Signal Processing Systems
Mekelweg 4,
2628 CD Delft
The Netherlands
https://sps.ewi.tudelft.nl/

Terahertz Sensing Group
Mekelweg 4,
2628 CD Delft
The Netherlands
https://terahertz.tudelft.nl/

M.Sc. Thesis

Photon Rate Estimation in MKIDs

B.Sc. Paul Julian Rozestraten

Abstract

Microwave kinetic inductance detectors (MKIDs) are superconducting detectors that are excellent candidates for astronomy in the far-infrared (FIR), roughly 100 GHz to 10 THz. Radiation in this part of the electromagnetic spectrum is particularly hard to detect compared to optical or near-infrared radiation. Furthermore, some sources in the FIR are so faint that the detectors are required to detect single photons to determine the incident photon rate. Recent MKIDs are highly sensitve and are capable of detecting single photons in the FIR, although detection of lower energy photons remains a challenge. Photons produce pulses in the output signal of the detector. As the pulse height is dependent on the photon energy, low energy photons are hard to distinguish from the noise.

This thesis presents a system model that is used in estimating the photon rate. The system model describes signal relations and noise characteristics, so that it provides a foundation for developing statistical estimation and detection algorithms. Based on this model, various estimators are proposed, e.g. a generalized matched filter. This shows the utility of the system model in deriving solutions for estimation problems.

This thesis represents a first step in advancing signal processing techniques for single photon detection in MKIDs designed for FIR astronomy.

Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

Photon Rate Estimation in MKIDs

Optimal Estimation of the Photon Rate in Microwave Kinetic Inductance Detectors via Statistical Digital Signal Processing Methods

---

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

ELECTRICAL ENGINEERING

by

B.Sc. Paul Julian Rozestraten
born in Delft, The Netherlands

This work was performed in:

Signal Processing Systems Section & Terahertz Sensing Group
Department of Microelectronics
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

Delft University of Technology

Delft University of Technology
Department of
Microelectronics

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled **“Photon Rate Estimation in MKIDs”** by **B.Sc. Paul Julian Rozestraten** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: March 14, 2025

Chairman: ____________________
prof.dr.ir. A.J. van der Veen

Advisor: ____________________
prof.dr.ir. J.J.A. Baselmans

Committee Member: ____________________
dr. R.T. Rajan

Contents

Introduction

1

1.1 Radio Astronomy using MKIDs

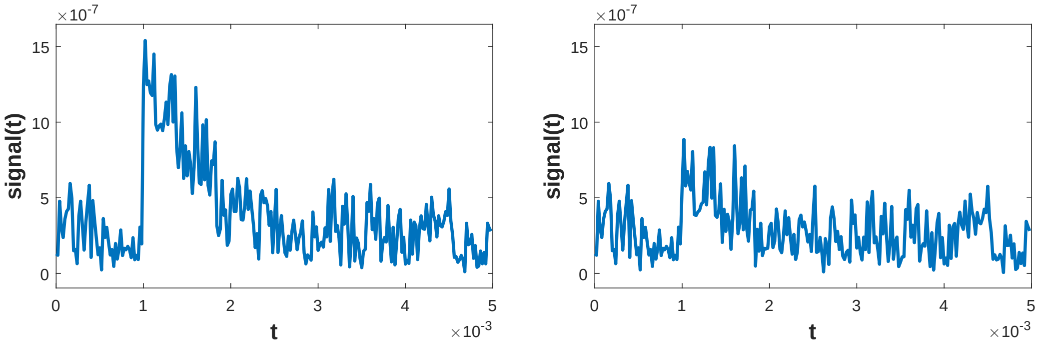
Objects in the night sky can be very dim. There are stars that we are able to see with the naked eye, but other objects, such as gas clouds, nebula and cold dust, remain hidden because they do not produce visible light. Radiation can come in at many frequencies and from vastly different sources, but these objects emit mainly in the far-infrared (FIR), roughly defined as radiation between 100 GHz and 10 THz. One of those sources is the cosmic microwave background, at frequencies around 160 GHz. This radiation is a remnant of the Big Bang. Another source of radiation at this frequency comes from exoplanets. Measuring the radiation coming in from these planets might tell us something about the molecules on the planet, and hence if life might exist there [1].

Optical and near-infrared radiation (0.1-1 μm), as well as radio frequencies (below 1 GHz), are relatively easy to detect. In the far-infrared it is much harder. This is because semiconducting detectors such as charge-coupled devices use silicon with a band gap of 1.1 eV, which implies that they only work for radiation at wavelengths shorter than about 1 μm. Radio receivers do not work within the FIR either, since the frequencies are simply too high. On top of that, sometimes the sources are so faint that single photons have to be detected.

This makes superconducting detectors necessary for the THz range. Cooling down the detectors to temperatures well below the superconducting transition temperature eliminates thermal noise so that the signal to noise ratio is sufficiently high enough to be able to measure the power from the astronomical sources [2]. At the moment there are three types of superconducting detectors that are able to achieve the necessary sensitivity: quantum capacitance detectors, transition edge detectors, and microwave kinetic inductance detectors (MKIDs) [2, 3].

MKIDs have been in development since the early 2000s [4–7]. One of the big advantages of MKIDs over the other types of detectors is that they can be passively multiplexed [7, 8]. By changing design parameters such as the resonator length of each MKID, up to thousands of MKIDs can be read out on a single read-out line [2]. This enables high-pixel imaging or high-resolution spectroscopy.

The MKIDs from the Terahertz Sensing Group are capable of sensing single photon hits at THz frequencies [2, 3]. By counting the number of photons in a time-stream, the power incident on the detector can be measured.



(a) High photon energy (b) Low photon energy

Figure 1.1: Comparison of simulated MKID signals with high photon energy vs. low photon energy.

1.2 Photon Counting in MKIDs

MKIDs have been able to detect single photons for some time now, especially for visible light and near-infrared this is rather easy due to the large photon energies [9]. They have the necessary sensitivity to detect higher energy photons. High energy photons produce a larger signal amplitude. But as the energy of the photons gets lower, the signal read-out also gets lower. This makes it harder to detect photons in the measured time-stream. The peaks in the signal from the photons are almost drowning in noise and are barely visible. Being able to also detect lower energy photons will open up parts of the THz spectrum that astronomers have not been able to measure.

An example of what an MKID signal in noise can look like is shown in Figure 1.1. When a photon hits the detector a peak is seen in the time-stream of the signal.

So far the Terahertz Sensing group at Delft has made some considerable efforts towards detecting single photons. Federica Facchin [3] tried to push single photon counting capabilities of MKIDs to 1.5 THz by improving the physical device design, which led to some pulse detections but at a very low detection reliability. This helped lead to an MKID design whose sensitivity is sufficient for use in future space telescopes [2]. The signal processing part of the photon counting consisted of bandpass filtering, downsampling, and applying a matched filter. The resulting output was then thresholded at 4 to 5 times the standard deviation. This early attempt at the signal processing is sufficient for high energy photons, but more sophisticated techniques are necessary to get the most out of the detectors. There are no optimality conditions given for the detection, and no system overview or statistical description of the measured signal is given.

Wilbert Ras is also working on processing photon pulses in MKID signals [1]. His work is concerned with the height of the peaks induced by the photons and the distribution of the height. These provide a measure of the energy of the photons hitting the detector. A type of optimal filter is used in the frequency domain, but there is still no system overview or statistical signal description from which optimality can be defined.

1.3 Thesis Goal and Outline

As of now there is no description of MKIDs in the signal processing framework. Likewise, there are no statistical descriptions of measurements or definitions of optimal detection or estimation. The goal of is thesis is, therefore, to define the problem of optimally estimating the photon rate incident on an MKID in the framework of statistical signal processing. If possible an optimal solution to the problem of estimating the photon rate will be given.

Therefore, a signal processing model is needed that describes the relation between the measured signal and the parameters to be estimated, which will aid in designing optimal detectors. There are three things necessary to derive that model:

1. A description of the physical mechanism happening in an MKID.
2. Stochastic descriptions of the signals and noise inside the system.
3. Data to help build and verify the model.

These subgoals are reflected in the structure of this work.

First, Chapter 2 gives a description of the physical processes inside an MKID. It starts with a quick introduction to superconductivity, after which it will explain how an MKID is able to sense changes in the conductivity when a photon hits the detector. It also contains a description of a new coordinate system that can be used to define the intermediate and output signals of an MKID.

After that, Chapter 3 will give the description of the system model. It starts with a continuous-time model with signals representing the physical quantities introduced in the previous Chapter and alters it to a discrete-time model where the noise is added as additive Gaussian noise. The distributions of the signals are described and the second order statistics of the noise are given as power spectral densities (PSDs). The model is verified by synthesizing MKID output data that has similar properties to actual MKID output. This model can then be used to derive solutions for estimating the photon rate from a measured time-stream.

After that, Chapter 4 presents an attempt at deriving an optimal solution for estimating the photon rate from the model of the previous Chapter. The optimal solution is sought via the probability density functions (PDFs) of the signals in the model. These PDFs arise from the signal descriptions of the previous Chapter. Unfortunately, an optimal solution is not found, because of an intractable expression of the PDFs. This leads to a simplification of the model, which is used to derive a generalized matched filter as a solution.

Finally, Chapter 5 presents the results of applying the generalized matched filter solution to experimental data. Experimental data is acquired from two chips, LT218 and LT366. The test set-up is described shortly in Chapter 2. The data contains measurements with photon rates varying from zero to hundreds per second.

Physical Mechanisms of MKIDs

2

An MKID is a superconducting radiation detector that can be optimized for many wavelengths ranging from mm-waves (100 GHz) to visible light. MKIDs for far-infrared wavelengths, roughly corresponding to wavelengths between 200-20 μm, typically operate as power integrators, whereas MKIDs in the optical spectrum can measure individual photons. However, recently fabricated MKIDs are capable of sensing single photons even at wavelengths of 200-20 μm [2, 3]. Power can then be measured by summing the energies of all the photons in a single time-stream. However, single photon peak heights have a natural spread in peak height, and current detectors only detect the highest peaks, while the lower ones remain hidden in the noise. This results in an underestimation of the photon rate from the astronomical source

This Chapter serves to give the necessary background on the physical processes happening in MKIDs in order to be able to build a system description in Chapter 3. First, a quick description of superconduction will be given where Cooper pairs and quasiparticles are introduced. Then, the process of a photon breaking a Cooper pair into quasiparticles will be explained. Then, a brief description of the resonator circuit will follow which is used to detect the pair breaking. Finally, a change to the coordinate system used to represent MKID measurements will be developed.

2.1 Superconductivity

In a normal conductor, charge transport is mediated by the movement of electrons. When the temperature T of a conductor drops to near zero, the resistivity will also go down but not to zero. In some materials, the resistivity suddenly drops to zero below a critical temperature T_c. This phenomenon is called superconduction and was first discovered by Kamerlingh Onnes in 1911.

The quantum-mechanical description of superconductivity was pioneered in 1957 [10], called the BCS theory. A very important follow up work was published by Mattis and Bardeen [11] on the behaviour of superconductors in high frequency electromagnetic fields. In a superconductor, there exist two types of additional particles, Cooper pairs and quasiparticles. Cooper pairs are formed by two electrons. At superconduction temperatures the electrons are able to fall into a lower energy state where they are bound to another electron [8]. Cooper pairs allow a conductor to have zero resistivity [12]. The conductor will behave as a perfect conductor with infinite conductivity.[1] Cooper pairs do not dissipate energy, but they do have an inductance associated with them [8]. This is because Cooper pairs have mass, and with that mass they resist a change in the current just like inductors do with their magnetic fields. The inductance

---

[1] Resistivity is the inverse of conductivity.

from the Cooper pairs is called kinetic inductance, which is where the name Kinetic Inductance Detector comes from [8].

Quasiparticles are excitations of a charge quantity from the superconducting ground state into a higher energy state. These excitations can be created by temperature or by radiation. Quasiparticle behave similar to electrons, they dissipate energy when moving through a metal and have a resistivity associated with them [8]. The quasiparticle density χ is the total number of quasiparticles in the solid over the volume. It is exponentially dependent on the temperature. At around $T < T_c/10$, there are therefore virtually no quasiparticles in the material.

The inductive part of the Cooper pairs together with the resistive part of the quasiparticles give rise to a complex conductivity

$$\sigma = \sigma_1 - j\sigma_2 \tag{2.1}$$

where σ_1 is the real and dissipative part from the quasiparticles, and σ_2 is the imaginary and inductive part from the Cooper pairs [8]. The behaviour of the conductance in a high-frequency electromagnetic field with angular frequency ω comes from Mattis-Bardeen theory [11]. The Mattis-Bardeen expression for the conductance is an integral over the possible energy states [13]. This can be simplified to an expression where the change in the conductance is approximately linear in the quasiparticle density [13]. This is only valid when the energy distribution of the quasiparticles is assumed to be given by a thermal quasiparticle distribution, i.e. a change in quasiparticle is equivalent to a change in temperature concerning σ. This happens at $\hbar\omega \ll \mathbb{E}\{_{\}}gap$ and $k_BT \ll E_{gap}$, where $\hbar$ is the reduced Planck constant, and k_B is the Boltzmann constant.

We find that the conductivity is an affine function of the quasiparticle density at a certain field frequency and can be written as

$$\sigma(\omega,\chi) = A(\omega)\ \ (\chi - \chi_0) + \sigma_0, \tag{2.2}$$

where σ_0 is the conductivity at equilibrium ($\chi = \chi_0$) and

$$A(\omega) = \frac{\mathrm{d}\sigma}{\mathrm{d}\chi} \tag{2.3}$$

is the slope of the affine equation (note that although σ is approximately affine in χ, it is not necessarily in ω). χ_0 and σ_0 are, respectively, the quasiparticle density and the conductivity at rest. Other quantities will also follow with the added subscript 0, for example $Q_{i,0}$ and $\omega_{r,0}$. These quantities will be introduced later. They are their respective quantities in the resting state or equilibrium state when $\chi = \chi_0$, which is defined here as the state of the detector at low temperatures ($T < Tc/8$) and no radiation being absorbed.

2.2 Quasiparticle Creation and Recombination

Cooper pairs can be broken up into quasiparticles. This pair breaking is the mode through which MKIDs work; they are pair breaking detectors. Photons can break Cooper pairs if their energy is high enough. For example for aluminium, photons with a frequency $\nu > 90$ GHz can break Cooper pairs [8]. This enables MKIDs to detect photons in the THz spectrum.[2] The photon energy needs to be higher than two times the gap energy E_{gap} of the quasiparticles.

The incoming photon energy is converted into a breaking energy ($2E_{gap}$) and into phonons. Phonons are essentially vibrations in the lattice of the superconductor and can transport energy out of the detector. The total amount of energy that goes into the excess quasiparticles is less than the photon's energy. The pair breaking efficiency is the ratio of the energy that is converted into the breaking of quasiparticles. Its value is typically in the range of 0.4 to 0.6 [8] and it depends on the exact details of the device we measure. For thicker aluminium films and for aluminium films on membranes the value is higher; for thin films on solid substrates this value is lower.

After the quasiparticles have been created, they will recombine again into Cooper pairs. This is a statistical process where two quasiparticles can meet and form a Cooper pair again. Even though the process is a statistical process, the mean of the number of recombined quasiparticles decays approximately exponentially in time [13]. The time scale of the decay is much larger than the time scale of the creation by a photon and can be measured at 50 kHz. The time constant τ of the exponential decay is constant in an MKID and is typically between 1 ns and 100+ μs.

When the quasiparticles density has decayed to the rest state again it can still fluctuate around the mean due to the generation-recombination [13]. This gives rise to quasiparticle noise. This noise is ignored in this work since it is only a minor noise source compared to other noise sources.

2.3 Resonator Read-Out and Behaviour

The behaviour of the quasiparticle density and the conductivity in a superconductor and their relation to incoming photons is now known. Therefore, if we can perform measurements from which we can infer the value of the conductivity or the quasiparticle density in the material, we will be able to sense photon hits. These measurements are performed at RF frequencies. Each MKID is a microwave resonator with its own resonance frequency. This is one of the strengths of MKIDs; they can be multiplexed in the frequency domain. This is achieved by varying the design parameters of the MKID, such as the resonator length.

An MKID chip can contain up to thousands of resonators that are connected to a single read-out line. The chip is fabricated from a superconducting material such as aluminium. The resonator circuit is excited by an RF field and the electrical properties of the circuit are measured. Figure 2.1 shows an image of a fabricated resonator circuit.

Another reason why the conductivity is measured using a resonator circuit is because the change in the conductivity is usually very small. A resonator circuit with a very high quality factor will give a very high response to small changes in the conductivity

---

[2] This is the second frequency that has been introduced. The photon frequency ν is different from the field frequency (also called the RF frequency or read-out frequency) $\omega = 2\pi f$. A third frequency F_s will be introduced in Chapter 3, the sampling frequency.

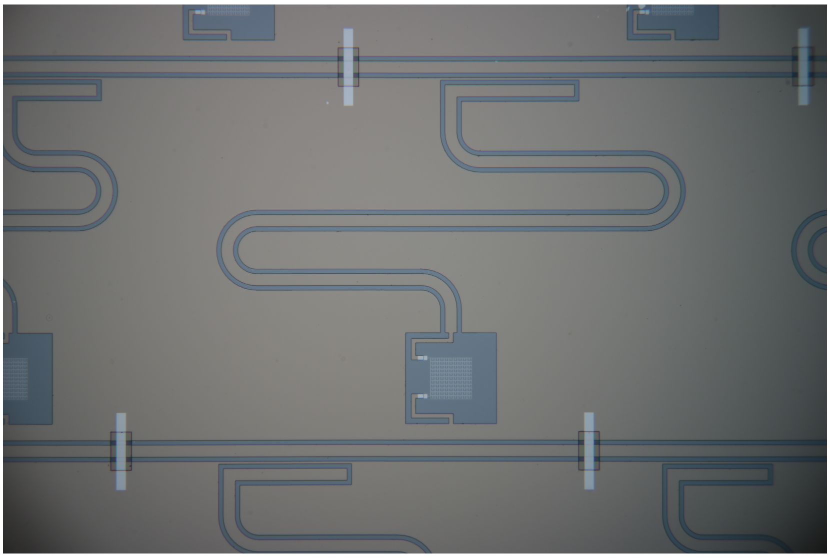


Figure 2.1: Image of resonator circuit on a fabricated chip. Chip LT366. The brown/gold surfaces are the superconducting material, while the grey surfaces are parts that are etched away.

[8]. This Section will describe the response of the resonator circuit to a change in the conductivity (and hence to a change in the quasiparticles and photon hits).

The resonator in Figure 2.1 is read out via the equivalent circuit of Figure 2.2. The two-port network in the circuit is the resonator. The impedance $Z_{line}(\omega, \sigma)$ is the impedance that changes due to a change in quasiparticles. The capacitance $C(\omega)$ couples the resonator to the feedline. The material of the resonator has a surface impedance $Z_s(\omega, \sigma)$ which is determined by the conductivity. At RF frequencies the surface impedance is given by [13]

$$Z_s(\omega, \sigma) = \sqrt{\frac{j\mu_0\omega}{\sigma}} \coth\left(d\sqrt{j\omega\mu_0\sigma}\right) = R_s(\omega, \sigma) + j\omega L_s(\omega, \sigma), \tag{2.4}$$

where μ_0 is the magnetic permeability of free space, d is the thickness of the film, R_s the resulting surface resistance, and L_s the resulting surface inductance. The derivation of Z_{line} from Z_s is not important here, but what is important is that Z_{line} is a function of ω and σ.

