University of Nottingham
UK | CHINA | MALAYSIA

Design and Validation of CMOS Atom Chips for Electronics and Sensor Integration in Cold Atom Devices

Thesis submitted to the University of Nottingham for the degree of
Doctor of Philosophy, August 2024.

Joshua New
10216160

Supervised by

Prof. Mark Fromhold
Dr. Thomas Fernholz
Dr. Christian Koller

Abstract

Cold atoms have proven to be a promising medium for numerous sensing [73] and timing [51] applications. Atomic clocks have provided the definition of the second since the 1960s [11] and are essential for navigation [28] and telecommunications [25]. Increasingly precise clocks are capable of measuring small changes in the gravitational potential [29], which can be used to detect seismic activity [10], and atomic clock networks have been suggested for detecting dark matter [19]. Cold atom-based [26] gyroscopes have already demonstrated stability comparable to light-based [60] gyroscopes, with the potential to exceed them [13].

For all of these applications, a compact and mobile system is required. Atom chips use microfabricated wires to produce magnetic traps that can confine atoms close to the surface of the chip using modest current [22,24,45]. We hope to combine the mature fields of semiconductor fabrication with atom chips to produce chip-scale cold atom systems with electronics, trapping structures, antennas, sensors, and optics all integrated into a substrate. Our current work focuses on the integration of sensors and electronics through the production of CMOS [6]-based atom chips.

The first CMOS atom chip was designed at the University of Applied Sciences of Wiener Neustadt (FHWN)[1] and features a Z-wire with an adjustable geometry and photodiodes. Part of this thesis described the design and progress in implementing an experimental system built to measure the performance of this first chip. ^{87}Rb atoms are accumulated and cooled in a magneto-optical trap, then sub-Doppler cooled to 70μK and loaded into an Ioffe trap located below the CMOS chip, generated by a copper Z-wire.

In parallel to the construction of this experimental system, further iterations of CMOS atom chips are being developed. Described in this thesis is the design of the second-generation chip, referred to as chip1. This chip is designed to perform an atom chip Mach-Zehnder interferometric scheme [9] using spatial modes, but with sensors and electronics integrated onto the chip. In addition to the modifiable Z-wire, this chip features avalanche photodiodes and single-photon avalanche photodiodes designed to detect very low light intensities, Hall sensors for magnetic field measurement, a direct-digital synthesiser for generating RF fields, and a serial interface for communication with the chip.

[1] The initialisation of FHWN comes from the German name, Fachhochschule Wiener Neustadt.

Declaration of Authorship

I, Joshua New, declare that this thesis titled "Design and Validation of CMOS Atom Chips for Electronics and Sensor Integration in Cold Atom Devices" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.

Signature _____

Date _____ / _____ / _____

Acknowledgements

I would like to start by thanking all of the technical and support staff at the university who were essential throughout my PhD. Specifically, I would like to acknowledge Thomas Napier, David Laird, and Sionnach Devlin from the mechanical workshop for producing many custom parts for the experiment; and perhaps saving me from my own clumsiness on one or two occasions. Additionally, Jake Greener and Jasbinder Chauhan for their help with wire bonding and all things related to the clean room - Jake, especially for his assistance in die bonding the chips to the PCBs.

Next, I would like to acknowledge the help, guidance, support, and not least, knowledge provided by my supervisors: Mark Fromhold, Thomas Fernholz, and Christian Koller. Mark for his moral support and assistance in navigating the "quirks" of Nottingham, especially the finance department. Thomas, for his scientific omniscience and uncanny skill with lasers. Christian, for his continued patience and guidance in my forays into the world of electrical engineering, advice, and company over more than a couple of beers.

Lastly, I would like to thank everyone in the cold atoms group. From my own group: Bethany Foxon, Jamie Johnson, Hayat Abbas, Rhys Morrison, and Freya Owen. From Lucia Hackermuller's group, Nathan Cooper for his generous help to anyone in cold atoms and from Christian's group, Bianca Giacomelli. My most grateful thanks, however, go to Villius Atkocius for his abundant and continuous help, mentorship, and friendship.

Contents

List of Tables

List of Figures

Physical Constants			
Name	Symbol	Value	Units
speed of light in vacuum	c	2.997 924 58	m/s
vacuum magnetic permeability	μ_0	$1.256\,637\,062\,12 \times 10^{-6}$	$\mathrm{N/A^2}$
Planck constant	h	$6.626\,070\,15 \times 10^{-34}$	J/Hz
reduced Planck constant	$\hbar$	$1.054\,571\,817 \times 10^{-34}$	Js
elementary charge	q	$1.602\,176\,634 \times 10^{-19}$	C
Bohr magneton	μ_B	$9.274\,010\,078\,3 \times 10^{-24}$	J/T
Rydberg constant	R_∞	10 973 731.568 160	$\mathrm{m^{-1}}$

Table 1: Constants are CODATA recommended values[86].

Chapter 1

Introduction

Cold atoms and molecules have found an important place in quantum computing, sensing, and timing applications. The most mature of these being timing, where the definition of the second has been based on the ground state hyperfine transition of caesium-133 since the late 1960s [11]. The development of more accurate and stable clocks has applications such as the redefinition of SI units [70], tests of fundamental physics [73,19,8], and chronometric geodesy [87,53,85]. Moreover, atomic clocks are already essential for time synchronisation in navigation [28] and telecommunications [25]. Despite neutral atom-based quantum computers being a new field compared to their superconducting qubit and trapped ion counterparts, they have shown exciting scalability with encouraging coherence times and gate fidelities [89]. Atom interferometers are also being proposed for the detection of gravitational waves, offering sensitivities that potentially exceed those of optical-based methods [35,20]. Likewise, atom interferometers have demonstrated very stable gyroscopes [32,26], which are crucial for inertial navigation.

Mobile platforms such as boats, submarines, and satellites are required for many of these applications, necessitating small, lower power devices. To this end, miniaturisation of cold-atom technologies is essential for their implementation. One possible path to this is the atom chip [22,24,45], in which magnetic fields are created by microfabricated wires that allow for the precise control and confinement of atoms, typically trapped at distances ranging from $100\text{nm} - 100\mu\text{m}$ from the chip's surface. In existing atom chips, only the wires and antenna are integrated onto the chip. This facilitates the creation of magnetic fields necessary for atom trapping, but additional components such as electronics and optics, which are readily integrated into semiconductor chips, have yet to be included. To this end, we are exploring the feasibility of CMOS-based atom chips for the integration of sensors and electronics, where CMOS is the most prevalent semiconductor fabrication process and is used to produce integrated circuit (IC) chips such as processors and sensors such as photodiodes.

In the rest of this chapter, we discuss clocks and gyroscopes in more detail as motivation, then further discuss the concept of an integrated atom chip.

Clocks

At the time of writing, Aeppli et al.[2] reported the lowest fractional frequency uncertainty of an optical lattice clock to be 8×10^{-19}, with other atomic clocks[56,85] having demonstrated uncertainties on the order of 10^{-18}. If we have two clocks, at heights h_1 and h_2 with frequencies ν_1 and ν_2. The gravitational potential difference, and consequently the height difference, between the clocks can be measured[85] from the frequency difference, $\Delta\nu = \nu_2 - \nu_1$, by

$$\Delta h = h_2 - h_1 = \frac{c^2}{g}\frac{\Delta\nu}{\nu_1} \approx 9.2 \times 10^{17}\frac{\Delta\nu}{\nu_1}\,\text{cm}, \tag{1.1}$$

where c is the speed of light in vacuum and g is the gravitational potential at sea level. Thus, we see that the uncertainties of $\sim 10^{-18}$ are sufficient to resolve the height difference of 1 cm. In the static case, this allows us to produce precise maps using a technique known as chronometric levelling[87], and in the dynamic case this can be used to study geologic phenomena such as tectonic plate movements potentially being used to predict earthquakes[10]. The field of measuring changes in gravitational potential is known as geodesy[18,29,53].

Precise clocks can also be used for the study of fundamental physics. For example, the fine structure constant α and the ratio of proton to electron mass $\mu = m_p/m_e$ are suggested to exhibit temporal variations in some theories[73]. These are dimensionless quantities, but they will modify the atomic energy levels, resulting in a detectable signal in a sufficiently precise atomic clock. There are also proposals for the measurement of dark matter using networks of correlated atomic clocks[19].

For navigation, atomic clocks are used for time synchronisation in global navigation satellite systems (GNSS). In satellite navigation, several satellites send a signal to the user which includes a timestamp, stating when the signal was emitted. The user then compares these times to when they received the signal to calculate the distances between them and the satellites. Precise clocks are necessary for this, and thermal atom clocks are already used in satellites to produce the time stamps[28]. These are periodically synchronised with an Earth-bound master clock to prevent precision drifts over time. Although clock precision is not currently a limiting factor in navigation, more stable clocks would require less frequent synchronisation with the master clock[51].

Inertial Sensors and Gyroscopes

GNSS are used to provide services such as the global positioning system (GPS), which is vital for navigation in commercial, personal, and military settings. However, access to the GNSS is not always available due to electromagnetic interference, whether intentional or incidental, and locations such as being deep underwater. GNSS disruption is potentially damaging, and removing reliance is essential to provide stable positioning and navigation capabilities. In order to do this, inertial navigation systems

(INS) are used to keep track of acceleration and thus change in position and orientation[28]. Therefore, knowledge of the starting position is all that is required for knowledge of the final position. Importantly, INS is self-contained, meaning that no signals are required that can be disrupted.

In the case of gyroscopes, there is potential for the sensitivity of matterwave rotation sensors to far exceed that of optical rotation sensors, employing visible or near-infrared photons. The ratio of the phase shifts measured by a matterwave detector and optical detector is given by

$$\frac{\phi_{\text{atoms}}}{\phi_{\text{light}}} = \frac{mc^2}{\hbar\omega} \approx 10^{10} \tag{1.2}$$

where m is the atomic mass, for example, $m \approx 10^{-25}$ for ^{87}Rb. c is the speed of light in vacuum, $\hbar$ is the reduced Planck constant and ω is the angular frequency of the light used in the interferometer. It should be noted that this assumes an equal rotation rate Ω, interferometric area A and count rate. In practice, optical gyroscopes can have interferometric areas on the order $A_{\text{light}} \approx 10^3\text{m}^2$ compared to $A_{\text{atoms}} \approx 10^{-4}\text{m}^2$ for typical atomic gyroscopes. Optical gyroscopes also have a photon flux on the order of 10^{19} photons/s for 1 W of power whilst atom interferometers operate below 10^7 atoms/s.

For navigation purposes, the most important characteristic is the stability of the gyroscope, given in units of rad/s or degrees per hour. Commercial fibre optic gyroscopes[60] have demonstrated stabilities of 2.28×10^{-11} rad/s. Comparably, the highest cold atom stability was reported by Geiger et al[26]. to be 3×10^{-10} rad/s.

Atom Chips

In the majority of the aforementioned applications, including GNSS timekeeping, INS, and geodesy, devices must be implemented on compact and portable platforms such as aeroplanes, satellites, and submarines. In addition, to aid in miniaturisation, we would like to replace free-fall with fully confined matterwave schemes. A potential solution exists in the form of atom chips, where atoms are confined by magnetic field gradients generated by current-carrying wires on the chip. With fully confined and guided waveguides, integrated antenna and trapping wires, and without the additional high-power lasers required for optical trapping, they are a promising option for mobile cold atom technologies.

The atom chip has several possible advantages over its optical and free-fall based counterparts. Since the atoms are trapped close to the wires, typically $< 100\,\mu$m, modest currents (<1 A) are capable of producing steep potentials resulting in tight confinement with low power. This in turn allows for efficient rethermalisation during evaporative cooling, resulting in high flux Bose-Einstein-Condensate (BEC) generation with much lower power requirements[72]. Although it should be noted that BEC is not always desired due to the strong atom-atom interactions, in the majority of cases BEC is necessary to maximise sensitivity.

Atom chips were first realised near the turn of the millennium[65,21,17] with the first BECs demonstrated soon after[37,44] and interferometry in the mid-2000s[36,76]. Even at this early stage, many tools were developed to give fine control over atom chip potentials including radio-frequency (RF) dressing for adiabatic potentials[61,76,36] and microwave (MW) dressing for state-dependent potentials[12]. In parallel with the lower current requirements for the static magnetic field, being able to integrate RF and MW antennas onto the chip greatly reduces RF and MW power requirements. Several other cold-atom schemes have been demonstrated on atom chips since. For example, a microwave clock[48,84] with a reported fractional frequency stability of 5.8×10^{-13} at 1 s. Spin-squeezing has also been implemented using BEC, which can improve sensitivities beyond the standard quantum limit (SQL)[68].

Integrated Circuit Atom Chips

Whilst atom chips alone are great for compactness in the near term, the long-term goal is for fully chip-scale cold atom sensors and clocks which would enable commonplace use wherever precision timing and sensing is required. This would also open up the possibilities for space-based gravimetry, magnetometry and navigation. Current atom chips have several disadvantages for reaching these limits. Most importantly, fabrication is often prohibitively expensive and requires specialist knowledge and facilities, all without the possibility to integrate beyond the trapping structures. To overcome these issues, commercial IC fabrication offers the production of chips with mixed functionality such as analogue and digital circuits, photodiodes, micro-electro-mechanical systems (MEMS) and more. Specifically, CMOS (complementary-metal-oxide-semiconductor) technology is the fabrication standard for global IC production, with over 90% of the world's chips made in some kind of CMOS technology. CMOS chips are built onto a semiconductor (typically silicon) substrate, with doped regions forming PN junctions which allow the construction of devices such as transistors and photodiodes. Metal layers are then added above the substrate, separated by insulating oxide layers, to form the necessary connections.

CMOS opens the possibility of integrating elements such as current sources, RF and MW sources and antennas, photodiodes, control circuits, Hall sensors, temperature sensors and more directly onto the chip[6], greatly enhancing the compactness of cold-atom devices. This is summarised in Figure 1.1. Figure 1.1a shows a traditional atom chip and lists some of the external systems required, such as current sources for the trapping wires, RF and MW source with amplifiers and cameras for fluorescence or absorption measurements. Figure 1.1b shows a simplified CMOS atom chip with all of these components included plus additional components such as Hall sensors which are not possible with current chips. The CMOS chip can act as a very simple microprocessor with the onboard components programmable via a serial interface, feedback to optimise trapping, sensors with analogue-to-digital converters (ADCs) and buffers to save the data which can be read through the serial interface.

To summarise, a CMOS based atom chip has the potential to make atom chip production a relatively

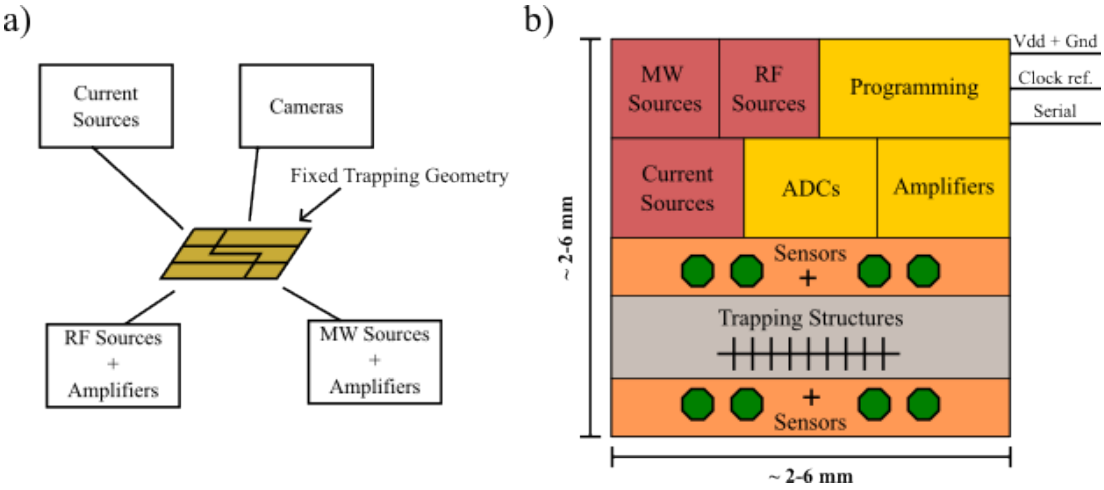


Figure 1.1: a) Shows a conventional atom chip where electronics, sensors, and current sources are all external. b) In the integrated chip concept, many of the previously external elements have been incorporated onto a single chip.

inexpensive commercial process where the resulting chips already contain many of the components required to run a cold atom experiment, simplifying the entire process. Moreover, these chips are very general and can be programmed to perform specific functions so that one device can be used as a clock, magnetometer or gyroscope. Although not included in the scope of this work, CMOS also opens up the possibility of integrated photonics, leading to the next stage of integration.

CMOS in Existing Quantum Technologies

Fortunately, the utility of CMOS has already been noticed and applied in the quantum technology world; meaning many of the principles and circuits necessary for a CMOS-based atom chip have already been proven.[4] Nuclear magnetic resonance (NMR) provides a tool for determining molecular structure, and whilst large NMR devices requiring cryogenically cooled superconductors are available, smaller and more portable NMR machines are desirable for broadening access to this technology. To this end, efforts have been made to miniaturise these devices by replacing the superconducting magnet with smaller, permanent magnets, and using CMOS ICs for the required radio-frequency (RF) magnetic field generation and readout circuits.[47,90] NMR involves placing the sample in a strong magnetic field B_0, and interrogating it with a resonant RF magnetic field (perpendicular to B_0) with frequency $f_0 = \gamma B_0/(2\pi)$, where γ is the gyromagnetic ratio of the nuclear spin of interest; this induces a precession of the nuclear spin about B_0 at frequency f_0, which can be measured using an RF receiver. Typical NMR uses the spin of the proton ^{1}H, which has $f_0 = 21.8$ MHz for a standard magnetic field strength of $B_0 = 0.51$ T; however, to expand the fingerprinting capabilities of an NMR system, other nuclear spins such as ^{19}F ($f_0 = 20.5$ MHz) and ^{2}H ($f_0 = 3.3$ MHz) can be used. Consequently, a wideband transceiver is desired and existing portable NMR devices have been constructed with RF transceivers functional at 1-100 MHz, integrated into CMOS chips with amplifiers and pulse sequencing.[47,90]

Another field in which CMOS has been applied is the integration of microwave generation/delivery and fluorescence-based spin detection in diamond nitrogen-vacancy (NV) center-based technologies. In these devices, a nitrogen atom is embedded in a diamond molecule adjacent to a vacancy, leaving two unpaired electrons which behave as a quantum system with multiple energy levels. The ground state and first excited state are triplet states, with an $|m_s = 0\rangle$ sub-level and two degenerate $|m_s = \pm 1\rangle$ states; in the presence of a magnetic field, this degeneracy is lifted according to the Zeeman effect. Within the ground state, with no magnetic field present, the sub-levels are separated by 2.87 GHz. The electrons can be excited from the ground to the excited states with green (532 nm) light, but decay to the ground state either by the emission of red light (625 nm); however, from the excited $|m'_s = \pm 1\rangle$ sub-levels there is a second, non-radiative decay path via a meta-stable singlet state. Thus, there is a reduction in the detected red light if the atoms start the excitation-decay cycle in either of the $|m_s = \pm 1\rangle$ sub-levels, due to the atoms decaying via the non-radiative pathway, which does not occur with decay from the excited $|m'_s = 0\rangle$ state. Scanning a magnetic field across the resonant 2.87 GHz frequency produces a spectrum, with a single dip in the detected red light intensity if there is no magnetic field present, and two dips if there is a static field splitting the $|m_s = \pm 1\rangle$ degeneracy. This allows NV centers to be used as magnetometers by measuring the distance between these dips in a process known as optically detected magnetic resonance (ODMR).[38] Alternatively, by assigning the $|m_s = 0\rangle$ and $|m_s = -1\rangle$ states as logical $|0\rangle$ and $|1\rangle$ states, the NV center ensemble can be used as a qubit.[58] For all of these applications, integration of the MW source and red light photodetector is desirable for both portability in the case of the magnetometer and scalability in the case of the qubit. Omirzakhov et. al.[58] reported a CMOS chip with an integrated LC-based voltage controlled oscillator (VCO) with a frequency range of 1.6-2.6 GHz, including an amplifier with a maximum gain of 12 dB, and a single-photon avalanche diode (SPAD) capable of measuring Rabi oscillations of the qubit.

Lastly, and the most relevant field for this work, is the incorporation of CMOS into atomic ion traps for quantum computing, where integrating elements of the system is desirable for scalability. In fact, the trapping of cold ions using commercial foundry CMOS chips has already been demonstrated in a 90 nm process technology provided by IBM.[54] However, atomic ion traps differ somewhat from neutral atom traps; namely, ions are trapped using electric, as opposed to magnetic, fields, and these devices are cryogenically cooled to around 4 K, adding a constraint to the CMOS designs. Much of the CMOS design in this field has focused on the development of DACs to generate the surface-electrode potentials used in the ion trapping.[83,59] Whilst not specifically fabricated using a CMOS foundry, fluorescence detection of trapped ions by integrated SPADs has also been demonstrated.[64]

1.1 Thesis Description

The work presented in this thesis can be divided into two separate tasks: the construction of an experiment to validate CMOS atom chips, specifically the existing chip called chip0, and the development of new iterations of chips. A complete chip, called chip1, was designed in collaboration with bachelor's and master's students at FHWN, and was fabricated by Austrian Microsystems (AMS): whilst this was primarily built as a test ground for several components while chip0 was validated, chip1 was designed to function as a double-well interferometer based on that of Berrada et al.[9]. Further chip iterations have also begun the design stage and will be discussed in Chapter 7. Descriptions of each thesis chapter are given:

1. Chapter 2: The theory of the trapping and cooling of neutral alkali atoms is discussed with a focus on ^{87}Rb as this is the species used in this work.

2. Chapter 3: This chapter serves as an introduction the CMOS technology, firstly the theory behind PN-junctions and MOSFETs and their simulation followed by the basics of digital design and a discussions of specific components used in our chip designs. Analogue design is introduced and conversion between analogue and digital signals is discussed before specific components are explained.

3. Chapter 4: Chapter 4 details the experimental setup, starting from CMOS chip0 and building the necessary infrastructure to support the full sequence of cooling, trapping and transporting the atoms up to loading the chip trap.

4. Chapter 5: The experimental sequence so far is introduced with results of the individual stages. We begin with loading of the MOT followed by laser cooling, magnetic trapping, transport and loading into a copper Z-wire Ioffe trap.

5. Chapter 6: Here the design of chip1 is detailed with simulations beginning with the trapping wires then the clock generation, serial interface, RF generation, ADC and lastly the sensors.

6. Chapter 7: A summary of the work conducted in this thesis is presented with suggestions on progress for the experiment. Then an outlook for CMOS atom chips is outlined before a discussion on further integration towards chip-scale cold atom devices.

Chapter 2

Cold Atoms Theory

In this chapter we explain the theory necessary to understand a standard ultra-cold atom experiment using ^{87}Rb.

2.1 The Atomic Structure of ^{87}Rb

For many neutral cold atom experiments, group 1A alkali metals are chosen as the subject atomic species due to a single valence electron in the outer shell which allows these atoms to be treated analogously to the hydrogen atom. Rubidium, for example, has an orbital structure of $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 4s^2 4p^6 5s$ using the notation nL^{n_e}, where n is the principal quantum number, n_e is the number of electrons in that orbital and L is the quantum number of the orbital angular momentum $\hat{\mathbf{L}}$; with the magnitude of $\hat{\mathbf{L}}$ is given by $\sqrt{L(L+1)}\hbar$. The letters are used to represent L with $s \equiv L = 0$, $p \equiv L = 1$, and $d \equiv L = 2$.

Coarse Structure

Ignoring relativistic effects and spin-coupling, the energy levels are given by the modified Bohr model

$$E_n = -hc\frac{R_\infty}{(n-\delta_s)^2}, \tag{2.1}$$

where δ_s is known as the quantum defect and accounts for the portion of the wavefunction present within the closed electron shells which are not present in a Hydrogen atom. For ^{87}Rb, $\delta_s = 3.19$[23]. R_∞ is the Rydberg constant, h is the Planck constant, and c is the speed of light. The energy levels given by this formula are known as the coarse structure of an atom and depend only on n.

Eigenstates of the orbital angular momentum are given by $|n, L, m_L\rangle$ and the eigenvalues of the

quantum numbers L and m_L are

$$\hat{\mathbf{L}}^2|n,L,m_L\rangle = \hbar^2 L(L+1)|n,L,m_L\rangle \tag{2.2}$$

$$\hat{L_z}|n,L,m_L\rangle = \hbar m_L|n,L,m_L\rangle. \tag{2.3}$$

Fine Structure

Electrons also possess an intrinsic property called spin described by the operator $\hat{\mathbf{S}}$ that can take the values $S \pm 1/2$. The resulting interaction between the orbital and spin angular momenta leads to an additional term in the Hamiltonian of the form $H_{so} \propto \hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$. We can define the total electron angular momentum as

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}, \tag{2.4}$$

where the quantum number J must lie in the range $|L-S| \leq J \leq L+S$. We have to use the total angular momentum as S and L are no longer good quantum numbers, since whilst their total magnitudes are conserved, their projections onto the z axis, m_S and m_L, are not. The definition of $\hat{\mathbf{J}}$ can be rearranged as

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \frac{1}{2}\left(\hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2\right) \tag{2.5}$$

from which the known expectation values for $\langle\hat{\mathbf{J}}^2\rangle$, $\langle\hat{\mathbf{L}}^2\rangle$ and $\langle\hat{\mathbf{S}}^2\rangle$ can be used to define the energy shift. The energy shift due to H_{so} is proportional to

$$\Delta E_{so} \propto [J(J+1) - L(L+1) - S(S+1)]. \tag{2.6}$$

Hyperfine Structure

At even smaller energy scales, the interaction between $\hat{\mathbf{J}}$ and the nuclear angular momentum $\hat{\mathbf{I}}$ must also be considered. For ^{87}Rb, the nuclear spin is $I = 3/2$. As before, we can define the total atomic angular momentum as $\hat{\mathbf{F}} = \hat{\mathbf{J}} + \hat{\mathbf{I}}$ with new quantum numbers F, where $|J-I| \leq F \leq J+I$, and m_F, which is the projection of $\hat{F}$ along z. In this case, the resulting Hamiltonian term is

$$H_{hfs} = A_{hfs}\hat{\mathbf{I}} \cdot \hat{\mathbf{J}} + B_{hfs}\frac{3(\hat{\mathbf{I}} \cdot \hat{\mathbf{J}})^2 + \frac{3}{2}(\hat{\mathbf{I}} \cdot \hat{\mathbf{J}}) - I(I+1)J(J+1)}{2I(2I-1)J(2J-1)} + C_{hfs}\mathscr{O}\left[(\hat{\mathbf{I}} \cdot \hat{\mathbf{J}})^3\right] \tag{2.7}$$

where A_{hfs} is the magnetic dipole constant, B_{hfs} is the electric quadrupole constant, C_{hfs} is the magnetic octupole constant which can be considered zero due to a lack of sufficiently precise measurements to distinguish it from zero[81], and $\mathscr{O}\left[(\hat{\mathbf{I}} \cdot \hat{\mathbf{J}})^3\right]$ represents the higher order term. From here the energy shift

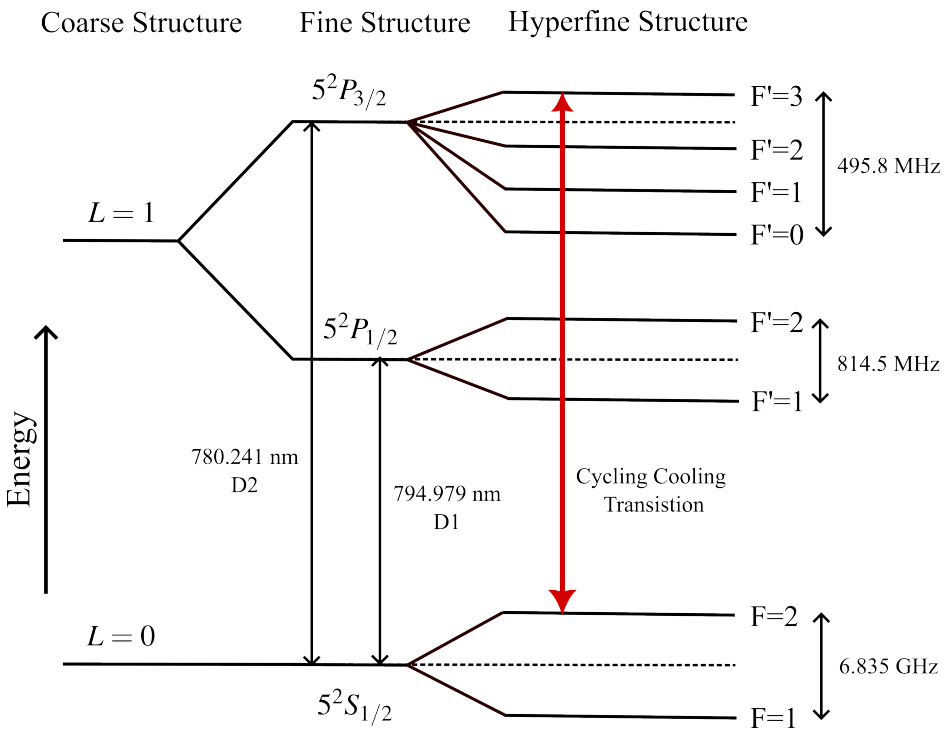


Figure 2.1: The energy level diagram for ground state and first excited state of Rb^{87} which arise from the course structure and the orbital angular momentum L. The excited $L=1$ state is then split to form the fine structure levels $J=1-1/2=1/2$ and $J=1+1/2=3/2$ where the transitions between the $S_{1/2}$ and $P_{1/2}$ is known as the D_1-line and the $S_{1/2}$ and $P_{3/2}$ is known as the D_2-line. Finally the hyperfine structure on the right arises from the coupling of the electron angular momentum to the nuclear angular momentum and each state is labelled by $|J-1| \leq F \leq J+1$. The cooling transition is indicated by the red arrow.

due to this term can be calculated to be

$$\Delta E_{hfs} = \frac{1}{2}K + B_{hfs}\frac{\frac{3}{2}K(K+1) - 2I(I+1)J(J+1)}{4I(2I-1)J(2J-1)}, \tag{2.8}$$

where

$$K = F(F+1) - I(I+1) - J(J+1). \tag{2.9}$$

This equation simplifies in the $5^2S_{1/2}$ state to be

$$\begin{aligned} \Delta E_{hfs} &= A_{hfs}\hat{\mathbf{I}}\cdot\hat{\mathbf{I}} \\ &= \frac{A_{hfs}}{2}[F(F+1) - I(I+1) - J(J+1)] \end{aligned} \tag{2.10}$$

since $B_{hfs} = 0$.

Selection Rules

We can use laser light or microwave radiation to couple different ground and excited states although not every transition is allowed due to conservation of angular momentum and parity. The coarse structure

allows transitions with $\Delta L = \pm 1$ and $\Delta m_L = 0$ for π-polarised light and $\Delta m_L = \pm 1$ for σ-polarised light. For the total electron angular momentum, the rules require $\Delta J = 0, \pm 1$ with the exception that $J = 0 \rightarrow J' = 0$ is not allowed, as this would imply $\Delta L = 0$, $\Delta m_J = 0$ for π-polarised light and $\Delta m_J = \pm 1$ for σ-polarised light. The rules are identical for the total atomic angular momentum $\hat{\mathbf{F}}$.

2.2 Interactions with Static Magnetic Fields

The magnetic moment, $\boldsymbol{\mu}$, of an atom interacts with an external magnetic field, leading to shifts in the energy levels. The Hamiltonian term describing the interaction between angular momenta and the magnetic field is given by

$$\begin{aligned} \hat{H}_{int} &= -\boldsymbol{\mu} \cdot \mathbf{B} \\ &= \frac{\mu_B}{\hbar}(g_S \hat{\mathbf{S}} + g_L \hat{\mathbf{L}} + g_I \hat{\mathbf{I}}) \cdot \mathbf{B} \\ &= \frac{\mu_B}{\hbar}(g_S S_z + g_L L_z + g_I I_z) B \end{aligned} \tag{2.11}$$

when the z-axis is defined along the direction of the magnetic field. g_S, g_L, and g_I represent the g-factors for electron spin, electron orbital momentum, and nuclear states, respectively. These g-factors are coefficients tied to their corresponding magnetic dipole moments. g_I is typically neglected because it is much smaller than the other g-factors[81].

Zeeman Effect (Weak Fields)

If the energy shift due to the magnetic field is small in comparison to the fine structure splitting, then $\hat{\mathbf{J}}$ is the relevant angular momentum and the interaction term can be re-written as

$$\hat{H}_{int}^{fs} = \frac{\mu_B}{\hbar}(g_L L_z + g_s Sz) B. \tag{2.12}$$

It can be treated as a perturbation of the fine structure Hamiltonian. The energy shift associated with this regime can be calculated using perturbation theory to first order by

$$\begin{aligned} \Delta E_B^{fs} &= \langle J, m_J | \hat{H}_{int}^{fs} | J, m_J \rangle \\ &= \mu_B g_J m_J B \quad \text{(Fine Structure Zeeman Effect)} \end{aligned} \tag{2.13}$$

where g_J is found to be

$$\begin{aligned} g_J &= g_L \frac{J(J+1) - S(S+1) + L(L+1))}{2J(J+1)} + g_S \frac{J(J+1) + S(S+1) - L(L+1))}{2J(J+1)} \\ &\approx 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \end{aligned} \tag{2.14}$$

where the approximation comes from taking the approximate values $g_S \approx 2$ and $g_L \approx 1$.

In the presence of weak magnetic fields, the Zeeman shift is significantly weaker than the hyperfine splitting. Therefore, the total angular momentum F is a good quantum number, and the interaction term becomes

$$\hat{H}_{int}^{hfs} = \frac{\mu_B}{\hbar}(g_J J_Z + g_I I_Z)B \tag{2.15}$$

and the corresponding energy shift is

$$\begin{aligned} \Delta E_B^{hfs} &= \langle F, m_F | \hat{H}_{int}^{hfs} | F, m_F \rangle \\ &= \mu_B g_F m_F B \quad \text{(Hyperfine Structure Zeeman Effect} \end{aligned} \tag{2.16}$$

with a hyperfine Lande g-factor of

$$\begin{aligned} g_F &= g_J \frac{F(F+1) - I(I+1) + J(J+1)}{2F(F+1)} + g_I \frac{F(F+1) + I(I+1) - J(J+1)}{2F(F+1)} \\ &\approx g_J \frac{F(F+1) - I(I+1) + J(J+1)}{2F(F+1)} \end{aligned} \tag{2.17}$$

where we have neglected g_I as stated above.

In most Rubidium experiments, the fine splitting Zeeman effect can be ignored as the magnetically trapped states lie in the L=0 manifold, so only the hyperfine shifts are relevant.

Paschen-Back Effect (Strong Fields)

If the field is weak compared to the fine structure splitting, but strong enough to dominate the hyperfine splitting, then the hyperfine splitting is treated as a perturbation of the interaction term and F is no longer a good quantum number. Instead the total electron angular momentum, $\hat{\mathbf{J}}$, precesses around the static magnetic field,**B**, and the eigenenergies are calculated as

$$\Delta E_{PB} = \langle J, m_J, I, m_I | \hat{H}_{hfs} + \hat{H}_{int}^{hfs} | J, m_J, I, m_I \rangle \tag{2.18}$$

to first order. Thus, the energy shifts are given by

$$\begin{aligned} \Delta E_{PB} \approx A_{hfs} m_I m_J + B_{hfs} &\frac{9(m_I m_J)^2 - 3J(J+1)m_I^2 - 3I(I+1)m_J^2 + I(I+1)J(J+1)}{4J(2J-1)I(2I-1)} \\ &+ \mu_B(g_J m_J + g_I m_I)B_z \end{aligned} \tag{2.19}$$

which can be further simplified in the $5^2S_{1/2}$ manifold to

$$\Delta E_{PB} \approx A_{hfs} m_I m_J + \mu_B(g_J m_J + g_I m_I)B. \tag{2.20}$$

This is known as the hyperfine Paschen-Back effect.

If the field is sufficiently strong to dominate the fine structure, the field couples to $\hat{\mathbf{S}}$ and $\hat{\mathbf{L}}$ individually and J is no longer a good quantum number. In this case the energy shift is given simply by

$$\begin{aligned}\Delta E_B^{fpb} &= \langle L, m_L, S, m_S | \frac{\mu_B}{\hbar}(g_S S_z + g_L L_z) | L, m_L, S, m_S \rangle \\ &= \mu_B (g_S m_S + g_L m_L) B.\end{aligned}$$

For atom chip experiments the field strengths are typically tens of Gauss and are far below this regime which becomes relevant at several hundred Gauss.

Breit-Rabi Formula (Intermediate Fields)

To calculate the intermediate fields, the Hamiltonian $\hat{H}_{hfs} + \hat{H}_{int}$ must be diagonalised which is typically done numerically. However, for the ground state of ^{87}Rb, where $L = 0$, $J = 1/2$ there is an analytical solution known as the Breit-Rabi formula which is given generally as

$$E_{|m_J, I, M_I} = \Delta E_{hfs} \left(-\frac{1}{2(2I+1)} + g_I \mu_B m B \pm \frac{1}{2} \sqrt{1 + \frac{4mx}{2I+1} + x^2} \right) \tag{2.21}$$

where

$$x = \frac{(g_J - g_I)\mu_B B}{\Delta E_{hfs}} \tag{2.22}$$

and $m = m_I \pm m_J = m_I \pm 1/2$ with the $\pm$ sign being the same as above. However, this can be simplified by realising that since this is the $5^2S_{1/2}$ state, the hyperfine splitting is simply given by $\hat{H}_{hfs} = A_{hfs} \hat{\mathbf{I}} \cdot \hat{\mathbf{J}}$ which implies that $\Delta E_{hfs} = 2A_{hfs}$, moreover, $g_I \approx 0$, $g_J \approx 2$ and $I = 3/2$ for ^{87}Rb yielding the final simplified equations

$$E_{|J, m_J, I, m_I\rangle} = \frac{3}{4} A_{hfs} \pm \mu_B B \qquad \text{for } m = \pm 2 \tag{2.23}$$

$$E_{|J, m_J, I, m_I\rangle} = A_{hfs} \left(-\frac{1}{4} \pm \sqrt{1 + mx + x^2} \right) \qquad \text{for } m = 0, \pm 1 \tag{2.24}$$

with

$$x \approx \frac{\mu_B B}{A_{hfs}}. \tag{2.25}$$

The + is used in the $F = 2$ manifold where $m_J = +1/2$ and the - is used for the $F = 1$ manifold where $m_J = -1/2$ and the Breit-Rabi formula is plotted in Figure 2.2. Initially we see that the splitting is linear when the total angular momentum couples to the static field but eventually splits into to group of four when the electron spin is coupled to the field. These groups simply correspond to $S = 1/2$ and $S = -1/2$.

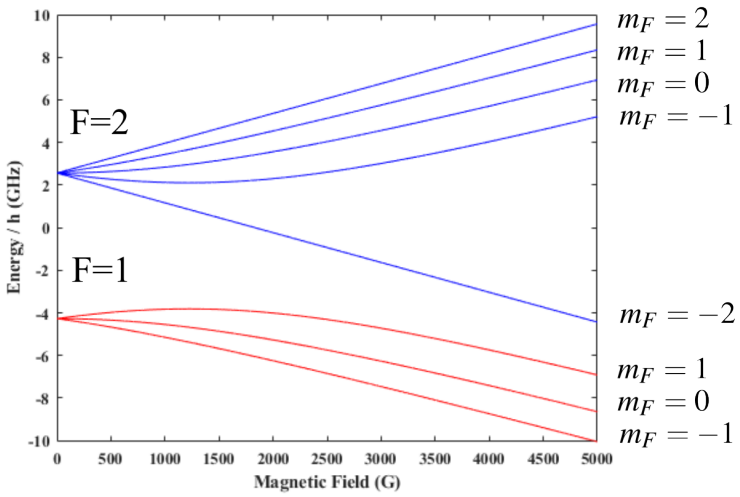


Figure 2.2: The energy splitting of hyperfine manifolds $F = 1$ (red) and $F = 2$ (blue) of the Rb^{87} $5^2S_{1/2}$ ground state in the presence of a static magnetic field is shown. The splitting is calculated using the Breit-Rabi formula.

2.3 Laser Cooling

The initial stages of cold-atom experiments involve the accumulation and cooling of atoms released from either a dispenser or an oven. This experiment makes use of two cooling stages before the atoms are magnetically trapped. Firstly, a 3D magneto-optical trap (MOT) followed by a compressed MOT (CMOT) stage, which involves Doppler cooling in conjunction with polarisation gradient cooling to increase the phase space density of the atoms. The atoms are then optically pumped in the stretched weak-field seeking $|F = 2, m_F = 2\rangle$ state before they are captured in a magnetic trap. Most of this section follows the textbook by Foot[23].

To derive the forces involved in laser cooling, it is essential to know the scattering rate for photons which dictates how much momentum can be imparted to the atoms over time. To find this, we can first note that the energy levels used are sufficiently far apart such that the ground and excited state can be treated as a two-level system. In reality, for ^{87}Rb decay is possible into two ground states, but a repumper beam excites atoms from the unwanted ground state to restore the two-level approximation. Writing the atomic states in the density operator formalism where $\rho = |\psi\rangle\langle\psi|$ and

$$\rho = \begin{pmatrix} \rho_{gg} & \rho_{ge} \\ \rho_{eg} & \rho_{ee} \end{pmatrix} \tag{2.26}$$

with ρ_{ee} and ρ_{gg} giving populations in the excited and ground states respectively and the off-diagonal elements are called coherences, which represent the effect of the driving field. Then exposing the atoms to a monochromatic light field of frequency ω, detuned from the transition frequency ω_0 by $\delta = \omega - \omega_0$, will lead to evolution of the state populations and coherences that are dictated by the optical Bloch

equations[81]

$$\partial_t \rho_{gg} = \frac{i\Omega}{2}(\tilde{\rho}_{ge} - \tilde{\rho}_{eg}) + \Gamma\rho_{ee}, \tag{2.27}$$

$$\partial_t \rho_{ee} = -\frac{i\Omega}{2}(\tilde{\rho}_{ge} - \tilde{\rho}_{eg}) - \Gamma\rho_{ee}, \text{ and} \tag{2.28}$$

$$\partial_t \tilde{\rho}_{ge} = -\left(\frac{\Gamma}{2} + i\delta\right)\tilde{\rho}_{ge} - \frac{i\Omega}{2}(\rho_{ee} - \rho_{gg}), \tag{2.29}$$

where $\Omega = -\mathbf{d}\cdot\mathbf{E_0}/\hbar$ is the resonant Rabi frequency, $\mathbf{d}$ is the dipole operator, $\mathbf{E_0}$ is the electric field amplitude and polarisation, $\Gamma = 1/\tau$ is the natural decay rate of the excited state, $\tilde{\rho}_{ge} = \rho_{ge}e^{-i\delta t}$, and $\tilde{\rho}_{ge} = \tilde{\rho}^*_{eg}$. The optical Bloch equations treat the system as a quantum harmonic oscillator, which is driven by the electric field and damped by spontaneous emission, which is added through terms including Γ.

Taking the steady state solution of these equations, we find the population in the excited state to be

$$\rho_{ee}(t\to\infty) = \frac{(\omega/\Gamma)^2}{1 + 4(\delta/\Gamma)^2 + 2(\Omega/\Gamma)^2}. \tag{2.30}$$

The rate of emission in this steady state is given by $\Gamma\rho_{ee}$ and must equal the rate of absorption which in turn is the scattering rate of photons, thus we find the final scattering rate to be

$$R_{sc} = \frac{\Gamma}{2}\frac{I/I_{\text{sat}}}{1 + I/I_{\text{sat}} + 4\delta^2/\Gamma^2} \tag{2.31}$$

where we have introduced the saturation intensity defined by $I/I_{\text{sat}} = 2(\Omega/\Gamma)^2$, which gives $I_{\text{sat}} = \pi hc\Gamma/3\lambda^3$ when on resonance. For ^{87}Rb, $\Gamma = 2\pi\times 6.07\,\text{MHz}$ and $I_{\text{sat}} = 1.67\,\text{mW/cm}^2$ for the $|F = 2, m_F = \pm 2\rangle \to |F' = 3, m'_F = \pm 3\rangle$ cycling transition used in cooling[80].

2.3.1 Optical Molasses

In the case of optical molasses, the mechanism for laser cooling is based on the scattering of light. Each photon carries a momentum of $\hbar k$, where $\hbar$ is the reduced Planck's constant and k is the wavevector given by $2\pi/\lambda$, which during spontaneous emission is emitted in a random direction, and so the momentum kicks from spontaneous emission average to zero. If we consider a source of resonant laser light from a single direction, then all of the momentum kicks from absorption are in the same direction, and the atom's net momentum change is in the direction of the laser. The force due to this net momentum change is given by $F_{sc} = \hbar k R_{sc}$ which is just the momentum of each photon multiplied by the rate of absorption. If the laser is now red-detuned from the transition by $\delta = \omega - \omega_0$, then the rate of absorption is lower for stationary atoms, but for atoms moving towards the beam at velocity v, the frequency is shifted by kv, which modifies the detuning to $\delta = \omega - \omega_0 + kv$, causing the force to become velocity-dependent such

that atoms moving towards the beam are excited more often, implying that the beam now acts to reduce the atom's velocity in one direction. If three pairs of opposing beams are placed along orthogonal axes, then this can cool the atoms by reducing their velocity in all directions.

The total scattering force along one direction is given by replacing δ with $\delta \pm kv$ in R_{sc} to find

$$\begin{aligned} F_{\text{total}} &= F_{sc}(\delta - kv) - F_{sc}(\delta + kv) \\ &\approx \left[F_{sc}(\delta) - kv\frac{\partial F_{sc}}{\partial \omega}\right] - \left[F_{sc}(\delta) + kv\frac{\partial F_{sc}}{\partial \omega}\right] \\ &= -2\frac{\partial F_{sc}}{\partial \omega}kv \\ &= -\alpha v, \end{aligned} \tag{2.32}$$

where

$$\alpha = 2k\frac{\partial F_{sc}}{\partial \omega} = 4\hbar k^2 \frac{I}{I_{\text{sat}}}\frac{-2\delta\Gamma}{[1+(2\delta/\Gamma)^2]^2}. \tag{2.33}$$

For damping to occur, α must be positive which implies a red-detuning of $\delta = \omega - \omega_0$. The steady state temperature, T, for optical molasses is given by

$$k_B T = \frac{\hbar\Gamma}{4}\frac{1+(2\delta/\Gamma)^2}{-2\delta/\Gamma}, \tag{2.34}$$

and has a minimum at $\delta = -\Gamma/2$ known as the *Doppler cooling limit*, T_D. This is given by

$$k_B T_D = \frac{\hbar\Gamma}{2} \tag{2.35}$$

which for ^{87}Rb is $T_D = 145.57\,\mu\text{K}$[80].

2.3.2 Magneto-Optical Trap (MOT)

Whilst the optical molasses provides a velocity-dependent force which cools the atoms, the atoms will still disperse over time. This can be eliminated by introducing a magnetic trap to lift the degeneracy of the hyperfine manifolds and introduce a position-dependent term to the scattering force. This force enables atoms to be trapped and cooled simultaneously. This configuration of laser cooling is referred to as a magneto-optical trap (MOT) and was first proposed by Dalibard and demonstrated using sodium atoms by Raab et al[62]. The standard six beam MOT is shown in Figure 2.3a with the magnetic field generated by two coils in anti-Helmholtz configuration and six opposing, circularly polarised laser beams. The polarisations are labelled in the convention that σ^+ excites the transition $m_F = 0 \to m'_F = 1$ and σ^- excites $m_F = 0 \to m'_F = -1$ using the simplified energy levels shown in Figure 2.3b. The scattering rate for the atoms increases as the light, which is red-detuned, is brought closer to resonance for a transition.

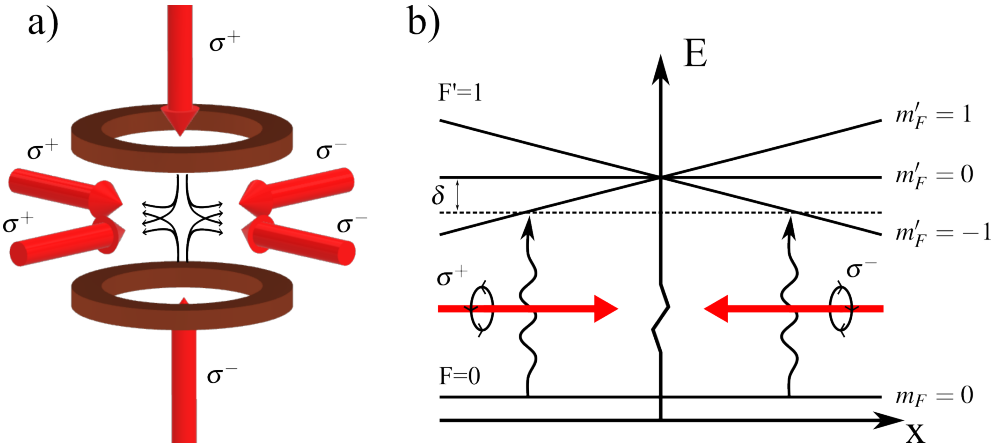


Figure 2.3: The six-beam MOT is shown. a) Two coils in anti-Helmholtz configuration (with current in opposing directions) produce a quadrupole trap with field lines shown in the center. Counter propagating circularly polarised laser beams red-detuned from a cooling transition are shown along each axis. b) The energy level diagram for a simplified cooling transition is shown. Due to selection rules σ^- polarised light is able to excite atoms from the $|F = 0, m_F = 0\rangle$ to the $|F' = 1, m'_F = -1\rangle$ state. The light is red-detuned from the transition to the $m'_F = -1$ state and is brought on resonance by a combination of the Doppler shift from an atom moving towards the beam and the energy shift of the transition due to the Zeeman effect. The scattering force from the absorption induces cooling and moves the atoms towards the trap centre.

As shown in Figure 2.3b as the atoms move along a spatially dependent magnetic field the Zeeman effect shifts the energy levels such that the light is nearer to resonance. The selection rules dictate that one polarisation only excites atoms on one side at any given time. Consequently, atoms which move away from the trap centre are excited by photons moving in the opposing direction and the resulting scattering force moves the atoms back towards the trap location.

To derive the new scattering force, firstly we note that the energy shift due to the weak field Zeeman effect over a small distance z is given by

$$\beta z = \frac{g_F |m_F| \mu_B}{\hbar} \frac{\partial B}{\partial z} z. \tag{2.36}$$

This equation is valid since the field is only a few Gauss. The new resonance condition for absorption to occur is $\omega_0 + \beta z$ for the $m_F = 1$ state and $\omega_0 - \beta z$ for the $m_F = -1$ state. Taking into account the selection rules, Zeeman shift, red-detuning and Doppler shift, the total scattering force is

$$\begin{aligned} F_{\text{MOT}} &= F_{sc}^{\sigma^+}(\omega - kv - (\omega_0 + \beta z)) - F_{sc}^{\sigma^-}(\omega + kv - (\omega_0 - \beta z)) \\ &\approx -\alpha \left(v + \frac{\beta}{k} z \right). \end{aligned} \tag{2.37}$$

The loading rate of a MOT is governed by the equation[79,33]

$$\frac{dN}{dt} = R - \Gamma_{\text{MOT}} N - \zeta \int n^2_{\text{MOT}}(r,t) dV, \tag{2.38}$$

where N is the atom number, the first term, R, is the *loading rate*, the second term is the losses due to background gas collisions, and the third term is the losses due to collisions with other trapped atoms. Γ_{MOT} is the loss rate due to background gas collisions, n_{MOT} is the atomic density in the MOT, and ζ is the loss rate due to collisions between trapped atoms. The last term is integrated over the trapping volume, V. In the analysis of the loading rate it is assumed any atom with velocity, v, less than the *capture velocity*, v_c, that passes through the trap is captured and the losses are dominated by collisions with background gas atoms of the same species. This allows us to derive an expression for the loading rate which is given by[33]

$$R = \left(\frac{8b^4 A}{\pi^2 m^2 v_{\text{th}}^3} \right) U_{\text{trap}}^2 n, \tag{2.39}$$

where v_{th} is the mean velocity of the background atoms, m is the mass of the atomic species, and n is the background gas density. U_{trap} and A are the depth and cross-sectional area of the trap respectively, and b is the ratio of the capture velocity to the escape velocity. From this we see that the loading rate is primarily determined by the background gas pressure, the trap depth, and the volume of the MOT. Lastly, assuming a constant atomic density in the trap which is reasonable with a sufficiently large MOT, Equation 2.38 can be integrated to obtain

$$N(t) = N_{\text{max}} \left(1 - e^{-t/\tau_{\text{MOT}}} \right). \tag{2.40}$$

N_{max} is the steady state atom number, which is determined by the equilibrium between the loading rate and collisional loss rates, and τ_{MOT} is the characteristic loading time. Again, this result assumes that

2.3.3 Compressed MOT

The CMOT stage follows the standard 3D MOT stage and is used to further reduce the temperature of the atoms and increase their phase-space density, based on requirements stated later. During this stage, the magnetic fields are ramped down to zero and the cooler light detuning is increased, which facilitates cooling mechanisms that work in a lower velocity regime than the MOT. The CMOT stage is considerably shorter than the MOT stage as there are no position-dependent forces preventing thermal expansion of the atomic ensemble, causing a longer CMOT stage to result in a sparse cloud.

In this experiment, the cooler beams consist of pairs of opposing circularly polarised light sources leading to the primary cooling mechanism for this stage being $\sigma^+ - \sigma^-$ polarisation gradient cooling. This effect is often confused with the Sisyphus cooling effect which employs linearly polarised beams. In this configuration, the cooling results from light shifts along the beam propagation axis; these shifts produce a wave with the Zeeman sub-levels being out of phase. An optical pumping process then ensures that atoms are always travelling up a potential gradient resulting in cooling. This might be present to some extent due to the complexity of polarisation in the superposition of all of the beams in a 3D MOT,

but it is not the primary cooling mechanism.

$\sigma^+ - \sigma^-$ Polarisation Gradient Cooling

The following section follows the theory by Dalibard and Cohen-Tannoudji [15] which was experimentally verified by Salomon et al [75]. In the $\sigma^+ - \sigma^-$ configuration, there are two opposing circularly polarised beams with opposite polarities and, assuming the same amplitude for both, these waves superimpose such that the electric field is linearly polarised and rotates around the propagation axis of the beams, as shown in Figure 2.4. In the rest frame of an atom, the polarisation of the light is seen to be linear and rotating about the z axis with a frequency determined by the velocity of the atom along z. Going from the rest frame of the atom to a rest frame in which the polarisation of the light is fixed along the y axis, this system can be treated as an atom exposed to a static (fictitious) magnetic field along z in addition to the electric field along y.

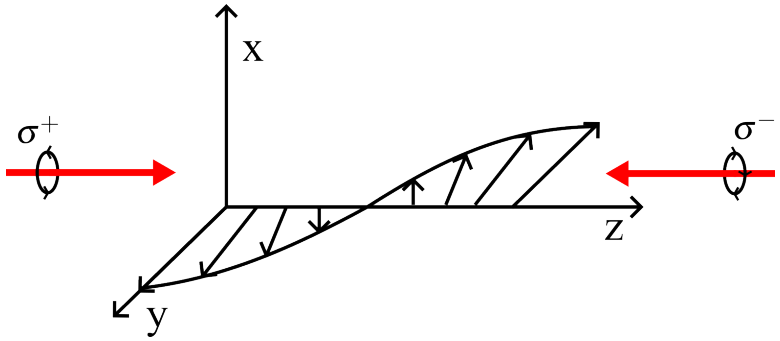


Figure 2.4: Counter propagating σ^+ and σ^- light superimposes to form linearly polarised field (black arrows) which rotates about the propagation axis.

The electric field results in a dc Stark shift of the Zeeman sub-levels for the angular momentum component along the y axis, lifting degeneracy, but with the light shift being constant along z. The effective Larmor precession about z resulting from the fictitious magnetic field causes coupling between the J_y Zeeman sub-levels, which in turn leads to a velocity dependence on the steady-state value of J_z, which is proportional to the difference in the populations of the J_z Zeeman sub-levels. It is this population difference which results in the imbalanced radiation pressure as the σ^+ (σ^-) radiation is more likely to be absorbed by atoms moving with velocity $v < 0$ ($v > 0$). It is a similar mechanism to the cooling process in a MOT, but with the absorption imbalance caused by population differences in the Zeeman levels instead of the combination of selection rules and Zeeman splitting bringing the light on resonance.

2.3.4 Optical Pumping

To optimise the number of atoms confined in a magnetic trap, they should all reside in the $|F = 2, m_F = 2\rangle$ state as this results in the largest potential energy gradient for a given magnetic field gradient. As-

suming a homogeneous magnetic field, the energy splitting for the atoms is independent of position and is equal between each m_F level as shown in Figure 2.5. If circularly polarised light, resonant with the $F = 2 \rightarrow F' = 2$ transition is shone parallel to the magnetic field axis, then this light is able to excite transitions dependent on the helicity of the polarisation. For σ^- polarised light $\Delta m_F = -1$ and for σ^+ polarised light $\Delta m_F = +1$, shown by the red arrows in Figure 2.5, however, the photons emitted due to spontaneous emission can take any polarisation and so for the emission $\Delta m_F = 0, \pm 1$ as shown by the blue arrows in Figure 2.5. Regardless of this, if the light is σ^+ polarised, then eventually the atoms will be 'pumped' to the $m_F = 2$ state since atoms will continue to be excited until they reside in this state due to the selection rules; once they are in the $m_F = 2$ state, they will no longer be excited by the light. This process is referred to as optical pumping.

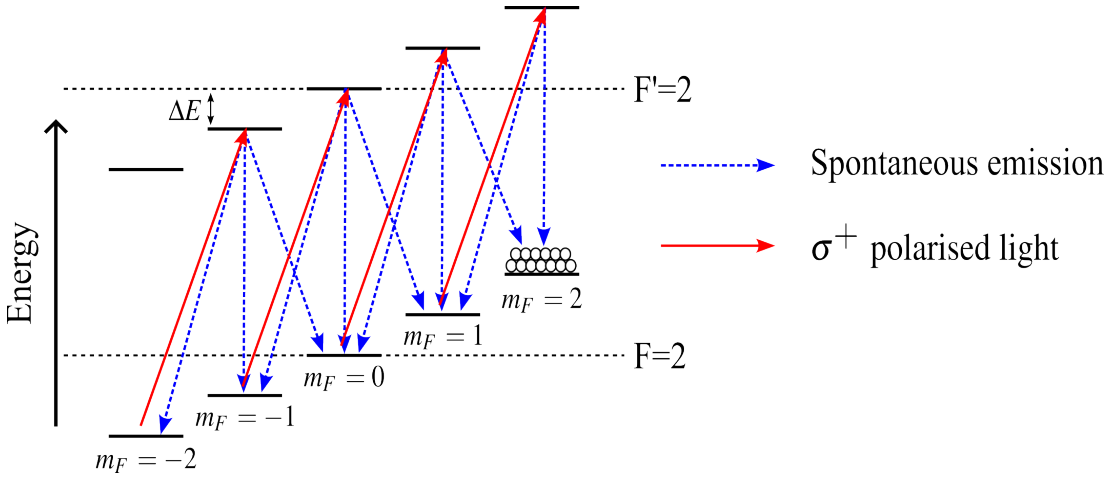


Figure 2.5: Atom can be pumped into the stretched $|F = 2, m_F = 2\rangle$ state with a pulse of σ^+ polarised light. The transitions due to absorption (red, solid) cause $\Delta m_F + 1$ and spontaneous emission (blue, dashed) results in $\Delta m_F = \pm 1, 0$. Thus on average m_F increases and after many cycles the atoms are pumped into the highest m_F state where the σ^+ light is unable to excite them further.

2.4 Magnetic Trapping

In this section, the principles of magnetic trapping are discussed. Typical magnetic field strengths for atom chip experiments are on the order of tens of Gauss, implying the weak Zeeman regime for calculating the energy shift.

The magnetic moment of an atom rotates about the quantisation axis of the static magnetic field at a frequency known as the Larmor frequency, which is defined as the energy spacing between Zeeman levels $\Delta E = \hbar\omega_L$ which gives

$$\omega_L = \frac{g_F m_F \mu_B}{\hbar} B_z. \tag{2.41}$$

Strong and Weak Field Seeking States

Referring back to Figure 2.2 we see that for some m_F states the energy shift is negative and for some it is positive. From Equation 2.16 the energy shifts are given by $\Delta E = g_F m_F \mu_B |\mathbf{B}|$ so whether or not the potential energy increases or decreases with an increasing magnetic field is determined by the sign of $g_F m_F$. States where $g_F m_F < 0$ are referred to as strong (or high) field seeking states and are attracted to increasing magnetic fields. Using static magnetic fields, Maxwell's equations forbid a source-free maximum magnetic field meaning we are unable to produce static magnetic traps for these states.

