Teemu Nieminen

High-sensitivity correlation spectrometer for shot noise measurements

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 20.5.2015

Thesis supervisor:

Prof. Pertti Hakonen

Thesis advisor:

M.Sc. (Tech.) Pasi Lähteenmäki

Aalto University
School of Electrical
Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Teemu Nieminen

Title: High-sensitivity correlation spectrometer for shot noise measurements

Date: 20.5.2015 Language: English Number of pages: 8+51

Low Temperature Laboratory, Department of Applied Physics, School of Science

Professorship: Engineering Physics

Supervisor: Prof. Pertti Hakonen

Advisor: M.Sc. (Tech.) Pasi Lähteenmäki

Shot noise measurement is an important characterization method in nanophysics research, although shot noise correlation measurement setups have gained little attention in the literature. Most of the previous setups either utilize low frequencies around 1 MHz or rely on solely analog circuitry.

This thesis presents a novel correlation spectrometer for low-temperature shot noise measurements. The setup utilizes a four-channel PCI-E digitizer card with a sample rate of 125 megasamples per second and 14 bit resolution. The setup captures both quadratures of two RF channels at 800 MHz which is mixed down using an IQ mixer designed and evaluated in this thesis. The digitized data is processed on a graphics processing unit of a desktop computer using a program developed using CUDA C, an extension to the C programming language. The program calculates cross- and autocorrelations of the signals in real time. CUDA C was chosen over Matlab implementation due to its significantly better performance. This enables the calculation of intensity correlation and other additional quantities simultaneously with cross- and autocorrelations. In addition, the modular hardware is upgradeable if more performance is required.

The sensitivity of the system was determined by measuring a stable noise source and calculating the standard deviation from 91 measurements. The determined sensitivity of the system, 810 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$, is close to both theoretical maximum sensitivity and best sensitivity reported in the literature, both 710 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$.

Keywords: shot noise, cross-correlation, spectrometer, GPU, CUDA

AALTO-YLIOPISTO
SÄHKÖTEKNIIKAN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Teemu Nieminen

Työn nimi: Korkean herkkyyden korrelaatiospektrometri raekohinan mittaukseen

Päivämäärä: 20.5.2015 Kieli: Englanti Sivumäärä: 8+51

Kylmälaboratorio, Teknillisen fysiikan laitos, Perustieteiden korkeakoulu

Professuuri: Teknillinen fysiikka

Valvoja: Prof. Pertti Hakonen

Ohjaaja: DI Pasi Lähteenmäki

Raekohinamittaus on tärkeä nanofysiikan tutkimusmenetelmä, mutta raekohinan korrelaation mittauskytkentöjä on kuitenkin käsitelty kirjallisuudessa vain vähän. Useimmat aiemmista kytkennöistä käyttävät matalia taajuuksia 1 MHz:n ympäristössä, tai perustuvat yksinomaan analogisiin piireihin.

Tässä diplomityössä esitetään uudenlainen korrelaatiospektrometri raekohinamittauksiin matalissa lämpötiloissa. Spektrometrissa hyödynnetään nelikanavaista PCI-E -väyläistä digitointikorttia, jonka näytteistysnopeus on 125 miljoonaa näytettä sekunnissa 14 bitin resoluutiolla. Spektrometri kaappaa kahden RF-kanavan molemmat kvadratuurit 800 MHz:n taajuudella, joka alassekoitetaan tässä työssä suunnitellulla IQ-sekoittimella. Digitoitu data prosessoidaan pöytätietokoneen näytönohjaimella käyttäen C-ohjelmointikielen laajennuksella, CUDA C:llä, toteutettua ohjelmaa, joka laskee signaaleiden risti- ja autokorrelaatiot reaaliajassa. CUDA C valittiin Matlab-toteutuksen sijaan sen paremman suorituskyvyn vuoksi, mikä mahdollistaa intensiteettikorrelaation ja muiden suureiden laskemisen samanaikaisesti risti- ja autokorrelaatioden kanssa. Lisäksi modulaarista laitteistoa on mahdollista päivittää, mikäli suorituskykyä vaaditaan lisää.

Spektrometrin herkkyys määritettiin mittaamalla stabiilia kohinalähdettä ja laskemalla 91 mittaustuloksen keskihajonta. Määritetty herkkyys, 810 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$, on lähellä sekä teoreettista maksimiherkkyyttä että parasta kirjallisuudessa raportoitua mittausherkkyyttä, molemmat 710 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$.

Avainsanat: raekohina, ristikorrelaatio, spektrometri, näytönohjain, CUDA

Acknowledgments

The research work in this thesis made use of the Aalto University Low Temperature Laboratory infrastructure. I would like to thank my supervisor, Prof. Pertti Hakonen, for advice concerning this thesis. I would also like to thank my advisor, M.Sc. (Tech.) Pasi Lähteenmäki, for numerous helpful discussions and support. Finally, I would like to thank all the members of the NANO group and the laboratory staff for the general support.

Otaniemi, 20.5.2015

Teemu J. Nieminen

Contents

Symbols and abbreviations

Symbols

B	bandwidth
e	elementary charge
$\mathcal{F}$	Fano factor or Fourier transform
f	frequency
G	gain
h	Planck's constant
i	imaginary unit
k_B	Boltzmann's constant
R	resistance
S_n	power spectral density
σ_e	variance
T	temperature
T_n	noise temperature
T_s	system noise temperature
τ	integration time or time variable
V	voltage
ω	angular frequency

Abbreviations

ADC	analog-to-digital converter
ASIC	application-specific integrated circuit
C	a general-purpose programming language
CPU	central processing unit
CUDA	Compute Unified Device Architecture, a parallel computing platform from Nvidia Corp.
DC	direct current
DFT	discrete Fourier transform
DMA	direct memory access
FFT	fast Fourier transform
FLOPS	floating point operations per second
FPGA	field-programmable gate array
GPU	graphics processing unit
HBT	Hanbury Brown and Twiss
He	helium
HEMT	high electron mobility transistor
IDFT	inverse discrete Fourier transform
IF	intermediate frequency
IQ	In phase - Quadrature phase
JPA	Josephson parametric amplifier
LNA	low-noise amplifier
LO	local oscillator
MS/s	megasamples per second
PCB	printed circuit board
QAM	quadrature amplitude modulation
QPC	quantum point contact
QPSK	quadriphase-shift keying
RAM	random access memory
RF	radio frequency
RMS	root mean square
S-parameters	scattering parameters
SEM	scanning electron microscope
SMA	SubMiniature version A, a coaxial connector type

Chapter 1

Introduction

The rapid development of electronics since the beginning of the 20th century has revolutionized our societies. One example is the integrated circuit, invented in 1958 [1], which has led to numerous applications, including fast computing, mobile communication and the Internet. Since the invention of the integrated circuit, the increase in the amount of transistors on a single chip has been exponential, doubling every two years on average. This phenomenon, known as Moore's law, has maintained the growth of the digital industry and is a major reason for the significant economic impact of semiconductor industry experienced globally during the past decades [2].

However, the growth rate is expected to be lower than estimated by the Moore's law in the near future due to the decrease in the size of the transistors as an increasing number of transistors are packed on a single chip. The decrease in transistor size has its limits owing to quantum mechanical phenomena experienced when the size of the transistors approaches the scale of individual atoms and molecules [3,4].

Therefore, novel technologies are required to replace the conventional semiconductors in order to keep up with the Moore's law. Nanoelectronics is one example of these technologies, developing new materials which enable further reduction in physical size compared to conventional semiconductors. These new materials include nanowires, graphene and carbon nanotubes. While some practical nanoelectronic devices have already been realized, including nanowire and carbon nanotube transistors [5,6], the full potential of graphene and related materials is yet to be realized [7] in addition to various other materials and structures. Nanoelectronics also covers other topics besides integrated circuits, for example sensors and high-frequency circuits. A long-term goal for nanoelectronics research is a practical quantum computer which operates on information stored in the quantum state of a system in order to solve various computational problems insolvable to classical computers [8]. Another aspect is the contribution of nanoelectronics and related quantum technologies to basic research, which deepens our understanding of the physical reality. For these reasons, nanoelectronics is expected to become one of the major technological advancements of the 21st century [9].

Research in nanoelectronics involves electronic measurements of a fabricated sample device in order to characterize its functionality and the properties of the material. One important measured quantity is electronic *shot noise* which is high-frequency electrical noise induced by DC currents. Shot noise measurement can be used to characterize the electrical and structural properties of the measured device. Accurate measurement of shot noise requires low temperature to suppress thermal noise, and therefore the measurements are conducted at temperatures close to the absolute zero. Moreover, the measurement system must be sensitive enough to detect weak shot noise signals and distinguish them from the noise induced by the measurement system itself. Such measurement systems operating at radio frequencies are referred to as *radiometers* or, when noise can be obtained as a function of frequency, *spectrometers*.

Radiometers and spectrometers are standard tools in radio astronomy, where they are used for example to characterize the anisotropy of cosmic microwave background radiation. However, these systems have gained less attention in nanophysics research, although they have become more popular during the recent years (For example in Refs. [10–14]). Yet most of the recent articles concentrate mainly on the measured physical phenomenon, with little attention paid to the measurement system itself. Some presented systems rely solely on analog circuitry [15], which limits the range of obtainable quantities compared to systems with digital data processing. Nevertheless, many of the digital systems have a limited measurement bandwidth of a few megahertz [10,16]. The measurement system in Ref. [17], although wide-band and digital, is intended for studying quantum microwaves rather than shot noise. The system is also relatively complex and expensive while simpler system would be more reasonable for shot noise measurements.

This thesis presents a novel type of spectrometer circuit based on a data acquisition card and data processing on a graphics processor unit (GPU) using a program implemented on CUDA C, an extension to the C programming language [18]. The relatively simple measurement program can be modified for measuring other quantities such as intensity correlation and higher moments of noise. The use of GPU for data processing follows the recent trend of general-purpose computing on GPUs which offer significantly more computational power with similar price compared to central processing units (CPUs) [19].

The aim of this thesis is to design and implement a two-channel correlation spectrometer and measure its functionality. The sensitivity of the spectrometer is determined and compared to literature along with tests of data processing performance. This thesis concentrates on the design of the IQ mixer device and on the development of an efficient data aqcuisition and processing software. Since the performance of the existing low-temperature parts was found adequate, they were included in the spectrometer as such, and hence the development of the low-temperature parts is omitted from this thesis.

This thesis is organized as follows. Chapter 2 introduces relevant theory of electronic noise, computational methods and radiometers as well as reviews existing

measurement setups from the literature. Chapter 3 presents the structure of the measurement system and describes the development of the data processing system. Chapter 4 provides measurement results of the individual parts as well as the whole system and presents the measurement methods. Finally, the implemented setup is evaluated in Chapter 5 together with proposals for further development of the system.

Chapter 2

Theoretical background

This chapter introduces previous research in the field and phenomena closely related to radiometers. Section 2.1 discusses the characteristics of the measured phenomenon, microwave shot noise. Section 2.2 introduces other types of noise present in the measurement environment. Sections 2.3 – 2.5 review different types of radiometer circuits and their main parts as well as previous radiometer implementations. The focus is on radiometers used in mesoscopic physics research, although some general results from radio astronomy are also discussed.

2.1 Shot noise

Shot noise results from the quantization of charge [20], although it should not be confused with quantum noise (Sect. 2.2.2). The movement of charge carriers in a conductor can be modeled as a Poissonian process with the probability distribution

$$P(X = k, T) = \frac{e^{-\lambda T}(\lambda T)^k}{k!},\ k \geq 0, \text{integer}, \tag{2.1}$$

where k is the number of events, T is length of a time interval and λ is a parameter of the distribution [21, p. 193]. However, electron-phonon interactions suppress this shot noise resulting in significantly less noise than calculated in normal metallic conductors. Such interactions are not present in systems with a potential barrier, namely tunneling and p-n junctions, and thus more shot noise can be observed [22, p. 432]. Measurement of shot noise can be used to characterize the charge transport and quasiparticle characteristics in a sample [23]. The power spectral density of shot noise is

$$S_n = 2eI_{DC}\mathcal{F}, \tag{2.2}$$

where e is the electric charge of electron, I_{DC} is the DC current and $\mathcal{F}$ is *Fano factor* which is unity for Poissonian statistics and differs from that in occurrence of suppression or enhancement in shot noise. However, the above formula omits thermal and quantum noise (Sect. 2.2) as well as frequency dependence, and therefore the complete expression for the noise of a tunnel junction ($\mathcal{F} = 1$) is

$$S_I(f, V, T) = \frac{2k_BT}{R}\left[\frac{eV + hf}{2k_BT}\coth\left(\frac{eV + hf}{2k_BT}\right) + \frac{eV - hf}{2k_BT}\coth\left(\frac{eV - hf}{2k_BT}\right)\right], \tag{2.3}$$

where f is frequency, V is DC bias voltage, T is physical temperature, k_B is the Boltzmann's constant, R is sample resistance and h is the Planck's constant [24].

2.2 Thermal and quantum noise

In addition to shot noise described in the previous section there are other noise types affecting the measurement system. This section introduces thermal noise and quantum noise, as well as the definition of noise temperature which combines the two noise types. While these topics are discussed in numerous textbooks, most of the books concentrate only on the phenomena relevant near room temperature. Therefore, this section focuses on microwave noise below 1 K and omits the general theory of the noise types.

2.2.1 Thermal noise

Thermal noise in electronic circuits is caused by thermal motion of charge carriers in a conductor [21, p. 130]. The available power spectral density from a resistor is commonly calculated using the Rayleigh-Jeans formula

$$S_n = k_BT, \tag{2.4}$$

where k_B is the Boltzmann's constant and T is the physical temperature in kelvins [25]. However, this formula is a low-frequency, high-temperature approximation that assumes photon energy to be significantly smaller than thermal energy ($hf \ll k_BT$). This Rayleigh-Jeans approximation is accurate enough for most practical applications at microwave frequencies, but not at very low temperatures. For example, at 800 MHz frequency photon energy equals thermal energy at 38 mK, a temperature reachable with ^{3}He/^{4}He dilution refrigerators (introduced in Section 3.2). Because of its inaccuracy, Eq. 2.4 is replaced by a more general Planck form

$$S_n = \frac{hf}{e^{hf/k_BT} - 1}, \tag{2.5}$$

where h is the Planck's constant and f is frequency [26].

2.2.2 Quantum noise

In addition to the common types of noise, there exists quantum noise also known as vacuum fluctuations. These fluctuations are caused by the quantum mechanical limit to measurement accuracy, known as the Heisenberg uncertainty principle [27]. The noise has a power spectral density of the form

$$S_q = \frac{hf}{2}. \tag{2.6}$$

Although quantum noise is present in all electrical systems, it is neglected in most microwave applications because of its small value compared to other sources of noise. However, at temperatures low enough, quantum noise becomes significant at microwave frequencies. At 800 MHz quantum noise has a noise temperature of 19 mK, while the base temperature of a ^{3}He/^{4}He dilution refrigerator is usually 10 mK.

