otimes \left(|0\rangle + e^{2\pi i 0 \cdot j_0 \cdots j_{n-1}} |1\rangle \right), \tag{1.40}$$

where $0.j_i \cdots j_n = (j_i \cdots j_{n-1})/2^n$.

This expansion into product form is not unique to the qubit register, and can also be done with exponentially more classical bits, where the components of the vector \mathbf{f} are assigned to the $N = 2^n$ bits themselves rather than the 2^n basis states:

$$\begin{aligned}
\text{Qubits:} \quad \mathbf{f} &= f_0 \underbrace{|00\rangle}_{j=0} + f_1 \underbrace{|01\rangle}_{j=1} + f_2 \underbrace{|10\rangle}_{j=2} + f_3 \underbrace{|11\rangle}_{j=3}, \\
\text{Bits:} \quad \mathbf{f} &= f_0 \underbrace{[0001]}_{j=0} + f_1 \underbrace{[0010]}_{j=1} + f_2 \underbrace{[0100]}_{j=2} + f_3 \underbrace{[1000]}_{j=3} = [f_3 f_2 f_1 f_0].
\end{aligned} \tag{1.41}$$

The exponential factors that we saw in eqs. (1.39) and (1.40) are now associated with whether we have the odd or even component of successive divisions of the problem,

$$\begin{aligned}
\mathcal{F}([j]) &= \frac{1}{\sqrt{4}} \sum_{k=0}^3 e^{2\pi i j k/4} |k\rangle \\
&= \frac{1}{2} \left(\underbrace{[0001]}_{j=0} + \underbrace{e^{\pi i j/2}}_{j \text{ odd}} \underbrace{[0010]}_{j=1} + \underbrace{e^{\pi i j}}_{\substack{[j/2] \text{ odd} \\ j \text{ odd}}} \underbrace{[0100]}_{j=2} + \underbrace{e^{\pi i j/2}}_{j \text{ odd}} \underbrace{e^{\pi i j}}_{\substack{[j/2] \text{ odd} \\ j \text{ odd}}} \underbrace{[1000]}_{j=3} \right).
\end{aligned} \tag{1.42}$$

This gives the idea of the divide and conquer approach to calculating the fast Fourier transform (FFT), where the problem is recursively divided into $\log_2 N = n$ sub-problems. Even though this is an exponential speedup in the calculation of the prefactors, the classical form still requires the evaluation of all $N = 2^n$ terms, leading to a total complexity of $\mathcal{O}(n2^n) = \mathcal{O}(N \log N)$.

The circuit for the quantum Fourier transform can be read almost directly from the product representation in eq. (1.40). Each of the qubits is on the equator of the Bloch

sphere, rotated by some angle:

$$\frac{1}{\sqrt{2}}(|0\rangle + e^{i\phi}|1\rangle) = \text{Bloch sphere diagram} \quad (1.43)$$

This angle, $2\pi i j_l / 2^m$, is a rotation 2^{-m} times around the z -axis of the Bloch sphere, controlled by qubit $|j_l\rangle$:

$$\frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i \overbrace{0.0 \dots j_l}^{m \text{ zeros}}} |1\rangle) \otimes |j_l\rangle = \sqrt{1/2}(|0\rangle + |1\rangle) \text{---} \boxed{2^m \sqrt{Z}} \text{---} |j_l\rangle, \quad (1.44)$$

where $2^m \sqrt{Z} = R_m$. This suggests a simple way of constructing the final state in eq. (1.40), shown in pseudocode in algorithm 1.2.

Algorithm 1.2: Quantum Fourier Transform (QFT).

```

1 foreach Qubit in Register:
2   rotate Qubit basis states to the equator
3   foreach ControlQubit in Register[Qubit.index+1::]:
4     target Qubit for a rotation
5     rotate by  $2^{-(\text{ControlQubit.index} - \text{Qubit.index})}$ 

```

We can now write the QFT in circuit form:

$$\begin{array}{ccccccc} |j_0\rangle & \text{---} H & \text{---} \sqrt[2]{Z} & \text{---} \dots & \text{---} \sqrt[2^n]{Z} & \text{---} \dots & |k_n\rangle \\ |j_1\rangle & \text{---} & \text{---} \bullet & \text{---} \dots & \text{---} H & \text{---} \dots & |k_{n-1}\rangle \\ \vdots & & & & & & \vdots \\ |j_n\rangle & \text{---} & \text{---} \bullet & \text{---} \dots & \text{---} \bullet & \text{---} \dots & |k_0\rangle \end{array} \quad (1.45)$$

All in all, since each qubit line is targeted by at most $n - 1$ other qubits, this only takes $\mathcal{O}(n^2)$ gates, exponentially fewer than the $\mathcal{O}(n2^n)$ required for the classical algorithm.

As we saw above, the quantum Fourier transform allows us to efficiently find the periodicities in a string of numbers. Shor discovered that period finding can in turn, with many steps in between, be used to calculate with high probability the prime factors of an integer [13], something exponentially hard classically²⁷. There really isn't room here to discuss all these steps²⁸, which can be found in popular form in Ref. [77] and in even more popular form on pp. 122–125 of Ref. [54]²⁹. The best-known direct implication of

²⁷The best classical algorithm is still the 2200-year old “sieve of Eratosthenes”.

²⁸This is a thesis about silicon transmons after all.

²⁹An anecdote oft told by the late Jonathan Dowling gives some color to the historical fact that quantum computing as a field of research used to be almost entirely funded by the American security agencies. See p. 121 of his 2013 book [54] or this webcomic: <https://www.smbc-comics.com/comic/jonathan-dowling>.

an efficient algorithm for prime factoring is the insecurity of RSA encryption [15], which relies on the hardness of this one-way problem (calculating $n = pq$ is easy given primes p and q , but finding those factors given n is hard). Though research on “post-quantum cryptography” is under way [3], the hardness of predicting whether a given NP problem is unsolvable on a quantum computer (the relation of BQP to NP is currently unknown) means that the reliability of any non-quantum method of encryption is questionable [54, 78].

By demonstrating an exponential speedup for a common classical problem, Shor’s algorithm provided strong evidence that quantum computers could give significant speedups in areas of practical use. The QFT and the closely related phase estimation and order finding algorithms are all used to efficiently find global properties of sets of numbers, something known to be hard classically. They are therefore starting points in the search for quantum solutions to classically hard problems, giving hope of further exponential speedups.

1.3.4 Grover’s algorithm

Often times a technical problem can be reduced to finding an element in an unstructured list. You may have a set of possible answers, only some of which are actual solutions, and you know how to recognize those elements once you see them [79]. If you know how to recognize the solution, then you could devise some function $f(x)$ that gives a phase shift iff x is a solution, and leaves the element unchanged if it isn’t. Grover found that if you use this function $f(x)$ as an oracle, you can extract the solutions from the list in time proportional to only the square root of the size of that list³⁰. This is a quadratic speedup relative to classical algorithms, where unstructured searches necessarily need to check a number of elements linear in the size of the list (on average, you’ll find all solutions half-way through if there’s one, two-thirds through if there are two, etc).

Grover’s algorithm uses two unitary operators, one of which is the operator mentioned above, which we can write as

$$U_{\text{oracle}} |\psi\rangle = e^{i\pi f(\psi)} |\psi\rangle \quad : \quad |\psi\rangle \equiv \boxed{U_f} \equiv e^{i\pi f(\psi)} |\psi\rangle . \quad (1.46)$$

The other is a “diffuser” unitary, which rotates the multi-qubit state around the uniform superposition of all states. This rotation can be understood as follows. Recall that we can construct the Pauli matrices from the cross product of their positive eigenvectors³¹,

$$\begin{aligned} |x\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, & \sigma_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 2|x\rangle\langle x| - 1, \\ |y\rangle &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, & \sigma_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = 2|y\rangle\langle y| - 1, \\ |z\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, & \sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 2|z\rangle\langle z| - 1, \end{aligned} \quad (1.47)$$

³⁰Unfortunately, for black-box problems like these, Grover’s quadratic speedup is the best you can ever get, even on quantum computers [80].

³¹Or the negative ones, using $1 - 2|-z\rangle\langle -z|$ etc.

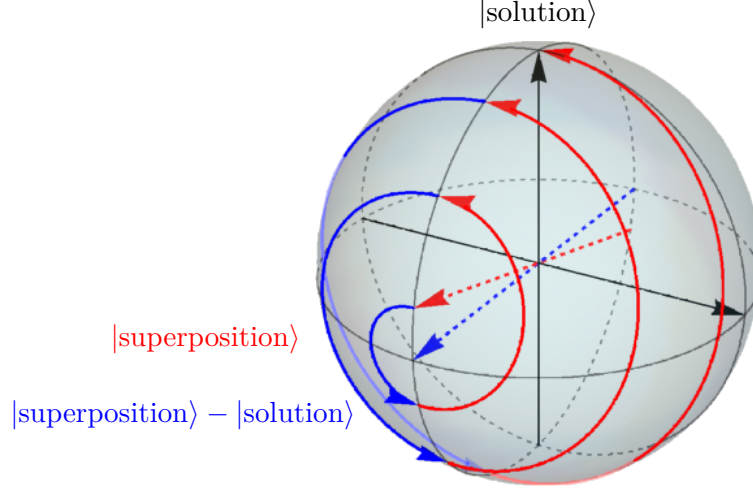


Figure 1.3: Grover's algorithm visualized: The Bloch sphere is a three-dimensional slice through Hilbert space, with the z -axis the superposition of solution states, the x -axis the states orthogonal to those, and the angle around the x -axis representing the global phase. The blue rotations are multiplications by Grover's oracle, while the red rotations are diffuser operations.

where σ_x flips bits that are pointing along the y and z -axes, σ_y flips those along the x and z -axes, etc. We can now construct a unitary $\sigma_{\text{superposition}}$ that flips multi-qubit states that point along the uniform superposition over all states:

$$|s\rangle = |\text{superposition}\rangle = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \sigma_s = 2|s\rangle\langle s| - 1. \quad (1.48)$$

This operation will rotate the qubit register in Hilbert space in such a way that its projection onto the uniform superposition is conserved, but any components orthogonal to it are flipped. As we will see below, this “diffuses” the wave function over the basis states by rectifying the phases relative to some tilted axis.

In Fig. 1.3 a kind of Bloch sphere is shown, where instead of the usual $|0\rangle$ and $|1\rangle$ along the z -axis and their superpositions with different phases on the sphere at some angle in between, the z -axis now represents the superposition of all solution states. Orthogonal to the solutions are all the non-solutions, the wrong answers to the problem, indicated in blue. At some angle between the two (a small angle if most states are not solutions), we can find the vector that represents the uniform superposition over all states in the two sets. You can now see that if the register is initialized in the uniform distribution, applying the oracle U_{oracle} will change only the phases of the solution states, effectively flipping its direction on the solution axis. Since the non-solution and uniform superposition axes are a bit off, applying the diffuser by rotating around the superposition then brings us closer to the solution state than we were before. After repeating the two operations on the order of $\mathcal{O}(\sqrt{N})$ times (the prefactor depends on the relative share of solution states), we will have a high probability of finding the solution when we perform a measurement.

$$|0\rangle^{\otimes n} \xrightarrow{n} H^{\otimes n} \xrightarrow{\times \mathcal{O}(\sqrt{N})} \boxed{U_f \quad U_s} \xrightarrow{\text{Measurement}} |\text{solution}\rangle \quad (1.49)$$

Algorithm 1.3: Grover’s algorithm.

```
1 Apply the Hadamard to each qubit // Create uniform superposition
2 foreach AmplificationStep in 0::sqrt(Register.length):
3     Apply the oracle // Flips phase of solution qubit
4     Apply the diffuser // Rotates around uniform superposition
5 Measure the state // Extract solution
```

1.3.5 Quantum annealing (QA) and the quantum approximate optimization algorithm (QAOA)

Quantum annealing is different from all the algorithms that we have discussed above in that it is not circuit based, i.e. it is not performed on quantum Turing machines, but on dedicated systems. The name is aptly chosen, as we can compare the process directly to the thermal annealing that we will discuss later on in chapter 3. When a material, such as the silicides that we study for Josephson field effect transistors, is first formed, it is not yet in its lowest-energy state. The Gibbs energy may be lowered significantly by aligning the crystal orientations in neighboring grains and removing the boundaries between them, but reaching that lower-energy state requires us to cross an energy barrier associated with all the chemical bonds that need to be broken. By annealing the material, sufficient heat is supplied to overcome this barrier, and the atoms are stimulated to move into configurations that are energetically favorable, lowering the Gibbs energy and leading to superior material properties such as higher superconducting critical temperatures.

In the problems solved by quantum annealing, we are faced with a similar issue: we want to find out what the ground state of a system is, but if we just immediately apply its Hamiltonian to a set of qubits, the resulting state is rarely in the global minimum. Instead, the annealing approach is to start out from a state that we know how to initialize, for example with all qubits in $|0\rangle$,

$$\mathcal{H}_{\text{init}} = \sigma_z^n. \quad (1.50)$$

The trick is then to heat up the system, such that the wave function spreads out from this initial state, while slowly changing the applied Hamiltonian to the one we want to find the ground state of,

$$\mathcal{H}(t) = u(t)\mathcal{H}_{\text{init}} + (1 - u(t))\mathcal{H}_{\text{problem}}, \quad (1.51)$$

after which the system is cooled down again. This can of course be done classically as well³², but the advantage of QA is that the escape from local minima is sped up by quantum tunneling. Current research focuses on what the ideal crossover function $u(t)$ is for specific problems [47, 83]. An especially interesting problem Hamiltonian is the Ising spin glass [84],

$$\mathcal{H}_{\text{Ising}} = \sum_{i < j} J_{ij} \sigma_z^{(i)} \sigma_z^{(j)}, \quad (1.52)$$

where finding solutions for $J_{ij} > 0$ (get as many neighbors as possible to have opposite spin) is in NP-complete. This class of problems is of special relevance since it contains the hardest problems in NP, a solution to any of which can be converted to a solution

³²The classical analog is called “simulated annealing”, and though it is sometimes used as a reference point by D-wave [81], it is not the fastest classical algorithm for finding approximate solutions to this problem [82].

to any other NP-complete problem in polynomial time. This means, for example, that if you can find an approximate ground state of the Ising model, you can find good solutions to the traveling salesperson problem. It is currently unclear what kind of speedup quantum annealing can offer [85], but it is believed that it is at most polynomial, likely quadratic [46, 86].

Quantum approximate optimization algorithms (QAOA) are a circuit-based equivalent to quantum annealing, where instead of slowly shifting from one Hamiltonian to another, the two Hamiltonians are applied sequentially for varying amounts of time [83, 87]. Again one starts out in an accessible ground state, but then instead of only slowly turning on the solution Hamiltonian, it is immediately applied for a relatively long stretch of time, ensuring that the system ends up at least close to some local minimum. Instead of heating up the system, the initialization Hamiltonian is then applied again to shake things up, and these two steps are repeated for a number of times. Perhaps an intuitive analogy is filling a cup with flour: after you tap a full cup on the counter a few times, it condenses until it is maybe only half full³³. The level to which the cup is filled is our cost function, the precise locations of all the wheat granules is the approximate solution to the problem, and the number of times we need to tap is the computational complexity.

One important result is that our final error, which we define as the difference between the energy of the final measured state and that of the actual ground state, goes down quadratically,

$$\text{Error}_{\text{QAOA}} = \langle \psi_{\text{final}} | \mathcal{H}_{\text{Ising}} | \psi_{\text{final}} \rangle - \langle \psi_0 | \mathcal{H}_{\text{Ising}} | \psi_0 \rangle \sim \frac{1}{n_{\text{steps}}^2}. \quad (1.53)$$

In other words, if you want n significant digits, the time complexity is $\mathcal{O}(\exp(n/2))$, while the optimal performance of classical computers is unknown [87, 88]. The catch is that you only get this optimal error scaling iff you use the optimal timing of applying the two Hamiltonians, which actually depends on the details of the problem [83]. It was also shown recently that QAOA even offers no quantum speedup at all if there are limits to the circuit depth (a real concern for NISQ devices) [88].

1.4 Physical implementations of quantum computers

“Any physical experiment can be regarded as a computation, and any computation is a physical experiment.”

– David Deutsch

The brief descriptions below serve only to illustrate the variety of platforms that quantum circuits can be implemented on, and to give a rough idea of how a quantum system can be turned into a qubit. For a more exhaustive, yet accessible overview, see e.g. pp. 153–171 of Ref. [54].

1.4.1 Ion traps

The physical implementation of quantum computing arguably started in ion traps, for which a two-qubit CNOT gate was proposed in late 1994 by Cirac and Zoller [89]. This

³³Always measure the weight of your ingredients, never the volume!

gate was quickly demonstrated by Wineland and Monroe [90], who had been working with ion traps for a long time in the context of atomic clocks. The development of theoretical understanding of gate operations in this physical system is of direct relevance to all other “platforms” (two of which will be discussed below), since *“if the math is the same, the physics must be the same!”* [54]. The cavity QED developed for gates based on ion-photon interactions [91–93] later formed the basis for circuit QED [94], which is the framework for superconducting qubits.

Ions are perhaps the most “natural” of the platforms, as the qubit states are encoded into the energy levels of a single electron in a half-occupied orbital of an ionized atom. These atoms are suspended by a rapidly oscillating electromagnetic field in ultra-high vacuum, and are operated on with laser beams so precise that they can be controlled almost down to individual photons [95]. The fact that electromagnetic trapping is even possible at all is perhaps counter-intuitive, as most freshman courses in electrodynamics will have you prove Earnshaw’s theorem [74, p. 115]:

“A charged particle cannot be held in a stable equilibrium by electrostatic forces alone.”

This is a direct consequence of the fact that the electric potential at any point in space is the spatial average of the potential at any fixed radius around it (Laplace’s equation, valid as long as there are no charges inside the sphere). The clever way around this that got Wolfgang Paul the Nobel prize in Physics in 1989 is to create an electric field that has a local minimum along one axis that simultaneously is a local maximum along another, and then rotate these axes [96].

Next we need to define the qubit levels in these trapped particles. The Pauli exclusion principle tells us that we cannot have multiple fermions like electrons occupying the same state. While the orbital quantum numbers uniquely define the orbits that electrons can be in, the spin degree of freedom ($|\uparrow\rangle$ and $|\downarrow\rangle$) still allows us to have two electrons in the same “place”. If we now make an ion by removing one of the outer electrons from an atom with an even atomic number (e.g. ^{12}Mg , ^{20}Ca , ^{48}Cd or IonQ’s ^{70}Yb), we are sure that the remaining electron is alone in its orbit. Once this ion is trapped inside the oscillating field, we can encode the qubit $|0\rangle$ and $|1\rangle$ states in its lone outer electron being excited or not. If the qubit were to decay from an excited $|1\rangle$ state, a photon would be emitted, which means that a superposition $\sqrt{1/2}(|0\rangle + |1\rangle)$ is shared with the electromagnetic cavity both containing a photon and not. Lasers can then be used both to excite the ions and to bring them back to the ground state. When the same laser beam acts on two ions, a photon can be in a superposition of having been absorbed by the one or the other, effectively entangling the two qubits.

This platform is especially well placed for integration with quantum communication systems, which as a long-distance technology necessarily operate with optical-range photons³⁴ [97]. While the photons carried by the resonators in superconducting circuits are on the order of a few GHz (times the Planck constant) [98], the energy levels in trapped ions are separated by energies on the order of hundreds of THz [99]. To quote Dowling and Milburn, *“at optical frequencies, the world is very cold ($hf \gg k_B T$)”* [100]. As a result, while superconducting circuits may need supercooled data links at ultra-high vacuum to couple distant chips [101], ion traps can be linked with simple room-temperature, ambient-pressure photonic interconnects [44, 99, 102].

³⁴At lower energies the thermal noise would be too strong, at higher frequencies you will have trouble building lenses.

1.4.2 Trapped spins in semiconductors

While atoms may be the “natural” candidate for qubits, as each of them is identical to — even indistinguishable from — any other of the same isotope, we are limited in our choice of parameters by the number of stable elements³⁵. This motivates the design of “artificial atoms”, quantum systems where the energy levels can be engineered, such as quantum dots defined by electrostatic gates in semiconductors, or superconducting circuits. Just like in ordinary atoms, qubit states can be defined by the occupied orbital or the spin of an electron on the dot, or even the total number of electrons residing there [63, 103].

Besides their tunability, the interest in quantum dots defined in semiconductors is motivated by the already present technological prowess in the fabrication of micro-electronics in these materials. This is important not just for the manufacturing of the qubits themselves. Fault-tolerant quantum computers will require real-time error correction, performed in part by classical logic and control circuits, which in a scalable quantum computer should be implemented as close to the qubits as possible [104]. The possibility of on-chip integration of classical CMOS circuitry [105] minimizes time delays, removes any losses of signal due to the coupling of long microwave lines to electromagnetic radiation from the environment, and reduces the hardware requirements on dilution refrigerators and room-temperature instruments.

In most physical implementations pursued today, the qubit degree of freedom is encoded in the spin orientation of an electron trapped on a quantum dot defined in semiconducting material by a gate or set of gates [106]. Microwave signals sent through a gate that is capacitively coupled to the dot can be used to perform operations on a single qubit, as well as to read out its state. Electrons on nearby dots can be brought closer together to create entanglement by changing the electrostatic environment, while they can be coupled over longer distances with microwave resonators.

1.4.3 Superconducting qubits

The focus of this thesis is a different kind of artificial atom, one where the energy levels are defined by the quantized resonance frequencies of an LC resonator.

In its most common form, a superconducting qubit consists of a capacitively shunted inductor, where the inductance is provided by a superconducting weak link. The oscillations in the resonator can be understood as a constant tension between on the one hand the desire of the charges on opposite sides of a capacitor to even out by flowing around through the inductor, and on the other hand the resistance of that inductor against any change in the amount of current that flows through it. By the time the capacitor is empty, the inductor will make sure the current keeps flowing, until an equal but opposite charge has built up again. While macroscopic circuits of this kind would slowly lose energy through heating and inductive coupling to the environment, the oscillations on microscopic superconducting circuits are protected by the same principle of energy level quantization that prevents the electrons in real atoms from spiraling into the core. It is these quantized levels that we use as our qubit states.

There are multiple degrees of freedom in these systems that can all be used as quantization axes for the qubit states, such as the number of Cooper pairs [107] or the superconducting phase [108] on one side of the capacitor, a combination of the two [109],

³⁵There are around 120 elements out there, only about 80 of which are stable. The number of options goes down further if we restrict ourselves to those with even atomic number, level spacings in the frequency range of existing lasers, non-toxicity, etc.

or even the flux through the split Josephson junction that makes up the inductor [110]. We will focus on designs where the qubit states are a superposition of different numbers of Cooper pairs having collected on either side of the capacitor [39]. In this limit relatively little energy is associated with a Cooper pair crossing the junction, $E_J \gg E_C$, and the energy levels are primarily defined by the Josephson operator. This operator couples charges states with different numbers of Cooper pairs [107],

$$\mathcal{H}_J = -\frac{E_J}{2} \sum_{n \in \mathbb{Z}} \left(|n\rangle \langle n+1| + |n+1\rangle \langle n| \right) = -\frac{E_J}{2} \begin{pmatrix} \ddots & & & \emptyset \\ & 0 & 1 & \\ & 1 & 0 & 1 \\ & & 1 & 0 \\ \emptyset & & & \ddots \end{pmatrix}, \quad (1.54)$$

such that any eigenstate is a symmetric superposition of Cooper pair occupation numbers around some average $\langle n \rangle$ defined by the electrostatic environment of the capacitor. The larger the capacitance, the broader the spread of the qubit eigenstates over the number of Cooper pairs. In general the circuit is designed to have a ratio $E_J/E_C \approx 50$, such that the qubit levels become almost entirely insensitive to fluctuations in occupation while still maintaining some anharmonicity [39].

The state of a superconducting qubit can be operated on by sending calibrated microwave pulses. Often the energy levels themselves can be changed as well by varying the Josephson energy E_J . In traditional designs, this energy level modulation is often done by splitting the Josephson junction into a superconducting quantum interference device (SQUID), and then tuning the external magnetic flux imposed on its surface with a current-carrying flux line nearby. Currents will be induced in the SQUID to bring the total flux that passes through it to an integer multiple of the flux quantum $\Phi_0 = h/2e$, affecting in turn the junction's effective coupling. This way the energy levels of nearby qubits can be brought close together, allowing for lossless energy exchange between them, and entangling their states with the *i*SWAP gate discussed in section 1.2.2. Some designs omit this individual tunability to improve coherence and reduce the number of cables, relying entirely on microwave signals to perform both single and two-qubit gate operations [40, 62]. Once a circuit has been executed, the state of a qubit can be read out by interrogating the same resonator that was used to operate on it, which will give a phase shift to a passing readout pulse that depends on the state of the qubit.

Quantum computers of this sort have had some widely publicized early successes, achieving quantum simulations [111] and computations [112] that outperform classical machines. For truly large-scale quantum computers however, many improvements still need to be made in error rates, fabrication cost, size of control hardware etc. In the next chapter, we will discuss how designing qubits around a Josephson junction made from a CMOS field effect transistor can bring us closer to realizing these ambitions.

Bibliography

- [1] Alan Mathison Turing. “Computing machinery and intelligence”. In: *Mind* 59.236 (1950), p. 433.
- [2] Jack S Kilby. *Miniaturized electronic circuits*. US Patent 3,138,743. 1959-02-06.
- [3] Eric Schmidt. “Eric Schmidt on the Early Strategic Decisions and Future of Quantum Computing”. Q2B20. 2020.
- [4] Alan Mathison Turing. “On computable numbers, with an application to the Entscheidungsproblem”. In: *Proceedings of the London mathematical society* 2.1 (1937), pp. 230–265.
- [5] Alan Mathison Turing. “On computable numbers, with an application to the Entscheidungsproblem. A correction”. In: *Proceedings of the London Mathematical Society* 2.1 (1938), pp. 544–546.
- [6] Warren S McCulloch and Walter Pitts. “A logical calculus of the ideas immanent in nervous activity”. In: *The Bulletin of Mathematical Biophysics* 5.4 (1943), pp. 115–133.
- [7] Matthew Cobb. *The idea of the brain: The past and future of neuroscience*. Profile Books Ltd, 2020. ISBN: 9781782832256.
- [8] Ethan Bernstein and Umesh Vazirani. “Quantum complexity theory”. In: *SIAM Journal on Computing* 26.5 (1997), pp. 1411–1473.
- [9] E Bernstein and U Vazirani. “Quantum complexity theory”. In: *Proceedings of the 25th Annual ACM Symposium on Theory of Computing* 11 (1993).
- [10] Richard P Feynman. “Simulating physics with computers”. In: *International Journal of Theoretical Physics* 21.6/7 (1982).
- [11] Stuart J Freedman and John F Clauser. “Experimental test of local hidden-variable theories”. In: *Physical Review Letters* 28.14 (1972), p. 938.
- [12] Scott Aaronson. *Quantum computing since Democritus*. Cambridge University Press, 2013. ISBN: 9780521199568.
- [13] Peter W Shor. “Algorithms for quantum computation: discrete logarithms and factoring”. In: *Proceedings 35th annual symposium on foundations of computer science*. Ieee. 1994, pp. 124–134.
- [14] Seth Lloyd. “Computational capacity of the universe”. In: *Physical Review Letters* 88.23 (2002), p. 237901.
- [15] Ronald L Rivest, Adi Shamir, and Leonard Adleman. “A method for obtaining digital signatures and public-key cryptosystems”. In: *Communications of the ACM* 21.2 (1978), pp. 120–126.

- [16] Ewin Tang. “A quantum-inspired classical algorithm for recommendation systems”. In: *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*. 2019, pp. 217–228.
- [17] Iordanis Kerenidis and Anupam Prakash. “Quantum recommendation systems”. In: *arXiv preprint arXiv:1603.08675* (2016).
- [18] David Deutsch. “Quantum theory, the Church-Turing principle and the universal quantum computer”. In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 400.1818 (1985), pp. 97–117.
- [19] A Chi-Chih Yao. “Quantum circuit complexity”. In: *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*. IEEE. 1993, pp. 352–361.
- [20] Leonard Susskind. *The black hole war: My battle with Stephen Hawking to make the world safe for quantum mechanics*. Little, Brown & Company, 2009. ISBN: 9780316016414.
- [21] Erik Verlinde. “On the Origin of Gravity and the Laws of Newton”. In: *Journal of High Energy Physics* 2011.4 (2011), pp. 1–27.
- [22] Subir Sachdev and Jinwu Ye. “Gapless spin-fluid ground state in a random quantum Heisenberg magnet”. In: *Physical Review Letters* 70.21 (1993), p. 3339.
- [23] Alexei Kitaev. “A simple model of quantum holography”. In: *KITP strings seminar and Entanglement*. Vol. 12. 2015, p. 26.
- [24] David J Griffiths. *Introduction to quantum mechanics, second international edition*. Pearson Prentice Hall, 2005. ISBN: 97871111182948.
- [25] Adam R Brown et al. “Quantum gravity in the lab: teleportation by size and traversable wormholes”. In: *arXiv preprint arXiv:1911.06314* (2019).
- [26] Sabine Hossenfelder and Tobias Zingg. “Analog models for holographic transport”. In: *Physical Review D* 100.5 (2019), p. 056015.
- [27] Daniel Gottesman. “The Heisenberg representation of quantum computers”. In: *arXiv preprint quant-ph/9807006* (1998).
- [28] Luke E Heyfron and Earl T Campbell. “An efficient quantum compiler that reduces T count”. In: *Quantum Science and Technology* 4.1 (2018), p. 015004.
- [29] A Yu Kitaev. “Quantum computations: algorithms and error correction”. In: *Russian Mathematical Surveys* 52.6 (1997), p. 1191.
- [30] Christopher M Dawson and Michael A Nielsen. “The Solovay-Kitaev algorithm”. In: *arXiv preprint quant-ph/0505030* (2005).
- [31] David Deutsch. *The fabric of reality*. Penguin, 1998. ISBN: 9780140275414.
- [32] Hugh Everett. “On the Foundations of Quantum Mechanics”. PhD thesis. 1957.
- [33] Sidney Coleman and Martin Greiter. “Sidney Coleman’s Dirac Lecture ‘Quantum Mechanics in Your Face’”. In: *arXiv preprint 2011.12671v1* (1994).
- [34] Sabine Hossenfelder. “A derivation of Born’s rule from symmetry”. In: *Annals of Physics* (2020), p. 168394.
- [35] Andrew M Childs et al. “Exponential algorithmic speedup by a quantum walk”. In: *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*. 2003, pp. 59–68.

- [36] Claude E Shannon. “A mathematical theory of communication”. In: *The Bell System Technical Journal* 27.3 (1948), pp. 379–423.
- [37] Michael A Nielsen and Isaac L Chuang. *Quantum Computation and Quantum Information*. 10th ed. Cambridge University Press, 2011. ISBN: 9781107002173.
- [38] Adriano Barenco et al. “Elementary gates for quantum computation”. In: *Physical Review A* 52.5 (1995), p. 3457.
- [39] Jens Koch et al. “Charge-insensitive qubit design derived from the Cooper pair box”. In: *Physical Review A* 76.4 (2007), p. 042319.
- [40] Philip Krantz et al. “A quantum engineer’s guide to superconducting qubits”. In: *Applied Physics Reviews* 6.2 (2019), p. 021318.
- [41] J Majer et al. “Coupling superconducting qubits via a cavity bus”. In: *Nature* 449.7161 (2007), pp. 443–447.
- [42] Edwin T Jaynes and Frederick W Cummings. “Comparison of quantum and semi-classical radiation theories with application to the beam maser”. In: *Proceedings of the IEEE* 51.1 (1963), pp. 89–109.
- [43] Alexandre Blais et al. “Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation”. In: *Physical Review A* 69.6 (2004), p. 062320.
- [44] Peter Chapman, Jungsang Kim, and Chris Monroe. *All clear to scale*. Q2B20 Conference. Oral presentation. Dec. 9, 2020.
- [45] Alexander Semenovitch Holevo. “Bounds for the quantity of information transmitted by a quantum communication channel”. In: *Problemy Peredachi Informatsii* 9.3 (1973), pp. 3–11.
- [46] Ryan Babbush. *Prospects for Error Corrected Quantum Applications*. Q2B20 Conference. Oral presentation. 2020.
- [47] Lucas T Brady et al. “Optimal Protocols in Quantum Annealing and QAOA Problems”. In: *arXiv preprint arXiv:2003.08952* (2020).
- [48] Leonardo Banchi, Nicolás Quesada, and Juan Miguel Arrazola. “Training Gaussian boson sampling distributions”. In: *Physical Review A* 102.1 (2020), p. 012417.
- [49] Brian Coyle et al. “The Born supremacy: Quantum advantage and training of an Ising Born machine”. In: *NPJ Quantum Information* 6.1 (2020), pp. 1–11.
- [50] Marco Cerezo et al. “Variational quantum algorithms”. In: *arXiv preprint arXiv:2012.09265* (2020).
- [51] Robert B Laughlin and David Pines. “The theory of everything”. In: *Proceedings of the National Academy of Sciences of the United States of America* 97.1 (2000), p. 28.
- [52] Seth Lloyd. “Universal quantum simulators”. In: *Science* (1996), pp. 1073–1078.
- [53] Ivan Kassal et al. “Polynomial-time quantum algorithm for the simulation of chemical dynamics”. In: *Proceedings of the National Academy of Sciences* 105.48 (2008), pp. 18681–18686.
- [54] Jonathan P Dowling. *Schrödinger’s Killer App: Race to Build the World’s First Quantum Computer*. CRC Press, 2013. ISBN: 978-1439896730.

- [55] Iulia M Georgescu, Sahel Ashhab, and Franco Nori. “Quantum simulation”. In: *Reviews of Modern Physics* 86.1 (2014), p. 153.
- [56] Alberto Peruzzo et al. “A variational eigenvalue solver on a photonic quantum processor”. In: *Nature Communications* 5.1 (2014), pp. 1–7.
- [57] Alexander Teplukhin et al. “Electronic structure with direct diagonalization on a D-wave quantum annealer”. In: *Scientific Reports* 10.1 (2020), pp. 1–11.
- [58] Finance & Economics section. “Quantum for quants: Wall Street’s latest shiny new thing: quantum computing”. In: *The Economist* (Dec. 16, 2020).
- [59] Bob Sorensen. *Surveying the quantum computing market landscape*. Q2B20 Conference. Oral presentation. 2020.
- [60] Austin G Fowler et al. “Surface codes: Towards practical large-scale quantum computation”. In: *Physical Review A* 86.3 (2012), p. 032324.
- [61] Norbert M Linke et al. “Experimental comparison of two quantum computing architectures”. In: *Proceedings of the National Academy of Sciences* 114.13 (2017), pp. 3305–3310.
- [62] Jay M Gambetta, Jerry M Chow, and Matthias Steffen. “Building logical qubits in a superconducting quantum computing system”. In: *NPJ Quantum Information* 3.1 (2017), pp. 1–7.
- [63] Richard P Feynman. “Quantum mechanical computers”. In: *Optics News* 11.2 (1985), pp. 11–20.
- [64] Umesh Vazirani. “On the power of quantum computation”. In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 356.1743 (1998), pp. 1759–1768.
- [65] Ryan Babbush et al. “Focus beyond quadratic speedups for error-corrected quantum advantage”. In: *arXiv preprint arXiv:2011.04149* (2020).
- [66] Richard Cleve et al. “Quantum algorithms revisited”. In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 454.1969 (1998), pp. 339–354.
- [67] David Elieser Deutsch. “Quantum computational networks”. In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 425.1868 (1989), pp. 73–90.
- [68] David Deutsch and Richard Jozsa. “Rapid solution of problems by quantum computation”. In: *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* 439.1907 (1992), pp. 553–558.
- [69] Hassi Norlén. *Quantum Computing in Practice with Qiskit® and IBM Quantum Experience®*. Packt Publishing Ltd, 2020. ISBN: 978-1-83882-844-8.
- [70] Daniel R Simon. “On the power of quantum computation”. In: *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science* (1994).
- [71] Daniel R Simon. “On the power of quantum computation”. In: *SIAM Journal on Computing* 26.5 (1997), pp. 1474–1483.
- [72] Erwin Kreyszig. *Advanced Engineering Mathematics*. 9th ed. John Wiley & Sons, Inc, 2006. ISBN: 978-0-471-72897-9.

- [73] Robert A Adams and Christopher Essex. *Calculus: a complete course*. 7th ed. Pearson Canada Inc, 2010. ISBN: 978-0-321-54928-0.
- [74] David J Griffiths. *Introduction to electrodynamics, third international edition*. Pearson, 2008. ISBN: 9780139199608.
- [75] E. Noether. “Invariante Variationsprobleme”. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* (1918), pp. 235–257.
- [76] Robert B Griffiths and Chi-Sheng Niu. “Semiclassical Fourier transform for quantum computation”. In: *Physical Review Letters* 76.17 (1996), p. 3228.
- [77] Scott Aaronson. *Shor, I’ll do it*. <https://www.scottaaronson.com/blog/?p=208>. Feb. 24, 2007. (Visited on 12/16/2020).
- [78] Scott Aaronson. *AMA With a Quantum Computing Superhero*. Q2B20 Conference. Oral presentation. Dec. 10, 2020.
- [79] *Jacobellis v. Ohio*, 378 U.S. 184. No. 11. Apr. 1, 1964.
- [80] Charles H Bennett et al. “Strengths and weaknesses of quantum computing”. In: *SIAM Journal on Computing* 26.5 (1997), pp. 1510–1523.
- [81] Alexander Condello. *Practical Quantum Computing: Resource Optimization*. Q2B20 Conference. Oral presentation. Dec. 8, 2020.
- [82] Sergei V Isakov et al. “Optimised simulated annealing for Ising spin glasses”. In: *Computer Physics Communications* 192 (2015), pp. 265–271.
- [83] Alexey Gorshkov. *Optimal Protocols in Quantum Annealing and QAOA Problems*. Q2B20 Conference. Oral presentation. Dec. 10, 2020.
- [84] James King et al. “Benchmarking a quantum annealing processor with the time-to-target metric”. In: *arXiv preprint arXiv:1508.05087* (2015).
- [85] Ilia Zintchenko, Ethan Brown, and Matthias Troyer. *Recent developments in quantum annealing*. <https://www.scottaaronson.com/troyer.pdf>. 2015.
- [86] Andrew Childs. *The Power of Quantum Computers*. Q2B20 Conference. Oral presentation. Dec. 8, 2020.
- [87] Adam Bouland. *What you need to know about quantum algorithms*. Q2B20 Conference. Oral presentation. Dec. 8, 2020.
- [88] Sergey Bravyi, David Gosset, and Ramis Movassagh. “Classical algorithms for quantum mean values”. In: *Nature Physics* (2021), pp. 1–5.
- [89] Juan I Cirac and Peter Zoller. “Quantum computations with cold trapped ions”. In: *Physical Review Letters* 74.20 (1995), p. 4091.
- [90] Chris Monroe et al. “Demonstration of a fundamental quantum logic gate”. In: *Physical Review Letters* 75.25 (1995), p. 4714.
- [91] Paul R Berman. *Cavity quantum electrodynamics*. Academic Press, Inc., Boston, MA (United States), 1994.
- [92] Thomas Pellizzari et al. “Decoherence, continuous observation, and quantum computing: A cavity QED model”. In: *Physical Review Letters* 75.21 (1995), p. 3788.

- [93] Tycho Sleator and Harald Weinfurter. “Realizable universal quantum logic gates”. In: *Physical Review Letters* 74.20 (1995), p. 4087.
- [94] Michel H Devoret et al. “Quantum fluctuations in electrical circuits”. In: *Les Houches, Session LXIII* 7.8 (1995).
- [95] Boris B Blinov et al. “Observation of entanglement between a single trapped atom and a single photon”. In: *Nature* 428.6979 (2004), pp. 153–157.
- [96] Wolfgang Paul. “Electromagnetic traps for charged and neutral particles”. In: *Reviews of Modern Physics* 62.3 (1990), p. 531.
- [97] Jonathan P Dowling. *Schrödinger’s Web: Race to Build the Quantum Internet*. CRC Press, 2020. ISBN: 9780367322311.
- [98] John M Martinis. “Saving superconducting quantum processors from qubit decay and correlated errors generated by gamma and cosmic rays”. In: *arXiv preprint arXiv:2012.06137* (2020).
- [99] Matthias Bock et al. “High-fidelity entanglement between a trapped ion and a telecom photon via quantum frequency conversion”. In: *Nature Communications* 9.1 (2018), pp. 1–7.
- [100] Jonathan P Dowling and Gerard J Milburn. “Quantum technology: the second quantum revolution”. In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 361.1809 (2003), pp. 1655–1674.
- [101] Paul Magnard et al. “Microwave quantum link between superconducting circuits housed in spatially separated cryogenic systems”. In: *Physical Review Letters* 125.26 (2020), p. 260502.
- [102] Raju Valivarthi et al. “Teleportation Systems Toward a Quantum Internet”. In: *PRX Quantum* 1.2 (2020), p. 020317.
- [103] RCC Leon et al. “Coherent spin control of s-, p-, d-and f-electrons in a silicon quantum dot”. In: *Nature Communications* 11.1 (2020), pp. 1–7.
- [104] Anthony Annunziata. *Advancing Quantum with Our Partners: A View of the Road Ahead from IBM*. Q2B20 Conference. Oral presentation. Dec. 9, 2020.
- [105] L Le Guevel et al. “Low-power transimpedance amplifier for cryogenic integration with quantum devices”. In: *Applied Physics Reviews* 7.4 (2020), p. 041407.
- [106] Anasua Chatterjee et al. “Semiconductor qubits in practice”. In: *arXiv preprint arXiv:2005.06564* (2020).
- [107] Audrey Cottet. “Implementation of a quantum bit in a superconducting circuit”. PhD thesis. Université Paris 6, 2002.
- [108] Matthias Steffen et al. “Measurement of the entanglement of two superconducting qubits via state tomography”. In: *Science* 313.5792 (2006), pp. 1423–1425.
- [109] Denis Vion et al. “Manipulating the quantum state of an electrical circuit”. In: *Science* 296.5569 (2002), pp. 886–889.
- [110] TP Orlando et al. “Superconducting persistent-current qubit”. In: *Physical Review B* 60.22 (1999), p. 15398.
- [111] Hannes Bernien et al. “Probing many-body dynamics on a 51-atom quantum simulator”. In: *Nature* 551.7682 (2017), pp. 579–584.

- [112] Frank Arute et al. “Quantum supremacy using a programmable superconducting processor”. In: *Nature* 574.7779 (2019), pp. 505–510.

Chapter 2

Complementary Metal Oxide Silicon gatemons

2.1 Introduction

We saw in the previous chapter that superconducting qubits are built around Josephson junctions, which are weak links between two superconducting leads. Typically these are made of insulating layers of aluminum oxide between superconducting strips of pure aluminum, and often come in parallel pairs that form superconducting quantum interference devices (SQUIDs). These SQUIDs are what allow for the tunability in tunable transmons: their coupling strength is highly sensitive to the out-of-plane magnetic field, something that can easily and precisely be set with a simple current carrying wire nearby [1]. The clear drawbacks to this design are that twice as many junctions are needed, a relatively large amount of current on the order of milli-Amps¹ linear in the number of qubits has to be supplied, fields intended for one qubit need to be compensated for in others², and any extra degree of freedom in your Hamiltonian also acts as an extra channel for noise. Some of these concerns have led to the adoption of fixed-frequency transmons [6, 7], though this is only one of many viable approaches. The idea pursued in this thesis is to stick to the tunable transmon design for all its advantages, but to replace the cumbersome SQUID by a single gate-modulated Josephson junction [8, 9].

Before getting into the details of the experimental work on Josephson Field Effect transistors, we will review qubit Hamiltonians in section 2.2, some of the principles of superconducting junctions in section 2.3, and some basics of transistors in section 2.4.

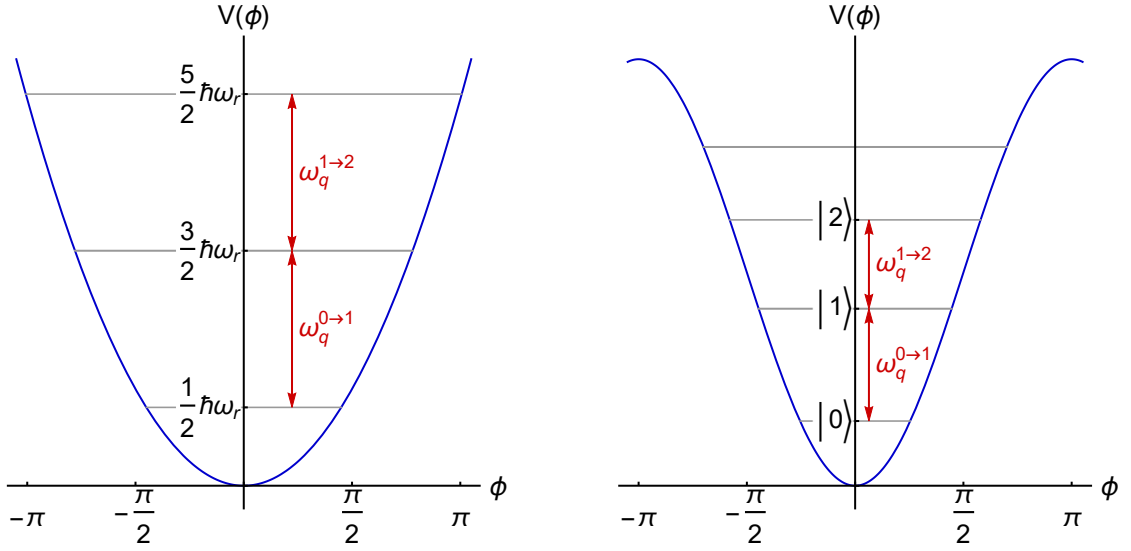


Figure 2.1: (Left) The quadratic potential energy of the quantum harmonic oscillator. On the ϕ -axis we have some variable, which can be anything from the angle of a pendulum or the displacement of a mass on a spring, to the phase across a Josephson junction. The y -axis is the energy, such that the gray lines represent levels where it is constant: as the system oscillates around $\phi = 0$, kinetic and potential energy are constantly interchanged. (Right) The almost-quadratic potential of an *an*harmonic oscillator. Here the level spacings are different, and we can isolate the bottom transition as our qubit.

2.2 The quantum harmonic oscillator

The only levels that we care about for quantum computational applications are the two with the lowest energy, which means that we always talk about local minima in the energy spectrum. And it turns out that at a local minimum, almost any potential can be well-approximated by a quadratic function, as can be seen by expanding some generic potential $V(\phi)$,

$$V(\phi) = V(\phi_0) + \partial_\phi V(\phi_0)(\phi - \phi_0) + \frac{1}{2} \partial_\phi^2 V(\phi_0)(\phi - \phi_0)^2 + \dots \quad (2.2)$$

We can forget about the first term as constant energies don't affect our system's dynamics, and we can ignore the second term since by definition of the “local minimum” it is strictly zero. As long as $\phi - \phi_0 \ll 1$ in the appropriate units, the higher-order terms $(\phi - \phi_0)^{n>2}$ will be negligible, and so all we are left with is the quadratic potential. This is what we call the harmonic oscillator, and for most practical purposes it is a good description of qubit systems.

¹For SQUIDS of about 10^2 square micron large [2], you need fields on the order of $B \approx \Phi_0/A \approx 10\mu\text{T}$. If the flux line is $\sqrt{A} = 10\mu\text{m}$ away, then Ampère's law tells us that we need currents of around

$$I_{\text{flux line}} = \frac{2\pi\Phi_0 A}{\mu_0 \sqrt{A}} = \frac{\pi h \sqrt{A}}{\mu_0 e} \approx 1\text{ mA}. \quad (2.1)$$

You could bring this number down by making the SQUIDS bigger, but then you'll end up with antenna arrays instead of qubits.

²The compounding interactions of many nearby components need not necessarily be intractable; similar concerns about crosstalk in classical circuits were ultimately resolved by simple design rules [3–5]. But this does of course not mean that exploring alternatives is not a good idea!

2.2.1 Qubit from It

In principle, any system with quantized energy levels could be used as a quantum bit, which is great, because if you look closely enough, *any* physical system exhibits quantum effects. Technically speaking however, for those bits to be qubits in any meaningful sense of the word, we also need them to have the following properties.

1. The system needs to be isolated well enough that it stays coherent much longer than it takes to perform a gate operation.
2. We need the spacings between the levels to be different, so that we can isolate a single transition between two levels as our degree of freedom.
3. We have to be able to perform the universal gate set, i.e. perform $\pi/4$ and $\pi/2$ rotations about different axes on a single qubit, and couple multiple qubits to each other.

We can break down how these conditions are all satisfied by transmons, the superconducting qubit type that we are aiming to build. The first of these conditions implies at the very least that the level spacing be greater than the temperature,

$$\hbar\omega_q \gg k_B T, \quad (2.3)$$

for otherwise we would lose the coherence in the thermal noise. To give an idea, a typical superconducting qubit has an angular frequency of about 3 to 5 GHz [10], while they are cooled to below 20 mK \approx 400 MHz. Apart from heat, there are a host of other ways that a superconducting qubit can couple to its environment. Charge qubits are most sensitive to fluctuations in charge on elements that it is capacitively coupled to, flux qubits easily couple to modes of the electromagnetic field that are available in inductive elements, and phase qubits decay in the presence of current fluctuations [11]. Transmons, whose energy levels are hybridized charge states [12], mostly decay due to dielectric losses to parasitic capacitances [13] and quasiparticle poisoning [14, 15].

The second condition limits how much we can achieve in the coherence time given by the first. Sending a signal to a many-level system in the ground state would first partially populate the $|1\rangle$ level, but instead of neatly oscillating back to $|0\rangle$, it would then spread into a superposition that also includes $|2\rangle$, $|3\rangle$, etc. We therefore need an *anharmonicity*³ α ,

$$\alpha = (\omega_{|2\rangle} - \omega_{|1\rangle}) - (\omega_{|1\rangle} - \omega_{|0\rangle}), \quad (2.4)$$

that is large enough that microwave signals will not excite level $|2\rangle$. Fourier transformation then tell us that any signal pulse shorter than some time τ will have energy components on the order of \hbar/τ , so if you want to have e.g. 100 ns pulses, you need $\alpha/\hbar > 10$ MHz. Obtaining a good ratio of gate operation to coherence time thus requires a large α . In practice, $\alpha \approx 100$ –300 MHz in transmons [12], so you could theoretically go down to pulse sequences with waveform features on the order of tens of nanoseconds.

The first part of the third condition, performing at least two orthogonal single-qubit gates, means that we should be able to couple to our transmon in such a way that we rotate its state about multiple axes. The fact that we can address different axes becomes clear when we consider the time dependence of any general state [16],

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r}) e^{iEt/\hbar}, \quad (2.5)$$

³Called so as it quantifies the degree to which the Hamiltonian is *not* like a harmonic oscillator, the perfectly quadratic potential that gives rise to equally spaced levels.

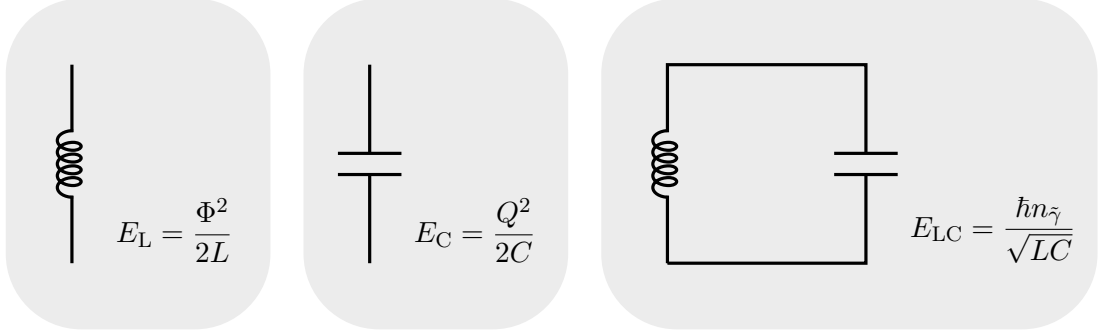


Figure 2.2: (Left) A classical inductor. (Middle) A capacitor. (Right) A simple LC circuit with a resonance frequency $\omega = 1/\sqrt{LC}$.

which directly implies that a qubit satisfies

$$|\Psi\rangle(t) = \alpha(t)|0\rangle + \beta(t)|1\rangle = e^{i(E_1 - E_0)t/\hbar} (\tilde{\alpha}|0\rangle + \beta(t)|1\rangle), \quad (2.6)$$

so that it always rotates around the quantization axis at a rate $\omega_q = (E_1 - E_0)/\hbar$. This means that we need to add a rotating component to any gate pulse that we send, but also that simply waiting for a time $1/4\omega$ will let you switch axes from x to y . It is perhaps less obvious that these are the axes that microwave signals act on through capacitive coupling, but this will become clear after we have discussed the transmon Hamiltonian in section 2.2.2. Two-qubit transmon gates can be performed in two different ways [7]: either by temporarily tuning nearby gatemons to similar frequencies (see the description of the i SWAP gate in section 1.2.2), or by permanently coupling the qubits such that the level separation of one depends on the state of the other, causing them to “control” each other’s rotation when excited by microwave pulses [17].

2.2.2 The transmon Hamiltonian

The transmon is an anharmonic LC resonator consisting of a Josephson junction (L) in parallel with a capacitor (C). Classically, the resonance of an LC circuit can be understood as follows. Energy is stored in the electric field between the capacitor’s plates as a voltage is built up across them,

$$E_C = \frac{1}{2} CV^2 = \frac{1}{2C} Q^2, \quad (2.7)$$

while the inductor stores energy in a magnetic field as current flows through it,

$$E_L = \frac{1}{2} LI^2 = \frac{1}{2L} \Phi^2. \quad (2.8)$$

This can only be a static situation if both energies are zero, as a fixed nonzero current would lead to a constantly growing charge on the capacitor, and a nonzero charge can only be fixed if the inductor has infinite inductance. The current and charge will therefore oscillate out of phase by an angle $\pi/2$: one will be maximum when the other has maximum derivative (and is zero), and vice versa. There is another way that we can state this phase shift, which is through their impedances Z_C and Z_L , which express their resistance to forming a current:

$$Z = \frac{V}{I} : \quad Z_R = R, \quad Z_C = \frac{1}{i\omega C}, \quad Z_L = i\omega L. \quad (2.9)$$

Here the phase shift of each element's current response to an applied voltage is expressed by the imaginary i , which represents a quarter turn ($\pi/2$) in the complex plane. Since the impedances of the capacitor and inductor have phase shifts in opposite directions, they are in total out of phase by π , as are their respective energies. The total impedance of the two elements then adds up to

$$\frac{1}{Z_{LC}} = \frac{1}{Z_C} + \frac{1}{Z_L} = \frac{i\omega L}{1 - \omega^2 LC}, \quad (2.10)$$

and is minimal when

$$\omega_{LC} = 1/\sqrt{LC}, \quad (2.11)$$

which means that the system will resist the least when it oscillates at this rate, giving us the resonance frequency.

When proportional⁴, these out-of-phase oscillations of the electric and magnetic fields are nothing else than photons⁵ [19, 20], and their number n_γ will be the quantization axis of the harmonic oscillator,

$$\mathcal{H} = \hbar\omega_{LC} \left(a^\dagger a + \frac{1}{2} \right) = \hbar\omega_{LC} \left(n_\gamma + \frac{1}{2} \right). \quad (2.12)$$

This photon number is but one of many quantization axes that can be chosen as qubit degree of freedom, and in general we will design the circuit to have energy levels that depend unevenly on electric and magnetic oscillations, to reduce sensitivity to particular kinds of noise. To better understand this general case, we go back to the Hamiltonian terms that we found earlier in equations (2.7) and (2.8) for the capacitor and inductor separately, which combine to

$$\mathcal{H} = \frac{1}{2C} Q^2 + \frac{1}{2L} \Phi^2 = 4E_C n^2 + \frac{1}{2} E_L \phi^2, \quad (2.13)$$

where Q and Φ are rewritten in countable units of the number of Cooper pairs and flux quanta [21],

$$n = \frac{Q}{2e}, \quad \phi = \frac{2\pi\Phi}{\Phi_0} = \frac{2e\Phi}{\hbar}. \quad (2.14)$$

Here ϕ represents the “gauge invariant phase difference” across the inductor, or equivalently, the phase difference between the macroscopic wave functions of the superconducting condensates on opposite sides of a Josephson junction, and n is the number of Cooper pairs built up at some node (in the case that one side of the loop is grounded, we choose the other as our node). These observables have their own operators that can be expressed in terms of the photon creation and annihilation operators [10, 22],

$$\hat{\phi} = \sqrt{\frac{4E_C}{\hbar\omega}} (a + a^\dagger) = \phi_{\text{zpf}} (a + a^\dagger), \quad \hat{n} = \sqrt{\frac{\hbar\omega}{16E_C}} (a - a^\dagger) = n_{\text{zpf}} i(a - a^\dagger). \quad (2.15)$$

⁴The virial theorem states that a perfectly harmonic oscillator has on average equal amounts of energy stored in the kinetic and potential energies, $\langle T \rangle = \langle V \rangle$, or in our case $\langle Q^2/2C \rangle = \langle \Phi^2/2L \rangle$. Since pure undressed photons satisfy the relation $|\mathbf{B}|^2 = \mu_0\epsilon_0|\mathbf{E}|^2$, this requires that $LC = \mu_0\epsilon_0 A^2/d^2$, with A the area of the inductor and d the spacing of the capacitor.

⁵This is best appreciated in the lumped-element approximation, where the wavelength is much longer than the circuit dimensions, clearly valid for $\lambda = 2\pi c/4 \text{ GHz} \approx 50 \text{ cm}$. In practice, the plasmon oscillations that define the qubit levels are never engineered to be pure photon, flux, or charge states, but some dressed combination of these. This way, the quantization axis can be chosen such that the levels are more isolated from the environment than the individual components it is made from, as has been demonstrated e.g. by dressing spin states with microwave photons [18].

Since ϕ and n are each other's conjugate, they satisfy the appropriate commutation relation

$$[\hat{\Phi}, \hat{Q}] = i\hbar, \quad [\hat{\phi}, \hat{n}] = i. \quad (2.16)$$

An analogy can be drawn to the position and momentum operators in the traditional derivation [16],

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(a^\dagger + a), \quad \hat{p} = i\sqrt{\frac{\hbar m\omega}{2}}(a^\dagger - a), \quad (2.17)$$

where the “mass” or inertia is now proportional to the capacitance, $m \triangleq C/4e^2$. In this new formulation in terms of charge and flux, the excitations of the circuit are separated by the geometric mean of their energies,

$$\hbar\omega = \hbar/\sqrt{LC} = \sqrt{8E_L E_C}, \quad (2.18)$$

where flux and charge can each appear as approximate quantization axis for a qubit system.

