Research and Development of a Hybrid Photo Sensor for a Water Cerenkov Detector

KUSAKA Akito
Depertment of Physics, University of Tokyo

February 5, 2004

Abstract

A new photo sensor, HPD(Hybrid Photo-Detector), has been developed for next generation water Cerenkov detectors. We describe the performance of a 5 inch-diameter prototype. A new readout system for HPD has also been developed, including, 1) a preamplifier with wide bandwidth and low noise, 2) Analog Memory Cell (AMC) ASIC as a waveform sampling device, and 3) techniques of digital signal processing (DSP).

We have two configurations of the DSP. One of them is optimized for the signal to noise ratio and the timing resolution of the photo sensor system. With the 5 inch prototype HPD and a prototype of the readout system using the S/N optimized DSP, we have achieved the following signal to noise ratio, S/N, and timing resolution, σ_t, for single photoelectron input:

$$S/N = 21, \quad \sigma_t = 0.4[\mathrm{ns}] \, .$$

Another DSP configuration is optimized to achieve short dead time, with tradeoffs: smaller signal to noise ratio and worse timing resolution. With the prototype HPD and the prototype readout system using the dead time optimized DSP, we have achieved the following performances for single photoelectron input:

$$S/N = 6.4, \quad \sigma_t = 1.1[\mathrm{ns}], \quad T_{\mathrm{dead}} = 7.5[\mathrm{ns}] \, ,$$

where T_{dead} is the dead time of the whole sensor system including the readout system.

The development of 13 inch HPD is in progress.

Contents

1 Introduction **3**
- 1.1 What is a neutrino? . . . 3
- 1.2 The neutrino mass and the neutrino oscillation . . . 3

2 The motivation and the goal of our R&D **5**
- 2.1 J-PARCν experiment . . . 5
  - 2.1.1 The physics goal . . . 5
  - 2.1.2 The features of the experiment . . . 5
- 2.2 The Water Cerenkov Detector as a near detector of JHFν . . . 7
  - 2.2.1 The ring-imaging water Cerenkov detector . . . 7
  - 2.2.2 The role of near detectors . . . 7
  - 2.2.3 The dominant background in the search of $\nu_\mu \to \nu_e$ appearance . . . 8
  - 2.2.4 The features of the Water Cerenkov Detector in the near site . . . 8
  - 2.2.5 The requirements to the photo sensor system . . . 9
- 2.3 For the future... (Hyper Kamiokande experiment) . . . 10

3 The R&D of HPD(Hybrid Photo-Detector) **11**
- 3.1 Overview of HPD . . . 11
- 3.2 The R&D of HPD . . . 12
  - 3.2.1 The evaluation of 5 inch prototype HPD . . . 12
  - 3.2.2 The R&D status of 13 inch prototype HPD . . . 18

4 The R&D of readout system **21**
- 4.1 Design concept of readout system . . . 21
  - 4.1.1 The requirement to the readout system . . . 21
  - 4.1.2 The design concept of the readout system . . . 21
- 4.2 Preamplifier . . . 22
  - 4.2.1 Conceptual design . . . 22
  - 4.2.2 Circuit simulation . . . 23
  - 4.2.3 Selection of the input transistor . . . 32
  - 4.2.4 Performance evaluation . . . 40
- 4.3 Waveform sampler . . . 46
  - 4.3.1 Principle of AMC . . . 46
  - 4.3.2 Comparison of FADC and AMC . . . 47
  - 4.3.3 The R&D of AMC . . . 47
- 4.4 Digital signal processing . . . 52
  - 4.4.1 Design concept . . . 52
  - 4.4.2 Assumed signal and noise . . . 58
  - 4.4.3 Filtering algorithms . . . 58
  - 4.4.4 Processing after the filtering . . . 60
  - 4.4.5 Performances . . . 61

5 The connection test of HPD and readout **66**
5.1 Setup . . . 66
5.2 Data before analisys . . . 66
5.3 Performance evaluation . . . 71
5.4 The measurement of 13inch HPD prototype . . . 76

6 Summary and future prospects **78**
6.1 Summary . . . 78
6.2 Future prospects . . . 78

A The calculations on the electronics **80**
A.1 The noise of electronics . . . 80
A.1.1 The noise of the bipolar junction transistor (BJT) . . . 80
A.1.2 The propagation of the noise in a negative feedbacked amplifier . . . 81
A.2 Estimation of base spreading resistance using S parameter . . . 83

B The algorithms for the digital filter **85**
B.1 The algorithms for the slow filter . . . 85
B.1.1 Matched filter . . . 85
B.1.2 Optimal matched filter . . . 87
B.1.3 On timing resolution . . . 87
B.2 The algorithms for the fast filter . . . 89
B.2.1 Deconvolution filter . . . 89
B.2.2 Wiener optimal filter . . . 89
B.2.3 On the normalization of the Wiener optimal filter . . . 91

C Avalanche Diode **94**
C.1 Avalanche multipllication . . . 94
C.2 Avalanche Diode . . . 96
C.2.1 Operation in proportional mode . . . 96
C.2.2 Operation in Geiger mode . . . 96

D The history of water Cerenkov detectors and their photo sensors **98**
D.1 IMB . . . 98
D.2 KamiokaNDE . . . 98
D.3 Super Kamiokande . . . 98
D.4 1 Kt detector in K2K . . . 99
D.5 SNO . . . 99

Chapter 1

Introduction

1.1 What is a neutrino?

Neutrinos are the fundamental particles, as is the electron, which make up the Universe. They are one of the most abundant particles in the universe. Because they have very little interaction with matter, however, they are incredibly difficult to detect. Consequently, they are also one of the least understood.

Neutrinos are similar to the more familiar electrons, with one crucial difference: neutrinos do not carry electric charge, which fact is the primary reason that they are difficult to detect. The nuclear forces treat electrons and neutrinos identically; neither participates in the strong nuclear force, but both participate equally in the weak nuclear force. Since neutrinos are proved to have masses, they also interact gravitationally with other massive particles.

Three types, or "flavors," of neutrinos are known, and there is strong evidence that no additional types exist, unless their properties are unexpectedly different from the known types. Each flavor of neutrino is related to a charged particle; electron-neutrino is associated to electron, mu-neutrino is to muon, and tau-neutrino is to tauon. "Lepton" is the name of the class of particles that consists of these six particles, and the three pairs of "leptons" are called "generations." There is another class of particles, "quark," that also have the three generations. "Lepton" and "quark" make up the Universe, with the "gauge boson," which is the class of the particles that mediate forces. (Table 1.1)

Table 1.1: The elementary particles that make up the Universe.

Generation	1	2	3
Leptons	e (electron)	μ (muon)	τ (tauon)
	ν_{e} (electron-neutrino)	ν_μ (mu-neutrino)	ν_τ (tau-neutrino)
Quarks	u (up)	c (charm)	t (top)
	d (down)	s (strange)	b (bottom)

Gauge bosons	g (Gluon)	$W^\pm$, Z^0 (Weak bosons)	γ (Photon)	G(?) (Graviton)

1.2 The neutrino mass and the neutrino oscillation

The question, whether the neutrinos have masses or not, had been opened for a long time since their discovery, 1956. There had been no clear evidence of the neutrino mass, until 1998, the discovery of the neutrino oscillation by the Super-Kamiokande experiment.[1]

The neutrino oscillation is the phenomenon that neutrinos change their flavor, e.g., ν_μ transforms into $\nu_{\rm e}$. The phenomenon can occur only when at least one of the neutrinos has non-zero mass, as described in the following.

Assuming the neutrinos have masses, their mass eigenstates, ν_1, ν_2 and ν_3, are written using their flavor eigenstates, $\nu_{\rm e}$, ν_μ and ν_τ, as

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} . \tag{1.1}$$

The matrix $\mathbf{U}$ is an unitary matrix called “NMS matrix.” The Shrödinger equation, the equation of motion of the neutrinos, is as follows in free space.

$$i\frac{d}{dt}\nu_i(t) = H_i\ \nu_i(t) \quad (i = 1, 2, 3) \tag{1.2}$$

Note that the H_i is diagonal by the definition of ν_i. The solution of the equation is

$$\nu_i(t) = \nu_i e^{-iE_i t} \quad \left(E_i = \sqrt{p^2 + m_i^2}\right) , \tag{1.3}$$

where m_i and p are the mass and the momentum of ν_i, and $\nu_i(0) = \nu_i$ by definition. To be simple, here we consider only two neutrino flavors, ν_e and ν_μ and two mass eigenstates, ν_2 and ν_3. Then the equation 1.1 is written with a single mixing angle θ:

$$\begin{pmatrix} \nu_\mu \\ \nu\tau \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_2 \\ \nu_3 \end{pmatrix} . \tag{1.4}$$

Equations 1.3 and 1.4 give the time evolution of the wave function of a neutrino, $\nu(t)$, with initial condition of $\nu(0) = \nu_\mu$:

$$\begin{aligned} \nu(t) &= \cos\theta\nu_2(t) + \sin\theta\nu_3(t) & (1.5) \\ &= \left(\cos^2\theta e^{-iE_2 t} + \sin^2\theta e^{-iE_3 t}\right)\nu_\mu - \sin\theta\cos\theta\left(e^{-iE_2 t} - e^{-iE_3 t}\right)\nu_\tau\ . & (1.6) \end{aligned}$$

Thus, a neutrino produced in an electroweak eigenstate, ν_μ, will be observed at a time t in another electroweak eigenstate, ν_e, with the following possibility $P(\nu_\mu \to \nu_\tau)$.

$$P(\nu_\mu \to \nu_\tau) = \sin^2 2\theta\ \sin^2\left[\frac{(E_3 - E_2)t}{2}\right] , \tag{1.7}$$

Assuming ${m_i}^2 \ll p^2$, the equation gives the following equation:

$$P(\nu_\mu \to \nu_\tau) = \sin^2 2\theta\ \sin^2\left(\frac{\Delta m^2}{2E}t\right) , \tag{1.8}$$

where $\Delta m^2 \equiv {m_3}^2 - {m_2}^2$ and $E \equiv p \simeq E_2 \simeq E_3$. Equation 1.8 shows that the oscillation can occur only when at least one of the mass eigenstates ν_i is non-zero.

Chapter 2

The motivation and the goal of our R&D

2.1 J-PARCν experiment

In this section, we summarize the physics motivations and the features of J-PARCν experiment. A detailed description for J-PARCν experiment can be found elsewhere.[2, 3]

2.1.1 The physics goal

Studies of atmospheric and solar neutrinos have shown that neutrinos have masses and have large mixing.[1, 4, 5] Some experiments have confirmed the results using artificial neutrinos.[6, 7]

The existence of the mass and mixing makes the neutrino physics very interesting. The neutrino mass and mixing can be one of a few possible windows of physics at the GUT scale. One can search for the CP violation in the lepton sector by the comparison of neutrino and antineutrino oscillations. The CP violation in the lepton sector, with that in the quark sector, may lead us to the understanding of the cosmological baryon/anti-baryon asymmetry.

J-PARCν experiment approaches these issues with its high luminosity neutrino beam and Super-Kamiokande detector. The aim of the experiment is, roughly, as follows.

1. The precise measurement of ν_μ disappearance.
2. The first observation of $\nu_\mu \rightarrow \nu_e$ appearance.

The precise measurement of ν_μ disappearance will confirm our current understanding on the neutrino. The observation of ν_e appearance will lead us to the future measurement of the CP violation in the lepton sector.

2.1.2 The features of the experiment

Overview

J-PARCν experiment will use the high intensity proton beam from J-PARC 50GeV proton synchrotron (PS) at Tokai, and Super-Kamiokande, at Kamioka, as a far detector. The distance between Tokai and Kamioka is 295km. Figure 2.1 shows the schematic view of the experiment.

The neutrino beam

The proton beam is extracted from the 50GeV PS and transported to the production target. The design intensity of the PS is 3.3×10^{14} protons/pulse (ppp) at a repetition rate of 0.285Hz (3.5 second period). Table 2.1 shows some specifications of the proton beam of the J-PARCν experiment.[8, 9] We also show the specifications of the KEK PS beam,[8, 10] which is used in K2K experiment, in the table as a reference.

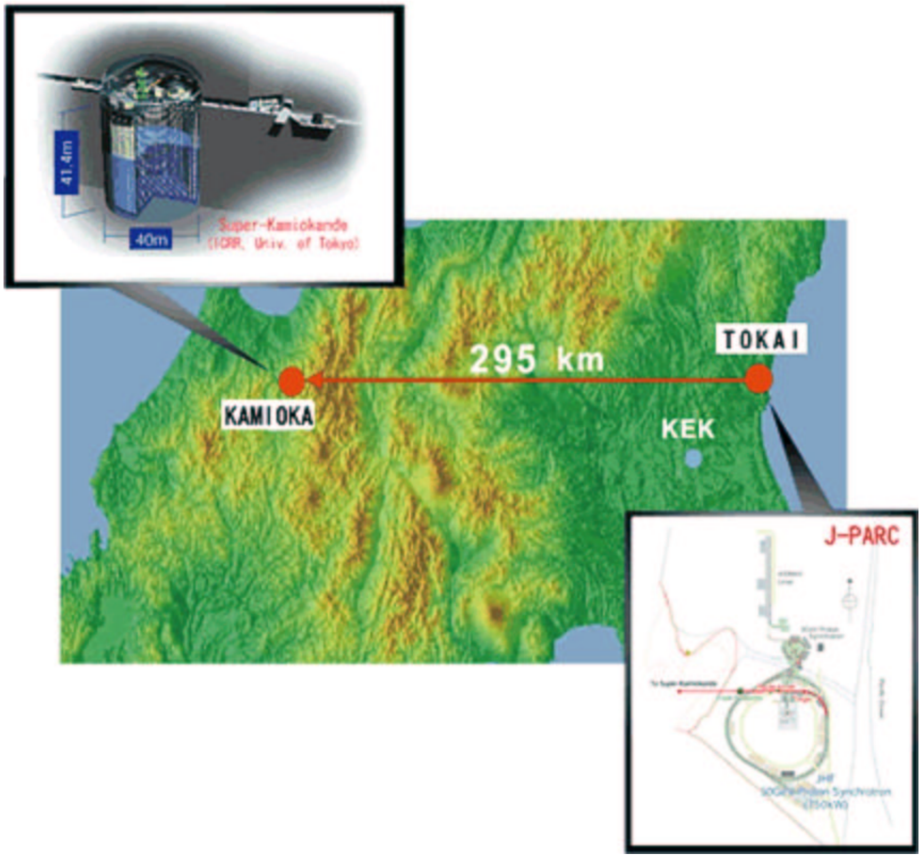


Figure 2.1: The schematic view of the J-PARCν experiment.

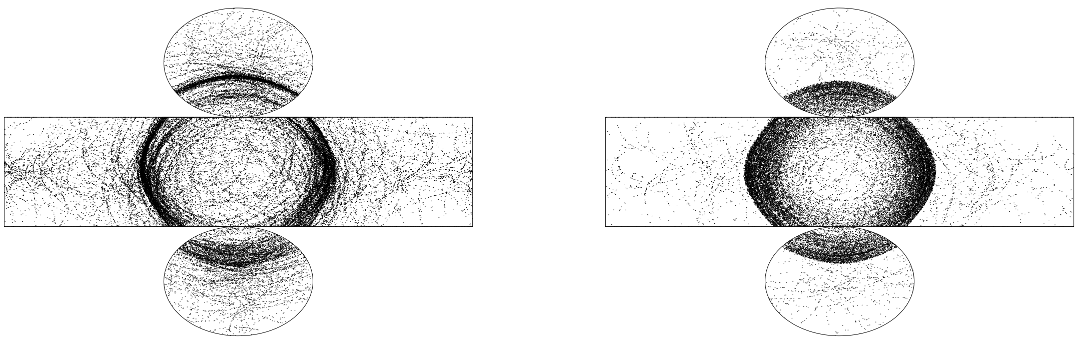
Table 2.1: The specifications of the proton beams for the J-PARCν experiment and the K2K experiment

	J-PARCν	K2K
Energy(GeV)	50	12
Intensity(ppp)	3.3×10^{14}	6×10^{12}
Power(kW)	750	5.2
Number of bunches / pulse(=spill)	8	9
Bunch width(ns)	58 (full width)	—
Bunch spacing(ns)	598	—
Spill width(μs)	~ 5	1.1
Cycle(sec)	3.53	2.2
Beam configuration	OAB 2°	WBB

2.2 The Water Cerenkov Detector as a near detector of JHFν

2.2.1 The ring-imaging water Cerenkov detector

The ring-imaging water Cerenkov detector detects charged particles with Cerenkov light produced by the particles. When a particle moves with a speed greater than the speed of light in water, it generates the Cerenkov light. The ring-imaging water Cerenkov detector consists of a water tank and photo sensors around the tank. When a charged particle goes through the water tank, it generates the cone of the Cerenkov light and the ring pattern is formed on the wall of the tank, where the photo sensors are equipped. The photo sensors detect the position and the timing of the light. The shape and timing of the ring pattern enable us to calculate the track position and the energy momentum vector of the particle. The detector can also discriminate the electron-like particle that makes an electromagnetic cascade shower from the muon-like particle that do not make the shower, using the difference of the Cerenkov ring pattern. (Figure 2.2)



(a) An electron-like event. (500MeV electron)

(b) A muon-like event. (500MeV muon)

Figure 2.2: The Cerenkov ring simulated with Geant4.