Quantum noise can be combined with thermal noise by adding Eq. 2.6 to Eq. 2.5. Then the formula for the spectral density of available noise power from a resistor gets the form

$$S_n = hf\left(\frac{1}{e^{hf/k_BT}-1} + \frac{1}{2}\right), \tag{2.7}$$

known as the Callen-Welton form [26]. This formula includes all the noise terms present in an ideal resistor at given temperature and frequency.

2.2.3 Definition of noise temperature

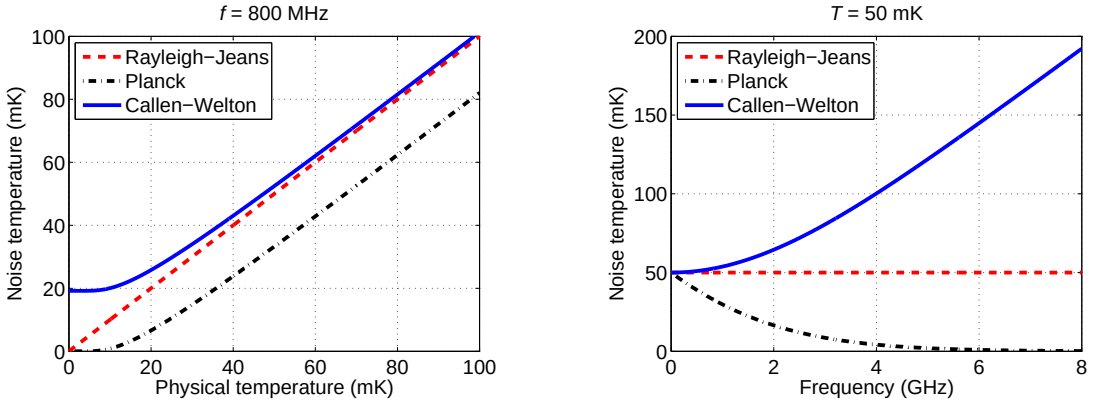
The noise temperature of a resistor can be defined in two different ways: by noise power or by physical temperature. In the noise power definition the noise temperature equals to the noise power spectral density divided by the Boltzmann's constant. In the physical temperature definition the noise temperature is equal to the physical temperature of a resistor that would produce the given spectral density [28]. These definitions can be considered equal when the Rayleigh-Jeans approximation ($hf \ll k_BT$) applies. However, as shown in Section 2.2.1, this approximation does not apply at microwave frequencies and low temperatures. Therefore a choice should be made between these two definitions.

The power definition is recommended for microwave radiometry [29] and it is most commonly used in the field of microwave engineering [28]. Therefore, the power

definition of noise temperature is also used throughout this thesis. The noise temperature of a resistor is then

$$T_n = \frac{S_n}{k_B}, \tag{2.8}$$

where S_n can be taken from Eqs. 2.4, 2.5 or 2.7. The differences between these definitions are illustrated in Fig. 2.1, where the noise temperature of a resistor is presented as functions of physical temperature and frequency. It should be noted that the expressions differ significantly from each other at the frequency and temperature ranges used in low-temperature noise measurements. The Planck expression (Eq. 2.5) is widely accepted as the correct formula, although there is no agreement whether or not the vacuum fluctuations should be included, in which case the Callen-Welton form (Eq. 2.7) would be the correct one [28]. While the choice is affected by the application, the Callen-Welton form is the most reasonable for resistors measured at low temperatures. This is because it includes all the noise terms present in an ideal resistor in a single formula. Therefore, the Callen-Welton form is chosen for this thesis.



(a) (b)

Figure 2.1: Noise temperature of a resistor using Rayleigh-Jeans law (Eq. 2.4), Planck law (Eq. 2.5) and Callen-Welton law (Eq. 2.7) (a) as a function of physical temperature at 800 MHz frequency, (b) as a function of frequency at 50 mK temperature.

This section reviewed thermal and quantum noise at low temperatures, as well as the definition of noise temperature. It was shown that the commonly used Rayleigh-Jeans formula is based on approximations which are invalid at the frequency and temperature ranges used in this thesis. In addition, quantum noise becomes significant under these conditions and thus it should be included in the noise formula. It was also pointed out that noise temperature should be defined as the noise power spectral density divided by Boltzmann's constant. Moreover, the noise temperature of a resistor should be calculated using the Callen-Welton formula which combines both thermal and quantum noise.

2.3 $1/f$ noise

Low-frequency fluctuations, namely $1/f$ noise (also called flicker noise), is present in all electronic systems [30]. Its power spectral density is proportional to $f^{-\alpha}$ where the value of α may vary between 0.5 and 1.5 [31]. This spectrum has been observed in electrical conductors and also in various other systems, including sea level, loudness of a radio broadcast and ocean current velocities [32]. $1/f$ noise of a conductor has a corner frequency ranging from 100 Hz to 1 MHz above which shot noise and thermal noise become dominant [33, p. 205]. Therefore, shot noise measurements are typically conducted above this corner frequency. A phenomenon related to $1/f$ noise is amplifier gain and noise temperature variations, also known as drift, which limit the stability of high-frequency shot noise measurements (see Sect. 2.5).

2.4 IQ mixer

Any signal can be expressed as a sum of sinusoidal waveforms of the form

$$x(t) = \mathrm{A}\cos(\omega t + \phi), \tag{2.9}$$

where A is signal amplitude and ω and ϕ denote angular frequency and initial phase of the signal, respectively. Alternatively, the above equation can be expressed as a sum of sine and cosine components with

$$x(t) = \mathrm{A}\cos(\phi)\cos(\omega t) + \mathrm{A}\sin(\phi)\sin(\omega t). \tag{2.10}$$

This representation separates the in-phase (cosine) and quadrature phase (sine) components of the signal. The initial phase term can be removed by including it to the amplitudes of the individual components by

$$x(t) = \mathrm{A_I}\cos(\omega t) + \mathrm{A_Q}\sin(\omega t). \tag{2.11}$$

IQ mixers extract both the in-phase (cosine) and quadrature phase (sine) components of the input signal and convert the radio frequency (RF) signal down to intermediate frequency (IF). This is done by multiplying the RF signal with a sinusoidal local oscillator (LO) signal.

IQ mixers are used in wireless communication systems to demodulate QPSK (quadri-phase-shift keying) and QAM (quadrature amplitude modulation) signals. Therefore IQ mixer is commonly known as IQ demodulator [34, p. 223]. However, the signal in noise spectrometry is not modulated and thus the term IQ mixer is used instead [17].

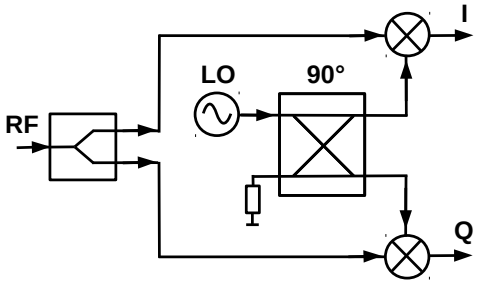


Figure 2.2: Funtional schematic of IQ mixer.

Figure 2.2 shows the functional structure of IQ mixer. The incoming RF signal is split to two mixers. These mixers share a local oscillator (LO) signal which is split with a 90° hybrid. Therefore the RF signal is mixed with LO signal of the form $\cos(\omega_0 t)$ in I branch and $\sin(\omega_0 t)$ in Q branch. The time-domain I signal at mixer output is then

$$\begin{aligned} \mathrm{I}(t) &= [\mathrm{A_I}\cos(\omega t) + \mathrm{A_Q}\sin(\omega t)] \cdot \mathrm{A_{LO}}\cos(\omega_0 t) \\ &= \frac{\mathrm{A_I A_{LO}}}{2}\left[\cos\left((\omega-\omega_0)t\right) + \cos\left((\omega+\omega_0)t\right)\right] \\ &\quad + \frac{\mathrm{A_Q A_{LO}}}{2}\left[\sin\left((\omega-\omega_0)t\right) + \sin\left((\omega+\omega_0)t\right)\right], \end{aligned} \tag{2.12}$$

where $\mathrm{A_I}$ and $\mathrm{A_Q}$ denote the amplitudes of the RF in- and quadrature-phase components, respectively and $\mathrm{A_{LO}}$ is the amplitude of the local oscillator signal. Similarly, the Q signal is given by

$$\begin{aligned} \mathrm{Q}(t) &= [\mathrm{A_I}\cos(\omega t) + \mathrm{A_Q}\sin(\omega t)] \cdot \mathrm{A_{LO}}\sin(\omega_0 t) \\ &= \frac{\mathrm{A_I A_{LO}}}{2}\left[-\sin\left((\omega-\omega_0)t\right) + \sin\left((\omega+\omega_0)t\right)\right] \\ &\quad + \frac{\mathrm{A_Q A_{LO}}}{2}\left[\cos\left((\omega-\omega_0)t\right) + \cos\left((\omega+\omega_0)t\right)\right]. \end{aligned} \tag{2.13}$$

Since the intermediate-frequency terms are of primary interest in down-conversion, the terms with $(\omega+\omega_0)$ are removed with low-pass filtering. The output signals after low-pass filtering are given by

$$\begin{aligned} \mathrm{I_{LP}}(t) &= \frac{\mathrm{A_{LO}}}{2}\left[\mathrm{A_I}\cos\left((\omega-\omega_0)t\right) + \mathrm{A_Q}\sin\left((\omega-\omega_0)t\right)\right] \\ \mathrm{Q_{LP}}(t) &= \frac{\mathrm{A_{LO}}}{2}\left[-\mathrm{A_I}\sin\left((\omega-\omega_0)t\right) + \mathrm{A_Q}\cos\left((\omega-\omega_0)t\right)\right]. \end{aligned} \tag{2.14}$$

It should be noted that both IF signals include terms proportional to both the I and Q components of the input signal. However, the IF signals have 90° phase difference and thus they maintain the orthogonality of the I and Q components.

2.5 Sensitivity

Spectrometer and radiometer circuits are prone to random fluctuations in the measurement value over time. This sets a limit for the maximum integration time in a measurement. At some point the fluctuations start to degrade the effective sensitivity of the measurement. Thus, there exists an *optimal integration time*, at which the maximal sensitivity of the measurement is obtained.

The sensitivity of a measurement is determined by the standard deviation which is given by

$$\Delta T = \frac{T_s}{\sqrt{B\tau}}, \tag{2.15}$$

where T_s denotes the system noise temperature, B is the frequency bandwidth and τ is the integration time [35, pp. 30–35]. In digital sampling systems, the highest detectable frequency is half of the sampling rate [36]. This leads to the replacement of the product $B\tau$ with the amount of digitized samples, N. The sensitivity is improved by a factor of $\sqrt{2}$ by using complex correlation with separate digitizer channels for I and Q components of the signal [11]. The complex correlation is analogous to doubling the amount of data points, and hence Eq. 2.15 gets the form

$$\Delta T = \frac{T_s}{\sqrt{2N}}. \tag{2.16}$$

The system noise temperature, T_s, has a significant effect on the sensitivity in the above equations. However, since the improvement of the system noise temperature is out of scope of this thesis, it is useful to introduce relative sensitivity of a noise measurement, defined as

$$\frac{\Delta T}{T_s} = \frac{1}{\sqrt{2N}}. \tag{2.17}$$

This separates the noise temperature from the other terms making it easier to compare the sensitivities of systems with different noise temperatures. An example is direct comparison between cryogenic and room-temperature measurements of a system since the latter typically involves significantly higher system noise temperature.

2.6 Radiometer types

The spectrometer circuit designed in this thesis shares similarities with radiometer circuits used for determining noise power mainly in radio astronomy or remote sensing applications. There are several types of radiometer circuits all aiming at improving the stability when compared to direct power measurement. However, here we restrict our treatment to three most important radiometers in mesoscopic physics experiments: the Dicke radiometer and cross-correlation radiometer.

2.6.1 Total power radiometer

Total power radiometer is the simplest type of radiometer, consisted of a single amplifier chain connected to the measured noise source. The sensitivity of an ideal total power radiometer system is given by Eq. 2.15. However, in practical situation the maximum achievable sensitivity is restricted by gain fluctuations, which can be included in the sensitivity by

$$\Delta T = T_s \sqrt{\frac{1}{B\tau} + \left(\frac{\Delta G}{G}\right)^2}, \tag{2.18}$$

where ΔG denotes the RMS value of gain fluctuations and G is the total gain of the system [34, pp. 252–253]. The presence of gain fluctuations leads to a lower bound of sensitivity after which longer integration time does not result in better sensitivity of the system, thus limiting the usefulness of total power radiometers.

2.6.2 Dicke radiometer

The Dicke radiometer is based on periodic switching between a reference noise source and the sample [37]. This improves the sensitivity over total power radiometers by eliminating gain fluctuations occurring at time periods longer than the switching time. The periodic switching reduces the sensitivity by a factor of two, which leads to

$$\Delta T = \frac{2T_s}{\sqrt{B\tau}}. \tag{2.19}$$

In mesoscopic physics experiments several practical factors limit the use of a microwave switch. The switch must be situated in the beginning of the signal chain and thus in the mixing chamber at 10 mK temperature. Few commercial switches are functional at such low temperatures. Also the power consumption of the switch is restricted by the cooling power of the cryostat, which is in the order of microwatts. Another considerable factor in mechanical switches are the voltage transients which

may damage the sensitive sample. With all these limitations, the microwave switch is commonly problematic in mesoscopic physics experiments.

Besides operating a microwave switch, the Dicke switching can also be performed by alternating the DC bias of the sample. Zero bias would then provide the noise reference to which measurement values are compared. The switching time can not be made arbitrarily short, due to relaxation time required for the sample to change its state to the new bias value. Therefore the factor of two in Eq. 2.19 is greater than two in practical systems where relaxation is present.

2.6.3 Cross-correlation radiometer

Another approach to overcome the sensitivity problem encountered in total power radiometers is the use of cross-correlation by connecting the noise from the sample to two separate amplifier chains and multiplying the two signals together as described in Sect. 2.8.2. This separates the correlated part of the noise resulting from the sample and the uncorrelated part due to noise of the amplifiers. Since amplifier gain variation is typically uncorrelated between the two amplifiers, cross-correlation measurement significantly reduces the effect of gain variations.

In nanophysics, cross-correlation enables the observation of other interesting phenomena besides shot noise power. The noise can be probed from two separate electrodes instead of splitting the noise coming from a single electrode. Then, a third electrode can be used to control the electrical connection between the two electrodes, varying the correlated part of the measured shot noise. In addition, cross-correlation is an essential tool in determining interference effects, including Hanbury Brown and Twiss (HBT) effects in multiterminal diffusive conductors [38].