It is important to note that the Hamiltonians described above have exactly equal level spacing between any two consecutive numbers of excitations: $\Delta E = \hbar/\sqrt{LC}$ for each level that we go up. As we saw in section 2.2.1, we need to introduce some degree of *anharmonicity* $\alpha = \omega^{1 \rightarrow 2} - \omega^{0 \rightarrow 1}$, which we can achieve by replacing the inductor by a Josephson junction with energy

$$E_J = \frac{I_c \Phi_0}{2\pi} = \frac{I_{c,0} \Phi_0}{2\pi} \cos(\phi). \quad (2.19)$$

The energy contributed by this junction is proportional to the coupling strength between its two superconducting condensates, which in turn depends on their phase difference ϕ . It is this sinusoidal dependence on ϕ that gives the Hamiltonian its non-quadratic anharmonicity,

$$\begin{aligned} \mathcal{H}_{LC} &= 4E_C n^2 + \frac{1}{2}E_L \phi^2 \\ &\quad \downarrow \\ \mathcal{H}_{\text{transmon}} &= 4E_C n^2 - \frac{I_{c,0} \Phi_0}{2\pi} \underbrace{\cos(\phi)}_{\substack{\phi^2 & \phi^4 \\ \mathbf{1} - \frac{\phi^2}{2!} + \frac{\phi^4}{4!} - \dots}}, \end{aligned} \quad (2.20)$$

where we discard the first term in the expansion since it is constant. For small ϕ , which is our usual operating point⁶ (the spread in n is much larger in transmons), higher orders in the cosine's Taylor expansion can be discarded, and we find an added term $\propto \phi^4$, which turns out to shift the second level by an amount $\alpha = -E_C$. See Fig. 2.1 for a visual representation. While circuits with substantial E_C have historical importance [1, 24], in what follows we will focus on designs with strong capacitive coupling and larger Josephson energies, such that $E_J \gtrsim 50E_C$ [12].

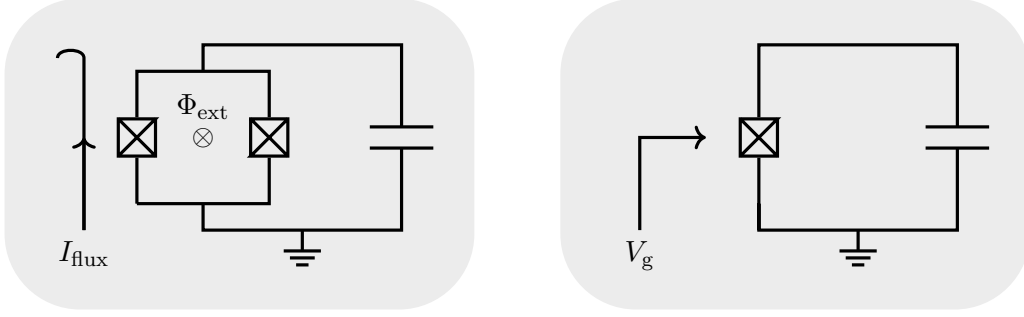


Figure 2.3: (Left) The traditional flux-tunable transmon design [12], where the current I_{flux} in a nearby superconducting loop is used to tune E_J/E_C . (Right) The “gatemon” approach [8, 9], where a gate voltage is used to adjust E_J instead.

2.2.3 Gatemons

We discussed earlier in section 1.2.2 that two superconducting qubits can be capacitively coupled to each other when their impedances ($Z = \sqrt{8E_C/E_J}$) are made to match for a limited amount of time⁷. To achieve this, either the capacitance or the Josephson coupling of one or both of the qubits needs to be changed. While the capacitance is literally hardwired into the circuit, we can quite easily control the effective Josephson coupling of two parallel junctions by changing the magnetic field that is imposed on the area between them, see Fig. 2.3 (left). In such a SQUID geometry, the superconducting condensate will enforce quantization of the penetrating flux towards the nearest integer multiple of the flux quantum by generating clockwise or anti-clockwise currents. So while the maximum critical current $I_{c,0}$ is fixed by the physical dimensions of the two junctions, the effective critical current $I_c = I_{c,0} |\cos \phi_{\text{ext}}|$ and therefore the effective coupling can be tuned [25]:

$$\mathcal{H}_{\text{transmon}} = 4E_C n^2 - \frac{I_{c,0}\Phi_0}{2\pi} \cos(\phi + \phi_{\text{ext}}). \quad (2.22)$$

This design has its drawbacks however, as flux lines of nearby qubits can interact and need to be corrected for each other, and currents linear in the number of qubits will need to be supplied to the chip (as we calculated before, kilo-Amps in the case of megaqubit devices). This is why some groups [8, 9] have started exploring the option of replacing the SQUID and flux line by a single, non-dissipative gate-tuned semiconducting junction, shown in Fig. 2.3 (right). This gatemon design is what is pursued in this thesis as well.

While the above-mentioned implementations have been very successful in demonstrating proof of principle with relatively long coherence times, both relied on InAs nanowires for their semiconducting junctions. This choice is well-motivated by the absence of any Schottky barriers and the resulting high transparency at the S/Sm interfaces [26, 27], especially when the Al/InAs interface can be grown epitaxially [28], but fabrication tech-

⁶For ϕ to be a reasonable quantum variable, the Josephson coupling needs to be stronger than other energy terms, e.g. [23],

$$E_J > k_B T \quad \Rightarrow \quad I_{c,0} > \frac{2\pi k_B T}{\Phi_0}, \quad (2.21)$$

which for fridge temperatures of around $T = 50$ mK means that the critical current needs to be larger than 2.1 nA. Typical transmons [10] with $\omega = 3\text{--}6$ GHz and charging energies on the order of 100–300 MHz ($C \approx 0.1\text{--}0.4$ pF) will need critical currents between 1.2 and 14 nA.

⁷In more general terms, an impedance mismatch is when two things take a different amount of time to compress and then relax, which leads to a less efficient collision.

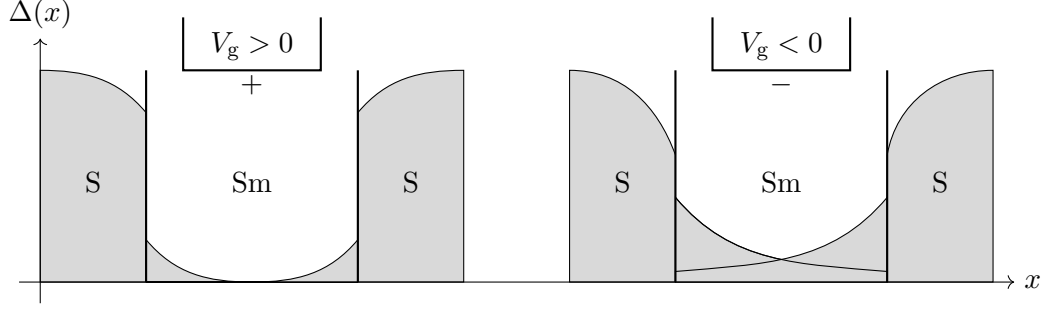


Figure 2.4: A cartoon illustration of the effect of applying a gate voltage to a semiconducting barrier between two superconductors. Shown is the superconducting order parameter, or effective gap, throughout the junction: its induction in the semiconducting (Sm) channel is called the proximity effect, while the suppression within the superconducting leads is the inverse proximity effect. There are two separate processes through which the proximity effect can be enhanced [23]: the barrier transparency can be increased, or the conductive and diffusive properties of the channel itself can be changed.

niques involving nanowires have so far been difficult to scale. This barrier to scalability can be overcome by lithographically patterning the semiconducting junctions in planar geometries instead [29, 30].

As shown in Fig. 2.4, the electrostatic field generated by the gate voltage (here shown for a p-doped channel) will attract charge carriers (holes in this case). In short junctions, where the critical current is proportional to the normal-state conductance [23], this increase in carrier density will strengthen the Josephson coupling by $E_J \propto G_N \propto n^{2/3}$. In longer junctions, where the proximity effect decays exponentially over a length scale roughly proportional to the carrier density [31], this dependence can be even stronger. By attracting carriers, the electric field also lifts up or pushes down the Fermi level inside the semiconductor, changing the mismatch in work function with the metallic superconducting leads, and thus modifying the height of the Schottky barrier. Increasing the carrier density will also reduce the width of the barrier by more effectively shielding the charges on the metallic side. This provides an additional means of modifying the critical current by reducing the tunnel barrier in non-Ohmic Schottky-barrier devices. Taken together, these various effects mean that though phenomenological and qualitative understanding exists, the exact relationship between the applied voltage and the critical current is device-specific, and can at best be determined only experimentally. We therefore give the gatemon Hamiltonian only in its general form,

$$\mathcal{H}_{\text{gatemon}} = 4E_C n^2 - \frac{\Phi_0}{2\pi} \mathbf{I}_c(\mathbf{V}_g) \cos(\phi), \quad (2.23)$$

and will explore in the coming sections these diverse phenomena of superconducting transport that together determine $I_c(V_g)$.

2.3 Superconducting transport across semiconducting junctions

Within the superconductors on either side of the Josephson junction, there is a phonon-mediated attractive force between electrons that provides entangled states [32] with lower

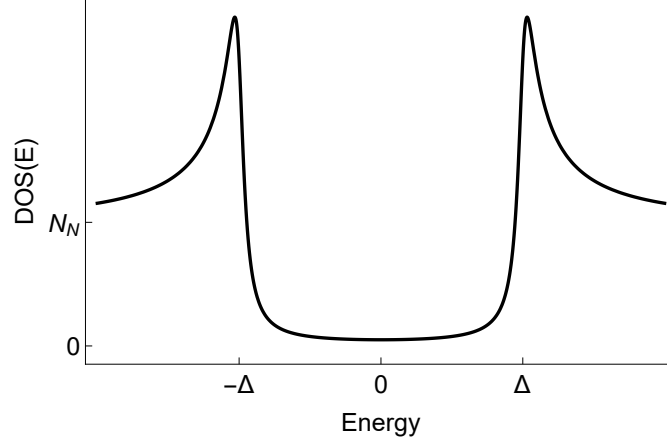


Figure 2.5: The quasiparticle density of states in a superconductor as a function of energy relative to the Fermi level, calculated from eq. (2.27). The Dynes parameter Γ is set at 5% of the superconducting gap.

energy [33, 34]. It was shown that due to this attractive interaction, the elementary quasiparticle excitations (“Bogoliubons”⁸) acquire energies of

$$E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2 + |\Delta|^2}, \quad (2.24)$$

where $\xi_{\mathbf{k}}$ is the kinetic plane-wave energy (usually measured from the Fermi level), and Δ the energy associated with the coupling, that also determines the energy band within which these quasiparticles are expected to form Cooper pairs [37]. Since the number of states per unit momentum is the same whether the material is superconducting (S) or not (N, for “normal”), we can write

$$\frac{dN_S}{dk} = \frac{dN_N}{dk}, \quad (2.25)$$

and because in the normal state $\Delta = 0$, we can also assume that $E_{\mathbf{k}} = \xi_{\mathbf{k}}$. Then, after multiplying both sides by dk/dE and integrating with respect to energy, we find

$$\frac{dN_S}{dE} \frac{dE}{dk} = \frac{dN_N}{d\xi} \frac{d\xi}{dk} \Rightarrow N_S(E) = N_N(E) \frac{d\xi}{dE} = N_N \frac{d\sqrt{E^2 - \Delta^2}}{dE}. \quad (2.26)$$

We can thus derive the density of quasiparticle states per unit energy, where we’ll introduce the Dynes parameter Γ [38], such that $E \rightarrow E - i\Gamma\Delta$, to account for inelastic scattering events⁹:

$$N_S = N_N \left| \operatorname{Re} \left(\frac{E - i\Gamma}{\sqrt{(E - i\Gamma)^2 - \Delta^2}} \right) \right|. \quad (2.27)$$

This density of states, shown in Fig. 2.5, has a energy gap of $2\Delta = 3.53k_B T_c$ at the Fermi level [33, 39, 40], which means that no states for single quasiparticles are available.

⁸Evidence for the existence of these combined particle-hole excitations [35] comes from Cooper pairs decaying not just into two electrons, but into both electrons *and* holes [36].

⁹In case you’re curious about the appearance of the absolute value: this is to make sure that we have a positive number of states at energies below the Fermi level. As far as I can tell, the origin of the missing sign is in the substitution of $\sqrt{h_k^2} \rightarrow h_k$ in eq. (2.41) of Ref. 33.

When a metal, where the density of states is to good approximation flat around the Fermi energy, contacts this superconductor, it is therefore impossible for single electrons to enter individually. Instead, transport across the interface happens through a process called *Andreev reflection* [41], where an incoming electron is reflected back into the metal as a hole that coherently traces back the electron’s path, while a Cooper pair that emerges on the other side ensures that charge is conserved. This process effectively leads to a “leaking” of the superconductivity into the normal metal, referred to as the proximity effect, where it induces a smaller *minigap* [42, 43].

If the metal is contacted on the other side by a second superconductor, there will now be two distinct ways that charge can be transported across the junction. Either the coherence of the electron and hole persist throughout, in which case the probability of entering the superconductor on the other side depends on the phase difference between the two condensates [43, 44], or the electron-hole pair decoheres, and the current is partially carried by regular individual quasiparticles. While the former creates the desired Josephson coupling between the two superconducting leads [1], the latter could add an undesirable energy relaxation mechanism. Additionally, at higher temperatures it becomes possible for quasiparticles to enter the superconductor directly without Andreev reflection, causing a different type of decoherence called quasiparticle poisoning [14, 15].

In the doped silicon channels of the Schottky-barrier MOSFETs that we study, there are relatively many impurities, and we expect the transport through the junction to be diffusive rather than ballistic ($\ell_{\text{mfp}} \approx 10 \text{ nm} \ll L_{\text{channel}} \approx 100 \text{ nm}$). In this case the spatial decay of the proximity effect depends on how the diffusion time of a charge across the junction compares to the decoherence timescales of the many disruptive physical processes that cause inelastic scattering. Before we discuss this in more detail, we will first briefly review the principles of Andreev reflection and diffusion.

2.3.1 Andreev reflection and the proximity effect

After a condensate has formed inside a superconductor, no single-particle states near the Fermi level will be available anymore. This seemingly poses a problem for an incoming electron or hole from an adjacent normal material, which at low temperatures necessarily moves near the Fermi energy. Luckily, a second-order process is possible, whereby the incident electron can form a Cooper pair as it enters the superconductor by picking up an additional electron near the interface, at the cost of back-scattering a hole with opposite spin into the normal lead. This is illustrated in Fig. 2.6, which assumes that the voltage bias across the insulating barrier (I) is less than the pair potential inside the superconductor (S), such that the incoming electron has an energy $E < \Delta$.

As an historical note, it is interesting that the notion of a proximity effect in fact predates this description. The leaking of the pair potential across the interface can already be derived with a phenomenological approach¹⁰, which also gives rise to different coherence lengths on the normal and superconducting sides, over which the proximity effect and its inverse decay [49]. The microscopic description was found later when Andreev solved the Gor’kov equations [50]¹¹ for this system [41]. He assumed solutions

¹⁰By “phenomenological” we mean assuming that whatever causes superconductivity, it is described by a wave function, and leave it at that. This simple approach [46, 47] recovers many of the relevant features [48].

¹¹The Schrödinger equation with interactions only in a narrow band around the Fermi energy, i.e. the “quasiclassical” approximation that is in general valid for superconductors.

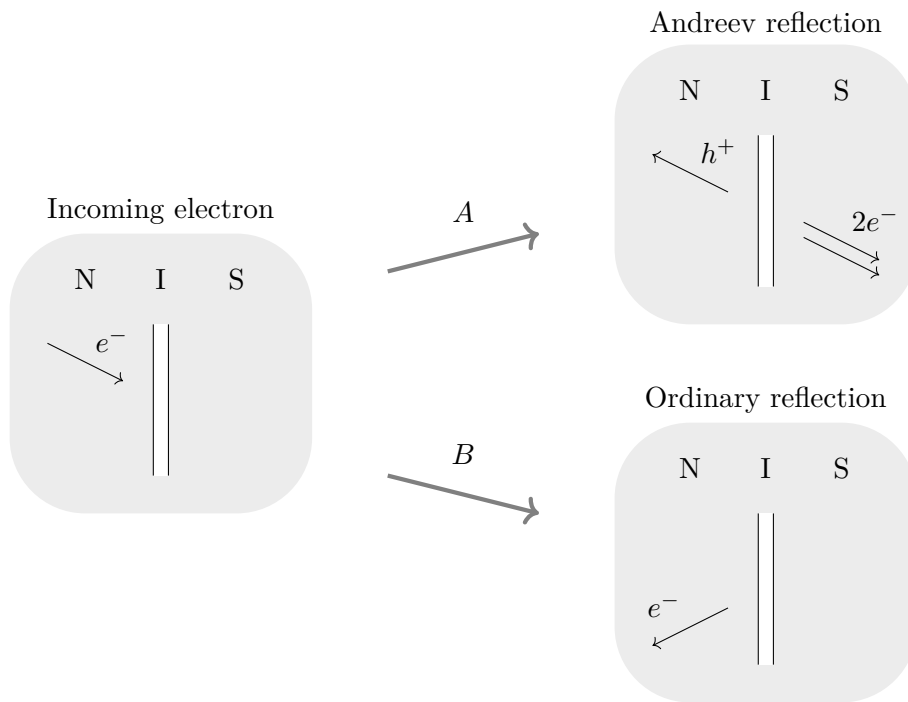


Figure 2.6: For incident particles with energies below the gap, $E < \Delta$, two processes can occur: Andreev reflection (A) and ordinary reflection (B). Their respective probabilities are a function of E/Δ and the barrier strength Z [45]. Note that in the former case, the hole is **not** reflected specularly, but instead traces back the path of the incident electron [41].

of the form

$$\begin{aligned} f &= e^{i(\mathbf{k} \cdot \mathbf{r} - Et/\hbar)} \eta(\mathbf{r}) \quad (\text{particle}), \\ \varphi &= e^{i(\mathbf{k} \cdot \mathbf{r} - Et/\hbar)} \chi(\mathbf{r}) \quad (\text{hole}), \end{aligned} \quad (2.28)$$

where η and χ are the envelopes of the wave packets [16], and arrived at kinetic equations of the form

$$\begin{aligned} \left(\frac{i\hbar v_F}{m} \mathbf{n} \cdot \nabla + E \right) \eta - i\Delta(\mathbf{r})\chi &= 0, \\ \left(\frac{i\hbar v_F}{m} \mathbf{n} \cdot \nabla - E \right) \chi - i\Delta(\mathbf{r})\eta &= 0, \end{aligned} \quad (2.29)$$

with \mathbf{n} the direction of the incident particle. What is remarkable, is that when these are solved on the normal side, where $\Delta(\mathbf{r}) = 0$, the solution is

$$\begin{pmatrix} \eta \\ \chi \end{pmatrix} = A \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{i\mathbf{k}_N \cdot \mathbf{r}} + B \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-i\mathbf{k}_N \cdot \mathbf{r}}, \quad (2.30)$$

which are a particle and a hole moving with precisely opposite momenta,

$$\mathbf{k}_N = \mathbf{n}E/\hbar v_F. \quad (2.31)$$

In the superconductor we use instead $\Delta(\mathbf{r}) = \Delta_0$, and obtain

$$\begin{pmatrix} \eta \\ \chi \end{pmatrix} = \frac{C}{\sqrt{2}} \begin{pmatrix} \sqrt{1 + \hbar v_F \mathbf{n} \cdot \mathbf{k}_S/E} \\ -i\sqrt{1 - \hbar v_F \mathbf{n} \cdot \mathbf{k}_S/E} \end{pmatrix} e^{i\mathbf{k}_S \cdot \mathbf{r}}, \quad (2.32)$$

with a momentum (+ for an incident particle, – for an incident hole)

$$\mathbf{k}_S \cdot \mathbf{n} = \frac{\pm 1}{\hbar v_F} \sqrt{E^2 - \Delta_0^2}. \quad (2.33)$$

This is a Cooper pair with a momentum that in the plane of the interface is identical to that of the incident particle, but that in the orthogonal direction depends on the energy of that particle relative to the gap Δ . Instead of seeing charge transmission across the interface forbidden by the lack of available states, Andreev found that it could in fact be enhanced to above unity!

Although it was not yet possible within this treatment to find the incidence, reflection and transmission coefficients A , B and C for arbitrary energy E , it did establish the now-famous mechanism that is truly quantum mechanical in nature. The predicted non-specular reflection (“back-scattering” or “retroreflection”) of a hole, and the resulting doubling of the current [51], was confirmed by an experiment with a quantum point contact on the back of a silver crystal (N) coated with lead (S) [52]. Later simulations of this experiment, and variants thereof, further corroborated the tracing out by the hole of the incident electron’s path, and found that this back-focusing is independent of the angle of incidence, any particularities of the interface, and even of the number of impurities in the normal material [53, 54]. A more quantitative description of the scattering rates was given in 1982 by Blonder, Tinkham and Klapwijk [45], which we will treat in more detail in chapter 4.

The situation is somewhat more complicated in semiconductors, for which there is no single overarching description [55, 56]. While highly (“degenerately”) doped

semiconductors may behave as metals [57]¹², albeit with a much shorter coherence length [55], materials with lower doping levels will be insulating after the dopants have frozen out [23, 58]. An additional mechanism may provide a means of transport at intermediate doping. At sufficiently high impurity densities and low temperatures, paths across periodically spaced dopants will generate what are called “resonance-percolation trajectories” [59], which can become the dominant contribution to coherent transport [60]. In general, the supercurrent through a semiconducting weak link will have to overcome both imperfect transmission at the interfaces, and decoherence in the channel itself [61]. Intuitively, for short junctions or high doping, the critical current will be limited by the interfaces and depend mostly on the properties of the superconductors, while for longer junctions or lower doping it will depend mainly on the properties of the weak link [31, 61, 62].

The clear advantage of semiconductors is that, unlike both insulators and metals, their properties can be altered with an electrostatic field; by accumulating carriers [30], the diffusion can be enhanced [31] and the effective Schottky barrier width at the interfaces can be reduced [62]. Since Andreev reflection is a second-order process, this reduction of the interface barrier strength has an even greater impact on superconducting transport than it does on normal transmission.

2.3.2 Diffusion

Diffusion is important not just in the context of the proximity effect in dirty semiconductors, but underlies all physical processes where scattering happens on length scales much smaller than the system size. This is the case for solid state reactions such as silicidation and oxidation (to be discussed in chapter 3), where atoms move erratically over distances of a few ångström, as well as for electronic transport in most devices, where carriers typically travel tens of nanometers between collisions (chapter 4).

To illustrate this process, consider a large collection of particles that at time $t = 0$ are placed at the origin, and are allowed to perform random walks for some time, such that they slowly spread out over space. A typical distribution is drawn on the left in Fig. 2.7, where the section of a sphere indicates a plane of constant density. Since the density is constant along the indicated contour, we can assume that there is no net flow of particles to neighboring regions above and below, nor to those in and out of the page. The change in the density ρ_n per time step Δt can now be calculated from the densities in cells $n - 1$ and $n + 1$. If we choose our Δt such that there is enough time for the density at point n to become the average of what the surrounding densities were at the previous time step, then the change in density is

$$\Delta_t \rho_n = \frac{1}{2} [(\rho_{n+1} - \rho_n) - (\rho_n - \rho_{n-1})]. \quad (2.34)$$

These steps in time and space are of course derivatives,

$$\frac{\Delta_t \rho}{\Delta t} = \frac{1}{2} \underbrace{\frac{(\Delta x)^2}{\Delta t}}_D \frac{\Delta_x^2 \rho}{(\Delta x)^2}, \quad (2.35)$$

¹²“Degenerate” refers to the overlap of orbitals of nearby donors or acceptors at high doping levels ($\gtrsim 1 \times 10^{18} \text{ cm}^{-3}$), and the resulting hybridization of their energy levels. Conductance *increases* with temperature in non-degenerate semiconductors, while it *decreases* in degenerate ones (which is why they are often referred to as “metallic”).

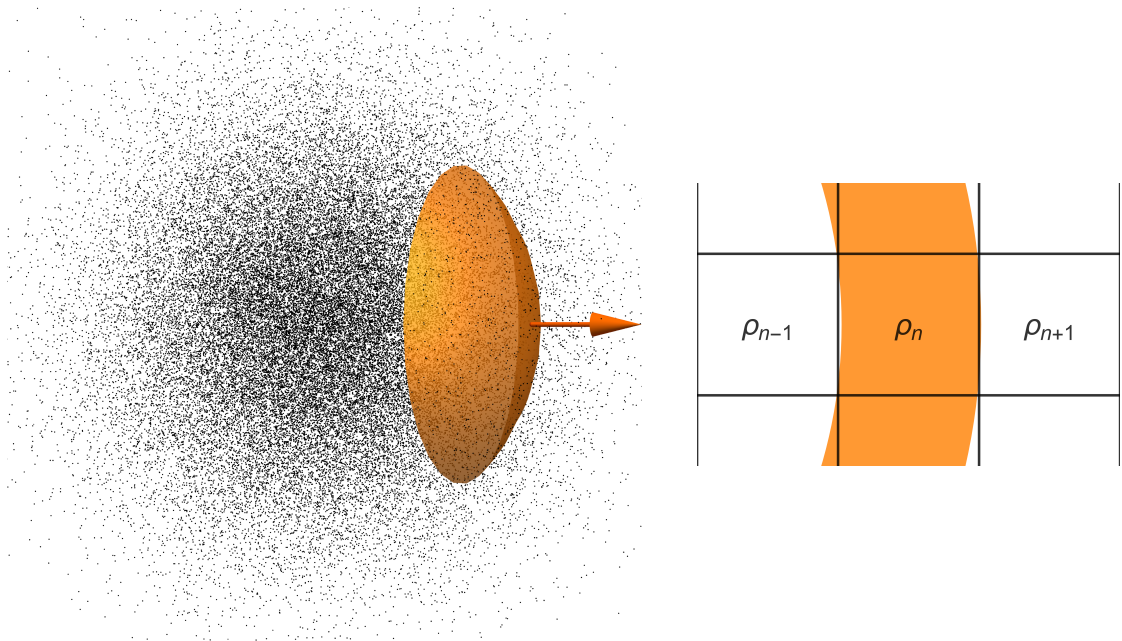


Figure 2.7: **(Left)** When a collection of particles are allowed to diffuse from the origin, the density profile will be spherically symmetric. The concentration will be constant along some surface at fixed distance from the origin, and we can define a diffusion direction orthogonal to it. **(Right)** We can now figure out the time evolution of the system by analyzing a one-dimensional array of positions along the diffusion direction.

where the prefactor D on the right-hand side is called the *diffusion constant*. We can rewrite this with continuous variables by letting $\Delta t, \Delta x \rightarrow 0$, such that¹³

$$\partial_t \rho_n = D \nabla^2 \rho. \quad (2.36)$$

For electrons we have

$$D = \frac{v_F \ell_e}{3}, \quad (2.37)$$

where v_F is the Fermi velocity and ℓ_e the mean free path.

The important thing to notice in this equation is that it contains a *first* order derivative to time, but a *second* order derivative to space: the change in the local density of particles is proportional to the *difference in difference* in density between pairs of cells. We can thus immediately guess that the time it takes to diffuse a particle a certain distance L is proportional to the *square* of that distance, $t \propto L^2$, and vice versa; that the spread of a collection of particles grows only as the square root of time. A more hand-waving and intuitive argument can be made based on the statistics of random walks alone. After a particle has taken N steps of size ℓ , it will have covered a path of total length $N\ell$, and on average it will end up right where it started. But since the uncertainty in the average of N uncorrelated values is proportional to $1/\sqrt{N}$, our estimate of the *sum* of all the steps

¹³Note that the only difference in form between the diffusion equation and Schrödinger's equation for a free particle is a factor i in front of the time derivative, accounting for the latter's phase evolution:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi.$$

in the particle's path will be accurate only to the order of $N\ell/\sqrt{N} = \sqrt{N}\ell$. Roughly speaking, we thus expect the particle to travel a net distance $\sqrt{N}\ell$ during N steps in time¹⁴: $L \propto \sqrt{t}$. Inversely, if the step size ℓ were to decrease, the time needed to travel a distance L would go up as $t \propto (L/\ell)^2$.

We will also encounter this relationship in the formation of silicides during heating: by checking if the growth rate of a new phase is linear in, or proportional to the square root of the annealing time, we can determine if the reaction is limited by the nucleation of the new compound at the interface or by the diffusion of the atoms through the forming silicide.

2.3.3 Decoherence

While an electron tracing back its own path would reverse any phase shift it had acquired, the back-scattered hole will compound it, so that after a path of length L covering an area¹⁵ A orthogonal to the field \mathbf{B} , the electron and hole phases would be off by [66]¹⁶

$$\Delta\phi = \frac{2EL}{\hbar v_F} + 2\pi \frac{\mathbf{B} \cdot \hat{\mathbf{n}} A}{\Phi_0}, \quad \text{with} \quad \Phi_0 \equiv \frac{h}{2e}. \quad (2.38)$$

Here we assume for simplicity that both are at energies $E \ll \Delta$, and ignore the small phase shift acquired during the reflection itself [53, 54]. If only a single trajectory is involved, the coupling could be enhanced again as $\Delta\phi$ approaches 2π . More realistically, many different paths contribute, each with different lengths L for the same depth into the semiconductor, and this relative dephasing will cause the effective coupling to decay exponentially inside the semiconductor, where the order parameter will evolve as

$$|\psi|^2 = \psi_{h+}^\dagger \psi_{e-} \propto e^{-\langle \Delta\phi \rangle}. \quad (2.39)$$

Apart from pure dephasing, decoherence can be caused by a broadening of the individual energy distributions of either particle due to a coupling to other degrees of freedom.

For the purpose of what we discuss here, we can reasonably approximate the electrons and holes that travel through the semiconducting channel as plane waves [16],

$$\Psi(\mathbf{r}, t) = e^{i(\mathbf{k} \cdot \mathbf{r} - E(\mathbf{k})t)}, \quad (2.40)$$

which are eigenstates of the momentum operator. Each of these momentum states has this same form $e^{i\phi}$, which Euler taught us to decompose into a sine and a cosine,

$$e^{i\phi} = \sum_{n=0}^{\infty} \frac{(i\phi)^n}{n!} = \sum_{n=0}^{\infty} \underbrace{(-1)^n \frac{\phi^{2n}}{(2n)!}}_{\cos(\phi)} + i \sum_{n=0}^{\infty} \underbrace{(-1)^n \frac{\phi^{2n+1}}{(2n+1)!}}_{\sin(\phi)}, \quad (2.41)$$

¹⁴Random walks and their square laws are surprisingly ubiquitous in science and math, appearing even in the remote field of Economics. For example, two Bayesians A and B who have disagreeing prior estimates X_A and X_B of some $X \in [0, 1]$ because they each have different information, will each update their estimates upon hearing the other's (and thus learning about their information) in a random-walk fashion, such that they need $1/\epsilon^2$ updates to agree within $|X_A - X_B| \leq \epsilon$ [63]. Even more interestingly, *quantum* random walks *do not* follow this square law, and so can explore graphs in linear time, providing a quadratic speedup compared to classical algorithms for stochastic processes [64, 65].

¹⁵Here, A would be the net area enclosed by the path after the start and end points are connected by a straight line. To see why, imagine instead of the hole tracing back the same path, the electron moving back along the trajectory mirrored in that line. Clearly that path would enclose $2A$, and half the phase shift would be acquired along the trajectory we are considering.

¹⁶The prefactor 2π , rather than 4π as used in the cited work, is due to our choice for the flux quantum of $\Phi_0 = h/2e$ instead of $\Phi_0 = h/e$.

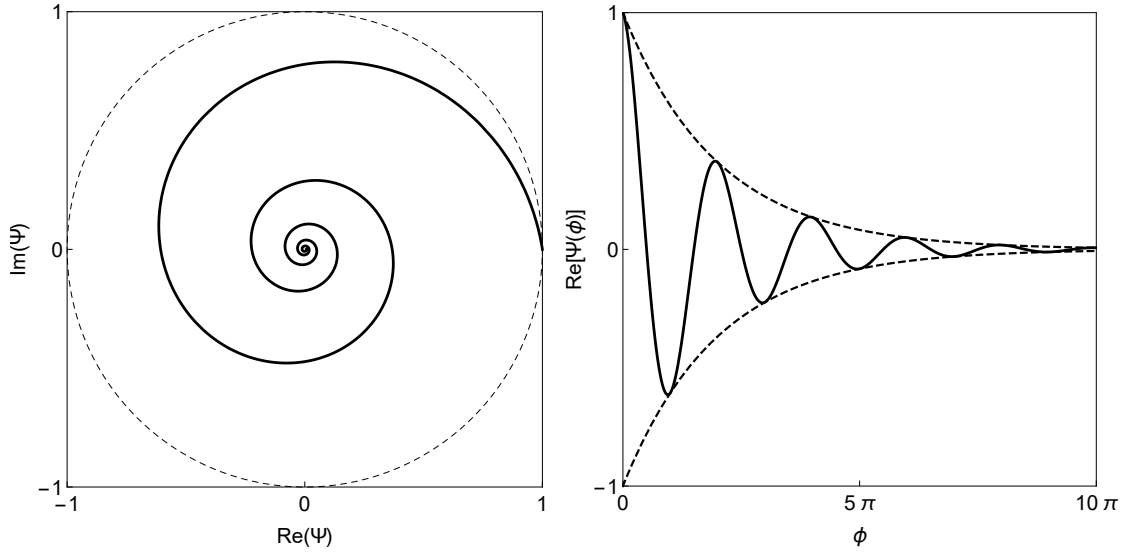


Figure 2.8: Decoherence of a wave function $\Psi = e^{i(\phi+i\Gamma)}$. **(Left)** the rotation of the function in the complex plane, which can be also taken as the equatorial plane of the Bloch sphere for a dephasing qubit. **(Right)** The decay of the real part of the wave function with growing ϕ .

which oscillate out of phase and orthogonal to each other in the complex plane. If the energy term in eq. (2.40) is completely real, it will only determine the rotation in the complex plane, but if there are imaginary terms such that $\mathbf{k} \cdot \mathbf{r} + Et = A + iB$, we find that

$$e^{i(A+iB)} = e^{iA}e^{-B} \quad \text{and} \quad |\Psi|^2 = (e^{iA}e^{-B})(e^{-iA}e^{-B}) = e^{-2B}, \quad (2.42)$$

and the wave function will be suppressed exponentially. This is shown graphically in Fig. 2.8.

Elastic scattering events, such as the majority of those causing diffusion in a silicon channel (collisions off heavy objects such as dopants or grain boundaries), add only real terms to the energy, making it possible for a particle to change momentum direction hundreds or thousands of times before coherence is lost [66]. The decoherence comes from *inelastic* scattering events, where carriers lose some of their kinetic energy as they interact with lighter objects (such as other carriers [38]). In such inelastic interactions, two quantum systems couple their energy degrees of freedom, speeding up interactions with the rest of the environment, which randomize their wave functions' phases. Decoherence due to phase randomization is hard to appreciate in wave function notation, so consider the simple density matrix of a superposition

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle = a|0\rangle + be^{i\phi}|1\rangle, \quad (2.43)$$

$$\rho = |\Psi\rangle\langle\Psi| = \begin{pmatrix} a^2 & ab e^{-i\phi} \\ ab e^{i\phi} & b^2 \end{pmatrix}.$$

Once the phase ϕ randomizes, the off-diagonal elements average to zero, and all that remains is a classical probability distribution on the diagonal: no wave function can reconstruct a purely diagonal density matrix with multiple nonzero entries. Moreover,

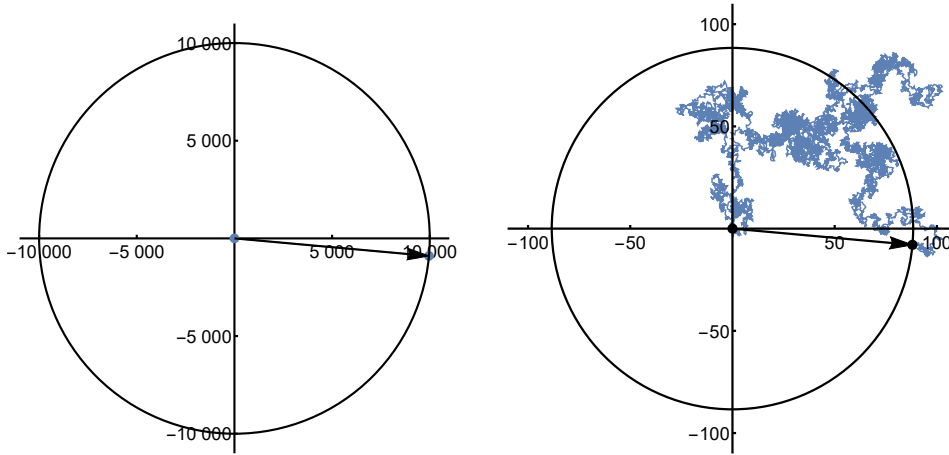


Figure 2.9: The motion of a particle at a speed of one unit distance per unit time, over 10 000 units of time. **(Left)** Ballistic transport is free of scattering, so the distance traveled is proportional to the time t . **(Right)** In the diffusive limit, the direction of motion is randomized many times, and the average distance from the origin is proportional to \sqrt{t} (there is a statistical prefactor < 1). The same power law holds for diffusion in three dimensions.

since it is the off-diagonal components of the density matrix that describe the time evolution, the particle has ceased to oscillate between its basis states.

In short, the electron-hole correlations will decohere due to an accumulating phase difference even if their constituent quasiparticles remain coherent. Inelastic scattering events that couple the degrees of freedom of the quasiparticle to the environment through energy exchange add an additional source of decoherence [67, 68] that prevents them from interfering constructively. Together, dephasing and inelastic scattering limit coherent transport across the junction.

2.3.4 Energy, length and time scales

If no scattering occurs at all before a particle traverses the channel, then the total distance that it covers is simply proportional to the time:

$$L_{\text{ballistic}} = v_F t. \quad (2.44)$$

At the other extreme, if the scattering occurs over distances much smaller than the channel length and movement becomes Brownian, then the particle is said to diffuse, and the length it is expected to have traveled is proportional to the *square root* of the time:

$$L_{\text{diffusive}} = \sqrt{D t}, \quad (2.45)$$

where the diffusion constant is defined as $D = v_F \ell_e / 3$. The difference between these two limits is illustrated in Fig. 2.9. Consequently, the shorter the mean free path ℓ_e , the longer it takes to cover some distance L , one of the concerns when increasing the doping.

As the wavefunction spreads out through the channel, it interacts with different potentials at different positions, which each change the effective $E(t)$ of the part of the wavefunction that passes there. In the case of coherent transport, the spread in phase between the paths that this causes should be less than π ,

$$\Delta\phi = \frac{t \delta E}{\hbar} \lesssim \pi. \quad (2.46)$$

This allows us to directly relate the time that the particle spends in the channel to the energy required to affect its coherence, the Thouless energy E_{Th} . In the ballistic case, since the time spent in the channel is simply proportional to its length,

$$E_{\text{Th,ballistic}} = \frac{\hbar v_F}{L_{\text{channel}}}, \quad (2.47)$$

while in the diffusive limit the transfer time goes as the *square* of the length [69]:

$$E_{\text{Th,diffusive}} = \frac{\hbar D}{L_{\text{channel}}^2}. \quad (2.48)$$

This provides a characteristic energy scale for the junction: any kind of interaction on the order $E > E_{\text{Th}}$ will cause decoherence before the particle can reach the other end¹⁷. Conversely, we can associate any scattering energy with a coherence length [49],¹⁸

$$\xi_{\text{E,ballistic}} = \frac{\hbar v_F}{2\pi E}, \quad \xi_{\text{E,diffusive}} = \sqrt{\frac{\hbar D}{2\pi E}}. \quad (2.49)$$

Finally, we can think of these processes in terms of coherence times τ , and compare them to the time it takes to cross the channel:

$$\tau_E = \frac{\hbar}{E}, \quad t_{\text{ballistic}} = \frac{L}{v_F}, \quad t_{\text{diffusive}} = \frac{L^2}{D}. \quad (2.50)$$

In the end, of course all these ways of looking at it are equivalent; interactions will suppress superconducting transport exponentially if they cause decoherence over length scales shorter than the channel, or if they do so in less time than it takes to cross it, and the transport will be limited by the Thouless energy. If the channel is traversed faster than this, transport will instead be limited by the properties of the interface and the magnitude of the superconducting gap. This was captured most clearly in 2001 with an experimental and theoretical study of SNS junctions of varying length [70], the first figure of which is reproduced in Fig. 2.10.

2.4 Transistors

For a qubit platform to be “scalable” means in broad terms that the resources needed for both the operation and the fabrication of circuits does not grow too quickly with the total number of physical qubits. In section 2.2 we discussed how operation at scale has led to the choice of gate-tunable transmons, while fabrication could be scaled by moving towards an all-CMOS qubit. CMOS technology has a strong track record: without much change in the total amount of time spent in fabricating a single die, the transistor count has increased exponentially from 4 in the late 50’s to more than 10^{10} at the time of writing [71]. If gatemonos could be fabricated using this same technology, not only could large numbers be patterned simultaneously and reproducibly, it would also allow for the on-chip integration of the classical logic circuits that are necessary for real-time error correction.

The central idea in this approach is to use a modified transistor as gate-tunable Josephson junction, an example of which is shown in Fig. 2.11. Below we will discuss some of their principles.

¹⁷Here E represents the energy *difference* between the different paths along the junction. If all paths are affected by identical energy shifts, no broadening of ϕ will occur and so neither will any decoherence.

¹⁸A factor 2π is introduced to scale ξ_E such that $|\psi(L)|^2 \propto e^{-\xi/\xi_E}$ [70].

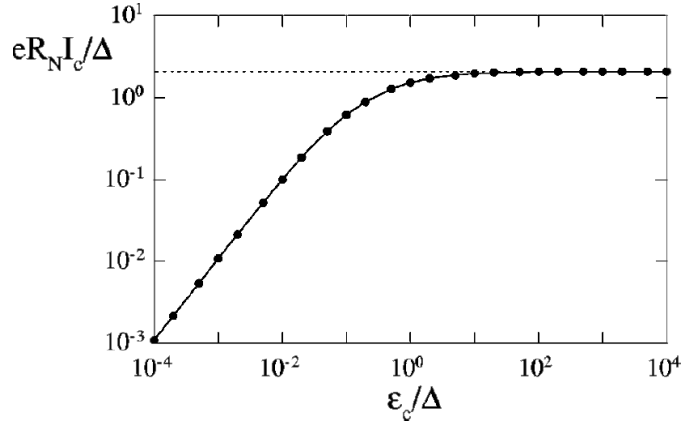


Figure 2.10: Fig. 1 from Ref. 70: on the x -axis it the normalized Thouless energy $\varepsilon_c/\Delta = E_{\text{Th}}/\Delta$, on the y -axis the normalized $I_c R_N$ product. This product, the voltage that would have to be applied in the normal state to get $I_N = I_{S,c}$, and a common experimental measure of the quality of a junction, is limited by the Thouless energy if the weak link is long ($E_{\text{Th}} < \Delta$), but by the gap for short junctions ($E_{\text{Th}} > \Delta$).

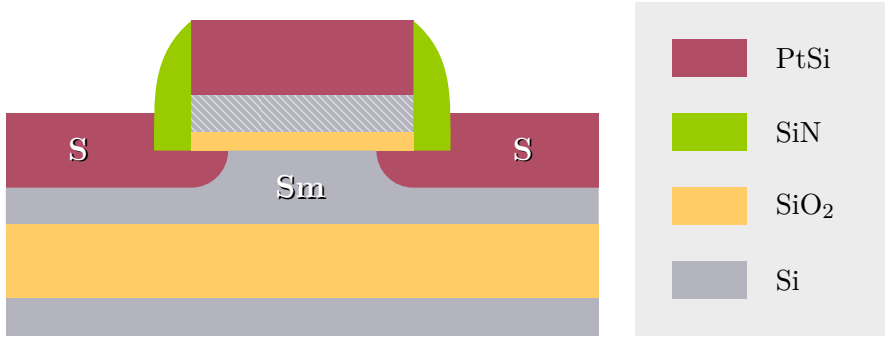


Figure 2.11: A Josephson field effect transistor (JoFET) has superconducting source and drain, here made of PtSi. It functions like a transistor, except that the current that is modulated is a resistanceless supercurrent.

2.4.1 The field effect

To understand the effect that an electrostatic field has on the current through a junction, consider first the simplified case in which there are no barriers to overcome, and an infinite reservoir of carriers is available to be attracted underneath the gate oxide. The total current that will flow from source to drain is then proportional to the charge accumulated in the channel, times the rate at which that charge moves across the device [72],

$$I_d = \underbrace{C(V_g - V_{\text{th}})}_{Q/A} \overbrace{W \langle v \rangle}^{A/t}. \quad (2.51)$$

Applying a gate voltage V_g then attracts charges proportional to the oxide's capacitance, which will linearly increase the current in the case that the drift and diffusion rates remain constant.

In reality this behavior is only observed above some threshold gate voltage V_{th} where all the carriers are of the same sign. Below this voltage, small numbers n_e and n_h of both

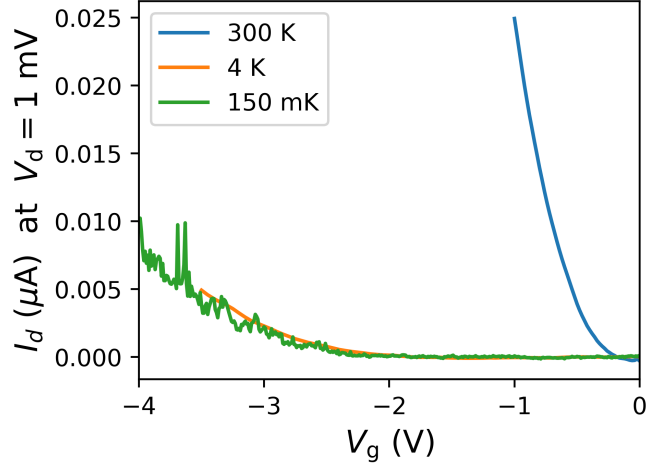


Figure 2.12: The field effect measured at different temperatures in a 50 nm long, 2.5 μm wide SBMOSFET with PtSi contacts. We observe a shift in the threshold voltage V_{th} due to the freezing out of intrinsic carriers, as well as the charging of individual dopants at low temperatures.

electrons and holes are present, the product of which is constant at a fixed temperature, and equal to the square of the intrinsic carrier density n_i of the semiconductor,

$$n_e n_h = n_i^2. \quad (2.52)$$

This number of intrinsic carriers in turn depends on the temperature of the material [73],

$$n_i^2 = C_n T^3 e^{-E_g/k_B T}, \quad (2.53)$$

where C_n is some constant such that the intrinsic concentration n_i of silicon is 10^{10}cm^{-3} at room temperature ($T = 300 \text{K}$) [74], and E_g the band gap of silicon. When a material is doped, the Fermi level is moved towards the conduction or valence band for (n and p doping, respectively), changing the electron and hole concentrations. For example, in the case of a p-doped (p for positive) channel with boron doping [75],

$$n_h = n_B, \quad n_e = \frac{n_i^2}{n_B}, \quad (2.54)$$

where a wide range of n_B from $1 \times 10^{13} \text{cm}^{-3}$ to $1 \times 10^{18} \text{cm}^{-3}$ has been used in industry [76]. This relationship $n_h = n_B$ does not hold at all temperatures however, since at low temperatures dopants can freeze out, which is an especially serious concern for holes in silicon [77].

Below the threshold voltage, when the channel is essentially depleted, charge can still cross the channel through thermionic emission proportional to $\exp(eV_g/k_B T)$. This gives an exponentially decaying tail to the drain current as a function of gate voltage,

$$I_d(V_g) \propto \exp\left(\frac{eV_g}{k_B T(1 + \alpha)}\right). \quad (2.55)$$

Here α is called the “level-arm parameter”, and quantifies the reduction in the field effect of the gate due to a parasitic capacitance between the channel and the substrate,

$$\alpha = \frac{C_{\text{channel-substrate}}}{C_{\text{channel-gate}}}. \quad (2.56)$$

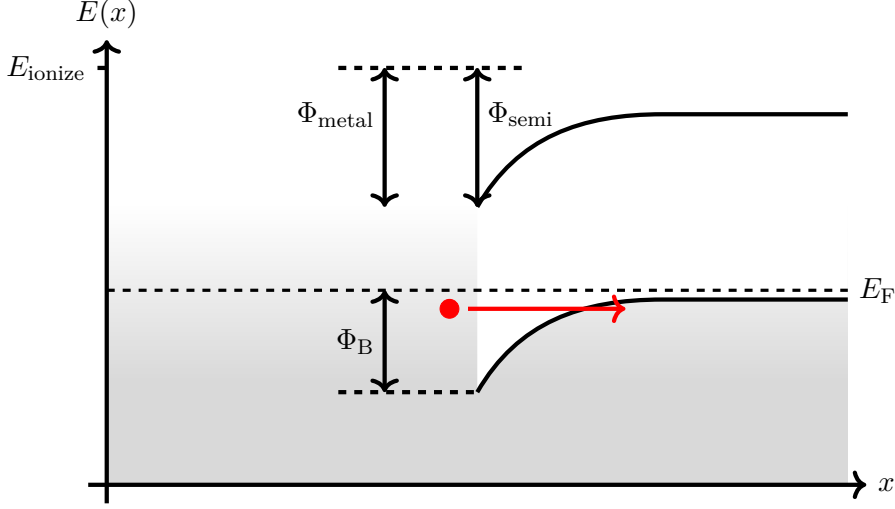


Figure 2.13: The formation of a Schottky barrier for holes with a height Φ_B at the interface between a metal and a p-doped semiconductor. Shown in red is a hole that tunnels through the barrier.

In general it is desirable to have only a narrow range of gate voltages where a sub-threshold current I_{off} is observed (and thus a small α). The degree to which this is achieved in a device is often expressed in the “sub-threshold swing”, which is the increase in gate voltage required to multiply the drain current ten-fold. We calculate this as the inverse of the slope of the logarithm of the drain current versus the gate voltage,

$$S = \left(\frac{d \log_{10}(I_d)}{d|V_g|} \right)^{-1} = \ln(10) \frac{k_B T}{e} (1 + \alpha). \quad (2.57)$$

In the ideal case where $\alpha = 0$, this gives a sub-threshold swing of 59.6 mV dec^{-1} at 300 K. At low temperatures the freezing out of dopant carriers and the suppression of thermionic emission mean that both the total current is reduced, and the threshold voltage is shifted. This is shown in Fig. 2.12 for a device that we will discuss further in chapter. □.

2.4.2 Schottky barriers

A phenomenon that becomes especially relevant at low temperatures is the formation of a Schottky barrier at the interface between a doped semiconductor (such as a p-doped channel) and a metal (such as PtSi contacts). The appearance of this tunnel barrier can be understood as follows. Since the Fermi level is the energy up to which the states are occupied (give or take $k_B T$), it needs to be constant across a device with finite resistivity. When two materials with different Fermi levels are brought together, a particle occupying a higher-energy state on the one side could gain energy by moving to the other, which ensures that the maximum occupied energy on either side of the interface equals out (the Fermi levels are “pinned” to each other). At the same time, since the two materials are now conductively coupled, a charge from either side of the interface can be moved through the other, which means that the total energy of removing a charge in either material (the work function Φ) also needs to match close to the interface. We saw above that doping the semiconductor will shift the Fermi level within the band gap closer to either the

valence or conduction band, while the work function of the material remains unchanged. For both the Fermi levels and the work functions to match at the interface, the valence and conduction bands thus have to be bent [78], as shown in Fig. 2.13. Different heuristics for predicting the degree and direction of band bending have been proposed, taking into account charge buildup at the interface and focusing on the work functions [79] or including also surface states that effectively pin the Fermi level [80], neither of which fully predict Schottky barrier heights at metal/semiconductor interfaces [81].

At higher temperatures, the Schottky barrier can be overcome by thermionic emission, but as the device is cooled down, $k_B T$ becomes smaller than the Schottky barrier height (SBH) Φ_B , and it will act as a tunneling barrier for the electrons or holes. Since this barrier needs to be crossed twice in the case of Andreev reflection, while it is traversed only once by an individual quasiparticle, Schottky barriers suppress supercurrent more than they do normal transport, and are therefore usually undesirable in JoFETs where coupling to quasiparticles should be minimized.

2.4.3 JoFET considerations

Since we can always use the gate voltage to completely deplete the channel and switch off the Josephson coupling, it is in general our goal to fabricate devices with large critical currents I_c . Note that since we also need a very large capacitance C in a transmon qubit, we are not in principle limited by the size of the device, and could achieve this by making junctions arbitrarily wide. Assuming a dielectric constant of $\kappa = 3.9$ for SiO_2 [82], a gate oxide between 5 and 20 nm thick and a target charging energy of 100 to 300 MHz (65 to 200 neV), we find that channels can be up to $60 - 700 \mu\text{m}^2$ large. There are other reasons why you may want to avoid too large devices though, such as flux noise sensitivity, gate leakage, and losses to parasitic two-level systems in the oxide [12, 13]. It is therefore preferable to derive the charging energy from a separate shunting capacitance made of a material with a lower loss tangent [12] (proportional to the density of two-level systems [13]) than SiO_2 . In practice then, we are looking for transistor devices with as large a critical current as possible, while limiting its size. This implies a preference for high-mobility semiconducting materials [83] and transparent interfaces between the superconductor and channel [29].

It is less straightforward to determine what level of doping should be aimed for in a specific system. On the one hand, the supercurrent carrying capacity of a junction is closely related to its normal-state conductance (especially for short junctions, where $I_c \propto G_N \Delta$) [49, 84, 85], which in turn is the product of mobility and carrier concentration, suggesting that increasing the doping level of the channel could improve its behavior. Even in longer junctions one can expect diffusion to be aided by higher carrier densities [31], despite a reduction in mean free path [61]. Moreover, dopants could provide a resonance-percolation trajectories [59] when other modes of transport have already been suppressed [60]. Higher doping concentrations would in turn also reduce the width of the Schottky barrier [86, 87], exponentially increasing its tunneling probability at low temperatures. For devices with silicon channels, it may therefore be necessary to have higher doping, especially near the interface with the metal contacts. On the other hand, the scattering sites that this introduces may in fact bring down the total transmission through the channel by reducing the mobility [88].

A different set of concerns is raised by the choice of contacting metal. Just as the coupling between two transmons depends on the matching of their impedances, so

does the transmission probability of a quasiparticle through an interface depend on the matching of its inertia on either side, given by the electron effective mass [87]. Differences in lattice spacing between the two materials will also induce a shift in momentum as the incoming and outgoing particles occupy different Bloch states, both creating an effective potential and narrowing the recombination window by causing inelastic scattering across the interface. Other factors that can limit the transparency are the Schottky barrier height that we discussed above, as well as the general quality of the interface in terms of grain boundaries, impurities and oxide layers. The application of the transistor for superconducting transport adds the requirement that the contacting metal have a high superconducting critical temperature. In the next chapter we will see how all of these aspects are addressed by fabricating the contacts with superconducting silicides.

Bibliography

- [1] Audrey Cottet. “Implementation of a quantum bit in a superconducting circuit”. PhD thesis. Université Paris 6, 2002.
- [2] Yu Chen et al. “Qubit architecture with high coherence and fast tunable coupling”. In: *Physical Review Letters* 113.22 (2014), p. 220502.
- [3] Carver Mead and Lynn Conway. *Introduction to VLSI systems*. Addison-Wesley Reading, MA, 1980.
- [4] Lynn Conway. “Reminiscences of the VLSI revolution: how a series of failures triggered a paradigm shift in digital design”. In: *IEEE Solid-State Circuits Magazine* 4.4 (2012), pp. 8–31.
- [5] Jonathan P Dowling. *Schrödinger’s Killer App: Race to Build the World’s First Quantum Computer*. CRC Press, 2013. ISBN: 978-1439896730.
- [6] David C McKay et al. “Universal gate for fixed-frequency qubits via a tunable bus”. In: *Physical Review Applied* 6.6 (2016), p. 064007.
- [7] Jay M Gambetta, Jerry M Chow, and Matthias Steffen. “Building logical qubits in a superconducting quantum computing system”. In: *NPJ Quantum Information* 3.1 (2017), pp. 1–7.
- [8] Thorvald Wadum Larsen et al. “Semiconductor-nanowire-based superconducting qubit”. In: *Physical Review Letters* 115.12 (2015), p. 127001.
- [9] G De Lange et al. “Realization of microwave quantum circuits using hybrid superconducting-semiconducting nanowire Josephson elements”. In: *Physical Review Letters* 115.12 (2015), p. 127002.
- [10] Philip Krantz et al. “A quantum engineer’s guide to superconducting qubits”. In: *Applied Physics Reviews* 6.2 (2019), p. 021318.
- [11] John M Martinis, Michel H Devoret, and John Clarke. “Energy-level quantization in the zero-voltage state of a current-biased Josephson junction”. In: *Physical Review Letters* 55.15 (1985), p. 1543.
- [12] Jens Koch et al. “Charge-insensitive qubit design derived from the Cooper pair box”. In: *Physical Review A* 76.4 (2007), p. 042319.
- [13] John M Martinis et al. “Decoherence in Josephson qubits from dielectric loss”. In: *Physical Review Letters* 95.21 (2005), p. 210503.
- [14] Gianluigi Catelani et al. “Relaxation and frequency shifts induced by quasiparticles in superconducting qubits”. In: *Physical Review B* 84.6 (2011), p. 064517.
- [15] L Sun et al. “Measurements of quasiparticle tunneling dynamics in a band-gap-engineered transmon qubit”. In: *Physical Review Letters* 108.23 (2012), p. 230509.

- [16] David J Griffiths. *Introduction to quantum mechanics, second international edition*. Pearson Prentice Hall, 2005. ISBN: 9787111182948.
- [17] GS Paraoanu. “Microwave-induced coupling of superconducting qubits”. In: *Physical Review B* 74.14 (2006), p. 140504.
- [18] Kevin C Miao et al. “Universal coherence protection in a solid-state spin qubit”. In: *Science* 369.6510 (2020), pp. 1493–1497.
- [19] DI Schuster et al. “Resolving photon number states in a superconducting circuit”. In: *Nature* 445.7127 (2007), pp. 515–518.
- [20] Andrew Addison Houck et al. “Generating single microwave photons in a circuit”. In: *Nature* 449.7160 (2007), pp. 328–331.
- [21] Michel H Devoret et al. “Quantum fluctuations in electrical circuits”. In: *Les Houches, Session LXIII* 7.8 (1995).
- [22] S.E. Rasmussen et al. “The superconducting circuit companion – an introduction with worked examples”. In: (2021). arXiv preprint: <https://arxiv.org/abs/2103.01225>.
- [23] TD Clark, RJ Prance, and ADC Grassie. “Feasibility of hybrid Josephson field effect transistors”. In: *Journal of Applied Physics* 51.5 (1980), pp. 2736–2743.
- [24] Denis Vion et al. “Manipulating the quantum state of an electrical circuit”. In: *Science* 296.5569 (2002), pp. 886–889.
- [25] Leonardo DiCarlo et al. “Demonstration of two-qubit algorithms with a superconducting quantum processor”. In: *Nature* 460.7252 (2009), pp. 240–244.
- [26] Yong-Joo Doh et al. “Tunable supercurrent through semiconductor nanowires”. In: *Science* 309.5732 (2005), pp. 272–275.
- [27] Mark Ebel et al. “Supercurrent-phase relationship of a Nb/ In As (2 DES)/ Nb Josephson junction in overlapping geometry”. In: *Physical Review B* 71.5 (2005), p. 052506.
- [28] P Krogstrup et al. “Epitaxy of semiconductor–superconductor nanowires”. In: *Nature Materials* 14.4 (2015), pp. 400–406.
- [29] Florian Vigneau et al. “Germanium quantum-Well Josephson field-effect transistors and interferometers”. In: *Nano Letters* 19.2 (2019), pp. 1023–1027.
- [30] Kushagra Aggarwal et al. “Enhancement of Proximity Induced Superconductivity in Planar Germanium”. In: *arXiv preprint arXiv:2012.00322* (2020).
- [31] Anatoly F Volkov and Hideaki Takayanagi. “Effect of gate voltage on critical current in controllable superconductor–normal-metal–superconductor Josephson junctions”. In: *Physical Review B* 53.22 (1996), p. 15162.
- [32] Lukas Hofstetter et al. “Cooper pair splitter realized in a two-quantum-dot Y-junction”. In: *Nature* 461.7266 (2009), pp. 960–963.
- [33] John Bardeen, Leon N Cooper, and John Robert Schrieffer. “Theory of superconductivity”. In: *Physical Review* 108.5 (1957), p. 1175.
- [34] John Bardeen, Leon N Cooper, and J Robert Schrieffer. “Microscopic theory of superconductivity”. In: *Physical Review* 106.1 (1957), p. 162.