Now that we have the impedance $Z_{line}(\omega, \sigma)$, we can find the response of the resonator. The response of the resonator that will be measured is the scattering parameter S_{21}. It is the ratio of the outgoing voltage wave at port 2 over the ingoing voltage wave at port 1 given a matched load, or [14]

$$S_{21}(\omega, \sigma) = \frac{V_2^-(\omega, \sigma)}{V_1^+(\omega)}\bigg|_{V_2^+(\omega,\sigma)=0}. \tag{2.5}$$

The S_{21} parameter is a complex number. It is sometimes also called the forward voltage gain. The ingoing and outgoing voltage waves as well as the ports are annotated in Figure 2.2. Scattering parameters are the standard parameter to measure in RF

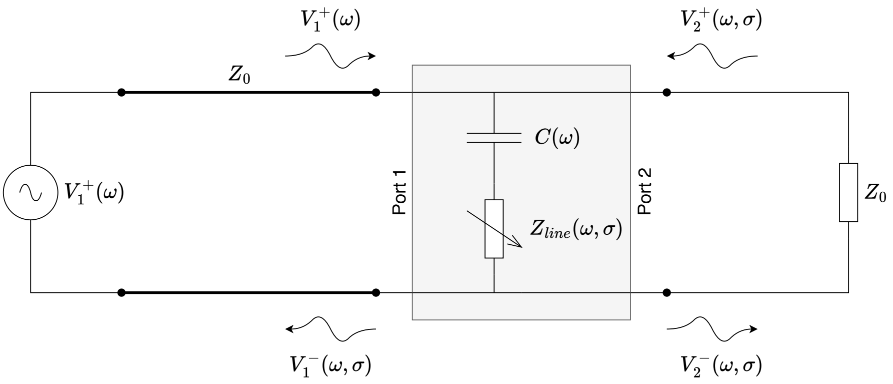


Figure 2.2: Read-out circuit. An RF voltage wave is sent through the resonator via transmission lines with a characteristic impedance Z_0. The grey two-port network is the resonator circuit. $C(\omega)$ is the capacitor coupling the half-wave resonator to the feedlines. $Z_{line}(\omega, \sigma)$ is the impedance of the half-wave resonator.

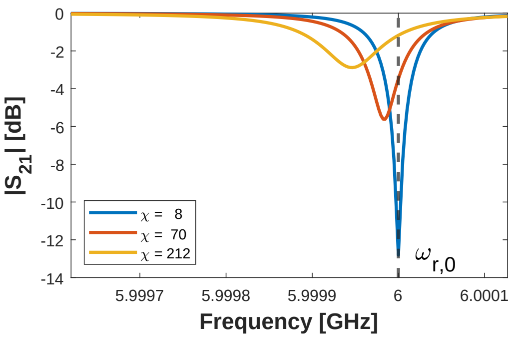


Figure 2.3: Frequency sweep over a simulated MKID near resonance frequency with varying quasiparticle density. A higher χ shifts the resonance frequency ω_r to the left, and the peaks become less sharp as the total quality factor Q lowers.

engineering because they can be measured at high sampling frequencies and accuracies by in-phase and quadrature mixers (IQ mixers) or by vector network analyzers (VNAs). An expression for S_{21} can be found by analyzing the series impedance of the capacitance $C(\omega)$ and the resonator line $Z_{line}(\omega, \sigma)$. This expression is given in [13].

Rather than giving the expression for S_{21} in terms of the impedances in the circuit, it is usually given in terms of two other quantities that are frequently used in resonator physics. These are the quality factor Q and the resonance frequency ω_r. Figure 2.3 shows a simulated example of a frequency sweep for multiple quasiparticle densities χ (and thus σ). These quantities are easily recognizable features in a frequency sweep of S_{21}, in contrast to the impedance Z_{line}. They are also very commonly used and recognized in RF engineering. That is why we will now work towards an expression of S_{21} in terms of Q and ω_r.

The quality factor is a measure of how little energy is dissipated in the system. Alternatively, it is a measure of the sharpness of a dip (or peak) when swept over the frequency. There are three quality factors involved in MKIDs which contribute to the total quality factor

- The coupling quality factor Q_c from the coupling capacitor
- The internal quality factor $Q_i(\sigma)$ due to the quasiparticle dissipation losses
- The limiting quality factor Q_{lim} which limits the maximum achievable quality factor

The internal quality factor is dependent on σ. This is because when there are more quasiparticles in the system, they will dissipate more energy. The three quality factors are added in parallel to form the total quality factor $Q(\sigma)$, so

$$\frac{1}{Q(\sigma)} = \frac{1}{Q_c} + \frac{1}{Q_{lim}} + \frac{1}{Q_i(\sigma)}. \tag{2.6}$$

The internal quality factor at equilibrium is $Q_{i,0} = Q_i(\sigma = \sigma_0)$.

The resonance frequency is the frequency at which the dip occurs. In Figure 2.3 the frequency and resonance frequency are given in Hz instead of in rad/s. This will not matter too much since angular frequency and regular frequency can be used interchangeably with a factor of 2π in this work. The resonance frequency is dependent on the conductivity and as such we write

$$\omega_r(\sigma).$$

This is because the resonance frequency is dependent on the kinetic inductance which becomes lower as more Cooper pairs are broken into quasiparticles. The resonance frequency at equilibrium is $\omega_{r,0} = \omega_r(\sigma = \sigma_0)$

The S_{21} parameter can now be written in terms of the quality factor and the resonance frequency.

$$S_{21}(\omega, \sigma) = \frac{Q/Q_i + 2jQ\frac{\Delta\omega}{\omega_r}}{1 + 2jQ\frac{\Delta\omega}{\omega_r}}, \tag{2.7}$$

where $\Delta\omega = \omega - \omega_r$ is the absolute difference between the read-out frequency and the resonance frequency of the resonator and

$$\frac{\Delta\omega}{\omega_r} = \frac{\omega - \omega_r}{\omega_r} \tag{2.8}$$

the relative difference. The derivation of (2.7) can be found in [13]. From (2.7) it is not immediately clear, but Q_i is dependent on σ and $\frac{\Delta\omega}{\omega_r}$ is dependent on both ω and σ. The next Section will make this dependence more explicit.

The resonator is read out at a fixed frequency $\omega_{r,0}$. This is the resonance frequency of the resonator at equilibrium. Therefore, it is reasonable to think of S_{21} as only a function of σ, or equivalently as a function of χ. However, this depends on the context. In Chapter 3, the relation of S_{21} to σ and χ are both important. In the following Section, the relation of S_{21} to ω at a varying χ is important, as is the relation of S_{21} to χ at the read-out frequency ($\omega = \omega_{r,0}$). Therefore, three new notations are introduced.

First of all, S_{21} will change from a dependence of ω and σ to a dependence on ω and χ, so

$$S_{21}(\omega, \chi) = S_{21}(\omega, \sigma(\omega, \chi)). \tag{2.9}$$

Second of all, when the MKID is read out at the equilibrium resonance frequency, the dependence on ω can be left out.

$$S_{21}(\chi) = S_{21}(\omega = \omega_{r,0}, \chi) \tag{2.10}$$

and is thus shorthand for the full function notation. Similarly,

$$\sigma(\chi) = \sigma(\omega = \omega_{r,0}, \chi). \tag{2.11}$$

Both $S_{21}(\chi)$ and $S_{21}(\omega, \chi)$ will be used depending on context.

2.4 Coordinate System

A common problem when discussing MKIDs is what coordinate system to choose to represent the measured S_{21} values, and hence which variable is defined as the time-series signal of interest. Common options are

- the amplitude of forward transmission $|S_{21}|$,
- the amplitude of the normalized KID circle R,
- the phase of the normalized KID circle θ,
- a nameless Smith chart-like coordinate system [15],
- a phase-dissipation coordinate system [16].

In this work, a new coordinate system is proposed, which is congruent with [15] and [16]. The relation of the $(\mathfrak{Re}\{S_{21}\}, \mathfrak{Im}\{S_{21}\})$ coordinate system to the new coordinate will be described, as will the coordinate system of the normalized KID circle. These are important to relate this work to the field.

2.4.1 MKID Coordinate Systems

When performing a frequency sweep and discussing quality factors and frequency shifts, we tend to look at $|S_{21}|$ as a function of ω as in Figure 2.3. It shows the quality factor and resonance frequency quite well as features.

The S_{21} parameter can also be looked at spatially. So, plotting the real part vs. the imaginary part. This is just S_{21} represented in the complex plane, which is also sometimes called the IQ plane after the in-phase and quadrature components. When performing a frequency sweep over S_{21}, a circle will appear in the complex plane. This is called a resonance circle. We can also do a sweep over the quasiparticle density. This traces out a different curve in the IQ plane, which is (not incidentally) part of a circle. This is called the quasiparticle response or quasiparticle curve. The resonance circle and quasiparticle response are shown in Figure 2.4. In this work we will call

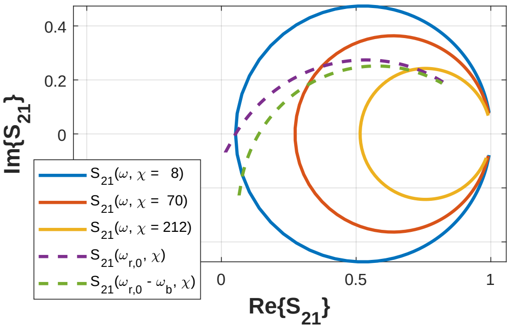


Figure 2.4: The real vs. the imaginary parts of simulated S_{21} parameters when sweeped over frequency and over quasiparticle density. The full lines are frequency sweeps while the dashed line is the quasiparticle curve.

such plots between real and imaginary parts spatial plots. They will appear later when introducing the suggested coordinate system and when showing scatter plots of the noisy measurements. This is used to exemplify the relation between the two parts. Notice how in Figure 2.4 $|S_{21}|$ is minimal at $\omega = \omega_{r,0}$ when the MKID is measured out $\chi = \chi_0$, where the dashed purple line and the full blue line cross.

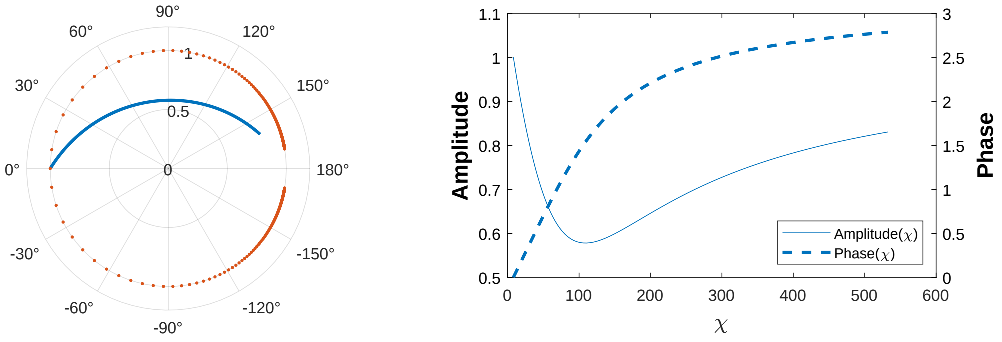
The spatial plot of a normalized KID circle looks similar to a spatial plot of the real and imaginary parts of S_{21}, only the coordinates are given in (amplitude, phase) referenced on the resonance circle at equilibrium. Otherwise it is just a translated and scaled version of the same plot. An example is shown in Figure 2.5. For asymmetric MKID frequency responses the mapping from the S_{21}values to a normalized KID circle is not as straightforward [17–20]. Asymmetric frequency responses will not be discussed in this work. The phase as a function of χ is also shown in Figure 2.5 as is the amplitude. The curves reveal a rather complex relation to χ for a read-out in this coordinate system. The amplitude is also not a one-to-one mapping of the quasiparticle density. The coordinates can be used to map to a quasiparticle density but they distort features. For example, an exponential decay in the quasiparticle density would not be exponential anymore when only examining the phase response.

This is one of the reasons the new coordinate system was developed. When deriving the system description in Chapter 3 from a photon to a response in S_{21}, the relation between the response and the quasiparticles is necessary. The new coordinate system allows for a straightforward mapping between $S_{21}(\chi)$ and the quasiparticle density.

2.4.2 Proposed Coordinate System

We develop the coordinate system from Equation (2.7), which reads

$$S_{21}(\omega,\chi) = S_{21}(\omega,\sigma(\omega,\chi)) = \frac{Q/Q_i + 2jQ\frac{\Delta\omega}{\omega_r}}{1 + 2jQ\frac{\Delta\omega}{\omega_r}}$$



(a) Spatial plot (b) QP response

Figure 2.5: Normalized KID circle in polar coordinates (R, θ). Blue shows a simulated response to an increasing quasiparticle density at $\omega = \omega_{r,0}$.

In this equation, the relation to the quasiparticle density can not be readily seen. Therefore, a change to (2.7) is proposed:

$$S_{21}(\omega, \chi) = \frac{\zeta_{lim} + \zeta_i(\chi) + j\xi(\omega, \chi)}{\zeta + j\xi(\omega, \chi)} = \frac{\zeta_{lim} + r(\omega, \chi)}{\zeta_c + \zeta_{lim} + r(\omega, \chi)}, \tag{2.12}$$

where $\zeta_i(\chi)$, ζ_c and ζ_{lim} are damping ratios, $\xi(\omega, \chi)$ is the frequency-dependent part, and

$$r(\omega, \chi) = \zeta_i + j\xi. \tag{2.13}$$

The real and imaginary parts of $r(\omega, \xi)$ form the proposed coordinate system.
(2.12) and (2.7) are related through

$$\xi(\omega, \chi) = \frac{\Delta\omega}{\omega_r}, \tag{2.14}$$

$$\zeta = \frac{1}{2Q}, \qquad \zeta_i(\chi) = \frac{1}{Q_i(\chi)}, \qquad \zeta_c = \frac{1}{Q_c}, \qquad \zeta_{lim} = \frac{1}{Q_{lim}} \tag{2.15}$$

and

$$\zeta = \zeta_c + \zeta_{lim} + \zeta_i. \tag{2.16}$$

In this coordinate system (ζ_i, ξ), the response to χ is affine. In this Section, we will show affineness of the components in χ, by showing affineness in σ. This is sufficient, because from Section 2.1 we know σ is approximately affine in χ.

The function

$$\begin{aligned} g(r; \zeta_c, \zeta_{lim}) &= \frac{\zeta_{lim} + r}{\zeta_c + \zeta_{lim} + r} \\ g^{-1}(S_{21}; \zeta_c, \zeta_{lim}) &= \frac{(\zeta_c + \zeta_{lim})S_{21} - \zeta_{lim}}{-S_{21} + 1} \end{aligned} \tag{2.17}$$

allows for easy mapping between the coordinate system (ζ_i, ξ), and $(\mathfrak{Re}\{S_{21}\}, \mathfrak{Im}\{S_{21}\})$. ζ_c and ζ_{lim} are parameters of the function.

Damping Ratio Components

Damping ratios are the inverse of quality factors. They describe the same information of a resonator. They can be understood as a measure of the power loss in a dampened resonator. A resonator with a high damping ratio will have a shorter ring time.

The subscripts of the damping ratios in (2.12) are the same as the ones from the quality factors in Section 2.3. We therefore have the following relations:

$$\zeta = \frac{1}{2Q}, \qquad \zeta_i(\chi) = \frac{1}{Q_i(\chi)}, \qquad \zeta_c = \frac{1}{Q_c}, \qquad \zeta_{lim} = \frac{1}{Q_{lim}}$$

and

$$\zeta = \zeta_c + \zeta_{lim} + \zeta_i.$$

This is convenient because from [13] we find that a change in σ leads to a linear change in the inverse of the quality factor. More precisely,

$$Q_i(\chi) = \frac{2}{\alpha_k \beta} \frac{\sigma_2(\chi)}{\sigma_1(\chi)} \tag{2.18}$$

which we can rewrite to

$$\zeta_i(\chi) = \frac{\alpha_k \beta}{4} \frac{\sigma_1(\chi)}{\sigma_2(\chi)} \tag{2.19}$$

using (2.15). Here α_k is the ratio of the kinetic inductance over the total inductance, β is the phase constant, and $\sigma_1(\chi)$ and $\sigma_2(\chi)$ are the real and imaginary parts respectively of the complex conductivity. This relation is approximately linear in σ_1, and therefore also in χ, because $|\sigma_2| \gg |\sigma_1|$ and $|\frac{\mathrm{d}\sigma_1}{\mathrm{d}\chi}| \approx |\frac{\mathrm{d}\sigma_1}{\mathrm{d}\chi}|$ [13]. However, note that ζ_i is usually not 0 at equilibrium, since χ is also not 0 at equilibrium; there are always some quasiparticles dissipating energy.

Frequency-Dependent Component

The frequency-dependent component is defined as

$$\xi(\omega, \chi) = \frac{\Delta\omega}{\omega_r} = \frac{\omega - \omega_r}{\omega_r}. \tag{2.20}$$

To show the affineness of this in χ, we use that the relative change in the resonance frequency is affine in χ [13],

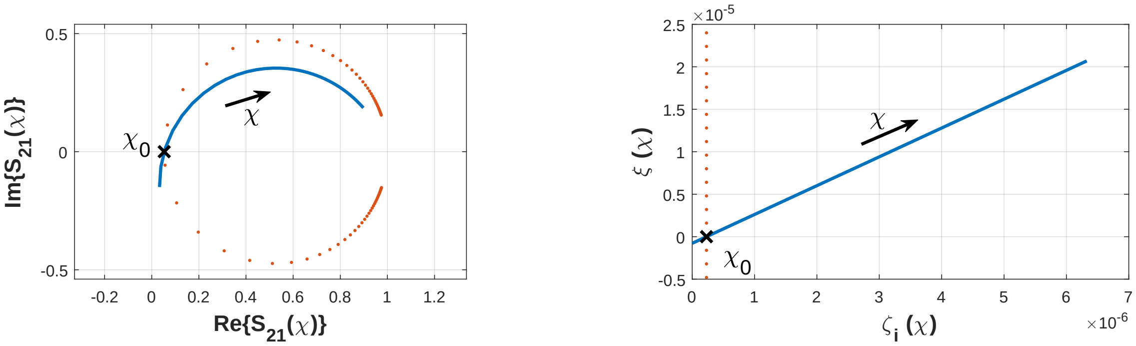
$$\frac{\Delta\omega_r}{\omega_{r,0}} = \frac{\omega_r - \omega_{r,0}}{\omega_{r,0}} = \frac{\alpha_k \beta}{4} \frac{\sigma_2(\chi) - \sigma_{2,0}}{\sigma_{2,0}}. \tag{2.21}$$

Plugging this quantity into (2.20) we get a not-so-simple equation

$$\xi(\omega, \chi) = \frac{\omega - (\frac{\Delta\omega_r}{\omega_{r,0}} + 1)\omega_{r,0}}{(\frac{\Delta\omega_r}{\omega_{r,0}} + 1)\omega_{r,0}}. \tag{2.22}$$

This does not look like it is affine in χ at first glance, but it is for the small-signal approximation.

Note here that although ξ is affine χ, $\xi = 0$ for $\chi = \chi_0$ and $\omega = \omega_{r,0}$.



(a) S_{21}-space

(b) r-space

Figure 2.6: Blue shows a simulated response of an MKID to an increasing QP density. The red dots shows the resonance circle for reference. χ_0 denotes the response at equilibrium quasiparticle density, note how $\zeta_i(\chi_0) > 0$ and $\xi(\chi_0) = 0$. (a) shows the response in the S_{21}-space, while (b) shows the same response in the transformed r-space.

2.4.3 Features of the Proposed Coordinate System

In the new coordinate system of $r = \zeta_i + j\xi$, the response to a change in χ looks approximately like a straight line, as shown in Figure 2.6. This realization is very important for the later signal processing.

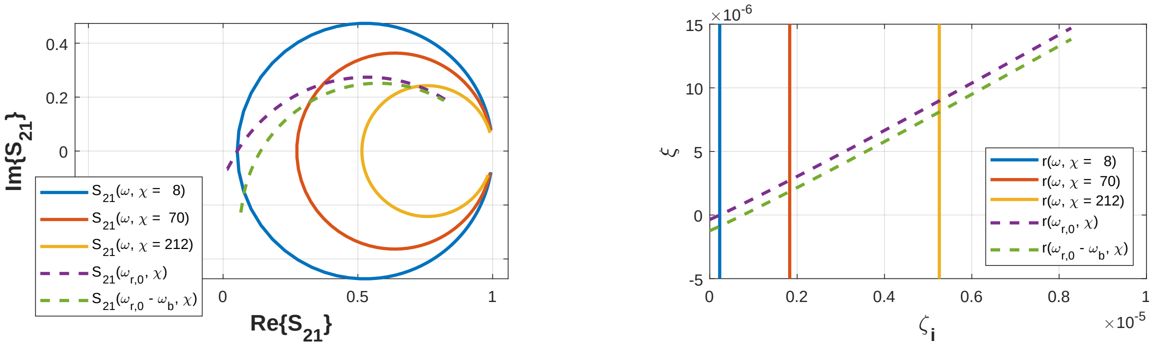
In Figure 2.6 is also visible that r is not 0 at $\chi = \chi_0$. r does not pass through the origin. This is because ζ_i is linear in χ, but ξ is not. On the other hand, ξ is zero for $\chi = \chi_0$, but not for $\chi = \zeta_i = 0$. In Section 3.3, it will be important to find linear relations instead of affine. From Figure 2.6, it is clear that at equilibrium $r_0 = r(\omega_{r,0}, \chi_0) = \zeta_{i,0}$, and $\xi = 0$.

