Emerging devices and materials for nanoelectronics

Silicon spin qubits and MoS$_2$ thin-film transistors

Davide Rotta

A thesis presented for the degree of
Doctor of Philosophy

Dipartimento di Scienza dei Materiali
Università di Milano Bicocca
Milan, 23 February 2015

Tutor: Prof. Marco Fanciulli
Coordinator: Prof. Gianfranco Pacchioni
Sommmario

Questa tesi analizza la possibile implementazione di due tipologie di dispositivi elettronici con funzionalità innovative: dispositivi per la computazione quantistica e transistors a film sottile.

Negli ultimi decenni l’industria dei semiconduttori ha portato alla realizzazione di circuiti integrati con milioni di transistors e performance sempre migliori a costi contenuti. Tuttavia, questo processo di miniaturizzazione è giunto a un punto tale che i dispositivi elettronici sono ora composti da pochissimi atomi e ridur- ne ulteriormente le dimensioni sta diventando sempre più difficile. L’International Technology Roadmap of Semiconductors (ITRS) suggerisce due vie alternative per migliorare le caratteristiche dei dispositivi a partire dalla Front-End-Of-Line. La prima si avvale di nuovi dispositivi sulla base di architetture innovative o dell’utilizzo di diverse variabili di stato (Emerging Research Devices), mentre la seconda punta all’utilizzo di nuovi materiali (Emerging Research Materials).

Questa tesi esamina due possibili candidati in quest’ottica: i dispositivi su architettura Complementary Metal-Oxide-Semiconductor (CMOS) per la computazione quantistica e i transistors a film sottile basati su un semiconduttore bidimensionale come il MoS$_2$. Da un lato, l’integrazione della computazione quantistica su Si sfrutterebbe il background tecnologico dell’industria dei semiconduttori per implementare su larga scala un nuovo protocollo di computazione dotato di un potenziale enorme e ancora inesplorato. D’altra parte il disolfuro di molibdeno (MoS$_2$) è intrinsecamente scalabile, in quanto può essere esfoliato fino allo spesso- re di un singolo strato atomico. Per questo motivo potrebbe essere un semicondu- tore ideale per dispositivi elettronici ultrascalati, così come per applicazioni nella sensoristica, nell’optoelettronica e nell’elettronica flessibile.

Questo lavoro mostra l’attività svolta al Laboratorio MDM-IMM-CNR nell’ambito del corso di dottorato in Nanostrutture e Nanotecnologie all’Università di Milano Bicocca. Lo sviluppo e l’utilizzo di processi di fabbricazione della nanoelettronica, in particolare la litografia a fascio elettronico (EBL), sono stati
parte integrante dell’attività sperimentale dedicata alla realizzazione di dispositivi CMOS-compatibili per la computazione quantistica e per l’integrazione di film sottili di MoS$_2$ in strutture Metal-Oxide-Semiconductor Field-Effect Transistor (MOS FET).

I necessari passi di processo sono stati adeguatamente calibrati e ottimizzati in modo da ottenere dispositivi quantistici basati su Quantum Dots (QD) con dimensioni características inferiori a 50 nm. Tali dispositivi sono stati sviluppati con tecnologia Silicon-On-Insulator (SOI), mantenendo così la compatibilità con lo standard della tecnologia CMOS.

Dispositivi a singolo donore e con QD di silicio sono stati poi caratterizzati elettricamente a temperature criogeniche (fino a 300 mK). Impulsando i potenziali di gate in modo controllato, è stato possibile studiare fenomeni di tunneling di singoli elettroni su un donore in alti campi magnetici (8T). In modo analogo è stato dimostrato il controllo dello stato di carica di QDs di Si. In particolare, si è osservato l’insorgere di rumore telegrafico associato al movimento di un singolo elettrone tra due QDs. Infine è stato condotto uno studio di fattibilità per l’integrazione su larga scala di un’architettura di computazione quantistica (il cosiddetto hybrid spin qubit) basata su doppi QDs di Si.

Sul secondo fronte sono stati realizzati dei MOS FETs a film sottile basati su frammenti di MoS$_2$, ottenuti per esfoliazione meccanica e contattati elettricamente tramite litografia EBL. Tali transistors sono poi stati caratterizzati elettricamente, con particolare attenzione alle proprietà di trasporto di carica e alla spettroscopia delle trappole all’interfaccia con l’ossido.
Abstract

This work of thesis explores two emerging research device concepts as possible platforms for novel integrated circuits with unconventional functionalities.

Nowadays integrated circuits with advanced performances are available at affordable costs, thanks to the progressive miniaturization of electronic components in the last decades. However, bare geometrical scaling is no more a practical way to improve the device performances and alternative strategies must be considered to achieve an equivalent scaling of the functionalities. The introduction of conceptually new devices and paradigms of information processing (Emerging Research Devices) or new materials with unconventional properties (Emerging Research Materials) are viable approaches, as indicated by the International Technology Roadmap of Semiconductors (ITRS), to enhance the functionalities of integrated circuits at the Front-End-Of-Line.

The two options investigated to this respect are silicon devices for quantum computation based on a classical Complementary Metal-Oxide-Semiconductor (CMOS) platform and standard Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs) based on MoS$_2$ thin film. In particular, the integration of Quantum Information Processing (QIP) in Si would take advantage of Si-based technology to introduce a completely new paradigm of information processing that has the potential to outperform classical computers in some computational tasks, like prime number factoring and the search in a big database. MoS$_2$, conversely, can be exfoliated up to the single layer thickness. Such intrinsic and extreme scalability makes this material suitable for end-of-roadmap ultrascaled electronic devices as well as for other applications in the fields of sensors, optoelectronics and flexible electronics.

This work reports on the experimental activity carried out at Laboratory MDM-IMM-CNR in the framework of the PhD school on Nanostructures and Nanotechnology at Università di Milano Bicocca. Electron Beam Lithography (EBL) and mainstream clean-room processing techniques have been intensively utilized to
fabricate CMOS devices for QIP on the one hand and to integrate mechanically exfoliated MoS$_2$ flakes in a conventional FET structure on the other hand.

After a careful calibration and optimization of the process parameters, several different Quantum Dot (QD) configurations were designed and fully realized, achieving critical dimensions under 50 nm. Such device architectures were developed on a Silicon-On-Insulator (SOI) platform, in order to eventually access a straightforward integration into the CMOS mainstream technology.

Si-QDs and donor-based devices have been then tested by electrical characterization techniques at cryogenic temperatures down to 300 mK. In detail, single electron tunneling events on a donor atom have been controlled by pulsed-gate techniques in high magnetic fields up to 8T, providing a preliminary characterization for the initialization procedure of donor qubits. The control of the charge states of Si-QDs have been also demonstrated by means of stability diagrams as well as the analysis of random telegraph noise arising from single electron tunneling between two QDs. Finally, a feasibility study for the large scale integration of quantum information processing was done based on a double QD hybrid qubit architecture.

On the other side, MoS$_2$ thin film transistors have been made by mechanical exfoliation of crystalline MoS$_2$ and electrodes definition by EBL. Electrical characterization was performed on such devices, with a particular focus on the electrical transport in a FET device and on the spectroscopy of interface traps, that turns out to be a limiting factor for the logic operation.
# Contents

Introduction

1 Emerging Research Devices: Silicon qubits
   1.1 Quantum Information Processing
      1.1.1 DiVincenzo Criteria for QIP
      1.1.2 Principal implementations
   1.2 Donor atoms in silicon
      1.2.1 Donor-based qubits in silicon
      1.2.2 Charge control and single electron dynamics on a donor atom
      1.2.3 Conclusions
   1.3 Silicon double quantum dots
      1.3.1 Double quantum dot qubits
      1.3.2 Design
      1.3.3 Fabrication: CMOS-compatible process integration
      1.3.4 Quantum transport characterization
      1.3.5 Conclusions
   1.4 CMOS-compatible Large Scale Integration of hybrid spin qubits
      1.4.1 Introduction
      1.4.2 Fundamental logic gates in the hybrid qubit architecture
      1.4.3 CMOS implementation of the hybrid qubit architecture
      1.4.4 Quantum computing on a large scale
      1.4.5 Conclusions

2 Emerging Research Materials: MoS\textsubscript{2}
   2.1 2D Materials: Flatlands Beyond Graphene
   2.2 MoS\textsubscript{2}
      2.2.1 Deposition of MoS\textsubscript{2}
      2.2.2 MoS\textsubscript{2} electronic devices
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3 MoS₂ MOSFET</td>
<td>65</td>
</tr>
<tr>
<td>2.3.1 MOS FET fabrication</td>
<td>65</td>
</tr>
<tr>
<td>2.3.2 Electrical transport in single gate MOS FETs</td>
<td>67</td>
</tr>
<tr>
<td>2.3.3 Admittance spectroscopy of interface traps</td>
<td>72</td>
</tr>
<tr>
<td>Conclusions</td>
<td>79</td>
</tr>
<tr>
<td>Appendices</td>
<td>82</td>
</tr>
<tr>
<td>A Complete process flow for Si-DQD devices</td>
<td>83</td>
</tr>
<tr>
<td>A.1 Detailed recipe</td>
<td>83</td>
</tr>
<tr>
<td>A.2 UV lithography</td>
<td>86</td>
</tr>
<tr>
<td>A.3 Electron Beam Lithography</td>
<td>87</td>
</tr>
<tr>
<td>A.4 Wet etching and cleaning</td>
<td>89</td>
</tr>
<tr>
<td>A.5 Atomic Layer Deposition</td>
<td>90</td>
</tr>
<tr>
<td>B Alternative layouts for Si-DQD devices</td>
<td>91</td>
</tr>
<tr>
<td>C List of acronyms</td>
<td>93</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>95</td>
</tr>
<tr>
<td>Bibliography</td>
<td>96</td>
</tr>
</tbody>
</table>
Introduction

The progressive scaling of integrated circuits is going to face its ultimate limits imposed by the atomic size of semiconductors. In the last decades, the continuous miniaturization of electronic components made it possible to realize electronic devices with unprecedented functionalities at affordable costs. Actual laptops and smartphones, for example, are now globally widespread and their actual size roughly resembles that of the very first transistor. However, such miniaturization has led semiconductor industry to a point where devices are composed by few atoms and further reduction of their critical dimensions is hardly possible. As a consequence, new methods have to be considered in order to overcome this physical limitation or to maintain performances without further miniaturization. This approach is also naturally open to introduce novel functionalities in a unique integrated circuit through the implementation of non-conventional materials and devices.

The International Technology Roadmap of Semiconductors (ITRS) indicates two different ways to enhance the functionalities of integrated circuits by improving the elementary components at the Front-End-Of-Line: introducing new materials with anomalous properties (Emerging Research Materials) or designing conceptually new devices and paradigms of information processing (Emerging Research Devices). The panorama of proposed materials and devices as candidates for the integration in the next technological nodes is as wide as variegated and research centers as well as industry are carefully considering all the available options. This thesis is situated in such a dynamic context and it aims at exploring two possible evolutions of Si-based Complementary Metal-Oxide-Semiconductor (CMOS) technology for future electronics, according to the presented schemes of Emerging Research Devices and Materials. The two options investigated in this respect are quantum devices for quantum computation based on a classical Si-CMOS platform and standard Metal-Oxide-Semiconductor Field-Effect-Transistors (MOSFETs) based on MoS$_2$ thin film as the channel material.
Consequently, this thesis describes the activity carried out at Laboratory MDM-IMM-CNR in the framework of the PhD school on Nanostructures and Nanotechnology at Università di Milano Bicocca. It shows the experimental results obtained for the realization of Si-MOS devices for quantum computation and MoS$_2$ thin film MOSFETs by means of Electron Beam Lithography and mainstream clean-room processing. Electrical characterization of such devices has been carried out from room temperature down to 300 mK to explore and assess their effective functionalities.

**CMOS Scaling of integrated circuits**

The foundations of microelectronics dates back to the discovery of the transistor by J. Bardeen, W. H. Brattain and W. B. Shockley in 1947 and the invention of integrated circuits by L. S. Kilby in 1958-1959, that have been awarded with the Nobel prize in 1956 and 2000 respectively for the technological impact of their works. The realization of the first integrated circuit, in particular, can be considered as the first milestone in microelectronics. The idea of monolithic integration became a key concept to build up complex circuits by combining few device types, like resistors, diodes and transistors, in order to obtain novel functionalities.

In the same years, the realization of the first Metal-Oxide-Semiconductor Field-Effect-Transistor (MOSFET) set the basis for the development of the Complementary Metal-Oxide-Semiconductor (CMOS) technology. CMOS has become in the years the dominant standard in semiconductor industry, due to the reliability and versatility of the fabrication process and the scalability of low power logic gates.

In 1965, Moore predicted an exponential trend for the reduction of device dimensions with the incoming years driven by an economical principle of cost reduction. Such scaling law has been respected up to about 2000, guiding microelectronics in its first era of scaling, characterized by a geometrical scaling of the device dimensions.

In the 90s it became clear that, for physical constraints, further geometrical scaling of devices could not be possible with the existing technology. The main challenges for ultrascaled MOSFET under 100 nm were tunneling currents through the oxide, short channel effects and drain-induced barrier lowering, besides the demanding requirements for low power consumption and high operating frequency.

Such challenges were addressed in this second era by leveraging the equivalent scaling approach (introducing high-$k$ gate oxides) and by integrating mobility
boosters (strained Si).

At present, analogous challenges for a further reduction of device dimensions are an obstacle towards the next technological nodes. As a consequence, the International Technology Roadmap for Semiconductors (ITRS) prospects a variegated landscape of the possible evolutions of microelectronics, with yet unpredictable outcomes in the medium-long term. Three different approaches are now considered to add value and functionality to the actual integrated circuits.

The first one is referred to as More Moore and relies on the continuous geometrical scaling of the conventional CMOS platform by taking advantage of 3D integration and technology boosters (the already cited high-\(k\) oxides, high mobility channels and ballistic gate electrodes).

The second method is called More than Moore and springs from the concurrent integration of different technologies to merge different functionalities (electronic, optical, photovoltaics, sensing) in a unique chip, like for example RF components, biological lab-on-chips and sensors or actuators, like Micro Electro Mechanical Systems (MEMS).

The third strategy to enhance the functionalities of integrated circuits is Beyond Moore. This route is based on the exploitation of new information processing paradigms involving different state variables like spin, spin waves and orbital states of electrons.

Each of the mentioned approaches offers exciting opportunities as well as technological challenges regarding fabrication and lithographic processes. Such an active field of nanoscience and nanotechnology is the context of this thesis, that aims at investigating the potential impact in microelectronics of a class of emerging devices (Si-MOS quantum devices for QIP) and of an emerging material (MoS\(_2\) employed as a channel material).

This thesis is organized as follows: after a brief introduction about the main topics of this work, the following three chapters are devoted to the presentation of experimental details and scientific results.

Chapter 1 is dedicated to Si quantum devices for quantum information processing: single atom devices and Si quantum dots. The fabrication process will be described in detail, as well as the main results obtained from the electrical characterization of the devices. In addition, the implementation of quantum computing on large scale circuits is examined in a feasibility study of CMOS-compatible implementation of the hybrid spin qubit.
Chapter 2 is committed to the realization and electrical characterization of thin film transistors based on MoS$_2$. The performances of such devices have been analyzed by standard techniques of electrical characterization, focusing on the electrical transport and the spectroscopy of interface traps in thin film transistors.

Finally, the main results of this work will be summarized in Chapter 2.3.3, drawing specific conclusions on the basis of the obtained results.
Chapter 1

Emerging Research Devices: Silicon qubits

Quantum computers have the potential to outperform classical ones for some specific tasks like prime number factoring, that is at the basis of the broadly used RSA encryption system, or searching records in a database. Some examples of quantum computers have been presented in literature up to now, controlling few qubits registers and performing easy tasks like the factorization of number 15 with a quantum algorithm. However, a full exploitation of the huge potential of quantum computing would be possible in systems with a large scale integration of such qubits. To this extent, a CMOS-compatible implementation of QIP would be extremely promising, since this approach could take advantage of the technological background developed by semiconductor industry. On the other hand, the paradigm of QIP is also a revolutionary scheme of information processing in the perspective of a Beyond Moore evolution of the ITRS roadmap.

This chapter is committed to the evaluation of a a CMOS-compatible implementation of QIP in Si. The theoretical basis of QIP and its main implementations will be presented in Section 1.1, with a particular attention to Si-based implementations. In particular, experimental results will be presented regarding the design, the realization and the electrical characterization of devices conceived for the implementation of two QIP architectures that are CMOS-compatible (donors and DQD qubits will be considered in Section 1.2 and Section 1.3 respectively). Finally, the LSI of a DQD architecture will be considered in Section 1.4, analyzing the time and space requirements for the implementation of fault-tolerant quantum computing with a hybrid spin qubit architecture.
1.1 Quantum Information Processing

The pioneering concept of QIP was introduced by the revolutionary genius of Richard Feynman in 1985: since the physics of nanoscopic systems is so complex, why not using quantum computers to simulate and understand quantum systems? Since then, QIP science made big steps and the applications of quantum computers overcame this first picture envisioned by Feynman, including metrology and the study of quantum physics foundations as well as quantum simulators [12].

Two-level systems, in particular, can be considered as a quantum counterpart of classical bits and much of the concepts and paradigms of classical information processing can be transposed to such quantum bits, namely qubits. However a crucial difference from classical bits is that, according to the laws of quantum mechanics, a qubit state can be found in each of the infinite complex linear superpositions of its 0 and 1 state. As a result, qubits are capable of a much larger information content than classical bits. In fact, when logical operations are carried out on qubits, such logical gates are actually performed on the 0 and the 1 components of the qubit state at the same time. A huge vector space can be consequently explored by performing unitary transformations on a set of N entangled qubits to find the right solution to a specific problem. An appropriate algorithm (i.e. an appropriate sequence of unitary transformations acting as logic gates) selects the desired results by quantum interference of the qubit quantum states, generating constructive interference for right solutions and destructive for wrong solutions [31]. Some computational tasks, that are impractical with classical computers, become feasible thanks to the intrinsically parallel computation performed by quantum algorithms, like the Shor’s algorithm for prime number factoring.

This extraordinary potential is also technologically relevant if contemplated in view of a Beyond CMOS extension of the ITRS roadmap. In this framework, QIP constitutes an alternative paradigm of information processing that in principle has the potential to outperform all the current classical computers without suffering from the geometrical limitations that electronic devices will meet at the end of roadmap. As a matter of fact, even large companies (IBM, Intel, HP, Hitachi) have a research department dedicated to QIP and great interest was recently gained by the Canadian D-Wave, that actually sold the first example of computer based on quantum annealing for the optimization of linear problems.

Concluding, QIP is an extraordinary example of the encounter between scientific and technologic interests. This encounter started an intriguing gold-rush in both fields with unpredictable evolutions in the next years. Few qubit circuits
have been demonstrated indeed, but scalability to practical multi-qubit systems is an issue still unaddressed to some extent. For these reasons, the next Sections will be devoted to QIP and in particular to Si-based QIP. Section 1.1.1 will outline the main requirements for the implementation of QIP, whereas Section 1.1.2 will explore the principal physical systems complying with such requisites and suitable for QIP.

1.1.1 DiVincenzo Criteria for QIP

David DiVincenzo displayed five key requirements for a quantum computer, also known as DiVincenzo criteria [32], with two additional requirements for quantum communication, that are now widely recognized as a metrics to compare different quantum systems for quantum information processing.

1. Scalable qubits A qubit must really be a "quantum bit" i.e. a quantum system defined by two states, that are generally identified as the $|0\rangle$ and $|1\rangle$ state of the qubit. Such two states span a two-dimensional Hilbert space, where a general qubit state can be written in the form:

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (1.1)$$

with $\alpha$ and $\beta$ complex numbers such as $|\alpha|^2 + |\beta|^2 = 1$. The qubit logical space has a convenient graphical representation in Figure 1.1, where the $|0\rangle$ and $|1\rangle$ states are placed at the poles of the Bloch sphere and all the qubit state can be addressed in terms of its polar coordinates $\phi$ and $\theta$.

A quantum system has generally more than two energy levels and, in addition, such levels can be degenerate in energy, complicating the definition of the logical basis of the qubit. A qubit can be obtained anyway by selecting two non-degenerate energy levels of the quantum system. In this framework, the system can be operated in particular conditions so that unwanted degeneracies are lifted and the interactions with levels other than the $|0\rangle$ and $|1\rangle$ states are significantly suppressed.