- [35] John Bardeen and JR Schrieffer. “Chapter VI Recent Developments in Superconductivity”. In: *Progress in Low Temperature Physics*. Vol. 3. Elsevier, 1961, pp. 170–287.
- [36] Malcolm Beasley. “Michael Tinkham (1928–2010)”. In: *Nature* 468.7325 (2010), pp. 766–766.
- [37] Leon N Cooper. “Bound electron pairs in a degenerate Fermi gas”. In: *Physical Review* 104.4 (1956), p. 1189.
- [38] RC Dynes et al. “Tunneling study of superconductivity near the metal-insulator transition”. In: *Physical Review Letters* 53.25 (1984), p. 2437.
- [39] RE Glover III and M Tinkham. “Transmission of superconducting films at millimeter-microwave and far infrared frequencies”. In: *Physical Review* 104.3 (1956), p. 844.
- [40] RE Glover III and Ms Tinkham. “Conductivity of Superconducting Films for Photon Energies between 0.3 and 4 0 k T c”. In: *Physical Review* 108.2 (1957), p. 243.
- [41] AF Andreev. “The thermal conductivity of the intermediate state in superconductors”. In: *Soviet Physics JETP* 19.5 (1964), pp. 1228–1231.
- [42] Wolfgang Belzig, Christoph Bruder, and Gerd Schön. “Local density of states in a dirty normal metal connected to a superconductor”. In: *Physical Review B* 54.13 (1996), p. 9443.
- [43] Wolfgang Belzig et al. “Quasiclassical Green’s function approach to mesoscopic superconductivity”. In: *Superlattices and Microstructures* 25.5-6 (1999), pp. 1251–1288.
- [44] AF Andreev. “Electron spectrum of the intermediate state of superconductors”. In: *Soviet Physics JETP* 22.455 (1966), pp. 18–23.
- [45] GE Blonder, M Tinkham, and TM Klapwijk. “Transition from metallic to tunneling regimes in superconducting microconstrictions: Excess current, charge imbalance, and supercurrent conversion”. In: *Physical Review B* 25.7 (1982), p. 4515.
- [46] LD Landau and VI Ginzburg. “K teorii sverkhrovodimosti (On the theory of superconductivity)”. In: *Zhurnal Eksperimental’noi i Teoreticheskoi Fiziki (Journal of Experimental and Theoretical Physics)* 20.1064 (1950), pp. 546–568.
- [47] L. D. Landau and V. L. Ginzburg. “On the Theory of Superconductivity”. In: *On Superconductivity and Superfluidity: A Scientific Autobiography*. Springer Berlin Heidelberg, 2009, pp. 113–137. ISBN: 978-3-540-68008-6. DOI: 10.1007/978-3-540-68008-6_4.
- [48] Michael Tinkham. *Introduction to Superconductivity, Second Edition*. Dover Publications, 2004. ISBN: 978-0486435039.
- [49] PGf de Gennes. “Boundary effects in superconductors”. In: *Reviews of Modern Physics* 36.1 (1964), p. 225.
- [50] LP Gor’Kov. “On the energy spectrum of superconductors”. In: *Soviet Physics JETP* 7.505 (1958), p. 158.
- [51] CWJ Beenakker. “Quantum transport in semiconductor-superconductor microjunctions”. In: *Physical Review B* 46.19 (1992), p. 12841.

- [52] PAM Benistant et al. “Angular relation and energy dependence of Andreev reflection”. In: *Physical Review B* 32.5 (1985), p. 3351.
- [53] H De Raedt, K Michielsen, and TM Klapwijk. “Andreev reflection in nanoscale metal-superconductor devices”. In: *Physical Review B* 50.1 (1994), p. 631.
- [54] H De Raedt, K Michielsen, and TM Klapwijk. “Andreev reflection”. Animations of the results discussed in Ref. 53, obtained from first author upon request.
- [55] KK Likharev. “Superconducting weak links”. In: *Reviews of Modern Physics* 51.1 (1979), p. 101.
- [56] Harry Kroger. “Josephson devices coupled by semiconductor links”. In: *IEEE Transactions on Electron Devices* 27.10 (1980), pp. 2016–2026.
- [57] CL Huang and T Van Duzer. “Josephson tunneling through locally thinned silicon wafers”. In: *Applied Physics Letters* 25.12 (1974), pp. 753–756.
- [58] Eddy Simoen et al. “Freeze-out effects on NMOS transistor characteristics at 4.2 K”. In: *IEEE Transactions on Electron Devices* 36.6 (1989), pp. 1155–1161.
- [59] IM Lifshitz and V Kirpichenkov. “Tunnel transparency of disordered systems”. In: *Zhurnal Eksperimental’noi i Teoreticheskoi Fiziki (Journal of Experimental and Theoretical Physics)* 77.3 (1979), pp. 989–1016.
- [60] LG Aslamazov and MV Fistul. “Resonant tunneling in superconductor-semiconductor-superconductor junctions”. In: *Soviet Physics-JETP* 56.3 (1982), pp. 666–670.
- [61] LG Aslamazov and MV Fistul. “Temperature dependence of the critical current at superconductor-semiconductor-superconductor junctions”. In: *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki* 81 (1981), pp. 382–397.
- [62] AW Kleinsasser et al. “Crossover from tunneling to metallic behavior in superconductor-semiconductor contacts”. In: *Applied Physics Letters* 57.17 (1990), pp. 1811–1813.
- [63] Scott Aaronson. “The complexity of agreement”. In: *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*. 2005, pp. 634–643.
- [64] Lov K Grover and Ben Reichardt. “Quantum Search”. In: *Encyclopedia of Algorithms* (2008). DOI: 10.1007/978-3-642-27848-8_317-2.
- [65] Andrew Childs. “The Power of Quantum Computers”. Q2B20 Conference. Dec. 8, 2020.
- [66] BJ Van Wees et al. “Excess conductance of superconductor-semiconductor interfaces due to phase conjugation between electrons and holes”. In: *Physical Review Letters* 69.3 (1992), p. 510.
- [67] H Dieter Zeh. “On the interpretation of measurement in quantum theory”. In: *Foundations of Physics* 1.1 (1970), pp. 69–76.
- [68] Maximilian Schlosshauer. “Quantum decoherence”. In: *Physics Reports* 831 (2019), pp. 1–57.
- [69] JT Edwards and DJ Thouless. “Numerical studies of localization in disordered systems”. In: *Journal of Physics C: Solid State Physics* 5.8 (1972), p. 807.
- [70] P Dubos et al. “Josephson critical current in a long mesoscopic SNS junction”. In: *Physical Review B* 63.6 (2001), p. 064502.
- [71] Max Roser and Hannah Ritchie. “Technological Progress”. In: *Our World in Data* (2013). <https://ourworldindata.org/technological-progress>.

- [72] JR Hook and HE Hall. *Solid State Physics*. 2nd ed. John Wiley & Sons, Ltd, 2010. ISBN: 978-0471-92805-8.
- [73] Martin A Green. “Intrinsic concentration, effective densities of states, and effective mass in silicon”. In: *Journal of Applied Physics* 67.6 (1990), pp. 2944–2954.
- [74] AB Sproul and MA Green. “Improved value for the silicon intrinsic carrier concentration from 275 to 375 K”. In: *Journal of Applied Physics* 70.2 (1991), pp. 846–854.
- [75] Malgorzata Chrzanowska-Jeske and Richard C Jaeger. “BILOW-simulation of low-temperature bipolar device behavior”. In: *IEEE Transactions on Electron Devices* 36.8 (1989), pp. 1475–1488.
- [76] J-P Colinge and Cynthia A Colinge. *Physics of semiconductor devices*. Springer Science & Business Media, 2005.
- [77] B Lengeler. “Semiconductor devices suitable for use in cryogenic environments”. In: *Cryogenics* 14.8 (1974), pp. 439–447.
- [78] Walter Schottky. “Vereinfachte und erweiterte Theorie der Randschicht-gleichrichter”. In: *Zeitschrift für Physik* 118.9 (1942), pp. 539–592.
- [79] Nevill Francis Mott. “The theory of crystal rectifiers”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 171.944 (1939), pp. 27–38.
- [80] John Bardeen. “Surface states and rectification at a metal semi-conductor contact”. In: *Physical Review* 71.10 (1947), p. 717.
- [81] J Tersoff. “Schottky barrier heights and the continuum of gap states”. In: *Physical Review Letters* 52.6 (1984), p. 465.
- [82] Angus I Kingon, Jon-Paul Maria, and SK Streiffer. “Alternative dielectrics to silicon dioxide for memory and logic devices”. In: *Nature* 406.6799 (2000), pp. 1032–1038.
- [83] Lucas Casparis et al. “Superconducting gatemon qubit based on a proximitized two-dimensional electron gas”. In: *Nature Nanotechnology* 13.10 (2018), pp. 915–919.
- [84] PG de Gennes. “Superconductivity of metals and alloys”. In: (1966).
- [85] AW Kleinsasser. “Critical currents of semiconductor-coupled Josephson weak links”. In: *IEEE Transactions on Magnetism* 27.2 (1991), pp. 2589–2593.
- [86] Nevill Francis Mott. “Note on the contact between a metal and an insulator or semi-conductor”. In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 34. 4. Cambridge University Press. 1938, pp. 568–572.
- [87] Walter A Harrison, Andreas Goebel, and Paul A Clifton. “Effective-mass theory of metal-semiconductor contact resistivity”. In: *Applied Physics Letters* 103.8 (2013), p. 081605.
- [88] Amir Sammak et al. “Shallow and undoped germanium quantum wells: a playground for spin and hybrid quantum technology”. In: *Advanced Functional Materials* 29.14 (2019), p. 1807613.

Chapter 3

Silicides

3.1 Introduction

We saw in the previous chapter that silicon-based transistors could serve as the weak links in superconducting qubits. Certainly silicon is not the *ideal* material for Josephson junctions, with its significant Schottky barriers at any useful doping level, low mobility and early onset of dopant freeze-out. It is, however, *compatible*. A device based on silicon can, within some constraints on materials and fabrication techniques, readily be integrated alongside existing cryo-CMOS circuits [1–3], and leverage fabrication techniques that have been honed over decades of innovation. The conventional MOSFET design relies on silicides, either by depositing pure metal that then reacts with the underlying silicon, or co-depositing both metal and silicon in the right stoichiometry [4–6]. While the former has the benefit of self-aligning the formed silicide with the exposed patches of silicon (known as the SALICIDE process) and diffusing any impurities introduced at the metal-silicon interface, the latter allows to select silicide phases that are otherwise unlikely to form, and often requires lower formation temperatures [7]. In either case, the resulting compound’s metal content ensures high conductance, while the incorporation of silicon generally provides a low contact resistance to the channel. By selecting silicides that also happen to be superconducting, these methods can be applied directly to the fabrication of JoFETs. The two materials that we will focus on in this chapter are V_3Si , for its superconducting properties, and $PtSi$, which boasts established integration processes.

We will begin with V_3Si , which is interesting first of all for its promise in JoFET applications; it has the highest superconducting critical temperature of any silicide that we are aware of [8–10], and has excellent lattice matching to (111) Si [11] — off by only 0.4%, much better than the best that can be achieved with $PtSi$ (3.0%) [12]. Second, it is also interesting for the highly unusual dependence of its thermal, conductive and superconducting properties on the strain that it is under [13–15]. Third, in systems of vanadium and silicon, this is the most elusive of the congruently forming phases, as it is neither the first phase to form [16–18], nor does it have the greatest enthalpy of formation (and is therefore not thermodynamically favorable) [19]. The first of these justifies our efforts, while the challenge of the other two makes it all worth our while.

Two approaches were tried in forming V_3Si thin films for our studies. The first took place at Uppsala Universitet, where the readily installed vanadium target allowed us to perform depositions on a variety of substrates. Reports in the literature had suggested that the formation of the right phase could be triggered by methods that can broadly be

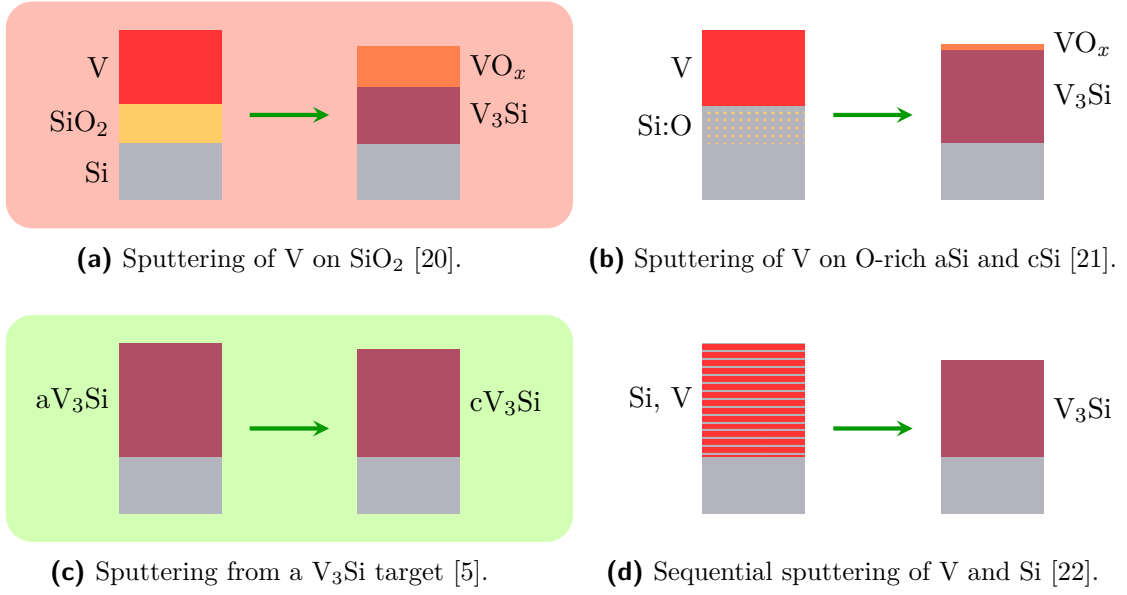


Figure 3.1: A selection of methods described in the literature to form V_3Si thin films, one of which has been reproduced without (indicated in red) and another with success (green).

grouped into three mechanisms. Number one would see silicon, the dominant diffusing species in VSi_2 during its formation, captured with oxygen either in the form of a native oxide [17, 21, 23, 24] (see Fig. 3.1a), or with impurities at a concentration of a few %at in deposited silicon [21] (see Fig. 3.1b), effectively decoupling the silicide from the Si reservoir and thus allowing the otherwise much more slowly diffusing vanadium to gain a foothold. This result of inverting the relative mobilities of the two species this way may be interpreted in the context of the Cu_3Au effect: “*the first phase to nucleate is the phase rich in the high-mobility constituent*”, or vice versa, “*the mobility is higher for the majority constituent*” [25]. In the second method, the Si substrate is amorphized prior to deposition [21, 26], likely aiding V_3Si formation by lowering the energy barrier to nucleation (see also Fig. 3.1b). The third would overcome the limitation posed by vanadium’s slow diffusion by immediately distributing the two species in the right stoichiometry [5, 27] or at least approaching this with thin alternating layers of Si and V [22] (see Figs. 3.1d and 3.1c). Since the target available in Uppsala was pure metal, and the second (and last) target slot in the deposition tool was to be used for titanium to allow for in-situ capping with TiN, the first campaign was limited to deposition of pure V on various substrates.

3.2 V_3Si formation by pure metal deposition

Perhaps an equal amount of time was spent preparing and analyzing the samples where pure vanadium was deposited (the “Uppsala samples”), as was used for those with V_3Si deposition from the compound target (the “Grenoble samples”). However, because it was ultimately concluded that we may only have successfully formed V_3Si on a single sample from this first batch (one that was annealed within hours of deposition), little of scientific value came out, and so this section will be short. Yet, precisely because this turned out to be a dead end, and because so much time was lost, it is worth tracing our

Table 3.1: A short summary of the wafer types analyzed for pure vanadium deposition in Uppsala. Treatments in parentheses are only performed on some samples.

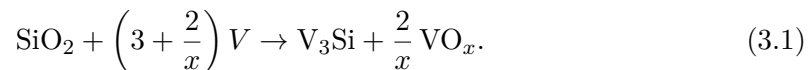
Wafer	Substrate	Treatment	Deposition
A	bulk Si	(HF)	V (+ TiN)
B	Si + 20 nm thermal SiO ₂	(HfO ₂ ALD)	V (+ TiN)
C	amorphized SOI with O doping (0.1–5%)	HF	V + TiN
D	Si + 40 nm Si ₃ N ₄	None	V + TiN
E	SOI with O doping (0.1–10%)	HF	V + TiN

steps, and we will conclude with some valuable lessons about methodology.

A total of 14 silicon wafers were prepared in Grenoble, 8 of which with a diameter of 300 mm and 6 of 200 mm, each with different surface preparations, giving us access to a range of different substrate parameters. This set was intentionally designed to be broader than that which could be fully analyzed, such that a series of quick tests in Uppsala could narrow down the wafers of interest. Eventually a total of five wafers were selected for deposition, shown in Table 3.1, allowing us to study the effects of oxide (both native and thermal), oxygen implantation, substrate pre-amorphization and a combination of the latter two. A wafer with silicon nitride was included to test the viability of this material as a spacer in CMOS transistors with V₃Si contacts (any sign of reaction with the silicide would rule it out). Since vanadium is known to oxidize extremely rapidly, it was necessary to cap the metal in-situ, for which we chose a 15 nm layer of TiN. This material was chosen not just for its availability in Uppsala, but also because consultation of binary phase diagrams for V-N and V-Ti [28] indicated stability up to extremely high temperatures of 3290 °C for TiN, while V_xN_{1-x} and V_xTi_{1-x} were only stable up to 2340 and 850 °C, respectively. This relative stability of TiN was taken to be an indicator of thermodynamic favorability, and thus unlikelihood of reaction with the underlying V, which assured us of its suitability. This conclusion was further supported by its routine use as a capping material for the silicidation of other metals, notably Pt, which will be discussed in section 3.4. As we will see later, our confidence in this choice was disastrously misplaced.

3.2.1 Early encouragement

Based on multiple reports in the literature, it was expected that vanadium deposited onto a 20 nm layer of thermal SiO₂ would upon thermal processing lead to the formation of a V₃Si layer. Whether the vanadium was deposited onto heated substrates [22], or at room temperature and annealed later [24], a maximum superconducting critical temperature was obtained at deposition/annealing temperatures of around 840 °C. During this reaction, the oxygen is transported to the top of the vanadium layer much faster than the V₃Si forms [17], forming there a vanadium oxide layer of about the same thickness as the consumed SiO₂,



If $x = 1$, which it will approach when the reaction is limited by the availability of SiO₂ [17], then 20 nm of SiO₂ will react with 36.8 nm of V to form 28.0 nm of V₃Si and 20.5 nm of VO. After the SiO₂ is fully consumed, the reaction may continue by consuming the underlying Si substrate [24], though it will slow down [29].

Table 3.2: Sheet resistance values and detected XRD peaks for samples that were annealed at 840 °C for 2 minutes under a N₂ atmosphere within hours or days after deposition. All samples were capped with 15 nm TiN.

Substrate	V (nm)	as-dep R_{\square}	annealed R_{\square}	XRD peaks detected
Si (HF)	80	6.06(28) Ω	3.95(11) Ω	39.6 (VTi or VSi ₂), 42.2 (V or VSi ₂), 42.8 (VTi or V ₃ Si), 49.2 (VSi ₂)
Si (no HF)	80	8.81(29) Ω	4.87(27) Ω	39.6 (VTi or VSi ₂), 42.2 (V or VSi ₂), 42.8 (VTi or V ₃ Si, much smaller), 49.2 (VSi ₂ , smaller)
20 nm SiO ₂	80	8.05(16) Ω	11.39(25) Ω	38.0 (VN, V ₃ Si), 44.0 (VN), 81.0 (VN, V ₃ Si)
20 nm SiO ₂ + 10 nm HfO ₂	100	11.86(11) Ω	15.3(41) Ω	No XRD performed
40 nm Si ₃ N ₄	100	6.06(17) Ω	3.95(11) Ω	37.9 (VN, V ₃ Si), 44.0 (VN)

To reproduce these results, layers of 80 nm of vanadium, capped with 15 nm of TiN, were deposited onto oxidized wafers, and immediately (within an hour) annealed during two minutes at 840 °C, with a high ramp rate of 20 °C/s. Sheet resistance measurements were taken before and after deposition (see Table 3.2), showing an increase from 8.05(16) Ω to 11.39(25) Ω . Sheet resistance values taken from calibration samples where 5, 10, 20, 40 and 80 nm of V was deposited, each followed by 15 nm of TiN, indicated that neither the oxidized substrate, nor the as-deposited TiN layer was highly resistive ($R_{\square} \gg 100 \Omega$), so that these values can be taken to be accurate for the layers of vanadium and its reaction products. In the most optimistic case, where the entire vanadium layer was consumed during reaction with the substrate (consuming first the oxide, then around 21 nm of silicon), around 83 nm of V₃Si would have formed, in which case the measured sheet resistance would translate to a resistivity of 94(2) $\mu\Omega$ cm. Though this is only a rough estimate, it is within the expected range for crystalline V₃Si, which for high-quality samples can be as low as 77 $\mu\Omega$ cm [30–33].

XRD analysis in a Bragg-Brentano configuration (a first scan in the 35–50° range, a second between 72 and 98°) indicated two peaks at 38.0 and 81.0, that could correspond to the (200) and (400) peaks of V₃Si, expected at 38.0 and 81.4. The absence of other XRD peaks of V₃Si in the $2\theta = 35$ –50° range, notably the (210) and (211) peaks at 42.7° and 47.1°, could be interpreted as strong epitaxial alignment of V₃Si with the substrate after full consumption of the SiO₂ layer. The large error in the position of the second peak was at the time ignored, while there are in fact many vanadium nitride compounds with different stoichiometries, two (VN and VN_{0.35}) with peaks right at 39.8 and 81.4 (PDF reference codes 00-025-1252 and 00-006-0624), and another with a peak at 81.0 (N_{0.9}V₂, reference code 00-030-1420). As mentioned earlier, TiN is well established as a capping material, with an extremely stable chemistry, so while binary compounds of V and N were indeed cross-checked during the XRD analysis in Uppsala, our confirmation bias (based on the V₃Si formation literature) led us to discard the hypothesis that vanadium nitrides had formed.

A comparison was made with samples where 80 nm of vanadium was deposited onto either HF-cleaned silicon (samples A42A,B,C,D, see Appendix. A.1), or silicon with remaining native oxide (A32A,B,C,D). After annealing at 840 °C, both showed three distinct peaks that could be attributed to VSi₂, two of which (42.8 and 49.2°) were smaller on the sample with native oxide than they were on the one that was cleaned

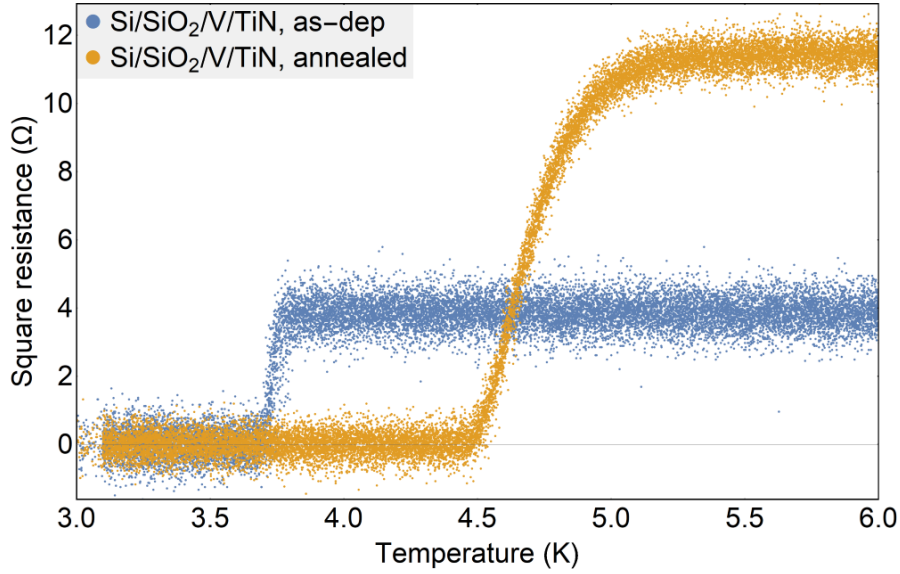


Figure 3.2: The resistance measured during slow heating from 3 to 6 K on an as-dep sample (blue) and a sample that was annealed in Uppsala at 840 °C (orange). The critical temperature is found to increase after annealing.

with HF, suggesting that HF limits VSi_2 growth. On all samples, a broad peak centered around 42.50 associated with pure vanadium was identified prior to annealing, which disappeared afterwards, confirming that the V was consumed. Furthermore, drops in sheet resistance after annealing (see Table 3.2) were also consistent with the formation of VSi_2 . Around 232 nm would form by complete consumption of the vanadium, giving a resistivity of $93(3) \mu\Omega \text{ cm}$, close to the $80 \mu\Omega \text{ cm}$ reported elsewhere for films cured for 30 min at 800 °C [34]. A smaller drop in sheet resistance on the uncleaned sample is consistent with the smaller XRD peaks observed and could indicate that the VSi_2 formation is slowed down by oxygen [21]. In short, the evidence points towards VSi_2 forming on a silicon substrate, and the inhibition of this reaction by the presence of a native oxide.

More apparent evidence of indeed having grown V_3Si on samples where V was deposited onto SiO_2 came in the form of low-temperature measurements performed later in Grenoble. Monocrystalline vanadium has a superconducting critical temperature of around 5.4 K [35–37], while that of high-quality TiN optimized for its superconductivity can be up to 4.8 K [38, 39], though polycrystalline thin films of either are expected to have critical temperatures well below these values. Two samples with 80 nm of V deposited onto 20 nm of oxide were measured at low temperature. One of these was as-deposited, while the other had been annealed at 840 °C for 2 minutes (with a ramp rate 20 °C/s) under an N_2 atmosphere. The critical temperatures were determined by following the resistance of $4 \times 10 \text{ mm}^2$ pieces during both cooling and heating at cryogenic temperatures (see Fig. 3.2). The first of these samples (blue curve) showed a critical temperature of 3.70 K with a transition width¹ of 0.09 K, which increased to 4.69 K with a width of 0.44 K for the second.

There are many ways that this increase in T_c can be explained. Critical temperatures

¹The temperature difference between the points at which 10 and 90% of the normal-state resistance is lost.

of around 5 K have been reported even for unannealed e-beam evaporated vanadium films of around 80 nm thick [40], so it is possible that the vanadium simply improved in quality through grain growth. The presence of the broad V peak at 42.5° on as-deposited samples indicates that the film is already crystalline before annealing, though the grain size is still on the order of the X-ray wavelength (a few times $\lambda = 1.54 \text{ \AA}$). The same could be true for the TiN capping layer, and of course it is possible that a vanadium nitride film has formed. However, reports in the literature suggest that the reaction of V with SiO_2 to form V_3Si should occur at 650 [24] or 700°C [16], and should be fast enough to consume at least half of the 80 nm film at 840°C [17]. Moreover, it is a priori not likely that a thermodynamically stable material like TiN would react at a temperature in this range at all. In short, with sheet resistance, XRD and critical temperature results all in principle consistent with the formation of V_3Si , there was good reason to be hopeful that we had indeed been successful in obtaining this phase.

3.2.2 Signs of trouble

Stimulated by the early results on the samples that were processed in Uppsala, we then moved on to evaluate a wider range of annealing temperatures in Grenoble. Five samples with SiO_2 substrates were selected: two with 40 nm of deposited vanadium, and three with 80 nm. The thinner films were annealed for 2 minutes under a nitrogen atmosphere at 700 or 800°C , while the thicker ones were annealed at 800 , 900 or 1000°C , after which the critical temperature of each was determined. Neither of the samples with 40 nm of vanadium exhibited superconductivity, while critical temperatures of 2.68 K (width of 0.54 K) and 4.00 K (0.91) were obtained after annealing at 900 and 1000°C (see Fig. 3.3). These results are clearly different from those obtained on the sample annealed in Uppsala, and could be explained either by different annealing conditions, or aging effects.

To gain more insight into the reactions that occur in this system, four more samples with 40 nm of vanadium were annealed, at temperatures both lower (500 and 600°C), and higher (900 and 1000°C) than the previous two. An increase in sheet resistance relative to that of the as-deposited samples was already observed at 500°C (see Fig. 3.4), far lower than 650°C , the lowest temperature at which a reaction between V and SiO_2 has been recorded [17, 20, 24]. No such increase in sheet resistance would be expected if the deposited vanadium would remain unreacted, and only increase its grain size.

In a separate set of experiments, reactions of vanadium on amorphized silicon were studied. It has been reported that both amorphizing the silicon, and introducing small amounts of oxygen at less than 10%at could trigger the growth of V_3Si instead of VSi_2 [21]. To reproduce this, a set of thinly oxidized 300 mm Si wafers were prepared in Grenoble, where oxygen was implanted to a final atomic concentration of 0, 0.1, 0.5, 1, 5 or 10% just below the native oxide. A second round of implantations followed with heavier Ge atoms, which reached a maximum concentration of 1.4% at a depth of 15 nm, amorphizing it down to 18 nm below the wafer surface. Layer of vanadium and TiN were later deposited in Uppsala, after which they were thermally processed in Grenoble. As on the SiO_2 substrates, no superconductivity was observed after annealing at 800°C , while the critical temperature increased with processing temperature up to a maximum of 8.17 K (see Fig. 3.5). This T_c is well above that reported for either vanadium or TiN, although it unfortunately does not surpass that of VN_x , which can reach 9.25 K when x approaches unity [41].

The sheet resistances of samples with amorphized silicon with various oxygen con-

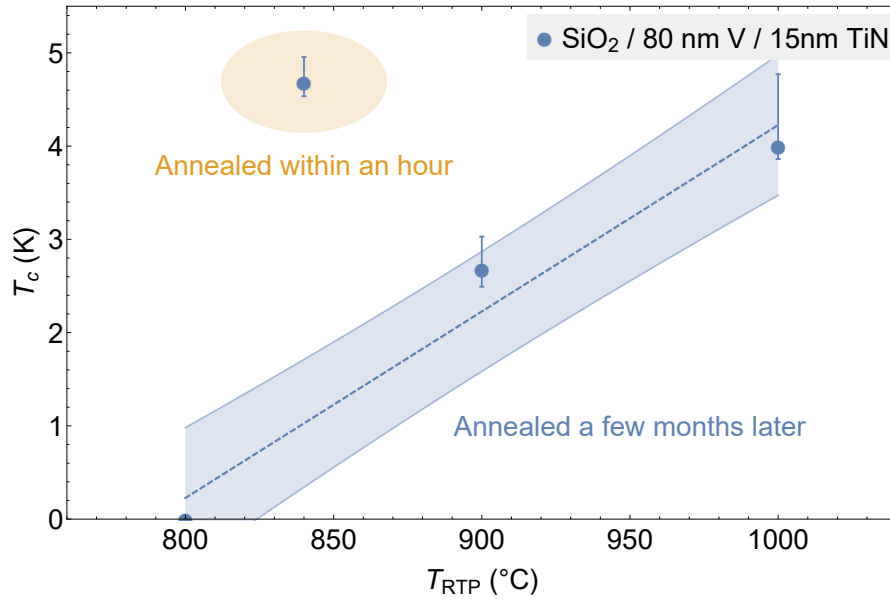


Figure 3.3: While samples that were stored for a few months (blue) also showed an increase in T_c after annealing, these critical temperatures were in stark contrast to the result obtained earlier.

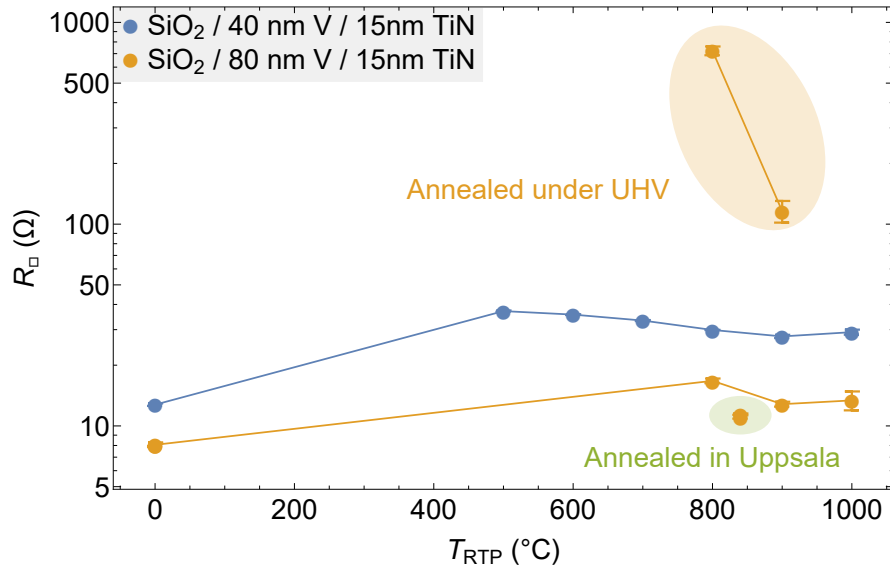


Figure 3.4: Sheet resistance of samples where 40 or 80 nm of V was deposited on 20 nm of SiO_2 . Three groups are indicated: samples annealed under N_2 atmosphere within an hour in Uppsala (green background), samples annealed under N_2 atmosphere a few months later in Grenoble (no colored background), and samples that were annealed under ultra high vacuum (orange background).

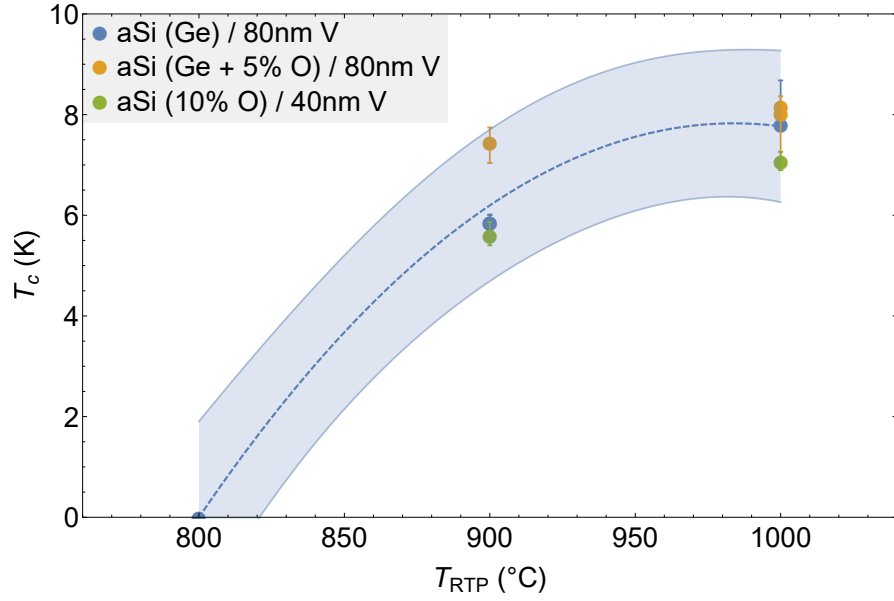


Figure 3.5: Critical temperatures measured on samples where the substrate was amorphized by germanium and/or oxygen implantation.

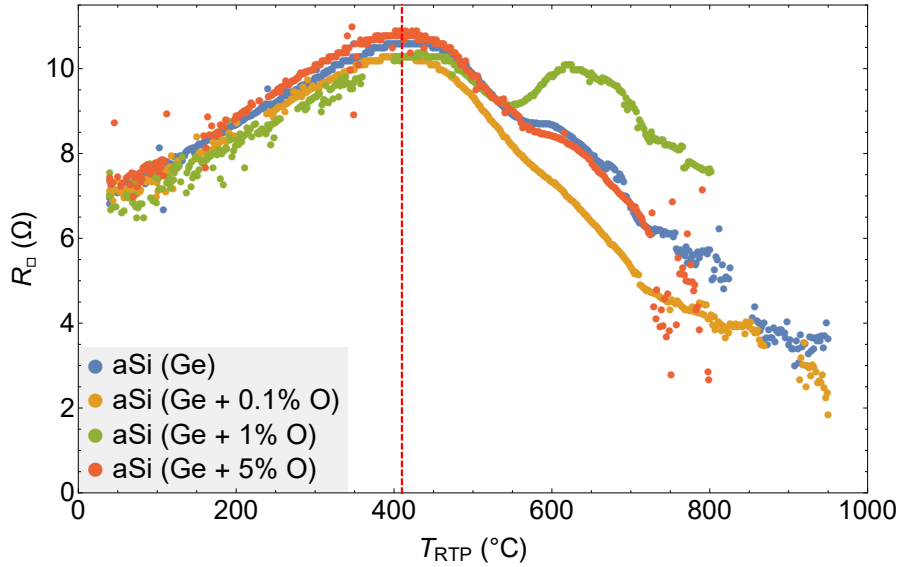


Figure 3.6: In-situ sheet resistance R_{\square} measured by four-point probe during the heating of $2 \times 2 \text{ cm}^2$ samples. No clear relationship is found between the oxygen content in the substrate, and the variation in R_{\square} .

centrations were measured in-situ during gradual heating with a ramp rate of 1 °C/s by master student Reda Alwaradi (see Fig. 3.6). It was found that while there were variations in sheet resistance development between these samples, no clear correlation with the oxygen content could be deduced. More importantly, the maximum sheet resistance² is reached already at 410 °C, far below³ the temperature at which V₃Si is expected to form by reacting with an aSi substrate [21]. Later XRD analysis showed that on amorphized samples with anywhere between 0 and 10%at. of O, each annealed at 1000 °C, both VN and TiVN₂ had consistently formed.

3.2.3 Lessons learned the hard way

After it was found that the superconductivity on samples with amorphized silicon was due to the presence of VN, other samples were annealed under ultra-high vacuum instead

²The initial rise is due to an increase in dissolved atoms (in this case, N), as both the solubility limit and the diffusion rate grow with time.

³Note that these measurements are done at a high ramp rate, which means that when the reaction starts, the diffusion time through the thin film is longer than the time scale over which the diffusivity itself grows. Since [17]

$$D(T) = D_0 e^{-\Delta E/k_B T}, \quad (3.2)$$

an increase by a factor e in ΔT would occur over a change in temperature ΔT such that $D(T + \Delta T)/D(T) = e$ (usually on the order of 10–10²K),

$$\Delta T = \frac{\Delta E k_B T}{\Delta E - k_B T} - k_B T \approx \frac{(k_B T)^2}{\Delta E}, \quad (3.3)$$

which for a ramp rate dT/dt gives a characteristic time τ ,

$$\tau \approx \frac{(k_B T)^2}{\Delta E} \left(k_B \frac{dT}{dt} \right)^{-1}. \quad (3.4)$$

We will thus have a time lag in detecting the start of the reaction when the diffusion time is longer than τ ,

$$t_{\text{diffusion}} = \frac{d_{\text{film}}^2}{D(T)} > \frac{(k_B T)^2}{\Delta E} \left(k_B \frac{dT}{dt} \right)^{-1}. \quad (3.5)$$

Typically, this leads to a few tens of degrees shift between the point at which an in-situ measurement detects a reaction [42], and the temperature at which one would see it in samples that were annealed at constant temperatures for a few minutes, making the above observation even more at odds with V₃Si formation.

Table 3.3: The results of XRD analysis on various samples. No V₃Si was formed anywhere, but a rich variety of other compounds was.

Sample	Substrate	V thickness	Cap	T_{RTP}	RTP atmosphere	XRD
A33B1	Si	40	TiN	800	N ₂	VS ₂ , TiN
A33B2	Si	40	TiN	800	UHV	VS ₂ , TiN
A33C1	Si	40	TiN	900	N ₂	VS ₂ , TiN
A33C2	Si	40	TiN	900	UHV	VS ₂ , TiN
B56I1A	SiO ₂	100	—	800	N ₂	V ₂ O ₅
B74A	SiO ₂	80	TiN	800	UHV	not V ₃ Si
B74D	SiO ₂	80	TiN	900	UHV	not V ₃ Si
C26C	aSi (Ge)	80	TiN	800	N ₂	VN, TiVN ₂
C65C	aSi (Ge + 5% O)	80	TiN	800	UHV	VN, TiVN ₂
E55C	aSi (10% O)	40	TiN	900	N ₂	VN, TiVN ₂

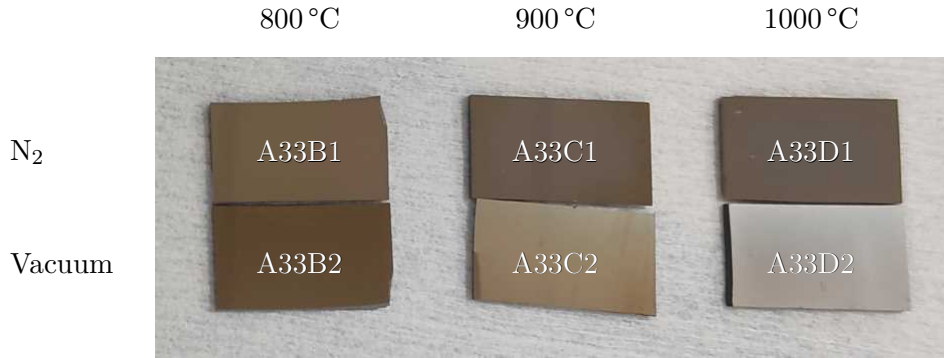


Figure 3.7: Samples with silicon substrates, 40 nm V, capped with 15 nm TiN, annealed under N_2 (top) or vacuum (bottom). Even from just looking at the color of the samples, it is clear that the annealing atmosphere matters. Note that XRD analysis did not pick up on such change, and is thus at most an indicative, rather than a conclusive method of characterization (see Table 3.3).

of under N_2 flux (see Table 3.3). No V_3Si was found on any of these samples either, suggesting that the capping layer may already have let through gases during storage, which would be consistent with the difference between the sample that was annealed in Uppsala immediately after deposition, and those that were annealed later in Grenoble. It is clear however, from the large change in sheet resistance upon switching from N_2 to UHV (see Fig. 3.4), that at least *some* N_2 diffusion through the TiN layer occurs during annealing under N_2 flux. Though this may be due to the particular composition of the TiN used in Uppsala (calibrated to around 1:1 Ti:N), and may not apply to TiN with different stoichiometries, the sheer ubiquity of this material as a capping layer in the semiconductor industry makes this fact worth exploring further. XRD analysis on other samples, where no capping layer was deposited at all, showed that V_3O_5 had formed. It is thus imperative to find a different material that can be used to cap deposited layers of vanadium, before V_3Si can be formed by metal deposition.

Though the compounds detected by XRD (see table 3.3) are all thought to be products of a reaction between vanadium and either the TiN capping layer or N_2 in the storage and annealing atmospheres, the substrate is found to be of some influence. As can be seen in Fig. 3.8, there are large differences between the samples where the vanadium was deposited on amorphous silicon, and those with 20 nm of SiO_2 .

A few clear lessons can be drawn from our experience with this project, in which disproportionately much time was spent on what later turned out to be a dead end:

1. **Identify first, *then* characterize.** Early results were consistent with the formation of V_3Si , though none actually confirmed it. Characterization of various physical properties on a wide range of samples continued for more than a year before identification methods (XRD) became available. This time could have been better spent.
2. **Verify.** It was assumed that TiN would not react with the vanadium or the atmosphere, and we expected that repeating the methods described in the literature would reliably produce V_3Si . Fabrication is a complex process, with many unknowns.
3. **Diversify.** Ultimately, we were able to recover lost ground thanks to the arrival of a 3:1 V:Si deposition target.

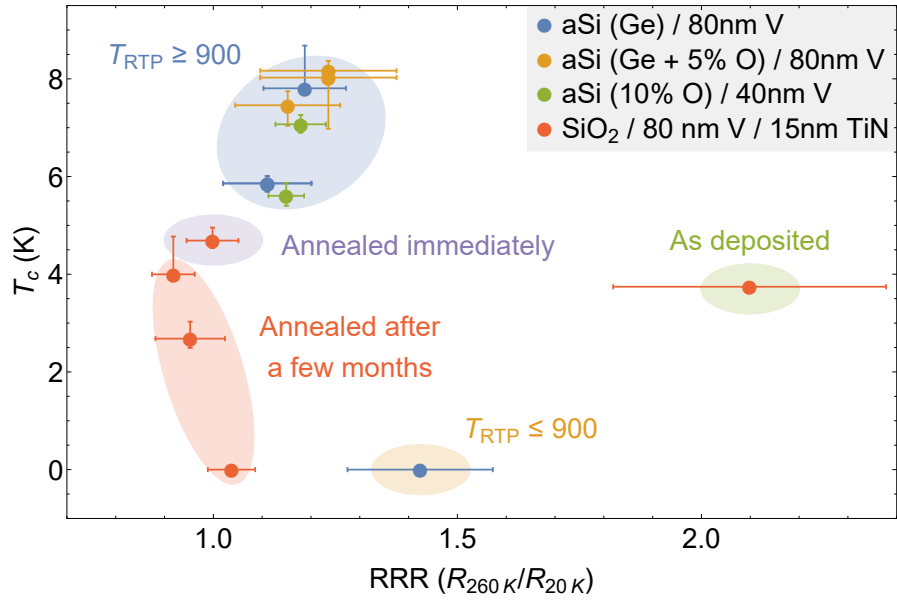


Figure 3.8: The resistance ratio (RRR) vs the critical temperature (T_c) of a selection of samples that were measured at low temperature. Based on analysis of XRD, sheet resistance and low-temperature data, five groups of similar material properties are identified.

3.3 V₃Si formation by deposition from a compound target

Though silicidation, the reaction of a deposited metal with the underlying silicon, is a common technique for silicide formation in CMOS devices (notably through the SALICIDE technology), we saw in the previous section that it did not work out for V₃Si. The main difficulty is that this phase is unlikely to form in the traditional scenario of metal deposition onto the monocrystalline exposed source/drain silicon, which instead leads to VSi₂. Besides lowering the activation energy of V₃Si, or slowing down the dominant diffusing species in VSi₂ formation, a third option is to directly prepare the deposited film with the right atomic ratio (see Fig. 3.1c). Thanks to a target provided by JX Nippon, this last option was put into practice to great success.

3.3.1 Strain, sound velocity and a drop in critical temperature

One important fact that has become central in our work, is that strain has a large effect on the physical properties of V₃Si, just as it does on other A-15 compounds. This topic was researched extensively in the 60s and 70s [43–48], mainly at Bell Labs [13–15, 49–56], and ultimately relates to a sudden weakening of the restoring force for shear deformation at low temperatures [50]. Since the potential energy stored in a spring is proportional to this restoring force (the “spring constant”), a weakening is associated with a change in the lowest-energy configuration of the atoms, and the *a priori* cubic V₃Si becomes tetragonal, with two of the axes (henceforth *a* and *b*) contracting, while the other (*c*) extends [49]. A nice corollary is that the spring constant also determines the resonant frequency (see section 2.2), and thus the rate at which displaced atoms bounce back, allowing for the study of the mechanical properties of the material by sound velocity measurements.

Without doing any justice to the profundity of Testardi’s analysis, which involves arguments about the thermodynamic, electrical and superconducting properties of A-15 compounds like V₃Si (references above for the avid reader), I will now boil down the relation between strain and a change in superconducting critical temperature to four hand-waving lines of equations. First, let us introduce the *stiffness parameters* c_{ij} , which are components of the elastic tensor that relate the stresses σ_{ij} to the strain ϵ_{ij} , which due to symmetries can be reduced to a matrix [57]⁴,

$$\sigma_{ij} = c_{ij}\epsilon_j, \quad c_{ij} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & & & \\ c_{21} & c_{22} & c_{23} & & & \\ c_{31} & c_{32} & c_{33} & & & \\ & & & c_{44} & & \\ & & & & c_{55} & \\ & & & & & c_{66} \end{pmatrix}. \quad (3.6)$$

Like an ordinary one-dimensional spring constant, they are the factor of proportionality between the force and the displacement⁵, the latter of which in the case of crystals corresponds to the strain $\epsilon = (d - d_0)/d_0$. Zooming in on a single unit of volume (to get

⁴The $i, j = 1, 2, 3$ terms represent the coupling of normal stress components, while the $i, j = 4, 5, 6$ are the shear components. The zeros imply that normal stresses lead to normal strains, and shear stresses lead to shear strains [58].

⁵To be more precise, in this three-dimensional case between the *force per unit area* or *pressure* and the displacement.

the right units for the argument),

$$\vec{F}_i = -\sum_j c_{ij}\epsilon_j, \quad V = -\int \vec{F} \cdot d\vec{r} = \frac{1}{2} \sum_{ij} c_{ij}^2 \epsilon_j \quad \Rightarrow \quad c_{ij} = \left(\frac{\partial^2 V}{\partial \epsilon_i \partial \epsilon_j} \right) \bigg|_T, \quad (3.7)$$

relating the stiffness parameters to the energy stored in the mechanical strain. In superconductors, the energy associated with condensing into the superconducting state is proportional to $E_{\text{cond}} \propto (T_c - T)^2$ [59], which means that

$$\partial_T E_{\text{cond}} \propto T_c \quad \Rightarrow \quad \frac{\partial^2 T_c}{\partial \epsilon_i \partial \epsilon_j} \propto \frac{\partial}{\partial T} \left(\frac{\partial E_{\text{cond}}}{\partial \epsilon_i \partial \epsilon_j} \right) = \frac{\partial c_{ij}}{\partial T} = \rho \frac{\partial v_{\text{sound},ij}}{\partial T}. \quad (3.8)$$

So if a reduction in T_c with strain $\vec{\epsilon}$ is indeed linked to a change in crystal stiffness, then sound velocity measurements should match the prefactors in the series expansion of $T_c(\epsilon)$,

$$T_c(\epsilon) - T_c(0) = \sum_i \Gamma_i \epsilon_i + \frac{1}{2} \sum_i \sum_j \Delta_{ij} \epsilon_i \epsilon_j + \mathcal{O}(\epsilon^3), \quad (3.9)$$

with e.g. for the parallel sound velocity $\rho \partial_T v_{\text{sound},\parallel}^2 = \Delta_{11}$. Measurements confirmed this hypothesis [14, 50, 52, 53], providing values for these prefactors of $|\Gamma| < 50$ K, $\Delta_{11} = -2.4 \times 10^5$ K, $\Delta_{12} = -5 \times 10^4$ K and $\Delta_{44} = -1 \times 10^4$ K, while all other $\Delta_{ij} = 0$ [13].

3.3.2 Substrate-induced thermal stress

One way that strain can be introduced into the silicide, is by attaching it to a material with a different thermal expansion coefficient. As the substrate-film couple is heated during crystallization annealing, several sequential stress developments occur that depend strongly on the expansion of the film relative to the substrate (more on this later), which are compounded by later cooling to cryogenic temperatures. This strain development was studied by depositing 200 nm V_3Si films from a compound target onto substrates made of sapphire and silicon with 20 nm of SiO_2 , the details of which are published elsewhere [60]. As can be seen in Fig. 3.9, the differences in accumulated strain between these two substrates leads to a relative reduction in T_c on silicon. The absence of any relationship between the residual resistance ratio and the critical temperature provides evidence that this relatively low T_c is not due to either a chemical or morphological dependence on the substrate⁶.

All of these interactions with the substrate are simplified a bit by symmetry: both the silicon and sapphire were monocrystalline and oriented such — (100) and (0001) respectively — that their expansion was isotropic in the plane. Furthermore, though V_3Si also expands when heated [54], at any fixed temperature its volume remains constant under stress [13, 51], which means that an in-plane compression is associated with a

⁶Early on in this project, it was already known from the literature that no detrimental chemical reaction should occur between V_3Si and SiO_2 , and that improvements in crystallinity and thus T_c due to lattice matching [61] are unlikely to appear on hexagonal sapphire (V_3Si is cubic, and the numbers don't add up even for weak matching once every few cells). Nonetheless, arguments by yours truly that strain should be responsible for the differences in T_c between the silicon and sapphire substrates took a while to be taken seriously, and models that later turned out to align with those of Testardi were initially disregarded in favor of “interface effect” interpretations. Given the overwhelming evidence that strain *is* the most relevant variable in this system [13–15, 43–56, 60], this now-irrelevant concern will not be further discussed here.

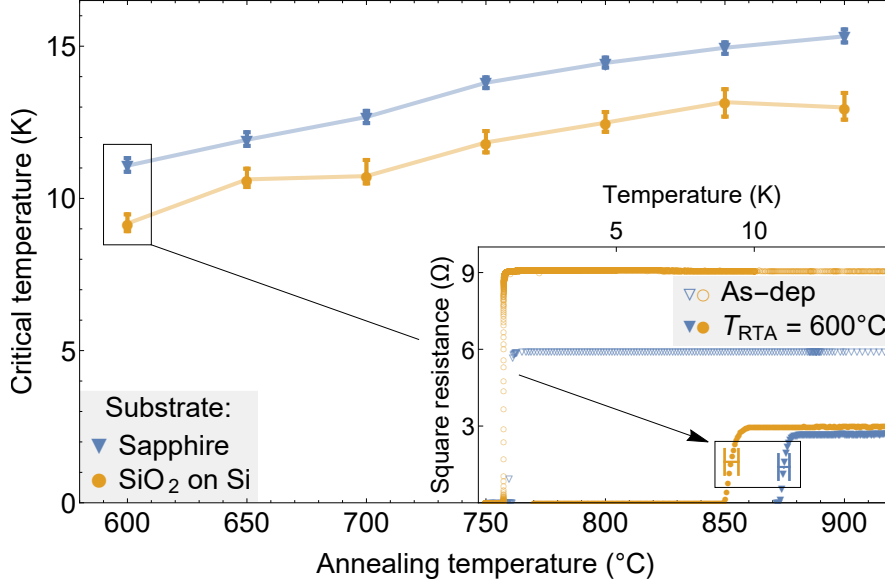


Figure 3.9: The critical temperature (T_c) of samples with a 200 nm thin film of V_3Si is plotted versus the temperature at which each sample was annealed, with error bars indicating the temperature at which t10 and 90% of the normal-state resistance was lost. **Inset:** $R(T)$ measurements on the first two samples annealed at 600°C, as well as on an unannealed sample for each substrate. The horizontal error bars in the inset correspond to the vertical bars in the main graph.

Table 3.4: The estimated contractions of sapphire [62], silicon [63] and V_3Si [48, 54] from 300 to 16 K (valid both in-plane and out-of-plane), and the induced out-of-plane expansion of V_3Si .

Material	Contraction	Relative to V_3Si	Induced $\Delta\epsilon_1$
Sapphire	$-6.2(8) \times 10^{-4}$	$+5.8(13) \times 10^{-4}$	$-1.2(3) \times 10^{-3}$
Silicon	-2.33×10^{-4}	$+9.7(10) \times 10^{-4}$	$-1.9(2) \times 10^{-3}$
V_3Si	$-1.2(1) \times 10^{-3}$	—	—

predictable out-of-plane expansion. Given some in-plane strains $\epsilon_1 = \epsilon_2$, we can thus directly compute

$$\epsilon_3 = \frac{-2\epsilon_1 - \epsilon_1^2}{(1 + \epsilon_1)^2} \stackrel{\epsilon_1 \ll 1}{\approx} -2\epsilon_1. \quad (3.10)$$

This then gives a strain vector of

$$\vec{\epsilon} = \left(\epsilon_1, -\frac{\epsilon_1}{2}, -\frac{\epsilon_1}{2}, 0, 0, 0 \right), \quad \text{where} \quad \epsilon_1 = \frac{2}{3} \left(\frac{c}{a} - 1 \right), \quad (3.11)$$

with $a = b$ and c the in-plane and out-of-plane lattice parameters, respectively, as mentioned in section 3.3.1. We can now simplify the general expression that we found in eq. (3.9), to

$$T_c(\vec{\epsilon}) - T_c(0) = \frac{3}{4} \epsilon_1^2 (\Delta_{11} - \Delta_{12}), \quad (3.12)$$

such that measurement of the strain along a single direction is enough to determine the expected reduction in T_c .

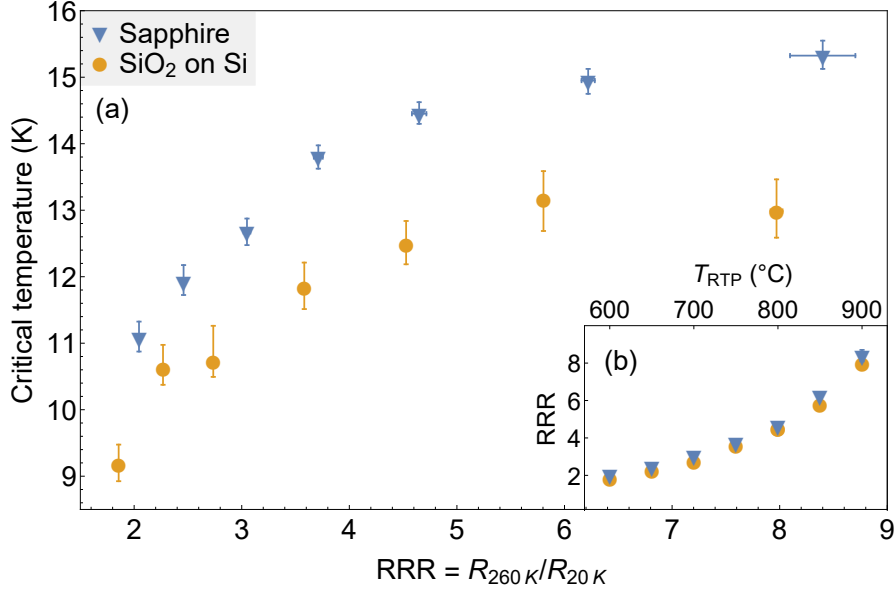


Figure 3.10: (a) The residual resistance ratio $RRR = R_{260K}/R_{20K}$ is a measure of material quality, where any resistance that is *residual* after the thermal electron-phonon scatterings are suppressed, is assumed to be due to impurities and defects. For any given RRR, there is a large difference in T_c between the two substrates, indicating that it is affected by something other than the quality of the V₃Si. (b) A near-perfect match between the RRRs measured on sapphire and silicon shows that the substrate has little influence on the quality of the V₃Si film.