Further looking at the equilibrium point r_0, it is possible to choose a different read-out frequency than $\omega = \omega_{r,0}$. Biasing the read-out tone to be a little higher or a little lower results in the full lines in Figure 2.7. ξ is simply shifted up or down. Therefore, we find for r_0

$$r_0 = \zeta_{i,0} + j\xi_{bias}. \tag{2.23}$$

In this work we will not do measurements at biased read-out tones and biasing will not be mentioned further, since it is not practical to do so. Although later from Section 4.3 it should become clear that biasing at a slightly lower read-out tone will result in a slightly lower noise variance.

The reason why r traces out a line while S_{21} traces out a circle is because Equation (2.12) is a Möbius transformation. Möbius transformations map generalized lines to generalized lines, and a circle is a generalized line in the extended complex plane. Something similar happens in Smith charts and z-transforms. In Smith charts, complex impedances, represented in a Cartesian coordinate system, map to complex reflection coefficients inside the unit circle. In z-transforms, poles and zeros of the transfer function of a continuous system, represented in the Laplace domain, map to poles and zeros of a discrete system inside or outside the unit circle.



(a) S_{21}-space

(b) r-space

Figure 2.7: Circles in the S_{21} plane trace out lines in the (ζ_i, ξ) plane. Full lines are frequency sweeps at different quasiparticle densities while the dashed line are quasiparticle curves. A quasiparticle curve is shown for ω at resonance frequency and for a biased read-out frequency.

System Model of MKIDs

3

The process of a photon hitting the MKID and the resulting S_{21} can be modeled as a system. The system without noise is shown in Figure 3.1, and Table 3.1 gives a description of the signals inside.[1]

The input of the system is a time-dependent photon stream which is modeled as a delta train where the height of each delta encodes the photon energy E_ϕ. The delta train is modeled as

$$d(t) = \sum_i \delta(t - t_i), \tag{3.1}$$

where t is time, $\delta(t)$ is the continuous-time unit impulse function, and t_i is the arrival time of each photon. Photon arrivals are modeled as a Poisson process. As such, the number of photons in a time window with fixed length is Poisson distributed with parameter λ, the photon rate. The goal of this thesis is to find an estimator $\hat{\lambda}$ for the photon rate.

The output of the system is the sampled complex S_{21} scattering parameter as defined in Section 2.3, and is sampled in time. As mentioned there, a frequency sweep will result in a circle being drawn in the complex plane of S_{21}. This resonance circle serves as a reference in the coordinate system of $(\mathfrak{Re}\{S_{21}\}, \mathfrak{Im}\{S_{21}\})$.

In this Chapter, a canonical description of the system will be given. This is the canonical way an MKID is analyzed, formulated in a signals and systems way. The proposed coordinate system from Section 2.4 leads to a novel way to define the signals in the MKID. It is a new interpretation of the measurements of an MKID where many variables in the system become affine or linear, while still being formulated in the

Table 3.1: Signal Descriptions of Continuous-Time System

Signal	Description
$d(t)$	Delta train where each photon hit is modeled as a delta spike.
$\phi(t) = E_\phi d(t)$	The incoming photon stream, with E_ϕ the energy of the photons.
$\chi(t)$	The quasiparticle density.
$\sigma(t)$	The complex conductivity.
$Z_s(t)$	The surface impedance of the resonator.
$S_{21}[n]$	The forward transmission of the resonator circuit.

---

[1] The system model and the coordinate system were designed from the assumption that the output of the system, S_{21}, is a function of the read-out frequency and the quasiparticle density in the system. This assumption leaves out the read-out power, but it is nice because a lot of things become linear or affine. Next the read-out frequency is set to be a constant, since it does not change during a measurement and it can be viewed more as a known input parameter rather than a signal. This makes it apparent that the signals in the system should be defined in reference to the quasiparticle density.

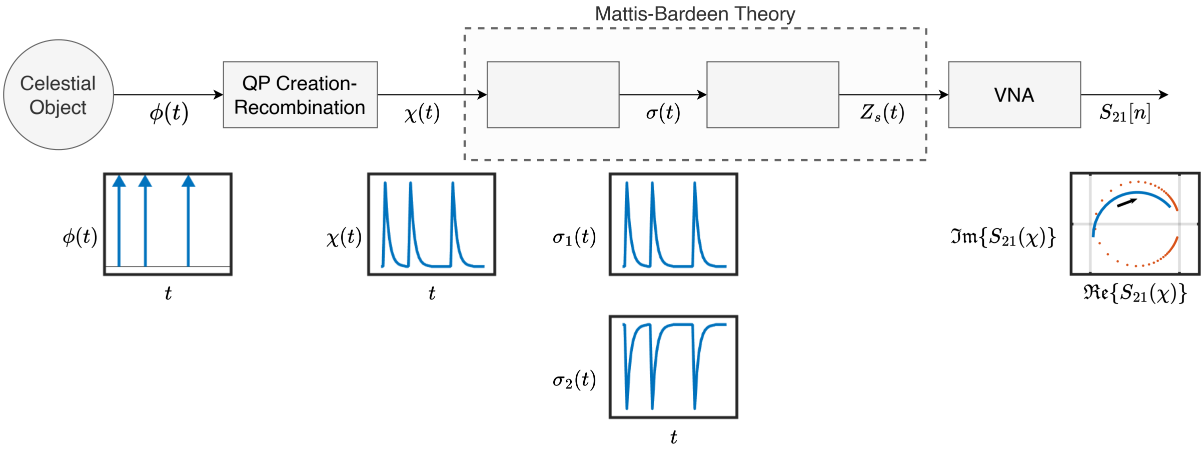


Figure 3.1: Continuous-time system model of an MKID. The inset blue plots are simulated examples showing the typical form of the time-series signals. The right circle plot shows a spatial plot of $S_{21}(\chi)$. The orange dots show a resonance circle for reference.

understanding of microwave resonators. Section 3.3 will use the new coordinate system to reformulate the system description, Section 3.4 gives a system description with added noise, and Section 3.5 gives the noise characteristics.

This model derived in this Chapter will form the basis of the estimation techniques proposed Chapter 4.

3.1 Canonical System Description

The photons break up Cooper pairs into quasiparticles. The density of created quasiparticles is assumed to be linear in the photon energy, $\chi_c(t) = \frac{\eta}{E_{gap}V}\phi(t)$, where η is the pair-breaking efficiency, E_{gap} the gap energy of the conductor, and V the volume (see Section 2.2). $\chi_c(t)$ looks like a delta train, just like the photon stream. Next the quasiparticles will decay back to the equilibrium quasiparticle density. This is modeled as an LTI system with an impulse response that looks like a decay function. So,

$$\chi(t) = \chi_c(t) * h(t) + \chi_0, \tag{3.2}$$

where $*$ denotes convolution and

$$h(t) = \begin{cases} e^{-t/\tau}, & \text{for } t \geq 0 \\ 0, & \text{for} t < 0 \end{cases} \tag{3.3}$$

After that comes the complex conductivity from Mattis-Bardeen theory. In our case, this is an approximately affine function of the quasiparticle density (see (2.2)). This is the last affine step in the system. After this step a VNA or an IQ mixer measures the resulting S_{21} at field frequency ω with a sampling frequency of F_s.

The system is shown in Figure 3.1. The individual blocks in the system are not necessarily LTI systems as is common in other signal processing problems. The system model will therefore be converted into a discrete system model with LSI subsystems.

3.2 Discretization

This Section will convert the continuous-time system model to a discrete model, such that there are only discrete-time signals. The discretization occurs in the VNA. This Section will therefore go back from the VNA to the photon stream to convert the model.

The discretization of a signal $x(t)$ is given by

$$x[n] = x(nT_s), \tag{3.4}$$

where $T_s = 1/F_s$ is the sampling time. Therefore,

$$S_{21}[n] = S_{21}(t = nT_s). \tag{3.5}$$

This notation may be ambiguous, as S_{21} was previously a function of frequency. One might mistakenly apply an inverse Fourier transform to $S_{21}(\omega, \chi)$ to obtain $S_{21}(t)$, but this does not yield the original voltage signals. Instead, S_{21} is sampled to measure the system's electrical properties changing in time.

Here, ω represents the instantaneous frequency, with $\omega \gg F_s$. A more precise notation is $S_{21}(\omega, \chi(t))$. During a measurement of the quasiparticle density, the frequency remains fixed at $\omega = \omega_{r,0}$, and the VNA measures only $S_{21}(\chi(t))$.

So, in order to get $S_{21}(t)$, we write

$$\sigma(t) = \sigma(\chi(t)) \tag{3.6}$$

$$Z_s(t) = Z_s(\sigma(t)) \tag{3.7}$$

and thus

$$S_{21}(t) = S_{21}(\chi(t)). \tag{3.8}$$

The discretizations of these signals are thus simply

$$\sigma[n] = \sigma(\chi[n]) \tag{3.9}$$

$$Z_s[n] = Z_s(\sigma[n]) \tag{3.10}$$

$$S_{21}[n] = S_{21}(\chi[n]) \tag{3.11}$$

and

$$\chi[n] = \chi(nT_s). \tag{3.12}$$

An exact equation for $\chi(t)$ can be found by writing out the convolution (3.2)

$$\chi(t) = \int_{-\infty}^{\infty} \chi_c(t')h(t-t')\ \mathrm{d}t' + \chi_0 \tag{3.13}$$

$$= \chi_0 + \frac{\eta E_\phi}{E_{gap}V} \int_{-\infty}^{\infty} d(t')h(t-t')\ \mathrm{d}t' \tag{3.14}$$

$$= \chi_0 + \frac{\eta E_\phi}{E_{gap}V} \int_{-\infty}^{\infty} \left(\sum_i \delta(t'-t_i) \right) h(t-t')\ \mathrm{d}t' \tag{3.15}$$

$$= \chi_0 + B \sum_i h(t-t_i), \tag{3.16}$$

where $B = \eta E_\phi / E_{gap} V$ is a constant. This reveals $\chi(t)$ to be an affine function of a sum of pulses with the shape of $h(t)$. This was already known but now we have a mathematical expression for $\chi[n]$.

This expression shows no relation to the photon rate λ however. We will therefore use an approximation for $\chi[n]$, where

$$\chi[n] = \chi_0 + Bk[n] * h[n] = \chi_0 + B \sum_{m=-\infty}^{\infty} k[m]h[n-m], \tag{3.17}$$

where $k[n]$ is the number of photons arriving in the interval nT_s to $(n+1)T_s$ or

$$k[n] = \int_{nT_s}^{(n+1)T_s} d(t)\ \mathrm{d}t = \sum_i 1 \quad \text{for} \quad i \in \{i | nT_s \le t_i < (n+1)T_s\}, \tag{3.18}$$

$$h[n] = h(nT_s), \tag{3.19}$$

and $*$ denotes here discrete-time convolution. Since $d(t)$ is a Poisson process, $k[n]$ should be Poisson distributed

$$k[n] \sim \mathcal{P}(\lambda). \tag{3.20}$$

The discretized incoming photon stream is also defined as $k[n]$. The photon energy that was encoded in $\phi(t)$ is included in the constant B.

3.3 System Definition as Linear Shift Invariant Systems

In this Section we will change the system until we have a system to which we can add noise and which is usable for our estimation techniques.

The first intermediate system model is found in Figure 3.2. The canonical system model is first defined in the coordinate system of Section 2.4 $r[n] = (\zeta_i + j\xi)[n]$. $\chi[n]$ follows from $\sigma[n]$ as because of Mattis-Bardeen theory, and $r[n]$ follows from $\chi[n]$. In Section 2.4.3 the measurements in the S_{21}-space were found to be a Möbius transformation of r, namely $g(r)$ which maps the response to a circle.

The second intermediate system model is found in Figure 3.3. The linear and affine mappings in Figure 3.2 can be combined into a single LSI system with impulse response $h[n]$. But in order to do that, all the affine mappings need to be made explicit as in Figure 3.3. The affine mappings are simply a gain followed by an offset specified in the parameter of the block.

The third intermediate system model is found in Figure 3.4. This is the final system model to which noise will be added in the next Section. This is done by simplifying Figure 3.3. First all the constants are taken, moved further along the system, and scaled them by the linear gains. This results in a single added constant r_0 which can be used as an extra parameter in $g(r)$.This is the same r_0 as from Section 2.4.3. The function is from this point on defined as

$$\begin{aligned} g(r; r_0, \zeta_c, \zeta_{lim}) &= \frac{r_0 + \zeta_{lim} + r}{r_0 + \zeta_c + \zeta_{lim} + r} \\ g^{-1}(S_{21}; r_0, \zeta_c, \zeta_{lim}) &= \frac{(r_0 + \zeta_c + \zeta_{lim})S_{21} - r_0 - \zeta_{lim}}{-S_{21} + 1} \end{aligned} \tag{3.21}$$

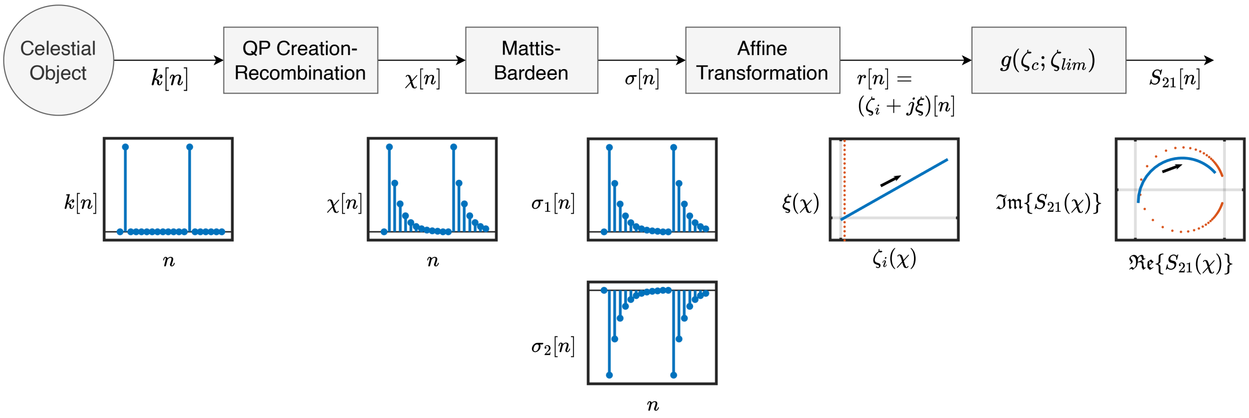


Figure 3.2: Discrete-time system defined with reference to signal $r[n]$. The inset blue stem plots show simulated examples of the typical form of the time-series signals. The spatial plots show the signals as a function of quasiparticle density in their respective coordinate system in blue. The orange dots represent the resonance cirlce for reference. The function g corresponds to the function g in Section 2.4, responses that lie on lines map to circles in the complex measurement plane of S_{21}.

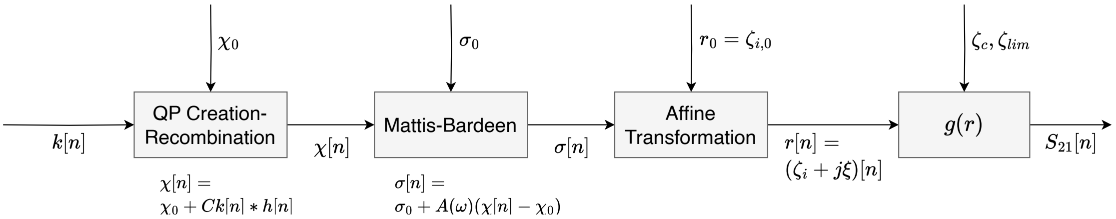


Figure 3.3: System overview where all affine transformations are noted so they can be more easily turned into an LSI system.

All the constant gains such as E_ϕ and B are also taken out of the input signal and put into a single complex gain C.

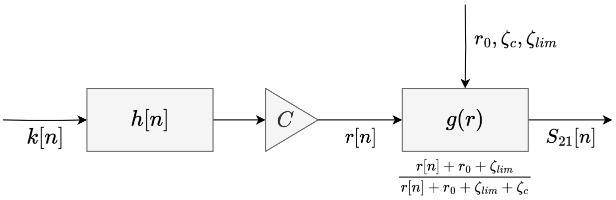


Figure 3.4: System overview where all affine transformations have been combined into a single impulse response $h[n]$, a gain C, and an extra offset in $g(r; r_0, \zeta_c, \zeta_{lim})$.

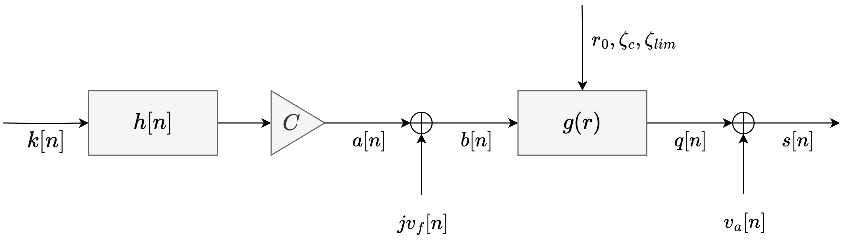


Figure 3.5: Final system overview with noise added at the appropriate places. This is the system that will be used in the estimation of the photon rate λ in Chapter 4.

3.4 System with Noise

Gao et al. [21] identify two types of noise in their paper. The first noise is amplifier noise from the read-out system. It is complex-valued noise which is added in the S_{21}-domain. The second is phase noise. They find in their paper that the phase noise lies perfectly on the resonance circle. The reason it is called phase noise is because it mostly shows up in the phase component of the (R, θ) coordinate system. It is also sometimes called frequency noise, because it lies along the resonance circle. It is as if either the read-out frequency or the resonance frequency has noise added to it. The source of the noise is from quantum mechanical two-level systems (TLS) existing in the dielectric substrate [21], although the source does not matter for this analysis.

In this interpretation of the phase noise, it can be added to the system as additive noise in σ_2 or in ξ. The amplifier noise can be added as additive complex noise to S_{21}. The final resulting system with noise can be seen in Figure 3.5. This is the model that will be used in Chapter 4. The data model of the measurement $s[n]$ is

$$s[n] = q[n] + v_a[n], \tag{3.22}$$

where

$$q[n] = g(b[n]; r_0, \zeta_c, \zeta_{lim}) \tag{3.23}$$
$$b[n] = a[n] + jv_f[n], \tag{3.24}$$

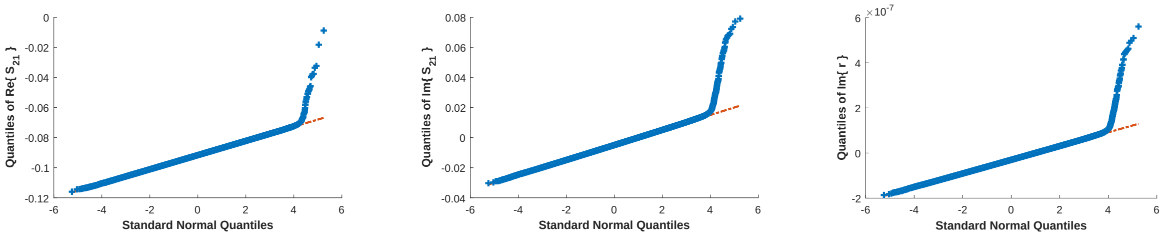
and

$$a[n] = Ck[n] * h[n]. \tag{3.25}$$

Here, $v_f[n]$ is real valued frequency noise that only lives in the imaginary part of the r-space. It is therefore added as $jv_f[n]$. $v_a[n]$ is the complex valued amplifier noise.

3.5 Noise Characteristics

The noise characteristics can be analyed by performing a measurement at a low blackbody temperature. At low temperatures the blackbody sends out almost no radiation and thus no photons. The measurement will only contain noise.



(a) $\mathfrak{Re}\{S_{21}\}$ (b) $\mathfrak{Im}\{S_{21}\}$ (c) $\mathfrak{Im}\{r\}$

Figure 3.6: Quantile-quantile plots of data from a measurement at black body temperature of 9.993 K of KID18 of chip LT218. The quantiles of the input data are almost distributed like a Gaussian distribution, except for a tail at higher values. The tails come from photons leaking in and cosmic rays.