Another type of cross-correlation measurement involves correlations of the noise intensities (power) instead of noise voltages. This can be done by squaring the measured noises of the two channels and multiplying them together. The sign of this intensity correlation changes according to the charge carrier properties of the sample. A positive correlation indicates that bosonic particles or quasiparticles act as charge carriers, whereas negative correlation indicates fermionic charge carriers. This way one can make a distinction between bosonic Cooper pairs and fermionic normal electrons in a superconductor. Yet another form of correlation measurement involves higher order cumulants of shot noise which can be calculated from digitized data [17] or obtained from an analog measurement setup [12, 39]. All these variants of correlation radiometer make it a versatile tool in nanophysics experiments.

2.7 Quantization noise

Representing an analog quantity with a finite number of discrete levels involves quantization error due to rounding to the nearest quantization level. These quan-

tization errors are observed as noise in digital continuous sampling systems. This quantization noise has variance of the form

$$\sigma_e^2 = \frac{\Delta^2}{12}, \tag{2.20}$$

where Δ is the quantization step size [21, p. 55]. It should be noted that the variance, as well as the noise power, is proportional to the square of the quantization step size. The spectral density of quantization noise is approximately uniform, and thus it can be considered as white noise.

However, the quantization noises of two separate channels are not correlated, and thus quantization noise is less significant in correlation receivers. Table 2.1 shows the required integration times for different bit depths in order to achieve the accuracy of an ideal analog correlator. It can be seen that a 1-bit correlator requires 2.43 times the integration time of the ideal correlator. The difference in integration times becomes insignificant as the bit depth increases, being only 2 % for a 4-bit correlator.

Table 2.1: Correlator integration time relative to that of ideal analog correlator with different bit depths and corresponding level counts [40].

Bits	Levels	Rel. int. time
1	2	2.43
2	4	1.30
3	8	1.10
4	16	1.02

Because of the relatively small difference in integration times, the early digital radiometers were equipped with one- or two-bit correlators [41, 42]. More recently, low-bit-depth correlators have been used in multichannel applications and when high sample rates are required. The drawbacks of low-bit-depth correlators include the requirements of linearization of the results and accurate adjustement of the digitizer threshold levels. It should also be noted that the results of Table 2.1 apply only for Gaussian signals [40]. Power- or V^2-correlations also require bit depth of more than one. In addition, autocorrelation of a single channel is affected by the additive quantization noise. Because of these reasons, a low-bit-depth system may experience unexpected performance degradation when the measurement conditions differ from the above criteria. These problems can be overcome by employing higher bit-depth correlators accompanied by computing power which are made affordable due to the advance in technology.

2.8 Computational methods

This section reviews the mathematical operations and computing environment relevant for the system developed in this thesis. At first, the Fourier transform and its discrete-time variant and their relevant properties are discussed. Next, the section introduces the cross-correlation function together with its implementation and relation to Fourier transform, as well as methods for normalization. Finally, graphics processing units, being the hardware environment used in this thesis, are reviewed together with existing implementations of cross-correlation from radio astronomy.

2.8.1 Discrete Fourier transform

A time-domain function can be presented in frequency domain using Fourier transform. The Fourier transform of a time domain function $f(t)$ is given by

$$\mathcal{F}[f(t)] = F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i\omega t}dt, \tag{2.21}$$

where ω and i denote angular frequency and imaginary unit, respectively [43, p. 9]. The Fourier transform can be applied to discrete-time series which are present in sampling receiver systems. The *discrete Fourier transform* (DFT) of a sampled sequence is

$$F\left[\omega\right] = \sum_{t=0}^{N-1} f\left[t\right] \exp\left(\frac{-2\pi i\omega t}{N}\right), \tag{2.22}$$

where N denotes the number of points in the series [44, p. 5]. It should be noted that variable symbols t and ω, which are chosen here to improve clarity, represent sampled data rather than time or angular frequency in strict definition. The DFT results in frequency-sampled data series $F[\omega]$ with N points. Both discrete and continuous Fourier transforms can be inversed by applying an inverse transform with the sign of the exponent reversed.

It follows from Eqs. (2.21) and (2.22) that both Fourier transform and DFT (and their inverse) are linear operators, which leads to the identity

$$\mathcal{F}(f + g) \equiv \mathcal{F}(f) + \mathcal{F}(g). \tag{2.23}$$

The DFT can be calculated efficiently using the *fast Fourier transform* (FFT) algorithm which reduces the computational complexity of the DFT. The FFT functions best when the number of points is an integer power of two, although effective algorithms have been implemented for several small integer factors, $N = 2$, 3, 4, 5, 7, 8,

11, 13 and 16 [45, pp. 1015–1016]. Several implementations of FFT with optimized algorithms are available as libraries for programming.

The sampled signal is a finite-size piece of an infinite sequence. This makes the frequency components in the DFT of the signal to appear as non-periodic in the sequence. This broadens the line features in the spectrum, which causes measurement error due to leakage of a spectral feature to the adjacent DFT frequency bins. This effect can be reduced by multiplying the time-domain sequence with a window function prior to DFT [46]. Various DFT window functions are presented in Fig. 2.3 and their Fourier transforms in Fig. 2.4. It can be seen that the rectangular window, which is equivalent to not having any window, has more leakage to the adjacent frequency bins compared to other windows. In contrast, lower frequency leakage correlates with broader main lobe width, and hence reduced frequency sensitivity. Therefore, different window functions are suitable for different applications [47, p. 286].

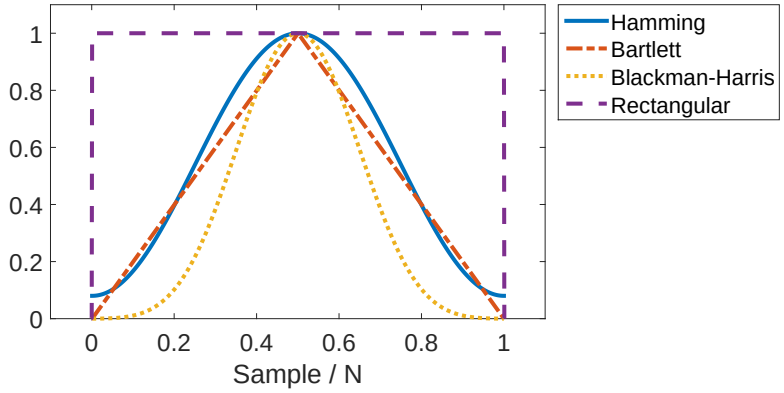


Figure 2.3: Various windowing functions for DFTs in time domain (based on [48, p. 447]).

The use of window functions is less important when the amount of points in the sequence is very large, since the effect of spectral leakage is determined by the number of adjacent DFT frequency bins rather than by absolute frequency. The frequency bin spacing is very dense in large sequences, and hence the leakage almost entirely affects the frequencies in the close proximity of the initial frequency. However, the use of window functions should be considered if the measurement involves closely spaced spectral features.

2.8.2 Cross-correlation function

The cross-correlation is a figure of similarity of two signals. The function is defined in discrete time domain as

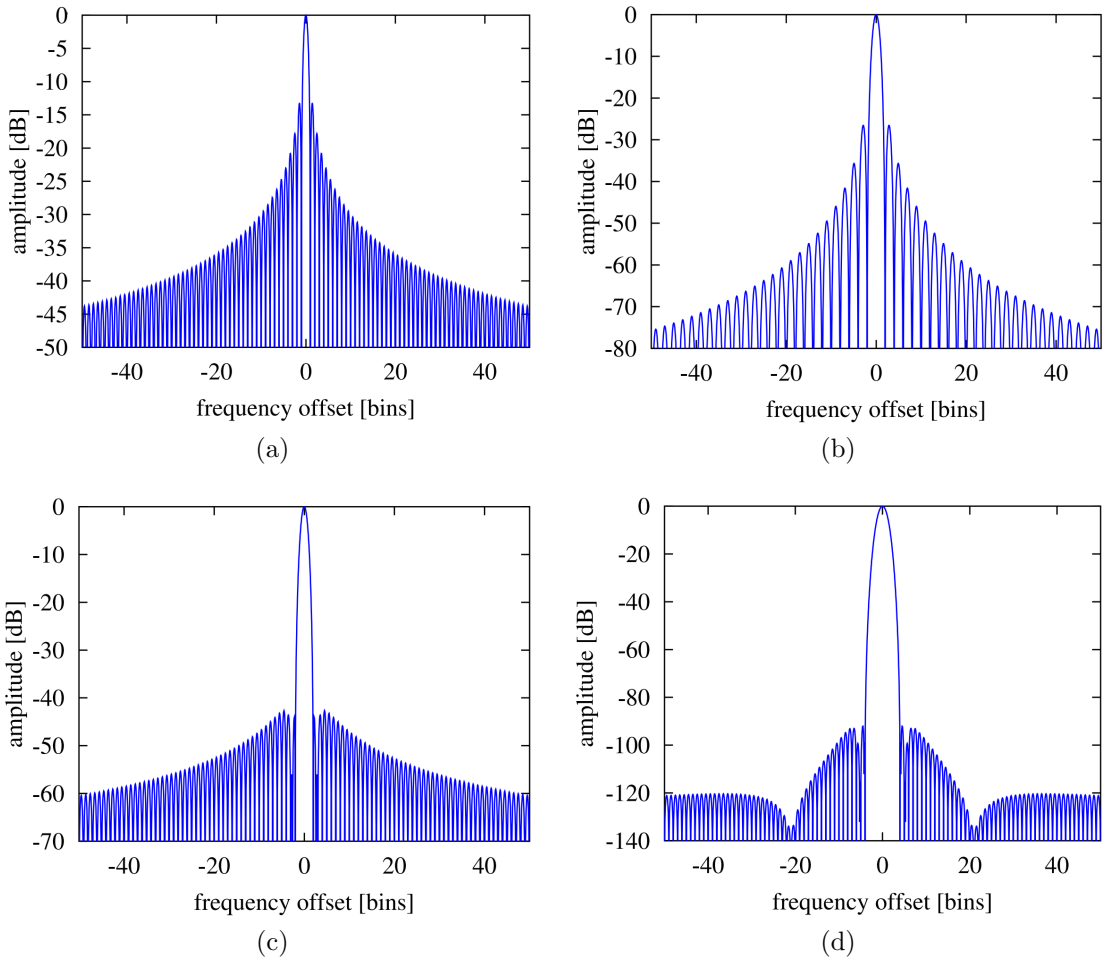


Figure 2.4: Fourier transforms of four DFT windowing functions, showing the main-lobe widths and sidelobe levels. (a) Rectangular, (b) Bartlett, (c) Hamming, (d) Blackman-Harris [49].

$$z_{cor}\left[t\right] = \sum_{\tau=-\infty}^{\infty} \overline{f\left[\tau\right]}\, g\left[t+\tau\right], \tag{2.24}$$

where t and τ are time variables and the overline denotes complex conjugate [48, p. 65]. The cross-correlation of two sequences with N samples produces a sequence of $2N+1$ points. The correlation of a signal with itself gives the autocorrelation, which at zero time lag is equal to the signal power. The cross-correlation operation reduces to multiplication in the frequency domain and therefore the cross-correlation can be obtained via Fourier transform by

$$z_{cor}\left[t\right] = IDFT\left[\frac{1}{N} X^{*}\left[f\right] \cdot Y\left[f\right]\right], \tag{2.25}$$

where X^* denotes the complex conjugate of X, and $IDFT$ denotes inverse discrete Fourier transform [44, p. 25]. The result from the above formula for time-domain sequences of N samples has likewise N points, compared to $2N + 1$ points for the time-domain formula (Eq. 2.24). This difference can be overcome by zero-padding the input sequences to length $2N$. Zero-padding reduces the error in spectral lines occuring in the middle of two cross-power spectrum bins. However, zero-padding increases the computation time, and it is not necessary if the reduction in spectral accuracy can be tolerated [47, p. 293].

It is usually beneficial to normalize the cross-correlation with the signal powers (autocorrelations) to obtain a comparable quantity not affected by the differences in signal level. The normalization also reduces the effect of amplifier gain variation in time. The normalized cross-correlation in frequency domain is called *coherence*, which is given by

$$|R(f)|^2 = \frac{s_{v12}^2(f)}{s_{v1}(f)s_{v2}(f)}, \tag{2.26}$$

where $s_{v1}(f)$ and $s_{v2}(f)$ are the autospectral densities (square root of autocorrelation spectrum) of signals in channels 1 and 2, and $s_{v12}^2(f)$ denotes the cross-spectral power density between the two channels. The coherence gets values between zero and unity, i.e. between no and full correlation between the two channels [50, p. 22]. The formula for normalized cross-correlation in time domain is then

$$z_{norm}[t] = \frac{z_{cor}[t]}{\sqrt{\sigma_{xx}^2 \sigma_{yy}^2}}, \tag{2.27}$$

where σ_{xx}^2 and σ_{yy}^2 are the autocorrelations at zero time lag.

2.8.3 Graphics processing units

Traditionally, all general-purpose computations in a computer were processed in the central processing unit (CPU), while the graphics processing units (GPUs) were reserved solely for generating and displaying graphics. However, the GPUs have gained a lot of attention during the last decade because of their high performance in certain general-purpose computational tasks. This is mainly due to computer graphics being highly parallel in nature, and hence also other problems involving parallelism gain performance advantage because of the optimized architecture. The parallel architecture enables more efficient use of the transistors on the GPU chip compared to that of CPUs, which results in significantly higher computing power measured in floating point operations per second (FLOPS) [51]. The FFT is an example of such parallel problem that can be efficiently implemented on the GPU [52].

There are several extensions to programming languages for developing programs running on the GPU, including NVIDIA CUDA which is also used in this thesis. CUDA includes an extension to the C programming language and several libraries including CUFFT for efficient calculation of FFTs on the GPU. CUDA also offers the developer a capability of running multiple (up to thousands) instances of a function in parallel, each instance operating on a different part of the input data.

CUDA has been applied to calculation of cross-correlation in radio astronomy [53, 54]. The cross-correlation is calculated according to Eq. (2.25) by first computing the DFTs of the individual sequences and multiplying the results. These applications benefit from the ease of the implementation on the GPU over application-specific integrated circuits (ASICs) and field-programmable gate arrays (FPGAs) traditionally used in radio astronomy correlators. This is because the program code running on the GPU is easily modifiable, and the hardware is upgradeable if higher performance is required. The drawback of the GPU implementation is the increased power consumption which may become an issue in radio telescopes installed in remote areas.

2.9 Previous setups

This section describes several existing correlation detection setups from the literature. Each setup is presented separately instead of comparing the implementations part-by-part. This is due to the fact that the setups are designed for different purposes and it is thus beneficial to treat each implementation as a whole.

2.9.1 Paris setup

The Paris setup is used for meoscopic physics experiments in low temperatures [15]. The frequency range of the setup is around 1.5 GHz, and the setup is based on analog electronics and lock-in detection. An overview of the setup is presented in Fig. 2.5.