It was found (see Fig. 3.9) that the superconducting transition occurs at temperatures 1.9(3) K lower on a silicon substrate than it does on sapphire. In the most naive interpretation (useful to get a sense of the orders of magnitude), where only one of the two substrates imposes a strain, this would correspond to an out-of-plane deformation of $3.7(3) \times 10^{-3}$, or about 0.4%. The reality is more complex however, and we will need to take into account both the strain built up during the thermal processing, and that induced by the relative contraction of the silicide during subsequent cooling to cryogenic temperatures. The last cooling step from 300 K to 16 K is in fact not that important, as there is little difference in this temperature range between the thermal strains induced by the sapphire and silicon substrates (see Table 3.4). Instead, more attention should be paid to the intricate pattern of stress developments during thermal processing, which was studied in detail with two separate experiments.

The first of these was the slow, step-wise heating of two *a priori* unannealed samples (one for each substrate) under secondary vacuum up to 1000 °C while in-situ XRD scans were performed (see Figs. 3.11, 3.12 and 3.13). On each $\theta/2\theta$ scan, Lorentzian distributions were fitted using Mathematica around the expected 2θ values of the (200), (210) and (211) peaks (drawn superposed in Fig. 3.11), giving fit parameters of both the position and width of each peak (shown in Fig. 3.12). The slow shifts in the position of the three peaks were then weighted by peak intensity (number of counts), and used to estimate the crystallite expansion Δd out of plane (Fig. 3.13).

No significant difference in grain growth was observed between the two substrates (estimated from crystallite size, see Fig. 3.12), consistent with the earlier conclusion from RRR measurements that there is no dependence of the material quality on the substrate.

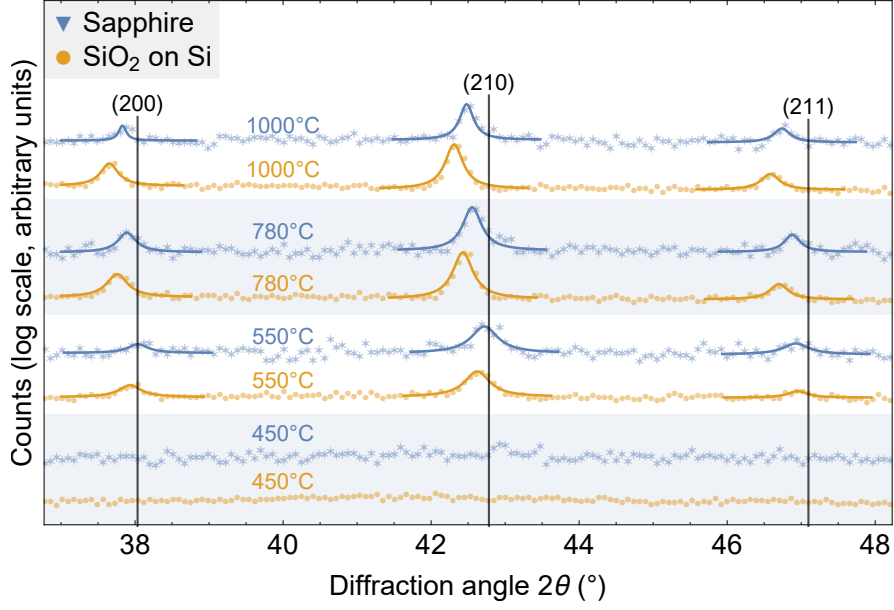


Figure 3.11: A selection of XRD $\theta/2\theta$ curves, where each peak corresponds to a reflecting plane of the cubic V_3Si crystal (hkl indices for expected peak positions at 300 K indicated above). No peaks are observed at all below 500 °C, indicating that the silicide is amorphous, while they become higher and sharper as the temperature is increased. A leftward shift towards smaller angles is due to the thermal expansion of the crystal.

However, as shown in Fig. 3.13, there is a large difference in the out-of-plane strain that was developed. Most importantly, the strain difference inverted after the samples were cooled down, leading to an out-of-plane strain of only $\epsilon_1 = +0.2(4) \times 10^{-3}$ on sapphire, and $\epsilon_1 = -3.3(1) \times 10^{-3}$ on silicon. Note that this implies that a stress relaxation occurred on sapphire at some point during the cooling from 1000 °C to room temperature, since the higher thermal expansion coefficient of V_3Si ($\alpha_{V_3Si} \approx 7.5 \times 10^{-6} \text{ K}^{-1}$ [54], while $\alpha_{sapphire \perp (0001)} \approx 5.0 \times 10^{-6} \text{ K}^{-1}$ [66]) would otherwise have caused the out-of-plane lattice parameter to shrink *faster* than free-standing V_3Si , giving its trajectory a larger slope than the indicated gray lines in Fig. 3.13.

When extrapolating these strains down to cryogenic temperatures using the values in Table 3.4, the following estimates are obtained:

$$\begin{aligned} \epsilon_1^{\text{sapphire}} &= -0.9(5) \times 10^{-3}, \\ \epsilon_1^{\text{silicon}} &= -5.2(2) \times 10^{-3}. \end{aligned} \tag{3.13}$$

If the strain would be distributed perfectly homogeneously across crystallite orientations (which it isn't, see Fig. 3.14b), and the films behaved entirely thermoelastically on either substrate (which they don't [13, 49, 51]), then eq. (3.12) tells us that such strains would cause reductions in T_c of between 0 and 0.3 K on sapphire, and between 3.5 and 4.2 K on silicon. We would then expect a relative reduction in T_c on silicon of anywhere between 3.2 and 4.2 K. Instead, as we saw in Fig. 3.9, the critical temperature is reduced on silicon only by about 1.9 K, a factor two less. Vice versa, the measured reduction in critical temperature should be associated with a difference in strain of *at most*⁷ $3.7(3) \times 10^{-3}$.

⁷There is a quadratic relation between T_c and ϵ , so the required $|\epsilon_1^{\text{sapphire}} - \epsilon_1^{\text{silicon}}|$ is smaller when $\min(|\epsilon_1^{\text{sapphire}}|, |\epsilon_1^{\text{silicon}}|)$ is larger.

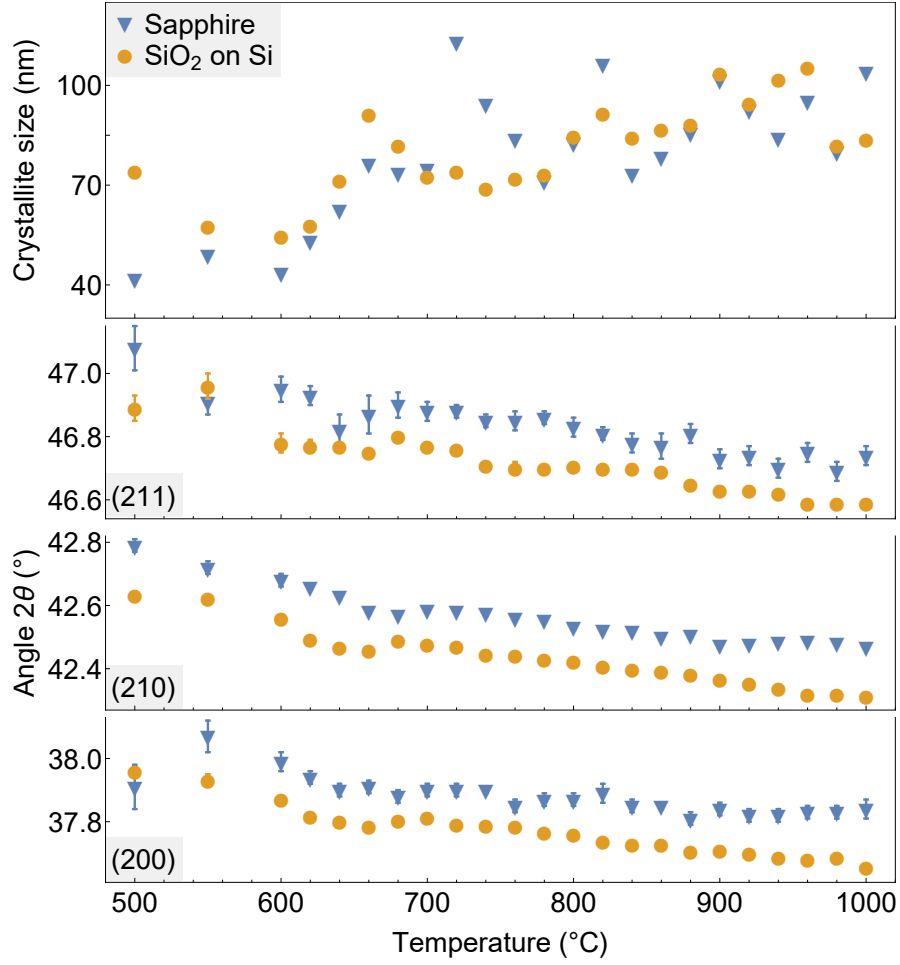


Figure 3.12: **(top)** A rough estimate of the crystallite size can be obtained from the peak width using the Scherrer equation [64, 65]. The measurement on the sapphire sample was performed with a slightly smaller slit size, leading to less instrumental line broadening and thus a larger apparent grain size at high temperatures. **(bottom panels)** Plotted is the shift in the peak position with temperature (see Fig. 3.11), from which the out-of-plane lattice parameter can then be extracted (see Fig. 3.13).

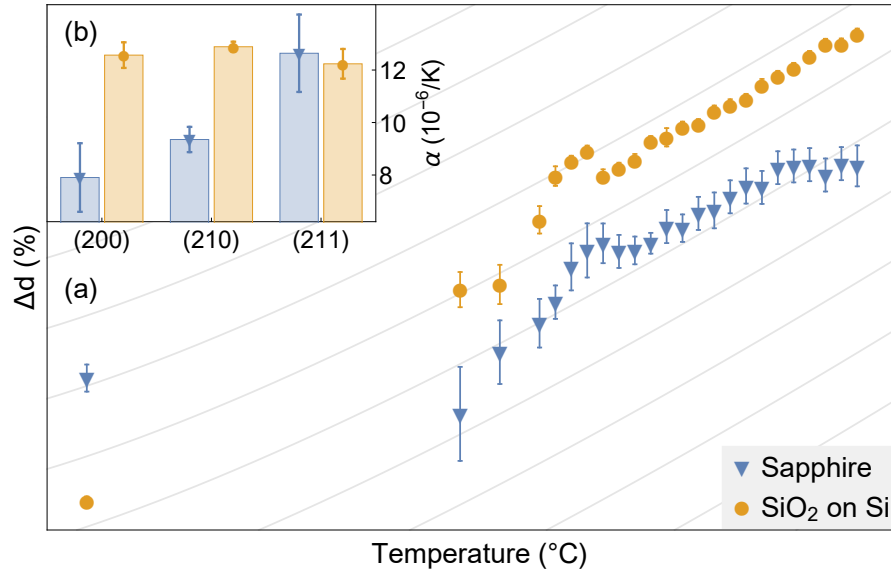


Figure 3.13: (a) Using data from Fig. 3.12, the change in out-of-plane lattice parameter during heating is plotted versus the temperature. Another data point was taken after the samples had cooled down to 30 °C. Gray inclined lines in the background show the thermal expansion that would be observed on free-standing crystalline bulk V₃Si [54], with different offsets. A clear difference can be observed between the two substrates: the silicide starts out more strained on sapphire right after crystallization at 500 °C, while it ends up more strained on silicon at the end of the heating cycle. (b) Effective out-of-plane thermal expansion coefficients are extracted for each XRD peak from linear regressions between 700 and 1000 °C on the peak positions in Fig. 3.12. While the expansion rate is independent of crystallite orientation on silicon, tilted planes expand faster on sapphire.

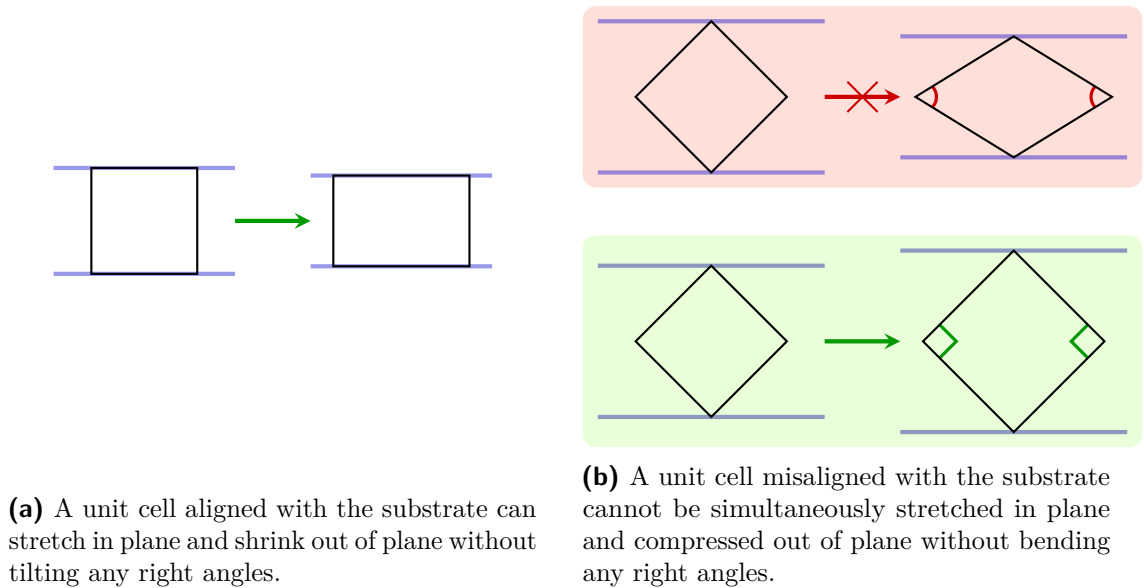


Figure 3.14: The more the unit cell is tilted, the less able it is to conserve the unit cell volume. From this it follows that (200) planes should expand less out of plane than (210) and (211), which explains the behavior seen in Fig. 3.13b.

This is where V_3Si becomes truly fascinating: it turns out that the weakening of the restoring force for shear deformation that we discussed earlier also causes the crystal to undergo a Martensitic transformation⁸. It has been observed on free-standing pieces of V_3Si that the *a priori* cubic crystal undergoes a tetragonal deformation at temperatures of around 20–21 K, just above the superconducting transition [13, 49, 51]. This ever-so-slight change is associated with a “strain”⁹ of up to $\epsilon_1 = -2\epsilon_{2,3} = +1.7 \times 10^{-3}$ (expansion along one axis, contraction along the other two), with higher strain values on samples with higher RRR [51].

This Martensitic transformation could also occur in polycrystalline thin films, in which case its effect on the critical temperature of a grain would depend on the orientation of the crystal relative to the substrate. As shown in eq. (3.13), the V_3Si is expected to be compressed out of plane due to thermal strain. Grains that have their crystal structure aligned with the substrate would thus (before the Martensitic transition) be subject to tensile strain along two axes, and compressive strain along the other. A tetragonal deformation with expansion along the out-of-plane axis could compensate this thermal strain, lessening the reduction in T_c . If this deformation would instead provide an expansion along one of the in-plane directions, the critical temperature would be further reduced. Since the Martensitic transition is thought to occur precisely because the crystal is more stable under a certain amount of strain, it is likely that the transformation would predominantly stretch crystals out of plane where they were strongly compressed before. On the other hand, it is possible that the transformation would instead align itself to *increase* the total strain in grains that are strained *less* than the stable equilibrium. This provides a mechanism for the strain on the two substrates to converge towards a common value, thus explaining the reduction in T_c on silicon of only 1.9 K relative to the sapphire substrate. Moreover, since compensation of in-plane tensile stress by tetragonal deformation is only possible in grains that are aligned with the substrate, such a mechanism should cause the shift in T_c to be smaller in tilted grains, and thus result in an overall broadening of the superconducting transition. As can be seen in Fig. 3.9, broader transitions are indeed observed on silicon, on which the strain is expected to be larger than on sapphire.

Plans to measure the strain in-situ by cooling V_3Si samples in a cryostat to liquid helium temperatures (4 K) were abandoned once it transpired that the cryo-enabled XRD setup had broken down while moving the laboratory to a new building.

3.3.3 Strain development during thermal processing

To gain a better understanding of the stress development during thermal processing, a second set of in-situ XRD measurements was performed, using again samples where 200 nm of V_3Si had been deposited onto oxidized Si wafers with 20 nm of thermal SiO_2 , and sapphire. This time, a different setup was used where peak positions could be tracked more accurately both in plane and out of plane, though no good vacuum could

⁸“Martensitic” refers to a change in local crystal structure where the individual atoms move around by less than the lattice parameter (it is diffusionless). To test the reversibility of such a transition (if any occurred in our films), a sample with 200 nm of V_3Si on silicon with thermal oxide was annealed at 900 °C under N_2 flow for two minutes, and then cycled six times in a cryostat between 11 and 30 K, during which no change in T_c was observed.

⁹Sticking to our earlier definition for the benefit of consistency, $\epsilon = (1 - d_0)/d_0$ along any direction, with d_0 the spacing expected for *cubic* V_3Si . It may be more technically correct to use the post-transition tetragonal structure as zero-strain reference point, but for our purposes that would be as correct as it would be useless.

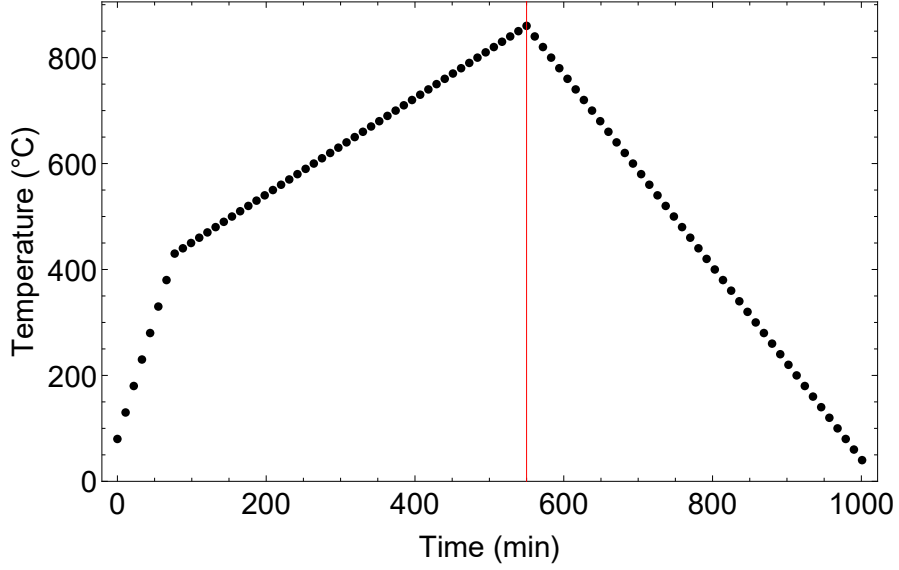


Figure 3.15: Temperature profile versus annealing time during the second set of in-situ XRD experiments. The red line in this and the following graphs in this section indicates the maximum temperatures reached, after which cooling starts.

be obtained (the previous experiment was performed under secondary vacuum). To avoid oxidation, a purified nitrogen atmosphere was therefore used, at a pressure slightly above ambient, i.e. 1.25 bar. Fearing nitridation, as well as reactions with residual traces of other gases, it was decided to stop the temperature ramp at 860 °C¹⁰, after which the samples were slowly cooled down (see Fig. 3.15).

The XRD measurements were optimized for the detection of only the position of the (210) peak, which is the largest in the V₃Si spectrum (see Fig. 3.11). Distinct behaviors were observed during the out-of-plane (see Fig. 3.16) and in-plane (see Fig. 3.17) measurements, showing a more non-linear development of the peak position out of plane. Combining the data from the in-plane (IP) and out-of-plane (OOP) measurements, it is then possible to extract the stress using the $\sin^2 \psi$ methodology [67] (assuming stress isotropy in the plane),

$$\epsilon_\psi = \frac{(d_\psi - d_0)}{d_0} = \frac{1}{2}S_2(hkl)\sigma \sin^2 \psi + 2S_1(hkl)\sigma. \quad (3.14)$$

Here the stress-free (210)-plane spacing d_0 is first calculated from a weighted mean of the detected OOP and IP spacings d_\perp and d_\parallel ,

$$d_0 = \frac{d_\perp - Ad_\parallel}{1 - A}, \quad \text{where} \quad A = \frac{4S_1(hkl)}{S_2(hkl) + 4S_1(hkl)}, \quad (3.15)$$

with $A \approx -1$ a negative number that gives the relative weights of d_\perp and d_\parallel , and which depends on the (hkl) index-dependent elastic constants S_1 and S_2 of V₃Si [68, 69]. It

¹⁰A trial run up to 1000 °C with a silicon substrate sample caused a visible degradation of the surface. A similar reaction may have occurred during an in-situ sheet resistance measurement (see Fig. 3.6), where the furnace was opened and air was let in before the cooling was completed, which caused the surface of the sample to turn into a fine dust within seconds.

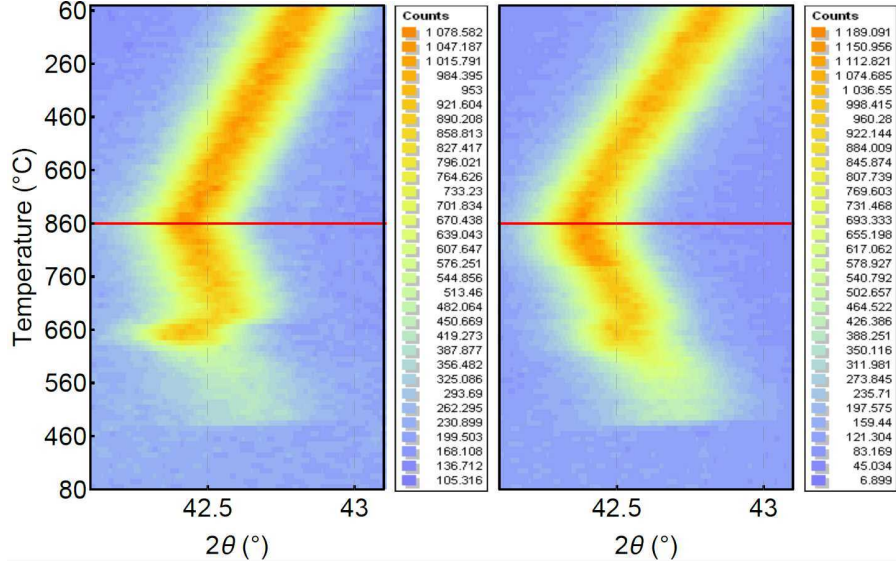


Figure 3.16: Contour maps of the *out-of-plane* (210) peak observed on samples with 200 nm of V_3Si deposited onto substrates of (left) sapphire and (right) silicon.

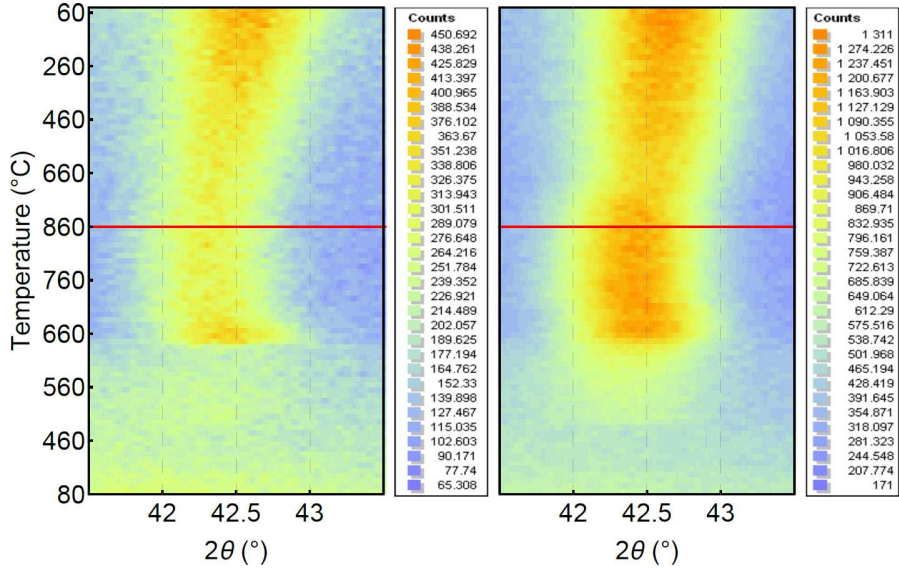


Figure 3.17: Contour maps of the *in-plane* (210) peak observed on the same samples as in Fig. 3.16.

was calculated that for the (210) plane, these constants are

$$\begin{aligned} S_1(210) &= -1.66 \times 10^{-6} \text{ MPa}^{-1}, \quad \text{and} \\ S_2(210) &= +1.338 \times 10^{-7} \text{ MPa}^{-1}, \end{aligned} \tag{3.16}$$

respectively, which gives $A = -0.985$ and thus $d_0 \approx (d_{\perp} + d_{\parallel})/2$. This then allows us to calculate the in-plane stress σ from eq. (3.14),

$$\sigma = \frac{2}{S_2(hkl)} \left(\frac{d_{\parallel} - d_{\perp}}{d_0} \right), \tag{3.17}$$

which is plotted in Fig. 3.18a.

To illustrate the various processes that occur during the heating, a simplified plot of the stress development on the sapphire substrate is shown in Fig. 3.19. Initially, as the *a priori* amorphous V_3Si crystallizes around 500°C , it becomes more compact while the substrate maintains its volume, leading to an in-plane tensile stress. During step ①, this stress relaxes due to a combination of thermally activated plastic deformation and a mismatch in thermal expansion coefficients, which around room temperature are $7.5 \times 10^{-6} \text{ K}^{-1}$ for V_3Si [54] and $5.0 \times 10^{-6} \text{ K}^{-1}$ in-plane for (0001) sapphire [66]. Once the grains become too large to allow for relaxation by plastic deformation around 650°C , a second increase in tensile stress occurs during step ② due to grain growth. This volume reduction due to grain growth slows down at around 700°C as the film nears its maximum packing density, giving way to ③ a second compression due to thermal strain. Neither plastic deformation nor grain growth occur after the maximum temperature of 1000°C has been reached (see Fig. 3.18c,d), and ④ during cooling the in-plane stress then develops proportional to the thermoelastic strain imposed by the substrate (see Fig. 3.18b).

Note that not all of these steps (①,②,③) will have occurred on samples annealed by rapid thermal processing (RTP) at lower temperatures, such as those reported on in Figs. 3.9 and 3.10. Specifically, the fact that the dip in in-plane stress around 650°C (see both Figs. 3.13 and 3.18) is deeper on sapphire than it is on silicon, could explain why there is less of a difference in T_c between the sapphire and Si samples annealed at this temperature.

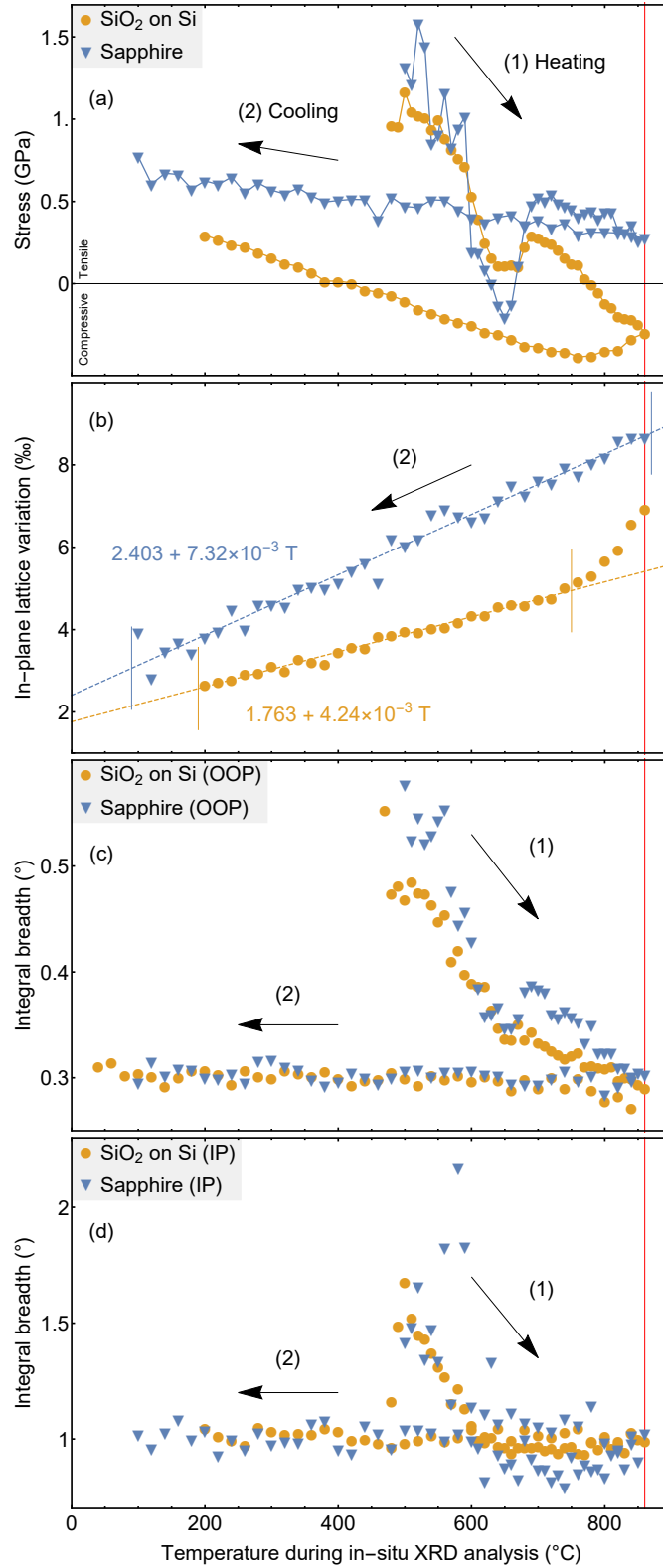


Figure 3.18: (a) Stress evolution versus annealing temperature for both the sapphire and the silicon substrates. (b) The strain of the V₃Si film versus the annealing temperature, shown only for the part of the temperature ramp where the samples are cooling. (c,d) OOP and IP integral breadth of the V₃Si (210) peak versus temperature.

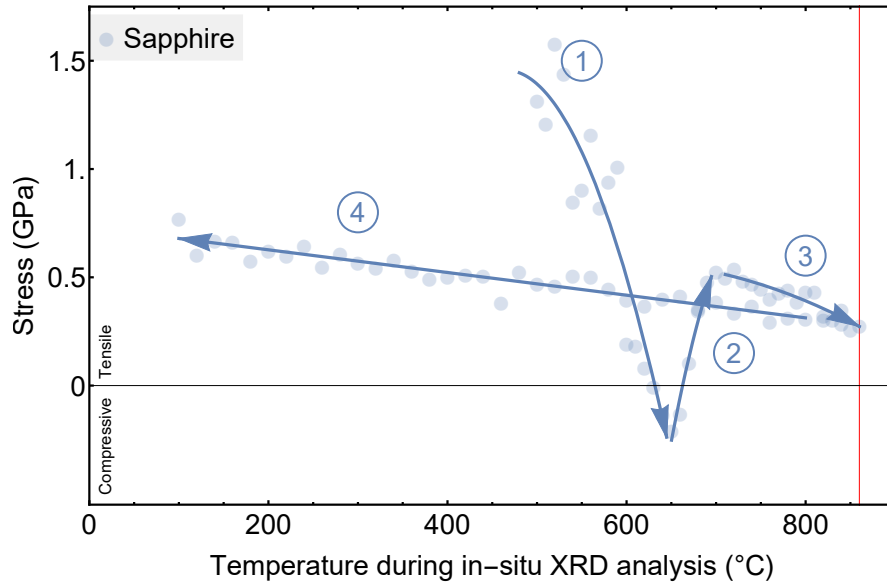


Figure 3.19: Four stages can be identified in the IP stress development on the sapphire substrate.

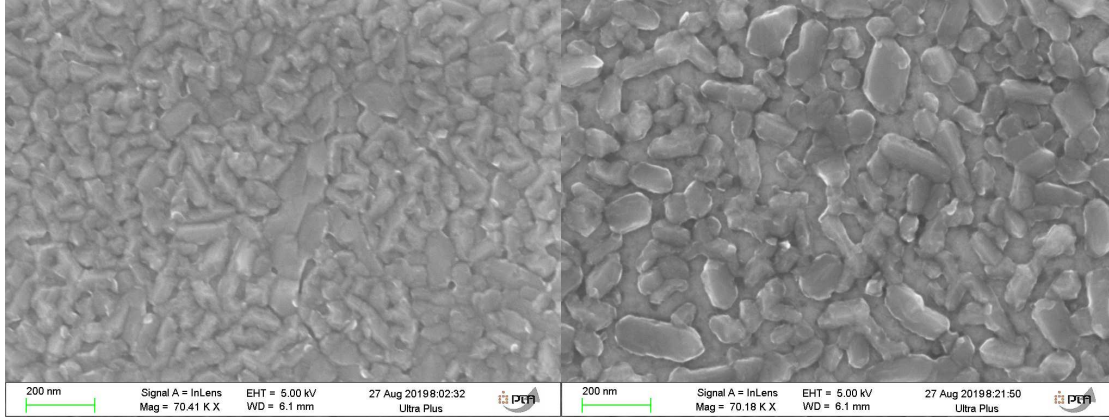
3.3.4 Stability of V_3Si thin films during thermal processing

Boundaries between grains, like the silicide-vacuum and silicide-substrate interfaces at the top and bottom of the film, have a positive surface energy associated with them. When enough energy is provided in the form of heat to activate the diffusion of atoms, the system will generally move towards a state in which the total surface energy is minimized, leading to a competition in area reduction between the different interfaces.

Shown in Fig. 3.20 are planar view scanning electron microscope (SEM) images of samples with a 17 nm V_3Si film on SiO_2 annealed under vacuum at different temperatures. As can be seen from the gradual island formation with increasing temperature, the surface energy density of the SiO_2 /vacuum interface is smaller than the sum of the energy densities of the SiO_2/V_3Si and V_3Si /vacuum interfaces [70, 71], causing it to de-wet. We will see more of this in the context of $PtSi$ island formation in section 3.4.

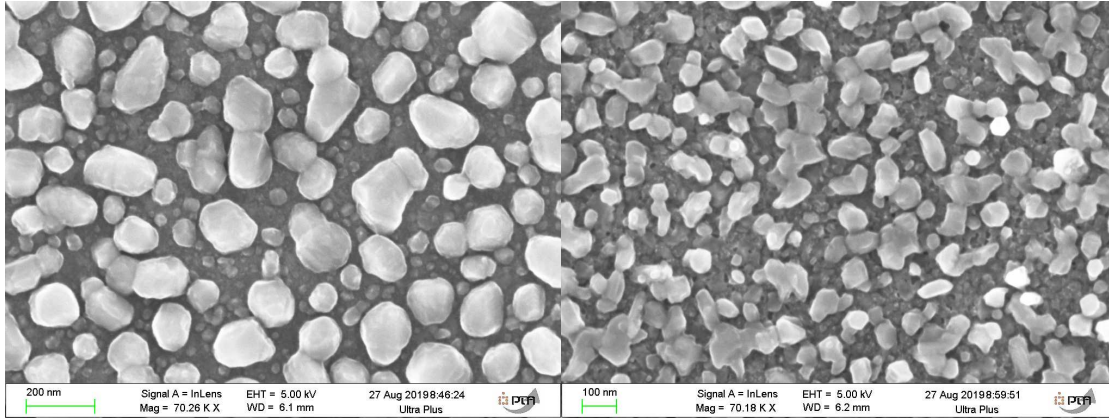
The temperature up to which a film of V_3Si remains stable depends on the thickness, since higher diffusion rates are required to overcome the energy barrier of increasing the film/vacuum surface area before the substrate can be exposed [71]. Fig. 3.21 shows a collection of SEM cross-sections of samples with various thicknesses of V_3Si on SiO_2 substrates that were annealed under vacuum, with in the top-left the normalized sheet resistances versus annealing temperature. A divergence in the sheet resistance indicates the formation of gaps or even islands in the V_3Si film, which occurs at higher temperatures on thicker films. Since the quality of the film improves with annealing temperature (see Figs. 3.9 and 3.10), the film thickness thus determines the thermal budget and limits the maximum attainable critical temperature. It is important to note that in all these samples the V_3Si was deposited onto an oxide, rather than HF-cleaned silicon, which likely gives a different surface energy density than a V_3Si/Si interface would.

A second type of thermodynamic instability occurs when V_3Si is deposited onto a silicon substrate. Whereas V_3Si does not react with SiO_2 due to the strong binding between silicon and oxygen, VSi_2 is likely to form at a V_3Si/Si interface (see Fig. 3.22). Though this new phase will likely nucleate at higher temperatures even at a perfectly



(a) 600 °C, $R_{\square} = 150(5) \Omega$.

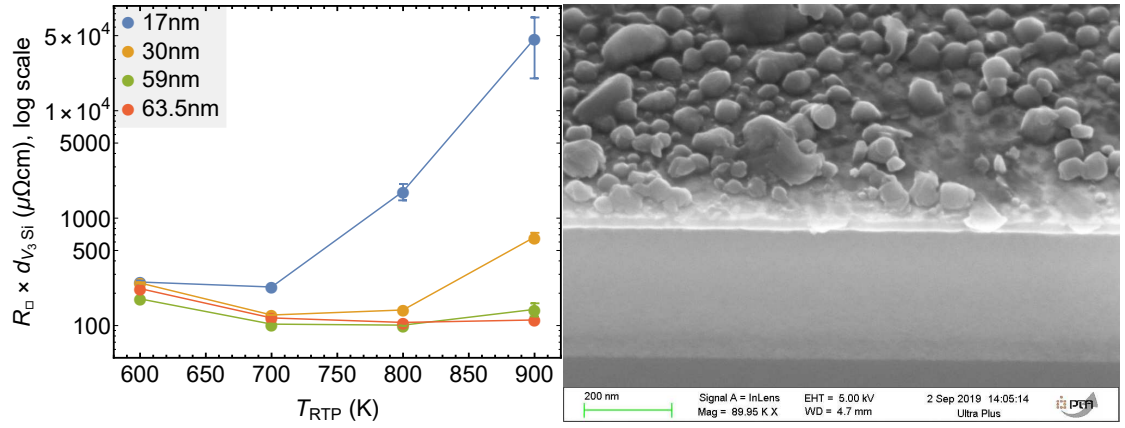
(b) 700 °C, $R_{\square} = 135(3) \Omega$.



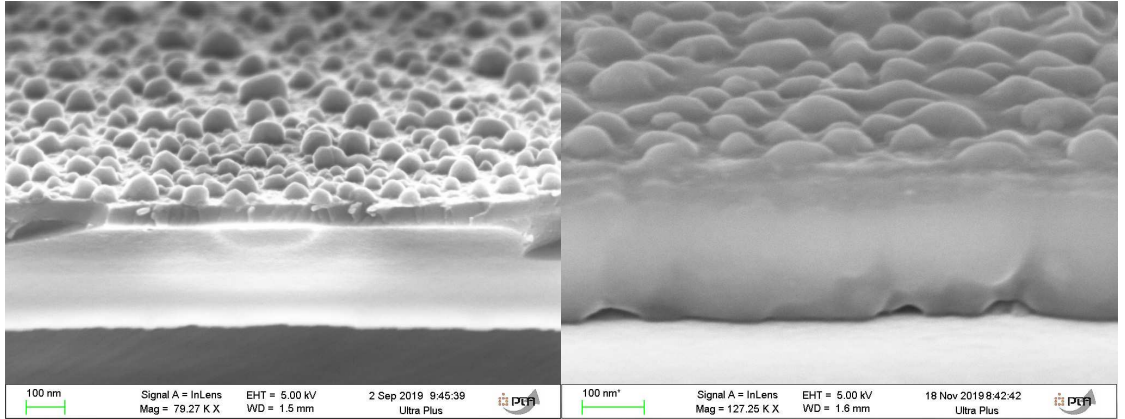
(c) 800 °C, $R_{\square} = 1.04(18) \times 10^3 \Omega$.

(d) 900 °C, $R_{\square} = 2.8(16) \times 10^3 \Omega$.

Figure 3.20: Scanning electron microscope images of thermally processed samples with a 17 nm layer of V_3Si deposited onto 300 nm of SiO_2 . The samples were annealed at the four indicated temperatures under vacuum during two minutes, after which the sheet resistance was measured. The deposition conditions were different from those used for the rest of the samples discussed in this chapter: the layers were RF sputtered (as opposed to DC) at a pressure of 100 μ bar. The sputtering power (200 W) and argon flow (50 sccm) were the same as those used elsewhere.



(a) Normalized sheet resistance vs T_{RTP} for selected thicknesses. (b) 30 nm V_3Si on 300 nm SiO_2 , $T_{RTP} = 800^\circ\text{C}$.



(c) 63.5 nm V_3Si on 300 nm SiO_2 , $T_{RTP} = 800^\circ\text{C}$. (d) 200 nm V_3Si on 300 nm SiO_2 , $T_{RTP} = 800^\circ\text{C}$.

Figure 3.21: (a,b) A combination of grain growth and de-wetting causes films to become discontinuous after annealing, leading to a divergence in sheet resistance (normalized by deposited V_3Si thickness). Thicker films remain continuous up to higher temperatures. (c,d) On thicker films of 63.5 and 200 nm, only a thin surface layer with protrusions appears after annealing at 800°C .

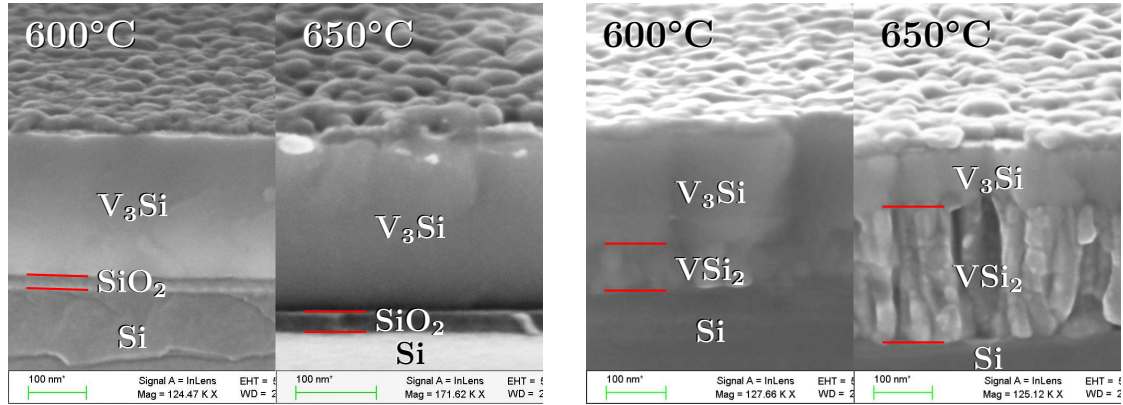
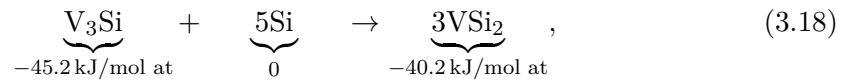


Figure 3.22: SEM images of cleaved samples with 200 nm of V_3Si deposited onto either oxidized or HF-cleaned silicon, and annealed at 600 or 650 °C. **(Left)** A silicon substrate with 20 nm of thermal oxide. No formation of VSi_2 observed. **(Right)** When an equal thickness of V_3Si is deposited on oxide-free silicon and annealed at 600 °C and 650 °C, a VSi_2 layer with distinct columnar morphology appears. It was confirmed by XRD analysis that V_3Si is present after deposition, and that VSi_2 appears after annealing.

sharp interface between crystalline V_3Si and Si, its formation is greatly aided by the conditions under which the V_3Si is sputtered onto the film.

As the V and Si atoms arrive on the Si surface, their kinetic and condensation energies together with the heating by the argon plasma cause an intermixing layer to form that typically depends in thickness on the total time of deposition. In this mixed zone there will be a smooth distribution in V and Si atomic concentrations, from 25% Si within the deposited layer, to 100% Si in the substrate. If V_3Si were the energetically most favorable state with the highest energy gain per bound V atom, then there would be a lower chemical potential for these atoms closer to the V_3Si layer, providing a mechanism for asymmetric diffusion. Alas, although V_3Si ($V_{0.75}Si_{0.25}$) has a larger effective heat of formation (EHF) from pure V and Si per mole of atoms involved *in total* [19], the energy gain *per vanadium atom* is greater for VSi_2 ($V_{0.33}Si_{0.67}$), of which three times as many molecules can be formed for a fixed amount of vanadium. Therefore, unless the intermixing layer is so thin as to make the appearance of a new phase energetically unfavorable due to the disproportionately large surface energy, it is to be expected that VSi_2 will form in such an intermixing layer.

Once VSi_2 has formed, it will begin to compete with the formation of V_3Si . Given that a reservoir of Si is present in the substrate, while no more V exist than what is deposited in the amorphous layer of $V_{0.75}Si_{0.25}$, the relevant energy is the heat of formation per atom of vanadium, which is minimized in the following reaction:



where an energy of 181 kJ/mol is gained per mole of V_3Si that is transformed. Since we are aiming to fabricate a JoFET with V_3Si source and drain contacts (see Fig. 3.23), it is imperative that if this reaction cannot be prevented, it is at least controlled.

To find a process window within which V_3Si can be crystallized to improve its electrical and superconducting properties, while also limiting the formation of VSi_2 , a set of samples was prepared where different thicknesses of V_3Si (20, 50, 100 and 200 nm) were deposited

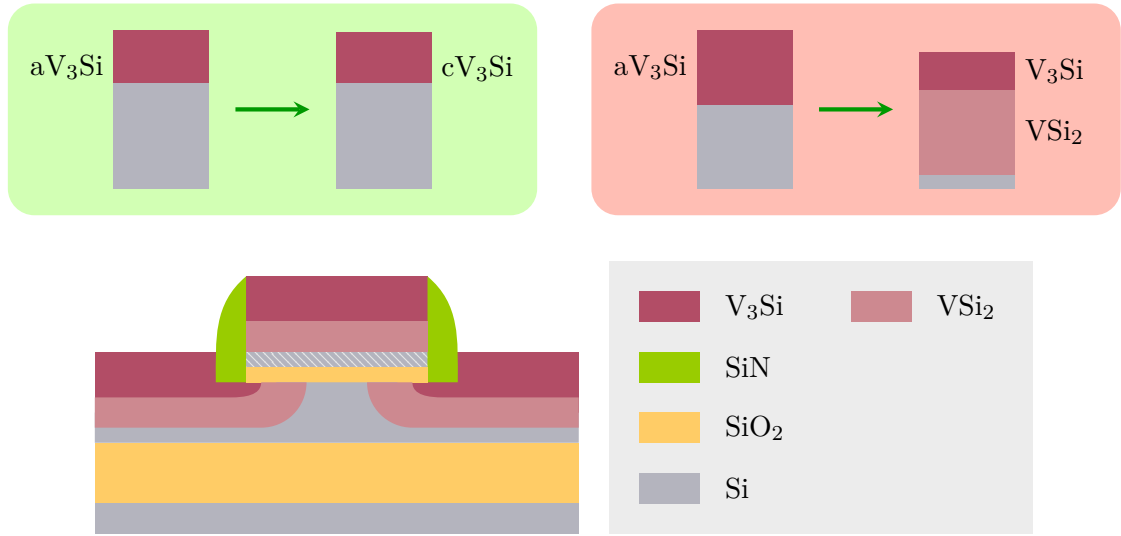


Figure 3.23: The annealing trade-off: while the V_3Si is crystallized and improves in superconducting properties, a VSi_2 layer is formed below until no V_3Si is left and superconductivity vanishes.

onto HF-cleaned substrates. Though removal of the native oxide is advisable in general to obtain good ohmic contacts, it may also aid VSi_2 formation (which is slowed down by oxygen [21]). This is only a minor concern however: superconductivity disappears between 650 and 700 °C on samples with 200 nm V_3Si on Si, whether it was cleaned with HF or not.

For each thickness of V_3Si , six samples were annealed at temperatures between 500 and 750 °C, after which their sheet resistance, residual resistance ratio (RRR) and critical temperature were determined (see Fig. 3.24). Samples with thinner layers (≤ 100 nm) annealed at 500 °C showed a resistivity of 180(10) $\mu\Omega$ cm, similar to that of as-deposited amorphous V_3Si , while the sheet resistance of the 200 nm layer remained out of range for our instruments up to annealing temperatures of 600 °C. Indicated in Fig. 3.24a is a red dashed line at 77.1 $\mu\Omega$ cm, which corresponds to the resistivity measured on a sample with 200 nm of V_3Si on a sapphire substrate, annealed at 900 °C under vacuum. Since this is the lowest resistivity that we have measured for V_3Si on any sample where we are sure that no chemical reaction had occurred with the substrate (XRD confirmed that V_3Si was the only V-rich compound), any value below this dashed line is taken to indicate the presence of VSi_2 .

As shown in Fig. 3.24b, at high annealing temperatures the RRR is higher for thicker layers, which can be explained by thickness-limited grain growth and is consistent with the lower quality of thinner films reported elsewhere [5]. An interesting inversion of the ordering of the RRR values for the three first thicknesses occurs at 550 °C, where the 20 nm film shows the highest RRR. This could be explained by the early homogenization of the forming VSi_2 film, which would imply a Si diffusivity of around $1.8 \text{ nm}^2 \text{ s}^{-1}$ through the VSi_2 layer [18], a value that has been reported elsewhere only for annealing temperatures of 650 °C [16].

The crystallization of V_3Si is observed on thicker films of at least 100 nm by critical temperature measurements (Fig. 3.24c). It is important to note that the critical tempera-

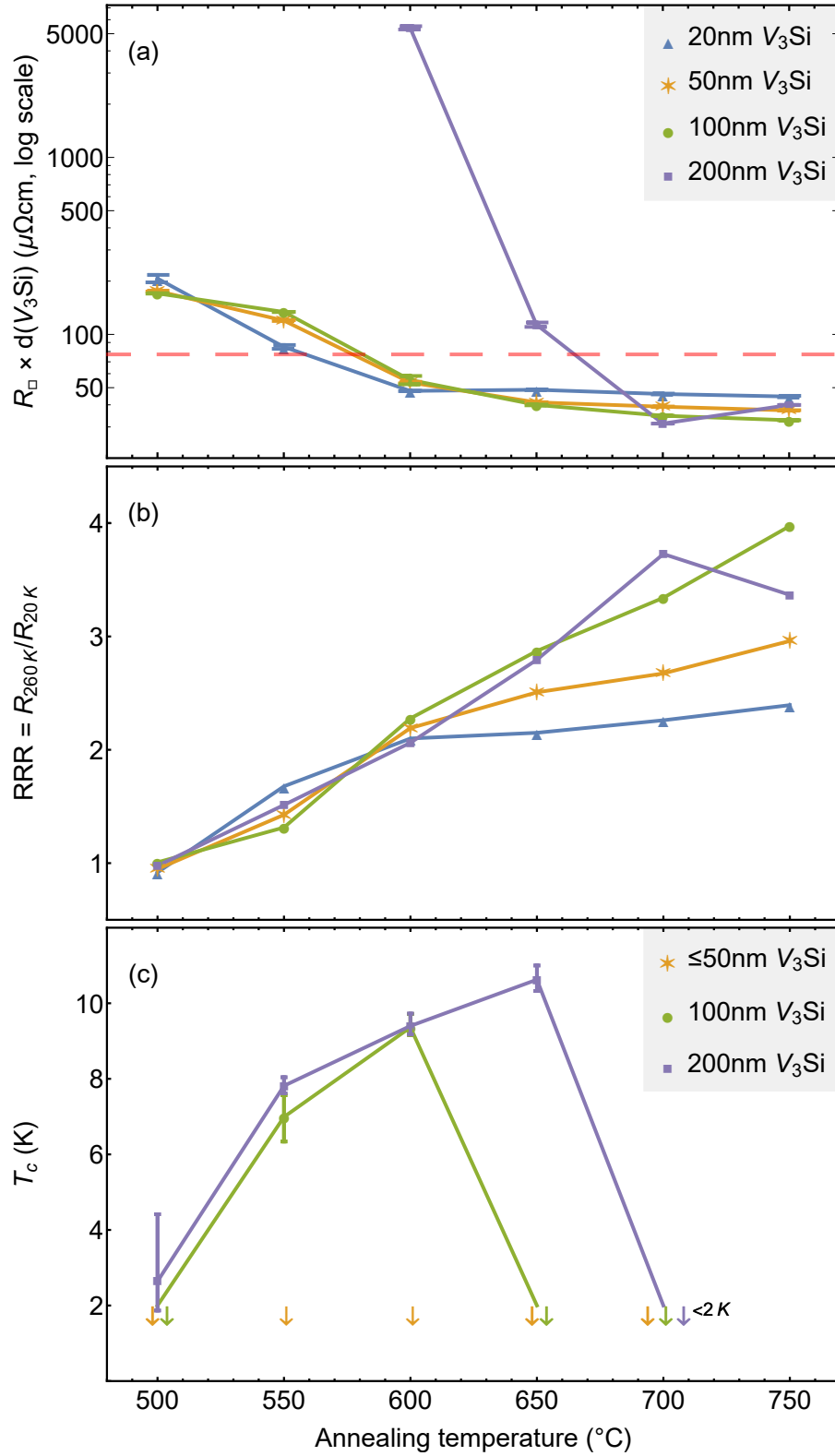


Figure 3.24: (a) The resistivity vs annealing temperature for each film thickness, the red dashed line indicates the lowest resistivity expected for V_3Si (see main text). (b) The residual resistance ratio (RRR). (c) The critical temperature. Arrows down (\downarrow) indicate that the critical temperature (if any) is below 2 K.

tures obtained for 200 nm films on HF-cleaned Si are within experimental error identical to those obtained on SiO₂ (see data points for 600 and 650 °C in Fig. 3.9), which means that the critical temperature does *not* depend on the thickness of the V₃Si film itself, but rather on the initial thickness of the as-deposited layer.

3.3.5 Conclusion

V₃Si is certainly a promising material for the CMOS integration of JoFETs, with a relatively low Schottky barrier to n-doped Si [29], a critical temperature that can reach over 15 K in thin films, and near-perfect lattice matching. Two challenges to its integration have however been identified; a negative impact of strain that is built up during thermal processing and compounded during cooling to cryogenic temperatures, and the appearance of VSi₂ at the V₃Si/Si interface at higher temperatures. Luckily, both of these can be mediated with the appropriate choice of annealing time and temperature.

The strain built up in the V₃Si thin film appears to depend mainly on the maximum temperature reached during annealing, at which point the atoms are most mobile. Since any solid-state reaction or diffusion is enhanced exponentially with temperature, relatively little plastic deformation will occur during and after cooling, such that the system is deformed only elastically and essentially retains a “memory” of this highest-temperature moment. Similarly, any reaction with the underlying silicon will be limited by atomic diffusion rates, and can therefore be controlled with thermal processing. Since the total distance that needs to be traveled by either species (V or Si) inside the initially amorphous V₃Si film to arrange itself in a crystalline manner is much smaller than that which needs to be crossed to consume the film by VSi₂ formation, it should be possible to find some time and temperature range within which to optimize the trade-off between V₃Si crystallinity and VSi₂ formation. It is not yet clear whether it is preferable to aim for some minimal critical temperature, and then minimize the thickness of the VSi₂ layer within that constraint, or if more robust superconductivity is required. As-deposited V₃Si already superconducts on a silicon substrate at 0.9 K (and at 1.2 K on sapphire), similar to what is observed for traditional aluminum thin films. Nor is it clear that the thickness of the VSi₂ layer should be minimized at all. For longer or more diffusive (highly doped) silicon channels, the precise details of the contact materials will not be relevant anymore, and the critical current will mainly be limited by the length of the semiconducting part of the channel. Schottky barriers will appear at the VSi₂/Si interface, rather than at the initial V₃Si/Si boundary, and will move underneath the gate electrode, where their effective width can be controlled (see chapter 4). Even when taking into account that every nanometer of consumed Si will give way to 2.3 nm of VSi₂, the proximity effect will likely be suppressed less in such a metallic region.

In the end, whichever combination of VSi₂ formation and V₃Si crystallization turns out to be optimal in JoFET devices, the sensitivity of this reaction to heat enforces a stringent thermal budget on subsequent processing steps. Such constraints can be avoided entirely by selecting a superconducting silicide that actually *is* the thermodynamically favored phase within its system, and therefore stable. We find this property in the more conventional PtSi, a material that has already been used in MOSFET fabrication for decades, and which we will discuss in the next section.

3.4 PtSi

While V_3Si is exciting for its novelty in the context of CMOS technology, PtSi is important precisely because it has already been successfully integrated in multiple generations of VLSI devices. What is unknown however, is how this material can be best used in superconducting applications such as JoFETs. To study this aspect, we need to move from unpatterned “blanket” thin films to actual devices. By patterning transmission line measurement (TLM) structures, we can measure such things as the contact and channel resistance, and ultimately determine the superconducting properties of the PtSi/Si interface at low temperatures.

The main concern in JoFET fabrication is that the proximity effect will be suppressed, part at the superconductor/semiconductor (S/Sm) interface and part within the semi-conducting channel, to the point that Josephson coupling can no longer occur between the superconducting leads. While the precise details of electron-hole pair decoherence in silicon depend on such experimental variables as doping level, surface oxide scattering, and the interaction with charging levels of dopants and the channel itself, we can to first order assume that it decays exponentially with the length L of the junction [72],

$$I_c \propto e^{-L/\xi}. \quad (3.19)$$

Expecting ξ to be on the order of tens of nanometers, we therefore aim for devices with the shortest channel possible, so that we will be limited only by the interface transparency. One way to push this length down, is by having the contact silicide encroach underneath the gate spacers (see Fig. 3.23). Since dopants in the channel can aggregate at the PtSi/Si interface [73], this gives some control over the Schottky barrier. By bringing the S/Sm interface within reach of the electrostatic gate field, it further provides control of the barrier by the gate.