The merits of the water Cerenkov detector are its large interaction volume and its hermeticity. Due to the large interaction volume, it suits the detection of rare events such as the neutrino interactions and the proton decay. The hermeticity reduces the possibility to miss signals from a particle, and thereby enhances the efficiency of the event identification. The high efficiency is useful in various analysises, e.g., it highly suppresses backgrounds from misidentifications in the rare decay search.

2.2.2 The role of near detectors

The role of the near detectors is to provide predictions of the expected neutrino flux at the far detector (Super-K). The rough idea of the method is as follows. We measure the ν_μ flux in the near site, F_{near}, and in the far site, F_{far}. Then we define the solid angles of the detectors measured from the decay volume, where the pions generated from proton beam decay into neutrinos, as Ω_{near} and Ω_{far} for the near detector and the far detector, respectively. Using the solid angles and the neutrino fluxes, we calculate the decrease of neutrinos, c_d:

$$c_d = \frac{F_{\text{far}}}{\Omega_{\text{far}}} \cdot \frac{\Omega_{\text{near}}}{F_{\text{near}}} \,. \tag{2.1}$$

When there is no oscillation, c_d equals unity. If we replace F_{far} with the ν_e flux at the far site, c_d is the ratio that ν_μ is transformed into ν_e, and equals zero if there is no $\nu_\mu \to \nu_e$ appearance. This is the basic idea, though it is not so simple in practice.

Note that the near detectors also have to measure the background in the neutrino beam. First, the detectors have to measure the ν_e contamination in the neutrino beam. From the measured ν_e flux in the near site, we estimate the flux of the beam induced ν_e in the far site, in the same way as we estimate ν_μ flux in the far site. Second, the detectors have to measure the cross sections of the neutrino interactions that can be the backgrounds, e.g., π^0 yielding events for the ν_e search. Since the neutrino beam with the *J-PARCν spectrum* is not available in other experiments, the near detectors only have chance to measure the cross sections with high statistics.

2.2.3 The dominant background in the search of $\nu_\mu \to \nu_e$ appearance

In the search of ν_e appearance, the dominant backgrounds are the ν_e contamination of the ν_μ beam, and the particle miss-ID in the event reconstruction, mainly due to the neutral pions (π^0).[2] We have to measure, therefore, the flux of the beam induced ν_e and the rate of π^0 yielding events with the near detectors, as described above.

2.2.4 The features of the Water Cerenkov Detector in the near site

The merits

The merits of the water Cerenkov detector are as follows.

1. The target material is water, which is the same with that of the far detector, Super Kamiokande.

2. It has 4π sr hermetic sensors around the neutrino target, the water tank.

We can measure the neutrino flux without any uncertainty due to neutrino interaction models, due to the first item. The reason is as follows. When we use the different materials as the neutrino target in the near and the far detectors, we have to adopt some models to translate the flux with the target of the near detector to the flux with the far detector target. This translation can cause an uncertainty due to uncertainties of the models, since the cross section of the neutrino interaction has not been studied so closely. Using the same material, water, as the neutrino target in both the far and the near detector, we can avoid the uncertainty.

Due to the hermetic sensors, the second item, we can discriminate π^0 yielding events from electron-like neutrino events with high efficiency. The reason is as follows. The π^0 decays into two γ particles and we cannot discriminate a γ from an electron, both yielding an electromagnetic shower. Thus, the difference between the π^0 event and the electron-like neutrino event is the number of the electromagnetic showers, π^0 events yielding two showers and electron-like neutrino events yielding one shower. We misidentify a π^0 event as an electron-like neutrino event if we miss one of the two showers from π^0. The water Cerenkov detector, with 4π sr hermetic sensors around the target, has high detection efficiency of the electromagnetic showers, and thereby has good separation efficiency of the π^0 events and the electron-like neutrino events.

Thereby, the water Cerenkov detector in the near site has a potential to play an essential role in the ν_e appearance search of J-PARCν experiment.

The demerits

The demerit of a water Cerenkov detector is that it has *no segmentation*. That is, since a water Cerenkov detector consists of one water tank and photo sensors around it, sensors can detect photons from an event, independent of the vertex position of the event.[i] On the other hand, in highly segmented detectors, a sensor detects only particles that pass through the segment attached to the sensor.

The *no segmentation* can result in the following two demerits.

1. The low resolution of the neutrino interaction vertex.

2. The high occupancy under the high neutrino flux environment.

The first item, the low vertex resolution, is the source of the systematic error of the neutrino flux. For the water Cerenkov detector, the vertex resolution is determined by the reconstruction precision of Cerenkov rings, and the resolution is ∼30cm in the Super-Kamiokande experiment, for example.[11] Tn

[i] For large water Cerenkov detectors, like Super-Kamiokande, it is not always the case.

the K2K experiment, the dominant systematic error comes from the precision of the vertex reconstruction of the water Cerenkov detector in the near site.[12] For highly segmented detectors, on the other hand, the vertex resolution is determined by the size of the segmentation, and we can obtain good enough resolution by making the size small.

The second item, the high occupancy, can result in more serious problem. The definition of the occupancy P_{occ} is

$$P_{\text{occ}} \equiv r_{\text{seg}} \cdot \tau_{\text{dead}} = \frac{r_{\text{total}} \cdot \tau_{\text{dead}}}{n_{\text{seg}}} , \tag{2.2}$$

where r_{seg}, r_{total}, τ_{dead}, and n_{seg} correspond to the event rate of each segment, the event rate of the whole detector, the dead time of each segment, and the number of the segmentation, respectively. The occupancy P_{occ} must be small. One can make n_{seg} large, keeping r_{total} constant, for the highly segmented detectors. For a water Cerenkov detector, on the other hand, since $n_{\text{seg}}{\sim}1$ effectively[ii] and we cannot change the n_{seg}, we have to make τ_{dead} very small to make the water Cerenkov detector work properly with the high luminosity neutrino beam.

2.2.5 The requirements to the photo sensor system

The requirements to realize the water Cerenkov detector

To realize the water Cerenkov detector as a near detector of J-PARCν experiment, we must overcome the demerits described above. To this end we use high performance photo sensors.

The vertex resolution of Cerenkov ring reconstruction is much dependent on the timing resolution of the photo sensors around the water tank.[13] It is expected that we obtain the better vertex resolution with photo sensors of the better timing resolution. The timing resolution is limited, however, by the speed of light and the diameter of the photo sensors. The path length differences of photons from a vertex to a photo sensor are order of D, the diameter of the sensor's sensitive area or the photocathode for PMT (see the figure 2.3). Since we cannot determine the places in the photocathode where the incident photon hits, the limit of the timing resolution σ_{lim} can be estimated as:

$$\sigma_{\text{lim}} \sim \frac{D}{c'} , \tag{2.3}$$

where c' is the light speed in water. Considering the total number of channels in the whole detector system and the expected number of photons in one neutrino event, we have decided to use $D{\sim}30$cm. Using the value of D and $c'{\sim}20$cm/ns, we calculate $\sigma_{\text{lim}}{\sim}1.5$ns. Thus, it is optimal to use a photo sensor with the timing resolution better than the σ_{lim}, 1.5ns.

In order to keep the occupancy manageable, it is important to reduce the dead time, from the equation 2.2 and the argument above. Here we roughly estimate the event rate of the J-PARCν near detector to determine the goal of τ_{dead}. In the 1 kiloton water Cerenkov detector at the near site of the K2K experiment, the average number of events observed in the full detector is about 0.2/spill including background events due to cosmic rays or ν_μ induced muons from upstream.[12] The number of protons/bunch in the J-PARCν neutrino beam is about 7 times larger than the number of protons/spill in the K2K neutrino beam, from the table 2.1. The neutrino flux of a beam with the OAB2° configuration is about three seventh of that with the WBB configuration.[14] From the 1 kiloton event rate, the ratio of beam fluxes of K2K and J-PARCν, and the difference of the beam configurations, we roughly estimate the event rate to be 0.6/bunch in the water Cerenkov detector in the J-PARCν near site, assuming the same volume with the K2K 1 kiloton detector. Accordingly, 2 (≥ 3) neutrino events occur in about 30% (7%) of those bunches with $\geq$1 event. Since the widths of the bunches are about 60ns from the table 2.1, the requirement to the dead time of the sensor is, $\tau_{\text{dead}} \ll 30$ns, for $P_{\text{occ}} \ll 1$.

We also require the detection capability of a single photon. To summarize, the requirements to the photo sensor for the J-PARCν water Cerenkov detector at the near site are:

1. The timing resolution of $\sim$ 1ns,
2. The dead time of $\sim$ 5ns,
3. The capability of single photon detection.

[ii] This means that photons from one event hit most of the sensors around the water tank. It is the case for small water Cerenkov detectors like the K2K 1 kiloton detector.

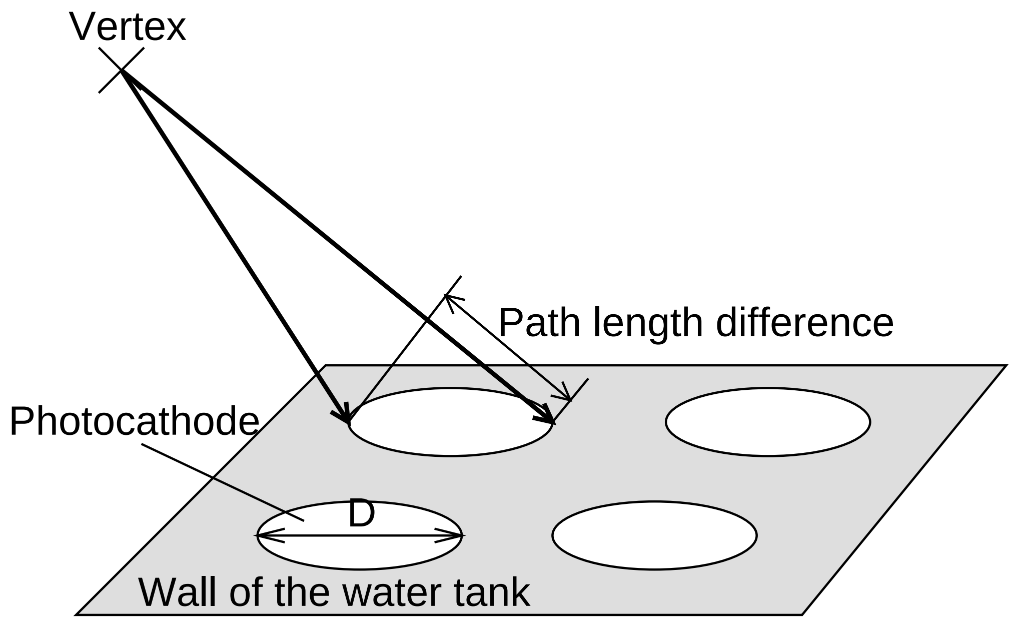


Figure 2.3: A schematic drawing to show the path length difference of photons from a vertex to a photo sensor.

The solution for the requirements: Hybrid Photo-Detector (HPD)

The HPD is a new detector under development, with unique characteristics. It has good time characteristics, a fast output signal and a low timing jitter. It can also detect a single photon. These characteristics exactly suit the requirements above, and hence we develop the HPD as a candidate of the photo sensor for the water Cerenkov detector.

2.3 For the future... (Hyper Kamiokande experiment)

The Hyper-Kamiokande has been proposed as a next generation water Cerenkov detector at Kamioka.[15] For the detector with a total mass of $\sim$ 1Mton, we need to develop large and low cost photo sensors. Our HPD is among the candidates of the sensor,[16] since it has a potential to be a low cost photo sensor, with its simple structure without dinode.

Chapter 3

The R&D of HPD(Hybrid Photo-Detector)

3.1 Overview of HPD

In this section, we briefly describe an overview of Hybrid Photo-Detector (HPD). A detailed description can be found elsewhere.[17]

The history of HPD

It was found that semiconductor devices are sensitive to ionizing particles, in the early 1960's. When a particle is driven in the depletion region of the reverse biased P-N junction of a semiconductor, the particle makes electron-hole pairs, and one can electrically read them.

If the particle is an electron and the energy of the incident electron is low enough that it is stopped in the semiconductor, the number of electron-hole pairs is proportional to the energy of the incident electron. Thus, we can consider the process as a multiplication process of electrons, where we call the gain of the multiplication *bombardment gain*. F. A. White and J. C. Sheffield proposed a possibility of a photon sensor made from the combination of a photocathode and a semiconductor device, using the multiplication process.[18] This type of photon sensor is called a "Hybrid Photo-Detector" (HPD) these days.

The potential of HPD was not seriously considered, though, until R. DeSalvo reinvented the HPD in 1987[19] to use as a photon sensor for particle physics experiments. After that, HPD made a steady progress, and nowadays, achieving the quality of practical use, some experiments plan to use HPDs.[20, 21, 22]

Hybrid Photo-Detector with an avalanche diode

An avalanche diode (AD) is a radiation sensor with intrinsic amplification owing to "avalanche multiplication." (See the appendix C.) HPD equipped with an avalanche diode (AD) has been studied due to its high gain, which is higher than HPD with a diode by one to two order(s) of magnitude.[i] The type of HPD is especially useful in single-photon detection at a high speed on the order of one nanosecond, and has a potential to replace the PMT. The HPD that we have studied is also classified in this type. Note that the avalanche type diodes are usually called avalanche photo diode (APD), although we call them avalanche diode (AD) since we detect electrons, not photons, with the diode.

In 1997, M. Suyama et al. reported on a practical HPD with AD.[23] N. Kanaya et al. made a detailed evaluation for this HPD.[24] The HPD with AD has been developed successfully, and it is now commercially available.[25] HPD with AD attracts, now, much attention in various fields such as gamma-ray telescopes,[26, 27] time-of-propagation (TOP) counter,[28, 29] and a detector for future linear colliders.[30]

The HPD with a multi-pixel AD, which has a position resolution, is also developed recently.[17, 31]

[i]This type of HPD may be sometimes called HAPD, distinguished from that with a diode.

The principle of HPD with AD

As described above, HPD with AD has two multiplication mechanisms. Figure 3.1 shows the principle schematically.

First, signals are multiplied by the bombardment, driving photoelectrons into the AD with a High Voltage (HV) applied between the photocathode and AD. The magnitude of the HV is usually about 10kV, and the bombardment gain is about 1,000 with the 10kV HV.

Second, the signals are multiplied by the multiplication in the avalanche diode (AD). The gain of the multiplication process, which gain we call *avalanche gain*, depends on the structure of the AD and the bias voltage applied on the AD. The AD in our HPD is an epitaxial type[32, 33] one, and the avalanche gain is about 50 with the maximum bias voltage.

With the two multiplication mechanisms, the HPD with AD realizes the gain of $\sim 10^5$.

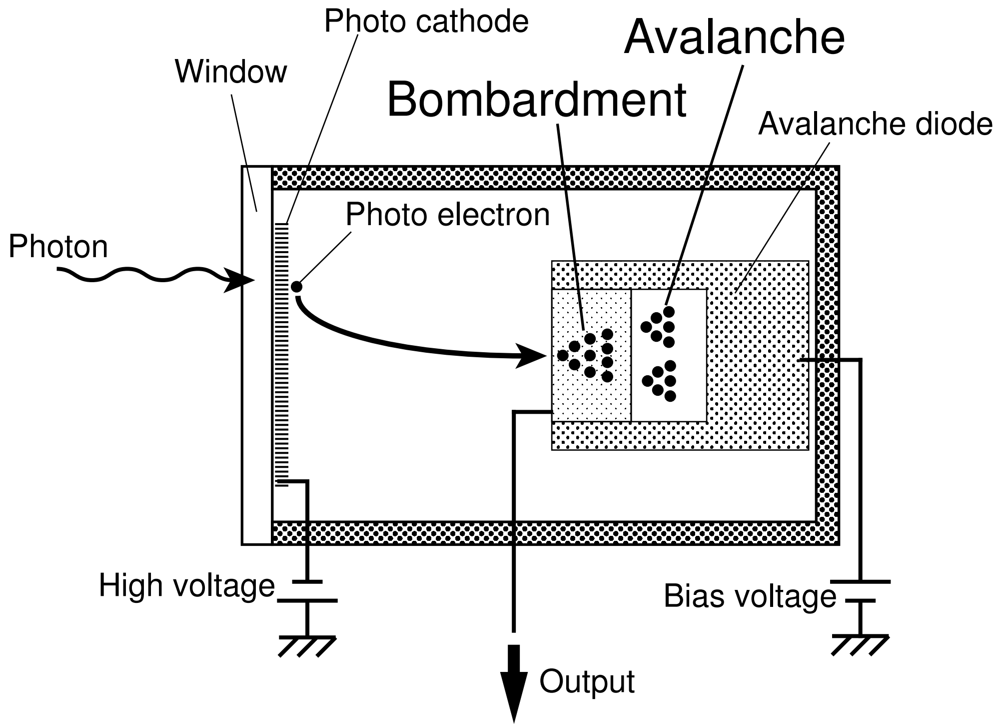


Figure 3.1: A schematic picture of HPD with AD.