States for which $g_F m_F > 0$ are known as weak (or low) field seeking states and are attracted to magnetic field minima and so are trappable. For ^{87}Rb, $g_F = -\frac{1}{2}$ for the $F = 1$ manifold and $g_F \approx \frac{1}{2}$ for the $F = 2$ manifold. Thus, the trappable states in the $5^2S_{1/2}$ manifold are $|F = 1, m_F = -1\rangle$, $|F = 2, m_F = 1\rangle$, and $|F = 2, m_F = 2\rangle$.

2.4.1 Quadrupole and Ioffe Traps

Magnetic traps can be divided into two categories, quadrupole and Ioffe-Pritchard type traps. The field around the minimum of a quadrupole trap is linear and can be expressed as

$$\mathbf{B} = B'_x x\mathbf{e}_x + B'_y y\mathbf{e}_y + B'_z z\mathbf{e}_z \tag{2.42}$$

where the primary characteristic of note is that the field crosses zero at the minimum. According to Gauss' law of magnetism, $\nabla \cdot \mathbf{B} = 0$ which leads to the restriction that $B'_x + B'_y + B'_z = 0$. The field strength is simply given by $B(\mathbf{r}) = \sqrt{(B'_x x)^2 + (B'_y y)^2 + (B'_z z)^2}$, where $B'_i = \partial \mathbf{B}/\partial i$ with $i = x, y, z$. A quadrupole trap is used in the MOT and in the case of the MOT in this experiment is produced using external coils as shown in Figure 2.3a in anti-Helmholtz configuration, where the current runs in opposing directions in each coil. In this case, the gradients are related by $B'_x = B'_y = -\frac{1}{2}B'_z$. If the coils are run in Helmholtz configuration, with the currents in the same direction, then a homogeneous field is produced.

However, the field zero leads to a loss mechanism for trapped atoms known as *Majorana Losses*. In this region, there is no longer any splitting between states in the hyperfine manifold, and atoms passing through the zero can spontaneously transition to either strong-field seeking states or less strongly confined weak-field seeking states, resulting in losses.

To prevent these losses, an Ioffe-Pritchard (usually shortened to Ioffe) type trap is generally used when the atoms are held for a long time. This is distinguished from the quadrupole trap by a raised trap minimum which prevents the transitions which occur at the zero field crossing. An Ioffe trap can be produced using four pairs of coils, as shown in Figure 2.6. Firstly, replacing the symmetric anti-Helmholtz coil pair used in a quadrupole trap with elongated rectangular coils, called Ioffe coils in Figure 2.6, produces a 2D quadrupole with no confinement along the z-axis. A pair of coils with

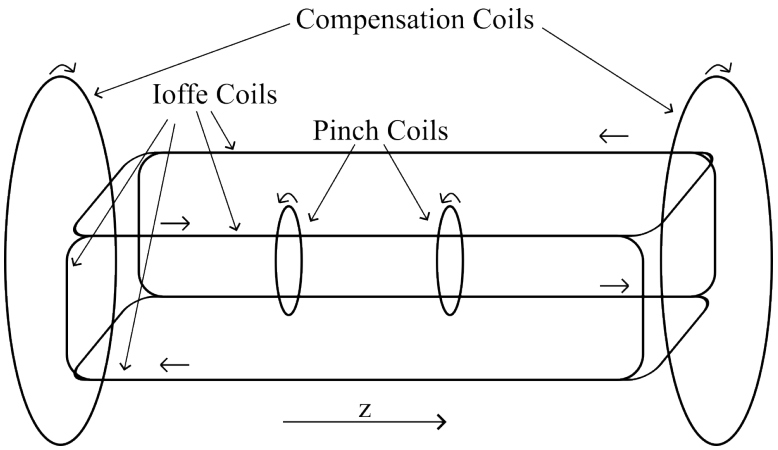


Figure 2.6: An Ioffe-Pritchard type trap can be formed by a combination of four pairs of coils, as shown, with short arrows indicating the direction of current. The Ioffe coil pairs having anti-parallel currents and form an elongated quadrupole trap along the z-axis. The pinch coils have parallel currents and produce an Ioffe field which raises the trap minimum whilst still providing trapping in the z-axis. Lastly, a pair of compensation coils can be added to oppose the pinch coils; this allows us to scale the trap minimum without modifying the gradient of the trap.

parallel current, called pinch coils, is then added with their circularly symmetric axis aligned with the z-axis. This produces an offset field which raises the trap minima and provides confinement in the z-axis since the field strength increases near the coils. Compensation coils, with currents opposing those in the pinch coils, can also be added to counter the Ioffe field. Since the bias field produced by these coils is homogeneous throughout the trapping region, it allows the trap minimum to be varied whilst keeping the gradient constant. The magnetic field produced by the Ioffe and pinch coils can be expressed as

$$\mathbf{B} = B' \begin{pmatrix} x \\ -y \\ 0 \end{pmatrix} + B_0 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \frac{B''}{2} \begin{pmatrix} -xz \\ -yz \\ z^2 - \frac{1}{2}(x^2+y^2) \end{pmatrix}, \tag{2.43}$$

where the first term gives the field produced by the 2D quadrupole and the second two terms describe the field from the 'pinch coils'.

2.4.2 Trapping Wire Geometries

Single Wire Waveguide

The simplest trapping geometry is the waveguide produced by a single wire and homogeneous bias field. This configuration is shown in Figure 2.7 with a wire along the y-axis with a positive current I such that a circularly symmetric field is produced with a field strength that decays quadratically along the radial axis. At any given point, the field is tangential to the wire, so a bias field, B_{bias}, along the x-axis is able to completely cancel the field due to the wire at some point along the z-axis. The strength of this bias field can be increased to move the zero towards, or decreased to move away from, the wire. The

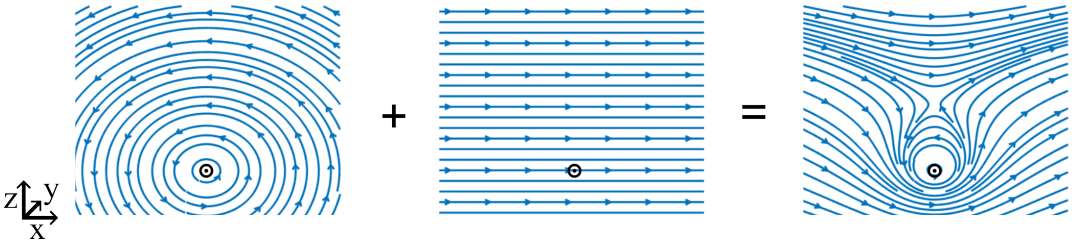


Figure 2.7: The superposition of the magnetic field due to a wire (left) and a homogeneous bias field (centre) can be used to produce a quadrupole trap (right) in the x and z axis at the height where the bias field exactly cancels the field from the wire.

resulting magnetic field, assuming a long, thin wire, is given by

$$\mathbf{B} = \begin{pmatrix} B_{\text{bias}} \\ 0 \\ 0 \end{pmatrix} + \frac{\mu_0 I}{2\pi(x^2+z^2)} \begin{pmatrix} -z \\ 0 \\ x \end{pmatrix} \tag{2.44}$$

which produces a quadrupole trap with confinement in the x and z-axis, but no confinement in the y-axis, thus forming a waveguide running along the wire. The height of the waveguide, z_0, is given by

$$z_0 = \frac{\mu_0 I}{2\pi B_{\text{bias}}} \tag{2.45}$$

and the gradient along z is

$$B' = -\frac{\mu_0 I}{2\pi z^2}. \tag{2.46}$$

The resulting magnetic field along the z and x-axis are shown in Figure 2.8 with the trap height, assuming the atoms are held sufficiently far from the surface of the chip, labelled B_{height}. For the atoms to be held against gravity, the magnetic potential gradient in the direction of gravity must exceed the gravitational potential gradient $\partial V_{\text{grav}}/\partial z = M_{\text{Rb}} g$. In the weak Zeeman effect regime which is typical for atom chip experiments, the energy shift from the magnetic field is given by Equation 2.16 and equating the gradient of this energy shift to that of the gravitational potential energy gives

$$B'_{\text{min}} = \frac{M_{\text{Rb}} g}{\mu_B m_F g_F}. \tag{2.47}$$

For ^{87}Rb in the $|F = 2, m_F = 2\rangle$ state the minimum magnetic field gradient is calculated to be ≈ 16 G/cm and ≈ 32 G/cm in the $|F = 1, m_F = 1\rangle$ and $|F = 1, m_F = -1\rangle$ states.

In addition to a minimum field gradient, the maximum potential height of the trap must be large enough to confine atoms of a given average kinetic energy, T. In a 3D harmonic potential, such as an

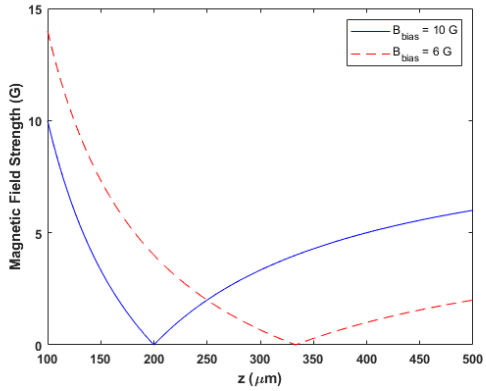


Figure 2.8: Shown is the variation of the magnetic field strength as the height away from the wire used to create a waveguide is increased. A current of 1 A is used and the resulting traps formed by both a 10 G bias field (blue, solid line) and 6 G bias field (red, dashed line) are shown. As the bias field is increased, the trap becomes deeper, steeper, and closer to the wire.

Ioffe trap, the kinetic energy of each atom is given by

$$K_{\text{Ioffe}} = \frac{3}{2} k_B T \tag{2.48}$$

according to the equipartition theory where k_B is the Boltzmann constant. Similarly, for a quadrupole trap this is given by

$$K_{\text{Quadrupole}} = 3 k_B T. \tag{2.49}$$

By equating this with the energy shift due to the magnetic field, given by Equation 2.16, we find that for a harmonic potential the minimum trap height,$B_{\text{H,min}}$, is

$$B_{\text{H,min}} \geq \frac{3}{2} \frac{k_B T}{\mu_B g_F m_F}, \tag{2.50}$$

and for a quadrupole trap, $B_{\text{Q,min}}$, is

$$B_{\text{Q,min}} \geq 3 \frac{k_B T}{\mu_B g_F m_F}. \tag{2.51}$$

However, to trap atoms for a significant time the actual trap height should be much higher than these values. Often a factor of five is used as a rule of thumb.

Z-Wire

Starting with a straight wire waveguide as described above, a simple method to provide trapping in the third dimension is to bend the wire at the ends to form either a "U" or "Z" shape. If the two arms are bent in the same direction to form the U-trap, then the currents flow in opposing directions such that the fields along the axial direction cancel, forming a zero field minimum, resulting in a quadrupole trap. This geometry is typically used to produce the quadrupole trap required for a MOT, which is known

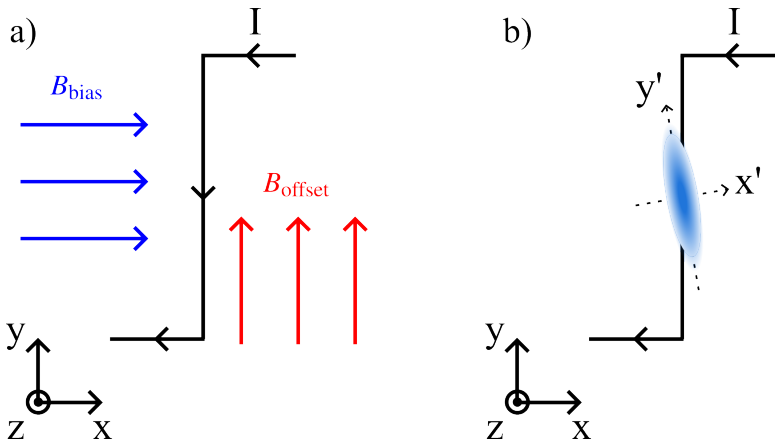


Figure 2.9: a) Shows the field configuration for forming a Z-wire trap. A bias field is perpendicular to the central wire and cancels the field above the wire to form the trap. Adjusting this field, one is able to adjust the distance of the trap from the wire. An offset field is parallel to the central wire and is able to vary the minima of the Ioffe trap. If the offset field is sufficiently strong it is able to split the Ioffe trap into two quadrupole traps. b) Shows the shape of the atomic cloud which is rotated away from parallel with the central wire. The trap frequency is significantly lower along the y'-axis forming a quasi-1D trap which is often described as cigar shaped.

as the U-MOT configuration. If the arms are bent in opposite directions forming a Z-trap such that the currents flow in the same direction, then the fields along the axial direction add, providing an Ioffe field with a non-zero minimum. Whilst U-traps are sufficient for thermal clouds, Ioffe-type Z-traps are required to prevent excessive losses from Majorana spin flips and so are the primary trapping geometry employed in atom chip designs where BECs and ultra-cold atoms are involved.

A Z-wire configuration is shown in Figure 2.9a with both a transverse bias field to control the trap height, as discussed in the case of a single wire waveguide, and an axial offset field which opposes the Ioffe field and can be used to control the trap minima. If this is set higher than the minimum of the Ioffe field then a quadrupole trap is formed which will split into two as the offset field is increased. The trap formed in this configuration is shown in Figure 2.9b and is often described as a "cigar" shape due to the elongation resulting from the significantly lower trapping gradient in the axial direction. The trap can be described using Equation 2.43, but from Figure 2.9b we see that the trap is actually rotated due to the asymmetry of the wire geometry and this equation is only valid in the new rotated axis x', y' and z'

Close to an Ioffe trap minima the fields can be approximated as a harmonic potential and these traps are often described in terms of the trap frequencies resulting from this description. The angular trap frequencies, ω_i, in each of the rotated axis are given by

$$\frac{\omega_i}{2\pi} = \sqrt{\frac{\mu_B g_F m_F}{M_{\text{Rb}}} \frac{\partial^2 |\mathbf{B}|}{\partial x_i'^2}}. \tag{2.52}$$

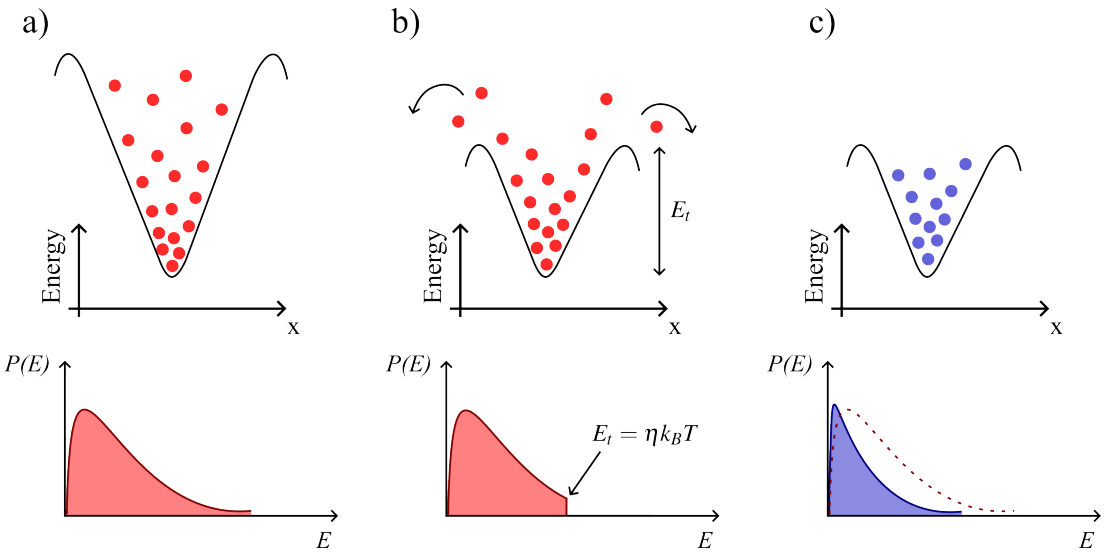


Figure 2.10: Starting with a potential well filled with hot atoms on the left which has a velocity distribution determined by the Boltzmann distribution shown below, the trap height can be lowered to E_t (middle images) allowing the hottest atoms to escape. This is often described as a knife cutting off the tail of the Boltzmann distribution. The atoms then rethermalise resulting in a lower average kinetic energy, ergo temperature.

2.5 Evaporative Cooling

To cool atoms to the point of BEC, laser cooling is insufficient as the photons begin to heat the atoms, preventing cooling below a certain point. We can introduce a parameter known as the phase space density, which is defined as

$$\rho_{ps} = n\lambda_{dB}^3 = n\sqrt{\frac{2\pi\hbar^2}{M_{\mathrm{Rb}}k_B T}} \tag{2.53}$$

which is simply the number of atoms per cube with sides equal to the de Broglie length and allows us to characterise the relevance of quantum statistics. n is the number density of atoms and λ_{dB} is the de Broglie length. Following laser cooling, the phase space density is typically on the order $\rho_{ps} \sim 10^{-6}$ and to form a BEC we require $\rho_{ps} = 2.6$. One method to cool atoms to this point is that of evaporative cooling, where we simply remove the hottest atoms and allow the cloud to rethermalise, thus reducing the average temperature. This method of cooling is one we experience every day in the form of steam in hot drinks. The process is illustrated in Figure 2.10 where the atoms are confined in a trap with an initial trap depth. This trap depth is then reduced, allowing the hottest atoms to escape, which is equivalent to cutting off the tail of the Boltzmann distribution governing the temperatures of the atoms and allowing the atoms remaining to rethermalise and reach a new equilibrium energy distribution.

To help characterise the evaporative cooling process we define another parameter known as the truncation parameter

$$\eta = \frac{E_t}{k_B T} \tag{2.54}$$

which quantifies the cut-off for the higher energy particles. So long as $\eta > 1$ the atoms being removed

have greater than average energy and the cloud is cooled.

2.6 Adiabatic RF-Dressed Potentials

An important tool in the development of atom chip technologies is the ability to manipulate the shape of potentials through the use of oscillating magnetic fields. This process essentially allows us to rotate the effective magnetic field coupled to the atomic angular momenta simply by varying the frequency, amplitude, and polarisation of the oscillating field. This allows for potentials that would be forbidden in the case of a pure static field, such as those with local maxima.

2.6.1 Linear Radio-Frequency (RF) Field

For this work we will cover the case of a linearly polarised RF field (σ-polarised with respect to the static field) used in conjunction with a static magnetic field as this is the relevant case for this experiment. This section closely follows the review by Perrin et al[61] where more information can be found.

The Hamiltonian resulting from the combination of a static and an RF magnetic field with frequency ω_{RF} and initial phase ϕ is

$$\hat{H} = \frac{g_F \mu_B}{\hbar} \mathbf{B}_0 \cdot \hat{\mathbf{F}} + \frac{g_F \mu_B}{\hbar} \mathbf{B}_1 \cdot \hat{\mathbf{F}} \cos(\omega_{\mathrm{RF}} t + \phi) \tag{2.55}$$

where the static field is given by $\mathbf{B}_0 = B_0 \mathbf{e}_z$ and the RF field is given by $\mathbf{B}_1(t) = B_1 \cos(\omega_{\mathrm{RF}} t + \phi)\mathbf{e}_x$. Here we have defined the z-axis to be along the direction of the static field and taken the RF field to be along x, which we can do since it is linearly polarised and it can be shown that the component along z has a negligible effect. In practice, this assumption means that we only consider the component of the RF field which is orthogonal to the static field. The Hamiltonian now simplifies to

$$\hat{H} = s \left[\omega_L \hat{F}_z + 2\Omega_1 \cos(\omega_{\mathrm{RF}} t + \phi) \hat{F}_x \right], \tag{2.56}$$

where we have introduced $s = g_F/|g_F|$ which is simply the sign of g_F, the Larmor frequency, ω_L, from Equation 2.41 and the Rabi frequency,

$$\Omega_1 = \frac{|g_F| \mu_B B_1}{2\hbar}. \tag{2.57}$$

From here it is convenient to move into a rotating reference frame which rotates around the z-axis at a frequency ω_{RF}, in a direction determined by s. To do this we can define the rotation operator

$$\hat{R} = \hat{R}_z[s(\omega_{\mathrm{RF}} t + \phi)] = \exp\left(-\frac{is}{\hbar} (\omega_{\mathrm{RF}} t + \phi) \hat{F}_z \right), \tag{2.58}$$

and introduce the rotated state $|\psi_{\text{rot}}\rangle = \hat{R}^\dagger|\psi\rangle$, where $|\psi\rangle = \hat{R}|\psi_{\text{rot}}\rangle$. The rotated state then follows the time-dependent Schrödinger equation $i\hbar\partial_t|\psi_{\text{rot}}\rangle = \hat{H}_{\text{rot}}|\psi_{\text{rot}}\rangle$, where $\hat{H}_{\text{rot}}$ is given by

$$\begin{aligned}\hat{H}_{\text{rot}} &= -i\hbar\hat{R}^\dagger[\partial_t\hat{R}] + \hat{R}^\dagger\hat{H}\hat{R} \\ &= -s\delta_{\text{RF}}\hat{F}_z + s\Omega_1\hat{F}_x + s\Omega_1\cos(2(\omega_{\text{RF}}t+\phi))\hat{F}_x + s\Omega_1\sin(2(\omega_{\text{RF}}t+\phi))\hat{F}_y\end{aligned} \tag{2.59}$$

and we have defined the detuning as $\delta_{\text{RF}} = \omega_{\text{RF}} - \omega_L$. In this reference frame, the first two terms are static whilst the last two rotate at the high frequency $2\omega_{\text{RF}}$. Assuming that both δ_{RF} and Ω_1 are small with respect to ω_L, we can apply the rotating wave approximation (RWA), where the oscillating terms can be neglected since they effectively average to zero. Thus, we find that

$$\hat{H}_{\text{rot}} \approx \left[-\delta_{\text{RF}}\hat{F}_z + \Omega_1\hat{F}_x\right], \tag{2.60}$$

which can be re-written as an effective Hamiltonian,

$$\hat{H}_{\text{eff}} = \Omega\hat{F}_\theta, \tag{2.61}$$

where the effective Rabi frequency is $\Omega = \sqrt{\delta_{\text{RF}}^2 + \Omega_1^2}$ and $\hat{F}_\theta$ is the projection along $\mathbf{e}_\theta = \cos\theta\mathbf{e}_z + \sin\theta\mathbf{e}_x$,

$$\hat{F}_\theta = \hat{\mathbf{F}}\cdot\mathbf{e}_\theta = cos\theta\hat{F}_z + sin\theta\hat{F}_x, \tag{2.62}$$

and

$$\theta = \arccos\left(-\frac{\delta_{\text{RF}}}{\Omega}\right) + \frac{s-1}{2}\pi. \tag{2.63}$$

We therefore find that the addition of a linear RF field orthogonal to the static field is equivalent to a rotation of the static field about the y-axis and $\hat{H}_{\text{eff}}$ is the Hamiltonian for an effective field along $s\mathbf{e}_\theta$ given by

$$\mathbf{B}_{\text{eff}} = \frac{\hbar}{g_F\mu_B}\sqrt{\delta_{\text{RF}}^2 + \Omega^2}(\cos\theta\mathbf{e}_z + \sin\theta\mathbf{e}_x). \tag{2.64}$$

The spin states $|m\rangle_\theta$ in the new basis given by the rotation of the quantisation axis about y are given by

$$|m\rangle_\theta = \hat{R}_y(\theta)|m\rangle_z = e^{-i\theta\hat{F}_y/\hbar}|m\rangle_z \tag{2.65}$$

and their eigenenergies are

$$E_m = m\hbar\Omega = m\hbar\sqrt{\delta_{\text{RF}}^2 + \Omega_1^2}. \tag{2.66}$$

It is important to note that this treatment is only valid when $\Omega, |\delta_{\text{RF}}| \ll \omega_L$, otherwise the RWA is no longer appropriate and instead a more involved approach based on Floquet analysis must be employed. In this approach a time-dependent Hamiltonian in matrix form is transformed into a time-independent,

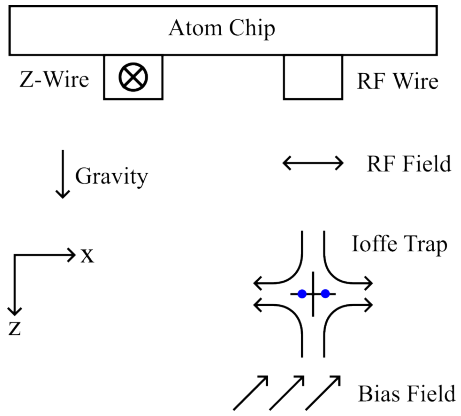


Figure 2.11: A scheme for introducting RF dressing as was implemented in[76] is shown. The atoms are confined in a Z-wire trap and shifted using a bias field to be directly below an RF wire such that the RF field is linearly polarised along x which results in the cloud being split along x. This results from the RF field being resonant with the Larmor frequency at some distance from the trap centre and the z axis confinement is due to the RF field and static field being parallel along this x resulting in a minima in the Rabi frequency.

but infinite matrix which can be cropped, allowing eigenvalues to be calculated.

2.6.2 Applying RF Fields to Ioffe-Traps

Let us consider the atom chip configuration shown in Figure 2.11, where the atoms are trapped in an Ioffe trap. A Z-wire trap is biased to position the atoms directly below an RF wire such that the RF field is linearly polarised in the horizontal axis. The adiabatic potential in this case is given by Equation 2.66, which can be explicitly written as

$$V_{\text{eff}}(\mathbf{r}) = m\sqrt{\underbrace{(\hbar\omega_{\text{RF}} - g_F\mu_B|\mathbf{B}_{\text{trap}}|)^2}_{\text{Resonant Term}} + \underbrace{\left(\frac{1}{2}g_F\mu_B B_{\text{RF},\perp}\right)^2}_{CouplingTerm}}, \tag{2.67}$$

where $B_{rf,\perp}$ is the magnitude of the RF field component orthogonal to the position dependent static field $\mathbf{B}_{\text{trap}}(\mathbf{r})$ and can be calculated as

$$B_{\text{RF},\perp} = \frac{|\mathbf{B}_{\text{trap}}(\mathbf{r}) \times \mathbf{B}_{\text{RF}}(\mathbf{r})|}{|\mathbf{B}_{\text{trap}}(\mathbf{r})|}. \tag{2.68}$$

The terms under the square root in equation 2.67 can be considered as a *resonant term* which is responsible for transitions between the Zeeman sublevels summed with a *coupling term* which lifts the degeneracy at the transitions and creates the avoided crosses between the adiabatic states. These effects are seen in Figure 2.12 where the bare states are first shown, with the $m_F = \pm 1$ states neglected for simplicity and the Larmor frequency, $\omega_{L,0}$, at the trap minima labelled. The resonant term then couples the bare states by allowing transitions wherever the $\omega_{\text{RF}} = \omega_L(\mathbf{r})$, which has the effect of producing a resonant shell encapsulating the entire Ioffe trap. A cross-section of which is shown in the middle

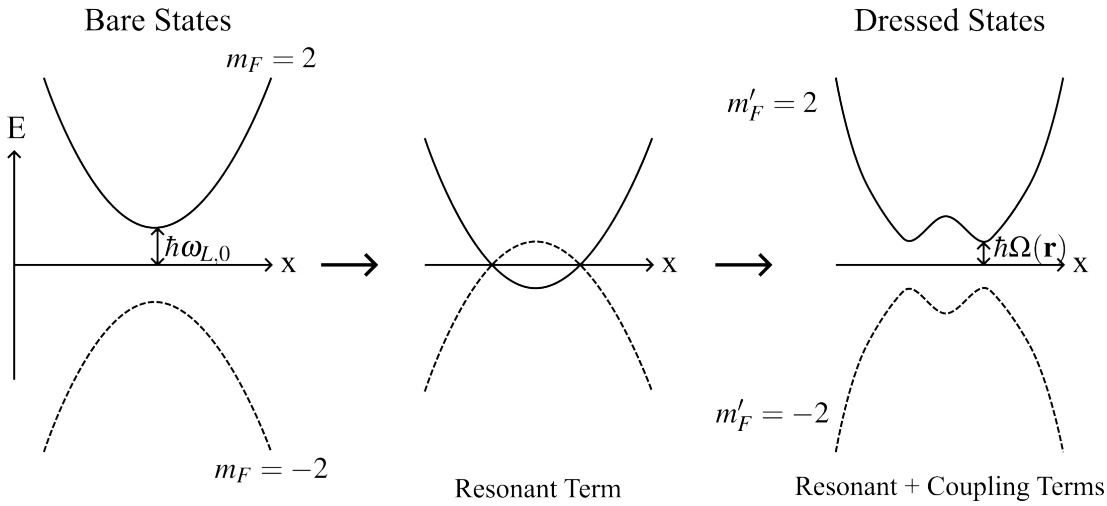


Figure 2.12: The formation of adiabatic potentials due to linearly polarised RF field on an Ioffe trap is shown. The initial bare states for the $F = 1$ manifold are shown on the left with the minimum Larmor frequency denoted $\omega_{L,0}$. In the centre shows the effect of the resonant term in Equation (2.67) which allows for transitions between the hyperfine levels, however as shown on the right, the coupling term in Equation (2.67) results in avoided crossings where atoms adiabatically follow the new potentials shown.

figure with the $m_F = \pm 2$ states being brought to cross at locations determined by the resonance of the RF frequency. The coupling term then gives the coupling between the RF field and static field which is responsible for repelling these levels and creating the avoided crossings that enable the adiabatic potentials. The shifting of these energy levels is given by the position dependent Rabi frequency, $\Omega(\mathbf{r})$, defined in Equation 2.66. Recalling the Ioffe trap described in Equation 2.43 and the RF polarisation shown in Figure 2.11, we see that along the y-axis the coupling strength is constant, resulting in a constant Rabi frequency. Along the z-axis, the coupling is minimal near the trap centre but increases with z, thus confining the atoms in the z-axis and resulting in a double-well potential as shown in both Figures 2.11 and 2.12 where the atoms are trapped at two points along the x-axis on the resonant surface.

This double-well potential allows for coherent splitting of a BEC or ultra-cold atomic cloud and so can be used to implement a Mach-Zehnder interferometer on an atom chip.

Chapter 3

CMOS Theory

Complementary metal-oxide-semiconductor (CMOS) is a fabrication technique for integrated circuit devices built on silicon wafers and is distinguished from other MOS technologies by the inclusion of both n-type and p-type transistors. Integrated circuits are made possible by the *MOSFET* (metal-oxide-semiconductor-field-effect-transistor) which has a simple design allowing for miniaturisation and relatively easy manufacturing. The earliest transistors were bipolar transistors developed at Bell Labs by Bardeen and Brattain in 1947[7] with a comprehensive description provided by Shockley in 1949[78] detailing the pn-junction. The MOSFET, which is the transistor used in the vast majority of integrated circuits, was introduced by Kahng and Atalla in 1960[16]and was built using silicon, thermally grown silicon oxide and metal contacts.

There are two types of MOSFET, N-channel (also called NMOS transistors) and P-channel (also called PMOS transistors) which are distinguished by the mobile charge carriers able to carry current through the device. MOSFETs have four ports, the *gate* which enables or disables current flow, the *source* and *drain* are the ports through which current flows and the *bulk* (or body) sets the potential of the well surrounding the MOSFET. Physically, the source and drain are symmetric with the source being whichever port is at a lower voltage for N-channel MOSFETs and a higher voltage for P-channel MOSFETs. Figures 3.1a-d show the schematic symbols used to denote MOSFETs, with a-b showing NMOS symbols, and c-d showing PMOS symbols. Each type has two variations, either the bulk connection is floating, to be connected explicitly, or it is connected to the source by default. Note that there are other variations in MOSFET symbols, having both the arrows showing the direction of current and the circle at the gate of the PMOS is redundant so often only one of these distinguishing features is used. Figure 3.1e shows an isometric view of a MOSFET. The gate is shown in green and sits in the gap between the source and drain (orange), but is separated from the substrate (blue) by a thin dielectric layer (white). Transistors are described by their width and length as labelled on e, with the length being the distance between the source and drain. Often CMOS technologies are described by the minimum gate length,

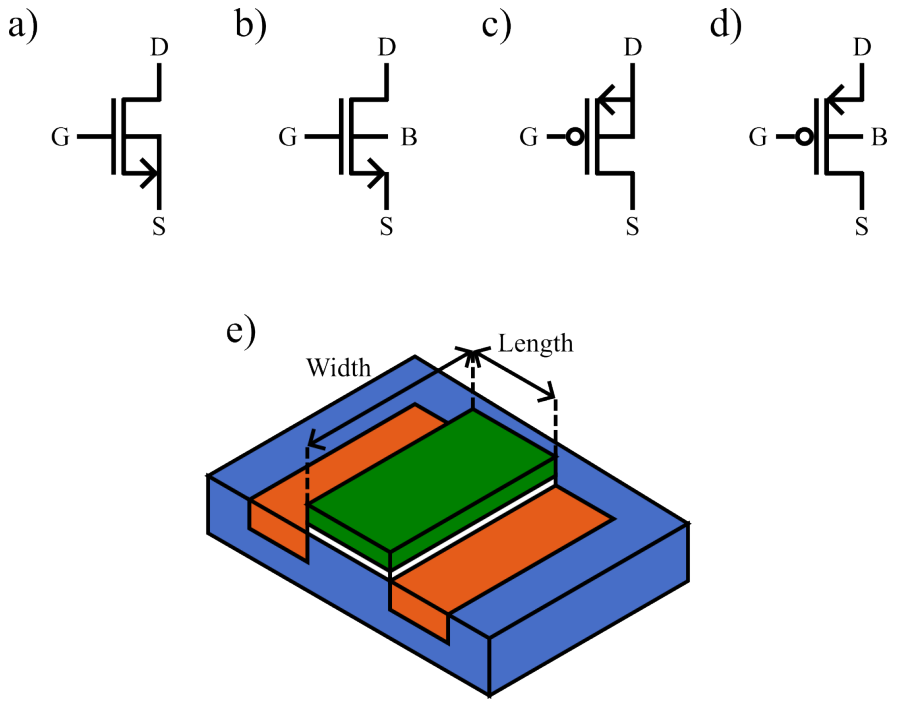


Figure 3.1: a) The schematic symbol for an NMOS with it's body connected to the source which is the lowest voltage. The arrow shows the flow of conventional current toward the source. b) Schematic symbol for an NMOS with the body as a separate connection. c) Schematic symbol for a PMOS with it's body connected to the source. The arrow shows the direction of conventional current from the highest voltage through the PMOS to the drain. d) A PMOS with the body left as a separate connection. e) An isometric view of a MOSFET with the substrate/well in blue, the source and drain shown in orange, the insulating oxide layer in white and the gate poly layer in green. A MOSFET is characterised by it's length, which is the distance between the source and drain, and the width.

for example a 350 nm technology is used in chip1, implying a minimum gate length of 350 nm.

3.1 Fabrication

Details of CMOS fabrication can be found in most textbooks on CMOS design[6 63 52 43], thus will only be briefly described here; however, it is still important to understand the structure of a CMOS chip. Starting with an Si wafer, which is typically between 0.5-1 mm thick with a diameter of 10-30 cm, many chips are fabricated over the wafer. For our purposes, an entire wafer would be prohibitively expensive, so a small area on a multi-project wafer is purchased. Typically, foundries charge per unit area.

A cross section of a CMOS chip is given in Figure 3.2 and shows the primary layers involved. At the bottom is the p doped Si *substrate* into which ions are implanted to form the separate p and n doped regions. Wells are deeper doped regions and the + sign is used to denote regions of higher doping concentrations such as the source and drain of MOSFETs. Devices are separated by thick oxide (SiO_2) barriers referred to as the *field-oxide* (FOX) and the process for producing this is known as *shallow trench isolation* (STI)[6]. Transistor gates are formed by a thin oxide layer that acts as a dielectric and a polysilicon (poly) layer to form the conductor, where *polysilicon* is simply an amorphous (non-

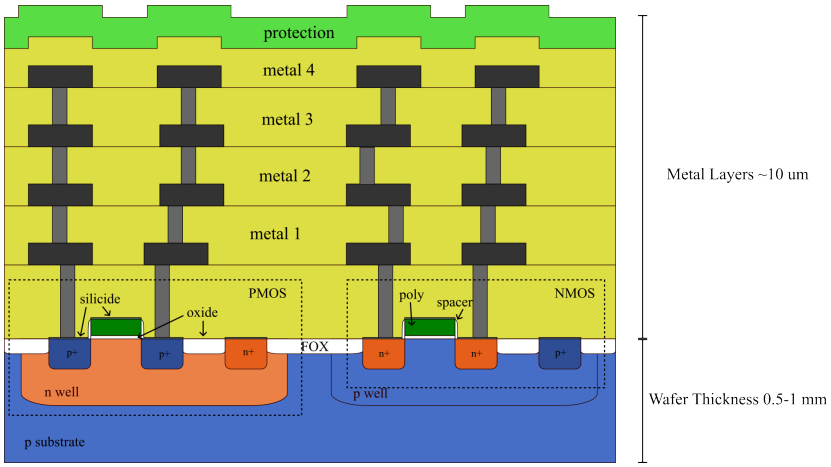


Figure 3.2: A simplified cross-section of a CMOS chip is shown (not to scale for clarity). In blue a the bottom is the p doped silicon substrate and both an NMOS and PMOS transistor are shown. P and N wells are deep doped regions and have a voltage set by a higher doped electrode. Shallow higher doped regions form the MOSFETs and are separated by thick oxide layers labelled FOX (field oxide). A thin oxide layer then separates the gate poly. Several aluminium layers labelled metal 1 to metal 4 are then added insulating oxide layers separating them. At the top an additional oxide layer is added to protect the top of the chip and is only opened for the pads used in wire bonding.

crystalline) form of Si. Since the resistance of poly and n doped regions is very high, everything is covered in a thin highly conductive layer known as *silicide* (typically either $TiSi_2$ called titanium silicide or tungsten) to reduce resistance. Metal layers are typically made of aluminium and are separated by insulating layers with vias connecting them. At the top, an additional oxide layer known as the protection layer or passivation layer is added to protect the chip.

3.1.1 Fabrication Techniques

There are several key fabrication techniques necessary to understand the fabrication process, which are detailed now.

Etching

Etching is the process of cutting away at materials with two categories, wet and dry. *Wet etching* uses chemicals to remove material and is typically used to clean and remove thin films. This can be combined with a polishing pad to perform *chemical mechanical polishing* (CMP) which is used to planarize. *Dry etching* involves bombardment with other particles; either plasma or sputtering can be used.

Deposition

There are two main types of deposition. *Physical vapour deposition* (PVD) in which the particles are directly deposited, either evaporated and then condensed on the wafer or sputtered. *Chemical vapour deposition* (CVD) is when a reactive gas is passed over the heated wafer and reacts with the surface to produce the desired layer.

Ion Implantation

Doping layers are formed by *ion implantation* where the wafer is bombarded with a high-energy focused beam of dopant atoms. The intensity and duration of the beam determine the doping profile. This process damages the crystal structure of the wafer, so it must be annealed to repair the lattice bonds by being heated to $1000°C$ for 15-30 minutes.

Oxide Growth

A layer of SiO_2 can be grown on the substrate by heating the wafer and exposing it to either gaseous O_2 (*dry oxidation*) or H_2O (*wet oxidation*).

Photolithography

To ensure that the above processes only affect the desired areas, a protective mask must be placed above the rest to act like a stencil. This is formed by first growing a thin protective oxide layer and then depositing a layer of *photoresist* over the material, this is a material which hardens when exposed to UV light forming a protective layer while remaining soft if not exposed. This layer is then covered in a transparent mask with opaque regions covering the target area for the deposition/etching/implantation and the wafer is exposed to UV light. The remaining soft photoresist is etched leaving the substrate beneath exposed.

3.1.2 Fabrication Steps

The fabrication steps are divided into two categories. *Front end of line* (FEOL) which includes everything up to the ion implantation, and *back end of line* (BEOL) processes, which include the metallisation. Fabrication starts with shallow-trench lithography (STL), where photolithography is used to protect areas where no FOX is to be present; then shallow trenches are etched into the wafer, ion implantation is used to increase the threshold voltage, and oxide is deposited with CVD. The threshold voltage is increased to prevent parasitic MOSFETs from forming, with the FOX acting as a gate. The remaining photoresist layer and protective oxide layer are removed, and new layers are added to expose the N wells.

The N and P wells are then formed by photolithography followed by ion implantation. To form the gates, a thin, gate-oxide layer is grown over the wafer, and a lithography mask is used to expose the gate areas. Ion implantation is then used to adjust the doping beneath the gate, which allows engineering of the threshold voltage, and a poly layer is deposited on top of the gate oxide. The unwanted gate oxide is then etched. Lastly, for the FEOL, the n+ and p+ doped regions are formed by separate lithography and ion implantation steps. Although the order of some FEOL stages might vary, ion implantation for the

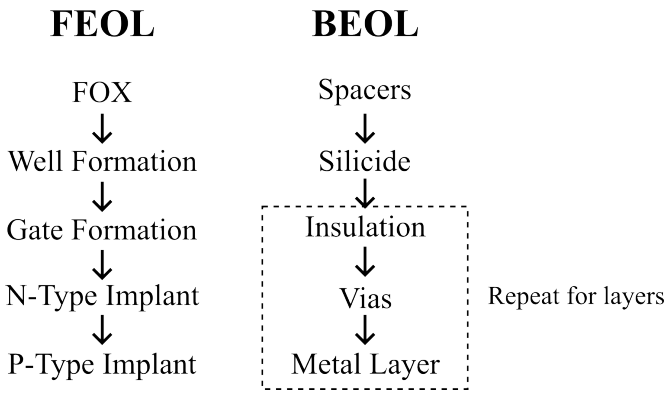


Figure 3.3: Fabrication steps in CMOS are categorised as either front-end of the line (FEOL) or back-end of the line (BEOL). FEOL includes everything up to MOSFET ion implantation and BEOL includes the metal layers.

source and drain is always performed after the gate to prevent the gate from being placed asymmetrically. These steps are summarised under the FEOL column in Figure 3.3.

Back-end of the line focuses on the metal layers and begins with oxide spacers being formed on the gates, as shown in Figure 3.2, to prevent shorts between the silicide on the gate and source/drain. Silicide is sputtered onto the active areas where no FOX is present. Thick oxide layers are then deposited over the entire wafer using CVD to form the poly-metal insulation layer and the areas for the metal contacts are etched. Note that many small contacts are used as opposed to large contacts because of a phenomenon called "spiking", where the metal cuts through the doped regions and potentially shorts the diode. The first metal layer is then sputtered over the wafer, filling the metal contacts and covering the wafer. This is then etched and the next oxide layer deposited. The process used to form the first metal layer is repeated for all of the metal layers with a final additional oxide layer deposited at the top to protect the chip. These steps are summarised in the BEOL column of Figure 3.3.

3.1.3 Passive Components

In addition to the MOSFETs used most ubiquitously in CMOS design, standard resistors, capacitors, and inductors can also be used; though inductors often require custom design and are not included as standard in the design library. There are several different variations of all these components which can be employed depending on the values of resistance, capacitance, or inductance required and will be introduced below. More details can be found in CMOS analogue design books[6,63].

Resistors

The two most common varieties of resistors are the silicide-block poly and n well resistors. Poly resistors are formed by a poly layer with metal contacts at the ends and without the silicide layer on top, resulting in a significantly higher resistance. They typically have a resistance of $1\Omega/\square$ where $\square$ denotes the

minimum sized square. N well resistors are formed by two n+ plugs separated by an n well. They have a significantly higher resistance of $1k\Omega/\square$.

Capacitors

Capacitors can be formed with a variety of combinations of conductor-oxide-conductor with the oxide as the dielectric. Poly-diffusion capacitors are formed from an n+/p+ region and poly layer separated by oxide. Poly-poly capacitors use two poly layers separated by an oxide layer. Metal-poly layers use a metal and poly layer. Metal-metal simply uses two metal plates. Each type has a different minimum size, capacitance, parasitic capacitance, and linearity, which vary between processes.

Inductors

Inductors are formed by metal wires forming loops. These are typically either formed by concentric squares/octagons on the same plane but can also include multiple layers.

3.2 The pn-Junction

In intrinsic silicon, electrons sit in the *valence band* with a small probability of being excited into the *conduction band* where they are able to move. The concentration of free electrons, n, in intrinsic silicon is known as the *intrinsic carrier concentration*, n_i, and for silicon is given by $n = p = n_i \approx 14.5 \times 10^9 \text{carriers}/\text{cm}^3$, where p is the concentration of holes. The concentration of silicon atoms is given by $N_{si} = 50 \times 10^{21} \text{atoms}/\text{cm}^3$.

In a p-type region, the silicon is doped with *acceptor impurities*, at a concentration N_A, which removes electrons from the valence band, allowing holes in the valence band to move. The hole concentration in this case is given by $p \approx N_A \gg n_i$. Similarly, silicon can be doped with *donor impurities*, at a concentration of N_D, which donates electrons to the conduction band. The concentration of free electrons in this case is given by $n \approx N_D \gg n_i$. Note that we assume $N_{si} \gg N_A$ and $N_{si} \gg N_D$. A helpful rule to remember is the *mass-action law* which dictates the relationship between holes and electrons in a doped semiconductor and tells us that $pn = n_i^2$.

One consequence of the modified carrier concentrations due to doping, is the change in the semiconductor *Fermi level*, which is defined as the energy level where the occupation of a free electron is 50%. As shown in Figure 3.4a, the intrinsic Fermi level for silicon, E_i, sits mid-way between the valence band, E_V, and the conduction band, E_C. In a p-type semiconductor, free electrons are far more likely to reside in the valence band, lowering the Fermi energy, E_{fp}, as shown in Figure 3.4b. The energy

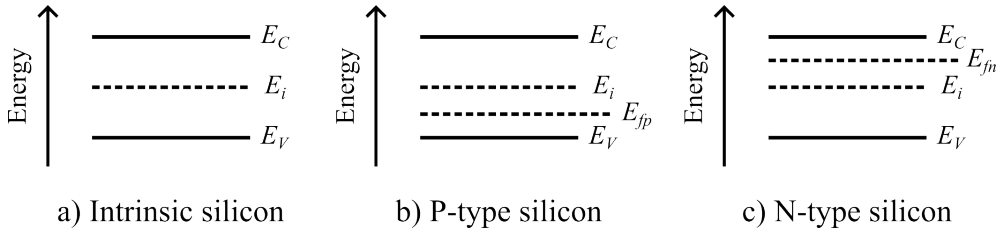
difference between the p-type Fermi level and intrinsic Fermi level is given by

$$E_i - E_{fp} = k_B T \ln\left(\frac{N_A}{n_i}\right), \tag{3.1}$$

where T is the temperature and k_B is Boltzmann 's constant. Conversely, in an n-type semiconductor, free electrons are more likely to be in the conduction band which raises the Fermi energy, E_{fn} as shown in Figure 3.4c. The energy difference in this case is given by

$$E_{fn} - E_i = k_B T \ln\left(\frac{N_D}{n_i}\right). \tag{3.2}$$

The effect of doping on the Fermi energy level is important for understanding equilibrium in a pn-junction.



a) Intrinsic silicon b) P-type silicon c) N-type silicon

Figure 3.4: Shown are the energy band diagrams for a) intrinsic silicon b) p-type silicon and c) n-type silicon. In each, E_C is the bottom of the conduction band, E_V is the top of the valence band, and E_i is the intrinsic Fermi level. E_{fp} is the Fermi level of p-type silicon and E_{fn} is the Fermi level of n-type silicon.

A pn-junction is formed at the interface between a p-type and an n-type semiconductor. Figure 3.5a shows the non-equilibrium case of a junction which has only just formed. On the p-type side, there are an equal number of stationary acceptor ions and free holes, whilst on the n-type side, there are an equal number of stationary donor ions and free electrons. Free charge carriers will immediately begin to diffuse across the boundary of the p and n-type regions to form a uniform distribution. This process reaches an equilibrium due to an electric field which repels further diffusion, formed by the excess positive ions on the n-type side of the boundary and the excess negative ions on the p-type side. The area where the free charge carriers have diffused is known as the *depletion region* and is typically approximated as a box of width W_D which isn't necessarily symmetric about the boundary.

The potential difference causing this electric field is known as the *built-in potential*, ϕ_{bi} of the pn-junction. It can be calculated by noting that in the equilibrium case, the Fermi level must be constant through the junction, resulting in the energy bands shown in Figure 3.6. This implies that the built-in potential is given by the difference between the p-type Fermi level and n-type Fermi level, $q\phi_{bi} = E_{fn} - E_{fp}$, where q is the electron charge. From Equations 3.1 and 3.2 we find

$$\phi_{bi} = \frac{k_B T}{q} \ln\left(\frac{N_A N_D}{n_i^2}\right), \tag{3.3}$$

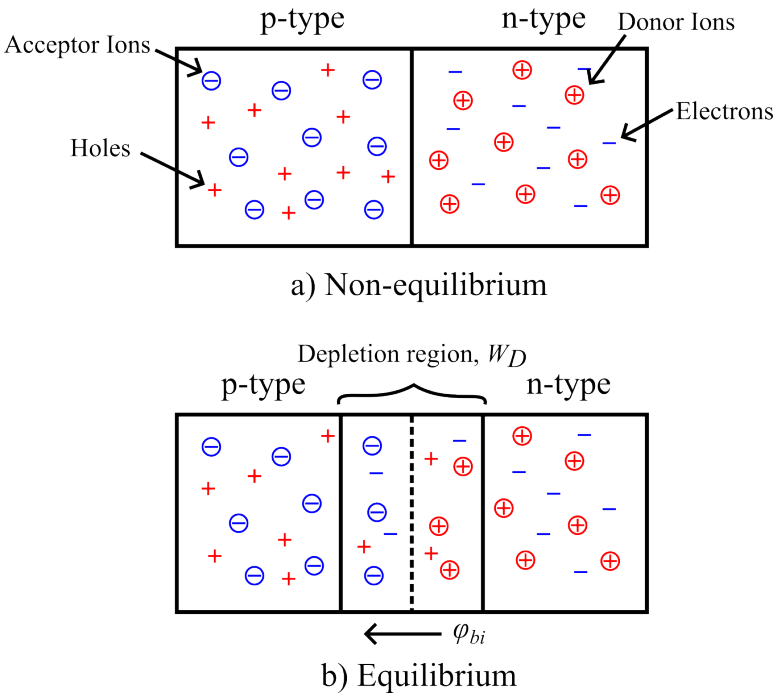


Figure 3.5: a) A pn-junction is formed at the interface of p-type and n-type regions of a semiconductor. Initially, both sides have equal numbers of fixed ions and charge carriers resulting in a neutral charge. b) In equilibrium, a depletion region of width W_D forms at the interface and a built-in potential, ϕ_{bi}, maintains the equilibrium.

where $k_B T/q$ is a quantity known as the thermal voltage.

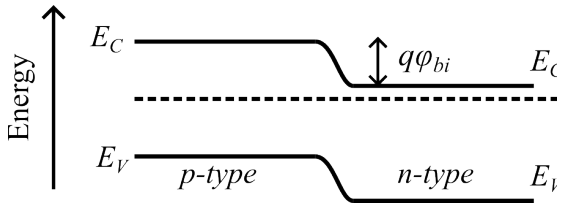


Figure 3.6: To maintain equilibrium, the Fermi level is constant throughout the interface resulting in the built-in potential.

3.3 Principles of MOSFET Operation

As introduced above, MOSFETs have 4 ports: the gate, source, drain, and bulk. The cross-section of an N-channel MOSFET is shown in Figure 3.7 with labelled ports and voltages. The NMOS consists of two n+ doped regions in a p well, separated by the area beneath the gate, which is electrically isolated from the p well by a thin oxide layer. The bulk voltage of the p well is controlled through a p+ plug and is typically set to ground for all NMOS, ensuring that no parasitic forward-biased pn-junctions appear.

The minimum gate-source voltage, V_{GS} necessary for current to flow is known as the threshold voltage and is determined by the fabrication process of the device. There are three modes of operation for the MOSFET transistor and are summarised for an N-channel MOSFET:

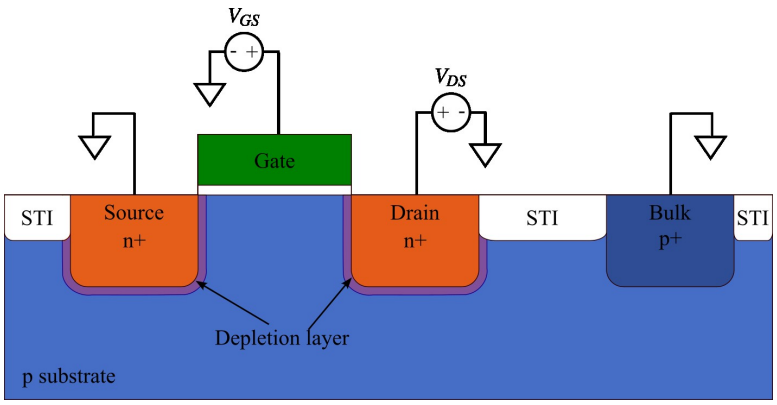


Figure 3.7: An NMOS is formed by two n+ doped regions (orange) in a p-doped well or substrate (lighter blue) seperated by a gate which is formed by a poly layer (green) isolated from the substrate by a thin oxide layer (white). The source and drain are interchangeable, and the names just refer to whichever one has the lowest and highest voltage, respectively. At the PN junction surrounding the n+ regions a depletion layer forms. The body voltage is controlled by a p+ electrode which modifies the potential of the surrounding substrate.

1. **Cut-off**: When $V_{GS} < V_{THN}$ the resistance between the source and drain is determined by the p-doped substrate so current flow is negligible.

2. **Triode**: When $V_{GS} \geq V_{THN}$ and $V_{DS} \leq V_{GS} - V_{THN}$ the transistor is in the *triode* or linear state where strong inversion of charge carrier concentration occurs below the gate, allowing current to flow and the drain current, I_D, to increase with V_{DS}.

3. **Active**: When $V_{DS} > V_{GS} - V_{THN}$, I_D approximately saturates and the transistor is in the *active*, or saturated, mode. The voltage $V_{D,\text{sat}} = V_{GS} - VTHN$ is known the *saturation voltage*.

Similarly, a PMOS transistor is shown in Figure 3.8, however, unlike the NMOS, the source and drain consist of p+ doped regions sitting in an n well. In this case, the source is at a higher potential than the drain and is often tied to VDD along with the bulk. The threshold voltage for a PMOS, V_{THP}, is negative with respect to the source. To avoid confusion with signs, all voltages in this chapter will be positive, with the order of the subscript letters determining the direction of the voltage drop such that $V_{SG} = -V_{GS}$. Hence, for a PMOS, the condition for strong inversion is $V_{SG} > V_{THP}$ and the saturation voltage is $V_{SD} = V_{SG} - V_{THP}$.

This section will introduce the operating principles of a MOSFET as initially described by Sah[74] then expanded upon by Shichman and Hodges[77] which is sometimes called the level-1 model or Shichman-Hodges model. Whilst sufficient for qualitative analysis and rough calculations, it is only accurate for long-channel technologies with lengths over 4μm. Going beyond this model will be discussed at the end of the section.

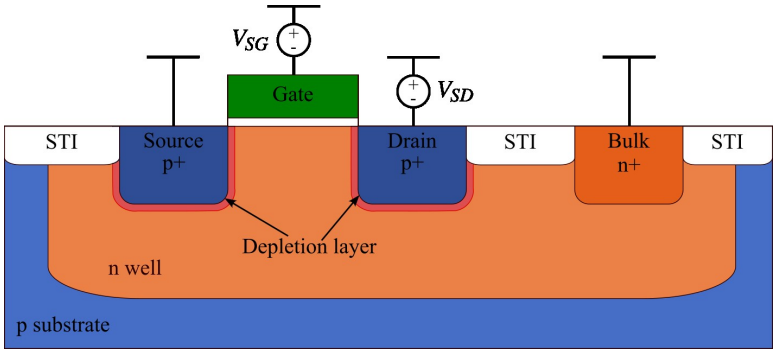


Figure 3.8: Since the substrate (light blue) is typically p doped, to form a PMOS an n doped region (light orange) known as a n well must first be implanted which has a voltage controlled by a p+ doped electrode (darker orange). The PMOS itself is formed by p+ doped regions (darker blue) seperated by a poly gate (green) which is isolated from the substrate by a thin oxide layer (white). The p doped region at a higher voltage is called the source and the one a lower voltage is called the drain. A depletion region forms at the PN junction surrounding the source and drain.

The Threshold Voltage

To calculate the threshold voltage, we first assume the source, drain, and bulk are all tied to ground and define ϕ_s as the electrostatic potential beneath the gate oxide. Initially, the gate voltage is $V_{GS} = 0$ V and $\phi_s = \phi_{fp}$, where ϕ_{fp} is the substrate Fermi potential given by

$$\phi_{fp} = -\frac{E_i - E_{fp}}{q} = -\frac{k_B T}{q} \ln\left(\frac{N_D}{n_i}\right), \tag{3.4}$$

using Equation 3.1. Weak inversion occurs when $\phi_s = 0$, at which point the area below the gate is depleted and the carrier concentration is $n = n_i$. As V_{GS} is increased further, $\phi_s = -\phi_{fp}$ and strong inversion occurs forming a channel of free electrons between the source and drain. The value of V_{GS} at which $\phi_s = -\phi_{fp}$ is defined as the threshold voltage.

Assuming the bulk is tied to ground, the threshold voltage is given by

$$V_{THN0} = -\phi_{ms} - 2\phi_{fp} + \frac{Q'_G}{C'_{ox}}, \tag{3.5}$$

where ϕ_{ms} is the electrostatic potential difference between the gate, ϕ_G, and bulk, ϕ_{fp} given by

$$\phi_{ms} = \phi_G - \phi_{fp} = \frac{k_B T}{q}\left[\ln\left(\frac{N_{D,\text{poly}}}{n_i}\right) + \ln\left(\frac{N_A}{n_i}\right)\right], \tag{3.6}$$

assuming a poly n-type doping concentration $N_{D,\text{poly}}$. Q'_G is the charge under the gate/unit area at the threshold voltage where

$$Q'_G = \sqrt{2q\varepsilon_{si}N_A - 2|\phi_{fp}|}, \tag{3.7}$$

and ε_{si} is the dielectric constant of silicon. Lastly, C'_{ox} is the gate oxide capacitance per unit area defined

by

$$C'_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}, \tag{3.8}$$

where we have introduced the oxide dielectric constant, ε_{ox}, and the oxide thickness, t_{ox}. Note that for Q'_G and C'_{ox} we have introduced a notation where the prime indicates that these values are per unit area, and the lack of a prime indicates the total value.

So far we have assumed that the bulk is connected to ground; if this is not the case, then the threshold voltage is modified by the *body effect*, and is instead given by

$$V_{THN} = V_{THN0} + \gamma\left(\sqrt{|2\phi_{fp}| - V_{BS}} - \sqrt{|2\phi_{fp}}\right), \tag{3.9}$$

where V_{BS} is the bulk-source voltage and γ is the *body effect coefficient* (also called body factor), defined as

$$\gamma = \frac{\sqrt{2q\varepsilon_{si}N_A}}{C'_{ox}}. \tag{3.10}$$

In most instances, the body and source are shorted, giving a constant threshold voltage; however, the body effect allows the threshold voltage to be adjusted for a specific application.

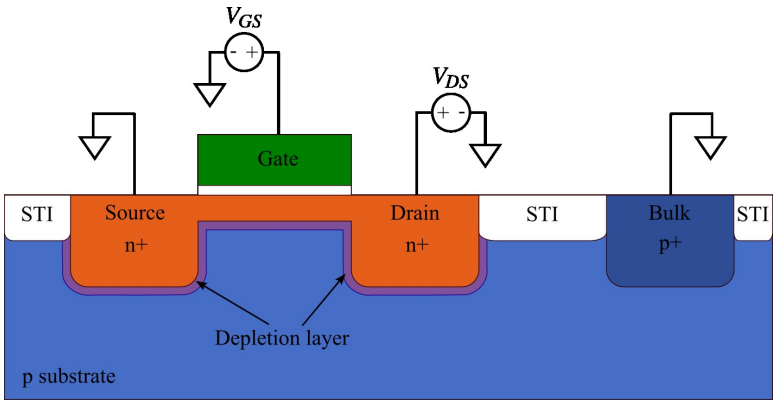


Figure 3.9: Shown is an NMOS operating in the triode region. As V_{GS} increases the positive charges beneath the gate are pushed away which inverts the p doped region to have an equal number of positive and negative charge carriers at the threshold voltage before having excess negative charge carriers forming an n bridge between the source and drain allowing electrons to flow.