Generalizing to the case of $N$ qubits, such extended system can be described in terms of an $2N$-dimensional Hilbert space. However, if $N$ qubits are entangled, so that their dynamics can’t be described independently from each other, then the whole system is more correctly described by a much larger $2^N$-dimensional Hilbert space.
Figure 1.1: The Bloch sphere is a graphical representation of all the possible quantum states of a qubit. The two eigenstates defining the logical basis are set at the poles, while each other possible superposition is identified by its polar coordinates $\phi$ and $\theta$.

Notably, such an exponential scaling of the system complexity (and of the quantum information contained within) comes from just a linear increase of the physical resources of the quantum computer. This property is at the heart of the huge potential of QIP, that in principle allows to perform specific tasks (such as prime number factoring) with an exponential speed-up compared to classical information processing algorithms.

2. Initialization of qubit states Qubits must be initialized to a known quantum state before starting computation, e.g. all qubits set in the $|0\rangle$ state. Then computation can be carried on, according to the required algorithm, provided that the qubit register involved in the computation is effectively initialized in the needed initial state, namely the $|000\ldots0\rangle$ state. If the register is not correctly prepared, computation is performed on a wrong state, producing erroneous outcomes. The initialization of a qubit can be carried out by manipulating the system in a specific configuration where the ground state is exactly $|0\rangle$ and waiting for a time long enough to let the system relax on its ground state. Alternatively, this process can be sped up by making the qubit state collapse onto the $|0\rangle$ state through a projection measurement.

3. Long decoherence times This requirement is at the core of quantum information processing: while other criteria are some kind of generalization of requirements for classical computation, this criterion has not a classical coun-
As noted before, quantum bits store much more information than classical ones, but such quantum information is more fragile (error probability in classical computers is almost null instead, being of the order of $10^{-17} - 10^{-18}$ per logic operation) and there are additional ways in which quantum information can be lost due to interactions with the surrounding environment [31, 68]. Besides "classical" bit-flip errors (due to the energy relaxation of the $|1\rangle$ state to the $|0\rangle$ state) coupling with the environment leads to an uncontrolled evolution of the quantum system, causing phase errors that are intrinsically quantum errors.

For example, a spin in a solid crystal initialized in the $|\uparrow\rangle$ state is expected to flip to the $|\downarrow\rangle$ in a timescale $T_1$ that is governed by the effective relaxation mechanisms. The time evolution of the spin state is driven by the intensity and direction of the magnetic field. Consequently, random fluctuations of the local magnetic field (due to the spin-flipping of the surrounding nuclei) alter the expected evolution of the system and limit the coherent evolution of the system to the characteristic time $T_2$. Besides this, the measurement of the qubit final state is normally repeated and averaged over many equivalent samples to perform a complete tomography of a qubit. Such an ensemble measurement experiences also the slow variations of the quantum oscillator frequency, that don’t affect the qubit coherence over a single measurement but after multi-sample averaging results in an apparent dephasing on a $T_2^*$ timescale [68].

For this reason, computational running time must be significantly shorter than the relevant dephasing time in order to reduce the probability of errors occurring during the computation. Nonetheless, Quantum Error Correction (QEC) techniques are generally required at any rate, in order to correct for quantum errors occurring during the computation while preserving the coherence of quantum information [134, 28]. The effectiveness of QEC schemes is strongly dependent on the specific encoding scheme and physical implementation of the qubit [110, 127, 130]. Standard QEC schemes tolerate an error threshold of about $10^{-5} - 10^{-4}$ errors per logic gate, but higher values ($\sim 1\%$) are allowed by topological codes [28].

4. **Universal set of quantum gates** A set of unitary transformations $U_1, U_2, U_3,...$ acting on few qubits must be available to perform a general quantum al-
gorithm. In other words, a complete set of logic operations should allow to explore the entire $2^N$-dimensional vector space identified by N qubits. Fortunately, this doesn’t mean that each of these transformation must necessarily act over the entire qubit register. Indeed, such a unitary transformation is normally decomposed in a set of unitary transformations acting on smaller qubit subsets, made of one to three qubits usually. As an example, 1-qubit (arbitrary rotations over a qubit Bloch sphere) and 2-qubit logic operations (like the Controlled-NOT, or CNOT) between all the couple of qubits suffice to obtain universal quantum computation.

5. **Read-out of qubit states** In the end of an algorithm, the outcomes of the computation must be made available. Actually, the read-out of a qubit state is more complicated than in the classical case, because the qubit state is generally a superposition $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$ of two possible results. In this case a perfect measurement (100% fidelity) should give $|0\rangle$ with a probability $|\alpha|^2$ and $|1\rangle$ with a probability $|\beta|^2$.

6. **Interconversion between stationary and flying qubits** The coherent transmission of quantum information i.e. quantum communication is a pretty different task compared to quantum computing and set different requirement for the implementation of flying qubits. As a consequence, generally the physical implementation of flying qubits devoted to the transmission of quantum information can be quite different from that of stationary data qubits, where the computation is performed [68]. For this reason, it is necessary to realize a module for such hybrid systems to interconvert stationary qubits (e.g. solid state qubits) and flying qubits (photons are ideal to this scope) preserving the coherence of quantum information [147].

7. **Transmission of flying qubits** The ability to perform the coherent transfer of quantum information over long distances is a key requirement for quantum computation as well as for the secured transmission of encrypted data.

1.1.2 **Principal implementations**

Several two-level systems have been proposed as physical implementations of qubits. Each of them has to face DiVincenzo criteria and, according to this metrics, their respective peculiar nature can reveal its strength as well as its weakness in view of the implementation of a QIP architecture.
Some review articles have been devoted to present and analyze the different architectures in detail [68, 11, 131, 3, 29, 61, 93]. Here, the main implementations of QIP will be briefly presented in order to highlight their principal properties and compare the results obtained by the different approaches.

**Photons** Quantum information can be encoded in the polarization of single photons [100, 101, 68]. Photon qubits are an almost decoherence-free system and in principle it can be operated at room temperature.

In 2001 the KLM scheme (after the inventors Knill-Laflamme-Milburn) introduced a novel concept for photon-based QIP, where the operation of arbitrary photonic circuits is obtained just with single photon sources and detectors and linear optical components [62]. This conceptual breakthrough was a boost for this technology, leading to the execution of small algorithms in this framework [106]. Photon losses in waveguides are actually the main source of errors.

Great effort is focused also on the optimization of single photon detectors and sources. Notably, besides the direct employment of photons as qubits, they are also promising candidates as long-distance buses for quantum communication in hybrid systems and quantum cryptography, exploiting their unrivalled speed and coherence properties [101].

**Neutral atoms and ion traps** Neutral atoms are among the most stable systems in nature and can be successfully employed as qubits due to the long decoherence times that make them suitable also for metrological applications [68, 11].

Atoms in ultra-high vacuum can be trapped in an optical lattice defined by laser standing waves. Such system is very well-isolated from the environment and can be cooled down to nK temperatures retaining quantum coherence for seconds. Ion traps are a system very similar to neutral atoms except that atoms are charged, so that a stronger atom-atom interaction is present due to Coulomb force and that atoms can be controlled also with electric fields.

Manipulation of neutral atoms is performed with laser radiation, while in the case of ion traps further control is achieved by applying electrostatic fields. In both cases decoherence times are much longer than manipulation timescales and the operation performances can be tuned with the laser beam.
parameters. Actually scalability to hundreds qubits presents technical issues due to the cooling of a larger system, that impacts coherence, and the control of mutual interactions between atoms. Nonetheless, entanglement between 14 ions have been successfully observed and an architecture for multi-qubit arrays have been proposed where qubits can be shuttled across the atomic array [88].

**Superconductors** Superconductor qubits, differently from the architectures, are macroscopic circuits, like classical parallel RLC circuits, that under particular conditions manifest a quantum behaviour.

When operated at cryogenic temperatures (i.e. under the superconductor critical temperature) the circuit resistance vanishes (R=0) turning the system into an ideal LC lossless circuit. The magnetic field flux in the inductance $\Phi$ and the charge on the capacitor $Q$ have the commutator $[\Phi, Q] = i\hbar$ and define an harmonic oscillator potential substituting the position and the momentum respectively [68, 11]. A Josephson junction (an insulating barrier between two superconductor wires), connected in parallel, introduces an anharmonic potential. This perturbs the periodic sequence of energy levels of the harmonic potential and allows quantum computation in various architectures [29]. Charge and flux qubits encode quantum information in the number of Cooper pairs in a Cooper pair box and in the direction of current in a superconducting loop respectively. Phase qubits are based in the phase difference introduced by a current-biased Josephson junction on the superconductor wavefunction.

Coupling between different qubits can be either inductive or capacitive, but can be also mediated by a photon in a resonant cavity [147]. Actually coherence times in the order of $\sim 10\mu$s can be obtained and further studies are focused on the comprehension of the decoherence mechanisms. Recently, a fidelity of more than 99% (that is over the fault-tolerant threshold of surface codes) was obtained for one and two-qubit operations in a register of five superconducting qubits by the Martinis group, after a partnership between Google and the University of California at Santa Barbara [5].

**Solid state qubits** This class of qubits includes semiconductor QDs and point defects in solids [68, 11, 93, 3]. Semiconductor QDs are confined nanostructures where single electrons can be bound and they can be electrostatically defined or self-assembled. In the first case electrostatic barriers are built up
in a two-dimensional electron gas by controlling the voltages of gate electrodes defined by lithographic processes that allow initialization and read-out of the charge state by spin-to-charge conversion [49]. In the latter, QDs are generally grown as semiconductor heterostructures after a bottom-up process, like Si-Ge or III-V semiconductors heteroepitaxy depositions. Several architectures employ one spin confined in a QD that can be manipulated by microwave or laser pulses. Other implementations make use of more than one electron (e.g. singlet triplet qubits) and can take advantage of the electron exchange interaction to obtain a faster operation time [81, 61]. The main sources of decoherence are electrical and nuclear noise generated by the random fluctuations in the surrounding nuclear spin bath [103, 9]. To this extent, group IV semiconductors, like Si and C, bring the advantage of a lower concentration of active nuclear spins in their natural lattice.

Donor atoms in silicon, such as P and As, can store a single spin as well and their sharp confining potential is a well-isolated system that is suitable for QIP [58, 138]. Quantum information can be stored either in the nuclear or the electron spin. The operating principle is analogous to a QD-based single spin qubit, but longer coherence times were measured exceeding 200 µs [105].

The negatively charged NV centre (NV−) is another solid state paramagnetic centre that can be employed for QIP. It is composed of a substitutional nitrogen atom close to a carbon vacancy in diamond. Notably, this architecture allows for a fast and reliable operation even at room temperature [68].

**Nuclear Magnetic Resonance** Another important implementation of QIP is based on liquid phase Nuclear Magnetic Resonance (NMR). This implementation takes advantage of the fast rotational motion of large molecules in a solution to obtain coherence times on the scale of seconds [140, 68]. Single nuclear spins can be reliably addressed by their specific resonant frequency, while two-qubit gates are mediated by the electron molecular orbitals binding different atoms.

Promising results have been obtained with this architecture [141], that however exhibits relevant scalability issues. Nonetheless, this architecture is actually noteworthy in the development of QIP for scientific as well as technological reasons, since it set the background for most theoretical concepts and experimental techniques involved in the manipulation of spin in solid state.
Concluding, it must be noted how each of these architecture has its own strengths and drawbacks. Moreover, besides the manipulation of single qubits, a realistic QIP implementation should include also a reliable quantum communication protocol to bring in contact a large number of "data" qubits (Section 1.1.1). Indeed, quantum communication is, at least, as important as quantum computation and the last two requirements must be equally considered. As a consequence, hybrid architectures have been proposed in literature, combining the functionalities of different implementations used for specific tasks (e.g. photons or superconductors as flying qubits and electron or nuclear spins for quantum memory) to take the best of each system [68, 11, 147].

1.2 Donor atoms in silicon

Nowadays the progressive miniaturization in the field of microelectronics is facing its ultimate limits as devices are fabricated with a smaller and smaller number of atoms. For this reason a deeper understanding of the physical properties of quantum confined systems is fundamental to investigate solid state physics at the nano-scale and realize nanoscaled devices for classical and quantum information processing. This section is dedicated to the characterization of quantum transport through a quantum dot (QD) coupled to a single donor as a preliminary study of single charge devices and qubits based on donors.\(^1\)

1.2.1 Donor-based qubits in silicon

Donors in silicon are a suitable environment to perform quantum computing and have been deeply investigated to this extent [40, 36, 6, 108, 133]. Their atomic size, combined with a deep confining potential, turns out to be a well-isolated environment, reaching coherence times in the order of seconds in purified \(^{28}\)Si at cryogenic temperatures [137, 138]. Gated donors in proximity to the Si-oxide interface have been extensively studied as well in a variety of device architectures [52, 73, 86, 117, 146]. Some implementations have been considered to realize

\(^1\)This Section has been readapted and published in Ref. [117]. The article is focused on the spin-dependent recombination in P-doped silicon nanowires and on quantum transport in a quantum dot interacting with a single donor. The author prepared and characterized the nanowire devices (not reported in this thesis) and performed the charge sensing measurements on a single donor atom reported in this Section.
charge and spin qubits in doped silicon [58, 144, 50, 90, 64], and the coherent manipulation of quantum states have been demonstrated in electron and nuclear spin qubits [92, 105].

The donor spin qubit architecture employs the two spin states of an electron bound to a donor as the logical basis of the qubit [90]. The donor charge state is monitored by a near charge sensor (generally a Single Electron Transistor), so that initialization and read-out can be performed through spin-charge conversion techniques: spin-dependent selection rules are applied for electron tunneling between the donor and the SET by properly controlling their electrochemical potential [52]. The spin degeneracy is lifted by the Zeeman splitting introduced by an external magnetic field and spin manipulation is carried out by means of microwave pulses inducing coherent Rabi oscillations between the two spin states [90].

1.2.2 Charge control and single electron dynamics on a donor atom

Fabrication of the single donor device

![Fabrication of the single donor device](image)

We performed the study of the charge dynamics on single donors on a bulk silicon n-MOSFET equipped with two lateral gates defined by EBL and Al deposition (fig.1.2a). The substrate is a lightly doped Si wafer, with a phosphorous concentration \( n_P = (1 - 5) \cdot 10^{15} \) cm\(^{-3} \). The main gate partially overlaps the degenerate doping regions at the source-drain contacts and controls the carriers accumulation in the channel, switching the electrical transport at cryogenic temperatures. In the
central part of the device, the channel width is strongly reduced to form a constriction 40 nm wide and 300 nm long, in order to operate this nano-MOSFET at low temperatures as a Single Electron Transistor (SET) and exploit its powerful features as a high sensitivity electrometer to detect the charge state of nearby donor atoms. A stability diagram of the SET is shown in fig.1.2b, where two conductive paths are evident at different gate voltages. In fact, charging energies are similar (roughly 5-6 meV), however the different lever-arm factor are associated to the red (0.11 eV/V) and black (0.21 eV/V) are clearly associated to distinct centers of confinement inside the SET channel. According to the known doping density, we expect a number of 1-10 P atoms interacting with the channel of the SET; moving the potential of the two auxiliary gates we gain enhanced control on the SET and on the charge state of the few neighboring donors [86].

**Donor stability diagram and single charge dynamics**

We studied the electrical transport of an electrostatically defined Si-QD interacting with a near donor atom at cryogenic temperatures (from 4 K down to 300 mK) controlling the gate voltages of the device in fig.1.2 and monitoring its conductance with a cryogenic amplifier thermally anchored to the 1K-pot of the utilized 3He cryostat. In fig.1.3 we report a stability diagram of the device, where the dashed lines denote the first Coulomb peaks of the SET, occurring when its chemical potential enters the bias window. Solid lines, instead, indicate the charge transitions of adjacent donors, clearly recognizable as abrupt steps of the current peaks when crossing such lines. These steps are caused by the capacitive coupling between the SET and near P atoms and they are the fingerprint of the charging/discharging process of a single electron on these atoms [86]. Controlling the potential of the auxiliary gates, we could explore different configurations of the system: in particular, moving along the arrow labelled as ε in fig.1.4a, we could detune the system and induce a charge transition on a donor while leaving the chemical potential of the SET constant. The charge on the SET induced by capacitive coupling with the donor is \( \Delta q \sim 0.2e \), which is an indication of an intermediate distance between the donor and the SET. This condition ensures an optimal compromise between the maximum sensitivity of the charge sensor and a reasonable tunneling rate to the donor. In this configuration, we expect a tunneling time of the order of milliseconds [91], which can be adequately measured with our experimental setup (amplifier bandwidth of 34 kHz, rise time of \( \sim 10\mu s \)).

Moving along the detuning axis \( \varepsilon \) from the initial state A to the final state B
Figure 1.3: Stability diagram of the SET at a base temperature $T = 800$ mK. $V_{ds} = 0.305$ mV, $V_{gate} = 2.12$ V. SET current is reported in pA; solid and dashed lines determine the boundaries of regions with constant electron occupation inside both the SET and the donor. In particular, one electron is added or removed from the SET (the donor) when crossing one of the dashed (solid) lines.

(and viceversa), it is possible to reveal the emission (capture) of an electron by a donor as an increase (decrease) in the SET current. Then we performed time-resolved measurements of the SET conductance along this "detuning" direction on a wide range of applied magnetic fields, in order to measure the capture and emission times of an electron by a P atom. After averaging over many scans, as shown in fig.1.4b, we could estimate the two tunneling times from a simple exponential fit of the SET current during the detuning operation. We found that the tunneling time is about 1 ms for the capture and the emission process as well; furthermore, both times are independent of the applied magnetic field up to 8 T (fig.1.5). These results are consistent with a tunneling process between a SET state and a donor state with the same spin projection along the field axis [86].

1.2.3 Conclusions

We observed the quantum transport behavior of a QD interacting with few donors. In details, we observed the tunneling process between the QD and a donor in the time domain, measuring tunneling times of the order of 1 ms independent of the applied magnetic field up to 8T. These results represent a solid platform for the initialization and read-out of the charge state of a single donor atom, which open up to various applications in QIP and single atom electronics [133, 118, 89, 73, 109].
Figure 1.4: (a) Stability diagram of the SET device (SET current is reported in pA) in proximity of a charge transition of a donor. The arrow shows the detuning direction \( \varepsilon \) followed during time-resolved experiments to unload an electron bound to a donor. (b) Time-resolved measurement (averaged over 100 samples) of the SET current along the detuning axis \( \varepsilon \) defined in (a). Electron tunneling times between the SET and the donor are extracted by fitting data to an exponential function.

1.3 Silicon double quantum dots

This section is focused on silicon Double Quantum Dots (DQD) and their device applications for QIP.

A brief introduction in Section 1.3.1 will examine the main QD-based architectures for QIP in order to compare different approaches and underline the principal scientific results and their respective potentials and drawbacks. The following sections will be dedicated to the design (Section 1.3.2), the fabrication (Section 1.3.3) and characterization (Section 1.3.4) of Si-DQD devices conceived for QIP on a SOI-MOS platform.

1.3.1 Double quantum dot qubits

Semiconductor quantum dots (QD) are a viable alternative for solid state QIP, as more relaxed bounds are imposed to the critical sizes of devices compared to donor qubits [81, 61]. Known as artificial atoms, they can be designed and made up with a great control over many important degrees of freedom, such as dimensions and inter-dot distances, which govern the physical properties and the dynamic performances of the resulting qubits. This is true for electrostatic QDs in particular, which can be reliably fabricated and finely tuned over a wide range of operating
regimes, presenting a larger potential for LSI than self-assembled QDs. Furthermore, a QD can easily store more than one electron. As a result, such versatility opens up to a variegated landscape of different architectures, employing more than one electron spin, with different operating principles and constraints.

**Charge qubits** The logical basis is defined by the dot occupied by an electron in a DQD system. This qubit can be detuned between the two logic states by moving the potential of metal electrodes, so as to control the chemical potentials of the two dots and the inter-dot tunnel coupling. Very fast operation (in the GHz range) is possible, while coherence times are limited to few ns by electrical noise [45, 14].

**Spin qubits** Spin qubits employ the spin states of a single trapped electron and they resemble donor spin qubits in many aspects, with the difference that the electron spin is bound to an artificial atom (a QD). Manipulation is carried out by microwave pulses at the spin resonance frequency, while initialization and read-out are performed by spin-to-charge conversion [65].

