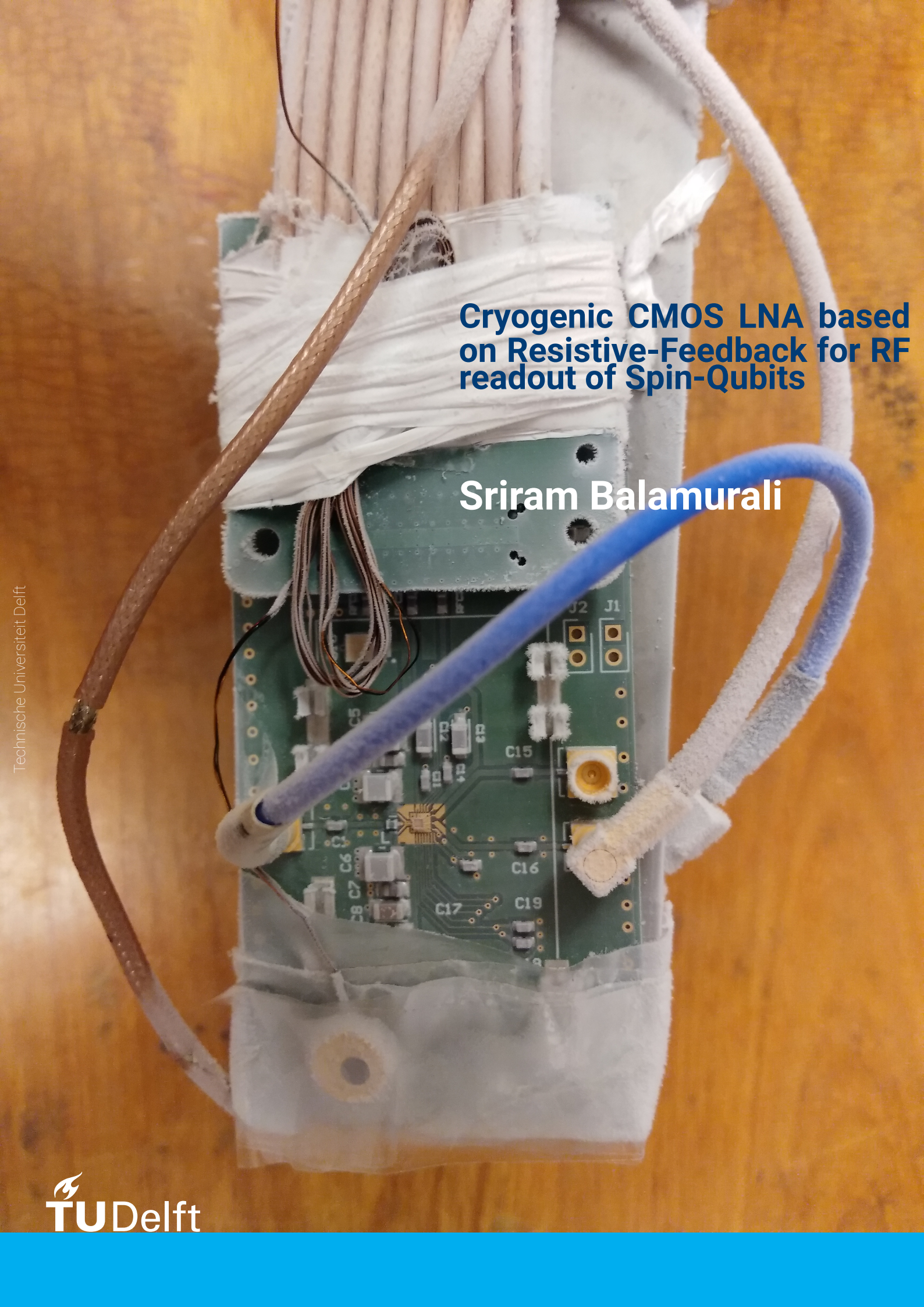

Cryogenic CMOS LNA based on Resistive-Feedback for RF readout of Spin-Qubits

Sriram Balamurali

TU Delft

List of Figures

Cryogenic CMOS LNA based on Resistive-Feedback for RF readout of Spin-Qubits

by

Sriram Balamurali

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday Nov 28, 2018 at 3:00 PM.

Student number:	4622510	
Project duration:	Sept 1, 2017 – Nov 28, 2018	
Thesis committee:	Prof. Edoardo Charbon,	TU Delft
	Dr. Fabio Sebastiano	TU Delft ,supervisor
	Dr. Morteza Alavi,	Tu Delft

This thesis is confidential and cannot be made public until Nov, 2023.

TU Delft Delft University of Technology

Acknowledgment

I would first like to thank my parents and my brother for their support and continuous encouragement during two years of my masters.

I would like to thank Dr.Fabio Sebastiano for accepting me as his master student and his assistance during the project. I would also like to thank Prof Edoardo Charbon for allowing me to work with such a wonderful group. I learned a lot from both the PhD and the fellow master students in the group. I also appreciate Dr. Masoud Babaie's help and feedbacks in the design review and the PCB. Thanks also to Dr. Morteza Alavi for accepting the proposal to be in my thesis committee and taking the time to read my thesis.

I would also like to especially acknowledge Milad and Rosario for their help in the design, Harald for helping me with the measurement setup and Gerd for his support in the review of the thesis. I also appreciate the support I got from Atef during the tapeout as well as Zu Yao and Lukasz for their help with the PCB. I thank my fellow master students Job, Jaco, and Edwin for all the stimulating discussions we had in circuit design and device physics.

Finally, I appreciate the time spent with my friends from the microelectronics Sarthak, Amitava, Rushil, Ashish and Sunil. I also learned a lot from them through discussions. I would especially like to thank Amitava for allowing me to stay at his home during the last month of my thesis work.

Last but not the least I would like to thank all my professors and lecturers here at TU Delft for the learning's I would take from here.

Sriram Balamurali
Delft, November 2018

Abstract

The promise that quantum computing is going to solve intractable computational problems has drawn tremendous interest and motivation in research communities throughout the world. However, operating a large-scale quantum computer requires classical electronics. In the present experimental setups, classical electronics is placed at room temperature rather than at cryogenic environment where qubits operate, hence requiring a huge amount of interconnections. Considering future scaling of qubits, classical electronics for readout and control must be placed close to these qubits. This thesis project explores the readout of spin qubits which are one of the most promising candidates for a large-scale quantum computer. The readout chain is one of the essential component for operating the large-scale quantum computer. Baseband and RF readout techniques for readout of spin qubits were explored in this thesis project. Based on this, the decision was taken to implement amplifier for RF readout as this gives the option for multiplexed readout of multiple qubits over a single wire connecting the quantum processor and the control electronics thus minimizing the number of readout lines from 20mK to the 4K stage where cryogenic electronics is envisioned to be operating. Limited scaling of the transistor drain noise at cryogenic temperatures was explored. Since many qubit channels are to be multiplexed broadband LNA is required. Resistive-Feedback was selected as a suitable topology to be implemented in TSMC 40nm CMOS. This is due to its simplicity in providing input match and robust broadband noise performance. The test chip was fabricated and tested both at the room and cryogenic temperatures. When operating at the room temperature the LNA achieved minimum noise temperature of 27K at 200MHz which increases up to 45K at 1GHz with a power consumption of 67mW. The S11 is greater than -30dB at 200MHz and increases upto -10dB at 800MHz. Preliminary measurement results at 5K show the LNA achieving minimum noise temperature close to 4K at 400MHz which increases to 6K at 1GHz with power consumption of close to 37mW. Following table shows the comparison with the prior art [15] noise cancelling CMOS LNA meant for cryogenic applications designed in the group. Their are no other existing broadband cryogenic LNA's in CMOS. The S parameter measurements are yet to be taken at cryogenic temperatures. Assuming 32 qubit channels with channel spacing 25MHz, this corresponds to 1.2mW power per qubit for readout. Low noise temperature of the LNA would enable higher readout bandwidths. Also, broadband low noise temperature enabled by resistive feedback will increase the number of channels to be multiplexed and will result in the reduction in interconnections.

WorK	Minimum Noise temperature(300K)	Core LNA Gain(300K)	S11 (300K)	Power Core LNA
JSSC[15]	60K	26.5dB	>-10dB	46mW
This Work	27K	42dB	<-10dB till 800MHz	67mW

Work	Minimum Noise temperature	Core LNA Gain(4K)	S11 (4K)	Power Core LNA
JSSC[15]	7K(4K base temperature)	NA	>-10dB	55mW
This Work	4K(5K base temperature)	NA	NA	37mW

Contents

1 Introduction

1.1. Quantum Computing and Quantum Error Correction

The idea of quantum computing dates back to 1981 when physicist Richard Feynman in his seminal paper first suggested a new paradigm in computing to enable the simulation of quantum systems [1]. Following Feynman, David Deutsch worked on the computing device that could simulate any arbitrary physical system which ultimately resulted in a computing paradigm based on the principles of quantum mechanics [2]. However, the first real computational problem for which quantum computers can outperform classical computers was shown by Peter Shor in 1994 [3]. Shor proposed how one can efficiently find prime factors of an integer using quantum mechanical principles like superposition and entanglement. This ultimately caused the academic interest in the field to explode. All classical computers process information using classical bits which are represented as 1 and 0. Its quantum mechanical counterpart is called a quantum bit, or qubit. Unlike a classical bit, a qubit can be in a superposition of both 0 and 1, which can be written using the Dirac notation as $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ with $|\psi\rangle$ the qubit state. After the qubit is measured, its state collapses to either state $|0\rangle$ or $|1\rangle$ with probability $|\alpha|^2$ and $|\beta|^2$, respectively, where α and β are complex amplitudes. The qubit operates in a Hilbert space, i.e. a complex vector space. Similary, two qubits can be represented as $\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$. More generally, to represent n qubits one needs 2^n complex amplitudes. A few simple operations on a quantum state effects all the 2^n complex amplitudes in the Hilbert space. This is the primary reason for the exponential speed up of quantum computers. Using this parallelism together with quantum interference, one can achieve tremendous computational power for certain class of problems. Imagine a quantum circuit U operating on an quantum register $|x,0\rangle$ to give $|x, f(x)\rangle$ [4]. If there are many inputs x_1,x_2 to x_n a classical computer in principle has to perform the computation n times. However, using the principles of quantum mechanics we can prepare quantum superposition of these n states

$$\psi = \frac{1}{2^{\frac{n}{2}}}(|x_1,0\rangle + |x_2,0\rangle .... |x_n,0\rangle) \tag{1.1}$$

Now applying the quantum circuit U only once, we will modify all the superpositions in the Hilbert space to

$$\psi = \frac{1}{2^{\frac{n}{2}}}(|x_1, f(x_1\rangle + |x_2, f(x_2)\rangle .... |x_n, f x_n\rangle \tag{1.2}$$

Hence one can see there is inherently a speed-up due to the quantum parallelism. Unfortunately, measurement will cause the resulting state to be selected randomly out of one of these states. Hence, in order to ripe the full benefits of quantum parallelism one has to use quantum interference. As an example, suppose we want to compute the function $G = f(0) \oplus f(1)$ [4]. Classically we need to compute $f(0)$ and $f(1)$ twice and then do the the comparison to do the computation. However, by using the principles of quantum mechanics one can do this in one step. First, one needs a quantum circuit which computes $|x, b \oplus f(x)\rangle$ from $|x,b\rangle$. After preparing the quantum superposition of the two input states $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, applying U gives

$$\frac{1}{2}(|0,0 \oplus f(0)\rangle - |0,1 \oplus f(0)\rangle + |1,0 \oplus f(1)\rangle - |1,1 \oplus f(1)\rangle) \tag{1.3}$$

This can be written as follows.

$$\frac{1}{2}(|0, f(0)\rangle - |0, f(\bar{0})\rangle + |1, f(1)\rangle - |1, f(\bar{1})\rangle) \tag{1.4}$$

Here $f(\bar{x})$ denotes the complement of the result $f(x)$ e.g. if $f(x) = 1$ then $f(\bar{x}) = 0$ and vice versa. Again applying quantum transformation on first qubit $|0\rangle$ to $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|1\rangle$ to $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, state of first qubit ends up at $|f(0) \oplus f(1)\rangle$, which is exactly the function we wanted to compute, but we did not need to do two separate computations as in the classical computer. The final state which results from this operation is the superposition of the following states.

$$|0, f(0)\rangle + |1, f(0)\rangle - |0, f(\bar{0})\rangle - |1, f(\bar{0})\rangle + |0, f(1)\rangle - |1, f(1)\rangle - |0, f(\bar{1})\rangle + |1, f(\bar{1})\rangle \tag{1.5}$$

Intuitively it can be seen from the superposition given in equation 1.5, when $f(0) = f(1)$ some of the possibilities of having the first qubit in state 1 interfere destructively to cancel out each other while possibilities having the first qubit state 0 interfere destructively when $f(0) = \bar{f}(1)$. This is the essence of quantum engineering and algorithms, i.e. to reinforce probability amplitudes only where desired. A full-scale quantum computer could have applications ranging from quantum chemistry which involves simulation of molecules for drug synthesis to hard optimization problems and encryption. Efficient quantum algorithms utilize large-scale quantum interference which is very fragile and very sensitive to noise [5]. Even though the final outcome is digital after the measurement, the computational process involves interference of probability waves which is analog and susceptible to noise. This makes large-scale quantum computing without error correction practically impossible. Hence, like classical error correction, quantum error correction is required for robust fault tolerance, and thus to have one logical qubit, one needs many physical qubits. Extra redundant qubits are needed for error correction. The discovery of power full error correction methods [6, 7] put quantum computing from mere academic curiosity to practical possibility. The class of problems that quantum computers can solve are still being explored. The most advanced quantum computer in 2018 is a 72-bit quantum processor by google [8] named bristlecone made of superconducting qubits which is still being tested. Big investments in the field by corporate giants like Google, IBM and Intel has boosted the research in the field such that the experts believe that we could have a full scale fault tolerant workable quantum computer working on practical problems within the next few decades.

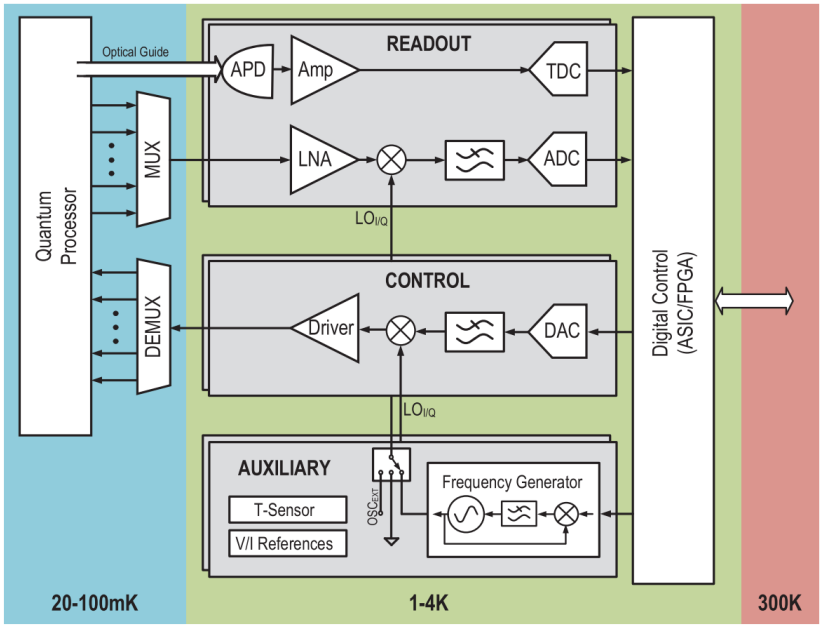

Figure 1.1: Classical control electronics for QEC showing readout and control (taken from [12])

1.2. Classical electronics for Quantum computation

As stated in the previous section that the quantum states are very fragile and are sensitive to noise. Hence quantum error correction (QEC) is an absolute necessity for the working of the quantum computer [5–7].

The minimum number of physical qubits needed to correct single-qubit errors is of the order of 5 [9]. Hence to emulate only a few logically error corrected qubits, more than 5 times the number of physical qubits are required. In the near term, more than 150 error corrected qubits are required for applications such as quantum chemistry and machine learning [10]. This requires close to equivalent 1000 physical qubits. In the present state of the art setups, physicists are able to control ≤ 20 physical qubits with the electronics at room temperature using long cables. This presents scalability issues when qubits are scaled to a few thousand in the near term future. Hence there is a need for the control electronics needed for the error correction loop to run in close proximity to qubits at cryogenic temperatures. The classical control unit working in close proximity to qubits provide that control [11–13]. Figure 1.1 shows the envisioned classical control [12]. The qubits operate at 20mK base temperature. The electronics needed for qubit readout and control is envisioned to be operating close to qubits at 4K base temperature. Due to 1 watt of cooling power of the dilution refrigerator at the 4K stage, this base temperature is envisioned to operate control circuitry needed for error correction. CMOS in this regard provides an advantage in the complex system on chip integration which will eventually be required when the number of physical qubit scales up. Spin qubits are many of the promising candidates for full-scale FTQC. In future the spin-qubits are expected to be more closely integrated with control electronics and the possibility of this is further explored [14]. Hence with respect to silicon qubits, CMOS can prove advantageous with respect to co-integration.

1.3. Thesis outline and objective

As seen in figure 1.2, LNA (Low Noise Amplifier) is used in classical readout chain. The present thesis concerns with LNA (Low Noise Amplifier) design for readout of spin-qubits. Present day setups use off the shelf electronics for readout and control at 300K with a lot of physical connections coming from 20mK where qubits are kept. This will not be practical when the number of physical qubits scales up. Hence the LNA designed in this thesis will be part of the readout chain used to readout qubit's spin signature while operating at the base temperature of 4K.

Chapter 2 will discuss the fabrication and readout of spin qubits. Also, readout specifications are derived. Chapter 3 will discuss various low noise amplifier topologies step by step to elucidate the reason for choosing the adapted architecture. Chapter 4 will discuss the design of the low noise amplifier and the performance of the amplifier in the simulations. Chapter 5 discusses the measurement setup of the adapted LNA and the measurement results. Conclusion and the future directions are discussed in the last chapter.

2

Readout of Spin Qubits

2.1. Spin Qubit implementation and readout

Spin qubits are one of the many promising candidates for fault-tolerant quantum computation. The information here is encoded in the electron spin which is controlled using microwave pulses [15]. Other competing implementations are superconducting, topological, and trapped ion qubits [16]. Spin qubits are much smaller in physical size and have long coherence times. The fabrication of spin qubits also highly resemble modern day semiconductor fabrication process making them amenable to scaling. Also, spin qubits have the potential to operate at 1-4K as opposed to 20mK and hence having the potential to operate much closer to the control electronics [14, 17].

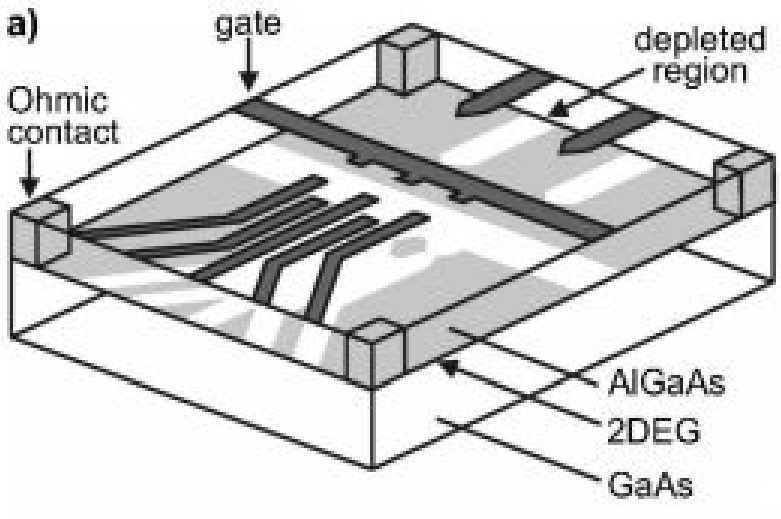

Figure 2.1: Formation of Quantum dot in AlGaAs and GaAs heterojunction showing how a single electron can be isolated from the sea of electrons by applying suitable gate potential (reproduced from [18])

Figure 2.1 shows how a single electron can be isolated and controlled using a quantum dot to form the spin qubit [18]. In Figure 2.1, the grey region indicates 2D electron gas formed when AlGaAs is deposited on top of GaAs. The bandgap difference between the two heterojunctions results in the formation of excess electrons on the surface. Applying a negative gate potential on the electrodes can deplete electrons from the region underneath the gate. This is indicated as the white region in the Figure 2.1. The gate voltages can be controlled to form a quantum dot which is shown in the figure as an isolated gray island. Inside the dot, a single electron can be controlled and manipulated. Quantum information in the spin-qubits is stored in the form of the electron spin. However, magnetic moment of a single electron is too low to be detected. Hence

before the readout electron spin information is converted to charge information. This is then read out using charge sensors like QPC (Quantum point contact) or SET (Single-electron transistor). QPC and SET have to be integrated close to the quantum dot to maximize cahrge sensitivity. This poses some integration challenges. RF-SET which is one of the most sensitive electrometers is used in this thesis to derive the specifications. To overcome the problems with integration DGS (Dispersive gate sensing) readout was developed which makes use of gates which are already in use to define quantum dots. The change in quantum capacitance is then detected using reflectometry [19].

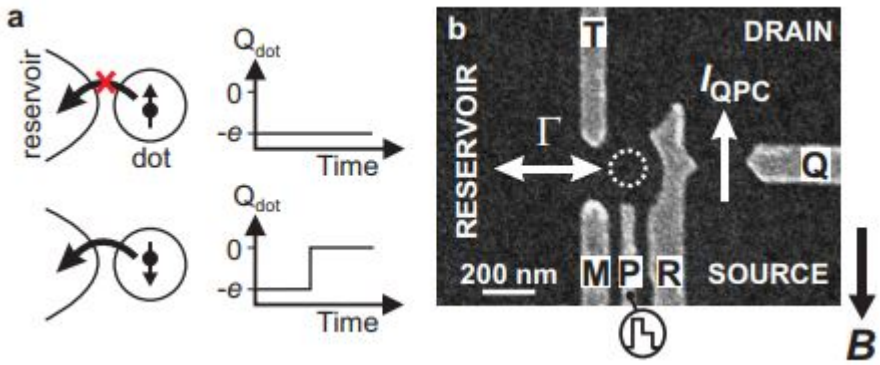

Figure 2.2: Quantum dot coupled to QPC acting as spin to charge sensor. (reproduced from [18])

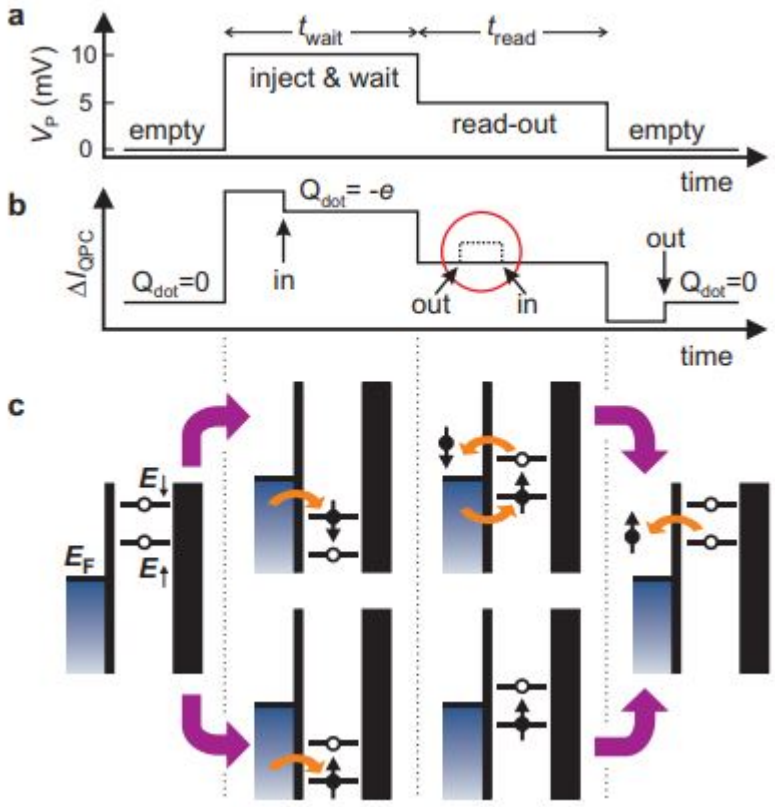

Figure 2.3: Elzerman Readout showing how momentarily the electron with the spin down state leaves the dot. This brief absence of an electron from the dot changes the channel potential of QPC ultimately changing the current briefly. (reproduced from [20])

To gain the basic intuition into the readout scheme one can consult Figure 2.2 and 2.3 which demonstrates the Elzerman readout using QPC [20]. The gates M, R, and T are used to create a dot. The potential of the dot can be controlled such that it contains either no electron or one electron. As seen in Figure 2.3

the energy state of the electron is first split by applying an external magnetic field. The electrons now align either parallel or anti-parallel to magnetic field depending upon the spin. This phenomenon is called the Zeeman's splitting and causes the energy difference in the two spin states. The resulting two states encode quantum information as $|0\rangle$ and $|1\rangle$. First, the gate potential is set such that potential inside the dot is more than the Fermi level of the reservoir so that electrons can tunnel. After the tunneling of one electron the potential inside the dot reduces resulting in Coulomb blockade and hence no electron can now enter the dot. In the read phase the potential of the well is again changed to a higher potential using gate electrodes. The potential is tuned such that if the electron that entered the dot was spin up then the electron cannot tunnel further and no change will happen. However, spin down electron having higher energy will have a tendency to tunnel through the barrier into the reservoir. This will be followed by spin up electron taking its place. Hence for a brief moment, the electron leaves the dot. The potential of the dot as shown in the figure 2.2 is electrostatically coupled to QPC which results in small change current ΔI_{QPC}. Instead of QPC more sensitive charge sensor SET (Single-electron transistor) can be used [21, 22]. Here the dot potential is coupled to channel of the transistor. Information on the spin is then converted into the resistance of the SET. The qubits and the sensors are kept at 20mK and integrated close to each other. Heisenberg uncertainty principle ($\Delta E \Delta T \geq h \implies \frac{e^2 RC}{2C} \geq h \implies R \geq \frac{2h}{e^2}$) puts lower bound on SET resistance for single electron transport hence nominal SET resistance value is 100k. The resistance changes by approximately 10% depending upon the spin [23]. The change in the resistance of these charge sensors can be read either using RF reflectometry or baseband DC readout which detects the change in the resistance. In the next few sections the motivation for using RF readout will be elucidated.

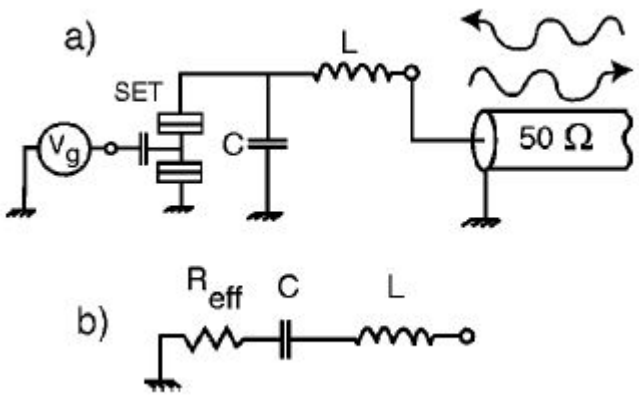

Figure 2.4: RF Reflectometry, The figure shows how SET resistance is matched to 50ohm line resistance using L match network (reproduced from [21])

2.2. RF Reflectometry

Reflectometry is one of the most widely used readout method for single-shot readout for spin qubits. As described in the previous subsection, the information on spin is encoded in the value of resistance. SET or QPC resistance is matched to 50 ohms using the L match network. As discussed in the previous section, electron tunnels through the dot momentarily when the spin state of the electron is down. This modulates the resistance momentarily which results in a change in reflection coefficient and hence changes the magnitude of the reflected wave. In this thesis, the Low Noise Amplifier will be designed with the purpose to detect small changes in the reflected wave. As shown in figure 2.4, the parasitic capacitance C_p and the inductance L are tuned to get a 50-ohm match. The parasitic capacitance is of order of 100 to 300fF. The power that one can drive SET is limited to -99dbm to prevent any incident signal from capacitively coupling to qubits. This also sets the maximum voltage V_{max} that can be driven across the SET. At resonance, the matching circuit acts as an impedance transformer. The incident voltage wave is boosted by a factor Q at the SET. Hence if V_{max} is the voltage incident on the SET then voltage wave incident on the matching network is given by V_{max}/Q. Where Q or quality factor of the transformation and is given by

$$Q = \sqrt{\frac{R_{SET}}{R_s} - 1} \tag{2.1}$$

where R_{SET} is the SET resistance and $R_s = 50$ is the line resistance. For eg taking a typical case of Cp=0.15p, matching frequency can be calculated using $\omega_o = Q/CpR_{SET}$. This gives $f_o = 474MHz$. The inductance can be calculated as $L = CpR_SR_{SET} = 750nH$. These large value of inductors are usually implemented using superconducting materials. The value of the inductor is what limits the scalability of this approach and is one of the drawback of using this method. For the matching QPC having 25K ohm inductance required is 190nH assuming same parasitics. $V_{max}^2/2R_{SET} = P_{max}$ one can get the maximum voltage that is incident of SET as 158uV. The reflection coefficient is given by $(Z_L - Z_o)/(Z_L + Z_o)$ i.e $\Delta R_S/2R_s$. The approximate input impedance looking into the matching network can be derived using the relation $P_{SET} = P_{in}$ which gives

$$Z_{in} = \frac{R_{SET}}{Q^2} = \frac{1}{\omega_o^2 Cp^2 R_{SET}} \tag{2.2}$$

Using 2-1 one gets

$$\frac{\Delta R_S}{R_S} = \frac{\Delta R_{SET}}{R_{SET}} \tag{2.3}$$

If V_{max} is the maximum voltage incident on SET then the reflected incident at the input of the LNA is given by

$$V_{reflected} = \frac{V_{max}\Gamma}{Q} = \frac{V_{max}\Delta R_S}{2QR_S} \tag{2.4}$$

Since the relative change in 10 percent, this results in $V_{reflected} = 177nV$. SET shot noise current is given by [23]

$$\frac{2eV_j}{R_j} coth(\frac{eVj}{2KT}) = 3 * 10^{-28} A^2/Hz \tag{2.5}$$

Here V_j is the junction voltage across the two set junctions. Current shot noise converted into voltage noise gives $1.73nV/\sqrt{Hz}$ at the SET end. Equivalent noise at input of LNA is $19pV/\sqrt{Hz}$. The measurement time determines the noise bandwidth. Hence only taking SET noise contribution, the signal reaching input of the LNA has SNR =6.4dB considering 10MHz noise bandwidth and SNR=16.37dB in 1MHz bandwidth.