The distribution of the noise sources can be verified via a quantile-quantile plot. A quantile-quantile plot shows the quantiles of one distribution against another distribution. This can be from a function description such as for a standard Guassian or from the distribution of sampled data. If the quantile-quantile plot shows an approximately straight line the data can be assumed to come from the same distribution. In this case the quantiles of sampled data are compared to a standard Gaussian. Figure 3.6 shows the quantile-quantile plot of a measurement of only noise in the S_{21}-domain and in the s-domain. The fact that it matches in the S_{21}-domain implies that the amplifier noise is Gaussian distributed; and the fact that it matches in the r-domain implies the frequency noise is also Gaussian distributed. The quantile-quantile plots show that in the sampled data there are more high values compared to a standard Gaussian. These are from cosmic rays which the MKID also responds to.

The general shape of the noise spectra is also given in [21]. When synthesizing MKID data in Section 3.6 it is important that the spectral shapes match that of actual MKIDs.

The amplifier noise is white noise with a $1/F$ component at low frequencies. The power spectral density of $v_a[n]$ can be modeled as

$$P_a(F) = \sigma_a^2(1 + \frac{F_l}{|F|}), \tag{3.26}$$

where σ_a^2 is the noise floor power and F_l is the frequency where the power of the $1/F$ component starts to dominate over the white noise component. Note that this frequency is not the same $f = \omega/2\pi$ but instead is the frequency of the discrete-time Fourier transform (DTFT).

The frequency noise has a $1/F^{1/2}$ power spectrum with a roll off at the resonator bandwith ΔF. The resonator bandwidth can be calculated from the resonance frequency and the quality factor as $\Delta F = \omega_r/2\pi Q$. This is usually specified as the ring time $\tau_r = 1/\Delta F$. This spectrum can be modeled as

$$P_f(F) = \sigma_f^2|F|^{-1/2}\frac{1}{1 + (F\tau_r)^2}. \tag{3.27}$$

Two simulated examples are shown in Figure 3.7.

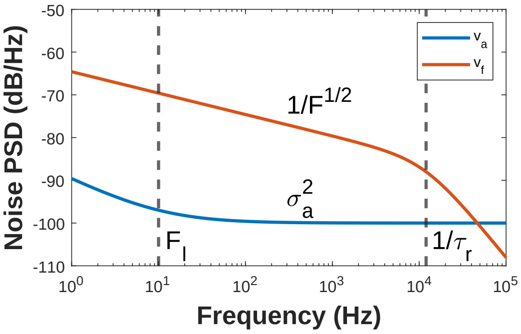


Figure 3.7: Single-sided power spectral densities of the two noise sources. Blue shows the PSD of the amplifier noise. It features a noise floor and a $1/F$ component below $F = F_l$. The noise power is divided between the real and imaginary terms of v_a. The frequency noise v_f features a $1/F^{1/2}$ slope, and a roll of at $F = 1/\tau_r$. Actual dB values are not necessarily characteristic due to normalization.

The noise spectra of Figure 3.7 do not not match the spectra of [21] exactly because there the noise is given in the (R, θ) coordinates, which are the most commonly used coordinates in the field. It is assumed that the amplifier noise has the same shape PSD as for the R-coordinate, and that the frequency noise is transformed by adding a noise floor and the $1/F$ component from the amplitude noise. This noise shows up in the θ-coordinate. This gives the spectra

$$P_A = P_a = \sigma_a^2(1 + \frac{F_l}{|F|}) \quad \text{and}$$

$$P_\theta = P_f(1 + \frac{F_l}{|F|}) + P_a,$$

of which the shape can be found in Figure 3.8. This shape lines up with the shape found in [21]. In the next Section it is important that the PSDs of the synthesized data match with this shape.

3.6 Verification via Data Synthetization

The model can be verified by showing that it can be used to synthesize data with the same properties as actual experimental data. The model used to synthesize the data is the one in Figure 3.2 with added noise. This Section will look at three properties to assess the validity of the model. First a spatial plot of measurements, second a timestream of pulses in phase coordinates, third a PSD estimation based on the generated data.

Figure 3.9 shows a spatial plot of synthesized S_{21} data. Data points without photon peaks lie around the minimal S_{21} value. Data points from peaks shoot along the quasiparticle curve. The quasiparticle curve curves a little more inward than is normal. Usually it lies more along the resonance circle. The noise blob looks like an ellipsoid that curves around the resonance circle. Noise power can be estimated by eye from the size of the noise blob and is normal for an MKID. All these features except the quasiparticle curve match regular data from MKIDs.

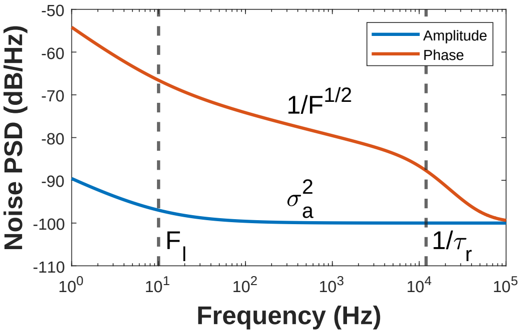


Figure 3.8: Single-sided power spectral densities of the two noise sources in (R, θ)-coordinates. Blue shows the PSD in the A-coordinate. It features the same noise floor and $1/F$ component as the amplitude noise. The frequency noise shows up in the θ-coordinate, which is shown in orange. The shape is the same as the frequency noise except for the noise floor at high frequencies and the $1/F$ component at low frequencies.

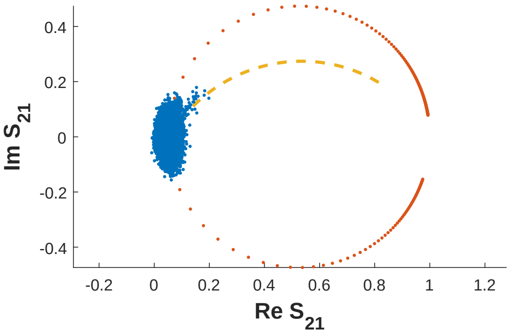


Figure 3.9: Spatial plot of synthesized data. Blue scatter plot is the measurements, orange dots are the resonance circle, yellow dashes are the quasiparticle curve.

Figure 3.10 shows a part of the timestream of synthesized data in (R, θ) coordinates. The timestream shows temporally colored noise with decay pulses in them where a photon hit the detector.

Figure 3.11 shows an estimated PSD via Welch's method [22] of data with noise only in the (R, θ) coordinates and of data with photon pulses.
In the PSD with only noise the $1/F^{1/2}$ behaviour of the θ coordinate is visible in the middle frequencies. At lower frequencies the $1/F$ behaviour is more prominent. At higher frequencies it rolls of towards the noise floor. The A noise is dominated by the amplifier noise and shows a noise floor with $1/F$ behaviour at lower frequencies. When photons hit the detector the PSDs start showing a roll-off at $1/\tau$. These PSDs match with those from [21] and Figure 3.8.

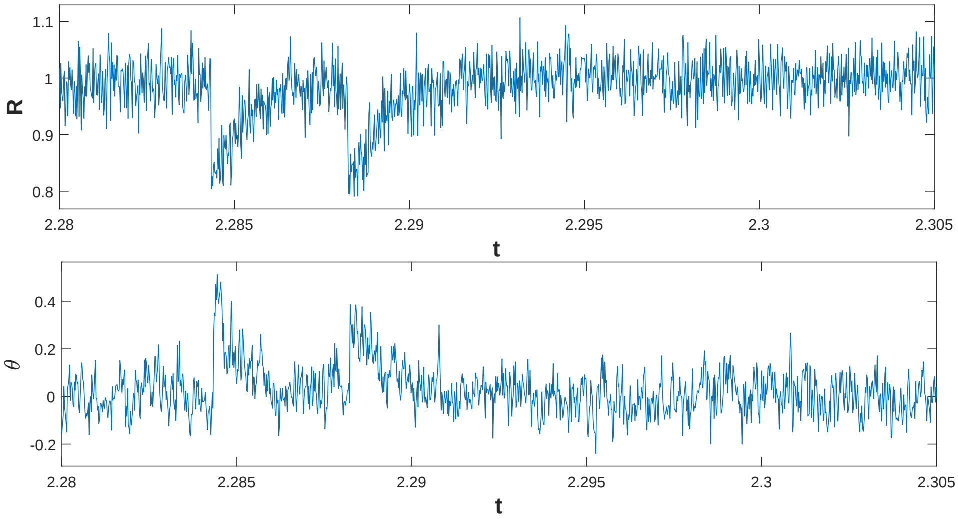
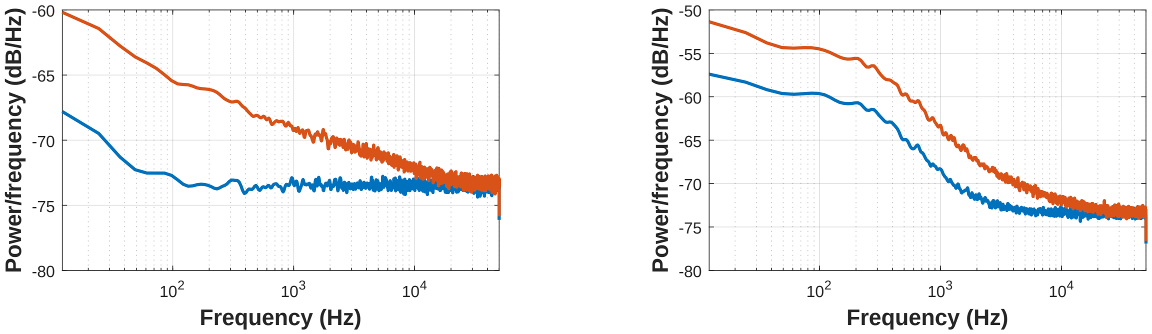


Figure 3.10: Part of a timestream of synthesized data in (R, θ). Two pulses that are close together are visible and the rest of the timestream shows colored noise. The discrete signal points have been connected with a line for easier viewing.



(a) Noise-only

(b) Photons

Figure 3.11: Power spectral densities of synthesized S_{21} data. Blue show the PSD of the R coordinate, orange shows the PSD of the θ coordinate. (a) shows the PSD of synthesized data with only noise (photon rate is 0), (b) shows the PSD of with photon pulses. Roll-off frequency for the θ noise is 1.2 kHz, quasiparticle decay time is 1 ms, photon rate is 100 /s.

Optimal Estimation Techniques

4

This Chapter will explore three different tactics of arriving at an optimal solution for finding an estimator for the photon rate on the MKID. The underlying signal model is the one from Figure 3.5 and the signals in the system are defined in Section 3.4 as.

The first approach entails direct derivation. So writing a statistical description of the measured data in terms of the actual photon rate, and deriving an optimal estimator from that. This approach does not work, because of an integral that does not have an exact solution. It does however, lead to an important insight that if an unbiased estimator of the photon stream is found, that estimator leads to an unbiased estimator of the photon rate.

Following this insight, two other tactics are proposed: one based on a matched filter, and another based on a Wiener filter. The matched filter tries to detect pulses in the noisy measurement, and reconstructs the photon stream from those detections. The Wiener filter approach tries to reconstruct the photon stream by assuming the photon stream consists of Gaussian random variables. This does not make sense physically, but it does lead to an unbiased estimator from which the photon rate can be determined. (It is not computable however, since it requires the photon rate also as input. This can perhaps be tackled by doing an alternating minimization scheme.)

The matched filter approach is the one that has been implemented in the end. And this is the one that will be tested and verified in the following Chapters.

Both of these approaches require some simplifications to the model derived in Chapter 3. These are explored in Section 4.3.

4.1 Notation

This Chapter introduces some notation for writing probability density functions and probability mass functions. These may be confusing for readers outside of the statistical signal processing field. Notations are mostly based on the books by Kay [23–25], and Hayes [22].

For a discrete random variable K, the cumulative mass function (CMF) is written as

$$\text{Prob}(K \leq k),$$

where $\text{Prob}(\cdot)$ denotes probability. The probability mass function (PMF) is written as

$$\text{Prob}(K = k) = p_K(k).$$

Similarly, for a continuous random variable X the cumulative density function (CDF) is

$$\text{Prob}(X \leq x).$$

Taking the derivative of the CDF leads to the probability density function (PDF)

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{Prob}(X \leq x) = p_X(x).$$

Usually in this work, random variables come in uncapitalized letters. This leads to the notations $p_k(k)$ and $p_x(x)$, which do not make sense in the $\mathrm{Prob}(\cdot)$ notation. It does not lead to trouble however, because the random variable X and indexing variable x never arise in the same expression.

For ease of notation, the subscript is sometimes also dropped so that

$$p_x(x) = p(x).$$

In contexts where this might lead to confusion, such as for a change of variables, the subscript is kept.

Parametrized probabilities are given by $p_x(x;\theta)$, where θ is the parameter. Conditional probabilities are given by $p_{x|y}(x|y)$, where y is another random variable. Joint probabilities are given by $p_{x,y}(x,y)$. Mixes of these three are also possible. For complex random variables $z = a + jb$, the PDF is defined as a joint probability, so $p_z(z) = p_{a,b}(a,b)$.

For a time-series measurement, we get the sequence of values $x[n]$ for $n = 1, \ldots, N$. For ease of notation, vectors are introduced where

$$\mathbf{x} = \begin{bmatrix} x[1] \\ \vdots \\ x[N] \end{bmatrix}.$$

The joint PDF of all $x[n]$ is then

$$p_{x[1],\ldots,x[N]}(x[1],\ldots,x[N]) = p_{\mathbf{x}}(\mathbf{x}).$$

The marginal probability

$$p(x) = \int_y p(x,y)\ \mathrm{d}y$$

is written with a single integral, whether it used for single random variable or a random vector. So,

$$\begin{aligned} p_{\mathbf{x}}(\mathbf{x}) &= p_{x[1],\ldots,x[N]}(x[1],\ldots,x[N]) \\ &= \int \cdots \int p_{x[1],\ldots,x[N],y[1],\ldots,y[N]}(x[1],\ldots,x[N],y[1],\ldots,y[N])\ \mathrm{d}y[1]\ldots\ \mathrm{d}y[N] \\ &= \int p(\mathbf{x},\mathbf{b})\ \mathrm{d}\mathbf{b}. \end{aligned}$$

Marginal probabilities of complex random variables are also written as a single integral. So, for $x = a + jb$ and $y = c + jd$

$$\begin{aligned} p(x) &= \iint p(a,b,c,d)\ \mathrm{d}c\,\mathrm{d}d \\ &= \int p(x,y)\ \mathrm{d}y. \end{aligned}$$

and similarly for complex random vectors $\mathbf{x} = \mathbf{a} + j\mathbf{b}$ and $\mathbf{b} = \mathbf{c} + j\mathbf{d}$

$$\begin{aligned} p(\mathbf{x}) &= \iint p(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})\ \mathrm{d}\mathbf{c}\,\mathrm{d}\mathbf{d} \\ &= \int p(\mathbf{x}, \mathbf{b})\ \mathrm{d}\mathbf{b}. \end{aligned}$$

Finally, the $*$ operator is denoted as a convolution, also for vectors.

4.2 Direct Derivation of Optimal Solution

For an optimal solution we should first find the statistical description of the measured data $s[n]$ parametrized by the photon rate λ, or $p(\mathbf{s}; \lambda)$. This is a classical estimation approach where λ is viewed as a constant, and not a random variable. We will borrow some techniques from Bayesian estimation however, by looking at conditional probabilities such as

$$p(\mathbf{s}; \lambda) = \sum_{\mathbf{k}} p(\mathbf{s}|\mathbf{k}; \lambda) p(\mathbf{k}; \lambda), \tag{4.1}$$

where $\mathbf{k}$ is the discrete random vector containing the number of photons that hit the MKID at sample n.

In the first Subsection, $p(\mathbf{s}; \lambda)$ will be written out in conditional PDFs of the signals in the system model. The second and third Subsections will try to find expressions for those conditional PDFs. They will build from the simplest case where we have exact knowledge of the photons, increasing the complexity in the system with each step. This approach is cut short very quickly unfortunately, because the second step in this approach already leads to an intractable expression. Nevertheless, the final Subsection will explain an important result that comes from this exploration of the PDFs.

4.2.1 Solution Based on Complete Model

In order to find an expression for $p(\mathbf{s}; \lambda)$, we resort to conditional PDFs. We use the signal model found in Section 3.4. To repeat:

$$s[n] = q[n] + v_a[n]$$

with $\mathbf{v}_a \sim \mathcal{CN}(0, \mathbf{\Sigma}_a)$ and thus $p(\mathbf{v}_a; \mathbf{\Sigma}_a)$ is known;

$$q[n] = g(b[n])$$

with g being a one-to-one function;

$$b[n] = a[n] + jv_f[n]$$

with $\mathbf{v}_f \sim \mathcal{N}(0, \mathbf{\Sigma}_f)$; and

$$a[n] = Ck[n] * h[n]$$

with $k[n] \sim \mathcal{P}(\lambda)$. $\mathcal{N}$ is the normal distribution, $\mathcal{CN}$ is the complex normal distribution, and $\mathcal{P}$ is the Poisson distribution. The matrices Σ_f and Σ_a are the autocovariance matrices of the noise. They can be estimated from measurements without photons, just like the PSD can be estimated. Estimation of noise characteristics from experiments will be described in Section 5.1.

From here on out the probabilities will be written without parametrization on λ, since almost every probability contains λ as a parameter except the ones introduced by noise ($p(\mathbf{v}_f;\mathbf{\Sigma}_f)$ and $p(\mathbf{v}_a;\mathbf{\Sigma}_a)$).

$$\begin{aligned} p(\mathbf{s}) &= \int_{\mathbf{q}} p(\mathbf{q})p_{\mathbf{v}_a}(\mathbf{s}-\mathbf{q})\ \mathrm{d}\mathbf{q} \\ &= \int_{\mathbf{q}} p(\mathbf{q})p_{\mathbf{v}_a}(\mathbf{s}-g(\mathbf{b}))\ \mathrm{d}\mathbf{q} \\ &= \int_{\mathbf{b}} p(\mathbf{b})p_{\mathbf{v}_a}(\mathbf{s}-g(\mathbf{b}))\ \mathrm{d}\mathbf{b} \\ &= \int_{\mathbf{b}} p(\mathbf{b})p_{\mathbf{s}|\mathbf{b}}(\mathbf{s}|\mathbf{b})\ \mathrm{d}\mathbf{b}, \qquad (4.2) \\ \text{where}\quad & p(\mathbf{s}|\mathbf{b}) = p_{\mathbf{v}_a}(\mathbf{s}-g(\mathbf{b})) \end{aligned}$$