The Triode Region

For an N-channel MOSFET, the triode region occurs when $V_{GS} > V_{THN}$ and $V_{DS} > 0$, but $V_{DS} < V_{GS} - V_{THN}$, where $V_{GS} - V_{THN}$ is the excess of the gate-source voltage over the threshold voltage and is known as the *overdrive voltage* or *saturation voltage*. In this region, strong inversion has occurred, so there is a channel connecting the source and drain, and electrons can flow from the source to the drain. This case

is shown in Figure 3.9 and the drain current is given by

$$I_D = \mu_n C'_{ox} \frac{W}{L} \left[(V_{GS} - V_{THN}) V_{DS} - \frac{V_{DS}^2}{2} \right], \tag{3.11}$$

where μ_n is electron mobility and W and L are the width and length of the NMOS respectively. The ratio W/L is called the *aspect ratio* of a MOSFET and is the most important design parameter. Figure 3.11a shows the variation of the drain current with V_{DS} for an N-channel MOSFET with the triode mode indicated.

For a PMOS transistor, the drain current in the triode mode is similarly given by

$$I_D = \mu_p C'_{ox} \frac{W}{L} \left[(V_{SG} - V_{THP}) V_{SD} - \frac{V_{SD}^2}{2} \right] \tag{3.12}$$

with the notation $V_{SG} = -V_{GS}$.

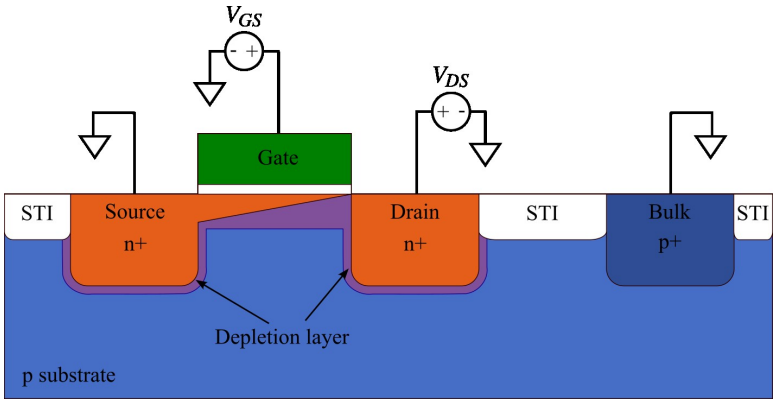


Figure 3.10: Shown is an NMOS operating in the active (saturated) regime. In this case V_{DS} is sufficiently high that negative charge carriers are attracted to the drain which causes the n channel to become pinched off, preventing the drain current from increasing significantly.

The Active Region

As the drain voltage V_{DS} increases to the saturation voltage $V_{D,sat} = V_{GS} - V_{THN}$ and beyond, charge carriers in the channel become attracted to the drain and the channel becomes pinched off as shown in Figure 3.10. At this point, the drain current saturates and is given by

$$I_D = \frac{1}{2} \mu_n C'_{ox} \frac{W}{L'} (V_{GS} - V_{THN})^2, \tag{3.13}$$

where L' is the pinched-off length of the depletion channel. The effect of the modified channel length is known as *channel-length modulation* and is a function of V_{DS}. Introducing the channel-length modulation coefficient λ, which is given by

$$\lambda = \frac{1}{L} \frac{dL}{dV_{DS}}, \tag{3.14}$$

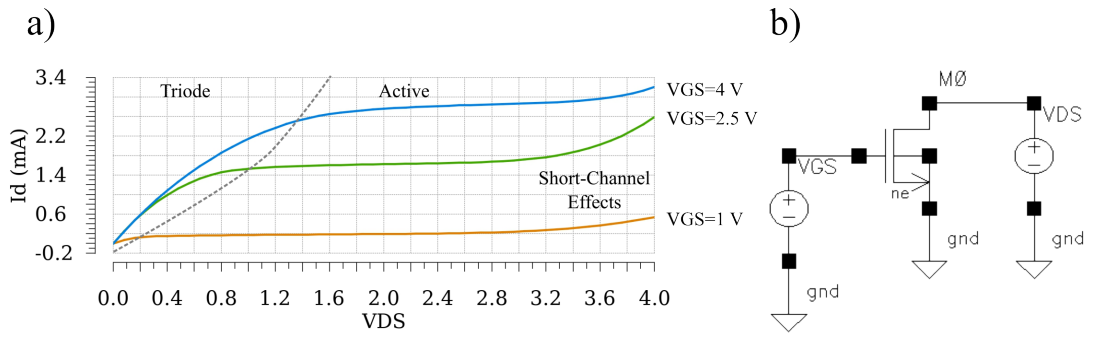


Figure 3.11: a) The drain current, I_d, is plotted with respect to V_{DS} for an NMOS with $V_{GS} = 1$ V (orange), $V_{GS} = 2.5$ V (green), and $V_{GS} = 4$ V (blue). Triode and active modes are indicated by the dotted grey curve. b) The schematic used to simulate the MOSFET.

we can approximate Equation 3.13 as

$$I_D \approx \frac{1}{2}\mu_n C'_{ox}\frac{W}{L}\left(V_{GS} - V_{THN}\right)^2\left(1 + \lambda V_{DS}\right). \tag{3.15}$$

The effect of λ can be seen in Figure 3.11 by the rising drain current in active mode. Assuming a constant channel length, the active drain current would be constant with V_{DS}.

For a PMOS transistor, the active mode drain current is given by

$$I_D \approx \frac{1}{2}\mu_p C'_{ox}\frac{W}{L}\left(V_{SG} - V_{THP}\right)^2\left(1 + \lambda V_{SD}\right). \tag{3.16}$$

3.3.1 Advanced MOSFET Models

As mentioned above, the MOSFET modelling discussed so far is only accurate for long-channel processes with lengths over 4 μm. When channel lengths decrease below this, additional effects such as velocity saturation, variation of threshold voltage, non-uniform doping, and more need to be considered. After the Shichman-Hodges model, the next most complex are the levels 2 and 3 models, which are able to accurately describe devices down to 1 μm. The current industry standards for submicron processes are the BSIM4[14] model for processes down to 20 nm and BSIM3[50] for $>$ 100nm processes such as those used in the design of chip1. Both models were developed by the BSIM group at the University of California in Berkeley.

3.4 Digital CMOS

We first discuss digital design in general terms before focusing on the specific components used in the designs of chip0 and chip1.

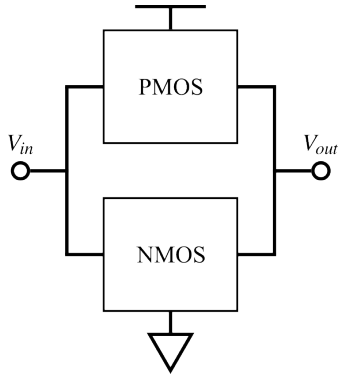


Figure 3.12: In static digital design, circuits are complementary such that either a PMOS network connects the output to VDD or an NMOS network connects the output to GND depending on if the output is a logical 1 or 0 respectively. This reduces static power dissipation by ensuring there are no low impedance paths from VDD to GND.

3.4.1 Complementary Design

As mentioned above, the advantage of CMOS over other MOSFET technologies such as pure NMOS is the complementary nature of PMOS and NMOS transistor blocks. Generally speaking, an input is connected to an NMOS block which is connected to ground and a PMOS block which is connected to the power rail as shown in Figure 3.12. One of these blocks is "off" and the other is "on", connecting the output to either ground or the supply rail, but never both at the same time. This minimises static power consumption and leads to CMOS chips being very efficient compared to other technologies. The PMOS network is often referred to as a pull-up network since it raises the potential at the output; similarly, the NMOS network is referred to as a pull-down network since it lowers the potential.

Another advantage of this design is that it allows the output to reach both power and ground. To clarify this, imagine an NMOS transistor connected to power with the input being the gate and the output then being the source. If the input is set to power and the output is initially low, then the transistor is "on" and current flows from power to the output, increasing the voltage at the output. However, once $V_{out} = V_{dd} - V_T$, $V_{GS} = V_{THN}$ causing the transistor to switch off, and the output to be limited to $V_{dd} - V_{THN}$. As this effect cascades, the high signal can be degraded to a level that causes errors. When a signal is degraded like this, it is said to pass a "weak" 1; alternatively, the NMOS is said to pass a "strong" 0 as it can output a true ground signal.

It might be assumed that having two blocks of logic will complicate the design of digital circuits; however, the pull-up and pull-down networks are complementary, meaning that once one is found, the other can be computed by interchanging AND and OR operations according to De Morgan's Law[43]. In the case of CMOS, this is equivalent to replacing series connections with parallel ones and vice versa, as will be discussed in the following section.

It should be noted that this style of logic, using pull-up and pull-down networks, is called static logic; however, this might not be ideal due to the large area required and the relatively slow perfor-

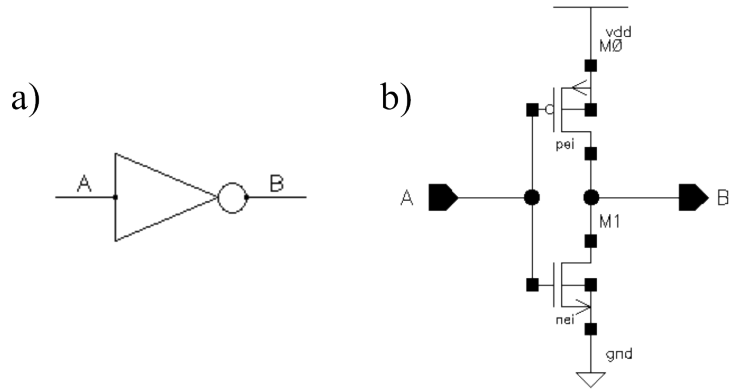


Figure 3.13: a) Circuit symbol for the inverter. b) Inverter schematic.

mance. Alternative systems of CMOS logic include pseudo-NMOS, clocked CMOS logic (C^2 logic), transmission gate logic, and domino logic, all of which have various trade-offs[43,6].

3.4.2 Static Logic

The basic inverter, NAND, and NOR gates which form the AOI (AND, OR, and inverter) logic family are detailed in this section. These gates form the majority of static logic.

The Inverter

The simplest of the logic gates is the inverter shown in Figure 3.13, with Figure a showing the circuit diagram component and Figure b showing the schematic of an inverter implemented in CMOS. The inverter consists of a single NMOS and PMOS connected in series between power and ground. The gates are connected as input, and the drains are connected as output.

The relationship between output voltage and input voltage can be seen in Figure 3.14, with VDD = 3 V, where we see that an input of 0 V results in $V_{GS,\mathrm{NMOS}} = 0$ V and $V_{SG,\mathrm{PMOS}} = 3$ V. This results in the NMOS being "off" since 0 V is less than the threshold voltage, thus providing a high resistance path to ground; conversely, the PMOS is "on" since VDD is greater than the threshold and there is a low resistance path to VDD. When both are partially on, the current can flow through both transistors, resulting in an intermediate voltage at the output.

For digital applications, the inverter is typically used with an input of either VDD or ground; however, for analogue applications, the region with the highest gain, where both transistors are on, is often used. It is also important to note that for a symmetric transfer function, with the midpoint being at VDD/2, the PMOS transistor should have approximately twice the width of the NMOS (depending on the exact technology) due to hole mobility in silicon being approximately half the electron mobility.

Inverters have a high impedance input (at low frequencies) due to the gate connections which act

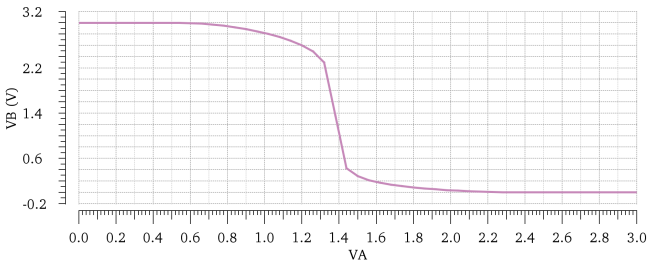


Figure 3.14: The transfer function for a CMOS inverter. VA is the voltage at the input and VB is the voltage at the output.

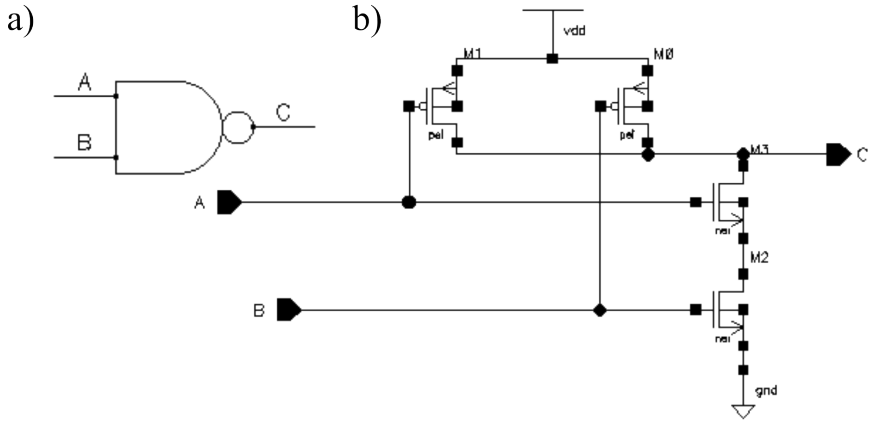


Figure 3.15: a) Circuit symbol for the NAND gate. b) Schematic for a CMOS implementation of a NAND gate.

as capacitors, and a low impedance output because of the low resistance of the transistor which is active. Due to this, they are often used as buffers, normally seen in pairs to not invert the signal. However, they are also often seen as basic amplifiers in analogue applications[6].

NAND Gate

If a gate has multiple inputs, they can be compared through the AND and OR functions; however, in CMOS it is much more natural to use the inverted NAND and NOR functions, as these gates are extensions of the inverter design. A 2-1 NAND gate (2 inputs, 1 output) is shown in Figure 3.15 with Figure a showing the circuit diagram symbol and Figure b showing the schematic.

Extending the principles of the inverter, we can see that if A or B are 0 V, then at least one of the NMOS transistors is "off", providing a high impedance on the path to ground. However, as long as one is low, then there is a low impedance route to power, and the output is high. In the case that both A and B are VDD, then there is one long low impedance path to ground and two high-impedance paths to the power rail, resulting in a low output. Thus, we find that the output is always high unless both inputs are high, providing the inverted AND gate.

This principle can be extended to an almost arbitrary number of gates by simply adding more NMOS transistors in series and more PMOS transistors in parallel. This is also where we first start to

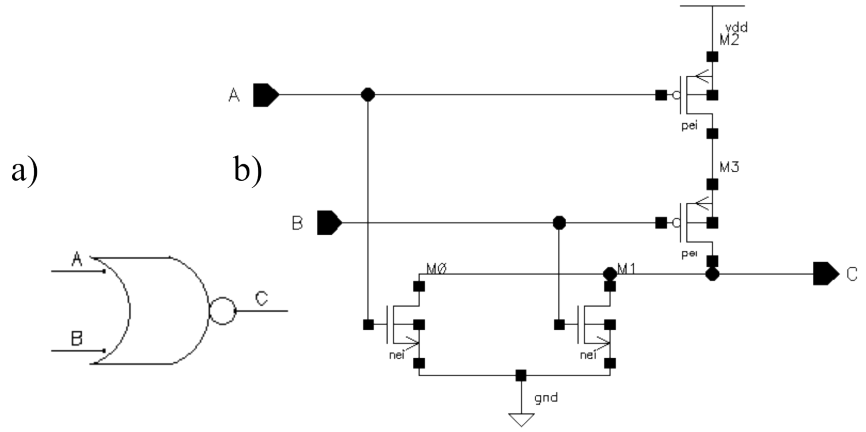


Figure 3.16: a) Circuit symbol for the NOR gate. b) Schematic for a CMOS implementation of a NOR gate.

see the idea of duality mentioned earlier with De Morgan's law, where for every NMOS pair connected in series, the corresponding PMOS pair is connected in parallel.

NOR Gate

The NOR gate is constructed very similarly to the NAND gate, but with the NMOS pairs in parallel and the PMOS pairs in series, as shown in Figure 3.16b. In this case, when the inputs A and B are both 0V, the NMOS transistors are set to a high impedance while both PMOS transistors are on, providing a low impedance path to the output from VDD. Thus, the output is high. However, in the case that either A or B is high, there is a low impedance path to ground and a high impedance path to VDD, causing the output to be low.

Transmission Gate

A somewhat adjacent, but equally important, gate is the transmission gate shown in Figure 3.17. This gate features parallel PMOS and NMOS transistors and behaves as a switch to either pass signals using a low impedance or block them with a high impedance. A digital switch signal S is used to control the gate voltage of the NMOS and its complementary signal Sn, produced by an inverter, controls the PMOS, such that both are "on" or "off" at the same time. Note that n in Sn is a common means of denoting an inverted signal.

The bodies of the PMOS and NMOS transistors are tied to VDD and GND, respectively, since the source of the transistors can flip between the two sides, resulting in the body sometimes being connected to the drain. Both PMOS and NMOS transistors are used to allow the passing of strong low and high signals.

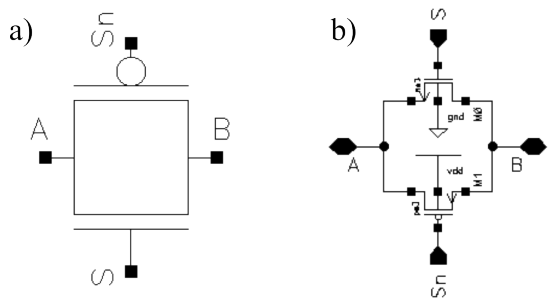


Figure 3.17: a) Circuit symbol for the Transmission gate controlled by the digital signal S and its inverted value Sn. b) Schematic for a transmission gate.

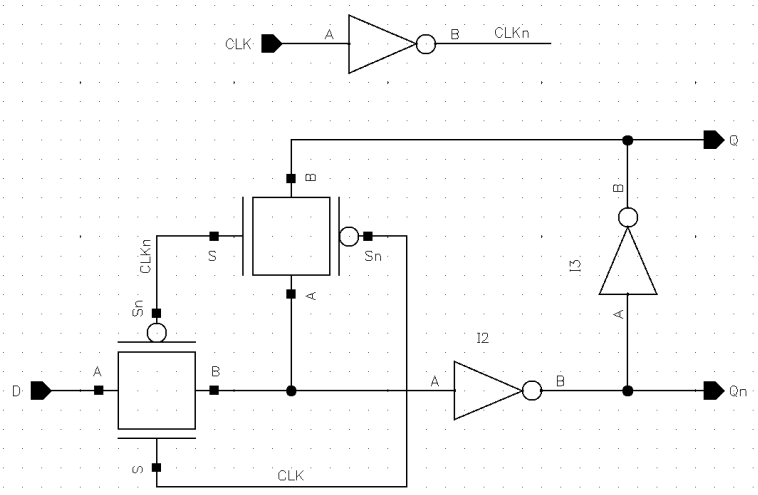


Figure 3.18: Schematic of a CMOS latch.

3.4.3 Sequential Logic

In order to prevent timing errors and keep track of the order of operations, circuits are typically clocked so that each stage happens in one period, with outputs passed to the next stage at the beginning of the next period. To do this, the signal is stored in a circuit known as a register where the input is only passed to the output at either a rising or falling edge on the clock signal. There are several different designs for this, but here we use a D-Flip-Flop based on a double latch configuration.

A typical latch is shown in Figure 3.18 and is based on a pair of transmission gates with opposite clock signals and a pair of inverters to latch the signal. The initial clock signal is inverted to produce a complementary signal called CLK_n, then CLK and CLK_n are used as the S and S_n control signals for a transmission gate, respectively, allowing the input signal D to pass whenever the clock signal is high. When the clock signal goes low, the second transmission gate is opened, forming a loop with the two inverters. A pair of inverters forming a loop in this way then latches whichever signals are present since any changes to these are strongly resisted by both nodes being connected to either VDD or GND. When the clock signal is in the opposite phase, the transmission gate breaks this loop, allowing the signal to be altered. The effects of this circuit are demonstrated in Figure 3.19 where we can see that whenever the

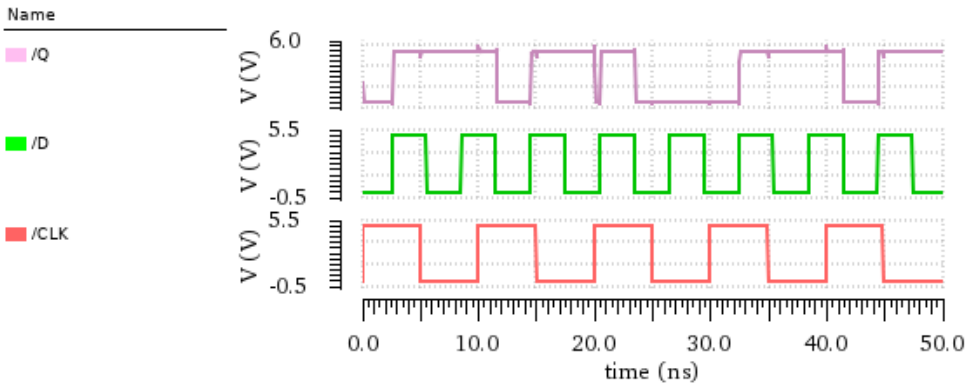


Figure 3.19: Output of a transient simulation of a latch. The inputs are the clock signal (red) and the data signal, D (green), with the corresponding output, Q (pink). Whilst the clock signal is high, the Q follows D, but retains its current value when the clock is low.

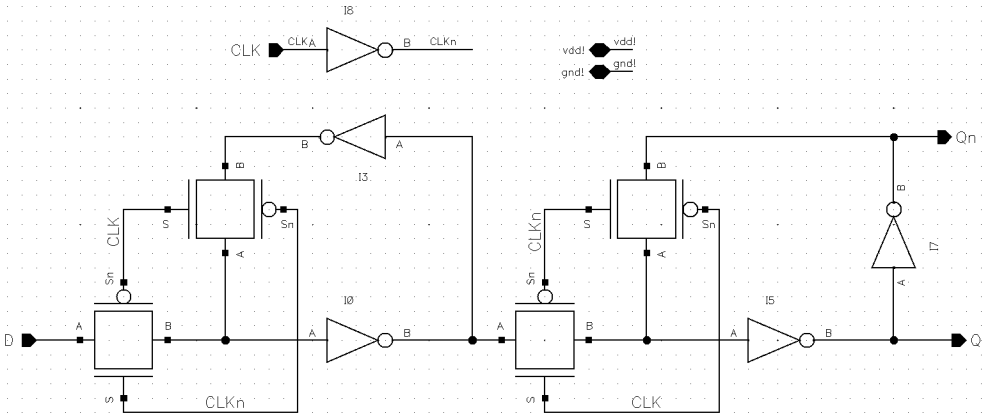


Figure 3.20: Circuit for a D-flip-flop.

clock signal is high, Q directly follows the input D since the latch is broken by the closed transmission gate, but when the clock signal is low the signal is latched and Q is maintained.

However, we only want the output of a stage to be passed during the rising edge of a clock signal to prevent intermediate values being passed. To do this, two latches can be combined as shown in Figure 3.20 which are set so that the first latch passes the input D whenever the clock signal is low and the second latch passes the signal when the clock is high, latching on the high and low phases respectively. Thus, as shown in Figure 3.21, at a rising edge of the clock signal the input is latched by the first latch and passed by the second. Importantly, the second latch latches the output during the low phase so that the output remains constant whilst the input is passed into the first latch and so only during the rising edge of the clock signal is the output changed. This circuit is known as a D-Flip-Flop (DFF) or register and acts as a buffer between steps in a circuit.

3.4.4 Digital Components

In this section, several digital components used in the design of chip1 are introduced.

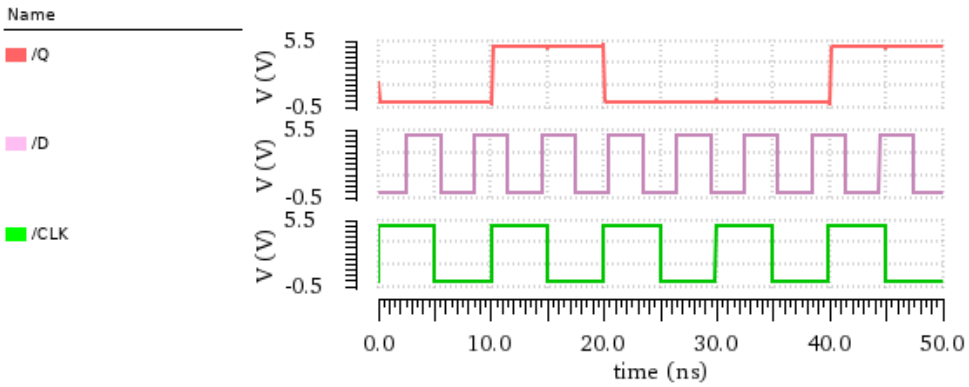


Figure 3.21: Output for the D-flip-flip where D (pink) is an arbitrary input, in this case a square wave and Q (red) is the output. The clock signal is shown in green. Only when the clock transitions from low to high does the Q change to match D.

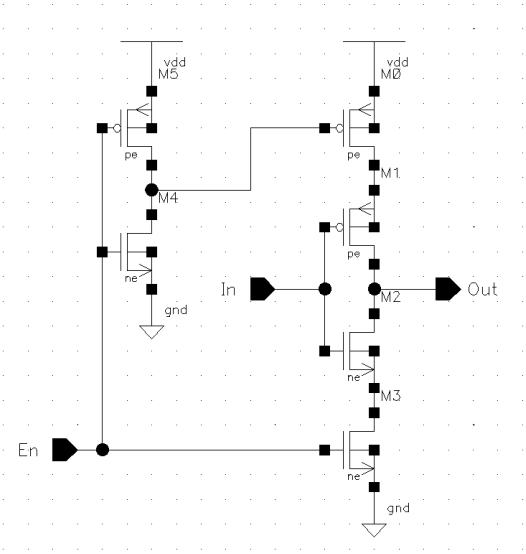


Figure 3.22: Circuit for producing a tri-state buffer where the output is floating if En is low.

Tri-State Buffers

Often, at the end of a sub-circuit, inverters are used as buffers since they have a high impedance input, a low impedance output, and a high gain, which prevents errors caused by intermediate voltages. However, the output is always connected to either VDD or GND through a low impedance, so if multiple inverters are connected to a single node they will fight each other, drawing current and producing an intermediate output. In some cases, such as the data bus on chip1, the outputs of several circuits must be connected to the same wire. It is therefore necessary for us to be able to set the output to floating on all but one of the outputs. This can be done with a tri-state buffer, such as that shown in Figure 3.22, where an inverter is nestled inside a second inverter. If the En pin is high then the inverter behaves like a normal inverter. If the En pin is low, the outside MOSFETs, M0 and M3, are "off" and the output is floating regardless of the value of In. A transmission gate could also be used to produce a floating output.

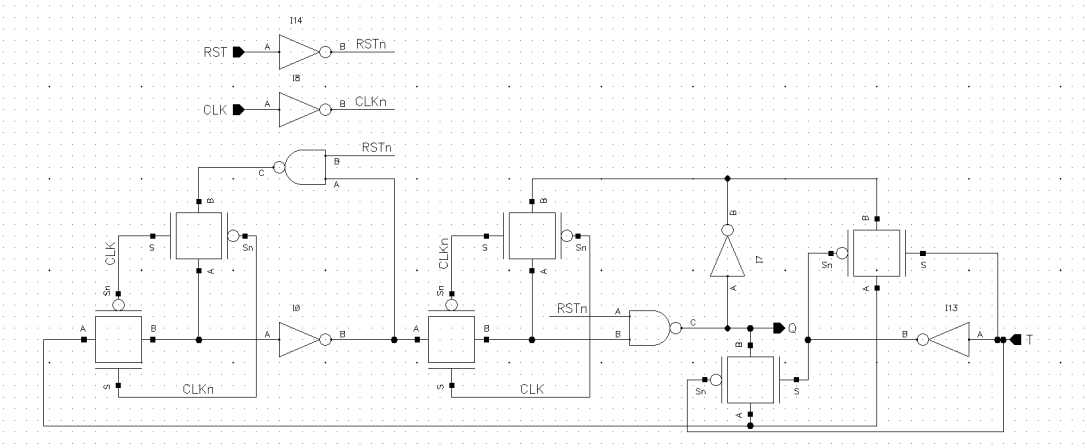


Figure 3.23: Circuit of a toggle-FF.

Multiplexers

Multiplexers enable several inputs to connect to a single output. There are several implementations, with the most common using either a transmission gate or tri-state buffer connecting each input to the output. Digital select ports then allow each gate/buffer to be activated individually by decoding the select signals using static logic.

T-Flip-Flop

Other types of flip-flops can be produced through modification to the DFF such as the toggle flip-flop (TFF), where the output of the flip-flop is toggled to the opposite value at every rising clock edge if the toggle input is high, and does nothing if it is low. Firstly, in Figure 3.23, the DFF has been modified by replacing one inverter in each latch with a NAND gate which allows us to add a reset pin (RST) to reset the current value of the DFF to GND, regardless of the input. We can then simply modify the output to toggle with every rising clock edge by connecting the inverted Qn output to the input. Conversely, we can disable the toggling by connecting the non-inverting output Q to the input. Thus, we can enable or disable the toggling by introducing transmission gates which switch the feedback loop to the input using the toggle input T. It should be noted that the DFF can be converted into any flip-flop using modified feedback so many component libraries (supplied by the foundry) only provide the DFF.

Counter

One of the more commonly used circuits in chip1 is the counter circuit. For example, this is used in the single-photon detector in chip1, to count individual photon detection events, and in the serial interface to determine when a phrase is complete. A counter is described by the number of bits such that a 2-bit counter is able to count to 4 and a 4-bit counter is able to count to 16. Shown in Figure 3.24 is an implementation of a 4-bit counter using TFFs where the output of each TFF is one bit of the output. In

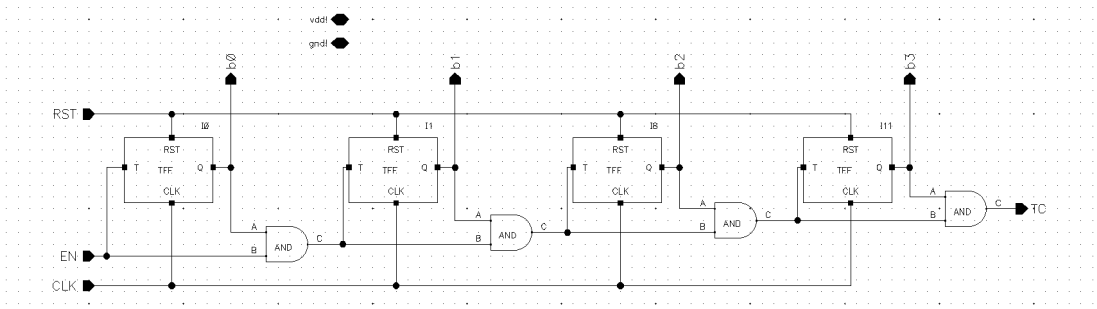


Figure 3.24: Circuit of 4-bit counter circuit

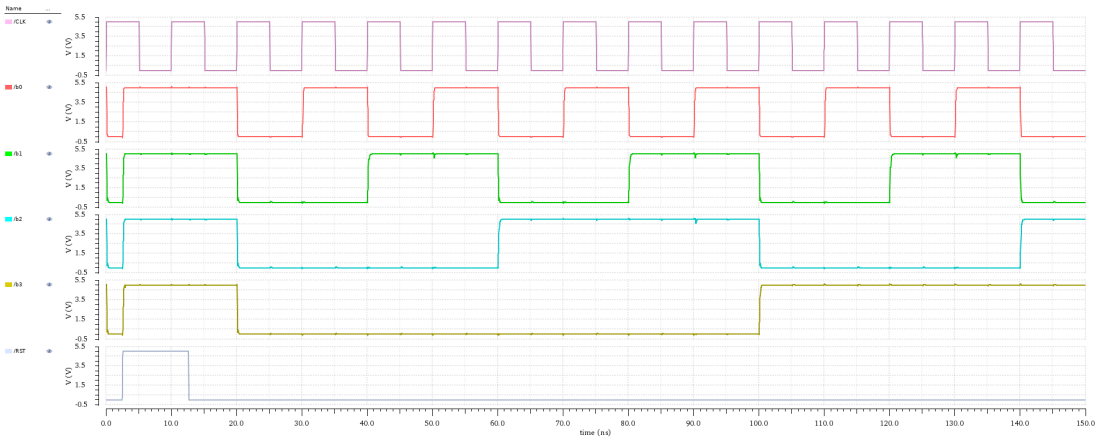


Figure 3.25: Simulation of a 4-bit counter circuit. The top trace (pink) is the clock signal and the bottom trace (grey) is the reset signal. Traces 2 to 5 are the outputs b0 (red), b1 (green), b2 (blue) and b3 (orange).

this circuit, all of the reset pins are tied together so that the counter can be reset at any time, and the clock signal (alternatively the signal that we wish to count) is connected to every TFF. When the enable (EN) pin is high, the circuit begins to count.

Figure 3.25 shows the simulation of the 4-bit counter, which is initially reset to 0. At the first rising edge of the clock signal, the output of the first TFF is toggled, bringing the output from 0 to 1. However, due to the AND gates at the inputs of the remaining TFFs, they are unaffected. With the second rising edge, the second TFF is toggled in addition to the first, bringing the count to 2. For the third TFF, the AND gate is connected to the toggle and output pins of the second TFF, so this is only toggled when both the first and second TFFs are high, therefore only triggering every 4 clock periods. Similarly, the fourth TFF is only toggled every 8. This process can then be extended to count to an arbitrary number of bits.

Adders

Adding circuits are a necessary part of the direct digital synthesiser[1] (DDS), where changes in the phase of the output wave are determined by adding a phase step to the current DDS phase, once per clock

[1] See section 6.3 for details and an introduction to direct digital synthesis.

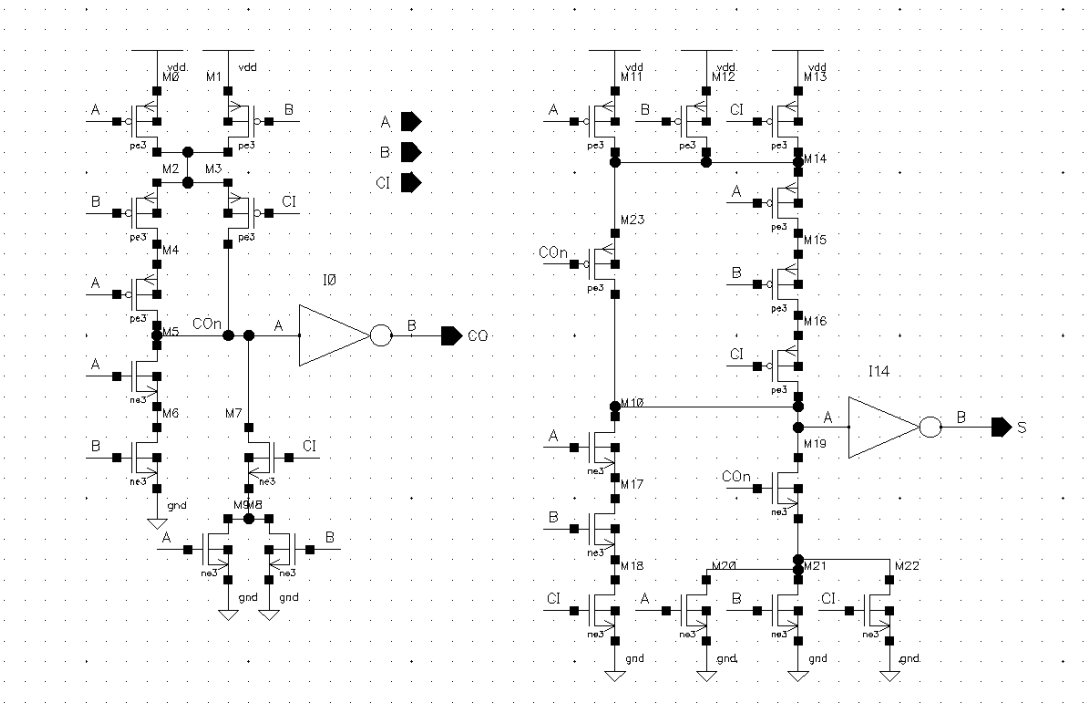


Figure 3.26: A full-adder circuit which adds two bits A and B and a carry over bit CI. The circuit on the left calculates the carry over bit for the addition whilst the circuit on the right calculates the sum.

cycle. A half-adder is a circuit which computes the sum of two bits A and B, outputting both the sum and carry-over. However, this can be extended to a full-adder by also including a carry-over bit in the input, which allows full-adders to be combined such that an arbitrary number of bits can be added together, allowing larger sums to be carried out. Figure 3.26 shows the circuit of a full-adder, which can be split into the left-hand circuit, which computes the carry-over, and the right-hand circuit, which calculates the sum.

3.5 Analogue CMOS

Whilst digital electronics are important for computation and communication, to be able to interface with the atoms we require analogue signals, which are used for sensing and magnetic field generation. The most common analogue parts found in CMOS are oscillators, phase-locked loops, current/voltage sources, and various types of amplifiers such as op-amps, buffers, and drivers. For the requirements of atom chips, oscillators are necessary for clocks, RF, and MW signals, amplifiers are needed to amplify sensor outputs and drive trapping wires, and phase-locked loops are essential for frequency stabilisation. For digital design, we consider maximal voltage swings between VDD and GND where MOSFETs are either off or active. However, analysing analogue circuits by considering the *large-signal* behaviour of MOSFETs is often intractable, so analysis is done by assuming all MOSFETs to be biased in one of the three modes, with the signal treated as a small perturbation about this operating point. This

type of analysis is referred to as *small-signal analysis* and is essential for deriving characteristics such as frequency response, gain, and resistance. Whilst it might initially seem short-sighted to simplify analysis in this way, circuits are typically designed to bias MOSFETs in either triode or active mode such that the small-signal analysis is valid for most inputs. In the following section, some useful small-signal MOSFET parameters are introduced. Comprehensive introductions to analogue circuit design can be found in Baker[6] and Razavi[63].

3.5.1 Analogue MOSFET Parameters

If $V_{DS} \ll 2(V_{GS} - V_{THN})$ then the N-channel MOSFET is in the deep triode region, and I_D is approximately linear with respect to V_{DS}. In this case

$$I_D \approx \mu_n C'_{ox} \frac{W}{L}(V_{GS} - V_{THN})V_{DS} \tag{3.17}$$

and the behaviour of the MOSFET is analogous to that of a variable resistor with resistance

$$R_{\text{on}} = \frac{1}{\mu_n C'_{ox} \frac{W}{L}(V_{GS} - V_{THN})} \tag{3.18}$$

which can be adjusted by the overdrive voltage.

In saturation, MOSFETs can be used as a current source with current I_D which is controlled by the overdrive voltage. A useful parameter in this context is the *transconductance* g_m defined as

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \mu_n C'_{ox} \frac{W}{L}(V_{GS} - V_{THN})(1 + \lambda V_{DS}) \tag{3.19}$$

which can also be expressed as

$$g_m = \sqrt{\frac{2\mu_n C'_{ox}(W/L)I_D}{1 + \lambda V_{DS}}} = \frac{2I_D}{V_{GS} - V_{THN}}. \tag{3.20}$$

The transconductance is a measure of how effective the MOSFET is at translating a change in the overdrive voltage to a change in drain current.

3.5.2 Analogue Components

Current Mirrors

One of the most essential components of analogue circuit design is the current source. The simplest implementation of a current source is a MOSFET in saturation; however, any instability in the bias voltage leads to instability in the current. A better way of doing this is to have one stable reference where the current is then copied wherever needed, which can be implemented using the current mirror

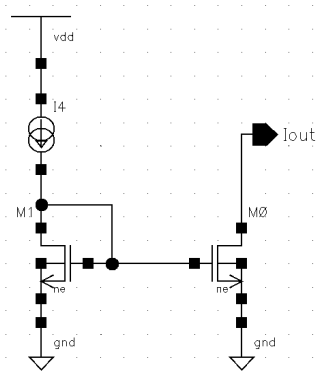


Figure 3.27: Circuit for a current mirror.

circuit in Figure 3.27. With both NMOS acting in saturation

$$I_{\text{ref}} = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_1 (V_{GS} - V_{THN})^2 \tag{3.21}$$

$$I_{\text{out}} = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_0 (V_{GS} - V_{THN})^2 \tag{3.22}$$

leading to

$$I_{\text{out}} = \frac{(W/L)_0}{(W/L)_1} I_{\text{ref}}. \tag{3.23}$$

which is solely dependent on the device dimensions, which are well controlled.

Voltage Follower

The voltage (or source) follower is a circuit which is primarily used as a buffer at the output of a circuit due to its high input impedance and moderate output impedance. The circuit for a voltage follower is shown in Figure 3.28, where Figure a shows the voltage follower with biasing from a current source, and Figure b replaces the ideal current source with a current mirror.

A DC simulation of the circuit in Figure 3.28b shows the output voltage, drain current of M3 and resistor current as a function of Vin. Whilst Vin is below V_{THN} for M1, the output is zero. However, once Vin$> V_{THN}$ current is able to flow and both M1 and M3 quickly saturate, at which point the output voltage increases linearly. The primary disadvantages of the voltage follower are the non-linearity at low input voltages and the loss of voltage headroom since the highest output voltage is VDD-V_{THN}. Although this loss in voltage actually makes the voltage follower useful for changing voltage levels.

3.6 Design Flow

In this section we shall discuss the steps involved in designing a CMOS chip based on the standard design flow within CMOS and experience from chip1. Figure 3.30 summarises the design flow. We start with the requirements for the chip, which in the case of an atom chip may involve the trapping

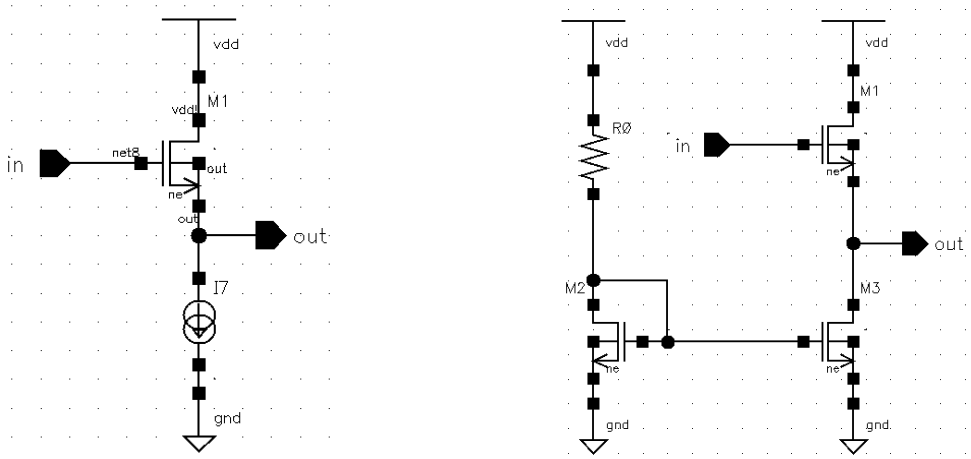


Figure 3.28: a) Circuit for a voltage (or source) follower biased with an ideal current source. b) The current source is replaced with a current mirror.

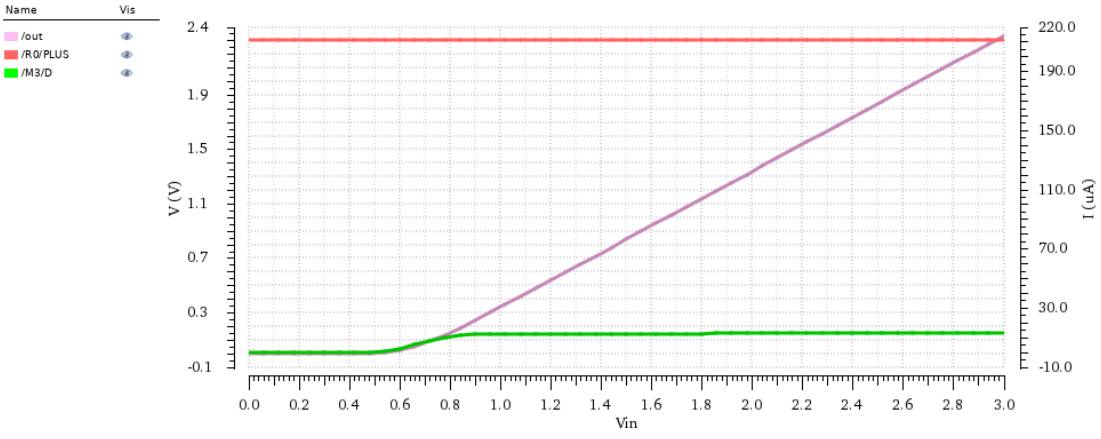


Figure 3.29: Simulation results for the voltage follower biased with a current mirror given in Figure 3.28b. The resistor current is given in red, the the drain current of M3 in green and the output voltage in pink. The simulation is a DC scan of Vin.

wires, RF and MW sources, photodetectors and serial interface. Further to this we need to know the specific requirements such as the RF frequency range, amplitude range and polarisation. From here high level block diagrams can be made to show how everything might be interfaced and what circuits are required for each component. For example, the RF source is a DDS which must be able to produce an output from ~500 kHz to ~20 MHz, and involves a phase-accumulator, registers, phase-to-amplitude converter, complementor, and digital to analogue converter (DAC). This process is repeated until the blocks are described at the transistor level. So far we have focused on the planning side of design as shown in Figure 3.30 where we use a top down approach. However, for the actual design a bottom up approach works best for the design of atom chips since many of the components require physics simulations and analogue components where it is best to quickly figure out what is feasible so that any design modifications can be made.

At this point, the design process splits for an atom chip as many of the components, such as antennae, photodiodes, and trapping wires, have to be simulated in other software such as COMSOL Multiphysics or MATLAB. From these simulations, we produce layouts and circuit models that approximate the device using current/voltage sources and simulated impedance, and must be used in the circuit simulations. The CMOS design begins with schematics, where literature and hand calculations are used to find the starting architecture and MOSFET sizes, and simulations are performed iteratively to aid in the design. This process will be detailed below. Once the design has been finalised, it must be converted into a physical layout where the masks for each layer are defined and is verified by three different checks called the design rule check (DRC), the electrical rule check (ERC) and layout versus schematic check (LVS). From here, physics simulations calculate parasitics and the layout can be modified to remove these. Typically, this process is done in blocks starting with the simplest of sub-circuits such as flip-flops and amplifiers, which can then be combined in a modular fashion to build up to the final chip.

3.6.1 Schematic and SPICE

In practice, hand calculations are only used to gain a qualitative understanding of circuits since analytical calculations of circuits, including modern MOSFETs, are intractable. To gain a more accurate understanding, a simulation tool known as *SPICE*, or "Simulation Program with Integrated Circuit Emphasis", is used. SPICE is an open-source simulation tool for modeling all electronic circuits; however, the most recent official release was in 1993, so alternatives such as the free LTSPICE from Analog Devices or HSPICE and SPECTRE from Cadence Design Systems are more commonly used. The original SPICE1 was developed at the University of California, Berkeley, during the 1970s and was based on the tool "Computer Analysis of Non-linear Circuits, Excluding Radiation" (CANCER), the final clarification being added to emphasize that it is not to be used for the development of nuclear weapons. SPICE models of MOSFETs use the BSIM models introduced above.

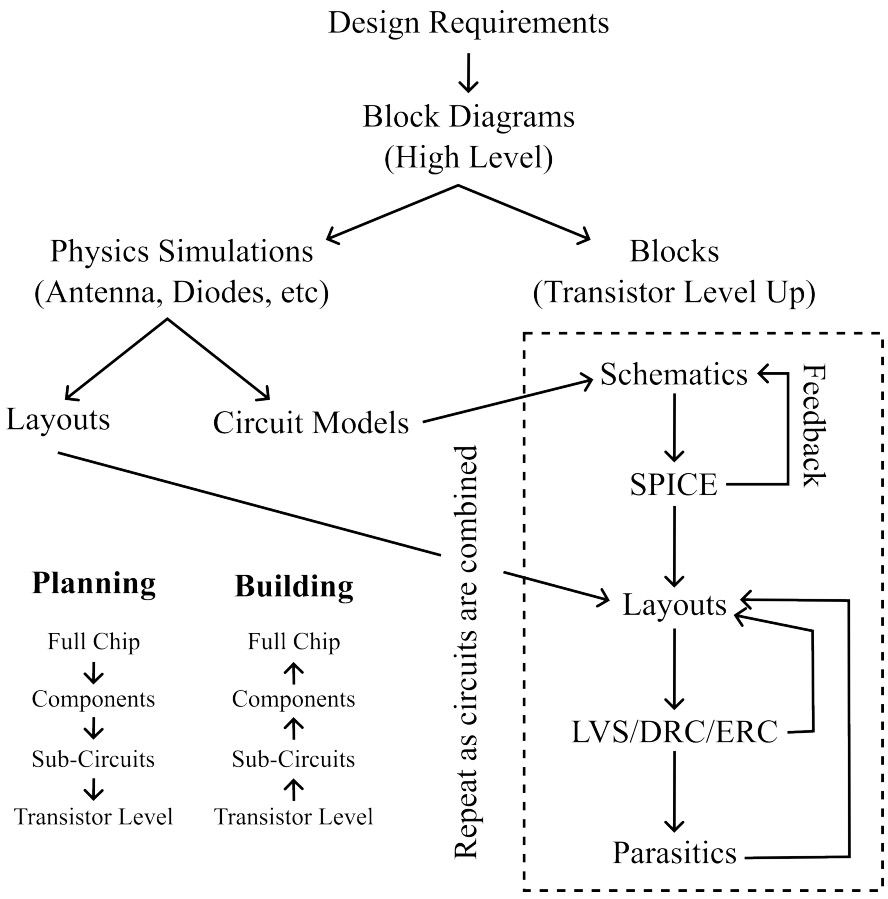


Figure 3.30: This flowchart summarises the design process for a CMOS atom chip. Planning starts from the full chip level, focusing on the design requirements and working down to decide which components are required. Circuit design on the other hand starts at the transistor level and builds up to the completed chip.

Once a schematic has been verified to not have any open or short circuits, the first step of a simulation is to select the type of analysis to be performed. The most common types of analysis are:

1. Operating Point (.op): Bias point parameters are found with all variables static.
2. DC (.dc variable start stop stepsize): An operating point analysis is performed with a scanned variable.
3. AC (.ac dec/lin/oct npoints fstart fstop) This is used to find the frequency response of a circuit by scanning the frequency of a since wave input around a bias point.
4. Transient (.tran tstep tstop) The analysis is performed with time being the scanned variable to analysis how the circuit changes with time.

Once the analysis has been selected, the schematic must be converted into a format which is understood by code and is known as a *NETLIST*. This also includes any initial values and inputs. Following this, the analysis is performed and the outputs are plotted, which can be used to improve the circuit design.

3.6.2 Layouting and Validation

Once a schematic has been completed, it must be converted into physical layers. Each process has a series of *design rules* which dictate how these layers can be placed. For example, the minimum poly width, the minimum separation, and the maximum metal area. Some of these rules are a result of limitations in the fabrication process, such as the minimum poly width, whilst others, such as the maximum metal area, are to prevent potential damage to the chip. Specifically, metal areas connected to gates have to be limited in size due to the *antenna effect*, where ions accumulate at the metal during etching in the fabrication, raising the gate voltage and causing the oxide layer to break down.

In order to verify the layout, there are three essential checks which are performed. Firstly, the *electrical rule check* (ERC) validates the connectivity of the circuit. This is where problems such as floating nodes, shorts, and opens are detected. Secondly, the *layout versus schematic check* (LVS) compares the layout with the schematic to verify that the circuit matches. Lastly, the *design rule check* (DRC) is specific to the fabrication process and verifies that the layout obeys the design rules. Further to these essential checks, optional parasitic simulations can be performed to help minimize parasitic effects throughout the layout.

3.6.3 Floorplanning and Wire-Bonding

Floorplanning refers to the overall layout of the chip and considers how power and signals are distributed throughout the chip. This becomes especially important in atom chip design, where minimizing noise is essential. At the small scale, layouts are arranged into *standard cells* with fixed heights.

Chapter 4

Experimental Setup

In this chapter, we discuss the experiment which has been built to test the CMOS atom chips with atoms. Although the experiment was designed to work generally with any chips, it was primarily designed around the first chip which was built prior to my involvement in the project and is known as chip0.

4.1 Chip0

Chip0 was designed at FHWN by Philip von Neumann[55], Dr. Christian Koller, and Dr. Alexander Nemecek. This fabrication run was funded by Europractice as part of a design challenge with the aim of investigating integrated photodiodes. Consequently, use as an atom chip is a secondary feature, and the Austrian-Microsystems (AMS) 180 nm technology used was not specifically selected for this work. The physical chip consists of three separate devices, chips A, B, and C, as shown in Figure 4.1a. Chips A and B are for testing individual components, with A featuring an ADC and B consisting of several photodiodes with different geometries. These devices are purely for testing at FHWN and are not used in this thesis. .

Chip C is the relevant half of the chip and is shown in more detail in Figure 4.1b. It consists of a Z-wire with two sets of arms: passive arms forming a longer Z-wire separated by 1700 μm and highlighted with a solid yellow line, and switchable arms forming a shorter Z-wire, separated by 950μm and shown with a solid orange line. These switchable arms are controlled by transmission gates and the trapping wire has a consistent width of 100 μm, using a thick top metal layer of height 2.5 μm. To either side of the Z-wire are two sets of four square photodiodes that are shown bordered with white lines and have a width of 199 μm. The white squares around the outside of the chip are 100 μm square metal pads used in wire bonding and appear white from reflecting the microscope light. Other metal areas do not reflect as much light due to the oxide insulation and protection layer at the top of the chip, which attenuates and diffuses light in the visible spectrum. The metal squares throughout the chip are capacitors.

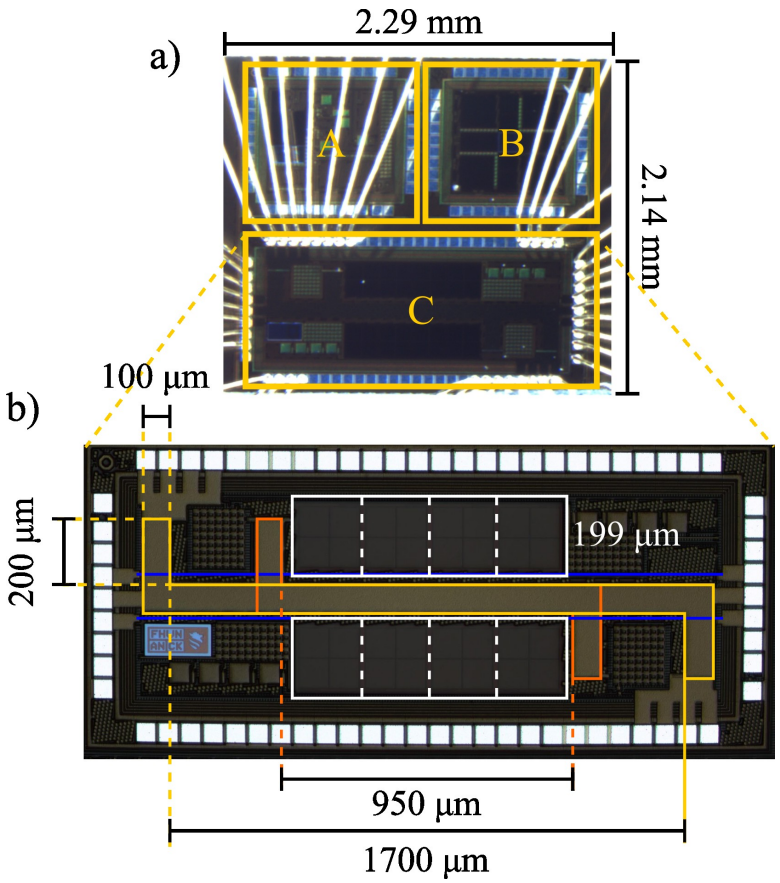


Figure 4.1: a) Shows a photograph of the top of chip0 with each of the three independent devices A, B and C labelled. A features an ADC, B includes several different photodiode geometries with amplifiers and C is the atom chip. b) A higher resolution image of chip C with the long Z-wire highlighted by a yellow solid line, the shorter Z-wire highlighted with an orange dashed line and the photodiodes surrounded with a white box. The white squares are the Al bonding pads reflecting light.

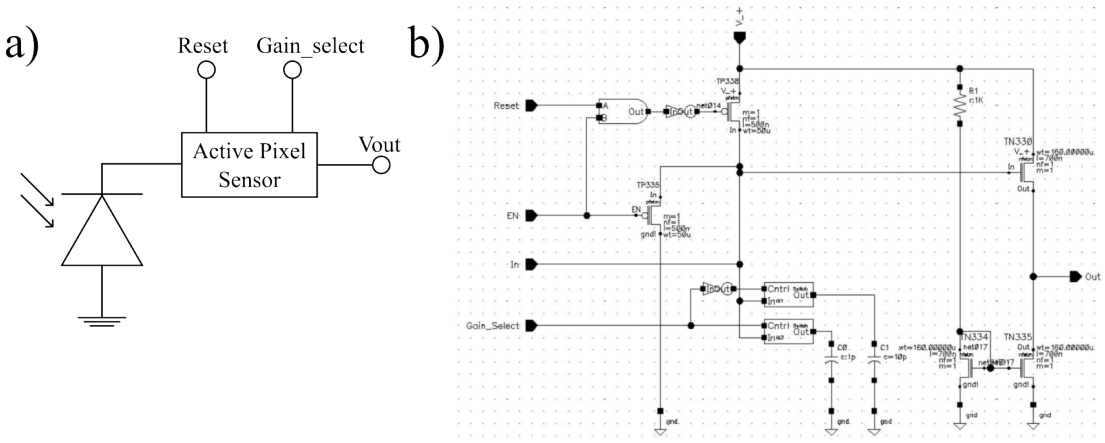


Figure 4.2: a) Shows a block diagram for the photodetector design in chip0 which is detailed in[55]. The photodiode is reverse biased with the cathode connected to ground and the anode connected to an active pixel sensor. The active pixel sensor has two inputs, Reset and Gain_Select and an output voltage. b) Shows the schematic for the active pixel sensor. In this circuit a capacitor (there are two and the capacitors are selected between using Gain_select) is charged to VDD when Reset is high and discharged by the photocurrent generated in the photodiode when Reset is low. The rate of discharge is determined by the light intensity and the capacitor voltage is measured at the output with a source follower acting as a buffer. The En pin is able to enable and disable the circuit.

4.1.1 Photodiodes

The photodetectors on chip0 use an active pixel sensor architecture to measure the photocurrent generated in the photodiodes, thus the incident light intensity. A block diagram for the scheme is given in Figure 4.2a with a reverse-biased photodiode acting as a current source, connected to the active pixel sensor, which is shown in Figure 4.2b. In the active pixel sensor, a capacitor is charged to VDD when both the Reset and EN pins are high, at which point the anode of the diode is at VDD and the cathode is at ground, resulting in the reverse bias necessary for photocurrent generation. If the reset pin is set low, the capacitor is discharged by the diode current. The voltage across the capacitor is measured at the Out pin via a source follower. The source follower has a high input impedance, preventing a low impedance at the output from affecting the charge across the capacitor and impacting the measurement. The bias current for the source follower is generated by a current mirror with a resistor. Lastly, there are actually two capacitors, a 1 pF and a 10 pF capacitor, which are in parallel and are isolated by transmission gates. Using the Gain_Select digital signal, the transmission gates can be alternately activated to select either capacitor. This allows the gain of the active-pixel sensor to be modified to optimise for the light intensity involved in the measurement.