2.3. FDMA or Multiplexed Readout

Frequency division multiplexing is used for simultaneous readout of multiple qubits [24]. Figure 2.5 shows FDMA readout of multiple qubits in parallel. The inductances are tuned to the different resonant frequencies. If T is the measurement time then fundamentally each of the qubits can be spaced $1T$ frequency apart. However, the spacing between the center frequency of the matching networks is determined by the interaction between the matching networks. In other words, ideally the matching network should present infinite impedance outside its resonant frequency but there is always some stray inductive/capacitive impedance which is added in parallel outside the resonant bandwidth of the matching network. The matching network also filters noise of each of SET's or QPC's outside the matching bandwidth that is $\frac{\omega_O}{Q}$. Since there is finite attenuation outside the bandwidth of the matching network spacing should be carefully chosen such the noise interference from other SET's or QPC's is limited.

The matching network required has a very high Q since very high impedance value (100K) is transformed to 50 ohms. In case of QPC 25K is transformed to 50ohms. Inductance value at the readout frequency ($L = C_pR_s^2Q^2$) required for impedance matching is few hundreds of nH. Here C_p is parasitic capacitance and $R_s = 50ohms$. This limits the scalability of this technique. Some other drawbacks include high Q nature of the matching network which is sensitive to mismatch and the interaction between the matching networks. However, this is the only technique which can be used for multiplexed readout of qubits in present setups. Multiplexing readout of qubits is necessary since in the present setups readout electronics are not at the same base temperature as the qubits. Hence, if multiplexing is not done each of the qubits will require separate lines.

2.4. Baseband SET readout

Figure 2.6 shows the frontend amplifier for baseband readout. The drawback of this approach is parasitic pole caused due to C_L or transmission line capacitance at lower frequencies. This limits the bandwidth and

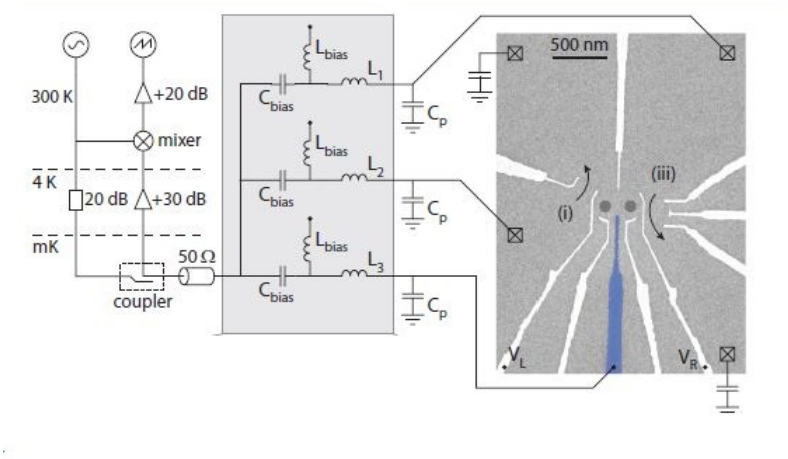

Figure 2.5: FDMA readout of multiple Qubits (reproduced from [24])

noise performance of the amplifier at higher frequencies. Also to filter high-frequency noise of the amplifier coupling into the SET large capacitance is added at the input [25]. Another limitation of this approach is the low frequency $\frac{1}{f}$ noise and charge noise which affects the performance of the readout [25].
However, as far as the effective input signal is concerned the transimpedance readout has much larger effec-

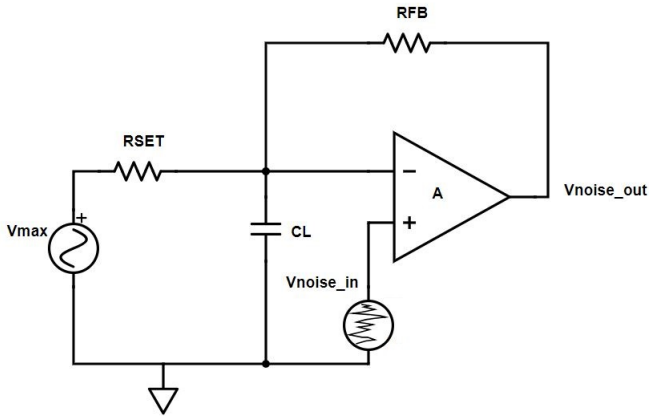

Figure 2.6: Baseband readout showing how resistance change in QPC or SET can be converted to current which can be then readout using transimpedance amplifier [25]

tive signal as compared to RF readout. This is because of high Q impedance transformation of the matching network which is required for RF readout. The input signal to LNA in case of reflectometry setup as given by

$$V_{in} = \frac{V_{max}\Delta R_{SET}}{2QR_{SET}} \tag{2.6}$$

This is the reflected wave due to the impedance mismatch which is described in the previous section. The output signal is given AV_{in} where A is the LNA gain. Similarly the rms output noise spectral density for broadband LNA's like resistive feedback is given by

$$Vnoise_{out} = \frac{Vnoise_{in}A}{1 + A\beta} \tag{2.7}$$

Here $Vnoise_{in}$ and $Vnoise_{out}$ is the input and output referred noise of the amplifier. As will be seen in the chapter 4 the loop gain for perfect matching condition is 1 in case of feedback amplifier like resisitive feedback and hence this expression reduces to

$$Vnoise_{out} = \frac{Vnoise_{in}A}{2} \tag{2.8}$$

This expression for output noise is the same for any feedback and noise cancelling LNA's. Similarly in transimpedance readout effective output voltage is given by incremental current change times the transimpedance gain.

$$V_{in} = \frac{V_{max}\Delta R_{SET} R_F}{R_{SET}^2} \tag{2.9}$$

Input referred noise due to transistors is again given by equation 2.7 which for large gain A needed for low input resistance turns out to be

$$\frac{Vnoise_{in} R_F}{R_{SET}} \tag{2.10}$$

where R_F is the feedback resistor which determines transimpedance gain. Using the above equations the following conclusions can be drawn

$$SNR_{reflectometry} = \frac{V_{max}\Delta R_{SET}}{QR_{SET} Vn_{in}} \tag{2.11}$$

$$SNR_{transimpedance} = \frac{V_{max}\Delta R_{SET}}{R_{SET} Vn_{in}} \tag{2.12}$$

The input impedance requirement for such a transimpedance amplifier can be of the order of few kiloohms rather than 50 ohms since no matching is required with the input line. As a typical example, a transimpedance amplifier having the gain of 20dB requires a feedback resistance R_F of 1Mohms. A single pole amplifier can easily have a gain of 40dB giving input impedance of $10Kohms$ which is sufficient considering high impedance of the SET. Hence from these two expressions, one can see that the effective signal in transimpedance type amplification is Q times greater than that in reflectometry type readout. Integrated input referred noise $Vnoise_{in}$ in case of baseband readout is much greater due to the presence of low-frequency noise both of the amplifier and SET which can in principle be subdued by design techniques. However, baseband readout of multiple qubits requires multiple lines coming from 20mK stage to 4K stage for each of the qubit. FDMA cannot be done here due to bandwidth limitations posed by line capacitance. Also, frequency multiplexing line as shown in figure 2.7 will add noise of each of the SET's on top of each other. If switches could be integrated close to qubits one option is to do TDMA like in figure 2.8. Comparing with FDMA with N SET's multiplexed, each of the SET's in TDMA have to be read N times faster. With the line going from 20mK to 4K this is not possible because ultimately bandwidth of the readout is determined by the parasitic capacitance as well as the capacitance added to filter high-frequency noise coupling to the SET. If qubits and switches can be integrated close to control electronics this could indeed be a good proposition. Because of the lesser noise requirement, one can operate readout cycle at higher rate provided low-frequency noise is subdued by suitable design techniques.

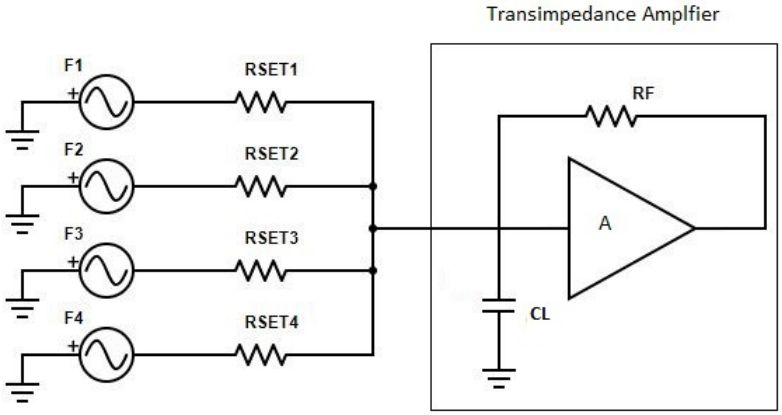

Figure 2.7: Transimpedance readout using FDMA

In this thesis, Low noise amplifier is designed for RF reflectometry as it enables multiplexing of readout of qubits in present setup used by the physicist. Specifications are derived based on RF-SET. The LNA can also be used for dispersive readout of spin qubits. The application space for cryogenic LNA's also include Radio astronomy, Deep space communications and many other applications in instrumentation physics.

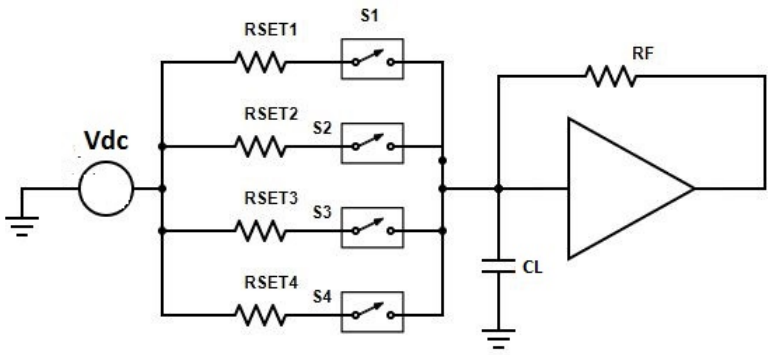

Figure 2.8: Transimpedance readout using TDMA

2.5. Specifications

Using equation 2.4 one can arrive at a conclusion that the input signal at the LNA is around 177nV. The system can be thought of as OOK(On off shift keying) where the pulse train is modulated by a carrier signal. The momentary absence of the electron from the dot modulates the reflected wave. The fault tolerance error threshold needed for one error correction cycle is approximately 0.1% which corresponds to bit error rate of 1e-3 [26]. Using the error function $Q(\frac{V_{pp}}{2\sigma})$ one can estimate the input referred rms noise required for this achieving this BER. Using $V_{pp} = 177nV$ one gets rms noise as $\sigma = 29.5nV$. If the readout bandwidth is BW then input noise power spectral density requirement for the LNA is

$$Vn_{in}^2 = \frac{\sigma^2}{BW} \tag{2.13}$$

The noise requirement of the LNA is

$$Vn_{in}^2 = \frac{\sigma^2}{BW} - Vn_{SET}^2 \tag{2.14}$$

$$T_n = \frac{\frac{\sigma^2}{BW} - Vn_{SET}^2}{KR_S} \tag{2.15}$$

Where T_n is the noise temperature of the LNA, K is the Boltzmann constant and R_S is 50 ohms standard resistance. Vn_{SET}^2 is the noise of the SET which is referred to the input of the LNA. This is discussed in section 2.2 and is equal to $19pV/\sqrt{Hz}$. From Equation 2.6 it is evident that the noise temperature specifications become more stringent with higher readout bandwidth. Following table elucidates this

Readout Bandwidth (MHz)	LNA noise ($V^2/\sqrt{Hz}$)	LNA (Noise Temperature)
1 MHz	5.1e-22	0.7
0.8 MHz	7.27e-22	1
0.5 MHz	1.38e-21	1.97
0.2 MHz	4e-21	5

In the next section, the limited scaling of transistor noise due to shot noise will be discussed. It is expected that the mobility will increase by a factor of 2 based on the previous measurements. For approximately constant transconductance i.e $gm = \sqrt{2\beta I}$, the current is expected to decrease by a factor of 2 and hence the shot noise component in the current i.e $F2qI$. Here F is called the fano factor. The discussion on the channel noise of the transistors will be done in the next chapter. Hence only a factor 2 scaling in drain channel noise of the transistors will be assumed in deriving the specifications. Assuming the spacing between the qubit channels as 25MHz approximately 800MHz of bandwidth is required to multiplex 32 qubits. Keeping some margin LNA will designed for close to 1GHz RF bandwidth. The main challenge in the LNA design is to achieve ultra-low noise level while delivering broadband noise performance($\geq 800MHz$) so that as many channels can be multiplexed. Another challenge in the design will be to keep gain as high as possible. Since the signal level is extremely low the gain of the LNA has to be sufficiently high to reduce contributions of next stages. All this by

keeping the power consumption as low as possible. The LNA will be designed to reach 10K transistor noise (drain/shot) contribution at room temperature. Hence LNA is expected to reach a noise temperature of $\leq 5K$ at 4K. The linearity is not a concern in the design as the input levels are extremely low. At room temperature, the LNA will be designed for 10K noise contribution of the transistor drain noise. However, the final LNA noise temperature at 300K will also have resistive channel noise contribution. The specifications are kept as 25K at room temperature which is close to what is required for many room temperature applications like radio astronomy.

Base Temperature	Broadband Noise temperature	Bandwidth	Gain	S11
$300K$	$\leq 25K$	$\geq 1GHz$	$\geq 40db$	$\leq -10db$
$4K$	$\leq 5K$	$\geq 1GHz$	$\geq 40db$	$\leq -10db$

3

A Survey on Low Noise Amplifiers

3.1. Introduction

As discussed in the previous chapter, the LNA needs to be designed achieving less than 5K noise temperature at 4K temperature. The discussion in this chapter concerns the selection of suitable topology to be designed for RF readout of spin-qubits. This chapter will go through all the common topologies in the literature of low noise amplifiers. The next section will discuss the limited scaling of transistor drain noise as one goes from room temperature to cryogenic temperature. This is to emphasize the point that the drain current shot noise will be the primary source of noise at cryogenic base temperature. Only the drain current noise of the transistor is taken into account in the analysis. Noise due to the parasitic resistance and intrinsic gate resistance is expected to be negligible at 4K.

3.2. Theory of shot noise and Cryogenic performance of CMOS LNA's

In long channel devices the thermal drain noise current in the channel is given by

$$idn = 4KT\gamma gd_0 \tag{3.1}$$

where gd_0 is the gate to drain transcondutance at zero bias and the factor γ is called the excess noise factor. In short channel devices, there is a significant increase in the γ as pointed out by many experimental pieces of evidence [27, 28]. The prevalent theory which explains this is the presence of shot noise. Many prior works in the literature indicate shot noise signatures in short channel devices [27, 28]. The presence of shot noise point towards limited scaling of transistor noise from 300K to 4K. The excess noise can be understood intuitively by understanding the underlying mechanisms of generation of shot and thermal noise. Figure 3.1 shows the potential energy diagram for carrier transport in the MOSFET. There are basically three regions of electron transport. At the end 1, carriers are transported from the region 1 to 2 through diffusion. The electrons that have high enough energy cross the barrier to region 2. In the low electric field region 2, transport of electron takes place through drift. After region 2, electrons enter region 3 which is a high electric field region where the electron velocity saturates. Carrier transport in the region 1 or the source is governed by diffusion of carriers at the source end which is Poisson due to the discrete nature of the charges. After electron enters the region 2 it will undergo scattering if the channel length is longer as compared to mean free path l_c of electrons. In region 2 electron transport is determined by the drift and since electron undergoes scattering it exchanges energy with surrounding and comes under thermal equilibrium. Hence carriers in case of long channel devices behave like that of the resistor. In region 3, electrons do not interact as they are velocity saturated and statistics of carrier interval is mostly determined by region 2. Hence in long channel noise is mostly thermal in its origin. While in short channel devices carrier transport approaches ballistic transport where most of the electrons do not undergo scattering as the channel length is less than mean free path. The statistics of current is determined by statistics of carrier entry in region 1. Hence the current noise approaches the shot noise limit. Depending upon the channel length the transistor shows full or suppressed shot noise. The suppression of shot noise can be thought of as a feedback mechanism. If more or fewer carriers are in the channel than its equilibrium value, it changes the potential profile such that it suppresses randomness. Hence noise in the shot channel MOSFETs can be approximately modeled as

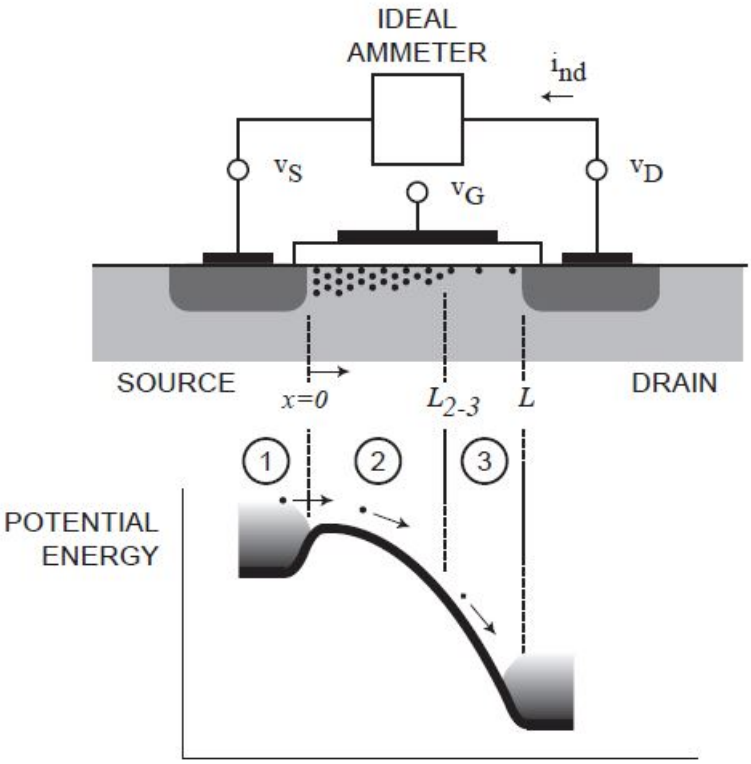

Figure 3.1: Mosfet Energy Diagram showing carrier transport of electrons in mosfets (reproduced from [29])

$$I^2_{short-channel} = F2qI \tag{3.2}$$

Where F is called the Fano factor which takes into account the short channel suppression. The BSIM model still models the noise as thermal noise which has been questioned in several papers in the literature due to the limited scaling of noise both in HFET(High electron mobility FET) and CMOS [30–32].Pospieszalski model which is used in the majority by Cryogenic HFET and SiGe amplifier design community models the two fundamental noise sources FET's as drain T_d and gate temperature T_G [30–32]. T_d represent drain current noise and T_G represents the noise due to intrinsic gate resistance which is responsible for gate induced noise. The experimental evidence in many publications suggest limited scaling of T_d which supports shot noise theory and scaling of T_G with ambient temperature hence meaning that gate induced noise or noise due to intrinsic gate resistance scales with temperature. Hence gate induced noise is expected to be negligible due to scaling of T_G.

3.3. CMOS LNA's

Figure 3.2 shows commonly used narrow-band Low noise amplifier using source degeneration [33]. The input impedance of this LNA is given by

$$Z_{in} = s(L_s + L_g) + \frac{1}{sC_{gs}} + \frac{gmL_s}{C_{gs}} \tag{3.3}$$

At resonance frequency, the inductive and capacitive part can be made to cancel each other leaving only the real part of impedance which can be controlled to be 50 Ω. Due to the bandpass nature of the input network, the matching is narrow band. At resonance the effective gate to source voltage is boosted by a factor of $Q = 1/(2CgsR_s\omega_o)$ where ω_o is frequency of impedance match. This gives the effective transconductance as

$$G_m = gmQ = \frac{\omega_t}{2R_s\omega_o} \tag{3.4}$$

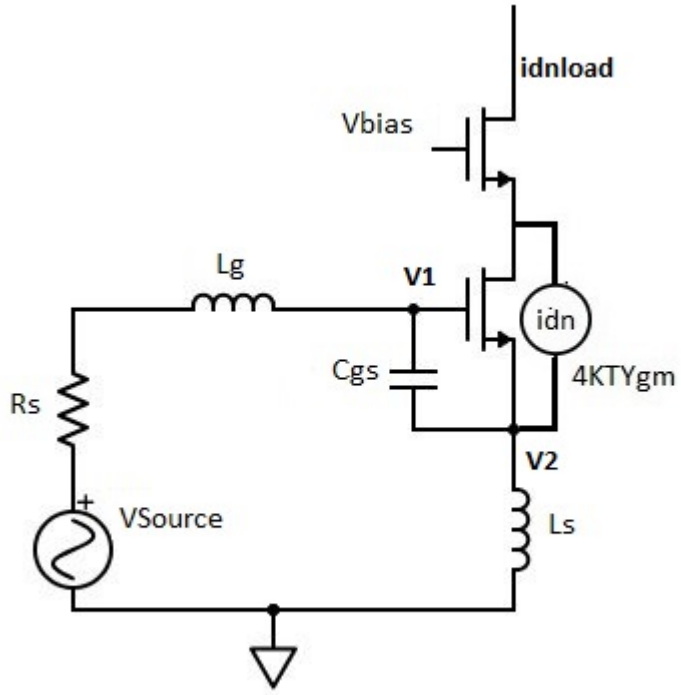

Figure 3.2: Commonly used narrow band LNA using source degeneration [33]

This is only dependent of ω_t(unity gain frequency) of the transistor and is independent of number of fingers. As one decreases the number of fingers the gm of the transistors decreases but the Q of the input match increases leading to more passive amplification which compensates for decreased transconductance. Effectively making transconductance or G_m constant. As far as the output drain noise is concerned some amount of drain current noise current flow through the C_{gs}. Following equations describe the dynamics of the noise current in the source degenerated LNA.

$$V_1 = V_2 \frac{R_S + sL_g}{sL_g + 1/sC_{gs} + R_S} \tag{3.5}$$

$$gm(V_1 - V_2) + idn = \frac{V_2}{sL_S} + (V_2 - V_1)sC_{gs} = idn_{load} \tag{3.6}$$

Using 3.5 and 3.6 the load drain noise current is given by

$$id_{load} = \frac{i_d}{(1 + ((gmL_s)/(C_{gs}R_s))} \tag{3.7}$$

which for matching condition $R_s = \omega_t L_s$ is given by $i_d/2$. One can concur from this that making number of fingers or width of transistor smaller and smaller can make noise figure to approach 0db. This is because decreasing number of fingers leads to decrease in output noise keeping the transconductance constant. However, at the cost of very narrow bandwidth due to high Q input match. However this implication is incorrect since there is another source of noise which is related to thermal fluctuations at the gate called the gate induced noise with power spectral density $4KT(\omega^2 C_{gs}^2)/(5gd0)$. The factor $1/5gdo$ represents the gate resistance. Figure 3.3 shows the origin of the gate induced noise . As one decreases the transistor width for same matching frequency, the $Q = 1/\omega_o C_{gs} R_s$ of the matching increases. Higher gate inductance is now required for matching due to which effective impedance seen by induced gate current noise current increases. This aggravates the contribution of gate noise at the output. This is the fundamental reason of existence of F_{min} of the transistor. At F_{min} both drain and gate noise have equal contribution to overall noise figure. Since this architecture is narrow-band as one goes away from center frequency noise figure degrades very sharply depending on the quality factor of the input match and hence it is not suitable for our application. For our

readout we need broadband LNA rather than narrow band LNA to enable frequency multiplexing using a single cable coming from 20mK stage to 4K stage. Narrowband source degenerated LNA could deliver ultra-low noise levels if the input matching is made high Q. This is possible since the noise due to the gate resistance which determines the Fmin level goes down. However, then it means single LNA per channel hence requiring multiple lines to readout multiple qubits.

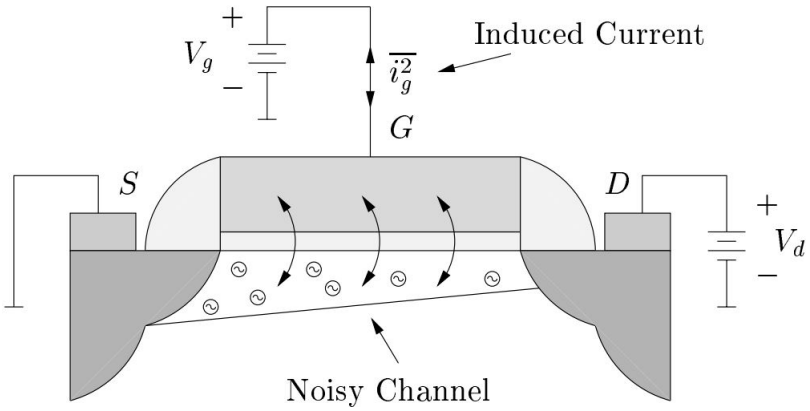

Figure 3.3: Gate induced noise caused by thermal fluctuations in the channel getting capacitively coupled to gate (reproduced from [33])

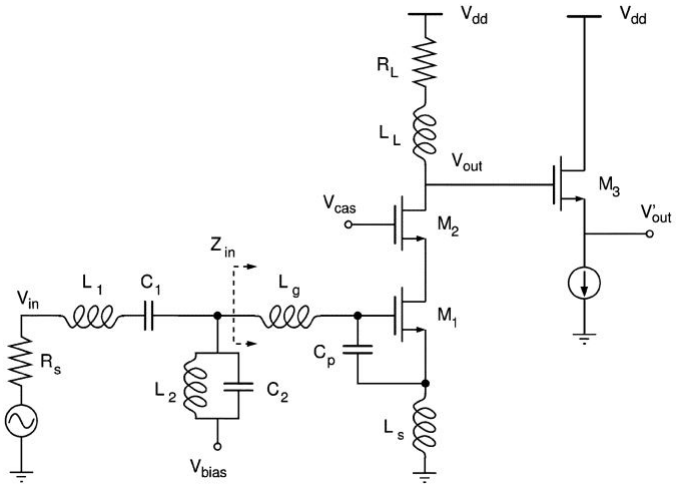

Figure 3.4: Broadband input match using LC ladder matching network at the input (reproduced from [34])

To make the input match broadband numerous architectures have been proposed. One such is shown in Figure 3.4. All these structures [34, 35] use matching networks like Chebyshev or other LC ladder structures for a broadband match. However, these amplifiers architectures have the drawback that they use the lossy matching network at the input and second, the effective signal between gate and source decreases with frequency. One can see from figure 3.4 inband impedance seen by the signal source is $2R_s$ when LNA is matched to $Rs = 50ohms$. Thus the current through C_p(Figure 3.4) under matched conditions will be

$$I_{in} = V_{in}/2R_s \tag{3.8}$$

which corresponds to gate to source voltage across the transistors as

$$V_{gs} = V_{in}/(2R_s\omega C_p) \tag{3.9}$$

As can be seen from the previous equation gate to source voltage decreases with frequency and hence the noise figure. Another architecture[36, 37] which achieved record noise figure in CMOS LNA's is shown in the

figure 3.5. For 50 ohm matching this LNA achieves record noise figure of 0.35 dB achieving broadband noise performance upto 700MHz. The input matching network here can be seen as parallel combination of two resonant networks as can be seen in the figure 3.6. One RLC branch is due to source degeneration and other is due to parasitic feedback effect due to C_{gd} and load capacitance C_L. First one resonates at the higher band edge and the second one resonates at lower band edge to achieve broadband match. However in this LNA, matching is dependent on a delicate relationship between the parasitic capacitances, gate inductance, and source inductance. Since the parasitics are not well modeled at 4K, this architecture was not chosen.

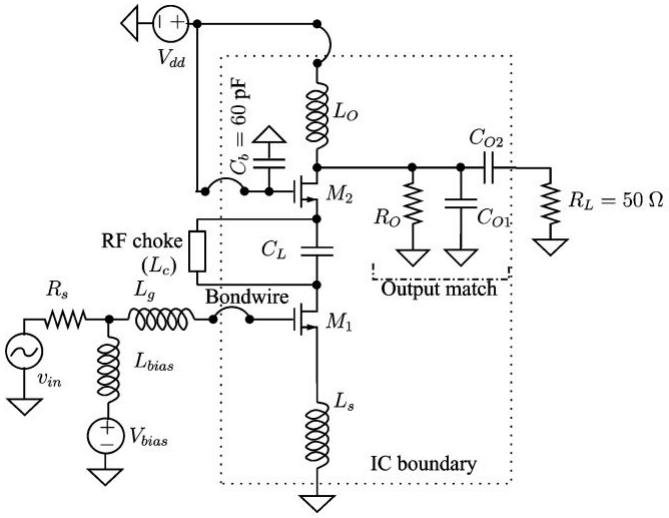

Figure 3.5: Broadband LNA utilizing conventional source degeneration and feedback of gate to drain capacitance (reproduced from [36])

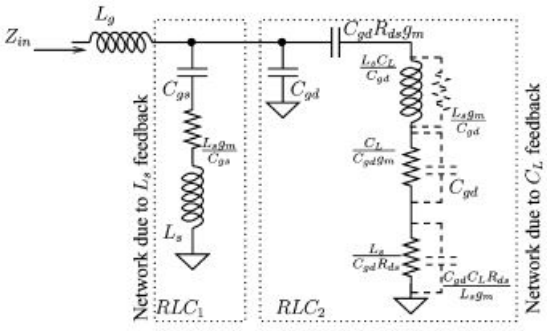

Figure 3.6: Input impedance of the LNA which is described in the figure 3.5 (reproduced from [36])

Noise cancelling[38–40] and Resistive feedback[41–43] are commonly used broadband LNA architectures used in wireless communication. Matching is simpler to implement in these circuits however they are not used in low noise figure applications at room temperature like in radio astronomy. The reason for this is the resistive contribution of noise in both of the circuits. Figure 3.7 shows the Noise-Canceling architecture. Here Common gate transistor is used for 50-ohm line matching. The description here is general of all noise cancelling LNA's whether differential or single ended. A common gate amplifier has a noise factor of

$$NF = 1 + \frac{\gamma}{gmR_S} \tag{3.10}$$

The condition for an input match of $R_s = 50ohms$

$$Z_{in} = R_s = \frac{1}{gm} \tag{3.11}$$

The common gate amplifier has noise factor greater than 2 corresponding to noise figure greater than 3dB. However, 50-ohm matching is simple to achieve. The noise canceling is a clever technique which utilizes

simplicity of matching of common gate transistor and low noise performance of common source transistor. As can be seen in the figure 3.7, some of the drain noise current flows through the source resistor. If V_n is the input referred noise of the common gate transistor then the noise current due to common gate transistor at the positive output is given by

$$Vout_{n+} = -gm_{CG}R_{CG}\frac{V_n}{2} \tag{3.12}$$

The noise due to common gate transistor at the negative differential output

$$Vout_{n-} = -gm_{CS}R_{CS}\frac{V_n}{2} \tag{3.13}$$

. It can be seen from the equations 3.12 and 3.13 that the noise of the common gate transistor is correlated at both the inputs hence will be canceled when the differential output is taken. The noise canceling condition is given by

$$gm_{CG}R_{CG} = gm_{CS}R_{CS} \tag{3.14}$$

and input matching condition is given by $Z_{in} = 50 = 1/gm_{CG}$. So there are two paths for amplification one is noisy common source path which determines the noise figure and one is noiseless common gate path since it's noise is canceled. As will be discussed in the next chapter the resistive contribution of the noise is dominated by R_{CG}.