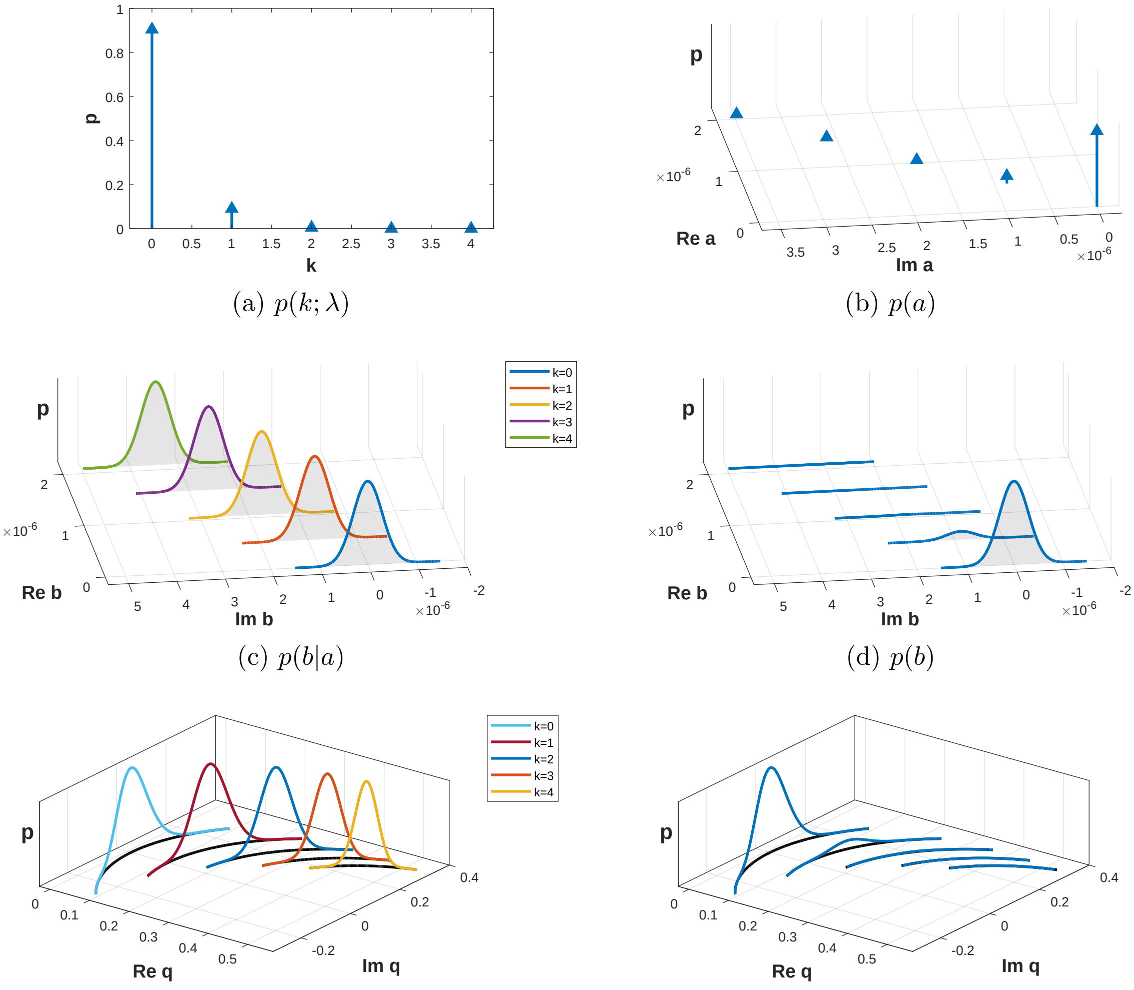
To get $p(\mathbf{b})$ we similarly do

$$\begin{aligned} p(\mathbf{b}) &= \sum_{\mathbf{a}} p(\mathbf{a})p_{j\mathbf{v}_f}(\mathbf{b}-\mathbf{a}) \\ &= \sum_{\mathbf{a}} p(\mathbf{a})p_{\mathbf{v}_f}((\mathbf{b}-\mathbf{a})/j) \\ &= \sum_{\mathbf{k}} p(\mathbf{k})p_{\mathbf{v}_f}((\mathbf{b}-C\mathbf{k}*\mathbf{h})/j) \\ &= \sum_{\mathbf{k}} p(\mathbf{k})p_{\mathbf{b}|\mathbf{k}}(\mathbf{b}|\mathbf{k}) \qquad (4.3) \\ \text{where}\quad & p(\mathbf{b}|\mathbf{k}) = p_{\mathbf{v}_f}((\mathbf{b}-C\mathbf{k}*\mathbf{h})/j). \end{aligned}$$

Here a sum over $\mathbf{a}$ is used instead of an integral, because $\mathbf{k}$ is a discrete random vector and the convolution $C\mathbf{k}*\mathbf{h}$ must then also be a discrete random vector. This result requires that we can find $p_{\mathbf{v}_f}((\mathbf{b}-C\mathbf{k}*\mathbf{h})/j)$ from $p_{\mathbf{v}_f}((\mathbf{b}-\mathbf{a})/j)$ [26] (i.e. $\mathbf{k}$ can be found exactly from $\mathbf{a}$), which is not necessarily true.

Figure 4.1 shows examples of what the PDFs and conditional PDFs could look like for MKIDs. The PDF of $\mathbf{s}$ would look similar to Figure 4.1f, excepts for a 2-D convolution of $p(a)$ with a circular Gaussian PDF because of the noise introduced by v_a. This results in a smearing along the real and imaginary axis.

Now, we have expressions for $p(\mathbf{k})$, $p(\mathbf{v}_f)$, and $p(\mathbf{v}_a)$; and we need to find an expression $p(\mathbf{b})$ and $p(\mathbf{s})$ by evaluating (4.2) and (4.3). We will start by working up from simple the simple case where $\mathbf{b}=\mathbf{k}$ and increasing complexity until we have $p(\mathbf{b})$ given $\mathbf{b}=C\mathbf{k}*\mathbf{h}+j\mathbf{v}_f$ in the next Subsections. This process reaches a dead end however in Section 4.2.3.



(a) $p(k;\lambda)$ (b) $p(a)$

(c) $p(b|a)$ (d) $p(b)$

(e) $p(q|a)$ (f) $p(a)$

Figure 4.1: Examples of PDFs and conditional PDFs of random variables in (4.2) and (4.3). The PDFs shown here are of single random variables instead of random vectors for easier viewing. The convolution $\mathbf{a} = C\mathbf{k} * \mathbf{h}$ is then also collapsed to $a = Ck$. Notice the delta spike behaviour of $p(b|a)$ and $p(b)$ in the $\mathfrak{Re}\{b\}$-direction.

4.2.2 Estimation of Poisson Parameter

Let's first look at a single Poisson random variable k with PMF

$$p_k(k;\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}. \tag{4.4}$$

The score function of this distribution is

$$\text{score}(k;\lambda) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\log(p_k(k;\lambda)) = \frac{k}{\lambda} - 1 \tag{4.5}$$

and the Cramér-Rao lower bound (CRLB) is given by $\text{Var}\{\hat{\lambda}\} \geq 1/I(\lambda) = \lambda$, where $I(\cdot)$ is the Fisher information. Since the score function can be written as $\text{score}(k;\lambda) =$

$I(\lambda)(g(k)-\lambda) = \frac{1}{\lambda}(k-\lambda)$, the minimum variance unbiased (MVU) estimator is simply $\hat{\lambda} = k$. An MVU estimator is optimal in that it reaches the CRLB while being an unbiased estimator [23].

For a timestream $k[n]$ of IID random variables, the Fisher information $I(\lambda) = Ni(\lambda)$, where N is the number of random variables and $i(\lambda)$ is the Fisher information for each random variable. The CRLB is thus given by $\mathrm{Var}\{\hat{\lambda}\} \geq \lambda/N$ and the MVU estimator is given by $\hat{\lambda} = \frac{1}{N}\sum k[n]$. This is simply the time average of the number of photons

4.2.3 Estimation of Poisson Parameter in Noise

Seemingly still a simple case, adding noise to the previous problem already makes the problem quite complex. The data model is now $b[n] = k[n] + v[n]$, where $v[n]$ is real Gaussian noise with mean 0 and variance σ_v^2. Since $b[n]$ is the sum of two random variables, the PDF of $b[n]$ is the convolution of the PDFs of $k[n]$ and $v[n]$. For a single random variable $b = k + v$,

$$p_b(b;\lambda,\sigma_v^2) = \int_{-\infty}^{\infty} p_k(b-b')p_v(b')\ \mathrm{d}b'.$$

There is one problem, which is that p_k defines a discrete PMF. We overcome this issue by treating it as a continuous PDF with Dirac delta functions at the integer values. p_b thus becomes

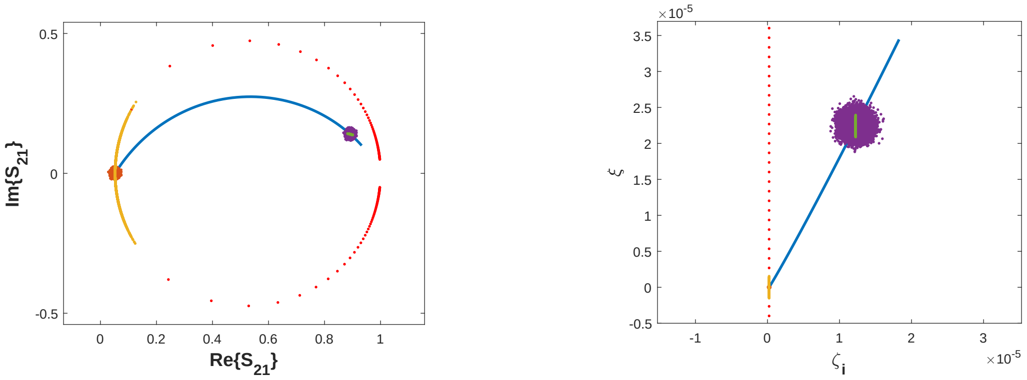
$$p_b(b;\lambda,\sigma_v^2) = \int_{-\infty}^{\infty} p_v(b-b')\sum_{i=0}^{\infty}\frac{e^{-\lambda}\lambda^{b'}}{b'!}\delta(b'-i)\ \mathrm{d}b' \tag{4.6}$$

$$= \sum_{i=0}^{\infty}\frac{e^{-\lambda}\lambda^{i}}{i!}p_v(b-i), \tag{4.7}$$

which is an infinite sum of Gaussian bells centered around the nonnegative integers and scaled by the Poisson PMF. No closed-form expression via a series simplification or by plugging it into Mathematica was found. This PDF was therefore not useful in finding the CRLB or an MVU. The tactic of directly finding an optimal solution thus also reaches a dead end.

4.2.4 Unbiased Estimator of Photon Hits

The form $b[n] = k[n] + v[n]$ can be further analyzed. Let's say we are able to find an estimator of the number of photons at each sample. For example via a detection scheme such as in Section 4.4 or via directly estimating $k[n]$ via deconvolution. This can be written as the actual number $k[n]$ plus a residual or error $e[n]$, $\hat{k}[n] = k[n] + e[n]$. If $\hat{k}[n]$ is an unbiased estimator, so $\mathbb{E}\{\hat{k}[n]\} = k[n]$, then the sum of $\hat{K} = \sum \hat{k}[n]$ is also an unbiased estimator of the total number of photons in the measurement $K = \sum k[n]$. The average number of estimated photons $\hat{\lambda} = \frac{1}{N}\sum \hat{k}[n]$ is then also an unbiased estimator of the photon rate. Bringing down the variance of the error should bring $\hat{\lambda}$ closer to the MVU estimator of the photon rate.



(a) S_{21}-space (b) r-space

Figure 4.2: Scatter plots of simulated noise on an MKID circle in S_{21}-space and in r-space. The red dots show the resonance circle and the blue line shows the QP-response for reference. Orange shows simulated amplifier noise around a measurement at low χ vs. purple around a measurement at high χ. Yellow shows simulated frequency noise around a measurement at low χ vs. green around a measurement at high χ.

This forms the mathematical basis for the fact that detecting the photon pulses in an MKID is a valid way to estimate the photon rate. Furthermore, any unbiased estimator of $k[n]$ can be used to estimate the photon rate, even without a photon pulse detection scheme.

4.3 Estimation in r-space

Now that directly finding an estimator via inspection of the PDF of the measurement does not work out, we will have to resort to simplifications of the model to make it easier to work with. One approach is to transform the measurement to r-space and see what the model then looks like. Transforming to r-space makes the read-out linearly dependent on the $\chi[n]$ signal, which in turn makes it linearly dependent on $k[n]$. However, transforming the measurement to r-space distorts the noise. Figure 4.2 shows an example of this distortion by showing simulated noise for a measurement at low vs. at high quasiparticle density. In the S_{21}-space, the variance of the amplifier noise is independent of the $\chi[n]$, but the frequency noise is dependent on the $\chi[n]$. In r-space, it is the other way around and the amplifier noise blows up at higher signal read-outs. In the example in Figure 4.2, the amplifier noise variance is 821 to 831 times higher at high χ than at low χ.

We will continue this Section by checking if the added amplifier noise can modeled as additive Gaussian noise in r-space. Figure 4.3 shows the model after transforming the measurement $s[n]$ to r-space. It introduces the new signal

$$c[n] = g^{-1}(s[n]). \tag{4.8}$$

Figure 4.4 shows the model assuming $c[n]$ can be written as

$$c[n] = g^{-1}(s[n]) \approx a[n] + jv_f[n] + v_a'[n]. \tag{4.9}$$

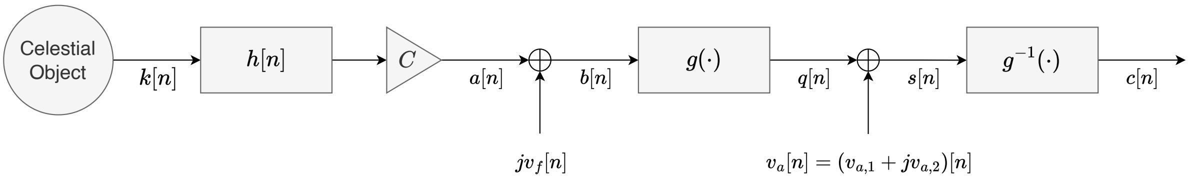


Figure 4.3: System model after transforming measurement to r-space.

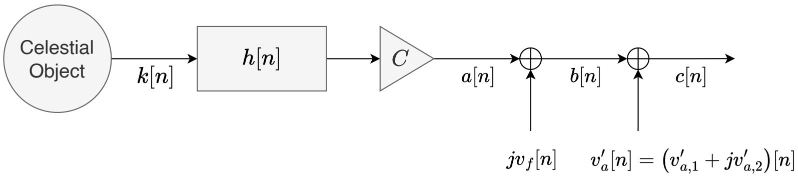


Figure 4.4: Model assuming $v_a[n]$ can be modeled as additive Gaussian noise in r-space.

To verify if we can model $v_a[n]$ as additive Gaussian noise as $v'_a[n]$ in r-space, we will first check if the noise stays Gaussian at low and at high signal read-out. We will do so again by checking quantile-quantile plots. Figure 4.5 shows the noise stays Guassian even at high signal read-out. The quantile-quantile plots show relatively straight lines, which imply Gaussian distribtutions. There is some distortion, but not enough to discard a Gaussian model.

We have now found that $v'_a[n]$ can be modeled as additive Gaussian noise, but the noise variance is still dependent on χ, or equivalently on $|a[n]|$. To find out the dependence of the noise power on $|a[n]|$, we will check what happens a point one standard deviation σ_a away from the quasiparticle response curve. Or in other words, what is the relation between the noise strength of the orange and the purple block in Figure 4.2b when compared to $|a[n]|$. We will use the following metric

$$\text{Noise Scaling} = \frac{|g^{-1}(S_{21}(\chi) + 1\sigma_a) - g^{-1}(S_{21}(\chi))|}{(S_{21}(\chi) + \sigma_a) - S_{21}(\chi)} = \frac{|g^{-1}(|S_{21}(\chi) + 1\sigma_a) - r(\chi)|}{\sigma_a} \tag{4.10}$$

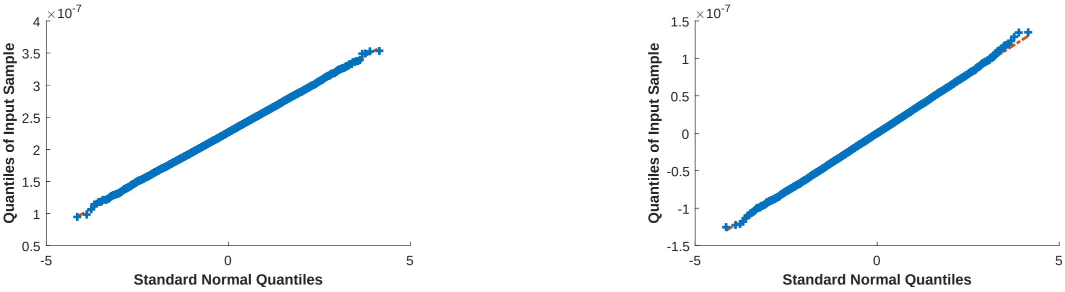
to inspect the relation. Figure 4.6 shows how far a point that has moved $1\sigma_a$ away from a point on the χ-response curve moves when transformed to r-space. A 2^{nd} order polynomial fits the found points almost perfectly. This implies we can write our data model as

$$\begin{aligned} a[n] &= Ck[n] * h[n] \\ c[n] &= a[n] + jv_f[n] + v'_a[n], \end{aligned} \tag{4.11}$$

where $v'_a[n] = (D|a[n]|^2 + E|a[n]| + G)v''_a[n]$, and $v''_a[n]$ is colored Gaussian noise.

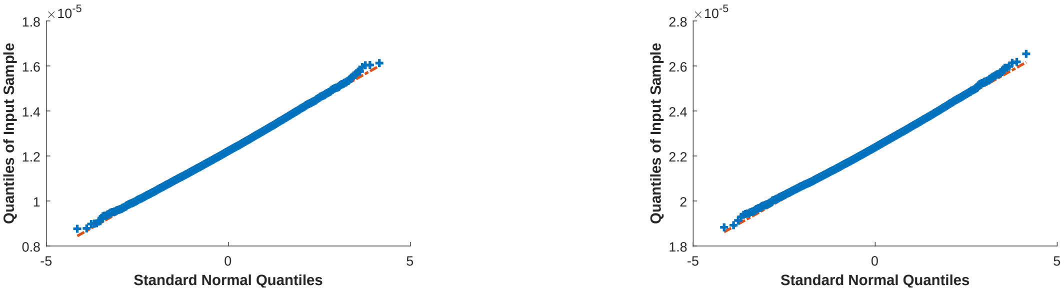
Now assuming we only have small signal strengths, i.e. $|a[n]|$ is fairly small, we can take $D|a[n]|^2 + E|a[n]| + G \approx G$. We are now left with

$$\begin{aligned} a[n] &= Ck[n] * h[n] \\ c[n] &= a[n] + v[n], \end{aligned} \tag{4.12}$$



(a) $\mathfrak{Re}$ amp-noise @low χ

(b) $\mathfrak{Im}$ amp-noise @low χ



(c) $\mathfrak{Re}$ amp-noise @high vs. low χ

(d) $\mathfrak{Im}$ amp-noise @high vs. low χ

Figure 4.5: Quantile-quantile plots of simulated amplifier noise in r-space at low and at high signal read-out. (a-b) show quantile-quantile plots of the noise at low signal read-out compared to a perfect Gaussian. (c-d) show quantile-quantile plots of the noise at high signal read-out compared to perfect Gaussian.

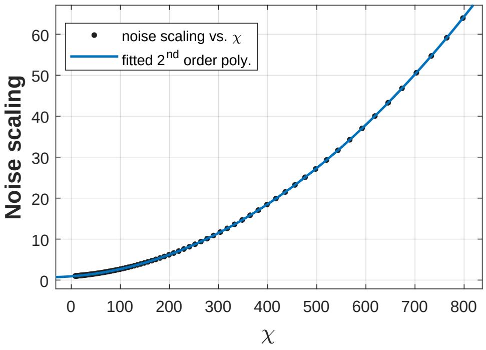


Figure 4.6: Scaling of simulated amplifier noise as a function of signal read-out strength.

where $v[n] = jv_f[n]+v'_a[n]$ is additive temporally colored Gaussian noise, with a stronger variance in the imaginary direction than in the real direction. The temporal correlation in $v[n]$ results from $jv_f[n]$ and $v'_a[n]$ both also being temporally colored noise.

4.4 Detection Using Matched Filter

In this Section we will explore reconstructing $k[n]$ by detecting pulses in the r-space. As shown in Section 4.3, we can model measured data according to

$$c[n] = k[n] * Ch[n] + v[n]. \tag{4.13}$$

Detecting whether a signal is present in colored Gaussian noise leads to an optimal detector called the generalized matched filter [25, Algorithm 10.3]. In our case the signal to detect is the exponential decay $h[n]$, since $k[n] * h[n]$ can be seen as $h[n]$ arriving at multiple time instances n_i.

The generalized matched filter as described in [25, Algorithm 10.3] will produce a test statistic for a only single time n. It is a measure for the likelihood that a photon had arrived at that time n. To apply this filter to multiple n, the generalized matched filter will be combined with an optimal estimation of time delay algorithm [25, Algorithm 9.4], which is a matched filter implemented as a convolution. The output of the filter is then a timestream of the likelihood of a photon pulse being present in the signal. In order to find $\hat{k}[n]$, a CLEAN-like algorithm should be applied to iteratively find the pulse locations from the filter output [27].

This Section will build up an implementation of a matched filter and end with a short description of what a CLEAN-like algorithm would look like. This matched filter implementation works on real data, whereas S_{21} data from an MKID is complex. Section 5.3 will explain how to manipulate the data so it can be used in this 1-dimensional matched filter. In this buildup generic signal names such as $x[n]$ and $y[n]$ will be used.