The mesoscopic sample is connected to one of the input ports of the 90° hybrid coupler. A thermal 50 Ω reference load is connected to the other input. The hybrid divides the signals equally to the two amplifier chains with 90° phase shift between the two arms. The signals are then amplified with cryogenic and room-temperature amplifiers. The signal in one of the arms is periodically phase-modulated between 0° and 180° by a biphase modulator at 2.7 kHz frequency. The signals are then fed to the second 90° hybrid coupler. The signal power at the outputs of the second hybrid alternates by 20 dB, depending on the phase modulation as shown in Fig. 2.6.

The output ports of the second hybrid are connected to Schottky detector diodes converting the RF power to DC voltage. The difference of the detector output

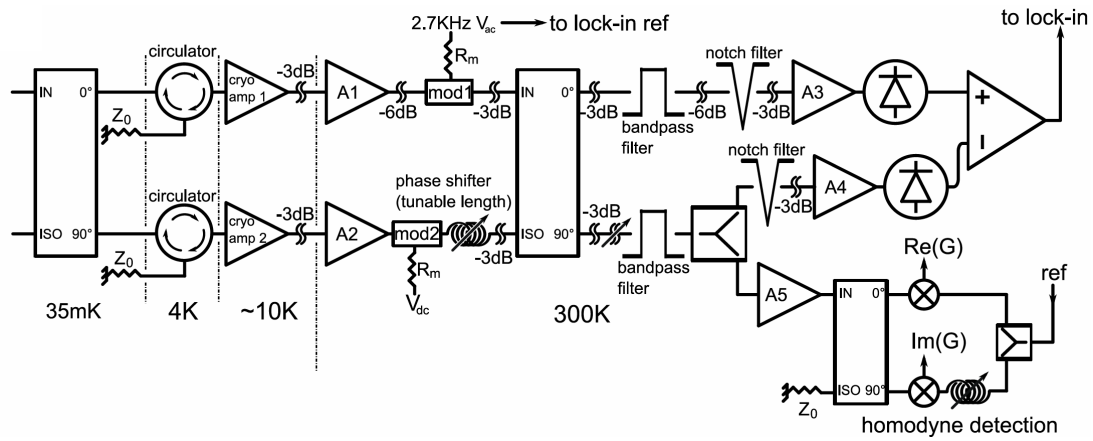


Figure 2.5: Detailed schematic of the Paris setup [15].

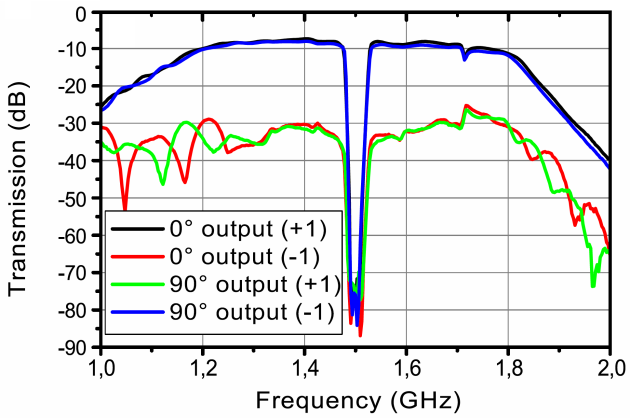
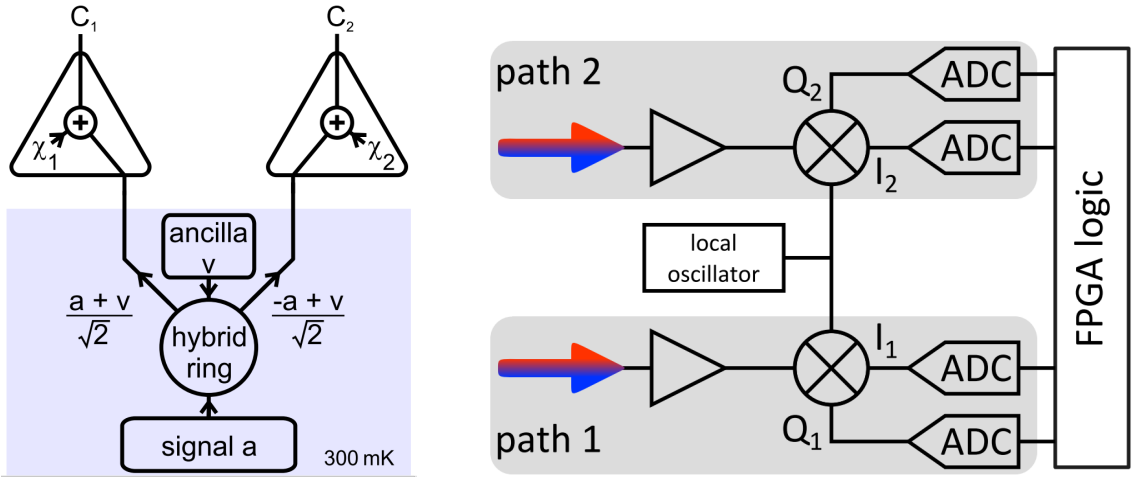


Figure 2.6: Magnitude of S_{21} parameter from one input port of the first hybrid to the outputs of the second hybrid with 0° (+1) and 180° (-1) phase modulation in one amplifier arm. 1.5 GHz notch filter used to suppress RF input signal [15].

voltages is finally detected with a lock-in amplifier with 2.7 kHz reference frequency from the biphase modulator.

The Paris setup utilizes analog electronics, which enables broad detection bandwidth from hundreds of MHz up to a few GHz depending on the center frequency. The detection bandwidth is thus significantly broader than in digitizer-based setups, in which the band is limited by the sampling rate of the digitizer. The setup is also capable of continuously collecting data points in contrast to digital setups, in which the duty cycle is often far from 100 %.

The 2.7 kHz biphase modulation is an integral part of the setup. The phase modulation enables continuous comparison between sample noise and reference noise, thus effectively removing the effects of amplifier gain and noise temperature fluctuations. The fluctuations occur below the $1/f$ noise corner frequency around 1 kHz, and therefore 2.7 kHz is sufficiently high in order to remove these drifts.



(a) (b)

Figure 2.7: (a) Schematic of the cryogenic parts of the Munich setup, with the signal (a), terminator noise (v) and noise of the amplifiers (χ_1 and χ_2). The amplified RF signals C_1 and C_2 are fed to room temperature for detection. (b) Room-temperature part of the setup. The signals C_1 and C_2 are further amplified and fed to the IQ mixers, downconverting the signals to 11 MHz and extracting both quadratures of the signals. Resulting four quadratures are detected by a digitizer employing four analog-to-digital converters (ADCs). FPGA logic is used for data processing. [17]

The drawbacks of the setup are related to the limited flexibility of the analog detection. The noise power difference is recorded a few times per second at maximum. The setup allows neither numerical correction of the data nor the extraction of other parameters from the measured data without hardware modifications.

The setup achieves temperature resolution of 10 μK in a two-hour measurement, corresponding to a sensitivity of 0.71 mK/$\sqrt{\text{Hz}}$ or 3.3×10^{-28} A^2/Hz/$\sqrt{\text{Hz}}$ on a 120 Ω load. These values are determined by periodic switching between two gate voltages, measuring one gate point for 10–20 s and subtracting the measured reference, equivalent to Dicke switching scheme. Thus, the given sensitivity is determined by measuring noise power difference with mean value of $\sim$ 20 μK instead of absolute noise power of 13.5 mK.

2.9.2 Munich setup

The Munich setup [17] is used for studying propagating quantum microwaves at low temperatures. The detection is based on downconverting and digitizing both quadratures of the RF signal on two signal paths using a four-channel digitizer. The desired quantities are then extracted from the digitized data on a field programmable gate array (FPGA). A schematic of the setup is presented in Fig. 2.7.

The low-temperature parts of the setup include a 180° ring hybrid dividing the measured signal into two amplifier chains with 180° relative phase difference. The other input port of the hybrid is connected to a 50 Ω terminator, referred to as *ancilla* state.

The system operates at 5.6 GHz RF frequency. The RF signals are downconverted using IQ mixers to intermediate frequency (IF) of 11 MHz, determined by the frequency difference of local oscillator and system input. The operation principle of IQ mixers is discussed in Sect. 2.4.

The four IF signals are digitized at sample rate of 150 MS/s with 16 bit resolution. A sample rate much higher than the Nyquist rate results in good time resolution. The sample rate corresponds to 625 MB/s data rate with 54.6 % duty cycle. The data is pre-processed in the FPGA by correcting the gain and phase inbalances and narrowing the measurement bandwidth to 982 kHz by digital band-pass filtering. All the noise moments up to fourth order are calculated in the FPGA and the results are transferred to a computer for further processing. The data output rate from the FPGA is a few megabytes per second.

The Munich setup is used for studying quantum microwaves amplified by a Josephson parametric amplifier (JPA). The nature of these measurements is the reason for the narrow measurement bandwidth when compared to noise measurement setups. The data pre-processing and moment calculation are computationally demanding, which justifies the use of FPGA rather than a computer for data processing.

The setup includes several factors improving the performance. Ten circulators in total are placed at different stages of the amplification chains to reduce crosstalk between the signal chains. However, the necessity of having such a large number of circulators is neither discussed nor justified with measurements. The stability of the setup is improved by reducing the temperature fluctuations. This includes closed-cycle cooling water system for the room temperature amplifiers with ±0.1 °C temperature accuracy and air-conditioning system for the laboratory. However, the authors provide no information on the effectiveness of these methods.

2.9.3 Harvard setup

The Harvard setup [55] employs a two channel digitizer (National Instruments PCI-5122) with 10 MS/s sample rate and 14 bit resolution. The operation frequency is 2 MHz, which is thus low enough for direct digitizing without downconversion in frequency. The signals are processed on a desktop computer which calculates the cross spectrum and the individual power spectra of the two signals. The calculations are performed almost in real-time: data collected in 10 s is processed in 10.8 s.

The fast Fourier transformations (FFTs) are calculated from blocks of 10 368 samples, which are multiplied by a Hanning window function to avoid end effects in the spectrums (see Sect. 2.8.1). This windowing is required mainly due to the resonant

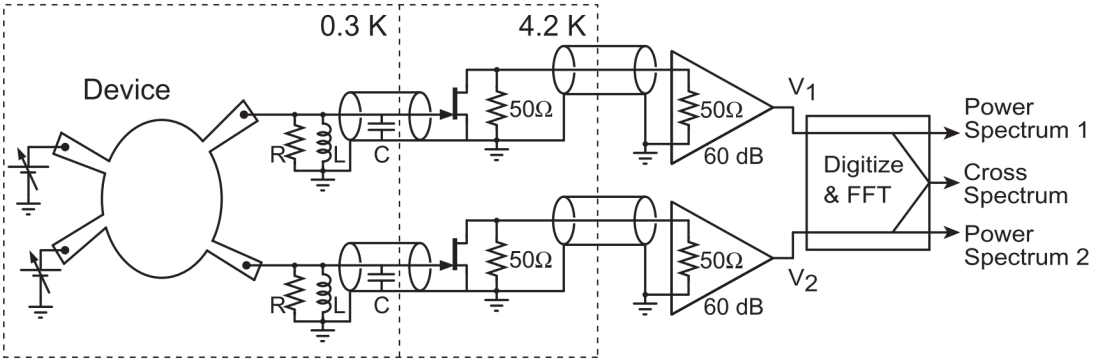


Figure 2.8: Schematic of the Harvard setup showing the measured device, the RLC tank circuit, the cryogenic and room temperature amplifiers and digitizer [55].

behaviour of the measured noise spectrum, as shown in Fig. 2.9. The resonance is caused by the RLC tank circuit located near the quantum point contact (QPC) device.

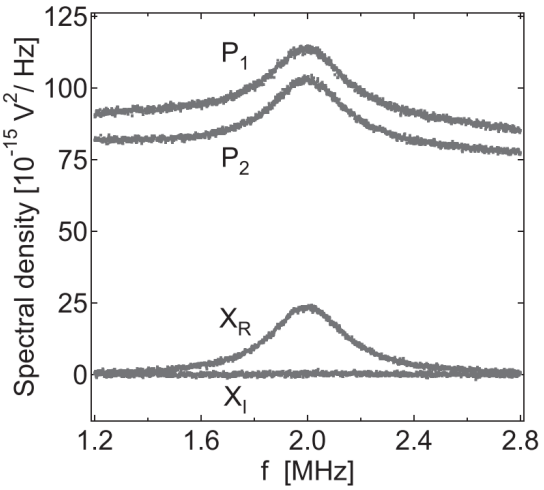


Figure 2.9: Power spectral densities at both input ports (P_1 and P_2) and the real and imaginary parts of the cross spectral density (X_R and X_I) with 20 s integration time at 290 mK temperature with the QPC device pinched off [55].

The current noise spectral sensitivity of the setup is 2.8×10^{-29} A^2/Hz for 10 s integration time. This value is comparable to the sensitivity of the Paris setup (3.3×10^{-28} A^2/Hz/$\sqrt{\text{Hz}}$) described in section 2.9.1 when multiplied by $\sqrt{10\text{ s}}$, resulting in sensitivity of 8.9×10^{-29} A^2/Hz/$\sqrt{\text{Hz}}$. This indicates that the Harvard setup performs better than Paris setup, but the wider measurement bandwidth of Paris setup results in better noise power sensitivity. The sensitivity as a function of integration time obeys the theoretical sensitivity formula up to 10 min integration time, thus demonstrating the good stability of the setup.

2.9.4 Tokyo setup

The Tokyo setup is used for cross-correlation measurements of quantum shot noise in the 200–500 kHz frequency band [16]. The setup includes two homemade cryogenic transimpedance amplifiers at 4 K stage of a dilution refrigerator. The signals are digitized by a two-channel digitizer (National Instruments PXI-5922) at 2 MS/s sample rate. The data is obtained in 20 000 data point blocks, which are converted to frequency domain with FFT. Then, the cross spectral density is calculated.

Despite the use of FFT, authors have omitted the discussion of window functions. This might be due to the white noise nature of the signals, thus justifying the use of rectangular window. However, the block size of 20 000 points is not optimal for FFT, as best performance is achieved when block size is an integral power of two (and $20\,000 = 2^5 \times 5^4$).

The authors concentrate on decreasing the effect of crosstalk between the two signal lines. The authors claim that crosstalk is caused by parasitic capacitance between signal wires and the high input impedance of the amplifiers compared to the sample impedance. This problem is solved by designing cryogenic transimpedance amplifiers with low input impedance.

The authors have noted that that longer averaging results in lower noise floor; an obvious result in microwave noise measurements. This effect is illustrated in Fig. 2.10. The authors provide no information whether this increase in noise resolution obeys or is lower than predicted by theoretical formulas (Sect. 2.5). This limits the possibility to compare the performance of the system to other setups.