The performance of the photodiodes was measured by Bianca Giacomelli at FHWN and the results are shown in Figure 4.3. The primary characteristic of performance being the responsivity, R_{pd}. This quantifies the efficiency of the photodiode to convert incident optical power into photocurrent and is given by

$$R_{\text{pd}} = \frac{I_{\text{pd}}}{P_{\text{op}}} \tag{4.1}$$

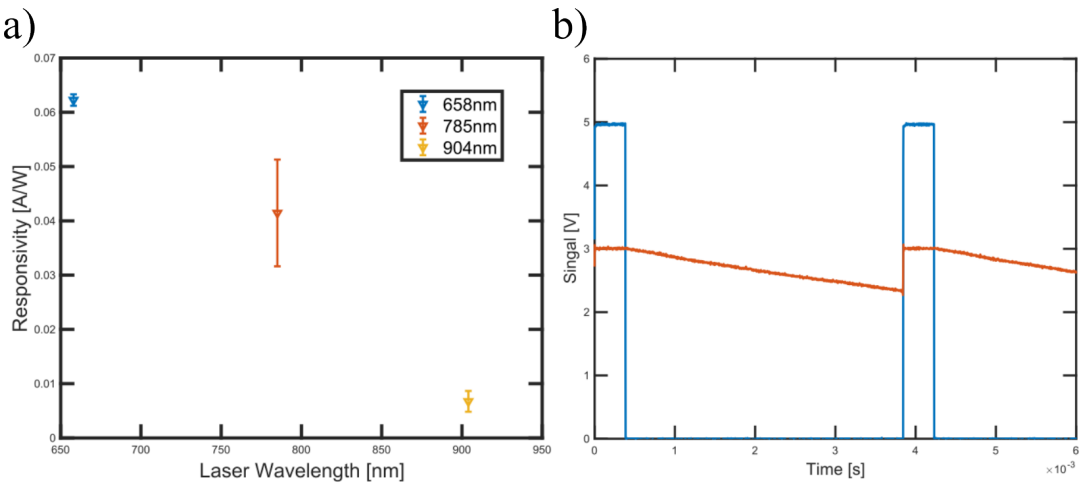


Figure 4.3: Measured performance of chip0 photodiodes. a) Responsivity spectrum of the photodiodes. b) Measured voltage across active pixel sensor 1 pF capacitor (red) incident with 0.89 mW of 658 nm light and the reset signal (blue). (Courtesy of Bianca Giacomelli)

where I_{pd} is the produced photocurrent and P_{op} is the optical power at a specific wavelength. Since the photocurrent is inferred through the measurement of a voltage, V_C, across a capacitor, it is calculated as $R_{\text{pd}} = \Delta V_C C_{\text{aps}} / P_{\text{op}}$ where C_{aps} is the capacitance of the capacitor. To measure this, the photodiodes were reverse biased with the maximum voltage of VDD = 5 V, subjected to a 260 MHz reset signal, and the analogue voltage drop across the 1 pF capacitor was measured. The laser light was produced by a Multi-Channel Fibre-Coupler Laser Source[1] and the power was measured.

Figure 4.3b shows the reset signal (blue) and the measured voltage drop (red). At higher powers and longer reset intervals, the capacitor will discharge exponentially; however, at the beginning, this can be approximated as a linear decay and allows us to estimate the light intensity from the gradient. The power incident on the photodiode was measured directly so that the responsivity could be calculated and is shown in Figure 4.3a for three wavelengths, 658 nm, 785 nm, and 904 nm. The responsivity is determined by the doping concentrations, profiles, and depths of the photodiodes, which are characteristics of the fabrication procedure and usually cannot be changed. Unfortunately, for 780 nm, the photodiodes on chip0 show fairly poor responsivity, but this can be changed by selecting a different CMOS technology, ideally one that allows doping to be modified.

4.1.2 Z-Wire Simulations

Before simulating the Z-wires, the maximum current in a vacuum must be measured. A four-wire measurement was used to record the resistance variation as the wire heats up and was performed by Felix Wenzl (FHWN) as part of his master thesis[88]. The results for a continuous current are shown in Figure 4.4 where the voltage across the central bar of the Z-wire was recorded with a continuous current and is measured in vacuum.

[1] MCLS1, ThorLabs

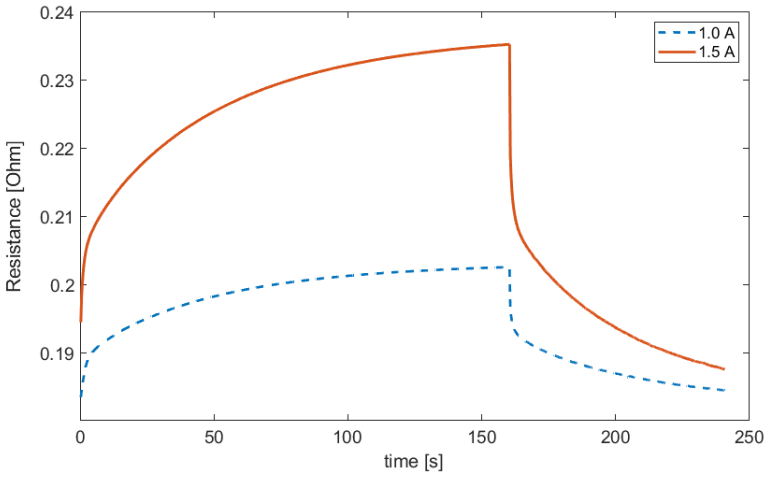


Figure 4.4: Wire heating measurements of the long Z-wire for 1 A (blue,dotted) and 1.5 A(red, solid).

To better understand the CMOS chip trap, it was simulated in COMSOL Multiphysics using the dimensions of the original chip and a current of 1.5 A through the long Z-wire. The simulated system is illustrated in Figure 4.5, where Figure 4.5a gives an artist's impression of the shape of the atomic cloud relative to the chip and defines the axis used in the simulations, the origin of which is set as the center of the Z-wire, at the surface of the metal. A trap is formed at a distance of z_0=200 μm below the chip with a bias field of 15.36 G (a bias field of 10.24 G is required per amp of current to maintain a trap at z_0) and a cross-section of the magnetic field in the x-y plane is shown in Figure 4.5b. Figure 4.5c shows the variation of the magnetic field along the z-axis with the minimum at 200 μm being 0.418 G, which can be adjusted by an offset field along the y-axis. It is useful to express the trap in the form of potential energy, and by rearranging equation 2.50, we find

$$U = \frac{2}{3}\frac{m_F g_F \mu_B}{k_B}|\mathbf{B}((r))|, \tag{4.2}$$

where $\mathbf{B}((r))$ is the magnetic field. This yields an effective temperature given by the potential energy divided by k_B, and is given in Figure 4.5c for the trap along z, assuming the atoms to be in the $|F = 2, m_F = 2\rangle$ state. This also includes the gravitational potential, $-mgz$, taking the trap minimum to be the zero. From this, we see that the maximum trap depth is approximately 0.5 mK, which implies that atoms should be cooled to tens of μK to be trapped for a significant time.

Recall from Figure 2.9b that the atomic cloud is rotated slightly about the z-axis. Figure 4.6 shows a cross section of the magnetic field in the x-y plane at the trap height. From this, the angle is rotated by 1.78° yielding the transformed coordinates x' and y'. The z-axis is unchanged, but for consistency when the transformed coordinates are used it will be referred to as z'. Lastly, using Equation 2.52, the trap frequencies in the rotated coordinates are found to be $\omega_{x'} = 2\pi \times 1.72$ kHz, $\omega_{y'} = 2\pi \times 348$ Hz, and $\omega_{z'} = 2\pi \times 2.22$ kHz.

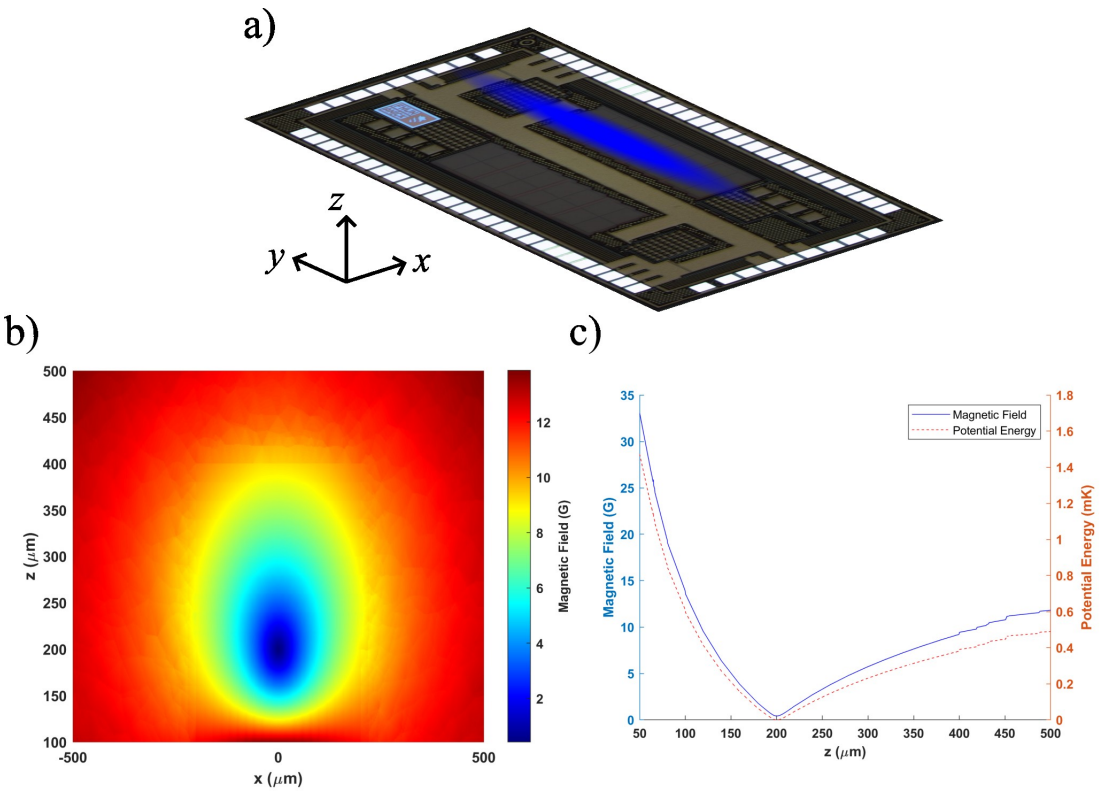


Figure 4.5: a) Illustration of the atomic cloud trapped using the CMOS Z-wire with an external bias field along the x-axis. b) A cross section in the x-z plane of the magnetic field which results in the trap. c) The magnetic field along z (solid blue line) and the corresponding potential energy in temperature units, including the gravitational potential (dashed red line).

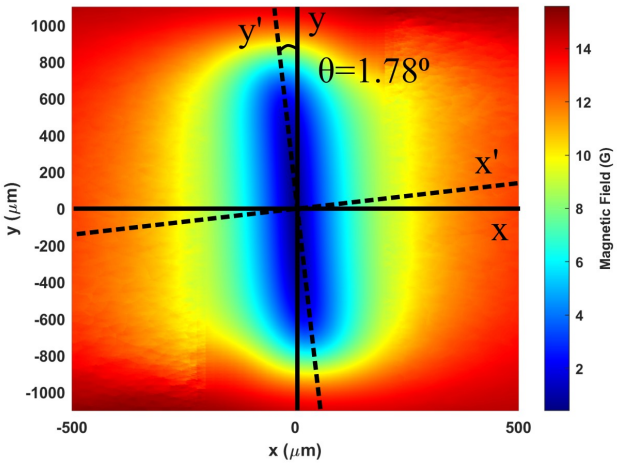


Figure 4.6: A cross section of the magnetic field in the x-y plane at the trap height of 200 μm with the rotated axis x' and y'. The rotation angle is found to be 1.78°.

4.2 Printed Circuit Board (PCB)

In order to interface with the CMOS chip and transport atoms from the MOT, an intermediate chip is required, which is implemented as a vacuum-compatible PCB. There are three primary requirements for the PCB:

1. Provide a quadrupole waveguide with a field gradient of at least 32 G/cm in the vertical axis to confine the atoms. However, due to the need to compress the cloud before loading into the CMOS trap, a target of 100 G/cm was chosen.

2. Provide longitudinal trapping with a trap that can transport the atomic cloud from the MOT position to the CMOS trap position.

3. Provide all of the necessary bonding pads to interface with the CMOS chip and act as a breakout board allowing these to be connected to the high density sub-D connectors on the flange.

To understand how the PCB was designed, it is necessary to know how a PCB is structured and built. PCBs are made up of three different kinds of layers: copper, core, and prepreg (from pre-impregnated). The core and prepreg layers are both electrically insulating, typically made of glass fibre and epoxy resin, and allow layers of copper to be stacked on top of each other. To make each layer, copper is placed on the top and bottom of a core layer and etched, leaving the required pattern of copper on each side. To remove unwanted copper, the desired copper is covered in a protective layer of hardened photo resist, and a copper solvent is administered to remove the unprotected copper. Once this process has been repeated for every core layer (every core layer is sandwiched by copper), these layers are then stacked on top of each other, with the prepreg layer placed in between to prevent contact between the copper layers. The resulting stack is shown in Figure 4.7.

In the final PCB, the layers are connected by vertical copper channels known as vias. Firstly, holes are drilled through the board at the locations of these vias, and copper is deposited on the inside of these vias by a series of chemical baths, leaving a wall of copper running along the inside of the drilled hole. The ends are then plated closed. These types of vias are known as plated through-hole vias and were the only type of vias used in the PCB, as they prevent air from being trapped on the inside of the board. Vias are referred to as blind when they end at an internal layer.

Since individual layers of copper used in PCB construction tend to be thin, 35 μm in our case, several layers are required for the necessary current densities used to trap atoms. An efficient way of simulating a thicker wire is to recycle current; by starting on one layer and looping downwards, similar to a spiral staircase. Because of this, simulations are conducted using several thin wires stacked above each other, as opposed to a single thicker wire.

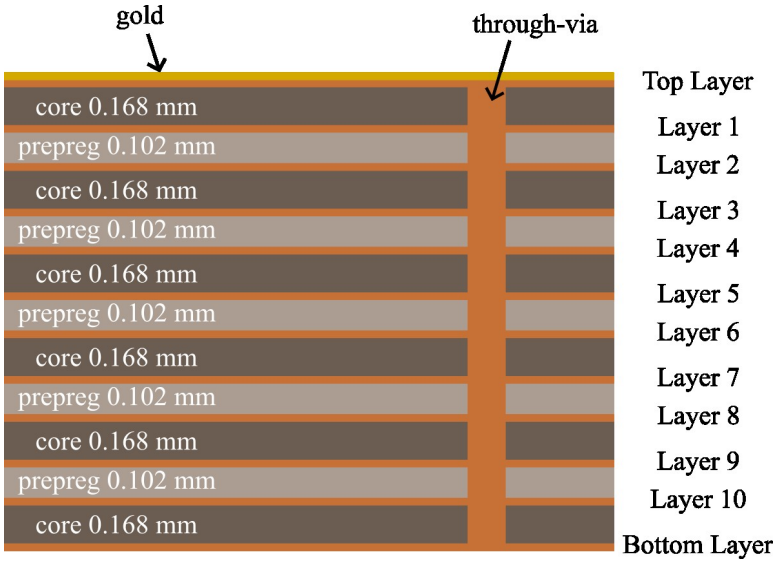


Figure 4.7: Shown is a simplified cross section of the PCB with all twelve layers used in the final design. The core and prepreg layers are electriclaly

4.2.1 Quadrupole Waveguide

The magnetic trap produced by the PCB can be analysed as two separate traps: a waveguide which provides confinement in the transverse axis, and a transport trap which confines the atoms along the direction of travel. To create a chip-based waveguide, there are several possible configurations such as the single wire waveguide in Figure 2.7, as well as self-biased configurations such as the three wire waveguide[24,22]. However, the self-biased four wires configuration allows us to recycle current easily, thus it was chosen for this design. This is implemented using four wires that run in parallel with alternating currents (alternating in this case meaning anti-parallel with consecutive wires), as shown in Figure 4.8. In reality, this is constructed with two rectangular coils, an inside coil and an outside coil where the alternating current results from the continuity of current around the coil. Each wire is comprised of 8 layers, corresponding to layers 3 to 10 in Figure 4.7.

To determine the optimal wire widths and currents, a cross section of the wires is simulated using COMSOL Multiphysics with varying widths and currents, but with a constant trap height and gradient. This process is outlined in the Ph.D. thesis of Amruta[3]. The gradient is set at 100 G / cm vertically and the trap height is set at $Z_0 = 1.2$ mm. The widths of the two inner (outer) wires are set to be equal with constant separation, and we calculate the necessary currents. To do this, first simulate with the inner coil current, i_c, set to 1 A and the outer coil current, i_o, set to 0 A, to find the magnetic field $B_{c,1A}(\boldsymbol{r})$. Then repeat with $i_c = 0$ A and $i_o = 1$ A to obtain $B_{o,1A}(\boldsymbol{r})$. To form a trap at Z_0, $B_{c,1A}(Z_0)$ must be scaled to equal $B_{o,1A}(Z_0)$. Noting that $B \propto I$, we can define the ratio of field strengths at Z_0 as

$$\frac{B_{o,1A}(Z_0)}{B_{c,1A}(Z_0)} = \frac{i_c}{i_o} = \alpha. \tag{4.3}$$

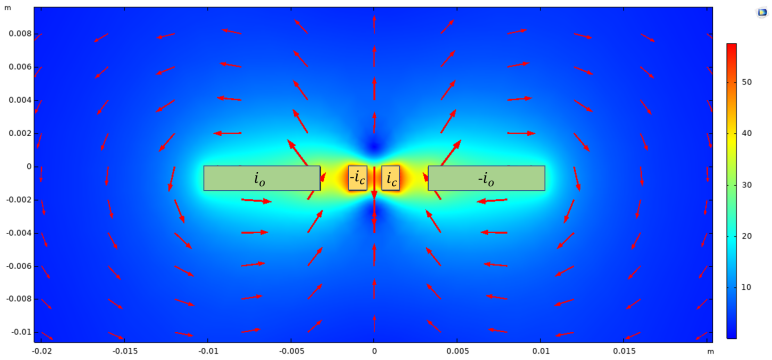


Figure 4.8: Four wires in parallel with alternate currents, shown by I, can produce a quadrupole trap which is demonstrated by the surface plot which is in units of G. A field zero is present between the central wires both above and below at a distance known as the trap height. The arrows show the direction of the field.

Setting the currents as $i_c = \alpha$ A and $i_0 = 1$ A results in a field zero at the trap height and they can then be scaled linearly to achieve the desired gradient. The necessary ratios and currents for inner coil widths varying from 0.5 to 2.5 mm and outer coil widths varying from 3 to 12 mm are shown in Figures 4.9a and 4.9b respectively. In all of the simulations, the separation between the inner wires was 1 mm and the gap between an inside wire and outside wire was 1.5 mm.

Once the necessary currents have been determined, the power per metre dissipated by the wires can be calculated. The power dissipated by a wire with resistance R and current i is $P = i^2R$. The resistance of a wire can then be expressed as $R = \rho_{\mathrm{Cu}}L/A$ where ρ_{Cu} is the resistivity of copper, L is the length and A is the cross sectional area. Setting the length to one metre, we get the expression

$$P = \rho\frac{i^2}{A}. \tag{4.4}$$

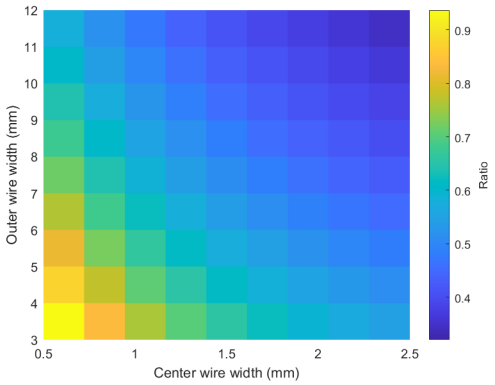
Thus, for the quadrupole trap wires

$$P = 16\rho_{\mathrm{Cu}} \times \sum_{j=c,o} \frac{i_j^2}{(35\times 10^{-}6)\times w_j}, \tag{4.5}$$

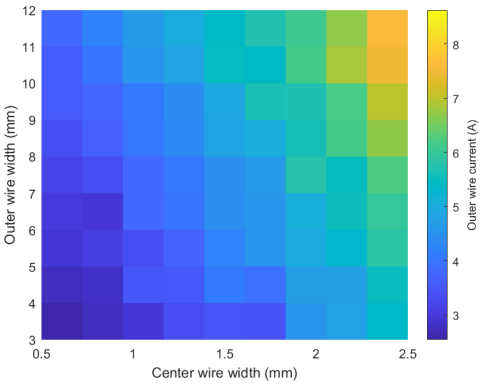
where the subscript j differentiates between the inner and outer coils and w_j is the width of each coil. Multiplying by 16 accounts for the 8 layers and the 2 sides of the coil.

In the simulations in which the current ratio and scaling were calculated, the power was also calculated and the results are shown in Figure 4.9c, from which we see that optimal performance occurs with an inner wire width of 1.6 mm and an outer wire width of 11 mm. The corresponding currents are then 2.4 A and 6.2 A with a power dissipation of 55 W/m. Unfortunately, due to an underestimate of the trap height early in the design phase, the widths of 1 mm and 7 mm were used in the final design. The corresponding currents are 2.3 A and 4.65 A with a power dissipation of 58 W/m.

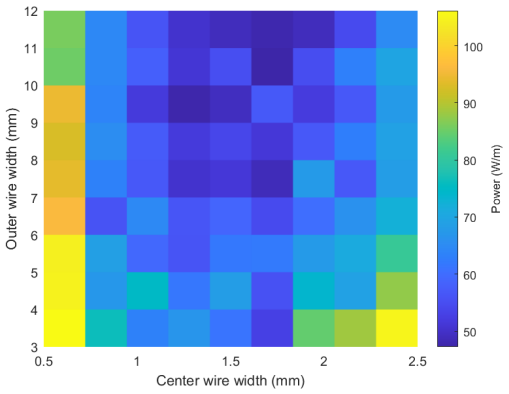
The final trap simulations are shown in Figures 4.10a and b, where a shows the trapping along the



(a) Ratio of outer wire current to inner wire current.



(b) Current required in the outer wire to result in the target trap gradient.



(c) Power dissipation due to the calculated currents.

Figure 4.9: Figures summarising the results of calculating the necessary currents and corresponding power dissipation when iterating over the widths of the inner and outer coils when generating a four wire quadrupole trap.

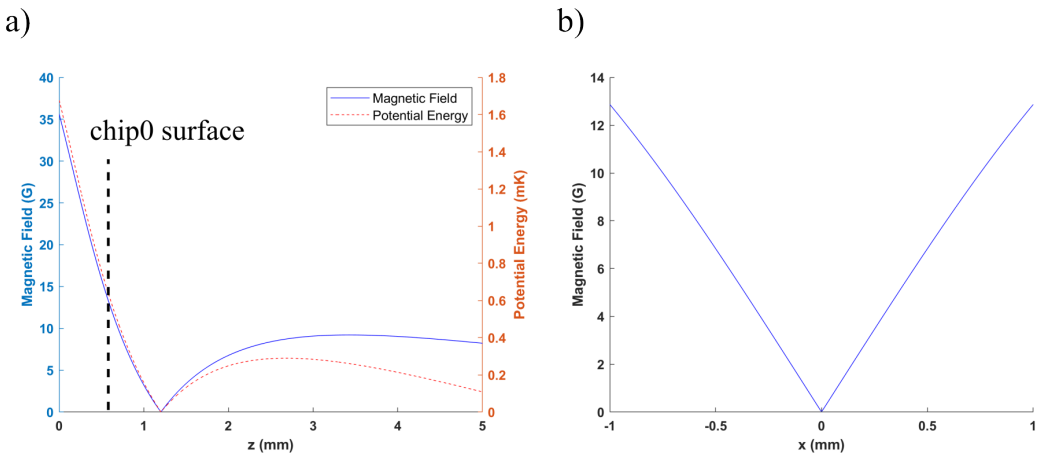


Figure 4.10: a) Trapping along the z-axis in both G and as potential energy including the effects of gravity. Indicated is the surface of chip0. b) The magnetic field along the x-axis.

z-axis and b shows the trapping along the x-axis. In Figure a, the trapping is also given as potential energy in temperature units, according to Equation 4.2, including the gravitational potential. We see that for larger traps the effect of gravity is much more pronounced and lowers the maximum depth to 0.3 mK, which is the limiting depth of the PCB trap. The surface of chip0 is also indicated in this figure to highlight the trap depth before atoms are lost into the surface. In the x-axis, the trap depth is at least 13 G, which corresponds to 0.58 mK.

4.2.2 Transport Wires

The longitudinal trapping is simply generated by pairs of parallel wires with currents running in the same direction, which results in magnetic fields that cancel in between, producing a deep trap. However, as the waveguide is located at some height above the PCB, the fields do not cancel completely, thus leaving a non-zero field at the bottom and converting the trap to an Ioffe trap. This results in the trap shown in Figure 4.11, where two layers are used in these wires. Inclusion of a bias field will then allow us to shift the trap along the waveguide whilst the currents are consecutively ramped to the next pair of wires.

4.2.3 Design

The final design of the PCB is demonstrated in Figures 4.12a-d. Figure a shows the top layer of the PCB that acts as the breakout board for the CMOS chip, which is fixed to the square near the centre of the board. The four thicker traces provide current to the Z-wires. This is around 1.5 A and is the highest current on this layer, hence the wider traces to reduce heating. The remaining traces are for a mixture of analogue, digital, and RF signals which are all low power. To minimise cross-talk, the traces have been spread out as much as possible with grounded planes in-between to act as shielding. Figure b shows the top layer of wires used to produce the Ioffe field, with the second layer of wires on the layer below which is shown in Figure A.3 in the Appendix. Figure c presents the implementation of the four wires

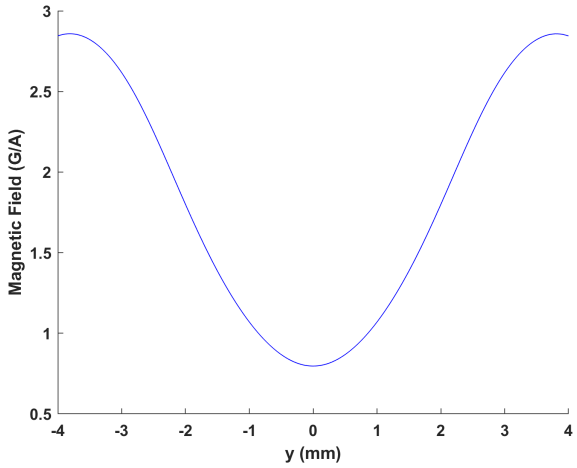


Figure 4.11: Magnetic field per amp produced by the longitudinal trapping wires on the PCB.

used to create the quadrupole trap, including dimensions. These wires are repeated for a total of eight layers, with the current recycled to the next layer by the vias on the right-hand side of the board. Lastly, Figure d gives the dimensions of the inner coil, and details of the wires used to generate the trapping are summarised in Table 4.1.

The PCB has a total of 60 pins, with four being used for the central coil, four for the outer coil, eight for the transport wires, and the remaining 44 being free for the CMOS chip. Given that the current chip only requires 29 pins, there are 15 additional pins available for future chips. The pins used are the 4366-0-00-21-00-00-03-0 model of press-fit pins supplied by Mill-Max and are used in existing experiments within the group. There are also four mounting pins at each corner of the PCB with a diameter of 2.54 mm (100 mil). The PCB was fabricated and assembled by Total Circuit Solutions, based in Peterborough.

Name	Width (mm)	Current (A)
Outer Coil	7	4.56
Inner Coil	1	2.3
T1-4	2	4

Table 4.1: Summary of PCB wires used to generate trapping.

As the PCB is used in an ultra-high vacuum (UHV) experiment, the outgassing characteristics of the materials used are an important consideration in maintaining adequate vacuum pressure. To this end, for the core and prepreg layers we use RogersCorp 4350 and 4450f, respectively, due to their low outgassing, which is validated in previous experiments. The vias are also filled purely with copper, as opposed to the standard mixture of copper and resin, which is easier to fabricate at the cost of worse outgassing. Finally, any polymer layers such as soldermask and silkscreen are excluded and the board is coated in soft gold to protect the copper traces and improve conduction. In this final step, nickel is often used instead of gold in the form of alloys which, whilst cheaper, may negatively impact the magnetic

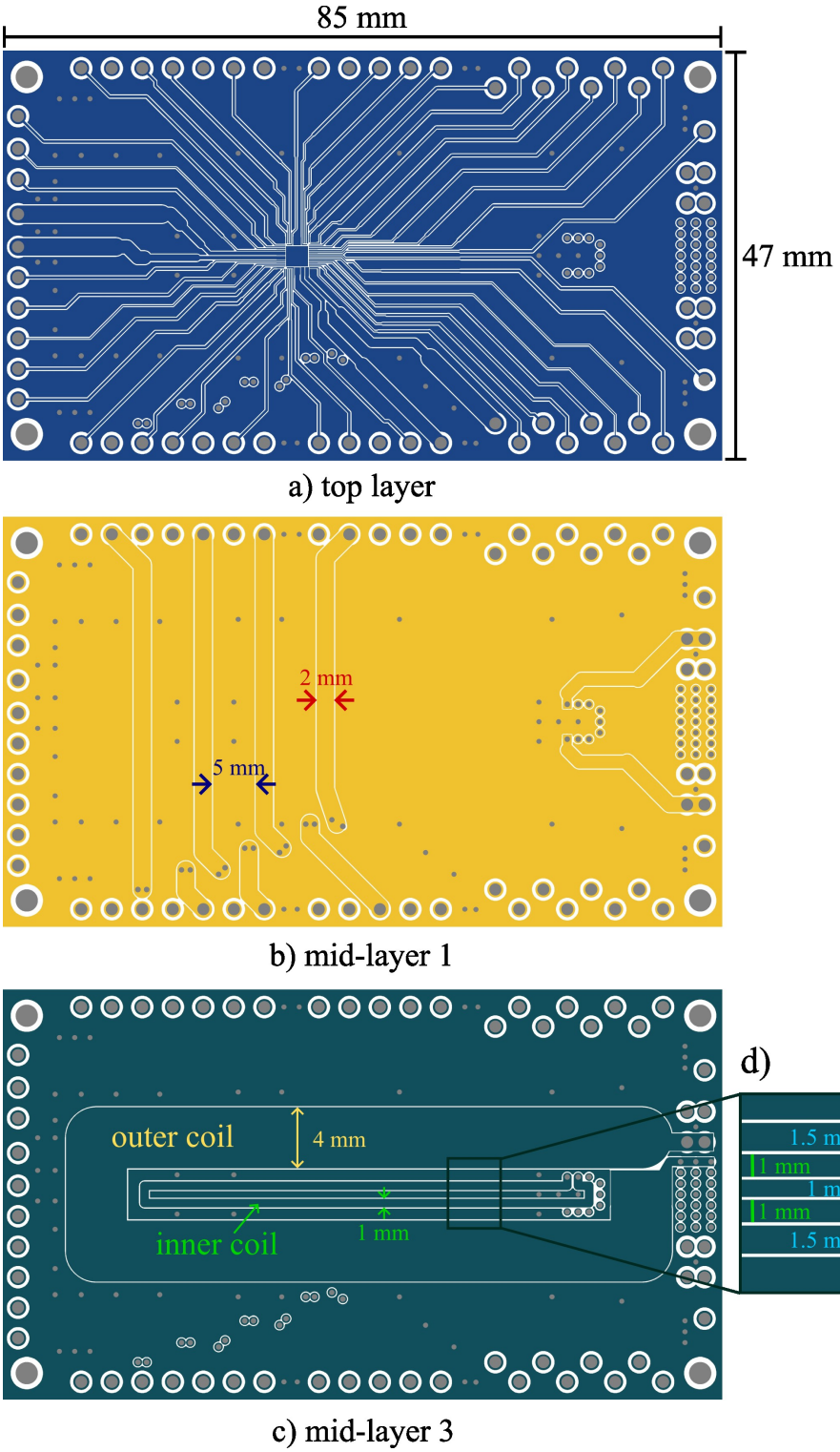


Figure 4.12: Shown are the labelled CAD layouts for three out of the twelve PCB layers.

trap.

4.2.4 Wire Bonding

As discussed above, the PCB is designed to mount the CMOS chip and act as a breakout board for the small 100 x 100 μm pads on the chip. As seen in Figure 4.12a, there is a 2.5 mm x 2.5 mm square located near the center of the top layer of the PCB which serves as a location marker for the chip. This is surrounded by 44 traces that connect to pins that can be connected to the vacuum chamber feedthrough. The CMOS chip is attached to the PCB using epoxy[2] and was accomplished using a micropositioner. The chip is held by the top face using a gel pack, and a small amount of the epoxy is placed on the location marker. The micropositioner is then used to place the chip on the marker, and the PCB is heated to cure the epoxy. Following this, 25 μm gold wire is used to wire bond the chip pads to the traces on the PCB as shown in Figure 4.13. In wire bonding, a combination of heat, pressure, and ultrasonic vibrations is used to weld the wire to a surface. Initially, balls are formed at the end of the wire by using arcing from a high-voltage arm which is quickly swung beneath the wire. This ball is then used to form a starting bond, referred to as a ball bond, from which the wire can be moved in any direction before being welded onto the location of the second bond, forming what is known as a fishtail bond, due to looking like a fishtail when viewed from above. Four wire bonds are used for each of the Z-wire traces to ensure a sufficiently high current is viable.

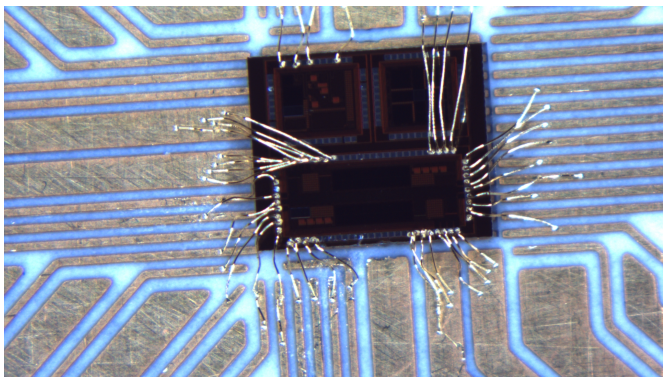


Figure 4.13: Shown is a photograph of chip0 wire bonded to the PCB. The wider PCB traces are for the Z-wires, the top left bonds on chip A are practice bonds and the rest are the analogue, digital, power and ground connections for chip C.

4.3 Science Chamber

The primary chamber of the experiment holds the CMOS chip, with necessary mounting structures, the magnetic trapping structures, and the optics necessary for collimating, polarising, and reflecting the cooler, repumper, optical pumping, and imaging beams. A computer-aided design (CAD) model for the

[2]EPO-TEK 353ND

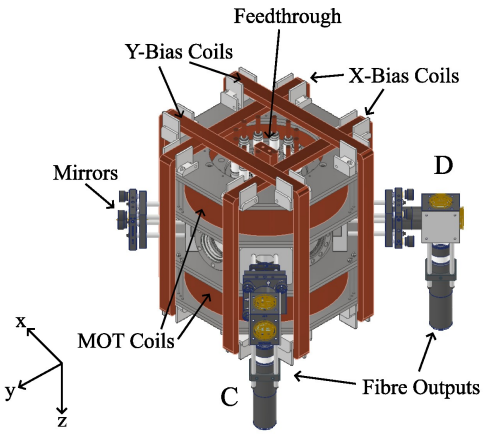


Figure 4.14: Shown is CAD model for the science chamber including the MOT coils, bias coils, feedthrough, water cooling and chamber mounted optics. The chamber mounted optics provide the horizontal cooler and repumper beams. The MOT coils are in anti-Helmholtz configuration to produce a quadrupole trap and are sandwiched between water cooled cold-plates which are detailed in[27]. The bias coils are mounted to the cold-plates using 3D printed mounts and are in Helmholtz configuration to provide homogenous field. Lastly the feedthrough provides the electrical and mounting infrastructure for the chip and PCB.

outside of the science chamber is shown in Figure 4.14 with the bias coils, MOT coils, feedthrough, and chamber-mounted optics labelled.

To provide an interface and mounting point for the chip and PCB, a feedthrough flange is used. The flange consists of a copper block with two M6 threaded holes for mounting which serves as a heat sink for the PCB; two 26-pin high density sub-D connectors and 12 copper rods for high current connections. This flange is connected to an octagonal vacuum chamber separated by a spacer, and the chamber has 8 DN40 CF flanges for viewports and other connections. Opposite the feedthrough is a DN100 viewport.

4.3.1 PCB Mount

Inside the science chamber there are various mounting structures and electrical connections necessary for interfacing with the CMOS chip and for magnetic trapping. The complete mounting sequence, excluding the wires connecting the PCB to the sub-D connectors, is shown in Figure 4.15. Initially, the CMOS chip is bonded and wire-bonded to the PCB as described above, a waveplate with reflective coating is bonded next to the chip using a UV curable epoxy. The waveplate is necessary for forming a MOT, although other geometries such as the gMOT[57] exist and may be explored in future iterations of the experiment. The PCB is then held to a copper mount, shown in Figure 4.15b, using vented bolts at each corner. The copper mount is designed to have as much contact with the PCB bottom layer as possible to act as a heat sink for the magnetic trapping wires on the PCB; to help with this, a UHV compatible thermal paste is added to increase the efficiency of this heat transfer. This mount is then fixed to the copper stem by two M6 bolts. Gaps must be left around the pins and small channels are added to the mount beneath the vias on the PCB to prevent shorts when the thermal paste spreads out

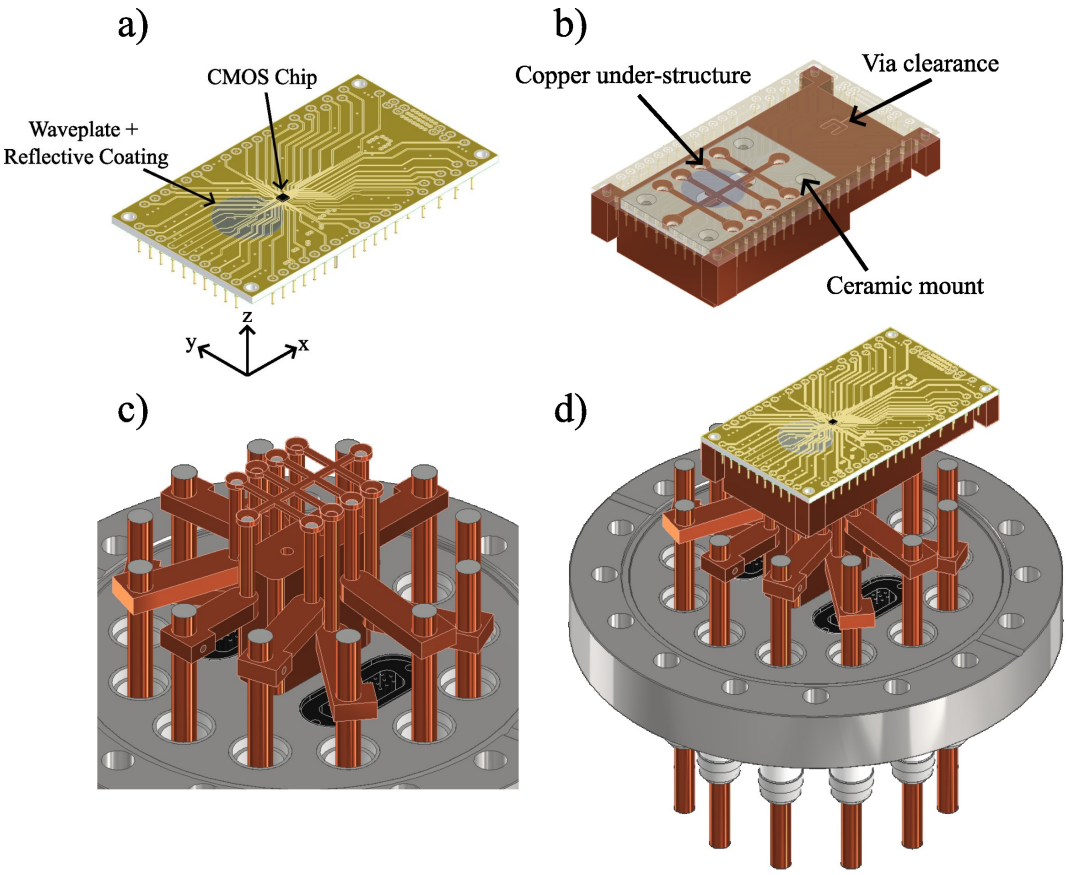


Figure 4.15: CAD models for the PCB and mounting to the feedthrough is shown. In a) chip0 is bonding the PCB with a neighbouring $\lambda/4$-waveplate with a reflective coating on the underside which is used in the MOT formation. b) Shows the PCB mount which is designed to act as thermal sink for the PCB in addition to housing a copper wire under-structure which essentially a repeated H-wire and was added as an alternative to the PCB for the intermediate magnetic trap where the atoms are compressed and cooled before being loaded into the CMOS chip trap. c) Shows the how the copper under-structure is connected to an external power supply via the copper rods in the feedthrough. d) Shows the completed mounting.

during baking. In addition to the magnetic trapping structures on the PCB, a copper under-structure was added to the mount as an alternative. The under-structure features cascaded H-wire structures which can be used to form U-wire traps, Z-wire traps, or a single-wire waveguide. The copper under-structure is mounted onto a ceramic base to electrically isolate it from the mount, and the top of the under-structure is electrically isolated from the PCB by a layer of kapton foil. The copper under-structure is connected to the copper rods in the feedthrough by threaded copper rods and horizontal copper clamps, as shown in Figure 4.15c before being connected externally to a power supply[3]. The complete mounting structure is shown in Figure 4.15d with the PCB mount held onto the copper block in the feedthrough with vented M6 bolts. In the mount design, it is crucial to incorporate channels which ensure that no air pockets are formed, as these could lead to artificial leaks in the vacuum. Moreover, UHV compatible materials must be used such as oxygen-free copper, UHV compatible epoxies, and thermal paste.

[3]Hewlett Packard 6671A

4.3.2 Chamber Optics

The fibre output and retro-reflectors for the horizontal cooler and repumper beams are also fixed to the science chamber as shown in Figure 4.16. To form a MOT, cooler beams of the correct polarisation with a sufficient beam diameter must be produced. Note that Gaussian beams have no clear end to their intensity in the radial axis, so the diameter of these beams is defined using the radius at which the intensity drops to $1/e^2$ of the maximum intensity in the centre. To calculate the diameter of a collimated fibre output, we start with the divergence angle of a Gaussian beam starting with waist w_o, which is 5.3 μm for the fibres[4] used in the experiment and is known as the *mode field diameter* of the fibre, calculated by

$$\theta = \frac{2\lambda}{\pi w_0} \tag{4.6}$$

where λ is the wavelength of the laser. The diameter of the beam using a lens with focal length f is then given by

$$d = 2\theta f = \frac{4f\lambda}{\pi w_0}. \tag{4.7}$$

Thus, for the outputs of the cooler beams where a focal length of 125 mm is used, we find a beam diameter of 23.4 mm.

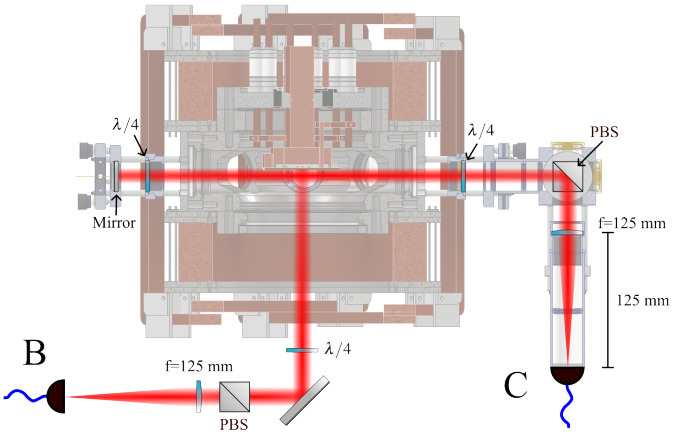


Figure 4.16: Shown are the chamber optics required to generate the cooler and repumper beams. At all of the fibre outputs the beams are collimated to a beam diameter of 23.4 mm using lenses with a focal length of 125 mm before PBSs are used to clean the polarisation of the beams. At this stage the beams are linearly polarised so quarter waveplates are used to circularly polarise the light for the MOT. For the opposing beam, a quarter waveplate and mirror are used to reflect the beam and reverse the polarisation. Not shown in this figure is the waveplate with reflective coating that is present on the PCB.

Following the collimation, a PBS is used to clean the polarisation of the beam by ensuring that only the reflected component is aligned with the chamber. The fibre can be rotated to maximise the light reflected as polarisation-maintaining fibres are used, and the input light is ensured to be linear and aligned with the slow axis of the fibre. A quarter waveplate is then used to form circularly polarised light for the MOT. At the opposing window, a quarter waveplate and mirror are added such that the

[4] Thorlabs P3-780PM-FC-5

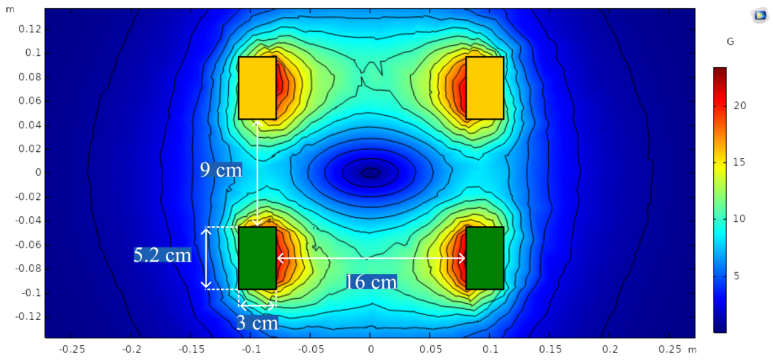


Figure 4.17: A COMSOL Multiphysics simulation of the MOT coils with 1 A is shown with the dimensions of the coils and the distance between them. cross sections of the coils are shown in yellow (top coil) and green (bottom coil).

MOT beams is reflected back along the same path with a σ-polarisation of the opposite helicity.

The fibre outputs and retro-reflective stages are fixed to the chamber using custom window mounts which fit to the windows using set screws and mount the optics offset from the centre of the window such that the beams align with the waveplate on the PCB whilst being low enough to avoid being blocked by the PCB and mount.

4.3.3 Coils

External coils are mounted to the science chamber to provide the necessary trapping and bias fields. The two largest innermost coils are the MOT coils, which are held between pairs of water-cooled cold plates. These are used to prevent the chamber from significantly heating up as the coils run at a high current during the magnetic capture stage and can reach temperatures of 90°C without cooling. The cold plates are detailed in the Ph.D. thesis of Jamie Johnson[42]. Although 32 G/cm is a sufficient magnetic field gradient to hold the atoms against gravity, since the MOT coils are used to catch falling atoms during the magnetic capture stage, they are designed to have a higher maximum trap gradient of 50 G/cm using the available power supplies[5] which have a maximum output of 25 A. Using wire with a 4 x 2 mm cross-sectional area, a coil geometry of 13 turns at the same radius stacked outward 15 times radially, leading to 195 turns total, was chosen as this was able to produce the highest field gradient of 2.7 G/cm/A based on the COMSOL Multiphysics simulation shown in Figure 4.17. The inner radius of 16 cm and the separation of 9 cm are restrictions due to the vacuum chamber and the cold plates. Assuming perfect construction of the coil, it would have a width of 3 cm and a height of 5.2 cm; however, due to imperfections such as small kinks or irregular turning radii, the actual width and height are measured to be 3.4 cm and 5.6 cm respectively.

Helmholtz configuration coils are added in the horizontal axis to provide bias fields. These coils are fitted to the outside cold plates using 3D printed custom mounts designed so that the coils can be

[5]Hewlett Packard 6652A DC Power Supply

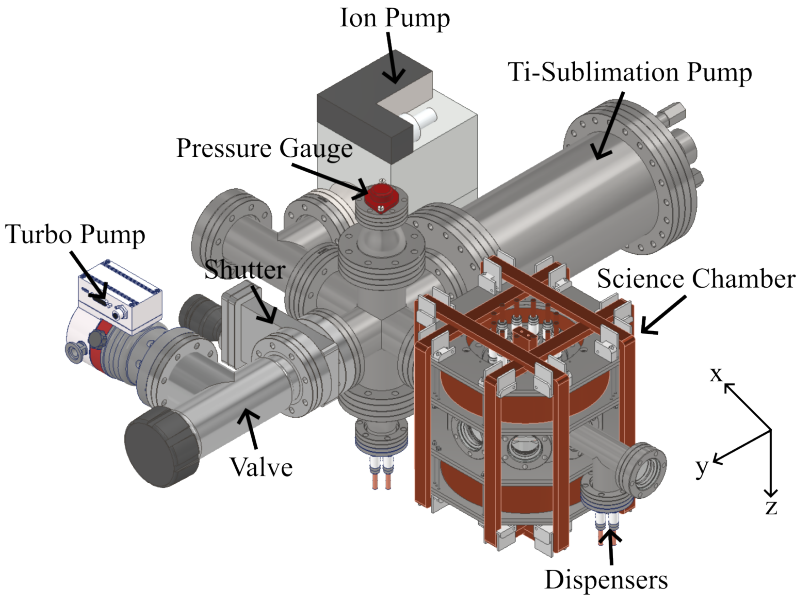


Figure 4.18: A labelled CAD model of the entire vacuum setup is presented. The science chamber as detailed above is connected to the pumps on one side and the dispensers on the other to prevent atoms from being pumped before making it into the chamber. The pressure gauge is located in the 6-way piece which also leads to the titanium-sublimation pump (TSP), ion pump and turbo pump via a shutter and valve which are closed during the bakeout.

easily removed when the chamber is opened. Both sets of coils have a separation of 8 cm and are again simulated in COMSOL Multiphysics. They both have 20 turns of wire with a 1 mm diameter.

4.4 Vacuum Setup

Ultracold atoms experiments require UHV background pressures on the order of $< 10^{-10}$ mbar to reduce background gas collisions sufficiently to produce BECs. High atom numbers and long trap lifetimes are necessary for this process. Reducing pressure to this level is not trivial and requires several pumps in addition to a construction that is able to prevent leaks such that the chamber is sufficiently isolated from the environment. A CAD model for the entire chamber is shown in Figure 4.18 with each of the components labelled. There are three different pumps: a turbo pump which is used at higher pressures and, once low enough pressures are reached, is isolated using a shutter and a valve and turned off; an ion pump which is run continuously; and a titanium-sublimation pump (TSP) which is used periodically. A pressure gauge is located in the centre of the setup and can measure pressures down to 1×10^{-11} mbar. The experiment itself is performed within the science chamber which is detailed above. Lastly, rubidium dispensers are located on the opposite side of the science chamber to the pumps. The setup is constructed using UHV compatible parts made of stainless steel, with ConFlat (CF) flanges that cut into a copper gasket to prevent leaks.

It is important to note that everything inside the vacuum must be UHV compatible, including the

PCB, epoxies, pins, etc. Details of this are found in the relevant sections within this work, but some basic rules are that nothing organic can be used, as these tend to outgas significantly (this is particularly important with the PCB) and everything must be thoroughly cleaned before being placed in the vacuum. The general procedure used was to clean with water and detergent (if required), place in an ultrasonic bath, then finally clean with an organic solvent such as ethanol. Disposable gloves and, ideally, a mask must be worn whilst handling these components, and they should only be placed down on fresh aluminium foil.

4.4.1 Producing UHV

To produce the required vacuum pressure, several stages are required. Firstly, a diaphragm pump (roughing pump) is connected to the turbo pump with all valves open and is used to pump the chamber down to $\sim 10^{-3}$ mbar. It is important that the turbo pump is always run in conjunction with the roughing pump. At this stage, the turbo pump can be turned on; it will initially require a large current to pump due to air resistance, but this will decrease as the pressure decreases to below 10^{-6} mbar. We know the chamber has reached this point when the turbo pump current stops decreasing; this is when we can turn on the pressure gauge without damaging the filaments. The entire setup is now slowly heated using heating belts to around 120°C (a limit set by the materials used in the PCB) to begin a process known as bakeout.

During the bakeout, the entire set-up is held at a high temperature to remove moisture and other contaminants that have diffused into the metal or other in-vacuum parts whilst they have been exposed to air. It also increases the pressure inside of the chamber to allow the pumps to work more efficiently. Typically, the chamber is held at the maximum temperature for anywhere between a few days and a couple of weeks, usually depending on the temperature used and the time taken for the pressure to sufficiently reduce. During this stage, we can also degas filaments and wires by running currents through them to remove contaminants which might degrade the vacuum later on.

Once the pressure has reached $\sim 10^{-7}$ mbar we can switch on the ion pump and pulse the TSP. The TSP has three filaments that are heated to sputter titanium over the outside of the pump chamber; particulates are then able to react with the titanium and form a stable compound which is stuck to the pump. Initially, the TSP is pulsed to degas contaminants stuck to the filament before only being pulsed periodically to replenish the titanium layer.

At $\sim 10^{-9}$ mbar or below, the chamber can be slowly cooled down to room temperature, decreasing the pressure of the heated gas inside. After cooling, the pressure should be $\sim 10^{-10}$ mbar, at which point the chamber can be sealed and the turbo pump switched off. With the ion pump and TSP, the pressure should continue to decrease to $< 5 \times 10^{-11}$ mbar before stabilising.

4.5 Optics

Several different frequencies of laser beams are required to address different transitions on the ^{87}Rb D_2 line, as summarised in Figure 4.19. The different frequencies are derived from three lasers, a Toptica DL Pro which is amplified by a Toptica BoosTA and a TA Pro. The DL Pro is locked to the $F = 1 \rightarrow F' = 1 \times 2$ crossover transition (laser locking and crossover transitions are explained below) and is then split to be amplified by the BoosTA and form the reference in the TA Pro beat lock. The amplified DL Pro beam is passed through a single-pass acousto-optical modulator (AOM) to produce the repumper. AOMs are detailed below. The TA Pro is beat locked to the DL Pro which means that it is locked to the frequency of the DL Pro plus a fixed frequency difference; in this case 6.7 GHz to account for the difference between the ground states $F = 1$ and $F = 2$, and the difference between the 1x2 and 1x3 crossover peaks. The beat lock puts the TA Pro resonant with the $F = 2 \rightarrow F' = 1 \times 3$ crossover transition, but has the advantage that it can be locked arbitrarily in steps of 10 MHz and so can be detuned from a transition. From here, the cooler, imaging, and optical pumping frequencies can be derived using AOMs.

Acousto-Optical Modulators

AOMs are an important tool for small frequency shifts in optics, between 70 MHz and 230 MHz in our case. In an AOM, the beam is passed through a crystal, and a high power rf signal is used to generate a sound wave in the crystal, which acts as a diffraction grating for the beam, with the beam diffracting into several beams at different angles, with frequency shifts equal to integer multiples of the RF frequency due to conservation of momentum between the phonons and photons.

AOMS can be operated in either a single-pass or a double-pass configuration; in a single-pass configuration, the beam passes through the AOM once and is diffracted at an angle proportional to the frequency shift with the unwanted orders blocked. The advantage of the single-pass AOM is a higher efficiency, as power is lost in each pass, the disadvantage being that scanning the frequency shift changes the diffraction angle, which affects alignment after the AOM. In a double-pass configuration, the diffracted beam is reflected back into the AOM and it is shifted again so that it now aligns with the input beam. This causes double the frequency difference and so is useful for larger frequency shifts; moreover, since the diffraction angle is cancelled, the frequency shift can be scanned without significantly impacting alignment.

The RF signal is generated in a homemade AOM driver which features a voltage-controlled oscillator (VCO) that can generate sine waves between 70-130 MHz, an attenuator, a switch, and finally an amplifier which has a maximum output power of 3 W. All of these can be controlled externally through digital and analogue control signals or internally with potentiometers and toggle switches. The ability to adjust the power of the RF signal is important for maximising efficiency without damaging the AOM

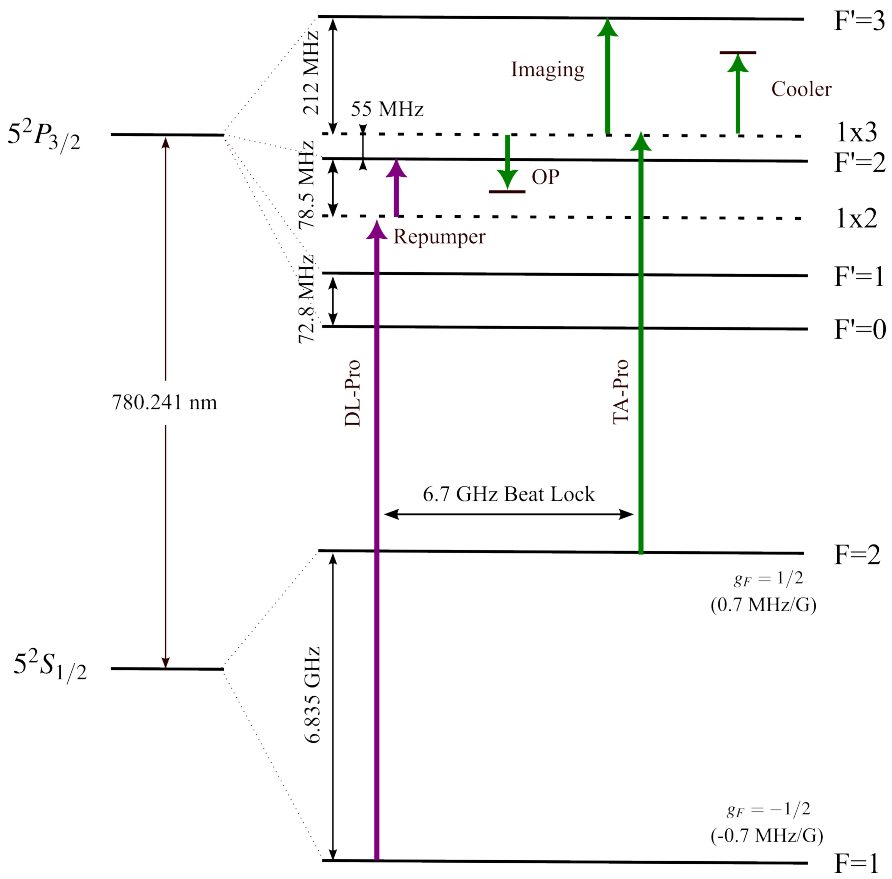


Figure 4.19: The laser frequencies generated are shown with respect to the energy levels in the D_2-line of Rb^{87}. The purple lines are derived from the DL Pro which is locked to the $F = 1 \rightarrow F' = 1 \times 2$ crossover peak, where $F' = 1 \times 2$ indicates that the crossover peak is mid-way between the $F' = 1$ and $F' = 2$ levels. The repumper is then brought on resonance with an 80 MHz AOM. Green lines are derived from the TA Pro which is beat locked 6.7 GHz below the DL Pro frequency which brings it approximately on-resonance with the $F = 2 \rightarrow F' = 1 \times 3$ crossover. From here the optical pumping (OP), imaging and cooler frequencies can be derived with a single pass 80 MHz AOM for the optical pumping and double-pass 110 MHz AOMs for the imaging and cooler. The optical pumping is red-detuned from the $F = 2 \rightarrow F' = 2$ transition by 20 MHz, the imaging is resonant with the $F = 2 \rightarrow F' = 3$ transition and the cooler is red-detuned from the $F = 2 \rightarrow F' = 3$.

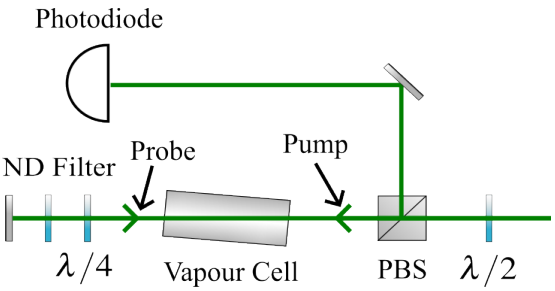


Figure 4.20: saturated absorption spectroscopy (Doppler free) is produced using the optics setup shown. A beam known as the pump beam is passed through an atomic vapour cell exciting atoms in the cell. This beam is the passed through a filter to attenuate it, a quarter waveplate to adjust the polarisation and thus the power reflected to the photodiode and a reflected off a mirror. The reflected beam is referred to as the probe beam and is weaker than the pump beam. If the probe beam excited atoms in the vapour cell then it is attenuated producing a weaker signal on the photodiode which occurs near resonance, however on resonance both beams address atoms with near zero velocity and so the probe is not attenuated since the atoms are already excited by the pump beam.

crystal.