**Singlet-triplet qubits** In singlet-triplet qubits the logical basis is defined in the $S_z = 0$ subspace of two electrons in a DQD, namely the S and the T$_0$ states. Time evolution of the qubit is driven by the exchange interaction and by the presence of a magnetic field gradient between the QDs [103, 85, 128]. The qubit is initialized with two electron in a QD in the $|S\rangle$ state, then $\hat{x}$ and $\hat{z}$ rotations are performed by a pulsed gate control of the exchange interaction between the two spins. On the other hand, the magnetic field can’t be completely controlled due to random fluctuations of the nuclear spin bath, which

![Figure 1.5: Donor capture and emission times at different magnetic fields.](image)

Figure 1.5: Donor capture and emission times at different magnetic fields.
is mainly responsible for qubit dephasing [56, 9]. The final state is read-out thanks to Pauli spin-blockade.

**Exchange-only qubits** Exchange-only qubits have been designed and demonstrated by using three electrons either in a double or a triple QD [70, 125]. Manipulation is restricted to a selected spin subspace and is driven by exchange interaction, that allows a rapid and all-electrical control of the qubit state.

For many years the scene of solid state QIP has been dominated by electrostatic QDs defined in III-V semiconductors, namely GaAs [35, 103, 65, 70]. This choice takes advantage of the technological improvements in the epitaxial growth processes, yielding high quality substrates with very low density of defects and high electron mobility that are a suitable environment for QIP [67, 139, 49]. However, weaker spin-orbit coupling and hyperfine interactions are expected in silicon [49, 93, 163, 138, 149]. In fact, isotopically purified 28Si can be grown with a concentration of \( \sim 100 \) ppm 29Si, favoring longer spin coherence times. The coherent manipulation of a spin qubit has been recently demonstrated in a Si QD with a coherence time \( T_2 = 28 \)ms and a fidelity of 99.6% [142]. The controlled manipulation of qubit states has been obtained also for a singlet-triplet qubit [85] and an exchange-only hybrid qubit [126, 59] in SiGe heterostructures. In both cases, controlled qubit rotations at GHz frequencies were obtained by pulsed gate control as in similar GaAs devices. Conversely, a higher number of coherent Rabi oscillations could be observed due to the longer coherence time provided by the silicon environment (360 ns for the Si-based singlet-triplet qubit in [85], that is about to two orders of magnitude longer than in the GaAs counterpart studied in [103]).

### 1.3.2 Design

The devices studied in this thesis have been designed for the implementation as DQD qubits, exploiting the remarkable versatility of this system, that in principle can host any of the architecture in the previous Section 1.3.1, from charge to spin and exchange qubits.

No matter the chosen material and technology, this system requires two QDs with tunable inter-dot tunnel coupling and individual control. In addition, a Single Electron Transistor (SET) or a Quantum Point Contact (QPC) needs to be placed in close proximity to the DQD to achieve a sufficient capacitive coupling to work as a single charge sensor of the DQD itself and perform the read-out of the qubit state.
Silicon has been selected as the host material in view of the supreme spin coherence properties on a side and for the evident technological potential for LSI on the other side. Vertical confinement is obtained by employing a thin Silicon-On-Insulator substrate. Lateral confinement in the DQD and the SET charge sensor is provided by defining ultrascaled Si wires in a SOI process, following an alternative approach with respect to the more popular strategy based on the use of depletion gates to define a DQD from a 2-Dimensional Electron Gas (2DEG) [35, 65, 85, 87, 59]. Further confinement in the third dimension is obtained by field effect on the framework of a Metal-Oxide-Semiconductor capacitor. Metal gates are consequently defined on top of the Si wire to be used either as depletion gates (inter-dot barriers) or as plunger gates (building up the QD potential well and modulating its chemical potential).

This technological framework has been already employed for the study of CMOS-compatible QDs [104, 107] and presents relevant advantages in terms of scalability, as will be explained further in Section 1.4. Preliminary Spin-dependent Density Functional Theory simulations (SDFT) were performed on this system, suggesting an optimal width = 60 nm for the Si island and the DQD [27]. On the other side, the calibration of Electron Beam Lithography indicated a 100 nm minimum pitch for the metal gates to guarantee a reliable lift-off process and an acceptable reproducibility. The device layouts and the corresponding lithographic masks have been consequently designed in this framework.

1.3.3 Fabrication: CMOS-compatible process integration

Figure 1.6 schematically displays the process flow that was followed to fabricate Si-DQD devices for QIP in a CMOS-compatible way.

At first, lithographic reference markers must be defined on the SOI substrate to guarantee a correct alignment of the multiple lithographic patterns. A micron-sized set of marker points has been etched in the Si device layer for low resolution optical lithography processes, like source and drain contacts formation and bonding-pads deposition. A second set was created by EBL and Ti/Pt deposition to further enhance the alignment precision for high-resolution EBL exposures, namely the creation of the Si islands and of the metal gates in the active area of the device.

Degenerate doping under source and drain contacts is then supplied through thermal diffusion during high temperature RTP. The Si islands incorporating the DQD and the SET charge sensor are created from the SOI device layer after EBL and Si wet etching. The gate oxide is then grown by RTO and ALD and opened
thereafter at the source and drain contacts.

Metal electrodes are finally deposited simultaneously through a lift-off process performed by EBL exposure and thermal evaporation of Al. In some devices the last two steps have been repeated in order to make a second level of metal electrodes and increase the gate surface density in the active area of the device (see Section 1.3.3).

All these steps have been previously calibrated and optimized for this process flow, as will be described in the following. Additional details on the process parameters and the complete process flow are reported in Appendix A.

As a final remark, it is worth noting that, although some techniques and materials invoked are normally out of the standard procedures of VLSI technologies, the general outline of this process flow is fully CMOS-compatible and flexible enough to be adapted from our research lab to state-of-the-art nanoelectronics factories [97, 135]. On the one hand, RTP and ALD are processes that are broadly used in semiconductor industry. On the other hand, dopant diffusion is carried out with similar results through ion implantation (that was not available for this work) and even better results are obtained through dry-etching rather than wet-etching and lift-off processes. Furthermore, deep-UV lithography could easily replace the slow EBL processes while still meeting the resolution requirements and making Pt markers (of difficult integration, if not completely out of question), totally useless.

This opens up to an effective scalability of this technology, in the perspective
to build up multi-qubit circuits based on Si-MOS DQDs.

**Ohmic contacts**

Metal-semiconductor contacts usually constitute a Schottky junction, resulting in rectifying behaviour that degrades the performances in most of nanoelectronic devices [132]. In the present case, ohmic source and drain contacts are required for a low-noise read-out of the DQD charge state.

Strongly doped semiconductors (at a density above the Anderson-Mott transition) provide a high density of free electrons, decreasing the width of the depletion region at the Schottky junction and yielding a pseudo-ohmic tunnel junction with low contact resistance [132]. Degenerate doping was obtained by thermal diffusion of dopants, namely phosphorus, from a solid source, according to the following procedure. A Spin-On-Glass (SOG) with high P concentration is uniformly spun and baked onto the substrate. Then, substrate doping is driven by a high temperature RTP through a mechanism of thermal diffusion from the highly concentrated SOG film to the undoped substrate [97].

![Figure 1.7](image.png)

**Figure 1.7** Electrical test after thermal diffusion processes. a) Electrical transport in silicon doped by thermal diffusion. RTP have been performed for 120 s in oxygen atmosphere at different temperatures. b) Check of the reliability of the SOG-based hard mask for thermal diffusion. n++ Si refers to transport through unmasked regions, that have been consequently doped. npn junction indicates transport between two doped region across material that was not doped indeed, due to the mask protection.

Figure 1.7a shows resistivity measurements after thermal diffusion at three different temperatures. An optimal processing temperature was selected at 800°C, after comparison with processes at 700°C, where insufficient thermal energy is supplied to reach an acceptable doping, and 950°C, where on the contrary an excessive doping level degrades the substrate due to doping clustering and surface
roughening.

The active area of devices, including the DQD and charge sensor region were left undoped by protecting the sample with a SiO₂ hard mask before the deposition of the doped SOG. The hard mask is made of undoped SOG, that was spun on the substrate and selectively etched at the source and drain region after optical lithography patterning. The effectiveness of the hard mask selectivity over dopant diffusion is demonstrated by the electrical measures in Figure 1.7b, that compares electrical transport between doped and undoped areas according to the hard mask pattern.

Reference markers for EBL

Precise alignment markers need to be defined in order to precisely align the high resolution EBL patterning exposures (Si island definition and multiple metal electrodes depositions). The chosen design of the devices requires a stitching accuracy better than 50 nm, in particular to take advantage of the design with two metalization levels (which are defined with a 50 nm half pitch). Full integration in the process flow also requires the markers to be stable to all the etching and high temperature annealing processes to be performed on the sample afterwards. As a consequence, proper reference markers must be chemically and thermodynamically stable, sharply defined and easily detectable from a SEM image, so different strategies were considered and compared in Figure 1.8.

An easy way to introduce alignment markers is by defining them within the first of the high resolution patterning processes, namely the Si island definition. The resulting structure is another Si island placed in a sacrificial region that can be used as a reference marker for the next exposures. This approach is not actually an ideal solution in very thin SOI substrates, as we can see in Figure 1.8a. In fact, the thin Si layer is almost transparent to high energy electrons, providing poor contrast with the surrounding regions where Si has been etched away.

Another approach consists in defining dedicated structures through a dedicated process. On the one hand, an additional process step is needed to this purpose. On the other hand, alternative materials are available to define more evident reference markers. Chemical contrast in SEM images is due to the enhanced electron backscattering arising from materials with high atomic mass, like Pt, that results in bright features where such materials are present. Moreover, Pt is inert to all the chemical etching solutions normally employed in Si-CMOS processing, (namely BOE and TMAH Si wet etching) and the high melting point of Pt (1414°C) in
Figure 1.8: SEM images of different markers for EBL. a) Etched silicon markers. b) Pt test markers (after RTP at T = 900K). c) Ti/Pt test markers (after RTP at T = 900K). d) Markers used for high precision alignment in the definitive process flow. The inset shows a typical intensity scan of scattered electron across a Ti/Pt marker and data post-processing by the automatic alignment procedure.

Actually, pronounced dewetting occurred after high temperature annealing (T = 900°C) due to the poor adhesion of Pt (Figure 1.8b). Best results have been obtained with Ti/Pt bilayer markers, thanks to the better adhesion guaranteed by a thin Ti adhesion layer (Figure 1.8c).

The dimensions of the markers were finally reduced, in Figure 1.8c, in the final process flow to reach a stitching accuracy of about 20 nm. Remarkably, in the first method the alignment procedure was necessarily manual (due to the poor contrast of Si markers) affecting the stitching accuracy and dramatically slowing down the overall patterning process. Conversely, Ti/Pt markers exhibit a sufficient contrast to allow post-processing of SEM linescans through the ordinary threshold detection algorithms and write-field alignment procedures to obtain fast and precise
automatic alignment for all the tens of devices processed on each sample.

**Definition of silicon islands**

Si islands hosting the DQD with the corresponding charge sensors have been defined by Si wet etching in hot Tetra-Methyl-Ammonium-Hydroxide (TMAH) after negative-tone EBL patterning on a SiO$_2$ hard mask.

Figure 1.9: a) Nanowire with the minimum feature size obtained by EBL. b) Detail of the lithographic mask for the SOI island etching. Color code indicates the the EBL dose multiplier. c) SEM image of a sample after Si etching.

Figure 1.9a shows the maximum resolution obtained by EBL with the negative resist (ARN7500) after optimal focusing and calibration of the electron beam dose. The small minimum feature size (60 nm) and the peculiar T-shape design of the DQD island required an additional careful analysis of the lithographic masks to correct for the proximity effect. The results of such optimization process are shown in the corrected mask in Figure 1.9b and in a physical device after Si etching in Figure 1.9c.

A SiO$_2$ hard mask, 3.5nm thick, was previously grown from the Si substrate by RTO at 900°C, while the pattern transfer from the resist layer onto the hard mask was carried out by chemical wet etching in dilute HF (1:50). TMAH-based Si etching is strongly anisotropic, since etching rate is much faster on the $\langle 100 \rangle$ surface than on the $\langle 111 \rangle$, while etching of SiO$_2$ is negligible. As a result, etching over linear patterns parallel to the $\langle 110 \rangle$ produces linear structures with trapezoidal section with height equal to the Si device layer (~12 nm).

**Gate oxide**

A good gate oxide is at the heart of a working field-effect device. Good oxide points to negligible leakage currents, that is closely related to the physico-chemical integrity of the material, and high quality at the interface with the semiconductor
low density of interface traps \( D_{it} \). Such requisites become even more urgent when applied to single charge devices, since single electron traps and few electrons tunneling through the oxide substantially affect their accurate operation.

Table 1.1: Comparison of the \( D_{it} \) extracted by differential admittance characterization performed on planar capacitors with oxide stacks and processing procedures. Thermal SiO\(_2\) (3.5 nm) was grown by RTO (30s at 900°C in 100% O\(_2\) atmosphere). Thicker SiO\(_2\) (30 nm) layers were deposited by e-beam evaporation of some test samples with a controlled deposition rate (0.1 nm/s). Furnace annealing was performed in a tubular furnace for 1h at 1000°C in N\(_2\) flux. Forming Gas Annealing (FGA) is performed by RTP for 20 minutes at 450°C in a 5% mixture of H\(_2\)/N\(_2\). Al\(_2\)O\(_3\) (40 nm) was grown after 500 ALD cycles at 300°C.

<table>
<thead>
<tr>
<th>Sample</th>
<th>( D_{it} ) [eV(^{-1}) cm(^{-2})]</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO(_2) thermal + evaporation</td>
<td>3.6 x 10(^{11})</td>
</tr>
<tr>
<td>SiO(_2) thermal + evaporation + furnace annealing</td>
<td>2.3 x 10(^{11})</td>
</tr>
<tr>
<td>SiO(_2) thermal + evaporation + EBL (10 kV)</td>
<td>3.4 x 10(^{11})</td>
</tr>
<tr>
<td>SiO(_2) thermal + evaporation + EBL (20 kV)</td>
<td>3.5 x 10(^{11})</td>
</tr>
<tr>
<td>SiO(_2) thermal + evaporation + furnace annealing + FGA</td>
<td>1.2 x 10(^{11})</td>
</tr>
<tr>
<td>SiO(_2) thermal + Al(_2)O(_3) ALD</td>
<td>2.0 x 10(^{11})</td>
</tr>
<tr>
<td>SiO(_2) thermal + Al(_2)O(_3) ALD + FGA</td>
<td>6.8 x 10(^{10})</td>
</tr>
</tbody>
</table>

For this reason different options have been considered, taking as a quality benchmark the \( D_{it} \) extracted from differential admittance measurement reported in Table 1.1. 3.5 nm thick thermal Si dioxide has proven to be an optimal choice, providing \( D_{it} \sim 10^{11}\) ev\(^{-1}\)cm\(^{-2}\). Such trap density is further lowered after high temperature furnace annealing and FGA in particular. Notably, the effect of high energy electron beam irradiation doesn’t produce a dramatic increase of the traps density during the EBL processes; anyway a lower value can be restored after FGA.

Unfortunately a thin SiO\(_2\) layer is not sufficient to completely avoid tunneling through the oxide and guarantee a stable operation of the device. Moreover, a deeper oxidation is not possible because of the limited substrate thickness, so other strategies are needed to obtain a robust gate oxide.

The final choice lead to a double layer stack made of SiO\(_2\), providing the high quality interface previously characterized, and Al\(_2\)O\(_3\). Aluminium dioxide is a high-\( k \) oxide (\( k \sim 9 \)) and is grown through a well-consolidated conformal ALD process at 300°C with Tri-Methyl-Ammonium-Hydroxide (TMA) and H\(_2\)O precursors.
Metal electrodes

Aluminium electrodes have been employed to define metallic control gates. The electrodes have been deposited by a lift-off process that consists in a high resolution EBL exposure followed by Al thermal evaporation and removal of the excess material in hot solvent. The resist is a bilayer Poly-Methyl-Meta-Acrylate (996k-350k molecular weight) exposed at 20 kV accelerating voltage, yielding a good compromise between high resolution and precise lift-off. This is probably the most critical process, due to the high resolution and the complex layout reported in Figure 1.10a-b, that in addition requires a stitching accuracy of about 20 nm with respect to the underlying Si islands.

![Figure 1.10: Definition of metal gate electrodes. a) Example of a lithographic mask for Si island formation (orange) and the metal gate definition (grey). b) The same mask of picture a) with a color-code corresponding to the electron beam dose factor. c) SEM image of a device after the deposition of metal gates. The underlying Si island is clearly visible in the background, successfully aligned with the metal gates.](image)

Besides this, a good lift-off process is crucial to obtain well-defined metal lines reproducing accurately the EBL pattern. Lift-off was carried out in n-methyl-2-pyrrolidone (NMP) heated at 80°C carefully aided by low power ultrasonic agitation and repeated gentle blowing with a laboratory pipette. Typical results of an optimal process are shown in Figure 1.10c.

Gate: second level

The second level gate oxide has been deposited by ALD with analogous parameters as for the first gate level (TMA and H₂O precursors at temperature T = 300°C). This is necessarily a low temperature process due to the presence of Al as the first level of gate electrodes, which is not stable at temperatures higher than 500°C. For this reason ALD-grown Al₂O₃ was favored over other available options, such as evaporated SiO₂, requiring a higher thermal budget to reach ac-
ceptable electrical reliability. Other possible choices have been considered, like oxidized Al by plasma treatment or RTP [2, 76], however yielding unacceptably low breakdown voltages. As a further point, ALD growth of Al₂O₃ is a well-established and calibrated process, ensuring conformal growth with good film uniformity and electrical reliability with the advantages of a stoichiometric high-\( k \) material.

**Final devices: different layouts**

![Figure 1.11: SEM micrographs of final devices with different layouts. a) Low magnification SEM image of a typical SOI-QD device. Labels with different colors indicates different levels and materials composing the device: degenerate doping region at source and drain ohmic contacts (pink), Ti/Pt EBL reference markers (yellow) and metal gate electrodes (light blue). The green box is magnified in b) and c) highlighting the device active region for device design A and B respectively. b-c) High magnification SEM images of the active area in devices with layout A and B respectively. The charge sensor (SET) is defined in the Si constriction by either the fork-shape gate \( V_{\text{down}} \) in b) or the two independent barriers BL and BR in c). The 2 levels of metal electrodes of layout B device are labeled with different colors in c). Lower gates (light blue) can be negative-polarized for use as electrostatic barriers. Higher gates (blue) can be used accordingly as plunger gates for the DQD.

Figure 1.11 displays different devices prepared during this thesis with different layouts. Figure 1.11a shows a typical device, highlighting the different materials
and lithographic levels used during the fabrication process. In Figure 1.11b the Si island can be recognized under the main gate and the three plunger gates that are used to form a DQD in the Si vertical finger underneath. This device layout will be further referred as layout A.

Better electrostatic control and gate density is obtained by stacking two levels of metal electrodes in layout B, as in the device represented in Figure 1.11c. A more complex process flow is needed to build up such device, involving an additional EBL exposure. Such exposure is a high resolution pattern that also requires to be adequately aligned to the Si-island pattern and the first metal level with an accuracy better than 50 nm. The lithographic levels has been aligned with errors smaller than 20 nm in Figure 1.11c. As a result, an effective gate pitch of 50 nm is obtained, that is half than in the device in 1.11b with one metal level.

Some additional devices have been manufactured with two different layouts, which are described in Appendix B.

1.3.4 Quantum transport characterization

This Section reports the electrical measures of quantum transport in selected devices that have been extensively characterized at liquid helium temperature (T = 4.2K). Gate potentials are controlled by a National Instrument NI-PXI-6733 module with a resolution of 305 µV, while the drain-source current is amplified by a trans-impedance room temperature amplifier (bandwidth=1.6kHz) and recorded by a NI-PXI-4071 multimeter. Operation of a SET as a charge sensor is assessed in order to obtain control of the charge state in adjacent QDs. Transitions involving single electrons are controlled and recognized by electrostatic gate control in the static as well as the dynamic regime.²

Tunable quantum transport in a nanoscaled MOS FET

The first requisite to perform charge sensing on a DQD is, obviously, a good charge sensor. Consequently, the electrical characterization of the SET to be employed as a charge sensor is the first step to achieve full control of the charge and spin states of a semiconductor DQD [67, 49].