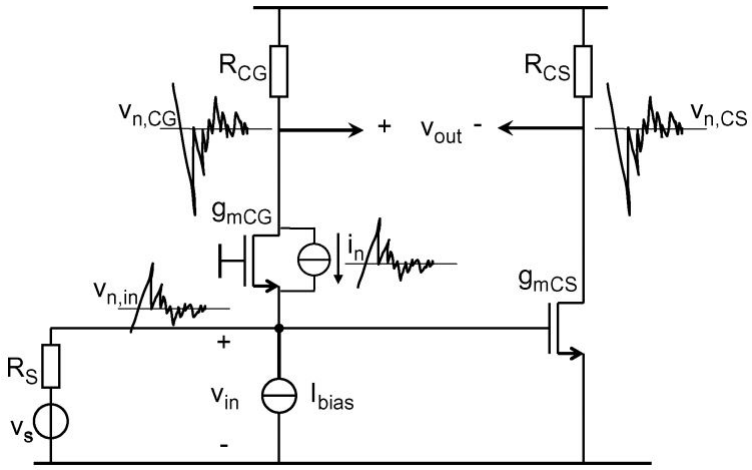

Figure 3.7: A broadband Noise-Canceling balun LNA (repoduced from [39])

Distributed amplifiers shown in Figure 3.8 are another class of broadband amplifier [45, 46]. The input capacitance of the amplifier is absorbed in the transmission line hence it can overcome tradeoff between NF and bandwidth. In general, Distributed-Amplifiers have very high noise figures and have high power consumption. The main reason is not only the loss due to transmission line but also the higher contribution of transistor noise. The noise of matching resistor R_g does not effect LNA noise figure due to reverse isolation. Intuitively what happens is that noise waves injected at port 3 interfere destructively at port 2. Hence the circuit can be thought of an open loop amplifier where matching is done by a 50 ohm noiseless resistor at the input. For such an amplifier the input referred noise is $\frac{4KT\gamma}{gm}$ which is four times higher that other broadband amplifiers like Noise-Cancelling and Resistive-Feedback i.e $\frac{KT\gamma}{gm}$ .The input referred noise for these two topologies will be derived in the next chapter. As discussed in the previous section due to the limited scaling of transistor noise keeping transistor noise minimum is the priority. Figure 3.9 shows the Resistive-Feedback amplifier. The input is matching is achieved by tuning the gain of the amplifier and the feedback resistor. The input matching condition is given by

$$Z_{in} = \frac{R_F}{gmR_L} \tag{3.15}$$

The common source is used as the main transconductor to reduce the noise. The other main contributor to noise is the feedback resistor. Both of the architectures have significant resistive noise contribution making

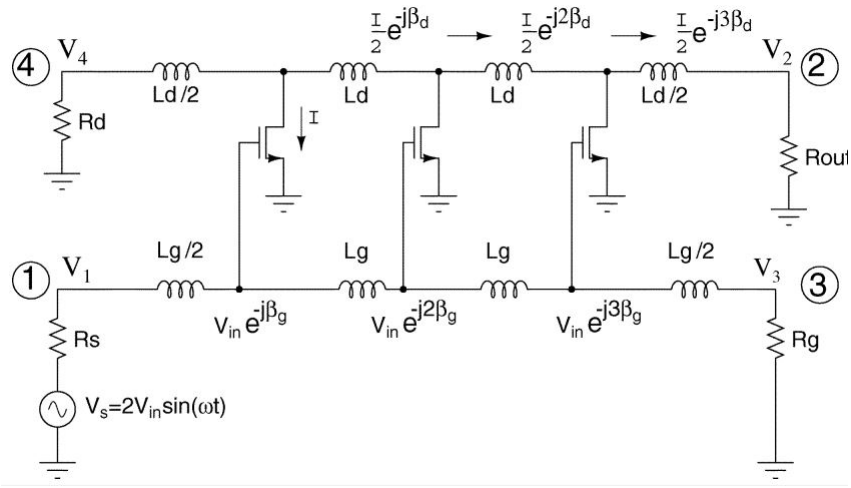

Figure 3.8: A CMOS distributed amplifier used to overcome noise bandwidth trade-off (reproduced from [45])

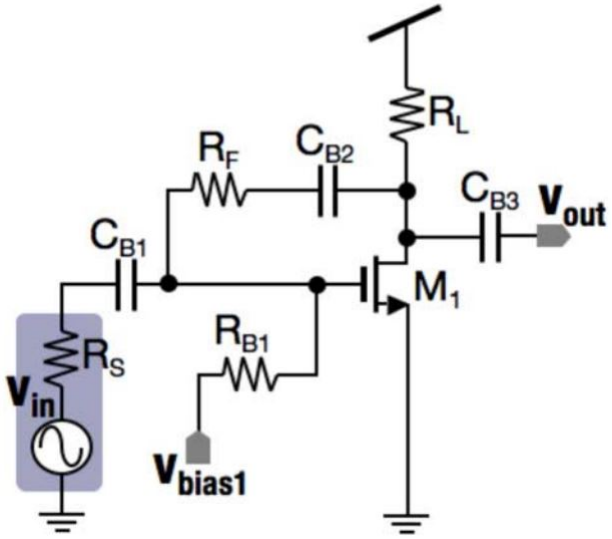

Figure 3.9: A Simplified broadband Resistive-Feedback LNA(reproduced from [42])

them not a desired candidate for ultra-low noise room temperature amplification. However, for cryogenic applications, they would be the ideal candidate because of the simplicity of matching and scaling of thermal noise which makes the resistive contribution of noise negligible. Figure 3.10 shows capacitive feedback amplifier. The capacitive feedback amplifier can achieve very good power efficiency for certain noise figure at room temperature since there is no resistive contribution to noise at room temperature. The input impedance of the capacitive feedback amplifier is given by parallel combination of the real impedance

$$Yin_{real} = gm(1 + C2/C1)) \tag{3.16}$$

and the imaginary capacitive component

$$Yin_{img} = C_1C_2/s(C_1 + C_2) \tag{3.17}$$

At lower frequencies the capacitive part is high and the input impedance is mainly determined by the real part. This is good assumption if the parallel capacitance due to the input match is low. For that both C1 and C2 has to be low. However, to achieve low noise levels, the transconductance has to be high resulting in low output impedance. To make the transconductor current to flow through load capacitor C_2 very high capacitance value is required especially at lower frequencies. This will not only add parasitic capacitance at the output but directly contradicts our previous requirement lowering the capacitive part of input impedance

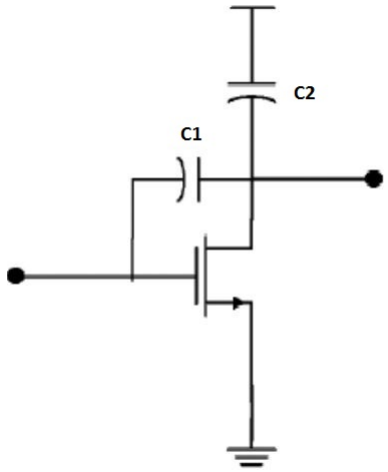

Figure 3.10: A Capacitive-Feedback low noise amplifier (reproduced from [44])

$C_1//C_2$ which will add in parallel to C_{gs}. Hence Resistive-Feedback and Noise-Canceling emerge as suitable candidates due to the simplicity of matching. The noise analysis of both of these amplifiers will be discussed in the next chapter.

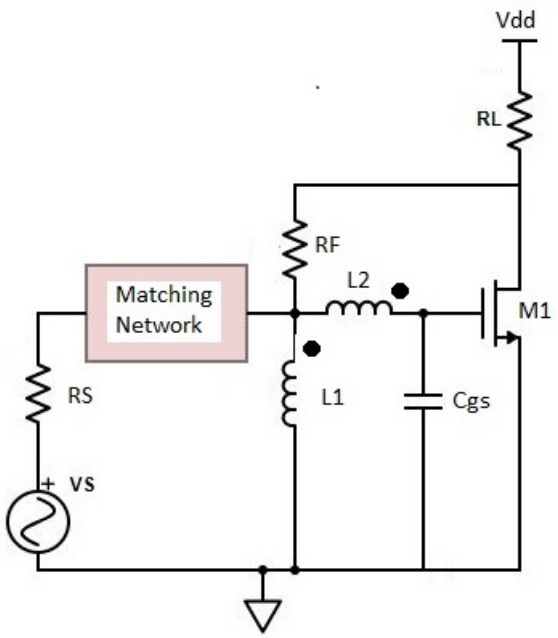

Figure 3.11: Resistive feedback amplifier with passive amplification using transformers and input matching network for broadband performance

To overcome the noise figure and power trade-off one can either use passive amplification using transformers [47] or parametric amplification [48, 49] which require lossy passives at the input. The loss is especially high at lower frequencies since these passives are bulky at lower frequencies. Hence they can be more suitable for superconducting qubit readout which occurs at higher frequencies. In architectures like [47] going away from center matching frequency, the noise figure degrades rapidly. This happens because of two reasons first signal transfer decreases and second output noise increases due to input parasitic pole. This will become clearer in the next chapter. In order to mitigate this an input bandpass matching network is required. Figure 3.11 shows how the network can be designed using the Resistive-Feedback amplifier. Since the loss is not well modeled at cryogenic temperatures for this tapeout passive amplification using matching networks is not chosen but can be considered as future work. Hence Noise Cancelling and Resistive-Feedback with-

out passive amplification seem to be viable options for low-frequency low-noise amplification and will be compared in the next chapter.

In the literature, CMOS is not used generally by the cryogenic community for low noise amplification. SiGe(Silicon Germanium) amplifiers due to their high β or the current gain and HEMT's(High Electron Mobility Transistors) due to their high electron mobility achieve low noise operation at much lower power. In other words, they have low input referred noise due to high transconductance value. The following table shows the state of the art which are all SiGe amplifiers. As can be seen, the state of the art Si-Ge LNA's achieves low noise temperature levels at very low power. There is no available art in CMOS LNA's that even come close to Si-Ge in their noise performance at cryogenic temperatures. However, one can see from the table below that at room temperatures CMOS LNA's dominate the state of the art. The objective of this thesis is also to check the weather noise temperatures close to that of Si-Ge can be achieved by ordinary CMOS at cryogenic temperatures. The objective of this thesis is also to study the low noise levels that CMOS can achieve at cryogenic temperatures. However, power consumption in CMOS LNA's to achieve low noise levels will be fundamentally higher due to transistor physics. However, CMOS due to maturity in the process can really provide cost-effectiveness as compared to Si-Ge.

Publication	Physical temperature	Noise Temperature	Bandwidth	Technology	Power
IMS [50]	16	2.8K	0.3-3GHz	SiGe	32mW
IMS [51]	16	3.4-4K	2-4GHz	SiGe	3mW
IMS [52]	18	5-8K	4-8GHz	Si-Ge	760uW
IMS [53]	15	3.7-4.3	0.1-5GHz	Si-Ge	20mW
IEEE RWS [36]	290	$\leq 25K$	700MHz-1.4GHz	CMOS	45mW
JSSC [51]	290	$\leq 14K$	700MHz-1.4GHz	CMOS	45mW
JSSC [54]	270	$4K$-$10K$	2GHz-8GHz	HEMT	1W

4

LNA Design and Architecture

4.1. Introduction

In this chapter both the Noise-Canceling and Resistive-Feedback are compared for their expected broadband noise performance at 4K base temperature. The motivation for selecting the adapted Resistive-Feedback LNA is elucidated. The complete design of each of the circuit blocks including the bias network is explained. It is followed by pre and post layout simulations to verify the performance of the LNA.

4.2. Noise-Canceling vs Resistive-Feedback

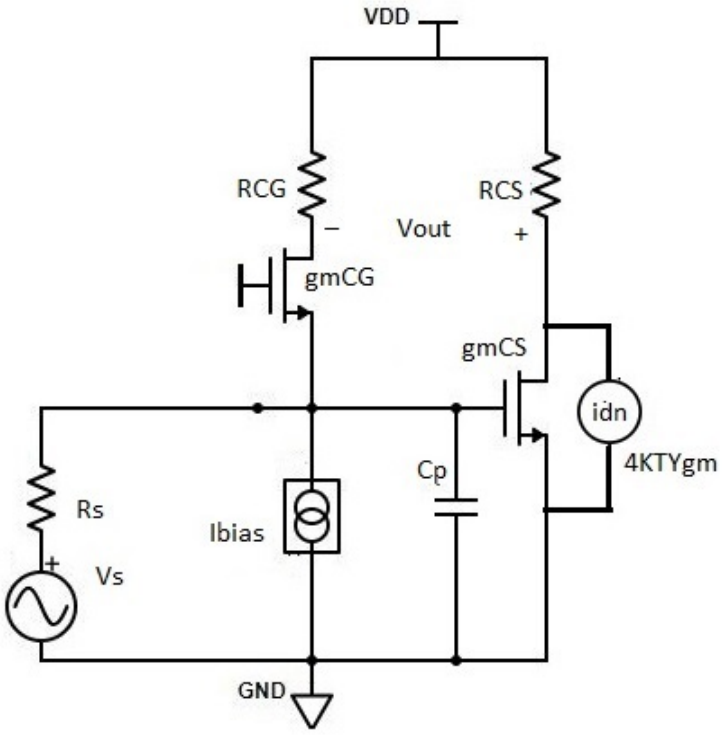

Figure 4.1: Noise Cancelling balun LNA [39]

As discussed in the previous chapter, experimental evidence indicates that the noise contribution of the transistors does not scale with temperature. Literature points out towards the presence of shot noise at cryogenic temperatures. However, still, it is an open question whether the limited scaling is due to shot noise or due to self-heating. However, noise contribution due to resistors which is thermal in origin is expected to

scale with temperature. In this chapter to gain intuition, the calculations are made using the thermal noise assumption. The same can be proved taking the shot noise assumption. Noise-Canceling, Resistive-Feedback, and Capacitive-Feedback are the simplest architectures giving broadband matching. Capacitive-Feedback will give the best noise performance at room temperature, however, each of these is expected to give similar performance at cryogenic temperature. This is because the resistive noise contribution in both the Noise-Cancelling and Resistive-Feedback is expected to scale with temperature. In other words, the resistive noise is expected to be way less than the drain current noise in the transistor channel. Input matching condition in the capacitive feedback is very difficult to meet when the limited output impedance of the transistors is taken into account as seen in the previous chapter. Figure 4.1 shows the Noise-Canceling amplifier. As described in the previous section, the noise of the common gate transistor which is used for 50-Ω matching is canceled using a common source transistor. The input impedance of the Noise-canceling amplifier is given by $Z_{in} = 1/(gm_{CG}) = Rs$ and the noise cancelling condition is given by $gm_{CG}R_{CG} = gm_{CS}R_{CS}$. Under the noise canceling condition correlated noise of the common gate transistor which is present at two differential outputs is canceled. The amplifier gain is given by $A_v = 2gm_{CG}R_{CG} = 2gm_{CS}R_{CS}$ The total output noise in matched condition is approximately given by

$$Vn^2_{out} \approx 4KTR_{CG} + 4KT\gamma gm_{CS}R^2_{CS} \tag{4.1}$$

The noise of the common source resistor is neglected. For achieving low noise levels, the common source transistor must have high transconductance or $gm_{CG} << gm_{CS}$. And hence to meet the noise canceling condition $R_{CG} >> R_{CS}$ holds. From above, the low-frequency input referred noise can be derived and related to gain of the amplifier as

$$Vn^2_{in} = \frac{2kTR_s}{A_v} + \frac{KT\gamma}{gm_{CS}} \tag{4.2}$$

Here R_s is the source resistance that is 50 ohm and A_v is the gain of the amplifier. The first term in the above expression comes from the resistive contribution of the noise which is dominated by the load resistor of the common gate amplifier and is expected to scale with temperature. The only noise source that is expected to be dominant at cryogenic temperatures is that of the common source stage. However, at higher frequencies the parasitic capacitance C_p which is mostly due to C_{gs} and C_{gd} of the common source transistor impact the noise canceling condition. Intuitively it can be seen that any mismatch in the noise-canceling condition will result in a significant contribution of the noise of the common gate transistor at the output. This is especially true if we are trying to achieve ultra low noise. The uncancelled noise of the common gate will be significant and will dominates the total noise. The high-frequency input referred noise contribution due to common gate transistor can be easily derived to be

$$Vn^2_{inCG} = \frac{KT\gamma}{gm_{CG}} * |H_{CG}(s)|^2 \tag{4.3}$$

Where

$$H_{CG}(s) = \frac{SC_pR_s}{2 + SC_pR_s} \tag{4.4}$$

It can be seen from the equation 4.3 and 4.4 that the noise due to the common gate transistor has high pass response or in other words, the input referred noise increases with an increase in the frequency. $H_{CG}(s)$ is the transfer function which determines the high-frequency response for the input referred noise. The total input referred noise contribution due to transistors is given by

$$Vn^2_{in} = \frac{KT\gamma}{gm_{CG}} * |H_{CG}(s)|^2 + \frac{KT\gamma}{gm_{CS}} \tag{4.5}$$

This can be rewritten in the form as

$$Vn^2_{in} = \frac{KT\gamma}{gm_{CS}} |H_{NC}(s)|^2 \tag{4.6}$$

Here the first factor is noise due to the common gate amplifier and second due to the common source amplifier. At low frequency, $H_{CG}(s) = 0$ or in other words common gate noise is canceled but as the frequency increases common gate noise begins to dominate.

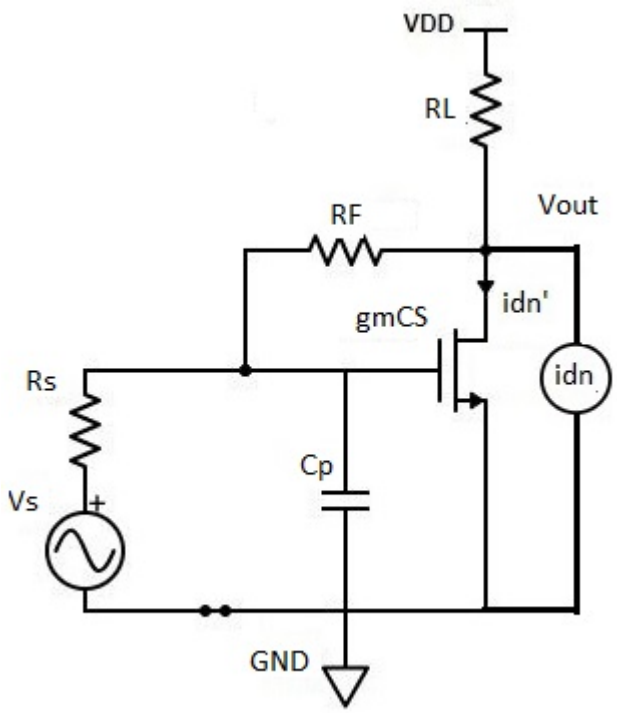

Figure 4.2: Resistive feedback LNA [42]

Figure 4.2 shows the basic Resistive-Feedback low noise amplifier. The input impedance is approximately given by $Z_{in} = (R_F)/(gmR_L)$ assuming the gain of the amplifier is high. The noise current at the output is approximately half the drain noise of the transistor which is idn. This is because of feedback action. The feedback action happens because of resistors R_F and R_S in the feedback loop. Suppose idn is the drain noise current of the transistor and idn' is the feedback noise current. Then

$$idn' = -(idn + idn')R_L \frac{R_S}{R_S + R_F} gm_{CS} \tag{4.7}$$

For high LNA gain $R_F >> R_S$ and using $R_S = (R_F)/(gm_{CS}R_L)$ this equation reduces to

$$idn' = \frac{-idn}{2} \tag{4.8}$$

Hence total drain current noise going through the load is given by $idn + idn'$ which is $idn/2$ and input referred drain noise voltage is given by

$$Vn_{in}^2 = \frac{idn^2}{4gm_{CS}^2} \tag{4.9}$$

Putting $idn = 4KT\gamma gm$ we get

$$Vn_{in}^2 = \frac{KT\gamma}{gm_{CS}} \tag{4.10}$$

The result is same as that of Noise-Canceling case. Hence as far as low-frequency transistor noise is concerned Resistive-Feedback and Noise-Canceling have the same input referred noise. The noise of the load resistor when input referred is attenuated by a factor A_v hence the contribution is minimal. Using small signal analysis one can easily prove that the noise contribution of feedback resistor is approximately given by $(4KTR_F)/(2gm_{CS}R_L^2)$ which can be rewritten in terms of the gain and the source resistor as $(KTR_S)/A_v$. Hence total low-frequency input-referred noise is given by

$$Vn_{in}^2 = \frac{KTR_s}{A_v} + \frac{KT\gamma}{gm_{CS}} \tag{4.11}$$

One can see from both equations 4.2 and 4.11 that as far as the resistive contribution is concerned, the Noise-Canceling contributes twice as much noise for the same gain. However, that is not a concern for us as the resistive contribution is expected to be negligible at the cryogenic temperature of 4K. One more subtle point that is to be noted here is that in the Noise-Canceling amplifier, the resistor R_{CG} will be carrying DC current while in the Resistive-Feedback topology the feedback resistor will not carry any DC current. Hence the noise of the resistor R_{CG} in the Noise-canceling amplifier may not scale due to self-heating. Next point to be noted is how the input referred noise varies for Resistive-Feedback as compared to Noise-Canceling for high frequencies. Intuition tells us that the variation must be much smaller. Unlike the Noise-Canceling, no noise canceling condition in the Resistive-Feedback has to be perfectly met. In the Resistive-Feedback, the presence of the parasitic poles reduces the feedback factor and hence causes the drain noise to increase at higher frequencies. One can now write

$$-(idn+idn')R_L\frac{Z}{Z+R_F}gm_{CS}=idn' \tag{4.12}$$

where Z is $R_S//sCp$. Using this one can arrive at the conclusion that $idn'=-idn/(1+Z/R_s)$. Total output noise current is given by $id(Z/(Z+R_S))$. The high-frequency input referred noise can be writtern as

$$Vn_{in}^2=\frac{KT\gamma}{gm_{CS}}H_{RF}(s)^2 \tag{4.13}$$

where $H_{RF}(s)=(2+2SC_PR_S)/(2+SC_PR_S)$ Comparing equation 4.6 and 4.13, it is the high-frequency noise response $H_{NC}(s)$ and $H_{RF}(s)$ in Noise-Canceling and Resistive-Feedback that makes the difference. As will be seen later, an approximate transconductance of order 1A/V will be required to reach a 10K transistor contribution to noise temperature using partial current reuse. The input parasitics for such a large transconductance value comes out to be close to 3pF. In all of the above analysis, the gate induced noise is neglected as it is expected to scale way more than drain current noise. The Figure 4.3 shows Matlab results for transfer functions $H_{NC}(s)$ and $H_{RF}(s)$. As seen in the figure 4.3, the high-frequency transistor noise in the Noise-Canceling amplifier increases significantly as compared to the Resistive-Feedback. This is expected as there is no delicate noise canceling condition that is required to be met in the Resistive-Feedback. Hence, due to robustness and superior high-frequency noise performance of Resistive-Feedback it is chosen as topology to be implemented. The benefits that balun Noise-Canceling amplifier will give with respect to Resistive-Feedback is better linearity for same gain since the output is differential. In our application linearity is not a primary issue since the input to the LNA will be close to 177 nV per channel and hence even for 20 to 30 qubit channels the peak to peak input is expected to be less than 5-6uV and the output to be less than 1 mV. Hence, the linearity is not expected to be an issue.

4.3. Architecture Resistive-Feedback

4.3.1. Effect of inductance at the gate

The effect of the parasitic pole due to Cp also causes the attenuation of the signal at the input of the amplifier. The signal transfer to the input is given by $V_{in}=V_S/(2+sC_PR_S)$. The parasitic pole also modifies the noise transfer to the output. The input referred noise as seen in the previous section is given by equation 4.13. Hence the signal to noise ratio taking only transistor noise into account is given by

$$SNR_{withoutinductor}=\frac{V_S^2*gm}{4KT\gamma}|H(s)|^2 \tag{4.14}$$

where $H(s)=1/(1+SC_PR_S)$.

Hence it can be seen that the parasitic pole causes SNR to degrade at high frequencies. Adding the inductor at the gate as shown in the figure 4.4 neutralizes the gate capacitance at high frequencies. It can be easily seen that

$$SNR_{withinductor}=\frac{V_S^2*gm}{4KT\gamma}|H_{ind}(s)|^2 \tag{4.15}$$

where $H_{ind}(s)=1/(1+SC_PR_S+S^2LC_P)$. We choose the inductor L for some in-band peaking in the SNR. L can be optimally chosen such that the band in which signal to noise ratio remains above $(V_S^2*gm)/(4KT)$ is maximum. The frequency at which the SNR is same as that of DC can be found by solving $|H_{ind}(s)|=1$. This gives

$$(1-\omega^2LC)^2+\omega^2C_P^2R_S^2=1 \tag{4.16}$$

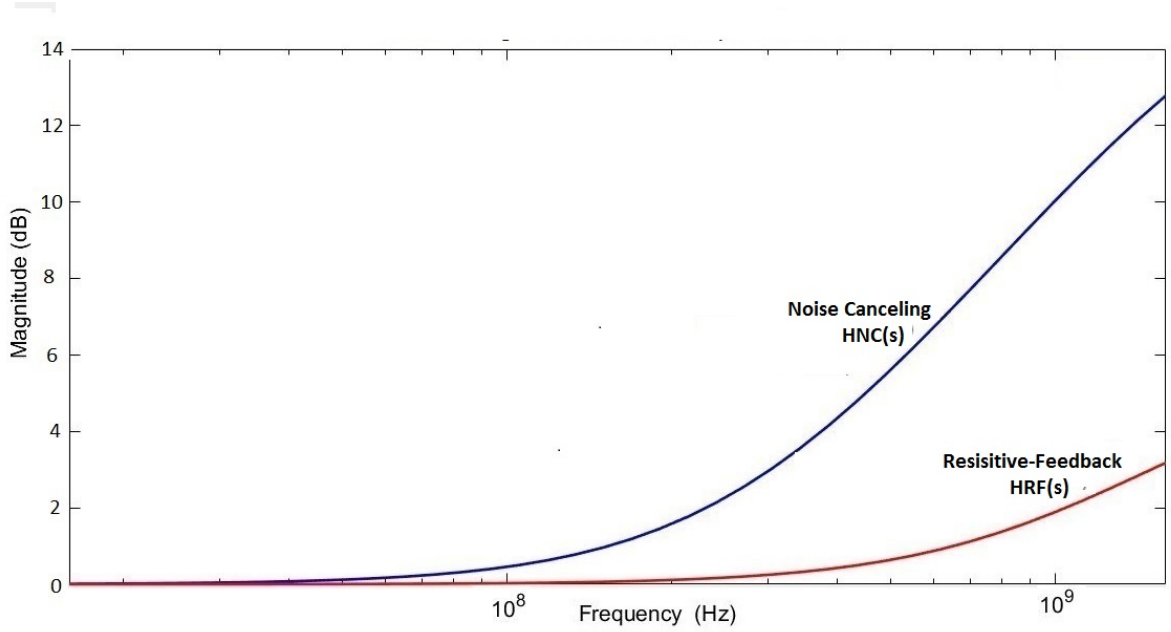

Figure 4.3: Noise Cancelling vs Resistive Feedback high frequency noise performance. The plot shows how the high frequency noise performance of the Noise-Canceling amplifier degrades rapidly with frequency.