4.5.1 Saturated Absorption Spectroscopy

Experimentally we can resolve these states through a technique known as Doppler-free spectroscopy. For standard spectroscopy, a probe laser beam is shone through an atomic vapour cell and the frequency of the laser is scanned, typically by scanning the cavity length within the laser. When the laser is off-resonance with a transition, then the atomic gas appears transparent; however, when the laser frequency is on-resonance the probe beam is absorbed and the intensity of the laser after the cell is decreased, resulting in a dip in intensity in the probe beam. This dip is then broadened by the Doppler effect, where the velocity distribution of the atoms is translated into the probe being on-resonant for different atoms at different frequencies. However, this Doppler broadening prevents us from resolving the hyperfine splitting. To do this, a pump beam first excites the atoms in the vapour cell; this beam is then attenuated to produce the probe beam and reflected back through the cell. If the atoms are moving toward the pump beam, then the pump frequency is shifted upward, and a red-detuned beam is able to excite the atoms. The probe beam, passing in the opposite direction, has the same frequency, and thus excites atoms moving in the opposite direction, and so is also attenuated. The same is true when the light is blue-detuned. However, if the atoms are stationary, both beams are on-resonance simultaneously, and the probe is no longer attenuated since the atoms are already excited by the pump, resulting in sharp peaks at resonant frequencies. This is shown in Figure 4.21.

In addition to the resonant peaks, peaks at the exact midpoint between hyperfine states are also present. These peaks are known as *crossover peaks* and result from the pump and probe beams addressing atoms with the same velocity regardless of the direction of the beam.

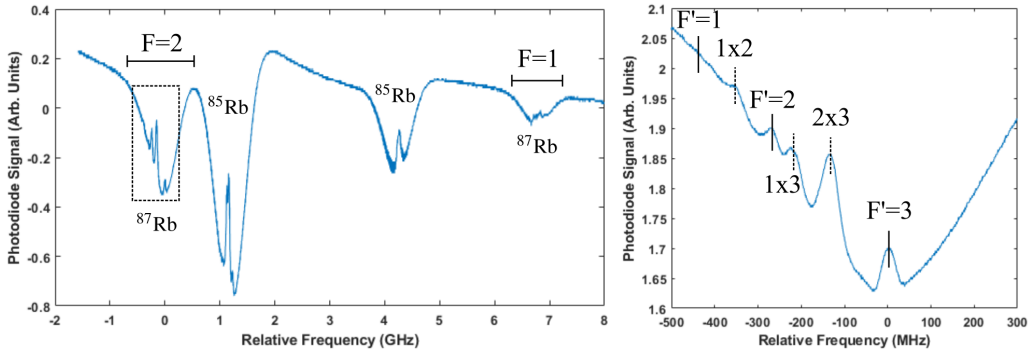


Figure 4.21: Labelled spectrum for the ^{87}Rb D_2 – line. On the left is a full scan of transitions from both the F=1 and F=2 ground states and on the right is an expanded view of the F=2 transitions.

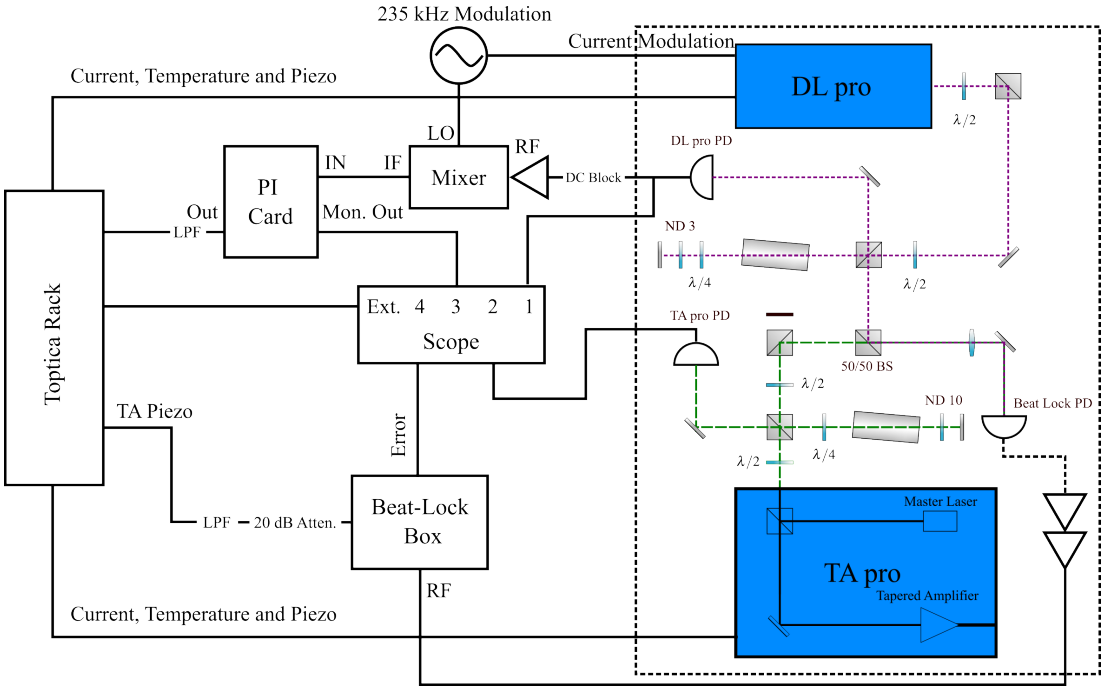


Figure 4.22: A schematic for the frequency stabilisation of both the DL Pro and TA Pro is shown. The outputs of both lasers are initially used to produce a Doppler free spectrum and 5 mW of each is overlapped at a fast photodiode for the beat lock. The DL Pro is locked to the $F = 1 \rightarrow F' = 1 \times 2$ crossover peak by modulating the current at 235 kHz and feeding both this signal and the spectrum (amplified with it's dc component removed) to a mixer followed by a low pass filter to produce an error signal. The error signal is proportional to the derivative of the spectrum thus can be used to lock to a peak. This error signal is fed back to the DL Pro piezo control via a PI card which amplifies and integrates the error signal to prevent it settling at a non-zero point. The fast photodiode for the beat locked detects the frequency difference of the two lasers which is divided and compared to a 10 MHz reference to produce an error signal determining if the target difference is met. The error signal, following an attenuator and low-pass filter, modulates the TA Pro piezo to maintain the target frequency difference.

4.5.2 Laser Locking

In order to maintain a stable laser frequency, the lasers can be "locked" to the peaks corresponding to a transition or a halfway point between transitions produced in saturated absorption spectroscopy and shown in Figure 4.21. This type of locking scheme is called a dither lock. In order to do this, a derivative

of the spectrum known as the *error signal* is produced such that the peak is zero-crossing with one side positive and one side negative, which breaks the symmetry of the peak and allows feedback to counter the change in frequency of the laser. To produce the error signal, the laser current is modulated with a low amplitude, A_m, sine wave of frequency ω_m, which in turn modulates the photodiode signal used to detect the spectroscopy. The output of the photodiode, S_{pd}, at a carrier laser frequency f is given by

$$S_{pd}(t) = I(f + A_m \sin(\omega_m t)) \approx I(f) + A_m \sin(\omega_m t)\frac{dI}{df} + \ldots, \tag{4.8}$$

where the second equality is derived from a Taylor expansion about f. The first term is then removed using a DC block, amplified and the remaining signal is mixed with a demodulation signal with the same frequency as the modulation signal. The resulting mixed signal is the multiplication of the inputs and is given by

$$\begin{aligned} S_{\text{mix}} &= A_{LO}\sin(\omega_m t + \phi_{dm}) \times A_{RF}\sin(\omega_m t + \phi_{pd})\frac{dI}{df} \\ &= B\frac{dI}{df}\left[\cos(\phi_{pd} - \phi_{dm}) - \cos(2\omega_m t + \phi_{pd} + \phi_{dm})\right] \end{aligned} \tag{4.9}$$

where we include phases for the modulation, ϕ_{pd}, and demodulation, ϕ_{dm} and $B = A_{LO} \times A_{RF}$. The higher frequency term is removed with a low-pass filter leaving the error signal which is shown in Figure 4.23a for the $F = 1$ ground state transitions, where we see that each peak is now given by a slope crossing zero, allowing feedback to stabilise the frequency of the laser. This error signal is passed to a proportional-integral controller (PI card), which is a control circuit necessary to optimise the feedback, before being fed back to the piezo-control for the laser.

The TA Pro is then beat locked (also called offset locked) to the DL Pro using a home-built beat lock box based on the design in[5]. The TA master oscillator output is split to produce a Doppler-free spectrum and 5 mW is overlapped with 5 mW from the DL Pro at a fast photodiode. The superposition of the two frequencies includes frequency terms at both the sum and difference of the initial frequencies, with the sum being too fast for the electronics to process, so only the difference is detected. A low-pass filter ensures that higher frequency terms are not present to add noise to the beat lock, and the photodiode is amplified before being input to the beat lock box. The frequency difference signal is divided down by a multiplier (670 for a 6.7 GHz difference) and phase-frequency locked to a 10 MHz reference signal, provided by the experiment control system[6], by an optical phase-locked loop (OPLL) in the beat lock box. At the output is a digital signal which is high if the frequency difference is below the target and low if the frequency difference is above. This digital signal is low-pass-filtered and attenuated before being fed back to the piezo control for the TA Pro.

[6]This is a PXI system provided by National Instruments

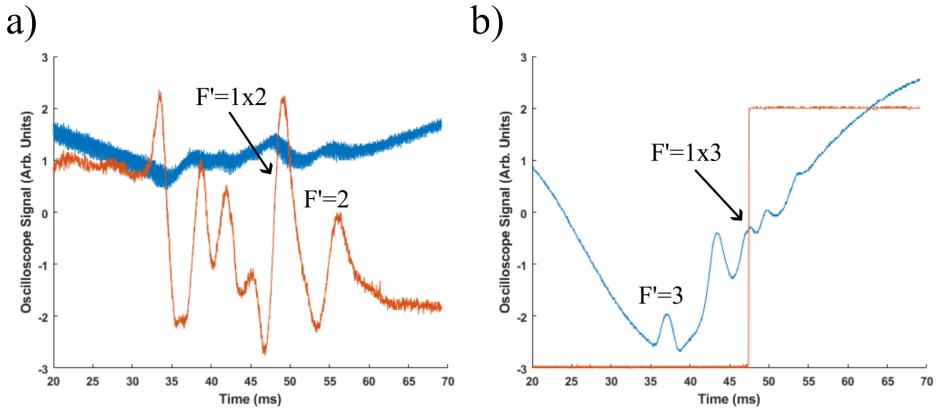


Figure 4.23: a) Absorption spectrum for the DL Pro (blue) zoomed into the F=1 transitions with the derived error signal (orange) which is used for frequency stabilisation. The spectrum looks noisy due to the modulation used in producing the error signal. b) Spectrum for the TA Pro (blue) and beat lock error signal (orange).

4.5.3 Optics Table

The full optical table for the experiment is shown in Figure 4.24. As described above, the DL Pro output is split with around 40 mW fiber coupled to the BoosTA and 10 mW being used for laser locking. The 10 mW is split again with 1 mW used for spectroscopy in a standard saturated absorption setup, and the remainder is overlapped with the TA Pro master laser output and aligned with a photodiode for the beat lock.

The TA Pro has two outputs, one from the master laser and one from the amplifier. Similarly to the DL Pro, the TA master output is split with 1 mW being used for spectroscopy, which is only used for reference in this case as opposed to locking, and the rest overlapped with the DL Pro output for the beat lock.

The TA amplifier output is split into two paths, 50 mW is used for imaging and optical pumping, and the remaining 600 mW is used for cooling. To derive the imaging laser, a double-pass configuration is used with a 110 MHz AOM to blue-shift the light on to resonance with the $F = 2 \rightarrow F' = 3$ transition, although the frequency shift of this AOM is often changed depending on when the imaging is occurring as magnetic fields shift the resonance with Zeeman splitting. The optical pumping is red-shifted with a single-pass AOM as we want it to be close to the $F = 2 \rightarrow F' = 2$ transition and the smallest AOMs we had available have a centre frequency of 80 MHz resulting in a minimum shift of 70 MHz. It might be assumed that on resonance is best for optical pumping; however, if the cloud is very dense, then it might be too opaque for a brief pulse of resonant light, and often detuned light is better for optical pumping. Once the frequencies have been derived, the optical pumping and imaging beams are overlapped at a polarising beam-splitter (PBS) with half-waveplates, allowing us to adjust the power of the beams before they are fibre-coupled and fed to the chamber with polarisation-maintaining fibres. Prior to all of the fibre in-couplers, which are just a lens to focus the light into the fibre, there are quarter and half

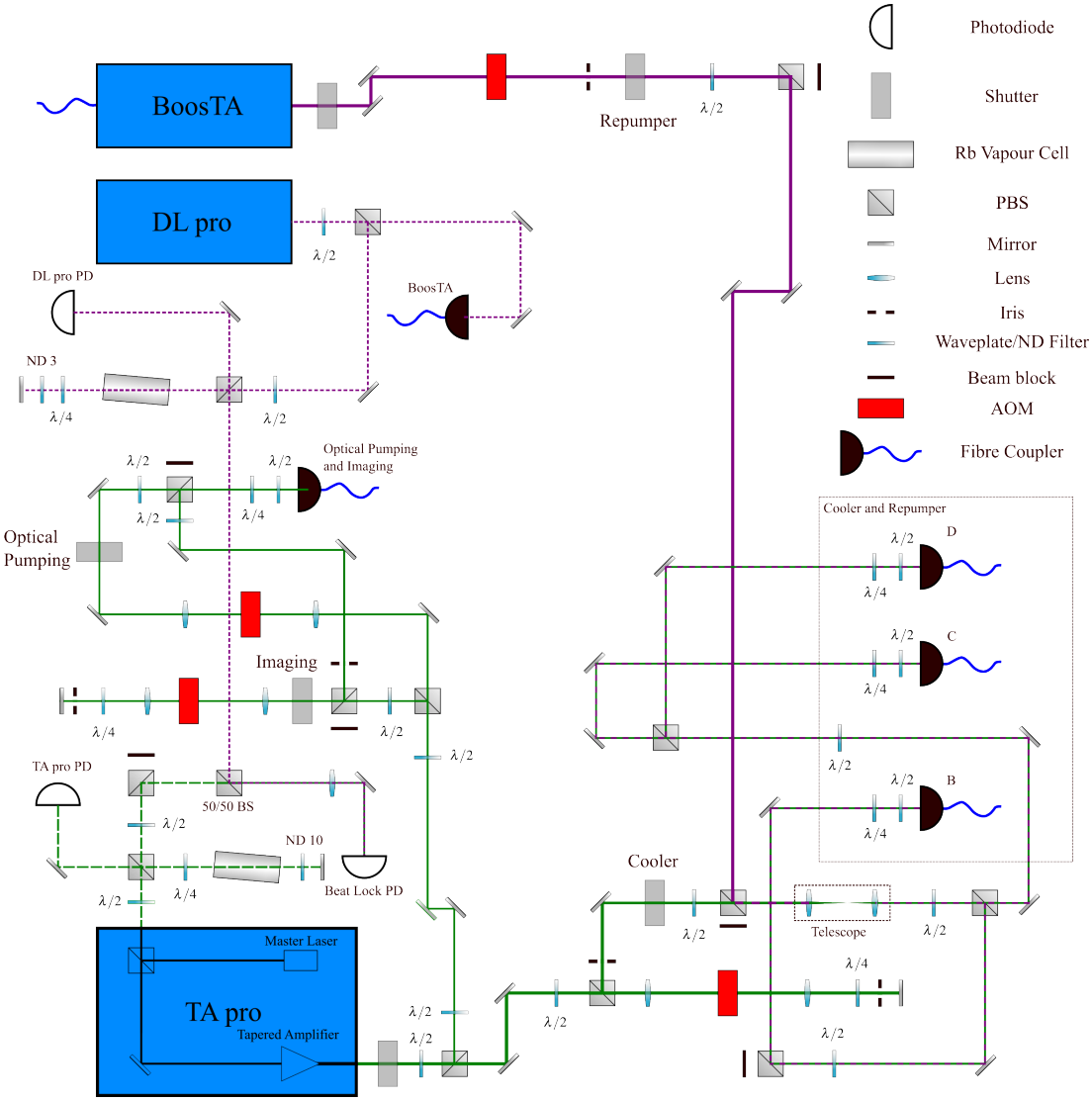


Figure 4.24: The generation of the correct frequencies and subsequent power distribution and fibre coupling is shown. The DL Pro output (purple dashed) is split, with 1 mW used to produce a saturated absorption spectrum, 5 mW going to the beat lock and 30 mW seeding the BoosTA. The TA Pro master oscillator output (green dashed) is split between producing a spectrum (1 mW) and the beat lock (5 mW). The BoosTA output (purple solid) is used to derive the repumper frequency with a single pass 80 MHz AOM before being overlapped with the cooler beam at a PBS. The TA output (green solid) is split between the imaging and optical pumping path and the cooler, where less than 1 mW is required per beam for the imaging and optical pumping so significantly more power is distributed along the cooler path. The imaging frequency is derived by a double-pass 110 MHz AOM, the optical pumping by a single pass 80 MHz AOM and both signals are overlapped and coupled independently to a fibre. The cooler frequency is derived using a double-pass 110 MHz AOM and is overlapped with the repumper beam before a telescope reduces the beam diameter and two PBSs are used to adjust the power distribution between the three fibres B, C and D. Shutters are able to block the imaging, optical pumping, repumper and cooler beams independently.

waveplates to allow us to align the polarisation of the light with the fast axis of the fibre.

The remaining light from the TA is used to derive the cooler beam with a double-pass AOM to blue-shift the light to be red-detuned from the $F = 2 \rightarrow F' = 3$ transition, as required in both the MOT and CMOT stages. Even if a single-pass AOM was able to produce the required frequency shift, a double-pass is essential here, as the detuning must be varied between the MOT and CMOT stages, and the change in diffraction angle of a single-pass AOM would ruin the alignment of the derived beams. The cooler beam is then overlapped with the repumper at a PBS and the beam size is reduced using a telescope to improve the fibre-coupling efficiency. Following this, the beam is split by a pair of PBSs which determine the ratio of power along each of the three MOT axes, B, C, and D. The first PBS splits the power between B and the C, D pair, and the second then splits the power between C and D. In the end, there are around 120 mW in B, 30 mW in C, and 60 mW in D, measured after the fibres.

The light from the BoosTA is passed through a single-pass AOM to be blue-shifted and brought on resonance with the $F = 1 \rightarrow F' = 2$ transition to be used as the repumper beam. This is coupled into the same fibres as the cooler beam by being overlapped at a PBS before the power distribution network. Note that since the repumper polarisation is orthogonal to that of the cooler (due to the PBS) the power distribution is opposite to that of the cooler.

4.5.4 Chamber Beam Paths

The beam paths of the generated cooler, repumper, imaging and optical pumping beams are shown in Figure 4.25 where the cooler and repumper are overlapped before the fibres and distributed together. It is essential for the MOT that the cooler beams are perpendicular to one another and correctly polarised with respect to the quadrupole trap. In the case of a quadrupole trap formed using two anti-Helmholtz coils, the field is cylindrically symmetric about the coils axis, so the horizontal beams can be positioned at any angle so long as they are perpendicular to each other. The vertical beam must be aligned with the central axis of the coils. Similarly, the imaging and optical pumping beams are overlapped before the fibre. They are aligned with the y-axis so that the entire sequence can be imaged without having to adjust the focus of the camera as the atoms move along the x-axis.

4.6 Absorption Imaging

The atoms are imaged on a charge-coupled device (CCD) type camera[7] using the shadow cast by the atoms on the resonant light. This allows us to measure the attenuation of the laser light, and thus the density of the cloud. As shown in Figure 4.26 the imaging light is emitted from a fibre and collimated using a lens with a focal length of 125 mm to produce a beam with a diameter of 23.4 mm, calculated

[7] The Imaging Source DMK 21BF04.H

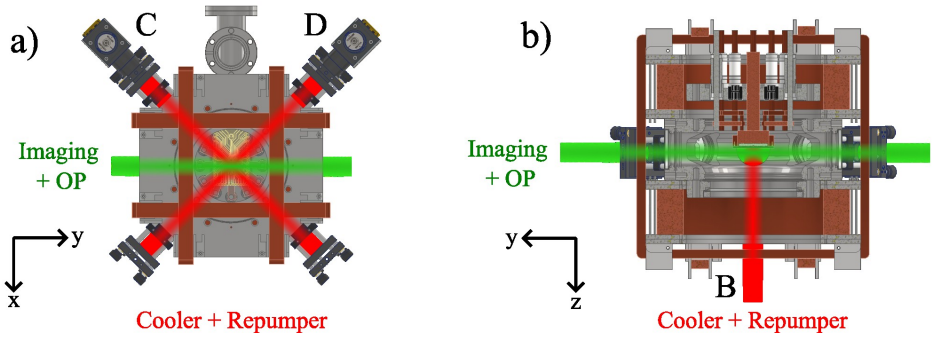


Figure 4.25: The paths of the lasers throughout the chamber are shown with the cooler and repumper beam in red and the imaging and optical beam (OP) in green. Details of the chamber optics for the cooler and repumper are found in section 4.3.2. a) Shows a bottom-up view with the imaging and OP horizontal along the y-axis and the cooler and repumper beams at 45° from the y-axis, but perpendicular to each other. b) Is a side-on view showing the imaging and OP beam again, but with the vertical cooler and repumper beam.

using Equation 4.7. The beam is made as wide as possible, whilst not reflecting off the chamber since the imaging and optical pumping are superimposed. The imaging occurs near the chip and the pumping near the position of the MOT, which is approximately 15 mm horizontally and 15 mm vertically away from the chip. A quarter waveplate is included following the lens so that the optical pumping beam can be correctly polarised.

On the opposite side of the chamber, a pair of lenses is used to shrink the imaging light before it is directed onto the CCD. The first lens has a focal length, f_1, of 150 mm and a diameter of 50.8 mm, whilst the second lens has a focal length, f_2, of 40 mm and a diameter of 25.4 mm. This pair results in a magnification, $M = f_2/f_1 = 40/150 = 0.267$. Finally, the light is shone onto the CCD which has a pixel size of 5.6 μm $\times$ 5.6 μm, thus each pixel images an area of 21 μm $\times$ 21 μm.

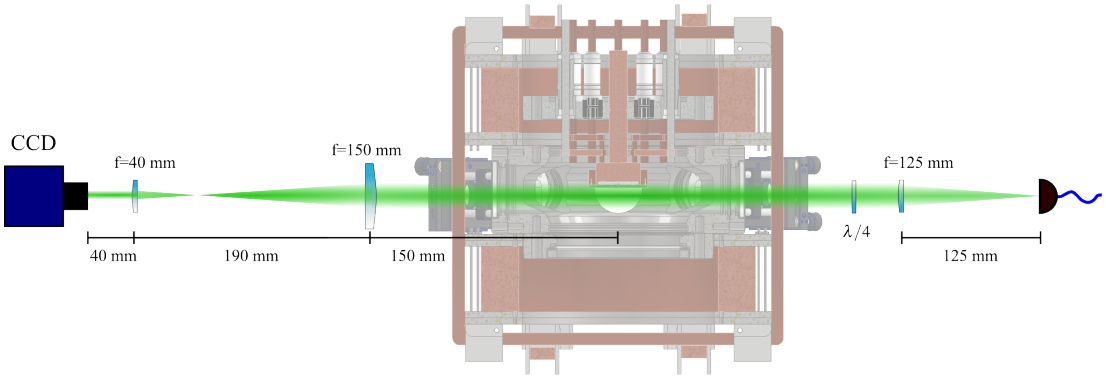


Figure 4.26: The imaging light is first collimated using a lens with focal length 125 mm, then a pair of lens with focal lengths 150 mm and 40 mm magnify the image before it is incident on the charge-coupled device (CCD) camera.

To estimate the atom number from the imaging, we can use the procedure described by Reinaudi

et al.[66]. This begins with defining *optical density*, *OD*, from the Beer-Lambert Law which states that

$$I(x,y) = I_o(x,y)e^{-OD(x,y)}, \tag{4.10}$$

where I is the final intensity and I_o is the initial intensity. The optical density thus tells us how light is attenuated through the atomic cloud. To estimate the initial and attenuated light intensities, images are taken without and with the atoms respectively. The image without atoms is called the light image, I_{light}, and the image with the atoms is called the atoms image, I_{atoms}. A dark image I_{dark}, is also taken and subtracted from the previous two images to remove noise and ambient light. Using these images, the optical density is found to be

$$OD(x,y) \propto \frac{\log(I_{\text{atoms}}(x,y) - I_{\text{dark}}(x,y))}{\log(I_{\text{light}}(x,y) - I_{\text{dark}}(x,y))}. \tag{4.11}$$

The atom number can then be obtained by summing over the optical density for each pixel, scaling by the magnified area of each pixel, and dividing by the resonant absorption cross section, σ_o, which can be found in Steck[81]. This gives the total atom number

$$N_{\text{atom}} = \frac{x_{px}^2}{M^2\sigma_0}\sum_{x,y} OD(x,y), \tag{4.12}$$

where x_{px} is the pixel length and M is the imaging magnification. Figure 4.27 shows the combination of actual atom, light and dark images to produce a final processed absorption image.

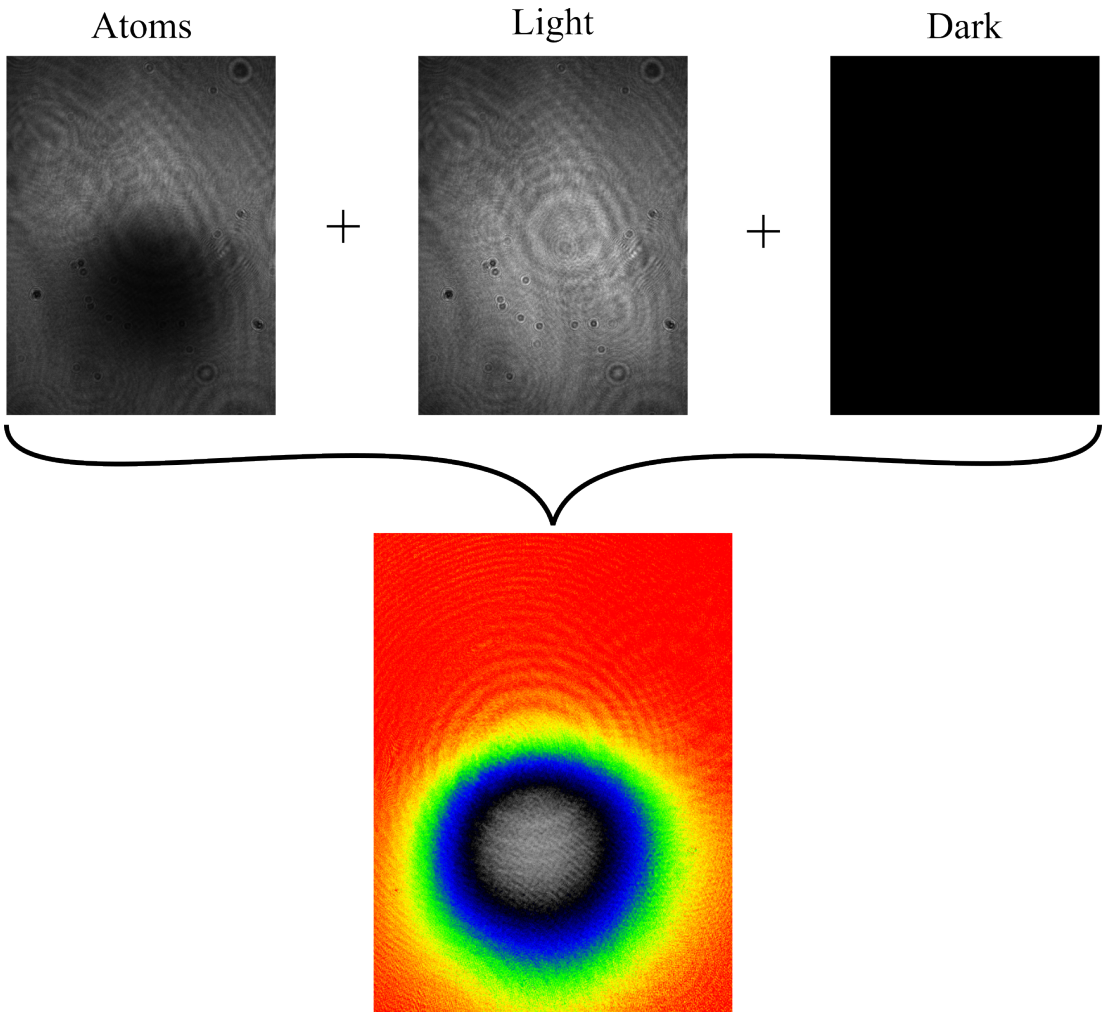


Figure 4.27: Absorption images of the atomic cloud are formed by three separate images. An atoms image shows the shadow cast by the atomic cloud. A light image shows the imaging beam once the atoms have been removed. Finally a dark image is used to remove background light and dead pixels.

Chapter 5

Experimental Sequence

Whilst not complete, progress has been made on the experimental sequence required to load the chip0 Z-trap. In this chapter, the results so far are discussed. The sequence, as of writing, is summarised in Figure 5.1 and ends with the loading of atoms into an Ioffe trap formed by the copper understructure. However, at this point, there are several problems which will be introduced.

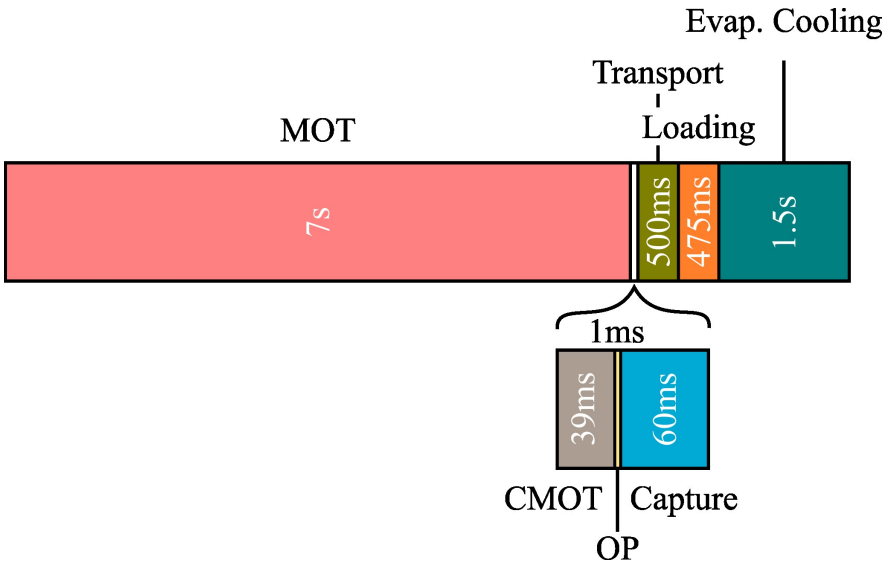


Figure 5.1: The experimental sequence is summarised consecutively with approximate durations for each stage.

As shown in Figure 5.1, the sequence begins with the MOT stage where atoms are accumulated, trapped and cooled, which lasts for 7 s and results in $\sim 4 \times 10^8$atoms. The magnetic trap is then ramped down whilst the cooler detuning is increased to further cool the atoms via polarisation gradient cooling for 35 ms in the CMOT, leaving the atoms at $\sim 67\mu$K. The y-bias field is ramped up to produce a well-defined quantisation axis and, once the cooler beam has been turned off, the atoms are optically pumped for 1 ms into the stretched $|F = 2, m_F = 2\rangle$ state whilst falling. We then quickly turn on the MOT coils to form a magnetic trap and catch $\sim 5 \times 10^7$atoms, implying an efficiency of 10%. However, this is

actually an underestimate in the atom number since the imaging occurs with the atoms magnetically trapped, shifting resonance at different positions. The captured atoms are then transported to below the chip where $\sim 4 \times 10^6$atoms are loaded into a copper Z-wire trap. This atom number is, again, an underestimate, but the transfer efficiency of this stage is still very poor. In addition, the lifetime in this trap is found to be 3.7 s which is surprisingly short for this sort of trap where a lifetime in the tens of seconds is expected.

5.1 Laser Cooling

As described above, the first task is to capture thermal atoms released from a dispenser and cool them sufficiently for magnetic trapping. Starting with a background pressure of $\sim 3.5 \times 10^{-11}$ mbar, which varies depending on how long the experiment has been idle, the dispensers are activated with a pulsed current of 4.5 A for 6.5 s during the MOT stage and a continuous current of 2 A. Pulsing is done to both maximise the available atoms during the MOT stage, whilst minimising the background pressure during trapped stages, and to maintain the temperature of the dispensers for faster activation. The background pressure during the pulsed stage typically peaks at $\sim 6 \times 10^{-11}$ mbar.

Several components are required for the six beam MOT as described in Section 2.3.2: circularly polarised cooler light of the correct helicity along three perpendicular axes, repumper light, atoms and a quadrupole trap. Whilst the MOT is relatively forgiving, any significant atom number requires careful optimisation of the component parameters. Figure 5.2 shows several of the optimisation stages where, with the correct cooler polarisation, the parameters to be optimised are the cooler power and detuning, the repumper power and detuning and the quadrupole trap gradient. In all of the scans, the atom number was measured 3 ms after the MOT coil currents are ramped down to 0 A with the cooler and repumper AOMs switched off. If the atoms are imaged without being allowed to expand, they are dense enough that the imaging light is very strongly attenuated over most of the cloud, resulting in an underestimate of the atom number. Moreover, in a trap, the imaging beam is detuned depending on the local magnetic field, which also results in an underestimate.

Figure 5.2a shows a scan of the total cooler power measured after the fibres, ranging from 32 mW to 280 mW. The atom number increases linearly until it begins to saturate at around 200 mW. 260 mW is used and an asymmetric distribution of power was found to be optimal with B = 150 mW, C = 40 mW and D = 80 mW (refer to Figure 4.25 to see which beams are labelled B, C and D). In Figure 5.2b the cooler detuning (AOM driver VCO) is scanned from 3.4 V to 7 V where we see a quick drop-off in atom number from the optimal at 5.18 V which corresponds to a red-detuning from the $F = 2 \rightarrow F' = 3$ by 28 MHz. Figure 5.2c is a scan of the quadrupole trap gradient by scanning the MOT coil currents. The control voltages are scanned from 0.8 V to 3.2 V which corresponds to 4 A to 16 A with a peak

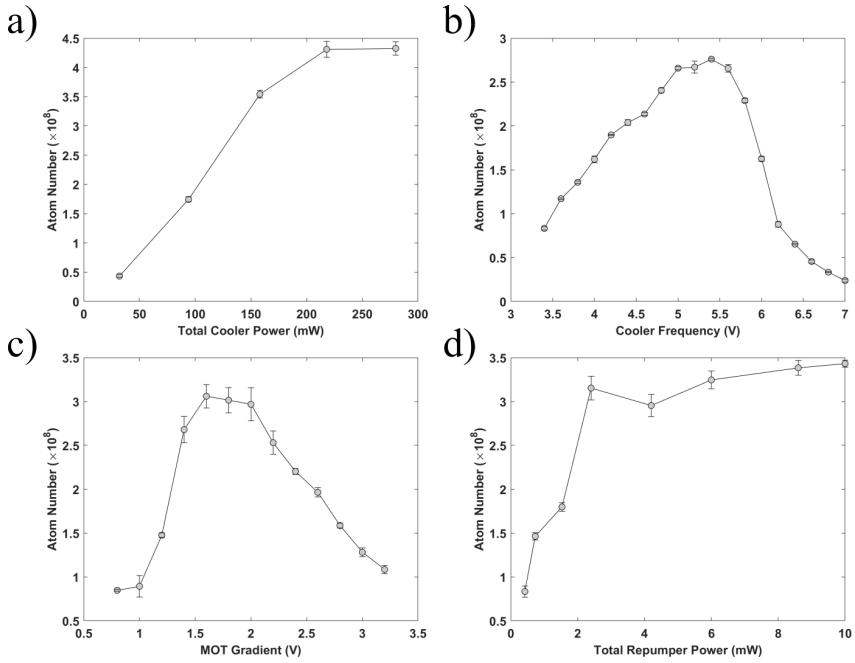


Figure 5.2: Optimisation of the MOT atom number where the atom number is measured after ramping down the magnetic fields, turning off the cooler and repumper AOMs and allowing the cloud to expand for 3ms. a) the total cooler power after the fibres is scanned from 32 mW up to 280 mW and 10 runs are averaged for each data point. b) the detuning of the cooler is scanned by varying the VCO voltage from 3.4 V to 7 V. c) the gradient of the quadrupole is scanned by varying the current in both MOT coils. The scanning range of 0.8 V to 3.2 V corresponds to 4 A to 16 A. d) the total repumper power after the fibres is scanned from 0.42 mW to 10 mW.

at 1.82 V, or 9.1 A. The simulated gradient is 2.7 G/cm/A resulting in an estimated field gradient of 25 G/cm. Lastly, Figure 5.2d is a scan of the repumper power from 0.42 mW to 10 mW. As shown, the atom number increases quickly over the first 3 mW after which it continues to increase slowly as the power increases. It was found during the optimisation that the cooler parameters and magnetic field had a much greater effect on the atom number than the repumper, where the atom number is relatively insensitive to changes in both power and frequency of the repumper.

In addition to these parameters, the duration of the MOT should also be optimised to maximise the atom number with the shortest loading time. Recall that the time evolution for the atom number in a MOT is given in Equation 2.40, and repeated as

$$N(t) = N_{\max}\left(1 - e^{-t/\tau_{\mathrm{MOT}}}\right). \tag{5.1}$$

The loading of MOT is measured in Figure 5.3a by loading the MOT for some time before switching off the lasers and MOT coils then imaging. The atom number was measured for loading times from 1.5 s to 13 s and it is seen that the atom number saturates with an 8 s loading time. However, it should be noted that the loading doesn't precisely follow Equation 2.40, with a more linear loading and a decrease in atom number after the peak at 8 s. This is due to the pulsing of the dispenser which is set high at the start of the MOT stage and lasts 6.5 s which causes the loading rate, given by Equation 2.39, to vary. Figures

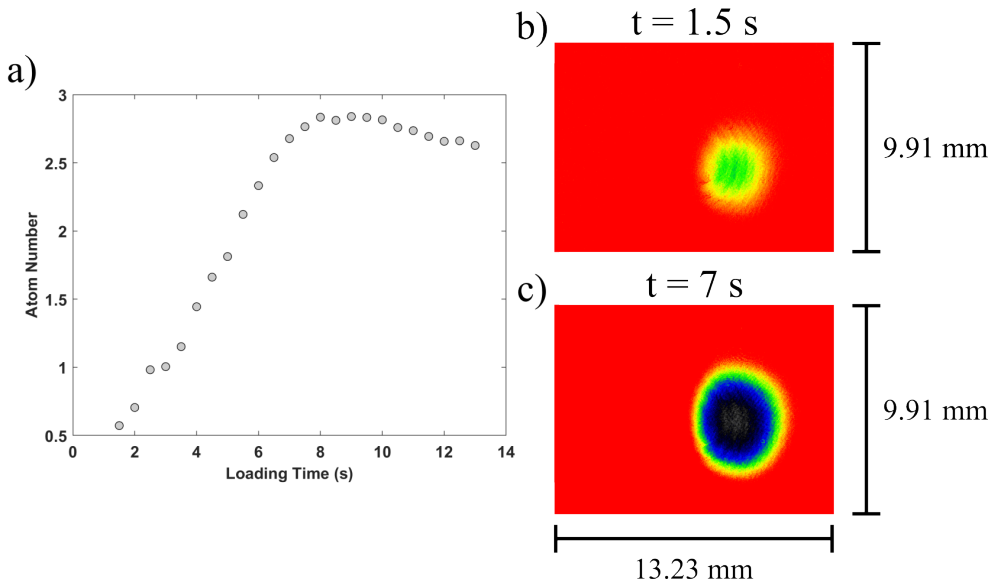


Figure 5.3: The MOT loading rate is measured by scanning the MOT duration from 1.5 s to 13 s. a) shows the atom number at each time. b) shows an absorption image of the cloud at 1.5 s. c) shows an absorption image of the cloud after 7 s. The atom number is measured with the fields off for 3 ms.

5.3b and c show absorption images of the atomic cloud at 1.5 s and 7 s. The average atom number at the end of the MOT is $3.6 \pm 0.04 \times 10^8$.

5.1.1 Sub-Doppler Cooling

Before magnetically trapping the atoms, they must be further cooled below the Doppler cooling limit using polarisation gradient cooling in the CMOT stage as described in Section 2.3.3. To measure the performance of this stage, a temperature estimate is made through the expansion of a free-falling atomic cloud. In 2D, the cloud has an approximately Gaussian distribution with an RMS width in each axis $\sigma_{i=x,y}$ that evolves as

$$\sigma_{i=x,y}(t) = \sqrt{\sigma_{i=x,y}(t=0) + \frac{k_B T}{M_{\mathrm{Rb}}} t^2}, \tag{5.2}$$

where T_i is the temperature along axis i. To estimate the temperature of the cloud, we switch off the lasers which confined the atoms and allow the cloud to expand as the atoms fall. Figure 5.4a shows the expansion of a free-falling cloud immediately following the MOT stage, with b and c showing the cloud at t = 0 ms and t = 5.5 ms respectively. In blue are the σ_x points, which are the measured RMS width in the x-axis, and the fitted Equation 5.2, which gives a temperature $T_x = 1.2 \pm 0.2$ mK. The y-axis equivalent is shown in orange and gives a temperature of $T_y = 0.73 \pm 0.1$ mK. It should be noted that the fit is quite poor, which is possibly due to eddy currents formed by the MOT coils affecting the initial expansion of the cloud.

To reduce the temperature further, the magnetic fields are ramped down quickly, whilst the cooler is further detuned to 58 MHz over 4 ms, and the cloud is left to cool for 31 ms. This stage implements

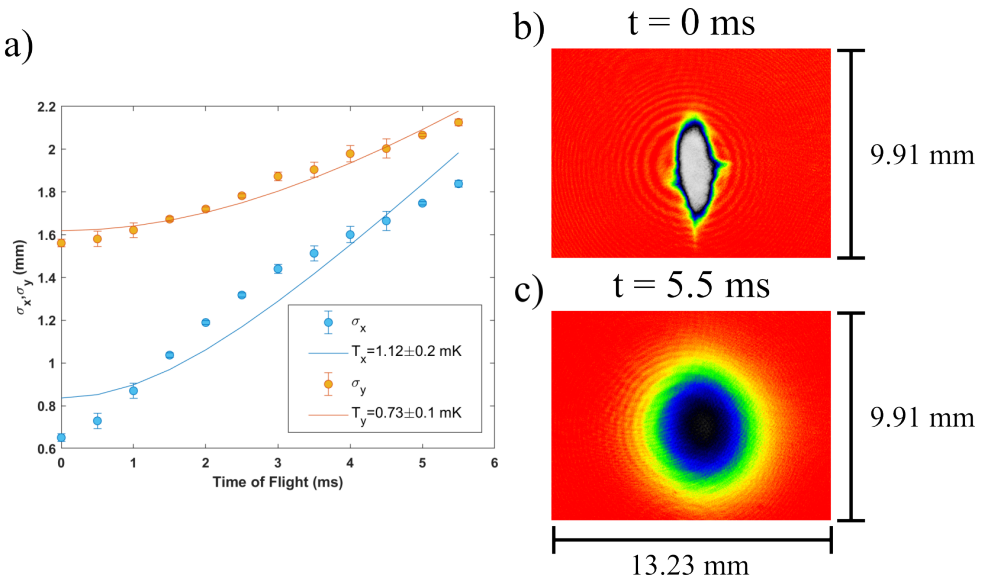


Figure 5.4: The temperature of the MOT is measured by ballistic expansion where the fields and cooler are turned off allowing the cloud to fall. In a) the data points show the measured RMS widths averaged over several runs and the lines show the fitted Equation 5.2 to attain the temperature estimates in each axis. b) and c) show images of the cloud at 0 ms and 5.5 ms respectively.

$\sigma^+ - \sigma^-$-polarisation gradient cooling as described in Section 2.3.3. Ballistic expansion measurements for the CMOT stage are shown in Figure 5.5a with images of the cloud shown in b and c at t = 0 ms and t = 25 ms respectively. Due to the significantly lower temperature, a much longer time of flight of 25 ms was possible. In blue are the measurements of σ_x and the fitted Equation 5.2, which gives an estimated temperature of $T_x = 66.79 \pm 2\mu K$. Similarly, orange gives the σ_y measurements and the estimated temperature of $T_y = 45 \pm 2\mu K$, implying a temperature reduction of over an order of magnitude. As seen by comparing Figures 5.4b and 5.5b, the atomic cloud has expanded during the CMOT stage. This results from the lack of magnetic fields which removes the position-dependent force present in the MOT, with the velocity-dependent force slowing the atoms but not confining them.

5.2 Optical Pumping and Magnetic Capture

At the end of the CMOT stage, the cooler and repumper AOMs are turned off, allowing the atoms to fall. The y-bias field is switched on to define a quantisation axis for the hyperfine states along the y-axis and the atoms are optically pumped to the $|F = 2, m_F\rangle$ state as described in Section 2.3.4. The optical pumping pulse lasts for 1 ms, has a detuning of 20 MHz, and a power of $\sim 500\mu$W. It was found that if the power is increased beyond 1 mW then the resulting scattering force is sufficient to accelerate atoms and significantly reduce the capture efficiency. Whilst optical pumping should be on-resonance in theory, a dense cloud can appear opaque to the pumping pulse, meaning atoms near the centre of the cloud are not reached. Detuning the optical pumping frequency will allow light to propagate throughout the entire cloud, and can often help to improve the pumping efficiency. In the F=2 manifold there are

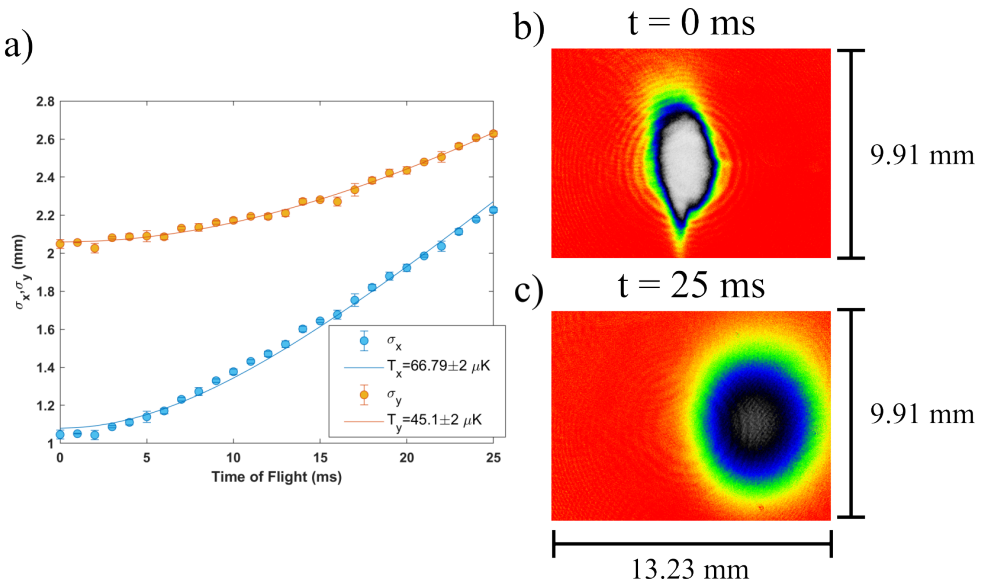


Figure 5.5: The temperature of the CMOT is measured by ballistic expansion where the cooler is turned off and the cloud is allowed to fall. In a) the data points show the measured RMS widths averaged over several runs and the lines show the fitted Equation 5.2 to attain the temperature estimates in each axis. b) and c) show images of the cloud at 0 ms and 25 ms respectively.

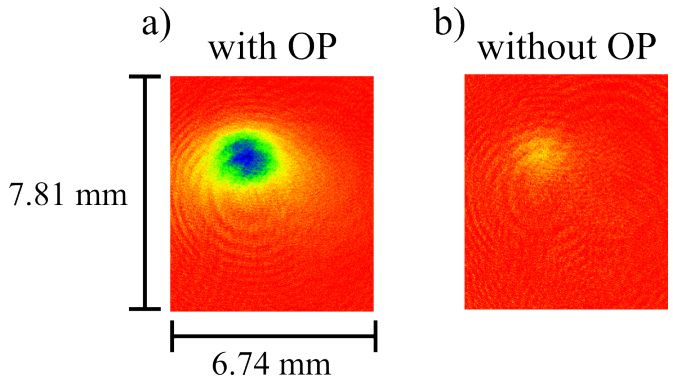


Figure 5.6: a) Absorption image of magnetically trapped atoms with optical pumping. b) Trapped atoms without optical pumping for comparison.

five sub-levels, and assuming an initial even distribution among all five sublevels, roughly 20% of the atoms are in the $|F = 2, m_F = 2\rangle$. Optical pumping should therefore increase the number of captured atoms by a factor of ~ 5. By measuring the captured atom number with and without optical pumping, we find a ratio of 4.54 which implies that the optical pumping is well optimised. Without optical pumping, some atoms in the weak field-seeking state $|F = 2, m_F = 1\rangle$ might also be caught, which reduces this ratio. Figure 5.6 shows absorption images of the captured atoms with optical pumping (Figure a) and without optical pumping (Figure b) for comparison. Including optical pumping, the average number of captured atoms is $4.502 \pm 0.14 \times 10^7$, which implies a capture efficiency of 12.6%. Atoms are caught in a quadrupole trap generated by the MOT coils.

The decay in atom number for a magnetic trap is expected to be governed by

$$N(t) = N(0)e^{-t/\tau}, \tag{5.3}$$

where τ is the characteristic lifetime of the trap and in the quadrupole trap, the primary loss mechanisms are Majorana losses, collisions with the background gas, and magnetic field noise. We can measure the lifetime in the quadrupole trap by measuring the atom number at different hold times and fitting Equation 5.3, which is shown in Figure 5.7. From this, the lifetime was estimated to be 6.657 s.

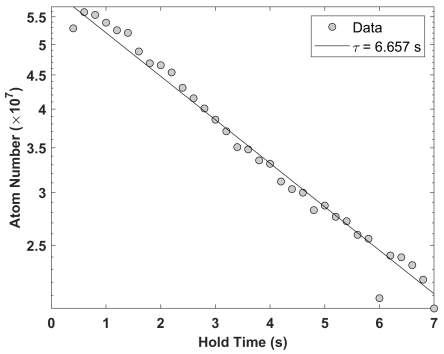


Figure 5.7: The remaining atom number is measured after the atoms are held in the quadrupole trap for some time.

A potential fault in the magnetic capture is the fact that the magnetic trap and CMOT are offset in the x-axis as shown in Figure 5.8, where the atoms fall along the z-axis. The centre of the CMOT and captured atom cloud are offset by $\sim$ 2mm which causes the atoms to be accelerated horizontally towards the magnetic trap, potentially causing unnecessary heating and losses when atoms are accelerated out of the trap. This arises from the fact that the x-bias field is directed to move the quadrupole trap towards the chip, which is in the opposite direction to the CMOT. To resolve this, an H-bridge can be used or a bipolar power supply purchased to switch the polarity of the x-bias coils.

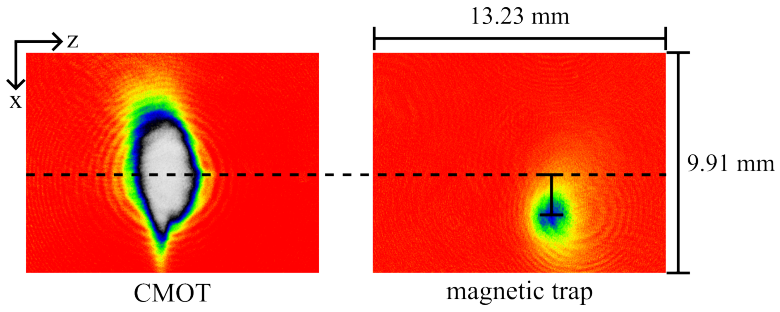


Figure 5.8: Optical density images for the CMOT (left) and atoms in the magnetic trap (right), where the z-axis is the direction of gravity. The dashed line shows the central x-coordinate of the CMOT which is offset from the central x-coordinate of the magnetic trap. The height of the images is 9.91 mm which is given for scale.

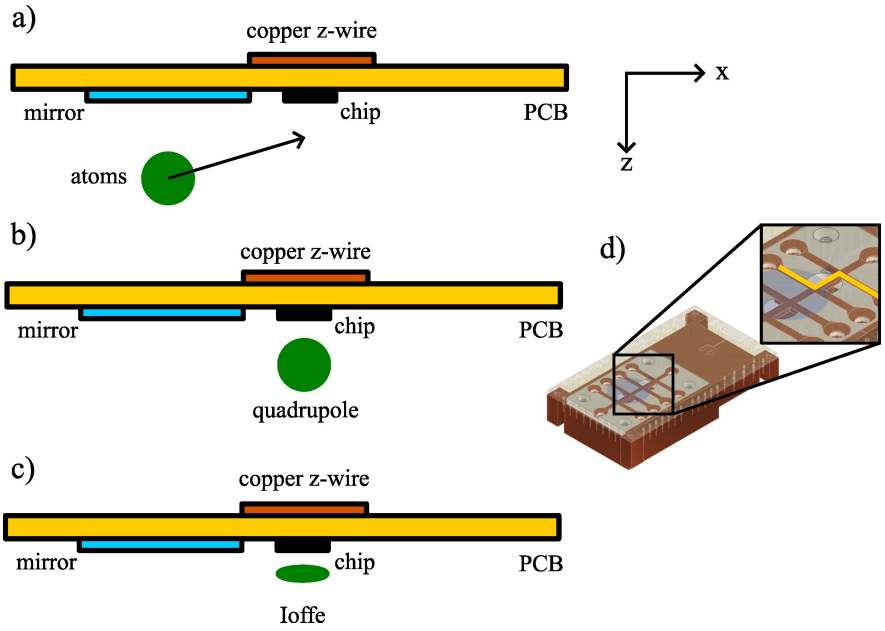


Figure 5.9: Simplified diagrams show the sequence of loading atoms into the Z-wire trap. a) Atoms start in the quadrupole trap at the position of the CMOT which is below the mirror. b) Bias fields are used to move this quadrupole trap to below the CMOS chip. c) The atoms are transferred from the quadrupole trap into an Ioffe trap generated by the copper Z-wire. d) Shows which path on the copper under-structure is used as the Z-wire (indicated by a solid, yellow line).

5.3 Copper Z-Trap Loading

In the original design, the PCB wire structures were meant to provide the intermediate trap necessary to transport the atoms to the chip and sufficiently compress them. Unfortunately, it was found that the outer and inner coils were shorted following the bakeout. The bare resistances for the inner and outer coils, taken directly to the PCB, were measured to be 0.3 Ω and 0.7 Ω respectively. However, the resistances measured at the feedthrough were 0.25 Ω and 0.42 Ω, suggesting shorts within the coils. Moreover, both were shorted to the copper mount. A possible explanation for this is that the thermal paste, used to improve the thermal conductivity between the PCB and mount, decreased in viscosity during the bakeout and spread out to short the vias at the bottom of the PCB. This would explain the shorting with the copper block and the decrease in resistance of the coils.

As an alternative to the PCB, the copper under-structure is used to produce the intermediate trap for compressing the atomic cloud before loading onto the chip. One of the paths in the copper under-structure, as shown in Figure 5.9d, forms a Z-wire positioned directly above the CMOS chip with the CMOS chip located in the middle of the central wire. So long as the polarity of the currents through the copper and CMOS Z-wires matches, then transfer from the copper trap to the CMOS trap should be efficient because of the matching geometries.

Atoms start roughly at the position of the CMOT, as shown in Figure 5.9a, then by ramping the x-bias field and the ratio of the MOT coil currents, this trap can be moved just below the CMOS chip, as shown in Figure 5.9b. The atoms can then be transferred from the quadrupole trap to the Ioffe trap

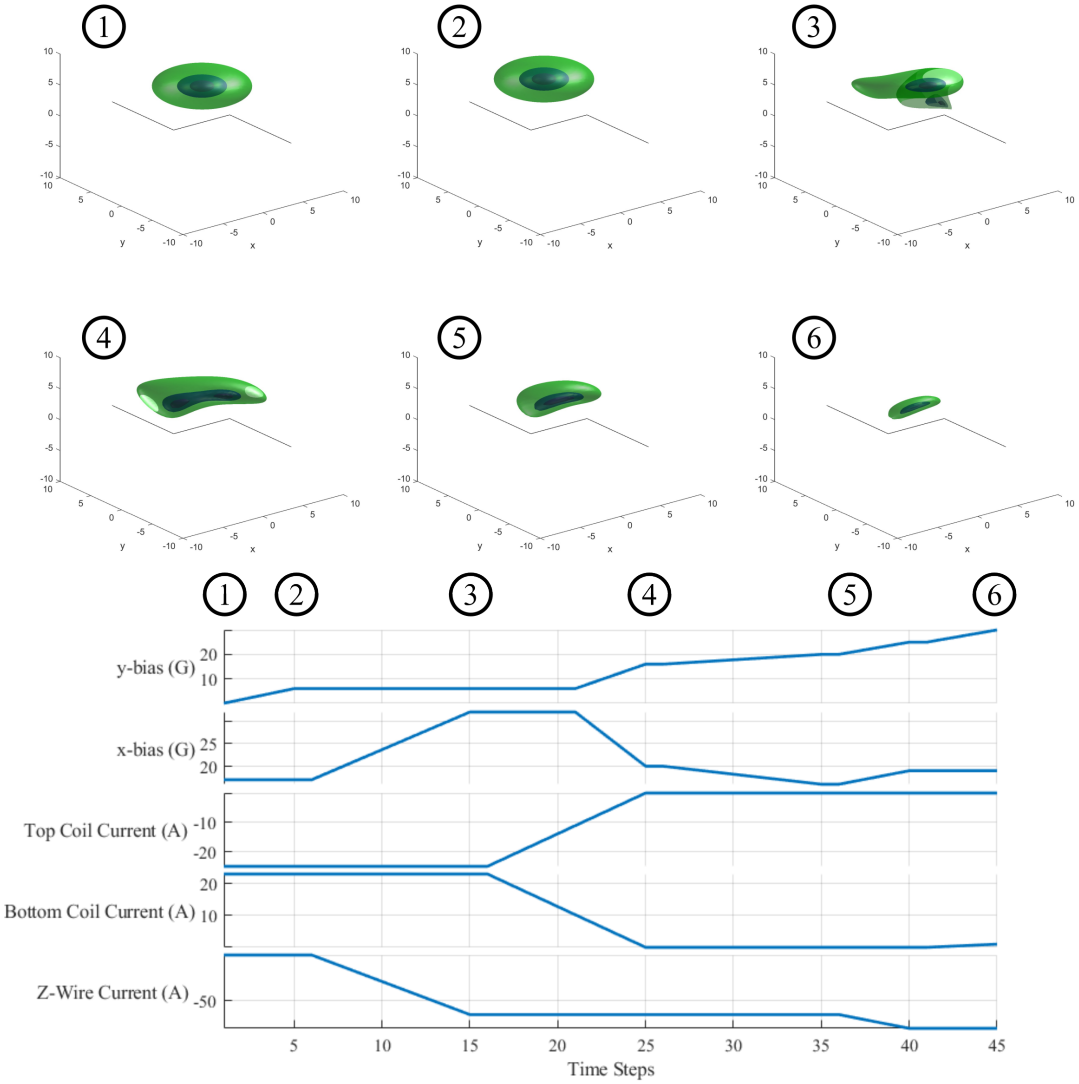


Figure 5.10: Sequence for transferring atoms adiabatically from the quadrupole trap to the copper Z-trap. 1) Atom start in quadrupole trap formed by MOT coils. 2) y-bias field is ramped up. 3) The copper Z-wire current is ramped up with the x-bias field. 4) The MOT coils are ramped down transferring the the atoms into the quadrupole trap formed at the corner of the Z-wire. 4) The x-bias field is ramped down, modifying the trap into an Ioffe trap. 5) The Ioffe trap is compressed and lowered to the surface.

generated by the copper Z-wire as shown in Figure 5.9c. This stage is typically done adiabatically, by quickly turning off the quadrupole trap and switching on the Ioffe trap at the same position. However, this requires precise alignment of the traps, is non-adiabatic so results in some inevitable losses, and without knowledge of the initial Ioffe position can take a long time to achieve. These issues, and the further constraints of an initially positive x-bias and a unipolar y-bias power supply, motivated the development of a novel transfer scheme which is summarised in Figure 5.10.