Figure 1.12b reports a transconductance plot calculated from the measured drain-source current $I_{ds}$ with respect to the drain potential $V_d$ of the SET. A SEM image of an equivalent device (layout B defined in Section 1.3.3) from the same

²Article in preparation.
Figure 1.12: Electrical characterization of quantum transport in the SET of device SOI16-10A5. a) SEM image of a device equivalent to the one characterized in this chapter. b) Stability diagram $V_d/V_g$ showing a diamond pattern arising from Coulomb blockade inside the SET. c) Stability diagram of the SET varying the potential of the two lower barrier gates $V_{BL}$ and $V_{BR}$. d) Stability diagram analogous to d) recorded during another sequential cycle.

Batch is reported in panel a), where two superimposed layers of metal (Al) gates are visible on top of the etched Si island. The final Si thickness is 11 nm. The first gate oxide is a bilayer stack composed of thermal SiO$_2$ (3.5 nm) and Al$_2$O$_3$ deposited by ALD as described in Section 1.3.3. The Single Electron Transistor (SET) has been characterized at cryogenic temperature ($T = 4.2$K) in a liquid He bath. Drain-source DC current $I_{ds}$ has been measured by varying the drain bias $V_d$ and the main gate potential $V_g$ in a standard characterization of quantum transport devices [66, 67]. Current flows horizontally through the small Si constriction by applying a small DC bias between the source and drain electron reservoirs induced by the main top gate. The two lower barrier gates are kept constant at a negative voltage to build up two tunnel barriers between the source and drain contacts and form a QD therein.

Figure 1.12b shows the typical diamond-like features ascribable to Coulomb blockade in a QD. This is the fingerprint of sequential tunneling of single electrons.
through the QD, that is here effectively operated as a SET. Notably, the conductance plot does not report a periodic sequence of Coulomb diamonds, as expected from an ideal QD defined by a harmonic potential. Indeed, the diamond height shows large variability, pointing to charging energies ranging from roughly 5 meV at high electron occupation number up to more than 30 meV in the few electron regime. This behavior is generally associated with an irregular potential landscape, featuring multiple confinement centers through the SET channel, resulting in a more complicated conductance plot such as that reported in Figure 1.12b.

Remarkably, the device architecture allows the independent operation of the two barrier gates, namely BL and BR, to finely tune the SET potential and the transparency of the respective tunnel barriers. As a result, a deeper insight of the SET behavior is gained from the $V_{BL}/V_{BR}$ stability diagrams in Figure 1.12c-d. The two plots have been recorded during different thermal cycles: as a consequence, different gate voltage ranges are likely due to the oxide electrical stress and charge reconfiguration processes after intensive electrical testing and exposure to air. Nonetheless, both measures show periodic series of parallel current peaks along the vertical and horizontal directions in the stability diagram. Such current peaks are clearly visible at low gate voltages and they are attributed to resonances inside the two barrier gates. At higher gate voltages (i.e. moving towards the upper right corner) diagonal modulations appears (indicated by black arrows in Figure 1.12c-d) related to a QD formed in the middle of the channel.

Figure 1.13: Tunability of a nano-FET (device SOI16-10A5) from a SET to the DQD regime. a) Few-electron regime and low inter-dot coupling. b) Strong inter-dot coupling regime.

Further measurements, reported in Figure 1.13, help to figure out what is happening inside the SET channel. At low gate voltage (Figure 1.13a) a DQD is
formed by effect of the positive bias applied to the barrier electrodes, that are here used as accumulation gates. As a result, charge transport is possible only at the DQD triple points (i.e. when both the QD chemical potentials are within the bias window) as expected from a series DQD in the few-electron regime [139]. Such triple points occur at the crossing of the horizontal and vertical lines corresponding to resonances with the levels in the right and left dot respectively. As a matter of fact, the left dot is strongly coupled to the left gate while almost decoupled from the right barrier gate. The same arguments are transposed to the right dot accordingly.

At $V_g = 6.5V$ (Figure 1.13b) the higher gate voltage effectively lowers the interdot potential barrier in the undoped Si region between the barrier gates. This perturbation enhances the inter-dot tunnel coupling and finally leads to the merging of the two QDs in a unique confining potential that actually hosts a single QD in the middle of the FET channel. This picture is supported by the barely diagonal slope of the current lines in Figure 1.13b, emerging from a single QD equally coupled to the $V_{BL}$ and $V_{BR}$ electrodes [139, 51].

Concluding, this preliminary experiment demonstrates the wide tunability of the nano-FET presented in Figure 1.13a. Such device has been fabricated in a CMOS-compatible process flow and characterized on a wide range of operating regimes, demonstrating great functional tunability from a DQD configuration to a SET regime. These results are a promising outcome in view of employing such device as a versatile charge sensor for an adjacent DQD induced in the vertical Si island. Moreover, the demonstration of tunable inter-dot coupling is also an important milestone for the exploitation of such SOI-MOS architecture for QIP applications [103, 70, 125, 63, 38].

Charge sensing of a SOI Quantum Dot

Figures 1.14a-b show the preliminary characterization of the SET in a device with two metal gate levels (layout B). Coulomb blockade features in Figure 1.14a arise from a single QD defined in the middle nano-FET channel. The measured value of the charging energy is 20 meV, ensuring a fairly robust operation of the device as a SET against thermal fluctuations ($k_BT = 350 \mu eV$ at 4.2 K). The quality of the SET is also confirmed by monitoring the SET current when varying the potential of the SET tunnel barriers. $V_{BL}$ and $V_{BR}$ were shorted for the measurement in Figure 1.14b and moved at the common potential $V_{down}$. As a result, strikingly clean and stable Coulomb blockade features are reproduced over a wide range of $V_{down}$, reflecting the behavior of a well-defined QD potential between source and
Figure 1.14: Charge sensing on device SOI16-13H1. a) Stability diagram of the SET as a preliminary characterization of the charge sensor. b) Stability diagram of the SET by varying the potential of the main gate $V_g$ and of the electrode $V_{\text{down}}$ controlling the electrostatic barriers to source and drain electron reservoirs as well as the SET chemical potential. c) SET current measurement with tunable $V_{B2}$ potential. d) Calculated derivative of the current measurement in panel c) with respect to $V_{B2}$. The white arrows in c) and d) indicate the current steps corresponding to charge transitions in a neighboring QD by effect of the potential $V_{B2}$.

Subsequently, such SET has been operated as a charge sensor for a QD in its close proximity that is formed by the potential $V_{B2}$. White arrows in Figure 1.14c indicate a series of linebreaks in the SET current peaks, that are better recognized after mathematical differentiation of the current direct measurement (Figure 1.14d) with respect to $V_{B2}$. $V_{B2}$ is actually working as a plunger gate for a QD formed in the Si vertical island adjacent to the SET. In fact, the observed linebreaks can be interpreted in the framework of the constant interaction model \[67, 49, 163\]. The SET current peaks are slightly shifted (as evident in Figure 1.14d in particular) to a higher $V_g$ potential when an additional electron is trapped in the QD due to its capacitive coupling to the SET. As a consequence, single electron trapping by an electrostatic QD can be monitored and detected as a shift of the current peak in the drain reservoirs.
stability diagram when crossing a charge transition line of the QD.

The nominal distance between the electrodes $V_g$ and $V_B2$ is 150 nm, corresponding approximately to the distance between the SET and the probed QD. Previous works on Si QDs demonstrated charge sensing capability at such large distances by exploiting either a SET or a QPC in planar CMOS devices [99, 148, 142]. In this experiment, analogous results have been obtained with a basically different device architecture, based on a SOI-MOS structure, validating this approach as a possible starting point to develop CMOS-compatible quantum computation.

**Charge sensing and dynamics in a SOI Double Quantum Dot**

Figure 1.15 reports a stability diagram recorded by varying the potentials of the side gates $V_{qd1}$ and $V_{qd2}$ in a device with layout A. Clearly visible are the diagonal current peaks that can be ascribed to the SET. Besides this, some deviations are visible from the model of simple transport through a SET [66, 67], that are better recognized in Figure 1.15c-d. Dotted lines in Figure 1.15c highlight the typical honeycomb pattern that is expected from a DQD, that in this case is composed by the SET and by a close QD. This side-coupled QD is electrostatically defined under the $V_{qd2}$ electrode, as proven by the almost horizontal slope of the QD charge transition lines. Current actually flows through the SET, but the SET current peaks are slightly shifted at higher potentials when an additional electron is bound to the QD due to inter-dot capacitive coupling [139, 163]. Notably, $V_{qd1}$ is here operated as an inter-dot electrostatic barrier and such anomalies are especially enhanced at higher $V_{qd1}$ gate voltages in according with theory. This configuration yields a lower inter-dot potential barrier that strengthen the inter-dot coupling.

The vertical derivative (from $V_{qd2}$) of current brings to light a supplementary modulation of the current measurement, occurring at smaller voltage scales, that is associated to a second QD in the Si island. A close inspection of the charge stability diagrams confirms the different origin of the two patterns, that are respectively addressed by different slopes as a consequence of the different lever-arm factors from the two electrodes [67, 163]. Multiple centers of confinement are expected due to roughness and defects at the Si-SiO$_2$ interface, that result in an irregular potential landscape featuring multiple potential minima.

Analogous behavior has been previously observed in triple QDs by charge sensing and quantum transport experiments in several configurations, demonstrating a good control over such system [70, 87, 10]. In particular, sequential electron tunneling between the three dots can be controlled by driving the gate potentials.
Figure 1.15: Single charge sensing in device SOI16-9A3: stability diagrams and Random Telegraph Noise. a) $V_{qd1}$-$V_{qd2}$ stability diagram. b) SEM image of the device and sketch of the tunneling processes under investigation in the f) panel. c) Detail of the stability diagram in a). Dashed lines indicate the charge transition lines of the SET and a side-coupled QD. d) Derivative of the measure in c) with respect to $V_{qd2}$ (along the vertical direction). e) $V_g$-$V_{qd2}$ stability diagram. Black and blue dashed lines define the charge stability regions of the SET and dot A. f) Representative sample of the time-resolved current measurements recorded along the black arrow in e).

in proximity of the degeneracy points where the electrochemical potential of the three dots is aligned and inside the SET bias window. Moreover, if SET current is monitored in real-time experiments, time-resolved single electron tunneling events
can be detected by charge sensing of the multi-QD system and useful information can be extracted [47, 48, 158].

Figure 1.15f reports representative time-resolved measurements of the charge sensor current from a set of 100 measurements of the duration of 3s recorded at variable $V_g$ and $V_{qD2}$ along the black arrow in Figure 1.15e. Reported measures display Random Telegraph Noise (RTN) unambiguously related to single electron trapping-detrapping events occurring at QDs close to the SET channel [22]. The black arrow in Figure 1.15e lies parallel to a SET current peak, so that SET potential is maintained constant across all the measurements. On the contrary, the electrochemical potential of side-coupled QDs is progressively lowered until an electron is trapped in one of the QDs when its potential is lower than that of the SET.

Inspecting the RTN traces, a gradual transition is noticed between three configurations of the triple QD system sketched in Figure 1.15b, reflected by three distinct levels of current. This system is effectively detuned, pushing an excess electron from the SET (low SET current state) to the QD labeled as dot B (high current state) through an intermediate state, corresponding to the electron bound to dot A. Random oscillations between these states are induced by thermal excitation, generating RTN. Other electrons are reasonably present at lower energy states of the three QDs, but they do not feature the observed tunneling processes. From now on, the triple QD charge states will be consequently defined just in terms of the excess electron, according to the notation in Figure 1.15f, that is $(N_{SET}, N_A, N_B)$ or $(N_1, N_2, N_3)$ equivalently.

**Single charge dynamics and Random Telegraph Noise statistics**

The RTN traces have been adequately post-processed to exclude slow drifts related to the electronic setup and enhance the signal to noise ratio and finally digitalized with a threshold algorithm to perform statistical analysis. The first key point in this analysis is the complete absence, all over the whole set of measures, of current steps between the high and the low current levels. This implies that electron tunneling from the SET to dot B is forbidden by the geometry of the sample and the electrostatic potential landscape created by metal gates in the experimental conditions. As a result, the scenario of possible tunneling events reduces to the sequential tunneling of a single electron between SET and dot A and between dot A and dot B.

Figure 1.16 displays the occupation probabilities of an electron in the first dot.
Figure 1.16: Random Telegraph Noise statistics. a) Event counts of tunneling events between SET and dot A (RTN 12) and between dot A and dot B (RTN 23). b) Occupation probability of an electron in the SET (N1), in dot A (N2) and in dot B (N3). c) Hystogram of the QD emission times of a single RTN trace for the extraction of tunneling rates. Red line is the exponential fit of experimental data (black dots). d) Fitting of the experimental rates for electron tunneling from dot A and dot B (Γ2 and Γ3 respectively).

(SET), the second (dot A) or the third one (dot B) depending on the tunable gate voltages (see the black arrow in Figure 1.15e). Occupation probabilities are extracted as the time fractions spent by an electron in each of the three charge states during a specific measurement. Previous works in literature analyzed analogous systems, featuring single electron tunneling in a DQD, as a two-level system in thermal equilibrium [30, 51]. In that case, the electron occupation in the left and right dot are derived in the framework of a rate equation model, resulting in an energy dependence ruled by the Fermi-Dirac distribution:

\[
N_L = f_{FD}(\Delta E) = \frac{1}{\exp \frac{\Delta E}{k_B T} + 1},
\]

\[
N_R = 1 - f_{FD}(\Delta E)
\]

where \(N_L\) and \(N_R\) are the occupation of the left and right dot respectively, \(k_B T\) is the thermal energy and \(\Delta E\) is the energy difference, or detuning, between the chemical potentials of the two dots. Detuning is defined as positive when the right potential
is higher than the left one.

In a series connected triple QD, two tunnel barriers are present. This system can be thought as a duplication of the aforementioned two-level system, where the chemical potential differences (between the SET and dot A and between dot A and dot B) govern the charge state of the triple QD. Consequently, experimental data were fitted to the following equations:

\[
N_1 = f_{FD} \left( \frac{\alpha_{12}}{k_B T} (n - n_{12}) \right), \\
N_2 = 1 - f_{FD} \left( \frac{\alpha_{12}}{k_B T} (n - n_{12}) \right) - \left[ 1 - f_{FD} \left( \frac{\alpha_{23}}{k_B T} (n - n_{23}) \right) \right], \\
N_3 = 1 - f_{FD} \left( \frac{\alpha_{23}}{k_B T} (n - n_{23}) \right)
\]

where \( n \) is the trace number, \( n_{12} \) and \( n_{23} \) indicate the positions of charge degeneracy in the two systems (SET-dotA and dotA-dotB respectively). Finally, \( \alpha_{12} \) and \( \alpha_{23} \) are the respective lever-arm factors, converting measurement numbers into energy scales and yielding the detuning between two adjacent measures. Two Lorentzian peaks have been added to equations 1.3 to take into account for the small bumps positioned roughly at trace 45 and 58. Such anomalies are recognized as the hallmarks of excited states of the SET [86]. Signatures of QD excited states are expected to show up as deviations in the tunneling statistics due to an increase of the available states for tunneling and an associated increase of random telegraph noise [47, 48, 158].

The fitting model shows remarkably good agreement with experimental data (Figure 1.16b). \( n_{12} = 64.9 \) and \( n_{23} = 72.1 \) are estimated from fit as the position of the charge degeneracy configurations, which are also reflected in the maximum of tunneling events, as displayed in Figure 1.16a. Lever arms \( \alpha_{ij} \) are roughly the same, yielding a value of 98 µeV assuming no deviations of the electronic temperature from the relatively high base temperature of the liquid He bath where the sample is immersed. From this value, the energy of excited states was estimated to be 644 µeV and 2.06 meV.

Average tunneling rates were then estimated by a time-average technique as shown in literature [47, 22, 86]. Figure 1.16c shows a representative hystogram of electron emission times from dot A. Exponential fit of the experimental data yields the tunneling rates reported in Figure 1.16d. \( \Gamma_3 \) is the electron emission rate from dot B to dot A, that is governed by the equation:

\[
\Gamma_3(E) = \Gamma_{3,0} \cdot \frac{1}{\exp \left( \frac{E_3(n - n_{23})}{k_B T} \right) + 1}
\]
\( \Gamma_3 \), however is defined generally as the emission time of an electron from dot A, that can equally tunnel into the SET or dot B. As a consequence, the composite emission rate arising from these two independent events is the sum of the respective tunnel rates, according to the following equation:

\[
\Gamma_2(E) = \Gamma_{21,0} \cdot \exp \frac{\alpha_{21}(\sigma - \sigma_{21})}{k_B T} + 1 + \Gamma_{23,0} \cdot \exp \frac{\alpha_{23}(\sigma_{23} - \sigma)}{k_B T} + 1
\]  

As a matter of fact, the unexpected increase of \( \Gamma_2 \) in the left side of Figure 1.16d is not related to tunneling to dot B, that is energetically hindered in this regime. It necessarily follows from the onset of tunneling processes towards the SET, that in this conditions hosts the lower energy state.

Fitting of the experimental data gave \( \Gamma_{3,0} = 142 \text{ s}^{-1} \) and \( \Gamma_{23,0} = 1100 \text{ s}^{-1} \) for the intrinsic tunneling rates between dot A and B. The difference between such low tunneling rates requires an explanation that only simulations of this system (which are still in progress) could give. However, this outcome is reasonably due to the extremely large inter-dot distance (the nominal distance between the \( V_g \) and \( V_{qd2} \) electrodes is approximately 150 nm). In particular, the strong asymmetry in the tunneling rates between dot A and dot B could be explained in terms of a different density of states between the two QDs occurring when two dots with different shape and dimension are defined by disorder [22, 158].

1.3.5 Conclusions

Concluding, several configurations of Si-QD have been characterized, developed on the background of SOI technology as described in Sections 1.3.2 and 1.3.3. A novel device layout has been proposed for the implementation of Si DQDs, based on a T-shaped Si island incorporating both the QDs and a SET. In this framework, the latter acts as an electron reservoir for the QDs as well as a single charge electrometer for the read-out of their charge state. This implementation is basically different from others proposed in literature, which are mainly based on a planar Si-MOS platform [98, 99, 95, 69] or on SiGe heterostructures [129, 85, 59]. The choice of a SOI platform is supported by recent works on ultrascaled MOS FET manufactured from SOI substrates [107, 104]. Moreover, this may be a winning strategy on account of the extremely high density of gates required when scaling up this architecture to multi-qubit systems, as will be pointed out in Section 1.4. As a result, several devices were realized by exploiting ultra-high resolution (\( \sim 10\text{nm} \)) EBL and deeply characterized, demonstrating charge sensing functionality with such platform. Disorder generated by surface roughness and charge traps
at the interface with the gate oxide appears to be a critical issue towards the target of coherent spin manipulation with this architecture. The oxide quality has been actually optimized by introducing a high-$k$ material (Al$_2$O$_3$) and a post-process treatment, namely FGA, but could be further enhanced by employing other insulators for a lower $D_{it}$ and dry etching processes to reduce the roughness [97].

Nonetheless, disorder-induced QDs has been recently probed as potential candidates to host spin qubits in Si [143]. In this work of thesis, the operation of the QD devices has been investigated also in the realistic situation of a disordered potential landscape, which defines multiple unintentional centers of confinement. Charge control have been demonstrated and single electron tunneling events have been monitored in real time between two QD by pulsed gate techniques. Random telegraph noise was observed due to thermal excitations at the relatively high temperature of 4.2K, activating inter-dot tunneling in agreement with theory.

These preliminary results about the charge control of Si QDs can be regarded as a first milestone in the roadmap towards the coherent manipulation of a Si spin qubit with the proposed architecture. Such ambitious task will require a further optimization of the device layout and of the fabrication process flow in order to obtain better performances with a good reproducibility. In fact, these requirements set the basic conditions to perform more sophisticated spectroscopy techniques, involving magnetic field and high frequency gate control, and reach a deeper understanding of the spin and valley states involved in quantum computation [56, 91, 75, 148, 150, 126].