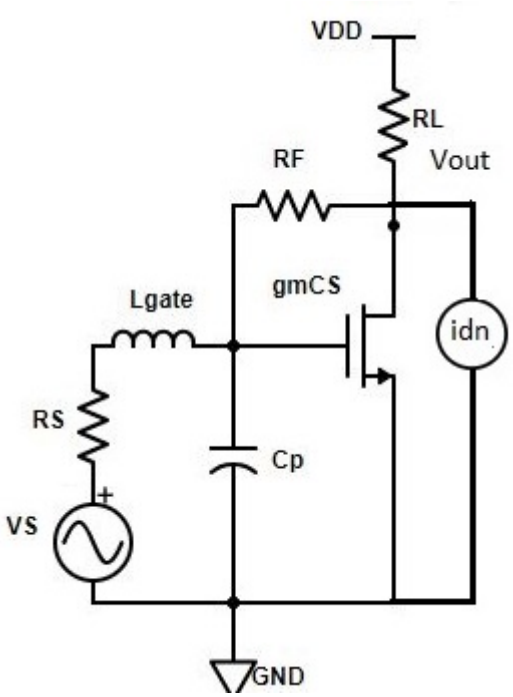

Figure 4.4: Gate inductance L_{gate} is added in the schematic to suppress degradation in high frequency NF

Solving the above equation one can arrive at

$$\omega^2 = \frac{2LC_P - C_P^2R^2}{L^2C_P^2} \tag{4.17}$$

$$\omega = 0 \tag{4.18}$$

For $L \leq (C_PR^2)/2$, only $\omega = 0$ is the solution hence $L = (C_PR^2)/2$ corresponds to maximally flat response. To maximize the bandwidth differentiate equation 4.17 with respect to L and putting $d\omega/dL = 0$ we get $L =$

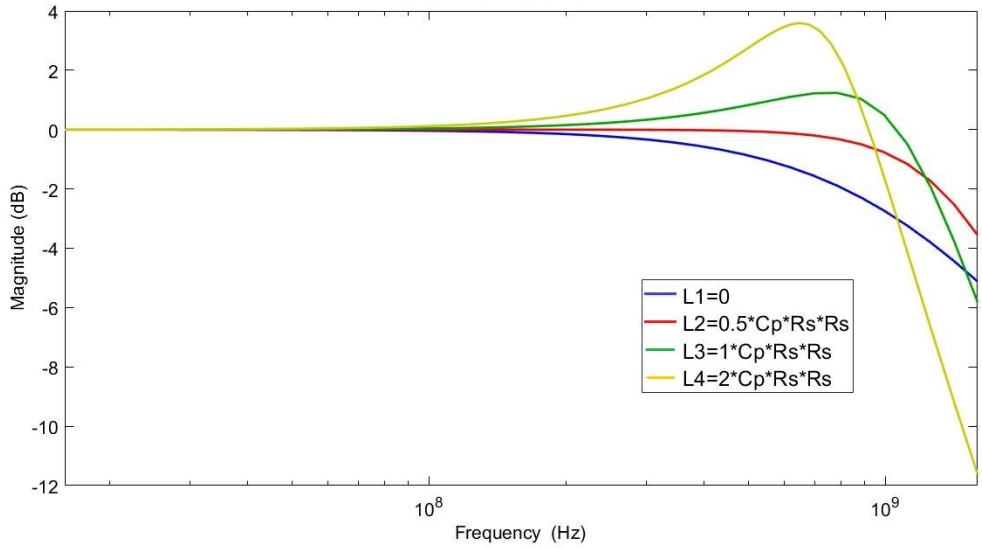

Figure 4.5: $H_{ind}(S)$ vs frequency for different values of the gate inductor. This figure shows how the signal to noise ratio signal to noise ratio varies for different value of gate inductance. The SNR dependence on $H_{ind}(S)$ is given by equation 4.15

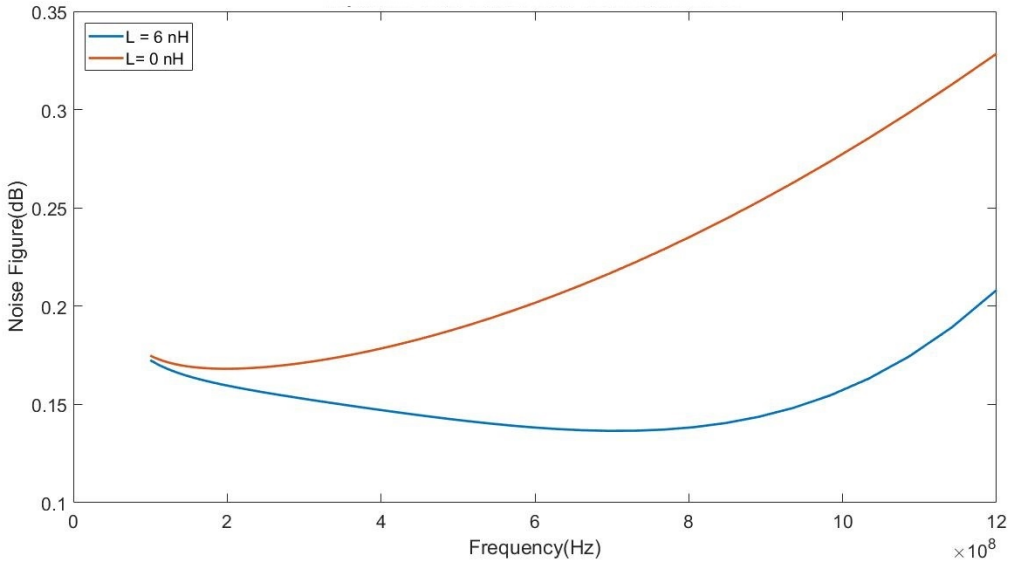

Figure 4.6: The Figure shows how the NF varies with and without inductance. The plot only shows transistor drain noise contribution.

$C_P R_S^2$. This is the approximate magnitude of the inductance that is needed at the gate. The figure 4.5 shows the transfer function $|H_{ind}(S)|$ vs frequency for different values of gate inductor. The later sections will show the final architecture of the LNA. The figure 4.6 shows the noise figure of the designed LNA with and without the optimized inductor at the gate. Transistor drain current noise is taken in this simulation as it is expected to dominate at 4K.

4.3.2. Cascaded LNA architecture

Figure 4.7 shows two LNA configurations with the Resistive-Feedback. It is desirable to achieve gain as large as possible in the LNA itself so that power consumption of succeeding stages can be reduced. For e.g, if we make the LNA having 10K noise contribution with 20 dB gain, the noise temperature of the second stage has to be less than 100k for its contribution to be less than 1k in the total input referred noise. In configuration A, the

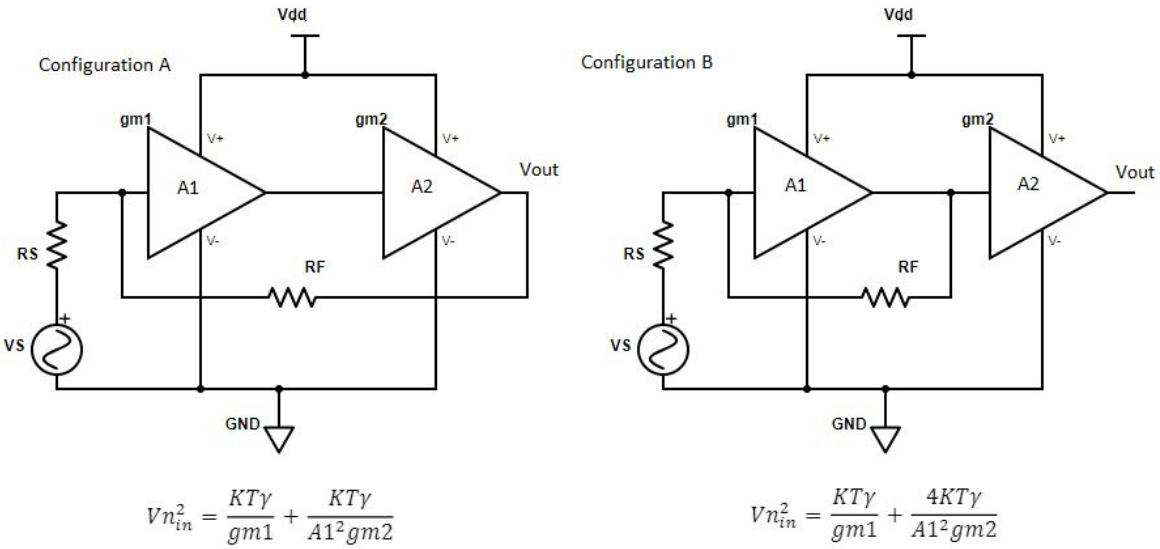

Figure 4.7: Cascading two LNA stages

second stage is kept inside the feedback loop. The input impedance, in this case, is given by $Z_{in} = R_F/A1A2$. The transistor contribution to input referred noise is given in this case is given as

$$Vn_{in}^2 = \frac{KT\gamma}{gm1} + \frac{KT\gamma}{A1^2 gm2} \tag{4.19}$$

Both the transconductors, gm1 and gm2 are in the feedback loop and hence as proved in the previous section the feedback suppresses drain noise power by a factor of 4. In configuration B the second amplification stage is outside the feedback loop hence the total input referred noise is given by

$$Vn_{in}^2 = \frac{KT\gamma}{gm1} + \frac{4KT\gamma}{A1^2 gm2} \tag{4.20}$$

One can see factor four increase in the second stage noise and hence second stage can cost four times the power if it is not kept inside the feedback loop. To elucidate the gravity of this situation taking previous example now one needs second stage LNA to be consuming power close to what is required for 25k LNA rather than 100K. Now one might argue for the stability of the circuit in configuration A since it consists of two stages and which will have multiple poles and may degrade the phase margin. The overall loop gain L of the circuit is given by the equation 4.21 for the condition of perfect match

$$L = A1A2\frac{R_S}{R_s + R_F} \tag{4.21}$$

This is approximately $A1A2R_S/R_F$. From the input impedance relation, one can see that the loop gain of the LNA in a perfectly matched configuration is 1. Even if there are two dominant poles lying on the same location, the loop gain would have dropped 6db as loop phase drops by 90 degrees. Now the total phase lag across the loop will be 270 degrees while the loop gain is -6db. Hence the condition for instability that is 360-degree loop phase and loop gain of 1 would never happen.

4.3.3. Need for differential LNA

For a robust design, it is necessary to keep the current loops inside the chip. Differential configuration keeps the current loops inside the chip. Single-ended configurations suffer from degeneration. Also, differential configuration enhances power supply rejection. However in the present design due to power considerations, the first stage has been kept single ended. This is because since the input is single-ended, power consumed to reach the same noise performance is approximately 4 times if the differential configuration is used. Also, each of the two transistors in differential configuration has to be twice as much size as in single-ended and

hence drastically impacting the bandwidth of operation. The input referred noise of a single ended common source amplifier is $4KT\gamma/gm$ and a differential amplifier with same current consumption approximately is $16KT\gamma/gm$. Since the first stage is single ended the current loop will extend beyond the chip i.e the current will flow through bond wires causing degeneration. The current loop for the first stage will be completed both through on-chip the decoupling capacitor and decoupling capacitors on PCB traces. The subsequent stages are differential in nature. The last stage or the output buffer is used to drive the 50-ohm load of the VNA and spectrum analyzer.

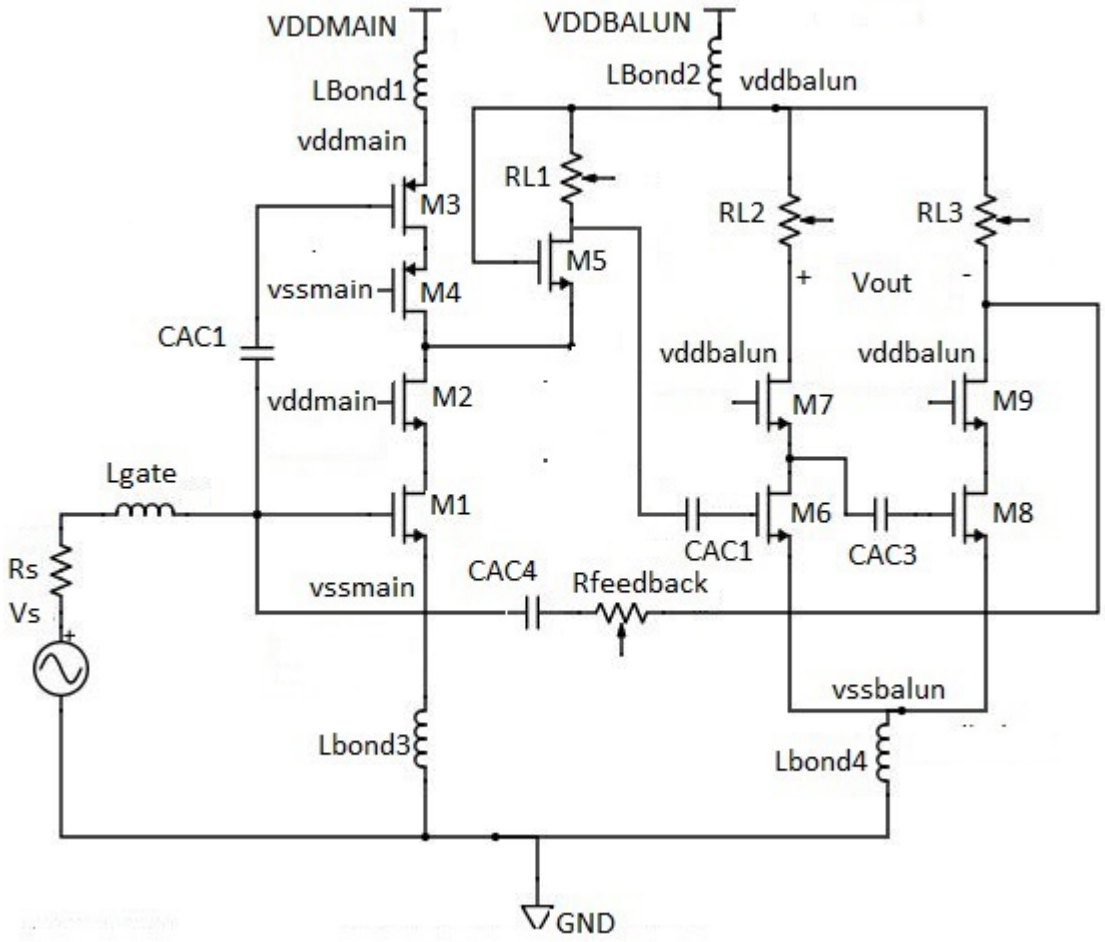

Figure 4.8: LNA architecture with parasitic and bondwires

4.4. LNA Design and layout

4.4.1. First Stage and Low impedance output buffer

Figure 4.8 shows the designed low noise amplifier. The first stage is a simple cascode amplifier with an output buffer M5. This is followed by a balun for single-ended to differential conversion. The feedback is taken from one of the outputs of the balun such that the overall feedback is negative. The gate inductor L_{gate} is used for enhancing the high-frequency performance. The cascode configuration is used in the first stage to reduce the effect due to C_{gd} else the effective input gate to drain capacitance would have been increased due to the Miller effect. To save power, partial current reuse is used. Both the PMOS and the NMOS are used for amplification where $gmp \approx gmn/2$. In this case, both PMOS and NMOS are of equal size and the current through them is approximately equal. Since broadband performance to 1GHz was required full current reuse was not used as then the size of PMOS would have to be approximately doubled leading to a 33 % increase in parasitic capacitance. The gm/Id ratio of NMOS used in the cascode is 10 and gm/Id of PMOS is 6. High gm/Id leads to greater power efficiency but increases parasitic capacitances. To decouple the bias points for both NMOS and

PMOS an AC coupling capacitor of value 15pF is used. Total parasitic input capacitance comes approximately to be 3pF which includes the C_{gs} and C_{gd} of the NMOS and the PMOS, bond pad capacitance of 400f and the parasitics of the ac-coupling capacitor. Both the top and the bottom plate capacitance of the AC coupling capacitor adds to the overall input parasitics. The impedance seen looking into the PMOS and NMOS cascode transistors should be low otherwise there would be signal gain from the gate of NMOS/PMOS(M1/M3) to the drain and hence increasing effective C_{gd} due to Miller effect. Hence the impedance seen at the drain of NMOS/PMOS cascode transistors should be low. An output buffer of approximate input impedance 10-ohm is used for this purpose. This corresponds to a transconductance of 100mS. The supply of the cascode amplifier of the first stage is kept different from the buffer to reduce the coupling. Otherwise, the ac current going through the buffer would flow through the bondwires and will disturb the source of PMOS M3. Making a single supply and adding decoupling capacitors could have kept the current loops inside the chip without causing any degeneration. In other words, the current would have flown through the decoupling capacitor rather than the bondwires. However, that would have required a huge amount of decoupling especially at lower frequencies around 100MHz($> 1nF$). Hence, the supply of the buffer and the cascode is kept different and some decoupling(250pF) is added from the supply to ground. In this way, the current flowing through the bondwires is kept the minimum. However, there is still some degeneration as not all the current loops are completed on the chip.

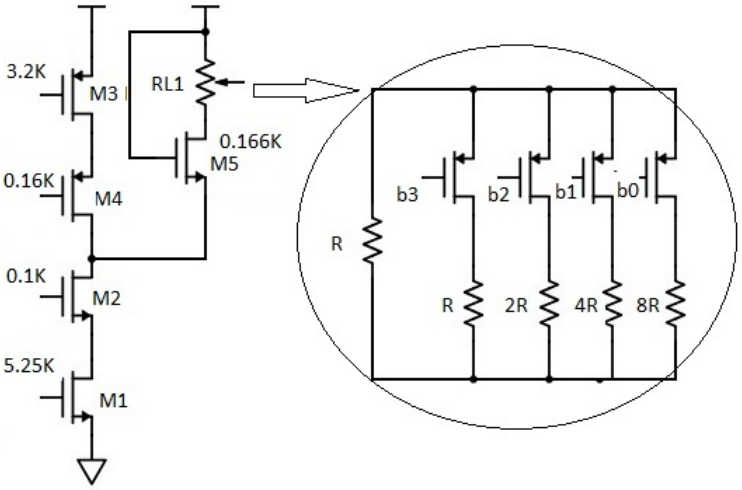

Figure 4.9: Noise summary of the first stage showing noise contribution of various transistors in kelvin(K)

Figure 4.9 shows the noise contribution of the transistors of the first stage at 500MHz. Both the cascode transistors and the output buffer are degenerate and hence contribute very little noise. The first stage low-frequency gain is close to 29dB. The 3dB bandwidth is dependent on both the supply and ground bond wire inductance. The first stage is single ended and hence both supply and ground bondwire inductance must be reduced. Otherwise, it would lead to degeneration of both the PMOS and NMOS transconductors impacting the 3db bandwidth. Signal degeneration due to bond wires would also cause an increase in the high-frequency noise contribution of succeeding stages. Hence a large number of supply and ground bond pads are used to reduce the inductance. The load resistance R_{L1} is kept programmable to accommodate process and temperature variations. Hence the load RL1 is a segmented resistive DAC. R=65 ohm is a fixed resistor while R,2R,4R,8R are programmable load resistors. For constant current the transconductance is expected to increase by a factor of 2 at 4K hence the load can be trimmed to half the value to keep gain constant since matching depends on feedback resistance R_F and gain A_v. In the buffer M5, the bulk is connected to the source to reduce the threshold and to keep first stage cascode transistors in saturation. The bulks of M2 and M4 are not connected to the source to increase the transconductance i.e 1/gm+gmb and hence to decrease input impedance seen at the source of cascode transistors. This will reduce the effective C_{gd}. The cascodes are then biased at supply rails to accommodate 1.1V operation. Fringe capacitor CAC1 is used to isolate bias points of the NMOS and PMOS transconductors. The Capacitor CAC1 will introduce parasitics

and hence top-most metal layers which can be used to synthesize the capacitors are used to minimize the parasitics. Out of the 7 metal layers available for TSMC-40nm, M3, M4, and M5 are top-most metal layers used to make the capacitors. Hence fringe capacitance is realized using these metal layers to minimize the parasitic capacitance. The parasitic capacitance to substrate ground will add to overall input capacitance.

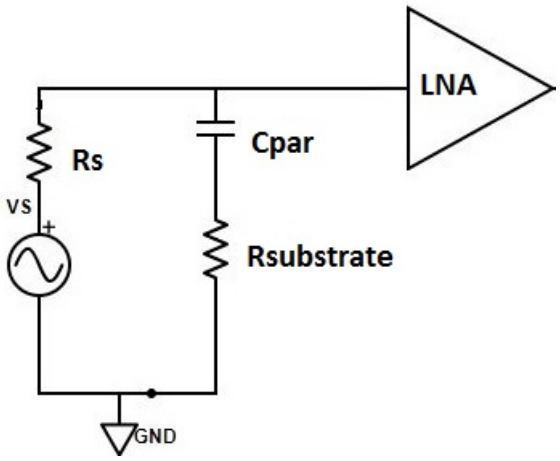

Figure 4.10: The figure shows how the substrate noise can couple to the input through parasitic capacitance Cpar. The Cpar here is the equivalent parasitic capacitance to the substrate

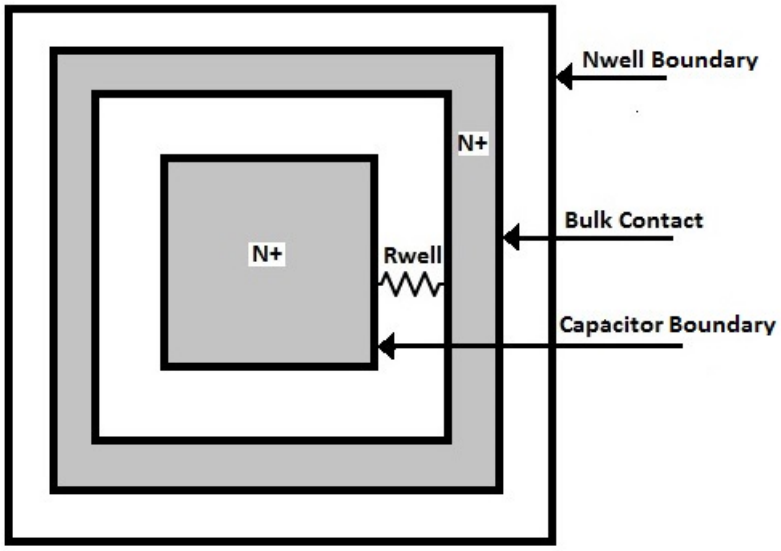

Figure 4.11: AC coupling capacitor layout showing nwell periphery between capacitor boundary and the actual bulk contact.

Another concern using the AC coupling capacitor CAC1 is the substrate noise. The capacitor layout is carefully chosen to reduce substrate noise. Measurements from the literature indicate that substrate resistivity explodes at 4K [11, 12]. Hence to reduce parasitics at 4K nonshielded RF MOM capacitors are used. However, the drawback is substrate noise introduced. Figure 4.10 shows how substrate noise gets coupled to the input. Here $Cpar$ is the parasitic capacitance to the substrate which can couple substrate noise at the input. The impact due to $R_{substrate}$ shows actually parabolic behavior. The noise that gets coupled has maximum values at some intermediate value of $R_{substrate}$. One can intuitively see this. Noise due to $R_{substrate}$ increases with increase in $R_{substrate}$. After a certain point, it reaches the maximum value. While increasing the $R_{substrate}$ further leads to the noise due to substrate resistance getting attenuated due to the resistive division. This can be seen in the figure 4.10. Hence layout in figure 4.11 ensures that room temperature substrate

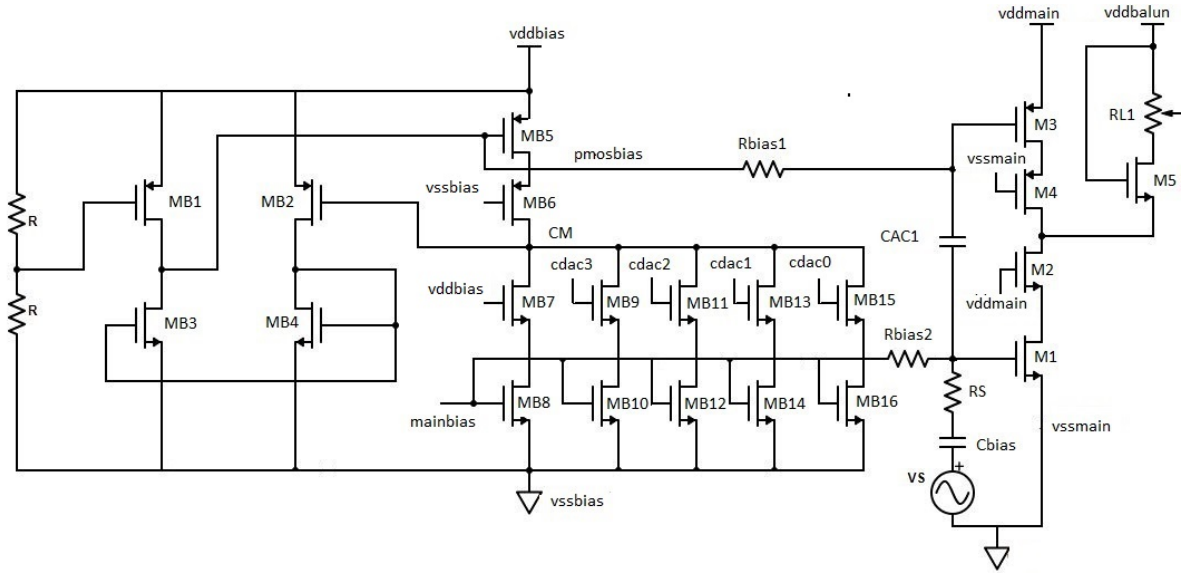

Figure 4.12: The figure shows replica bias used to bias the first stage PMOS and thus controlling buffer current.

noise is low enough. Rather than intrinsic silicon, the layer below the capacitor is chosen as N+. Between the actual bulk contact and the N+ layer below the capacitor, there is a nwell ring which acts as a small resistance of value 20 ohms which contributes negligible substrate noise. At 4K this resistance is expected to explode by a factor of 10^5 hence effectively killing the effect due to C_{par}.

In the layout, the PMOS is split into two parts. It was done to serve two purposes. Source and drain routing for the first stage was primarily done in the topmost metal 7 layer to reduce the voltage drop across the cascode. This is because the first stage cascode is expected to carry a substantial amount of current(60mA). The routing for the gate was done in the less thick M6 layer due to routing constraints. However, since the gate routing was done in M6 it adds gate resistance. Hence splitting the PMOS effectively helps in reducing gate resistance of the trace used for connecting the PMOS. Effective gate resistance added by routing traces is given by $(rp * gmp^2 + rn * gmn^2)/(gmn + gmp)^2$. Here rp and rn is the effective gate trace resistance used for connecting PMOS and NMOS. As will be seen later, the gate trace resistance will increase the noise by 1K after the layout. This is equivalent to addition of 0.16 ohms effective input resistance.

Also effectively twice as much decoupling capacitance can be connected close to PMOS trans conductor both on and off the chip. This can help subdue the effect of supply bondwires. The current flows in loops hence keeping decoupling capacitor as close as possible will help to shorten the current loops. The total power consumption of first-stage for achieving 10K noise contribution of the transistors is close to 66mW running on 1.1v supply at room temperature with the low-frequency gain of 29dB. Figure 4.11 shows the biasing for the first stage. It can be seen that AC coupling capacitor CAC1 is used to isolate the bias points. Both PMOS and NMOS are biased separately. Both bias points cannot be kept equal due to the low supply(1.1V) and the high threshold of the devices i.e 0.55V. Hence if the same bias-point is selected either one of them would be off or in weak inversion and hence would require very large size transistors to get the required transconductance. Bias for the all the NMOS transistors (mainbias) in the figure 4.12 comes from a constant gm circuit. The difference of currents in M3 and M1 flow through the buffer. Current through the buffer M5 is controlled by controlling current through PMOS M3. This is done using the current DAC made of transistors MB9, MB11, MB14 and MB16. All the transistors used in the first stage are short channel devices(50nm) to improve high-frequency performance hence suffer from significant shot channel effects like DIBL(Drain induced barrier lowering) which makes threshold dependent on the drain to source voltage of the transistor. It was seen that the noise increased drastically when one goes from 50nm to 40nm. The noise performance improves as one increases channel length further but on the contrary input capacitance increases. Hence 50nm was found as the optimum enabling good power efficiency with current re. Hence to ensure matching of the threshold voltages, replica bias is used in which a cascode replica of the circuit is used for controlling the current through M3. Hence pseudo differential amplifier comprising of the transistors from MB1 to MB4

are used as the level shifter to keep all the transistors in saturation and to have $vdd/2$ across each PMOS and NMOS cascodes of the bias. Not using the level shifter and hence diode connecting PMOS MB5 will leave NMOS transistors out of out of saturation. A four bit current DAC is used to control the buffer current through M5. The DAC is designed to compensate for any process variations and mismatch. If I is the current through NMOS M1, the buffer current can be controlled from 0 to I/4 with I/8 being typical. Typical current is set I/8 which is approximately 7.5mA corresponding to 10 ohm buffer impedance. The bias resistors R_{bias1} and R_{bias2} are of the order of 60KΩ. The bias resistors should be high to reduce their noise contribution. The high value of bias resistors serves a dual purpose first to attenuate noise of the bias circuit and second is to attenuate its own noise. From this logic, it may seem that the higher the value of R_{bias} better it is as the attenuation will be more. However after some value of R_{bias} parasitic capacitance of R_{bias} dominates. Higher the value of resistor higher will be parasitics introduced due to it which will be added directly at the input of the LNA. Hence the value was restricted to 60KΩ. Attenuation factor is given by 50/60000 which is approximately 1/1200. Hence the noise of the bias circuit will be attenuated approximately $60dB$ and hence will be negligible. The noise that the bias resistors will inject approximately at the input.

$$Vn^2_{bias} = \frac{4KTR^2_S}{R_{bias}} \tag{4.22}$$

Since this is at the input of the LNA this is equivalent to input referred resistance of R^2_S/R_{bias} which is approximately equal 1/24Ω . Two of such resistors (one for NMOS and other for PMOS) will add approximately 1/12Ω which corresponds to 0.6K in noise temperature at room temperature. The capacitor here C_{bias} is the external bias AC coupling capacitor.

4.4.2. Second stage Balun

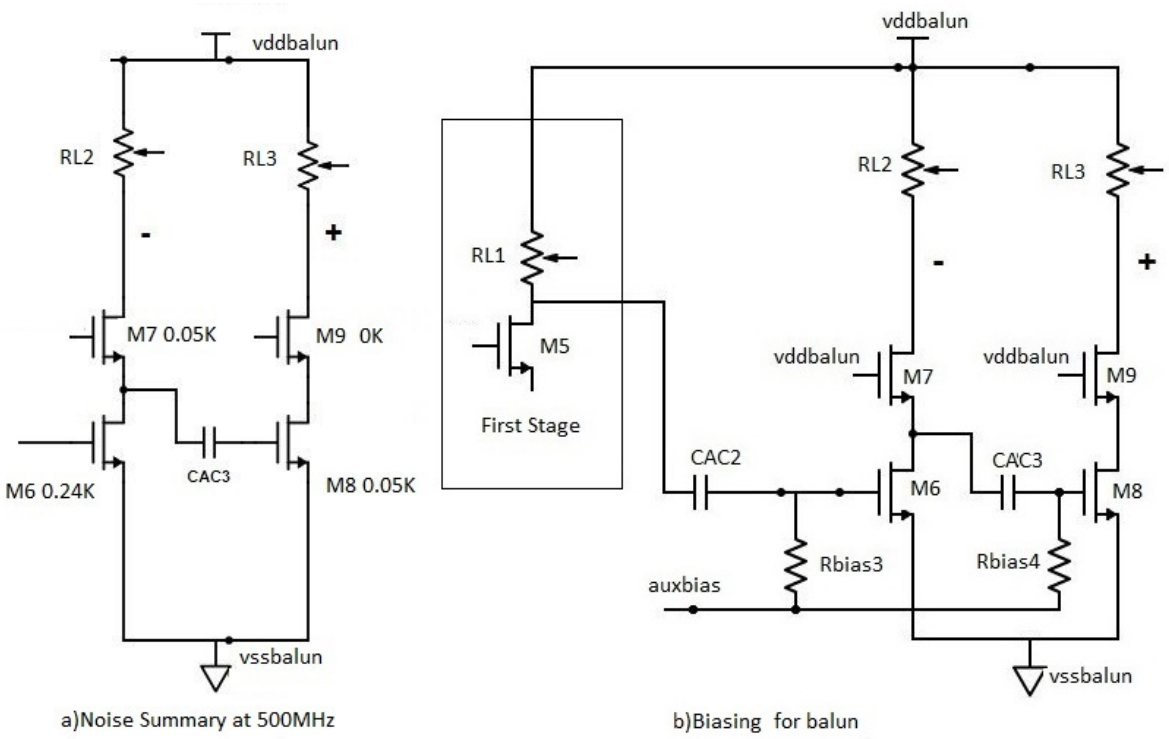

Figure 4.13: Balun noise summary and biasing

As discussed previously in order to mitigate interstage coupling single-ended to differential converter is used in the second stage. The second stage is designed for a differential gain of greater than 12db such that the current consumption of post-processing circuits will add a negligible amount of noise. The total gain of the LNA now exceeds 40dB. A Balun is used instead of a differential amplifier because of its low input referred noise. The input referred noise for a single-ended to differential conversion comes out to be $8KT\gamma/gm$ for a differential amplifier and $6KT\gamma/gm$ for balun for same input power. Hence noise contribution due to

balun is approximately 25% less than a differential amplifier. Figure 4.13 shows noise summary of the balun at 500MHz. To keep noise contribution below 0.5K approximately 7mW is spent. As discussed in the previous section on architecture, balun will be in the feedback loop to reduce the noise four times. If the balun is not inside the feedback loop, power consumption for 0.5K noise contribution will be close to 28mW. Balun and succeeding stages unlike the first stage are made using 100nm transistors to reduce noise as the bandwidth requirements of the succeeding stages are way relaxed than the first stage. This is because 100nm transistors have lesser drain noise level (20 percent) as compared 50nm transistors. This illustrates the significance of keeping balun in feedback loop as well as necessity of having high gain in the first stage to reduce the noise contribution from the second stage. Noise from M6,M7 and M8 is available at the + output of balun hence the feedback action causes suppression by a factor of 4. Also M6 is the highest contributor as its noise which is present at both differential outputs is correlated and out of phase by 180 degrees. Hence, its noise contribution is expected to be 4 times the other contributors. Again like the first stage the balun load resistance is made programmable to compensate for process and temperature variations. Balun like the first stage is also biased with high resistance bias resistors R_{bias3} and R_{bias4} as shown in Figure 4.13. The bias resistor R_{bias3} should be much greater than the series impedance of AC coupling capacitor CAC2 and RL1 as can be seen from the Figure 4.12. Similarly, the resistance R_{bias4} must be much greater than $1/gm3$ for significantly attenuating the noise of both bias source as well as bias resistors. Simulation results suggested the value of 25KΩ for both the resistors. The load resistor are also kept programmable in the balun to trim the gain for input matching.