Encroachment, in turn, brings along a new risk. In the self-aligned silicide (SALICIDE) process, the metal is deposited once the contact openings to the monocrystalline silicon have been defined, after which rapid thermal processing activates the silicidation reaction. In the case of PtSi formation, which is diffusion controlled [73], planar silicide layers grow with a thickness proportional to the square root of time. Encroachment starts once the silicide extends below the spacers¹¹, and then suddenly speeds up once the silicide reaches the buried oxide below the contacts. This sudden increase in horizontal silicide growth underneath the spacer and gate can be understood as follows. The reaction is limited by the rate at which Si atoms (the dominant diffusing species in PtSi formation [74]) arrive at the Pt or Pt_2Si reservoir, where they bind to Pt and form PtSi. This arrival rate then balances the rate at which Si atoms dissolve into the forming PtSi from the other side of the film, which means that in a planar configuration the Si/PtSi interface recedes downward at a pace proportional to the upward extension of the PtSi/ Pt_2Si or PtSi/Pt interface. Once the Si on the bottom of the contacts is fully consumed however, the surface area of the Si/PtSi interface is given only by the vertical sides of the Si channel on which the silicide is encroaching. Balancing the arrival rate of Si atoms at the large planar Pt reservoir now requires a much higher absorption rate of Si at the smaller, vertical Si/PtSi interface on the side of the channel. At this point, the distance of horizontal encroachment becomes the effective diffusion length that limits the reaction as Si atoms diffuse through this narrow estuary, and the $\ell \propto \sqrt{t}$ behavior begins anew.

¹¹This can be delayed by epitaxial growth of Si in the contact openings.

Table 3.5: An overview of the lots prepared for the study of PtSi formation.

Name	Reference	Wafer types	Split type
Lot 1	D16S0667A	blanket, patterned (TLM)	Pt thickness
Lot 2	D16S0667B	blanket, patterned (TLM)	RTP temperature, time
Lot 3	D19S0918	blanket	RTP temperature, # steps completed
Lot 4	D19S1723	blanket, patterned (TLM)	Pt thickness

To further expound on the difficulty of controlling encroachment, consider the respective numbers of Pt and Si atoms involved. While the thickness of planar PtSi films can be easily controlled by precisely tuning the amount of deposited Pt, any metal that remains after full consumption of the Si in the contacts will be incomparably large compared to the volume of silicon in the channel, providing a practically inexhaustible metal reservoir.

This brings us to the question that we will try to answer in this section: how can we optimize the formation of PtSi in JoFETs, such that we simultaneously obtain a high superconducting critical temperature, a moderate amount of encroachment, and highly transparent interfaces? Four lots¹² were prepared to answer this question, listed in Table. 3.5. In each of these, a *split* was made, in which the parameters of only a few of the process steps vary across the wafers. Since the 200 and 300 mm industrial-cleanroom fabrication technologies employed at the CEA LETI are highly reproducible, we can then with high confidence assume that any variations observed in the properties of these wafers are due to the differences designed in the split¹³.

3.4.1 Split 1: platinum deposition

The first lot, detailed in Table 3.6, considers a variation in deposited Pt thickness. Though perhaps at first sight a mundane parameter, its choice in fact has wide-ranging consequences for the properties of the final device. First, this parameter gives direct control over the superconducting critical temperature, which depends strongly on both the quality and thickness of the film. Second, the duration of the silicidation and thickness of the silicide is likely to affect the morphology of the interface [75], which in turn influences the S/Sm interface transparency. Third, the amount of available platinum determines the depth to which the silicon in the contacts will be consumed, and thus how far the superconducting contacts will be from the channel.

As discussed before in the context of V_3Si , resistivity is a good proxy for the overall material quality. By assuming that all the Pt had reacted in a 1:1 ratio with the underlying silicon to form PtSi, this resistivity can be calculated directly from the sheet resistance of the bulk wafers (P07 – P11 in Table 3.6). As shown in Fig. 3.25, consistently higher quality films are obtained by depositing thicker layers of platinum. Such a decrease in R_{\square} with film thickness has been observed before, with resistivities reducing from 59 to 32 $\mu\Omega\text{ cm}$ by going from 14 to 250 nm of PtSi [76], and has been attributed to increases in grain size. This quality is reflected also in the residual resistance ratio, indicating fewer non-thermal scattering events in thicker films. Such scattering impacts superconductivity

¹²In this context, a “lot” is a batch of 25 wafers. From a common Germanic root, the same word exists in French; in algorithm 1.1 you could find yourself assessing a *lot de 6 rouleaux*.

¹³To get a sense of the reproducibility of CMOS technology: for our TLM mask set it takes 111 process steps to make a “base wafer” ready for metallization, and a further 72 steps to finish the back-end. Still, broken devices are usually due only to the steps intentionally varied in the split. Even wafers with actual transistors, which require hundreds more steps, typically have 80–90% yields.

Table 3.6: Lot 1 (D16S0667A): A set of both patterned and blanket wafers are prepared, with deposited Pt thicknesses varying from 5 to 25 nm. A double is included for the patterned wafer with median thickness of 15 nm, to perform in-line XRD.

Lot 1	Wafer #										
	1	2	3	4	5	6	7	8	9	10	11
Blanket bulk Si wafer	○	○	○	○	○	○	●	●	●	●	●
Patterned SOI wafer	●	●	●	●	●	●	○	○	○	○	○
Deposit <> nm Pt + TiN	5	10	15	15	20	25	5	10	15	20	25
RTA 500 °C 120s	●	●	●	●	●	●	●	●	●	●	●
Etch TiN	●	●	●	●	●	●	●	●	●	●	●
Etch Pt	●	●	●	●	●	●	●	●	●	●	●
XRD	○	○	○	●	○	○	●	●	●	●	●
Complete back-end	●	●	●	○	●	●	○	○	○	○	○

in the thinner films [27], which is further affected by the finite thickness itself [77].

Low-temperature measurements on TLM structures are required to further investigate the superconducting behavior of PtSi/Si junctions. Unfortunately, it was found that none of the wafers in lot 1 (nor lot 2, for that matter) had any working devices. This was ultimately traced back to faulty base wafers, where the Si in the contacts had been over-etched during the last process step prior to metallization (see Fig. 3.26). A new lot with a split identical to lot 1 (patterned wafers only) was launched to correct for this mistake, which was analyzed by master student Axel Leblanc and is discussed briefly in section 4.2.2¹⁴.

¹⁴Industrial 200 and 300 mm cleanroom technology is certainly more reliable, and in many aspects more advanced than fabrication processes in academic cleanrooms. It is also much slower. After the conclusive FIB-SEM analysis shown in Fig. 3.26, it took from 2018-09-04 until 2020-01-27, a total of 1 year, 4 months and 23 days, before a replacement for lot 1 was finished (this replacement, lot 4, was in fact given priority in the cleanroom). This delay partially explains the relative emphasis on material studies in this manuscript.

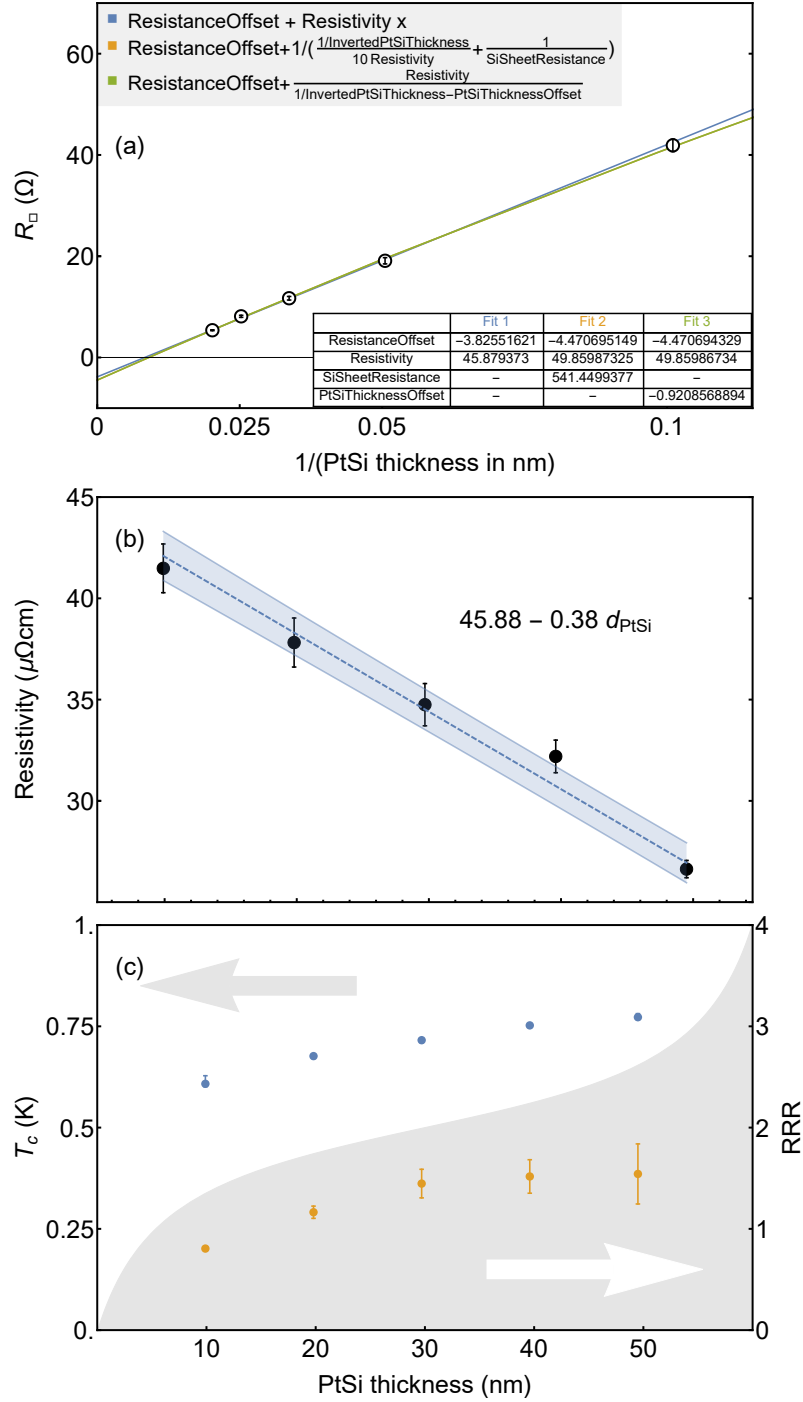


Figure 3.25: (a) It is customary to plot the sheet resistance versus the inverse of the film thickness (1.98 nm of PtSi is formed per nm of deposited Pt), such that the resistivity of the material is given by the slope. If this method is used, which assumes that the resistivity of the film is independent of the film thickness, we have to conclude that the measurement apparatus has an offset. (b) If instead, the resistivity is calculated for each individual film, and then plotted versus the thickness, we find a more likely explanation: the film quality improves with thickness. The blue dashed line is a linear regression with 2σ error margin. (c) Both the residual resistance ratio (RRR) and the critical temperature T_c improve with film thickness, in line with the improvement in room-temperature resistivity.

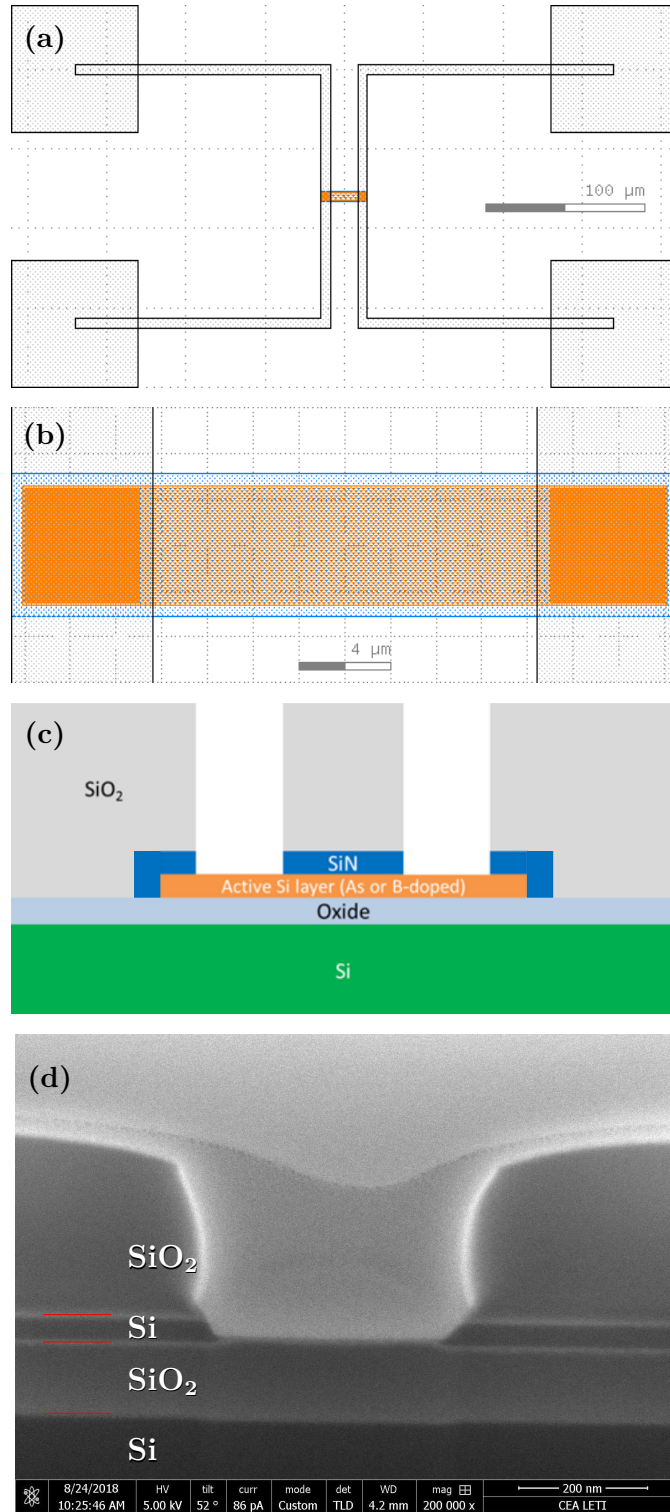


Figure 3.26: (a) The design of a TLM structure: in gray are shown the bonding pads and lines leading to the two PtSi contacts on either side of a silicon channel. (b) Zoom on the channel itself. (c) A cross-section of the structure at the stage that the wafer was analyzed: before platinum deposition. (d) A FIB-SEM image taken of wafer 20 from base lot D16S0667, the lot that the patterned wafers in lots 1 and 2 were taken from. Shown is a contact opening for a TLM structure, where it is clear that there is no remaining Si in the contact. This explains why no working devices were found.

3.4.2 Split 2: rapid thermal processing (RTP)

Like the thickness of deposited Pt, the thermal processing that follows can be used to control a wide range of parameters. In the context of JoFETs, the most important of these is perhaps again the encroachment. Annealing can be done in a single high-temperature step, immediately forming the final silicide phase, as was done in the split discussed in section 3.4.1. Since the grain size and thus the quality of the final silicide improves with temperature [76], it is necessary to go up to perhaps 450 or 500 °C. This is about 150 to 200 °C higher than the temperature at which PtSi nucleates at the silicon-metal interface, leading to full metal consumption within seconds. Needless to say, this makes it impossible to control the encroachment. One solution is to divide the annealing process into two: a first step at a lower temperature, where diffusion is slow enough to control the lateral formation, followed by the removal of unreacted metal and Pt₂Si by selective etch and a second heating at 500 °C to reach the desired material quality.

As shown in Table 3.7, six conditions were tested for the first annealing step, with temperatures of 300 or 350 °C, and durations between 10 and 60 s, after which all samples were annealed a second time at 500 °C for a longer period of 2 minutes. A relatively large amount of Pt (25 nm) was deposited, about 30% more than what is required to consume the 25 nm of Si on the bottom of the contact openings¹⁵, to ensure that the reactions would be limited by the temperature, rather than the available reagents. Previous studies on thicker (120 nm Pt) films with annealing times of multiple hours had shown that Pt₂Si forms at 260 °C, while PtSi appears around 312 °C [78]. Multiple phases are unlikely to appear simultaneously in metal-silicon systems, especially in thin films [79], and Pt₂Si usually occurs before the nucleation of PtSi [80], though this first phase may be skipped altogether in thinner films or higher temperatures. Short heating at 300 °C during the first RTP is thus expected to lead to relatively slow and controllable Pt₂Si formation by Pt diffusion into the Si [81], followed by Si diffusion into the Pt₂Si and the nucleation of PtSi during RTP 2 [82]. At a higher temperature of 350 °C during RTP 1, the Pt₂Si phase is expected to be either quickly transformed into PtSi, or skipped entirely, leading to the immediate appearance of the final phase. Since 350 °C is still far lower than the temperature used during single-step annealing (see Table 3.6), it is hoped that some degree of control over the Pt consumption remains.

The subsequent selective etch (SE) of the TiN capping layer and unreacted Pt is designed to leave behind both PtSi and Pt₂Si. Any Pt₂Si that formed during the first RTP (and remained after the SE), will be fully transformed into PtSi during later thermal processing at 500 °C, consuming a roughly equal volume of Si. In case PtSi already appeared (in which case no Pt₂Si is expected to remain [79]), no further Si consumption should follow during RTP 2.

As with lot 1, no working devices were found on the patterned wafers (see Fig. 3.26). Shown in Fig. 3.27 are sheet resistance measurements on the blanket wafers averaged over 49 equally spaced locations on the wafer, taken after both RTP 1 (and SE) and RTP 2. The large difference in sheet resistance after RTP 1 between the two sets of three wafers annealed at different temperatures suggests that different reactions occurred, likely the formation of Pt₂Si at 300 °C, and PtSi at 350 °C. The resistivity obtained after RTP 1 at 350 °C for 30 s on P06 is 35(2) μΩ cm, already close to the 26.6(4) μΩ cm obtained after

¹⁵Of course this estimate for the available volume of Pt is not accurate for trenches in patterned wafers, ignoring both the reduced Pt thickness on the bottom due to shadow effects, and the available volume of Pt on the side walls. In any case, 30% overhead is enough to assume that the reaction will not be limited by the amount of Pt available.

Table 3.7: The split in lot 2 (D16S0667B), with a variation in the temperature and time of the first annealing step.

Lot 2	Wafer #												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Blanket bulk Si wafer	●	●	●	●	●	●	○	○	○	○	○	○	○
Patterned SOI wafer	○	○	○	○	○	○	●	●	●	●	●	●	●
Deposit 25 nm Pt + 10 nm TiN	●	●	●	●	●	●	●	●	●	●	●	●	●
RTP 1: Temperature (°C)	300	300	300	350	350	350	300	300	300	300	350	350	350
time (s)	10	30	60	10	30	60	10	30	60	60	10	30	60
Etch TiN	●	●	●	●	●	●	●	●	●	●	●	●	●
Etch Pt	●	●	●	●	●	●	●	●	●	●	●	●	●
RTP 2: 500 °C 120s	●	●	●	●	●	●	●	●	●	●	●	●	●
XRD	●	●	●	●	●	●	○	○	○	●	○	○	○
Complete back-end	○	○	○	○	○	○	●	●	●	○	●	●	●

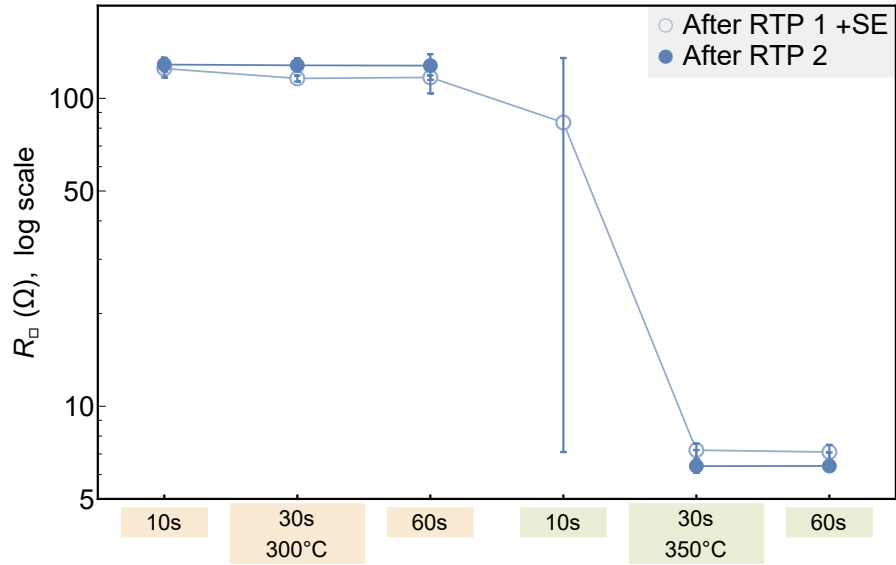


Figure 3.27: Sheet resistance measurements after RTP 1 and the selective etch (open symbols), and after RTP 2 (closed symbols). A clear distinction is visible between the wafers annealed at 300 °C, and those annealed at 350 °C. Wafer P04, annealed at 350 °C for 10 s, had a variation in color, with a darker spot near the center, where the R_{\square} was lower — logs of the RTP show a ~ 10 °C higher overshoot on the thermometer below the center of the wafer than on those near the edge.

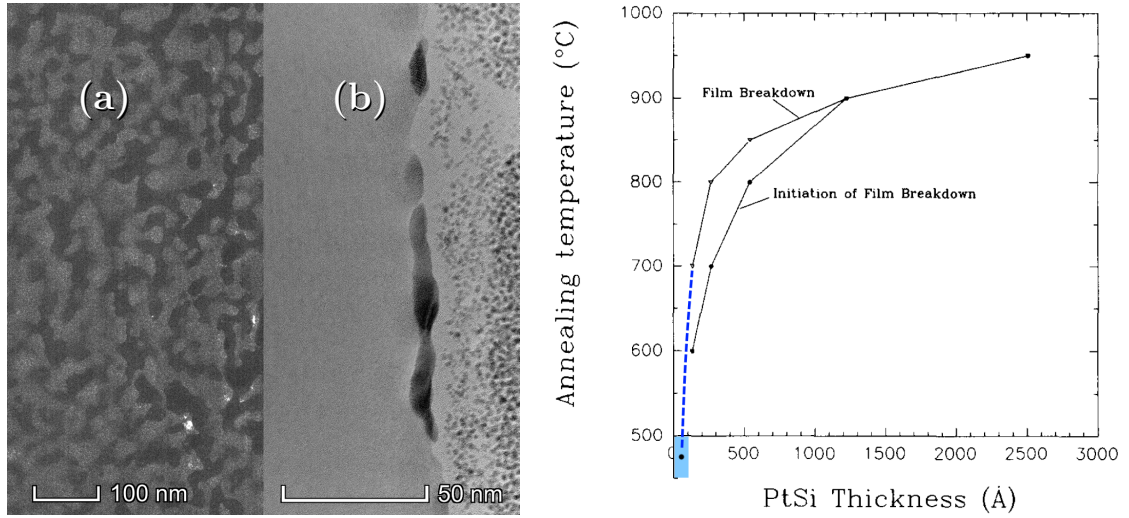


Figure 3.28: (Left, **a**) Plane view SEM and (**b**) Cross-section TEM of lot 2 (D16S0667B) P01. The lighter meanders on the left are PtSi (insulating Si appears darker under SEM), which is visible as darker regions on the right (heavier Pt is darker under TEM), as confirmed by EDS analysis. XRD analysis found a grain size of around 3–4 nm, on the same order as the size of the islands. (**Right**) Reproduced from Ref. 76, blue dashed line added. Shown is the temperature at which PtSi films of different thicknesses break down, which is attributed to Pt diffusion into Si, disintegration of PtSi layer and simultaneous formation of a Pt_3Si phase.

single-step annealing at 500 °C (see Table 3.6, wafer 11 and Fig. 3.25). The high sheet resistance after RTP 1 on the wafers annealed at 300 °C cannot be explained only by the formation of Pt_2Si , which has a similar resistivity to PtSi of around $32 \mu\Omega \text{ cm}$ [83]. Instead, we now know that the selective etch with aqua regia ($\text{HNO}_3\text{:HCl:H}_2\text{O}$ with volume ratios 2:2:4, at 60 °C) that was employed for Pt removal, is likely to have etched Pt_2Si as well, and may even have transformed some of the Pt_2Si at the interface with Si into PtSi [84].

Fig. 3.28 (left) shows SEM and TEM analysis of a sample annealed at 300 °C during 10 s, followed by SE, and then at 500 °C during 2 min. The film has likely broken up during the second RTP at 500 °C, which would be consistent with earlier studies on the stability of PtSi thin films [76]. As shown in the right panel of Fig. 3.28, though not reported before, the breakdown of a few-nanometer film into islands is likely to occur already at a temperature below 500 °C. Evidently, only a very small amount of Pt remains in the system, suggesting that most or all of the Pt_2Si that had formed after RTP 1 had indeed been etched by aqua regia as reported elsewhere [84].

The results of low-temperature measurements on the six blanket wafers in lot 2 are shown in Fig. 3.29. Samples that we annealed only at 300 °C during RTP 1 had high sheet resistances at room temperature, and show even higher sheet resistance below temperatures at which carriers introduced by dopants in Si freeze out ($\lesssim 150 \text{ K}$). These low-temperature resistance values are five orders of magnitude above that obtained on the samples annealed during RTP 1 at 350 °C, a difference that is expressed also in the residual resistance ratio. This is a behavior that is not expected of continuous metallic films, and indicates that at least part of the current path passes through insulating regions, likely the substrate itself.

Further, Fig. 3.30c shows that the samples with only meanders/islands of PtSi

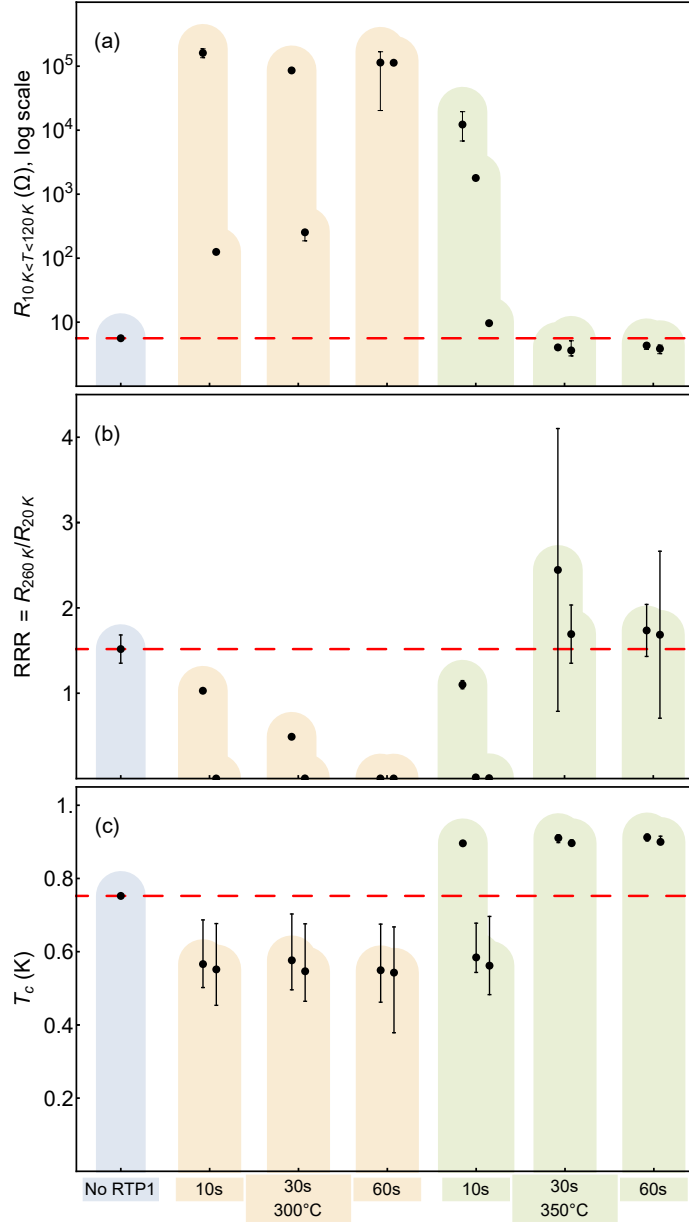


Figure 3.29: Low-T behavior of lot 2 (averaged between 10 and 120 K, see Fig. 3.30), in blue on the left lot 1 P11 (identical Pt thickness, annealed once at 500 °C, see Table 3.6). Multiple measurements performed on each wafer: first data point within weeks after RTP 2, second a few months later (degradation associated with Si depletion of PtSi by SiO₂ formation). Three data points on P04 (350 °C, 10 s) each during the first round of measurements, on samples taken at increasing distances from the dark spot. **(a)** Sheet resistance between 10 K and 120 K, in which all samples had relatively flat $R(T)$. **(b)** The residual resistance ratio. **(c)** The critical temperature, measured with a bias current on the order of 1 μA .

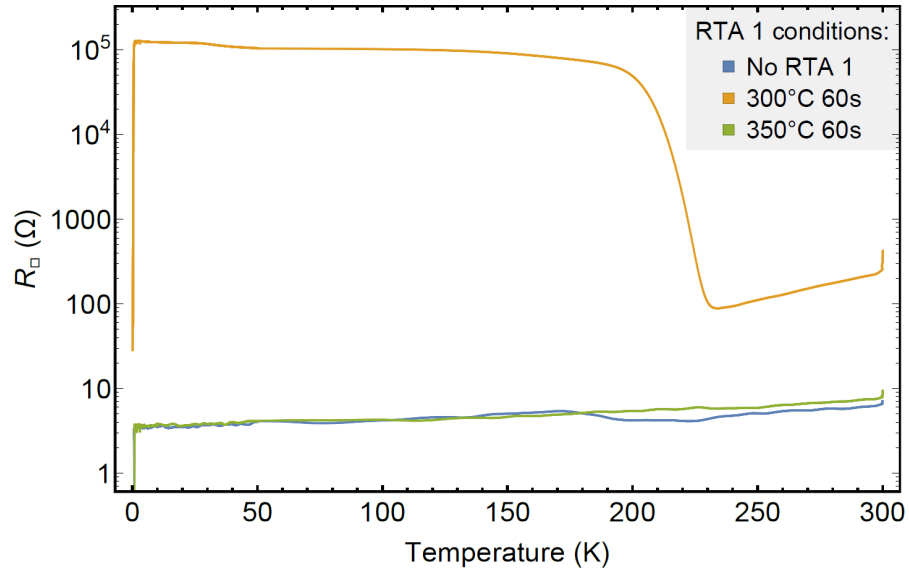


Figure 3.30: Logarithmic plot of the sheet resistance of three wafers, with either no low-temperature annealing step at all, RTP 1 at 300 °C or RTP 1 at 350 °C. The wafers that underwent RTP 1 subsequently had unreacted Pt and Pt₂Si removed by selective etch, after which all three underwent a final annealing step at 500 °C. The film on the **first** of these, only annealed once at 500 °C, has a metallic behavior: the resistance decreases as the wafer is cooled. This is consistent with the thick, 50 nm PtSi film that we expect. The **second** wafer underwent RTP 1 at a temperature too low to transform the entire Pt layer to PtSi, such that only a small amount of metal remained after selective etch. The ensuing high-temperature anneal then likely caused de-wetting of the film into disconnected islands, which explains the insulating behavior at temperatures below the freeze-out of intrinsic carriers in the silicon substrate. A slightly higher temperature during RTP 1, as experienced by the **third** sample, allows for full Pt consumption by PtSi formation during RTP 1, such that the resulting film remains stable during RTP 2, and metallic behavior is again recovered.

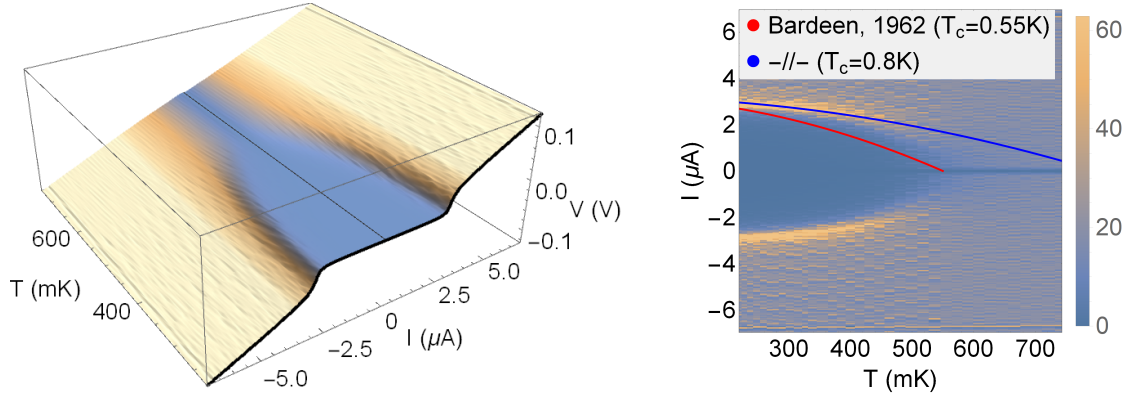


Figure 3.31: Critical (switching) current measured on lot 2 P01. **(Left)** Measured voltage as the bias current is swept from 0 to $\pm 7 \mu\text{A}$. The inflection points indicate the switching current. **(Right)** Differential resistance ($\text{k}\Omega$) extracted from the plot on the left, with the predicted temperature dependence of the switching current for thin films superposed in red, using $T_c = 800 \text{ mK}$ and $I_c(T = 0) = 3.2 \mu\text{A}$.

(RTP 1 at 300°C for any duration, or 350°C for 10s), have both lower T_c and broader superconducting transitions. This can be explained by the reduced thickness of the films (see Fig. 3.25), as well as their inhomogeneity. Measurements of the critical current between aluminum bonding wires with roughly $50 \times 50 \mu\text{m}^2$ contacts to the PtSi film on P01 from lot 2 are shown in Fig. 3.31. The temperature dependence of the critical current does not follow the usual $(1 - T/T_c)^{2/3}$ behavior for thin films [59, 77, 85], and instead drops off more sharply close to the effective $T_{c,\text{eff}} \approx 0.6 \text{ K}$. It is likely that the thicker regions in these films have higher critical temperatures of up to perhaps 0.8 K , while the effective T_c is given by the weakest links in any current path.

3.4.3 Split 3: A look under the hood

A third split was prepared (see table 3.8) to better understand what happened during the first annealing in split 2 discussed in the previous section. In addition to the temperatures used for RTP 1 in this earlier split, two more were added: 250°C and 325°C . To simplify the comparison, no further split in annealing times was used, and instead all wafers were either not annealed at all or for the same duration of 60s. One of the questions raised during the previous campaign was whether the island formation had occurred during RTP 2. To answer this, three wafers were prepared for each RTP 1 condition: one where only the protective capping layer of TiN was removed by selective etch, another where also the unreacted Pt and Pt_2Si were removed, and a last wafer that further underwent the final annealing step at 500°C during 2 min. Unfortunately, a mistake was made in the planning of the Pt deposition, and 10 nm was deposited instead of 25 , making direct comparison to lot 2 difficult. We will therefore mainly limit the discussion to comparisons between wafers of lot 3.

In Fig. 3.32a are shown the sheet resistances for each of the wafers listed in Table 3.8. The wafers are grouped on the x -axis by the temperature used during RTP 1, such that an consistent rise in sheet resistance during the fabrication process becomes apparent. Most striking is that for all RTP 1 temperatures, a large increase in sheet resistance by roughly an order of magnitude can be observed immediately after the selective etch with aqua regia (green curve), indicating the removal of most of the material present after

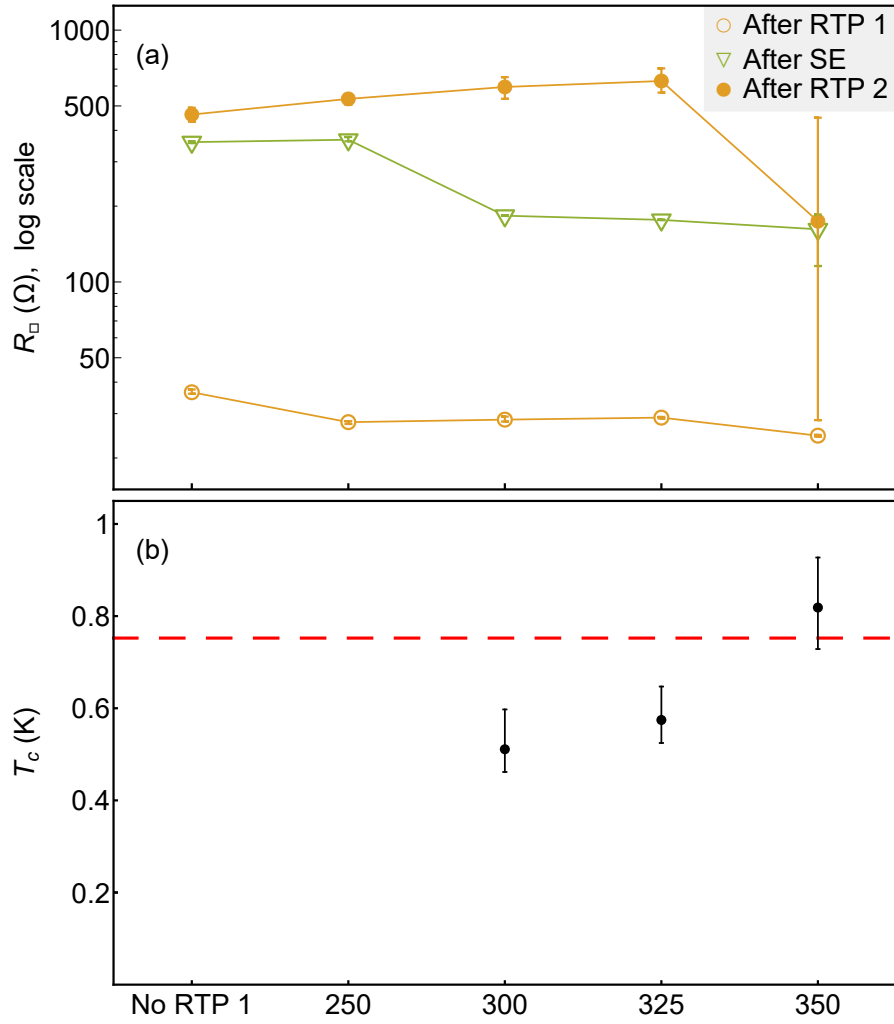


Figure 3.32: (a) Logarithmic plot of the sheet resistances of each wafer in Table 3.8, grouped by temperature during RTP 1. A large increase in sheet resistance is observed on all wafers after SE with aqua regia, with a subsequent increase after RTP 2. (b) Critical temperature measured on wafers P09, P12 and P15, showing a positive dependence on the temperature during RTP 1. The dashed red line indicates the T_c of reference sample P11 from lot 1, where 25 nm of Pt was deposited, and which was annealed once at 500 °C.

Table 3.8: The split for lot 3. Four annealing temperatures between 250 and 350 °C are evaluated for the first RTP. For each temperature, three wafers are stopped at different stages.

Lot 3	Wafer #														
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
Deposit															
10 nm Pt + 10 nm TiN	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Anneal 250 °C 60s	○	○	○	●	●	●	○	○	○	○	○	○	○	○	○
Anneal 300 °C 60s	○	○	○	○	○	○	●	●	●	○	○	○	○	○	○
Anneal 325 °C 60s	○	○	○	○	○	○	○	○	○	●	●	●	○	○	○
Anneal 350 °C 60s	○	○	○	○	○	○	○	○	○	○	○	○	●	●	●
Etch TiN	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Etch Pt (and Pt ₂ Si)	○	●	●	○	●	●	○	●	●	○	●	●	○	●	●
Anneal 500 °C 120s	○	○	●	○	○	●	○	○	●	○	○	●	○	○	●

RTP 2. This is consistent with the hypothesis that aqua regia also etches Pt₂Si [84]. While PtSi was already found after annealing at 350 °C for 60 s in lot 2, it is unlikely that the entire volume of Pt had been consumed to form this phase on the present lot. When 25 nm of Pt was deposited on the wafers of lot 2, a large difference was seen between the sheet resistances of wafers annealed at 300 °C and those heated to 350 °C (see Fig. 3.27). Further, PtSi films of 20 nm should be stable up to 600 °C, and not form any islands until 700 °C (see Fig. 3.28) [76]. However, when the wafer heated to 350 °C during RTP 1 was subsequently annealed at 500 °C, a large variation in sheet resistance across the wafer was observed (solid orange curve in Fig. 3.32), suggesting the onset of the film breakdown. All other wafers, annealed at 325 °C or less during RTP 1, saw a further large increase in sheet resistance after RTP 2, which is expected if de-wetting has occurred.

A combination of XPS (X-ray Photo-electron Spectroscopy), AES (Auger Electron Spectroscopy) and REELS (Reflected Electron Energy Loss Spectroscopy) analyses performed in Budapest by the team of Labar Janos László and Miklos Menyhard provided further insight into the reactions that occurred in these films. Shown in Fig. 3.33 are simulated depth profiles of the different phases of the Pt-Si system, which match the measured spectra of these measurements. Most importantly, we see that both silicide phases, Pt₂Si and PtSi, are already present after deposition (P01, Fig. 3.33a), which was later confirmed by TEM analysis (see Fig. 3.35). Not all of these phases are simultaneously crystalline, however. XRD analysis (shown in Fig. 3.34) indicates that only a single phase appears in crystalline form at the same time: (111)-oriented Pt after deposition (P01), then Pt₂Si after RTP 1 at 250 °C (P04), and untextured PtSi after annealing at 350 °C. Nor does the disappearance of a phase's crystalline form necessarily imply that it has disappeared entirely. This can be seen in Fig. 3.33a, where Pt₂Si was in fact still present in samples annealed at temperatures as high as 350 °C.

It is also notable that the PtSi that formed in P13, after annealing at 350 °C, shows no signs of any texture; all peaks from different (hkl) planes are present with roughly the same intensities as they are in reference powder samples (shown as gray vertical lines in the bottom bar). The fact that the PtSi does not have any texture in this last sample is an interesting result in itself, since “columnar” growth has been reported for all film thicknesses when PtSi forms immediately [76], as confirmed by XRD analysis performed on samples annealed only once at 500 °C (not shown). It is known that the preferred

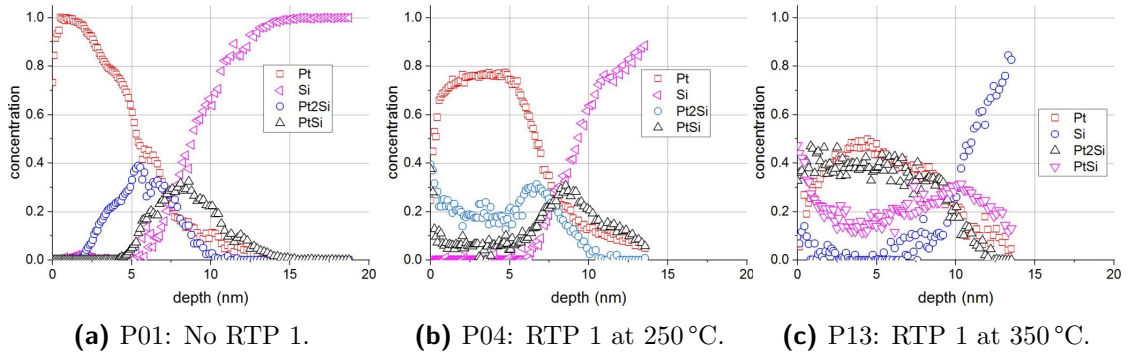


Figure 3.33: Combining plasmon spectroscopy and XPS measurements, it was possible to estimate the variation with depth in the relative densities of chemical bonds associated with Pt, Pt₂Si, PtSi and Si. It is clear that PtSi and Pt₂Si co-exist after RTP 1 up to 350 °C.

orientation depends on thickness and thermal history [86], and this disappearance of texture may be related to the increase in T_c relative to the single-step annealing process as observed in Fig. 3.29, since axiotaxitic PtSi with thin vertical columns would likely have a very small lateral grain size. One possible interpretation is that the intermediate appearance of untextured Pt₂Si prevents the inheritance of texture from the deposited Pt to the final PtSi. Since the interface between Si and epitaxial PtSi is often roughened by undulations and atomic steps due to the large lattice mismatch [87], removing texture could have the added benefit of smoothing the S/Sm interface to the channel.

Perhaps the most surprising result is that all phases already co-exist after deposition, as can be seen in the TEM image in Fig. 3.35. Intermixing during room-T deposition has been shown before [76, 84], as has an initial reaction between Pt and Si at room temperature leading to amorphous products [88, 89]. Previous reports [90] have also observed the same full stack of Pt, Pt₂Si, PtSi and Si after annealing at low temperatures. It remains unexplained however, how the crystallinity of one phase (such as Pt₂Si) can disappear during the formation of the next (e.g. PtSi), while its chemical bonds remain.

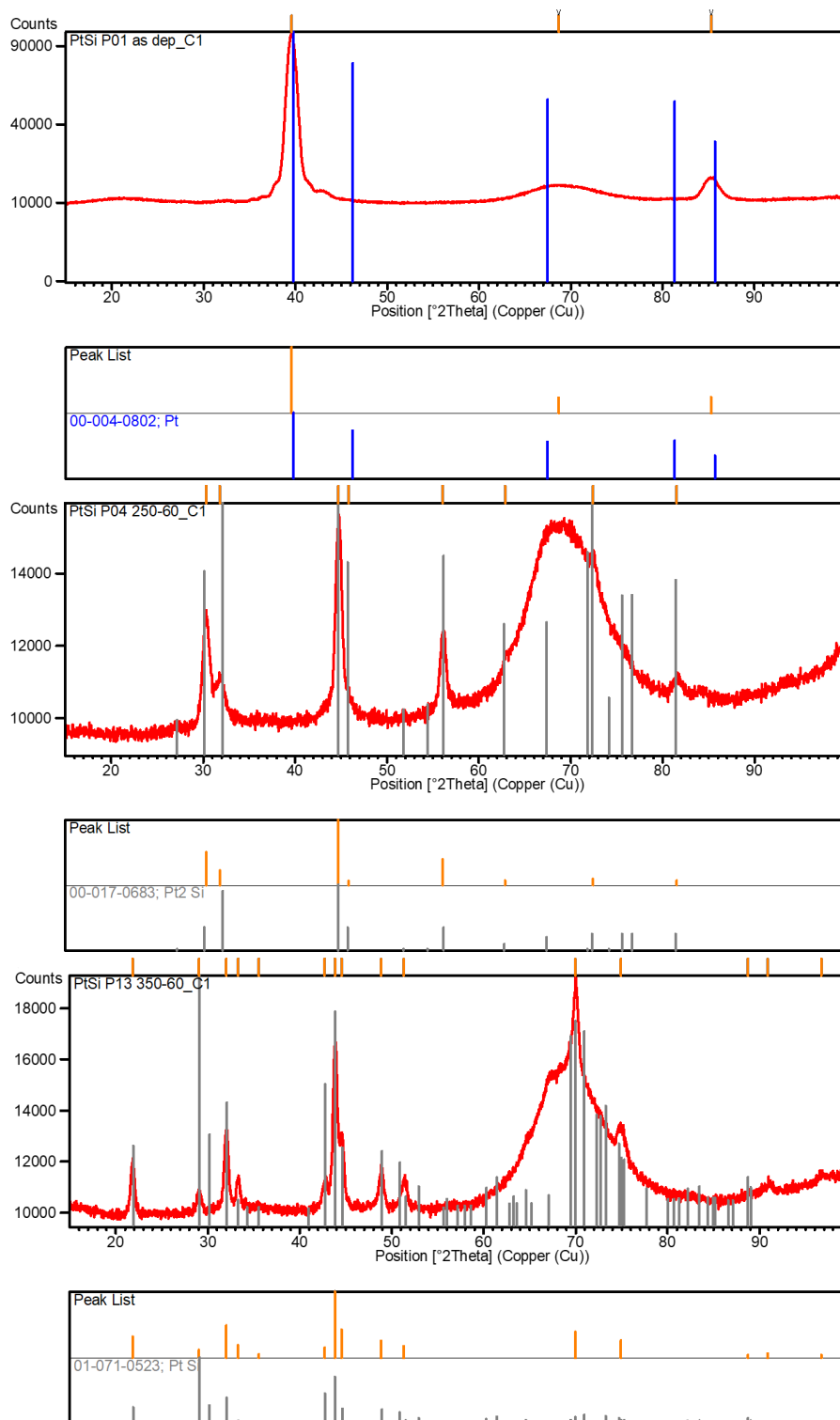


Figure 3.34: Out-of-plane $\theta/2\theta$ XRD analyses performed on wafers P01, P04 and P13 of lot 3 (see Table 3.8). **(Top)** The as-deposited wafer shows strongly textured (111)-oriented Pt. Additional reciprocal space mapping confirmed the presence of (220)-oriented grains in-plane. **(Middle)** Annealing at 250 °C leads to the formation of almost entirely untextured Pt₂Si. **(Bottom)** Annealing at 350 °C results in untextured PtSi.

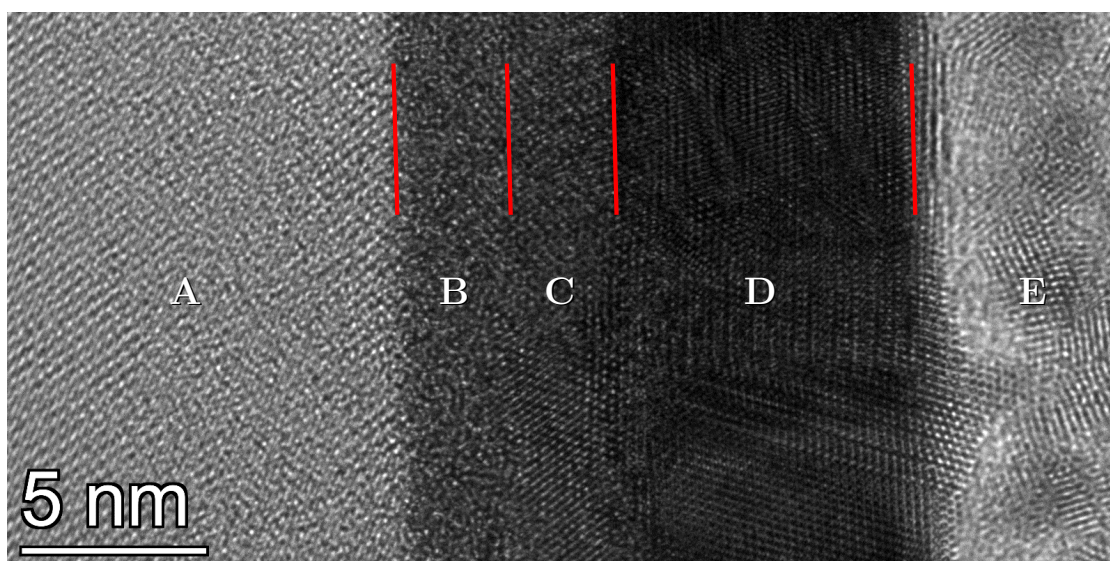


Figure 3.35: Transmission electron microscope (TEM) image of a FIB-milled lamella of P01, with five zones indicated. EDS analysis indicated that zone **A** is silicon with progressive concentrations of up to 10% of Pt dissolved close to the interface with zone **B**, which has a rather homogeneous concentration of around 50% Pt. Zones **C** and **D** contain 60–70 and nearly 100% of Pt, respectively. These results are consistent with the plasmon depth profile shown in Fig. 3.33a, which is likely broadened due to finite depth resolution.

3.4.4 Conclusion

PtSi has several advantages over V_3Si ; it is the thermodynamically favored phase within its binary system (which ensures its stability at higher annealing temperatures and allows for longer annealing times), has established process flows for its integration in MOSFET devices, and is compatible with the SALICIDE process. Furthermore, since it is formed after pure metal deposition, and thus consumes Si during its formation, its interface to the silicon channel can usefully be moved underneath the gate electrode. By introducing an additional annealing step at lower temperatures, around or above 350 °C, the quality of the PtSi can be improved, while also controlling the lateral encroachment into the channel. As we will see in the next chapter, such a reduction of the channel length, paired with the placement of the Schottky barrier within the gate electrostatic field, gives rise to a favorable Schottky-barrier MOSFET (SBMOSFET) configuration. In such a device it is possible, even at cryogenic temperatures after dopants have frozen out, to tune the effective width of the Schottky barrier, and thus modulate the proximity effect induced from the PtSi contacts.

Bibliography

- [1] Bishnu Patra et al. “Cryo-CMOS circuits and systems for quantum computing applications”. In: *IEEE Journal of Solid-State Circuits* 53.1 (2017), pp. 309–321.
- [2] L Le Guevel et al. “Low-power transimpedance amplifier for cryogenic integration with quantum devices”. In: *Applied Physics Reviews* 7.4 (2020), p. 041407.
- [3] Loïck Le Guevel et al. “A 110mK 295 μ W 28nm FDSOI CMOS Quantum Integrated Circuit with a 2.8 GHz Excitation and nA Current Sensing of an On-Chip Double Quantum Dot”. In: *2020 IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE. 2020, pp. 306–308.
- [4] Ho C Theuerer and JJ Hauser. “Getter sputtering for the preparation of thin films of superconducting elements and compounds”. In: *Journal of Applied Physics* 35.3 (1964), pp. 554–555.
- [5] Osamu Michikami and Hisataka Takenaka. “V₃Si Thin-Film Synthesis by Magnetron Sputtering”. In: *Japanese Journal of Applied Physics* 21.3A (1982), p. L149.
- [6] SP Murarka et al. “High temperature stability of PtSi formed by reaction of metal with silicon or by cosputtering”. In: *Journal of Applied Physics* 54.12 (1983), pp. 6943–6951.
- [7] Shyam P Murarka. “Silicide thin films and their applications in microelectronics”. In: *Intermetallics* 3.3 (1995), pp. 173–186.
- [8] George F Hardy and John K Hulm. “Superconducting silicides and germanides”. In: *Physical Review* 89.4 (1953), p. 884.
- [9] George F Hardy and John K Hulm. “The superconductivity of some transition metal compounds”. In: *Physical Review* 93.5 (1954), p. 1004.
- [10] WE Blumberg et al. “Correlations between superconductivity and nuclear magnetic resonance properties”. In: *Physical Review Letters* 5.4 (1960), p. 149.
- [11] A Zur, TC McGill, and M-A Nicolet. “Transition-metal silicides lattice-matched to silicon”. In: *Journal of Applied Physics* 57.2 (1985), pp. 600–603.
- [12] Hédi Ben Ghazlene, Pierre Beaufrère, and André Authier. “Crystallography of PtSi films on (001) silicon”. In: *Journal of Applied Physics* 49.7 (1978), pp. 3998–4004.
- [13] LR Testardi et al. “Unusual strain dependence of T_{c}^{extc} and related effects in A-15 superconductors”. In: *Solid State Communications* 8.11 (1970), pp. 907–911.
- [14] LR Testardi. “Unusual Strain Dependence of T_{c}^{extc} and Related Effects for High-Temperature (A-15-Structure) Superconductors: Sound Velocity at the Superconducting Phase Transition”. In: *Physical Review B* 3.1 (1971), p. 95.

- [15] LR Testardi et al. “Unusual Strain Dependence of T_c and Related Effects for High-Temperature (A-15-Structure) Superconductors: Elastic, Thermal, and Alloy Behavior”. In: *Physical Review B* 3.1 (1971), p. 107.
- [16] KN Tu, JF Ziegler, and CJ Kircher. “Formation of vanadium silicides by the interactions of V with bare and oxidized Si wafers”. In: *Applied Physics Letters* 23.9 (1973), pp. 493–495.
- [17] H Kräutle, M-A Nicolet, and JW Mayer. “Kinetics of silicide formation by thin films of V on Si and SiO_2 substrates”. In: *Journal of Applied Physics* 45.8 (1974), pp. 3304–3308.
- [18] WK Chu et al. “Identification of the dominant diffusing species in silicide formation”. In: *Applied Physics Letters* 25.8 (1974), pp. 454–457.
- [19] R Pretorius, TK Marais, and CC Theron. “Thin film compound phase formation sequence: An effective heat of formation model”. In: *Materials Science Reports* 10.1-2 (1993), pp. 1–83.
- [20] H Kräutle et al. “Reactions of Thin Metal Films with Si or SiO_2 Substrates”. In: *Applications of Ion Beams to Metals*. Springer, 1974, pp. 193–207.
- [21] RJ Schutz and LR Testardi. “The formation of vanadium silicides at thin-film interfaces”. In: *Journal of Applied Physics* 50.9 (1979), pp. 5773–5781.
- [22] S De Stefano et al. “Preparation and properties of V_3Si thin films for superconducting electronics”. In: *IEEE Transactions on Magnetics* 21.2 (1985), pp. 878–879.
- [23] WK Chu and KN Tu. “Identification of diffusion species in V- SiO_2 reactions”. In: *Applied Physics Letters* 33.1 (1978), pp. 83–85.
- [24] Gin-ichiro Oya et al. “Superconducting transition temperatures of thin V_3Si layers formed by the interaction of V films with thinly oxidized Si wafers”. In: *Journal of Applied Physics* 53.2 (1982), pp. 1115–1121.
- [25] FM d’Heurle and P Gas. “Kinetics of formation of silicides: A review”. In: *Journal of Materials Research* 1.1 (1986), pp. 205–221.
- [26] PA Psaras, M Eizenberg, and KN Tu. “Sequential silicide formation between vanadium and amorphous silicon thin-film bilayers”. In: *Journal of Applied Physics* 56.12 (1984), pp. 3439–3444.
- [27] Carlo Ferdeghini et al. “Superconducting Properties of V_3Si Thin Films Grown by Pulsed Laser Ablation”. In: *IEEE Transactions on Applied Superconductivity* 19.3 (2009), pp. 2682–2685.
- [28] M Enomoto. “The N-Ti-V system (nitrogen-titanium-vanadium)”. In: *Journal of Phase Equilibria* 17.3 (1996), pp. 248–252.
- [29] James Jeffrey Hugunin. “The superconductor/semiconductor interface of V_3Si and Si”. Master thesis. Massachusetts Institute of Technology, 1995.
- [30] H Taub and SJ Williamson. “Electrical resistivity of transforming and non-transforming V_3Si ”. In: *Solid State Communications* 15.2 (1974), pp. 181–184.
- [31] R Caton and R Viswanathan. “Normal state resistivity of single crystalline V_3Si as a function of neutron irradiation”. In: *Le Journal de Physique Colloques* 39.C6 (1978), pp. C6–385.