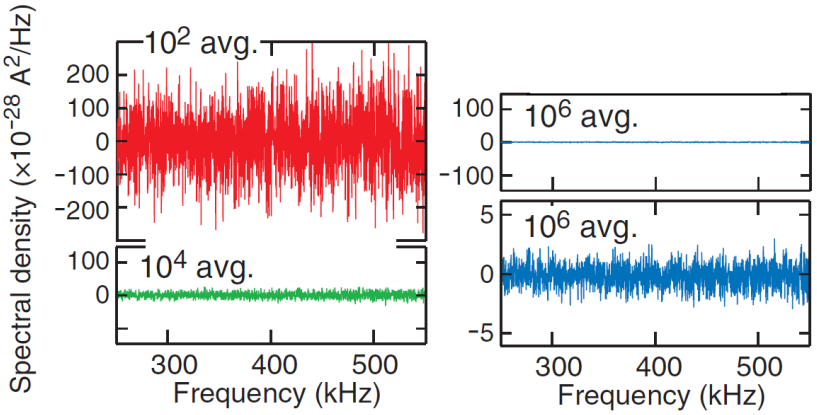


Figure 2.10: Noise power spectral densities with different averaging factors in the Tokyo setup [16].

2.9.5 Discussion on previous setups

The sections above presented four previous correlation measurement setups from the literature, with different approaches. The Paris setup is based on analog circuitry,

thus differing from the spectrometer designed in this thesis. However, its high sensitivity (710 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$) can be compared against the designed setup.

The Munich setup employs an IQ mixer and a digitizer with high sample rate and it is therefore closest to the spectrometer of this thesis in terms of hardware. However, the Munich setup is used for quantum microwaves rather than shot noise and hence its performance is generally not comparable to correlation spectrometers. The data processing is done on FPGA which makes the addition of new features more complicated compared to a solution based on a desktop computer. Moreover, the overall design of the Munich setup is rather complicated, while this thesis aims for a more straightforward solution.

The Harvard and Tokyo setups are based on digitizer and DFT for cross-correlation, although they have low sample rates on the order of few megasamples per second. The low sample rate limits the bandwidth to a few megahertz which reduces the amount of noise and hence the noise power resolution. Data processing at these sample rates is not very demanding for a typical desktop computer and therefore it has not been discussed in detail.

Chapter 3

Implementation of the setup

This chapter describes the structure and implementation of the measurement setup. This thesis concentrates on the room-temperature and software parts, although the cryogenic parts are discussed as well in this chapter for completeness.

3.1 Cryogenic parts

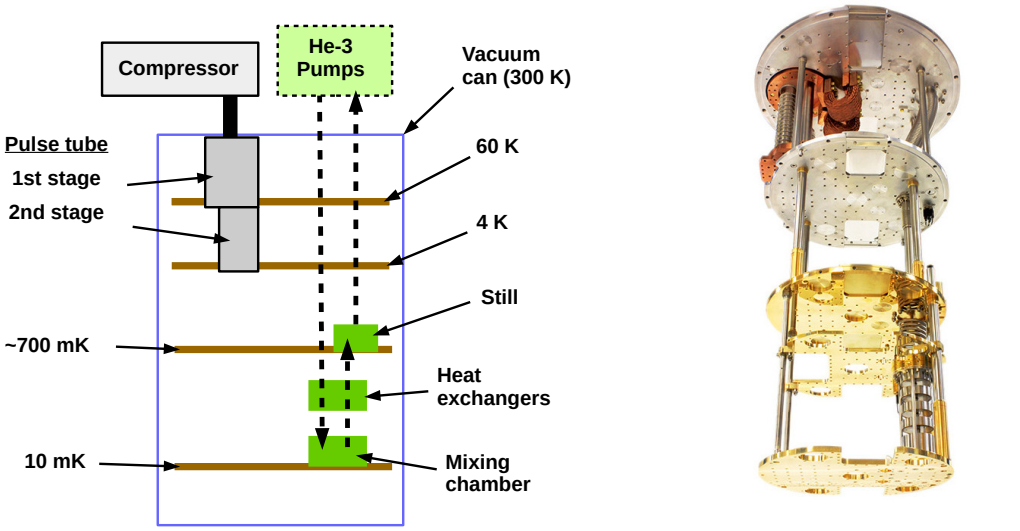
The system designed in this thesis measures weak shot noise signals, and therefore the environment temperature should be lowered to 10 mK using a cryostat which is presented in this section. This is followed by introducing the sample chip, cryogenic circulators and amplifiers.

3.1.1 Dry dilution cryostat

The samples are cooled down to 10 mK in a cryostat equipped with a pulse tube refrigerator and a ^{3}He/^{4}He dilution refrigerator. The pulse tube cooling does not require a continuous supply of liquid helium or other cryogenic liquid; hence the term "dry" cryostat. The cryostat is a crucial part of the operating environment of the noise measurement setup, and it is thus important to introduce the basic structure, operating principle and cooling power limitations of the cryostat. While the details of the structure and operation of the cryostat are beyond the scope of this thesis, the interested reader is referred to Ref. [56] for a thorough introduction to cryoengineering.

Figure 3.1 shows the structure of a BlueFors LD250 cryostat [57]. All the low-temperature parts are in a vacuum of 10^{-6} mbar, contained in the vacuum can. The cryostat is equipped with two cooling systems: A pulse tube refrigerator for cooling from room temperature down to 4 K and a dilution refrigerator from 4 K downwards. The pulse tube refrigerator has two cooling stages in series: first stage cools down to 60 K and the second stage to 4 K. The operation of a pulse tube

is based on periodic pressure variations of compressed helium gas (10–20 bar) in a tube [56, pp. 133-134]. The cryostat used in this thesis has a Cryomech PT410-RM pulse tube refrigerator with 0.9 W cooling power at 4 K stage [58].



(a) (b)

Figure 3.1: (a) Structure of the cryostat. (b) Photograph of a BlueFors LD series cryostat [57].

The dilution refrigerator is based on the heat of mixing of two helium isotopes: ^{3}He and ^{4}He. The mixing occurs in the mixing chamber, where the mixture is separated into two phases: the ^{3}He rich and dilute phases. The cooling power of this process at 100 mK is 250 μW (and less in lower temperatures) for the cryostat used in the thesis [57]. From the mixing chamber, the ^{3}He of the dilute phase moves upwards towards the still through superfluid ^{4}He. The ^{3}He evaporates in the still, which requires several milliwatts of external heating (or heat load from thermal anchoring of cables) in order to maintain the optimal temperature of 0.7 K. The evaporated ^{3}He is pumped from the still using a room-temperature turbomolecular pump backed with a scroll pump. The ^{3}He returns to the mixing chamber after pre-cooling in heat exchangers which transfer the heat to the still and to the ^{3}He flowing up from the mixing chamber [56, Ch. 7].

3.1.2 Sample

Figure 3.2 shows a measured sample mounted to a copper sample box with SMA connectors. The measured nanoscale device is patterned on top of a silicon or sapphire chip with its electronic contacts connected to a printed circuit board (PCB) with bonding wires. Both DC and RF are connected with coaxial cables to the SMA connectors of the box. It is also possible to apply DC bias and measure RF noise from the same connector using a bias-tee which separates the DC and RF signals.

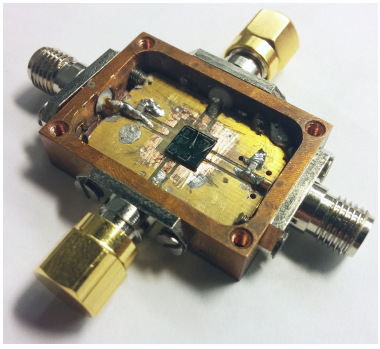


Figure 3.2: Sample box.

3.1.3 Circulators

A circulator is a passive, nonreciprocal, three-port microwave component which directs the microwave signal from one of its ports to the next port in one rotational direction. Therefore the device passes microwave power in only one rotational direction while the other direction is isolated. The operation of the device is based on ferrimagnetic material biased with a permanent magnet [59, p. 476].

In cryogenic measurement setups circulators are thermalized to the mixing chamber, typically in isolator configuration with one port terminated with a matched load. If the measurement requires an RF input signal, the port can be connected to an attenuated microwave input line instead of termination. In both configurations the circulator prevents amplifier back-action noise from reaching the sample. The insertion loss of a cryogenic circulator is approximately 1.5 dB while the isolation in reverse direction reaches 20 dB. The circulator operates at a finite bandwidth of maximum one octave, thus limiting the operating bandwidth of the measurement system. It should be noted that a special cryogenic circulator should be used instead of a regular room-temperature device since the behavior of ferrimagnetic materials change significantly when the device is cooled down.

3.1.4 Amplifiers

The weak microwave signal is amplified prior to detection using both cryogenic and room-temperature low-noise amplifiers based on high electron-mobility transistors (HEMTs). Being active semiconductor devices, the amplifiers add noise to the measurement system, and therefore amplifier noise becomes dominant in typical measurement scenarios. The amount of amplifier noise is expressed with *noise temperature* which is by definition the temperature of a matched load at the input of a noiseless amplifier that produces same amount of output noise as a noisy amplifier with no noise at its input [34, p. 207]. A typical cryogenic amplifier exhibits noise temperature of approximately 10 K while the noise temperature of the whole measurement system is also affected by other devices on the signal chain. The system noise temperature is given by

$$T_{tot} = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 G_2} + ... + \frac{T_N}{G_1 G_2 ... G_{N-1}}, \tag{3.1}$$

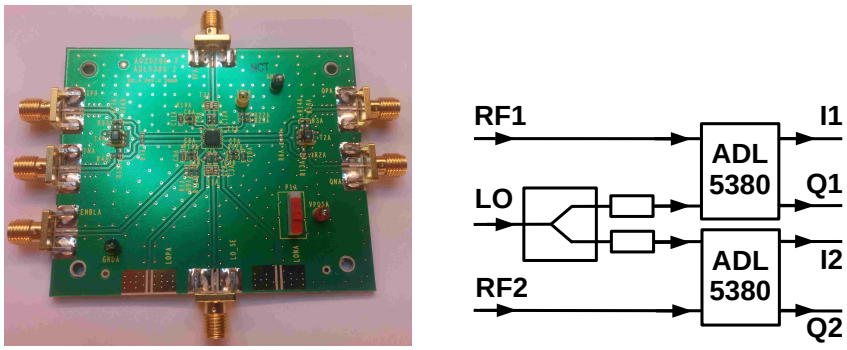
where T_i and G_i denote the noise temperature and gain of the i$^{\text{th}}$ amplifier, respectively [60]. It should be noted that the contribution of an amplifier is reduced by the gain of the preceding amplifiers. Moreover, lossy devices including cables and attenuators can be included in the above formula with less-than-unity gain. The noise temperature at the input of an attenuator is given by

$$T_L = T_F \left(L - 1 \right), \tag{3.2}$$

where T_F and L denote physical temperature and attenuation, respectively.

3.2 IQ mixer device

The amplified RF signals of the two channels are down-converted in frequency using two ADL5380 quadrature demodulators (IQ mixers) from Analog Devices [61]. The device is a single chip operating at 0.4 – 6 GHz RF bandwidth. The actual device is available mounted on a PCB with all the necessary DC bias and RF transformer components. The reference board presented in Fig. 3.3 (a) was used in this thesis instead of a time-consuming custom PCB design. The two devices share a split LO signal in order to maintain a constant phase reference between the mixers. A schematic of the two-channel IQ mixer device is presented in Fig. 3.3 (b).



(a) (b)

Figure 3.3: (a) Photograph of ADL5380 quadrature demodulator reference board. (b) Schematic of the two-channel IQ mixer device used in this thesis.

The design aims for low excess correlation level, which is obtained by good electrical isolation of the two channels. The mixer boards were mounted in separate diecast aluminium enclosures to prevent interference from the other board and from the environment. Small grooves were made for the coaxial cables between the box and the cover. Copper tape with conductive adhesive was used to improve the isolation

at the small openings. The DC bias for the demodulator was fed through a capacitive feedthrough filter. The aluminium enclosures were mounted in a 2U high, 19-inch rack enclosure. The rack enclosure includes feedthrough connections for all RF, LO and IF signals as well as for DC power supply. The 5 V DC supply is provided by two separate LM7805 regulator circuits with sufficient electrical filtering [62]. The regulator circuits are connected to separate 12 V lead acid batteries located outside the box. This separation of power supply circuitry prevents the mixer boards from coupling through the DC supply lines. A photograph of the IQ mixer device is presented in Fig. 3.4.

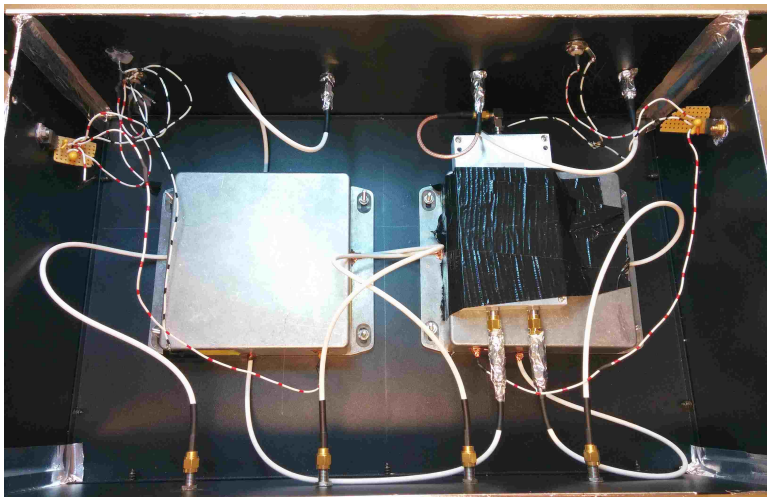


Figure 3.4: Photograph of the IQ mixer device with top cover open. Mixer boards mounted inside the two aluminium enclosures; power splitter situated on top of the right enclosure; DC regulators attached to both side walls; RF, LO and DC connectors on upper wall and IF connectors on lower wall.

The LO signal is split to the mixer boards using a ZN2PD2-63+ power splitter from Mini-Circuits. The device has a bandwidth of 0.35 – 6 GHz, thus covering the entire bandwidth of the IQ mixers. The isolation between the output ports is better than 17 dB throughout the bandwidth. A good isolation between the LO inputs of the mixer boards is necessary in order to prevent coupling and hence reduce the excess correlation level. However, the manufacturer-specified isolation level is valid only when ports are terminated with 50 Ω reference impedance, which is not neccessarily the case for the LO inputs of the mixer boards. Therefore, 7 dB attenuators were connected to the output ports of the splitter improving the impedance matching and increasing the attenuation between the LO input ports. The LO was fed to the mixer box from a signal generator at +10 dBm level to provide both mixers the manufacturer-specified LO input level, 0 dBm.