4.4.3. Differential gain stages for noise measurement

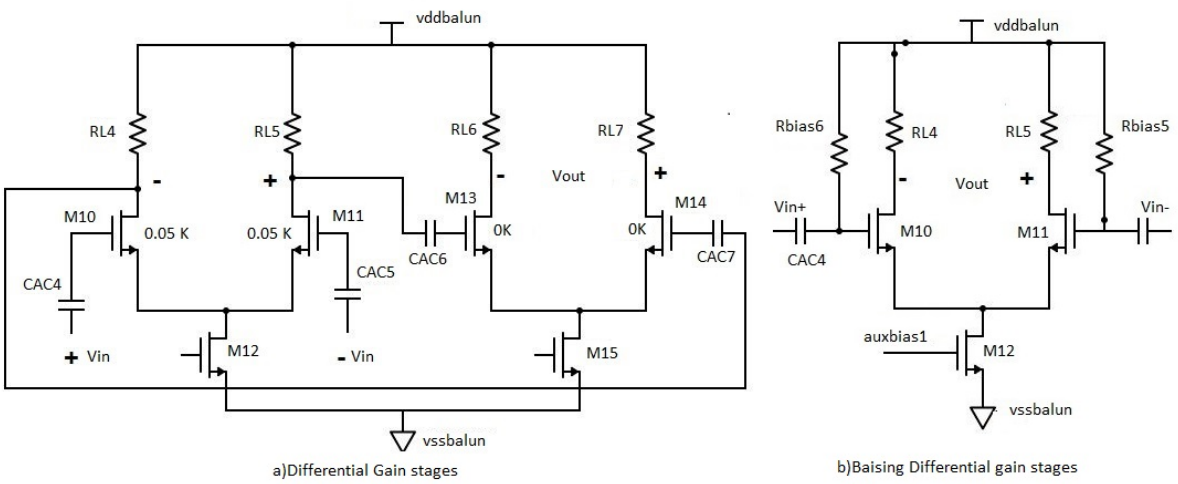

Figure 4.14: The differential gain stage to acting as a pre-amplifier to enable low power spectral density noise measurements. The figure also shows the how he amplifier is biased

A spectrum analyzer without pre-amplifier option was used for measuring noise both for room temperature and cryogenic measurements. The instrument has a finite resolution of -167 dBm/Hz. Power spectral density at room temperature due to the source is equal to

$$Pin_{300K} = KT = -174dbm/Hz \tag{4.23}$$

If one wants to measure with a resolution of 0.1k then power spectral density corresponding to 0.1K at the input

$$Pin_{0.1K} = -209dbm/Hz \tag{4.24}$$

Corresponding Output power spectral density is given by

$$Pout_{0.1K} = -209dbm/Hz + G \tag{4.25}$$

Here G is the gain or S_{21} of the LNA after post amplification. The output power change must be much greater

than the resolution of the spectrum analyzer.

$$Pout_{0.1K} \geq -167dbm/Hz \tag{4.26}$$

which gives $G \geq 42db$. At 4K the long cables going inside the dip stick to PCB will be used for measurements. Hence there would be cable losses which have to be accounted. Some margin is taken, therefore. Simple differential amplifiers are added for noise measurements to increase the gain from 43dB approximately to 54dB. Each of the differential amplifiers gives 9dB of gain. However at output their is 6 dB of loss due to output terminations. Figure 4.14 shows the schematic of used differential gain stages. Since the amplifier is differential in nature its supply is shared with balun and output buffer of the first stage. Also, all the amplifiers are AC coupled so that offset cannot propagate forward which could change DC bias points of one of the gain stages. This offset propagation could, in fact, ruin the measurements.

4.4.4. 50 ohm differential to single ended output buffer

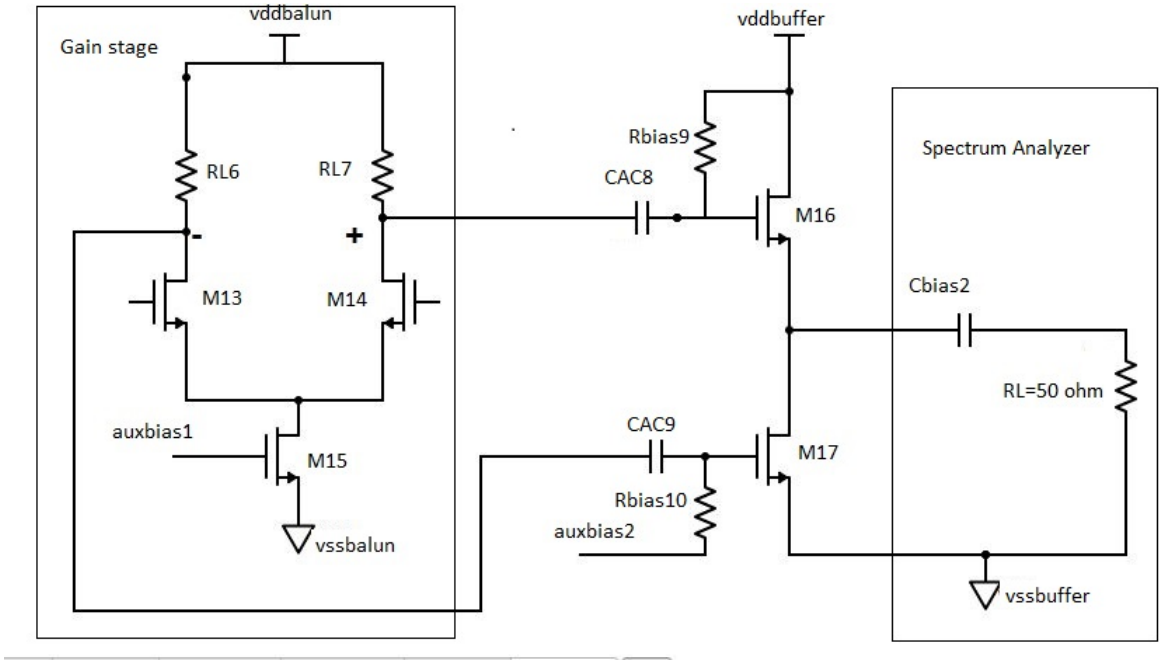

Figure 4.15: The 50-ohm output buffer to drive the VNA and the spectrum analyzer. The figure also shows the biasing of the output buffer.

Figure 4.15 shows differential to single ended output buffer. The differential gain of the buffer is given by

$$A_{d2s} = \frac{1}{2}\frac{gm_{16}RL + gm_{17}R_L}{1 + gm_{16}R_L} \tag{4.27}$$

The bulk and the source of the transistor M16 is connected together and hence $gm_{16} = gm_{17}$. The gm_{16}, which determines the output impedance of the buffer for 50Ω matching is selected as $gm_{16} = 1/R_L = 50\Omega$. Hence for an ideal case, the differential gain is equal to $\frac{1}{2}$ or a loss of -6db due to output matching. The 50Ω output buffer can be controlled using another constant gm bias. Also, the current of M17 is made programmable using current DAC. Again both the biasing resistors should be high enough so that their noise contribution is low due to attenuation.

4.4.5. Feedback Resistor

Feedback resistor of a typical value 4KΩ is used in the design. The differential gain of the core LNA is 43dB which corresponds to the single-ended gain of 37dB as seen on one of the differential outputs connected to the feedback resistor. This corresponds to an input impedance of 56 ohms or S11 of -25 dB. The feedback resistor can be trimmed to either 3KΩ, 4KΩ, and 5KΩ. With one of these three resistor values a decent $S11 (S11 < -10dB)$ can be achieved even if there is ±6dB change in the gain. As shown in the figure 4.16 one of

the resistance is switched as a feedback resistor using one of the PMOS switches. PMOS switches are implemented using 100nm transistors to reduce threshold which is approximately 0.5V. Typical room temperature source and drain potential of the PMOS switches are close to 0.9V corresponding to gate to source voltage of 0.9V. Switch resistance is close to 150 ohms typical. The bulk of the switches are connected to the supply rather than the source to mitigate parasitic bulk capacitance which would have been added to the input. At cryogenic temperatures both the threshold and mobility is expected to rise. Threshold of these switches are expected to be 0.7V based on the measurements However, if one takes into account the constant gm operation, current will also reduce by a factor of half at 4K. Hence the V_{ov} is expected to decrease from 0.4V to 0.3V. Along with this mobility will increase by a factor of 2. Hence, the switch resistance is expected to decrease at 4K. This can be seen from the following equation of R_{on}.

$$R_{on} \propto \frac{1}{V_{ov}\mu} \tag{4.28}$$

Noise due to the resistors that are not used, example the 3K and 5K resistors as shown in figure 4.16 can couple through parasitic C_{ds} to the input. Hence a large number of switches cannot be used for very fine trimming of feedback resistance. Also, care was taken care and simulations were performed since there is also a chance the noise from transistors driving the transistor MBF2 can couple at the input through C_{gd}. Hence the switch size was carefully chosen. However, the noise at the gate of the transistor MBF2 will appear at the input after attenuation because of the lower value of C_{gd}. There is a capacitive division between C_{par} of the total input capacitance and C_{gd} of PMOS used to switch feedback resistance. The lower the size of the switch the more will be attenuation of noise from the digital driver.

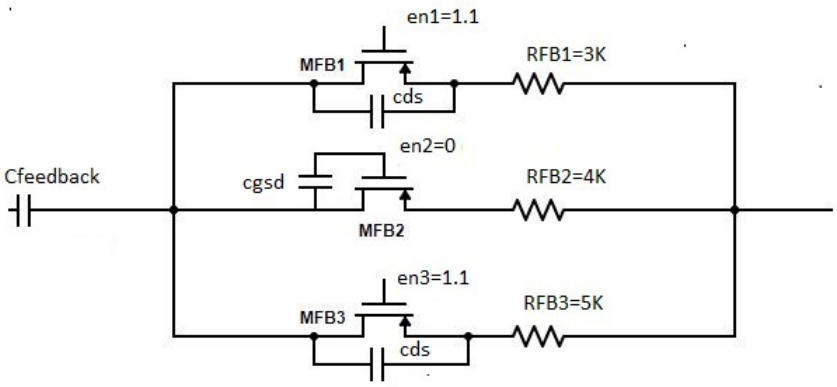

Figure 4.16: Feedback Resistor to control the input matching of the LNA

4.4.6. Constant gm bias

The gm expression for 1:2 ($W_{MG6} : W_{MG7}$) ratio of the core transistors in the constant gm bias circuit is given by the following equation.

$$gm = \frac{2-\sqrt{2}}{R} = \frac{0.585}{R} \tag{4.29}$$

Figure 4.17 shows the modified constant gm bias which is used to bias the first stage. The channel length is selected as 50nm which is the same as that of NMOS used in the first stage. This is done for matching and to mitigate any length dependent short channel effect such as threshold voltage dependence on length. The ratio between sizing M_{CG6} and M_{CG7} is kept 1:2 as discussed earlier. In contrast to traditional constant gm bias, the modified constant gm circuit is used [55]. It has another loop comprising of the transistors M_{CG3} and M_{CG5}. This loop is used for the dual purpose, first is to make $I_{MCG1} = I_{MCG2}$ and second is to make $Vth_{MCG6} = Vth_{MCG7}$. The additional feedback loop makes the gate voltage of transistor M_{CG1} approximately equal to drain voltage of M_{CG2} or node X. This ensures that $I_{MCG1} = I_{MCG2}$ since the gate and drain voltages of both the transistors are equal there is almost no current mismatch due to channel length modulation effect. Also, this ensures that drain voltages of transistors M_{CG6} and M_{CG7} are equal hence reduces threshold mismatch in the transistors M_{CG6} and M_{CG7} due to DIBL (Drain induced barrier lowering).

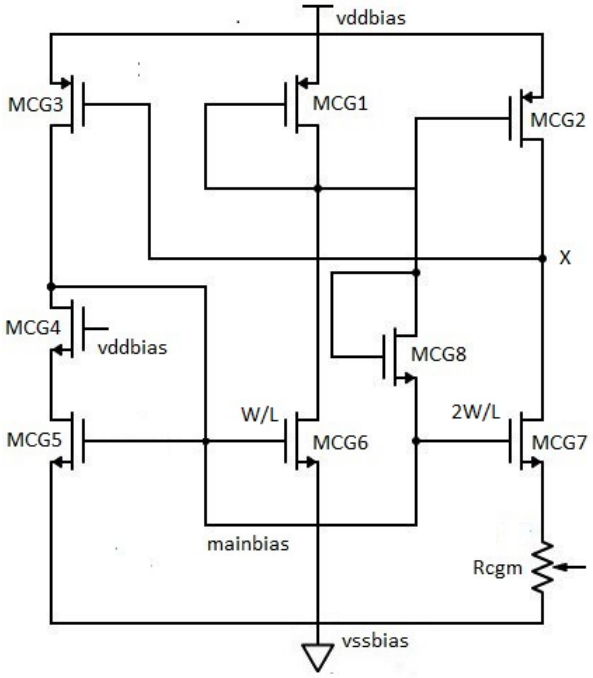

Figure 4.17: Constant gm bias circuit used to bias the NMOS transconductor of the first stage and NMOS transistors in the replica bias.

The threshold mismatch can add more temperature dependency in transconductance. The PMOS current sources are made of 100nm devices to further reduce current mismatch due to channel length modulation. Also using the deep nwell, body, and source of the transistor M_{CG7} are connected together for matching with the threshold of M_{CG6}. Intuitively the circuit feedback works as described. The aspect ratio of the PMOS transistors M_{CG1} ,M_{CG2} and M_{CG3} are the kept same. Also, the aspect ratio of NMOS transistors M_{CG5} and M_{CG6} are kept the same. Since the gate of M_{CG5} and M_{CG6} are connected the current through them is approximately same assuming they are in saturation. This further suggests that V_{SG} of transistors M_{CG1} and M_{CG3} has to be approximately the same since the current through them is equal. Hence the node X follows gate/drain voltage of transistor M_{CG1}. Drain voltage of M_{CG5} is made approximately equal to that of M_{CG6} and M_{CG7} by using NMOS M_{CG4} which replicates $vddbias - V_{SG}$ drop across PMOS M_{CG1} and M_{CG2}. Using the long channel approximation one can see that effective transconductance of transistor M_{CG6} is $gm_{MCG6} = gm$ while for M_{CG7} it is given by $gm_{MCG7} = \sqrt{2}gm/(\sqrt{2}gmR_{cgm}+1)$. This is due to degeneration of transistor M_{CG7}. Since in this circuit $gm = 0.5851/R_{cgm}$, the effective transconductance for M_{CG7} is given by $gm_{MCG7} = 0.77gm$. Seeing the loop starting from node X one can see that the negative feedback effect due to transistor gm_{MCG6} is greater than positive feedback due to M_{CG7}. One can effectively prove that if the transistors M_{CG6} and M_{CG7} are sized in the ratio $1:\alpha$ then transconductance of M_{CG7} is equal to $(\sqrt{\alpha}gm)/(2\sqrt{\alpha}-1)$ which is never greater than or equal to gm for α greater than one hence even with mismatch in our case with $\alpha = 2$ negative feedback will always be greater than positive feedback.

It is seen in the simulations that the gm of the circuit is close to the long channel expression even when short channel devices are used. The most probable reason could be low V_{ov} of the operating transistor. For good power efficiency, the transistors are operated with an overdrive of 70mV. Hence the channel pinch-off seems to occur before velocity saturation. The reason for choosing 1:2 ratio rather than 1:4 ratio which would have given $gm = 1/R$ is to prevent transistor $MCG7$ from entering the sub-threshold region. At 4K, it is expected that mobility will increase by a factor of two. Hence to keep the same transconductance, overdrive will be even less and the gm is expected to even close to the long channel expression. The shot noise component which is proportional to the current is expected to be twice as less than for constant gm. R_{CGM} is made programmable to first account for any process variation in the resistor and also to account for mismatch variations. The bias node main bias will directly bias the first stage amplifier using the biasing resistor. 6 bit segmented R DAC is used for trimming transconductance to account for any process and temperature variations. If gm is the nominal value of transconductance, the transconductance dynamic range varies from 0 to 3gm. The range of 0 to 2gm will be used at room temperature which can easily compensate for process

variations and mismatch. Monte-Carlo mismatch simulations showed sigma in gm of approximately 70ms for the mean value of 640ms. On top of this, there will be ± 20% process variation in resistance that has to be compensated. Hence trimming is added in order to compensate for this effect. Since the LNA is not fully differential, source degeneration could be a problem hence transconductance can be decreased if source degeneration turns out to be a problem or in other words, the estimated bond wire and PCB inductance turns out to be way greater. Figure 4.18 shows the traditional constant gm bias used to bias balun and gain stages. Both the PMOS and NMOS are made using 100nm transistors to match the transistors of balun and gain stages. The drain induced barrier lowering effect is a short channel effect and is more pronounced in 50nm devices. Hence for the balun,gain stages and the output buffer which uses 100nm transistors traditional constant gm bias used. The gm is made programmable in each of the bias circuits. In all the constant gm circuits one of the equilibrium states could be zero current state. In this state, the PMOS gate is high or close to Vdd and the NMOS gate is low or close to gnd. To mitigate this startup condition from happening a diode connected NMOS is added between gates of PMOS and NMOS. If such a state happens diode-connected transistor causes charge to flow from gate of PMOS to the gate of NMOS and hence helping the circuit to come to the operational state.

Figure 4.19 shows external bias control for the main stage.

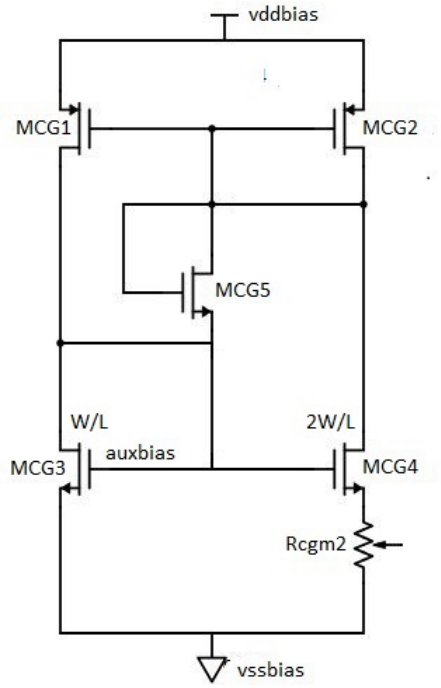

Figure 4.18: Constant gm bias for balun, gain stages and the output the buffer

4.4.7. Chip Layout with Padring

To enable digital control an SPI(serial peripheral interface) is used and synthesized. Figure 7.19 shows the overall chip layout with padring. The area of the chip is limited by the padring. The overall chip area is approximately $1 \times 1.1 mm^2$. The number of the pads used in the padring were 34. Approximately 6 pads are used for supply and 4 for ground in order to reduce bond pad inductance of the first stage. The chip will be placed inside a ground plane on the PCB. The length of ground wire-bond connection is expected to be shorterr as compared to supply wire-bond. Hence larger number of supply bond pads are used. An internal temperature diode is also put close to the amplifier to get an estimate of temperature.

4.5. Gate inductance and it's impact on S11

The gate inductance helps in improving broadband noise figure. In the design, an external gate inductance on PCB is used. This is because an external inductor will have much higher quality factor than an on-chip 6.6nH inductor. If an on-chip gate inductor is used it will dominate the overall noise temperature of the LNA at the

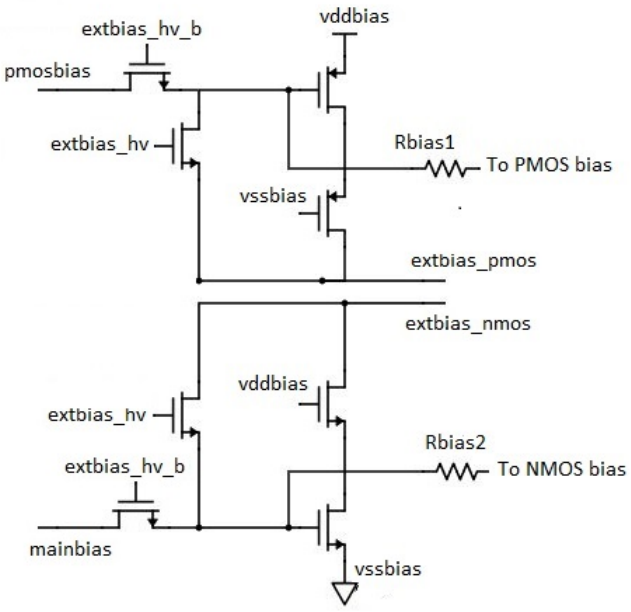

Figure 4.19: External bias control to enable external control of the LNA bias points in the case the internal bias malfunctions at cryogenic temperatures

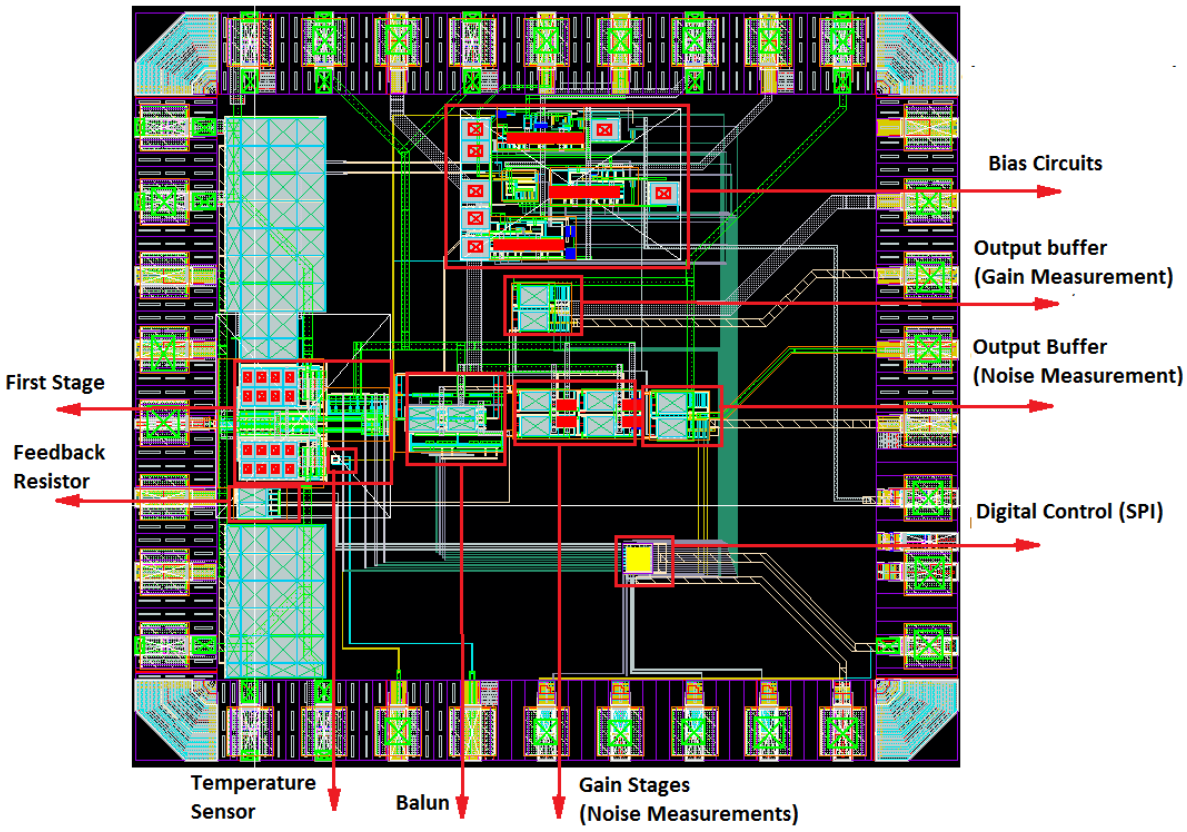

Figure 4.20: Top level layout with the padring showing the location of the LNA, gain stages, bias circuits, and SPI

room temperature. Hence to study the behavior of transistor noise for experimental purposes when one goes from room temperature to cryogenic temperature an external gate inductor is used so that room temperature performance is not impacted. Gate inductance also helps in neutralizing some input capacitance. Input

impedance without inductor is dominated by input capacitance.

$$Z_{in} = \frac{R_{in}}{1+\omega^2 C_{par}^2 R_{in}^2} - \frac{j\omega C_{par} * R_{in}^2}{1+\omega^2 C_{par}^2 R_{in}^2} \tag{4.30}$$

Inductance with $L = CR^2$ neutralizes some capacitance to improve $S11 \leq -10$ db bandwidth. The modified S11 now is

$$Z_{in} = \frac{R_{in}}{1+\omega^2 C_{par}^2 R_{in}^2} + j\omega(L - \frac{C_{par} * R_{in}^2}{1+\omega^2 C_{par}^2 R_{in}^2}) \tag{4.31}$$

4.6. Schematic and Post Layout results

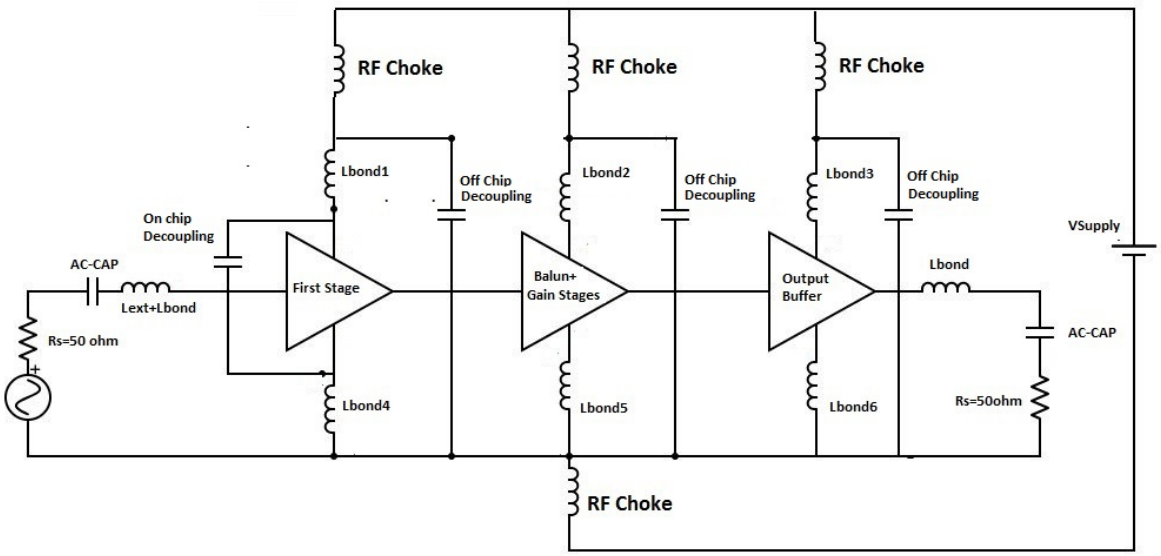

Figure 4.21: Top level Simulation Setup

For simulations, set up in the figure 4.21 is used for S parameter and noise figure simulations. The RF choke is connected to represent long wires which is particularly important for 4K measurements. On-chip decoupling capacitors are used in the first stage to reduce degeneration of the first stage PMOS and NMOS transconductors. The simulation setup also represents how the chip is connected on the PCB.

Figure 4.22 shows the schematic level simulations of S parameters with and without bond wires. As discussed in the previous section, 250pF of on-chip decoupling is used to prevent degeneration of NMOS and PMOS transconductors of the first stage. Total approximate value of on-chip decoupling used is equal 250 pF. However, still the 3db bandwidth reduces significantly from 1.5GHz to 1.2GHz when the effect of bond-wires is considered due to source degeneration. The ground bond wires of value 1nH are approximately taken in simulations and the supply bond wires are taken to be 2 nH. This is because the chip will be placed on the ground plane. Hence, because of the shorter length of the bondwires, the ground bondwire inductance is expected to be much lesser as compared to supply bondwires. Also it can be seen in the figure 4.22, band in which $S11 \leq -10dB$ increases. This is because source degeneration increases the effective impedance seen looking at source to gate capacitance i.e $(C_{gs} + L_s + gmL_s/C_{gs})$. This is the standard result from source degeneration amplifier. Hence it can be seen that the resistance gmL_s/C_{gs} decreases the current through C_{gs}. The gain shown in the S parameter simulations is of overall noise chain including the post-amplifiers required for noise measurement. The In-band gain including post amplifier is approximately 54dB. The gain after post amplification is 60dB. Approximately 6dB gain is lost due to output matching. Hence differential gain of the amplifier after post amplification is approximately 60dB. The differential gain of the core LNA(first stage+balun (4.8) is close to 42.5 dB which corresponds to single ended gain of 36.5dB. The default value of feedback resistance is 4KΩ. This corresponds to default input impedance of 60 ohms or S11 of $-21dB$. The reverse isolation also decreases due to bond wires however it is still negligible.