- [32] R Caton and R Viswanathan. “Analysis of the normal-state resistivity for the neutron-irradiated A15 superconductors V_3Si , Nb_3Pt , and Nb_3Al ”. In: *Physical Review B* 25.1 (1982), p. 179.
- [33] Wenrui Zhang et al. “Thin-film synthesis of superconductor-on-insulator A15 vanadium silicide”. In: *Scientific Reports* 11.1 (2021), pp. 1–7.
- [34] Filippo Nava et al. “Electrical characterization of alloy thin films of VSi_2 and V_3Si ”. In: *Thin Solid Films* 140.1 (1986), pp. 167–172.
- [35] Ray Radebaugh and PH Keesom. “Low-temperature thermodynamic properties of vanadium. I. Superconducting and normal states”. In: *Physical Review* 149.1 (1966), p. 209.
- [36] ST Sekula and RH Kernohan. “Magnetic properties of superconducting vanadium”. In: *Physical Review B* 5.3 (1972), p. 904.
- [37] IN Khlyustikov. “Surface Superconductivity of Vanadium”. In: *Journal of Experimental and Theoretical Physics* (2021), pp. 1–4.
- [38] Michael R Vissers et al. “Characterization and in-situ monitoring of sub-stoichiometric adjustable superconducting critical temperature titanium nitride growth”. In: *Thin Solid Films* 548 (2013), pp. 485–488.
- [39] Lei Yu et al. “Fabrication of niobium titanium nitride thin films with high superconducting transition temperatures and short penetration lengths”. In: *IEEE Transactions on Applied Superconductivity* 15.1 (2005), pp. 44–48.
- [40] NE Alekseevskii et al. “Superconducting properties of vanadium films”. In: *Physica Status Solidi (A)* 34.2 (1976), pp. 541–546.
- [41] BR Zhao et al. “Superconducting and normal-state properties of vanadium nitride”. In: *Physical Review B* 29.11 (1984), p. 6198.
- [42] SM Zhou et al. “Kinetics of platinum silicide formation followed in situ by spectroscopic ellipsometry”. In: *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films* 17.1 (1999), pp. 144–149.
- [43] TF Smith. “Pressure-Enhanced Superconductivity in V_3X Compounds”. In: *Physical Review Letters* 25.21 (1970), p. 1483.
- [44] TF Smith. “Variation of T_c with lattice parameter in V_3X compounds”. In: *Solid State Communications* 9.12 (1971), pp. 903–905.
- [45] E Fawcett. “Low-Temperature Thermal Expansion of V_3Si ”. In: *Physical Review Letters* 26.14 (1971), p. 829.
- [46] LJ Sham and TF Smith. “Thermodynamics of Pressure Effects in V_3Si and V_3Ge ”. In: *Physical Review B* 4.11 (1971), p. 3951.
- [47] RN Shelton and TF Smith. “Pressure dependence of T_c and phase stability for V_3Si ”. In: *Materials Research Bulletin* 10.10 (1975), pp. 1013–1016.
- [48] TF Smith, TR Finlayson, and RN Shelton. “Superconductivity and anharmonicity in V_3Si ”. In: *Journal of the Less Common Metals* 43.1-2 (1975), pp. 21–32.
- [49] BW Batterman and CS Barrett. “Crystal structure of superconducting V_3Si ”. In: *Physical Review Letters* 13.13 (1964), p. 390.
- [50] LR Testardi et al. “Lattice Instability of V_3Si at Low Temperatures”. In: *Physical Review Letters* 15.6 (1965), p. 250.

- [51] BW Batterman and CS Barrett. “Low-temperature structural transformation in V_3Si ”. In: *Physical Review* 145.1 (1966), p. 296.
- [52] LR Testardi and TB Bateman. “Lattice instability of high-transition-temperature superconductors. II. Single-crystal V_3Si results”. In: *Physical Review* 154.2 (1967), p. 402.
- [53] LR Testardi et al. “Lattice Instability of High-Transition-Temperature Superconductors. I. Polycrystalline A-15 Compounds”. In: *Physical Review* 154.2 (1967), p. 399.
- [54] LR Testardi. “Structural Instability, Anharmonicity, and High-Temperature Superconductivity in A-15-Structure Compounds”. In: *Physical Review B* 5.11 (1972), p. 4342.
- [55] LR Testardi. “Structural instability and superconductivity in A-15 compounds”. In: *Reviews of Modern Physics* 47.3 (1975), p. 637.
- [56] LR Testardi, JM Poate, and HJ Levinstein. “Anomalous electrical resistivity and defects in A-15 compounds”. In: *Physical Review B* 15.5 (1977), p. 2570.
- [57] HL Alberts. “Ultrasonic Wave Velocities in Solids: Elastic Properties, Temperature, and Pressure Dependence”. In: (2002).
- [58] Erik Eberhardt. *Stress & Strain: a review*. Lecture notes. 2017.
- [59] Michael Tinkham. *Introduction to Superconductivity, Second Edition*. Dover Publications, 2004. ISBN: 978-0486435039.
- [60] Tom Doekle Vethaak et al. “Influence of substrate-induced thermal stress on the superconducting properties of V_3Si thin films”. In: *Journal of Applied Physics* 129.10 (2021), p. 105104.
- [61] Tatsuya Shiino et al. “Improvement of the critical temperature of superconducting NbTiN and NbN thin films using the AlN buffer layer”. In: *Superconductor Science and Technology* 23.4 (2010), p. 045004.
- [62] JB Wachtman Jr, TG Scuderi, and GW Cleek. “Linear thermal expansion of aluminum oxide and thorium oxide from 100 to 1100 K”. In: *Journal of the American Ceramic Society* 45.7 (1962), pp. 319–323.
- [63] KG Lyon et al. “Linear thermal expansion measurements on silicon from 6 to 340 K”. In: *Journal of Applied Physics* 48.3 (1977), pp. 865–868.
- [64] P Scherrer. “Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen”. In: *Mathematisch-Physikalische Klasse* 2 (1918), pp. 98–100.
- [65] BD Cullity and SR Stock. *Elements of X-ray diffraction*. Pearson Education Limited, 2013.
- [66] Elena R Dobrovinskaya, Leonid A Lytvynov, and Valerian Pishchik. “Properties of sapphire”. In: *Sapphire*. Springer, 2009.
- [67] Ismail C Noyan and Jerome B Cohen. *Residual stress: measurement by diffraction and interpretation*. Springer, 2013.
- [68] Conal E Murray. “Equivalence of Kröner and weighted Voigt-Reuss models for x-ray stress determination”. In: *Journal of Applied Physics* 113.15 (2013), p. 153509.

- [69] Romain Gaillac, Pluton Pullumbi, and François-Xavier Coudert. “ELATE: an open-source online application for analysis and visualization of elastic tensors”. In: *Journal of Physics: Condensed Matter* 28.27 (2016). <http://progs.coudert.name/elate/mp?query=mp-2567>, p. 275201.
- [70] David J Srolovitz and Max G Goldiner. “The thermodynamics and kinetics of film agglomeration”. In: *JOM* 47.3 (1995), pp. 31–36.
- [71] F Leroy et al. “How to control solid state dewetting: A short review”. In: *Surface Science Reports* 71.2 (2016), pp. 391–409.
- [72] Anatoly F Volkov and Hideaki Takayanagi. “Effect of gate voltage on critical current in controllable superconductor–normal-metal–superconductor Josephson junctions”. In: *Physical Review B* 53.22 (1996), p. 15162.
- [73] H Takai, PA Psaras, and King-Ning Tu. “Effects of substrate crystallinity and dopant on the growth kinetics of platinum silicides”. In: *Journal of Applied Physics* 58.11 (1985), pp. 4165–4171.
- [74] JB Bindell et al. “An analytical study of platinum silicide formation”. In: *Thin Solid Films* 37.3 (1976), pp. 441–452.
- [75] Guo-Ping Ru et al. “Surface and interface morphology of CoSi_2 films formed by multilayer solid-state reaction”. In: *Materials Characterization* 48.2-3 (2002), pp. 229–235.
- [76] SR Das et al. “Thickness dependence of the properties and thermal stability of PtSi films”. In: *Thin Solid Films* 253.1-2 (1994), pp. 467–472.
- [77] Johannes Romijn et al. “Critical pair-breaking current in superconducting aluminum strips far below T_c ”. In: *Physical Review B* 26.7 (1982), p. 3648.
- [78] JT Pan and IA Blech. “In SITU study of film stress and kinetics of platinum silicide formation on silicon”. In: *Thin Solid Films* 113.2 (1984), pp. 129–134.
- [79] P Gas and FM d’Heurle. “Formation of silicide thin films by solid state reaction”. In: *Applied Surface Science* 73 (1993), pp. 153–161.
- [80] AA Naem. “Platinum silicide formation using rapid thermal processing”. In: *Journal of Applied Physics* 64.8 (1988), pp. 4161–4167.
- [81] KN Tu. “Selective growth of metal-rich silicide of near-noble metals”. In: *Applied Physics Letters* 27.4 (1975), pp. 221–224.
- [82] JE McLeod et al. “Marker and radioactive silicon tracer studies of PtSi formation”. In: *Journal of Applied Physics* 72.6 (1992), pp. 2232–2241.
- [83] Erik J Faber, Rob AM Wolters, and Jurriaan Schmitz. “On the kinetics of platinum silicide formation”. In: *Applied Physics Letters* 98.8 (2011), p. 082102.
- [84] S Jin et al. “Microstructural studies by transmission electron microscopy of the formation of ultrathin PtSi layers with novel silicidation processes”. In: *Journal of Materials Research* 14.6 (1999), pp. 2577–2587.
- [85] John Bardeen. “Critical fields and currents in superconductors”. In: *Reviews of Modern Physics* 34.4 (1962), p. 667.
- [86] AK Sinha et al. “Thermal stability of thin PtSi films on silicon substrates”. In: *Journal of Applied Physics* 43.9 (1972), pp. 3637–3643.

- [87] Hiroshi Kwarada, Iwao Ohdomari, and Shigeo Horiuchi. “Structural study of PtSi/(111) Si interface with high-resolution electron microscopy”. In: *Japanese Journal of Applied Physics* 23.10A (1984), p. L799.
- [88] L Ley et al. “Initial stages in the formation of PtSi on Si (111) as followed by photoemission and spectroscopic ellipsometry”. In: *Thin Solid Films* 270.1-2 (1995), pp. 561–566.
- [89] R Sinclair, TJ Konno, and D-H Ko. “Reactions at Semiconductor Metal Interfaces”. In: *Control of Semiconductor Interfaces*. Elsevier, 1994, pp. 247–254.
- [90] RA Donaton et al. “Formation of ultra-thin PtSi layers with a 2-step silicidation process”. In: *Microelectronic Engineering* 37 (1997), pp. 507–514.

Chapter 4

Josephson Field Effect Transistors

4.1 Introduction

“If you wish to make an apple pie from scratch, you must first invent the universe.”

– Carl Sagan

We are aiming to fabricate a transmon using only CMOS-compatible materials and processes. Before we can begin designing the actual LC circuit itself, it is necessary to thoroughly understand each of its components, not least the Josephson Field Effect Transistor (JoFET) that makes up the nonlinear inductance. In turn, we need to have some control over such properties as the superconducting critical temperature and coherence length in the source and drain, the transparency and contact resistance of the interface with the semiconductor, and the carrier mobility and scattering rates inside the channel, before it makes sense to design this JoFET. It is most efficient to extract each of these from simplified test structures, or even just blanket thin films, rather than spending time fabricating transistors immediately. To this end, both an industrial mask set, TASP (TLM for As and P-based wafers¹) and an academic mask set for test structures were used.

4.2 Test structures

4.2.1 Academic mask set: Python code for adjustable patterns

A mask set was designed with Hall bars, Greek crosses, meanders and Josephson junctions [1], for wafers where a superconducting layer (typically Al, Nb, TiN, NbTiN, ...) is deposited on top of a semiconductor (Si, SiGe, Ge, ...). This set can be used both on wafers where the superconductor has already been deposited, and on wafers where this layer will be defined by lift-off instead. The structures are fabricated with three lithography steps:

1. Contact pads with extensions towards the center of the cell are defined with optical wavelength laser lithography.

¹Any similarity to actual French words, *verlan* or otherwise, is purely coincidental.

2. A mesa is drawn, removing both the superconductor and semiconductor layers everywhere else.
3. Only the superconductor is selectively etched on part of the mesa, leaving the semiconductor exposed.

Although we did not have the opportunity to apply this process to silicon on insulator (SOI) wafers with V_3Si or PtSi films, the publicly available Python code may be useful to future studies. The main advantage of such “academic” fabrication processes are the speed with which samples can be produced, and the flexibility in adjusting the mask sets. One of these types of structure, for transmission line measurements (TLM), was studied also with the industrial mask set, as we will see in the following section.

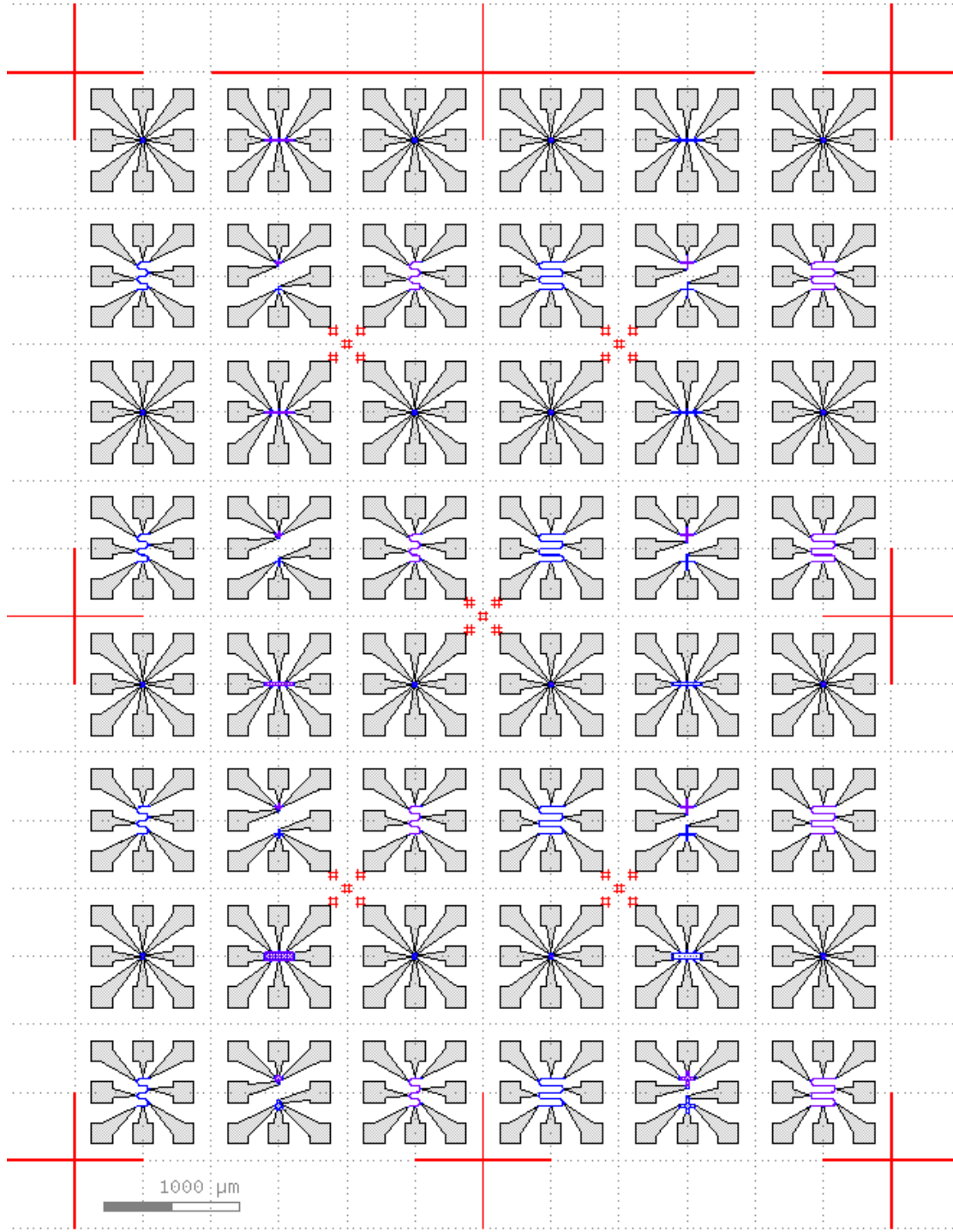


Figure 4.1: The mask set designed for a combination of laser and e-beam lithography. In this example, 6×8 cells are prepared, with 16 different TLM structures, 16 meanders (8 of Al/Ge, 8 Ge only), 8 Hall bars (4 of Al/Ge, 4 Ge only) and 16 Greek crosses (8 of Al/Ge, 8 Ge only). Each cell has different device parameters.

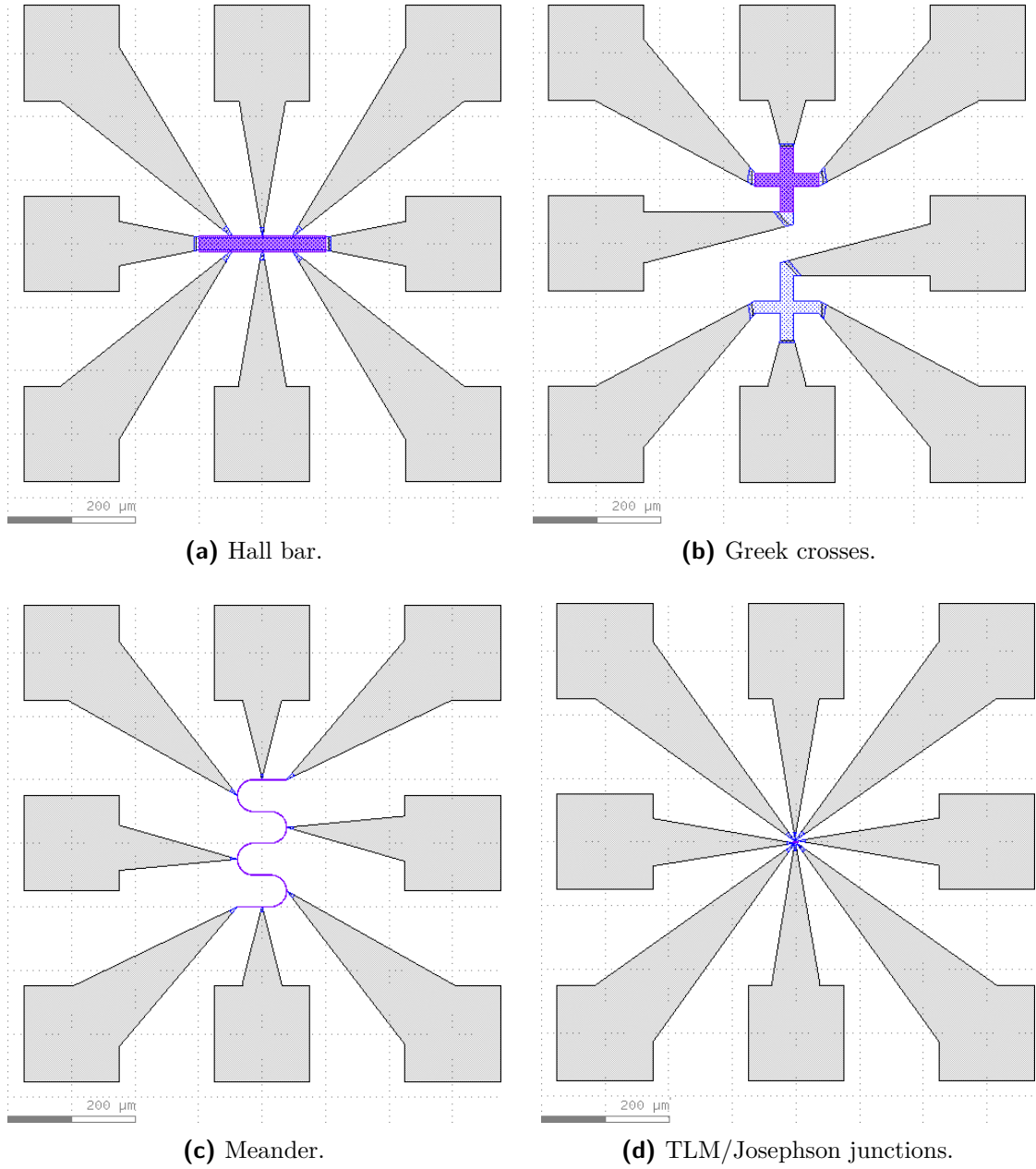


Figure 4.2: The four types of structure included in the mask set. (a) Hall bars can be used to determine both the sign and the density of carriers. (b) Greek crosses are convenient structures to accurately determine the sheet resistance. (c) Long and thin meanders to check homogeneity. (d) TLM/Josephson structures to characterize the superconductor/semiconductor interface and to study the Josephson effect in junctions of different lengths.

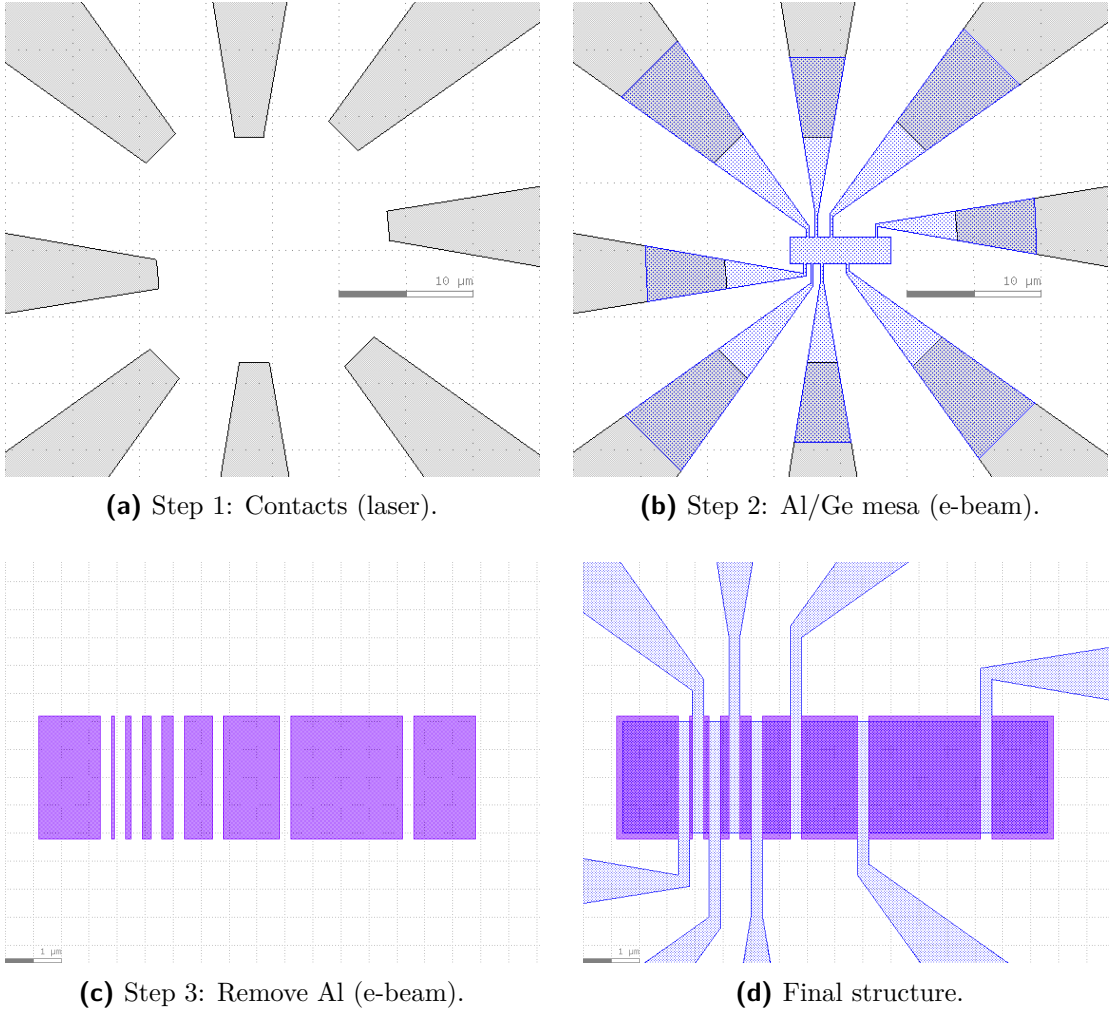


Figure 4.3: The TLM structure can be made with three lithography steps. **(a)** Contact pads with extensions towards the center of the cell are defined, either in a separate step by lifting off metal, or by choosing a resist where this laser lithography can be combined with the e-beam in the next step. **(b)** With some (adjustable) overlap, an Al/Ge mesa is drawn, removing both Al and Ge everywhere else. **(c)** Only the aluminum is selectively etched on part of the mesa, leaving Ge. **(d)** With 8 contacts, 7 different junction lengths can be measured on a single device.

4.2.2 Industrial mask set: TASP

In section 3.4.1 we discussed PtSi lot 1, a split in the thickness of deposited Pt. Unfortunately, we did not obtain any working devices on the patterned wafers in the first iteration of this lot, and a second lot (D19S1723) was launched that included only the patterned wafers, shown in Table 4.1. In Fig. 4.4 are shown the total resistances of the TLM structures in scribe RA0C1,

$$R_{\text{total}} = 2R_{\text{contact}} + \frac{L_{\text{channel}}}{W_{\text{channel}}} R_{\square, \text{Si}}, \quad (4.1)$$

for each wafer in this lot, which shows a linear dependence on the channel length for only a single wafer (P07, 5 nm of Pt). For wafers P10–P12, with deposited Pt thicknesses of 15–25 nm, the automatic probe station returned zero for each TLM structure on nearly every die, indicating either a short or no contact at all. We can understand what happened by looking instead at wafer P09 (green curve), where a reasonable resistance is obtained only for the longer channels: it is likely that the structures with $L_{\text{channel}} \leq 4 \mu\text{m}$ were shorted by lateral encroachment of PtSi through the entire channel. As discussed in section 3.4, this behavior is to be expected after all the silicon in the contacts has been consumed. Alternatively, the absence of any signal on wafers where thicker Pt layers were deposited could be attributed to the lifting of PtSi due to poor adhesion to SiO_2 [2].

A few scribes from wafer P07 were then cooled down by master student Axel Leblanc. Shown in Fig. 4.4 (left) is a similar plot, this time repeated at both room temperature and 4 K for structures with slightly wider, but shorter contacts: RB0C1, RB0C2, RB0C1RE and RB0C2RE, with $W_{\text{contact}} = 1.25 \mu\text{m}$ and $L_{\text{contact}} = 0.35 \mu\text{m}$. Although the silicon channel has been doped with boron only up to 10^{15}cm^{-3} , we see that the channel sheet resistance (the slope in this graph) in fact drops from 870Ω to 635Ω after cooling to 4 K. Such a drop suggests degenerate doping ($\gtrsim 10^{18} \text{cm}^{-3}$); the resistance values are still at least an order of magnitude above that expected for real metals, so we can be confident that we are not just measuring a bar of PtSi. More consistent with expectation is that the contact resistance increases from $2R_c = 300 \Omega$ at 300 K to $2R_c = 470 \Omega$ after cooling, as thermionic emission over the Schottky barriers at the PtSi/Si interfaces becomes suppressed. Note that this contact resistance is not simply proportional to the surface area of the contact: since the transfer length is less than one micron, only the edge of the larger contacts measured for Fig. 4.4 actually contribute to the transmission.

Table 4.1: Lot 4 (D19S1723): A set of patterned wafers are prepared, with deposited Pt thicknesses varying from 5 to 25 nm. Same as the first half of Table 3.6 (without any doubles).

Lot 4	Wafer #				
	7	9	10	11	12
Patterned SOI wafer	•	•	•	•	•
Deposit <> nm Pt + TiN	5	10	15	20	25
RTA 500 °C 120s	•	•	•	•	•
Etch TiN	•	•	•	•	•
Etch Pt	•	•	•	•	•
XRD	•	•	•	•	•
Complete back-end	•	•	•	•	•

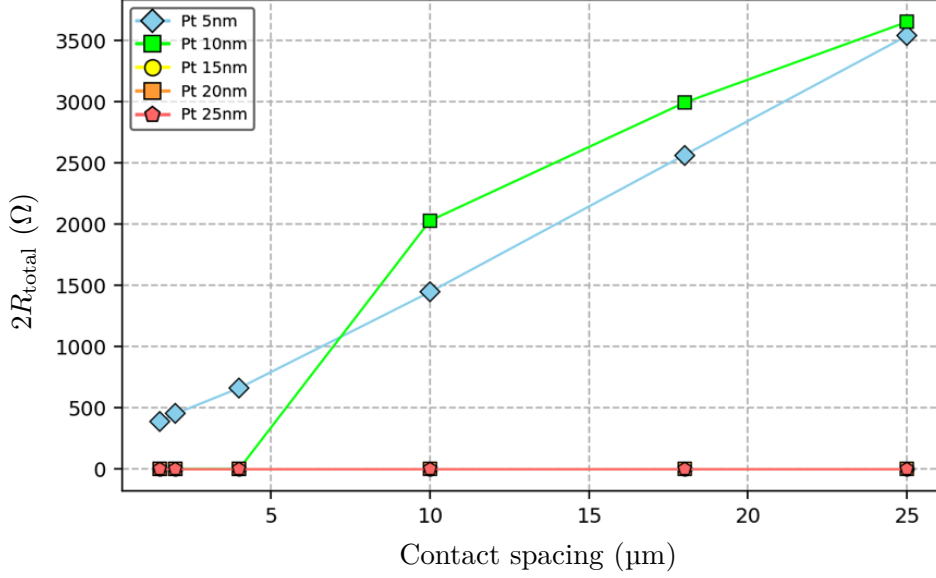


Figure 4.4: The resistances of the six TLM structures (see Fig. 3.26) on scribe RA0C1 were measured, which each have a contact width of $W = 1 \mu\text{m}$ and a contact length of $L = 5 \mu\text{m}$, but have varying channel lengths between these two contacts from $L_{\text{channel}} = 1.5 \mu\text{m}$ to $25 \mu\text{m}$. This scribe was tested on seven dice of each wafer, each square represents the average value obtained for a single channel length on a given wafer (5 nm = P07, 10 nm = P09, etc). The y -axis intercept for P07 corresponds to $2R_{\text{contact}} = 145(5) \Omega$, and an extrapolation of the linear regression gives a transfer length of $0.80(5) \mu\text{m}$.

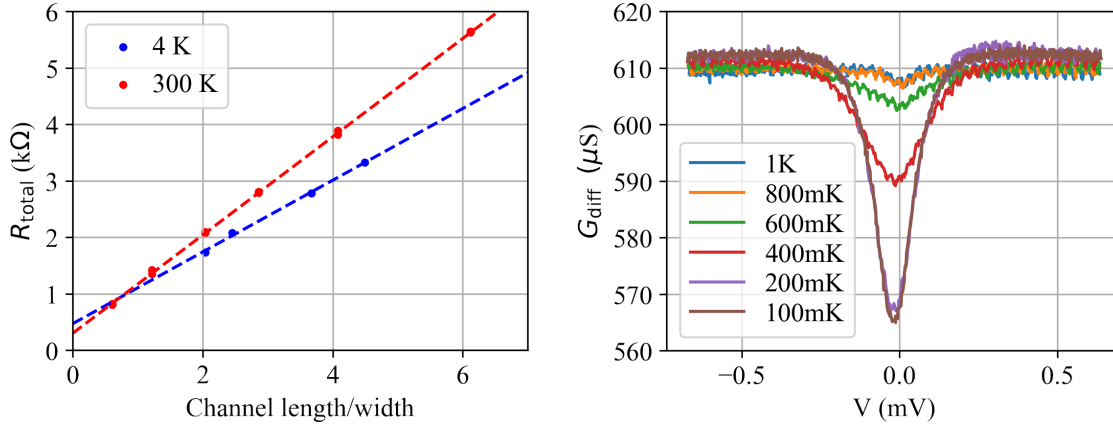


Figure 4.5: (Left) Scribes RB0C1, RB0C2, RB0C1RE and RB0C2RE, all with contacts $1.25 \mu\text{m}$ wide and $0.35 \mu\text{m}$ long, were measured on wafer P07, which had 5 nm Pt deposited (see Table 4.1). At room temperature (red curve) the slope indicates a silicon sheet resistance of 870Ω , which drops to 635Ω at 4 K (blue curve). The opposite trend is seen for the contact resistance, which increases from $2R_c = 300 \Omega$ at 300 K to $2R_c = 470 \Omega$ at 4 K. (Right) One of these structures, RB0C2RE5 (the fifth structure of scribe RB0C2RE, $L_{\text{channel}} = 25 \mu\text{m}$) was further cooled to below the superconducting critical temperature of PtSi. Shown is the differential conductance versus the voltage bias across the junction, for different temperatures from 100 mK to 1 K. It is clear that the transmission across the interface is partially suppressed at $E < \Delta$, although no clear feature of the superconducting gap can be observed. This is likely a combination of strong inelastic scattering ($\Gamma \approx 0.5\Delta \approx 16 \mu\text{eV}$) and a high electronic temperature inside the channel (note that there is no difference between the curves for 100 and 200 mK).

More insight into the PtSi/Si interface was gained by measuring the differential conductance of one of these junctions, RB0C2RE5 ($L_{\text{channel}} = 25\text{ }\mu\text{m}$), at different temperatures between 100 mK and 1 K, shown in the right of Fig. 4.5. Although a clear reduction in transmission can be seen around $|eV| < 2\Delta = 31\text{ mV}$, indicative of a low-transparency S/Sm interface, no sharp coherence peaks are observed. This suggests that the inelastic scattering length is shorter than $25\text{ }\mu\text{m}$, which is not unreasonable for a dirty material [3].

4.3 Proximity effect in a PtSi SBMOSFET

4.3.1 Description of the samples

The measurements in the coming sections were performed on transistors from a wafer provided by Laurie Calvet, manufactured in 1997–1998 at National Semiconductor in Santa Clara, California by Chinlee Wang and John Snyder [4, 5], with further details provided in Appendix B. A schematic drawing and a cross-section TEM micrograph of a representative device with a channel length of 27 nm is shown in Fig. 4.6. These Schottky-barrier MOSFETs were originally designed to improve the scaling of transistors [2]; introducing a Schottky barrier further away from the gate electrode in devices without encroachment prevents short-channel effects, while moving the contact edges and associated Schottky barrier into the channel allows for barrier modulation in addition to charge accumulation, and thus a steeper sub-threshold slope. These devices find a new application in JoFETs, with the advantage of short channels on the order of a few tens of nanometers thanks to the encroachment, and direct gate control of the superconductor/channel interface transparency.

The bulk silicon was implanted with boron to reach a p-doping that ranges from $5 \times 10^{15} \text{ cm}^{-3}$ in the bulk to 10^{19} cm^{-3} right below the 3.5 nm gate oxide [4]. After gate and spacer definition, 29(3) nm of Pt was DC sputter deposited onto boron-implanted polycrystalline silicon [2]. An estimated 55(6) nm layer of PtSi was then formed by a one-hour thermal processing at 450 °C under a N_2 atmosphere, after which unreacted Pt was removed with aqua regia (4:3:1 $\text{H}_2\text{O}:\text{HCl}:\text{HNO}_3$) at 85 °C for 10 min. Note that there was no buried oxide underneath the silicon in the contacts and channel, such that silicidation occurred isotropically, with no concern of a lateral speed-up underneath the gate as discussed in section 3.4. Nonetheless, a large variation in channel length was observed, with a reduction of up to 330 nm in channel length from the target [4], and a standard deviation of 3 nm between devices on the same quad [2]. Additional variations across the wafer are due to nonuniform Pt deposition, leading to a range in channel lengths from 30 to 70 nm on the shortest devices. In the discussions below, we will assume an average length for these transistors of 50 nm. The resulting PtSi layer had a resistivity of $40 \mu\Omega \text{ cm}$ [2], a factor $1.6\times$ higher than what would be expected by extrapolating from our own experiments on single-anneal thick PtSi layers (see Fig. 3.25). This had no negative impact on the superconductivity, as a relatively high critical temperature of 1.03 K was observed on these devices (see Fig. 4.14).

Transistors usually have counterdoping with opposite sign in the channel to suppress current in the OFF state. In these devices however, both the contacts and the channel were implanted with boron, leading to a high I_{off} at 300 K that would make them unsuitable for room-temperature logic operation. Such an OFF current disappears when the device is cooled down, as thermionic emission is suppressed and dopants in the channel freeze out. At the operational temperature of a transmon qubit (preferably below 50 mK), any transport across the interface will occur by quantum tunneling, at which point it becomes crucial to have a low potential barrier. Assuming a triangular barrier shape, the tunneling probability due to field emission will be suppressed super-exponentially as $\exp(-C\phi^{3/2}/E)$ (where E is the electric field) [6]². The low Schottky barrier of 0.16 eV [2] due to the all-p doping³ is thus essential for proper JoFET operation. For comparison,

²Thermionic emission, which scales as $\exp(-\phi/k_{\text{B}}T)$, is irrelevant at these temperatures.

³The cited thesis mentions that the low value of 0.16 eV, as compared to 0.19–0.25 eV in the literature, may be due to inaccuracies in the temperature measurement. However, a later report [7] addressed a

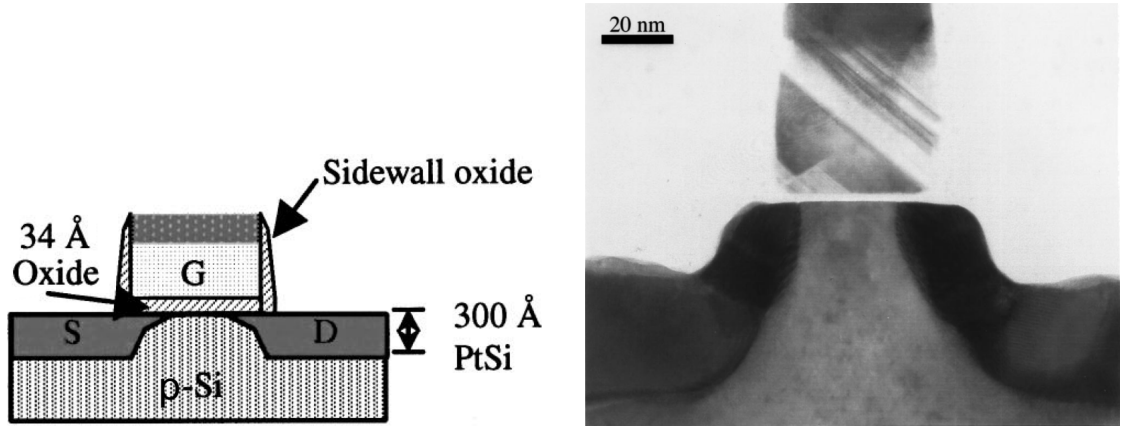


Figure 4.6: (Left) A schematic of the SB-MOSFETs that were measured. Adapted from Ref. 5. (Right) An XTEM of a device with a 27 nm channel length, reproduced from Ref. 4. Lateral encroachment occurred underneath the spacers during silicidation, after which a sidewall etch cut into the PtSi vertically, resulting in the upward “bend” in the silicide contacts.

an n-type channel would have led to a Schottky barrier of 0.81–0.88 eV [8]. A downside of the intentional channel doping is that reduced mobility was observed due to the high impurity concentration [2].

Measurements of the gate capacitance with gate voltage indicated a high density of interface states, which were attributed to contamination during gate oxidation and silicon deposition [2]. This is of special concern for the integration of these devices in superconducting qubits; although circuits in the transmon limit ($E_J \gg E_C$) are less prone to decoherence through the coupling of the qubit degree of freedom to these two-level systems [9] than other designs [10], fluctuations in Josephson coupling strength due to their charging will complicate qubit operation.

4.3.2 Room-temperature tests

An automatic probe station was used to perform rapid tests on a series of devices at room temperature. The station switched between devices by moving the wafer underneath the probes, allowing for a large number of transistors to be characterized simply by providing a spreadsheet of coordinates. Shown in Figs. 4.7 and 4.8 are measured currents through gate, substrate, drain and source probes for the four types of device most often observed. As described in the figure captions, such automatically generated graphs are useful to identify working or broken devices, and helped select transistors that would then be measured at low temperature.

It should be noted that a current from gate to substrate on the connected devices does not per se imply a broken oxide, since this path was shunted by a parallel diode (see Fig. B.4 in the appendices). This current will disappear at low temperatures as the substrate freezes out [11, 12]. Nor does a large source-drain current linear in V_d necessarily indicate that the channel has been fully consumed by silicidation; such large currents can be expected for the shortest devices due to the low Schottky barrier, and a

similar mismatch with the same “commonly accepted numbers” by fabricating a range of devices to remove short-channel effects and account for lateral transport and tunneling, and still found that only values as low as 0.14–0.145 eV matched the data.

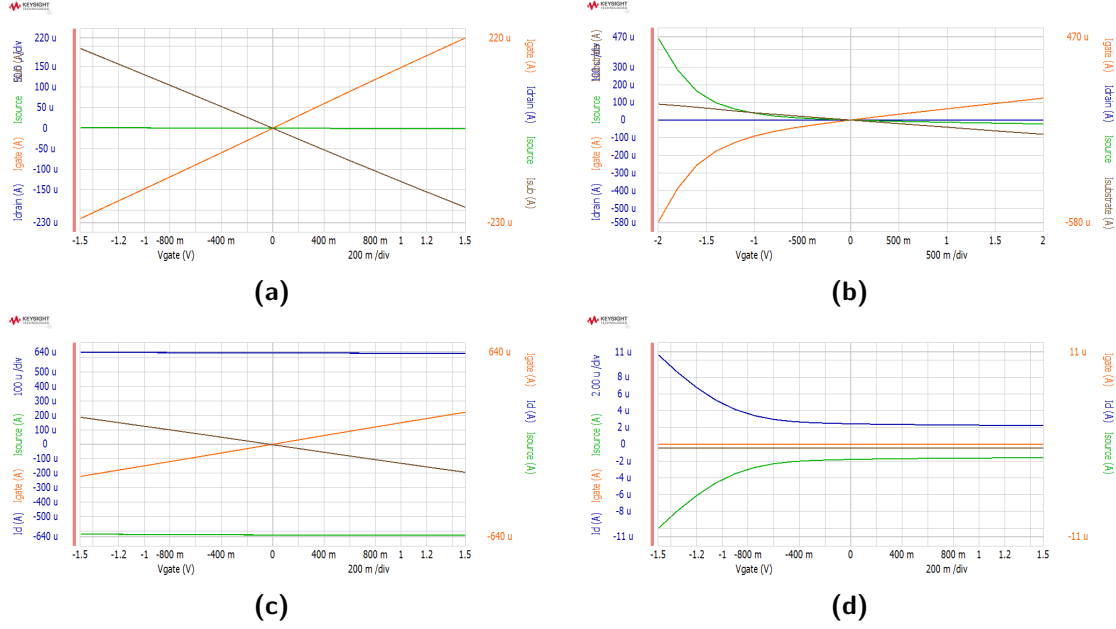


Figure 4.7: Currents through the **gate**, **substrate**, **drain** and **source** at $V_d = 1$ mV as V_g is varied, for different devices. (a) Leakage from gate to substrate, no field effect. (b) Field-enhanced leakage from gate to source. (c) Shorted from source to drain. (d) Field effect, part of the current injected at the drain flows through the substrate (as expected).

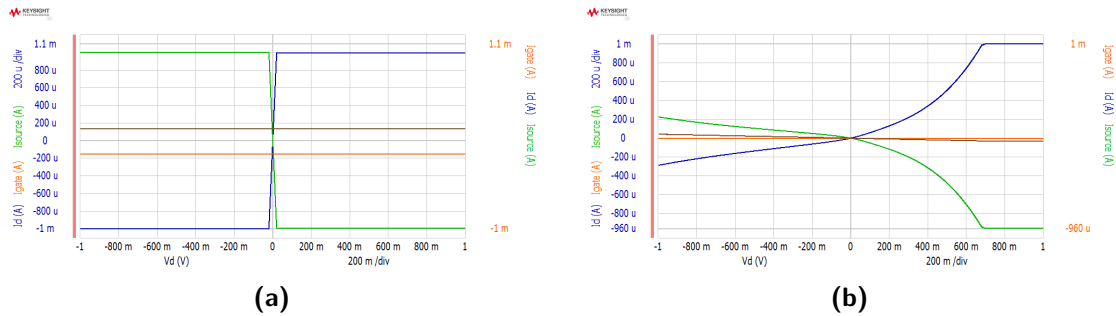


Figure 4.8: Currents through the **gate**, **substrate**, **drain** and **source** at $V_g = -1$ V as V_d is varied, for different devices. (a) Short from source to drain. (b) Nonlinear source-drain current due to the presence of a Schottky barrier (saturation of the instrument at $I_d = 1$ mA).

good field effect is often observed at cryogenic temperatures as soon as dopants have been frozen out and thermionic emission is reduced. As detailed graphically in appendix B, working devices were cut from the wafer, glued onto a sample holder and wire bonded for cryogenic measurements.

4.3.3 Cryogenic measurement setup

Shown in Fig. 4.9 is a schematic of the measurement setup used in the experiments discussed below. The channel of the transistor was biased with a DC voltage (V_{DC}), together with an AC oscillation from a lock-in amplifier (V_{AC}). These were connected using three resistors, $R_{\text{DC}} = 10 \text{ k}\Omega$, $R_{\text{AC}} = 100 \text{ k}\Omega$ and a shunt to the ground $R_{\text{ground}} = 10 \Omega$, effectively dividing the DC and AC voltages by 10^3 and 10^4 , respectively, to be able to use voltage steps smaller than the resolution of the voltage source. The resulting current was amplified using a trans-impedance amplifier that delivered 10^7 V/A , while a differential amplifier was used to measure the voltage across the device.

Low-pass RC filters were anchored to the mixing chamber at 30 mK, ensuring that their impedance did not vary during the measurement. The sample holder, on the other hand, had only a weak thermal coupling to the mixing chamber, such that it could be heated from 38 mK to 2 K.

Since the resistance of the filters is much greater than that of the shunt resistance to the ground ($2R_{\text{filter}} \approx 40 \text{ k}\Omega \gg R_{\text{ground}} = 10 \Omega$), we can safely assume that the effective applied voltage at the top of the fridge does not depend on the resistance of the device under test (DUT). In our case, that means that the AC and DC voltages at the second line entering the fridge in Fig. 4.9 are

$$\begin{aligned} V_{\text{applied,AC}} &= \alpha \frac{R_{\text{ground}}}{R_{\text{AC}}} V_{\text{AC,source}} \approx 10^{-4} V_{\text{AC,source}} \quad , \text{ and} \\ V_{\text{applied,DC}} &= \beta \frac{R_{\text{ground}}}{R_{\text{DC}}} V_{\text{DC,source}} \approx 10^{-3} V_{\text{DC,source}}, \end{aligned} \quad (4.2)$$

where $\alpha = 0.88$ and $\beta = 1.00$ are fitting parameters to take into account any instrumental inaccuracies. However, this does not equal the actual bias across the device itself, especially when R_{DUT} becomes smaller than $|Z_{\text{filters}}|$, which will be the case when the device becomes passing at large negative gate voltages. To take this into account, the voltage drop $R_{\text{filter}} I_{\text{DC}}$ on each side of the device is subtracted,

$$V_{\text{DC,DUT}} = V_{\text{DC,applied}} - 2R_{\text{filter}} I_{\text{DC}}. \quad (4.3)$$

Similarly, the filter impedance will limit the measured differential conductance, which is corrected for by calculating [13]

$$G_{\text{diff,DUT}} = \left| \left(\frac{\alpha}{G_{\text{diff}}} - 2Z_{\text{filter}} \right)^{-1} \right|, \quad \text{where} \quad Z_{\text{filter}} = \frac{R_{\text{filter}}}{1 + iR_{\text{filter}}C_{\text{filter}}\omega_{\text{lock-in}}}. \quad (4.4)$$

The two lines to the top operational amplifier in Fig. 4.9 that measured V_{DUT} directly were disconnected once the resistance and capacitance of the filters were extracted, to avoid introducing additional noise. In all the figures that follow, the drain voltage V_{d} will be bias across the device as defined in eq. (4.3), and the differential conductance G_{diff} will be that of the device only, as in eq (4.4).

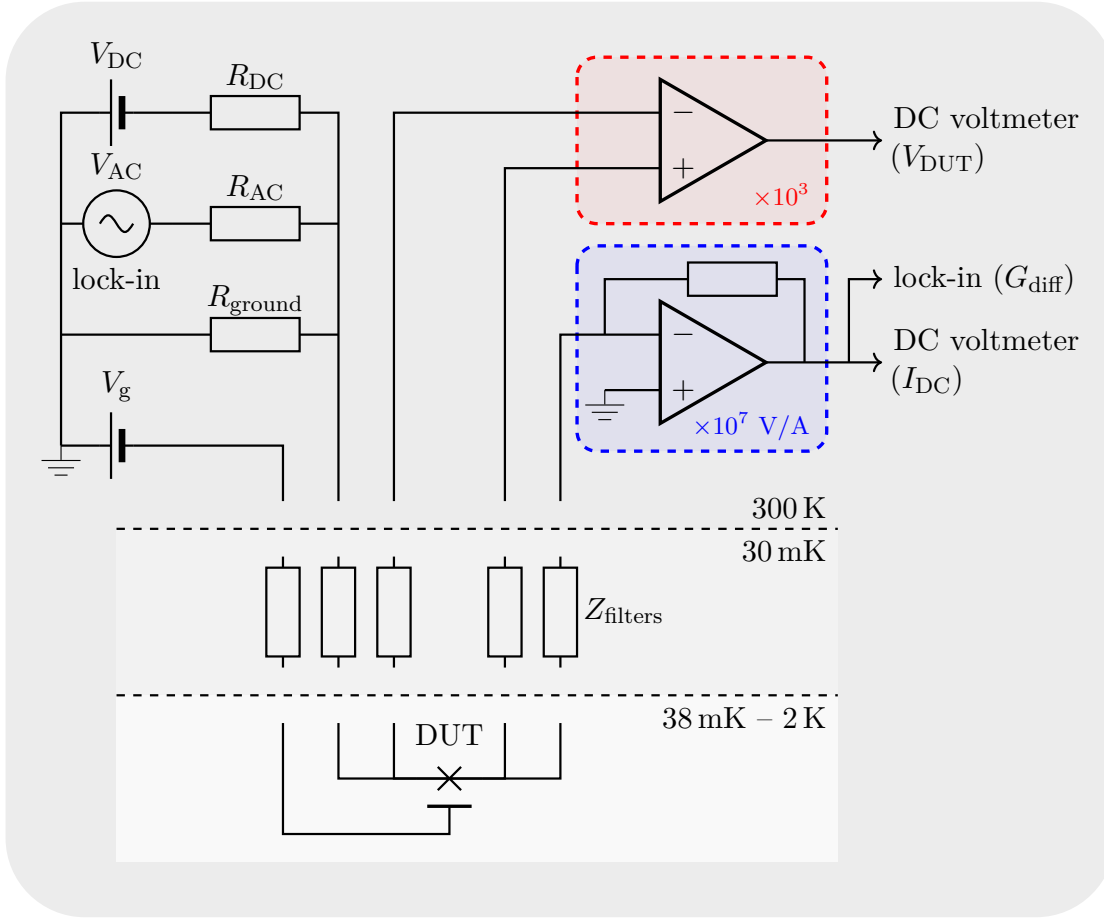


Figure 4.9: Schematic of the measurement setup.

Table 4.2: An overview of the devices where a gate effect was detected at low temperature, indicated in green those that showed a ZBCP.

Device	W (μm)	L (nm)	Scan in V_g		Scan at fixed V_g	
			$H = 0$	$H > H_0$	$H, V_g =$	$T, V_g =$
D61D1*	?	?	$[-4.0, -1.45]$	×	×	$-3.5, -3, -2.5, -2.25$
D61D2*	?	?	$[-3.5, -1.56]$	$[-1.5, -1.4926]$	-1.5	×
D61D4*	2.5	50	$[-4.15, -2.0]$	×	×	×
D61D5*	?	?	$[-3.0, -1.87]$	×	×	×
D63D1*	?	?	$[-3.5, -1.6]$	×	×	$-3.5, -2$
D63D3*	2.5	50	$[-4.9, -1.0]$	$[-4.5, -1]$	$-4.5, -2.7$	$-4.5, -4.0, -2.7$
D63D4*	?	?	$[-4.0, -1.0]$ [‡]	×	-3.9	×
D73D4*	2.5	50	$[-4.4, -2.4]$	$[-2.5, -2.494]$	×	×
D84D1 [†]	?	?	$[-4.5, -2.1]$	×	×	×
D84D3 [†]	?	?	$[-3.1, -2.05]$	×	×	×
D84D4 [†]	?	?	$[-5, -4.7]$	×	×	×
D84D5 [†]	?	?	$[-3.5, 0]$	×	×	×

*Measured in *Christophe's fridge*.

[†]Measured in the fridge *Bigoudène*.

[‡]Offset in magnetic field detected, $H \neq 0$.

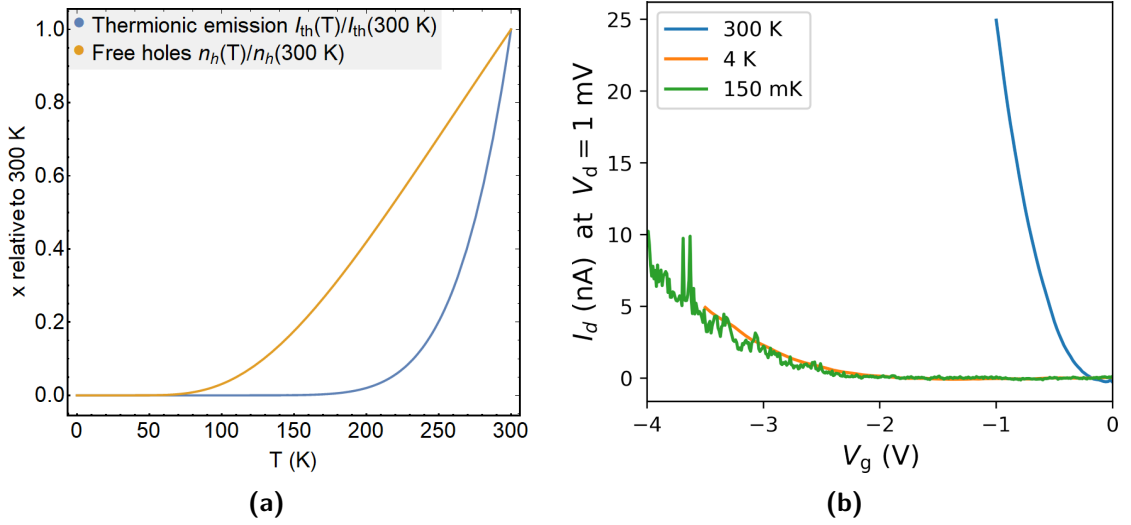


Figure 4.10: (a) Theoretical estimates of the thermal suppression of the thermionic emission (blue) and free carrier concentration (orange), after equations (4.5) and (4.8). (b) Device D84D4 conducts at negative gate voltages. Cooling the device from 300 to 4 K suppresses thermionic emission and freezes out dopant carriers, shifting the threshold voltage down. At lower temperatures of 150 mK, reproducible fluctuations appear due to charging effects.

4.3.4 Demonstration of the field effect at low temperature

As its name suggests, thermionic emission is activated by heat. This contribution to the current, due to the hot tail of the electron energy distribution that extends above the potential barrier, scales as [14–16]

$$I_{th} \propto T^2 e^{-\phi_{Schottky}/k_B T}, \quad (4.5)$$

where we will use $\phi_{Schottky} = 0.16\text{ eV}$ for our devices [2]. When a bias V_d is applied, the net current across a single barrier will then be [7, 17]

$$I_{th} \propto T^2 e^{-\phi_{Schottky}/k_B T} \left(e^{-qV_d/k_B T} - 1 \right). \quad (4.6)$$

In the case of pure thermionic emission ($qV_d \ll k_B T$, no field emission), the voltage drop is independent of the bias direction and (using a normalized $V' = qV/k_B T$),

$$I(\leftarrow\rightarrow) = \left(e^{-V'/2} - 1 \right) - \left(e^{+V'/2} - 1 \right) \stackrel{V' \ll 1}{\approx} e^{-V'} - 1 = I(\leftarrow), \quad (4.7)$$

such that we can approximate the SBMOS channel as a single diode. The key thing to note here is that the effective resistance of the two Schottky barriers in the channel sharply increases at low temperatures.

The channel itself is made of boron-doped⁴ Si, where the product of the free electrons and holes remains equal to the square of the intrinsic carrier density of pure silicon, as discussed in section 2.4.1. This means that in the absence of an accumulating electrostatic field, the number of majority carriers (holes) will be [18]

$$p \propto e^{(E_V - E_F)/k_B T}, \quad (4.8)$$

⁴Boron, a column to the left of silicon in the periodic table, has only 3 electrons in the outer shell.

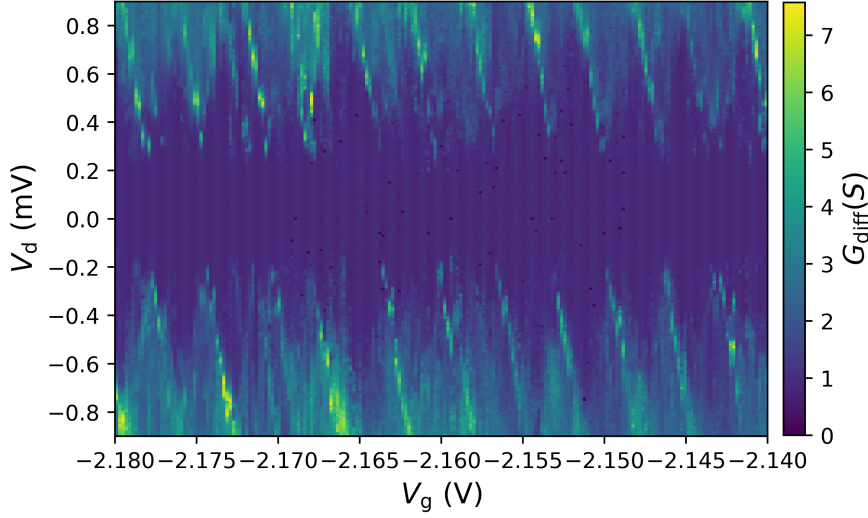


Figure 4.11: Data from device D63D4. Coulomb diamonds are most clearly visible in the differential conductance at small absolute gate voltages, when the only mode of transport is by resonantly tunneling through the charging centers.

where the negative E_V , measured from the gate-dependent Fermi level E_F , is the relative energy of the valence band edge. By introducing easily excited acceptors with energy levels at E_A , the boron pins the Fermi level at a distance of $E_A - E_V$ from the valence band, such that we can equivalently write

$$p \propto e^{(E_V - E_A)/k_B T}. \quad (4.9)$$

The intrinsic (i.e. field-independent) acceptor energy $E_A - E_V$ of boron depends on temperature, and decreases from around 62 to 45 meV when cooled from room temperature to anywhere between 100 K and 0 K [19]. At low temperature, holes introduced by the missing bond of boron atoms in the channel will be frozen out [20, 21], and accumulation can only be achieved by either ionizing these acceptors with the gate electrostatic field [22, 23], or by exciting carriers from metallic atoms in the contacts across the Schottky barrier.

As the device is cooled down, thermionic emission and dopant excitation will both be suppressed, as shown in Fig. 4.10a, leading to larger threshold voltages at lower temperatures, illustrated in Fig. 4.10b. In this last figure, an additional feature can be observed at 150 mK: reproducible peaks in conductance at fixed values of V_g . These can be better understood by scanning also the source-drain bias, as is done for a different device in Fig. 4.11. These Coulomb diamonds, spaced unevenly by around 4 or 5 meV, are likely⁵ not due to the charging of the channel itself. Since this device has a channel of 2.5 μm by 50 nm and a 3.5 nm gate oxide, we would expect its charging energy to be two orders of magnitude smaller,

$$E_{C,\text{channel}} = \varepsilon_0 \kappa_{\text{SiO}_2} \frac{A_{\text{channel}}}{d_{\text{oxide}}} \approx 65 \mu\text{eV}. \quad (4.10)$$

Instead, these should probably be attributed to the charging of surface states in the oxide detected earlier by capacitance measurements [2] (see section 4.3.1).