4.4.1 Building the Generalized Matched Filter

Matched Filter in White Noise

The first data model consists of

$$x[n] = p[n] + w[n], \tag{4.14}$$

where

$$p[n] = k[n] * h[n] \tag{4.15}$$

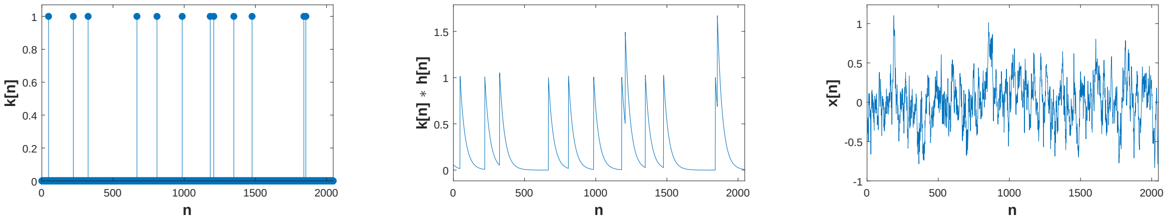
and $w[n]$ is white noise with variance σ_w^2. We take for $h[n]$ a finite impulse response (FIR) approximation. This is possible because $h[n]$ is an exponential decay and will go to zero. For example, for a τ of 0.7 ms and a sampling frequency F_s of 50 kHz, the length M of $h[n]$ will be 280 after 8τ. Figure 4.7 shows signals $x[n]$ consists of.

The matched filter is just a convolution of $x[n]$ with $h[-n]$.

$$y[n] = x[n] * h[-n] \tag{4.16}$$

For a single n this can regarded as an inner product. Introducing the vectors

$$\mathbf{x}_n = \begin{bmatrix} x[n] \\ \vdots \\ x[n+M-1] \end{bmatrix}, \tag{4.17}$$



(a) Photons (b) Pulses (c) Noisy signal

Figure 4.7: Examples of the components of a noisy signal. The discrete signal points have been connected with a line for easier viewing.

and

$$\mathbf{h} = \begin{bmatrix} h[0] \\ \vdots \\ h[M-1] \end{bmatrix}, \tag{4.18}$$

each $y[n]$ is the inner product of $\mathbf{x}_n$ and $\mathbf{h}$. The inner product is a test statistic for how well the data shape matches the signal shape. In order to compute the test statistic for each n you shift the input $x[n]$ and do the inner product. That is where the convolution comes from.

The output $y[n]$ is thus a test statistic at each n. It is given in standard deviations (although (4.16) is not scaled right) and can be converted to probability via the Q function. Normally the point with the highest test statistic score will be taken. Here we have an unknown number of photons. That's why an algorithm like CLEAN is necessary. This Section will continue with building the test statistic $y[n]$ without trying to find the photons themselves.

The test statistic is the optimal solution to maximizing the MLE.

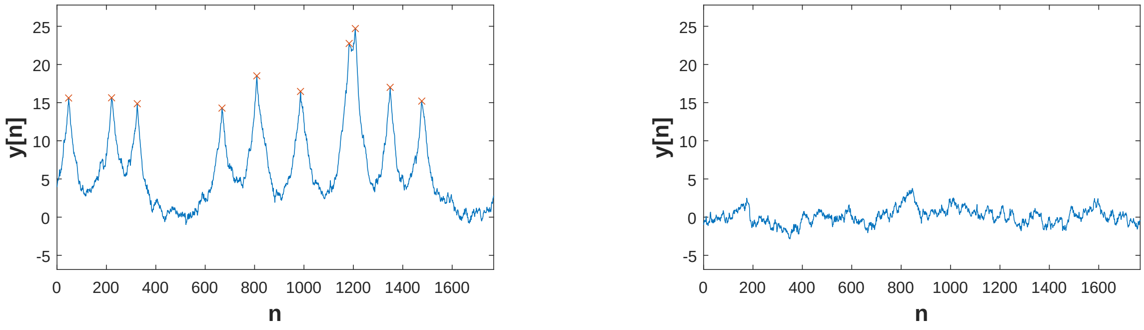
Figure 4.8 shows the output of a matched filter for a signal with photon pulses and a signal with white noise only. The filtered signal is the one from Figure 4.7c. When white noise is filtered with a matched filter it leaves colored noise with variance 1 (and thus a standard deviation of 1).

Matched Filter in Colored AR Noise

An MKID has temporally colored noise. Temporally colored noise will make the solution suboptimal. Especially in this case where both the noise and the signal have high power in low frequencies.

As a first example, an auto-regressive process of order 1 (AR(1)) is taken for the noise with its pole at 0.9. It has more power in the low frequencies just like MKID noise. An AR(1) is chosen because the second-order statistics are well known and easy to implement.

Figure 4.10 shows the filter output for white noise vs. AR(1) colored noise. The peaks are less pronounced and there are peaks in the output where no photon came in. The matched filter loses performance in AR(1) noise.



(a) With pulses

(b) Noise only

Figure 4.8: Examples of matched filter output. (a) shows the matched filter has peaks at exactly the time instances where a photon came in, the orange crosses mark the pulse locations. (b) show the matched filter output for noise only. The discrete signal points have been connected with a line for easier viewing.

Generalized Matched Filter in Colored AR Noise

In order to improve the test statistic, [25, Algorithm 10.3] can be used. It can be interpreted as a matched filter which takes in prewhitened data, and a prewhitened signal form.

The new equation for the filter output is

$$y[n] = \mathbf{x}_n^{\mathsf{T}} \mathbf{\Sigma}_w^{-1} \mathbf{h}. \tag{4.19}$$

This is not ready yet to use for larger data sequences. It can only compute a single $y[n]$. In [24] it is mentioned that $\mathbf{\Sigma}_w^{-1}$ can be decomposed into

$$\Sigma_w^{-1} = \mathbf{R}^{\mathsf{T}} \mathbf{R}\,. \tag{4.20}$$

The output then becomes

$$y[n] = (\mathbf{R}\,\mathbf{x}_n)^{\mathsf{T}} \mathbf{R}\,\mathbf{h}, \tag{4.21}$$

which can be performed as a convolution of a whitened $x[n]$ with a whitened $h[-n]$.

The whitening of $x[n]$ can be tricky in practice. Although $\mathbf{R}$ is here an M-by-M matrix, which is fine for $h[-n]$, to whiten the entire N long sequence $x[n]$ would require an N-by-N whitening matrix. The examples of the output in the Figure are therefore also only of a small N, but still with $M \gg N$.

Figure 4.11 shows the filter output for a regular matched filter and a generalized matched filter in colored AR(1) noise. The peaks in the signal are more pronounced again and their height is more equal. The noisy parts have a larger high-frequency component, it looks more jagged. There is a scaling error in the generalized matched filter, so the absolute height of the peaks does not matter here, only the shape.

Moving to a Frequency Domain Representation

A frequency domain representation is nicer than the matrix form because it requires much less calculations and memory. It also allows other signal processing techniques to be used such as the overlap-add or overlap-save methods.

The filter output is [24]

$$y[n] = \text{DTFT}^{-1}\left(\frac{X(F)\overline{H}(F)}{P_w(F)}\right), \tag{4.22}$$

where $X(F)$ is the discrete-time Fourier transform (DTFT) of $x[n]$, $\overline{H}(F)$ is the conjugate of the DTFT of $h[n]$, and $P_w(F)$ is the PSD of the AR(1) noise.

There lies a problem in the length of the DTFT. The longest length is that of $x[n]$ and is N. So in order to get $y[n]$ for all N samples the N-point DTFT would have to be applied to $h[n]$, and $P_w(F)$ would need to be calculated for N frequencies. This leaves a very inefficient calculation. This is where the overlap-add and the overlap-save method come in [28, 29]. This allows us to do the frequency domain calculations in blocks of size N_{block} with $M < N_{block} < N$. Doing multiple N_{block}-sized computations can be much faster than a single N-sized one. The example of the filter output is of a single block.

This example still uses the AR(1) because the PSD is readily given without having to know the autocorrelation function. This is helpful because later the PSD of the MKID noise can be easily estimated and verified, while the characteristics of the autocorrelation of the MKID noise are not well known.

Figure 4.12 shows the filter output when calculated via the autocovariance matrix vs. via the PSD. The two figures are almost not discernible from each other besides a scaling error. The two outputs are not exactly equal but they are equivalent.

Matched Filter in Colored MKID Noise

This is finally the implementation that is used on experimental data in Section 5. Here not an AR(1) spectrum will be used, but a spectrum that matches the noise of an MKID.

The filter output is the same as from the previous Section, but with a PSD from MKID noise and is

$$y[n] = \text{DTFT}^{-1}\left(\frac{X(F)\overline{H}(F)}{P_v(F)}\right), \tag{4.23}$$

where $P_v(F)$ is the PSD of the MKID noise.

The MKID noise can be found theoretically, although in practice it will be very hard to match it to the actual noise in an MKID. We have the frequency noise and amplifier added together, and the data points are rotated on top that (see Section 5.3). The actual noise power is also not known in advance from the MKID, because it depends on the design parameters, fabrication quality, and bath temperature. An estimation of the spectrum is therefore used in the actual computation. This can be done via Welch's spectrum estimation [22]. The windowing function used here is a Blackmann filter since the spectrum is rather smooth.

Figure 4.13 shows an example of a matched filter that uses an estimated PSD via Welch's method. Even though the signal is quite noisy, the generalized matched filter manages to produce quite prominent peaks. The regular matched filter shows some peaks on the right that are much lower than the ones on the left and those would have probably been missed.

There is still a scaling error in the output. The scaling error can be removed by normalizing it with the square root of the variance of the output of the generalized matched when only noise is put in.

4.4.2 CLEAN Algorithm

The generalized matched filter output consists of peaks at the photon pulse location. Extracting the peaks from the signal can be done via CLEAN [27]. CLEAN has not been implemented in this work, but it is mentioned because it would make for an easy next step to implement.

In short, a CLEAN algorithm works by iteratively removing pulses from the timestream that best match the pulse shape. So let's say $y[n]$ reaches the max at time n_{max}. Then if the expected pulse height is known, a copy of $h[n] * h[-n]$ would be removed from the filtered $y[n]$ at time n_{max}; and a copy of $h[n]$ would be removed from $x[n]$. This process is repeated until $x[n]$ looks like only noise. The exact metric when to stop is not exactly given and needs to be devised. It can be based on the PSD of $x[n]$ or on the quantiles of $x[n]$.

The removed pulses and timestamps are used to reconstruct the original signals $k[n]$ and $k[n] * h[n]$.

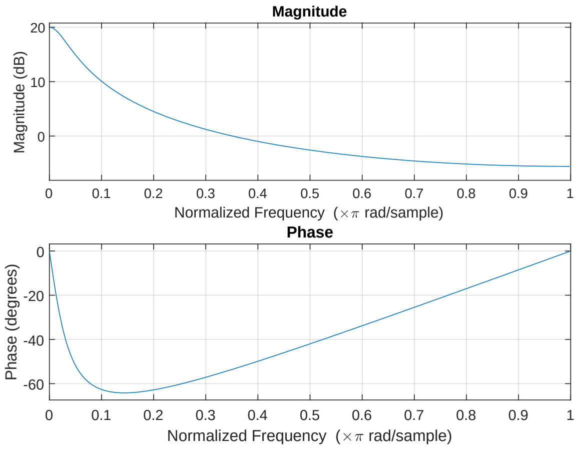
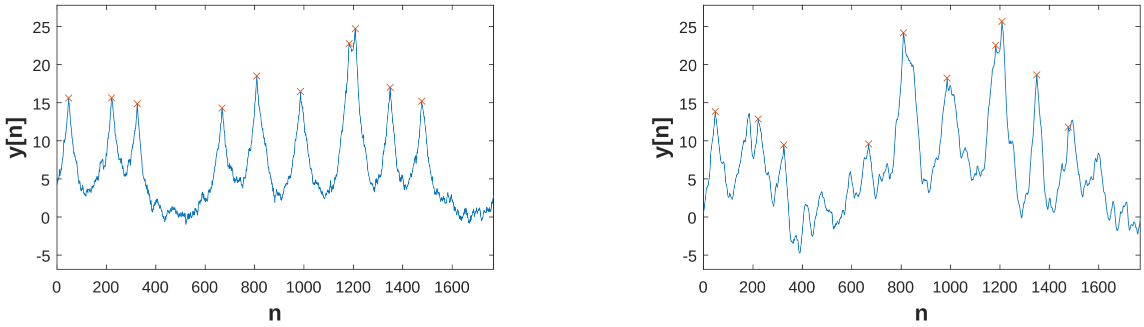


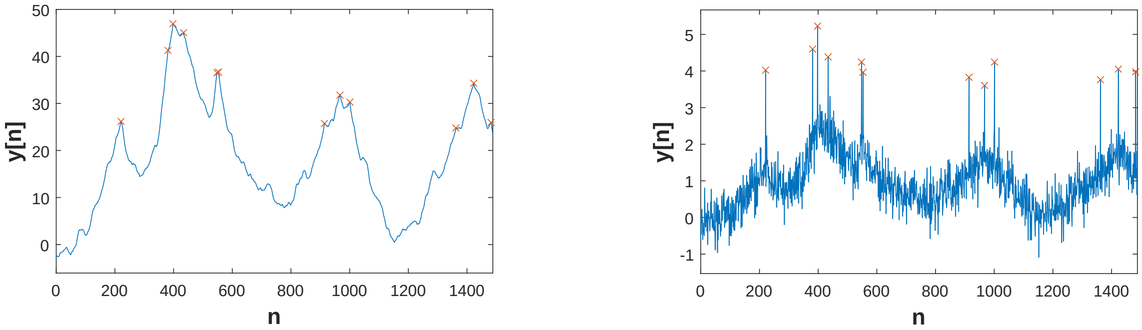
Figure 4.9: Power spectral density and phase of the filter used to create AR(1) noise.



(a) White noise

(b) Colored noise

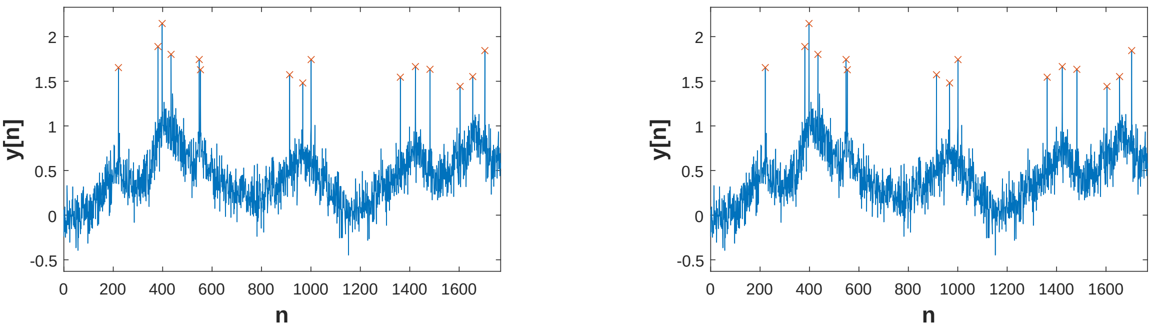
Figure 4.10: Examples of matched filter output with white added Gaussian noise vs. colored added Gaussian noise. Orange crosses mark the pulse locations. Noise power is the same in both cases. The discrete signal points have been connected with a line for easier viewing.



(a) Regular

(b) Generalized

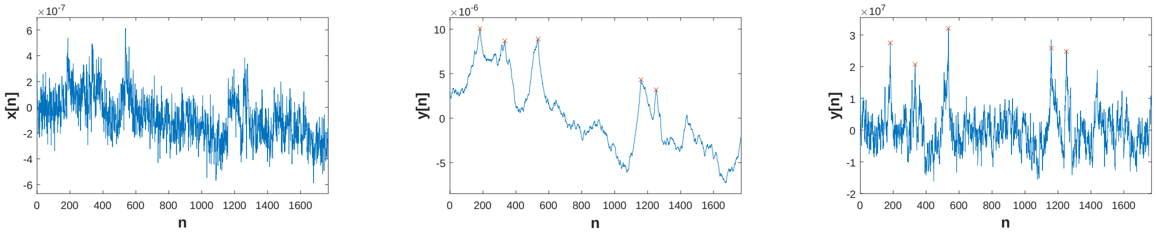
Figure 4.11: Examples of output regular matched filter vs. a generalized matched filter on pulses in AR(1) noise. Orange crosses mark the pulse locations. The discrete signal points have been connected with a line for easier viewing.



(a) Matrix

(b) PSD

Figure 4.12: Examples of output of generalized matched filter in AR(1) noise calculated via the autocovariance matrix vs. via the PSD. Orange crosses mark the pulse locations. The discrete signal points have been connected with a line for easier viewing.



(a) Time-stream

(b) Matched filter

(c) Generalized

Figure 4.13: Examples of input and output of generalized matched filter used on a synthesized MKID signal. For comparison the filter output of a regular matched filter is also added. The discrete signal points have been connected with a line for easier viewing.

Experimental Verification

5

This Chapter presents the results of applying the matched filter, introduced in Chapter 4, to experimental data. Before giving the results, it will describe the experimental set-up that has been used to gather the measurement data. Two datasets have been recorded, one of LT218 with data from six MKIDs, and one of LT366 with data from one MKID.

Following the description of the experimental setup, cosmic rays will be briefly discussed. These are events that are not modeled in previous Chapters, but they can be removed from the data rather easily.

Afterwards, the necessary pre-processing steps applied to the data before running the matched filter are discussed. These steps include flattening the complex data to real data, spatial noise whitening, and estimation of noise characteristics, and an estimation of the gain C. The pre-processing is necessary to reduce the noise as much as possible.

Finally, the outcomes of applying the generalized matched filter will be presented for a photon rate varying from zero to hundreds per second. There is no definitive performance metric to compare the old method to the one presented in this work. Therefore, a qualitative analysis of the time-stream data is performed.

5.1 Measurement Set-Up

The MKID can be tested in a cryostat. There are multiple MKIDs on a single chip that can be read out at the same time using suitable electronics. However, in our experiment we use a single tone read-out system measuring one detector at a time. The advantage of this method is that we can sample faster than the characteristic timescales (the resonator ringtime and the quasiparticle lifetime). The chip can be cooled down to around 120 mK in the cryostat [2]. A blackbody radiator is also placed in the cryostat and its temperature can be controlled between 3 K and 30 K [13]. The blackbody radiates on the MKID and this radiation is what is measured. A set of quasi-optical filters is placed between the blackbody and the chip that only lets through radiation in a narrow band around a specific frequency ν. When the blackbody temperature increases it will radiate more power which increases the photon rate.

This work examines the dataset of two chips:

1. LT218 from which six MKIDs have been measured,
2. and LT366 from which one MKID has been measured.

LT218 was measured with a 12 THz filter, and LT366 was measured with a 7 THz filter. The filters let through a range of frequencies around the center frequency, but the range is small enough that the photon frequency can be considered constant.

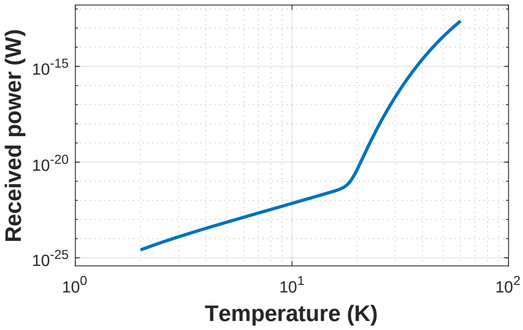


Figure 5.1: Received power from the blackbody as a function of the blackbody temperature. The blackbody radiation is filtered by a set of quasi-optical filters at a frequency of $\nu = 12$ THz.

This set-up allows measurements to be taken of only noise, done with the blackbody at 3K where there is negligible power emitted, and of a varying photon rate, which is achieved by increasing the blackbody temperature. The power falling on the chip can be calculated from Planck's law and the exact filter transmission [2]. This enables analysis of the noise characteristics and verification of estimation of the photon rate. Figure 5.1 shows the received radiation power from the blackbody after being filtered as a function of blackbody temperature.

5.2 A Word on Cosmic Rays

Cosmic rays are high energy particles that collide with the atmosphere and cause a cascade of lighter particles to rain down on earth. The cascade of particles includes particles that can excite quasiparticle from the ground state in the MKIDs. When the particles shower onto the MKID it can create a very high response in all MKIDs on the chip at the same time.

These cosmic rays are therefore very easily filtered out from the time-stream because they produce a response that can be several times higher than that of a photon. They also occur simultaneously in all MKIDs on the chip.