Figure 4.23 shows the post layout and schematic S parameter simulations taking bondwires into consideration. The low frequency gain of the noise chain decreases by about 2dB. The output of each stage is ac cou-

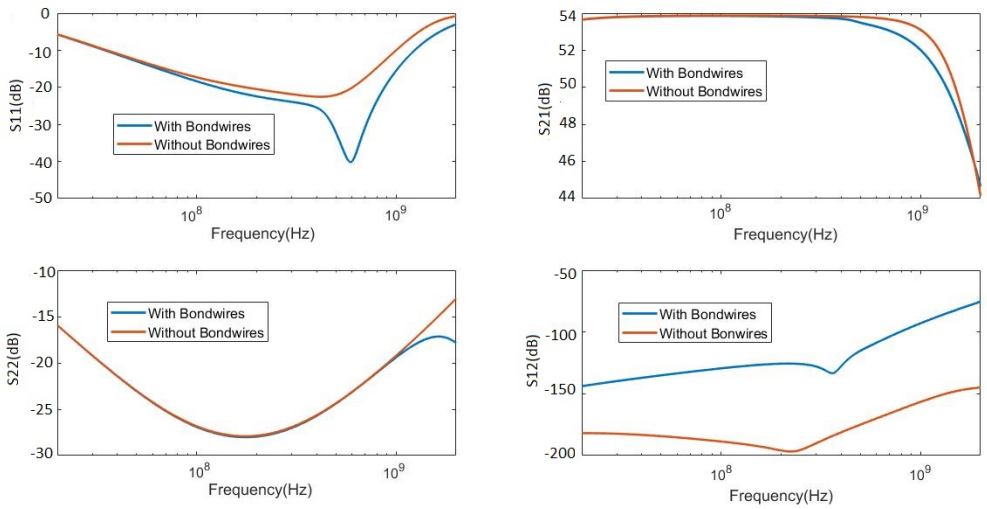

Figure 4.22: Schematic level S parameter simulation of the LNA after post amplification with and without taking the effect of bond-wires

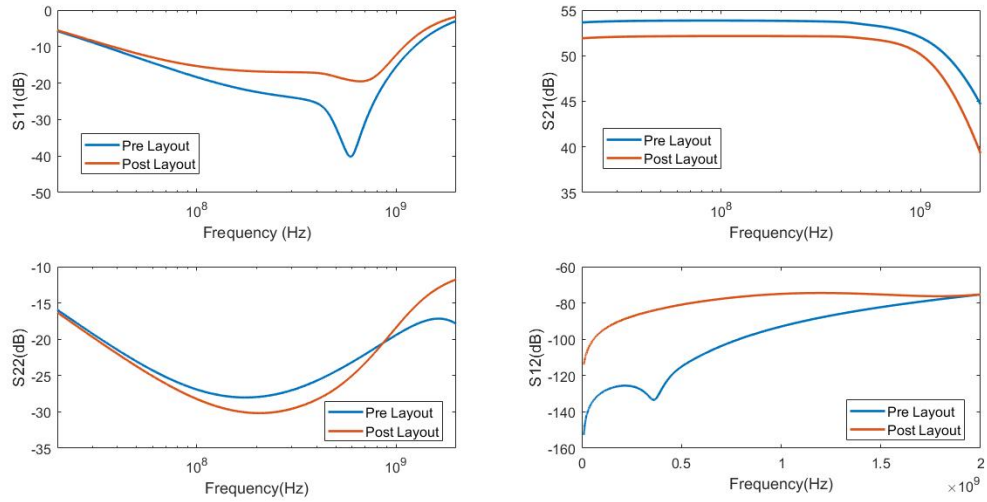

Figure 4.23: Simulated schematic and post-layout S parameter simulations of the LNA after post amplification including effect of bond-wires

pled to succeeding stages and hence increase in the parasitics causes some attenuation due to ac-coupling capacitance and input capacitance of succeeding stages. However still the gain level is high enough for accurate noise measurements. Figure 4.24 shows the top level small signal schematic diagram without the biasing and ac-coupling capacitors. Gain measurement buffer bypasses the post amplifiers used for noise measurement and is used to measure the differential gain of the core LNA. The gain of the balun is also kept limited because of loading due to the buffer for gain measurement and differential amplifier for noise measurement will limit the bandwidth of the balun. However, still the overall bandwidth is dominated by the input pole of the LNA. Figure 4.25 shows the post layout gain of the gain measurement buffer. The buffer is the copy of the buffer used in noise measurements. The S_{21} of the LNA is different from the open loop gain single ended gain A_v that is used to make the amplifier. The relationships can be derived using simple algebra. Suppose V_{in} is the incident wave to the input of the Resistive-Feedback amplifier then $V_{in}(R_{in}-R_S)/(R_{in}+R_S)$ is then reflected signal from the LNA. Effective input to the LNA is the superposition of these to waves and is given by $V_{input} = 2V_{in}R_{in}/(R_{in}+R_S)$. Using $R_{in} = R_F/A_v$ and $V_{out} = A_v V_{input}$ one can get

$$S_{21} = \frac{Vout}{V_{in}} = \frac{2R_F}{R_S + R_{in}} \tag{4.32}$$

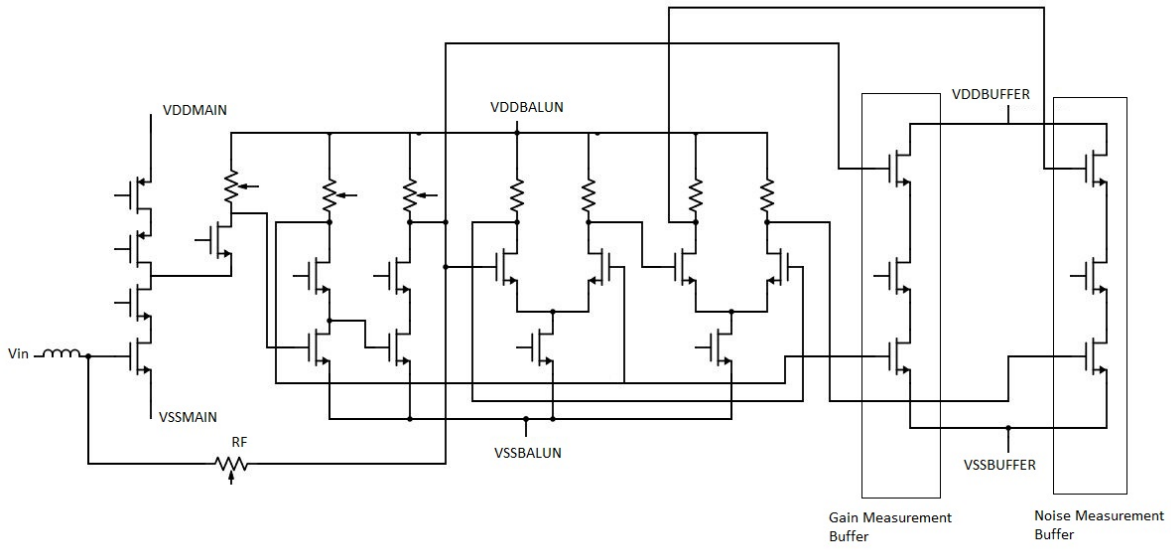

Figure 4.24: Small signal top level diagram of the LNA excluding the bias

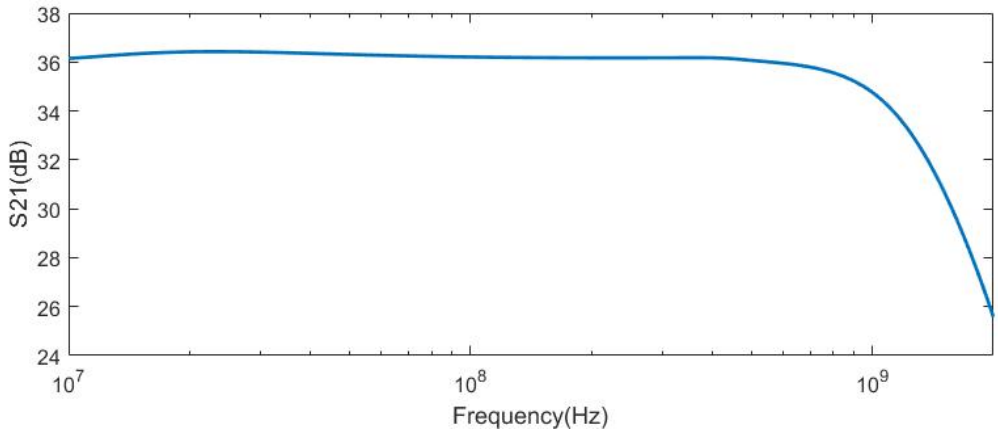

Figure 4.25: Simulated post-layout gain of the LNA core read from the gain measurement buffer

For the gain measurement buffer chain total $S_{21}(total)$ is given by

$$S_{21}(total) = \frac{2R_F A_{buffer}}{R_S + R_{in}} \tag{4.33}$$

Only in the case of perfect input match i.e ($R_s = R_{in}$), $S_{21}(total) = 2 * A_{buffer} A_v$. The factor 2 is added because differential gain is twice the single ended gain. The buffer gain is -6db or 1/2 for the output match. The direct S_{21} gain measurement using the buffer can also be used to predict low frequency S_{11} by using

$$R_{in} = \frac{2R_F}{S_{21}(total)} - R_S \tag{4.34}$$

The $36.5dB$ post layout S_{21} measurement of the core LNA indicates an input impedance of $63ohms$ which corresponds to low frequency S11 of $-16dB$ and input impedance of 70 ohms. It can be seen in the simulations that S_{11} is close to this value. On the silicon S_{11} will be directly measured. This was just to elucidate that the numbers make sense. Hence from pre-layout to post-layout default S11 worsens a bit. Also reverse isolation worsens due to increase in parasitic C_{gs} and C_{gd} of the transistors. The default value of the feedback resistor was not changed from pre-layout default value of 4K ohms because the S11 was still low enough

$\leq -15dB$ and their are enough controls to trim the S11 if it degrades. The gain is also expected to increase at 4K due to increase in the output impedance of the transistors. Hence the input impedance is selected a bit higher so that their is a possibility that matching can be achieved without digital trimming. The reverse isolation and buffer impedance are still low enough which is required for good stability and output match.

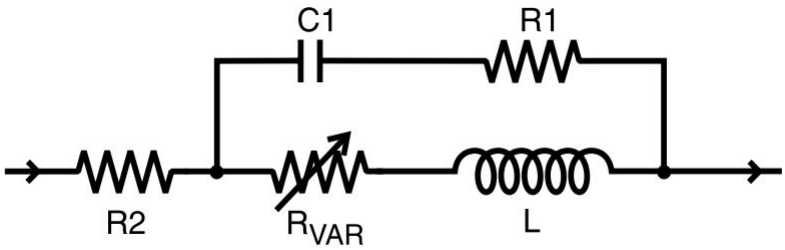

Figure 4.26: Model of the gate inductor including the parasitics

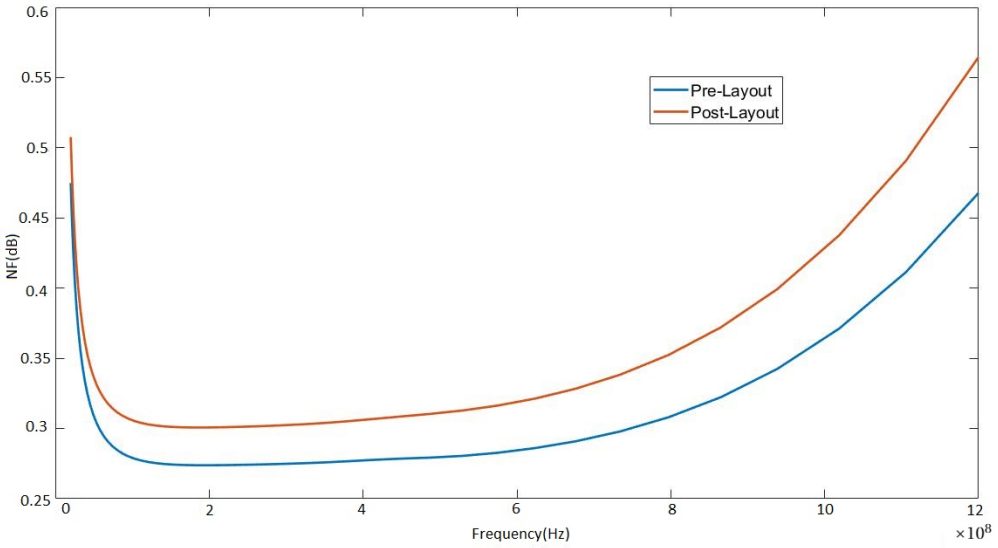

Figure 4.27: Room temperature post and pre layout noise figure simulations

Figure 4.27 shows both pre and post layout noise figure simulation results. The simulations results are at room temperature or 27°C. The gate inductance used in the simulations is a 6.6nH SMD coil-craft inductor. The package parasitics of the gate inductor are taken from coil craft website. The major contributor of the noise is the dc series resistance ($R2$) and the frequency dependent resistance which models the skin effect (R_{var}). The low frequency(100M-500M) noise figure at room temperature is close to 0.28dB or 20K. About 10-11K contribution comes from transistors. The effective input referred noise contributed by feedback resistance(R_S/A_v) is close to 1 ohm or 6K (300/50K). The rest comes majorly from gate inductance and some from bias resistance. As frequency increases the noise contribution of transistors and feedback resistor increases due to the input pole. The skin effect also increases the effective series resistance of the inductance and thus increasing noise figure at high frequencies. The pre-layout noise temperature at 1GHz is 25K or 0.35dB. The low-frequency post layout noise figure is close to 0.3 dB which is 26K. The 1K increase in noise temperature is primarily due to gate resistance. In the layout, both the drain and source connections of all the fingers of both the PMOS and NMOS transistors in the first stage is made in metal 7 to decrease the voltage drop. This is because at room temperature first stage is expected to carry about 55mA of current. Metal 7 is the topmost thick metal layer in the tsmc-40nm process and has least resistance. Hence due to routing constraints, the gate connections to the gate of main transconductors of PMOS and NMOS transistors are made in metal thick which has lower thickness compared to metal 7 and hence has higher trace resistance. 1K in-

crease in the noise temperature is primarily due to 0.16 ohm of effective gate resistance that is added post layout. At 1GHz post layout noise figure is 0.42dB or 28.5K which 3.5K higher than in pre-layout simulations. This can be accounted by the increase in the parasitic capacitance at the input by approximately 300f. This is close to 10% increase from prelayout case.

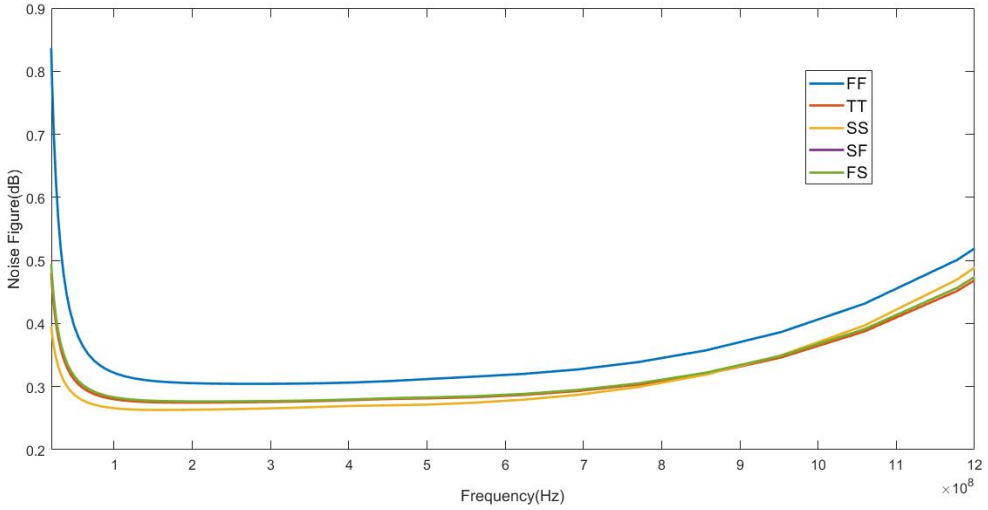

Figure 4.28: Figure showing process spread in noise performance of the LNA for constant transconductance(gm)

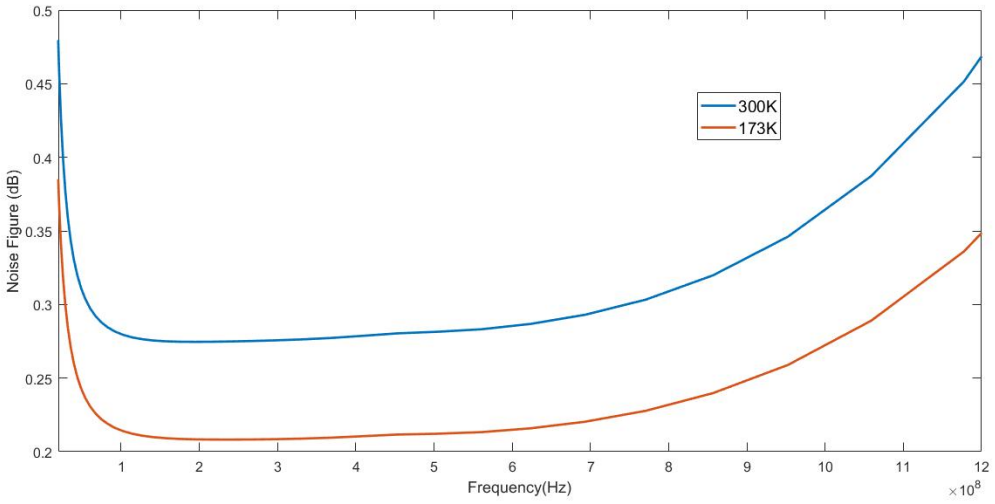

Figure 4.29: Simulations showing reduction of Noise-Figure at low temperature

Figure 4.28 shows the pre-layout corner simulations. The simulations are done keeping the transconductances of all the transistors constant by trimming the constant gm bias circuit. In other words, the gm of the first stage, balun and the gain stages are trimmed to their reach typical values across the process. The Noise figure can be seen relatively constant for constant transconductance except at the fast-fast corner where the noise temperature increases by 1-2 K. Figure 4.29 shows the reduction in noise as one goes from 300K to 173K.

4.7. Post layout simulation summary

The following table shows post layout simulation summary showing different performance parameters at room temperature.

Noise Temperature (200MHz)	20K
Noise Temperature (1GHz)	29K
S_{11}(dB)(100MHz-1GHz)	<-15dB
S_{22}(dB)(100MHz-1GHz)	<-20dB
S_{21}(100MHz-200MHz)	-52dB
S_{21}(3dB)	1.1GHz
S_{12}(100MHz-1GHz)	<-80dB
Power Consumption(1st stage)	60mW
Power Consumption(Balun)	7mW
Power Consumption(Core LNA)	67mW

4.8. Expected cryogenic performance

Out of 20K noise temperature of the LNA at 200MHz about 10K comes from the transistors 3.8. Again assuming constant gm works at 4K and mobility increases by a factor of 2 thus using the equation

$$gm = \sqrt{2\beta I} \tag{4.35}$$

One can see that the current needs to decrease by a factor approximately 2 to keep constant gm. Hence the power consumption of the LNA is expected to be close to 34-35mW. The shot noise component is expected to be half hence the noise temperature at 200MHz is expected to be close to 5K. At 1GHz since the simulated noise transistor contribution is 12K the 4K noise is expected to be close to 6K. The gain is relatively unknown but is expected to increase due to the increase in output impedance.

5

Measurements

5.1. Measurement Setup

Figure 5.1 and 5.2 shows the micrograph of the chip in TSMC 40nm CMOS and the area of the core LNA.

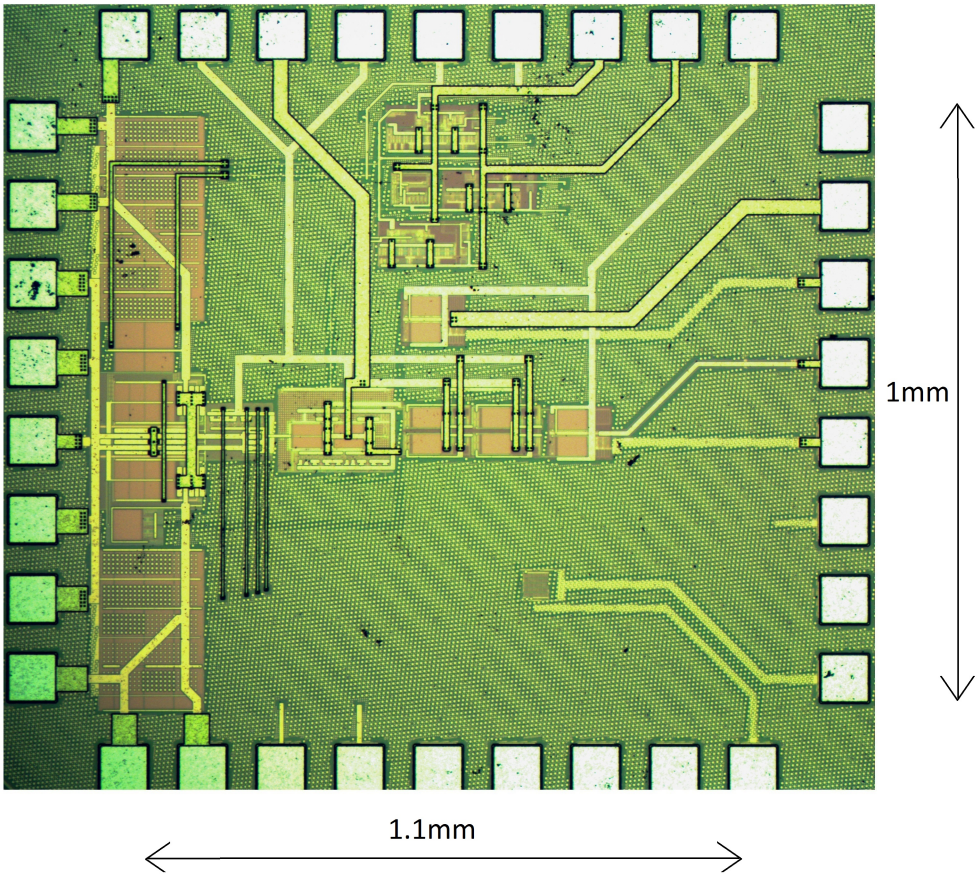

Figure 5.1: Micrograph of the Chip.

Figure 5.3 shows the top level measurement setup. The benchtop power supplies are very noisy switching supplies. In order to mitigate the noise coming from supplies, LDO boards made of ultra-low noise LDO's are desined and fabricated. The noise power spectral density achieved by these low noise LDO's is of the order of $1nV/\sqrt{Hz}$ at frequencies beyond 10MHz. As can be seen in the top-level design of the a LNA in the figure 4.8 any supply noise could appear at output through PMOS M3. Hence, care has to be taken that supply

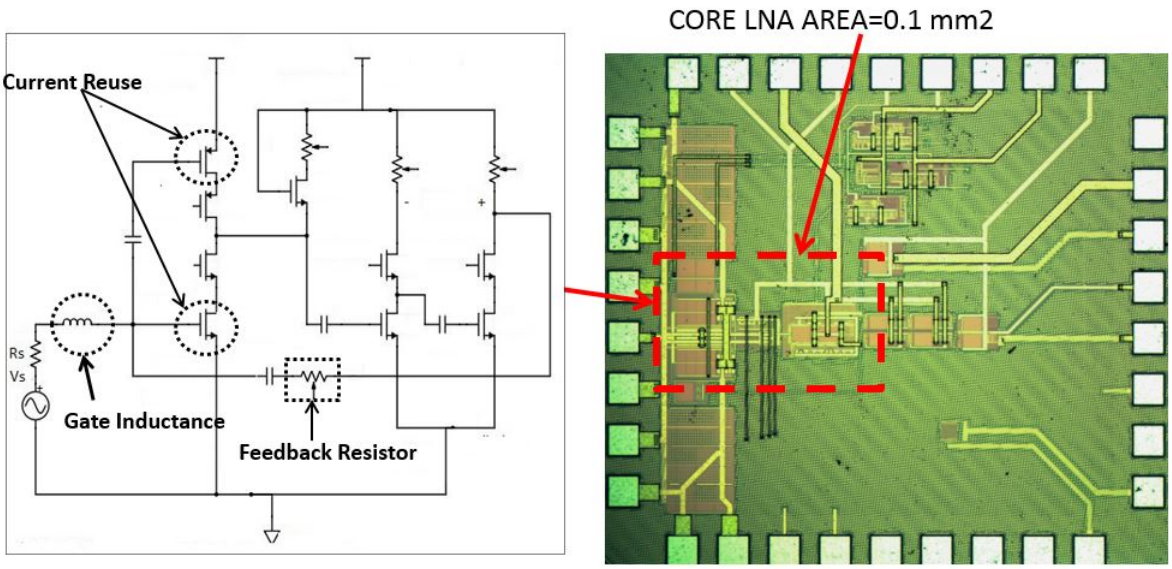

Figure 5.2: Figure showing the LNA core with it's area

noise is as low as possible. Some other measures are also taken in the PCB design to suppress the supply noise which will be discussed later in this chapter. The low noise LDO's which are used for board design are fixed LDO's LP5912 and programmable LDO's TPS7A9001DSKR by Texas Instruments. As can be seen in the figure 5.3 VNA (Vector network analyzer) is used for S parameter measurements. Rhode and Schwartz spectrum analyzer having a resolution of -167dbm/Hz is used to measure power spectral density needed for noise figure measurements. The SPI for digital control is controlled using the Nexys 4 DDR Artix-7 FPGA development board.

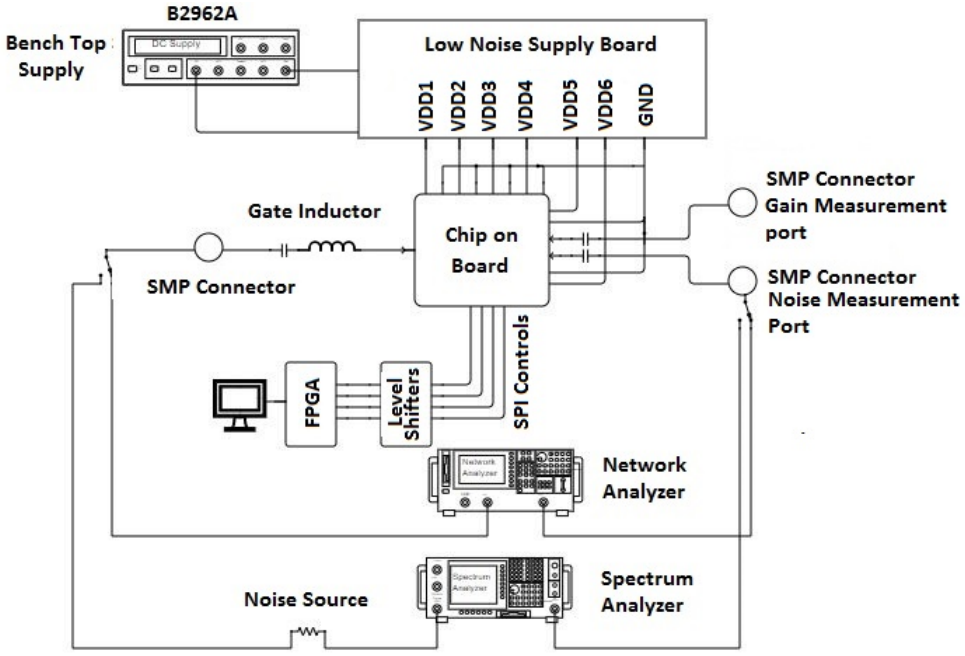

Figure 5.3: Schematic of the measurement setup used for room temperature measurements.

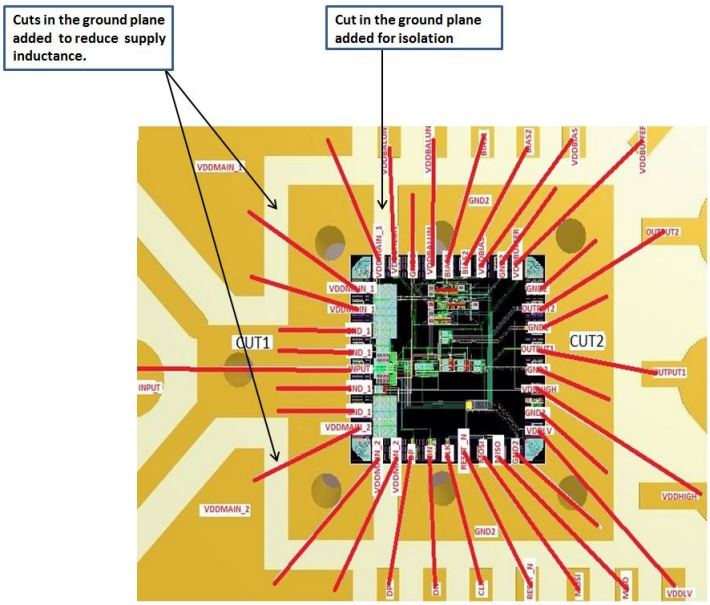

Figure 5.4: Bonding diagram of the chip on the PCB. The ground plane on the first layer is cut to reduce coupling between the first stage and succeeding stages.

5.2. PCB design

The PCB is made using FR-4 dielectric four layer board with a 150um pitch which was fabricated by eurocircuits. In the first layer signal routing is done and the components are placed. Most of the signal and power routing is done using this layer. All the SMD (Surface mount) components which include connectors, capacitors, and the gate inductor are placed using this layer. The second layer is the ground plane. The third layer is used as a power plane for power routing. The final layer is the ground layer acting as an interference shield from the bottom side. The four-layer board not only simplifies routing but also the distance between the ground plane and the first layer is much less in a four-layer board. This not only leads to more coupling between the ground and the signal but also reduces the trace width of the line such the SMA to line termination is gradual. As discussed in the previous section since the first stage design is a single-ended amplifier, it undergoes degeneration. Hence care has to be taken to minimize bond wire inductance. The small ground plane on which chip is placed is split into two. The first acts as the ground for the first stage and second acts as the ground for all succeeding stages, bias circuits and digital circuits. This is primarily done to reduce interstage coupling since this ground plane is not ideal due to due to multiple holes due to vias and the chip placed on top of ground plane. The vias of this split plane are connected to the ideal ground plane in the second layer. Cuts in the first layer ground plane are also added to reduce the supply inductance of the first stage. This can be seen in the bonding diagram. 50Ω SMP connectors is used at the input and output for S parameters and noise measurements. The chip is mounted on the ground plane using the chip on board method to reduce bond wire inductance to the ground as seen in figure 5.4. 50Ω CBCPW(Conductor backed Coplanar waveguide) lines are added both at input and output. The coupling between the signal and ground is more in CBCPW as compared to microstrip [57]. Also, the CBCPW has lesser loss as compared to microstrip as a larger portion of the fields also travels through the air rather than the lossy dielectric. The first stage is susceptible to power supply noise. The ferrite beads having greater than 70Ω impedance from 100MHz to 1GHz are used in supply path to mitigate this. DC impedance of the ferrite beads is approximately 0.02Ω which ensures low resistive drop. Figure 5.5 shows how the ferrite beads can help in power supply noise suppression. Decoupling capacitors of varying self-resonance frequencies are used to provide low impedance paths to the current loops. Capacitors of the type CG0/NP0 and tantalum are used since they are proven to be working at 4K temperature [56]. Also, an RF interference shield is added on the PCB to suppress RF interference. Figure 5.6 shows the complete PCB. Room temperature measurements are made using a breakout board.