⁵Unless something went wrong during fabrication and the wide channel is effectively broken up into many smaller parts due to variations in gate length.

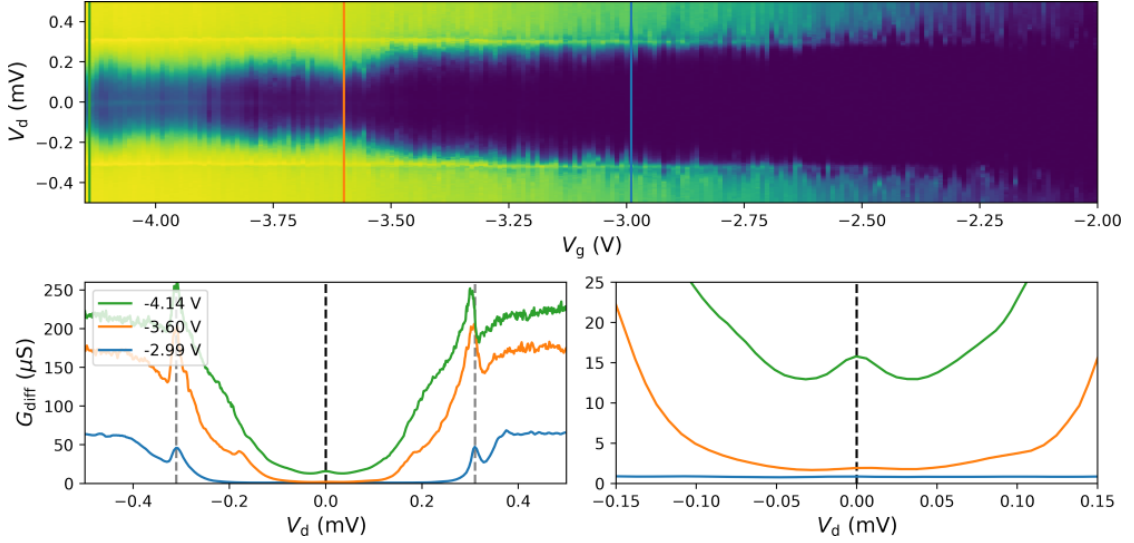


Figure 4.12: Data from device D61D4. Both source-drain and gate voltage were scanned while the differential conductance was measured. **(Top)** As larger negative voltages are applied to the gate, we see first the appearance of coherence peaks at $V_d = \pm 0.31$ mV around $V_g \approx -2.5$ V, then large sub-gap conductance from $V_g \approx -3.5$ V, and finally a clear zero-bias conductance peak at $V_g \lesssim -3.75$ V. **(Bottom, left)** Selected line curves from the density plot shown above. **(Bottom, right)** The same curves, in a smaller V_d range.

4.3.5 Evidence of superconductivity

An additional feature can be noticed in fig. 4.11: the diamonds are spaced vertically by 0.62 mV. This gap becomes clearer at more negative gate voltages, as can be seen in Fig. 4.12, and has characteristic coherence peaks. If this is due to superconductivity in the source and drain, then we expect the observed value to equal four times the superconducting gap, as explained graphically in Fig. 4.13. Since its magnitude corresponds to [24, 25]

$$0.62 \text{ meV} = 4\Delta_0 = 4\pi e^{-\gamma} k_B T_c, \quad \text{and thus} \quad T_c \approx 1.02 \text{ K}, \quad (4.11)$$

where $\gamma \approx 0.5772$ is Euler's constant [26], we can be confident that this is due to superconductivity in the PtSi source and drain. To our knowledge, no other superconductors were present in the contacts or vias.

A fit of the peak positions versus temperature to that predicted by BCS theory, shown in Fig. 4.14 for a measurement performed on D63D3, confirms that PtSi is a weakly coupled superconductor (such that $\Delta \approx 1.76 k_B T_c$ is valid), with a critical temperature of 1.03 K. This is slightly higher than what we found in section 3.4 for films of similar thickness, and implies a better quality of the PtSi obtained at National Semiconductor.

Additionally, at larger negative gate voltages in Fig. 4.12, we see the appearance of sub-gap conductance. Since no quasiparticle states are available in that energy range on either side of the channel, this must be due to a different mode of transport. When Andreev [29] solved the Gor'kov equations [30], he found that an incident particle with an energy below the gap could be reflected as a hole, and vice versa, resulting in the effective transmission of regular quasiparticles as superconducting Cooper pairs, thus contributing to current within the gap. We now refer to this process as Andreev reflection, which has the additional feature that the paired electron and hole inside the normal material

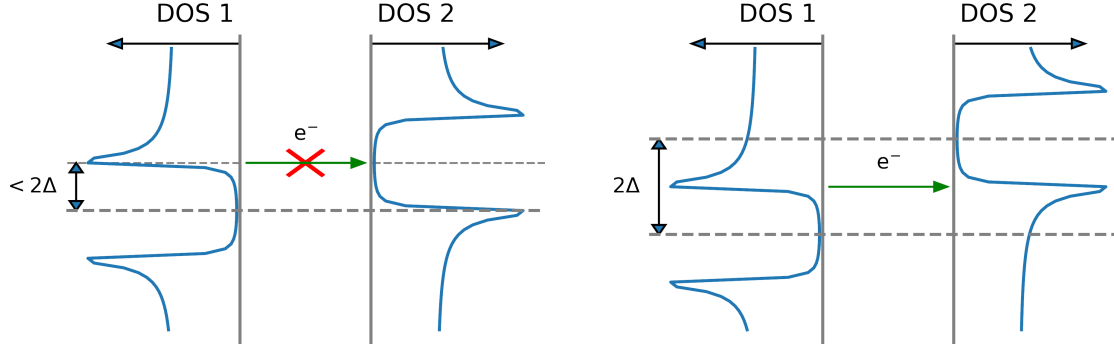


Figure 4.13: (Left) If the densities of states in the superconducting source and drain are biased by less than twice the gap, transport at the Fermi level will be prevented by the absence of available states. Since the electronic temperature in the device is far below Δ/k_B , no thermionic emission to higher-energy states will occur either. (Right) Single quasiparticles will be able to enter the superconductor when $eV_d \geq 2\Delta$, where they either contribute to the supercurrent immediately when $eV_d \approx 2\Delta$, or relax into the condensate within $\sim 10^{-9}$ s [27, 28].

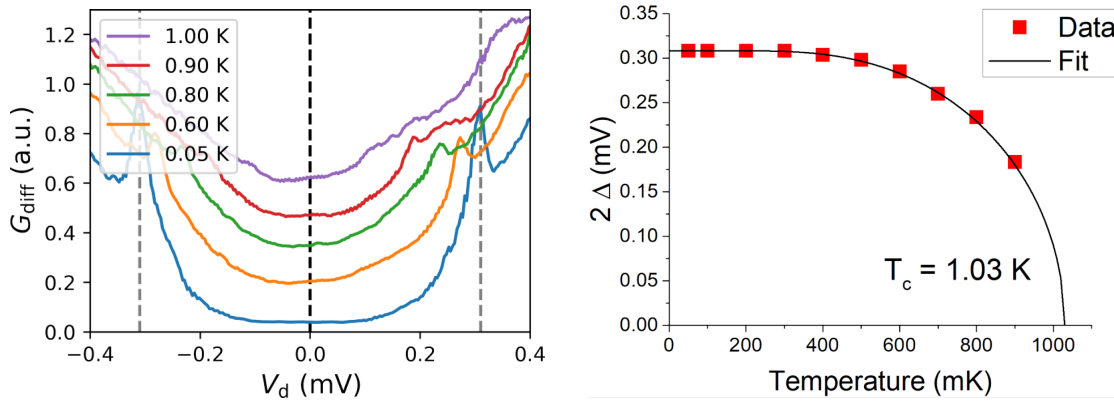


Figure 4.14: Data from device D63D3. (Left) Differential conductance at $V_g = -2.7$ V, at selected temperatures between 0.05 and 1.00 K. The coherence peaks (indicated by vertical dashed lines for $T = 0.05$ K) move towards smaller values as the device is warmed up. (Right) A fit of the peak positions from the graph on the left, to the expected BCS effective gap $\Delta_{\text{BCS}}(T)$.

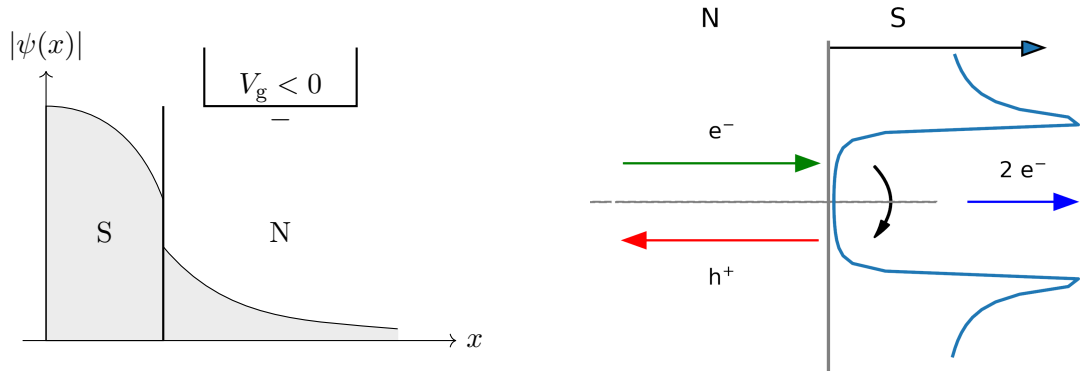


Figure 4.15: **(Left)** The sub-gap conductance that appears at negative gate voltages can be explained by the proximity effect: the “leaking” of superconductivity from the superconductor (S) into the normal material (N), here represented by the Ginzburg-Landau order parameter $\psi(x)$ [33–35]. **(Right)** Cartoon illustration of Andreev reflection: an incoming electron is reflected as a hole, introducing a Cooper pair inside the superconductor. The blue curve on the right represents the superconducting DOS.

briefly maintain their coherence. This is associated with a minigap in the density of states (DOS) inside the normal material [31] that decays over a distance that depends on the energy scale E of the decohering processes (see section 2.3.4),

$$\xi_{\text{ballistic}} = \frac{\hbar v_F}{2\pi E}, \quad \xi_{\text{diffusive}} = \sqrt{\frac{\hbar D}{2\pi E}}. \quad (4.12)$$

This opening of a gap close to the boundary, which in the case of a perfectly transparent interface smoothly connects to the gap inside the superconductor, is commonly referred to as the proximity effect [32]. The process is illustrated in Fig. 4.15, and explains the sub-gap conductance observed in Fig. 4.12.

Since the differential conductance of a single S/N interface in the absence of any scattering is exponentially suppressed at $V = 0$ [36], the zero-bias conductance peak (ZBCP) cannot be explained by Andreev reflection alone. Two different mechanisms that contribute to zero-bias conductance are relevant in the current situation. Either transmission is enhanced across each interface individually, which can be done by successive reflections inside the normal material [37, 38], or a zero-resistance channel is created by coherent transport between the two interfaces [39–41]. These two explanations for the ZBCP are shown in Figs. 4.16 and 4.17.

Scans in temperature and field were also used to investigate the nature of the ZBCP observed below $V_g = -3.75$ V. As can be seen in Fig. 4.18, this peak diminishes but survives up to $T = T_c$, while it is already suppressed with magnetic fields far below $H = H_{c,2}$. This is consistent with both reflectionless tunneling and coherent transport. While small magnetic fields are not expected to directly impact the proximity effect at each individual S/N interface [29], the total phase difference acquired by an electron-hole pair can be multiplied by repeated reflections [37, 38]. In the case of coherent transport, a single flux quantum would suffice to suppress an inhomogeneous supercurrent across the channel. Taking $W = 2.5 \mu\text{m}$ and $L = 50 \text{ nm}$ for the device considered in Fig. 4.18, we arrive at an equivalent field of

$$\frac{\Phi_0}{A_{\text{junction}}} = \frac{h/2e}{2.5 \mu\text{m} \times 50 \text{ nm}} = 17 \text{ mT}, \quad (4.13)$$

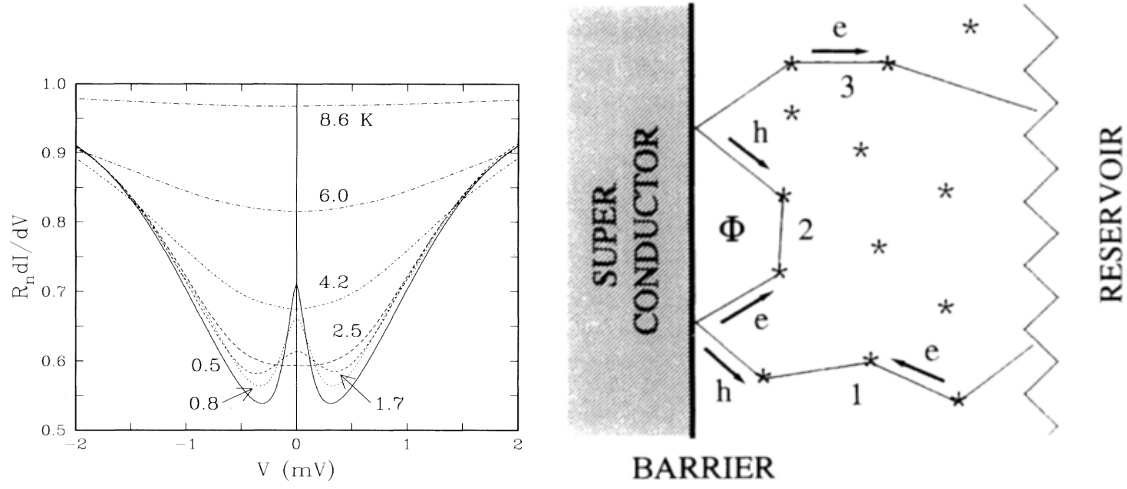


Figure 4.16: (Left) Kastalsky et al observed a zero-bias peak in an SN junction — i.e. a junction *without* a second superconducting lead, and attributed it to a pair current [37]. (Right) Four months later, a team in Groningen submitted their explanation to the same journal [38]: the conductance could be enhanced by interacting with the interface multiple times before coherence is lost, by scattering many times inside the dirty normal material. We now refer to this process as “reflectionless tunneling”.

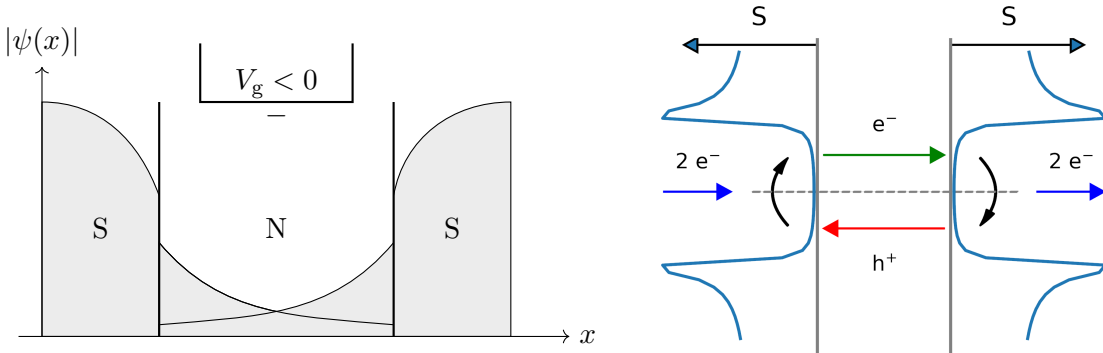


Figure 4.17: (Left) As explained by de Gennes [39], a Josephson coupling can be established when the induced pair potentials from the two superconductors overlap inside the normal material. (Right) A Cooper pair entering the normal metal is transferred as an electron-hole pair, before forming a new Cooper pair on the other side. Note that the hole is moving in the opposite direction: the hole component of the pair travels as an electron, but “backward in time” [42, 43]. To be clear: in terms of a single reflection, the hole traces back the electron path *after* the electron has been absorbed by the superconductor [44].

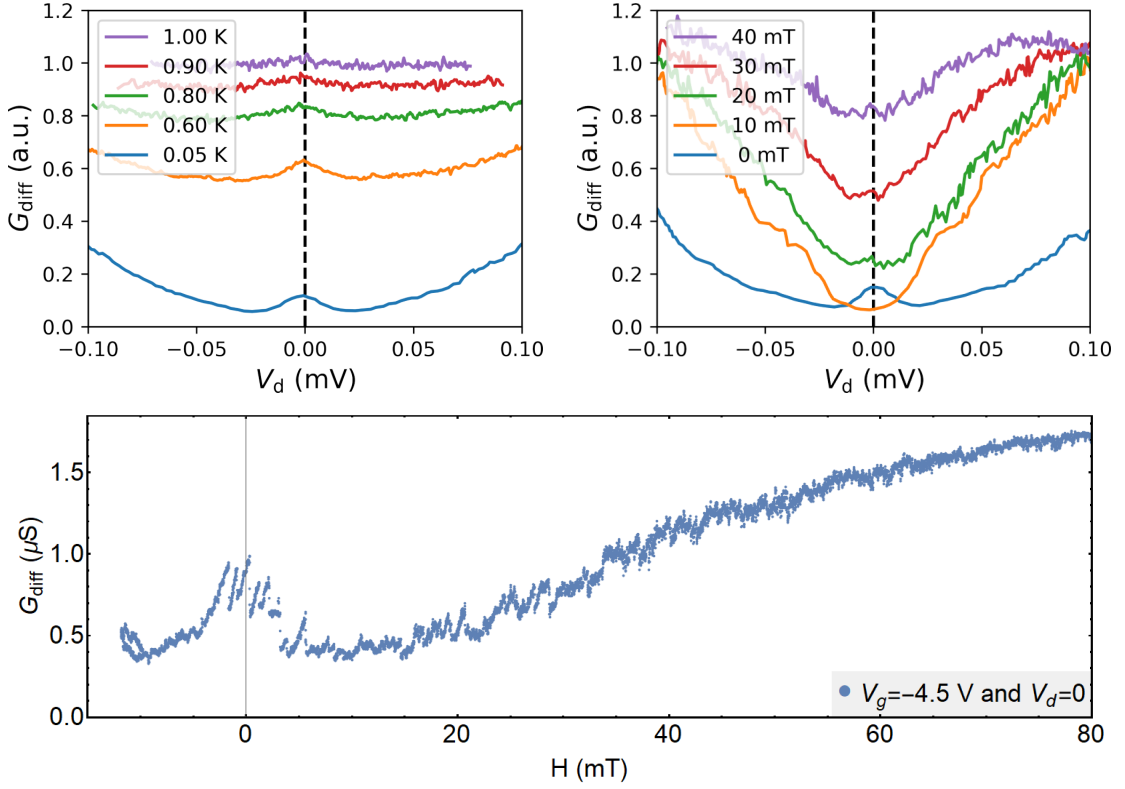


Figure 4.18: Data from device D63D3. **(Top left)** The differential conductance is plotted for a small V_d range within the gap, at $V_g = -4.0$ V. The zero-bias conduction peak disappears as the gap closes. **(Top right)** The differential conductance is shown for a range of applied magnetic field strengths, this time at $V_g = -4.5$ V. The critical field for thin films of PtSi (a type-II superconductor) was estimated to be 84 mT [45]. **(Bottom)** The differential conductance, again at $V_g = -4.5$ V, but now at fixed $V_d = 0$ mV versus the applied magnetic field. The zero-bias conductance peak is suppressed at around $H \approx 10$ mT.

in the right ballpark for suppression above 10 mT. Note that the channel length is not exactly known as discussed in section 4.3.1, and that flux focusing would increase the effective field in the channel. However, direct evidence of a supercurrent through the channel has not yet been observed. A back-of-the-envelope calculation, multiplying width (V) and height ($S=A/V$) of the ZBCP, suggests that such a measurement needs to be sensitive to 10^{-10} A if performed with the same setup, while improving noise filtering may further increase the critical current itself.

4.3.6 Gate modulation of the proximity effect

It is clear from Fig. 4.12 that the behavior of the device changes with gate voltage. We would like to make this more precise; find out whether the transport is limited by the channel or the transparency of the interfaces, how the Schottky barrier impacts that transparency, and then get an idea of what would need to be changed in the fabrication to make these transistors suitable for integration in transmons. Central to this analysis is the characterization of the interface, for which we will rely on the methodology introduced by Blonder, Tinkham and Klapwijk (BTK).

In their seminal 1982 paper [46], BTK provided, among many other things, a simple

method for extracting the interface barrier strength⁶ Z from the I - V characteristics of an S/N junction. To extend this to our case, we will assume that our S/N/S junction is symmetric, such that exactly half the voltage drop occurs on either side, and treat the system as having a single interface with twice the resistance and half the bias. We refer to the probability of transmitting a particle as the *transparency* of such a barrier, which at an S/N interface of course depends strongly on its energy — recall the blocked quasiparticles at $eV < \Delta$ in Fig. 4.13. We therefore cannot just put a single number on the transparency for superconducting transport, and instead we customarily state its value in the normal state, which we call Γ (C in BTK notation),

$$\Gamma = \frac{1}{Z^2 + 1} : \quad \lim_{Z \rightarrow 0} \Gamma = 1, \quad \lim_{Z \rightarrow \infty} \Gamma = 0. \quad (4.14)$$

The Josephson effect is of course due to the transfer of Cooper pairs around the Fermi level ($E \ll \Delta$), and not to the transfer of normal electrons outside the gap, so we may gain some insight by discussing also the probability that Andreev reflection occurs. This, in turn, is given by BTK's coefficient A ,

$$\begin{aligned} \lim_{E \rightarrow 0} A(Z) &= \lim_{E \rightarrow 0} \frac{\Delta^2}{E^2 + (\Delta^2 - E^2)(1 + 2Z^2)^2} = \frac{1}{(1 + 2Z^2)^2}, \\ \lim_{E \rightarrow 0} A(\Gamma) &= \left(1 + \left|\frac{2}{\Gamma} - 2\right|\right)^{-2}. \end{aligned} \quad (4.15)$$

Shown in Fig. 4.19 are the relationships between A at $V_d = 0$ and Γ , and the barrier strength Z .

In the simpler case (not ours) where the normal-state density of states has no additional features, we can at zero temperature express the current through an interface as

$$I_{\text{NS}}(T = 0, eV) = \frac{1 + Z^2}{eR_N} \int_0^{eV} [1 + A(E, Z) - B(E, Z)] dE, \quad (4.16)$$

which is plotted in the range of $0 \leq eV \leq 1.2\Delta$ for a selection of Z values in Fig. 4.20a. This current $I_{\text{NS}}(V)$ has a slope that always tends to $1/R_N$ at large voltages, and has an offset from V/R_N that depends only on Z , allowing one to quickly extract the transparency of the interface either by drawing a tangent that intersects with the I axis, or by subtracting V/R_N to get the excess current,

$$\begin{aligned} I_{\text{excess}} &= (I_{\text{NS}} - I_{\text{NN}})|_{eV \gg \Delta} \\ &= \frac{1 + Z^2}{eR_N} \int_0^{eV} \left[A(E, Z) - B(E, Z) + \frac{Z^2}{1 + Z^2} \right] dE. \end{aligned} \quad (4.17)$$

This excess current is plotted as a function of Z in **red** in Fig. 4.20b.

This method of extracting Z has a number of complications. In short, three conditions need to be met in order to reliably extract the interface transparency from the excess current:

⁶The model assumes a delta function potential $H\delta(x)$ at the interface, that collects all possible causes of normal (i.e. non-Andreev) reflection: Schottky barrier, mismatch in effective electron mass or lattice parameter, grain boundaries, oxides etc.

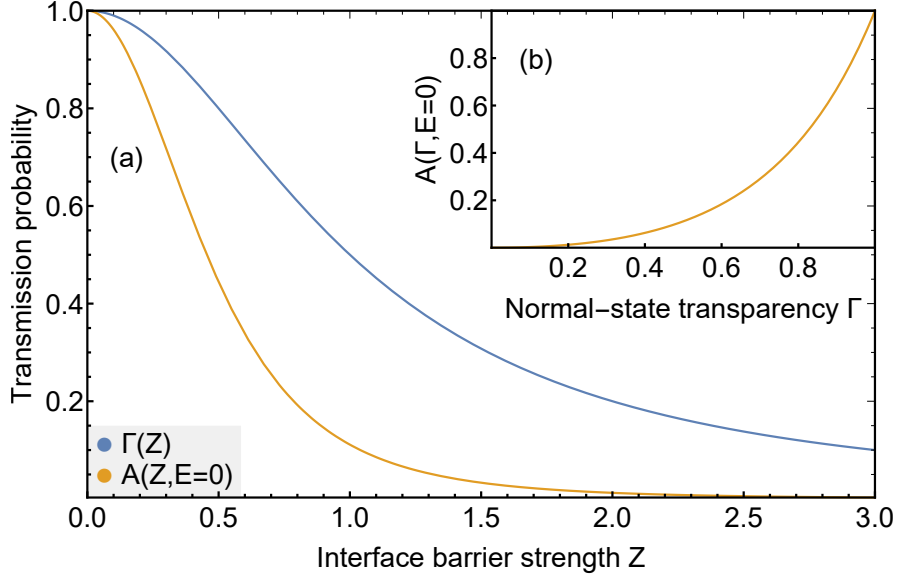


Figure 4.19: (a) The parameters Γ and A are both relevant in describing the interface. The normal-state transparency Γ gives the probability of regular electron transmission in the absence of superconductivity. Transfer of Cooper pairs is better described by the Andreev reflection probability A . (b) The parameter $A(E=0)$ at the Fermi level can be expressed directly in terms of Γ .

1. The integral in eq.(4.16) needs to be valid, for which the normal-state conductance needs to be independent of V_d around E_F ;
2. The error in I_{SN}/I_{NN} should be much smaller than the desired accuracy in Z (e.g., by a factor 6 when $Z \approx 2$)⁷; and
3. Both I_{SN} and I_{NN} should be known at $eV_d \gg \Delta$.

Unfortunately, none of these were met in our experiments. Fig. 4.21 shows the normal-state conductance for a similar device, which features a strong dip around zero bias, invalidating the BTK assumption of flat normal-state conduction. Second, as is clearest in Fig. 4.11, the conductance is highly sensitive to small variations in gate voltage, while hysteresis (likely due to trapped charges in the oxide) has been observed in the gate field. This means that we can never directly compare superconducting and normal-state scans at $H = 0$ and $H > H_c$, nor is it practical to fix a gate voltage and then change the field (you can only do this once, and then you need to heat up the coil recover $H = 0$ exactly). Third, we never did scans wide enough to reach $eV > 50 \times 2\Delta$ [46], at which point $I_{\text{excess}}/I_{\text{total}}$ would be smaller than our signal-to-noise ratio.

Instead, a qualitative method was developed that, like the excess current in this situation, does not per se give an accurate estimate for Z for a single given gate voltage, but that provides a way of comparing the Z values for different V_g 's with limited data. This can then give us an idea of whether we are able to tune the transparency with the gate, by estimating the barrier strength Z from only a single measurement in the superconducting state.

⁷See Fig. 4.20b: since I_{excess} falls off quickly with Z , a small error in the estimate of the normal-state resistance will lead to large uncertainty in Z , especially when the interface has a low transparency.

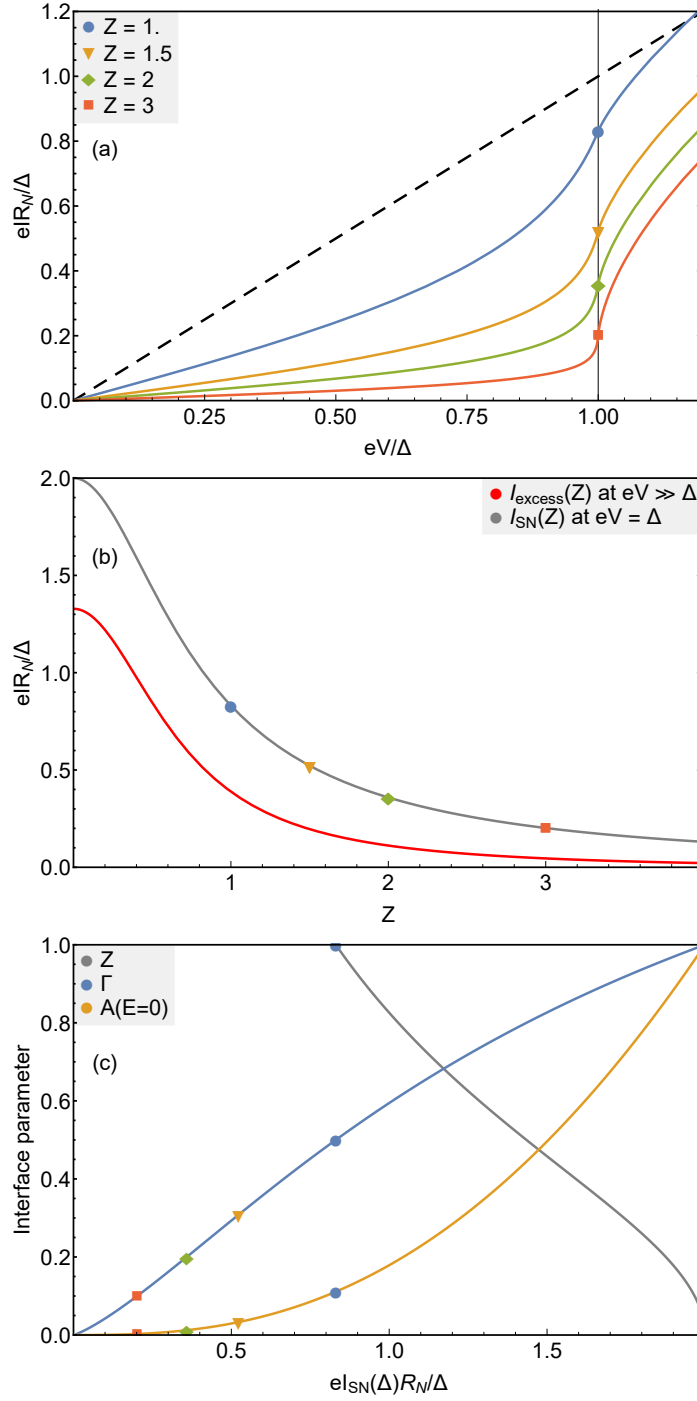


Figure 4.20: (a) Normalized current $eI_{SN}R_N/\Delta$ versus applied bias voltage for selected values of Z . Markers indicate the current at $eV = \Delta$. (b) The **excess current** at $eV \gg \Delta$ as described by BTK, as well as the **current at $eV = \Delta$** . The same markers are indicated. (c) Swapping the axes of the previous plot, we can find the interface parameters Z , Γ and A by solving $I_{SN}(Z)$ for Z . Note that for opaque interfaces, the Andreev reflection rate A grows much more slowly with $I_{SN}(\Delta)$ than the normal-state interface transparency Γ . E.g., a moderate $Z = 1.5$ gives $\Gamma = 0.31$, while $A(Z = 1.5, E = 0) = 0.03$.

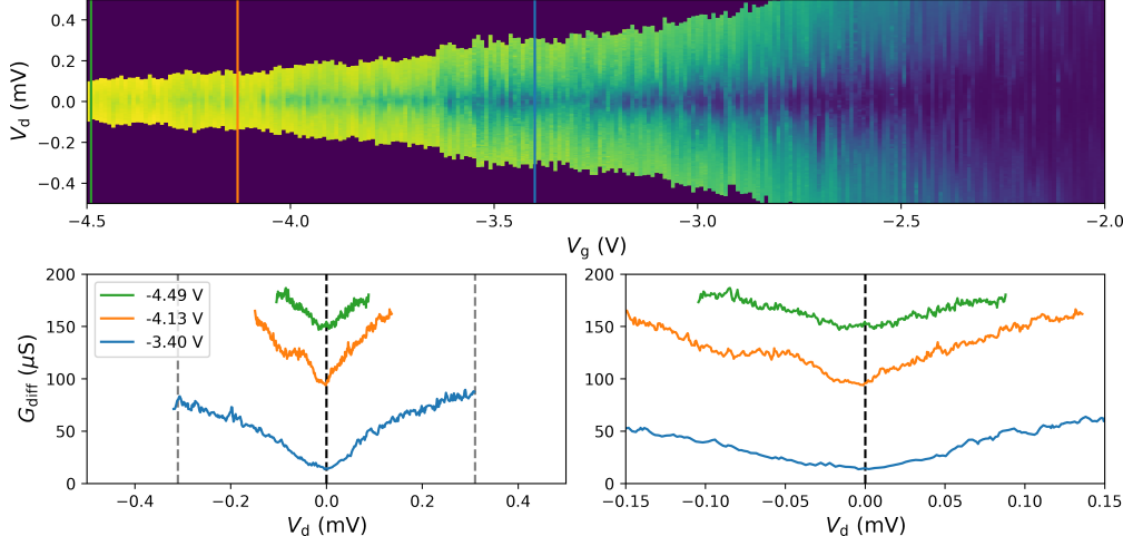


Figure 4.21: Data from device D63D3, $H = 100 \text{ mT} > H_0$. A clear non-linearity can be observed in the differential conductance even when the superconductors in the source and drain are in the normal state.

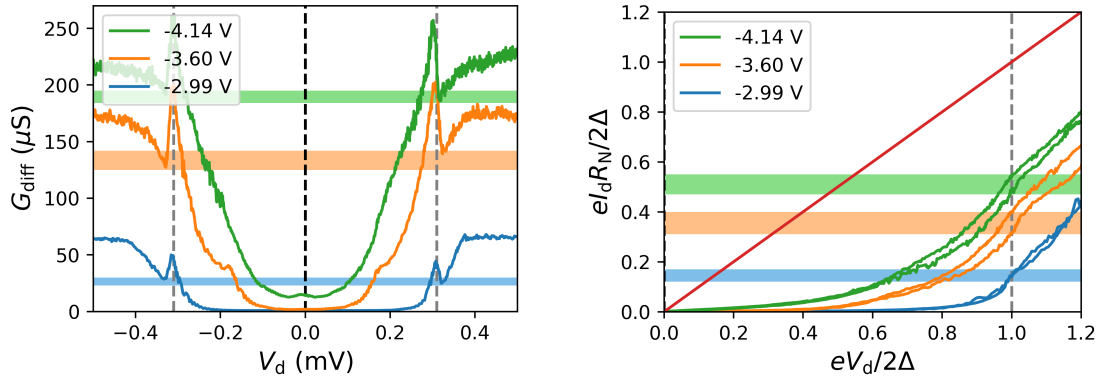


Figure 4.22: (Left) The normal-state conductance $G_{\text{diff,N}}$ at 2Δ can be approximated by taking the lowest value of G_{diff} outside the coherence peaks, from which $eR_N(2\Delta) = 2\Delta/G_{\text{diff}}$. Indicated in horizontal bars are the estimated ranges for $G_{\text{diff,N}}$. **(Right)** Using this estimate of R_N , we can now plot the normalized current I_{SN} (data for negative bias voltages is mirrored and added). In this figure, horizontal bars represent the estimates for the current at $eV = 2\Delta$.

Shown in in Fig. 4.22 are selected curves of the differential conductance and normalized current in device D61D4, from the same measurement as shown in Fig. 4.12. Sub-gap conductance increases with applied gate voltage, as expected for an increasingly transparent interface. Now, we know that expression (4.16) for the current at the gap (or twice the gap in our SNS junction) is not valid, since the current in both the superconducting and normal state is suppressed by the process⁸ that gives rise to the dip around $V = 0$ in Fig. 4.21. But what's important, is that we can be reasonably sure that it remains a smoothly varying one-to-one relationship on larger V_g scales⁹, even when we replace $R_N \rightarrow R_{N,\text{diff}}$,

$$I_{\text{NS}}(T = 0, eV) \approx \frac{1 + Z^2}{eR_{N,\text{diff}}} \int_0^{eV} [1 + A(E, Z) - B(E, Z)] dE. \quad (4.18)$$

We know that $R_N \neq R_{N,\text{diff}}$, shown in Fig. 4.23, but we will here make the approximation that $R_{N,\text{diff}}$ and I_{NS} are suppressed by about the same amount within the gap. Using $R_{N,\text{diff}}$ essentially allows us to approximately compensate for the depression in R_N . This will be an overcompensation, since $R_N G_{\text{diff},S}$ increases between $eV = 0$ and $eV = 2\Delta$ for any value of Z [46], such that conductance near the edge of the gap makes up a larger relative share of total current in the superconducting state than it does in the normal state. This means that the normalized $eIR_{N,\text{diff}}/2\Delta$ that we will calculate at $eV = 2\Delta$ will be a lower bound.

To estimate $R_{N,\text{diff}}$ from data in the superconducting state, we make one further approximation, justified by earlier data on PtSi transistors [47], that

$$G_{\text{diff},N}(eV = 2\Delta) \approx G_{\text{diff},S}(eV = 2\Delta + \delta), \quad (4.19)$$

where δ is a small shift in voltage from the coherence peak, such that we take a $G_{\text{diff},S}(V)$ value just outside the gap, at the point where it is smallest. For clarity, this is drawn in Fig. 4.22 as shaded horizontal bars that intersect the minimum of $G_{\text{diff},S}(V)$ outside the gap. Once we have this estimate for $R_{N,\text{diff}}$, we can proceed to normalize the measured current at $eV = 2\Delta$.

Integrating the modified expression in eq. (4.18) to $eV = \Delta$ (not $eV = 2\Delta$, BTK's formula is for a single interface) yields the following,

$$I_{\text{SN}}(T = 0, eV = \Delta) = \frac{\Delta}{eR_N} \frac{\sqrt{1 + Z^2}}{Z(1 + 2Z^2)} \operatorname{atanh} \left(\frac{2Z\sqrt{1 + Z^2}}{1 + 2Z^2} \right), \quad (4.20)$$

which is plotted in Fig. 4.20b together with the excess current. Given an estimate for I_{SN} at $eV = 2\Delta$, we can then solve for Z , as is detailed in algorithm 4.1. The last step, solving the equality, is done by minimizing the error using Brent's algorithm [48] with the built-in function `minimize_scalar()` from the Python library `scipy`, bounded between $Z = 0$ and $Z = 100$. Plotted in Fig. 4.24 are the resulting estimates for the normal-state transparency Γ and the probability of Andreev reflection at $eV = 0$, given by BTK's $A(eV = 0)$.

⁸At the moment, we suspect that this dip is due to a dynamical Coulomb blockade.

⁹See Figs. 4.21 and 4.11: charging effects on the order of $\Delta V_g \approx 5 \text{ mV}$ will affect our estimates of R_N and I_{NS} differently. This doesn't matter when we just want to know the trend over a wider range in V_g .

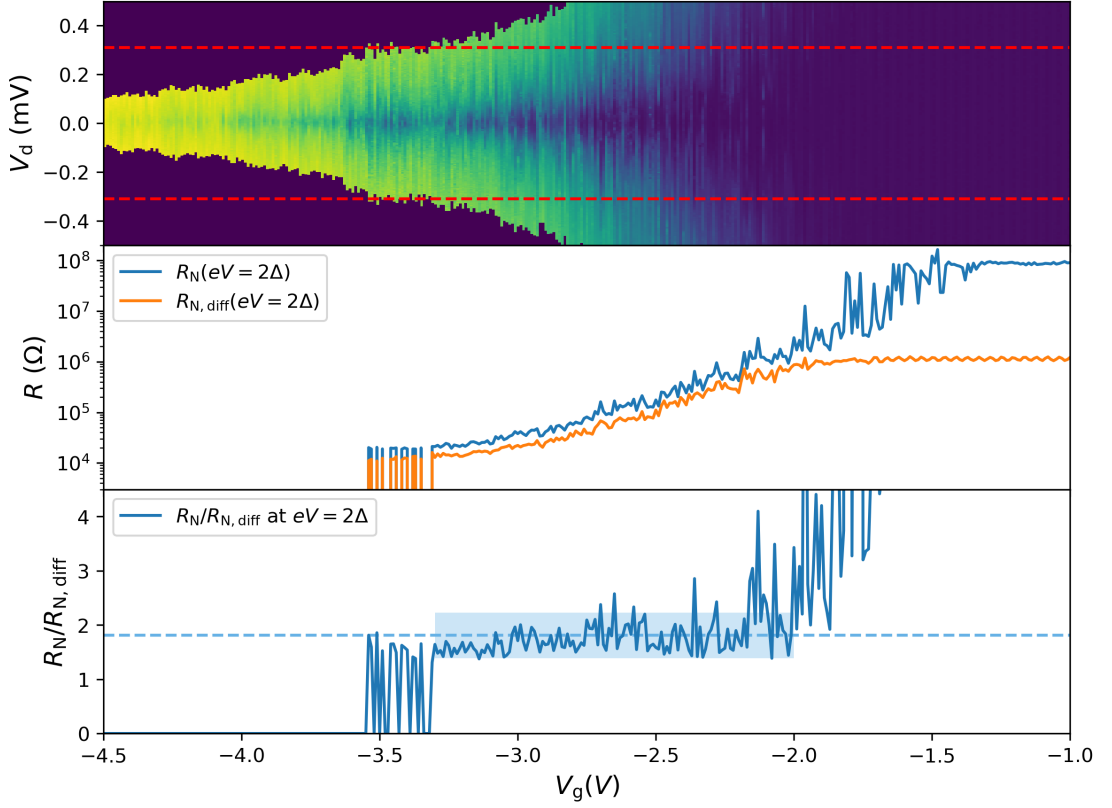


Figure 4.23: Data shown for D63D3, same measurement as shown in Fig. 4.21. **(Top)** The differential conductance versus V_g and V_d , red dashed lines indicate $V = \pm 2\Delta$. **(Middle)** The normal-state resistance at $eV = 2\Delta$ is calculated from the current, $R_N = \Delta/eI$, while the differential resistance is simply $R_{N,diff} = 1/G_{N,diff}$. **(Bottom)** In the range where we have data ($V_g \gtrsim -3.3$ V), and a decent signal-to-noise ratio ($V_g \lesssim -2.0$ V) (indicated in light blue), R_N and $R_{N,diff}$ are off by about a factor 1.8(4).

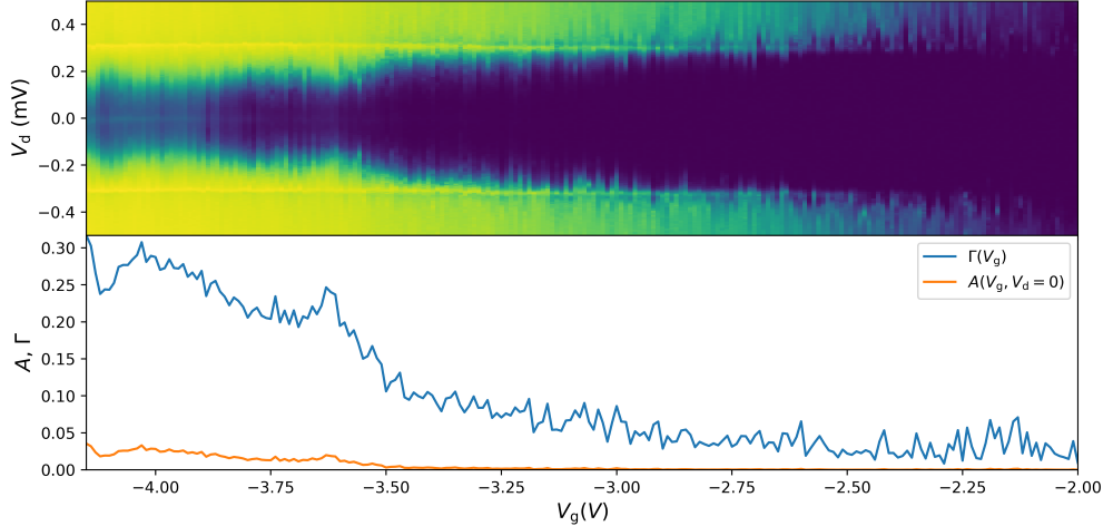


Figure 4.24: Data from device D61D04. **(Top)** The differential conductance of device D61D04 versus gate and drain voltage. On the left, around $V_d = 0$, a zero-bias conductance peak (ZBCP) can be observed. **(Bottom)** Around the same range in V_g , the extracted $A(E = 0)$ increases, consistent with this peak being due to the proximity effect.

Algorithm 4.1: Estimate Z from $G_{\text{diff}}(V)$ and $I(V)$.

```

1 function Estimate_RN ( Gdiff ):
2     GdiffWindowLeft    $\leftarrow$  select ( Gdiff ,  $-(1 + \delta)\Delta < eV < -\Delta$  )
3     GdiffWindowRight   $\leftarrow$  select ( Gdiff ,  $\Delta < eV < (1 + \delta)\Delta$  )
4     { RNLeft , RNRight }  $\leftarrow$  {  $1 / \min ( GdiffWindowLeft )$  ,  $1 / \min ( GdiffWindowRight )$  }
5     return mean ( RNLeft , RNRight )
6
7 function Estimate_Isn_at_Δ ( Isn ):
8     IsnLeft            $\leftarrow$  -select ( Isn ,  $eV = -\Delta$  )
9     IsnRight           $\leftarrow$  select ( Isn ,  $eV = +\Delta$  )
10    return mean ( IsnLeft , IsnRight )
11
12 function Theoretical_Isn ( Z ):
13    return  $\sqrt{(1+Z^2)/(Z(1+2Z^2))} \tanh(2Z \sqrt{(1+Z^2)/(1+2Z^2)})$ 
14
15 RN  $\leftarrow$  Estimate_RN ( Gdiff ( Vg ) )
16 Isn  $\leftarrow$  Estimate_Isn_at_Δ ( Isn )
17 solve Theoretical_Isn ( Z ) = Isn * RN / Δ for Z

```

4.3.7 Estimation of the Schottky barrier height

An additional way that we can analyze the dependence of the interface transparency on variations in the applied gate voltage is by extracting the Schottky barrier height. As long as the temperature is high enough, transmission at the interface will mostly occur by thermionic emission, as opposed to regular or field-assisted tunneling¹⁰. Using eq. (4.6) for the thermionic emission current, we can then relate the Schottky barrier

¹⁰Thermionic emission of course has no role to play in the JoFET operation itself. Both the temperature and the superconducting gap are much smaller than the Schottky barrier, and Josephson coupling depends on transport at $E = eV_d = 0$.

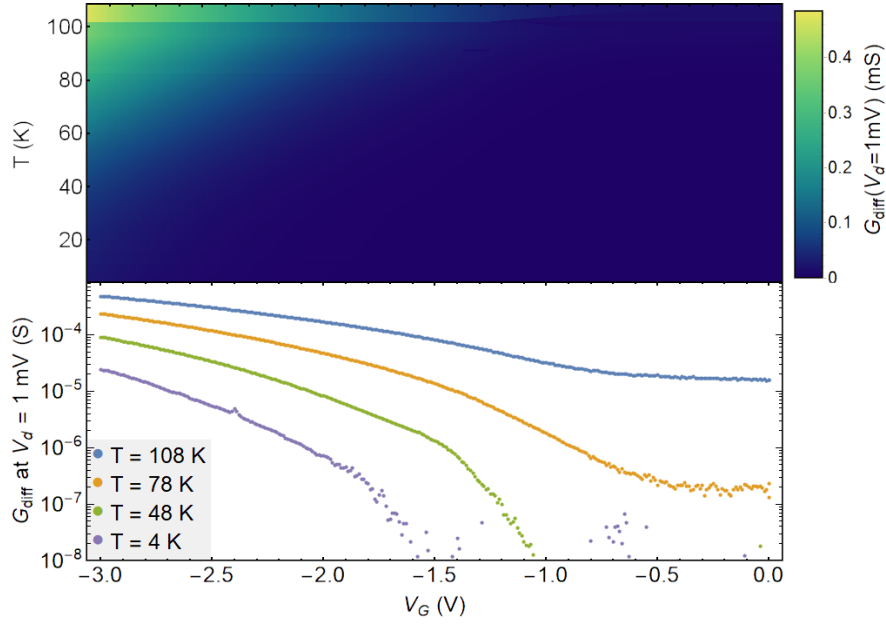


Figure 4.25: Data shown for device D84D3. The measurement was performed over a period of 1 week, and also gives some insight into the quality of the thermal insulation of the cryostat. **(Top)** The differential conductance at $V_d = 1$ mV for $-3 \text{ V} \leq V_g \leq 0 \text{ V}$ and temperatures from 4 K to 108 K. **(Bottom)** Selected curves $G_{\text{diff}}(V_g, V_d = +1 \text{ mV})$.

height ϕ_{Schottky} to the increase in current with temperature,

$$\phi_{\text{Schottky}} = -k_B T \ln \left(\frac{I_{\text{th}}}{T^2} \right) + C, \quad \text{or} \quad \ln \left(\frac{I_{\text{th}}}{T^2} \right) = -\frac{\phi_{\text{Schottky}}}{k_B T} + C', \quad (4.21)$$

such that ϕ_{Schottky} can be extracted as the slope in an Arrhenius plot of $\ln(I_{\text{th}}/T^2)$ versus $1/k_B T$. We will not attempt the more comprehensive analysis that takes into account the change in shape of the Schottky barrier with gate voltage [16].

The device for which data is shown in Fig. 4.24 was characterized only at 300 K, 4 K and base temperature. We will therefore take data from a similar device, for which a scan in V_g at fixed $V_d = 1$ mV was repeated continuously, after all the ^4He had evaporated from the cryostat. These data are shown in Fig. 4.25, and clearly indicate an increase in conductance with temperature.

The regime where thermionic emission dominates over tunneling depends on the width of the barrier [49], which in turn in the case of an SBMOSFET depends on the applied gate voltage. Heuristically, we can guess from linearity in the Arrhenius plot that thermionic emission dominates above ~ 80 K, similar to that used in other sources [2, 7]. At higher temperatures, perhaps $T > 150$ K for a typical PtSi SBMOSFET [7], resistance may be limited by the channel instead of the Schottky barrier. The magnitude of the tunnel/field-emission current depends only on the gate voltage and not the temperature, and so does not need to be taken into account. Nor do we have to worry about non-linearities due to the dependence of I_{th} on the source-drain bias V_d in eq. (4.6), since $eV_d = 1 \text{ meV} \approx k_B \times 12 \text{ K}$, far below the temperatures at which we perform the fit.

As can be seen in Fig. 4.25, the Schottky barrier is suppressed by the gate from 0.16 eV around the threshold voltage (equal to the value measured in a diode [2], see section 4.3.1), to nearly zero at $V_g = -3 \text{ V}$. Though we cannot directly compare these results to the improved interface transparency observed in Fig. 4.24 (the data are from

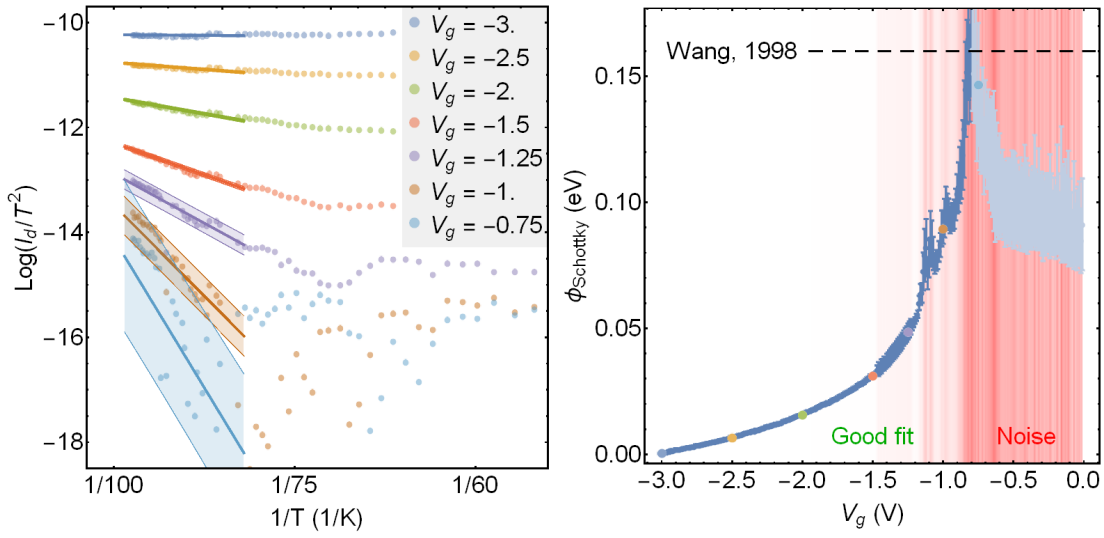


Figure 4.26: Data shown for device D84D3, same measurement as Fig. 4.25. **(Left)** By plotting $\log(I_d/T^2)$ at fixed source-drain bias V_d and gate voltage V_g versus the reciprocal temperature $1/T$, it is possible to extract the Schottky barrier height ϕ_{Schottky} using eq. (4.5). The linear regression fails when the detected current is on the same order as the noise, as is the case for e.g. $V_g = -0.75$ V (bottom curve). **(Right)** Extracted SBH versus V_g , colored dots for the curves shown on the left superposed. The red background shading is proportional to the error bar on the slope estimate. For $V_g > -0.8$ V (shown in lighter blue), the signal to noise ratio became too poor to perform good fits. The dashed horizontal line at $\phi_{\text{SBH}} = 0.16$ eV was extracted from a reverse-bias PtSi Schottky diode on a different wafer with the same doping [2].

different devices), we can be confident that the suppression of the Schottky barrier is responsible for at least part of the improvement in transparency. Transport then occurs by tunneling through this reduced barrier, though it is not yet clear what role is played by gate-stimulated emission.

4.3.8 Conclusion

Towards the end of this PhD we have begun measuring PtSi SBMOSFET devices with channel lengths on the order of 50 nm and doping levels that varied from $5 \times 10^{15} \text{ cm}^{-3}$ deeper in the bulk to 10^{19} cm^{-3} underneath the oxide. We have found that the source and drain are superconducting, and that this superconductivity can be induced by proximity effect inside the channel. More importantly, we saw that the extent of this proximity effect can be tuned with an electrostatic field, likely by tuning the Schottky barrier width even after dopants have frozen out, which suggests that the shielding charges that reduce the barrier width [50] are instead ejected by field emission from the PtSi contacts [6]. The appearance of a zero-bias conductance peak is consistent both with reflectionless tunneling and coherent transport across the junction, as is its suppression by small magnetic fields on the order of 10 mT. It remains to be seen if a supercurrent can be detected by moving to a cryogenic setup with better noise filtering.

As discussed in section 2.2.3, a Josephson junction for transmon applications only needs a critical current on the order of 1–10 nA, which corresponds to a current density on the order of 1–10 A/cm² for the devices of the dimensions that we have studied, still 3 to 4 orders of magnitude below what can be obtained on degenerately doped (7×10^{19} – $1.8 \times 10^{20} \text{ cm}^{-3}$) silicon [51]. Gate modulation has been demonstrated in devices with lower doping of $5 \times 10^{18} \text{ cm}^{-3}$ in an experiment where a gate electrode was introduced by etching a hole from the back of the wafer [52, 53]. The channel lengths in these experiments ranged from 40 to 200 nm, well within the reach of CMOS fabrication techniques. Such structures seem to fall within the intermediate regime, where critical currents are reported to be limited by both the interface [40, 54], and decoherence in the channel [55], a crossover between which occurs as the Schottky barriers are suppressed by the gate [56, 57].

Now that a two-step annealing process has been found for PtSi formation, whereby the encroachment underneath the gate can be controlled, and undulations in the PtSi/Si interface can be reduced (see section 3.4.2), we have begun to fabricate transistors with pre-encroachment gate lengths down to 40 nm. Since both the doping level and the channel length can be well controlled in CMOS technology, it should be possible to optimize these designs to obtain a gate-controllable supercurrent on the order of a few nano-amps, and then move on to designing a transmon circuit around it. With this, I wish the next student the best of luck.

Bibliography

- [1] Tom Doekle Vethaak. *Python code for lithographic mask sets*. <https://github.com/TomVethaak/SuperGe-Masks/>. (Visited on 04/01/2021).
- [2] Chinlee Wang. “A sub-0.1 μm PtSi Schottky source/drain MOSFET”. PhD thesis. University of Illinois at Urbana-Champaign, 1998.
- [3] MJ Black and V Chandrasekhar. “Influence of temperature-dependent inelastic scattering on the superconducting proximity effect”. In: *EPL (Europhysics Letters)* 50.2 (2000), p. 257.
- [4] Laurie Ellen Calvet. “Electrical transport in Schottky barrier MOSFETs”. PhD thesis. Yale University, 2001.
- [5] LE Calvet et al. “Suppression of leakage current in Schottky barrier metal–oxide–semiconductor field-effect transistors”. In: *Journal of Applied Physics* 91.2 (2002), pp. 757–759.
- [6] Ralph Howard Fowler and Lothar Nordheim. “Electron emission in intense electric fields”. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 119.781 (1928), pp. 173–181.
- [7] Emmanuel Dubois and Guilhem Larrieu. “Measurement of low Schottky barrier heights applied to metallic source/drain metal–oxide–semiconductor field effect transistors”. In: *Journal of Applied Physics* 96.1 (2004), pp. 729–737.
- [8] S Zhu et al. “A BEEM study of the temperature dependence of the barrier height distribution in PtSi/n-Si Schottky diodes”. In: *Solid State Communications* 112.11 (1999), pp. 611–615.
- [9] Jens Koch et al. “Charge-insensitive qubit design derived from the Cooper pair box”. In: *Physical Review A* 76.4 (2007), p. 042319.
- [10] John M Martinis et al. “Decoherence in Josephson qubits from dielectric loss”. In: *Physical Review Letters* 95.21 (2005), p. 210503.
- [11] TD Clark, RJ Prance, and ADC Grassie. “Feasibility of hybrid Josephson field effect transistors”. In: *Journal of Applied Physics* 51.5 (1980), pp. 2736–2743.
- [12] Eddy Simoen et al. “Freeze-out effects on NMOS transistor characteristics at 4.2 K”. In: *IEEE Transactions on Electron Devices* 36.6 (1989), pp. 1155–1161.
- [13] Hans Albrecht Bethe. *Theory of high frequency rectification by silicon crystals*. Tech. rep. 11. MIT Radiation Laboratory (US), 1942.
- [14] Hans Albrecht Bethe. *Theory of the Boundary Layer of Crystal Rectifiers*. Tech. rep. 12. Unpublished report 43-12 (November 23, 1942). MIT Radiation Laboratory (US), 1942.