3.2 The R&D of HPD

In this section, we describe the details of the R&D of the HPD. Here, we use two 5 inch prototype HPDs made by Hamamatsu K. K. Figure 3.2 shows one of the HPDs we use. The development of the 13 inch HPD, which is the goal of our R&D, is also going on, and its R&D status is described below.

3.2.1 The evaluation of 5 inch prototype HPD

Specifications

Table 3.1 shows the specifications of the two 5 inch prototype HPDs we use (HY0009 and HY0010). Figure 3.3 schematically shows the structure of them. Using these two HPDs, several feasibility tests have been done. Here we note that the HPDs we use are not final designed ones, and there still exist many rooms for the R&Ds. For example, the optics of electric field and quantum efficiency of the photocathode is not optimized yet.

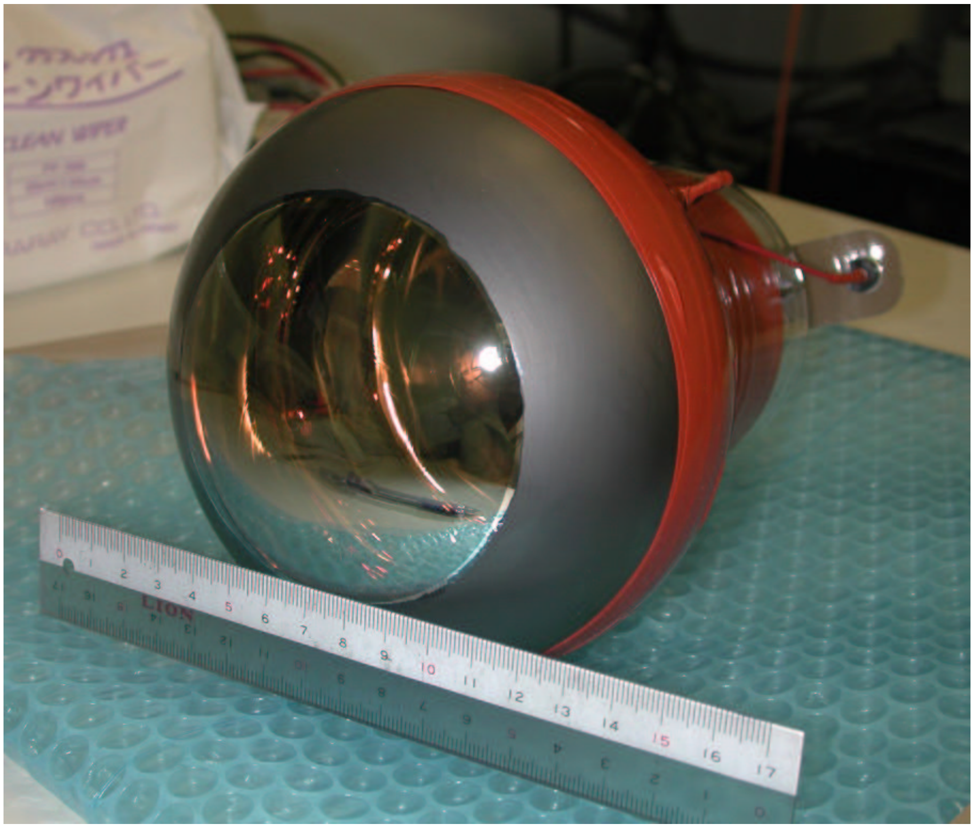


Figure 3.2: The 5 inch prototype HPD (HY0010).

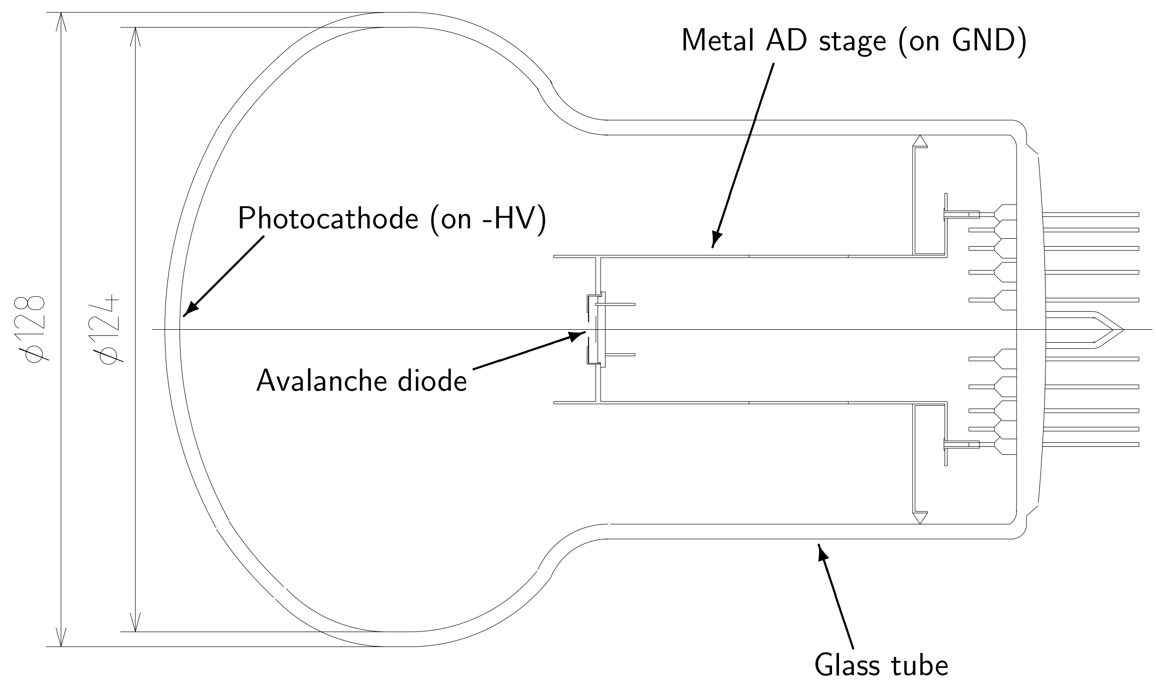


Figure 3.3: The structure of the 5 inch prototype HPD.

The main difference between the two HPDs is from the structure of the Avalanche Diode (AD). The avalanche diode used in HY0010, which we call *high speed type*, has a thicker depletion region than normal type. Due to the structure, HY0010 has low detector capacitance C_D and fast output signal shape.[ii] Note that the electric field strength of the depletion region is not so high even with its higher bias voltage on AD, owing to the thick depletion region. The gain of *high speed* type AD is, therefore, the same or small compared with that of normal type AD.

Table 3.1: The specifications of 5 inch prototype HPD.

Serial No.	Diameter	HV	AD size	AD type	AD Bias	Detector capacitance
HY0009	5 inch	> -8.5kV	3mmϕ	Normal	< 155V	~ 130pF
HY0010	5 inch	> -8.5kV	3mmϕ	High speed	< 350V	~ 30pF

Raw signal shapes

The raw signal shapes of two HPDs are shown in the figure 3.4. The HPD signal shape is measured with an oscilloscope[iii] by connecting with a 50Ω impedance BNC cable. HY0009 is measured with -8.5kV HV and 155V bias voltage on AD, and HY0010 with -8.5kV HV and 350V bias voltage on AD. The input signal is generated with a Picosecond Light Pulsar(PLP)[iv], and we adjust the strength of the input pulse light using an optical attenuator[v].

As shown in the figure 3.4, the raw signal widths are ∼10ns for HY0009 and ∼5ns for HY0010. The signal from HY0010, *high speed* type AD with less detector capacitance, is much faster than that from normal type.

Gain of Avalanche Diode(AD)

We measure the gain of AD (avalanche gain) as a function of the bias voltage on AD. Putting the light from PLP into HPD, we measure the output signal height of HPD. We observe the HPD output signal amplifying with the preamplifier, which we have developed (see chapter 4). Here we vary the bias voltage while keep High Voltage(HV) constant(-8.5kV). The setup is schematically shown in the figure 3.5.

The AD does not avalanche with low bias voltage: $\lesssim$40V for HY0009 and $\lesssim$60V for HY0010. At first, we measure signal heights with *no avalanche* mode as references, setting bias voltage of 40V and 60V for HY0009 and HY0010, respectively. Then to calculate the avalanche gain, we vary the bias voltage on AD and measure the signal heights with several bias voltage values.

In order to avoid the limitation of the preamplifier's dynamic range and to decrease the influence of its non-linearity, we measure the signal heights as follows. We adjust the input light strength by the optical attenuator so that the output signal height roughly stays 400∼1,000mV for various bias voltages. Then we convert the measured signal height, A_{mes}, using the attenuation coefficient, C_{att}, and calculate the avalanche gain G_{av} with the converted signal height:

$$G_{av} = C_{att} \times A_{mes}/A_{norm} , \tag{3.1}$$

where A_{norm} is the signal amplitude of no avalanche mode.

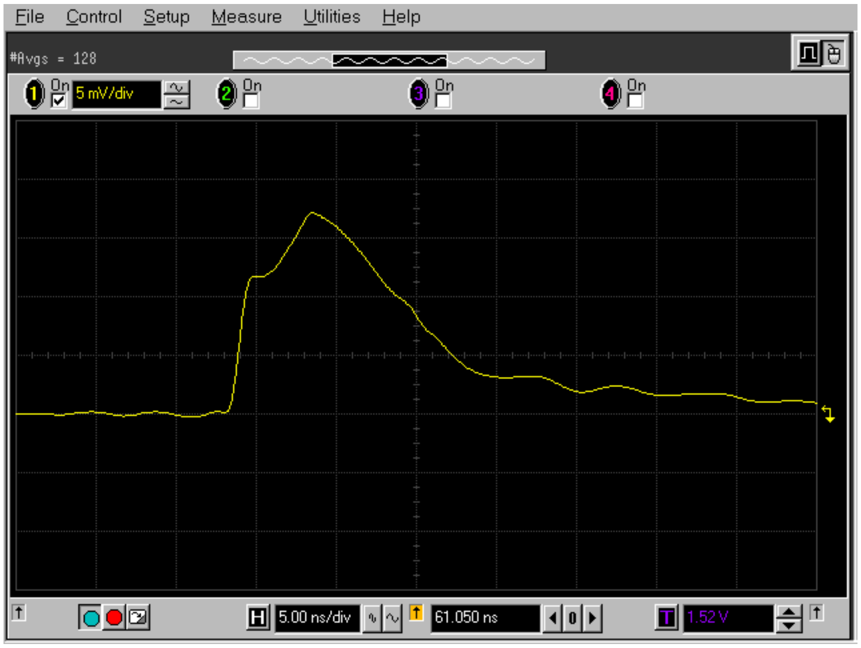
Figure 3.6 shows the result of the measurement. We obtain the avalanche gains of 58(HY0009, with 155V bias on AD) and 53(HY0010, with 350V bias on AD). The value of HY0009 agrees with the value measured by Hamamatsu Photonics K.K. (HPK). The value of HY0010, however, does not agree with the measurement of HPK, where the gain is ∼30. We understand the inconsistency as a result of the different methods of the measurement. The detector capacitance of AD decreases as the bias on AD increases, and the decrease is particularly large for HY0010, due to its low detector capacitance with

[ii] These two, the low C_D and the fast signal, are correlated, though the relation is not simple.

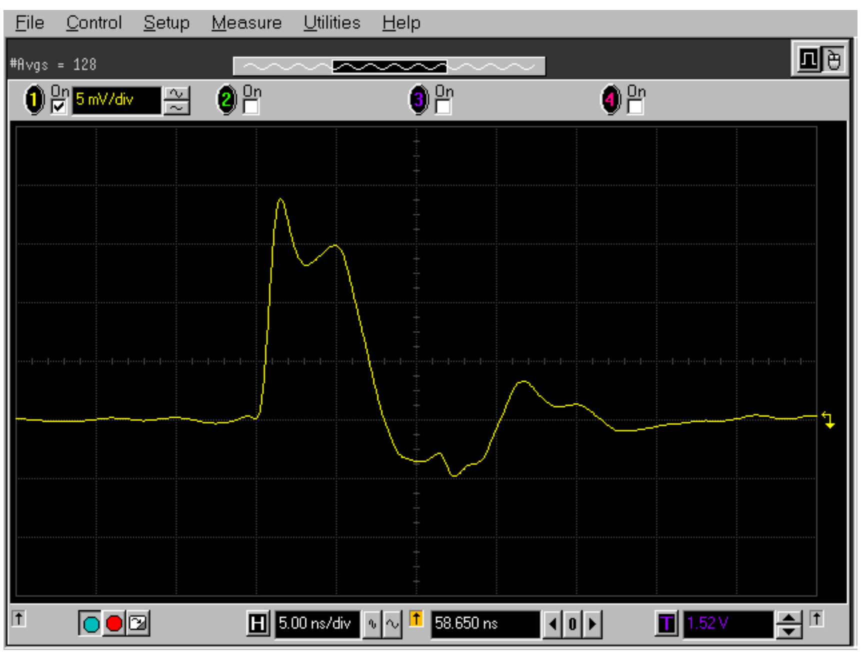
[iii] Infinium 54825A, by Hewlett Packard. Its bandwidth is 500MHz and its sampling rate is 2GHz.

[iv] PLP-02, by Hamamatsu Photonics. Its pulse width is less than 50ps. Its trigger jitter is less than 10ps, and its wavelength of light is 410nm.

[v] AQ-1227 Variable Optical Attenuator, by ANDO



(a) The raw signal shape of HY0009.



(b) The raw signal shape of HY0010.

Figure 3.4: The raw signal shapes of HPD. The scale of the horizontal axis is 5ns/division and that of the vertical axis is 5mV/division.

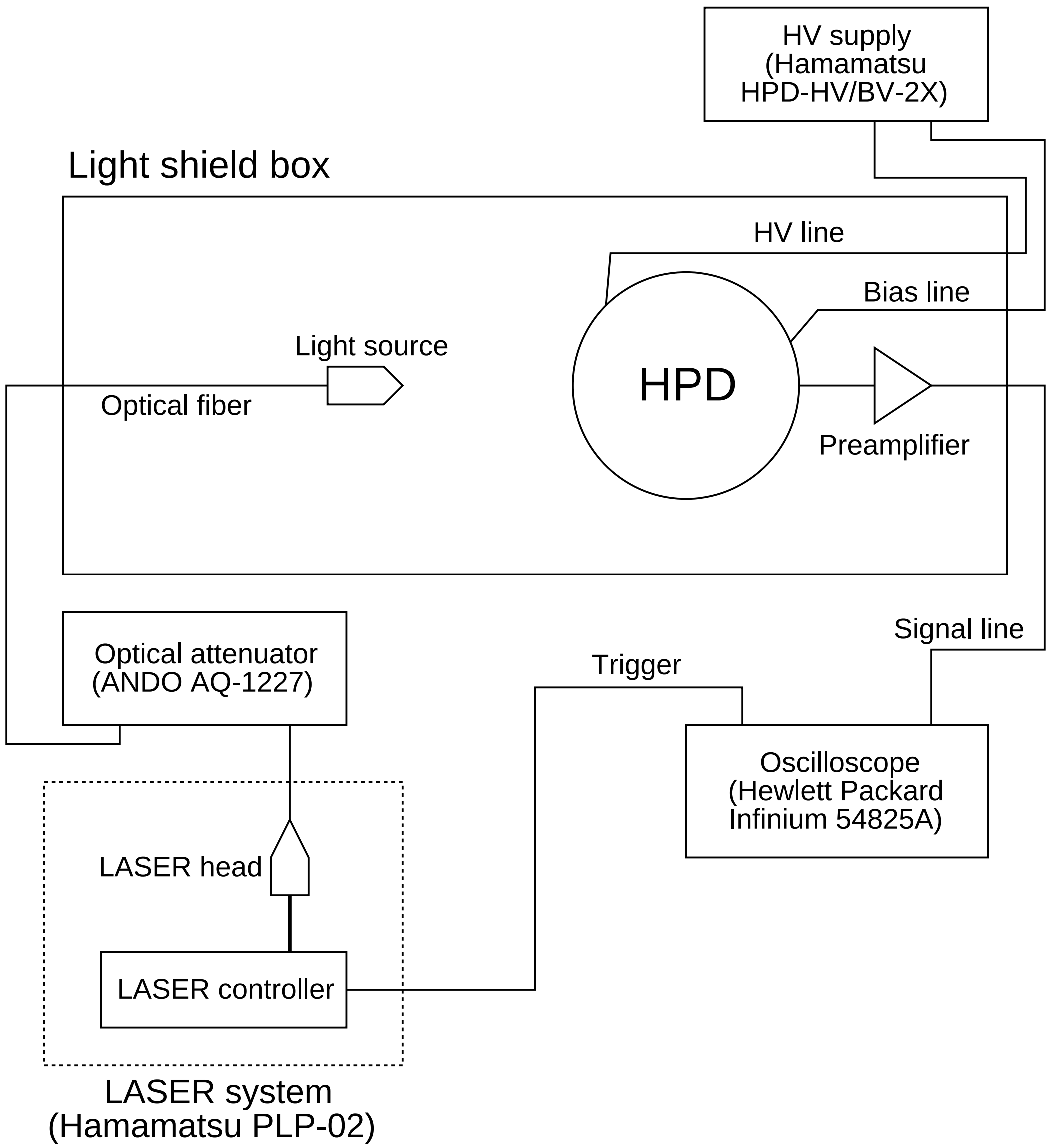


Figure 3.5: A schematic picture of the measurement setup.