1.4 CMOS-compatible Large Scale Integration of hybrid spin qubits

Several implementations have been considered for circuitual quantum information processing (QIP) [31, 68, 11, 93]. Nevertheless, scalability is still an issue for many architectures. In this framework, a Complementary Metal-Oxide-Semiconductor (CMOS) architecture based on silicon would take full advantage of the well-known physical properties of the material and of the mature technological improvements driven by the semiconductor industry. Here we provide a CMOS-compatible architecture for QIP in silicon and evaluate the performances of the fundamental logic gates and the physical constraints for their scalability to multi-qubit logic circuits. As a result, two important parameters are determined for this implementation: the maximum surface density of quantum information and the
characteristic time for quantum communication between two logic qubits.\textsuperscript{3}

1.4.1 Introduction

Implementations of charge and spin qubits in silicon have been explored in both quantum dots systems [93, 27, 107, 149] and single dopant atom transistors [58, 144, 50, 64, 73, 86, 117, 108, 133]. Coherent manipulation of quantum states have been demonstrated in both atomic systems [92, 105], as well as in semiconductor quantum dots (QD), which take advantages from more relaxed bounds on the device dimensions [81, 61]. Besides single spin and charge qubits [65, 14], that make use of a single electron in a double quantum dot (DQD), in the last decade several architectures have been explored by employing two electron spins (S-T\textsubscript{0} i.e. singlet-triplet qubit) [103, 128] or three spins in double [125] and triple QDs [70]. Although satisfactory results have been achieved mainly in III-V heterostructures, spin-orbit coupling and hyperfine interactions are weaker in silicon [138, 93, 149], suggesting that silicon itself could be a promising platform for QIP. Focusing on silicon DQD qubits, the controlled manipulation of qubit states has been obtained for a S-T\textsubscript{0} qubit [85] and a hybrid qubit [124, 59] in SiGe heterostructures. In particular, the hybrid qubit is an attractive candidate for a large scale integration of QIP, as it allows a fast and all-electrical manipulation of the qubit states, with no need for either a strong magnetic field gradient, like in S-T\textsubscript{0} qubits, or microwave antennas, required in single spin qubits. Besides the study of the electronic properties of Si-MOS QDs for QIP [104, 107, 27], we also derived the effective Hamiltonian for hybrid qubits, defining the pulse sequences to perform universal quantum computation with such architecture [38, 37, 26]. Here we calculate the maximum storage of quantum information processing allowed by a CMOS-compatible implementation of silicon hybrid qubits.

In Section 1.4.2 the logic basis of the hybrid qubit is presented, as well as the operation of 1-qubit and 2-qubit logic gates for universal QIP. In Section 1.4.3 their feasibility is discussed in a state of the art CMOS process: the physical requirements for the manipulation of quantum states are compared with the constraints imposed by the existing technologies and realistic devices are designed to implement data and communication qubits. Finally, in Section 1.4.4 the large scale

\textsuperscript{3}The content of this Section has been adapted and submitted for publication. Download is available from the online ArXiv [116]. The author addressed the technological limitations involved in the scalability of the hybrid qubit architecture. He designed the modular lithographic masks and their composition in multi-qubit circuits.
integration of silicon hybrid qubits is considered in multi-qubit networks capable of fault tolerant computation and quantum error correction. The maximum surface density of logic qubits per unit area is estimated, as well as the time load for quantum communication between two logic qubits.

1.4.2 Fundamental logic gates in the hybrid qubit architecture

This Section is devoted to the description of the main concepts underlying the hybrid qubit architecture. The main building blocks for such architecture are defined in terms of data and communication qubits. Data qubits perform quantum information processing, that is one and two qubit logic operations as well as initialization and read-out of individual qubit states. Communication qubits, conversely, are devoted to the communication of quantum information between distant data qubits.

In Subsection 1.4.2 the fundamentals of hybrid architecture are introduced, as well as the schematic design and the operation of the quantum logic gates with one and two interacting qubits. In Subsection 1.4.2 we show how quantum information can be transmitted between distant hybrid qubits through the sequential repetition of SWAP logic gates between adjacent communication qubits.

Data qubits: one and two qubit logic gates

An architecture that promises the best compromise among fabrication, fast gate operations, manipulation and scalability is the hybrid qubit proposed in Refs. [125, 124, 59, 39, 37, 26, 63]. It consists of two quantum states based on three electrons electrostatically confined in two QDs, with at least one electron in each. The convenience to use such an architecture is due to the possibility of obtaining fast gate operations with purely electrical manipulations. The exchange interaction, which is the dominant mechanism of interaction between adjacent spins, suffices for all the one and two qubits operations. In addition the three electrons spin system removes the need of using oscillating magnetic or electric fields or quasi-static Zeeman field gradient to realize full qubit control, which is required for instance in singlet-triplet qubits [103]. Starting from an Hubbard-like model we have derived in Ref. [38] a general effective Hamiltonian for the hybrid qubit in terms of only exchange interactions among the three electrons.

To define the logic basis for the hybrid qubit let’s first introduce some preliminary notions. The total Hilbert space of three electron spins has a dimension of 8 and the total spin eigenstates form a quadruplet with $S = 3/2$ and $S_z =$
where the square of the total spin is $\hbar^2 S(S+1)$ and the $z$-component of the total spin is $\hbar S_z$. The qubit is encoded in the restricted two-dimensional subspace with spin quantum numbers $S = 1/2$ and $S_z = ±1/2$, like in Ref. [125]. We point out that only states with the same $S$ and $S_z$ can be coupled by spin independent terms in the Hamiltonian. The logic basis $\{|0\rangle, |1\rangle\}$ used is constituted by singlet and triplet states of a pair of electrons in combination with the angular momentum of the third spin, that is:

$$|0\rangle \equiv |S\rangle \downarrow, \quad |1\rangle \equiv \sqrt{\frac{1}{3}}|T_0\rangle \downarrow - \sqrt{\frac{2}{3}}|T_-\rangle \uparrow$$

(1.6)

where $|S\rangle$, $|T_0\rangle$ and $|T_-\rangle$ are respectively the singlet and triplet states:

$$|S\rangle = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}, \quad |T_0\rangle = \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}}, \quad |T_-\rangle = |\downarrow\downarrow\rangle$$

(1.7)

in the left dot, and $|\uparrow\rangle$ and $|\downarrow\rangle$ respectively denote a spin-up and spin-down electron in the right dot.

Every logic operation starts with the initialization process, when all the variables are regulated through appropriate external electric and magnetic fields [125]. During this procedure, all the qubits composing the system are moved in the state corresponding to the 0 logic state. Starting from this condition it is possible to proceed further with the operations that are generally described by unitary matrices that finally lead to the desired logic gates.

In the following, an implementation of the hybrid qubit is presented. A sketch of the device is reported in Figure 1.17, where metal gates form two electrostatic QDs and control the energy barrier between them. However, additional structures are needed to inject electrons in the QDs. This can be achieved by fabricating a reservoir as source of electrons near the double QD and by controlling the height of an energy barrier between the reservoir itself and the double QD through an electrostatic gate. In addition, the fabrication of a charge sensor is needed for the readout of the qubit state which coincides with the read out of the spin state of electrons confined in the doubly occupied QD. To serve this purpose, a Single Electron Transistor (SET), which is a MOSFET where a QD is formed by placing additional lateral gates orthogonal to the channel, can be used to electrostatically sense the spin state of the electrons in the doubly occupied QD.

Once that the operations on the qubits are concluded, the next step is represented by the read out process, as is described in the following. When read out of the qubit starts, tunneling is allowed from the doubly occupied QD to the reservoir.
by a reduction of the interposed electrostatic barrier. When the electron pair is in a singlet state the corresponding wavefunction is more confined and the tunneling rate to the reservoir is lower than that of the triplet state, which has a broader wavefunction. When the electron tunnels, the electrostatic potential landscape changes and so does the current passing through the electrostatically coupled SET. The measurement of the time interval between the read out signal and the current variation in the SET is supposed to reveal the spin state of the electron pair [125].

By adopting the same approach as for the single-qubit logic gates, the extended effective Hamiltonian model for two interacting qubits is derived [37]. The total system, composed by six electron spins, is described within the subspace with total angular momentum operator $S = 1$ and $S_z = -1$ adopting the basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, where the logic state $|0\rangle$ and $|1\rangle$ are defined in Eq.(1.6). A possible layout for two hybrid qubit gates is sketched in Figure 1.18, where two data qubits are put in close connection by a controllable electrostatic barrier. The Controlled-NOT (CNOT) gate, for example, is obtained by using the sequence reported in Refs. [37, 26]. The gates in Figure 1.17 and 1.18 are sufficient to carry out arbitrary qubit rotations as well as general two qubit operations, providing a complete set of quantum gates for universal quantum computing.
Communication qubits: the SWAP chain

In this paragraph the problem of the communication among data qubits within an interconnecting circuit is analysed and an efficient strategy for an optimal transfer is presented. The model is based on hybrid qubits chains where the exchange interaction is exploited to transfer end to end the logic states through SWAP operations, where the SWAP gate operates an exchange between the states of two adjacent qubits.

In Figure 1.19 a scheme and a qualitative design of the hybrid qubit chain is shown where an even number of hybrid qubits are put into direct connection.

From a practical point of view it is necessary firstly to initialize each qubit, which is composed by $3 \times 2n$ electrons ($3$ is the number of electrons for each qubit and $2n$ is the total number of hybrid qubits). The transfer begins when the state of the system at the head of the chain has been exchanged through a SWAP operation with the state of the adjacent qubit. In this way the first qubit has received the state of the second qubit and vice versa. At the second step the same mechanism involves the second and the third qubit. At the end of the process, the information has been completely transferred to the last qubit. In this case the exchange is operated sequentially. An optimized control, as depicted in Figure 1.20, will allow to operate the exchange in parallel.
Figure 1.19: Top: schematic of a chain of an even number (2\(n\)) of hybrid qubits. Bottom: example of qualitative design for the implementation of quantum dot chains. The electrons confined in the QDs are highlighted in red. \(d_{ID}\) is the inter-dots distance, \(l_{CQ} = 2d_{ID}\) is the length of the communication qubit and \(d_{DQ}\) is the distance between the interconnected data qubits which corresponds to the distance between the head and the tail of the chain.

Instead of in sequence, optimizing the bidirectional transfer of the states as pictured in Figure 1.20.

In order to find the gate sequences necessary to generate a SWAP operation between two qubits we employed the same search algorithm used in [37]. The resulting pulse sequence to obtain the SWAP gate is reported in Figure 1.21, where the different \(J\) terms represents the exchange parameters [37].

Only a couple of electrons interact after tuning the tunneling parameters between the dots belonging to the same qubit or to different ones. Interactions between 1\(_a\) and 2\(_a\) and between 1\(_b\) and 2\(_b\) have been set to a constant value in the search algorithm, as they cannot be externally controlled [37]. More in detail, \(J_{1b2a} = J_{1a2b} = J/2\) where \(J\) is the maximum effective exchange interaction between the two dots.

In order to quantitatively design the qubit chain, the simulation results on the single qubit reported in Ref. [37] are used. The SWAP time, \(t_{SWAP}\), between the states of two adjacent hybrid qubits depends on the exchange interaction \(J\) that again depends on the tunneling rate \(t_r\) between the energy levels in the two QDs forming the qubit. \(t_r\) depends on the inter-dots distance \(d_{ID}\) and it is linked to the length of the communication qubit by \(l_{CQ} = 2d_{ID}\). The number of qubits 2\(n\) forming
A sequence of $2n - 1$ SWAP steps, with time duration $t_{SWAP}$, is applied when a chain of $2n$ hybrid qubits is considered. SWAP operations between qubits 1-2, 3-4, 5-6, etc. are required in odd time steps whereas qubit 2-3, 4-5, 6-7, etc. are swapped in even time steps. As a result, an independent control of all the gates in each SWAP step is not required. Gates can be grouped in two sets and driven alternatively, making the chain control easier. After $2n - 1$ SWAP steps, a bidirectional transfer of the states initially localized at the extremities of the chain is obtained.

The total time $t_{TOT}$ to transfer the information from one extremity to the other by successive SWAP operations is:

$$t_{TOT} = (2n - 1) \cdot t_{SWAP} = \left( \frac{d_{DQ}}{l_{CQ}} - 1 \right) \cdot t_{SWAP}(d_{ID}) = \left( \frac{d_{DQ}}{2d_{ID}} - 1 \right) \cdot t_{SWAP}(d_{ID}) \tag{1.8}$$

where $t_{SWAP} = t_{seq} \cdot \hbar / J$ and $t_{seq}$ is the duration of the sequence in units of $\hbar / J$. $J$ is estimated with $J = t_c^2 / \Delta E_{ST}$ where $\Delta E_{ST}$ is the singlet-triplet energy splitting [125]. In Figure 1.22 the total chain time $t_{TOT}$ is reported as a function of $d_{ID}$ for three different values of $d_{DQ}$. Total time $t_{TOT}$ increases exponentially as $d_{ID}$ raises.

### 1.4.3 CMOS implementation of the hybrid qubit architecture

In this paragraph we explore the limiting size of the discrete components to implement the hybrid qubit architecture imposed by a CMOS-compatible fabrication process.

The standard in semiconductor industry is set by silicon CMOS manufacturing, due to the capability to fabricate p- and n-channel devices on the same chip.
Figure 1.21: Waveforms of the effective exchange variables implementing a SWAP gate (up to a global phase) with fixed $J_{12} = J/2$ in both qubits. Times are in unit of $\hbar / J$.

and to build devices with a low power consumption [97]. Hence, the technologic constraints set by the 22 nm technologic node of CMOS nanoelectronics are examined in Section 1.4.3. According to such physical constraints, realistic data and communication qubits are designed and reported in Section 1.4.3, providing the building blocks for a complete implementation of the hybrid qubit architecture in silicon.
Figure 1.22: Graph of the total time SWAP chain $t_{TOT}$ as a function of the inter QD distance $d_{id}$ for three different distances $d_{DQ}$ between head and tail data qubits.

**Technologic constraints of CMOS manufacturing at the 22 nm node**

Figure 1.23 shows a schematic process flow for the realization of Si-MOS hybrid qubits on a Silicon-On-Insulator (SOI) platform.

Quantum gates are defined by selective etching of the SOI device layer after a lithographic patterning exposure. Then a gate insulator, for example Al$_2$O$_3$ or other high-$k$ oxides, is deposited on the silicon islands [97, 135]. Finally, the metal electrodes for source-drain leads and the electrostatic gates are deposited and patterned by a further lithographic exposure. The mentioned process flow can be completely adapted to an industrial one: all the steps can be performed with the main deposition and etching techniques employed in a common industrial production line, such as Chemical Vapor Deposition (CVD), Atomic Layer Deposition (ALD) and Reactive Ion Etching (RIE) [97].

The main critical issues concern the lithographic steps, because a few nanometer resolution is required. At the present times, most of the patterning processes in microelectronics and micromachining for Ultra Large Scale Integration (ULSI) are carried out through Deep Ultra-Violet (DUV) lithography, that makes use of ArF laser sources ($\lambda = 193$ nm) and is capable of high resolution as well as high throughput (100 wafers per hour) [97, 135]. According to the last International Technology Roadmap for Semiconductors (ITRS) Update, 22 nm is the minimum half-pitch of un-contacted Poly-Si in flash memories and it represents a significa-
Figure 1.23: Schematic process flow for a CMOS-compatible realization of semiconductor hybrid qubits on SOI wafers. A schematic top view of the resulting device is shown on the right bottom.

A minimum feature size of 20 nm is a reasonable design rule for a realistic implementation of the hybrid qubit architecture proposed in Section 1.4.2. All the masks for data and communication qubits have been designed accordingly and are compatible with the 22 nm node technology.

**Realistic design of data and communication qubits**

The mask for a single hybrid qubit is reported on the left of Figure 1.24. One lithographic level, that defines the SOI islands, is in blue, while the other levels correspond to four superposed metal levels for the electrical contacts, high doping for the electron *reservoirs* and vias to Back End Of Line (BEOL) levels. In Figure 1.24 we highlighted the silicon regions where the DQD qubit and SET charge sensor are defined. Grey and green electrodes (level 1 and 3) are used as inter-dot...
Figure 1.24: Lithographic masks for data qubits. Left: 1-qubit gate is composed of a DQD contacted with an electronic reservoir (2DEG) for the system initialization and read-out and a SET for charge sensing of the DQD. Right: 2-qubit gate. The mask is designed analogously to the 1-qubit gate to obtain full control over two independent qubits coupled by a tunable electrostatic barrier. A color code on the right side identifies the lithographic levels required for the fabrication process, corresponding respectively to the definition of silicon islands, degenerate doping at the source and drain contacts, four levels for metal gates and vias for electrical connections. The minimum feature size is 20 nm for both masks, whereas the total area is 300 x 500 nm$^2$ for one qubit and 380 x 500 nm$^2$ for two qubits.

barriers, while red and light blue ones (level 2 and 4) act as plunger gates defining the potential wells and controlling the chemical potential in the quantum dots. The minimum feature size for structures on the same level is 20 nm. As a result, the one-qubit gate in Figure 1.24 can be realized on a submicrometer area (300 x 500 nm$^2$).

A two-qubit gate can be obtained as a replica of the one-qubit mask with few adjustments. The mask on the right of Figure 1.24 covers an active area of 380 x 500 nm$^2$ and it’s composed of two DQDs with separated electron reservoirs and SET charge sensors for independent initialization and read-out of the two qubits. The doubly occupied dot is closer to the charge sensor in both qubits and can directly communicate with the electron reservoir to facilitate the read-out procedure. The gates in Figure 1.24 are sufficient to carry out arbitrary qubit rotations as well as general two qubit operations, like the Controlled-NOT (CNOT). As a results, such devices provide a complete set of logic gates and represent the starting point to perform universal quantum computation with the hybrid qubit architecture.
Figure 1.25: Communication qubits for the coherent transfer of quantum information in a 2-dimensional array of qubits. Chain modules enable quantum communication between adjacent qubits through a SWAP logic operation. The T module, on the right side, is a modified version of the chain module that makes it possible to bring in contact orthogonal qubit chains. The lithographic levels are depicted according to the color code in Figure 1.24. The shaded squares indicate the active area of the chain module (160 x 460 nm$^2$) and of the T module (1300 x 700 nm$^2$).

According to the discussion in Section 1.4.2, quantum communication can be accomplished by sequential SWAP operations across a qubit chain. We designed two modular structures, namely the "chain" and the "T" module, to be composed in arbitrary 2-dimensional arrays of communication qubits (Figure 1.25). The chain module consists of two qubits controlled by independent plunger and barrier gates with the only scope to carry out a SWAP operation between the states of the two qubits with no need to initialize and read-out the logic states. Analogously, the T-module is a modified and rearranged version of a multi-chain module and acts like a crossroad for flying qubits: orthogonal qubit chains are brought in contact through this module, creating the conditions for 2-dimensional arrays of qubits.

1.4.4 Quantum computing on a large scale

In this Section a possible integration on a large scale of silicon hybrid qubits is evaluated and the maximum quantum information density per unit surface is estimated. The occurrence of faulty logic gates and memory errors is taken into account and a Quantum Error Correction scheme is proposed to improve the effective gate fidelity in multi-qubit circuits. In this framework, two important figure of merit are estimated: the maximum density of logic qubits per unit area and the
Gate fidelity is an important metric in a QIP architecture, since errors are much more frequent in quantum computers than in their classical counterparts. In fact, information in qubits is rapidly corrupted by decoherence, i.e. the interaction with the surrounding environment. In hybrid qubits charge and spin noise are the main sources of decoherence and they induce unwanted rotations in the Bloch sphere. As a result, charge and spin noise are responsible for errors with a probability of about $10^{-3}$ errors per logic gate [63]. Generally, protection against errors in QIP is achieved through bit encoding and fault-tolerant computation [28].

A multi-qubit circuit is considered to this extent in Figure 1.26, where many data qubits are connected by T-modules on a bus structure. Here a logic qubit can be encoded by 7 physical qubits according to the [[7, 1, 3]] Steane code, allowing for fault-tolerant computation and quantum error correction (QEC) [28, 130]. In this framework, the information of a logic qubit is stored in a 2-dimensional subspace of the $2^7$-dimensional Hilbert space defined by 7 physical qubits. As a result, a logic qubit is less susceptible to single physical qubit failures, since a faulty gate can be revealed and corrected with standard QEC techniques maintaining the coherence of the logic qubit. Actually, QEC algorithms require some supplementary
Figure 1.27: Lithographic mask of a quantum register made of 8 logic qubits (A-H). Every logical qubit is composed of 20 double data qubits (see the right side of Figure 1.24) depicted as blue boxes and 20 T blocks (see Figure 1.25) colored in red. Connections between logic qubits are provided by chain modules (see Figure 1.25) colored in green and 8 additional T modules. Such logic quantum byte is composed of 1720 data qubits and 1400 communication qubits and covers an area of 307.502 µm² for the measurement of the syndrome of the logic qubit. The number of such extra-qubits, or ancillae, generally depends on the quantum code, the QIP architecture and the specific procedure of error detection and correction. In particular, 12 auxiliary qubits are sufficient for a complete QEC algorithm with the [[7,1,3]] code [28, 21].