5.3 Pre-Processing

The data that gets filtered with the generalized matched filter should go through some pre-processing beforehand. The filter as defined in Section 4.4 uses real data as input, whereas the data from the MKID is complex. The data is therefore rotated such that the signal lies along the real axis.

Another problem, which was not accounted for in the model, is that in actual MKID data the real and imaginary part of the noise are correlated in space. This correlation can be removed by whitening the real and imaginary part of the noise.

Furthermore, up until this point it was assumed that all the parameters in the system were known, except for the photon rate, but in reality some remain to be estimated beforehand. These parameters should stay the same however in between experiments.

These parameters are the gain C and the PSD of the noise. The time constant of the quasiparticle decay for each MKID has already been estimated and will not be discussed.

5.3.1 Rotating the Signal to the Real Axis

The signal of interest, $a[n] = Ck[n] * h[n]$ lies in the complex plane due to the complex gain C. The generalized matched filter implementation from Section 4.4 can be extended to be used on complex data directly. However, we will use a different approach here, since we know that the signature $h[n]$ is a real signal.

Both the $k[n]$ and $h[n]$ that $c[n]$ is composed of are real signals. Dividing the measurement $c[n] = a[n] + v[n]$ by C will rotate the $a[n]$ term back onto the real axis. Now taking the real part of $c[n]$ will remove all of the noise in the imaginary part of the signal. We can then use the generalized matched filter on this flattened data.

Alternatively, the $c[n]$ can also only be rotated, and not scaled, because in reality we do not know the magnitude of C exactly. The rotation angle is the phase ϕ of C. Then,

$$a_r[n] = a[n]e^{-j\phi} \tag{5.1}$$

is a purely real number. We then also get

$$c_r[n] = c[n]e^{-j\phi} = (a[n] + v[n])e^{-j\phi} = a_r[n] + v_r[n] \tag{5.2}$$

This introduces a problem where the noise $v[n]$ is not spatially uncorrelated anymore. This makes the noise variance in the $\mathfrak{Re}\{c_r[n]\}$ direction larger than if it were uncorrelated. We can spatially whiten it by shearing the data such that the ellipse defined by the eigendecomposition of the spatial noise covariance matrix has its major axis along the imaginary axis and the minor axis is along the real axis. We can then the real part and flatten our data from 2-D to 1-D. This results in losing the noise in the imaginary part of the signal.

5.3.2 Whitening Spatially Correlated Noise

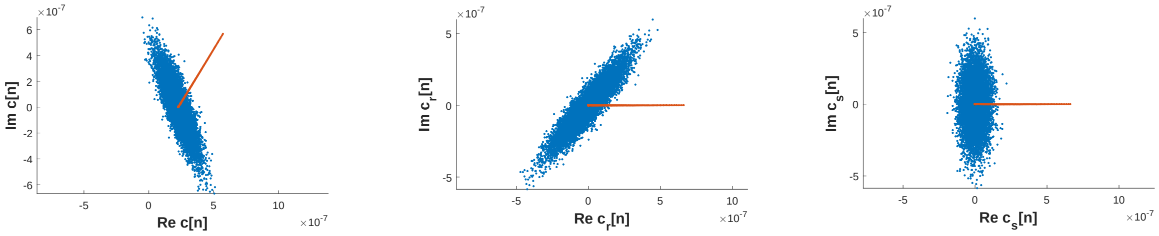
Another problem arises in real data where the major axis of the noise ellipse of $v[n]$ does not lie exactly in the imaginary direction. The reason for this is not exactly known. This problem can be easily overcome however by computing the spatial covariance matrix of a noise-only measurement and doing a Cholesky decomposition on the inverse of the covariance matrix [30]. We can lay out our data in

$$\mathbf{C}_r = \begin{bmatrix} \mathfrak{Re}\{c_r[1]\} \ldots \mathfrak{Re}\{c_r[N]\} \\ \mathfrak{Im}\{c_r[1]\} \ldots \mathfrak{Im}\{c_r[N]\} \end{bmatrix}, \tag{5.3}$$

and similarly for $\mathbf{A}_r$ and $\mathbf{V}_r$. The whitening matrix $\mathbf{W}$ of the 2-by-2 covariance matrix $\mathbf{\Sigma}_{spatial}$ of the real and imaginary part of $v_r[n]$ is then

$$\mathbf{L}\,\mathbf{L}^\mathsf{T} = \mathbf{\Sigma}_{spatial}^{-1} \tag{5.4}$$

$$\mathbf{W} = \mathbf{L}^\mathsf{T}, \tag{5.5}$$



(a) Unrotated (b) Rotated (c) Sheared

Figure 5.2: Simulated noise and signal response when rotated and sheared. Blue dots show a scatter of the noise, orange dots show a scatter of the signal response. Notice how in (a) the primary axis of the noise does not lie along the imaginary axis.

such that $\mathbf{W}$ is a shearing matrix and

$$\mathbf{W}\,\mathbf{C}_r = \mathbf{W}(\mathbf{A}_r + \mathbf{V}_r) = \mathbf{W}\,\mathbf{A}_r + \mathbf{W}\,\mathbf{V}_r\,. \tag{5.6}$$

$\mathbf{W}\,\mathbf{V}_r$ then defines spatially white noise and $\mathbf{W}\,\mathbf{A}_r$ is a scalar times $\mathbf{A}_r$, since $a_r[n]$ has no component anymore along the imaginary axis.

This process of the rotating and shearing the signal is graphically presented in Figure 5.2.

5.3.3 Estimating Noise Characteristics

The experiments for each MKID are performed at blackbody temperatures of roughly 10 K to 30K. At 10 K there are virtually no photons at the filtered photon frequency. That means that a measurement at low temperature can be used to estimate noise characteristics. The characteristics that are of interest here is the PSD of the noise, since that is necessary in the filter. The PSD is estimated of the 1-D real signal of Section 5.3.1 and Section 5.3.2.

Estimation is performed via Welch's method. The windowing function that is used is a Blackman window. A Blackman window smooths the PSD estimate in the frequency domain. This is preferable since from Section 3.5 the PSD is known to be rather smooth. Care has to be taken to not take the window length too short, this leads to power density at low frequencies to be estimated to high.

Alternatively, the PSD can be estimated parametrically using (3.26) and (3.27), although this has not been implemented.

5.3.4 Estimating the Quasiparticle Curve

The quasiparticle curve (and thus the phase of the gain C) can be estimated by looking at the signals from all the measurments together. The measurements will show photon pulses along the quasiparticle curve. However, the pulses in some of these measurements are so low that estimation by looking at photon pulses only can give wrong results.

An alternative is to look at the response from the cosmic rays. These produce responses along the quasiparticle curve, but can be filtered out from the noise much more easily. The angle ϕ is then found by taking the mean of the angles of measurements that belong to the cosmic rays events.

The magnitude of the gain C can not be estimated in this way. Fortunately, this is not necessary for the matched filter. It will only produce a scaling error, but that can be removed later on. The magnitude might be a necessary parameter for other estimation schemes and should then be estimated as a nuisance parameter.

5.4 Generalized Matched Filter Results

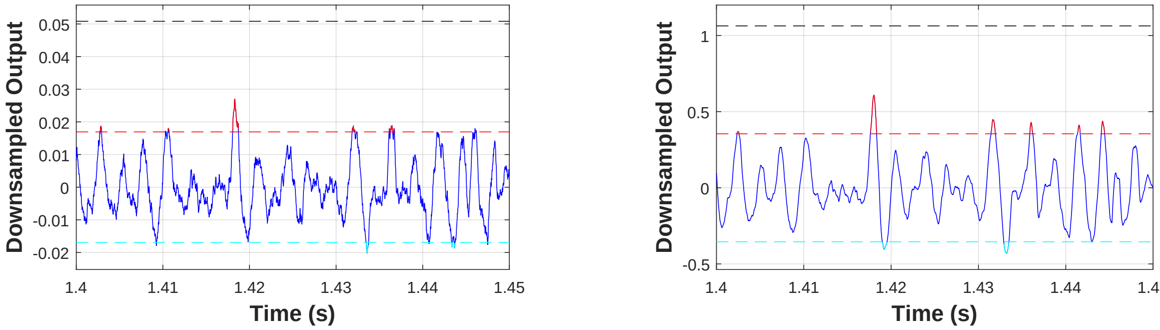
The generalized matched filter has been applied to the datasets of LT218 and LT366. It has been applied for all temperatures and all available MKIDs. One of the most direct ways of comparing this filter approach to previous approaches would be to apply a counting algorithm (such as CLEAN), and comparing the resulting photon rates. The photon rates can then also be compared to the theoretical photon rate that should be radiated by the blackbody.

However, the photon rate is not considered a reliable performance metric as of now. There was no counting algorithm applied that makes full use of filtered output due to time constraints. Instead, the filtered output is post-processed via the previous counting method.

In the previous method, the filter output is first low-pass filtered and downsampled. It is downsampled with a factor of 1.2τ [3], the time constant of the exponential decay. It calculates the standard deviation (SD) of the measurement at the lowest blackbody temperature, after removing the cosmic rays. This way the SD is of only the noise. In order to count the photons, this signal is then thresholded at 4 to 5 standard deviations (SD) above 0, depending on the dataset. The Figures in shown in this Chapter are all with thresholds of 4.5 SD. This counting method was applied both to the previous filtering method (a bandpass filter followed by a matched filter), and to the newly presented generalized matched filter. Instead of comparing the photon rate, this Section will compare the time-streams of the previous filtering approach and the newly presented filter, after being low-pass filtered and downsampled, but before being thresholded. These time-streams will be referred to as the downsampled signals of the previous filtering method, and the downsampled signals of the new filtering method.

5.4.1 LT366, KID5

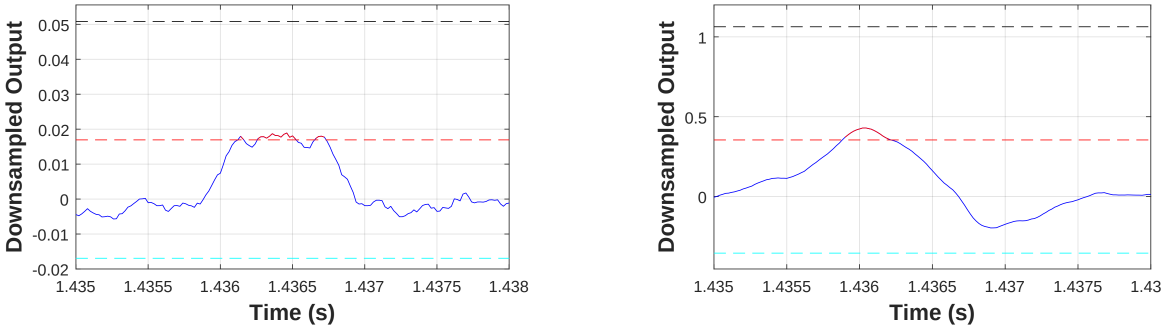
Figure 5.3 shows parts of a time-stream where multiple peaks are visible. The time-streams are comparable, except for extra high frequency noise for the previous filtering method. Figure 5.4 shows the same data but zoomed in on one peak. In the output from the previous filtering method there are not one, but three peaks. These will all get counted in determining the photon rate. With the new filtering method this is reduced to just a single peak. It is not possible to say for sure that it was one photon and not three since that data does not exist for experimental data. However, the multiple threshold crossings appear to come from the high frequency noise and not actually from three seperate photons.



(a) Previous

(b) New

Figure 5.3: Comparison of parts of a time-stream showing multiple peaks in the downsampled output. Dashed lines show thresholds, red +4.5 SD, light blue -4.5 SD, black 13.5 SD for cosmic rays. Dataset: LT366, KID5, 18.0 K.



(a) Previous

(b) New

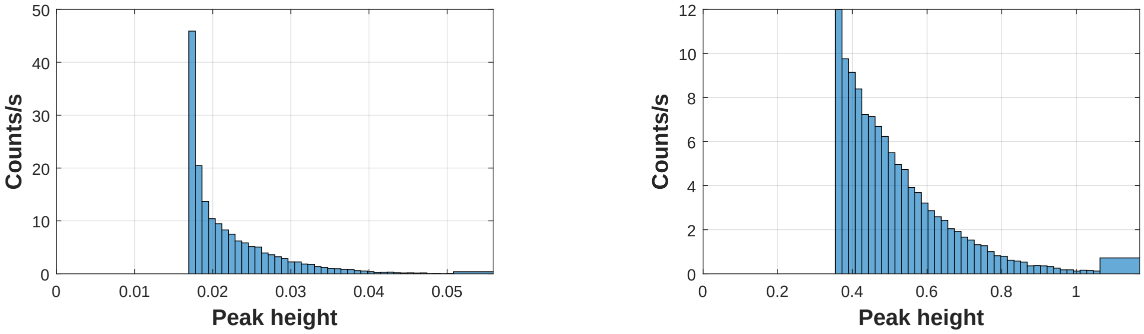
Figure 5.4: Comparison of zoom in on time-stream of a single peak. Dashed lines show thresholds, red +4.5 SD, light blue -4.5 SD, black 13.5 SD for cosmic rays. Dataset: LT366, KID5, 18.0 K.

The output from both filtering methods show a low frequency component coming through. This is present at both high blackbody temperatures and at low blackbody temperatures. It is also present in the non-downsampled filter output. It appears to be at a frequency of 244 Hz.

Figure 5.5 shows a histogram of shows a histogram of a time-stream of downsampled data after thresholding at 4.5 SD. The underlying distribution is not visible, although it looks like it could be the right tail of a Guassian distribution.

5.4.2 LT218, KID9

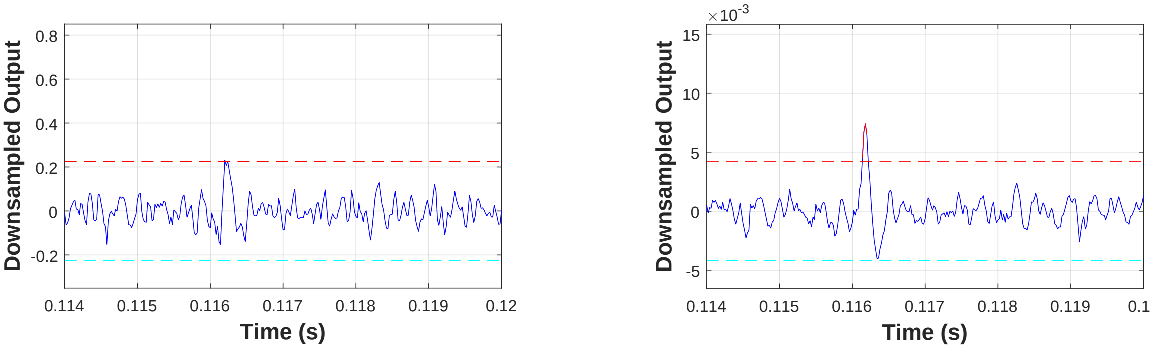
Figure 5.6 shows a time-stream comparison of a photon pulse that is missed with the previous filtering method, but is picked out with the new filtering method. Figure 5.7 shows a longer time-stream with multiple peaks. The peaks from the new filering method are higher and there are more peaks. Figure 5.8 shows a histogram of the same data after thresholding at 4.5 SD. The histograms show two bell curves. One is from the noise and one is from the photon pulses. The two bell curves are more separated



(a) Previous

(b) New

Figure 5.5: Comparison of histograms of the thresholded and downsampled data. Only samples above 4.5 SD are counted. Dataset: LT366, KID5, 18.5 K.

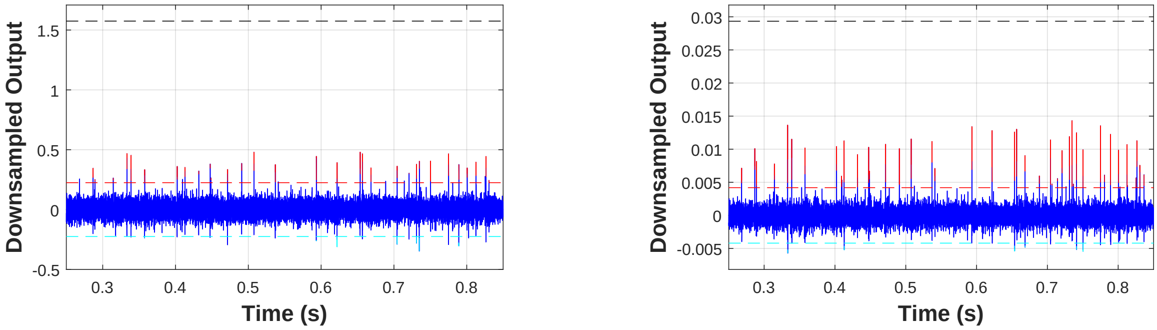


(a) Previous

(b) New

Figure 5.6: Comparison of time-stream where a peak is missed with the previous method. Dashed lines show thresholds, red +4.5 SD, light blue -4.5 SD, black 13.5 SD for cosmic rays. Dataset: LT218, KID9, 23.5 K.

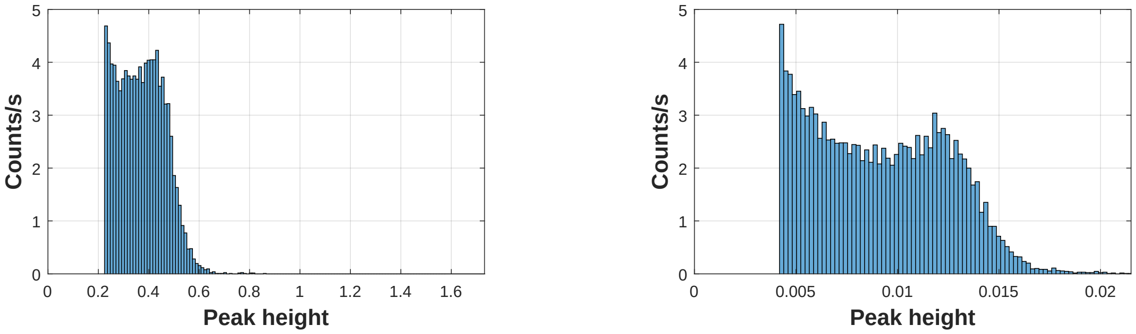
for the new filtering method compared to the previous filtering method.



(a) Previous

(b) New

Figure 5.7: Comparison of part of a time-stream showing a difference in peak height and number of peaks. Dashed lines show thresholds, red +4.5 SD, light blue -4.5 SD, black 13.5 SD for cosmic rays. Dataset: LT218, KID9, 25.0 K.



(a) Previous (b) New

Figure 5.8: Comparison of histograms of the thresholded and downsampled data. Only samples above 4.5 SD are counted. Dataset: LT218, KID9, 25.5 K.

Discussion

6

6.1 System Model

The goal stated in the introduction was to define the problem of optimally estimating the photon rate incident on an MKID in the framework of statistical signal processing. This led to a system model being made in Chapter 3 based on the description of the physical processes in MKIDs in Chapter 2. The model was verified by succesfully synthesizing data that matches the output of actual MKIDs.

The use of the system model is directly showcased by using it to derive an estimator for the photon rate via applying a generalized matched filter. The output of the matched filter is used to detect the photon hits and from that estimate the photon rate. This method is similar to what has been tried before in [1, 3].

The derivation of the generalized matched filter shows that in order to use a type of matched filter solution, certain assumptions have to be made about the model. Simply applying a matched filter on the measured phase coordinate on a normalized KID circle yields results, but nothing can be said as to why that would be a good idea. A form of generalized matched filter is probably the best way forward, but it is only after simplifying the model and assuming the variance of the amplifier noise stays constant that it comes about as a type optimal solution.

Assuming the slightly more complex model, the one that results from Chapter 3, will probably yield better results. No closed-form expression for the PDF of the measured data parametrized on the photon rate was found. However, an expression could possibly be found via numerical methods. If this expression is found, it might lead to better estimators, as well as bounds on the performance of these estimators.

One other method that was not tried in finding an expression for the PDF was assuming a Bernoulli distribution instead of a Poisson distribution for $k[n]$ (the number of photons arriving at sample n). This is a valid assumption since the photon rate is low enough that it is highly unlikely for more than one photon to arrive during one sample. Besides, two photons arriving at the same time would probably lead to the sample being labeled as a cosmic ray, and therefore being discarded.