- [15] CR Crowell. “The Richardson constant for thermionic emission in Schottky barrier diodes”. In: *Solid-State Electronics* 8.4 (1965), pp. 395–399.
- [16] VL Rideout and CR Crowell. “Effects of image force and tunneling on current transport in metal-semiconductor (Schottky barrier) contacts”. In: *Solid-State Electronics* 13.7 (1970), pp. 993–1009.
- [17] CR Crowell and SM Sze. “Current transport in metal-semiconductor barriers”. In: *Solid-State Electronics* 9.11-12 (1966), pp. 1035–1048.
- [18] Malgorzata Chrzanowska-Jeske and Richard C Jaeger. “BILOW-simulation of low-temperature bipolar device behavior”. In: *IEEE Transactions on Electron Devices* 36.8 (1989), pp. 1475–1488.
- [19] RG Pires et al. “Carrier freezeout in silicon”. In: *Cryogenics* 30.12 (1990), pp. 1064–1068.
- [20] F Balestra, L Audaire, and C Lucas. “Influence of substrate freeze-out on the characteristics of MOS transistors at very low temperatures”. In: *Solid-State Electronics* 30.3 (1987), pp. 321–327.
- [21] Richard M Dickstein, Stephen L Titcomb, and Richard L Anderson. “Carrier concentration model for n-type silicon at low temperatures”. In: *Journal of Applied Physics* 66.6 (1989), pp. 2437–2441.
- [22] J Frenkel. “On pre-breakdown phenomena in insulators and electronic semiconductors”. In: *Physical Review* 54.8 (1938), p. 647.
- [23] NS Saks and A Nordbryhn. “Time dependence of depletion region formation in phosphorus-doped silicon MOS devices at cryogenic temperatures”. In: *Journal of Applied Physics* 50.11 (1979), pp. 6962–6968.
- [24] John Bardeen, Leon N Cooper, and John Robert Schrieffer. “Theory of superconductivity”. In: *Physical Review* 108.5 (1957), p. 1175.
- [25] David C Johnston. “Elaboration of the α -model derived from the BCS theory of superconductivity”. In: *Superconductor Science and Technology* 26.11 (2013), p. 115011.
- [26] Leonhard Euler. “De progressionibus harmonicis observationes”. In: *Commentarii Academiae Scientiarum Petropolitanae* (1740), pp. 150–161.
- [27] John Clarke. “Experimental observation of pair-quasiparticle potential difference in nonequilibrium superconductors”. In: *Physical Review Letters* 28.21 (1972), p. 1363.
- [28] M Tinkham and John Clarke. “Theory of pair-quasiparticle potential difference in nonequilibrium superconductors”. In: *Physical Review Letters* 28.21 (1972), p. 1366.
- [29] AF Andreev. “The thermal conductivity of the intermediate state in superconductors”. In: *Soviet Physics JETP* 19.5 (1964), pp. 1228–1231.
- [30] LP Gor’Kov. “On the energy spectrum of superconductors”. In: *Soviet Physics JETP* 7.505 (1958), p. 158.
- [31] Wolfgang Belzig, Christoph Bruder, and Gerd Schön. “Local density of states in a dirty normal metal connected to a superconductor”. In: *Physical Review B* 54.13 (1996), p. 9443.

- [32] Pierre Gilles De Gennes. “Superconductivity of Metals and Alloys”. In: (1999).
- [33] LD Landau and VI Ginzburg. “K teorii sverkhrovodimosti (On the theory of superconductivity)”. In: *Zhurnal Eksperimental’noi i Teoreticheskoi Fiziki (Journal of Experimental and Theoretical Physics)* 20.1064 (1950), pp. 546–568.
- [34] L. D. Landau and V. L. Ginzburg. “On the Theory of Superconductivity”. In: *On Superconductivity and Superfluidity: A Scientific Autobiography*. Springer Berlin Heidelberg, 2009, pp. 113–137. ISBN: 978-3-540-68008-6. DOI: 10.1007/978-3-540-68008-6_4.
- [35] Kristian Fossheim and Asle Sudbø. *Superconductivity: physics and applications*. John Wiley & Sons, 2004. ISBN: 978-0-470-84452-6.
- [36] Michael Tinkham. *Introduction to Superconductivity, Second Edition*. Dover Publications, 2004. ISBN: 978-0486435039.
- [37] A Kastalsky et al. “Observation of pair currents in superconductor-semiconductor contacts”. In: *Physical Review Letters* 67.21 (1991), p. 3026.
- [38] BJ Van Wees et al. “Excess conductance of superconductor-semiconductor interfaces due to phase conjugation between electrons and holes”. In: *Physical Review Letters* 69.3 (1992), p. 510.
- [39] PGf de Gennes. “Boundary effects in superconductors”. In: *Reviews of Modern Physics* 36.1 (1964), p. 225.
- [40] AW Kleinsasser. “Critical currents of semiconductor-coupled Josephson weak links”. In: *IEEE Transactions on Magnetism* 27.2 (1991), pp. 2589–2593.
- [41] P Dubos et al. “Josephson critical current in a long mesoscopic SNS junction”. In: *Physical Review B* 63.6 (2001), p. 064502.
- [42] H De Raedt, K Michielsen, and TM Klapwijk. “Andreev reflection in nanoscale metal-superconductor devices”. In: *Physical Review B* 50.1 (1994), p. 631.
- [43] HFC Hoevers et al. “Determination of the energy gap in a thin YBa₂Cu₃O_{7-x} film by Andreev reflection and by tunneling”. In: *Physica C: Superconductivity* 152.1 (1988), pp. 105–110.
- [44] H De Raedt, K Michielsen, and TM Klapwijk. “Andreev reflection”. Animations of the results discussed in Ref. 42, obtained from first author upon request.
- [45] K Oto et al. “Superconductivity in PtSi ultrathin films”. In: *Journal of Applied Physics* 76.9 (1994), pp. 5339–5342.
- [46] GE Blonder, M Tinkham, and TM Klapwijk. “Transition from metallic to tunneling regimes in superconducting microconstrictions: Excess current, charge imbalance, and supercurrent conversion”. In: *Physical Review B* 25.7 (1982), p. 4515.
- [47] Mike Schwarz et al. “Physics and Technology of Schottky-Barrier Source-Drain Contact Field Effect Transistors”.
- [48] Richard P. Brent. “An algorithm with guaranteed convergence for finding a zero of a function”. In: *The Computer Journal* 14.4 (1971), pp. 422–425.
- [49] James Jeffrey Hugunin. “The superconductor/semiconductor interface of V₃Si and Si”. Master thesis. Massachusetts Institute of Technology, 1995.

- [50] LG Aslamazov and MV Fistul. “Temperature dependence of the critical current at superconductor-semiconductor-superconductor junctions”. In: *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki* 81 (1981), pp. 382–397.
- [51] CL Huang and T Van Duzer. “Josephson tunneling through locally thinned silicon wafers”. In: *Applied Physics Letters* 25.12 (1974), pp. 753–756.
- [52] T Nishino et al. “Three-terminal superconducting device using a Si single-crystal film”. In: *IEEE Electron Device Letters* 6.6 (1985), pp. 297–299.
- [53] T Nishino, E Yamada, and U Kawabe. “Carrier-concentration dependence of critical superconducting current induced by the proximity effect in silicon”. In: *Physical Review B* 33.3 (1986), p. 2042.
- [54] Th Becker, M Mück, and Ch Heiden. “Transport properties of niobium-silicon superconductor-semiconductor junctions”. In: *Physica B: Condensed Matter* 204.1-4 (1995), pp. 183–189.
- [55] M Okamoto et al. “Gate-voltage dependence of the critical current of the superconducting field effect transistor”. In: *Applied Physics Letters* 61.20 (1992), pp. 2473–2475.
- [56] AW Kleinsasser et al. “Superconducting InGaAs junction field-effect transistors with Nb electrodes”. In: *Applied Physics Letters* 55.18 (1989), pp. 1909–1911.
- [57] AW Kleinsasser et al. “Crossover from tunneling to metallic behavior in superconductor-semiconductor contacts”. In: *Applied Physics Letters* 57.17 (1990), pp. 1811–1813.

Conclusion

Quantum computers are more powerful than their classical counterparts [1], in the sense that they can solve certain classes of tasks more efficiently. For example, there exist n -bit problems that would take $\mathcal{O}(\exp(n))$ time on a classical machine, but that can be solved to arbitrary accuracy within $\mathcal{O}(n^k)$ (with $k \in \mathbb{N}$) on a quantum computer. Other problems benefit from more modest reductions by polynomial factors, or can be solved on a quantum computer while the classical computation cost is not yet clear. It is important to point out that this feat is achieved not by “performing exponentially many calculations in parallel”, but by creating interference patterns in the collective wave function of the entangled qubits, with the probability density in the final state concentrated near solutions. This gives an advantage only in the asymptotic scaling of the computation time; performing a single fault-tolerant two-qubit operation can actually take up to 10^7 times longer than the simple switch of a transistor [2]. For most problems, a quantum advantage can therefore only be achieved when large numbers of qubits are involved, at which point logical qubits need to be encoded by error correction in collections of perhaps thousands of physical ones [3].

To control entanglement of a few million qubits, it is necessary that both their operation and fabrication become scalable. Much of the progress in the past one and a half decade has been achieved with the transmon design [4, 5], which as a lithographically defined “artificial atom” has the advantage that nearly any aspect of it can be engineered. This has allowed scalable operation by moving from flux-controlled SQUID-based transmons to circuits made with gate-tunable Josephson junctions [6, 7], while mass production may be facilitated by developing a fabrication process that is fully compatible with the technologies of the semiconductor industry. Our approach is to modify existing transistor designs by using superconducting materials for the source and drain, such that the gate voltage can be used to tune the magnitude of a small Josephson current on the order of a few nano-Amperes. The resulting JoFET is then connected to a planar superconducting capacitor to form a transmon, and can be directly integrated alongside cryo-CMOS control electronics [8, 9].

Silicon is however not the most natural candidate material for the weak link in a Josephson junction, and the main challenge is to achieve a sufficiently large supercurrent while maintaining a degree of gate control. The first concern is the transparency of the superconductor/semiconductor interface, since the relevant mechanism, Andreev reflection, is a second-order process that scales with the square of this parameter. Depending on the length of the channel L and the diffusion constant D inside it, the supercurrent is then limited by either the superconducting gap Δ of the contacts, or the Thouless energy $E = \hbar D/L^2$ of the channel (whichever is smaller). Both the transparency of the interfaces and the diffusion between them depend on the number of carriers that are attracted by the gate [10, 11], such that the limiting factor for a given device can change as the gate

voltage is varied. To maximize supercurrent, we thus start by selecting contact materials with high transparency to silicon. Luckily, similar concerns for regular transistors have already driven the microelectronics industry to develop materials with low contact resistance, which has led to a thorough understanding of the fabrication and behavior of silicides. We can take advantage of this by selecting those that are superconducting, and then further optimize the fabrication to increase the superconducting gap and Thouless energy. It is not yet clear which of these two energy scales will typically be smallest in devices that can be fabricated with conventional CMOS technology, so we hedge our bets by studying both V_3Si , the silicide with the highest known critical temperature, and PtSi, which during its formation offers good control over the channel length.

The integration of V_3Si in transistors comes with the challenge that it is neither the first, nor the final stable phase to form in the presence of a silicon reservoir. One way to overcome this challenge is to abandon the SALICIDE procedure in favor of direct sputtering of the silicide in the right stoichiometry, followed by crystallization annealing. This thermal processing in turn presents the challenge of VSi_2 formation at the interface with the underlying silicon, with the risk of either consuming the entire channel, or all available superconducting V_3Si . Before either occurs, it could however also bring down the Thouless energy due to faster carrier diffusion in this metallic material, and move the Schottky barrier within range of the electrostatic gate field. This encroachment underneath the spacers, or even the gate itself, is also a central part of PtSi-based JoFETs. At the temperature where this final phase in the Pt-Si system obtains its desired properties, both species of atoms diffuse so fast that it becomes impossible to control the location of the silicide-silicon interface by timing of the thermal processing alone. Since in a silicon-on-insulator (SOI) design the planar contact openings have a vastly larger surface area than the vertical sides of the channel edges, it is also challenging to moderate the encroachment by tuning the metal deposition. A two-step annealing process provides a solution, in which the slower growth of either Pt_2Si or PtSi during a first heating step at lower temperatures is used to define the channel, and a selective etch of Pt and Pt_2Si before the final anneal prevents further silicon consumption.

Measurements on a PtSi Schottky-barrier MOSFET confirmed that when the silicide extends underneath the spacers, the interface transparency can be modulated with the gate voltage. This allows for control of the proximity effect, the first step towards a semiconductor-based tunable Josephson junction. Further reduction of the Schottky barrier width and enhancement of the carrier diffusion could be achieved with higher levels of doping, where a second boron implantation after gate stack definition would minimize the barrier height while ensuring semiconducting behavior below the gate. Once a critical current on the order of 10 nA is obtained, and can be modulated with the gate, a surrounding planar capacitor can be lithographically defined to form a gatemon. A fully CMOS-compatible design may then be achieved by defining both this capacitor and the surrounding control and readout circuits in a superconducting silicide thin film, possibly on the same level as the JoFET contacts. The main strength of this approach is that by leveraging decades of industrial experience, a high degree of reproducibility can be achieved, with little variation in device performance within a single die, or even across the wafer. Overcoming this bottleneck [12] in scalable fabrication may enable the development of larger NISQ circuits, and could be a stepping stone towards hardware suitable for error correction, a necessary component of fault-tolerant quantum computing.

Bibliography

- [1] Ethan Bernstein and Umesh Vazirani. “Quantum complexity theory”. In: *SIAM Journal on Computing* 26.5 (1997), pp. 1411–1473.
- [2] Ryan Babbush. *Prospects for Error Corrected Quantum Applications*. Q2B20 Conference. Oral presentation. 2020.
- [3] Austin G Fowler et al. “Surface codes: Towards practical large-scale quantum computation”. In: *Physical Review A* 86.3 (2012), p. 032324.
- [4] Jens Koch et al. “Charge-insensitive qubit design derived from the Cooper pair box”. In: *Physical Review A* 76.4 (2007), p. 042319.
- [5] Frank Arute et al. “Quantum supremacy using a programmable superconducting processor”. In: *Nature* 574.7779 (2019), pp. 505–510.
- [6] Thorvald Wadum Larsen et al. “Semiconductor-nanowire-based superconducting qubit”. In: *Physical Review Letters* 115.12 (2015), p. 127001.
- [7] G De Lange et al. “Realization of microwave quantum circuits using hybrid superconducting-semiconducting nanowire Josephson elements”. In: *Physical Review Letters* 115.12 (2015), p. 127002.
- [8] L Le Guevel et al. “Low-power transimpedance amplifier for cryogenic integration with quantum devices”. In: *Applied Physics Reviews* 7.4 (2020), p. 041407.
- [9] Loïck Le Guevel et al. “A 110mK 295 μ W 28nm FDSOI CMOS Quantum Integrated Circuit with a 2.8 GHz Excitation and nA Current Sensing of an On-Chip Double Quantum Dot”. In: *2020 IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE. 2020, pp. 306–308.
- [10] Anatoly F Volkov and Hideaki Takayanagi. “Effect of gate voltage on critical current in controllable superconductor–normal-metal–superconductor Josephson junctions”. In: *Physical Review B* 53.22 (1996), p. 15162.
- [11] AW Kleinsasser et al. “Crossover from tunneling to metallic behavior in superconductor-semiconductor contacts”. In: *Applied Physics Letters* 57.17 (1990), pp. 1811–1813.
- [12] JM Kreikebaum et al. “Improving wafer-scale Josephson junction resistance variation in superconducting quantum coherent circuits”. In: *Superconductor Science and Technology* 33.6 (2020), 06LT02.

Appendices

Appendix A

Samples for material studies

A.1 Samples prepared by pure V sputtering

Table A.1: Samples prepared for the study of V_3Si formation by V deposition.

#	Sample	Substrate/cap (TiN = cap)	V deposition (nm)	RTA (°C)	atmosphere	R_{\square}	$E(R_{\square})$
1	A32A	15nm TiN	80	840.	N2	4.78	0.046366
2	A32B	15nm TiN	80	0.	N2	8.94	0.286974
3	A32C	15nm TiN	80	840.	N2	5.16	0.087204
4	A32D	15nm TiN	80	0.	N2	8.78	0.15365
5	A33B1	15nm TiN	40	800.	N2	11.1724	0.08196
6	A33B2	15nm TiN	40	800.	UHV	8.61458	0.0130839
7	A33C1	15nm TiN	40	900.	N2	8.4348	0.0613132
8	A33C2	15nm TiN	40	900.	UHV	7.99213	0.00692336
9	A33D1	15nm TiN	40	1000.	N2	8.2248	0.0114063
10	A33D2	15nm TiN	40	1000.	UHV	8.33357	0.00523356
11	A35B	15nm TiN	20	0.	N2	20.1782	0.0197563
12	A35A	15nm TiN	20	500.	N2	55.8833	0.670025
13	A35C	15nm TiN	20	600.	N2	30.0367	0.070946
14	A36A	15nm TiN	20	700.	N2	23.8267	0.0929157
15	A36C	15nm TiN	20	800.	N2	21.8767	0.292973
16	A35D	15nm TiN	20	900.	N2	16.7983	0.186969
17	A36B	15nm TiN	20	1000.	N2	16.084	0.242474
18	A42A	HF dip + 15nm TiN	80	840.	N2	4.11	0.02055
19	A42B	HF dip + 15nm TiN	80	0.	N2	6.04	0.281464
20	A42C	HF dip + 15nm TiN	80	840.	N2	3.95	0.00237
21	A42D	HF dip + 15nm TiN	80	0.	N2	6.07	0.226411
22	A66A		100	300.	N2	14.9569	0.402849
23	A66B		100	400.	N2	20.5318	0.197563
24	A66C		100	500.	N2	21.6951	0.566701
25	A66D		100	600.	N2	13.5972	0.197563
26	A66E		100	700.	N2	10.2885	0.119916
27	A66F		100	800.	N2	6.58709	0.145696
28	A66G		100	900.	N2	5.30895	0.0911377
29	A66H		100	1000.	N2	4.42211	0.287882
30	A66I		100	0.	N2	11.3612	0.232585
31	B34B	20nm SiO ₂ + 15nm TiN	40	0.	N2	12.7436	0.175792
32	B34A	20nm SiO ₂ + 15nm TiN	40	500.	N2	37.0733	0.275379
33	B34C	20nm SiO ₂ + 15nm TiN	40	600.	N2	35.62	0.155885
34	B23B	20nm SiO ₂ + 15nm TiN	40	700.	N2	33.28	0.18735
35	B23D	20nm SiO ₂ + 15nm TiN	40	800.	N2	29.8967	0.023094

Table A.1: (Continued) Samples prepared for the study of V_3Si formation by V deposition.

#	Sample	Substrate/cap (TiN = cap)	V deposition (nm)	RTA ($^{\circ}C$)	atmosphere	R_{\square}	$E(R_{\square})$
36	B34D	20nm SiO ₂ + 15nm TiN	40	900.	N2	27.771	0.514728
37	B23C	20nm SiO ₂ + 15nm TiN	40	1000.	N2	29.158	0.841264
38	B54A	20nm SiO ₂ + 15nm TiN	80	840.	N2	11.44	0.088088
39	B54B	20nm SiO ₂ + 15nm TiN	80	0.	N2	8.12	0.1624
40	B54C	20nm SiO ₂ + 15nm TiN	80	840.	N2	11.08	0.212736
41	B54D	20nm SiO ₂ + 15nm TiN	80	0.	N2	8.	0.1472
42	B55A	20nm SiO ₂ + 10nm HfO ₂ + 15nm TiN	100	840.	N2	20.9	3.98354
43	B55B	20nm SiO ₂ + 10nm HfO ₂ + 15nm TiN	100	0.	N2	11.88	0.105732
44	B55C	20nm SiO ₂ + 10nm HfO ₂ + 15nm TiN	100	840.	N2	15.15	0.67266
45	B55D	20nm SiO ₂ + 10nm HfO ₂ + 15nm TiN	100	0.	N2	11.85	0.103095
46	B56A	20nm SiO ₂	100	300.	N2	17.8879	0.291393
47	B56B	20nm SiO ₂	100	400.	N2	23.6138	0.522702
48	B56C	20nm SiO ₂	100	500.	N2	24.4901	0.228126
49	B56D	20nm SiO ₂	100	600.	N2	18.885	0.228126
50	B56E	20nm SiO ₂	100	700.	N2	19.6253	0.119916
51	B56F	20nm SiO ₂	100	800.	N2	11.6181	0.228126
52	B56G	20nm SiO ₂	100	900.	N2	11.2857	0.252353
53	B56H	20nm SiO ₂	100	1000.	N2	14.1109	0.188699
54	B56I	20nm SiO ₂	100	0.	N2	–	–
55	B73A	20nm SiO ₂ + 15nm TiN	80	800.	N2	16.7291	0.448868
56	B73B	20nm SiO ₂ + 15nm TiN	80	900.	N2	12.8161	0.359815
57	B73C	20nm SiO ₂ + 15nm TiN	80	1000.	N2	13.3812	1.43864
58	B73D	20nm SiO ₂ + 15nm TiN	80	0.	N2	8.05906	0.173023
59	B74A	20nm SiO ₂ + 15nm TiN	80	800.	UHV	723.371	36.1826
60	B74D	20nm SiO ₂ + 15nm TiN	80	900.	UHV	116.029	14.2681
61	C26A	aSi (Ge) + 15nm TiN	80	800.	N2	11.5168	0.216467
62	C26B	aSi (Ge) + 15nm TiN	80	900.	N2	8.36832	0.1638
63	C26C	aSi (Ge) + 15nm TiN	80	1000.	N2	8.54811	0.447562
64	C26D	aSi (Ge) + 15nm TiN	80	0.	N2	7.42377	0.194495
65	C33A	aSi (Ge) + 0.1% O + 15nm TiN	80	800.	N2	11.0243	0.322353
66	C33B	aSi (Ge) + 0.1% O + 15nm TiN	80	900.	N2	8.73242	0.275111
67	C33C	aSi (Ge) + 0.1% O + 15nm TiN	80	1000.	N2	7.31983	0.404604
68	C33D	aSi (Ge) + 0.1% O + 15nm TiN	80	0.	N2	7.3168	0.099541
69	C53A	aSi (Ge) + 1% O + 15nm TiN	80	800.	N2	11.5425	0.193164
70	C53B	aSi (Ge) + 1% O + 15nm TiN	80	900.	N2	9.04818	0.274101
71	C53C	aSi (Ge) + 1% O + 15nm TiN	80	1000.	N2	7.63105	0.136073
72	C53D	aSi (Ge) + 1% O + 15nm TiN	80	0.	N2	7.06752	0.140894
73	C65A	aSi (Ge) + 5% O + 15nm TiN	80	800.	N2	11.3642	0.156308
74	C65B	aSi (Ge) + 5% O + 15nm TiN	80	900.	N2	8.36137	0.136723
75	C65C	aSi (Ge) + 5% O + 15nm TiN	80	1000.	N2	7.69148	0.249364
76	C65D	aSi (Ge) + 5% O + 15nm TiN	80	0.	N2	7.68091	0.203218
77	D32A	40nm SiN + 15nm TiN	100	840.	N2	11.58	0.26055
78	D32B	40nm SiN + 15nm TiN	100	0.	N2	10.69	0.132556
79	D32C	40nm SiN + 15nm TiN	100	840.	N2	11.46	0.121476
80	D32D	40nm SiN + 15nm TiN	100	0.	N2	10.62	0.174168
81	D33A	40nm SiN + 15nm TiN	100	800.	N2	14.6638	0.304189
82	D33B	40nm SiN + 15nm TiN	100	900.	N2	10.9896	0.300657
83	D33C	40nm SiN + 15nm TiN	100	1000.	N2	10.355	0.208507
84	D33D	40nm SiN + 15nm TiN	100	0.	N2	10.0272	0.130288
85	E55A	10% O + 15nm TiN	40	800.	N2	19.8444	0.177999
86	E55B	10% O + 15nm TiN	40	900.	N2	21.5002	1.07972
87	E55C	10% O + 15nm TiN	40	1000.	N2	19.654	0.838755
88	E55D	10% O + 15nm TiN	40	0.	N2	20.0045	0.65509

Table A.2: Superconducting critical temperature (T_c), residual resistance ratio (RRR) and low-temperature resistance ($R(20\text{ K})$) of some of the samples listed in table A.1.

#	Sample	T_c	$E(T_c)$ (10,90%)	RRR	$E(\text{RRR})$	Normalized $R(20\text{K})$	Normalized $E(R(20\text{K}))$
–	Ti10nm	0	{0, 0}	1.41449	0.00838401	30.1146	0.0431351
31	B34B	3.64205	{-0.222872, 0.0725929}	1.63019	0.0441452	7.81723	0.239143
34	B23B1A	0	{0, 0}	1.04629	0.000958445	31.8075	0.181416
35	B23D1A	0	{0, 0}	1.04629	0.000958445	28.5739	0.0342406
38	B54A	4.68876	{-0.154002, 0.2653}	0.998304	0.0530377	11.4594	0.623714
39	B54B	3.74038	{-0.0409306, 0.0395684}	2.09872	0.279515	3.86903	0.56646
55	B73A	0	{0, 0}	1.03719	0.0484716	16.1292	0.880549
56	B73B	2.68131	{-0.188784, 0.347914}	0.952515	0.0711176	13.455	1.0919
57	B73C	3.99827	{-0.135969, 0.772515}	0.918061	0.0437025	14.5755	1.71775
61	C26A	0	{0, 0}	1.42372	0.148817	8.08926	0.911885
62	C26B	5.85612	{-0.16484, 0.145038}	1.11062	0.0904084	7.53485	0.645378
62	C26B	5.86252	{-0.169211, 0.151172}	1.11062	0.0904084	7.53485	0.645378
63	C26C	7.80634	{-0.0813111, 0.872293}	1.18734	0.0844744	7.19937	0.646353
74	C65B	7.45921	{-0.418516, 0.283958}	1.15235	0.107453	7.25593	0.711175
75	C65C	8.02589	{-1.04534, -0.125081}	1.23567	0.139476	6.22453	0.7722
75	C65C	8.16673	{-0.181696, 0.197914}	1.23567	0.139476	6.22453	0.7722
86	E55B	5.60622	{-0.203318, 0.261955}	1.14946	0.0365311	18.7047	1.11395
87	E55C	7.06924	{-0.173071, 0.189863}	1.17895	0.0516317	16.6708	1.02769

A.2 Samples prepared by compound V_3Si sputtering

Table A.3: Before the processing of lots D19S2345–7, sputtering tests were performed on wafers A through L. Pieces were annealed, listed below.

#	Sample	Substrate	V_3Si deposition (nm)	RTA	R_{\square}	$E(R_{\square})$
1	C01	300 nm SiO_2	59.	600.	30.0196	0.811203
2	C02	300 nm SiO_2	59.	700.	17.48	0.159173
3	C03	300 nm SiO_2	59.	800.	17.1023	0.228126
4	C04	300 nm SiO_2	59.	900.	23.9764	3.33185
5	C06	300 nm SiO_2	59.	0.	34.0685	0.488855
6	D01	300 nm SiO_2	17.	600.	149.569	4.5324
7	D02	300 nm SiO_2	17.	700.	134.461	2.61678
8	D03	300 nm SiO_2	17.	800.	1042.45	181.296
9	D04	300 nm SiO_2	17.	900.	27606.8	15840.9
10	D06	300 nm SiO_2	17.	0.	133.479	2.47296
11	E01	300 nm SiO_2	63.5	600.	34.7786	0.340182
12	E02	300 nm SiO_2	63.5	700.	18.5526	0.145696
13	E03	300 nm SiO_2	63.5	800.	16.8454	0.369143
14	E04	300 nm SiO_2	63.5	900.	17.7519	0.321556
15	E06	300 nm SiO_2	63.5	0.	37.4678	0.41126
16	F01	300 nm SiO_2	30.	600.	83.0336	0.891628
17	F02	300 nm SiO_2	30.	700.	41.8038	0.0943494
18	F03	300 nm SiO_2	30.	800.	46.5629	1.71594
19	F04	300 nm SiO_2	30.	900.	221.937	21.4959
20	F06	300 nm SiO_2	30.	0.	67.6989	0.900036
21	G01		200.	600.	5.52953	0.119916
22	G02		200.	700.	1.58634	0.0207701
23	G03		200.	800.	2.49131	0.0301782
24	G04		200.	900.	2.66203	0.0228126
25	G06		200.	0.	9.06631	0.0228126
26	H01		200.	600.	3.11527	0.0346168
27	H02		200.	700.	601.298	49.925
28	H06		200.	0.	4.62456	0.0525315
29	I02		56.	700.	20.4109	0.265574
30	I03		56.	800.	17.8425	0.171594
31	I05		56.	0.	36.7427	0.308514
32	J02		111.	700.	8.36983	0.214193
33	J05		111.	0.	18.4469	0.353992
34	K02		223.	700.	4.36319	0.123072
35	K03		223.	800.	4.0429	0.0437089
36	K05		223.	0.	9.0784	0.233452
37	K10		223.	800.	—	—
38	K11		223.	800, 5 min	—	—
39	L01		54.	0.	34.5	—
40	L02		54.	0.	34.5	—

Table A.4: Superconducting critical temperature (T_c), residual resistance ratio (RRR) and low-temperature resistance ($R(20\text{ K})$) of some of the samples listed in table A.3.

#	Sample	T_c	$E(T_c)$ (10,90%)	RRR	$E(\text{RRR})$	Normalized $R(20\text{ K})$	Normalized $E(R(20\text{ K}))$
2	C02B	10.9133	{-0.176991, 0.355359}	3.10273	0.0417793	2.30049	0.0267165
3	C03B	11.7983	{-0.194216, 0.363596}	4.31187	0.0846629	1.48415	0.027792
12	E02	9.20809	{0.120734, -0.734057}	3.23786	0.0430004	2.32326	0.0278059
13	E03	8.51099	{0.205673, -0.0168385}	4.76806	0.0831607	1.63769	0.0274138
17	F02	10.0402	{-0.210082, 0.40793}	3.3369	0.0265089	4.54692	0.0270185
18	F03	10.9782	{-0.19664, 0.39647}	5.12474	0.0570779	3.0604	0.0296731
29	I02	10.6875	{-0.229926, 0.492383}	3.16905	0.0338262	2.58731	0.0249651
30	I03	11.7113	{-0.203699, 0.382762}	4.46587	0.0819486	1.57074	0.0260369
34	K02	7.84126	{-0.121149, 0.270682}	3.03805	0.125333	0.609604	0.0242081
37	K10	12.7378	{-0.274433, 0.374017}	4.98353	0.0520608	0.867358	0.00652876
38	K11	13.1118	{-0.473417, 0.475283}	5.08479	0.0680129	1.10151	0.0130036
39	L01	0.918545	{-0.00185573, 0.00262192}	0.965418	0.000770539	33.5135	0.0267525
40	L02	0.918545	{-0.00185573, 0.00262192}	0.965418	0.000770539	33.5135	0.0267525

Table A.5: V₃Si compound deposition.

#	Sample	Substrate	V ₃ Si deposition (nm)	RTA (°C)	R_{\square}	$E(R_{\square})$
1	B2P03A	Si + HF	20	500	103.43	10.0436
2	B2P03B	Si + HF	20	550	42.5054	2.12527
3	B2P03C	Si + HF	20	600	23.973	0.153255
4	B2P03D	Si + HF	20	650	24.3273	0.259712
5	B2P03E	Si + HF	20	700	23.0238	0.259712
6	B2P03F	Si + HF	20	750	22.2728	0.431383
7	B2P04A	Si + HF	50	500	35.1378	1.0697
8	B2P04A	Si + HF	50	500	35.1378	1.0697
9	B2P04B	Si + HF	50	550	23.9447	0.649281
10	B2P04B	Si + HF	50	550	23.9447	0.649281
11	B2P04C	Si + HF	50	600	10.6972	0.534848
12	B2P04C	Si + HF	50	600	10.6972	0.534848
13	B2P04D	Si + HF	50	650	8.26021	0.088482
14	B2P04E	Si + HF	50	700	7.80682	0.098162
15	B2P04F	Si + HF	50	750	7.43844	0.112459
16	B2P05B	Si + HF	200	600	269.909	114.647
17	B2P05C	Si + HF	200	650	5.67447	3.25166
18	B2P05C	Si + HF	200	650	5.67447	3.25166
19	B2P05C	Si + HF	200	650	5.67447	3.25166
20	B2P05D	Si + HF	200	700	1.56703	0.0171784
21	B2P05E	Si + HF	200	750	1.99634	0.00649281
22	B2P05F	Si + HF	200	800	2.1196	0.0129856
23	B2P05G	Si + HF	200	850	2.24003	0.0542673
24	B2P05H	Si + HF	200	900	2.48798	0.0249059
25	B2P05J	Si + HF	200	500	–	–
26	B2P05K	Si + HF	200	550	–	–
27	B2P05L	Si + HF	200	500, 5 min	–	–
28	B2P05Q	Si + HF	200	500, 5 min	–	–
29	B2P05R	Si + HF	200	500, 1 s	–	–
30	B2P06A	Si + HF	100	500	17.0447	0.490196
31	B2P06B	Si + HF	100	550	13.3892	0.212527
32	B2P06C	Si + HF	100	600	5.5257	2.97538
33	B2P06D	Si + HF	100	650	3.99551	0.224917
34	B2P06E	Si + HF	100	700	3.47127	0.0245405
35	B2P06F	Si + HF	100	750	3.27292	0.0425054
36	B2P08B	Si + 20nm SiO ₂	200	600	205.726	68.6275
37	B2P08C	Si + 20nm SiO ₂	200	650	4.05218	0.578656
38	B2P08D	Si + 20nm SiO ₂	200	700	4.16553	0.112459
39	B2P08D	Si + 20nm SiO ₂	200	700	4.16553	0.112459
40	B2P08E	Si + 20nm SiO ₂	200	750	4.00968	0.088482
41	B2P08F	Si + 20nm SiO ₂	200	800	3.85382	0.088482
42	B2P08G	Si + 20nm SiO ₂	200	850	3.85382	0.088482
43	B2P08H	Si + 20nm SiO ₂	200	900	3.90908	0.0446475
44	B2P08I	Si + 20nm SiO ₂	200	900	–	–
45	B2P08P	Si + 20nm SiO ₂	200	800	–	–

Table A.5: (Continued) V₃Si compound deposition.

#	Sample	Substrate	V ₃ Si deposition (nm)	RTA (°C)	R_{\square}	$E(R_{\square})$
46	B2P09A	Si + 20nm SiO ₂	50 + 10 nm Si	500	–	–
47	B2P09B	Si + 20nm SiO ₂	50 + 10 nm Si	550	–	–
48	B2P09C	Si + 20nm SiO ₂	50 + 10 nm Si	600	–	–
49	B2P09D	Si + 20nm SiO ₂	50 + 10 nm Si	650	–	–
50	B2P09E	Si + 20nm SiO ₂	50 + 10 nm Si	700	–	–
51	B2P09F	Si + 20nm SiO ₂	50 + 10 nm Si	750	–	–
52	B2P09G	Si + 20nm SiO ₂	50 + 10 nm Si	800	–	–
53	B2P09H	Si + 20nm SiO ₂	50 + 10 nm Si	850	–	–
54	B2P09I	Si + 20nm SiO ₂	50 + 10 nm Si	–	–	–
55	B2P17B	Si	200	600	280.111	136.952
56	B2P17C	Si	200	650	2.81952	0.789884
57	B2P17D	Si	200	700	1.58545	0.0297538
58	B2P17E	Si	200	750	2.2117	0.0371365
59	B2P17F	Si	200	800	2.39447	0.0507104
60	B2P17G	Si	200	850	2.5149	0.0401745
61	B2P17H	Si	200	900	2.71184	0.0449835
62	B3P11A	Si + Amorphisation + HF	200	600	8.50108	–
63	B3P11B	Si + Amorphisation + HF	200	650	–	–
64	B3P11C	Si + Amorphisation + HF	200	700	20.686	5.14181
65	B4PP02	Si + PS5	200	700	3.91758	0.26573
66	B4PP03	Si + PS5	200	800	3.84107	0.0538215
67	B5PSA	Sapphire	200	600	5.6381	–
68	B5PSB	Sapphire	200	650	–	–
69	B5PSC	Sapphire	200	700	–	–
70	B5PSD	Sapphire	200	750	–	–
71	B5PSE	Sapphire	200	800	–	–
72	B5PSE	Sapphire	200	800	–	–
73	B5PSE	Sapphire	200	800	–	–
74	B5PSF	Sapphire	200	850	–	–
75	B5PSG	Sapphire	200	900	3.90908	–
76	B5PSJ	Sapphire	200	–	9.2095	–

Table A.6: Superconducting critical temperature (T_c), residual resistance ratio (RRR) and low-temperature resistance ($R(20\text{ K})$) of some of the samples listed in table A.5.

#	Sample	T_c	$E(T_c)$ (10,90%)	RRR	$E(\text{RRR})$	Normalized R(20K)	Normalized $E(R(20\text{K}))$
16	B2P05B	9.39975	{-0.225125, 0.3252}	2.06268	0.00234758	130.854	55.5819
17	B2P05C	10.6252	{-0.300275, 0.37477}	2.7949	0.0111599	2.03029	1.16346
18	B2P05C	3.21457	{-0.149301, -0.374715}	2.79629	0.0070044	2.02929	1.16286
19	B2P05C	8.42608	{-1.12277, 0.744664}	—	—	—	—
25	B2P05J	2.64123	{-0.773577, 1.77471}	0.976846	0.000912291	14.4724	0.011331
26	B2P05K	7.81542	{-0.1994, 0.226192}	1.51343	0.00666554	5.72973	0.0247937
27	B2P05L	7.06625	{-0.201275, 0.323417}	1.19422	0.000981507	9.37858	0.00750012
28	B2P05Q	8.0154	{-0.1737, 0.250225}	1.62802	0.00389435	0.00563819	7.126726579471347* ⁻⁶
31	B2P06B	6.99141	{-0.650431, 0.574157}	1.31487	0.000501889	10.1829	0.16168
32	B2P06C	9.36222	{-0.202775, 0.353863}	2.27716	0.0102915	2.42658	1.30667
36	B2P08B	9.17528	{-0.250425, 0.299725}	1.85604	0.00841305	110.841	36.9786
37	B2P08C	10.6248	{-0.25025, 0.350187}	2.26837	0.0145972	1.78639	0.255358
38	B2P08D	10.7321	{-0.238435, 0.529402}	2.73546	0.0176014	1.52279	0.0422633
39	B2P08D	8.33437	{-0.460382, 0.461381}	—	—	—	—
40	B2P08E	11.8377	{-0.324225, 0.374675}	3.58144	0.0238104	1.11957	0.0258034
41	B2P08F	12.4874	{-0.300166, 0.350533}	4.52791	0.0320953	0.851126	0.0204526
42	B2P08G	13.1619	{-0.474246, 0.425721}	5.80453	0.0228682	0.663934	0.0154681
43	B2P08H	12.9878	{-0.40045, 0.475787}	7.97082	0.0562111	0.490424	0.006596
44	B2P08I	12.6124	{-0.2502, 0.4626}	6.5462	0.0676341	1.01451	0.00928348
45	B2P08P	3.0011	{-0.00183321, 0.149609}	4.50756	0.0197293	1.68563	0.00723219
47	B2P09B	7.91395	{-0.201025, 0.301025}	1.60018	0.00911362	19.233	0.108694
48	B2P09C	8.76345	{-0.2242, 0.293588}	1.97441	0.0173875	14.1185	0.123979
49	B2P09D	9.61453	{-0.200125, 0.352287}	2.2364	0.0267886	12.6645	0.152503
50	B2P09E	11.8152	{-0.301651, 0.400862}	3.09353	0.061513	10.0559	0.202928
51	B2P09F	12.3402	{-0.350916, 0.449673}	3.76488	0.0987009	8.19155	0.219781
52	B2P09G	11.941	{-0.324925, 0.475233}	4.75268	0.0394714	6.20588	0.0419948
53	B2P09H	12.241	{-0.274217, 0.450363}	5.53965	0.0538192	4.34384	0.0359899
55	B2P17B	9.61425	{-0.201925, 0.325233}	2.03642	0.00230517	137.551	67.2518
56	B2P17C	10.6754	{-0.30025, 0.424525}	2.88818	0.0234438	0.97623	0.273604
62	B3P11A	9.2303	{-0.26265, 0.286975}	1.85695	0.00508018	4.57798	0.0125477
63	B3P11B	10.9663	{-0.30055, 0.27475}	2.6311	0.00418009	1.90483	0.0028325
65	B4PP02	11.0002	{-0.2753, 0.49984}	2.76095	0.0130342	1.41893	0.0964786
66	B4PP03	12.6254	{-0.375513, 0.474538}	4.66919	0.0372196	0.822642	0.0132642
67	B5PSA	11.0747	{-0.200125, 0.250125}	2.04521	0.0167994	2.75673	0.0227421
68	B5PSB	11.9254	{-0.200475, 0.24955}	2.45945	0.0168331	2.10274	0.00982798
69	B5PSC	12.6747	{-0.199525, 0.1997}	3.04876	0.0317685	1.58516	0.0140517
70	B5PSD	13.7998	{-0.175025, 0.175225}	3.71176	0.0416517	0.00171005	0.0000161791
71	B5PSE	14.4515	{-0.1525, 0.173175}	4.64903	0.0672126	0.00138091	0.0000182585
74	B5PSF	14.9495	{-0.1992, 0.17575}	6.21999	0.0615928	0.786249	0.00536839
75	B5PSG	15.3249	{-0.199625, 0.224892}	8.39993	0.30363	0.46537	0.0174079
76	B5PSJ	1.19738	{-0.243652, 0.431044}	0.955918	0.000396282	9.63419	0.00399455

Appendix B

Devices measured at low temperature

Some tips for future students:

- Don't go above $|V_g| = 2\text{ V}$ at room temperature.
- Devices that leak through the gate at room temperature, will still do so when cooled down. However, devices that seem shorted from source to drain often come back to life at 4 K (remember: we are *supposed* to have enormous OFF current in a device with all-p doping).
- Use silver paste to glue the sample.
- Ground yourself, the bonding machine and the sample holder to a common ground before bonding anything.
- Always bond the substrate contact first, and the gate last. Bond all substrate pins to the back plate, so that all devices can be grounded together without the RC delay of the filters.
- Keep the substrate grounded while cooling down. If no LED is used, the substrate will be frozen out at low temperatures, and any accumulated charges can no longer be evacuated.
- Only apply a magnetic field once all the zero-field measurements are finished. It is absolutely impossible to get the coil back to $H = 0$ without warming it up, and the measurements are extremely sensitive to the field.

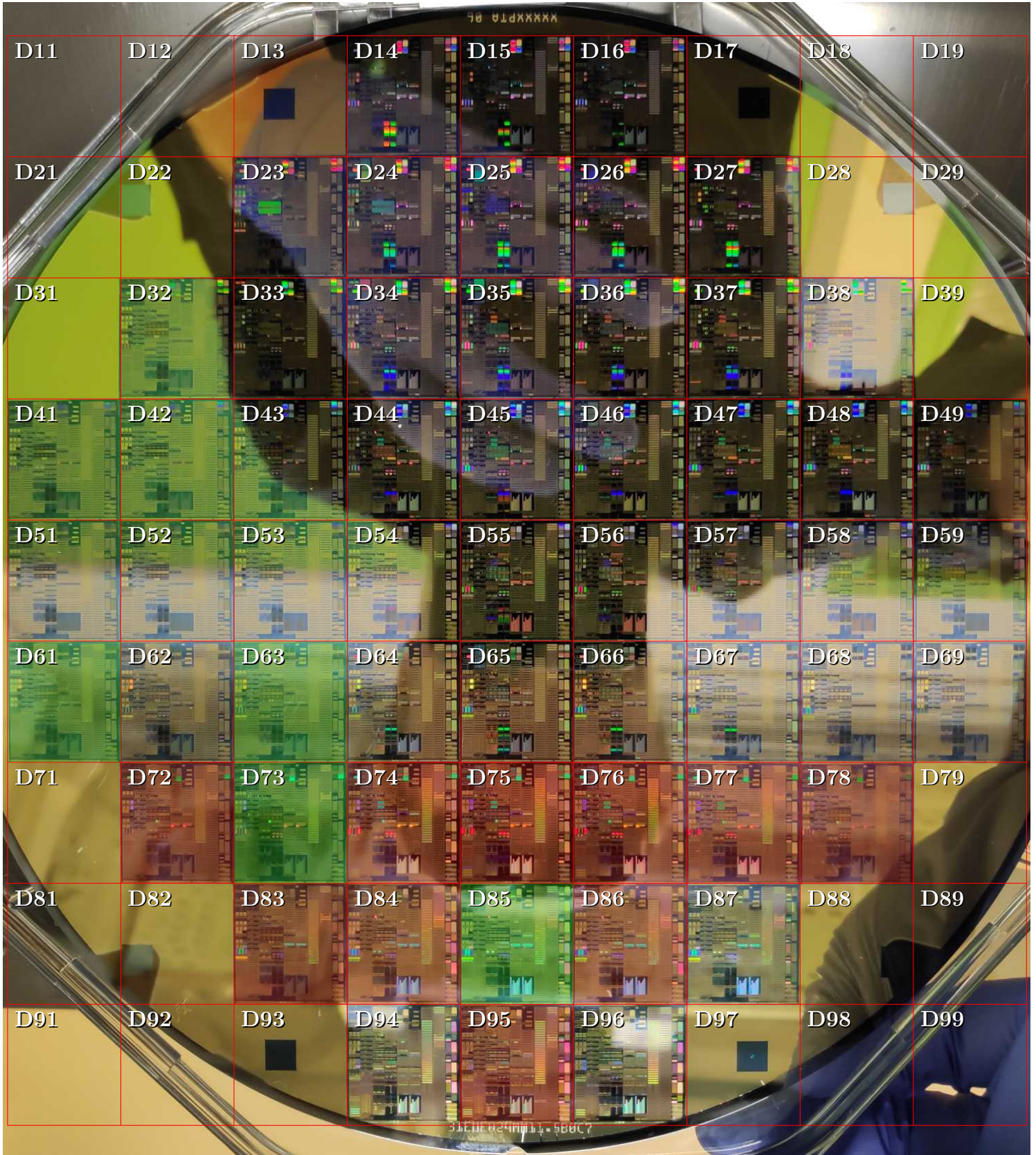


Figure B.1: Wafer PTA06 provided by Laurie Calvet, manufactured by Chinlee Wang and John Snyder at National Semiconductor. Dies with successful measurements at low temperature indicated in green, dies with only shorted/leaking devices in red.

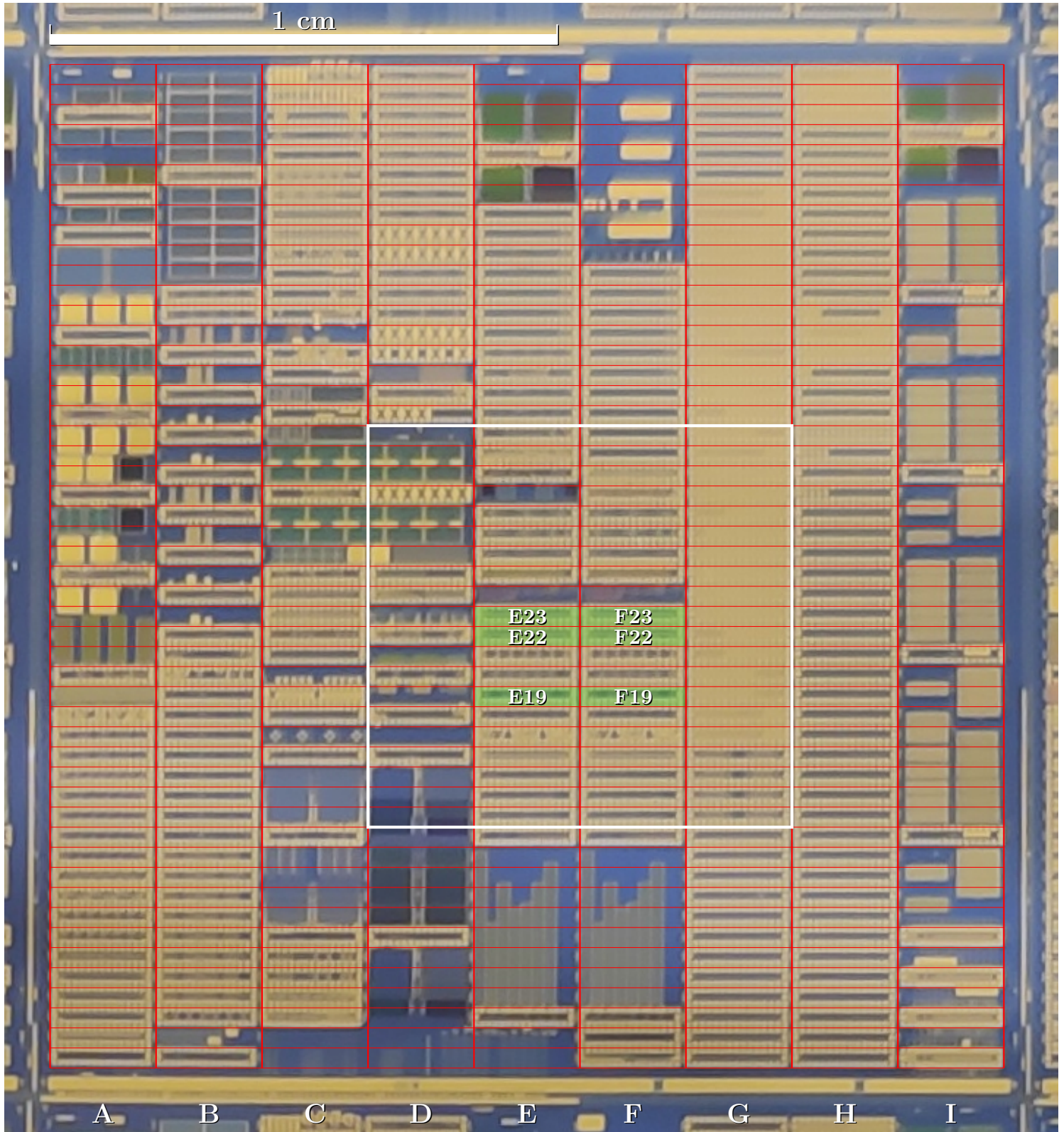
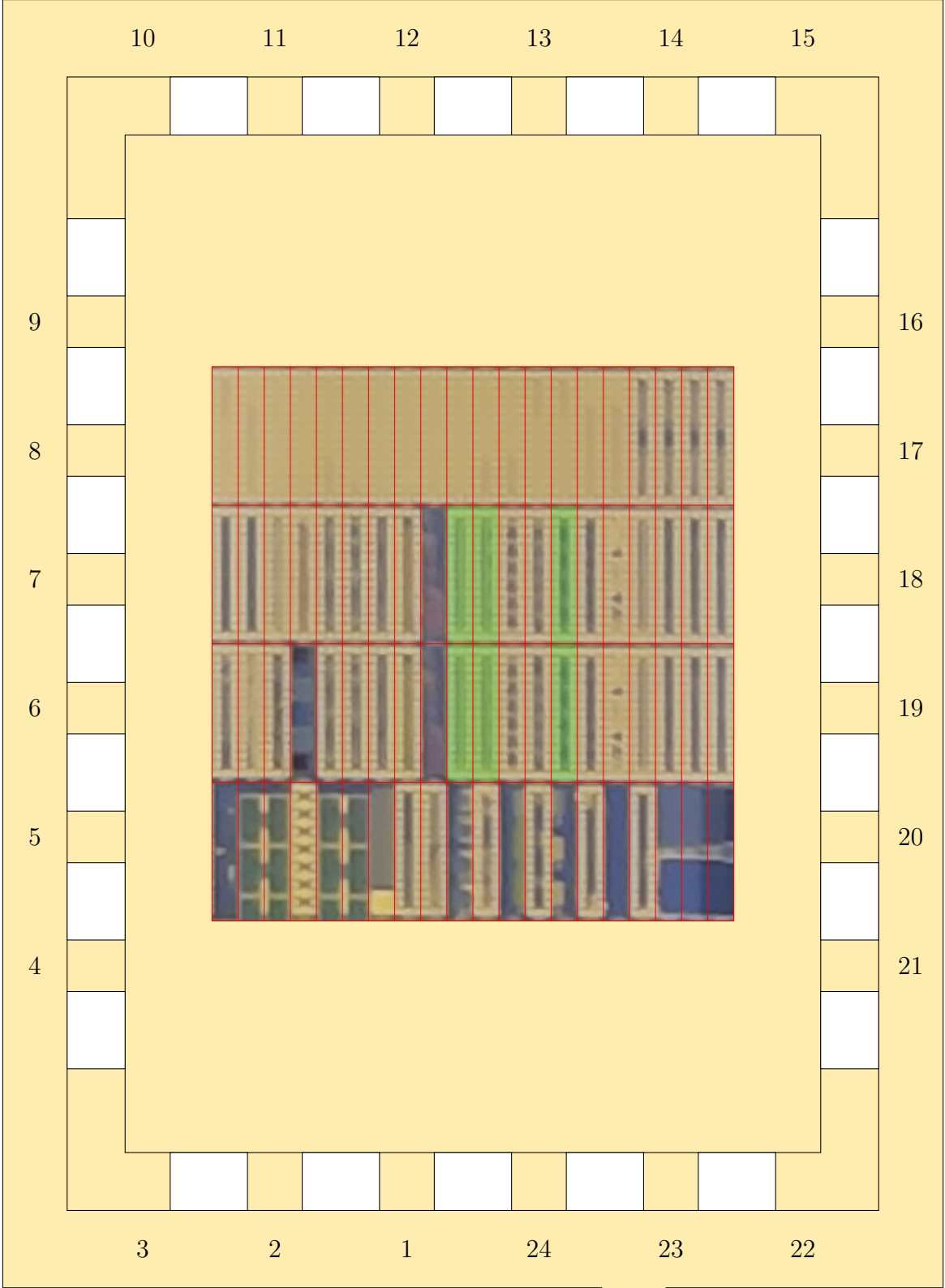


Figure B.2: Each die has 9 columns (letters) and 50 rows (numbers). Scribes E,F:19,22,23 contain the SBMOSFETs that were measured. The white square was cut out with a dicing saw and glued to a KYOCERA sample holder with silver paste. A bonding template is provided on the next page.



	Sub			
Gate	31	32	1	Gate
Drain	30		2	Drain
Source	29		3	Source
Gate	28		4	Gate
Drain	27		5	Drain
Source	26		6	Source
Gate	25		7	Gate
Drain	24		8	Drain
Source	23		9	Source
Gate	22		10	Gate
Drain	21		11	Drain
Source	20		12	Source
Gate	19		13	Gate
Drain	18		14	Drain
Source	17	16	15	Source

All Transistors are 20/0.05

		Source connected transistors			Drain of connected transistor:	Drain of connected transistor:	Source connected transistors			Drain of connected transistor:	
20/0.05	{	Drain	31	32	1	0.2 / 0.24	0.3 / 0.24	31	32	1	0.26 / 0.24
		Source	30		2	0.2 / 0.24	0.6 / 0.21	30		2	0.2 / 0.21
		Gate	29		3	0.24 / 0.24	3 / 0.24	29		3	0.20 / 0.24
2.5/0.05	{	Drain	28		4	0.28 / 0.24	0.3 / 0.3	28		4	0.20 / 0.27
		Source	27		5	2 / 5	3 / 0.3	27		5	20 / 0.3
		Gate	26		6	1 / 5	0.3 / 0.38	26		6	20 / 0.38
5/0.3	{	Drain	25		7	0.5 / 5	0.6 / 0.5	25		7	20 / 5
		Source	24		8	0.4 / 5	0.3 / 0.5	24		8	20 / 1
		Gate	23		9	0.3 / 5	0.3 / 1	23		9	20 / 1.5
10/0.7	{	Drain	22		10	0.28 / 5	0.3 / 1.5	22		10	20 / 2
		Source	21		11	0.26 / 5	3 / 1.5	21		11	20 / 5
		Gate	20		12	0.24 / 5	0.3 / 2	20		12	20 / 20
20/1.7	{	Drain	19		13	0.22 / 5	0.6 / 5	19		13	0.6 / 20
		Source	18		14	0.2 / 5	3 / 20	18		14	0.3 / 20
		Gate	17	16	15	Sub	0.26 / 20	17	16	15	Sub
		Gate connected transistors					Gate connected transistors				

(a) Quads E,F:19.

(b) Quads E,F:22.

(c) Quads E,F:23.

Figure B.3: Layouts of quads E19 through F23, adapted from Ref. 1 (note that on wafer PTA06, **all** quads have PMOS devices).

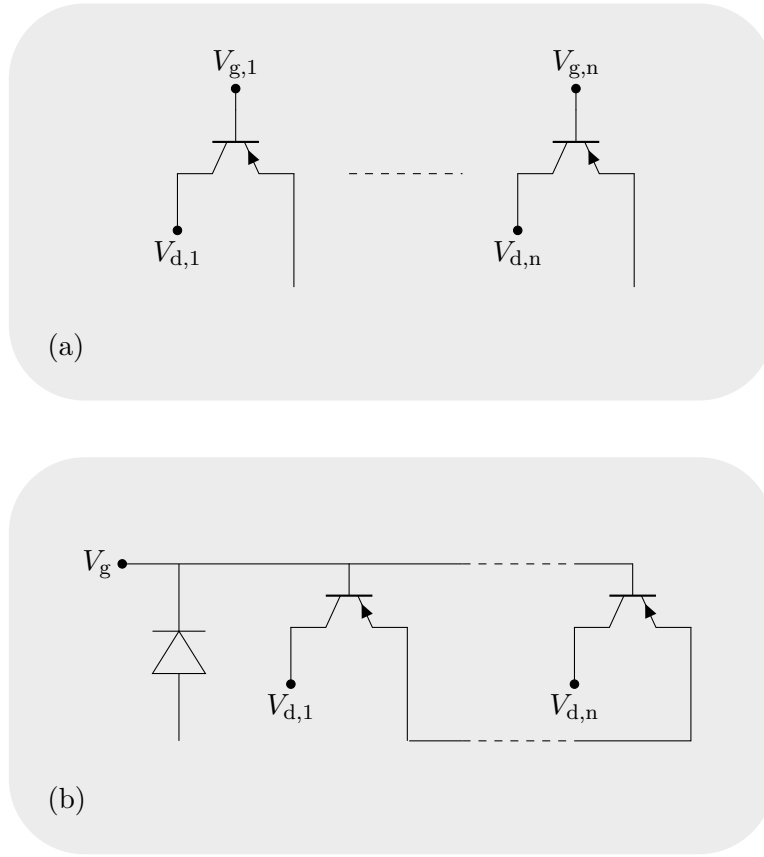


Figure B.4: (a) The separate PMOS transistors (all devices on E,F:19; drains 19–31 on E,F:22) each have their own source, drain and gate contact pads. A voltage was systematically applied to the drain, and the current measured from the source. (b) The connected PMOS transistors (drains 1–14 on E,F:22, all devices on E,F:23) share the source and the gate, each has its own drain contact. A parallel reverse biased ESD diode prevents large currents from flowing across the gate.

Bibliography

- [1] Laurie Ellen Calvet. “Electrical transport in Schottky barrier MOSFETs”. PhD thesis. Yale University, 2001.