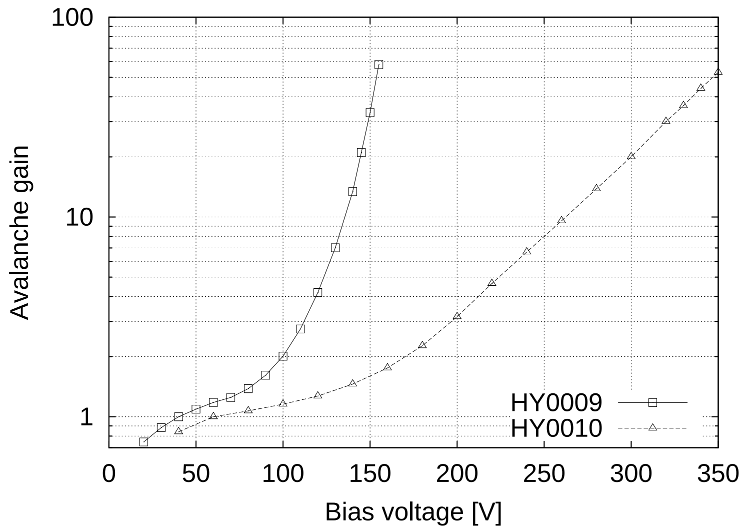


Figure 3.6: The avalanche gain of 5 inch prototype HPD.

maximum AD bias. Consequently, the speed of the pulse signal gets slower with low AD bias, due to the large detector capacitance. The slower signal results in the lower output signal amplitude, in pulse mode. In HPK, the AD gain is measured with DC input light signals, while we measure with the pulse input light signals. Thus, in our measurement, where the input signal is pulse-like, we may underestimate the signal amplitude of no avalanche mode, and accordingly, we may overestimate the avalanche gain with maximum AD bias.

Bombardment gain

We measure the bombardment gain as a function of applied HV. The measurement setup is the same as the measurement of the avalanche gain. Here we vary HV with constant AD bias voltage of 155V(350V) for HY0009(HY0010).

Differing from the avalanche gain measurement, we have no way to get the normalization of the

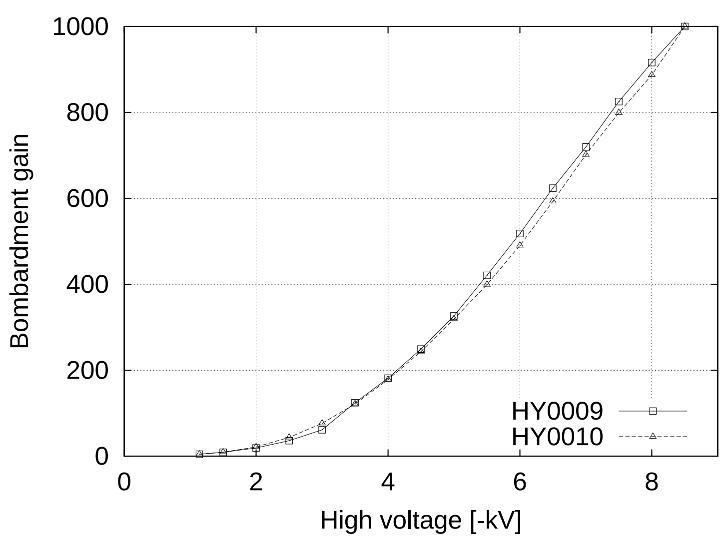


Figure 3.7: The bombardment gain of 5 inch prototype HPD. (Normalization should be confirmed.)

bombardment gain directly. Hence, we can only measure the relative gain here. Here, we normalize the bombardment gain so that the gain is 1,000 when applied HV is -8.5kV, and thus, the absolute value of the gain here has no meaning. We adjust the input signal using the optical attenuator in the same way as in the avalanche gain measurement.

The result of the measurements is shown in the figure 3.7. The gain rises at HV ~ -3.5kV. The reason is that AD has an insensitive layer on the avalanche region and the energy loss of a driven-in electron in the layer is about 3.5keV.

On the saturation of the avalanche gain

The avalanche gain of AD is saturated with an input signal larger than a specific value. We consider this is due to a space-charge effect in the avalanche region. The saturation mainly determines the intrinsic dynamic range (or non-linearity) of HPD itself.

HPK staffs have done a measurement on the saturation effect. They put LED pulse light with a width of 50ns into an HPD, and measure the relation between the output current and the avalanche gain. The measured HPD has the normal type AD. They observe the gain is ∼3% lower than the ideal (linear) behavior when the output current is 10mA and the bias voltage on AD is 153V. The 10mA current with a 50ns width corresponds to the output charge of

$$10[\text{mA}] \times 50[\text{ns}] = 500[\text{pC}]\,. \tag{3.2}$$

From the result, we consider the avalanche gain is saturated (by $\sim$ 3%) for the output charge of 10pC/1ns. Since the width of the output signal is about 10ns for a narrow ($\sim$ 50ps) pulse input (see the figure 3.4(a)), the avalanche gain begins to be saturated, for the narrow pulse input, when the output charge is 100pC. Assuming an avalanche gain of 50 and a bombardment gain of 1,000, the 100pC output corresponds to $100[\text{pC}]/(50 \times 1000) = 12,500$[photo electron] input for the HPD. This dynamic range is large enough, compared with the requirement from physics ($\lesssim$1,000[photo electron]). Even if we assume a bombardment gain of 4,000, the dynamic range is $\sim 3,000$[photo-electron] and still large enough. We expect the bombardment gain of the 13 inch HPD with 20kV HV is smaller than the 4,000 and hence the 13 inch HPD also has a large enough dynamic range.

3.2.2 The R&D status of 13 inch prototype HPD

Specification

The R&D of the 13 inch HPD is also going on. Figure 3.8 schematically shows the structure of the 13 inch prototype HPD. We show specifications of the 13 inch HPD in the table 3.2. There are several differences between the specifications of 5 inch and 13 inch HPDs.

The 13 inch prototype is modified from the 5 inch prototypes as follows.

13 inch HPD works with positive HV mode.

This is because 13 inch HPD is developed recently, and has more realistic design to work in water. In order to work in water, the photocathode, which contacts with water, must be electrically on the ground level, and hence HPD must work with the positive HV mode. This change brings new difficulties, as follows.

1. We have to apply the bias voltage for AD *on* the +20kV HV.
2. We have to insulate electrically the readout preamplifier at the ground level from AD at +20kV, using high voltage resistant capacitors. Otherwise, using a floating ground power supply system, we have to have the preamplifier work *on* +20kV voltage.

For the second item, we select the former option: insulate the preamplifier from AD using capacitors.

We increase the absolute value of the high voltage from 8.5kV to 20kV.

This is to achieve the higher collection efficiency of photoelectrons. It is expected from electric field optics calculations that the higher voltage between photocathode and anode, where AD is placed, gives the better collection efficiency. It is necessary to apply 20kV HV to realize the collection efficiency of 97.5% with the 13 inch photocathode and 5mmϕ AD.

The higher HV also has advantage because the bombardment gain increases as the HV gets higher. The bombardment gain is roughly proportional to (V_H − 4kV), where V_H is applied HV. We therefore expect the gain 3.5 times larger gain for 20kV HV than 8.5kV case.

13 inch HPD has larger size AD (5mmϕ).
Since the larger AD provides the better collection efficiency, we have decided to use the 5mmϕ AD.

The test result of 13 inch prototype HPD is described in the chapter 5.

Table 3.2: The specifications of 13 inch prototype HPD.

Diameter	HV	AD size	AD type	AD Bias	Detector capacitance
13 inch	$\lesssim +20$kV	5mmϕ	High Speed	$\lesssim 360$V	~ 70pF

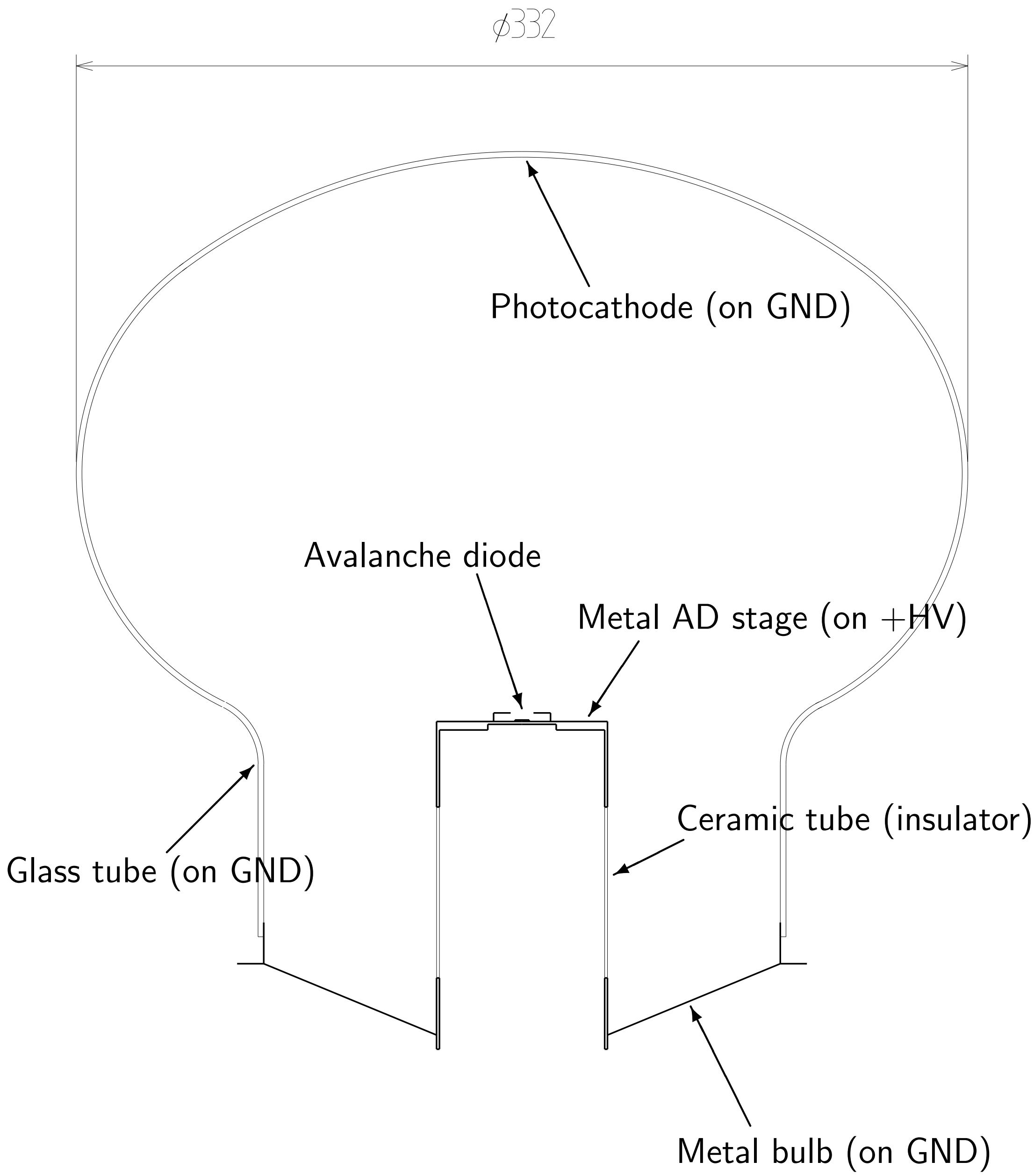


Figure 3.8: The structure of the 13 inch prototype HPD.

Chapter 4

The R&D of readout system

4.1 Design concept of readout system

4.1.1 The requirement to the readout system

As described in the section 2.2.5, the readout system for the HPD must meet the following requirements.

Low noise In order to detect a single photon using HPD with the gain of $\sim 10^5$, which is one or two order(s) of magnitude lower than that of PMT, the readout system must have a good noise figure. The low noise system is also required for the 1ns timing resolution.

Good timing resolution The readout system has to have the timing resolution better than 1ns. Since the dispersion of the photoelectron orbits is another dilution factor of the resolution, the resolution of the readout system should be *better* than 1ns to achieve the 1ns resolution in the total sensor system.

Short dead time The dead time of the system has to be very short. Taking account of the signal width of HPD, we aim at the dead time of ∼5ns.

Here the backend of DAQ system is not needed to be so fast, even requiring of the short dead time. This is because the signals are expected to come in burst-like sequence, and by putting a buffer of sufficient length, we can relax the requirement to the speed of the DAQ system after the buffer.

4.1.2 The design concept of the readout system

Figure 4.1 schematically shows the signal flow of our readout system. We put the signal from HPD into the fast and low noise preamplifier. The rise time of the output from the preamplifier strongly effects in the dead time of the total system, as we describe in the section 4.4. The preamplifier, therefore, has to be fast enough to follow the fast signal from HPD. Then we sample the output from the preamplifier with the sampling rate of ∼ 1GHz. We do this with a fast FADC or an Analog Memory Cell(AMC), as described in the section 4.3. Note that we put the output from the preamplifier into the waveform sampler directly without any shaping amplifier, because the digital filter after the waveform sampler plays the same role as the shaping amplifier. We also deduce the information of signal amplitude and signal timing using the FPGA or DSP for the digital filtering.

The main points of our idea are to use the fast, low noise preamplifier, and to adopt the waveform sampling. The reasons we adopt the waveform sampling are as follows.

1. We can achieve the very short dead time with the fast waveform sampling. Using the waveform sampler as a buffer, we can acquire multiple signals with ∼10ns intervals. It is very difficult to achieve the high rate DAQ (∼100MHz at peak) with the traditional ADC+TDC solution.
2. We can adopt digital filters only using the waveform sampling. The performances of digital filters are much superior to that of analog filters. Thus, we can highly enhance the signal to noise ratio and the timing resolution with them.

In order to achieve our requirement for the fast read out system, the first item is very important.

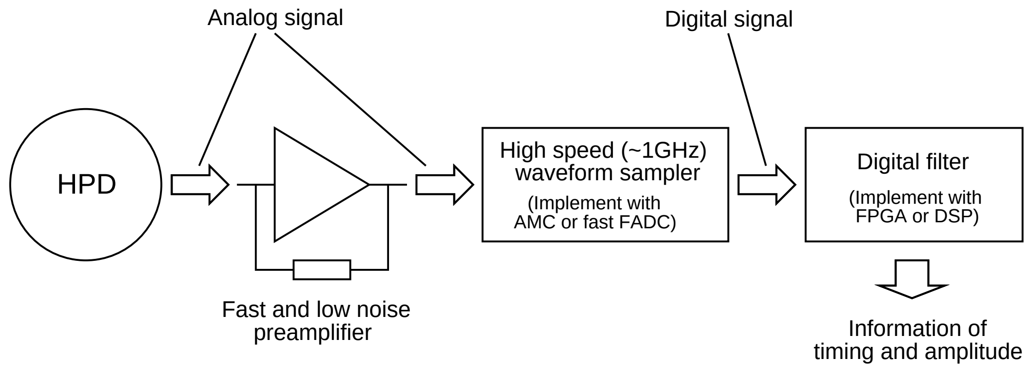


Figure 4.1: The signal flow in our readout system.

4.2 Preamplifier

We describe the development of the preamplifier in this section. The development proceeds as follows:

1. Conceptual design
2. Circuit simulation
3. Selection of the input transistor
4. Performance evaluation

4.2.1 Conceptual design

Requirements to the preamplifier

The requirements to the preamplifier are as follows.

The noise performance should be good. Since the gain of HPD is $\sim 5 \times 10^4$, the Equivalent Noise Charge(ENC) must be $\lesssim 5 \times 10^3$.

The rise time should be fast. The dead time of the total system is determined by the rise time of the output signal from the preamplifier, and the rise time is determined by the shape of the raw signal from HPD and the speed of the preamplifier. Thus, the preamplifier have to be fast enough, or have wide enough band, to catch up with the fast signal from HPD.

The input impedance of the preamplifier should be low. This requirement is derived from the two requirements above. If the input impedance is not low enough, the speed of the signal from HPD to the preamplifier becomes slow, and hence the amplitude of the output signal gets lower and the signal speed gets slower. Since the noise from the preamplifier is constant here, the lower output signal amplitude means the smaller signal to noise ratio.

We design the preamplifier to meet the requirement above.

Design concept

We have developed the design of a preamplifier referring the preamplifier for the Belle CsI calorimeter.[34] The design of our preamplifier is based on the traditional amplifier configuration: *cascode* type configuration with negative feedback. *Cascode* is a common-emitter-common-base configuration, in bipolar form. The cascode configuration has two merits. First, it increases output impedance. With the high output impedance, we can achieve high open loop gain by connecting the output to a high impedance node. The high open loop gain is essential to achieve the low input impedance for the amplifiers with the negative feedback. Second, it reduces unwanted feedback capacitance. It is essential to achieve the wide bandwidth.