The system models of Chapter 3 can also be used in other estimation projects. For example, take the goal to be the estimation of the energy of an incoming photon, such as in the work of Ras [1]. In this case the photon model becomes $\phi(t) = \sum_i E_{\phi,i}\delta(t-t_i)$ and the goal should be to estimate each $E_{\phi,i}$. Or if the number of quasiparticles is the parameter of interest, the goal should be to estimate $\chi(t)$. Or it could be useful in the case of de Rooij [31]. In this work the single photon response is modeled in time including non-linearities. When formulating such new usecases for the system model, it is important to specify what is the parameter of interest vs. what are nuisance parameters.

6.2 Generalized Matched Filter Output

One method of estimating the photon rate was implemented. This was via the generalized matched filter.The estimated photon rate can not be compared properly because of a missing algorithm to do the actual photon counting. Therefore the previous methods and the presented methods are compared via a qualitative analysis. The time-streams seem to contain less errors with the new filering method compared to the previous method. Examples include:

- A missed photon that was now detected;
- a single photon that was first detected as being three photons was now detected as a single photon.

Furthermore, the histograms of the peak height show more separation between the underlying distributions. The shape of the distribution was not very different for KID5 of LT366, but for KID9 of LT218 two seperate bell curves can be recognized. This implies that there are two different Gaussian distribution at play, one where there is only noise and one where there is noise plus a photon pulse. This matches detection theory from [24] such that a threshold can be set based on a binary hypothesis problem. In this view the two underlying bell curves should have as much separation as possible to maximize the performance of the detector. The histograms of KID9 show more separation using the new filtering method vs. the previous filtering method and thus marks a preference for the new filtering method over the previous methods.

6.3 Deconvolution and Counting Algorithm

A counting algorithm is still needed to determine the photon rate from the filter output. A promising candidate for this is the CLEAN algorithm, and developing such an algorithm would be a good next step for further research. A clean algorithm is similar to what is already being done, but many examples exist already in use today. In a CLEAN algorithm the photon pulses are iteratively removed from the filter output until it only looks like noise. The found photon pulses are recorded during the process and can then be counted afterwards to determine the photon rate.

CLEAN can be seen as a type a type of deconvolution algorithm. It is a useful solution because it can be combined with the generalized matched filter that has already been implemented. Other types of deconvolution should also be considered. In that case the problem should be reformulated as finding $k[n]$ from the measurement samples. One example could be a minimization of the sparsity of $k[n]$. Often this is achieved via an ℓ_1-norm minimization. Another possibility is assuming the $k[n]$ are independent Gaussian random variables with mean 0 or 1. Deconvolution might then lead to a Wiener filter or an alternating least squares solution.

6.4 Coordinate System

The coordinate system of Section 2.4 came forth as a byproduct of building the system model. Going through the equations and deriving the signals in the system it became clear that the quasiparticle density was affinely related to a certain undefined quantity. This quantity was then used in the relation of the S_{21} output of the detector. Defining this quantity as the variable $r = \zeta_i + j\xi$ then led to the coordinate system of (ζ_i, ξ), and simultaneously a function to map between $(\mathfrak{Re}\{S_{21}\}, \mathfrak{Im}\{S_{21}\})$ and (ζ_i, ξ). The new coordinate system helped in finding the relation of the quasiparticle density and the measurement signal. It is therefore of great use in estimating the quasiparticle density as a signal by identifying how noise can be removed from the system.

In the Terahertz Sensing Group there are ideas going around of what is a useful coordinate system to represent the measured values in. Multiple people have independently come to a coordinate system similar to this one. This work adds to that by offering a new usecase, exemplifying its utility.

One limitation of the coordinate system is that it needs a good estimate of the ζ_c and ζ_{lim} quantities, or alternatively of Q_c and Q_{lim}. This ties in with analysis of asymmetric KID dips. Fitting asymmetric dips leads to estimates of these quantities. So a better fit then also leads to less error when mapping to the new coordinate system. In a sense, the coordinate system can be seen as following from a simplification of an asymmetric dip model.

Improving understanding of asymmetric KID dips will thus also lead to a better understanding of how a time-series measurement can be used to detect photon hits.

6.5 Related Works

[32–34] report photon pulse processing for microcalorimeters in the x-ray and gamma-ray range for high photon rates. The photon pulses have a similar shape to the one discussed here, an exponential decay, and the methods employed in these articles is also similar. The articles do not mention a system model such as the one mentioned here, although they do mention signal models. The filter they first arrive at is an "orthogonal filter" to find multiple overlapping pulses [32]; then a similar filter with noise whitening [33]; and after that they report the result of filtering pulses when taking into account a pulse height model where the pulse height is not linearly dependent on the photon energy [34].

They call the filter they arrive at an "orthogonal" or "tangent" filter, because the convolution shape is orthogonal to a preceding pulse in the case of overlapping pulses. Although the filter is not exactly equivalent to a matched filter, its function is quite similar. It produces a symmetric peak that seperates the noise from the pulses. The resulting peak height is used to estimate the energy of the incoming photons.

The works of these articles resemble the work done here in more ways. In [32] a similar approximation from a continuous-time delta train to a signal $k[n]$ is used. Furthermore, in [33] a similar problem is encountered where for noise whitening a large autocovariance matrix needs to be converted. The problem is solved there by modeling the noise as an auto-regressive moving-average (ARMA) and making use of the structure of the autocovariance matrix of such a process.

The techniques used in this thesis and the ones in these articles can be compared as future work. An analysis of the similarities and differences of the filters should help refine the signal processing further.

Conclusion

7

The goal of this thesis was to define the problem of optimally estimating the photon rate incident on an MKID within the framework of statistical signal processing. This was achieved through the development of a system model that represents the behavior of the system in response to photon hits. In the initial, canonical system model, the physical quantities in the detector are represented as continuous-time signals. The model was then simplified via discretization and linearization. This model represents the detector without noise. Since the noise is an essential obstacle in estimating the photon rate, the stochastic nature of the noise is analyzed and added to the appropriate signals.

The system model developed in this thesis represents the main contribution to the field. While the foundational knowledge for the model already existed, this work consolidates that knowledge and restructures it so that it is more accesible for people who have a good understanding of signal processing, but have no background in the field of MKIDs. The model should simplify further signal processing and lead to a better understanding of the application of optimal filers. Additionally, the model's utility extends beyond this specific problem; it provides a structured framework for other estimation projects which helps quantify what needs to be estimated and how noise affects those estimations.

However, the model is not a definitive solution. Further refinement is necessary to make it more comprehensible to both signal processing experts and those working in MKID design. A key step in this process is comparing this model with similar works to provide additional validation and insights. These insights could lead to an improvement in practical applications from MKID theory.

Another notable contribution of this thesis is the introduction of a novel coordinate system. The coordinate system was necessary to be able to define the signal model. Multiple version of the coordinate system already exist, but this application adds a new perspective to the body of works that use such similar coordinate systems. The definition of the coordinate system in terms that should be familiar in resonator physics highlights the validity of the coordinate system.

Additionally, a function is provided that maps between S_{21} measurements and the newly introduced coordinates. The function also includes a set of parameters that need to be found to accurately map between the two coordinate systems.

The system model is then used to try and find an estimator for the photon rate. This represents a continuation of the thesis goal, as deriving a system model is only a one step towards designing effective estimation algorithms. The ultimate objective, beyond the scope of this thesis, is dependent on the application of the required estimation algorithm. One example would be in an on-satellite application, where the signal processing is required to be performed before sending information down towards earth. Only transmitting the found photon rate instead of entire streams of data would greatly

reduce the amount of transmission bandwidth.
However, before reaching such advanced applications, more off-line algorithms and signal processing applications need to be developed and refined. If these prove to be effective they can be adapted for more advanced purposes. This makes this research an early step in laying down the foundation for future signal processing.

The utility of the system model is exemplified by deriving a generalized matched filter. This resulted from a necessary simplification, as the full system model was too complex to yield an optimal solution directly. This does not mean that the optimal solution can not be found. Future research can possibly be focussed on finding an actual optimal solution, although for now the generalized matched filter implementation is the best option available.
Unfortunately, the matched filter implementation could not be compared quantitively to older methods. Instead, the two are compared qualitatively, because the methods are quite similar. The qualitative analysis shows that the method likely contains less errors. Nonetheless, further assessment is necessary to strengthen this conclusion.

This research leaves several open questions for future work. There are still many avenues to be explored, but deciding which ones to take ultimately depends on the expected performance gains and the time it will take to achieve them. From this perspective, refining the generalized matched filter implementation and developing a CLEAN algorithm for photon counting would provide the most immediate benefits. However, there is also a deeper understanding to be gained from further development of the theoretical model and discussions of the optimality of estimators. For this deeper understanding, other approaches should be explored such as deconvolution algorithms. Additionally, an expression for an optimal estimator should be explored, one that is based on the full system model rather than the simplified representation in r-space. These results should then be compared with other works on photon pulse processing to gain a broader perspective and verify the approaches.

In conclusion, this thesis lays the groundwork for future research in photon rate estimation for MKIDs. While the system model and initial filtering methods provide valuable results, further refinement, validation, and exploration of other solutions will be essential for further insights and potentially for optimal estimators.

Bibliography

[1] W. Ras, “Microwave Kinetic Inductance Detectors For The Mid-Infrared,” 2022. Accessed: May 29, 2024. [Online]. Available: https://repository.tudelft.nl/islandora/object/uuid%3A8cf9cabb-7348-4eb6-b93e-7dde0b143091.

[2] J. J. A. Baselmans *et al.*, “Ultra-sensitive THz microwave kinetic inductance detectors for future space telescopes,” *Astronomy & Astrophysics*, vol. 665, A17, Sep. 1, 2022, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202243840. Accessed: May 27, 2024. [Online]. Available: https://www.aanda.org/articles/aa/abs/2022/09/aa43840-22/aa43840-22.html.

[3] F. Facchin, “Single THz Photon Detection with Microwave Kinetic Inductance Resonators,” M.S. thesis, Delft University of Technology, Delft, Aug. 2020.

[4] B. A. Mazin, P. K. Day, J. Zmuidzinas, and H. G. Leduc, “Multiplexable kinetic inductance detectors,” *AIP Conference Proceedings*, vol. 605, no. 1, pp. 309–312, Feb. 5, 2002, ISSN: 0094-243X. DOI: 10.1063/1.1457652. Accessed: Mar. 4, 2025. [Online]. Available: https://doi.org/10.1063/1.1457652.

[5] B. A. Mazin, P. K. Day, H. G. LeDuc, A. Vayonakis, and J. Zmuidzinas, “Superconducting kinetic inductance photon detectors,” in *Highly Innovative Space Telescope Concepts*, vol. 4849, SPIE, Dec. 18, 2002, pp. 283–293. DOI: 10.1117/12.460456. Accessed: Mar. 4, 2025. [Online]. Available: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/4849/0000/Superconducting-kinetic-inductance-photon-detectors/10.1117/12.460456.full.

[6] P. K. Day, H. G. LeDuc, B. A. Mazin, A. Vayonakis, and J. Zmuidzinas, “A broadband superconducting detector suitable for use in large arrays,” *Nature*, vol. 425, no. 6960, pp. 817–821, Oct. 2003, ISSN: 1476-4687. DOI: 10.1038/nature02037. Accessed: May 22, 2024. [Online]. Available: https://www.nature.com/articles/nature02037.

[7] B. A. Mazin, “Microwave Kinetic Inductance Detectors,” Ph.D. dissertation, California Institute of Technology, 2005. DOI: 10.7907/GZ72-V784. Accessed: May 22, 2024. [Online]. Available: https://resolver.caltech.edu/CaltechETD:etd-10042004-120707.

[8] J. J. A. Baselmans, “EE4635 Superconducting Astronomical Instrumentation, Lecture 7: Multiplexing and Microwave Kinetic Inductance Detectors,” 2023.

[9] B. A. Mazin, B. Bumble, S. R. Meeker, K. O’Brien, S. McHugh, and E. Langman, “A superconducting focal plane array for ultraviolet, optical, and near-infrared astrophysics,” *Optics Express*, vol. 20, no. 2, pp. 1503–1511, Jan. 16, 2012, ISSN: 1094-4087. DOI: 10.1364/OE.20.001503. Accessed: Mar. 8, 2025. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?uri=oe-20-2-1503.

[10] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, “Theory of Superconductivity,” *Physical Review*, vol. 108, no. 5, pp. 1175–1204, Dec. 1, 1957. DOI: 10.1103/PhysRev.108.1175. Accessed: Mar. 3, 2025. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.108.1175.

[11] D. C. Mattis and J. Bardeen, “Theory of the Anomalous Skin Effect in Normal and Superconducting Metals,” *Physical Review*, vol. 111, no. 2, pp. 412–417, Jul. 15, 1958. DOI: 10.1103/PhysRev.111.412. Accessed: May 29, 2024. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.111.412.

[12] S. Sarangi, S. P. Chockalingam, R. G. Mavinkurve, and S. V. Bhat. "Experimental Evidence for Zero DC Resistance of Superconductors." arXiv: cond-mat/0506426, Accessed: Mar. 3, 2025. [Online]. Available: http://arxiv.org/abs/cond-mat/0506426, pre-published.

[13] P. J. De Visser, "Quasiparticle dynamics in aluminium superconducting microwave resonators," 2014. Accessed: May 21, 2024. [Online]. Available: https://repository.tudelft.nl/islandora/object/uuid%3Aeae4c9fc-f90d-4c12-a878-8428ee4adb4c.

[14] D. M. Pozar, *Microwave Engineering*, Fourth edition. Hoboken, NJ: John Wiley & Sons, Inc, 2012, 1 p., ISBN: 978-0-470-63155-3 978-1-118-21363-6.

[15] K. Kouwenhoven *et al.*, "Resolving power of visible-to-near-infrared hybrid $\beta-$Ta/Nb$-$Ti$-$N kinetic inductance detectors," *Physical Review Applied*, vol. 19, no. 3, p. 034 007, Mar. 2, 2023. DOI: 10.1103/PhysRevApplied.19.034007. Accessed: Feb. 28, 2025. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevApplied.19.034007.

[16] N. Zobrist *et al.*, "Improving the dynamic range of single photon counting kinetic inductance detectors," *Journal of Astronomical Telescopes, Instruments, and Systems*, vol. 7, no. 1, p. 010 501, Feb. 2021, ISSN: 2329-4124, 2329-4221. DOI: 10.1117/1.JATIS.7.1.010501. Accessed: Jun. 12, 2024. [Online]. Available: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes-Instruments-and-Systems/volume-7/issue-1/010501/Improving-the-dynamic-range-of-single-photon-counting-kinetic-inductance/10.1117/1.JATIS.7.1.010501.full.

[17] M. S. Khalil, M. J. A. Stoutimore, F. C. Wellstood, and K. D. Osborn, "An analysis method for asymmetric resonator transmission applied to superconducting devices," *Journal of Applied Physics*, vol. 111, no. 5, p. 054 510, Mar. 8, 2012, ISSN: 0021-8979. DOI: 10.1063/1.3692073. Accessed: Jun. 14, 2024. [Online]. Available: https://doi.org/10.1063/1.3692073.

[18] C. Deng, M. Otto, and A. Lupascu, "An analysis method for transmission measurements of superconducting resonators with applications to quantum-regime dielectric-loss measurements," *Journal of Applied Physics*, vol. 114, no. 5, p. 054 504, Aug. 6, 2013, ISSN: 0021-8979. DOI: 10.1063/1.4817512. Accessed: Feb. 22, 2025. [Online]. Available: https://doi.org/10.1063/1.4817512.

[19] S. Probst, F. B. Song, P. A. Bushev, A. V. Ustinov, and M. Weides, "Efficient and robust analysis of complex scattering data under noise in microwave resonators," *Review of Scientific Instruments*, vol. 86, no. 2, p. 024 706, Feb. 17, 2015, ISSN: 0034-6748. DOI: 10.1063/1.4907935. Accessed: Feb. 22, 2025. [Online]. Available: https://doi.org/10.1063/1.4907935.

[20] D. Rieger, S. Günzler, M. Spiecker, A. Nambisan, W. Wernsdorfer, and I. Pop, "Fano Interference in Microwave Resonator Measurements," *Physical Review Applied*, vol. 20, no. 1, p. 014 059, Jul. 26, 2023. DOI: 10.1103/PhysRevApplied.20.014059. Accessed: Feb. 22, 2025. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevApplied.20.014059.

[21] J. Gao, J. Zmuidzinas, B. A. Mazin, H. G. LeDuc, and P. K. Day, "Noise properties of superconducting coplanar waveguide microwave resonators," *Applied Physics Letters*, vol. 90, no. 10, p. 102 507, Mar. 5, 2007, ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.2711770. Accessed: May 21, 2024. [Online]. Available: https://pubs.aip.org/

apl/article/90/10/102507/332953/Noise-properties-of-superconducting-coplanar.

[22] M. H. Hayes, *Statistical Digital Signal Processing and Modeling.* New York: John Wiley & Sons, 1996, 608 pp., ISBN: 978-0-471-59431-4.

[23] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory* (Prentice Hall Signal Processing Series). Upper Saddle River, NJ: Prentice Hall PTR, 1993, 595 pp., ISBN: 978-0-13-345711-7.

[24] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory.* Upper Saddle River, NJ: Prentice Hall PTR, 1998, 560 pp., ISBN: 978-0-13-504135-2.

[25] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume III: Practical Algorithm Development.* Upper Saddle River, NJ: Pearson Education, Inc., 2013, 1 p., ISBN: 978-0-13-280803-3 978-0-13-280805-7.

[26] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4. ed., internat. ed., Nachdr. Boston, Mass.: The McGraw-Hill Companies, Inc, 2002, 852 pp., ISBN: 978-0-07-112256-6 978-0-07-122661-5.

[27] A.-J. Van Der Veen and S. J. Wijnholds, "Signal Processing Tools for Radio Astronomy," in *Handbook of Signal Processing Systems*, S. S. Bhattacharyya, E. F. Deprettere, R. Leupers, and J. Takala, Eds., New York, NY: Springer New York, May 2013, pp. 421–463, ISBN: 978-1-4614-6858-5 978-1-4614-6859-2. DOI: 10.1007/978-1-4614-6859-2_14. Accessed: Feb. 14, 2025. [Online]. Available: http://link.springer.com/10.1007/978-1-4614-6859-2_14.

[28] A. V. Oppenheim and R. W. Schafer, Eds., *Discrete-Time Signal Processing* (Prentice Hall Signal Processing Series), 3. ed. Upper Saddle River, NJ: Pearson Higher Education, Inc., 2010, 1108 pp., ISBN: 978-0-13-198842-2.

[29] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 3rd ed. Upper Saddle River, New Jersey: Prentice-Hall, Inc., 1996, ISBN: 0133943389.

[30] A. Kessy, A. Lewin, and K. Strimmer, "Optimal Whitening and Decorrelation," *The American Statistician*, vol. 72, no. 4, pp. 309–314, Oct. 2, 2018, ISSN: 0003-1305. DOI: 10.1080/00031305.2016.1277159. Accessed: Jul. 12, 2024. [Online]. Available: https://doi.org/10.1080/00031305.2016.1277159.

[31] S. A. de Rooij, "Quasiparticle Dynamics in Optical MKIDs: Single Photon Response and Temperature Dependent Generation-Recombination Noise," M.S. thesis, Delft University of Technology, Delft, Jun. 2020. [Online]. Available: http://resolver.tudelft.nl/uuid:1d935dc8-58d1-438f-995b-872e9d5f077e.

[32] B. K. Alpert *et al.*, "Note: Operation of gamma-ray microcalorimeters at elevated count rates using filters with constraints," *Review of Scientific Instruments*, vol. 84, no. 5, p. 056 107, May 15, 2013, ISSN: 0034-6748. DOI: 10.1063/1.4806802. Accessed: Oct. 14, 2024. [Online]. Available: https://doi.org/10.1063/1.4806802.

[33] J. W. Fowler *et al.*, "MICROCALORIMETER SPECTROSCOPY AT HIGH PULSE RATES: A MULTI-PULSE FITTING TECHNIQUE," *The Astrophysical Journal Supplement Series*, vol. 219, no. 2, p. 35, Aug. 2015, ISSN: 0067-0049. DOI: 10.1088/0067-0049/219/2/35. Accessed: Oct. 14, 2024. [Online]. Available: https://dx.doi.org/10.1088/0067-0049/219/2/35.

[34] J. W. Fowler *et al.*, "When "Optimal Filtering" Isn't," *IEEE Transactions on Applied Superconductivity*, vol. 27, no. 4, pp. 1–4, Jun. 2017, ISSN: 1558-2515. DOI: 10.

1109/TASC.2016.2637359. Accessed: Oct. 14, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7778170.