The mixer box is situated outside the shielded room of the cryostat. The RF and LO input connectors of the mixer box were equipped with DC-blocks to break the galvanic connection of the inner and outer conductors of the coaxial lines. This is important for preventing low-frequency interference from entering the shielded room

via coaxial cables. The IF ports were connected directly to the input ports of the data acquisition card mounted on a desktop computer. This means that the enclosure of the mixer box shares its ground with the computer, therefore making it more important to separate it from the ground of the cryostat and shielded room.

3.3 Data acquisition

The IF signals are captured to a desktop PC with ATS9440 digitizer from AlazarTech [63] which is presented in Fig. 3.5. The digitizer card is capable of capturing four channels at 125 MS/s (megasamples per second) with 14 bit resolution. The captured data is transferred from the card to the computer random access memory (RAM) with direct memory access (DMA) transfer. The data is stored in 16-bit unsigned integer format with the samples from different channels interleaved (ABCD-ABCD... instead of separating the channels). The ability to transfer data directly to GPU memory is not enabled in the device drivers without extra charge, and therefore techniques for these direct transfers (NVIDIA GPUDirect RDMA [64] or AMD DirectGMA [65]) are not utilized in this thesis. Moreover, these techniques would not increase the performance of the setup, since data transfer is not the limiting factor to performance.

Figure 3.5: AlazarTech ATS9440 capture card with PCI Express interface [63].

3.4 Data processing

The data is captured in buffers of $2^{22} \approx 4$ million samples per channel, corresponding to 34 ms of capture time at 125 MS/s sample rate. The cross-correlation is then calculated for each buffer, and the final result is obtained as an average over multiple (up to several hundreds) buffers. The challenge is that the processing has to occur in real time, meaning that the previous buffer has to be processed before the next buffer is transferred to RAM. This section describes the implementation of a high-performance data processing algorithm running on GPU. A more

compact mathematical description of the program functionality is provided in Appendix A.

3.4.1 Choice of implementation

At first, the data processing was implemented with Matlab which supports GPU acceleration of various operations, including FFT. This was beneficial since all other measurement devices are controlled with Matlab, and the capture card manufacturer provided example programs written for Matlab. However, it was found out that the speed of data processing on Matlab was not sufficient to keep up with the high data output rate of the capture card. This was mainly due to inefficiency in data transfer operations. Therefore, more control over the transfer operations was required in order to speed up the data processing.

Implementation with the C programming language was chosen due to its high performance and ability to control memory processes at a low level. In addition, the capture card manufacturer provided example programs to speed up the development. C is also supported by NVIDIA CUDA (Sect. 2.8.3) with language extensions for GPU programming as well as CUFFT library, a ready-made implementation of FFT.

The source code of example programs from NVIDIA and AlazarTech were used as a starting point for the development process. The NVIDIA example program was a CUFFT-based implementation of complex convolution, an operation similar to complex correlation. The AlazarTech example included adjusting the settings of the digitizer card, starting a capture operation and transferring the captured data to RAM. Combining these example programs provided a basis for the core functionality of the data processing program.

3.4.2 Data transfer

Transfer of the raw data from RAM to GPU memory forms a time-consuming part of the whole process, which can be seen in Fig. 3.6. The total amount of data is 2^{25} bytes = 32 MB for each buffer. The transfer process is implemented with DMA transfer, which allows processing of the previous buffer on the GPU while simultaneously transferring the next buffer to the GPU memory. The DMA transfer requires allocation of *pinned memory* from the RAM, in contrast to virtual memory normally allocated by the operating system.

When the raw data is transferred to the GPU memory, it consists of 16-bit unsigned integers in an interleaved array. This array is converted to two arrays of complex numbers in 32-bit floating point format. This function is implemented as a *CUDA kernel*, which is run in multiple instances simultaneously, each of which operates on a different part of the input array. The function converts the unsigned integers to signed floating point format, moving the value 0 to correspond with zero input

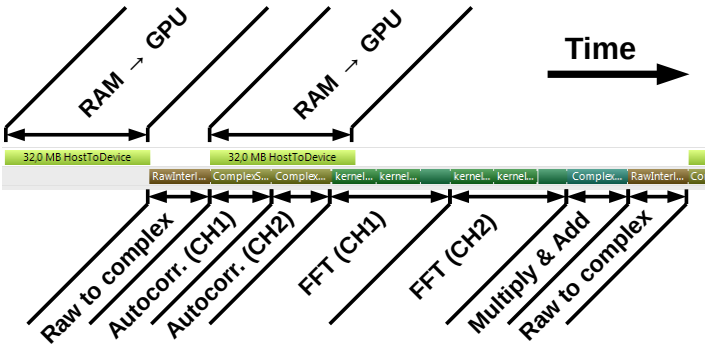


Figure 3.6: Profiler timeline of data processing run (without capture) showing relative durations of functions.

voltage level. In addition, the function divides the values by four, since the raw values are stored in the 14 most significant bits of the 16-bit integer. The function is the only part of the data processing which can not be run while data transfer is in progress, since the data would be transferred to the array processed by the function. Therefore, the data transfer from RAM to GPU is triggered to start immediately after the function is completed.

3.4.3 Autocorrelation

The zero-time-lag autocorrelations are needed for normalization of the cross-correlation (Eq. 2.27) as well as for diagnostic purposes. The autocorrelation is calculated using Eq. (2.24), which reduces to dot product of the vector with itself when only zero-time-lag is required. The multiplications of the dot product can be calculated in parallel while the summation requires a tree-like structure to benefit from parallel architecture. This sum-reduction tree is illustrated in Fig. 3.7 which presents the summation of eight integers. Each step consists of individual summations, whose results are used in the next step. The sum-reduction tree requires synchronization between different instances of the CUDA kernel in order to avoid the next level of the tree from running before the previous level is completed. It should be noted that the degree of parallelism reduces as the calculation advances, with only single calculation during the last step. Therefore, the sum-reduction tree causes slightly inefficient use of the GPU, which reduces the performance increase compared to CPU implementation. This can be seen in the profiler timeline presented in Fig. 3.6 as the duration of autocorrelation is relatively long compared to the low complexity of the operation.

3.4.4 Cross-correlation

The calculation of cross-correlation is implemented by multiplication in frequency domain (Eq. 2.25) to enable frequency-domain filtering of the result. Since the final

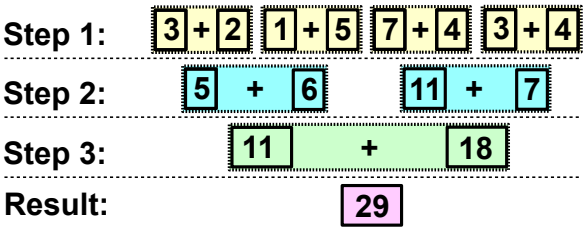


Figure 3.7: Sum-reduction tree for eight integers.

result is obtained as an average over multiple buffers, the processing would involve cumulative summation of the time-domain correlations for each buffer, in which case the correlation would be calculated as

$$z_{cor}[t] = \frac{1}{M}\sum_{k=1}^{M} IDFT\left[\frac{1}{N}X_k^*[f]\cdot Y_k[f]\right], \tag{3.3}$$

where M is the number of buffers. However, due to the linearity of Fourier transform (Eq. 2.23), the above equation can be written as

$$z_{cor}[t] = \frac{1}{M} IDFT\left[\sum_{k=1}^{M}\frac{1}{N}X_k^*[f]\cdot Y_k[f]\right], \tag{3.4}$$

meaning that averaging can be done in frequency domain. This reduces the number of required inverse FFTs from M to one which is applied when the averaging is completed. During each averaging round the complex FFTs of the two channels are calculated with a simple kernel call of the CUFFT library which automatically uses optimal algorithm for calculating the FFTs. This step is followed by point-wise multiplication of the complex arrays and addition of the result to a cumulative result vector. The step also includes the complex conjugation of one of the arrays and scaling of the result with the geometric mean of the aurocorrelations.

After the final averaging the 20 000 lowest frequency points are removed from the frequency-domain correlation array in order to reduce excess correlation caused by low-frequency interference. Following this, the result is converted to time-domain by applying inverse FFT. 100 points around the zero-time-lag are transferred from the GPU memory to RAM and written to a data file along with the averaged autocorrelation values. This file is then read to a Matlab program which acts as an interface between measurement scripts running on Matlab and the compiled C program.

3.4.5 Possible modifications

The above description concentrated solely on voltage auto- and cross-correlations omitting V^2 correlations, also known as power or intensity correlations. However,

V^2 correlations are important in certain measurement schemes and thus they are included in the data processing program as an optional feature. This involves pointwise squaring of the input vectors and reduction of the mean from the squared vectors prior to FFT.

FFT windowing (Sect. 2.8.1) should be applied prior to Fourier transform when observing measurement results as a function of frequency. This is important when measuring sharp spectral features in resonating structures. The windowing operation involves pointwise multiplication of the input arrays with the desired window function. However, the implementation of FFT windowing was omitted from this thesis in order to narrow down the topic.

It should be noted that the calculation of the above mentioned modifications increases the amount of required computational operations compared to the standard configuration, which might lead to performance issues. However, in such circumstances the GPU and other hardware can be upgraded to increase computational power while no modifications to the software are required. Additionally, modern chipsets support multiple GPUs which can be utilized if the performance of a single GPU is not sufficient. Thus, the designed implementation can be scaled according to the demands of the measurement.

The motherboard and chipset of the computer were found to significantly affect the system performance. Initially, the system was incapable of achieving the full 125 MS/s data capture rate, and hence 100 MS/s was used instead. This problem was most probably caused by data transfer between the capture card and computer RAM, and it was fixed by upgrading the motherboard and CPU of the computer. However, a majority of the results presented in this thesis were already measured with the old system.

Chapter 4

Results

This chapter reports measurement results of the spectrometer designed in this thesis. The chapter starts with the comparison of data processing performance between the C and Matlab implementations. In addition, the chapter presents results of phase and amplitude imbalance between the two channels and their I- and Q-branches. This is succeeded by measurement of excess correlation level when non-correlated noise is connected to the inputs. Following this, the sensitivity of the system is evaluated and compared against theoretical formula. Finally, measurement results of a graphene sample are presented to demonstrate the functionality of the system with a real nanoscale sample.

4.1 Data processing performance

One of the goals of this thesis is the implementation of a high-performance data processing program for correlation measurements. The performance of the program can be measured by the amount of data it can process at a given time period. A reasonable choice for the unit is megasamples per second (MS/s) which is easily comparable to the capture rate of the digitizer card (125 MS/s). The data processing speed has to be greater than the data capture rate to enable arbitrary long averaging periods.

The test was conducted on the measurement computer with Intel Core i5-3570K processor and Nvidia GeForce GTX 760 graphics processing unit. The processed data consisted of an interleaved array of random 16-bit unsigned integers in computer RAM, read repeatedly during each averaging round. The use of the digitizer card would fix the data rate at 125 MS/s and it is thus not suitable for performance testing. The absence of the digitizer card may lead to higher performance than in the actual measurement. The developed C program was compared against Matlab implementations of the same functionality both with and without GPU-based functions. The Matlab scripts were optimized to achieve near-maximum performance,

although even better performance could have been achieved with further optimization. However, the results give an approximate picture of the data processing speeds of different implementations.

The results are shown in Fig. 4.1. As it can be seen, neither of the Matlab implementations achieve the speed required for data capturing which is 125 MS/s, although the use of GPU speeds up the calculation by 4.5 times. The C program achieves 567 MS/s which is clearly sufficient, thus enabling the calculation of additional quantities described in Sect. 3.4.5 while maintaining the required performance. The performance increase of the C implementation is 13 times compared to Matlab utilizing GPU and 60 times compared to Matlab utilizing CPU only. These results indicate that the better controllability of data transfers in the C program over the Matlab implementations result in significant incerease in performance.

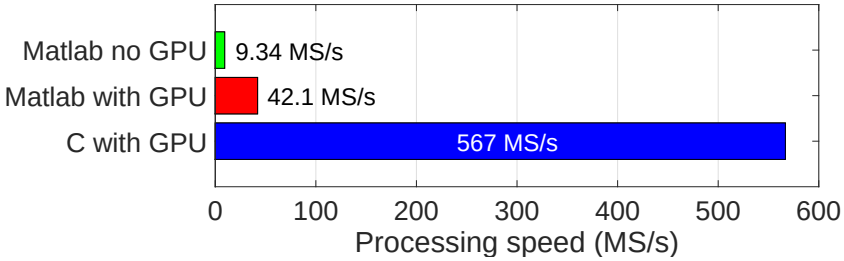


Figure 4.1: Measured speeds of different implementations of complex cross-correlation.

Another figure of merit for digital correlation setups is the relation of data capture time to total measurement time, namely *duty cycle*. The setup of this thesis captures data continuously without gaps during an averaging round, and therefore the duty cycle is limited by the time required for memory allocation and data output before and after the capture operation. Since the duration of these operations is independent of the amount of captured buffers, the duty cycle improves as the integration time is increased. This constant delay is on the order of one second, which leads to 80 % duty cycle with the typical integration time of four seconds. This duty cycle is already higher than the 54.6 % of the Munich setup (Sect. 2.9.2). The duty cycle could be improved by running the program continuously on the background instead of restarting the program for each averaging round. However, this would require considerable programming effort relative to the resulting improvement of duty cycle.

4.2 Channel imbalance

The four captured channels are assumed to have the same amplitude and phase characteristics, except from the 90° phase separation between I and Q channels caused by the IQ mixers. The amplitude and phase imbalance caused by the IQ

mixer device and the digitizer card was tested using the connection illustrated in Fig. 4.2. A sine wave was fed to the LO port of the IQ mixer at frequency f_0 while the RF ports shared a split sine wave at frequency $f_0 + \Delta f$. In ideal case the IF signals at frequency Δf should all have the same amplitude and the phases of I and Q should differ by 90°. The test was conducted by capturing a time trace of all four channels simultaneously and fitting a sine function to the data. One set of these time traces are plotted in Fig. 4.3 along with the sine fits which can be seen to follow closely the captured data. The test was conducted for several values of f_0 and Δf.

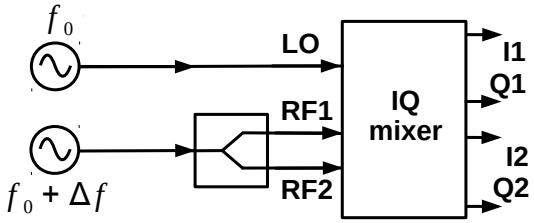


Figure 4.2: Setup used for channel imbalance measurement.