The transfer scheme is split into six primary stages, beginning with the quadrupole trap, and modifies the x-bias, y-bias, MOT coils, and Z-wire current to finish with the atoms in a compressed Ioffe trap. Initially, the quadrupole trap is located at the centre of the Z-wire and the y-bias field is ramped up, pushing the quadrupole trap to the side of the central wire. It should be pointed out early that the correct

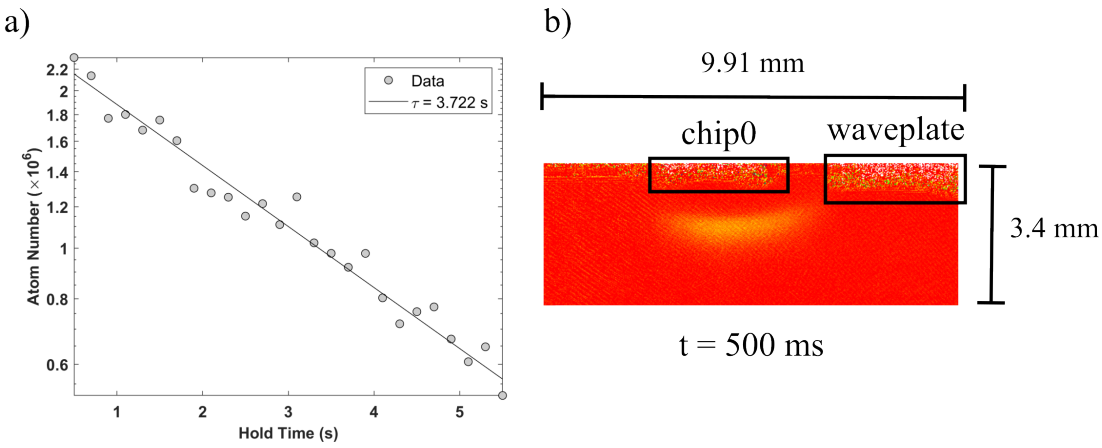


Figure 5.11: a) Lifetime measurement for atoms in trapped using the copper Z-wire, where the lifetime is estimated to be 3.7 s. b) Optical density image of atoms in the copper Z-wire trap.

polarity of the fields and currents is essential for this scheme to work. Explicitly, a positive y-bias should correctly bias the Z-wire to form an Ioffe trap, and a positive x-bias should reduce the Ioffe field. The quadrupole trap should be in the sense that the coils produce fields that point away from the opposite coil. After stage two, the Z-wire current is ramped up to 65 A, which is simulated to be sufficient for a trap depth of over 20 G at a trap height of 1.5 mm from the surface of the PCB. Simultaneously, the x-bias field is ramped up, which pushes the quadrupole trap to the corner of the Z-wire. During this stage, a second trap forms at the surface of the PCB; it is therefore important that the x-bias and y-bias are tuned to maintain sufficient distance between these two quadrupole traps.

The key idea of the scheme is to overlap the quadrupole trap formed by the coils with one of the two quadrupole traps formed by offsetting the Ioffe field produced by a Z-wire. From here the MOT coils can be switched off, and if the traps are well matched, this transfer will be adiabatic. This is stage four of the scheme. Between stages three and four, the y-bias is increased and the x-bias is decreased to maintain the position of the atoms. To load the atoms into the Ioffe trap, the x-bias field is simply lowered to below the x-component of the Z-wire field. Lastly, the trap is compressed and lowered by increasing the Z-wire current, y-bias, and x-bias.

Using this scheme, atoms were successfully loaded into the copper Z-trap as shown in Figure 5.11 with a measured atom number in the trap of $3.55 \pm 0.53 \times 10^6$ atoms. This was measured in the magnetic trap, implying an underestimate in the atom number. The lifetime of atoms in the trap was measured to be 3.7 s as shown in Figure 5.11a.

Both the low lifetime and poor transfer efficiency seem to arise from the outgassing of the copper Z-wire. When a high current passes through it, the vacuum pressure rapidly increases, reaching its peak at 8×10^{-10}mbar when a current of 80 A is continuously passed for 3 s. This increase in background pressure will significantly limit lifetimes in the trap, making evaporative cooling at this stage untenable. Moreover, the necessity to minimise the Z-wire current forces the transfer scheme and final trap position

to be very close to the surface of the CMOS chip, limiting the trap depth on this side and allowing losses into the surface.

Chapter 6

The Design of Chip 1

The second generation chip, referred to as chip1, expands on the concept of integrated sensors whilst adding a programmable RF source and beginning to simplify interfacing with the chip through the addition of a serial interface, called the SPI as shorthand for serial peripheral interface. The design goal of this chip was primarily to prototype a variety of components that are necessary to produce a Mach-Zender interferometer, such as that found in[9]. The atoms are evaporatively cooled to form a BEC which is then coherently split using adiabatic potentials, equivalent to a beam-splitter in the optical interferometer and the atoms are left to accumulate a phase. The atoms are then partially recombined by abruptly reducing the spacing which allows the separate wavepackets to interfere and the relative population difference between the spatial modes is measured to estimate the phase difference. To accomplish this, we begin with a Z-wire trap with two RF wires positioned as shown in Figure 6.1 where switches are added to the shorter Z-wire path as in chip0 to allow in situ modifications to the trap geometry. To produce the RF currents required for the evaporative cooling and RF dressing, a direct-digital synthesiser (DDS) is included. The DDS produces a voltage output that is converted to a current with a high amplitude (400 mA along the RF wires) using a transconductance amplifier. This is connected to a switch which allows us to select between the RF current from the DDS or from an external source. Ideally, we would be able to program the RF amplitude; however, this was not implemented due to time constraints.

With RF control, the next step is to be able to measure the relative population differences between the spatial modes, which requires precise, position-dependent fluorescence measurements. To this end, two kinds of photodetectors have been added to the chip: firstly avalanche-photodiodes (APDs), which are able to detect and measure low-light intensities through the avalanche effect, where a strong electric field at a PN junction causes electron-hole pairs generated by an incident photon to produce additional pairs, amplifying the produced current. Secondly, single-photon-avalanche diodes (SPADs) use a higher electric field so that the gain is sufficiently high for a single photon to be detected, triggering a counter. Having photodiodes on either side of the trapping wires allows us to measure the population difference

through the different light intensities generated by fluorescence imaging of the two spatial modes.

The design of chip1 was split among several people: from FHWN, Lukas Rennhofer designed the photodetectors and readout circuits, Luca Reider and Mario Beisteiner designed the Hall probes and DAC, Prof. Helmut Frais-Kölbl provided extensive advice, and Dr. Christian Koller added the pins, Z-wires (including switches), and helped throughout. The author was primarily responsible for the SPI, DDS, PLL, and clock. Lastly, Julius Hansen from Sarcura (https://sarcura.com) provided advice, and staff funding for Lukas Rennhofer.

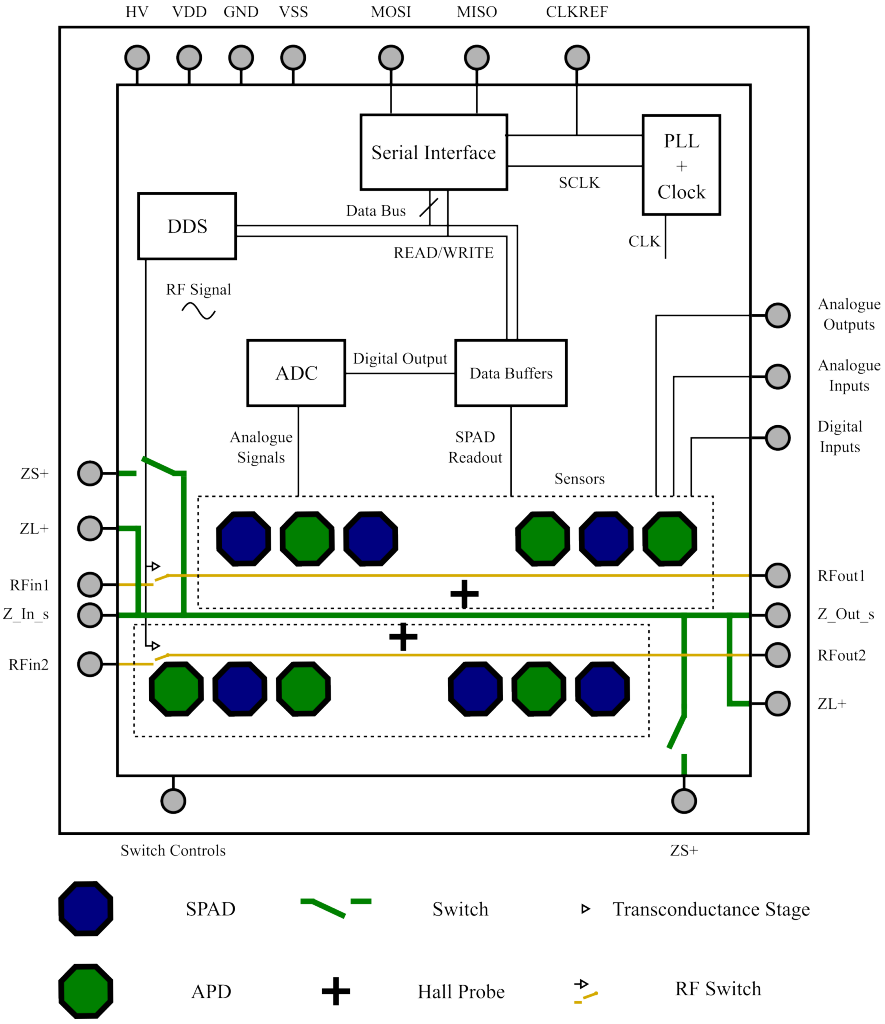


Figure 6.1: Simplified overview of chip1: the Z wires are shown in green where the long Z wire is passively connected to the pads, but the shorter Z is connected vias transmission gate switches. The pads Z_IN_s and Z_Out_s allow four-wire measurements of the Z wires. Hall probes, SPADs and APDs are shown to either side of the chip where both the Hall proves and APDs produce an analogue output which is sent to the ADC, but the digital output of the SPADs is sent directly to the data buffers which interface with the serial interface. The serial interface allows for programming of the DDS and reading of data from registers on the chip. Everything is synchronised by an onboard clock which is stabilised by being phase-frequency-locked to an external reference which doubles as the serial communication clock. The DDS is able to produce an RF signal which is amplified at a high impedance buffer before connecting to the RF wires (yellow) via switches which allow a choice between using the DDS or and external RF source. There are four power connections required, power(VDD), ground (GND), -power (VSS) and high voltage (HV).

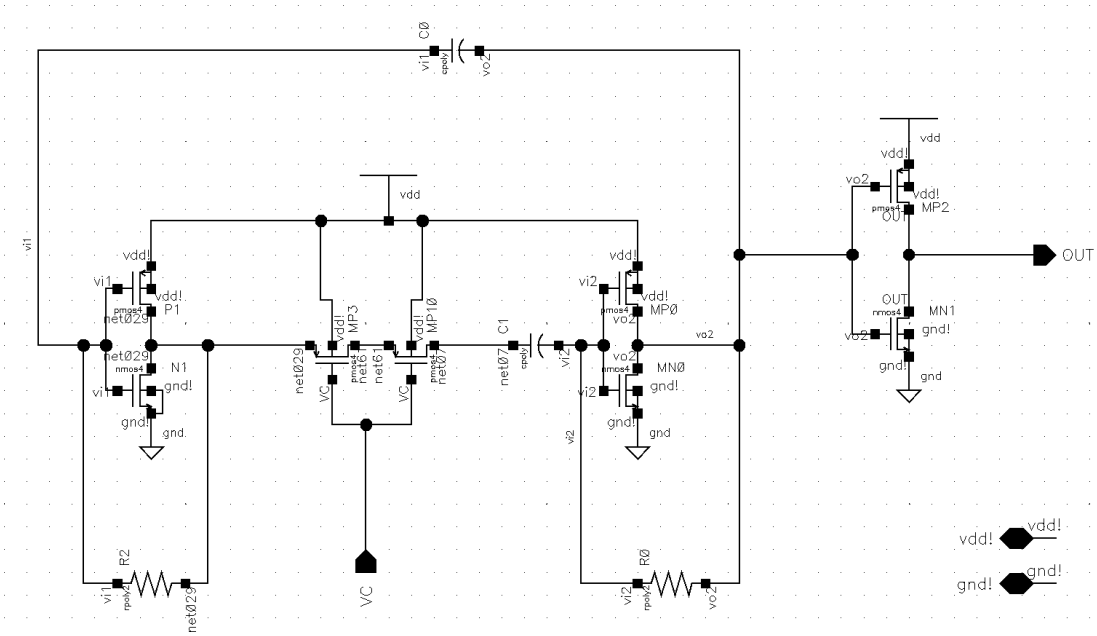


Figure 6.2: Schematic for the VCO first described by Zheng and Saavedra[91].

Chip1 is fabricated using an Austria-Microsystems (AMS) 360 nm, 120 V process. This indicates that transistors have a minimum length of 360 nm and a maximum voltage of 120 V. However, not all components are able to handle this high voltage without damage, and specialised MOSFETs must be used at voltages higher than 10 V. The chip is designed with a VDD of 3.3 V and must include a 0 V GND, a -3.3 V VSS (for some of the amplifiers) and a high voltage (<70 V) for biasing the photodiodes. The naming convention of VDD and VSS arises from the N-channel MOSFET, where the lowest potential is the source and the highest is the drain. Correspondingly, the lowest (highest) chip potential is labelled with an S (D), with the double letter indicating a power voltage. So VSS, also called V_{SS}, is a power rail and the lowest voltage on the chip.

The design was accomplished using the Cadence custom IC product suite (https://www.cadence.com/en_US/home/tools/custom-ic-analog-rf-design.html), more specifically Virtuoso for both circuit and layout design and Spectre for the simulation of the circuits.

6.1 Clock and PLL

A phase-locked loop (PLL) is a circuit that compares two clock signals: a reference signal and an internally generated signal with an adjustable frequency. It uses this comparison to provide feedback to the internal clock, adjusting its frequency and phase to match that of the reference clock. This is used on the chip to frequency stabilise the integrated clock by phase locking it to a much higher quality external reference. Phase locking also reduces the likelihood of errors in the serial interface by ensuring the transceivers at both ends are in phase.

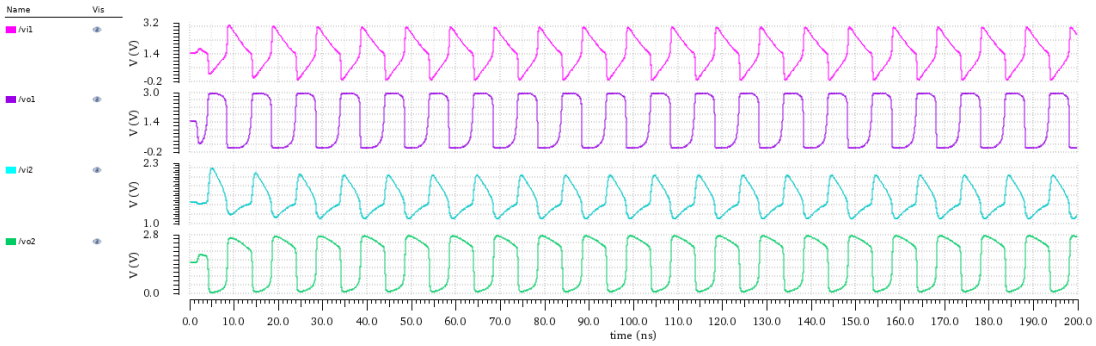


Figure 6.3: The voltages at the inputs and outputs of the inverters in the VCO.

VCO

The internal clock signal is generated by the VCO shown in Figure 6.2. This design was adapted from that of Zheng and Saavedra[91]. The core consists of two inverters with parallel resistors forming a loop, separated by capacitors, with a third inverter acting as an output buffer. Calling the left and right inverters A and B respectively, we see that if the inputs at A and B, see vi1 and vi2 in Figure 6.3, respectively, are initially 0 (0 V) and 1 (3.3 V), then the output of A, v01, is connected to VDD and the output of B, vo2, is connected to GND. The capacitor after A then begins to charge with current following the path from VDD to GND through the capacitor followed by the resistor, which is parallel to B. Similarly, the capacitor after B is charged through the path that includes the resistor parallel to A. Thus, after some time determined by the sizes of the capacitors and resistors, the input at B drops to the threshold voltage, and the output switches to VDD which connects the input at A to VDD and therefore the output drops to 0. The state at the output of B is then seen to oscillate between 0 and 1 at a frequency determined by the time constant of the RC channels, which is fed into an additional inverter to produce a square wave and to act as a buffer, preventing loads from affecting the frequency of the clock. The frequency of the clock is then tuned by the additional PMOS transistors in series with the capacitor after A, which have a resistance determined by the control voltage, V_C, allowing us to control the RC time constant for the charge/discharge paths. The voltages at the inputs and outputs of the inverters are summarised in Figure 6.3.

Sizing for the circuit was determined using simulations, with the frequency range centered around 105 MHz producing the simulated frequency range shown in Figure 6.4. As shown, the VCO has a minimum frequency of 100 MHz and a maximum output frequency of 113 MHz with a control voltage ranging from 0 to 450 mV. The sweep is roughly linear in this region.

The layout of the VCO is shown in Figure 6.5, where we can see that the clock is fairly compact due to the small capacitor and resistor sizes required for 105 MHz operation. In CMOS technology, capacitors are simply square metal plates separated by a silicon-oxide layer, with the capacitance determined by the area and placed in the bottom of the layout. Resistors are typically either paths of n-doped

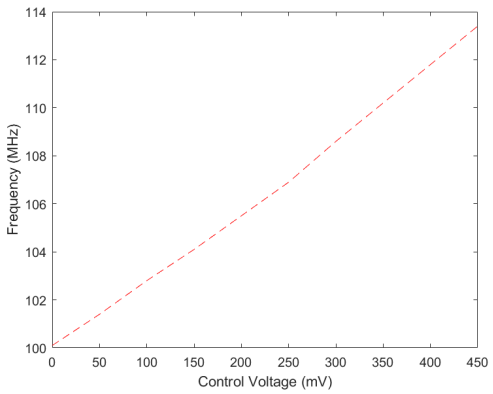


Figure 6.4: Simulated frequency range for the VCO.

silicon or polysilicon, where the resistance is determined by the length and width of this path. We see in Figure 6.5 the resistors, which are in the top part of the figure, snake from left to right so that a long path can fit in a smaller region, allowing for higher resistances. Finally, the transistors are placed in between the resistors and capacitors, with the PMOS transistors on top connected to the VDD (top blue wire) path and the NMOS transistors below connected to the GND (bottom blue wire) path.

We see from the design of the clock that the frequency is dependent upon having a stable control voltage set to the correct value. Moreover, the resistance and capacitance of analogue components will change depending on temperature, so a constant control voltage won't produce a constant clock frequency, which will cause problems such as frequency drift in the RF-source. Thus, the PLL is necessary to maintain a stable clock frequency.

Reference frequency

The reference signal to the PLL is set to 1.64 MHz, which is the same frequency used in the SPI, and is generated off chip. To compare the 105 MHz internal clock to the 1.64 MHz external clock, the internal signal must be divided down, which is accomplished simply using cascaded toggle-flip-flops (TFFs). To understand the principle behind this, consider a single TFF. Every time the input goes high, this toggles the output of the first TFF; however, as this only happens once per period, the output remains high/low for an entire input period; thus, the output has double the period (half the frequency) of the input. By cascading six TFFs, the frequency is halved six times, leading to a final frequency of $(105/2^6)MHz = 1.64MHz$.

Phase-frequency-detector

This divided signal is compared to the reference signal by the phase-frequency-detector (PFD) which is from[41] and is shown in Figure 6.6. The PFD has two inputs, F_{vco} and F_{ref}, and two outputs, UP and DN. When F_{vco} and F_{ref} are both low, UPb and DNb are charged to VDD and UP and DN are both

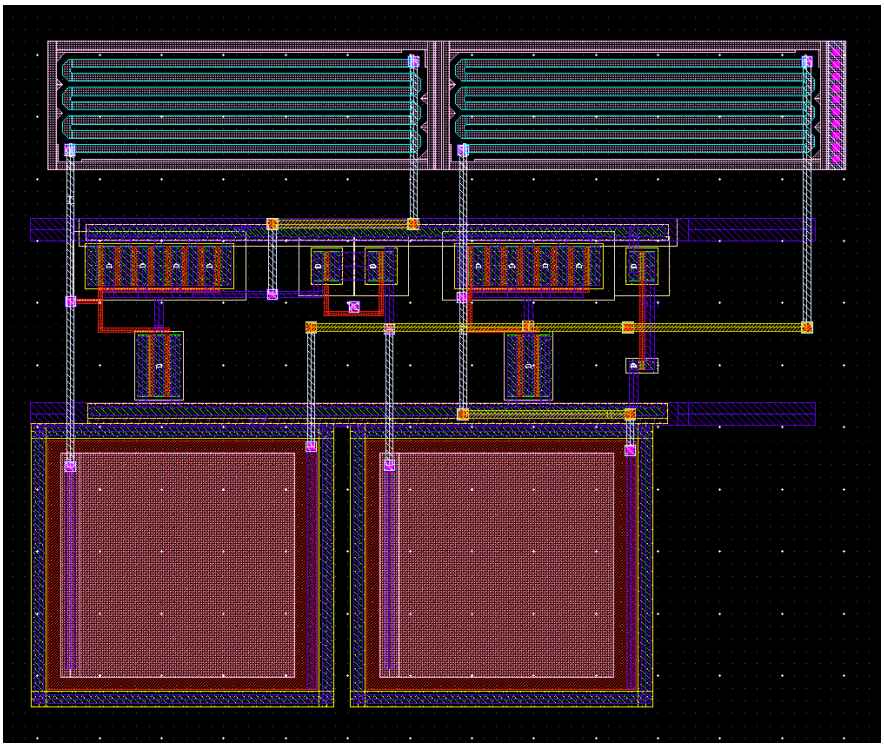


Figure 6.5: Layout for the VCO.

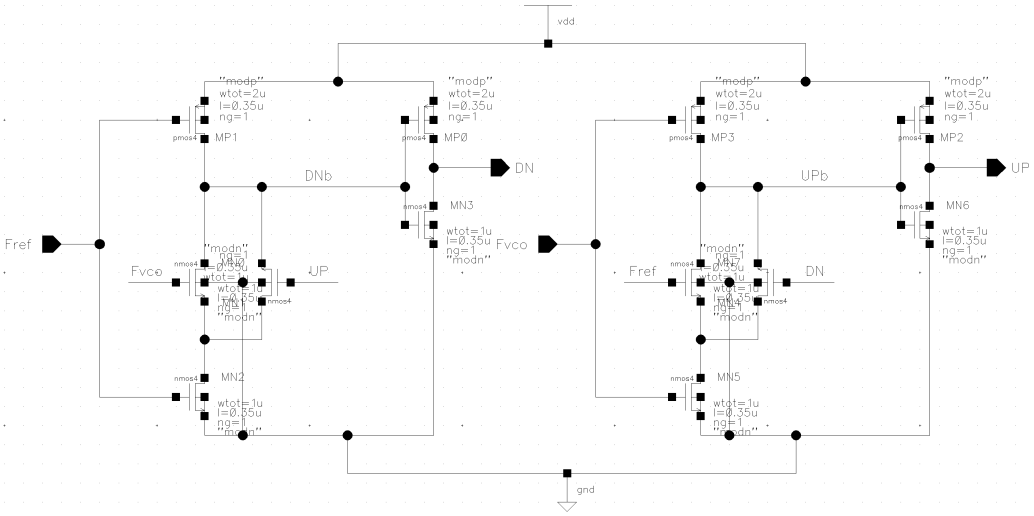


Figure 6.6: Schematic for the PFD.

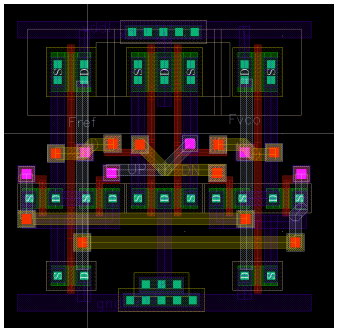
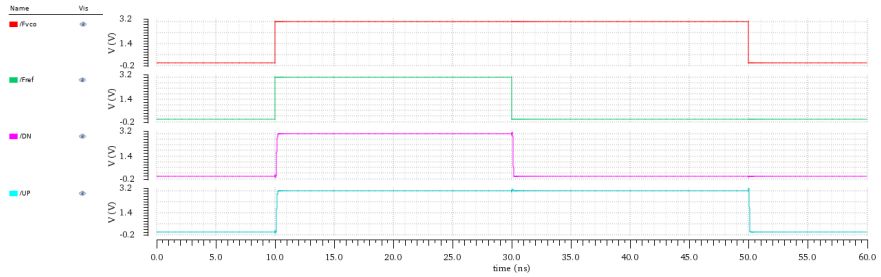
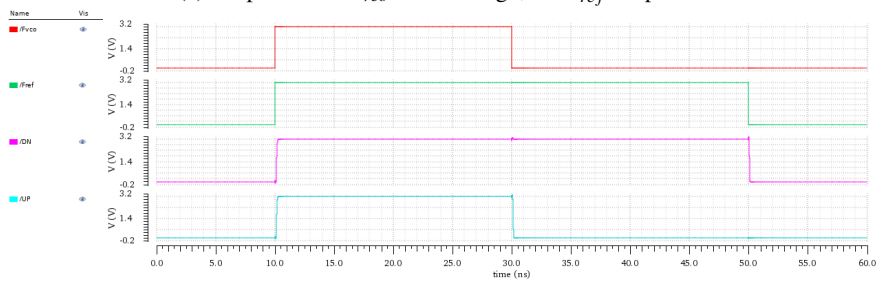


Figure 6.7: PFD layout.



(a) Outputs when F_{vco} remains high, but F_{ref} drops to 0.



(b) Outputs when F_{ref} remains high, but F_{vco} drops to 0.

Figure 6.8: Simulated PFD signals.

low. When F_{vco} and F_{ref} both go high, UPb and DNb are disconnected from VDD but not connected to GND so they remain high, meaning UP and DN stay low. If F_{vco} is high, but F_{ref} goes low, DNb is connected to VDD so DN stays high, but DN and F_{vco} are both high causing UPb to discharge to GND and UP to go high as shown in Figure 6.8 (a). The UP signal is then used to increase the control voltage through the charge pump, increasing the frequency and allowing it to "catch up" with the reference signal. Similarly, as shown in Figure 6.8 (b), if F_{ref} stays high when F_{vco} goes low, DN goes high which lowers the frequency of the VCO output. The layout for the PFD is shown in Figure 6.7.

Charge-pump

To make use of UP and DN signals, they are passed to a charge pump which is used to convert these signals into meaningful changes to the control voltage. The charge pump architecture is from the same work as the PFD,[41], and is shown in Figure 6.9. The underlying principle of charge pumps is to have a current source and current sink, which are controlled by UP and DN respectively, to charge/discharge a capacitor at the output. In the design, we see UP and DN are connected to transistors which act as switches to either disconnect or connect the current source/sink, which are implemented using current mirrors with the current set by the sizes of the PMOS transistors in the middle which are connected to VDD. Other transistors are added to improve matching between the source and sink.

At the output, there is an RC low-pass filter to prevent unstable oscillations of the clock frequency. This is implemented with a capacitor and a transmission gate with the transistor gates set to an intermediate voltage such that they operate in the non-saturated region and have a high resistance; this allows us to use a far smaller capacitor than would be required if a passive resistor had been used. Finally,

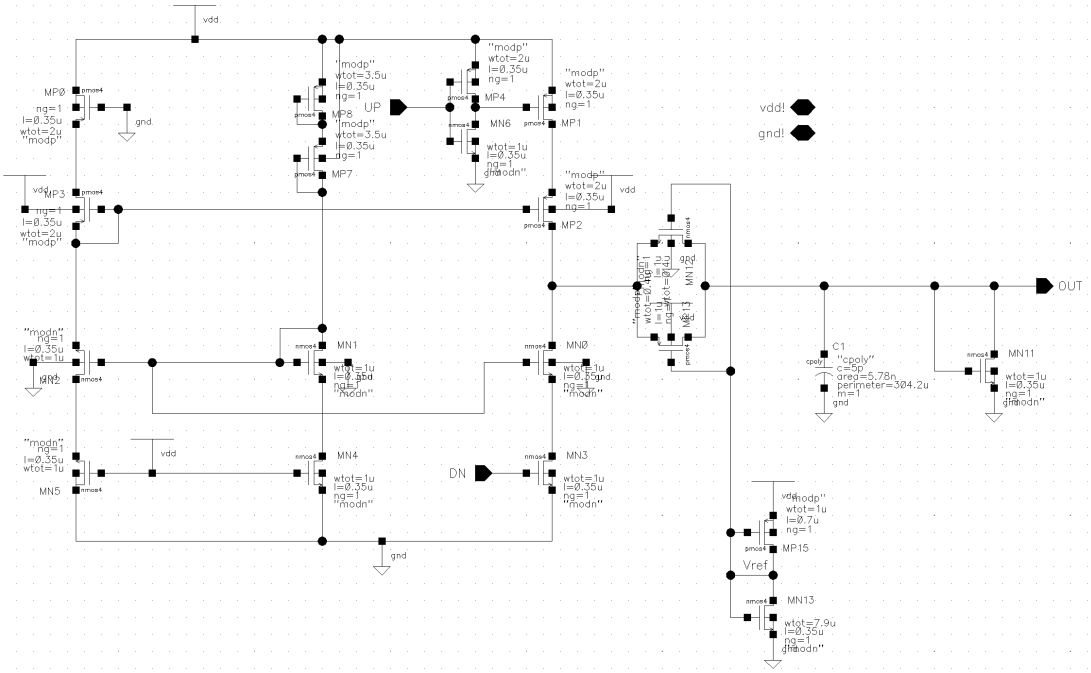


Figure 6.9: Charge-pump schematic.

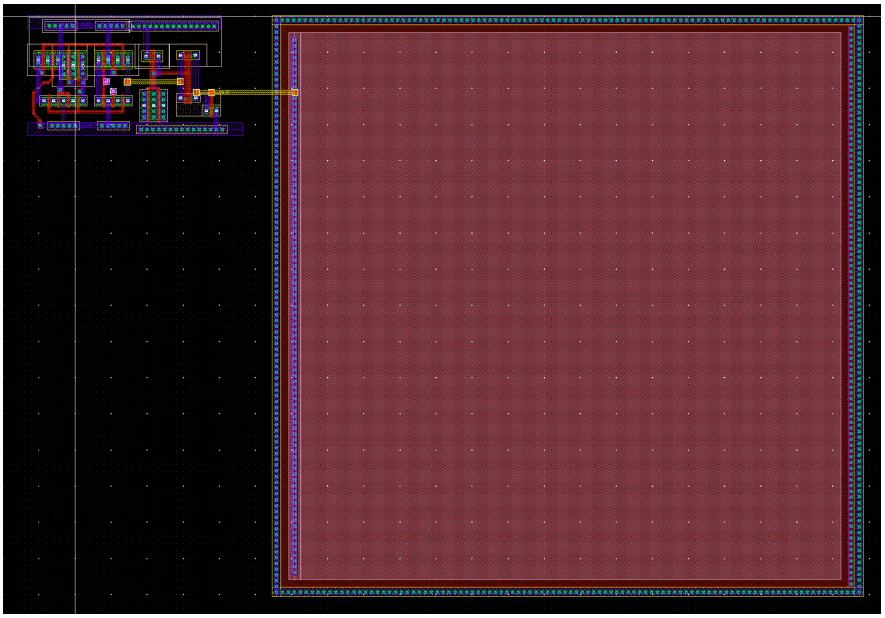


Figure 6.10: Charge-pump layout.

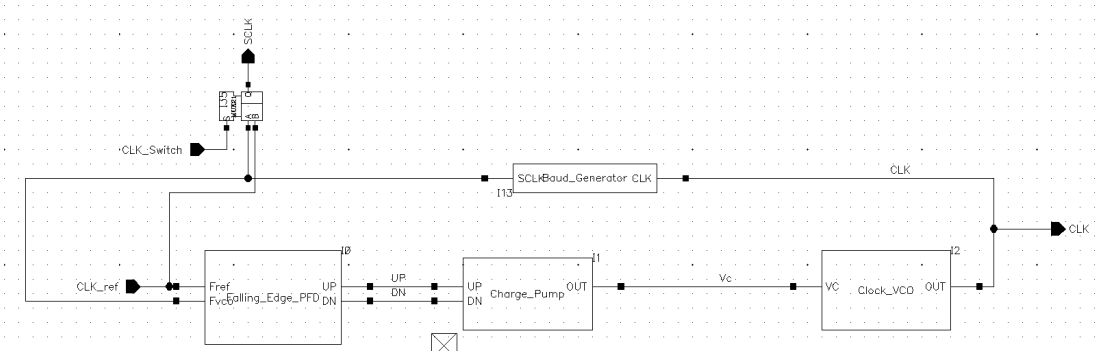


Figure 6.11: Block diagram for the PLL.

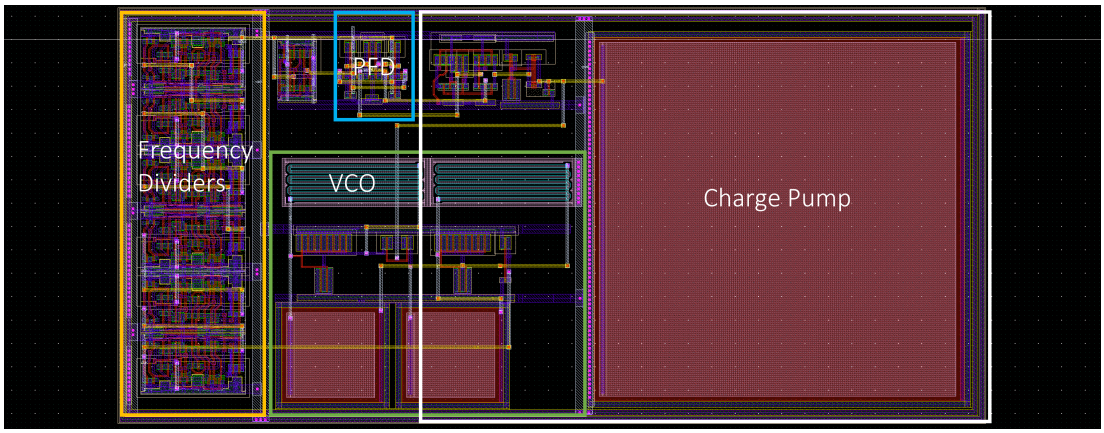


Figure 6.12: PLL layout.

a pull-down NMOS is used to limit the control voltage; this is to prevent the PMOS transistors in the VCO from turning off completely. The layout for the charge pump is given in Figure 6.10, where we can see that the capacitor is by far the largest part of the PLL, with the rest of the charge pump located in the top left corner.

Full PLL and locking

Finally, the full PLL is given in Figure 6.11, where we see the feedback loop for the generated clock signal with the frequency divider, PFD, and charge pump. Note that there is also a multiplexer at the SCLK output which enables us to switch between using the internally generated 1.64 MHz and the reference clock in the serial interface block. The full layout with labelling identifying each component is given in Figure 6.12.

An overall simulation of the locking process is given in Figure 6.13, with the reference clock in red at the top, the internal divided clock at the bottom in pink, and the control voltage in green in the middle. Over the course of 100μs, we see that the control voltage oscillates with a decaying amplitude to just under 210 mV. Unfortunately, the current matching between the source and the sink isn't perfect, so the voltage increases slowly while both UP and DN are high. This occurred due to time constraints and should be rectifiable in future chips.

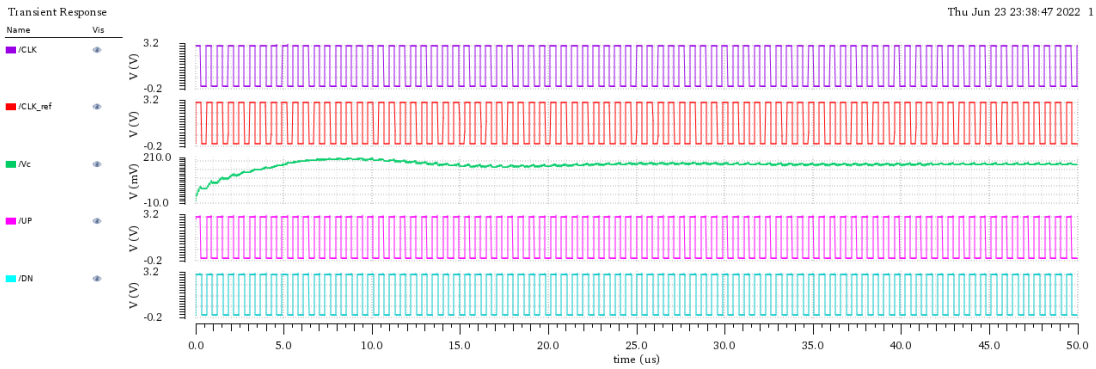


Figure 6.13: Signals during 100 μs simulation of the PLL with 1.64 MHz reference frequency.

Figure 6.14: Format for communication to and from the chip.

6.2 Serial Interface

Serial Peripheral Interface (SPI) is a serial protocol for interfacing between a master device and several slave devices. The protocol requires at least three connections, the clock, master-out-slave-in (MOSI) and master-in-slave-out (MISO). If multiple slave devices are used, a chip-select (CS) line is also needed to identify which slave is being communicated with. Communication is then based on digital words of a fixed length, where the first bit is indicated by a change in the default state of the wire, which signals for the receiver to begin reading the digital word up to the fixed length; after this, additional bits are either discarded or considered a second word.

In our protocol, the default state of the wire is set to VDD, with the transition to GND signalling the start of a digital word which has a length of 16 bits and has the format shown in Figure 6.14. The communication is very simple and only has two types of commands that can be issued to the chip: either writing to registers (programming the chip) or reading data from registers. Thus, as shown in Figure 6.14, the first bit is always 0, the second then distinguishes read (0) commands from write (1) commands. Bits 2-6 are used for the address of the register, which is summarised in Table 6.1; bits 7-14 represent the data that is either written to/read from the chip, and the final two bits are ignored. The bit rate for the communication is 1.64 MHz.

Address	Register	Read/Write
0000	DDS register A	Both
0001	DDS register B	Both
0010	Photodiodes ADC output	Read only
0011	Photodiodes output	Read only

Table 6.1: Addresses of available registers on the chip.

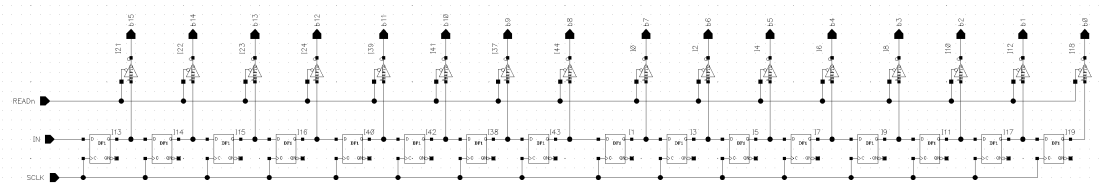


Figure 6.15: Serial-in parallel-out shift register which is used in the receiver.

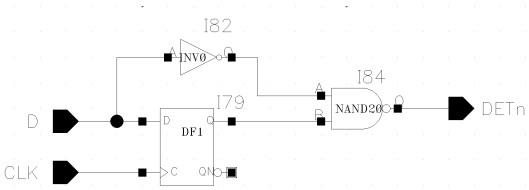


Figure 6.16: Falling-edge detection circuit.

Receiver

The receiver design is built around the 16-bit shift register shown in Figure 6.15, which is built using 16 cascaded DFFs with buffers connected to the outputs. The shift register takes a serial input from the MOSI connection and outputs a parallel word, pins b0-b15, which are latched by DFFs in the buffer-register block shown in Figure 6.25. If the READn pin is set to VDD, the buffers are set to a "null" output and the pins b0-b15 are left at their previous voltages. When READn is set to GND, the values in the DFFs are passed through to the output bits.

To control the READn pin, the start and end bits of the digital word have to be identified such that the READ signal is high whilst the word is passed into the shift register and goes low once the 16th bit is stored to output the full word. To accomplish this, a falling-edge-detect circuit is used in conjunction with a 16-bit counter, as shown in Figure 6.18.

The falling-edge-detect circuit is shown in Figure 6.16 and works as follows: there are two inputs, data (D) and clock (CLK), and one output (DETn). If D starts high, this is inverted at the A input to the NAND gate and when CLK goes high D is passed to the B input which sets DETn to VDD. When D goes low, the A input goes high but the B input is already high causing DETn to go low signalling that a falling edge has occurred. When the clock signal next goes high D is passed through the DFF and DETn is reset to VDD. During the low to high transition of D the input A and B start as VDD and GND respectively, until D goes high when A transitions to GND, which changes nothing at the final output. Thus, the falling-edge-detect output defaults to VDD but goes to GND when a falling edge is detected on the data line, resetting to VDD at the next rising edge on the clock.

As shown in Figure 6.18, the DETn output is then passed to a NAND gate with the READ signal at the other input and the output connected to a TFF. This combination means that whenever DETn and READ start high, but one of them drops low, the CLR node on the output of the TFF is toggled, which only occurs when a start bit is detected or when the counter reaches 16. CLR going high then triggers the

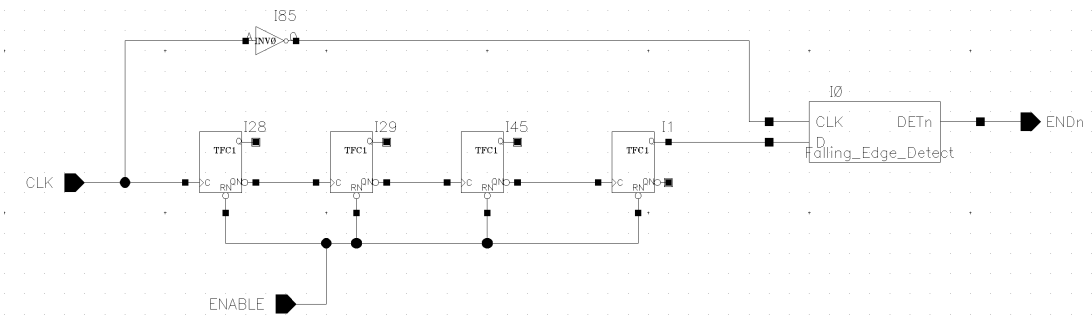


Figure 6.17: 16-bit counter with falling-edge detection to flag the end of one phase of counting.

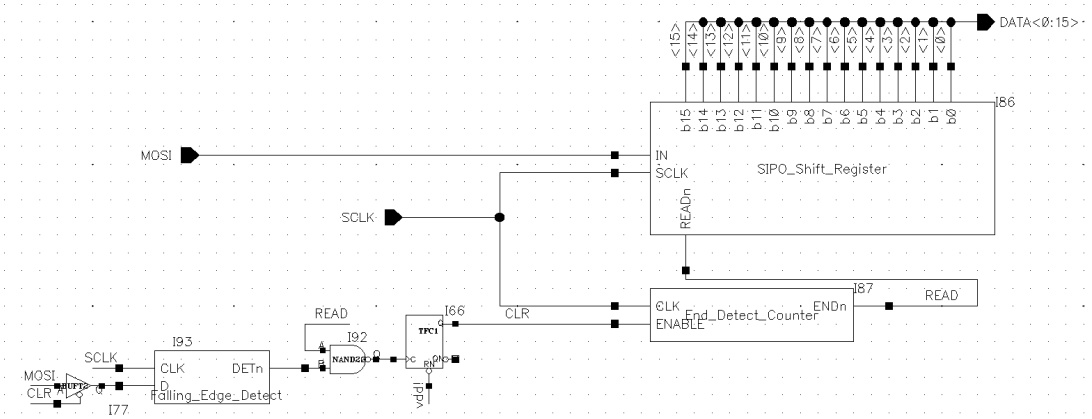


Figure 6.18: Schematic for the receiver.

counter and locks the MOSI signal from triggering the initial falling-edge-detect block through disabling the buffer at the block's data input pin, causing DETn to remain high. Thus, CLR can only be toggled by READ going low.

The counter circuit which produces the READ signal is shown in Figure 6.17 and is a 16-bit TFF counter, with the most significant bit connecting to the data pin of a falling-edge-detect block, which has the clock pin connected to an inverted copy of the least significant bit (clock signal). The counter is enabled/disabled by the ENABLE pin, which is connected to CLR. When CLR goes high, the counter is enabled and counts upwards from 0 to 15, at which point the inverted final bit will fall to GND and the falling-edge-detect will trigger, resulting in READ also going low, which causes the data stored in the shift register to be output and CLR to be toggled low, disabling the counter.

Bus and buffers

To program and store information on the chip, registers like that shown in Figure 6.19 are used in conjunction with an 8-bit data bus and 32 read/write bits. Each register can either be written to, if the WRITE pin is high, or read from, if the READ pin is high. The READ/WRITE pins on each register correspond to a unique bit on the read/write bus such that the 32 bits imply a maximum of 16 registers (more registers can be used if they're restricted to read or write) on the chip, which only has 4 registers

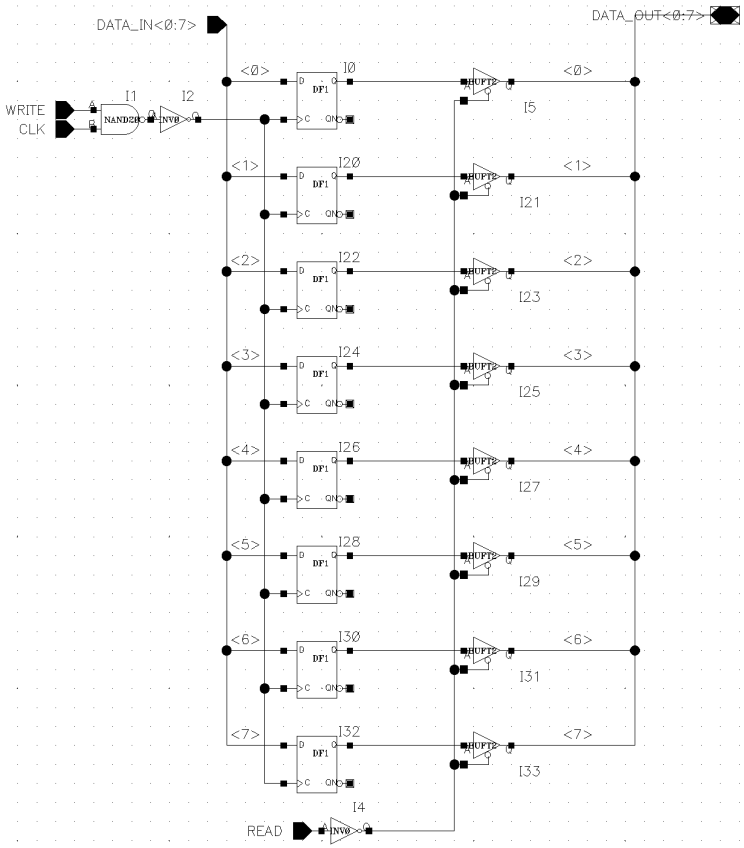


Figure 6.19: Schematic for the data buffer which allows information to be stored and written to the bus.

in this initial prototype. We see in Figure 6.19 that the register consists of 8 DFFs with buffers on the output. The READ pin is then connected to the buffers which have a "null" output when READ is low and thus do not affect the data bus, but allow the information in the DFFs to pass to the data bus when READ is high. Note that only one register should be read from at a time; otherwise, lines on the data bus may be connected to VDD and GND simultaneously.

To write to the register, the WRITE pin and clock pin are connected to an AND gate (a NAND gate followed by an inverter) such that only when WRITE is high, the clock is passed to the DFFs, where it enables data to be stored.

Figure 6.20 shows a variation of the register design where both the input and output are connected to the data bus, but there is a secondary output in between the DFF and the buffer so that the stored data may be read from without affecting the data bus. This design of the register is used when programming part of the chip, such as the DDS, where the information is accessed frequently. The first design is used to store and access digital output from the photodiodes, where the stored information is only occasionally accessed.

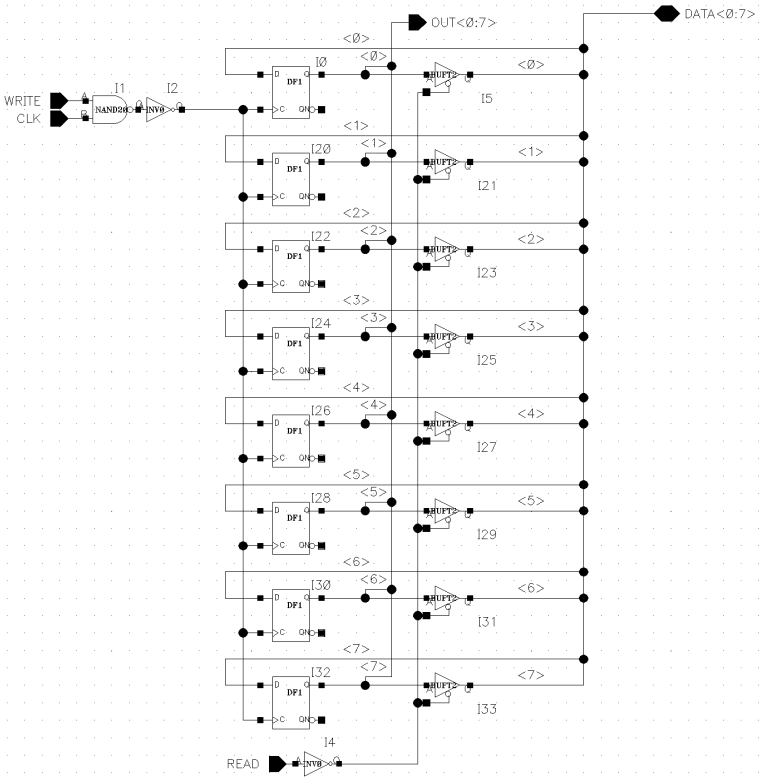


Figure 6.20: Register circuit with can be programmed from and read to the bus. Additional outputs allows information to be read from the register without affecting the bus.

Address decoder

Following the receiver and buffer, the input word is split, as shown in Figure 6.25, with bits 1-5 going to the address decoder where they are used to enable the read/write function of the target register. Given 5 address bits, there are $2^5 = 32$ possible combinations which must be distinguished and assigned to the read/write functions of each register on the chip. The address decoder is shown in Figure 6.21 and consists of only three components, an input stage which ensures the input is at VDD or GND, 2-bit demultiplexers and an output stage which does a similar job to the input stage. The demultiplexers are arranged in columns with 16 before the output stage, leading to 32 possible outputs, 8 in the second to last column, 4 in the column before and so on until there is 1 in the first column with its input connected to VDD. Each input then controls one column of demultiplexers such that each column halves the number of routes the VDD signal can take.

Transmitter

Similarly to the receiver, the transmitter is based on a shift register, shown in Figure 6.23. At the input of each DFF in the shift register is a multiplexer which allows us to load the DFF when the LOAD pin is set high and the clock transitions low. The loading sequence occurs when the clock is low as the low to high transition triggers the bits to be shifted to the next DFF, and the DFF can only be loaded while the

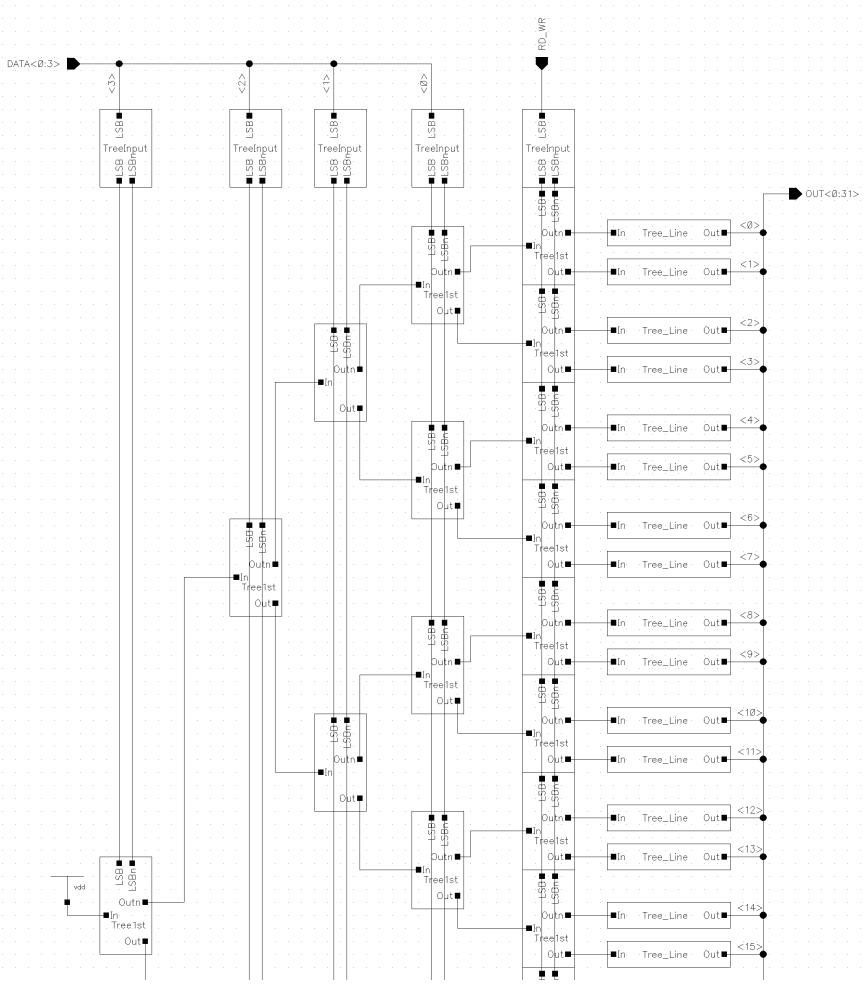


Figure 6.21: Address decoder schematic.

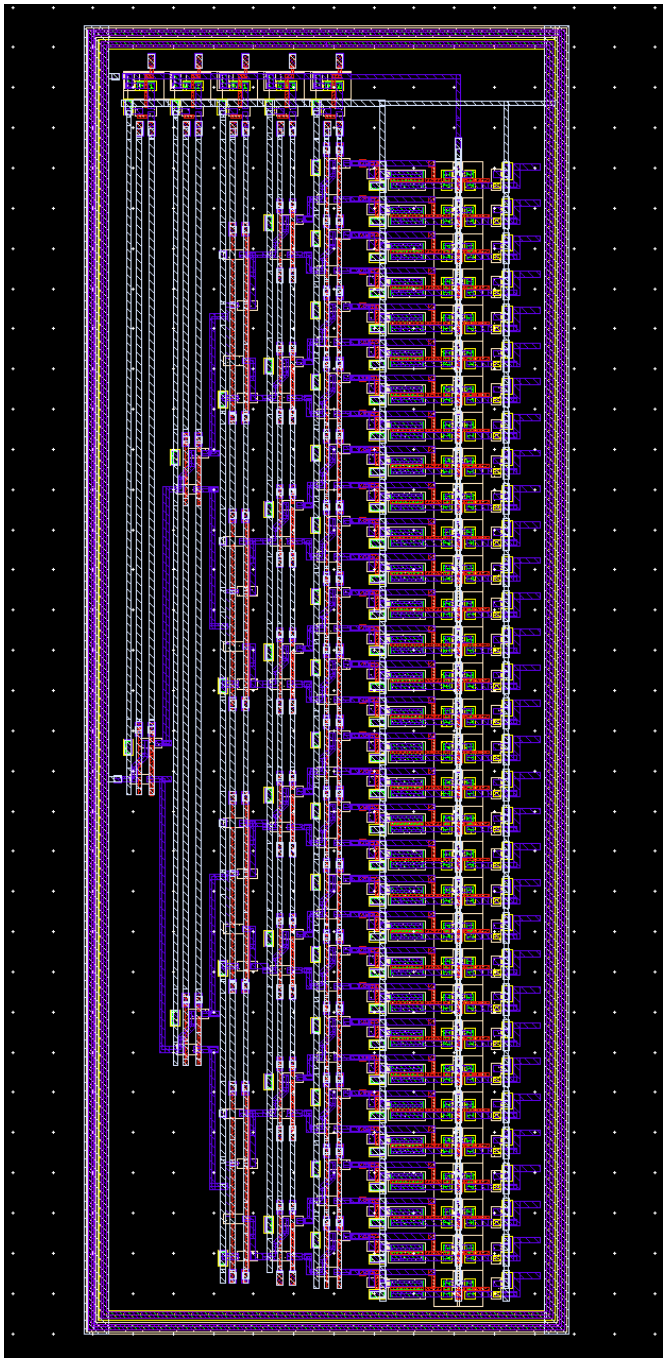


Figure 6.22: Address decoder layout.

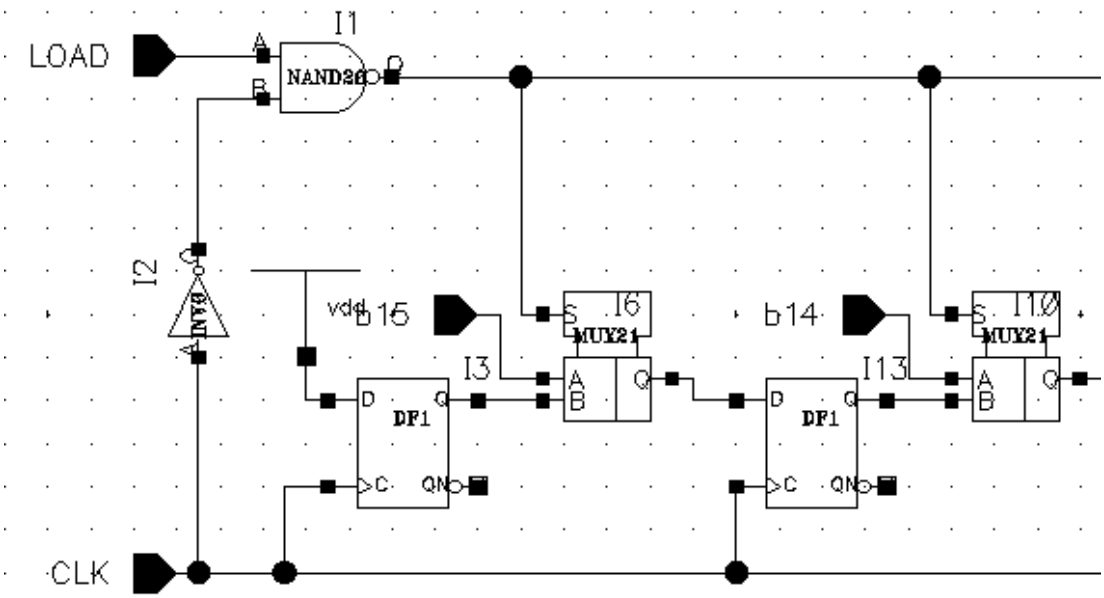


Figure 6.23: Schematic for the parallel-in serial-out shift register.

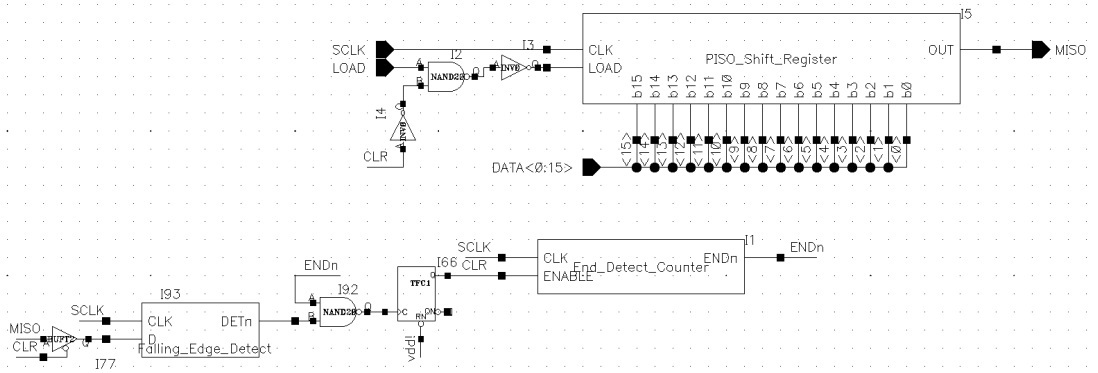


Figure 6.24: Transmitter circuit block diagram.

clock is low. Note that the input of the first DFF is connected to VDD so that the MISO pin is at VDD by default.

The timing for the transmitter is controlled using the same principles as in the receiver, with a 16-bit counter to prevent the LOAD pin on the transmitter from being triggered whilst the digital word is being output to the MISO line. If no word is currently being output, the LOAD pin is initially raised by the read command from an input word which is latched twice, as shown in Figure 6.25, to allow adequate time for the serial interface to read data from the selected register. Once the word has begun to transmit, the start GND bit triggers the counter to start and prevent the LOAD pin from being raised again until the word has finished being sent.