We have examined two variations of the *cascode* configuration: totem pole cascode and folded cascode. Figure 4.2(a) schematically shows the totem pole cascode configuration. The merit of the totem pole cascode configuration is that it does not need any PNP type bipolar transistor. PNP type transistors are slow, with $f_T \lesssim 1$GHz, compared with NPN type ones, for which $f_T > 5$GHz is available. Here, f_T is the gain bandwidth product defined as:

$$f_T = f \cdot h_{\mathrm{FE}}(f)\ , \tag{4.1}$$

where $h_{\mathrm{FE}}(f)$ is the forward current gain at the frequency f.[i] The demerit of the totem pole cascode configuration is that we need the larger power supply voltage to get the larger dynamic range. The larger power supply voltage results in the larger power dissipation. In addition, it needs *level shifter* in the signal path as an extra node, depending on the dynamic range requirement. The *level shifter* with some specific configurations, which is determined by the dynamic range requirement, can be an additional noise source. Figure 4.2(b) shows the folded cascode configuration. The merit of the folded cascode configuration is that it does not need *level shifter*, while the demerit is that we have to use a PNP type bipolar transistor.

The first transistor of the cascode configuration, which is the input device of the preamplifier, is the most important device, since it is the primary factor that determines the noise performance of the preamplifier. As an input transistor, we have decided to use the Bipolar Junction Transistor (BJT). The reasons we select BJT are, 1) its transconductance, g_m, is large, 2) its bandwidth is very wide, and 3) its noise performance is good for wide bandwidth application, compared with other types of transistors.

4.2.2 Circuit simulation

After the conceptual design, we simulate the circuit with SPICE. Since we cannot easily change the circuit configuration of the preamplifier in practice, we need to fix the circuit configuration with the circuit simulation to some extent. It is also important to confirm that the circuit works as we intend.

We simulate the preamplifier circuit to determine the circuit configuration, checking:

1. The stability and bandwidth
2. The influence of the detector capacitance
3. The noise of the preamplifier

The simulation tool: PSpice

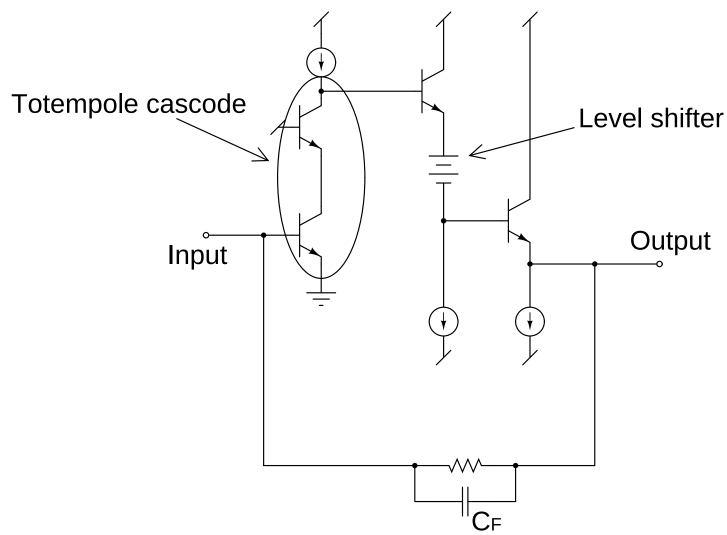
SPICE (Simulation Program with Integrated Circuit Emphasis) is a program that simulates electronic circuits on PCs, which is originally developed in U.C. Berkeley. PSpice is one of the many commercial derivatives of U.C. Berkeley SPICE. We simulate the preamplifier using Orcad PSpice version 10.0.

The limitations of the circuit simulation

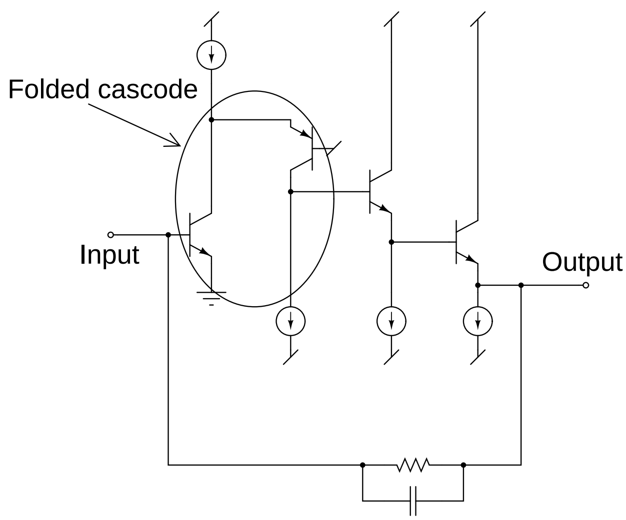
We have to note the limitations of the circuit simulation. One of the limitations is that we cannot estimate quantitatively the stability and the bandwidth, since PSpice does not emulate the parasitic capacitance and reactance of the conducting wire on printed circuits. Another is that we cannot always use the parts that we use in practice, since PSpice does not have sufficient parts libraries.

It is particularly serious for us that the PSpice library does not contain the fast PNP transistors with $f_T \sim 1$GHz, which are essential to simulate the folded cascode configuration. Accordingly, we emulate the folded cascode configuration by the totem pole cascode configuration, in the following way. As mentioned above, the differences between the two configurations are, 1) that the totem pole cascode configuration amplifier has *level shifter*, and 2) that a PNP transistor in the folded cascode configuration is slower than NPN transistors. Thus, we can substitute the totem pole cascode configuration for the folded cascode one, using an ideal *level shifter* and a slower NPN transistor for the second transistor. The ideal *level shifter*, which is available only in the simulation, contains no noise source and has an infinitely quick response.

[i] This definition is valid when $h_{\mathrm{FE}} > f_T/f$ and $f < f_T$, where h_{FE} is the forward current gain at $f \sim 0$.



(a) The totem pole cascode configuration.



(b) The folded cascode configuration.

Figure 4.2: Schematic pictures of the preamplifier configurations.

Simulation settings

Figure 4.3 shows the circuit we use in the simulation. We have simulated our circuit with three settings. The intention of each setting is as follows.

Setting 0 The ideal, reference setting.

Setting 1 The setting to examine the totem pole cascode configuration.

Setting 2 The setting to examine the folded cascode configuration.

Among the settings, the *level shifter* and the second transistor Q16 differs. We list the differences in the table 4.1. Table 4.2 shows the transistors we use in the simulation with some of their specs. We use

Table 4.1: The settings for the circuit simulation.

Setting #	Second transistor (Q16)	Level shifter
0	Normal NPN	Ideal
1	Normal NPN	Practical
2	Slow NPN	Ideal

Table 4.2: The transistors used in the circuit simulation.

Name	Role	f_T	h_{FE}
2SC4841	Bipolar(input), Normal NPN	7GHz	150
2SC4215	Slow NPN	0.8GHz	70

Note:f_T and h_{FE} are the values for $I_{C(D)}$=3mA.

the circuit configuration shown in the figure 4.4 as the *practical* level shifter. We have determined the values of the registers R10 and R11 to adjust the shift voltage to 8V. The shift voltage is determined only by the relative values of R10 and R11, and we can use arbitrary value in this view point. Large R10 or R11 results in, however, the large noise. Thus, we have decided to use the values shown. The capacitor C_D, in the figure 4.3, is not the part of the amplifier. It is a capacitor to emulate the detector capacitance.

Procedure of the simulation

First, we check the stability of each setting. If a setting is not stabilized, we modify the circuit configuration and stabilize it. After the stabilization, we check the bandwidth of the preamplifier. We set the detector capacitance, C_D, to be zero here. Second, we check the effect of C_D on the stabilized circuits, changing the value of C_D. Finally, we measured the noise of each setting. Here, C_D is set to be zero, again.

Stability and bandwidth

We check the stability and the bandwidth of each setting. We set $C_D = 0$ in this evaluation.

Figure 4.5 shows the amplitude and the phase of the current flow in the feedback capacitor C1, in the setting 0, as a function of the frequency. The amplitude of unity corresponds to the value of the input current. The phase 0 corresponds to the exact negative feedback and the phase $\pm\pi$ corresponds to the exact positive feedback. Thus, the amplitude equals to 1 and the phase equals to 0 if the amplifier is *perfect*, that is, the open loop gain of the amplifier is infinite and the response of the amplifier is infinitely fast. To stabilize the circuit, the amplitude must be small enough at the frequency in which the absolute value of the phase is near or larger than $\pi/2$.

We call the default configuration of feedback shown in the figure 4.3 *default feedback*. In the *default feedback* line of the figure 4.5, one can see the amplitude is large at the frequency of a few GHz, where

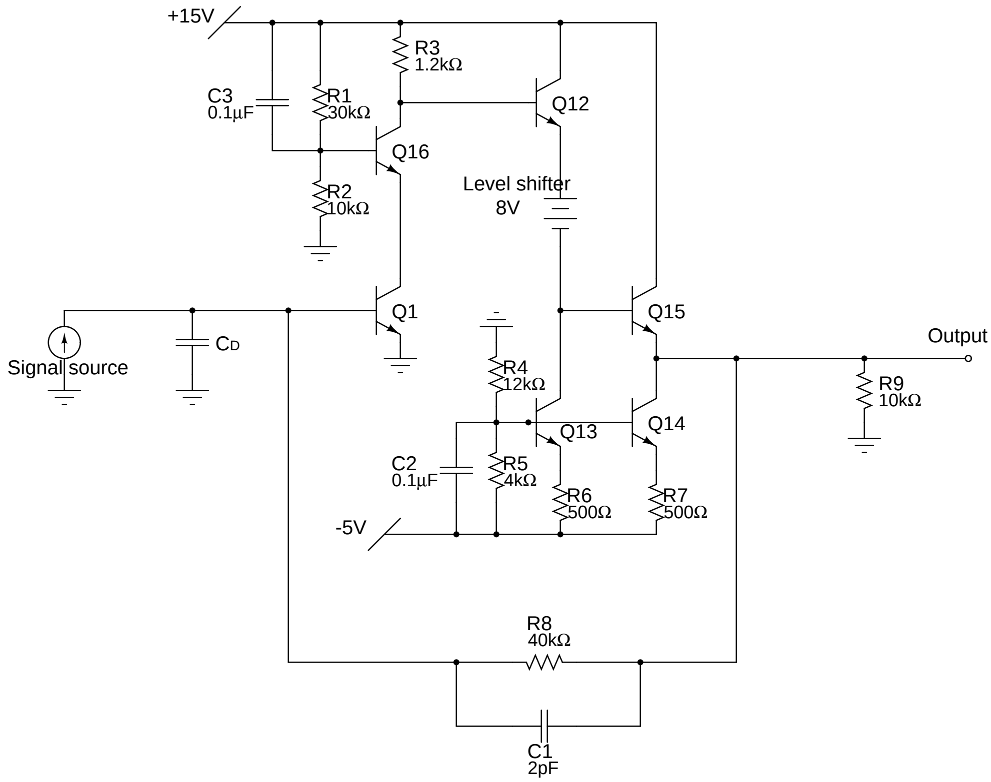


Figure 4.3: The circuit used in the simulation.

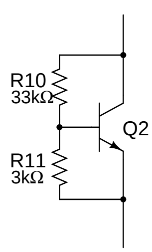


Figure 4.4: The practical configuration of the level shifter.

the phase is close to $-\pi/2$. This suggests the instability of the setting 0. In practice, there is a ringing at the beginning of the impulse response of the *default feedback* configuration, as shown in the figure 4.6(a).

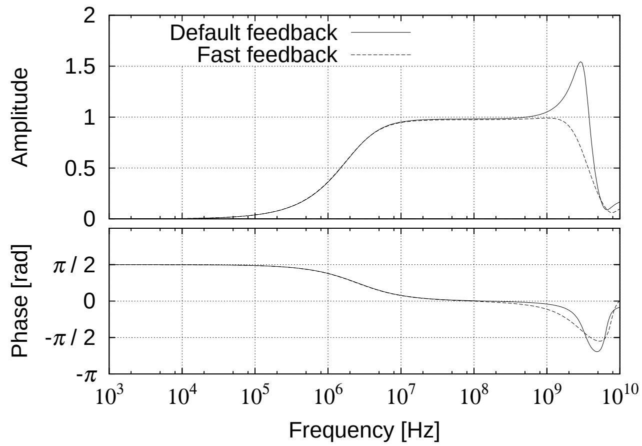
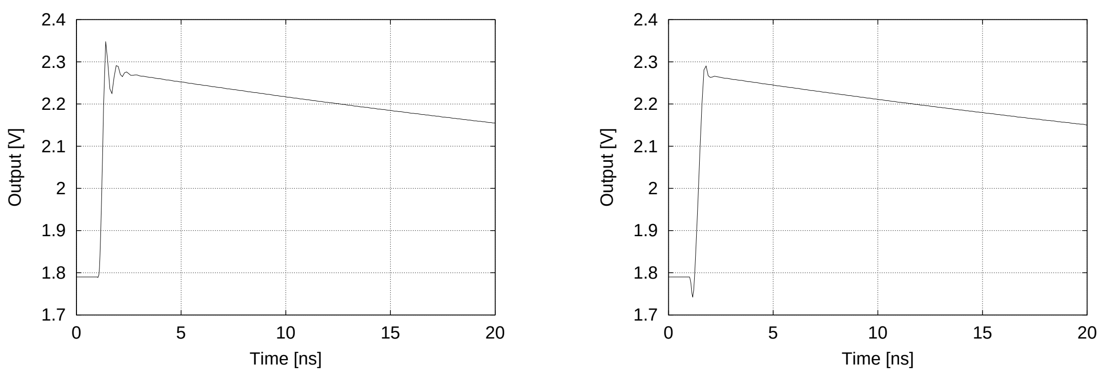


Figure 4.5: The amplitude and the phase of the feedback current. The vertical axis is normalized by the input current.



(a) The waveform of *default feedback*.

(b) The waveform of *fast feedback*. The dent before the rise is not essential here.

Figure 4.6: The beginning of the impulse response of the setting 0 circuit.

Accordingly, we have to stabilize this circuit. We use two methods to remove such instability. Figure 4.7(a) shows the first method. The method is called *compensation*. The strategy of the compensation is simple: to reduce the bandwidth of the amplifier, which method is often called *narrowbanding*, by adding a capacitor in the signal path. The additional capacitor corresponds to the C_C in the figure 4.7(a). By *narrowbanding*, we can decrease the feedback amplitude at the *instable frequency*, and then we can make the circuit stabilized.

The second method is shown in the figure 4.7(b). We call this method *fast feedback*. With this method, we can make the circuit stabilized in two ways. In one way, the feedback capacitor, C_F, in the figure 4.7(b) works as the capacitor for *compensation*, in the same way as C_C in the figure 4.7(a). This is because that one of the C_F's terminals of the input side is almost fixed at a constant voltage. In another way, the feedback current through C_F in *fast feedback* configuration is namely faster than that in the *default feedback* configuration, which is shown in the figure 4.2(a). This is because the former configuration bypasses the two emitter follower transistors. The *faster* feedback current corresponds to the phase that is *closer* to zero. Thereby, the circuit is stabilized.

Comparing the two methods, *compensation* and *fast feedback*, the *fast feedback* is better, since its functions include that of the *compensation*. The *fast feedback* has, however, a limitation that we cannot make C_F larger, while we can make C_C in *compensation* larger as needed. Thus, we use the two prescriptions in the following way. If the circuit is not stable, we use *fast feedback* method, first. If it is not stable yet, we use *compensation*, in addition to the *fast feedback*, and increase C_C until it gets stable. For example, we make the setting 0 circuit stabilized only with the *fast feedback*, which the figures 4.5 and 4.6(b) show the result, and we do not use the *compensation* for the setting.

Table 4.3 shows how we can stabilize the circuit in each setting. Only the setting 2 needs the additional *compensation* capacitor, while the other settings can be stabilized only with the *fast feedback*. Figure 4.8 shows the amplitude and the phase of the feedback current in all settings with the stabilized configuration. Note the bandwidth of the feedback current corresponds to the bandwidth of the amplifier. Accordingly, the result suggests that, in setting 2 with the slow transistor, the bandwidth is limited not only by the narrow band of the transistor but also by the compensation capacitance, which is needed to stabilize the amplifier.

Table 4.3: The way to stabilize each amplifier setting.

Setting	The way
Setting 0	Fast feedback only
Setting 1	Fast feedback only
Setting 2	Fast feedback and 4pF compensation capacitor

The influence of the detector capacitance C_D

The second item we check using PSpice is the open loop gain of the amplifier. We need the sufficiently large open loop gain to get the fast signal from the detectors with large detector capacitance like HPD, whose detector capacitance is typically ~50pF. We check the open loop gain of each setting, by varying the capacitor C_D in the figure 4.3. The open loop gain is not large enough if the amplitude of the output signal decreases significantly with increasing the capacitor C_D.

We input the current signal, having the shape shown in the figure 4.9. Since the integration of this current signal corresponds to the charge of 0.1fC, and the feedback capacitor, C_F, is 2pF, the output amplitude is expected as 0.1fC/2pF = 50mV if the amplifier is perfect, that is, open loop gain of the amplifier is infinite. Varying the capacitance of C_D, we measure the amplitude of the output signal for each setting.