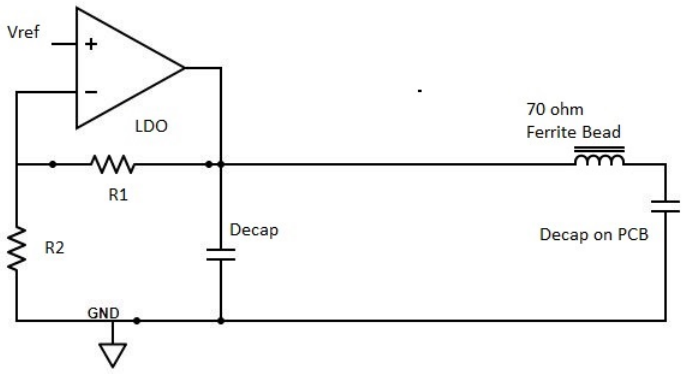

Figure 5.5: Ferrite beads connected after the LDO to filter power supply noise.

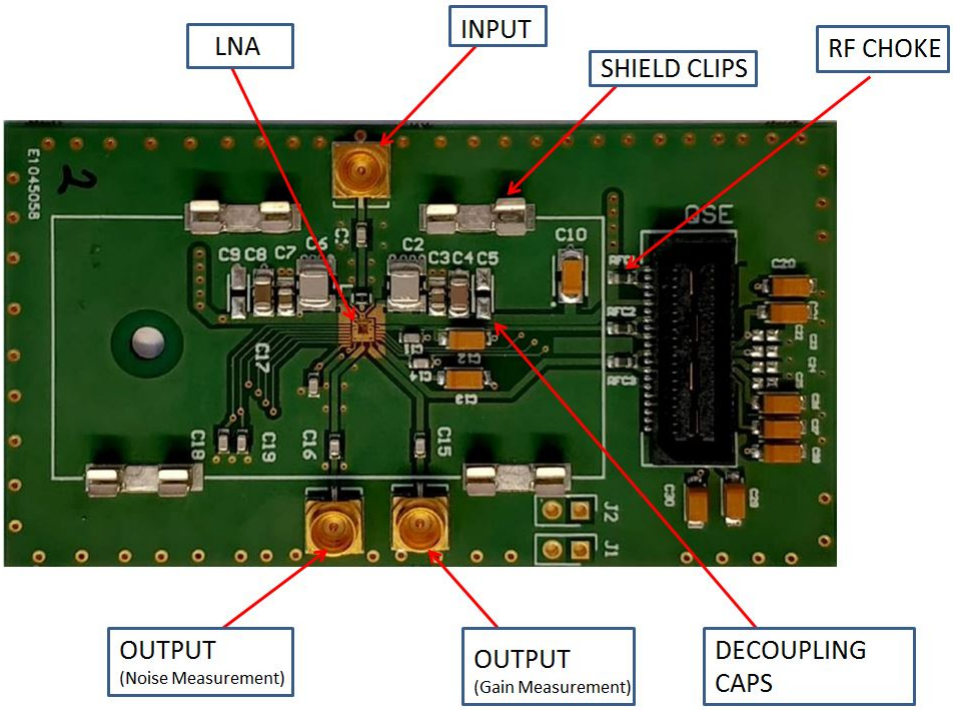

Figure 5.6: Complete LNA PCB board.

5.3. Measurements setup at room temperature

The most commonly used methods for measuring noise figure in the low noise amplifiers are Y factor method and the cold noise method[38, 58–60]. The measurement of S_{21} is made before the measurement of the output noise in the cold noise method. The output noise is given by

$$P_{noise} = K(T_O + T_N)|S_{21}|^2 \tag{5.1}$$

Here T_o and T_N are noise temperature of the resistor and the LNA respectively. In the cold noise measurement first the output noise power spectral density is measured and then the $|S_{21}|$ is measured using the spectrum analyzer. Accurate measurements of S_{21} needs large amount of calibration and also depends on the input power level and thus inaccurate. Hence Y factor measurement method is used. This involves measuring noise spectral density at two different input noise levels called as hot and cold points. For room temperature measurements low ENR(excess noise ratio) noise source is used. The ratio of the output power spectral density at the hot and cold point is called the Y factor. The ENR of the noise source varies from 5.46 at 100MHz to 5.29 at 1GHz. The impedance change from hot to cold point can be significant in high the ENR noise source. Hence it is not used or taking the measurements. The hot point noise of the noise source is given by

$$T_H = 290 \times (10^{ENR/10}) + T_O \tag{5.2}$$

Here T_O is the measured ambient temperature. The cold point noise of the noise source is given by

$$T_C = T_O \tag{5.3}$$

Here T_O is the ambient temperature at which noise source is kept. A small SMA cable with loss factor L_{Cable}(L<1) is used to connect the noise source and the LNA. The noise of the cable is given by the expression [58].

$$T_{Cable} = T_O(1 - 10^{-L_{Cable}(dB)/10}) \tag{5.4}$$

L_{Cable} is related to it's dB value by the following expression

$$L_{Cable} = 10^{-L_{Cable}(dB)/10}) \tag{5.5}$$

The effective hot point T_{HEFF} and cold point T_{CEFF} noise at the input of the LNA is given by the expression

$$T_{Heff} = T_H L_{Cable} + T_{Cable} \tag{5.6}$$

$$T_{Ceff} = T_C L_{Cable} + T_{Cable} \tag{5.7}$$

The Y factor is given as the ratio of power spectral density at hot and cold points and is independent of the gain. It is given by the following expression

$$Y = \frac{T_{Heff} + T_N}{T_{Ceff} + T_N} \tag{5.8}$$

The noise temperature of the LNA can be related to the Y factor as given in [58] as

$$T_N = L_{cable}\frac{(T_H - Y\,T_C)}{(Y-1)} - T_{Cable} \tag{5.9}$$

Instead of the VNA, the Y factor method can be used to calculate the gain. The gain or S_{21} of the LNA can be calculated after the measurement of noise temperature T_N. This is the actual gain seen by the signals of small amplitude. The input to the LNA cannot be less than -60dBm if the VNA is used. Since the circuit is designed to detect much lower signals it is better to predict the gain using the Y factor method otherwise lower gain will be measured due to gain compression.

$$P_{cold} = K T_{Ceff} |S_{21}|^2 \tag{5.10}$$

The effective gain of the LNA is given by subtracting $|S_{21}|$ by the loss of output cable.

5.4. Room Temperature Measurements

The figure 5.7 shows the S_{11} of the LNA with the default trimming. The low frequency S_{11} of the LNA is poor and close to -9dB. The S_{11} improves by decreasing the feedback resistor indicating that input impedance in much higher than 50Ω. It can be easily seen that 100Ω input impedance gives the S_{11} close to -9dB.

To improve the S_{11} of the LNA, the gain of the first and the second stage each is increased by approximately 1.25 times with respect to default gain from default trimming. As can be seen from the figure 4.9 segmented resistor DAC is used as the load resistor for the first stage. The same is used for second stage balun. By default the load resistor of R and 4R are connected in parallel equivalent to equivalent load resistance of 0.8R. Now, 4R resistor is switched off giving equivalent load resistance of R. Hence the gain of both the first and the second stage is increased by 1.25 times. The feedback resistor 3KΩ is selected instead of the 4KΩ resistor. The input impedance now should be close to $(0.75 \times 100)/(1.25 \times 1.25)$ or 48Ω. The 0.75 factor comes from decreasing the feedback resistor from 4KΩ to 3KΩ and division factor 1.25×1.25 comes from increasing the gain of the first stage and balun by trimming the load resistor. The whole point is to get almost the same gain as in the case of the simulations. 48 ohms corresponds to S_{11} of -34dB at low frequencies. The figure 5.8 shows the S_{11} of the LNA after trimming. The figure 5.9 shows the measured S_{22} of the output buffer used for the noise measurements. The measured S_{22} is less than -15dB in the band of interest that is 100MHz to 1GHz.

Figure 5.10 shows the observed noise temperature of the DUT. The minimum noise temperature that the LNA reaches is close to 27K at 200MHz and increases to 47K at 1GHz. The major uncertainty in the

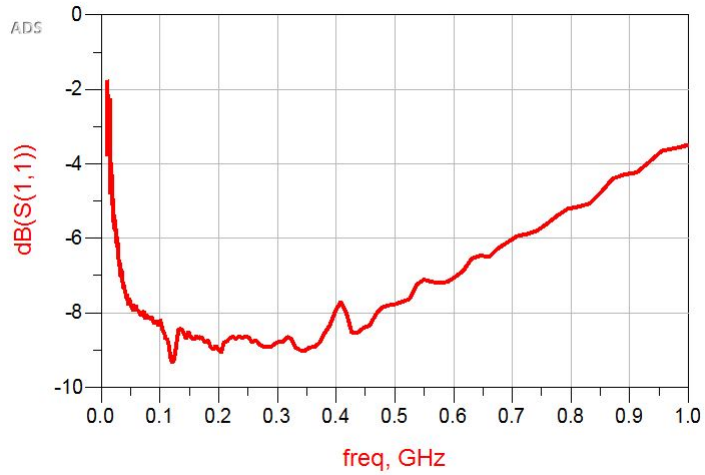

Figure 5.7: Figure shows poor untrimmed S11 of the LNA. This is due to lesser gain of the core LNA than in the simulations.

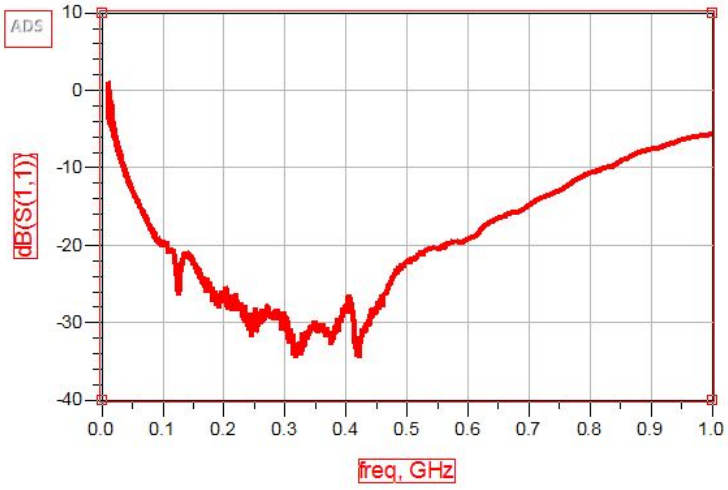

Figure 5.8: S11 of the LNA after increasing the gain of the core LNA and decreasing the feedback resistor.

Figure 5.9: Measured S22 of the output buffer used for the noise measurements.

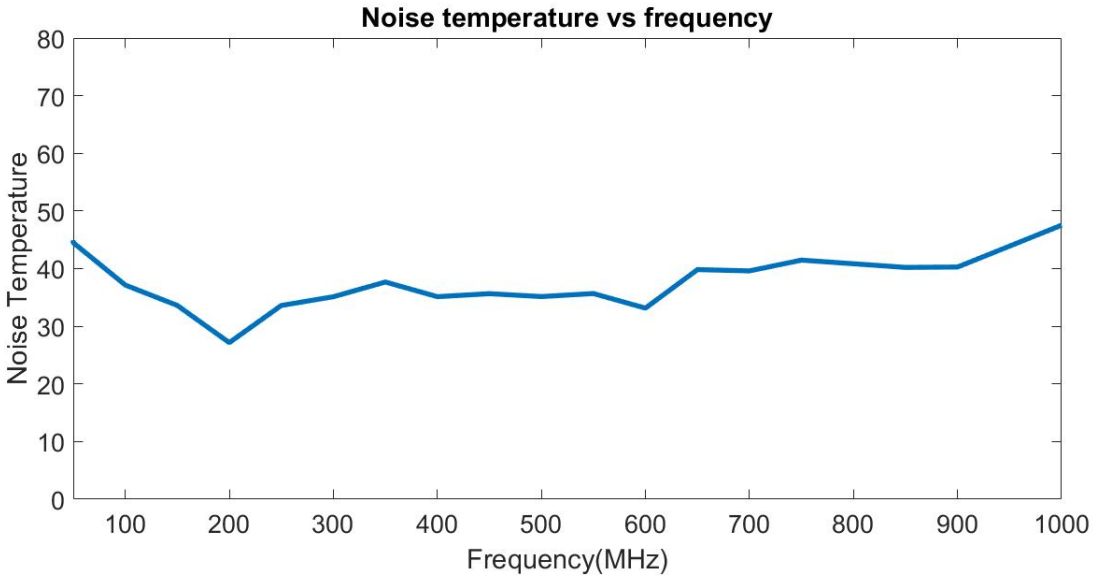

Figure 5.10: Measured LNA noise temperature.

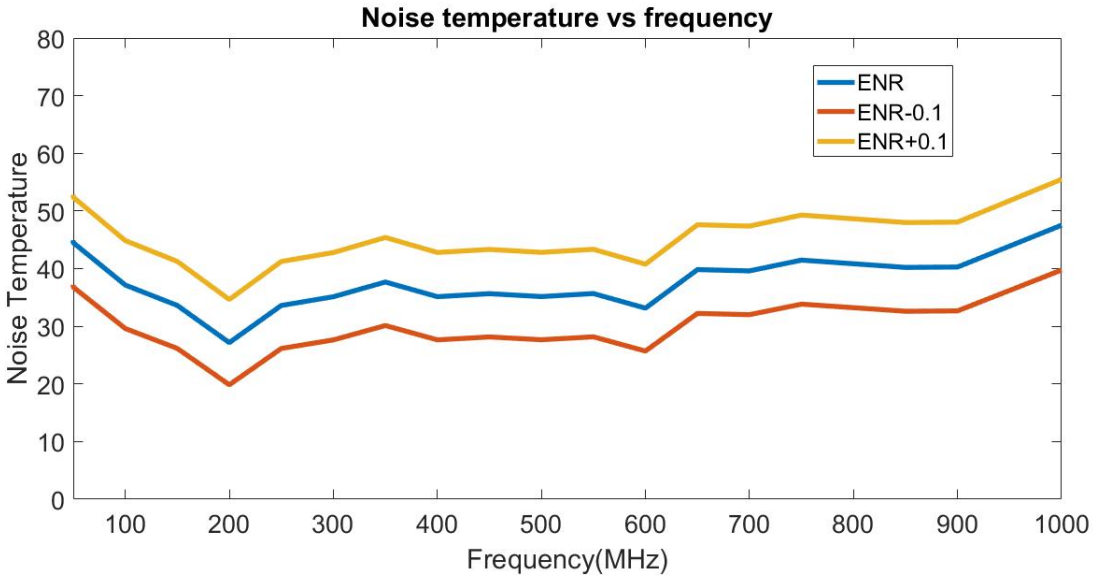

Figure 5.11: The figure shows how the uncertainty in the ENR results in the noise temperature uncertainty.

measurement arises from $\pm 0.1dB$ uncertainty in the ENR which can result in an approximate error of ±7K in the noise temperature as shown in figure 5.11. Figure 5.12 and 5.13 shows the comparison between the simulated and measured gain and noise temperature using the Y factor method. The S_{21} in the figure 5.2 is 0.7-0.8dB less than in simulations which primarily is due to the reduction of LNA gain. The simulations are compared at different setting than in simulations. The point of doing this is to compare when the gain is almost the same as that in simulations. The S_{21} contribution of the core LNA at lower frequencies can be calculated by the formula $2R_F/(R_S + R_{in})$. In the simulations the feedback resistor used was 4K and the low-frequency input impedance(100MHz) was close to 70 ohms giving S_{21} contribution of 36.5dB. While here the feedback resistor is close to 3K with the input impedance of 48 giving S_{21} contribution as 35.4dB. This is just an rough estimate and can be compared to the simulations in the figure 5.17. This approximately explains the decrease in S_{21}. The approximate 3dB bandwidth of the noise chain is close to 1GHz. The noise temperature is higher than that simulations. The most probable reasons could appear to be PCB loss. PCB loss of just 0.1-0.2 dB can add up to 10K of thermal noise. Figure 5.17 shows the directly buffered gain of the LNA measured using VNA. The bandwidth of the core LNA is higher than the bandwidth measured by the noise chain since

it also takes into account post-amplification amplifiers. Now 2dB bandwidth is close to 1GHz. Figure 5.15 shows noise temperature measurement results using 3dB attenuator at the input. This method is usually employed to see the impact of poor input match [58]. Poor input match can cause noise waves to be reflected back from the DUT. These noise waves can be again reflected from the source especially at hot point when the source impedance changes. The multiple reflections can lead to error in the noise measurement especially at higher frequencies where matching is not good. The measurement measures minimum noise temperature of 28K at 200MHz which reaches to about 50K at 1GHz. This is close to the one observed without attenuator. In all the measurements above gate inductor is 3.4 nH. It was assumed that bond wire will add about 2nH gate inductance and 500 pH will be added by the coupling capacitor. 500pH was calculated based on self the resonance frequency of the capacitor. Figure 5.16 shows the impact of gate inductor on noise temperature of the DUT. 6.6 nH gate inductor shows lower minimum noise temperature at 200MHz close to 20K. However, at high frequencies the performance is almost same as 3.4 nH gate inductor. 6.6 nH inductor adds more series and skin resistance but also reduces noise contribution of the transistors at high frequency. Two effects taken simultaneously could be resulting in a negligible change in the noise temperature of the DUT when the inductor value is changed. The power consumption of the core LNA was measured to be 60mW. Out of this 54mW is the power consumption of the first stage and approximately 6mW is the power consumption of the balun.

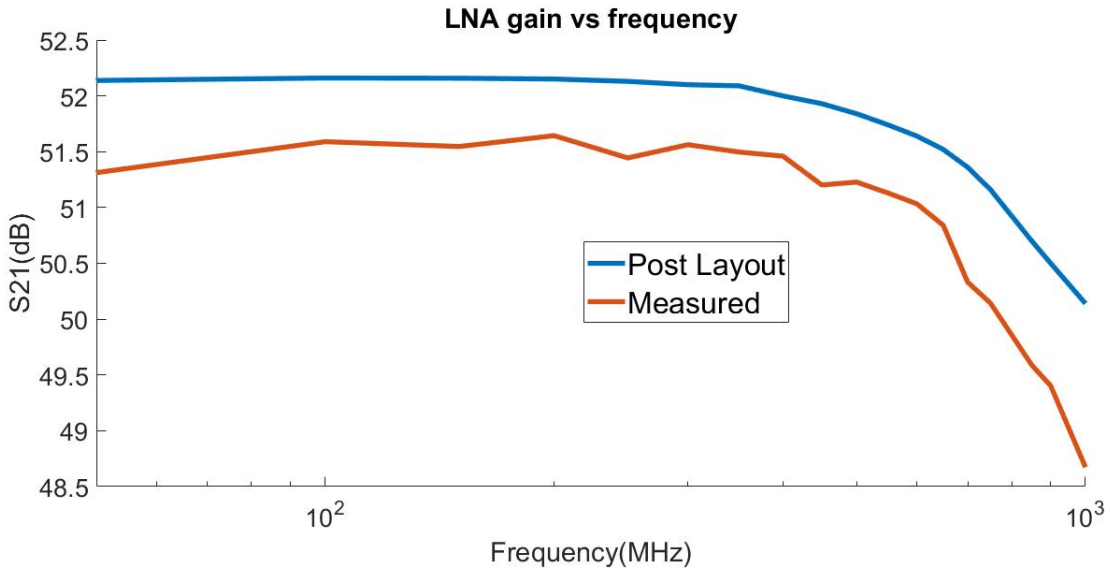

Figure 5.12: Plot showing the total measured S_{21} of the complete chain for noise measurements against the simulated.

5.5. Noise figure measurements at cryogenic temperatures

Cryogenic noise measurements involve dipping the sample in the liquid helium. Figure 5.17 shows how the setup works. For cryogenic noise temperature measurements, a 20dB attenuator is used at the input to provide hot and cold point noise at the input of the LNA. The measurement involves using long cables hence to do accurate measurements one need to calibrate the loss of the cables. Following procedure was used to derive the noise temperature of the LNA. First, the loss of both the input and the output cable which is used to connect the input and output of the LNA is calibrated. The input and output cables are shorted at the sample holder level. Then a factor of half is taken to predict the input or output cable losses. Before taking the measurements at the room temperature, the losses were measured individually at the room temperature to ascertain that the losses of input and the output cables are approximately equal. The second step involves measuring total loss after connecting the cryogenic attenuator. The loss of the attenuator and the input cable is then derived using these two data points. ENR uncertainty does not result in any significant measurement error as the noise of the noise source is attenuated at the LNA input. The hot point noise at the input of the LNA is given by

$$T_{Heff} = T_H L_{atten} L_{cable} + T_{Cable} L_{atten} + T_B \tag{5.11}$$

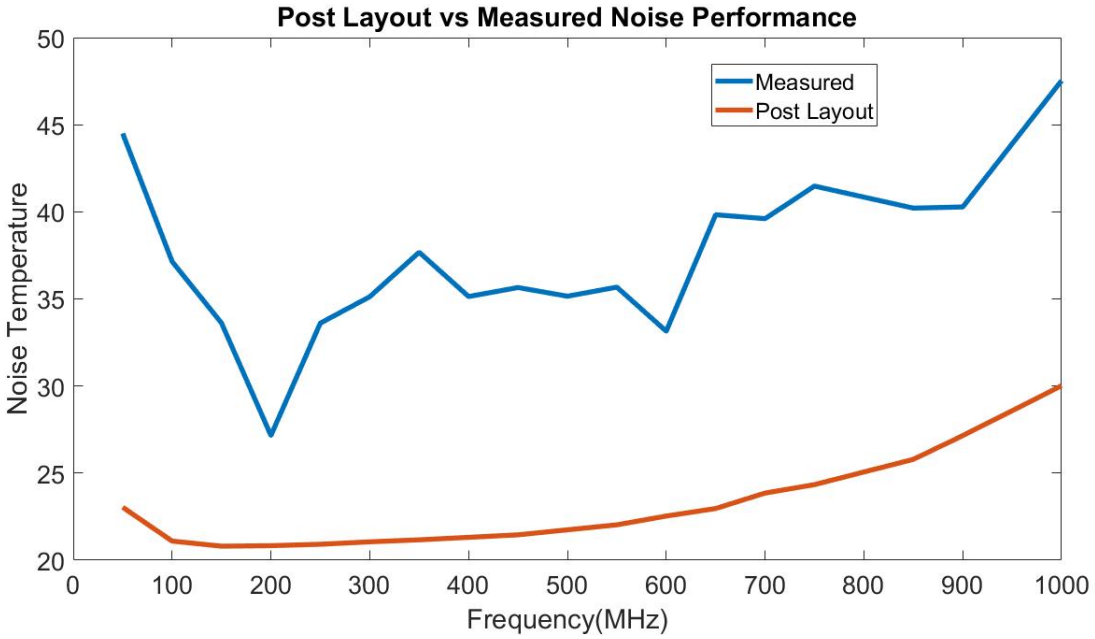

Figure 5.13: Plot showing the comparison of the simulated and the measured noise temperature of the LNA.

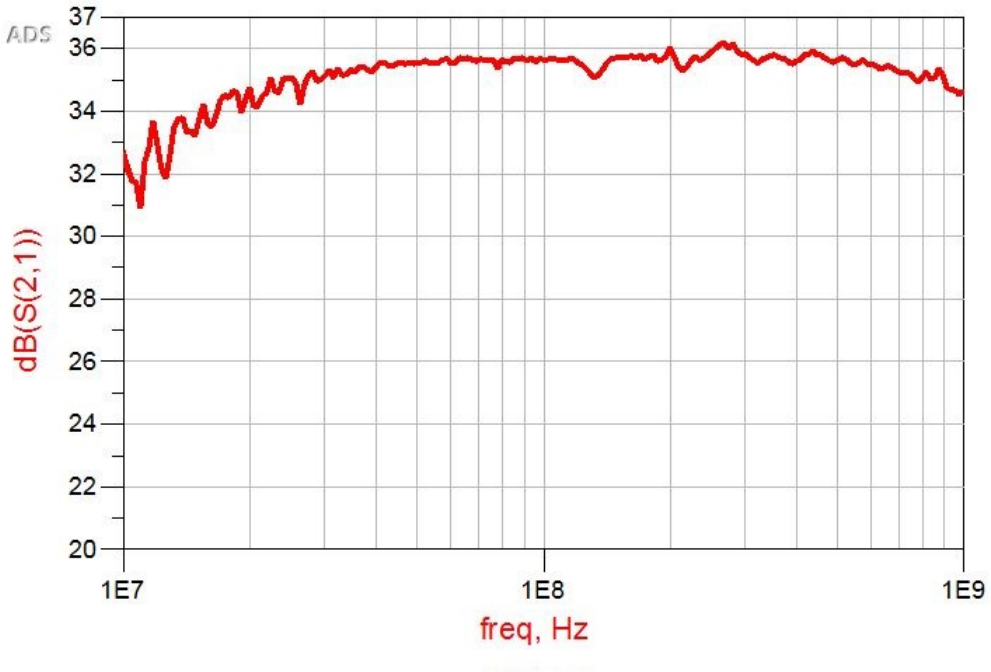

Figure 5.14: Measured gain or S_{21} of the LNA using the output buffer for gain measurements.

Here T_H is the hot point noise temperature of the noise source. T_B is the base temperature of the DUT where the attenuator is placed. L_{atten} and L_{cable} is the attenuation of the attenuator and the loss of the cable respectively. T_{Cable} is the thermal noise of the lossy cable which is given by the equation 5.4. The cold point noise temperature is given by

$$T_{Ceff} = T_C L_{atten} L_{cable} + T_{Cable} L_{atten} + T_B \tag{5.12}$$

Here T_C is the cold point noise temperature of the noise source. The noise temperature of the cable is selected midway between the base temperature of the source and room temperature to 150K. Since the noise of the

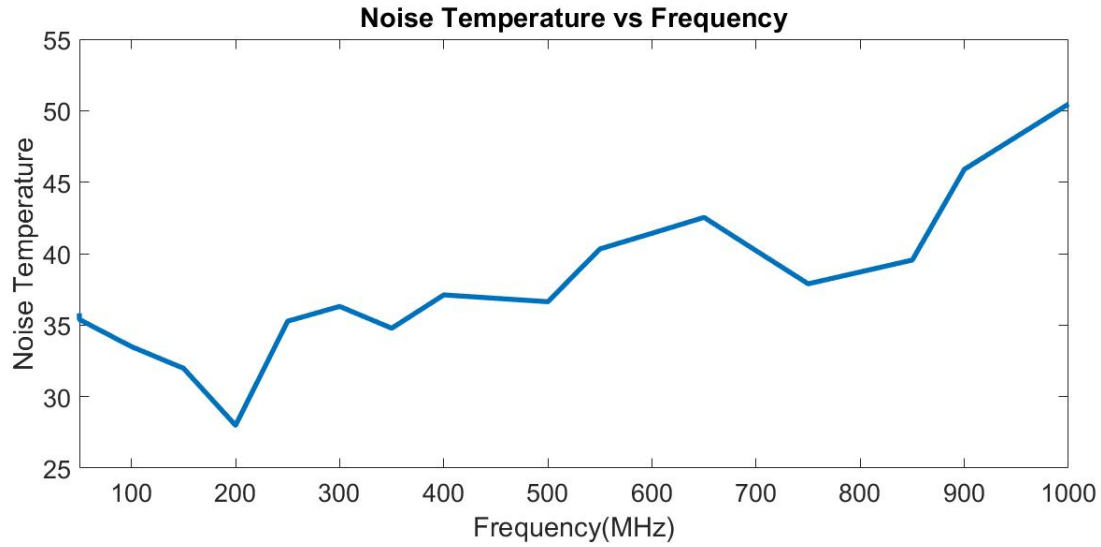

Figure 5.15: Noise temperature measurement using the 3dB attenuator at the input to see the impact of multiple reflections of the noise waves.

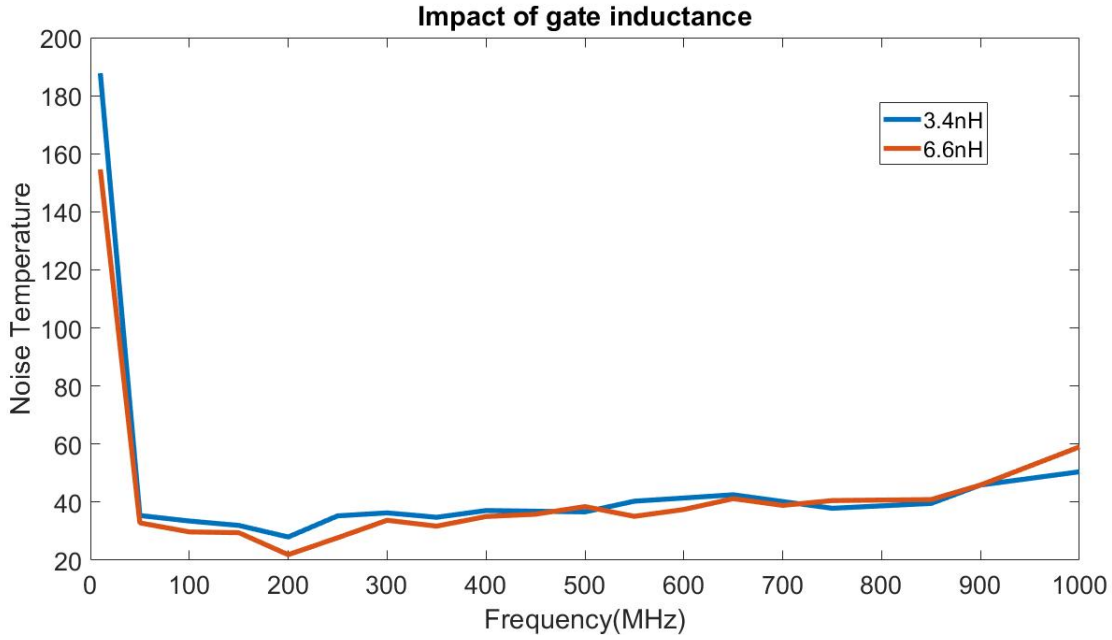

Figure 5.16: Impact of the gate inductor on the noise figure measurements.

cable is attenuated its contribution is minimal hence this approximation only leads to minimal uncertainty.