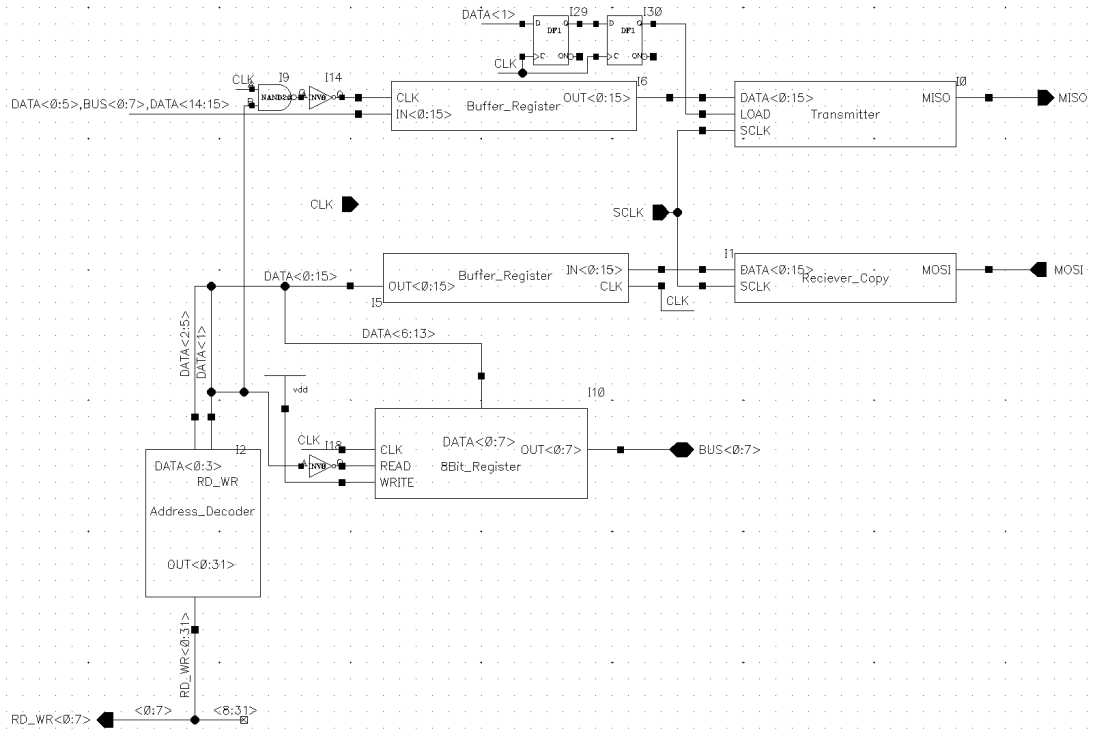


Figure 6.25: schematic for the full serial interface.

Full serial interface

The full serial interface is shown in Figure 6.25, where we see the main blocks which have already been discussed: the receiver, the transmitter, and the address decoder. Other important components are the buffer registers, used to store data output from the receiver or input to the transmitter, and the 8-bit register used as a buffer to the data bus. Note that the transmitter buffer's clock input is triggered by the output of the MOSI b1 AND CLK, meaning data is only loaded to this buffer when a read command has been received. The on-chip outputs from the serial interface are the data bus and read/write lines, the on-chip inputs are the SCLK and CLK and the off-chip connections are the MOSI and MISO lines.

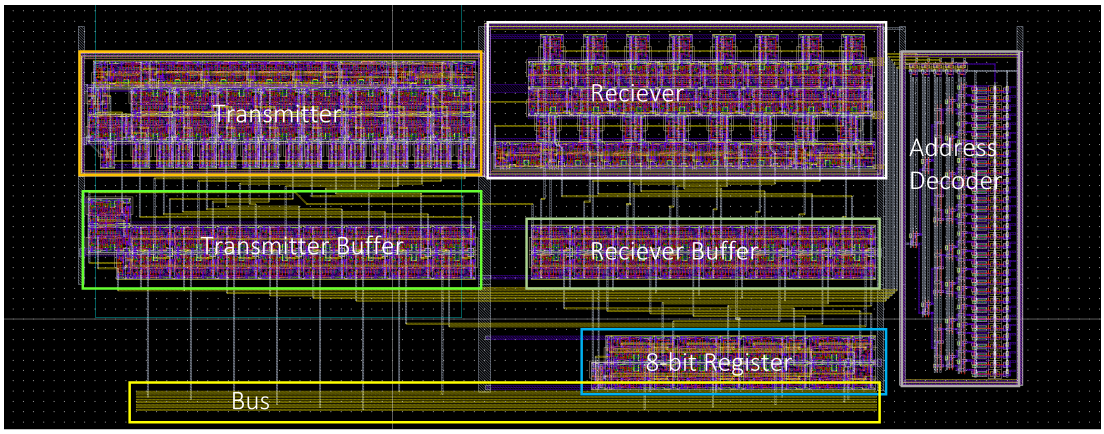


Figure 6.26: Labelled serial interface layout.

6.3 DDS

An essential part of nearly all cold atom experiments involving magnetic trapping is the ability to provide RF fields, which are used in evaporative cooling and RF dressing as described in Section 2.6.2. A direct digital synthesiser (DDS) was built onto the chip as an RF source to be used for evaporative cooling with the end goal of producing a BEC. A block diagram for the DDS is shown in Figure 6.27, where the inputs are a 16-bit digital word used to tune the frequency, known as the frequency tuning word (FTW), and the chip clock. The output frequency of the DDS is then given by

$$f_{out} = \frac{FTW \times f_{clock}}{2^{16}}, \tag{6.1}$$

where FTW is the base-10 value of the digital tuning word, $b_0 b_1 b_2 ... b_{15}$, which can be calculated as

$$FTW = b_{15} + b_{14}2 + ... b_1 2^{14} + b_0 2^{15}. \tag{6.2}$$

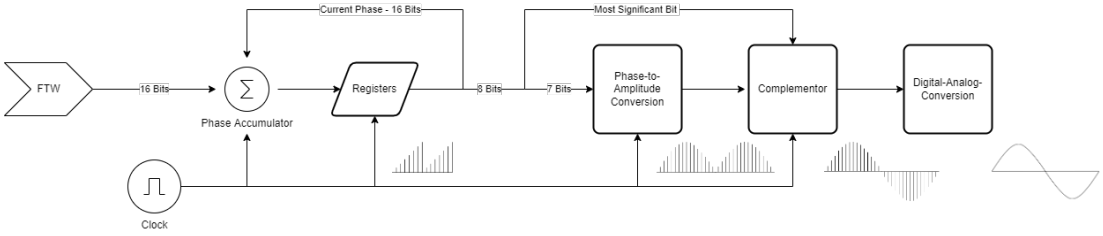


Figure 6.27: Block diagram for the DDS.

The basic principle behind the DDS, which is shown in the block diagram, is to step through a set phase angle (the FTW) every clock cycle such that doubling the phase angle means half as many clock cycles are required to cover one period and the frequency of the output is doubled. This is implemented in the phase accumulator, detailed below, which keeps track of the current phase (16-bits) and adds the FTW to this every clock cycle until it exceeds 360°, where it loops back around to a lower phase. This phase is then converted into a digital amplitude through a phase-to-amplitude table that stores the amplitude corresponding to bits 2-7 of the phase, with b_0 and b_1 reserved for digital operations. However, this table is quite large; so to maintain the resolution of the DDS without requiring too much space, this table only stores the amplitudes for the first quarter of the sine wave. These amplitudes can then be easily modified using digital operations to produce the remaining portions of the wave. Specifically, b_1 is used to invert bits 2-7 such that the phase-to-amplitude table is accessed in reverse to produce the second quarter and b_0 determines if the amplitude is positive or negative, completing the full period. Lastly, this final digital signal is converted to an analogue signal through a digital-to-analogue converter (DAC).

The full schematic for the DDS is shown in Figure 6.28, where we see the FTW is stored in registers

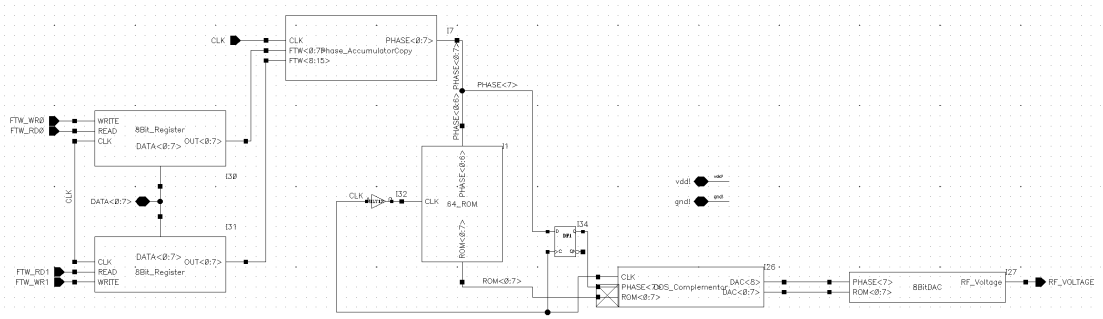


Figure 6.28: DDS schematic.

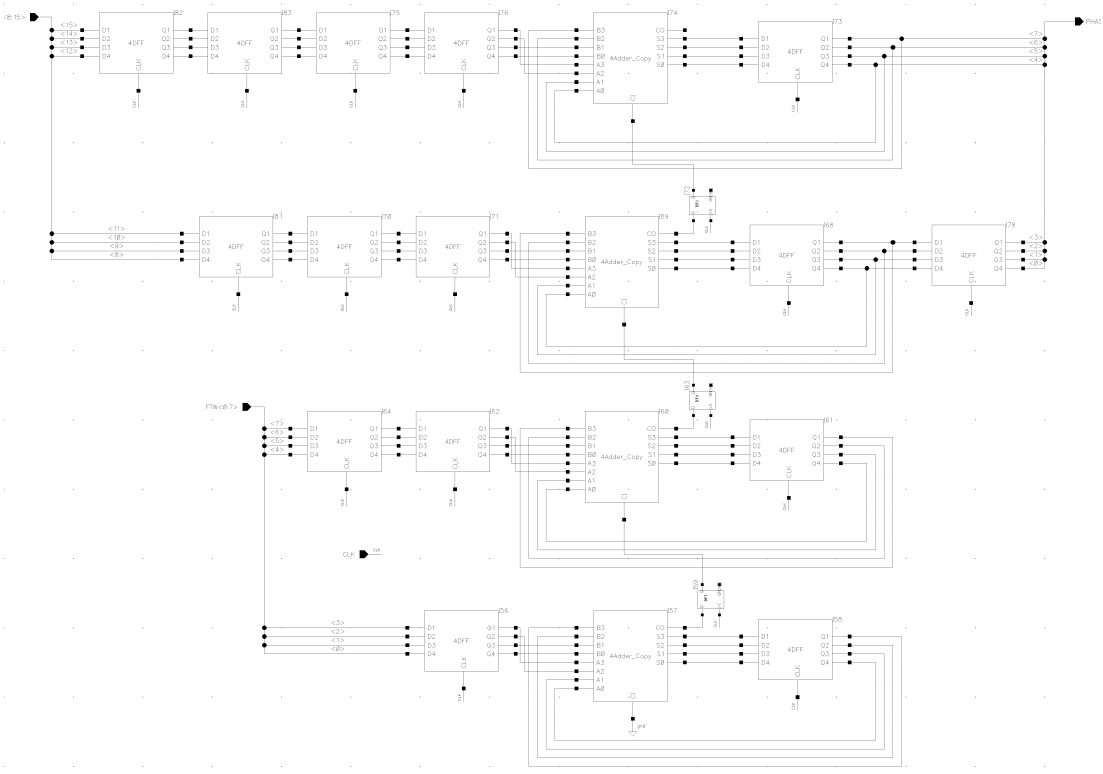


Figure 6.29: Phase-accumulator schematic.

which are programmed through the SPI and passed to the phase-accumulator. The phase then goes to the phase-to-amplitude table (64-ROM) and complementor, followed by the DAC. Note that b_0 is latched so that it remains in phase with the bits that pass through the ROM table.

Phase-accumulator

The phase-accumulator schematic is detailed in Figure 6.29 and is a common design which is detailed in . This phase-accumulator design uses 16 bits to represent the phase of the output signal, with a 16-bit input FTW but only the 8 most significant bits being output to the phase-to-amplitude blocks. The simplest phase accumulator would consist of a 2-bit adder and two d-flip-flops (DFFs); where the input bit would be stored in one DFF and the current phase in the second. The Q output of the input DFF would be connected to one input of the adder and the Q of the phase DFF is connected to the other. The input to

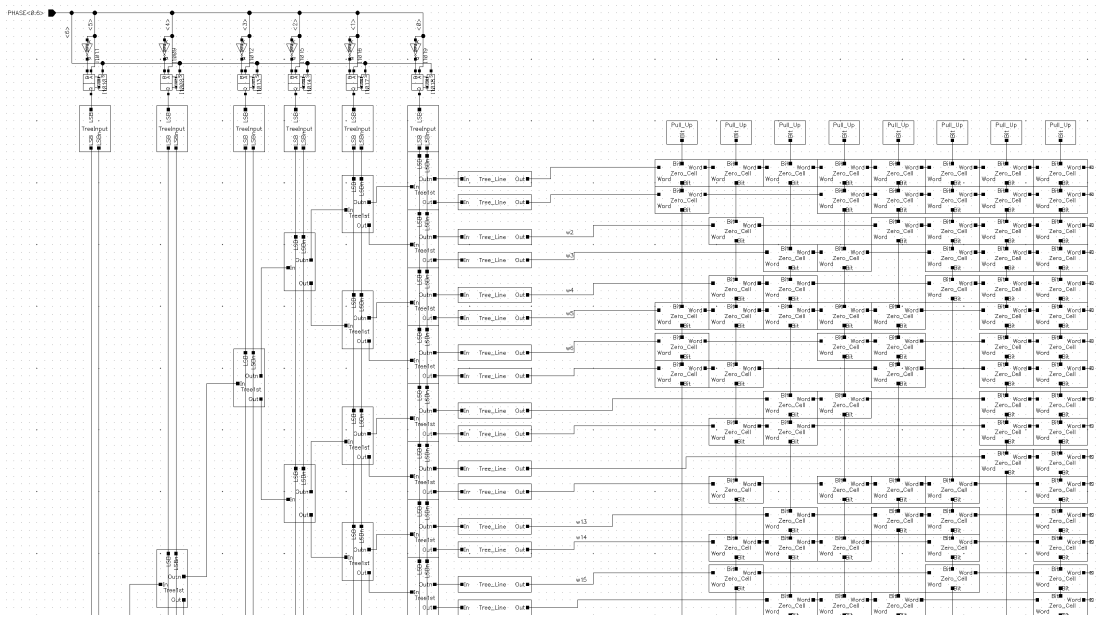


Figure 6.30: The top section of the schematic for the ROM table, which isn't shown in full for readability. Included is the input stage, part of the demultiplexer tree and the first 11 entries in the table.

the phase DFF is then connected to the output of the adder which can be seen in Figure 6.29. Note that the schematic has four DFFs and adders in a single block to help with readability. Thus, when the clock goes high the phase and input bits are passed to the adder, when the clock goes low the clock's output is then stored in the phase DFF. In the full design, we are unable to perform 16 additions in a single clock cycle so this process is split into four stages of four additions which each happen during different clock cycles. However, this is implemented by greatly increasing the space requirements of the phase-accumulator with the additional DFFs required to maintain the correct order of additions. Looking at Figure 6.29, we see that an additional latch is required per stage so that a change in FTW doesn't modify the additions still occurring from the previous word. Similarly, the outputs must be latched so that the output bits are consistent with the phase. Lastly, the carry bits are latched in-between successive adder stages to maintain consistency. The phase-accumulator is built using standard components which are supplied by AMS and the layout is included in Figure 6.34

ROM table

As we see in Figure 6.28, the 8 most significant bits are passed on to the phase-to-amplitude blocks, where bits 1-7 are passed to the ROM table. Figure 6.30 shows the input and general scheme of the read-only-memory (ROM) table used to produce the first quarter of the sine wave. Starting at the right, each column leads to an output bit and each row corresponds to one possible output configuration. At the top, "weak" PMOS transistors are used to charge each column up to VDD, and at each row there is either a wire, corresponding to a 1, or a "strong" NMOS transistor which can discharge the column to GND, corresponding to a 0. Here, weak and strong refer to the magnitude of the transistor's

transconductance. As a couple of examples, if row 0 is activated then the output is 0000-0000 (hyphen added for readability), and if row 1 is activated the output is 0110-0000. Jumping to the left of Figure 6.30 we see the address decoder, which is just a tree of demultiplexers with a VDD input (not shown) and an input bit as a select pin. From left to right, the input bits are b_2 to b_7 so an input of 000000 selects row 0 and an input of 000001 selects row 1. With 6 input bits, there are a total of $2^6 = 64$ rows. Following the address decoder, there are small blocks labelled tree-line which ensure that the row activation is either at GND or VDD such that only one row is activated at a time and no rows are partially activated. At the input of the address decoder, there is a tree-input block, which is just two inverters to ensure input is VDD or GND, and multiplexers with one input being an input bit, the other being the inverted input and the select pin being b_1. If $b_1 = 0$ the table is iterated through starting at 0000-0000 as output and ending at 1111-1111 as output, but if $b_1 = 1$ the table is iterated the other way, starting at 1111-1111 and ending at 0000-0000. This allows us to produce the first half of the sine wave. Note that 8 bits are chosen as a compromise between the precision of the amplitudes stored in the table and the size of the DDS. Lastly, the outputs are latched with DFFs (not shown) to ensure only pure digital signals are output from the table. In Figure 6.28 we see that b_0 must then also be latched to maintain synchronisation. The layout of the ROM table is shown in Figure 6.31.

Complementor

To complete the sine wave with the "negative" half (the sine wave is actually centered around VDD/2), a digital complementor is used to find the complement of the digital amplitude which is the digital equivalent of a negative number. To see how this works, consider a 3-bit digital word: above 100 is "positive" and below "negative". 101 then has the same magnitude as 011 and 111 has the same amplitude as 001. There is a simple algorithm for calculating the "negative" of a "positive" digital number, known as two's complement, which is to invert the "positive" number and add 1. For example, starting with 101, we invert to get 010 and add 1 to get 011 which is the answer we expected. This algorithm is used to take the "positive" amplitudes defined in the ROM table and reflect them below VDD/2. b_0 tells us if the phase is in the first or second half of the wave; if it's in the first, the complementor does nothing, but if $b_0 = 1$ the digital amplitude is inverted and 1 is added to produce the last part of the sine wave.

The schematic for this is shown in Figure 6.32, where we see the inversion stage is implemented in the same way as in the phase-accumulator, and the addition is implemented as a 2-bit adder on each bit, with the initial added bit being b_0.

Digital-to-analogue converter

The DAC used in the DDS is a basic resistor-ladder design with $R = 15k\Omega$ as shown in Figure 6.33. This type of current-scaling DAC was chosen due to being smaller, simpler, and faster than most other types,

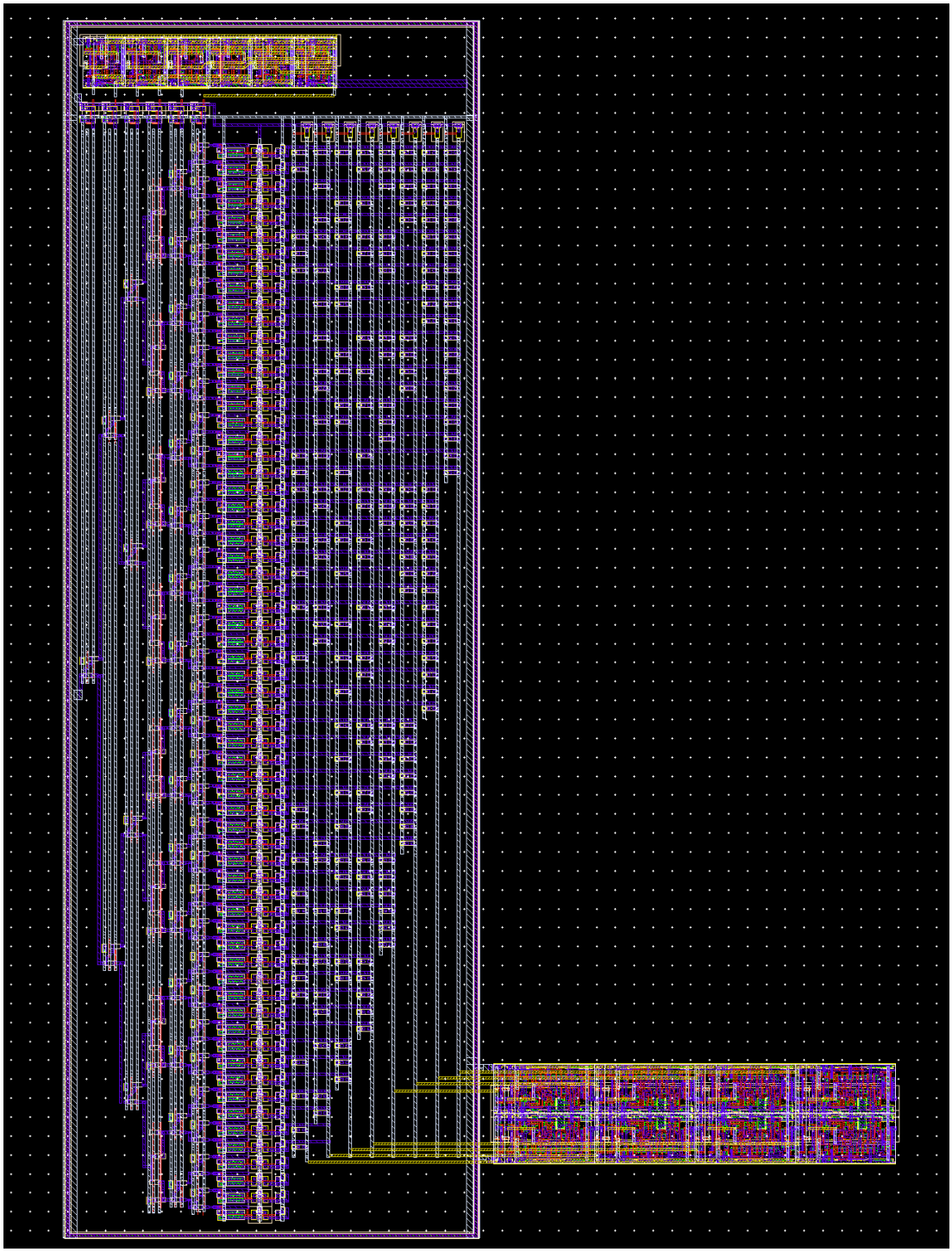


Figure 6.31: The full layout for the ROM table.

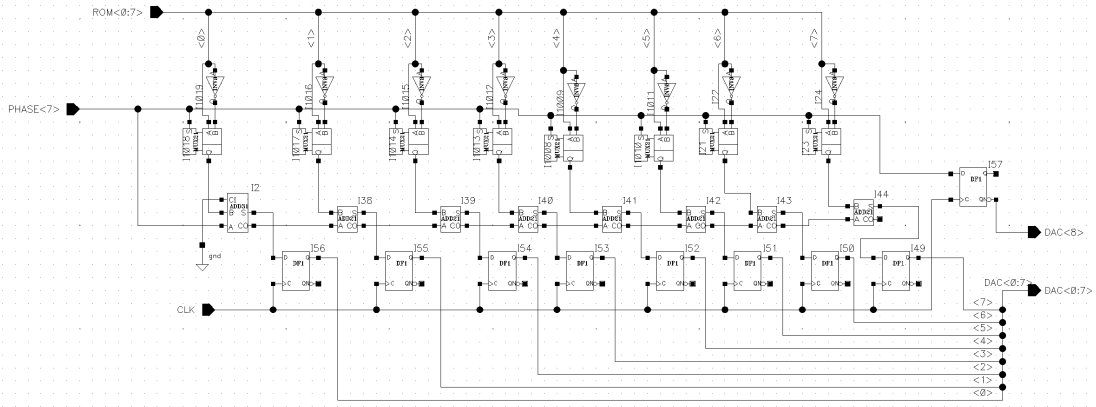


Figure 6.32: Schematic for the digital complementor circuit.

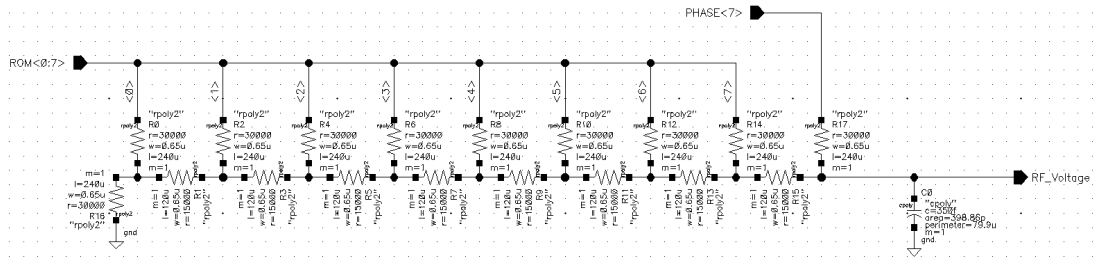


Figure 6.33: Digital-analog-converter schematic.

with the primary drawback being potentially poor matching of the resistors. As seen in the schematic, a capacitor was added at the end of the DAC to act as a low-pass filter and reduce the presence of higher-order harmonics in the final signal.

DDS layout and simulations

The full layout of the DDS is shown in Figure 6.34 with the individual blocks labelled.

The DDS was simulated at both 700 kHz, representing output at the lower end of required frequencies, and 30 MHz, which is around the maximum desired output. As seen in Figure 6.35, the sine wave at low frequencies where the full ROM table is read in full looks reasonable. However, when the ROM is skipped through ignoring many of the entries, we end up with a messy looking signal like that in Figure 6.36.

6.4 Photodetectors

There are two types of photodetectors found on the chip, avalanche photodiodes (APDs) which are able to detect very low-light intensities by generating a current which discharges a capacitor, and single-photon avalanche diodes (SPADs) which are able to detect individual photons. When a reverse bias is applied to a diode, incident photons are able to produce a current as shown in Figure 6.37 where we see that initially, a very small constant current is produced which arises from incident photons only

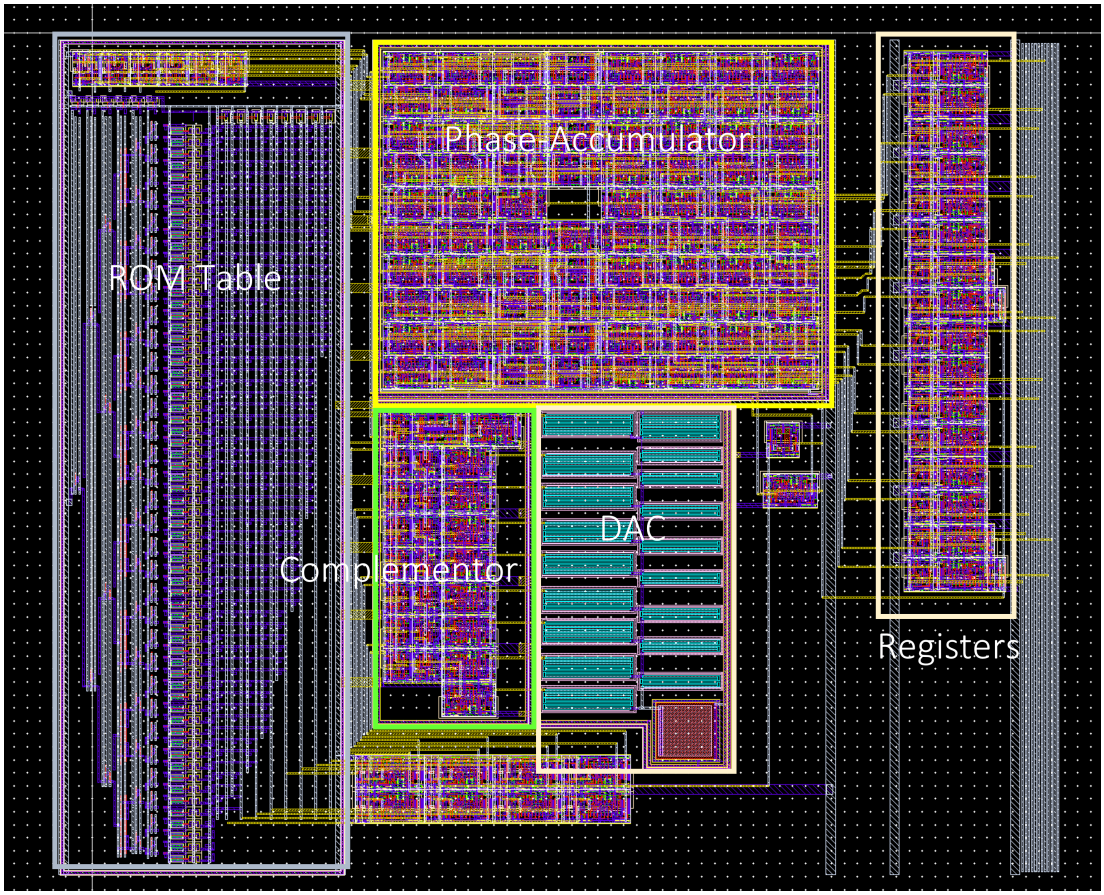


Figure 6.34: The labelled layout for the DDS.

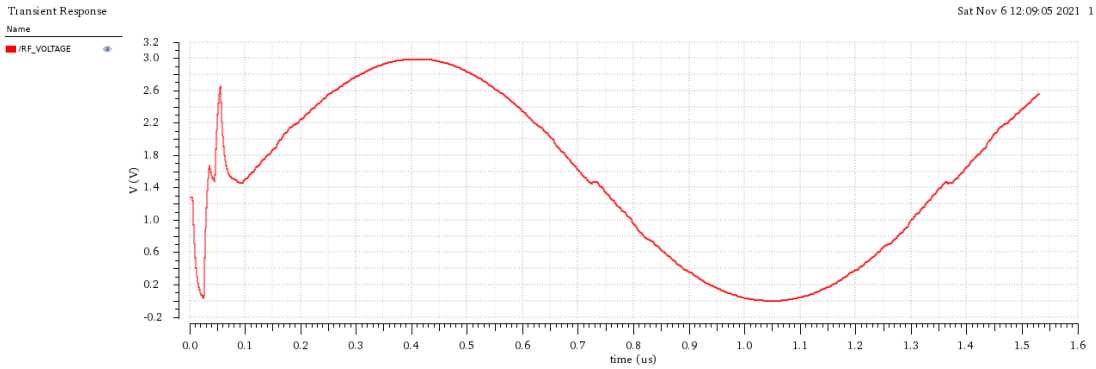


Figure 6.35: Simulated DDS output set to 700 kHz.

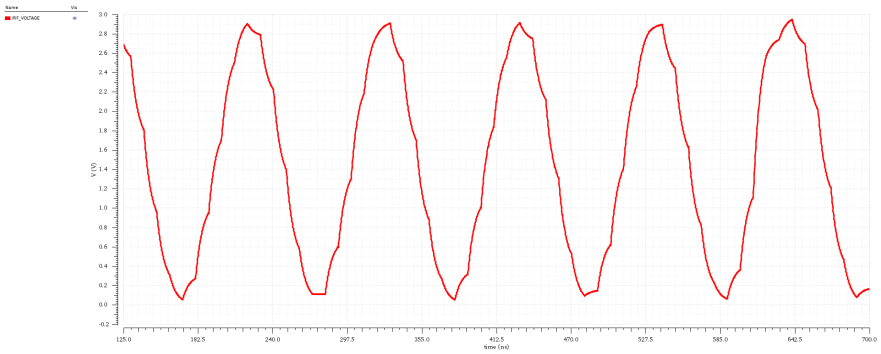


Figure 6.36: Simulated DDS output set to 10 MHz.

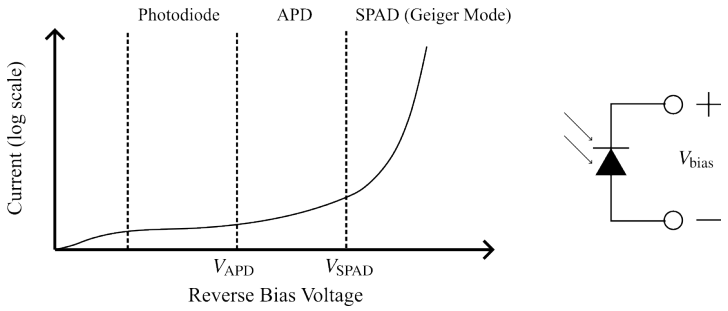


Figure 6.37: Reverse-bias operating regions for a photodiode.

being able to produce a single electron-hole (e-h) pair. Once the bias voltage increases above V_{APD} the avalanche effect begins to take effect with electrons able to produce further e-h pairs due to impact ionization from the electrons being accelerated by the electric field across the junction. However, only once the bias voltage increases to above the breakdown voltage V_{SPAD} are holes sufficiently accelerated to generate e-h pairs due to their higher effective mass[31]. Above the breakdown voltage, the photodiode is in the Geiger regime, and individual photons generate a runaway current which can be detected then quenched by disconnecting the bias voltage. The triggering of the quenching adds to a counter circuit to keep track of the photon number.

In this section we begin with the photodiode design then explain the readout circuits required for each type of photodetector. The photodetectors on chip1 were developed by Lukas Rennhofer as his Masters dissertation at FHWN where further details and simulations can be found[67].

Photodiode

The photodiode structure used in chip1 is adapted from that in[82] and has a cross-section which is shown in Figure 6.38. The cathode is connected to an n+ region with a diameter of 59.2 μm and a deep n-well, whilst the anode is connected to the p-substrate. A deep p-well beneath the n+ region forms the pn-junction. In this design, there is a high electric field between the n+ and deep p-well which allows the avalanche effect to occur; this is known as the multiplication region. If the cathode voltage is sufficiently high, the entire deep n-well becomes depleted and serves as a much larger region for the absorption of photons. This separation of multiplication and absorption region helps to improve the efficiency of the SPAD.

From simulations[67] it was found that the breakdown voltage of the diode is around 18 V and a maximum electric field of 5×10^7 V/m is present at the n+/deep p-well junction when a bias voltage of 17.5 V is used.

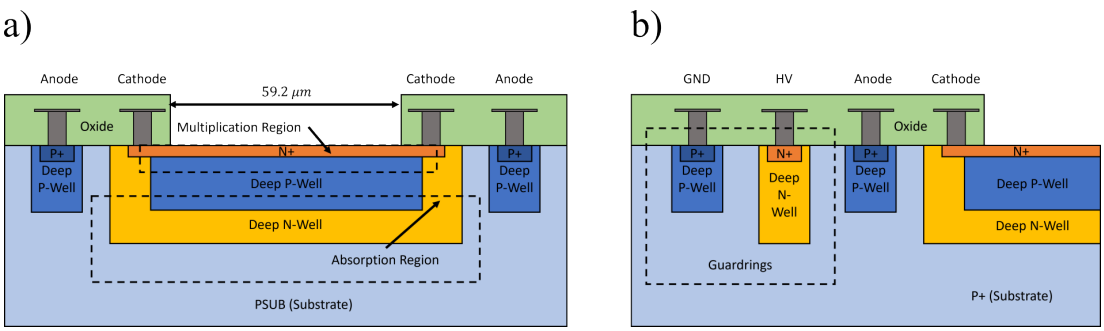


Figure 6.38: a) Cross-section of the photodiode geometry used in chip1. Photons are able to generate e-h hole pairs in the large depletion region from the n+ doped layer to the substrate. The avalanche effect occurs at the boundary between the n+ and deep p-well layers. b) Guardrings are added, surrounding each photodiode, to help isolate them and reduce substrate noise.

APD

By using a lower reverse bias voltage, photodiodes are able to detect low-light intensities through the avalanche effect without the gain being so high that the detector saturates after a single photon. The photodiode acts as a current source which is able to discharge a capacitor in a circuit known as an active pixel sensor, and the voltage across the capacitor is then measured as this is proportional to the light intensity. The active pixel sensor is based on that from chip0[55] and the schematic is shown in Figure 6.39. There are 5 inputs and 1 output, the output being separated from the capacitors by a source follower with a current mirror acting as a current source where the current is determined by a resistor. This is essential so that a low impedance at the output doesn't contribute to the discharge of the capacitor. The inputs are RES, EN, IN, HOLD, and select; IN is connected to the diode and provides a path for the current which discharges the capacitor. select allows us to choose which of the two capacitors is used in the active pixel sensors, thus controlling the sensitivity of the photodetector since a higher light intensity would very quickly discharge a small capacitor, but a lower light intensity might make an undetectable change to a larger capacitor. HOLD is able to disconnect the input from the capacitors, maintaining the current voltage across the capacitors. EN and RES both need to be high to connect the capacitors to VDD, allowing them to be initially charged; RES is then set low during the measurement phase when the capacitors are discharging. EN is used to enable or disable the entire sensor.

A block diagram from[67] shows the complete APD scheme for a single photodiode. Following the active pixel sensor are an output pin OUTA_A and an analogue multiplexer which connects all of the individual photodiodes and selects which is connected to the ADC at any time. The inputs SA0, SA1, and SA2 are responsible for selecting the output of the multiplexer, and the ADC requires a reference voltage Vref_ADC. The digital output from the ADC is stored in 8-bit memory, where the registers are addressed individually by a memory select circuit with 4 inputs: SA0, SA1, SA2, and WR_A. The three numbered inputs are responsible for selecting which memory is accessed, and the WR_A signal determines if the memory is written to or read from. Finally, a multiplexer is used to serialise the output

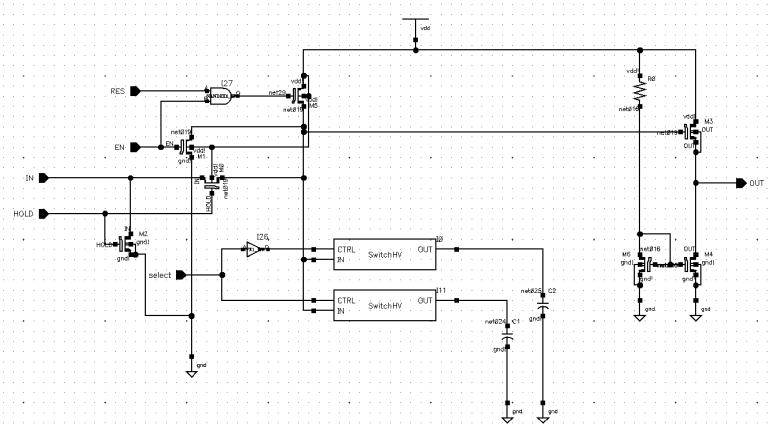


Figure 6.39: Active pixel sensor

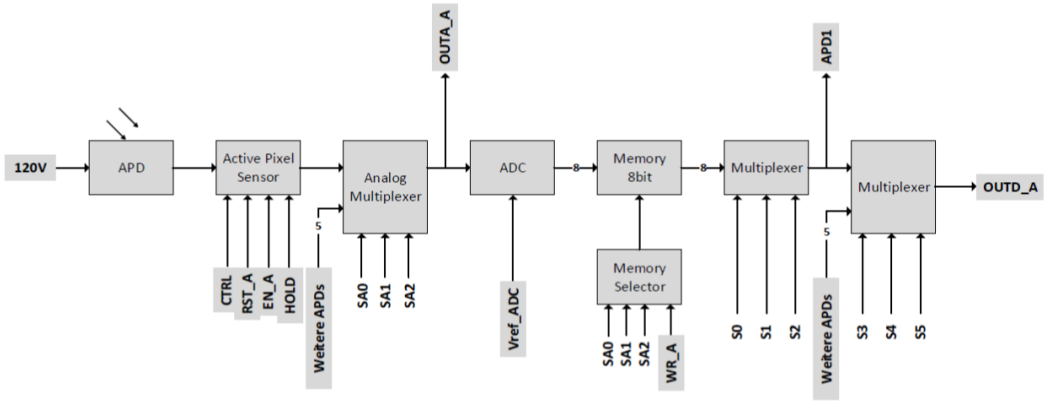


Figure 6.40: APD block diagram courtesy of Lukas Rennhofer[67].

from the 8-bit memories, which is connected to an output pin APD1 and a second multiplexer which selects from the other APDs and connects to an output pin OUTD_A, which is used to access the digital outputs from all of the APDs.

SPAD

SPADs are capable of detecting individual photons by operating the photodiode above the breakdown voltage so that a single photon produces a measurable current. Once the photodiode current has been detected, the diode must be "reset" by connecting the anode to VDD, which allows the anode voltage to quickly saturate and the current to decrease to zero. In order to do this, a quenching circuit based on that from[49] is used to detect the current, quickly quench it, then reset the anode voltage to 0 V. Photocurrent enters the circuit through IN and follows the path to GND through R1, raising the voltage at the inverter. Once the voltage is raised above the threshold of the inverter, the quenching circuit is activated and the IN node is connected to VDD, allowing the voltage to rapidly increase and the current to decrease. At the same time, the quenching voltage is passed through a second inverter to produce an output voltage, which is simultaneously used as a digital signal to announce the detection of a photon and is passed through a delay circuit where it then enables an NMOS connecting IN to GND and resetting the anode

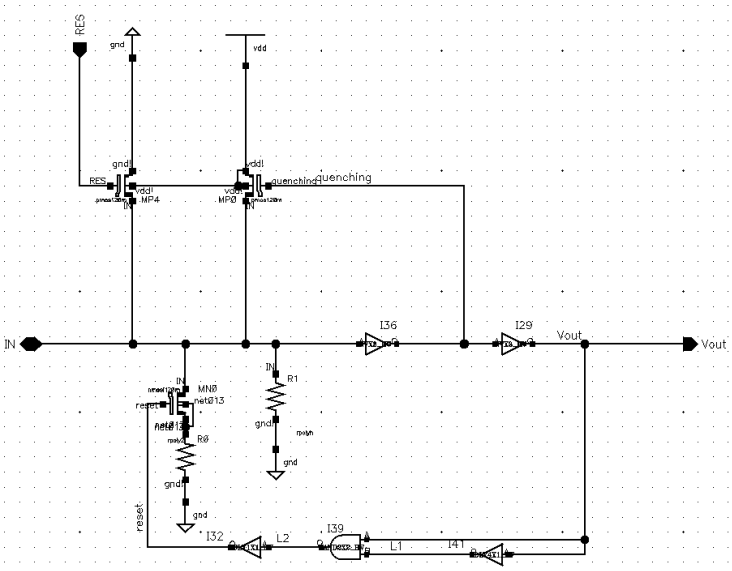


Figure 6.41: Quenching circuit

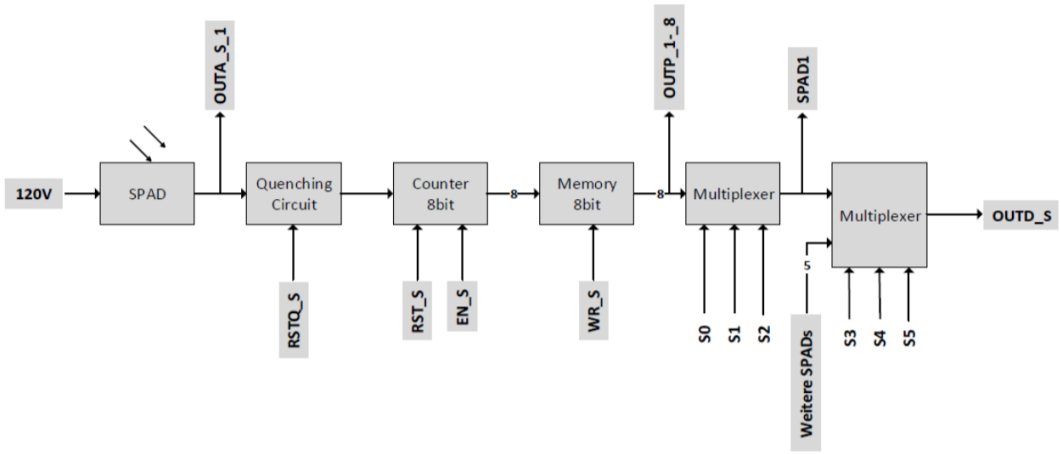


Figure 6.42: Simplified SPAD block diagram courtesy of Lukas Rennhofer[67].

voltage to 0 V so that another photon can be detected. Note that a reset pin labelled RES can be used to manually reset the device.

A full block diagram for a SPAD and readout circuit is shown in Figure 6.42 where the SPAD is connected to the HV bias voltage, an analogue output pad called OUTA_S_1 for a direct voltage measurement and the quenching circuit. The quenching circuit has a manual reset input called RSTQ_S and the output is connected to an 8-bit counter which has two further inputs. A reset pin called RST_S which sets the counter to 0 and an enable pin EN_S, both of which are controlled externally. The output from the counter is stored in an 8-bit memory block as in the APD readout, which is serialised by a multiplexer and connected to a pad labelled SPAD1. Finally, a second multiplexer connects all of the SPADs where the output is accessed at pad OUTD_S.

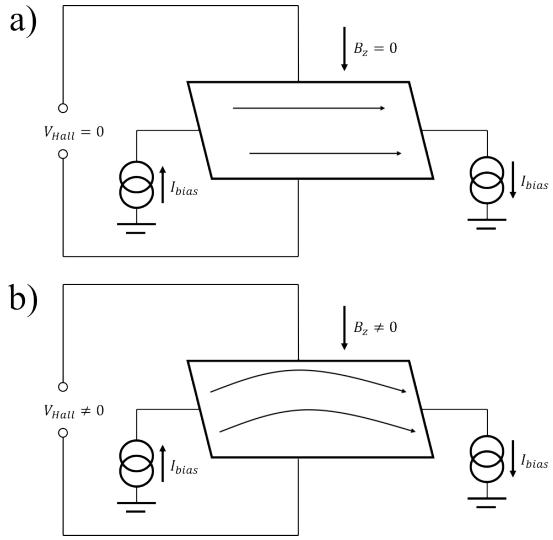


Figure 6.43: a) With no magnetic field, the current passes throught the conductor and no voltage is measured. b) When a magnetic field is present, charge carriers are pushed to the edge of the conductor, inducing a voltage across the sides on the conductor.

6.5 Hall Probes

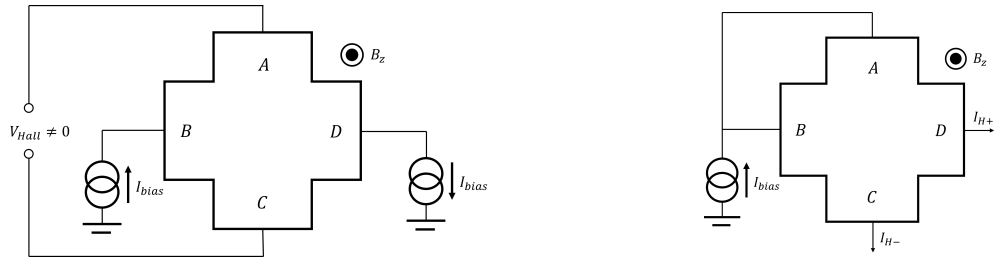
The Hall effect, discovered in 1879 by Dr. Edwin Hall, is a commonly used method for measuring the strength of magnetic fields. Figure 6.43 shows the effect in action, where current is passed through a conductor, from left to right, and the voltage drop across the top and bottom of the sheet is measured. In a, no magnetic field is present and charge carriers are able to move through the sheet directly resulting in no voltage across the top and bottom. However, in b, a magnetic field perpendicular to the sheet causes the Lorentz force to affect the charge carriers which results in positively and negatively charged particles being pushed to opposite ends of the sheet, meaning a voltage drop is now measured. This Hall voltage, V_{Hall}, is given in terms of the bias current, I_{bias}, by

$$V_{\text{Hall}} = \frac{Gr_H}{ent} I_{\text{bias}} B_z = S_I I_{\text{bias}} B_z, \tag{6.3}$$

where G is a geometrical correction factor, r_H is the Hall scattering factor, n is the carrier concentration, e is the electron charge, t is the thickness of the sheet, B_z is the magnetic field perpendicular to the sheet and S_I is known as the sensitivity. Much of this section follows the Ph.D. thesis by Heidari [34].

Current-mode Hall cross

Hall probes can be operated in both voltage and current modes, which is depicted in Figure 6.44. In the standard voltage mode, the Hall voltage is measured directly from the arms perpendicular to those used to run the bias current through the device, whereas in the current mode sensor the bias current is injected into two perpendicular arms, A and B, and the output currents in D and C are then measured. In the absence of a magnetic field, the currents measured in D and C should be equal; however, when a



(a) Voltage-mode Hall probe.

(b) Current-mode Hall probe.

Figure 6.44: A Hall cross operating in both voltage and current modes.

magnetic field is present, this causes an imbalance between the arms. The measured currents are

$$I_{H+} = \frac{I_{\text{bias}}}{2} + \frac{I_{\text{Hall}}}{2} \quad \text{and} \tag{6.4}$$

$$I_{H-} = \frac{I_{\text{bias}}}{2} - \frac{I_{\text{Hall}}}{2} \tag{6.5}$$

with the Hall current calculated in [?] to be

$$I_{\text{Hall}} = \mu_H \frac{w}{l} B_z I_{\text{bias}} \tag{6.6}$$

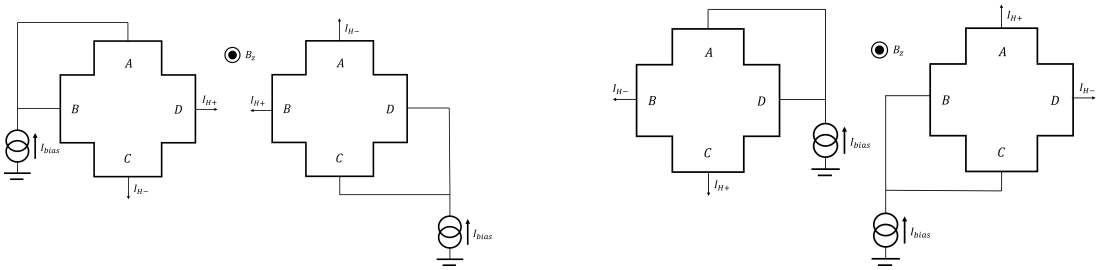
where μ_H is the Hall mobility of the majority carriers, w is the width of an arm and l is the length from injection to output.

There are several advantages to utilising a current-mode Hall probe over one operating in the voltage-mode. Namely, current-mode devices have improved sensitivity, both inherently and through the use of an integrator to scale the measurable current with time, and the option to apply noise reduction techniques, such as current spinning and orthogonal coupling, which will both be discussed in the following section.

Noise reduction techniques

Within the Hall probe, there are four primary sources of error: the offset, 1/f noise, thermal noise, and generation-recombination noise. Whilst there is little that can be done to improve the thermal noise and generation-recombination noise, the offset and 1/f noise can be mitigated through a combination of orthogonally coupled Hall crosses with current spinning.

Both of these techniques are demonstrated in Figure 6.45 where we see two Hall crosses with the bias currents entering orthogonally to each other and the output currents summed. In addition to this, the directions of the inputs and outputs are rotated based on a clock signal so that the effects of any structural asymmetry, such as inconsistent doping or geometric errors, in the crosses (which is the main cause of offset error) are reduced. Moreover, provided the switching occurs above the 1/f corner frequency, 1/f noise is also compensated for as this can be thought of as a slowly varying offset and at high enough



(a) Step one of current spinning.

(b) Step two of current spinning.

Figure 6.45: The first two stages of the current spinning technique with orthogonally coupled Hall probes.

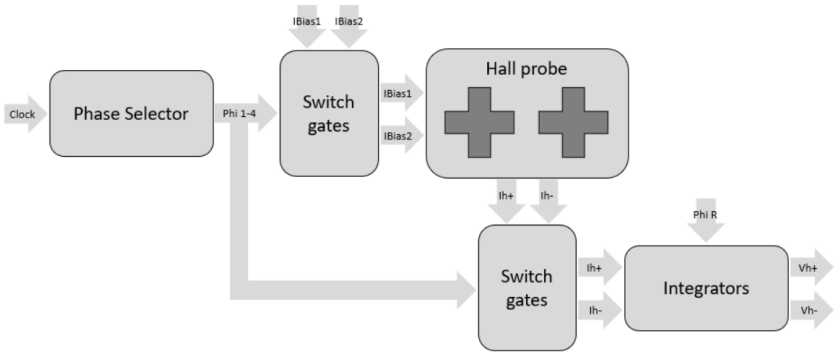


Figure 6.46: A block diagram describing the Hall sensors[69].

switching speeds the static and varying offsets are distinguishable.

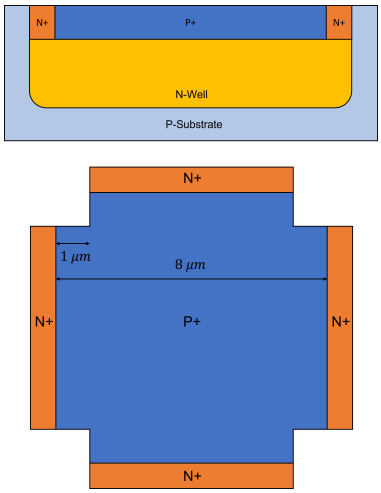
Implementation of Hall probes

A block diagram outlining the full Hall sensor circuit is shown in Figure 6.46 where we see that the bias currents are injected from external sources and are guided by switches which are in turn controlled by the phase generated by a phase selection circuit. The output from the Hall crosses is then integrated before being output directly or passed to an ADC.

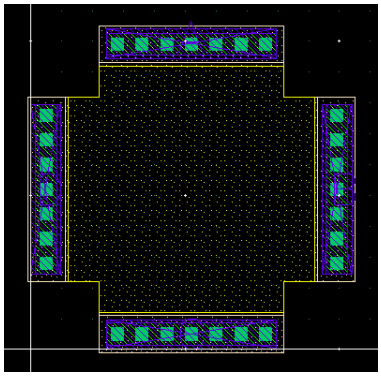
The Hall probe structure is given in Figure 6.47, where Figure (a) gives both the cross-section and top view of the Hall cross and (b) shows the layout in Cadence Virtuoso. Here, as shown in the cross-section, the Hall plate is built using the n-well layer with the metal contacts attached to the n+ borders. The p+ layer is added as it helps to reduce noise in the device. The dimensions of the Hall cross were shown by Heidari to present a favourable sensitivity[?] .

Next, we see how the Hall switches can be easily implemented using two transistors. Note that in a normal transmission gate design, a pmos and an nmos transistor are used to reduce the resistance across the device; however, two nmos transistors are used in this design as each path is controlled by a separate phase. In an alternative switch design used in the Hall probe circuit, there are two inputs which are only connected to a single nmos.

Figure 6.49 shows how the orthogonally coupled Hall crosses are connected with the required



(a) The cross-section is shown above and a top-down view is shown below.



(b) The layout of the Hall cross taken directly from Cadence.

Figure 6.47: CMOS implementation of the Hall cross.

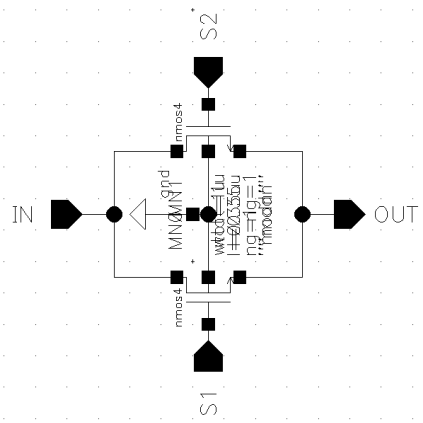


Figure 6.48: Switch schematic.

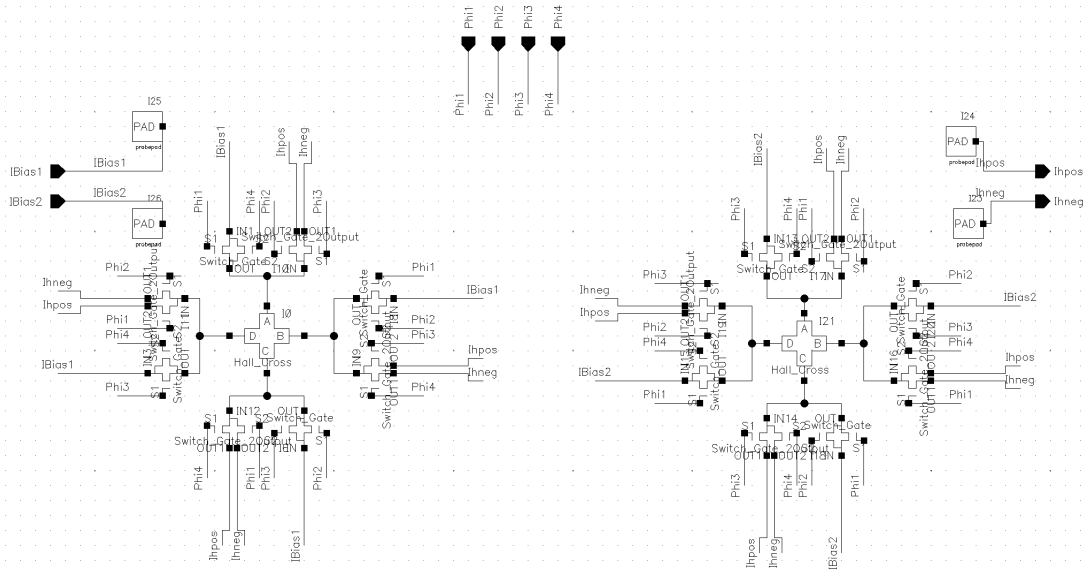


Figure 6.49: Orthogonally coupled Hall crosses which allow for current spinning based on the four input phases.

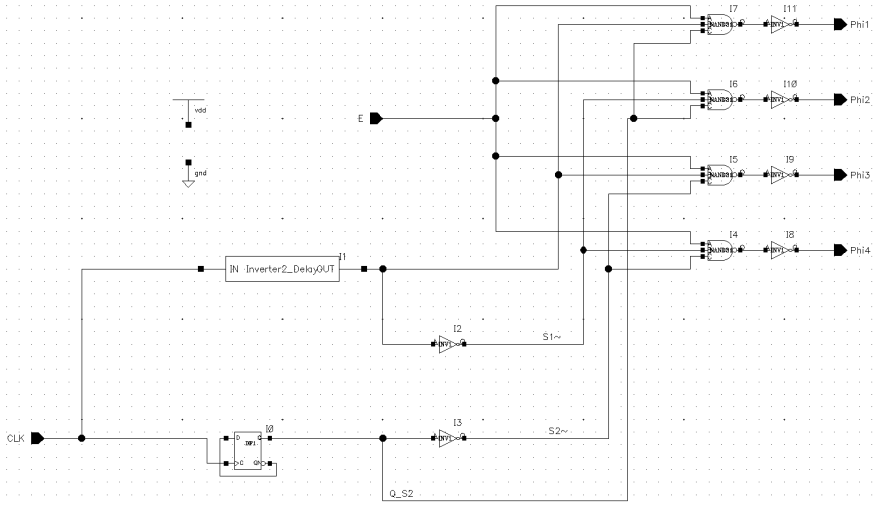


Figure 6.50: Phase selection circuit.

switches.

The phase selection circuit is presented in Figure 6.50 and is a modified demultiplexer. An input clock signal has its frequency halved by a DFF frequency divider, then this signal and the original clock signal are combined using AND gates to produce the phases shown in Figure 6.51. An enable pin allows us to toggle whether or not the phases are being generated.

Lastly, the integrator circuit is given in Figure 6.52. The op-amp is part of the default analogue components library supplied by the foundry. In the original design, a fully-differential op-amp is used here, but due to time constraints, this was not completed in time and the default op-amp had to be used. A capacitor is then added to the op-amp feedback to produce an integrator, and a reset pin is added which enables us to discharge the capacitor. There are two integrators used in the design, one for each

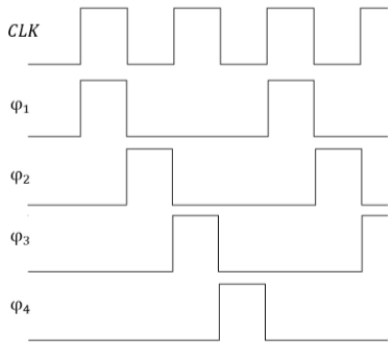


Figure 6.51: Timing for the current spinning used in the Hall sensor[69].