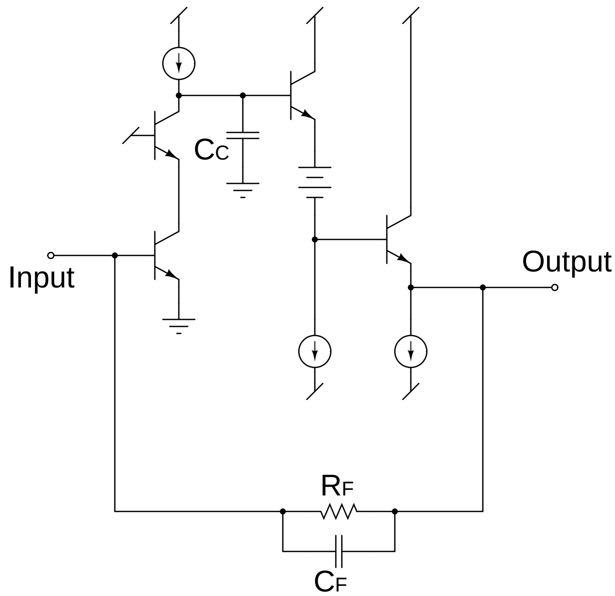
The results are shown in the figure 4.10. There is little difference among the three settings.

Noise estimation

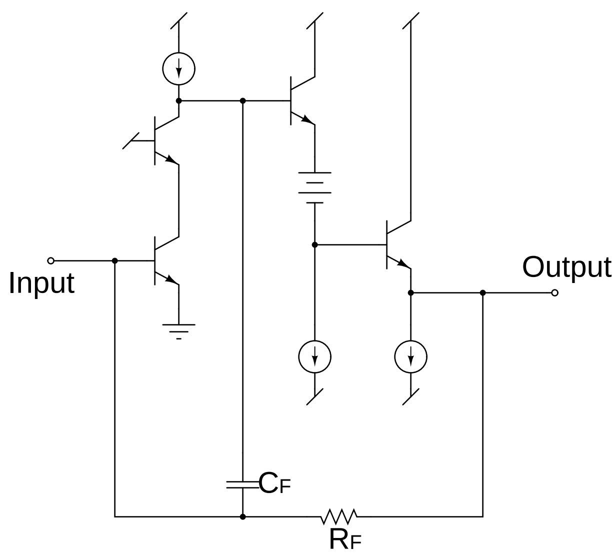
The noise of the circuit is the final item that we check using PSpice simulation. The primary noise source of an amplifier is usually, and *should be*, the input transistor. We check if there are any extra noise sources that are not negligible.

Figure 4.11 shows the spectra of the total noise and the noise from the input transistor, Q1, in the setting 0. Figure shows that the dominant noise source is Q1 only. This is almost the same for the setting 2.

On the other hand, setting 1 shows a little different tendency. Figure 4.12 shows the noise spectra, in the setting 1, for the total noise, the noise from Q1, the noise from the transistor used in the level shifter Q2 and the noise from the register used in the level shifter R11, which are shown in the figure 4.4. There is the extra noise from the level shifter nodes, Q2 and R11, and it is not negligible. As shown in the figure 4.12, the contaminations of the noise from Q2 and R11 become significantly large at the high frequency region. Since the noise performance in the high frequency region is very important in our application, the noise contamination is large disadvantage of the configuration of setting 1.



(a) The *compensation* method.



(b) The *fast feedback* method.

Figure 4.7: The circuit configurations to stabilize the amplifier.

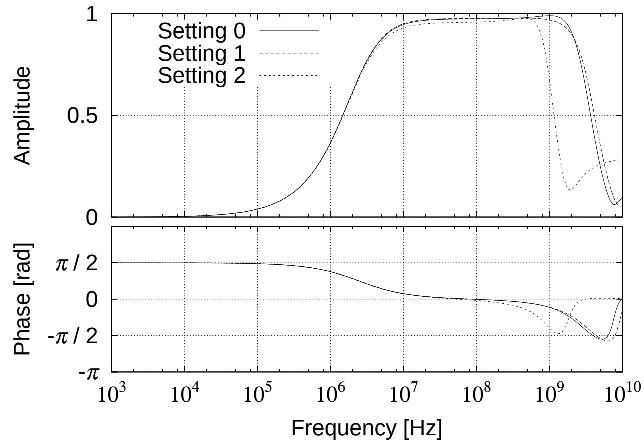


Figure 4.8: The amplitude and the phase of the feedback current.

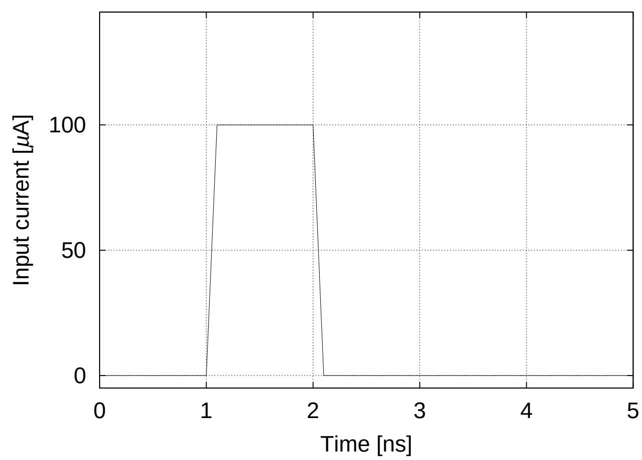


Figure 4.9: The input signal current to check the open loop gain.

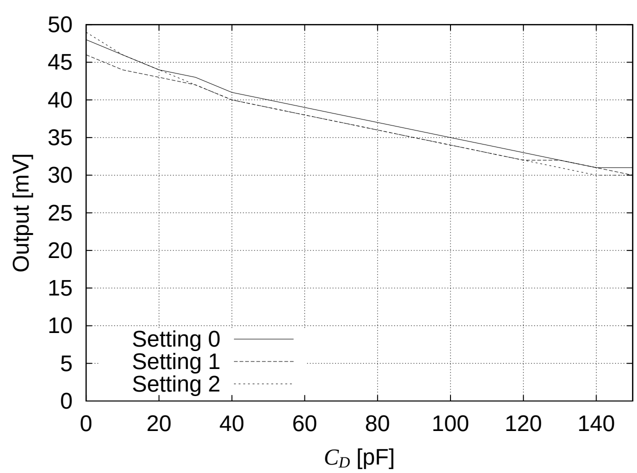


Figure 4.10: The output signal amplitude as functions of C_D.

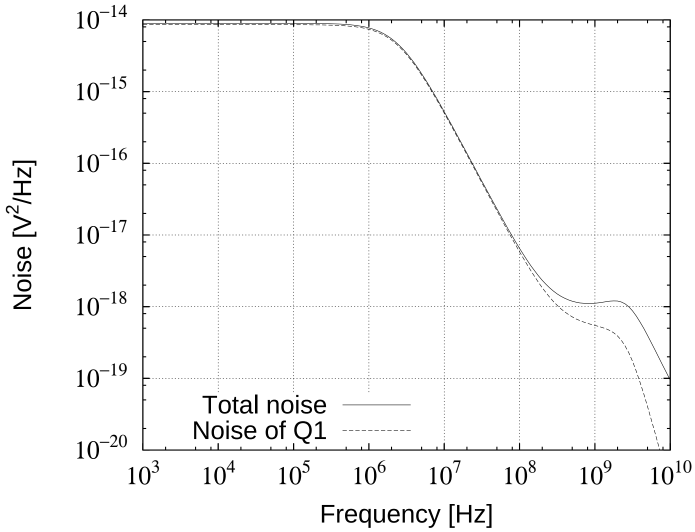


Figure 4.11: The noise spectra of the setting 0.

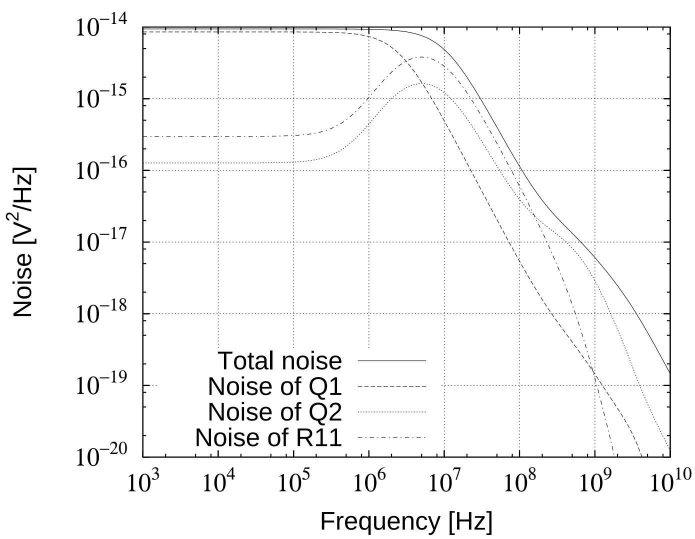


Figure 4.12: The noise spectra of the setting 1.

Summary of the simulation result

We summarize the results of the circuit simulation using PSpice, here.

- On the stability and the bandwidth.
  - We can stabilize the instable preamplifier using the *fast feedback* configuration and/or the *compensation*.
  - The bandwidth of the preamplifier using a slow BJT is limited, not only by the bandwidth of the transistor, but also by the capacitor for the *compensation* needed to stabilize the circuit.
- On the open loop gain
  - The open loop gains are large enough for all settings.
- On the noise
  - The level shifter can be the sizable noise source, in the setting 1.
  - Only the input transistor is the noise source that we have to pay attention, for the other two settings.

By the results of the simulation, we have decided to adopt the folded cascode configuration, which corresponds to the "setting 2" in the simulation above. The reasons of the decision are as follows.

1. The noise contamination from the level shifter in the "setting 1," which is shown in the figure 4.12, is a large disadvantage for our application.
2. The disadvantage of the "setting 2," the narrow bandwidth, is not significant, since the bandwidth shown in the figure 4.8 is ∼1GHz and still wide enough.

4.2.3 Selection of the input transistor

As described in the previous section, the input transistor is the dominant noise source in our preamplifier. Accordingly, we have to choose the transistor with low noise as the input transistor. Here we describe the selection of the input transistor.

First, we calculate the noise of the bipolar junction transistor (BJT) and its propagation in a feedback amplifier, in order to make out which noise source of the input transistor has the dominant effect on the noise performance of the preamplifier. The calculation is necessary to decide the selection criteria. Second, we do some preselections that can be done without any measurement. By the selection, we reduce the number of the candidates for real noise measurement. Third, we measure the noise characteristics of the candidate transistors and decide the transistor to use as the input transistor of the preamplifier.

The noise of the input transistor

Here, noises are expressed in terms of noise power, and represented in lower case, e.g., a current (voltage) noise power of the current I_x (voltage V_x) is expressed as ${i_x}^2$ (${v_x}^2$). Refer the appendix A.1 for more detailed description.

We briefly describe the noise of the bipolar transistor. For more detailed description, refer the appendix A.1.1. The input equivalent noise of the bipolar transistor is as follows:

$$v_n^2 = 4kTR_{bb'}\Delta f + \frac{2eI_c}{{g_m}^2}\Delta f + \left(2eI_b + K\frac{I_b^a}{f^b}\right){R_{bb'}}^2\Delta f \tag{4.2}$$

$$i_n^2 = 2eI_b\Delta f + K\frac{I_b^a}{f^b}\Delta f \tag{4.3}$$

Here, k, and e are the physical constants: the Boltzmann constant, and the electron charge, respectively. T, I_c, I_b, and f are the parameters determined externally: the temperature, collector current, base current, and the frequency, respectively. $R_{bb'}$, g_m, $K_{1/f}$, a, and b are the parameters that depend on the transistor type: the base spreading resistance, transconductance from base to collector, and constants that describes $1/f$ noise, respectively.

It may be convenient to show rough magnitudes of each term here. Assuming $R_{bb'} = 50[\Omega]$, $I_c = 1[\text{mA}]$, $h_{\text{FE}} = 100$, $a = b = 1$, $T = 300$ and $K \sim 10^{-16}$ as typical values, and using the following relation,

$$g_m \sim \frac{eI_c}{kT} \tag{4.4}$$

we have calculated typical value of each term in the equation 4.2 and 4.3, which values we list in table 4.4 and 4.5.

Table 4.4: Typical values of each term in voltage noise of BJT.

Term	$4kTR_{bb'}$	$2eI_c/{g_m}^2$	$2eI_b{R_{bb'}}^2$	$K\frac{I_b^a}{f^b}{R_{bb'}}^2$
Value [V^2/Hz]	1×10^{-18}	2×10^{-19}	1×10^{-20}	$(1/f[\text{kHz}]) \times 10^{-20}$

Table 4.5: Typical values of each term in current noise of BJT.

Term	$2eI_b$	$K\frac{I_b^a}{f^b}$
Value [A^2/Hz]	3×10^{-24}	$(1/f[\text{kHz}]) \times 10^{-24}$

The noise propagation from the input transistor to the preamplifier

Using the equation A.19, we estimate how the noises, v_n and i_n, of the input transistor effect in a negative feedback amplifier. We can regard the detector, AD, as a capacitor with the detector capacitance of C_D. Our feedback impedance consists of a capacitance C_f and a resistance R_f. Thus, we write Z_D and Z_f as:

$$Z_D = \frac{1}{2\pi f j C_D} \tag{4.5}$$

$$Z_f = \frac{R_f}{1 + 2\pi j f C_f R_f} \tag{4.6}$$

We then rewrite the equation A.19 as:

$${i_{\text{eq}}}^2 = \left[(2\pi f)^2 (C_f + C_D)^2 + \frac{1}{{R_f}^2}\right] {v_n}^2 + {i_n}^2 \tag{4.7}$$

This result shows that v_n is the dominant noise source in the high frequency region, while i_n is dominant in the low frequency region. If we assume $C_D = 50[\text{pF}] \gg C_f$, $R_f = 50[\text{k}\Omega]$ and typical BJT noise values in table 4.4 and 4.5, the critical frequency f_c at which the contribution from v_n becomes comparable with that from i_n, is about 100kHz. The frequency is much smaller than the bandwidth of our preamplifier, 100MHz$\sim$1GHz. Since we especially make use of the higher frequency region than the 100kHz, the noise we have to pay attention is v_n.

The selection criteria of the input transistor

Among the constants included in the equations 4.2 and 4.3, $R_{bb'}$ is the only one that we have to pay attention in the selection of transistors. The reasons are as follows.

- As described above, we have decided to use BJT as the input transistor. The $1/f$ noise term of BJT is small enough in general, while other types of transistors have larger $1/f$ noise. Consequently, we need not care about the $K_{1/f}$ so much.

- The effect of ${v_n}^2$ is more important than that of ${i_n}^2$ as described above. Table 4.4 shows that the first term is dominant, among the terms in the equation 4.2.
- The transconductance g_m has little variation among different transistors because it is determined mostly by the physics of the P-N junction. Thus, it cannot be a selection criterion.

Although the $1/f$ noise is expected to be negligible, as mentioned above, we also have to confirm really it is.

Accordingly, we decide the selection criteria for the input transistor as follows:

1. The bandwidth of the transistor must be sufficiently wide.
2. The base spreading resistance, $R_{bb'}$, should be small.

Besides the noise performance, the speed of the transistor is also important. The wide bandwidth, or fast speed, is needed to achieve the fast rise time. Since the width of the signal from HPD is less than 10ns, the bandwidth of the preamplifier must be $\sim$ 100MHz. Note that it is *not* sufficient to use the transistors with $f_T \sim$ 100MHz to achieve the 1GHz amplifier bandwidth. Here, the definition of f_T can be written as [ii]

$$f_T = f \cdot h_{\mathrm{FE}}(f) \tag{4.8}$$

where $h_{\mathrm{FE}}(f)$ is the forward current gain at the frequency f. Since h_{FE} must be larger than 1 for a transistor to work as a *transistor*, f_T must be much larger than 100MHz.

Preselection by the $\mathbf{R_{bb'}}$ estimation

In the transistor selection, we try BJTs of NEC Compound Semiconductor Devices, TOSHIBA Semiconductor Company, Renesas Technology and SANYO Semiconductor Company.

First, we have applied the criterion $f_T >$ a few GHz. There are a few hundreds or more types of NPN BJTs that passed this criterion. Then the base spreading resistance $R_{bb'}$ is the only characteristic for the further selection. Unfortunately, $R_{bb'}$ is commonly not listed in data sheets. It is not realistic, however, to evaluate a few hundreds types of BJTs in practice. Thus, we have done a preselection by the estimation of $R_{bb'}$ from S parameters, which are available in the web sites of the vendors or in the data sheets for RF devices.[iii] The method of the estimation is described in the appendix A.2. As is described there, the method assumes certain transistor model, which is in the figure A.4. Thus, the selection is proceeded by considering:

- If the behavior of the transistor matches with the model.
- The value of estimated $R_{bb'}$. The smaller is the better.

We have applied the selection to 192 transistors: 75(NEC), 18(TOSHIBA), 41(Renesas) and 58(SANYO). We also considered the commercial distribution features, then we have narrowed down the number of candidates to 13 with the preselection.