In the bus structure of Figure 1.27, one of the 8 branches is a logic qubit according to the [[7,1,3]] code and it is composed of 20 physical data qubits, including ancillae qubits. Tab. 1.2 reports the dimensions and compositions of the principal quantum gates and multi-qubit circuits presented in this work. In particular, the 8-qubit block in Figure 1.27 is the quantum analog of the classical unit of information (the byte) and can be taken as a first benchmark to estimate the maximum density of logic qubits per unit area. Such register covers an active area of 25.54 x 12.04 = 307.502 µm², that corresponds to a density of information of 2.6 Mqubit per cm².

A remarkable property of the [[7,1,3]] code is that the fundamental logic gates operate on logic qubits in complete analogy to logic gates on physical qubits. More in detail, the operation principle of a logic gate between two encoded qubits A
Figure 1.28: Operation of a two-qubit logic gate between encoded qubits. The reported mask refers to the box highlighted in Figure 1.27. A logical qubit (E) is transferred in proximity of a second logical qubit (F) by sequential SWAP operations through communication qubits. Then a two-qubit logic gate (*e.g.* CNOT) is operated between all the couples of physical qubits (*E*/*F*) within data qubit blocks, effectively carrying out a CNOT gate between the logic qubits (*E*/*F*). The target qubit F is modified by the CNOT gate according to the state of the control qubit E, resulting in a new quantum state *F*' in the end the two qubits are moved to other qubit registers for further logic operations.
and B is sketched in Figure 1.28. Logic qubit A is firstly transferred through a SWAP channel in proximity to qubit B, where the logic gate is carried out on a one-by-one basis between all the couples of physical qubits. Finally qubit A is brought back to the starting point or directed to another qubit site for the next step in the algorithm. Notably, all the logic operations are performed fault-tolerantly within this scheme, i.e. interaction between physical qubits in the same logic qubit never takes place. As a consequence, the propagation of errors inside a logic qubit is forbidden, preserving the possibility to perform quantum error correction over single faults.

Table 1.2: Physical dimensions and composition of data and communication qubits with \( d_{ID} = 40 \) nm. The last two rows report the dimensions of a logic qubit and of a register of 8 logic qubits respectively.

<table>
<thead>
<tr>
<th>Device</th>
<th>Dimensions ([\mu m^2])</th>
<th>Area ([\mu m^2])</th>
<th>Data/Comm. qubits</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-qubit</td>
<td>0.3 x 0.5</td>
<td>0.15</td>
<td>1/0</td>
</tr>
<tr>
<td>Two-qubit</td>
<td>0.38 x 0.5</td>
<td>0.19</td>
<td>2/0</td>
</tr>
<tr>
<td>Chain</td>
<td>0.16 x 0.46</td>
<td>0.0736</td>
<td>0/2</td>
</tr>
<tr>
<td>T</td>
<td>1.3 x 0.7</td>
<td>0.91</td>
<td>0/7</td>
</tr>
<tr>
<td>1 Log. qubit</td>
<td>11.38 x 2.52</td>
<td>28.6776</td>
<td>20/70</td>
</tr>
<tr>
<td>8 Log. qubits</td>
<td>25.54 x 12.04</td>
<td>307.502</td>
<td>1720/1400</td>
</tr>
</tbody>
</table>

In order to evaluate the physical performances of this architecture, we report in Tab. 1.3 the path length between two adjacent physical/logic qubits to estimate the corresponding time needed for quantum communication. According to the analysis in Subsection 1.4.2, \( t_{SWAP} = 6.47 \) ns for \( d_{ID} = 40 \) nm. As a result, the time needed to transfer quantum information through a SWAP chain ranges from 71.2 ns for communication between adjacent physical qubits to approximately 2 \( \mu s \) for coherent transfer between logic qubits within a 8-qubit register.

The estimated characteristic times for quantum information processing should be compared to the qubit coherence time \( T_2^* \). Although the first experimental works reported a short \( T_2^* \sim 20 \) ns for a silicon hybrid qubit [124, 59], the expected value from theoretical calculations is of the order of \( \mu s \) [125]. A fidelity of 99.999% seems to be within range for 1-qubit operations and could be further improved by tuning the singlet-triplet splitting to a good balance between operational speed and gate fidelity [125, 63]. Besides this, the effects of the principal sources of decoherence could be drastically reduced with several techniques, such as dynam-
Table 1.3: Time load for quantum computation and communication between distant data qubits with a 40 nm inter-QD distance. According to the analysis in Subsection 1.4.2, $t_{\text{SWAP}}$ is 6.47 ns. The minimum and maximum transfer times for quantum communication have been calculated considering the minimum and maximum distance between physical qubits in the same logic qubit (made of 20 physical qubits) and between different logic qubits in a byte (8 logic bits).

<table>
<thead>
<tr>
<th>Operation</th>
<th>Number of Qubits</th>
<th>Distance [$\mu$m]</th>
<th>Time [ns]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comm. 2 phys. qubits (min)</td>
<td>12</td>
<td>1</td>
<td>71.2</td>
</tr>
<tr>
<td>Comm. 2 phys. qubits (max)</td>
<td>138</td>
<td>11</td>
<td>886.4</td>
</tr>
<tr>
<td>Comm. 2 log. qubits (min)</td>
<td>192</td>
<td>15.4</td>
<td>1235.8</td>
</tr>
<tr>
<td>Comm. 2 log. qubits (max)</td>
<td>311</td>
<td>24.9</td>
<td>2005.7</td>
</tr>
</tbody>
</table>

We also note that the bus-structure reported in Figure 1.27 can be easily extended to higher order ramifications in order to introduce recursive coding techniques [28]. A recursive code of order $k - 1$ gives an error threshold of $(cp)^{2^k}/c$ where $p$ is the error probability of a logic operation and $1/c$ is the error threshold, i.e., the maximum error rate tolerated by a specific quantum code [28, 21]. As a result, recursive coding rises the error threshold by re-encoding logical qubits in a higher level logic qubit provided that $p < 1/c$. If this condition is satisfied, the error threshold is enhanced by an exponential law, whereas the circuit area and the computational times increase by a power law [21].

### 1.4.5 Conclusions

A CMOS-compatible design of the semiconductor hybrid qubit architecture has been proposed. Such architecture is suitable for large scale quantum computing, since it allows all-electrical manipulation of qubits on a nanosecond timescale.

One- and two-qubit gates have been designed for a Si-CMOS platform, complying with the technologic standards of semiconductor industry. The fundamental building blocks for quantum computation and communication have been proposed, and the feasibility of multi-qubit networks has been discussed. The requirements of fault-tolerant computation and the introduction of a quantum error correction
scheme based on the [[7, 1, 3]] Steane code have been taken into account. The time and space resources for universal quantum computation are estimated accordingly in a register of 8 logical qubits.

The calculated maximum surface density of logical qubits is 2.6 Mqubit/cm^2.
Chapter 2

Emerging Research Materials: MoS$_2$

The exploitation of Emerging Research Materials is a second approach, alternative to the concept of Emerging Research Devices, to support the continuous functional scaling in post-Si electronic devices.

Layered materials, also known as graphenic materials from their famous cousin namely graphene, naturally refer to this category for their unique properties arising from the intrinsically two-dimensional nature. Among the wide variety of two-dimensional materials, Transition Metal Dichalcogenides (TMD) and molybdenum disulfide in particular are extensively studied in the perspective of a massive utilization in future generations of electronic devices. Differently from graphene, they also match the requisite of a relatively wide band-gap to obtain a large on-off current ratio in logic circuits. As a consequence, the ultimate scalability provided by single layer MoS$_2$ is expected to significantly reduce the short channel effects in end-of-roadmap devices and become competitive with ultrathin Silicon-On-Insulator and III-V semiconductor channels. Furthermore, the peculiar properties of MoS$_2$ band structure are suitable for other applications in spin- and valleytronics, optoelectronics, sensoristics and flexible electronics.

In this chapter, MoS$_2$ integration as the channel material in standard MOS FET structures will be addressed. Section 2.1 will review the main properties and the potential applications of two-dimensional materials. Section 2.2 will go into further detail with MoS$_2$ mainly. The electrical characterization of MoS$_2$ MOSFETs will be reported in Section 2.3, with a particular attention on electrical transport and the spectroscopy of interface traps.
2.1 2D Materials: Flatlands Beyond Graphene

The Nobel prize awarded to A. K. Geim and K. S. Novoselov for the discovery of 2D graphite, namely graphene, fostered new interest in 2D materials science. Indeed, although the existence of these crystals was already known, the potential of such materials was still to be fully understood.

The 2D crystal of graphene sets the proper conditions for many physical phenomena that cannot be observed in bulk 3D materials [13]. For example, the linear dispersion of the electronic conduction band is a quite peculiar feature in material science and provide a powerful tool to study relativistic effects in a solid state environment. Furthermore, the record-breaking mechanical and electrical properties of graphene generated a sharp technological interest targeting multiple applications in nanoelectronics. The success of this material as the replacement of silicon for next generation devices is still under debate due to the lack of a band-gap, that is actually a key requirement for logic devices. Nonetheless, graphene has proven to be a golden material from a scientific point of view, as it can be noticed from the great number of graphene-related publications. A considerable effort has been done to set up theoretical and experimental methods to study this material and deposition techniques are available nowadays to grow high quality graphene on a centimeter scale.

2D material science took full advantage of this scientific and technologic breakthrough driven by graphene since most of the deposition and characterization techniques can be generalized and transferred to other materials. Some reasons for the interest around these materials, in fact, are inherently coupled to their 2D character, such as the absence of dangling bonds and inter-layer covalent interactions. Moreover, a variegated phenomenology arises from the specific crystal and electronic structure of such materials: massless fermions like in graphene, superconductivity, topological protected states and spin or valley polarization of carriers may arise depending on their physical and chemical composition.

Layered 2D materials include group IV X-enes (graphene and its "cousins" silicene and germanene) that arise from the sp² hybridization of C, Si, and Ge orbitals respectively and their H-terminated counterparts namely X-anes. Other important materials are MX-enes and MX₂-enes, where M and X generically indicate metal and chalcogenide atoms and much more like phosphorene, hexagonal boron nitride h-BN and BiTe. A particular mention goes to transition metal dichalcogenides (TMDs), that share the chemical formula MX₂ where M is a transition metal (e.g. Mo and W) and X is a chalcogen (e.g. S, Se, Te) [145, 17, 54]. They can be exfo-
liated maintaining a good stability as self-standing monolayer materials and they exhibit a wide range of electronic properties from metallic to wide gap semiconductors depending on the specific chemical composition. In the following I will focus on MoS$_2$, that is by far the most studied material among TMDs [41].

2.2 MoS$_2$

Molybdenum disulphide (MoS$_2$) exists in nature as the molybdenite crystal. For this reason, the presence of MoS$_2$ as a layered material was known since the 1960s, but only in the last decade it raised as an interesting material for microelectronics [113, 114, 72, 34, 112, 41].

Like other 2D materials, it is characterized by strong in-layer covalent bonds, while different layers are stacked by weak van der Waals forces [84]. As a result, MoS$_2$ is a layered material that can be easily exfoliated up to a single layer, remaining stable after exposure to atmospheric pressure and relatively high temperatures. It has a remarkable mechanical strength and maintains its electronic structure under exceedingly intense strain, making it an appealing candidate for applications in flexible electronics and microelectromechanical systems [8, 120, 15, 111, 119].

The vibrational properties of MoS$_2$ have been extensively studied by means of Raman spectroscopy, since four vibrational modes are Raman active modes [74, 161, 136]. The energy difference between the A$_{1g}$ breathing mode (out-of-plane vibrations) and the E$_{2g}$ shear mode (in-plane vibrations) vary with the number of layers and is a fast and powerful tool to determine the thickness of MoS$_2$ thin films.

Mo and S atoms are arranged with a trigonal simmetry where the chemical bond length and the layer thickness are 2.4 and 7 Å respectively [57, 41]. Bulk MoS$_2$ is a semiconductor with an indirect band-gap of 1.3 eV, however a metastable metallic phase can be obtained inducing a transition to octahedral simmetry through lithium intercalation [33]. Valence band maximum is located at the Γ point in the Brillouin zone, while conductance band minimum is about halfway towards the K point, with a pretty high electron effective mass $m^* = 0.48m_e$ [41, 152].

A peculiar property of MoS$_2$ is the thickness-dependent band-gap [84, 57, 55]. The direct gap at K is almost constant against the number of layers, while the Γ states are more sensitive to inter-layer interactions. As a result, reducing the crystal thickness the indirect gap increases from the bulk value (1.3 eV) up to the monolayer case. In single layer MoS$_2$ the indirect gap overcoming the 1.9 eV direct gap at the K point, and MoS$_2$ actually becomes a direct-gap material, with
strongly enhanced absorption, photoluminescence and photo-generation of carriers \cite{84, 145}. From this point of view, single layer MoS\(_2\) has unique properties in terms of thickness scaling compared to other materials for optoelectronics.

Furthermore, the strong spin-orbit coupling and the breaking of inversion symmetry in single layer MoS\(_2\) lift the spin degeneracy at the conduction band minima, allowing for spin polarization of carriers. Besides this, electrons have opposite spin-orbit coupling in different valleys due to time reversal symmetry, resulting in a coupling between spin and valley degree of freedom. As a result, it is possible to control the spin and valley polarization of carriers with a polarized light source, setting the background for spin- and valley-tronic devices based on MoS\(_2\) \cite{83, 154, 44}.

\subsection{Deposition of MoS\(_2\)}

The first method that have been reported to obtain single layer flakes of layered materials is the mechanical exfoliation, or "scotch tape method", that actually launched the gold-rush to graphene and 2D materials \cite{13}. Since inter-layer forces are much weaker than intra-layer covalent bonds, if a sticky scotch tape is applied on the surface of an MoS\(_2\) crystal, it is likely to remove small fragments of material that are very few layers thick. Single layer MoS\(_2\) can be eventually obtained by repeating this procedure and finally depositing the resulting fragments on a proper substrate. The exfoliation process can’t be properly controlled, consequently the size of exfoliated flakes is generally limited to few tens micrometers. The yield is not very high indeed, but mechanical exfoliation is the most popular method to exfoliate layered materials, since it is actually an effective way to obtain wide domains of high quality monolayer flakes with very low density of defects \cite{114, 113, 112}.

A newer method is the chemical exfoliation, where the inter-layer binding forces are further reduced by lithium ion intercalation or by immersion in organic solvents and following sonication \cite{20, 102, 23}. In both cases it is possible to reach a higher yield for the production of single layer MoS\(_2\) at the expense of material quality: such processes normally involve high temperature annealing and sonication processes that degrade the material reducing the grain dimensions to about 100 nm. Concluding, exfoliation methods are hardly controlled processes that are compatible with research purposes but not with large scale integration of electronic devices.

A possible answer to the demand of large scale growth of thin MoS\(_2\) is given by
CVD techniques [155, 78, 94, 119, 79, 156, 121, 41] based on the sulfurization of Mo precursors carried out in furnace at high temperature (400 - 1000°C). Common precursors are molybdenum trioxide (MoO$_3$), ammonium molybdate ((NH$_4$)$_2$MoS$_4$) and simple Mo thin films deposited on SiO$_2$ or sapphire substrates. Large area of monolayer MoS$_2$ has been growth with this technique, though with higher density of defects and smaller grain dimensions compared to exfoliated flakes.

### 2.2.2 MoS$_2$ electronic devices

The opening of a fairly wide band-gap is one of the strong point supporting the integration of MoS$_2$ as the channel material in logic devices [13, 41, 54]. In fact, MoS$_2$ has a much lower mobility compared to graphene and III-V semiconductors, however a band-gap of more than 1 eV sets the ground to build devices with high on-off current ratio and low power consumption [145, 113, 60, 15, 25, 114]. More precisely, the intrinsically 2D character of MoS$_2$ could be the key to outperform silicon in terms of short channel effects, enabling channel length scaling below 10 nm [152, 1]. From this point of view, the relatively high effective mass can become an additional advantage in reducing the tunneling current in such extremely scaled devices. In 2011, the operation of a FET based on top-gated single layer MoS$_2$ showed encouraging results: on-off current ratio higher than $10^8$, field effect mobility $\approx 200$ cm$^2$/Vs and a subthreshold slope of 74 mV/dec at room temperature [113]. The electrical properties of MoS$_2$ strongly depend on the substrate and on the ambient atmosphere for bottom-gated devices [4] as well as on the capping layer for top-gated devices [113]. The employment of a top gate, in particular, is beneficial to the channel mobility by mitigating the Coulomb scattering through dielectric screening [60, 112].

The direct band-gap in single layer MoS$_2$ leads to additional perspectives in optoelectronics [80, 19, 151, 71, 157, 115]. Optical absorption is dominated by the generation of bound excitons via inter-band absorption of photons. Photocurrent is then generated in photodiodes and phototransistors through an electric field that separates the electron-hole pairs. Such electric field is generally introduced as a voltage bias between the source-drain leads or as a built-in field in heterojunctions, like Schottky contacts.

Finally, another point of interest related to MoS$_2$ is the opportunity to build up heterojunctions with other 2D materials combining the properties of other materials like graphene, hexagonal boron nitride and different TMDs [18, 16, 123, 42]. This approach lead to the realization of tunneling FET, p-n TMD junctions and
elementary logic and memory devices of nanometre thickness [153, 145, 54].

Concluding, the intriguing device perspective presented in this Section still have some issues to be considered in view of a large scale integration of MoS$_2$ devices for high-performance logic circuits. On the one hand, reducing the contact resistance, finding a reliable doping method and a better control of work-function engineering are actually the major challenges to realize logic circuits in a CMOS framework [54]. On the other hand, the excellent mechanical properties of ultra-thin MoS$_2$ could be fully exploited in thin-film transistors and transparent/flexible electronics [15, 119]. The performance requirements are much less demanding for these applications and MoS$_2$ has a great potential compared with the main competitor materials, that are organic semiconductors, thanks to the better electronic properties and the extreme scalability.

2.3 MoS$_2$ MOSFET

MoS$_2$ MOSFET in the last few years exhibited great prospects for various applications that go far beyond bare nanoelectronics in a MoreMoore framework. A number of recent works already demonstrated the possible integration of MoS$_2$ in a MOS FET, including single layer MoS$_2$. Remarkable performances were obtained by optimizing the gate stack in a double gate configuration, which lead to improved electrostatic control and mobility. Nonetheless, a single gate configuration is preferable in a series of alternative applications, e.g. sensoristics and optoelectronics. This Section is dedicated to the fabrication and the electrical characterization of single gate MOS FET incorporating mechanically exfoliated MoS$_2$.

2.3.1 MOS FET fabrication

MoS$_2$ MOS FET were fabricated on two sets of commercial Si/SiO$_2$ wafers with different oxide thickness: 90nm or 285nm. Such a large thickness ensures a good optical contrast, as required to recognize few-layer fragments of MoS$_2$ that will be deposited on top. Si is strongly doped (n++) in order to use the substrate as a back-gate in a MOS FET configuration.