5.6. Cryogenic noise measurements

The measurement of DUT noise temperature first proceeds with the measurement of the losses of the cables and the measurement of the attenuator loss. Figure 5.18 A shows the measured combined attenuation of the input and output cable with frequency. Figure 5.18 B shows the combined attenuation of cables as well as the 20dB attenuator. This data was used to derive the attenuation of the input/output cable and the attenuator. Then hot and cold point noise at the input of the LNA was calculated using equation 5.11 and 5.12.

Figure 5.18 shows the initial Y factor measurements that are taken at 5K. The noise temperature goes as low as 4K at 400MHz and increases to about 6K at 1GHz. The value of low-frequency input impedance which is RF/A_v does not affect output SNR. However, the parasitic pole effects both SNR and matching as frequency

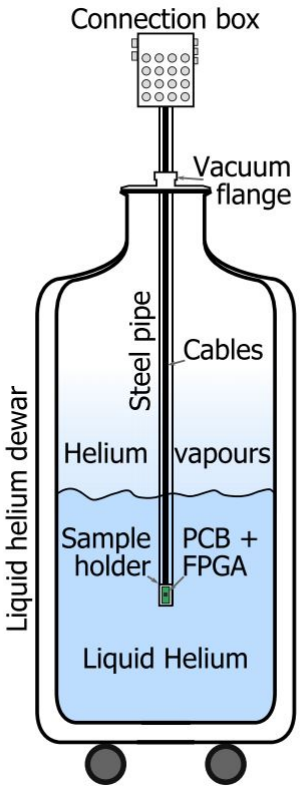

Figure 5.17: The figure shows how the LNA sample was dipped into liquid helium for the performance measurements at different base temperatures.

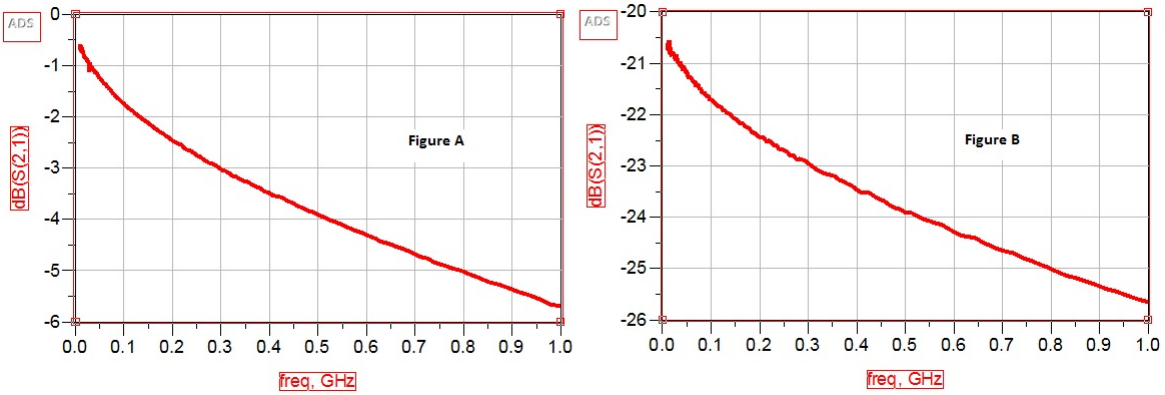

Figure 5.18: Characterization of the loss of the cable and the attenuator at 4K base temperature. Figure A shows the combined loss of both the input and output cable while figure B shows the total loss of the cables combined with the 20dB attenuator.

goes up. The impact of low-frequency input impedance on the output SNR can be seen from the below two equations.

$$idn_{load} = \frac{id}{1 + \frac{gmR_LR_S}{R_F}} \tag{5.13}$$

$$i_{Signal} = \frac{gm * Vin}{1 + \frac{gmR_LR_S}{RF}} \tag{5.14}$$

Here idn_{load} and i_{Signal} is the signal and noise current transfer to the output load. From this equation one can see that signal to noise ratio in the designed topology is unaffected by low-frequency impedance

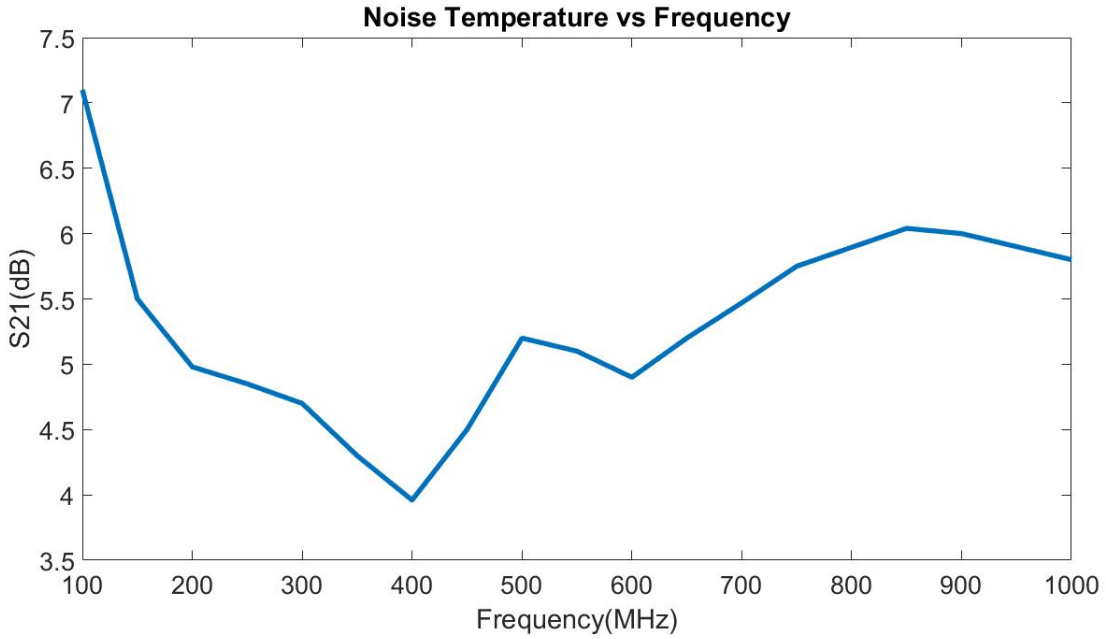

Figure 5.19: Figure shows the variation noise temperature of the DUT at 5K

match which depends on the gain of the amplifier and the feedback resistance. Hence, before the measurement of S parameters, noise measurements were done. Both the cryogenic noise measurements and S parameter measurements have to be taken again as the chip was damaged after the measurement.

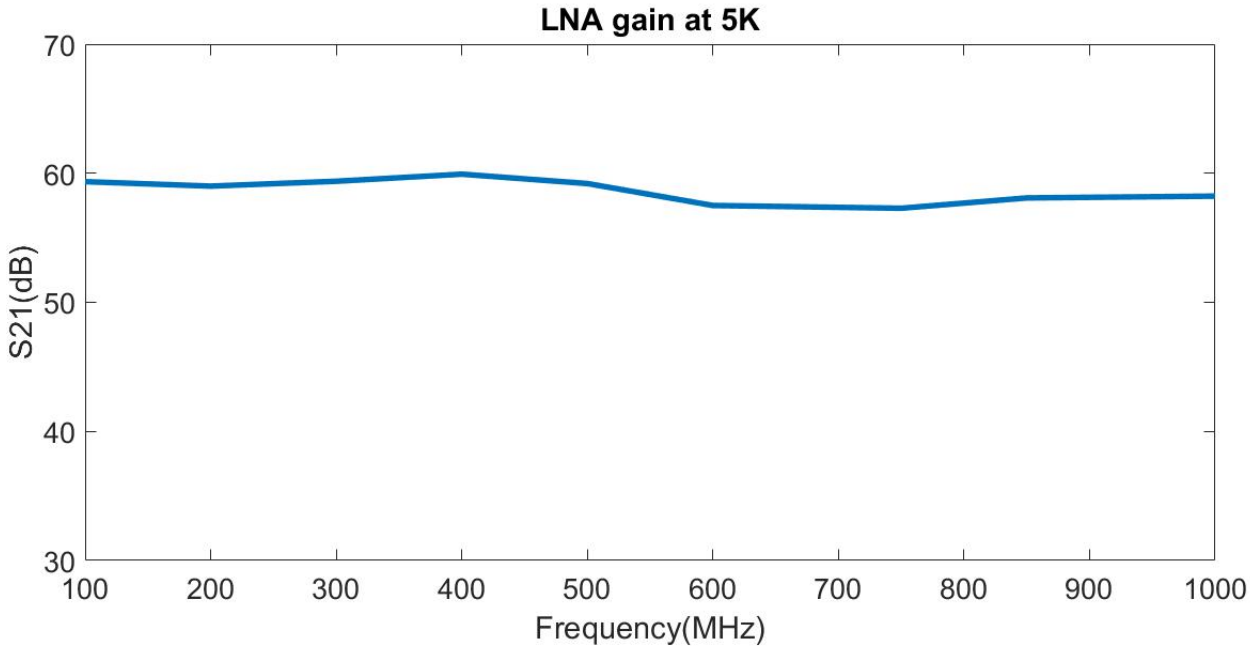

Figure 5.20: Measured S_{21} or gain of the DUT at 5K

Figure 5.20 shows the measured S_{21} of the chain used for noise measurements. The gain at 4K increased from room temperature measurement by greater than 6dB. Assuming that the transconductance is approximately the same as that at room temperature due to constant gm circuit, the most probable reasons seem to increase in the output impedance of the transistors and decrease of parasitic capacitances. Due to the increase in mobility at cryogenic temperatures the resulting overdrive voltages for keeping constant gm will be less hence resulting in more margins and higher output impedance. Also, parasitic capacitances can

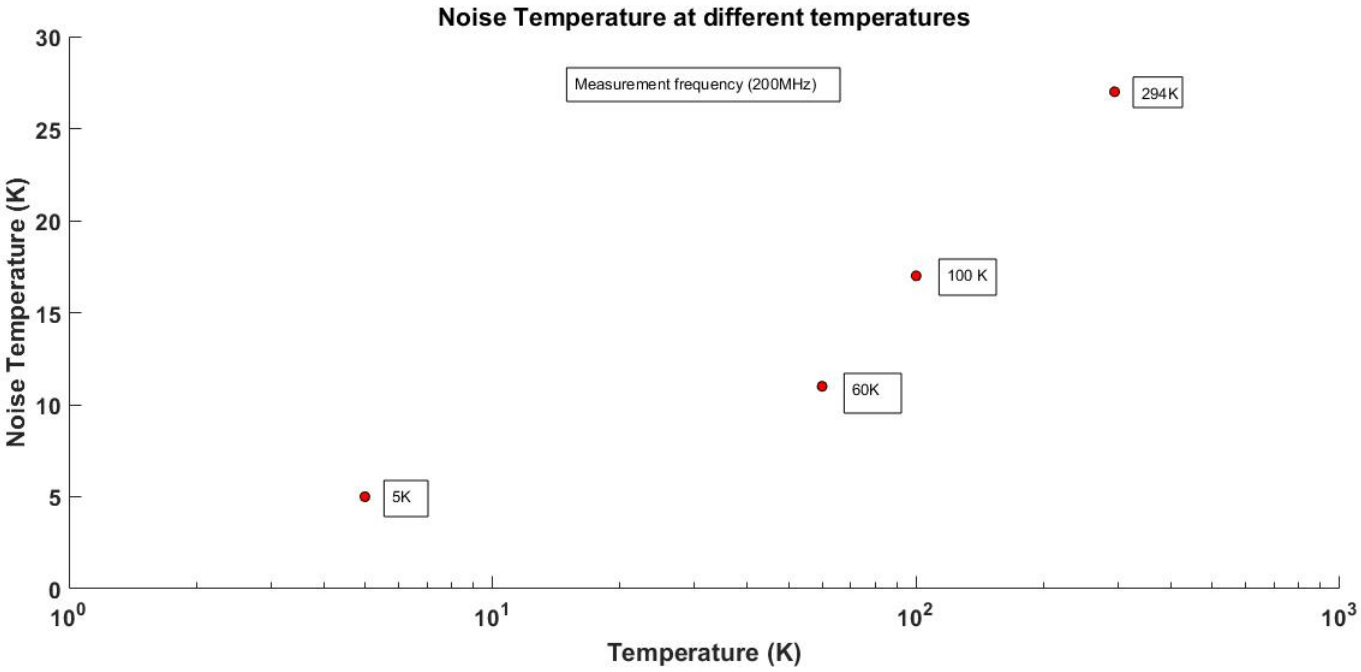

Figure 5.21: Measured noise temperature of the LNA

cause signal attenuation especially when the signal is ac coupled. Fewer parasitics mean lesser attenuation. For cryogenic measurements, SPI was not used to trim input impedance. Since at room temperature the impedance of the untrimmed sample is 100 ohms, the input impedance after the increase in the gain was expected to decrease. The sample measured uses an input inductance of 3.4 nH at the gate. The power consumption of the core LNA is about 60 percent the power consumed at room temperature i.e about 37mW. The increase in mobility enables such reduction in the current for constant gm. Figure 5.21 shows the Y factor noise measurements taken at the different temperatures. For all the cryogenic measurements i.e at 100K, 60K and 5K, the 20dB attenuator is used at the input to provide hot and cold point noise. Using attenuator also subdues the problem of multiple reflections if the matching is poor at the lower temperatures and enables accurate noise temperature measurements. For 100K and 60K measurement 15dB, ENR noise source is used since 5.46 dB ENR source is not enough to make accurate noise measurements. This is because hot and cold points are very close to each other if low ENR noise source is used. The decrease is steeper from 293K to 60K. This is because most of the noise at room temperature had thermal origins. At 60K transistor shot noise dominates the thermal noise from other sources. For a factor 12 change in temperature from 60K to 5K the noise level decreases only by a factor of 2. The transistor drain noise contribution as was simulated in schematic simulations was close to 10K as can be seen in the figure 4.5. The initial noise measurements show that the noise temperature at 5K is close to half the value of the drain noise in simulations. This is close to our prediction based on the shot noise assumption keeping transconductance constant. However, as was done in the room temperature measurements the results at cryogenic temperature needs to be validated with multiple times.

5.7. Summary of Performance at Room Temperature

Noise Temperature (200MHz)	27K
Noise Temperature (1GHz)	45K
S_{11}(dB)(100MHz-800MHz)	<-10dB
S_{22}(dB)(100MHz-1GHz)	<-15dB
S_{21}(100MHz-200MHz)[Noise Measurement Chain]	-51.5dB
S_{21}(3dB)	1GHz
S_{21}(100MHz-200MHz)[Gain Measurement chain]	-35.5dB
S_{21}(2dB)	1GHz
Power Consumption(1st stage)	54mW
Power Consumption(Balun)	6mW
Power Consumption(Core LNA)	60mW

5.8. Summary of Performance at Cryogenic Temperature

Noise Temperature (400MHz)	4K
Noise Temperature (1GHz)	6K
S_{21}(100MHz-200MHz)[Noise Measurement Chain]	59dB
S_{21}(2dB)	1GHz
Power Consumption(Core LNA)	37mW

6

Conclusion and future improvements

6.1. Conclusion

In this thesis, the Resistive-Feedback LNA for the RF readout of spin-qubit was designed and taped out in 40nm CMOS. The main objective of the thesis was to assess the cryogenic noise performance of the CMOS LNA's and whether it can meet the stringent noise requirement needed to readout spin-qubits. The LNA at the room temperature achieved the noise temperature varying between 27-45K in the band of 100MHz to 1GHz with the core LNA power consumption of 67mW. The S11 was measured to be less than -10dB till 800MHz. The gain of the core LNA was measured to be greater than 42dB. The noise measurement results at room temperature are 30-50 percent off from the simulations. The most probable reasons for such a discrepancy at room temperature could be ENR uncertainty or the PCB losses which were not de-embedded from the DUT. Some initial Y factor measurements at 5K were taken to estimate noise temperature of the DUT at the cryogenic environment. Initial measurement results look promising with the LNA achieving sub 6K performance from 150MHz-1GHz bandwidth with minimum noise temperature reaching 4K at 400MHz with power consumption of only 37mW. The results are also close to the prediction made assuming channel drain noise as shot noise. These results are the lowest reported noise temperatures achieved by any CMOS LNA at cryogenic temperature with noise performance close to Si-Ge LNA. However, as the chips were damaged during the measurements S parameter measurement were not taken.

The following table summarizes the results till taken at cryogenic temperatures.

Noise Temperature (400MHz)	4K
Noise Temperature (1GHz)	6K
S_{21} (100MHz-200MHz) [Noise Measurement Chain]	59dB
S_{21} (2dB) Noise Measurement Chain	1GHz
Power Consumption(Core LNA)	37mW

The following table shows the state of the art in CMOS and Si-Ge both at room and cryogenic temperatures.

Publication	Physical temperature	Noise Temperature	Bandwidth	Technology	Power
IMTT [50]	16	2.8K	0.3-3GHz	SiGe	32mW
IMTT [51]	16	3.4-4K	2-4GHz	SiGe	3mW
IMTT [52]	18	5-8K	4-8GHz	SiGe	760uW
IMTT [53]	15	3.7-4.3	0.1-5GHz	SiGe	20mW
IEEE RWS [36]	290	$\leq 25K$ (50 ohm match)	700MHz-1.4GHz	CMOS	45mW
JSSC [51]	290	$\leq 14K$ (non 50 ohm match)	700MHz-1.4GHz	CMOS	45mW
JSSC [54]	270	$4K$-$10K$	2GHz-8GHz	HEMT	1W
JSSC [17]	4K	7K(Spot noise)	Unknown	CMOS	55mW

As can be seen from the table that the preliminary noise temperature values of the designed CMOS LNA are close to the Si-Ge ones. However at the cost of power dissipation and bandwidth. This is primarily because of the high transconductance or gm of the SiGe amplifiers. However, silicon germanium is way more

expensive process than CMOS and in the envisioned quantum processor billions of transistors will be needed for the purpose of error correction hence CMOS is the only process which allows such a complex integration.

6.2. Future Work

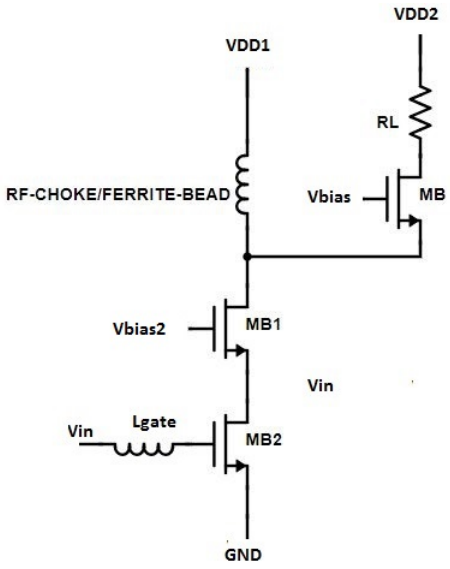

Figure 6.1: Modified first stage of Resistive-Feedback LNA showing how one can reduce power further by using NMOS only transistors

- S parameter measurements which includes measuring the matching or S11, gain of the core LNA or S21 needs to be taken at cryogenic temperatures.
- PCB line loss even though it is small should be dembedded for room temperature measurements using test structures. This is because even $0.1dB$ can lead to measurement error of 7K.
- Figure 3.11 shows the modified Resistive-Feedback LNA with passive amplification. Due to the limited scaling of drain noise of the transistor and direct scaling of the noise of passives some passive amplification using auto-transformer can be used. However, for the broadband noise performance input matching network needs to be evaluated. This is because as one goes away from the center frequency of the match, the signal and the noise transfer at the output degrades rapidly. However, the readout frequency then needs to be pushed higher. The matching networks used for matching QPC and SET's are then needed to be then evaluated at those frequencies.
- Another possibility which can reduce the noise requirement of the LNA is to design the LNA for higher impedance like 100 ohms. The input signal to the LNA as given by equation 2.1 and 2.4 and is inversely dependent on the quality factor of the matching network. The signal to the LNA will be higher by a factor $\sqrt{2}$ times and thus decreasing the noise requirement for the readout chain by a factor of two.
- The core LNA power can be pushed further down by pushing reducing the supply of the LNA 6.1. The ferrite beads used for power supply noise rejection can be used as an RF choke. Now the supply can be lowered to the limit that the cascode transistors are in saturation.

Bibliography

[1] R.P Feynman. *Simulating physics with computers,* Int. J. Theoret. Phys., 21 , pp.467-488, 1982.

[2] D Deutsch. *Quantum theory, the Church Turing principle and the universal quantum computer,* Proc. Roy. Soc. London Ser. A, 400 , pp.96-117 ,1985.

[3] P.W Shor. *Algorithms for quantum computation: Discrete logarithms and factoring* , in Proc. 35th Annual Symposium of Computer Science , IEEE Computer society press, Los Alamitos, CA, pp.124-134 ,1994.

[4] G.Brassard, I.Chuang, S.Lloyd and C. Monroe. *Quantum Computing* , in Proc. Natl. Acad. Sci., Irvine , CA, 95, pp.11032-11033 ,1998.

[5] A.M Steane. *A Tutorial on Quantum Error Correction,* Proc. Int. Sch. Phys. "Enrico Fermi", course CLXII "Quantum Comput. Algorithms Chaos", Amsterdam, 95, pp.1-32 ,2006

[6] P.W Shor. *Scheme for reducing decoherence in quantum computer memory,* Phys. Rev. A ,Vol 52, pp R2493-R2496,1995

[7] A.M Steane. *Error Correcting Codes in Quantum Theory,* , Phys. Rev. Lett.„ Vol 77 ,pp 793-797, 1996

[8] J. Kelly. *A Preview of Bristlecone, Googles New Quantum Processor,* https://research.googleblog.com/2018/03/ a-preview-of-bristlecone-googles-new.html, 2018

[9] Raymond Laflamme et.al *Perfect Quantum Error Correction Code,*arXiv:quant-ph/9602019

[10] A. Fruchtman and I. Choi. *Technical Roadmap for FaultTolerant Quantum Computing, Tech. Rep. (University of Oxford, 2016).*

[11] E. Charbon et.al. *Cryo-CMOS circuits and systems for scalable quantum computing,* in IEEE Solid State Circuits Conference ,pp 264-265, 2017

[12] B. Patra et.al. *Cryo-CMOS circuits and systems for quantum computing applications,* in IEEE Journal of Solid State Circuits ,Vol 53 ,pp 309-321, 2018

[13] J.van Dijk et al. *A co-design methodology for scalable quantum processors and their classical electronic interface* , in IEEE Design, Automation and Test conference, March, 2018

[14] D. Rotta, F. Sebastiano, E. Charbon, E. Prati. *Quantum information density scaling and qubit operation time constraints of CMOS silicon-based quantum computer architectures,* npj Quantum information, 2017

[15] J.I Colless, D.J Reilly. *Cryogenic High-Frequency Readout and Control Platform for Spin Qubits,*Review of Scientific Instruments, 023902, 2012

[16] Trais S Humble e.t.al. *Quantum Computing Circuits and Devices,* arxiv.org:1804.10648

[17] L.M.K Vandersypen et.al. *Interfacing spin qubits in quantum dots and donors—hot, dense, and coherent,* npj Quantum information, 2017

[18] R. Hanson et.al. *Spins in few-electron quantum dots,* Rev. Mod. Phy., Vol 79, pp 1217-1265,2007

[19] J. I. Colless et.al. *Dispersive Readout of a Few-Electron Double Quantum Dot with Fast rf Gate Sensors,* Vol 110, pp 046805 ,2013

[20] J.M Elzerman et.al. *Single-shot read-out of an individual electron spin in a quantum dot,* Nature , Vol 430, pp 431-435 , 2004

[21] R.J Schoelkopf et.al. *The radio-frequency single-electron transistor (RF-SET): A fast and ultrasensitive electrometer,* Science , Vol 280, pp 1238-1242 , 1998

[22] Henrick Brenning et.al. *An ultrasensitive radio-frequency single-electron transistor working up to 4.2 K,* Journal of Applied Physics , Vol 100, 2006

[23] L. Roschier et.al. *Noise performance of the radio-frequency single-electron transistor,* Journal of applied physics,Vol 95,No 3, pp 1274-1286 ,2004

[24] J.M Hornibrook et.al. *Frequency multiplexing for readout of spin qubits,* Appl. Phys. Lett 104, 103108 , 2014

[25] L.M.K Vandersypen et.al. *Real-time detection of single-electron tunneling using a quantum point contact,* Appl. Phys. Lett. 85, 4394, 2004

[26] J.M Auger. *Fault-tolerance thresholds for the surface code with fabrication errors,* Phys. Rev. A 96, 042316, 2017

[27] M. S. Obrecht, E. A Allam, and T. Manku. *Diffusion Current and Its Effect on Noise in Submicron MOSFETs,* IEEE transactions on electron devices, Vol. 49, No. 3, MARCH 2002

[28] S. Anderson and C Svensson. *Direct experimental verification of shot noise in short channel MOS transistors,* IET Electronics letters,Vol. 41,Issue 15,2005

[29] J. A. McNeill. *Noise in Short Channel MOSFETs,* IEEE Custom intergrated circuits conference ,2005

[30] M. W Pospieszalski. *On noise properties of transistors and amplifiers a critical review,* 20th International Conference on Microwaves, Radar and Wireless Communications ,2014

[31] M. W Pospieszalski. *On the Dependence of FET Noise Model Parameters on Ambient Temperature,* IEEE Radio and wireless symposium ,2017

[32] M. W Pospieszalski. *On the Limits of Noise Performance of Field Effect Transistors,* IEEE International microwave symposium (IMS), 2017

[33] D.K Shaeffer and T.H Lee. *A 1.5-V, 1.5-GHz CMOS low noise amplifier,* IEEE Journal of solid state circuits, vol 32, no 5, pp 745–759, 1997

[34] A. Bevilacqua and A. Niknejad. *An Ultrawideband CMOS Low-Noise Amplifier for 3.1–10.6-GHz Wireless Receivers,* IEEE Journal of solid state circuits, vol 39, no 12, pp 2259 - 2268, 2004

[35] A. Ismail and A.A Abidi. *A 3-10-GHz low-noise amplifier with wideband LC-ladder matching network,* , vol 39, no 12, pp 2269 - 2277, 2004

[36] L. Belostotski and J.W Haslett. *Wide Band Room Temperature 0.35-dB Noise Figure LNA in 90-nm Bulk CMOS,* IEEE Radio and wireless symposium, 2007

[37] L. Belostotski and J.W Haslett. *Sub-0.2 dB Noise Figure Wideband Room-Temperature CMOS LNA With Non-50 ohm Signal-Source Impedance,* IEEE Journal of solid state circuits,vol 2,no 11,2007

[38] F.Bruccoleri, E.A.M Klumperink and B. Nauta. *Noise cancelling in wideband CMOS LNAs,* in IEEE Solid state circuits conference,2004

[39] F.Bruccoleri,E.A.M Klumperink, D.M Leenaerts and B. Nauta. *Wideband Balun-LNA With Simultaneous Output Balancing, Noise-Canceling and Distortion-Canceling,* IEEE Journal of solid state circuits,vol.43,no.6,pp 1341-1350,2008

[40] W.H Chen, G Liu, B. Zdravko and A. M. Niknejad. *A Highly Linear Broadband CMOS LNA Employing Noise and Distortion Cancellation,* IEEE Journal of solid state circuits,vol.43,no.5,pp 1164-1176,2008

[41] J. Hong, C. Zhan, S. S. Taylor. *A 5GHz resistive-feedback CMOS LNA for low-cost multi-standard applications,* in IEEE Solid state circuits conference ,2006

[42] B.G Perumana et.al. *Resistive-Feedback CMOS Low-Noise Amplifiers for Multiband Applications,* IEEE Transactions on Microwave theory and techniques,vol.56,no.5,pp 1218-1225,2008

[43] C. W Kim, M.S Kang, P.T Anh, H.T Kim, and S.G Lee. *An Ultra-Wideband CMOS Low Noise Amplifier for 3–5-GHz UWB System,* IEEE journal on solid state circuits,VOL. 40, NO. 2, 2005

[44] E. Adabi and A.M Nicknejad. *CMOS Low Noise Amplifier with Capacitive Feedback Matching,* in Custom Intergrated Circuits Conference (CICC), 2007

[45] F. Zhang and P. R. Kinget. *Low-Power Programmable Gain CMOS Distributed LNA,* IEEE joural of solid state circuits ,VOL. 41, NO. 6, JUNE 2006

[46] J. Zhu, H. Krishnaswamy, P.R. Kinget. *A DC-9.5GHz Noise-Canceling Distributed LNA in 65nm CMOS,* in IEEE Radio Frequency Integrated Circuits Symposium, 2013

[47] S.C.Blaakmeer et.al *A wideband CMOS LNA exploiting transformer,* IEEE IEEE Radio Frequency Integrated Circuits Symposium , June 2006

[48] W. Lee and E Afshari. *Low-Noise Parametric Resonant Amplifier,* IEEE transactions on circuits and systems—I ,VOL. 58, NO. 3, March 2011

[49] W. Lee and E Afshari. *A CMOS Noise-Squeezing Amplifier,* IEEE transactions on microwave theory and techniques, VOL. 60, NO. 2, February 2012

[50] Su-Wei Chang and Joseph C. Bardin *A wideband cryogenic SiGe LNA MMIC with an average noise temperature of 2.8 K from 0.3–3 GHz,* IEEE transactions on microwave theory and techniques, June 2017

[51] Shirin Montazeri and Joseph C. Bardin *A 2-4 GHz Silicon Germanium Cryogenic Low Noise Amplifier MMIC* IEEE transactions on microwave theory and techniques, June 2018

[52] Shirin Montazeri and Joseph C. Bardin *A Sub-milliwatt 4–8 GHz SiGe Cryogenic Low Noise Amplifier* IEEE transactions on microwave theory and techniques, June 2017

[53] Shirin Montazeri and Joseph C. Bardin *A 0.1–5 GHz Cryogenic SiGe MMIC LNA* IEEE transactions on microwave theory and techniques, June 2009

[54] Kevin W. Kobayashi et.al *A Cool, Sub-0.2 dB Noise Figure GaN HEMT Power Amplifier With 2-Watt Output Power,* JSSC, oct 2009

[55] S. Nicolson and K. Phang *Improvements in biasing and compensation of CMOS opamps,* IEEE International symposium on circuits and systems (ISCAS), 2004

[56] Harald Homulle,Stefan Visser,Bishnu Patra, and Edoardo Charbon *FPGA Design Techniques for Stable Cryogenic Operation,* arXiv:1709.04190

[57] John Coonrod *Understanding PCB's for high frequency applications,* Application note, Retrieved from rogerscorp.com.

[58] *Noise Figure Measurement Accuracy: The Y-Factor Method,* Application note; literature number 5952-3706, Retrieved from Key Sight website.

[59] *10 Hints for Making Successful Noise Figure Measurements,* Application note; literature number 5980-0288, Retrieved from Key Sight website.

[60] Joe Kelly and Frank Goh *The Basics of Noise Figure Measurements for ATE,* Application note, Retrieved from advantest.com.