On the SiGe HBT

In addition, we also try silicon germanium (SiGe) Hetero-Junction Bipolar Transistor (HBT) as our candidates. There are some classes of BJTs. Normal Si transistors, GaAs transistors, and the SiGe transistors are representative. Among them, GaAs has large $1/f$ noise and does not suit our application, as well as MOSFET, HEMT, etc. Thus, normal Si transistors and the SiGe transistors are the only choices left for us.

SiGe is a new technology that offers significant improvements in operating frequency, current, noise and power capabilities. The characteristics of SiGe HBTs are widely noticed for the applications like high speed wireless communication and Global Positioning Satellite (GPS). Since they also meet our requirements and we are interested in the technology, we add 6 SiGe HBTs (2 from NEC and 4 from TOSHIBA) on our list of candidates.

[ii] This definition is valid when $h_{\mathrm{FE}} > f_T/f$ and $f < f_T$, where h_{FE} is the forward current gain at $f \sim 0$.

[iii] Here, we have the additional criterion: *the S parameter is available*, though this is nearly equivalent to "$f_T >$ a few GHz."

Table 4.6: The BJT candidates for actual noise measurements.

Vendors	Type
NEC	2SC5008, 2SC5011, 2SC5012, 2SC5508, 2SC5509, 2SC5677, 2SC5704
NEC (SiGe)	2SC5761(NESG2030M04), NESG2031M05
TOSHIBA	2SC5086, 2SC5319, 2SC5464, MT3S03T, MT3S04T
TOSHIBA (SiGe)	MT4S100T, MT4S101T, MT4S102T, MT4S104T
SANYO	2SC5781

$\mathbf{R_{bb'}}$ measurement

After the preselection, we carry out the noise measurement for the candidate 19 transistors listed in the table 4.6. As described above, the only parameter we have to measure is the base spreading resistance $R_{bb'}$.

We can estimate the $R_{bb'}$ by the measurement of noise. Figure 4.13 shows the schematic view of our noise measurement setup. The noise of collector current i_c can be estimated by the relation:

$$\frac{{i_c}^2}{\Delta f} = 2eI_c + {g_m}^2 \left\{ 4kT(R_{bb'} + R_b) + \left(2eI_b + K\frac{{I_b}^a}{f^b} \right) (R_{bb'} + R_b)^2 \right\} \tag{4.9}$$

Ignoring the $1/f$ noise, $R_{bb'}$ can be expressed as:

$$R_{bb'} + R_b = \frac{2A}{4kT + \sqrt{(4kT)^2 - 8eI_bA}} \tag{4.10}$$

$$\left(A \equiv \frac{{i_c}^2/\Delta f - 2eI_c}{{g_m}^2} \right) \tag{4.11}$$

where all parameters in the right-hand side are constants (k, e) or measurable ones (${i_c}^2/\Delta f$, I_c, I_b, g_m, T). Here, ${i_c}^2/\Delta f$ represents noise power of the noise floor, since we ignored the $1/f$ noise. Thus, we can estimate $R_{bb'} + R_b$ by the measurement of ${i_c}^2/\Delta f$, I_c, I_b, g_m and T[iv]. We define the estimated value of $R_{bb'} + R_b$ as $R_{b\mathrm{mes}}(R_b)$. In the ideal measurement, the following relation is realized:

$$R_{b\mathrm{mes}}(R_b) = R_{bb'} + R_b \tag{4.12}$$

We also define $R_{b\mathrm{mes}} \equiv R_{b\mathrm{mes}}(0)$.

We measure collector current noise power, ${i_c}^2/\Delta f$, and other parameters, I_c, I_b, g_m, and T, simultaneously using the NoisePro system. The NoisePro system consists of a low noise preamplifier, a noise analyzer (Celestry 9812B) which controls the biases on the transistor, a vector signal analyzer

[iv] We assumed $T = 300$[K] as an approximation in practice. The temperature of the room for measurement is maintained to be $\sim 20[C^\circ]$ and hence the precision of the approximation is good enough. ($\sim 2\%$)

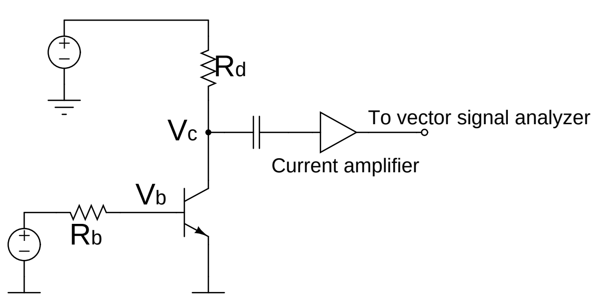


Figure 4.13: The schematic view of BJT's noise measurement setup.

(Agilent 89410A) which measures the noise spectrum, and a semiconductor parameter analyzer (HP 4156A) which measures the transistor parameters. A picture of the system is shown in the figure 4.14. An example of the noise measurement result is shown in figure 4.15. Here measured noise spectra have the common shape which consists of the constant noise floor and the $1/f$ noise component. The $1/f$ noise is prominent especially in the noise spectrum of $R_b = 10[\mathrm{k}\Omega]$. There is a decrement of noise power in the region $10^5[\mathrm{Hz}] > f$. This effect is common in all transistors we have measured. The decrement is probably from the bandwidth limitation of our measurement equipment. Noise floor and the $1/f$ noise increase as R_b increases, which result suits the expectation from equation 4.9. There exists the extra noise at the frequency region of $f \sim 10^3[\mathrm{Hz}]$ and $f \sim 10^5[\mathrm{Hz}]$. This feature is also common in all transistors. This extra noises are probably from our equipment, because they cannot be easily interpreted to belong to some of the noise components in the equations 4.2 and 4.3, and their contribution decreases as the noise increases, or R_b increases. In order to avoid the extra noise component, we use the values in the region of $5 \times 10^3 < f < 10^4[\mathrm{Hz}]$. We use the average of the noise power in the region as the measured value of ${i_c}^2/\Delta f$.

To calibrate the equipment, the following two measurements are carried out.

The measurement of $\mathbf{R_{bmes}(0)}$ with various $\mathbf{I_c}$ values.
: For varying I_c, the noise power ${i_c}^2/\Delta f$ also varies according to the equation 4.9. If the measurement is valid, we still get the same $R_{b\mathrm{mes}}$ from the equation 4.10. On the other hand, we get various $R_{b\mathrm{mes}}$ values if we have effects by the extra noise from outside of transistors.

The measurement of $\mathbf{R_{bmes}(R_b)}$ with various $\mathbf{R_b}$ values.
: This is the direct calibration, since R_b has the same effect as $R_{bb'}$, which fact the equation 4.9 shows. By the determination of the function shape of $R_{b\mathrm{mes}}(R_b)$, we can estimate the $R_{bb'}$ from the $R_{b\mathrm{mes}}(0)$.

We have done the measurements using 2SC5011. Figure 4.16 shows the result of the first measurement. $R_{b\mathrm{mes}}$ is not constant at the region $I_c \lesssim 2[\mathrm{mA}]$ in the figure 4.16. The accuracy of measurement with small I_c is lower than that with large I_c since we have large noise power ${i_c}^2/\Delta f$. Because we cannot measure with too large I_c due to the possible oscillation, we have decided to measure noise with $I_c = 2.5[\mathrm{mA}]$, where $R_{b\mathrm{mes}}$ is constant independent of I_c.

Figure 4.16 shows the result of the second measurement and that there is nearly linear relation between $R_{b\mathrm{mes}}$ and R_b with an error of $\sim 2[\Omega]$ for $R_{b\mathrm{mes}}$. The inclination of the fitted line is 0.44. From the result, we can roughly estimate $R_{bb'}$ by $R_{b\mathrm{mes}}$ as follows:

$$R_{bb'} = R_{b\mathrm{mes}}/0.44 \tag{4.13}$$

We have determined the following measurement conditions, taking account of the results of the measurement above.

- We measure three samples for each transistor type.
- For the first sample of the three, we set I_b so that I_c becomes about 2.5mA, in an accuracy of ~ 0.1mA. We set the same I_b for the second and third samples. This way is appropriate because forward current gain h_{FE} does not differ much in the same type of transistor.
- We measure noise with $R_b = 0[\Omega]$ and estimate $R_{b\mathrm{mes}}$ for all three samples.
- We measure noise with $R_b = 100[\Omega]$ for one of the three samples and check the size of $1/f$ noise.

The result of the $R_{bb'}$ measurement is shown in figure 4.18 with the measured value of $1/f$ noise coefficient $K_{1/f}$. Though the value shown is the raw value of the measurement, $R_{b\mathrm{mes}}$, the relation of large and small is the same as that of $R_{bb'}$. In measurement of $1/f$ noise, we have measured only one sample for each transistor type, have assumed the indices a and b to be $a = b = 1$ in the $1/f$ noise term of equation 4.9, and have assumed $R_{bb'} = 40[\Omega]$ for all transistors. Thus, the values of $K_{1/f}$ are reliable only in the *order of magnitude* accuracy. The accuracy is, however, still satisfactory, since our purpose of $K_{1/f}$ measurement is only to confirm that $K_{1/f}$ is small enough. Here, $K_{1/f} \lesssim 10^{-14}$ is small enough.[v]

[v] $K_{1/f} \sim 10^{-14}$ corresponds to $\sim$100kHz of the critical frequency, above which the $1/f$ noise is negligible. The derivation of the values in the table 4.5 shows the way to calculate the critical frequency.

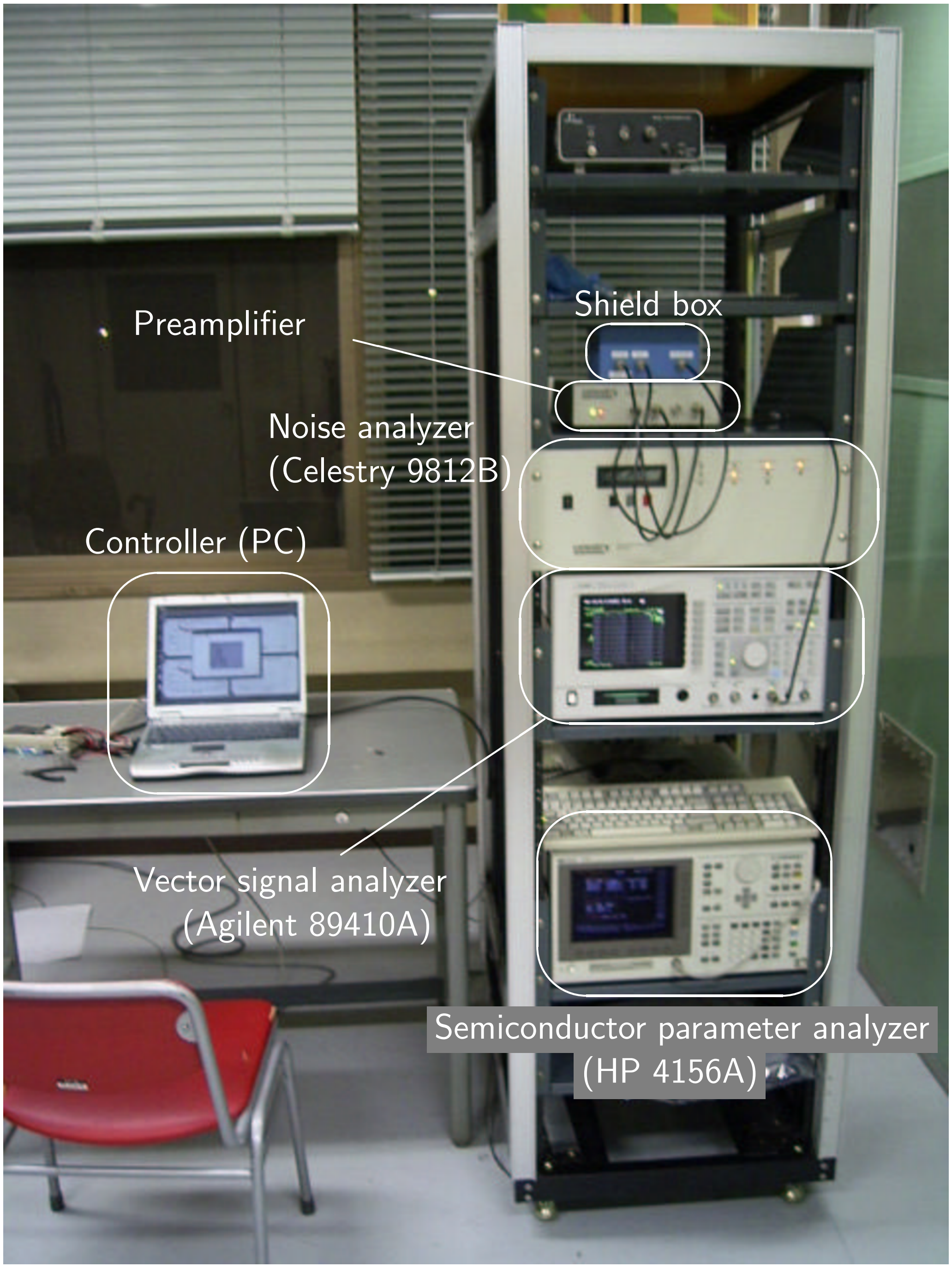


Figure 4.14: A picture of the NoisePro system.

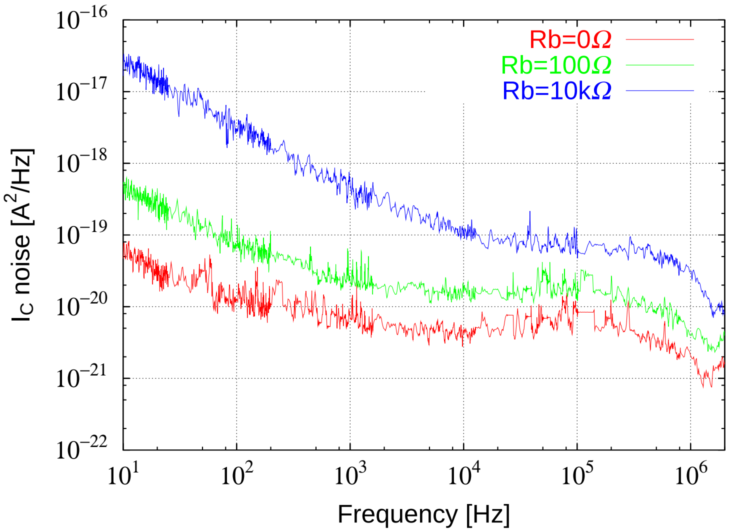


Figure 4.15: The result of noise measurement: 2SC5319.

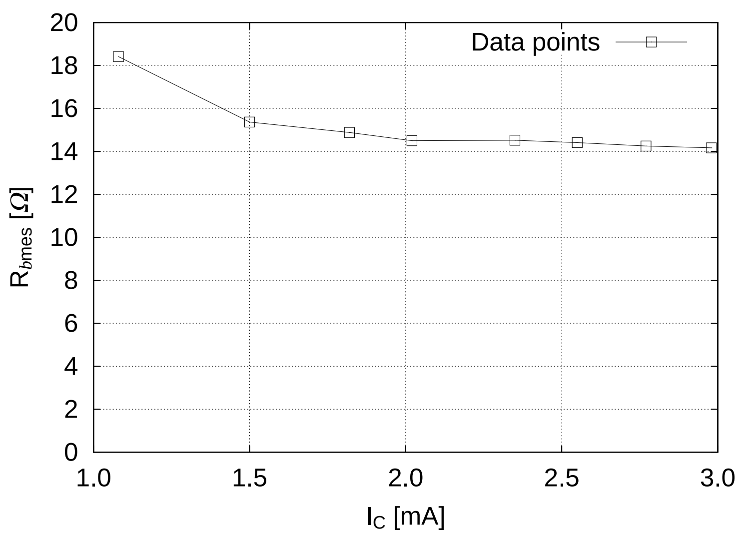


Figure 4.16: I_c v.s. $R_{b\mathrm{mes}}$.

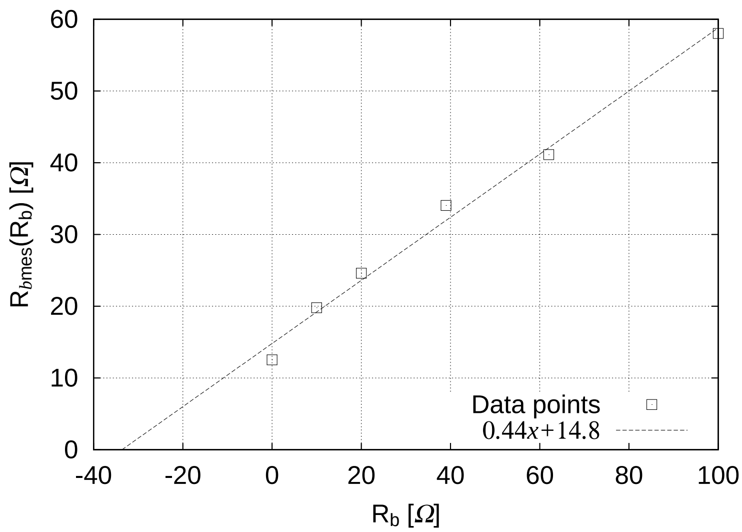


Figure 4.17: R_b v.s. $R_{b\mathrm{mes}}(R_b)$.