The substrate was patterned by EBL and subsequent SiO$_2$ etching in HF at first to define a reference grid. Such reference markers are fundamental for an easy addressing of the MoS$_2$ flakes and the eventual definition of electrical contacts, as will be explained later in this Section. MoS$_2$ micrometric fragments were deposited by mechanical exfoliation with the conventional scotch-tape technique. The thinnest
Figure 2.1: Individuation and validation of MoS$_2$ flakes by Raman spectroscopy. a) Comparison between the Raman spectra of a bulk and a 3-layer flake. The two peaks are associated with the E$_{2g}$ and A$_{1g}$ vibrational modes. b) Optical microscope image of a MoS$_2$ flake. Inset: Map of the intensity of the E$_{2g}$ peak in the area delimited by the dashed line (7.5 x 7.5 µm$^2$).

fragments were selected after a first inspection at the optical microscope: generally thinner flakes appear darker than bulk MoS$_2$ thanks to the optimized oxide thickness of the substrate. The quality of the material is validated by means of Raman spectroscopy. Two vibrational modes are Raman-active in MoS$_2$ [161], namely the E$_{2g}$ and the A$_{1g}$, which are sketched in Figure 2.1a. Notably, the energy difference between these two modes is thickness dependent up to 4 layers, giving an indirect measure of the flake thickness [74]. The good crystalline quality of the exfoliated material is assessed by the uniform surface map of the E$_{2g}$ Raman peak reported in Figure 2.1b.

Then, electrical contacts to selected flakes were defined by EBL and e-beam evaporation of Ti/Au onto the MoS$_2$ flakes. Figure 2.2a reports optical micrographs of a substrate after the procedure of mechanical exfoliation. The reference grid, previously defined with a 100 µm periodicity by EBL, is clearly visible as well as the abundance of MoS$_2$ flakes deposited on the substrate. The size of manageable MoS$_2$ flakes is in the range of few microns. A thin flake has been selected according to the thickness-dependent optical contrast and the Raman characterization to build up a MOS FET. The pictures in Figure 2.2a have been post-processed and imported in the EBL software to be aligned to the reference markers, depicted in blue in Figure 2.2b. Using this alignment procedure, metal electrodes could be defined on the flakes with reference to the pre-patterned marker grid, that are also easily recognizable by SEM imaging. EBL exposure was aligned accordingly and
Figure 2.2: Manufacturing of a MoS$_2$ MOS FET. a) Optical micrographs of the substrate after mechanical exfoliation of MoS$_2$. The flake inside the blue circle, magnified in the panel at the right top, has been selected to be incorporated in a FET structure. b) EBL mask for the metal electrodes. The morphology of the sample is depicted in red after, allowing alignment to the blue reference grid. The pattern of metal electrodes is represented in light blue. c) Image from the optical microscope of the FET designed in panel b). Another FET is shown in the inset, including a thicker and more evident MoS$_2$ flake.

Ti/Au contacts (80nm Au over 5nm Ti adhesion layer) were deposited by e-beam evaporation and following lift-off, yielding the 2.2c.

As a final remark, the material integrity is not affected by this process flow, as could be verified by Raman spectroscopy over a final device (Figure 2.3).

2.3.2 Electrical transport in single gate MOS FETs

Standard characterization techniques have been employed to study the electrical transport in MoS$_2$ FET. Room temperature characteristic (drain-source current $I_{ds}$ vs drain voltage $V_d$) and transcharacteristic curves (drain-source current $I_{ds}$ vs gate voltage $V_g$) were recorded with a HP-4140B pA-meter and an Agilent B1500 semiconductor device parameter.

The impact of metal-semiconductor contacts on electrical transport

Figure 2.4 compares the current characteristic curves of two representative devices with ohmic (4FET1.1) and strongly rectifying (4FET3.3) contacts. Variability in the device electrical properties is expected due to the built-in Schottky barrier at the metal-semiconductor interface. In particular, charge injection mechanisms at Schottky junctions are strongly dependent on the quality of the interface (roughness, defects) as well as the junction depth [132]. As a result, when asymmetric
contacts are built at the source and drain junctions, rectifying behavior occurs, as in device 4FET1.1. A second consequence of the inherent Schottky junctions is an extremely high contact resistance [77, 33], that ultimately limits charge transport in rectifying device 4FET1.1, as shown in Figure 2.4.

The fabrication of good electrical contacts to MoS$_2$ is actually a major issue in MoS$_2$ FET, fostering extensive studies on this topic [77, 33, 54]. Recent works compared different materials, indicating Sc as an optimal candidate thanks to a good matching to the work-function of MoS$_2$ [24]. Another strategy to achieve ohmic contacts is conceived in analogy with the degenerate doping employed for ohmic contacts in CMOS devices. In the specific case of MoS$_2$, high n-doping densities can be obtained with the introduction of K or Na. Furthermore, a phase transition to the metallic octahedral phase can be induced by lithium intercalation, providing the high carrier density required to reduce the Schottky barrier depth [17, 41].

In this work a conservative choice was made by utilizing Ti/Au contacts, which have been successfully employed in previous works. Several devices were produced on each sample and further analysis was focused on the devices exhibiting the best electrical properties.
Advanced analysis on multi-layer MoS$_2$ FETs

Figure 2.5 reports the electrical characterization of two representative FETs built on multi-layer MoS$_2$ flakes from the same substrate.

Here again the variability related to the properties of the contacts show up from the evident asymmetry of the direct and transfer curves of device 4FET7 (Figure 2.5c-d). Nonetheless, efficient current conduction is observed in both devices and a more detailed study of their electrical properties can be done by varying the gate potential V$_g$ and the drain-source bias V$_{ds}$ over a wide range. The important parameters in Table 2.1 have been extracted from the size features and the electrical transport measurements according to consolidated procedures in semiconductor devices characterization [132].

Current is limited by the contact resistance originated by Schottky barriers, that is of the order of M$\Omega$ for both devices. As a consequence, a direct measurement of the MoS$_2$ resistivity and mobility is not available in the two-contact configuration used in this experiment. Anyway, the field effect mobility $\mu_{\text{FET}}$ can be extracted according to equation:

$$\mu_{\text{FET}} = \frac{L}{W} \frac{G}{C_{\text{ox}} V_d}$$

(2.1)
where $L$ and $W$ are the channel length and width respectively, $V_d$ is the drain voltage, $C_{ox} = \varepsilon_{ox}/d_{ox}$ is the oxide capacitance and $G$ is the transconductance $dI_{ds}/dV_g$.

The inset of charge accumulation, and of current conduction as well, occur at a similar threshold voltage $V_{th}$ for the two devices. Another important parameter is the subthreshold swing $SS$, which is a measure of the switching efficiency of the FET under the effect of back-gate. This figure of merit is directly related to the density of interface traps, as will be explained later in Section 2.3.3.

Similar values were extracted for the field-effect mobility, the threshold voltage and the subthreshold swing, indicating a good reproducibility of the process integration developed for bulk MoS$_2$ FETs. Mobility is roughly one order of magnitude lower than the state of the art [159, 60, 4]. Such record values of the order $\sim 100$ cm$^2$V$^{-1}$s$^{-1}$, however, could be attained after the application of a top-gate and dedicated substrate engineering, which are well beyond the explorative scope of this work.
Table 2.1: Device parameters of transistors 3FET7 and 3FET8. See text for the description of parameters. The gate oxide thickness and dielectric constant are respectively $\varepsilon_{\text{ox}} = 3.9\varepsilon_0$ and $d_{\text{ox}} = 285\text{nm}$.

<table>
<thead>
<tr>
<th>Device</th>
<th>3FET7</th>
<th>3FET8</th>
</tr>
</thead>
<tbody>
<tr>
<td>L [$\mu$m]</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>W [$\mu$m]</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>$G/V_d$ [A]</td>
<td>$4.0 \cdot 10^{-8}$</td>
<td>$7 \cdot 10^{-8}$</td>
</tr>
<tr>
<td>$\mu_{\text{FET}}$ [cm$^2$V$^{-1}$s$^{-1}$]</td>
<td>1.1</td>
<td>4.6</td>
</tr>
<tr>
<td>$V_{\text{th}}$ [V]</td>
<td>-7</td>
<td>-5</td>
</tr>
<tr>
<td>SS [V/dec]</td>
<td>2.94</td>
<td>1.95</td>
</tr>
</tbody>
</table>

Figure 2.6: FET transconductance in the dark and under illumination. a) Two-layer MoS$_2$ device 3FET6. b) multi-layer MoS$_2$ device 3FET8.

Generation of photocurrent in a bi-layer MoS$_2$ FET

Bilayer MoS$_2$ lies at the borderline between the pure direct gap typical of monolayer crystal and the indirect gap of multi-layer MoS$_2$ [84, 41]. As a consequence, optoelectronic effects like the enhanced photogeneration of charge carriers under illumination observed in monolayer MoS$_2$ can occur in two-and three-layer MoS$_2$ as well [151, 115, 160, 157, 71].

The identification of a bilayer flake was validated by Raman spectroscopy: the energy difference between the $E_{2g}$ and the $A_{1g}$ modes for the flake 3FET6 is 22cm$^{-1}$, compatible with a thickness of two-layers [74]. Figure 2.6a compares the dark current of device 3FET6, incorporating bilayer MoS$_2$, with the FET current at the same bias conditions but under light irradiation.

Significative photocurrent generation is observed, approaching 100% of the dark value in weak accumulation. On the contrary, no photocurrent was measured...
in multi-layer devices under equivalent conditions (Figure 2.6b). The symmetric behavior with respect to the drain bias reasonably excludes side-effects related to contact junctions and allows a reliable identification of the origin of such photocurrent in the generation of charge carriers by photon absorption in the FET channel. Light was delivered to the sample from a LED source through the optical microscope integrated in the electrical characterization setup. Consequently, quantitative considerations on the quantum efficiency and the spectral response of this device are not possible. Nonetheless, the observed electronic and optoelectronic properties of bilayer MoS$_2$ were obtained from the simplest MOS FET configuration, with back-gate manipulation and without post-processing and optimal passivation by dielectric capping [113, 15]. These results confirm the potential of few-layer MoS$_2$ for the production of ultrascaled and efficient photodetectors [80, 19, 151, 157, 71].

### 2.3.3 Admittance spectroscopy of interface traps

Interface traps are one of the main factors limiting the performance of MOS FETs by degrading the channel mobility and increasing the subthreshold swing [132]. Numerous techniques have been conceived and refined in the last decade to investigate the physico-chemical nature and the impact of such traps in realistic semiconductor devices, like admittance spectroscopy and Deep Level Transient Spectroscopy [122]. However, the available literature on MoS$_2$ interface traps is not as rich as the background of Si-based devices. Recent works addressed the interface traps issue by electrical characterization techniques and Scanning Tunneling Microscopy (STM) [162, 60, 43, 82]. Here a study of admittance spectroscopy is carried out on multi-layer and 3-layer MoS$_2$ MOS FETs to investigate the presence and the density of electron interface traps.$^1$

Differential admittance spectroscopy measurements were performed with a HP 4284A LCR-meter and consequently analyzed according to the standard procedure widely described in [132, 96, 122].

**Subthreshold swing in MoS$_2$ FET**

Figure 2.7 reports the semilogarithmic plot of the $I_{ds}$ vs $V_g$ curves of the two transistors under investigation in this Section: a bulk (named 3VTF8) and a 3-layer (denoted as 4VTF2) MoS$_2$ FET. Quite different transcharacteristics arise from the different substrate and the significantly different nature of the two transistor chan-

$^1$Article in preparation.
nals. The changed material thickness, in particular, can lead to very different transport properties, since the band-structure itself is profoundly modified in few-layer MoS$_2$ [55], while the lateral dimensions generally introduce only a vertical shift in the semilog plot of current. Here the focus is on the subthreshold region of this graph, where the inverse of the subthreshold slope, namely the Subthreshold Swing (SS), is directly related to the density of electron traps at the interface between semiconductor and gate oxide and can be extracted by the equation:

$$SS = k_BT \ln(10 \left(1 + \frac{qD_{it}}{C_{ox}}\right))$$  \hspace{1cm} (2.2)$$

$D_{it} = 2.3 \cdot 10^{12}$ eV$^{-1}$cm$^{-2}$ and $D_{it} = 3.9 \cdot 10^{12}$ eV$^{-1}$cm$^{-2}$ were calculated from the linear fit of subthreshold current in Figure 2.7 for bulk (SS = 1950 mV/dec) and 3-layer (SS = 1020 mV/dec) respectively, giving a practical estimate of the $D_{it}$ consistent for both physical systems.

**Bulk MoS$_2$ FET - Floating pad configuration**

Admittance spectroscopy measurement was performed on device 3VTF8 (bulk MoS$_2$) by contacting the substrate in back-gate configuration and one contact connected at ground, while the second contact is left floating according to the procedure in Ref [60]. Raw data of differential capacitance and conductance are reported in Figure 2.8a at different frequencies from 500Hz to 100kHz.

Focusing on the capacitance measurements, at lower gate-voltages MoS$_2$ is
Figure 2.8: Admittance spectroscopy on a bulk MoS$_2$ FET. a) Measured differential capacitance and conductance (normalized by the frequency $\omega$) at different frequencies. b) Equivalent circuit in the depletion-inversion regime. c) Equivalent circuit in accumulation. d) Map of the measured conductance with respect to the gate voltage and the frequency of the ac modulation. Conductance unit is nF/cm$^2$.

fully depleted and the total capacitance is dominated by the metal-oxide-metal junction defined by the metal pad and the back-gate in the substrate. As a matter of fact, the measured total capacitance of 2.7pF is in good agreement with the expected capacitance of a SiO$_2$ dielectric with thickness $d_{ox} = 285$nm and area equal to the pad area ($2.4 \cdot 10^4 \mu m^2$). In accumulation, conversely, the capacitance is about twice the depletion value, which is not compatible with the oxide capacitance defined by the MoS$_2$ flake area ($530 \mu m^2$). In fact, at high gate voltages an electron gas is accumulated in MoS$_2$ at the interface with the gate oxide, bulding up an electrical connection with the floating pad. As a result, the equivalent circuit in these conditions is well described in Figure 2.8c. The sketched circuit makes clear the effect of two additional capacitors, arising from the MoS$_2$ conductive channel and from the second pad, which is now actually shorted with the polarized pad and gives explanation of the huge capacitance step in Figure 2.8a.

The two dimensional map of the measured conductance in Figure 2.8d displays a conductance peak with a clear frequency dispersion depending on the back-gate voltage. This behavior can be understood in the framework of a conventional MOS capacitor as the frequency-dependent response of interface traps [132, 96, 122]. Trapping-detrapping of interface states induced by the ac modulated potential shows up as a conductance peak due to the energy loss taking place in between the charge capture and emission inside the traps.
However, the observed conductance peak is measured in a complex circuit configuration, actually experiencing a transition between the circuit models in Figures 2.8b and 2.8c. As a result, the relatively high calculated value of $D_{it} \sim 2 \cdot 10^{13}$ eV$^{-1}$cm$^{-2}$ can be mainly associated with the inherent interface traps, however the possible impact of lateral conductance inside the channel cannot completely be excluded from a quantitative analysis in the present experimental conditions.

**Bulk MoS$_2$ FET - Common pad configuration**

Figure 2.9a displays admittance measurements recorded by connecting both contact pads (source and drain of the MOS FET) at ground. In this experimental conditions the two pads are equipotential, preventing lateral current flow through the FET channel and reconfiguring the system to a more conventional MOS structure.

At first sight, the capacitance off-set measured in depletion is more than doubled compared to the measure with a floating pad, because both the metal-insulator-metal junctions created by the contact pads are now biased and can be seemingly treated. As a consequence, the small conductance step observed here is now unambiguously ascribed to the electron accumulation in the MoS$_2$ flake. The measured capacitance increase of 70fF is comparable to the expected oxide capacitance under the area of MoS$_2$ flake $C_{ox} = 64fF$.

Furthermore, the extraction of the $D_{it}$ from the admittance can now be applied to the conductance curves in Figure 2.9a, which are the physical response of interface traps to the ac modulation of the gate bias. As a result, the density of interface traps have been calculated from the conductance data according to the well-known
The energy scaling was extracted by comparing the experimental data of capacitance with the ideal curve simulated by assuming a semiconductor doping of $5 \times 10^{13} \text{cm}^{-3}$ as the best fit parameter.$^2$

The $D_{it}$ spectrum obtained from admittance spectroscopy can be compared with the value obtained from electrical transport in Section 2.3.3. The average value of $D_{it} = 2 \times 10^{12} \text{eV}^{-1} \text{cm}^{-2}$ measured from the admittance is in remarkable agreement with the one estimated from the subthreshold swing of the corresponding FET.

In fact, the latter measurement is integrated over the ensemble of interface traps throughout the whole energy band-gap of the semiconductor whereas the former one selectively probes those traps with a specific energy position in the band-gap.

3-layer MoS$_2$ FET

Figure 2.10 reports the capacitance curves taken on the 3-layer MoS$_2$ transistor referred as 4VTF2. The gate oxide thickness is 90nm, while the sample geometry is analogous to the multi-layer MoS$_2$ case discussed in the previous section. Admittance measurements have been carried out in the floating pad configuration, i.e. with only one pad grounded. This choice is motivated by the much higher resistivity of 3-layer MoS$_2$ compared to bulk MoS$_2$, as highlighted in Figure 2.7. As a consequence, the electrical contact to the floating pad is highly inefficient and capacitance/conductance contributions arising from the second pad can be neglected.

Raw data have been offset-subtracted to remove the contribution of the capacitor under the contacted pad and subsequently normalized by a $102 \mu \text{m}^2$ area. The capacitor area was calculated a posteriori from the ratio between the capacitance in accumulation (that is assumed to be equal to the oxide capacitance under MoS$_2$) and the ideal oxide capacitance per unit area, that is $3.78 \times 10^{-2} \mu \text{F/cm}^2$.

The obtained area of the capacitor is much larger than the real size of the MoS$_2$ flake incorporated in the FET channel (flake area is about $10 \mu \text{m}^2$). This incongruence is due to the multitude of MoS$_2$ flakes deposited in the surroundings of the device: unwanted contacts have been likely defined on part of these fragments, which contribute to the admittance signal but not to the transport signal, that is undoubtedly associated with the only flake contacted by both electrodes. As a further remark, the error made in estimating the capacitor area has a significative impact only on the analysis of the present device, since the 3-layer flake area is three order of magnitude smaller than the surface of the bulk flake. Consequently, the result-

$^2$Courtesy of Stefano Paleari
Figure 2.10: Differential capacitance curves (normalized by the MoS$_2$ area) on a MOS FET incorporating 3-layer MoS$_2$. Frequencies range between 1 and 100kHz.

...ing error would be of the order of 1% for device 3VTF8, thus leaving the previous analysis almost unaffected.

Figure 2.11 reports differential conductance maps recorded at different temperatures on device 4VTF2. A clear frequency-dependent peak feature is visible in each plot at different temperatures, thus making evidence of an energy dependent $D_{it}$ distribution. In particular lower temperature measurements allowed us to probe shallow levels in the energy band-gap, thus extending the admittance spectroscopy of interface traps to a wider energy range in the band-gap.

At first sight, a significant increase of the conductance signal is observed in the measurements at 350K (at low frequency) and 120K (at high frequency) with respect to the almost uniform background. The intense signal at high temperature emerges from midgap traps related to sulfur and molybdenum vacancies for example [82, 41]. On the other side, the relatively high conductance measured at 120K can be rationalized in terms of band-edge tail states induced by disorder [43, 162]. The other regions feature a $D_{it}$ of $4 \cdot 10^{12} \text{eV}^{-1}\text{cm}^{-2}$ in average, that is consistent with the value estimated from the subthreshold swing.
Figure 2.11: Differential conductance map with respect to the back-gate voltage and the signal frequency at different temperatures. The common color-scale is reported in the top left plot.
Conclusions

In the present thesis, two possible evolutions for the future generations of electronic devices have been considered.

Great challenges have to be taken into account to improve the performances of present devices or to introduce novel functionalities in the consolidated background of CMOS-based integrated circuits. Prosecuting on the pathway of bare geometrical scaling is no more practical, due to the tremendous technological constraints that will be encountered at the forthcoming end-of-roadmap. As a consequence, smart strategies must be employed to integrate novel materials or revolutionary device architectures that still have an unexplored potential of innovation. This thesis originated from such a stimulating background and it was fostered by the belief that this condition is a unique opportunity for scientific breakthroughs to happen. More in detail, this thesis considered the integration of a new device architecture for quantum computation and an emerging material (molybdenum disulfide MoS$_2$) in electronic devices.

Quantum Information Processing (QIP) is a revolutionary information processing paradigm that has the potential to outperform the present classical computers in some specific tasks. Therefore, the present thesis addressed the problems inherent to the large scale integration of QIP from a technological point of view. The evaluation of the time and space resources for the implementation of solid state QIP lead to the proposal of a CMOS-based implementation complying with the actual technological constraints of semiconductor industry.