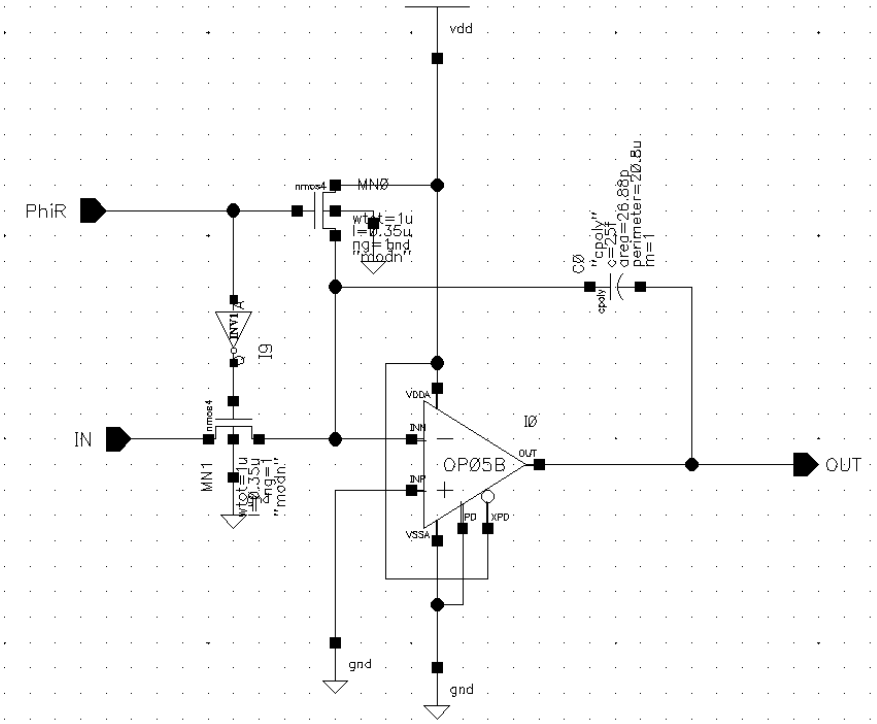


Figure 6.52: Integrator schematic.

of the two output currents.

6.6 Full Chip Layout

The layout for the full chip is given in Figure 6.53 with all of the components labelled. We see the Z wires used for trapping the atoms in the lower half of the chip, with switches that allow us to choose between a longer or a shorter Z wire. The Z wires are 100 μm wide, with the distance between the arms (measured between the centres of the arms) being 1.136 mm, for the shorter Z, and 1.69 mm for the longer Z. Parallel to the main length of the Z wires are the rf-wires, which are used to introduce rf-dressing to the chip. There are switches connected to these rf-wires that allow us to alternate between using an externally generated rf-signal and the internally generated one from the DDS.

The PLL, SPI, and DDS are placed far away from the Z wires to help alleviate the effects that low frequency noise might have on the atoms. These frequencies have the potential to cause the atoms to transition to strong-field seeking states, which will cause losses and a reduced atom number.

The photodiodes are placed next to the atoms in an interleaving pattern to maximise the amount of light present at both kinds of sensor.

The Hall probes are beneath the Z wires, so that the static field can be measured.

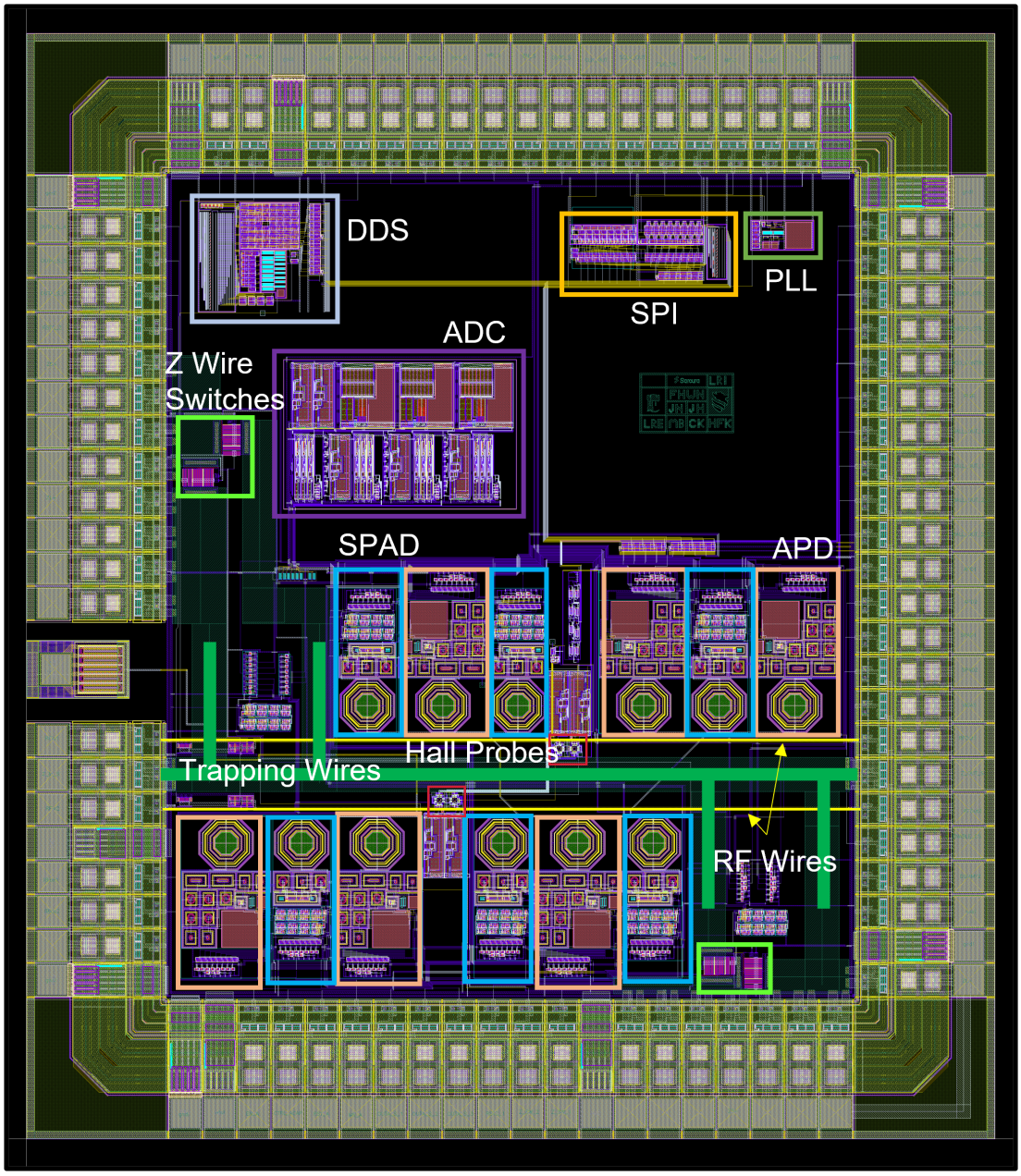


Figure 6.53: Layout of the full chip including labels for the different components. The overall chip is 2.828 mm × 3.228 mm, or 2.012 mm × 2.411 mm when the pads are excluded.

Chapter 7

Summary and Outlook

The work presented in this thesis can be divided into two parts: construction of the experimental system and the design of new CMOS chips. Progress in both of these areas is summarised below.

7.1 Experimental System

A cold atoms experiment was constructed to provide the infrastructure necessary to test the performance of CMOS chips within cold atoms experiments. With this experiment, we were able to initially trap 3.6×10^8 atoms during the MOT stage and cool the cloud to 66μK with polarisation gradient cooling. Approximately 4.5×10^7 atoms were then magnetically trapped using the MOT coils at an efficiency of 12%. Following this, 3.6×10^6 were then successfully transported and loaded into an Ioffe trap below the CMOS chip, generated by the copper understructure. However, there are several major issues with the experiment:

1. Shorts between several pins and vias on the PCB developed during the bakeout stage, including a short between VDD and GND on the CMOS chip.

2. Outgassing of the copper understructure limits lifetimes in the Ioffe trap to 3.7 s preventing adequate evaporative cooling ramps.

3. The magnetic capture is very inefficient.

Possible causes for these issues were discussed in Chapter 6.

The next step for the experiment is to open the vacuum chamber and verify the causes of the shorts and outgassing. In the best-case scenario, this will simply involve cleaning Kapton foil, copper understructure, and PCB, which could only take a few weeks. Whilst the experiment is dismantled, several other issues can also be addressed, such as adding an H-bridge or bipolar power supply to the x-bias coils to help address the inefficient magnetic capture. However, the vast majority of the groundwork

has been done for the experiment, and once the shorts and outgassing are resolved, very little needs to be done before measurements of the performance of chip0 are possible.

7.2 Chip1 and Further Iterations

Over the project, a second-generation CMOS atom chip was designed in collaboration with FHWN and fabricated in 350 nm high-voltage AMS technology. This chip was designed to act as a double-well interferometer with a DDS RF source, Hall probes for magnetic trap measurements, APDs and SPADs for the low-level light intensity measurements, and a serial interface for programming and readout. At the time of writing, the chip is in the process of being tested at FHWN, where the performance of the individual components will be measured before the chip is tested with atoms in the vacuum chamber. Whilst the highest priority has been to collect data on existing chips, work has begun on future iterations, which will be discussed here.

Chip2

The double-well interferometer used in chip1 allows for interferometry with external states, specifically spatial modes. However, control of internal states is necessary for both clocks[84] and logic gates[1]. Typically, the states used on an atom chip are the hyperfine "clock" states $|1\rangle = |F = 2, m_F = 1\rangle$ and $|0\rangle = |F = 1, m_F = -1\rangle$ which exhibit the same magnetic moment at a "magic" magnetic field of 3.23 G and consequently the same potential. The transition frequency between these states is 6.83 GHz; however, to conserve angular momentum, this is a two-photon transition, so a microwave source of 6.35 GHz is required in conjunction with the RF source already present on the CMOS chip to perform state transfer. When the RF and MW fields are applied, Rabi oscillations begin to occur between the clock states, and the length of the applied pulse determines the resulting state populations by ending the oscillations at different phases. This allows the implementation of a Ramsey sequence where a $\pi/2$-pulse produces a coherent superposition of $(|0\rangle + |1\rangle)/2$ from an atom in $|0\rangle$ which is then evolved for some time T resulting in the state $(|0\rangle + e^{i\theta}|1\rangle)/2$ which is subjected to another $\pi/2$-pulse, leaving all atoms in $|0\rangle$ if there is no phase and population imbalance; if some phase is acquired, this allows a measurement of the phase to be made. Further schemes can be implemented with additional π-pulses. Ramsey sequences form the basis of atomic clocks and internal state interferometers.

In addition to state transfer, MWs also allow for state-dependent MW potentials where states can be manipulated independently. This has been implemented on atom chips to entangle internal states and spatial modes[12], which is necessary for quantum information applications, and perform spin-squeezing[68], which allows for sensing below the standard quantum limit set by the uncertainty principle.

To summarise, with chip2 we hope to expand on chip1 by adding a MW source which will allow for

state-transfer and state-dependent potentials. Atom chips are particularly advantageous for this as near-field MW antennas require far less power than external antennas. Specifically for CMOS atom chips, since they are of a similar size to the wavelength of the required MW (0.44 mm), impedance matching and reflections are much less of a concern, resulting in far more efficient power transfer and less noise. Smaller changes will also be made in chip2 such as programmable amplitude and phase control for the RF. At the time of writing, work on the MW source and antenna has begun, but the design is still at an early stage.

Programmable Trapping Geometries

One of the potential advantages of IC-based atom chips is the ability to switch current paths using transmission gates. This would allow for both arbitrary trapping geometries to be produced on a single device such as U-wires, Z-wires, and ringtraps, whilst also allowing modifications in the polarisation of RF and MW fields and in-situ modifications to said geometry.

Integrated Current Sources

It is also feasible that current sources for the trapping wires are integrated into the chip. This would greatly simplify devices, allow for active feedback, and minimise noise picked up by the lead wires. Having the current source close to the trapping wires also minimises the power dissipated.

Feedback Systems

With the inclusion of Hall probes, temperature sensors, and photodiodes, it is possible to implement feedback systems directly onto the chip. For example, this could be used to actively minimise noise in the trapping wires while also maintaining the trap minima of an Ioffe trap at a desired value, which is necessary for any schemes using the clock states. Locking schemes for local oscillators, such as those necessary for a microwave clock, can also be included.

7.3 Summary

The existing applications of CMOS in quantum technologies suggest a positive outlook for CMOS-based atom chips, as many of the required elements have already been demonstrated in the context of other systems. Fluorescence detection with integrated photodiodes has been shown with both NV centres [58,38] and ions [64]; integrated RF generation and delivery from MHz to several GHz has been demonstrated in NMR [90,47] and NV centres, and ions have been successfully trapped on a chip produced in a commercial CMOS foundry. [54]

However, demonstrating trapping of neutral atoms on CMOS chips and measuring lifetimes and

coherence times is a priority before there can be too much optimism. The subsequent priority involves fluorescence detection, which may encounter challenges related to sensitivity and the influences of heat generated by the wires carrying current. To overcome these, APDs and SPADs may be necessary and the placement of the photodiodes will need to be optimised; shielding or heat sinking might also be required. Another issue with early CMOS chips might be the wire bonds which could block optical access and atom transport; to overcome this, through-substrate vias are a potential alternative.[30] Next, RF and MW generation and delivery are essential for interferometry and quantum information/clock schemes. Given the distance of atoms from the chip, power is unlikely to be an issue; noise, however, could be a concern. Dealing with this might just be a matter of more experienced design to protect sensitive circuits from noisy ones. Lastly, the digital circuits for pulse sequencing, programming, readout, and the interfacing between digital and analogue signals are vital for fully integrated chips. At this stage, space might start to become an issue, in which case multiple chips with interfacing could be required.

In summary, whilst there is still much to prove before integrated atom chips become a reality, there is a clear roadmap of existing technologies to guide their development. Moreover, adjacent to the Si wafer CMOS chips discussed in this work, chips based on other substrates could begin to integrate photonics[40,39], or have much higher current carrying capabilities.[46]

Bibliography

[1] M. a. Cirone, a. Negretti, T. Calarco, P. Krüger, and J. Schmiedmayer. A simple quantum gate with atom chips. *The European Physical Journal D*, 35:165–171, 6 2005. ISSN 1434-6060. doi: 10.1140/epjd/e2005-00175-8. URL http://www.springerlink.com/index/10.1140/epjd/e2005-00175-8.

[2] A. Aeppli, K. Kim, W. Warfield, M. S. Safronova, and J. Ye. A clock with $8 \times 10^{\{} - 19\}$ systematic uncertainty. *Physical Review Letters*, 133:23401, 2024. ISSN 10797114. doi: 10.1103/PhysRevLett.133.023401. URL http://arxiv.org/abs/2403.10664.

[3] G. Amruta. *A cold atom apparatus for the microscopy of thin membranes*. PhD thesis, University of Nottingham, 2017.

[4] J. Anders, M. Babaie, J. C. Bardin, I. Bashir, G. Billiot, E. Blokhina, S. Bonen, E. Charbon, J. Chiaverini, I. L. Chuang, C. Degenhardt, D. Englund, L. Geck, L. L. Guevel, D. Ham, R. Han, M. I. Ibrahim, D. Kruger, K. M. Lei, A. Morel, D. Nielinger, G. Pillonnet, J. M. Sage, F. Sebastiano, R. B. Staszewski, J. Stuart, A. Vladimirescu, P. Vliex, and S. P. Voinigescu. Cmos integrated circuits for the quantum information sciences. *IEEE Transactions on Quantum Engineering*, 4, 2023. ISSN 26891808. doi: 10.1109/TQE.2023.3290593.

[5] J. Appel, A. MacRae, and A. I. Lvovsky. A versatile digital ghz phase lock for external cavity diode lasers. *Measurement Science and Technology*, 20:055302, 2009. ISSN 09570233. doi: 10.1088/0957-0233/20/5/055302.

[6] R. J. Baker. *CMOS: Circuit Design, Layout, and Simulation*. Wiley & Sons, IEEE Press, 2010. ISBN 978-0-470-88132-3.

[7] J. Bardeen and W. H. Brattain. The transistor, a semi-conductor triode. *Physical Review*, 74: 230–231, 1948. ISSN 0031899X. doi: 10.1103/PhysRev.74.230.

[8] K. Beloy, M. I. Bodine, T. Bothwell, S. M. Brewer, S. L. Bromley, J. S. Chen, J. D. Deschênes, S. A. Diddams, R. J. Fasano, T. M. Fortier, Y. S. Hassan, D. B. Hume, D. Kedar, C. J. Kennedy, I. Khader, A. Koepke, D. R. Leibrandt, H. Leopardi, A. D. Ludlow, W. F. McGrew, W. R. Milner,

N. R. Newbury, D. Nicolodi, E. Oelker, T. E. Parker, J. M. Robinson, S. Romisch, S. A. Schäffer, J. A. Sherman, L. C. Sinclair, L. Sonderhouse, W. C. Swann, J. Yao, J. Ye, and X. Zhang. Frequency ratio measurements at 18-digit accuracy using an optical clock network. *Nature*, 591: 564–569, 2021. ISSN 14764687. doi: 10.1038/s41586-021-03253-4.

[9] T. Berrada, S. V. Frank, R. Bücker, T. Schumm, J. F. Schaff, and J. Schmiedmayer. Integrated mach-zehnder interferometer for bose-einstein condensates. *Nature Communications*, 4:1–8, 2013. ISSN 20411723. doi: 10.1038/ncomms3077.

[10] R. Bondarescu, A. Schärer, A. Lundgren, G. Hetényi, N. Houlié, P. Jetzer, and M. Bondarescu. Ground-based optical atomic clocks as a tool to monitor vertical surface motion. *Geophysical Journal International*, 202:1770–1774, 2015. ISSN 1365246X. doi: 10.1093/gji/ggv246.

[11] P. D. Breteuil. The 13th conférence générale des poids et mesures (cgpm). *Proceedings of the 13th CGPM*, 1967. URL https://doi.org/10.59161%2Fcgpm1967res1e.

[12] P. Böhi, M. F. Riedel, J. Hoffrogge, J. Reichel, T. W. Hänsch, and P. Treutlein. Coherent manipulation of bose-einstein condensates with state-dependent microwave potentials on an atom chip. *Nature Physics*, 5:592–597, 2009. ISSN 17452481. doi: 10.1038/nphys1329.

[13] A. D. Cronin, J. Schmiedmayer, and D. E. Pritchard. Optics and interferometry with atoms and molecules. *Reviews of Modern Physics*, 81:1051–1129, 2009. ISSN 00346861. doi: 10.1103/RevModPhys.81.1051.

[14] C. K. Dabhi, A. Dasgupta, H. Agrawal, N. Paydavosi, T. H. Morshed, D. D. Lu, M. V. Dunga, X. Xi, J. He, W. Liu, M. Cao, X. Jin, J. J. Ou, M. Chan, Y. S. Chauhan, A. M. Niknejad, and C. Hu. Bsim4 4.8.2 mosfet model user's manual, 2020. URL http://bsim.berkeley.edu/models/bsim4/.

[15] J. Dalibard and C. Cohen-Tannoudji. Laser cooling below the doppler limit by polarization gradients: simple theoretical models. *Journal of the Optical Society of America B*, 6:2023, 1989. ISSN 0740-3224. doi: 10.1364/josab.6.002023.

[16] K. Dawon. Electric field controlled semiconductor device. *U.S. Patent 3102230*, page 4, 1963.

[17] N. H. Dekker, C. S. Lee, V. Lorent, J. H. Thywissen, S. P. Smith, M. Drndic, R. M. Westervelt, and M. Prentiss. Guiding neutral atoms on a chip. *Physical Review Letters*, pages 2–5, 2000.

[18] P. Delva and J. Lodewyck. Atomic clocks: new prospects in metrology and geodesy. *Acta futura*, 7:67–78, 2013. doi: 10.2420/AF07.2013.67. URL http://arxiv.org/abs/1308.6766.

[19] A. Derevianko and M. Pospelov. Hunting for topological dark matter with atomic clocks. *Nature Physics*, 10:933–936, 2014. ISSN 17452481. doi: 10.1038/nphys3137.

[20] S. Dimopoulos, P. W. Graham, J. M. Hogan, M. A. Kasevich, and S. Rajendran. Atomic gravitational wave interferometric sensor. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 78:1–35, 2008. ISSN 15507998. doi: 10.1103/PhysRevD.78.122002.

[21] R. Folman, P. Kröger, D. Cassettari, B. Hessmo, T. Maier, and J. Schmiedmayer. Controlling cold atoms using nanofabricated surfaces: Atom chips. *Physical Review Letters*, 84:4749–4752, 2000. ISSN 10797114. doi: 10.1103/PhysRevLett.84.4749.

[22] R. O. N. Folman, P. K. R. Uger, P. Institut, and J. Denschlag. Microscopic atom optics :. *Advances in Atomic, Molecular, and Optical Physics*, 48:263–356, 2002.

[23] C. J. Foot. *Atomic Physics*. Oxford University Press.

[24] J. Fortágh and C. Zimmermann. Magnetic microtraps for ultracold atoms. *Reviews of Modern Physics*, 79:235–289, 2 2007. ISSN 00346861. doi: 10.1103/RevModPhys.79.235. URL http://link.aps.org/doi/10.1103/RevModPhys.79.235https://link.aps.org/doi/10.1103/RevModPhys.79.235.

[25] I.-T. G.8271.1/Y.1366.1. Network limits for time synchronization in packet networks with full timing support from the network, 11 2022.

[26] R. Geiger, D. Savoie, M. Altorio, B. Fang, and A. Landragin. Interleaved matter-wave gyroscope with 3 × 10-10 rad.s-1 stability. *CPEM 2018 - Conference on Precision Electromagnetic Measurements*, pages 1–2, 2018. doi: 10.1109/CPEM.2018.8501207.

[27] F. Gentile. *Towards an atomic Sagnac interferometer with full dynamical control of atoms in ring waveguides*. PhD thesis, University of Nottingham, 2019.

[28] M. Grewal, A. Andrews, and C. Bartone. *Global Navigation, Satellite Systems, Inertial Navigation and Integration*. Wiley & Sons, IEEE Press. ISBN 9781626239777.

[29] J. Grotti, S. Koller, S. Vogt, S. Häfner, U. Sterr, C. Lisdat, H. Denker, C. Voigt, L. Timmen, A. Rolland, F. N. Baynes, H. S. Margolis, M. Zampaolo, P. Thoumany, M. Pizzocaro, B. Rauf, F. Bregolin, A. Tampellini, P. Barbieri, M. Zucco, G. A. Costanzo, C. Clivati, F. Levi, and D. Calonico. Geodesy and metrology with a transportable optical clock. *Nature Physics*, 14: 437–441, 2018. ISSN 17452481. doi: 10.1038/s41567-017-0042-3. URL http://dx.doi.org/10.1038/s41567-017-0042-3.

[30] N. D. Guise, S. D. Fallek, K. E. Stevens, K. R. Brown, C. Volin, A. W. Harter, J. M. Amini, R. E. Higashi, S. T. Lu, H. M. Chanhvongsak, T. A. Nguyen, M. S. Marcus, T. R. Ohnstein, and D. W. Youngner. Ball-grid array architecture for microfabricated ion traps. *Journal of Applied Physics*, 117, 5 2015. ISSN 10897550. doi: 10.1063/1.4917385.

[31] S. Gundacker and A. Heering. The silicon photomultiplier: Fundamentals and applications of a modern solid-state photon detector. *Physics in Medicine and Biology*, 65, 2020. ISSN 13616560. doi: 10.1088/1361-6560/ab7b2d.

[32] T. L. Gustavson, A. Landragin, and M. A. Kasevich. Rotation sensing with a dual atom-interferometer sagnac gyroscope. *Classical and Quantum Gravity*, 17:2385–2398, 2000. ISSN 02649381. doi: 10.1088/0264-9381/17/12/311.

[33] M. Haw, N. Evetts, W. Gunton, J. V. Dongen, J. L. Booth, and K. W. Madison. Magneto-optical trap loading rate dependence on trap depth and vapor density. *Journal of the Optical Society of America B*, 29:475, 2012. ISSN 0740-3224. doi: 10.1364/josab.29.000475.

[34] H. Heidari. *Current-Mode High Sensitivity CMOS Hall Magnetic Sensors*. PhD thesis, University of Pavia, 2015.

[35] J. M. Hogan and M. A. Kasevich. Atom-interferometric gravitational-wave detection using heterodyne laser links. *Physical Review A*, 94:1–10, 2016. ISSN 24699934. doi: 10.1103/PhysRevA.94.033632.

[36] P. Hommelhoff, W. Hänsel, T. Steinmetz, T. W. Hänsch, and J. Reichel. Transporting, splitting and merging of atomic ensembles in a chip trap. *New Journal of Physics*, 7:0–17, 2005. ISSN 13672630. doi: 10.1088/1367-2630/7/1/003.

[37] W. Hänsel, P. Hommelhoff, T. W. Hänsch, and J. Reichel. Bose-einstein condensation on a microelectronic chip. *Nature*, 413:498–501, 10 2001. ISSN 00280836. doi: 10.1038/35097032. URL http://www.nature.com/articles/35097032http://www.ncbi.nlm.nih.gov/pubmed/11586353.

[38] M. I. Ibrahim, C. Foy, D. R. Englund, and R. Han. High-scalability cmos quantum magnetometer with spin-state excitation and detection of diamond color centers. *IEEE Journal of Solid-State Circuits*, 56:1001–1014, 3 2021. ISSN 1558173X. doi: 10.1109/JSSC.2020.3027056.

[39] M. H. Idjadi and F. Aflatouni. Integrated pound-drever-hall laser stabilization system in silicon. *Nature Communications*, 8:1–9, 2017. ISSN 20411723. doi: 10.1038/s41467-017-01303-y.

[40] A. Isichenko, N. Chauhan, D. Bose, J. Wang, P. D. Kunz, and D. J. Blumenthal. Photonic integrated beam delivery for a rubidium 3d magneto-optical trap. *Nature Communications*, 14, 2023. ISSN 20411723. doi: 10.1038/s41467-023-38818-6.

[41] N. M. H. Ismail and M. Othman. Cmos phase frequency detector for high speed applications. *2009 4th International Design and Test Workshop, IDT 2009*, pages 201–204, 2009. doi: 10.1109/IDT.2009.5404165.

[42] J. Johnson. *Atom-Chip Designs for a Trapped and Guided Matterwave Sagnac Interferometer*. PhD thesis, University of Nottingham, 2021.

[43] H. Kaeslin. *Digital Integrated Circuit Design: From VLSI Architecture to CMOS Fabrication*. Cambridge University Press, 2008. ISBN 978-0-511-45537-7.

[44] A. Kasper, S. Schneider, C. vom Hagen, M. Bartenstein, B. Engeser, T. Schumm, I. Bar-Joseph, R. Folman, L. Feenstra, and J. Schmiedmayer. A bose–einstein condensate in a microtrap. *Journal of Optics B: ...*, 5:143, 2003. URL http://iopscience.iop.org/1464-4266/5/2/372.

[45] M. Keil, O. Amit, S. Zhou, D. Groswasser, Y. Japha, and R. Folman. Fifteen years of cold matter on the atom chip: promise, realizations, and prospects. *Journal of Modern Optics*, 63:1840–1885, 2016. ISSN 13623044. doi: 10.1080/09500340.2016.1178820. URL http://dx.doi.org/10.1080/09500340.2016.1178820.

[46] T. Kimoto, J. A. Cooper, J. A. Cooper, and T. Kimoto. *Physical Properties of Silicon Carbide*, pages 11–38. Singapore: Wiley, 1 edition, 2014. doi: 10.1002/9781118313534.ch2.

[47] D. Kruger, A. Zhang, B. Aghelnejad, H. Hinton, V. M. Arnal, Y. Q. Song, Y. Tang, K. M. Lei, J. Anders, and D. Ham. A portable cmos-based spin resonance system for high-resolution spectroscopy and imaging. *IEEE Journal of Solid-State Circuits*, 58:1838–1849, 7 2023. ISSN 1558173X. doi: 10.1109/JSSC.2023.3274043.

[48] C. Lacroůte, F. Reinhard, F. Ramirez-Martinez, C. Deutsch, T. Schneider, J. Reichel, and P. Rosenbusch. Preliminary results of the trapped atom clock on a chip. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 57:106–110, 2010. ISSN 08853010. doi: 10.1109/TUFFc.2010.1385.

[49] M. D. Lakeh, J. B. Kammerer, W. Uhring, J. B. Schell, and F. Calmon. An ultrafast active quenching circuit for spad in cmos 28nm fdsoi technology. In *Proceedings of IEEE Sensors*, volume 2020-Octob. Institute of Electrical and Electronics Engineers Inc., 10 2020. ISBN 9781728168012. doi: 10.1109/SENSORS47125.2020.9278902.

[50] W. Liu, X. Jin, J. Chen, M. chie Jeng, Z. Liu, Y. Cheng, K. Chen, M. Chan, K. Hui, J. Huang, R. Tu, P. K. Ko, and C. Hu. Bsim 3v3.2.2 manual, 1999.

[51] A. D. Ludlow, M. M. Boyd, J. Ye, E. Peik, and P. O. Schmidt. Optical atomic clocks. *Reviews of Modern Physics*, 87:637–701, 2015. ISSN 15390756. doi: 10.1103/RevModPhys.87.637.

[52] J. Martin. *Analog Integrated Circuits Design*. John Wiley & Sons, Inc., 1997. doi: 10.1021/jo000589k.

[53] W. F. McGrew, X. Zhang, R. J. Fasano, S. A. Schäffer, K. Beloy, D. Nicolodi, R. C. Brown, N. Hinkley, G. Milani, M. Schioppo, T. H. Yoon, and A. D. Ludlow. Atomic clock performance enabling geodesy below the centimetre level. *Nature*, 564:87–90, 2018. ISSN 14764687. doi: 10.1038/s41586-018-0738-2. URL http://dx.doi.org/10.1038/s41586-018-0738-2.

[54] K. K. Mehta, A. M. Eltony, C. D. Bruzewicz, I. L. Chuang, R. J. Ram, J. M. Sage, and J. Chiaverini. Ion traps fabricated in a cmos foundry. *Applied Physics Letters*, 105:044103, 7 2014. ISSN 0003-6951. doi: 10.1063/1.4892061. URL https://doi.org/10.1063/1.4892061.

[55] P. Neumann. Design eines multifunktionalen integrierten optischen sensorchips für universelle anwendung in 0.18um cmos-technologie. Master's thesis, The University of Applied Sciences Wiener Neustadt, 2017.

[56] T. L. Nicholson, S. L. Campbell, R. B. Hutson, G. E. Marti, B. J. Bloom, R. L. McNally, W. Zhang, M. D. Barrett, M. S. Safronova, G. F. Strouse, W. L. Tew, and J. Ye. Systematic evaluation of an atomic clock at $2 \times 10^{\{} - 18\}$ total uncertainty. *Nature Communications*, 6:6896, 2015. ISSN 20411723. doi: 10.1038/ncomms7896.

[57] C. C. Nshii, M. Vangeleyn, J. P. Cotter, P. F. Griffin, E. A. Hinds, C. N. Ironside, P. See, A. G. Sinclair, E. Riis, and A. S. Arnold. A surface-patterned chip as a strong source of ultracold atoms for quantum technologies. *Nature Nanotechnology*, 8:321–324, 2013. ISSN 17483395. doi: 10.1038/nnano.2013.47.

[58] K. Omirzakhov, M. H. Idjadi, T. Y. Huang, S. A. Breitweiser, D. A. Hopper, L. C. Bassett, and F. Aflatouni. An integrated reconfigurable spin control system on 180 nm cmos for diamond nv centers. *IEEE Transactions on Microwave Theory and Techniques*, 71:4052–4063, 9 2023. ISSN 15579670. doi: 10.1109/TMTT.2023.3254600.

[59] S. Park, S. Song, K. Kim, H. Lee, M. Lee, and J. Y. Sim. A cryo-cmos 64-channel bias generator ic for surface ion trap. In *European Solid-State Circuits Conference*, pages 488–491. IEEE Computer Society, 2024. ISBN 9798350388138. doi: 10.1109/ESSERC62670.2024.10719474.

[60] Y. Paturel, J. Honthaas, H. Lefèvre, and F. Napolitano. One nautical mile per month fog-based strapdown inertial navigation system: A dream already within reach? *Gyroscopy and Navigation*, 5:1–8, 2014. ISSN 20751087. doi: 10.1134/S207510871401009X.

[61] H. Perrin and B. M. Garraway. Trapping atoms with radio frequency adiabatic potentials. *Advances in Atomic, Molecular and Optical Physics*, 66:181–262, 2017. ISSN 1049250X. doi: 10.1016/bs.aamop.2017.03.002.

[62] D. Pritchard. Atom interferometers. *Optics and Photonics News*, 2:28, 1991. ISSN 1047-6938. doi: 10.1364/opn.2.12.000028.

[63] B. Razavi. *Design of Analog CMOS Integrated Circuits*. McGraw Hill, 2001. ISBN 007-118839-8.

[64] D. Reens, M. Collins, J. Ciampi, D. Kharas, B. F. Aull, K. Donlon, C. D. Bruzewicz, B. Felton, J. Stuart, R. J. Niffenegger, P. Rich, D. Braje, K. K. Ryu, J. Chiaverini, and R. McConnell. High-fidelity ion state detection using trap-integrated avalanche photodiodes. *Physical Review Letters*, 129, 9 2022. ISSN 10797114. doi: 10.1103/PhysRevLett.129.100502.

[65] J. Reichel, W. Hänsel, and T. W. Hänsch. Atomic micromanipulation with magnetic surface traps. *Physical Review Letters*, 83:3398–3401, 10 1999. ISSN 10797114. doi: 10.1103/PhysRevLett.83.3398. URL http://link.aps.org/doi/10.1103/PhysRevLett.83.3398.

[66] G. Reinaudi, T. Lahaye, Z. Wang, and D. Guéry-Odelin. Strong saturation absorption imaging of dense clouds of ultracold atoms. *Optics Letters*, 32:3143, 2007. ISSN 0146-9592. doi: 10.1364/ol.32.003143.

[67] L. Rennhofer. Entwurf und simulation eines optoelektronischen systems zur detektion geringer lichtmengen. Master's thesis, The University of Applied Sciences Wiener Neustadt, 2021.

[68] M. F. Riedel, P. Böhi, Y. Li, T. W. H. Signnsch, A. Sinatra, and P. Treutlein. Atom-chip-based generation of entanglement for quantum metrology. *Nature*, 464:1170–1173, 2010. ISSN 00280836. doi: 10.1038/nature08988.

[69] L. Rieder. Implementation of an integrated hall sensor in a 0.35um cmos process. Master's thesis, The University of Applied Sciences Wiener Neustadt, 2022.

[70] F. Riehle, P. Gill, F. Arias, and L. Robertsson. The cipm list of recommended frequency standard values: Guidelines and procedures. *Metrologia*, 55:188–200, 2018. ISSN 16817575. doi: 10.1088/1681-7575/aaa302.

[71] Z. D. Romaszko, S. Hong, M. Siegele, R. K. Puddy, F. R. Lebrun-Gallagher, S. Weidt, and W. K. Hensinger. Engineering of microfabricated ion traps and integration of advanced on-chip features.

[72] J. Rudolph, W. Herr, C. Grzeschik, T. Sternke, A. Grote, M. Popp, D. Becker, H. Müntinga, H. Ahlers, A. Peters, C. Lämmerzahl, K. Sengstock, N. Gaaloul, W. Ertmer, and E. M. Rasel. A high-flux bec source for mobile atom interferometers. *New Journal of Physics*, 17, 2015. ISSN 13672630. doi: 10.1088/1367-2630/17/7/079601.

[73] M. S. Safronova, D. Budker, D. Demille, D. F. Kimball, A. Derevianko, and C. W. Clark. Search for new physics with atoms and molecules. *Reviews of Modern Physics*, 90:25008, 2018.

ISSN 15390756. doi: 10.1103/RevModPhys.90.025008. URL https://doi.org/10.1103/RevModPhys.90.025008.

[74] C. T. Sah. Characteristics of the metal-oxide-semiconductor transistors. *IEEE Transactions on Electron Devices*, XII:1365–1376, 1928.

[75] C. Salomon, J. Dalibard, W. D. Phillip, A. Clairon, and S. Guellat. Laser cooling of cesium atoms below μk. *Europhys. Lett.*, 12:683–688, 1990. ISSN 12864854. doi: 10.1209/0295-5075/12/8/003.

[76] T. Schumm, S. Hofferberth, L. M. Andersson, S. Wildermuth, S. Groth, I. Bar-Joseph, J. Schmiedmayer, and P. Krüger. Matter-wave interferometry in a double well on an atom chip. *Nature Physics*, 1:57–62, 9 2005. ISSN 1745-2473. doi: 10.1038/nphys125. URL http://www.nature.com/doifinder/10.1038/nphys125http://www.nature.com/articles/nphys125.

[77] H. Shichman and D. A. Hodges. Modeling and simulation of insulated-gate field-effect transistor switching circuits. *IEEE Journal of Solid-State Circuits*, 3:285–289, 1968. ISSN 1558173X. doi: 10.1109/JSSC.1968.1049902.

[78] W. Shockley. The theory of p-n junctions in semiconductors and p-n junction transistors. *The Bell System Technical Journal*, 28:435–489, 1949.

[79] A. M. Steane, M. Chowdhury, and C. J. Foot. Radiation force in the magneto-optical trap. *Journal of the Optical Society of America B*, 9:2142, 1992. ISSN 0740-3224. doi: 10.1364/josab.9.002142.

[80] D. A. Steck. *Quantum and Atom Optics*. available online at http://steck.us/teaching (revision 0.16, 26 April 2024)., 2007. URL http://steck.us/teaching%0Ahttp://www.opencontent.org/openpub/%0Ahttp://steck.us/teaching.

[81] D. A. Steck. Rubidium 87 d line data. *Journal of Geophysical Research*, 2009:31, 2010. ISSN 00301299. doi: 10.1029/JB075i002p00463. URL http://steck.us/alkalidata.

[82] B. Steindl, R. Enne, S. Schidl, and H. Zimmermann. Linear mode avalanche photodiode with high responsivity integrated in high-voltage cmos. *IEEE Electron Device Letters*, 35:897–899, 2014. ISSN 07413106. doi: 10.1109/LED.2014.2336678.

[83] J. Stuart, R. Panock, C. D. Bruzewicz, J. A. Sedlacek, R. McConnell, I. L. Chuang, J. M. Sage, and J. Chiaverini. Chip-integrated voltage sources for control of trapped ions. *Physical Review Applied*, 11, 2 2019. ISSN 23317019. doi: 10.1103/PhysRevApplied.11.024010.

[84] R. Szmuk, V. Dugrain, W. Maineult, J. Reichel, and P. Rosenbusch. Stability of a trapped-atom clock on a chip. *Physical Review A - Atomic, Molecular, and Optical Physics*, 92:012106, 2015. ISSN 10941622. doi: 10.1103/PhysRevA.92.012106.

[85] M. Takamoto, Y. Tanaka, and H. Katori. A perspective on the future of transportable optical lattice clocks. *Applied Physics Letters*, 120:140502, 4 2022. ISSN 00036951. doi: 10.1063/5.0087894.

[86] E. Tiesinga, P. J. Mohr, D. B. Newell, and B. N. Taylor. Codata recommended values of the fundamental physical constants: 2018. *Reviews of Modern Physics*, 93:25010, 2021. ISSN 15390756. doi: 10.1103/RevModPhys.93.025010. URL https://doi.org/10.1103/RevModPhys.93.025010.

[87] M. Vermeer. Chronometric levelling. *Reports of the Finnish Geodetic Institute*, pages 1–7, 1983.

[88] F. Wenzl. Master thesis test einer integrierten atomfalle test einer integrierten atomfalle. Master's thesis, The University of Applied Sciences Wiener Neustadt, 2021.

[89] K. Wintersperger, F. Dommert, T. Ehmer, A. Hoursanov, J. Klepsch, W. Mauerer, G. Reuber, T. Strohm, M. Yin, and S. Luber. Neutral atom quantum computing hardware: performance and end-user perspective. *EPJ Quantum Technology*, 10:1–26, 2023. ISSN 21960763. doi: 10.1140/epjqt/s40507-023-00190-1. URL http://dx.doi.org/10.1140/epjqt/s40507-023-00190-1.

[90] A. Zhang, D. Krüger, B. Aghelnejad, G. Yang, H. Hinton, Y.-Q. Song, and D. Ham. A portable wideband cmos nmr spectrometer for multinuclear molecular fingerprinting. *IEEE Journal of Solid-State Circuits*, pages 1–11, 7 2024. ISSN 0018-9200. doi: 10.1109/jssc.2024.3362808.

[91] Y. Zheng and C. E. Saavedra. Compact cmos vco using a transistor for frequency control. *Antem/URSI 2006 - 12th International Symposium on Antenna Technology and Applied Electromagnetics and Canadian Radio Sciences Conference, Proceedings*, pages 3–6, 2006.

Appendices

Appendix A

PCB Layers

All 12 layers of the PCB are shown

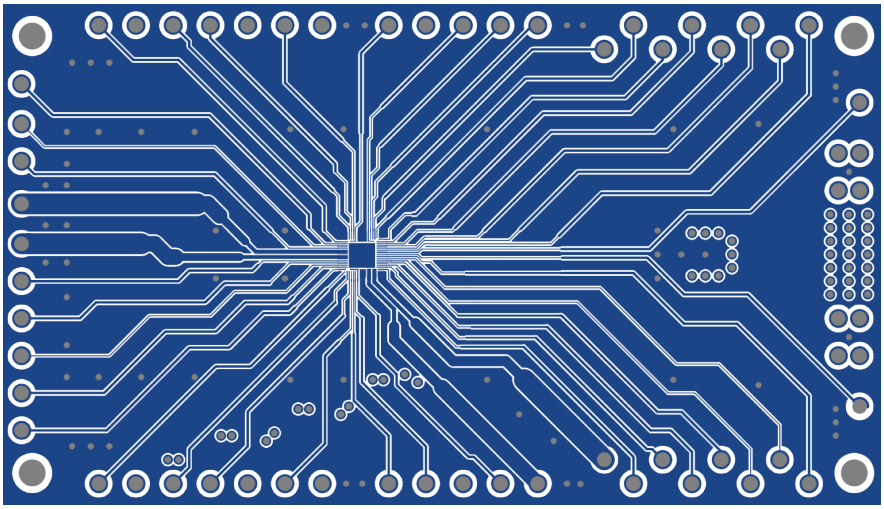


Figure A.1: Top Layer

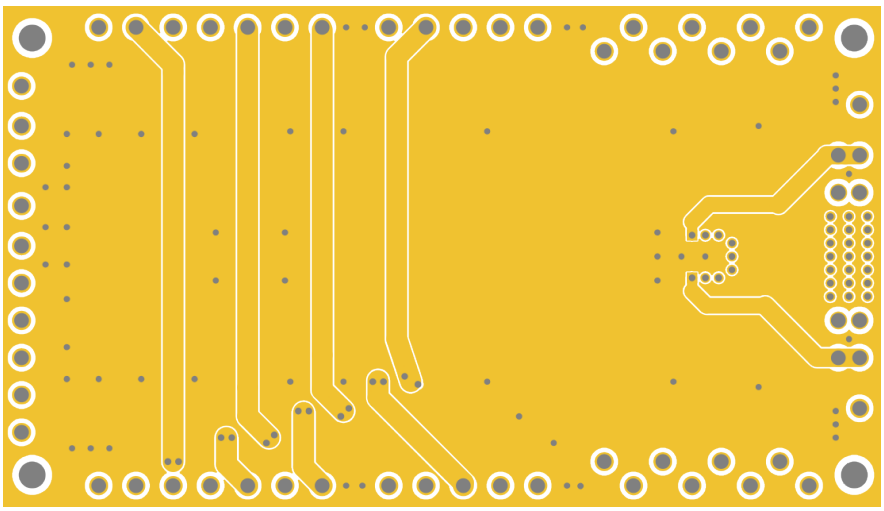


Figure A.2: Layer 1

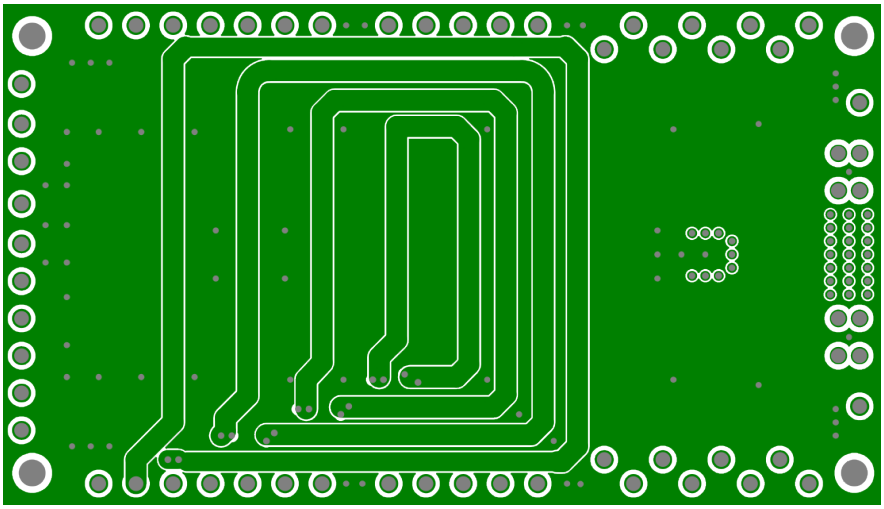


Figure A.3: Layer 2

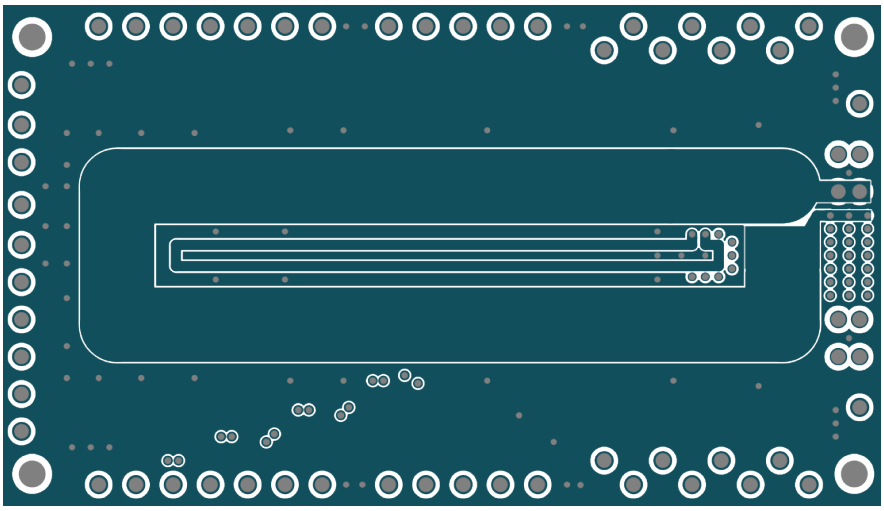


Figure A.4: Layer 3

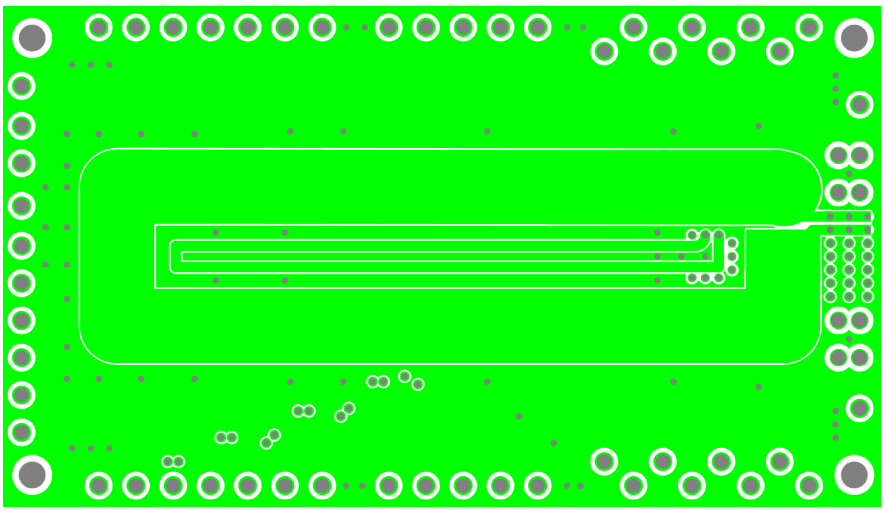


Figure A.5: Layer 4

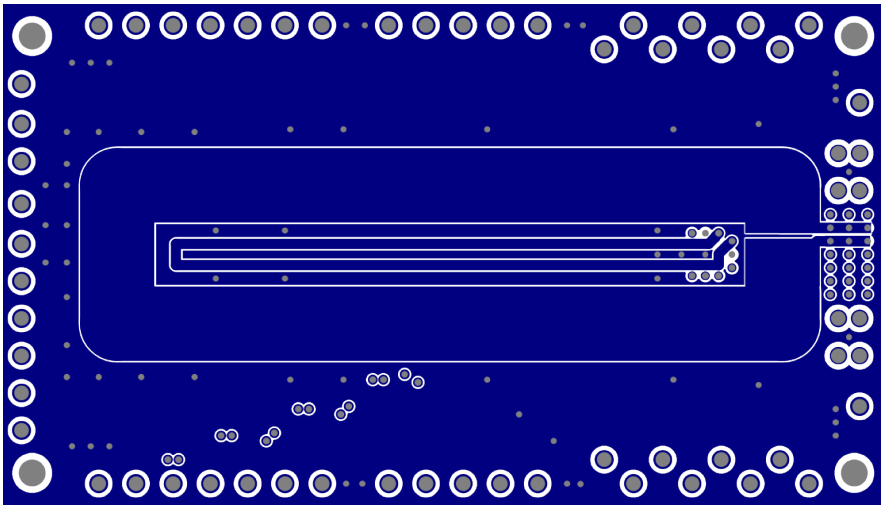


Figure A.6: Layer 5

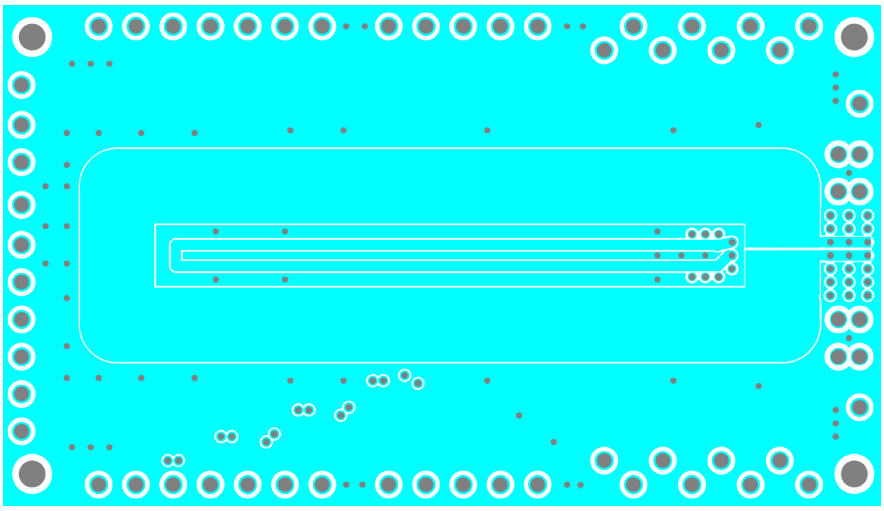


Figure A.7: Layer 6

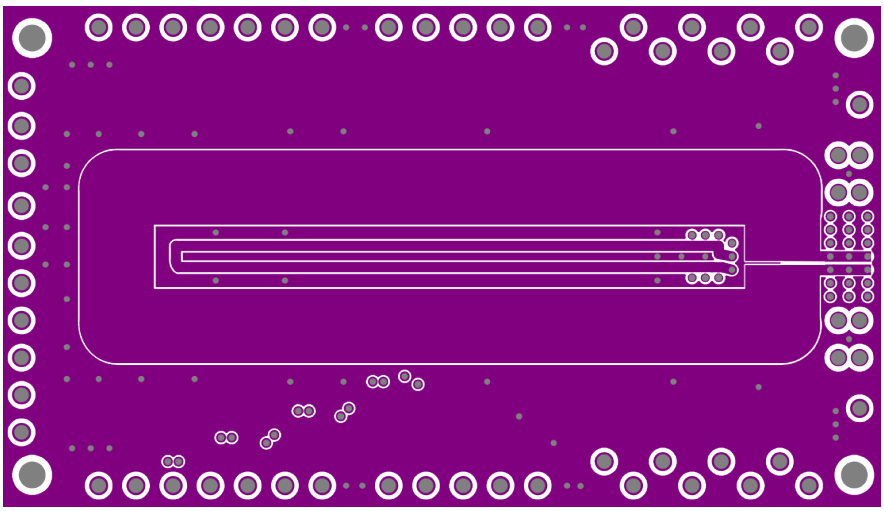


Figure A.8: Layer 7

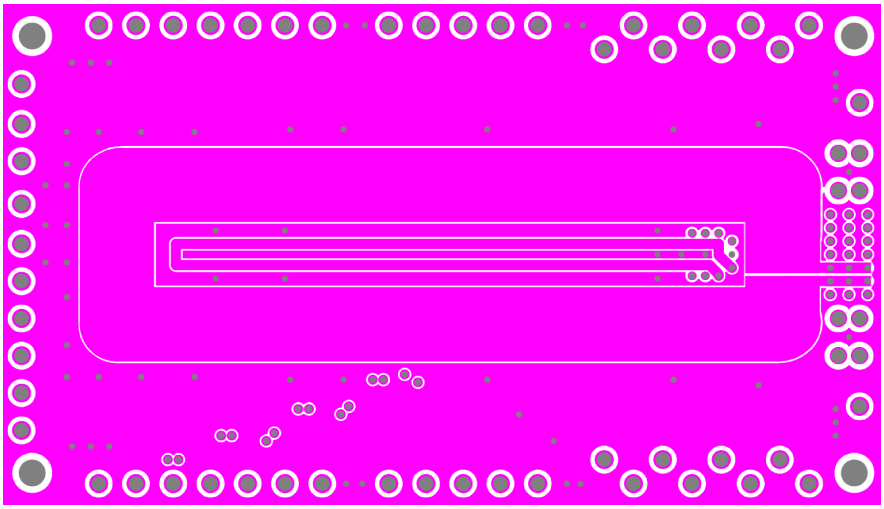


Figure A.9: Layer 8

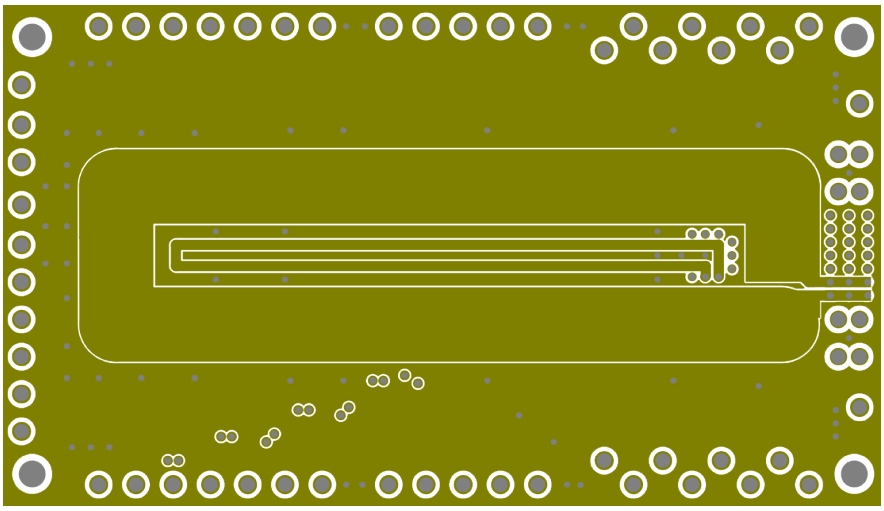


Figure A.10: Layer 9

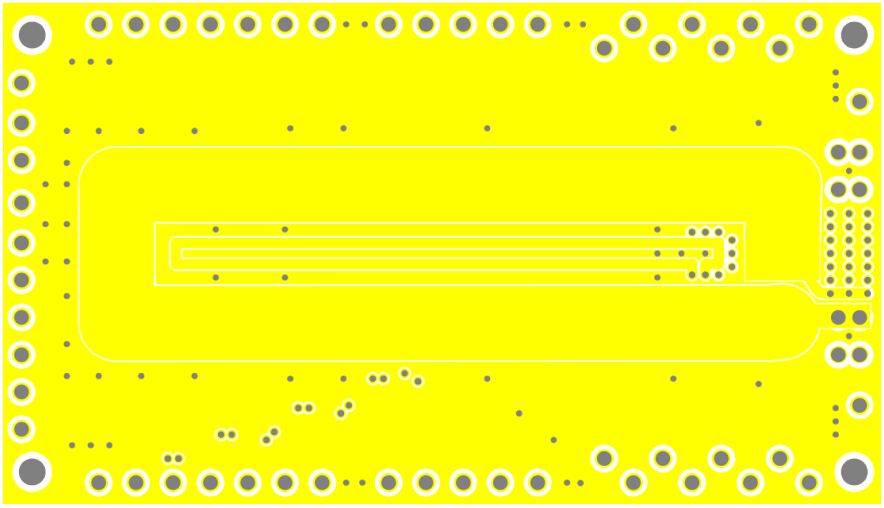


Figure A.11: Layer 10

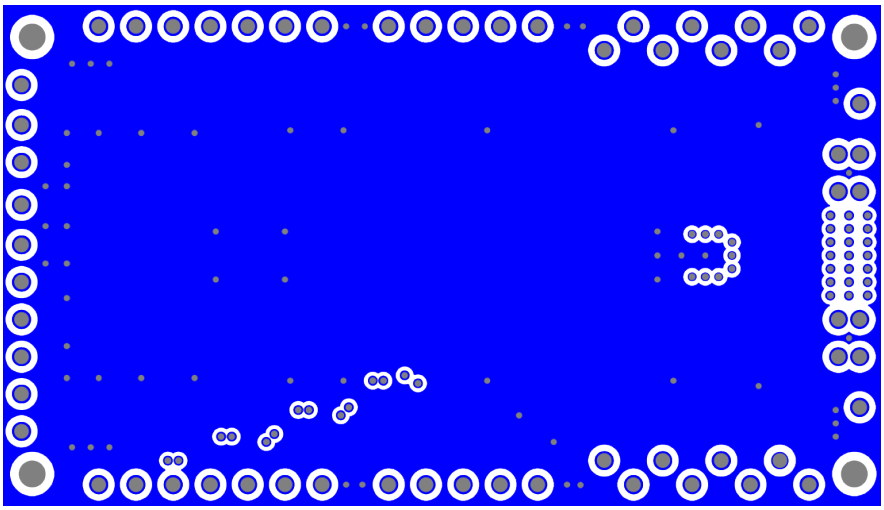


Figure A.12: Bottom Layer

Appendix B

Chip1 Pin-Out

In this chapter of the appendix, the pins for chip1 are given.

Top		Right		Bottom		Left	
Pin	Name	Pin	Name	Pin	Name	Pin	Name
1	vdd	20	gnd	44	ZL-	64	gnd
2	RF_SIG	21	V_REF_ADC	45	ZL-	65	DUMMY
3	RF_SW	22	DIVIDER_CTR_0	46	ZL-	66	IBias2
4	gnd	23	DIVIDER_CTR_1	47	ZS-	67	IBias1
5	OUTA_A	24	DIVIDER_CTR_2	48	ZS-	68	vdd
6	OUTD_A	25	VCO_CONTROL	49	ZS-	69	RFin2
7	RSTQ_S	26	PhiR	50	ZS-	70	Z_In_S
8	RST_S	27	HALL_OUT_1_1	51	vdd	71	RFin1
9	WR_S	28	HALL_OUT_1_2	52	Z_SW_1	72	vdd120
10	EN_S	29	HALL_OUT_2_1	53	Z_SW_2	73	ZL+
11	OUTD_S	30	HALL_OUT_2_2	54	RSTC_1	74	ZL+
12	CLK_OUT	31	APD1	55	OUTA_S_3	75	ZL+
13	SCLK_OUT	32	APD2	56	OUTA_S_2	76	ZL+
14	CLK_SW	33	SPAD1	57	OUTA_S_1	77	ZS+
15	MOSI	34	SPAD2	58	WR_A	78	ZS+
16	MISO	35	VBIAS	59	RST_A	79	ZS+
17	CLK_REF	36	RFOut1	60	CTRL_ACAP	80	ZS+
18	vss	37	Z_OUT_S	61	EN_A	81	RSTC_2
19	vdd	38	RFOut2	62	vdd	82	OUT_S_6
		39	ENABLE_CS	63	gnd	83	OUT_S_5

			40	DUMMY				84	OUT_S_4
			41	vss				85	gnd
			42	gnd					
			43	ZL+					