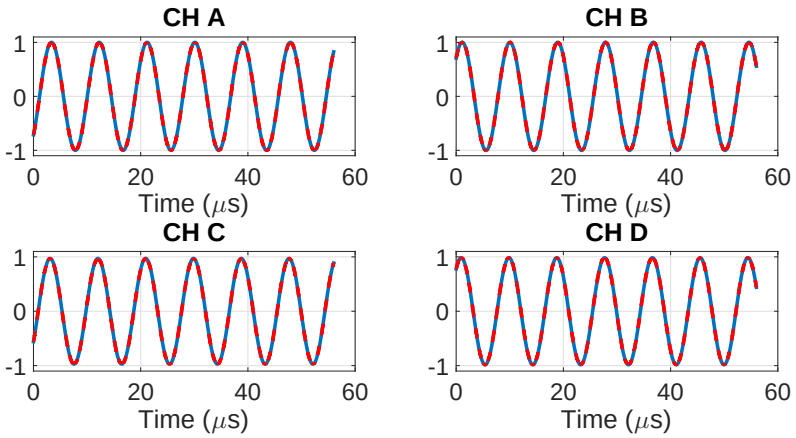


Figure 4.3: Time traces from a single capture of channel imbalance measurement (blue solid curve) and the fitted sine functions (red dashed curve). $f_0 = 1$ GHz, $\Delta f = 100$ kHz.

The amplitudes and initial phases of the fitted sine functions were compared against each other. The results indicate that individually the both IQ mixers have good balance between I and Q outputs with amplitude imbalance of less than 1 % and phase difference less than 2° away from the ideal 90°. The imbalance between the two mixers is higher, with amplitude and phase differing 3 % and 10°, respectively. This is most likely caused by the power splitter and coaxial cables inside the two-channel IQ mixer device. This imbalance is presumably not a problem in practical measurements since the imbalance caused by the amplifiers and cables of the RF lines is typically more significant. However, the imbalance can be corrected with adjustable phase shifters and attenuators if required for a measurement.

4.3 Excess correlation

Ideally, the correlated part of the observed noise should originate solely from the measured sample. However, there exists a finite amount of correlation due to improper separation of the two signal lines, referred to as *excess correlation*. It affects the measurement when sample noise is of the same order or lower than the level of excess correlation. In addition, it degrades the stability of the system due to possible drift in the power level of excess correlation. Therefore, the excess correlation limits the system performance and hence its magnitude should be determined via measurement.

In initial tests, the complete system showed a correlation level of 3×10^{-3} relative to total noise power when the sample was unbiased with no change observed when the input ports were terminated separately. The level of excess correlation was higher than the weakest measured shot noise signals, and therefore its origin was traced in order to find proper measures for reducing it.

The excess correlation level related to the capture card was tested by capturing noise generated by two function generators and calculating the correlation. However, no correlation peak above the background level was observed, and hence the capture card caused no excess correlation. After this, the role of the IQ mixer device was tested using two separate room-temperature noise sources to feed uncorrelated noise to the RF ports of the IQ mixer device. At first, the 3×10^{-3} correlation level was observed, although it disappeared with better separation of the noise sources and proper connection of grounding. This confirmed that the excess correlation originates from the two RF signal chains together with grounding issues and not from the IQ mixer or the digitizer card.

The exact origin of excess correlation was not traced in this thesis, although the problem is presumably caused by crosstalk between the inputs and outputs of amplifier chains. Another mechanism is the coupling of grounding interference to the IQ mixer device. However, in practical measurements the correlation peaks from excess correlation and measured sample occur at different indexes of the time-domain correlation array, and hence they are distinguishable from each other. Therefore, the presence of excess correlation does not prevent typical correlation measurements, although it should be eliminated eventually to enable more sophisticated measurements involving the shape of the time-domain cross-correlation function.

4.4 Sensitivity

The sensitivity of the setup was determined by calculating the standard deviation of a sequence of noise measurements. This assumes the noise power to stay stable during the measurement. The schematic of the test setup is presented in Fig. 4.4. The noise from a diode noise source with 10 000 K noise temperature (Micronetics

NS346B) is fed to the cryostat through a variable attenuation. Alternatively, a 50 Ω terminator is used as a noise source, producing room-temperature thermal noise at 290 K. The noise is attenuated in various temperature stages in the cryostat with a total attenuation of 45 dB. The noise is split to two channels with a power splitter thermalized to mixing chamber (10 mK). The splitter is followed by circulators in order to suppress excess correlation due to coupling of amplifier inputs. The remaining part of the signal chain is similar to that described in Ch. 3 with cryogenic and room-temperature amplifiers, IQ mixer and data acquisition.

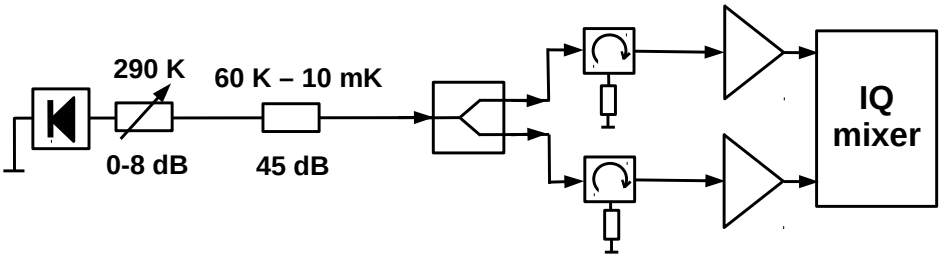


Figure 4.4: Simplified schematic of the configuration used for determining the sensitivity of the measurement setup.

The normalized cross-correlation is the ratio of correlated noise to total noise power, which is dominated by amplifier noise. The system is intended for measuring low ratios of correlated noise, ranging from zero to few percent of total power. One percent corresponds to 0.1 K when the system noise temperature is 10 K. The measured relative noise powers with different input noise levels are listed in Table 4.1.

Table 4.1: Relative power of correlated noise with respect to total noise measured.

Noise input	$P_{\mathrm{corr}}/P_{\mathrm{tot}}$ (%)
290 K terminator	1.61
Noise source, 8 dB attenuation	3.40
Noise source, 3 dB attenuation	7.05
Noise source, 0 dB attenuation	11.8

The measurement was repeated with integration times ranging from 0.04 s to 8 s in order to observe the effect of the amount of sampled points on system sensitivity. Longer integration times were not used due to limitations caused by the measurement computer. For the same reason, 100 MS/s sampling rate was used instead of the 125 MS/s maximum rate. The measurement was repeated several (91–560) times to obtain time series from which the standard deviation of the normalized cross-correlation was calculated. The determined quantity is standard deviation of temperature relative to system noise temperature. This was chosen over absolute standard deviation of temperature due to the uncertainty related to the exact value of system noise temperature.

The determined standard deviation of the measurement is presented in Fig. 4.5 together with the theoretical sensitivity obtained from Eq. 2.17. It can be seen

that the results agree well with the theory prediction at the shortest integration period while noticeable deviation occurs at longer integration times and higher input noise level. This is due to the fact that higher level of correlated noise has higher absolute variation of temperature. Hence the results with 5–8 dB attenuation and 290 K termination are the best to represent the sensitivity of the noise measurement setup. The relative sensitivity of 2.8×10^{-5} with 8.43 s integration time corresponds to 280 μK temperature sensitivity if the system noise temperature is 10 K. This translates to 813 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$ which is comparable to the sensitivity of the Paris setup (710 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$) (Sect. 2.9.1). The theoretical maximum sensitivity is 707 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$ for the 100 MS/s sample rate used in the measurement. If the results are scalable to full 125 MS/s rate, the measured and theoretical sensitivities would be translated to 727 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$ and 632 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$, respectively.

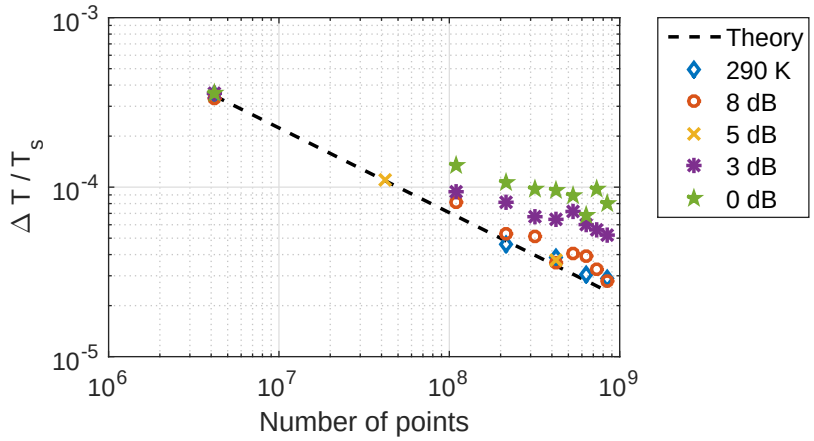


Figure 4.5: Measured relative sensitivity of the system with 50 Ω terminator at 290 K and noise source with 0–8 dB attenuation. Theoretical sensitivity from Eq. 2.17.

4.5 Shot noise of graphene sample

The setup was used in shot noise correlation measurement of a graphene sample and the results are presented in this section in order to demonstrate the functionality of the setup in real shot noise measurements. The measured sample was a box-shaped single layer graphene with four electrodes at the corners as shown in Fig. 4.6. The sample was fabricated in University of Würzburg, Germany by Shijin Babu. Shot noise was measured from terminals 2 and 3 while the two other terminals were used for DC biasing of the sample. The purpose of these measurements was to investigate Hanbury-Brown and Twiss type of interference in graphene by measuring the sample in different DC bias configurations [38]. More specifically, the purpose was to study the effect of high magnetic field perpendicular to the sample in addition to earlier results measured in another setup [66]. However, the sample was not fully functional due to increase in contact resistance compared to earlier measurements.

Nevertheless, the sample generated sufficient amount of shot noise for characterizing the system performance.

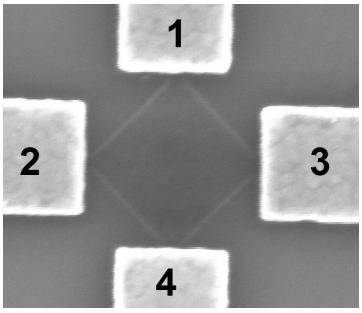


Figure 4.6: Scanning electron microscope (SEM) image of the sample with the four electrodes numbered, length of graphene edge is 360 nm.

The results of the shot noise measurement are shown in Fig. 4.7. The results are obtained in a configuration where V_{bias} is applied to terminal 1 while terminal 4 is grounded. The integration time was 4.19 s at 100 MS/s sampling rate. The effect of excess correlation was avoided since the correlation peaks from the measurement and excess correlation occured at different indexes of the time-domain correlation array. It can be seen from the figure that the correlation is $< 2 \times 10^{-4}$ at zero bias which is an order of magnitude lower than the excess correlation. The cross-correlation is suppressed as the amplitude of the magnetic field is increased from 0 T to 8 T.

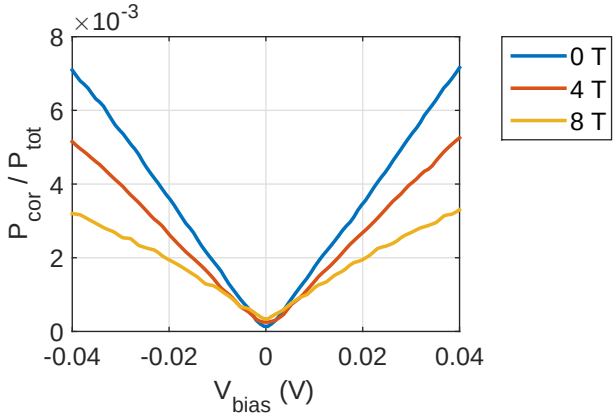


Figure 4.7: Normalized cross-correlation of shot noise from graphene sample as a function of bias voltage in magnetic field of different magnitudes.

Chapter 5

Conclusion

This thesis presented a novel type of correlation spectrometer intended for high-sensitivity shot noise measurements at low temperatures. The spectrometer digitizes both quadratures of two down-converted RF signals and processes the data on a GPU mounted in a desktop computer. The sensitivity of the spectrometer was found to be consistent with theoretical formula, and the high performance of GPU data processing enables the calculation of additional quantities in real time.

The spectrometer designed in this thesis includes a two-channel IQ mixer device consisting of two commercial IQ demodulator boards sharing a common LO signal. The device was designed to prevent interference to the signal paths with metallic enclosures and proper filtering. The amplitude and phase balance of the device were measured to be adequate, with imbalance of the remaining system being more significant.

The implemented system employs a data acquisition card which captures four channels at 125 MS/s with 14-bit resolution. The card is mounted in a desktop computer using PCI Express bus. The data is processed on a GPU, with a program implemented using CUDA C, an extension to the C programming language. The parallel hardware architecture of the GPU increases the computational performance, although programming becomes more challenging. However, the data processing required for this thesis involved mainly straightforward algorithms of multiplication, summation and data copying while the relatively complicated FFT calculation made use of the ready-made CUFFT library. Therefore, the amount of software development involved in this thesis remained at a reasonable level.

GPU was found to be a suitable data processing platform owing to the high degree of parallelism involved in the calculation. Moreover, the CUDA C implementation outperformed GPU-accelerated Matlab script due to more flexible memory handling and data transfers. The speedup was determined to be 13-fold which justifies the relatively demanding programming effort. In broader perspective, the use of GPU and CUDA C follows the ongoing trend of parallel computing in scientific applications.

The GPU-based data processing is highly modular in nature, since the program is not hardware-dependent. New features can be added to the program with the hardware performance being the limiting factor. In this case the GPU can be upgraded to a more powerful model without changes to the software. This modularity is in contrast to the FPGA-based data processing used in the Munich setup (Sect. 2.9.2). The proposed additions include intensity (V^2) correlations among higher moments of noise.

The performance of the implemented setup is limited by excess correlation resulting from the cryostat or amplifier crosstalk. The level of this excess correlation relative to total measured noise power is 3×10^{-3} which is of the same order as typical correlation level in shot noise measurements. However, the occurence of the excess correlation peak to a different index than the sample correlation enables most measurements in spite of the excess correlation. Nevertheless, the exact source of excess correlation should be traced and its level reduced in order to improve the performance of the setup.

The sensitivity of the system was measured up to 8.43 s integration time showing little deviation from the theoretical maximum sensitivity. The related temperature sensitivity was determined to be 810 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$ when the theoretical maximum is 710 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$ for the 100 MS/s sample rate used in the measurement. The proximity of the theoretical maximum narrows the possibilities for improving the setup, besides upgrading the capture card and lowering system noise temperature. Moreover, the values are close to the sensitivity of the Paris setup, namely 710 $\mu\mathrm{K}/\sqrt{\mathrm{Hz}}$.

To conclude, nanophysics research will benefit from the novel approach of digitizing and processing data instead of measuring with solely analog circuitry. This approach requires a cross-disciplinary combination of physics, radio engineering and computer science to develop new types of measurement systems. This, together with other advancements speeds up nanoelectronics research, helping to maintain the ongoing digital revolution started from the integrated circuit nearly six decades ago.

References

[1] J. S. Kilby, *Turning Potential into Realities: The Invention of the Integrated Circuit*. Nobel Lecture, 2000.