Analogous devices have been designed and developed for this ultimate scope, always keeping in mind the requirement of compatibility with the CMOS standard. In this framework, a careful optimization of the lithographic processes was done to meet the extremely demanding requirements concerning the device size. The poor device yield, due to challenging device specifications and the related complexity of the process flow, allowed a deep characterization of few devices, which lead to significant results anyway. In particular, the electrostatic control of the charge state...
of Si quantum dots was achieved, which is the starting point for more advanced characterization and eventually the operation of spin qubits, generally based on spin-to-charge conversion.

Technological improvements to the device fabrication and the development of dedicated control circuitry will be mandatory to reach the final goal of driving coherent rotations of a qubit state. The additional requirement of device reproducibility, which was not addressed in this thesis, will strongly impact the effective feasibility of multi-qubit systems. However, keeping in mind the technological readiness of Si-based integrated circuits and the recent results in solid state QIP, the huge potential of QIP is well worth the effort to meet the target. As a further remark, QIP turned out to be also an active environment for the investigation of fundamental science, quantum metrology and quantum phenomena, like quantum coherence and entanglement. For example, in this work the electrical characterization of quantum devices offered the opportunity to study a rich phenomenology of single electron tunneling events on donors as well as quantum dots.

On the other side, MoS$_2$ and other layered materials are actually a hot topic in the scientific research landscape, as confirmed by the great number of related articles published in the last years. The second part of this thesis has been dedicated to the process integration of MoS$_2$ in a MOS FET structure, that has been validated by standard electrical characterization techniques. The present analysis was focused on the charge transport properties of few-layer MoS$_2$ FETs and the impact of interface traps on the transistor performances.

Electrical transport characterization of MOS FET showed electrical properties that, at the moment, are not competitive with other mainstream semiconductors for ultrascaled devices e.g. Si, Ge or GaAs. However, specific technological issues have been identified in this work, namely the quality of metal-semiconductor contacts and the impact of interface traps. These issues, if properly addressed (for example by means of work-function engineering and top gate passivation) could lead to a substantial improvement of the device performances. Moreover, the unique mechanical and optical properties of MoS$_2$ turned out to be of great interest for applications in optoelectronics, sensoristics and flexible electronics.

In conclusion, this thesis was not meant to give an ultimate response to the demanding requirements of nanoelectronics, with the ambition to definitively introduce a new class of electronic devices in the short window of a three year PhD course. In fact, although many technological challenges remain still unaddressed, this work of thesis is rather an analysis of these problems from the personal point of
view of the author. If some original and interesting results were reached throughout this work, inspiring further scientific discussions, this thesis will have its mission accomplished by having made a little step forward in this exciting field of science.
Appendices
Appendix A

Complete process flow for Si-DQD devices

This Appendix synthetically reports the detailed process flow developed for the manufacturing of Si-DQD devices. This Section is completed by the detailed description of some important process steps, including lithography, wet etching and Atomic Layer Deposition (ALD).

A.1 Detailed recipe

1. Preparation and cleaning of the substrate

   • RCA clean

2. UV lithography markers

   • Hard mask (thermal SiO₂, thickness = 3.5 nm) by RTP 30 s at 900°C
   • UV lithography (Level: UV Markers - Positive tone)
   • Selective etching of the SiO₂ hard mask (60 s in dilute HF 1:50)
   • Resist strip and cleaning (ACE, IPA, DIW) in ultrasonic bath
   • Pattern transfer by Si etching (TMAH etching, 3s at 80°C)
   • Hard mask removal (60 s in dilute HF 1:50)

3. Source/drain n++ doping

   • Deposition of SOG 700A by spin-coating (stopper against lateral dopant diffusion)
• SOG baking on hot plate (20 min at 200°C)
• SOG curing by RTP (180 s at 700°C in N₂ atmosphere)
• UV lithography (Level: S/D doping - Positive tone)
• Selective etching of the SiO₂ hard mask (calibrated for 7 s in dilute HF 1:10)
• Deposition of SOD P508 by spin-coating (phosphorus-doped version of SOG)
• SOD baking on hot plate (20 min at 200°C)
• Thermal diffusion of dopant by RTP (120 s at 800°C in N₂ atmosphere)
• Removal of SOG and SOD (2 min in dilute HF 1:10)

4. Ti/Pt markers for EBL

• EBL (Level: EBL markers - Positive tone): EHT = 20 kV, Base dose = 160 μC/cm², Step = 6 nm
• Removal of resist excess residues: oxygen plasma ashing (25 s at 40W)
• Ti/Pt deposition by sputtering (Ti 5 nm + Pt 50 nm)
• Lift-off in hot ACE with ultrasonic agitation

5. SOI islands definition (mesa etching)

• Hard mask (thermal SiO₂, thickness = 3.5 nm) by RTP 30 s at 900°C
• EBL (Level: SOI islands - Negative tone): EHT = 20 kV, Base dose = 45 μC/cm², Step = 8-120 nm (for high-res and large area exposures respectively)
• Removal of resist residues: O₂ plasma ashing (180 s at 100 W)
• Etching of the SiO₂ hard mask (calibrated for 50 s in dilute HF 1:50)
• Resist strip and cleaning (ACE, IPA, DIW) in ultrasonic bath
• Pattern transfer by Si etching (TMAH etching, 3s at 80°C)
• Complete removal of resist residues: O₂ plasma ashing (240 s at 300 W)

6. Gate oxide deposition

• RCA clean
• Thermal SiO₂ (calibrated for thickness = 3.5 nm) by RTP 30 s at 900°C
• $\text{Al}_2\text{O}_3$ by ALD at 300°C (Precursors: TMA and $\text{H}_2\text{O}$ - 500 cycles for thickness = 40 nm)
• Interface passivation by FGA (20 min at 450°C)

7. Gate oxide opening at source/drain contacts
• UV lithography (Level: S/D open - Positive tone)
• Selective etching of gate oxide (calibrated for 180 s in dilute HF 1:10)
• Resist strip (10 min in hot NMP at 80°C) and cleaning (ACE, IPA, DIW) in ultrasonic bath

8. Deposition of metal electrodes 1
• EBL (Level: Metal 1 - Positive tone): EHT = 20 kV, Base dose = 150 $\mu$C/cm², Step = 6-60 nm (for high-res and large area exposures respectively)
• Removal of resist excess residues: $\text{O}_2$ plasma ashing (25 s at 40W)
• Aluminium deposition (20 nm) by thermal evaporation (rate = 1 nm/s, Pressure $\sim 10^{-6}$ mbar)
• Lift-off in hot ACE

9. Deposition and opening of the second gate oxide
• $\text{Al}_2\text{O}_3$ by ALD at 300°C (Precursors: TMA and $\text{H}_2\text{O}$ - 500 cycles for thickness = 20 nm)
• UV lithography (Level: bonding pads - Positive tone)
• Selective etching of gate oxide (calibrated for 40 s in dilute HF 1:10)
• Resist strip (10 min in hot NMP at 80°C) and cleaning (ACE, IPA, DIW)

10. Deposition of metal electrodes 2
• EBL (Level: Metal 2 - Positive tone): EHT = 20 kV, Base dose = 150 $\mu$C/cm², Step = 6-60 nm (for high-res and large area exposures respectively)
• Removal of resist excess residues: $\text{O}_2$ plasma ashing (25 s at 40W)
• Aluminium deposition (30 nm) by thermal evaporation (rate = 1 nm/s, Pressure $\sim 10^{-6}$ mbar)
Lift-off in hot ACE

11. Bonding pads

- UV lithography (Level: bonding pads - Positive tone)
- Aluminium deposition (100 nm) by thermal evaporation (rate = 1 nm/s, Pressure \(\sim 10^{-6}\) mbar)
- Lift-off in hot ACE

A.2 UV lithography

UV lithography has been used for low-resolution patterning of large structures, like bonding pads and \(\mu\)m-sized area of metal-semiconductor contacts at source and drain. UV exposures were carried out with a Suss MicroTec MJB4 mask-aligner. The used photoresist (AZ5214EIR supplied from Clariant) is a h-line (405nm) and i-line (365nm) positive resist capable of image reversal (negative tone patterning). The lithographic mask is a 4” quartz mask including 12 lithographic layers needed for the fabrication of devices with different designs.

1. Sample preparation

- Solvent clean: ACE, IPA, DIW
- Dehydrating bake: 5 min on hot plate at 150°C
- Deposition of adhesion promoter HMDS by spin-coating (30 s at speed 2000 rpm, giving \(\sim 1.5\) \(\mu\)m thickness)
- Deposition of resist AZ5214EIR by spin-coating (30 s at speed 4000 rpm)

2. UV exposure and development

Positive tone

- Pre-exposure bake: 60 s at 90°C
- UV exposure: 5s at power 300 W
- Development: 40s in KOH-based developer AZ400K (AZ400K : DIW = 1 : 4)

Negative tone
• Pre-exposure bake: 60 s at 110°C
• UV exposure: 3 s at power 300 W
• Inversion bake: 60 s at 110°C
• Flood (mask-less) UV exposure: 30 s at power 300 W
• Development: 40 s in KOH-based developer AZ400K (AZ400K : DIW = 1 : 6)

3. Post-process (only for etching processes)

• Hard bake to improve contact: 4 min at 120°C

A.3 Electron Beam Lithography

SEM images and EBL exposures were done with a Zeiss Supra 40 SEM equipped with a Raith Elphy Plus lithographic attachment. Zeiss Supra 40 is a field-emission SEM equipped with a secondary-electron detector and an in-lens detector reaching $\sim 10$ nm resolution. Raith Elphy Plus is a beam-blanker and pattern generator that remotely controls the SEM column during the EBL exposures. Lateral and vertical correction routines are available from the built-in software, allowing high precision focusing and placement of the electron-beam during the exposure. ARN 7500 by AllResist and PMMA are used as negative and positive tone electro-resist respectively. The main process parameters and the detailed procedure for resist processing, exposure and development are described in the following.

Positive tone EBL

1. Sample preparation

• Solvent clean: ACE, IPA, DIW
• Dehydrating bake: 5 min on hot plate at 150°C
• Deposition of low-res PMMA (low res: 350k molec. weight, 1.5% concentration) by spin-coating (30 s at speed 3000 rpm). Bake for 10 min at 200°C. Repeat the process for a total thickness $\sim 80$ nm.\[^{1}\]

\[^{1}\]Process optimized for the deposition of structures 20-30nm-thick. For thicker layers (e.g. bonding pads with thickness $= 80-100$nm) PMMA with 6% concentration is used, producing a resist-thickness of the order of 250nm.
Deposition of high-res PMMA (high res: 996k molec. weight, 1.5% concentration) by spin-coating (30 s at speed 4000 rpm). Repeat the process for a final thickness \( \sim 110 \) nm.

2. Electron beam exposure and development

3. Pre-exposure bake: 20 min at 180°C

4. EBL exposure at accelerating voltage \( V = 20 \) kV, dose \( \sim 150 \) µC/cm²
   - High resolution: Aperture diameter = 7.5 µm, Writefield size = 100 µm, Step = 6 nm
   - Large areas: Aperture diameter = 60-120 µm, Writefield size = 1 mm, Step = 60-120 nm

5. Development: 60s in developing solution (MIBK : IPA = 1 : 3). Rinse in IPA

**Negative tone EBL**

1. Sample preparation
   - Solvent clean: ACE, IPA, DIW
   - Dehydrating bake: 5 min on hot plate at 150°C
   - Deposition of adhesion promoter HMDS by vapour deposition
   - Deposition of resist ARN7500 by spin-coating (30 s at speed 4000 rpm)

2. Electron beam exposure and development
   - Pre-exposure bake: 60 s at 85°C
   - EBL exposure at accelerating voltage \( V = 20 \) kV, dose \( \sim 45 \) µC/cm²
     - High resolution: Aperture diameter = 7.5 µm, Writefield size = 100 µm, Step = 6 nm
     - Large areas: Aperture diameter = 60-120 µm, Writefield size = 1 mm, Step = 60-120 nm
   - Development: 120s in TMAH-based developer AR300-47 (AR300-47 : DIW = 3 : 1)

3. Post-process for etching processes
   - Hard bake to improve contact: 4 min at 120°C
A.4 Wet etching and cleaning

RCA clean

RCA clean is widely used as a wafer cleaning process to remove organic and metallic contaminants, which form deep recombination centers in Si electronic devices [132, 97].

- Cleaning from organic contaminants (S1)
  \[ \text{NH}_4\text{OH} : \text{H}_2\text{O}_2 : \text{H}_2\text{O} = 1 : 1 : 5 \]
  10 minutes at 90°C

- Oxide strip (HF)
  \[ \text{HF} : \text{H}_2\text{O} = 1 : 50 \]
  30 s at room temperature. Rinse 30 s in DIW

- Cleaning from metallic ions (S2)
  \[ \text{HCl} : \text{H}_2\text{O}_2 : \text{H}_2\text{O} = 1 : 1 : 5 \]
  10 minutes at 90°C

Si etching

Basic solutions of KOH and TMAH are known to selectively etch Si at much higher rates than SiO₂. They are generally used to transfer lithographic patterns in Si through a SiO₂ hard mask. Etching rate is more than 10 times faster on the \( <100> \) than on the \( <111> \) surface, that turns out to be an effective etch-stop. Such anisotropy results in the formation of nanostructures defined by \( <111> \) facets. Etching rate at 80°C is almost the same for KOH and TMAH (~ 8 nm/s). However, TMAH is preferable to avoid the presence of K⁺ metallic ions.

SiO₂ etching

SiO₂ and Al₂O₃ etching is carried out by dilute hydrogen fluoride (HF) at room temperature. Etching rate depends on the quality of the grown oxide and is tunable with the HF concentration, ranging from 0.7 Å/s at 1:50 concentration on a thermal oxide grown at 900°C up to few nm/s at 1:10 concentration on cured SOG. Buffered Oxide Etching (BOE) with the addition of ammonium fluoride (NH₄F) has been used during the production of MoS₂ transistors to define the marker grid in the SiO₂ substrate (Section 2.3.1). Etching rate is comparable, but the higher pH results in a
better stability of the PMMA film, which conversely experienced peeling-off when using HF aqueous solutions.

A.5 Atomic Layer Deposition

Atomic Layer Deposition (ALD) is a high pressure deposition technique providing a conformal growth of a variety of materials with high control over the grown layer thickness. In fact, ALD is a process occurring layer by layer and it is self-limited by the alternating introduction of different chemical precursors which saturates the chemically active sites on the surface [97]. ALD is employed in the process flow of Si-DQD devices for QIP to deposit Al₂O₃ as the gate oxide. This process results in a conformal covering of the SOI islands with a high-κ oxide with a breakdown voltage higher than 15V for a layer 40nm-thick. The deposition chamber is a Savannah 200 ALD reactor from Cambridge Nanotech. The process parameters are reported in the following.

- Cationic (Al) precursor: Tri-Methyl-Aluminium (TMA) - Pulse duration = 10ms, Exhaust = 8s
- Anionic (O) precursor: De-Ionized Water (DIW) - Pulse duration = 15ms, Exhaust = 8s
- Carrier gas: N₂
- Temperature T = 300°C
Appendix B

Alternative layouts for Si-DQD devices

Additional devices have been developed and fabricated with alternative designs to be compared with the devices actually manufactured and characterized during this thesis (Section 1.3).

Figures B.1a-c report the lithographic mask and a SEM image of a modified version of layout B, here referred as layout C. The process integration is completely analogous to that of design B, as well as the lithographic mask, that is quite similar indeed. However, in this case the Si finger has been horizontally shifted to achieve a direct access to the electron reservoir at the source contact. This adjustment actually raises the physical distance between the SET and the DQD, possibly lowering the charge sensor sensitivity. However it could be beneficial in terms of a faster electron loading-unloading process to the DQD in the vertical Si finger.

A totally different design (layout D) is reported in Figure 1.11d-f, which has been reproduced on some test samples according to a more conventional layout employed in many previous works [35, 103, 65, 70, 85, 125, 59]. In this case, the Si device layer has not been etched. All the gate electrodes are defined on the same layer and used as depletion gates to define a DQD within a 2 Degree Electron Gas (2DEG) and a Quantum Point Contact (QPC) for charge sensing. The 2 Degree Electron Gas (2DEG) is electrostatically accumulated in a planar Si-based MOS capacitor by applying a positive voltage to a top gate defined through UV lithography and covering the whole device.
Figure B.1: Implementation of possible layouts for Si-DQD devices: layout C (panel a-c) and layout D (panel d-f). a-b) Lithographic mask (1mm$^2$ area) for layout C and zoom-in of the active area. SOI islands are colored in orange, while grey and red areas correspond to the lower and upper metal gates respectively. c) SEM image of a final device with design C. d-e) Lithographic mask (1mm$^2$ area) for layout D and zoom-in of the active area. The main lithographic levels are source/drain electrodes (violet), depletion gates (grey) and the accumulation top-gate (green). Depletion gates in panel e) are pictured according to their respective EBL dose, which are properly modulated to correct for the proximity effect. f) SEM image of a layout D device before the deposition of the accumulation top-gate. Two circles (diameter = 150nm) roughly indicate the position of two QDs.
Appendix C

List of acronyms

- 2DEG - 2-Dimensional Electron Gas
- ACE - Acetone
- ALD - Atomic Layer Deposition
- CMOS - Complementary Metal-Oxide-Semiconductor
- CVD - Chemical Vapor Deposition
- DIW - De-Ionized Water
- DQD - Double Quantum Dot
- DUV - Deep Ultra-Violet
- EBL - Electron-Beam-Lithography
- EHT - Electron High Tension (EBL accelerating voltage)
- EUV - Extreme Ultra-Violet
- FET - Field Effect Transistor
- FGA - Forming Gas Annealing
- HMDS - Hexa-Methyl-Di-Silazane
- IPA - IsoPropylc Alcohol
- ITRS - International Technology Roadmap of Semiconductors
• LSI - Large Scale Integration
• MIBK - Methyl-IsoButyl-Ketone
• MOS - Metal-Oxide-Semiconductor
• NMP - N-Methyl-Pyrrolidone
• PMMA - Poly-Methyl-MethAcrylate
• QD - Quantum Dot
• QEC - Quantum Error Correction
• QIP - Quantum Information Processing
• QPC - Quantum Point Contact
• RIE - Reactive Ion Etching
• RTN - Random Telegraph Noise
• RTO - Rapid Thermal Oxidation
• RTP - Rapid Thermal Processing
• SET - Single Electron Transistor
• SOD - Spin-On-Dopant
• SOG - Spin-On-Glass
• SOI - Silicon-On-Insulator
• TMA - Tri-Methyl-Aluminium
• TMAH - Tetra-Methyl-Ammonium-Hydroxide
• TMD - Transition Metal Dichalcogenide
Acknowledgments

After three years spent at the MDM lab (four if we add my Master thesis!), there is a long series of people that I would like to acknowledge for sharing this beautiful experience and contributing, in many different ways, to this PhD thesis.

First of all, I would like to thank prof. Marco Fanciulli, who actually made it possible, offering the opportunity to work in the active environment of MDM lab and his scientific supervision as the tutor of my research activity.

Secondly, I gratefully acknowledge prof. David Jamieson, that I had the pleasure to meet recently in Turin, for accepting to review my thesis with useful advices and comments.

Special thanks are well deserved also by Enrico Prati and Alessandro Molle for their countless and valuable advices throughout my research activity: Enrico for his enthusiastic approach to quantum physics and research in general and Alessandro for his unbreakable determination in the research on 2D materials.

I would like to thank also my colleagues of the quantum physics group: Marco De Michielis and Elena Ferraro for theoretical (and friendly!) support, Marco L. V. Tagliaferri and Alessandro Crippa for the (too!) long time spent together at the batcave (the cryomagnet lab).

Then, many thanks to all those people who, in large or small measure, took part to this work: Matteo and Simone for scientific and technical support with quantum transport measurements, Mario and Milka for the clean-room activity, Elena and Marcello (ALD growth) Flavio (electrical characterization), Stefano (AFM), Eugenio (Raman spectra) and my PhD-mates Stefano and Carlo.

Last, but not least, I would like to gratefully acknowledge my former colleagues Antonio Vellei and Giovanni Mazzeo: all I learned on clean-room nanofabrication and SEM/EBL (and it’s a big credit!) comes from them.

At the very end, I would like to thank all the person who joined in this adventure, sharing happy as well as difficult moments: all my colleagues in MDM and Bicocca, my dearest friends and my family.
Bibliography


103


110