[2] D. E. Liddle, "The Wider Impact of Moore's Law," *Solid-State Circuits Newsletter, IEEE*, no. September, pp. 28–30, 2006.

[3] B. Rogers, S. Pennathur, and J. Adams, *Nanotechnology : The Whole Story*. Boca Raton, FL, USA: CRC Press, 2013.

[4] R. K. Yadav, *Discovering the Nano-Electronics*. Delhi, India: Mangalam Publishers, 2009.

[5] J. Xiang *et al.*, "Ge/Si nanowire heterostructures as high-performance field-effect transistors," *Nature*, vol. 441, no. 7092, pp. 489–493, May 2006.

[6] H. W. Postma *et al.*, "Carbon nanotube single-electron transistors at room temperature." *Science*, vol. 293, no. 5527, pp. 76–79, 2001.

[7] A. C. Ferrari *et al.*, "Science and technology roadmap for graphene, related two-dimensional crystals, and hybrid systems," *Nanoscale*, vol. 7, no. 11, pp. 4598–4810, 2014.

[8] D. McMahon, *Quantum Computing Explained*. Hoboken, NJ, USA: John Wiley & Sons, 2008.

[9] M. Joodaki, *Selected Advances in Nanoelectronic Devices: Logic, Memory and RF*. Berlin, Germany: Springer-Verlag, 2013.

[10] L. DiCarlo *et al.*, "System for measuring auto- and cross correlation of current noise at low temperatures," *Rev. Sci. Instrum.*, vol. 77, no. 7, p. 073906, 2006.

[11] R. C. Toonen, C. C. Haselby, and R. H. Blick, "An Ultrawideband Cross-Correlation Radiometer for Mesoscopic Experiments," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 12, pp. 2874–2879, 2008.

[12] J. Gabelli and B. Reulet, "Measurements of the Third Cumulant in Quantum Shot Noise at High Frequency," in *21st Int. Conf. Noise Fluctuations (ICNF), 2011*, no. 4, 2011, pp. 329–331.

[13] C. M. Wilson *et al.*, "Observation of the dynamical Casimir effect in a superconducting circuit." *Nature*, vol. 479, no. 7373, pp. 376–9, Nov. 2011.

[14] P. Lähteenmäki *et al.*, "Dynamical Casimir effect in a Josephson metamaterial." *Proc. Natl. Acad. Sci.*, vol. 110, no. 11, pp. 4234–4238, 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23404708

[15] F. Parmentier, "Short-time Hanbury-Brown and Twiss correlation of a single-electron beam in ballistic conductors," Ph.D. dissertation, l'Université Paris 6, 2010.

[16] M. Hashisaka *et al.*, "Cross-correlation measurement of quantum shot noise using homemade transimpedance amplifiers." *Rev. Sci. Instrum.*, vol. 85, no. 5, p. 054704, May 2014.

[17] E. P. K. Menzel, "Propagating Quantum Microwaves: Dual-path State Reconstruction and Path Entanglement," Ph.D. dissertation, Technische Universität München, 2013.

[18] NVIDIA Corporation, "About CUDA," 2015. [Online]. Available: https://developer.nvidia.com/about-cuda

[19] M. G. Knepley and D. A. Yuen, "Why Do Scientists and Engineers Need GPU's Today?" in *GPU Solut. to Multi-scale Probl. Sci. Eng.* Berlin, Germany: Springer-Verlag, 2013, ch. 1.

[20] W. Schottky, "Über spontane Stromschwankungen in verschiedenen Elektrizitätsleitern," *Ann. Phys.*, vol. 57, pp. 541–567, 1918.

[21] W. C. Etten, *Introduction to Random Signals and Noise*, 3rd ed. Chichester, United Kingdom: John Wiley & Sons, 2005.

[22] P. Horowitz and W. Hill, *The Art Of Electronics*, 2nd ed. Cambridge, UK: Cambridge University Press, 1989.

[23] Y. M. Blanter and M. Büttiker, "Shot noise in mesoscopic conductors," *Phys. Rep.*, vol. 336, no. 1-2, pp. 1–166, Sep. 2000.

[24] L. F. Spietz, "The Shot Noise Thermometer," Ph.D. dissertation, Yale University, 2006.

[25] H. Nyquist, "Thermal Agitation of Electric Charge in Conductors," *Phys. Rev.*, vol. 32, pp. 110–113, 1928.

[26] H. B. Callen and T. A. Welton, "Irreversibility and Generalized Noise," *Phys. Rev.*, vol. 83, no. 1, pp. 34–40, 1951.

[27] A. R. Kerr, "Suggestions for revised definitions of noise quantities, including quantum effects," *IEEE Trans. Microw. Theory Tech.*, vol. 47, no. 3, pp. 325–329, Mar. 1999.

[28] A. R. Kerr and J. Randa, "Thermal Noise and Noise Measurements — A 2010 Update," *IEEE Microw. Mag.*, vol. 11, no. 6, pp. 40–52, 2010.

[29] J. Randa *et al.*, "Recommended Terminology for Microwave Radiometry," *NIST Tech. Note*, vol. 1551, p. 32, 2008.

[30] G. Liu *et al.*, "Origin of 1/f noise in graphene multilayers: Surface vs. volume," *Appl. Phys. Lett.*, vol. 102, no. 9, p. 093111, 2013.

[31] E. Milotti, "1/f noise : a pedagogical review," 2002. [Online]. Available: http://arxiv.org/ftp/physics/papers/0204/0204033.pdf

[32] W. H. Press, "Flicker Noises in Astronomy and Elsewhere," *Comments Astrophys.*, vol. 7, no. 4, pp. 103–119, 1978.

[33] S. H. Kogan, *Electronic noise and fluctuations in solids.* Cambridge, UK: Cambridge University Press, 1996.

[34] A. Räisänen and A. Lehto, *Radiotekniikan perusteet*, 13th ed. Helsinki, Finland: Gaudeamus, 2011.

[35] B. F. Burke and F. Graham-Smith, *An Introduction to Radio Astronomy*, 3rd ed. New York, NY, USA: Cambridge University Press, 2010.

[36] H. Nyquist, "Certain Topics in Telegraph Transmission Theory," *Trans. AIEE*, vol. 47, no. 1, pp. 617–644, 1928.

[37] R. H. Dicke, "The Measurement of Thermal Radiation at Microwave Frequencies," *Rev. Sci. Instrum.*, vol. 17, no. 7, p. 268, 1946.

[38] Y. M. Blanter and M. Büttiker, "Shot-noise current-current correlations in multiterminal diffusive conductors," *Phys. Rev. B*, vol. 56, no. 4, pp. 2127–2135, 1997.

[39] J.-C. Forgues *et al.*, "Noise intensity-intensity correlations and the fourth cumulant of photo-assisted shot noise." *Sci. Rep.*, vol. 3, p. 2869, Jan. 2013.

[40] J. R. Piepmeier, A. J. Gasiewski, and J. E. Almodovar, "Advances in Microwave Digital Radiometry," in *IEEE Int. Geosci. Remote Sens. Symp.*, 2000, pp. 2830–2833.

[41] S. Weinreb, "Digital Radiometer," *Proc. IRE*, vol. 49, no. 6, p. 1099, 1961.

[42] B. F. C. Cooper, "Correlators with two-bit quantization," *Aust. J. Phys.*, vol. 23, no. 7, pp. 521–527, 1970.

[43] J. F. James, *A Student's Guide to Fourier Transforms*, 3rd ed. Cambridge, UK: Cambridge University Press, 2011.

[44] K. R. Rao, D. N. Kim, and J. J. Hwang, *Fast Fourier Transform: Algorithms And Applications*, 1st ed. Springer, 2010.

[45] E. W. Weisstein, *CRC Concise Encylopedia of Mathematics*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2002.

[46] F. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, 1978.

[47] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr., *Interferometry and Synthesis in Radio Astronomy*, 2nd ed. Weinheim, Germany: Wiley-VCH, 2004.

[48] A. V. Oppenheim and R. W. Schafer, *Discrete-time Signal Processing.* Prentice Hall, 1989.

[49] G. Heinzel, A. Rüdiger, and R. Schlilling, "Spectrum and spectral density estimation by the Discrete Fourier transform (DFT), including a comprehensive list of window functions and some new flat-top windows," Max-Planck-Institut für Gravitationsphysik, Hannover, Germany, Tech. Rep., 2002. [Online]. Available: http://holometer.fnal.gov/GH_FFT.pdf

[50] M. Kumar, "A study of electron scattering through noise spectroscopy," Ph.D. dissertation, Leiden University, 2012.

[51] J. D. Owens *et al.*, "A Survey of General-Purpose Computation on Graphics Hardware," *Comput. Graph. Forum*, vol. 26, no. 1, pp. 80–113, 2007.

[52] N. Govindaraju *et al.*, "High performance discrete Fourier transforms on graphics processors," *2008 SC - Int. Conf. High Perform. Comput. Networking, Storage Anal.*, pp. 1–12, Nov. 2008.

[53] M. A. Clark, P. L. Plante, and L. J. Greenhill, "Accelerating radio astronomy cross-correlation with graphics processing units," *Int. J. High Perform. Comput. Appl.*, vol. 27, no. 2, pp. 178–192, May 2012.

[54] R. B. Wayth, L. J. Greenhill, and F. H. Briggs, "A GPU-based Real-time Software Correlation System for the Murchison Widefield Array Prototype," *Publ. Astron. Soc. Pacific*, vol. 121, no. 882, pp. 857–865, 2009.

[55] L. DiCarlo, "Mesoscopic Electronics Beyond DC Transport," Ph.D. dissertation, Harvard University, 2007.

[56] F. Pobell, *Matter and Methods at Low Temperatures*, 3rd ed. Berlin, Germany: Springer-Verlag, 2007.

[57] BlueFors Ltd., *Helsinki, Finland.* [Online]. Available: http://www.bluefors.com/

[58] Cryomech Inc., *Syracuse, New York, USA.* [Online]. Available: www.cryomech.com

[59] D. M. Pozar, *Microwave engineering*, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, 2005.

[60] H. T. Friis, "Noise Figures of Radio Receivers," *Proc. IRE*, vol. 32, no. 7, pp. 419–422, 1944.

[61] Analog Devices Inc., *ADL5380 data sheet (Rev. A)*, 2013.

[62] Fairchild Semiconductor, *LM78XX data sheet*, 2014. [Online]. Available: https://www.fairchildsemi.com/datasheets/LM/LM7805.pdf

[63] AlazarTech, "ATS9440 product page," 2014. [Online]. Available: http://www.alazartech.com/products/ats9440.htm

[64] NVIDIA Corporation, "GPUDirect product page," 2015. [Online]. Available: https://developer.nvidia.com/gpudirect

[65] Advanced Micro Devices Inc., "DirectGMA product page," 2014. [Online]. Available: http://developer.amd.com/community/blog/2014/09/08/amd-firepro-gpus-directgma/

[66] T. Nieminen *et al.*, "Shot Noise Correlations in Graphene Box," in *Proc. 49th Annu. Meet. Finnish Phys. Soc.*, Helsinki, Finland, 2015.

Appendix A

Calculation of cross-correlation

This appendix presents all the mathematical operations the C program uses for calculating voltage cross-correlation and autocorrelation. The purpose of this appendix is to provide a detailed picture of the calculation from the viewpoint of the end-user, and therefore the details of the implementation are omitted.

The capture card captures 4 channels at 125 MS/s. An individual sample consists of voltage values corresponding to the two RF channels:

$$\begin{aligned} &\text{Channel 1}: \ x_1 \ , \ y_1 \\ &\text{Channel 2}: \ x_2 \ , \ y_2 \end{aligned} \tag{A.1}$$

where x is the I component and y is the Q component of the down-converted RF signal coming from the IQ mixer. The data is transferred to GPU as several *buffers* of 2^{22} samples ($\sim 4 \times 10^6$). Data is represented in two time-domain complex arrays of the form

$$\begin{aligned} \bar{C_1}(t) &= \bar{x_1} + i\bar{y_1} \\ \bar{C_2}(t) &= \bar{x_2} + i\bar{y_2}. \end{aligned} \tag{A.2}$$

Size of these arrays is 2^{22}.

Cross-correlation of one buffer

Cross-correlation of one buffer is obtained in frequency domain as pointwise multiplication of FFT($\bar{C_1}(t)$) and complex conjugate of FFT($\bar{C_2}(t)$). The resulting frequency-domain cross-correlation array is denoted by

$$\bar{C_{12}}(\omega) = \text{FFT}\left(\bar{C_1}(t)\right) \circ \overline{\text{FFT}\left(\bar{C_2}(t)\right)}, \tag{A.3}$$

where $\circ$ denotes elementwise product. Also this array consists of 2^{22} complex values.

Autocorrelation

Autocorrelation (or channel power) is given for channel 1 by

$$|\bar{C_1}(t)|^2 = \sum_{1}^{2^{22}} x_1 x_1 + y_1 y_1 \tag{A.4}$$

This is just a dot product of a vector with itself. The result is a positive real number.

Normalization of cross-correlation

The frequency-domain cross-correlation is normalized pointwise by

$$\bar{C_{12}}'(\omega) = \frac{\bar{C_{12}}(\omega)}{\sqrt{|\bar{C_1}(t)|^2 |\bar{C_2}(t)|^2}}. \tag{A.5}$$

This is done separately for each buffer.

Averaging of cross-correlation

The average is calculated in frequency domain by

$$\bar{C}_{12}^{\text{avg}}(\omega) = \frac{\sum_1^N \bar{C_{12}}'(\omega)}{N}, \tag{A.6}$$

where N is the number of buffers.

Final value of autocorrelation

The final output value of autocorrelation is given (for channel 1) by

$$|\bar{C}_1'(t)|^2 = \sum_{1}^{N} |\bar{C_1}(t)|^2. \tag{A.7}$$

This is equivalent to dot product of the vector with itself over all the samples in N buffers. Hence the buffer size doesn't affect the final value of autocorrelation in any way.

Neglecting low frequencies from cross-correlation

The kHz frequencies of cross-correlation include severe excess correlation peaks, most likely caused by interference in the IF lines. Therefore, they are neglected simply by making an array of $(2^{22} - 20\,000)$ points, from which the 20 000 lowest frequency values of $\bar{C}_{12}^{\mathrm{avg}}(\omega)$ are left out. This new reduced vector is denoted by $\bar{C}_{12}^{\mathrm{red}}(\omega)$.

Final output of cross-correlation

The averaged cross-correlation is converted to time domain by

$$\bar{C}_{12}^{\mathrm{red}}(t) = \mathrm{IFFT}\left(\bar{C}_{12}^{\mathrm{red}}(\omega)\right). \tag{A.8}$$

From this, the 100 points around zero time lag are written to a file that is read to Matlab. Then, one of these values (usually zero-time-lag) is chosen and outputted from the Matlab function as a scalar complex cross-correlation value.