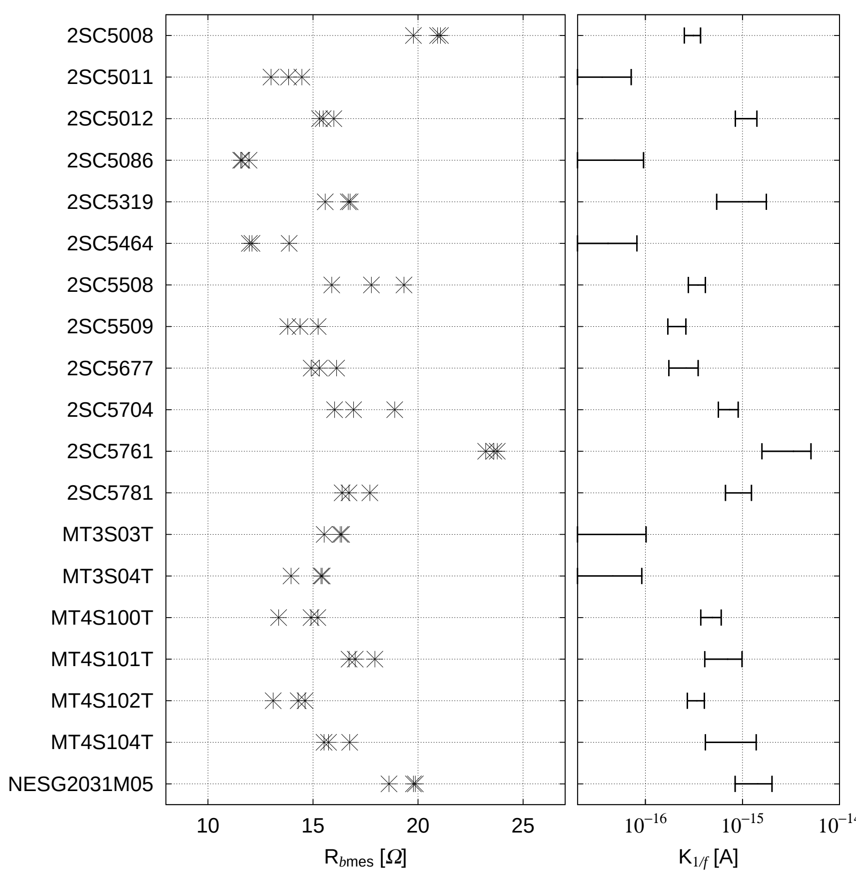


Figure 4.18: The measurement result of $R_{b\mathrm{mes}}$ and $K_{1/f}$.

Consideration on the measurement result

Figure 4.18 shows that the best transistor is 2SC5086, with its small $R_{bb'}$ and $K_{1/f}$. Thus, we decide to use 2SC5086 as the input transistor. Here, $\sim 25\Omega$ is the estimated value of the $R_{bb'}$ of 2SC5086.

4.2.4 Performance evaluation

Here we fix our amplifier configuration. We use folded cascode configuration, and set Cf=2pF, Rf=20kΩ. As the input transistor, we use 2SC5086. The circuit is shown in the figure 4.19, and the pictures of the preamplifier are shown in the figure 4.20.

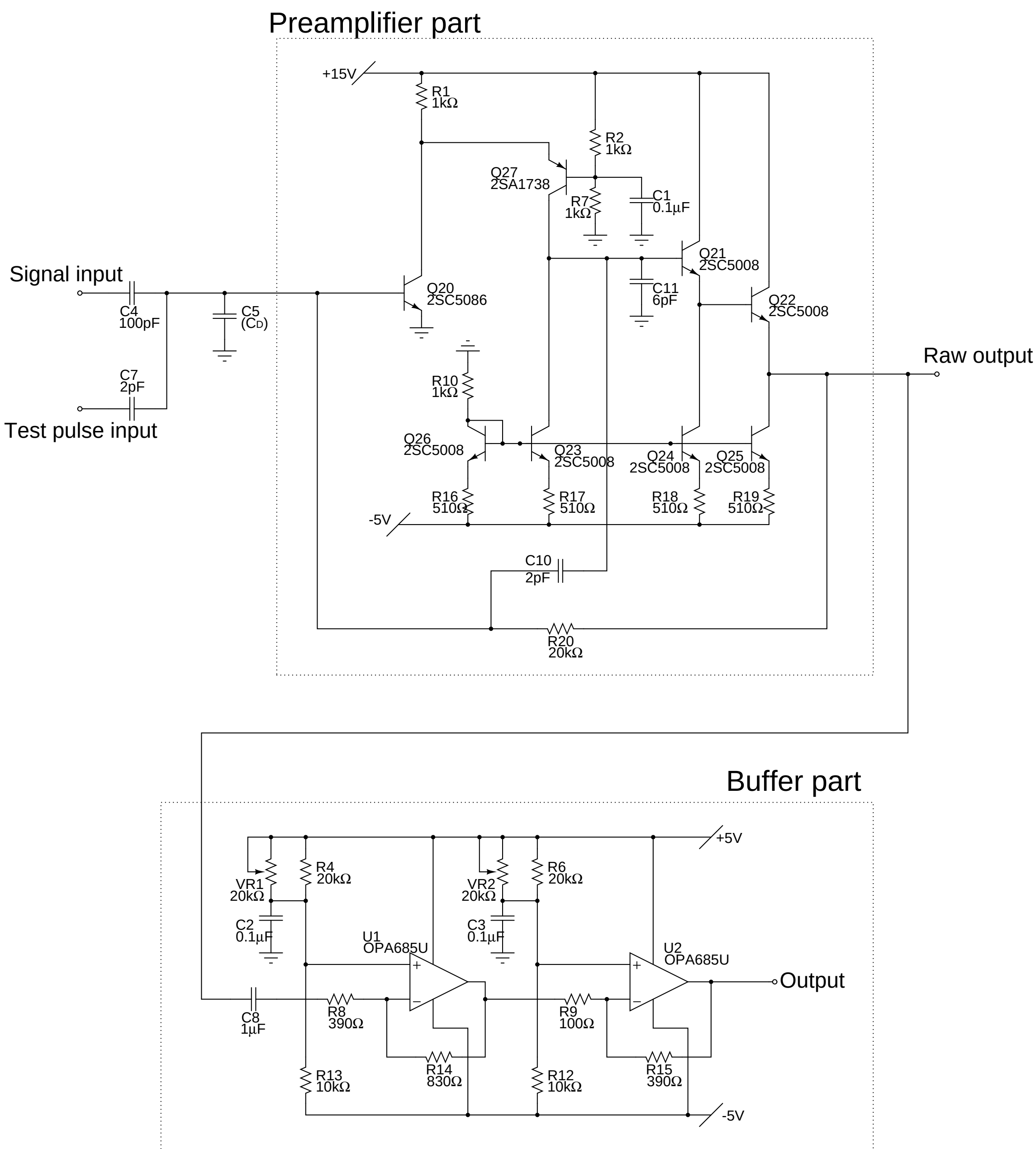


Figure 4.19: The final preamplifier configuration.

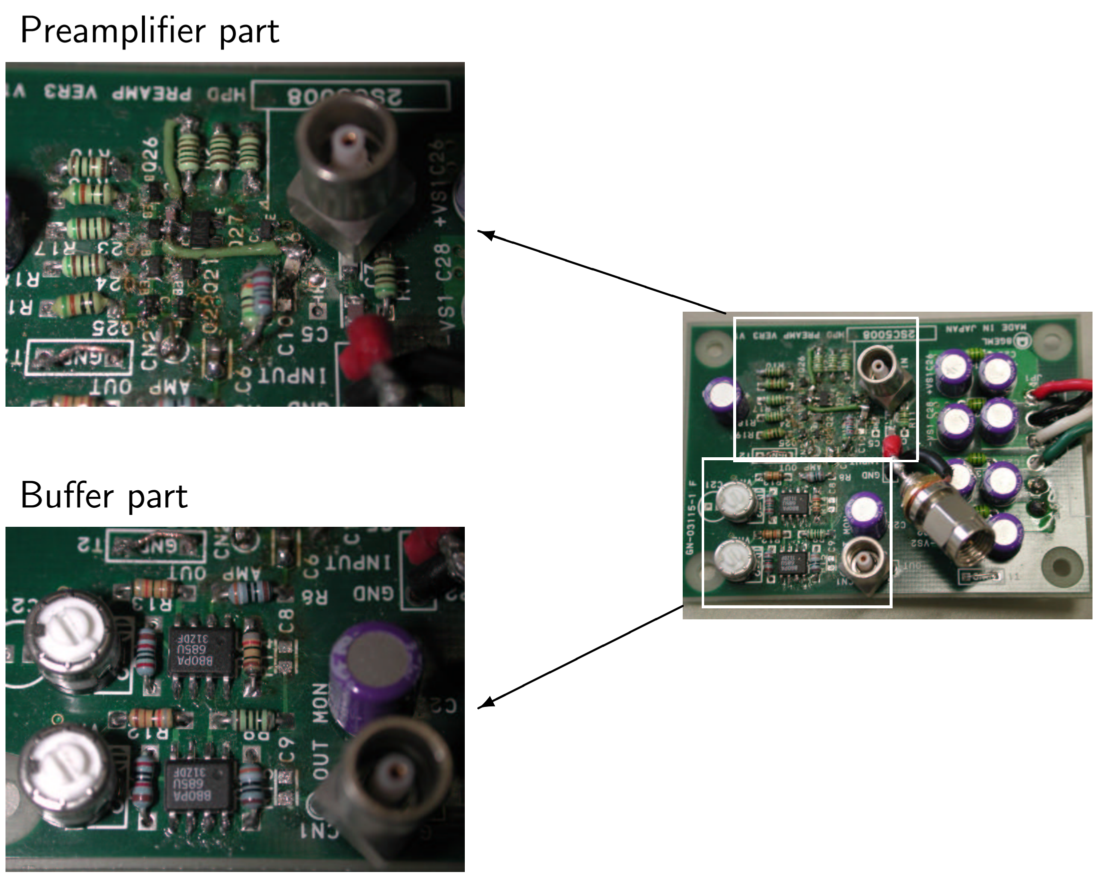


Figure 4.20: The pictures of the preamplifier.

Noise expectation of the preamplifier

First, we calculate the expected noise value of our amplifier using the transistor noise measurement in the previous section and the equations 4.2, 4.3 and A.20. We assume here $R_{bb'} = 25\Omega$, $I_c = 1$mA, $I_b = 25\mu$A and $g_m = 40$mS for the input transistor, and $C_f = 2$pF and $R_f = 20$kΩ for the feedback impedance. The detector capacitance $C_D = 70$[pF] is also assumed.

Then the expectation value of the noise power spectrum is calculated. The result is shown in figure 4.21. By integrating[vi] the noise power in the figure 4.19, we have estimated the noise to be ~ 0.34mV. Since the expected signal height for 10^5electron input charge[vii] is ~ 8mV, the Equivalent Noise Charge (ENC) is calculated to be $10^5 \times 0.34/8 \sim 4,300$[electron]. In the same way, we evaluate ENC for various detector capacitances. We also calculate the noise for setting the bandwidth of noise power integration to be 20MHz. The result is shown in the figure 4.22.

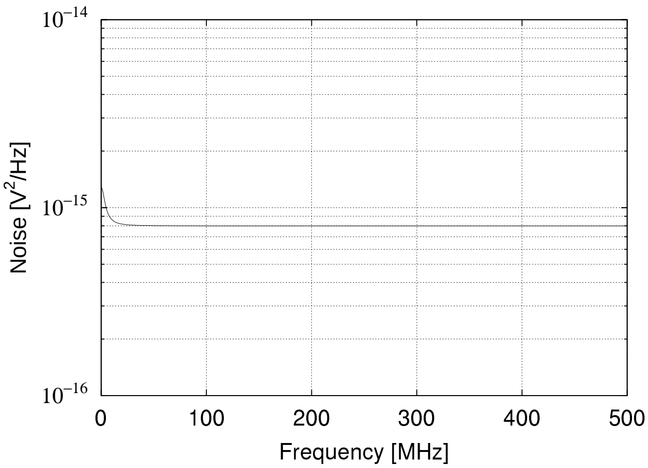


Figure 4.21: The expected noise spectrum.

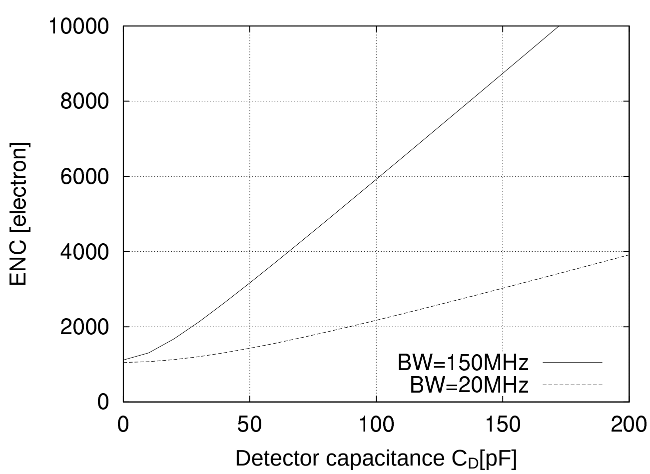


Figure 4.22: The expected relation between ENC and detector capacitance.

[vi]We limit the region of integration less than 150MHz, which corresponds to the bandwidth of our oscilloscope.
[vii]This corresponds to the single photoelectron when gain of HPD is 10^5.

Noise measurement

Then we measure the amplitudes of noise and signal height of our preamplifier. Note that we have buffer amplifiers to drive 50Ω cable and to avoid inessential noise sources. Consequently, only the relative values of amplitudes of noise and signal matters.

First, we measure the noise from the buffer part in the figure 4.19 with the preamplifier part disconnected, and the noise is 0.80mV in RMS. Second, we measure the noise with the preamplifier part connected, and the noise is 1.02mV for $C_D = 0$. The two measurements are done with an oscilloscope (Tektronix TDS5054A), with the bandwidth setting of 150MHz. By the measurements, the noise from the preamplifier is calculated to be $\sqrt{1.30^2 - 0.80^2} = 1.02$[mV]. On the other hand, the signal height is 95mV for a 18.5mV test pulse. Since the capacitance C_i in the test pulse input is 2pF, the 18.5mV test pulse corresponds to 18.5[mV]×2[pF] = 2.31×10^5[electron]. Thus, the ENC(Electron Noise Charge) of this amplifier is $2.31 \times 10^5 \times 1.02(\text{noise})/95 = 2.48\times10^3$electron.

We also measure the amplitudes of noise and signal for various C_D values and calculate ENC for each measurement. The result is shown in the figure 4.23. Figure shows the obtained ENC results for the 150MHz and 20MHz oscilloscope bandwidths as functions of noise on detector capacitance C_D. Noise for $C_D = 70$[pF] is about 3,500[electron] and is acceptable for the HPD gain of $\gtrsim 5\times10^5$. This measurement result is consistent with the expectation in the figure 4.22

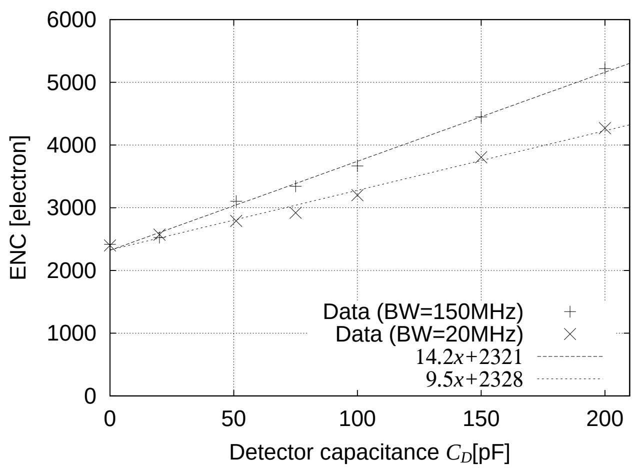


Figure 4.23: The result of preamplifier noise measurement: C_D v.s. ENC.

The speed

The raw waveform of our amplifier is shown in the figure 4.24. The rise time is about 4ns and is limited by the leading edge of the input signal. The width of the signal from HPD is larger than 5ns and thus the rise time of 4ns meets our requirement. Figure also shows that there is no ringing or oscillation, that is, our amplifier has good stability.

The dynamic range

In order to check the dynamic range and the linearity of our amplifier, we check the relation between input charge and amplitude of output.

We put a positive [viii] step pulse into the terminal *test pulse in* in the figure 4.19. Here the input charge Q is obtained by multiplying the test-pulse amplitude V_{tp} and the input capacitance C_i, $V_{\text{tp}} \times C_i$. We set the leading edge of the test pulse about 40ns, which is the signal width of the test-pulse input charge. Note that this is slow compared with that of HPD. We chose the slow signal to divide the effect from its slew rate and dynamic range. By choosing the slow signal as input, we can measure dynamic

[viii]A positive step pulse generates a positive charge signal. Signal polarity of Avalanche Diode (AD) is also positive.

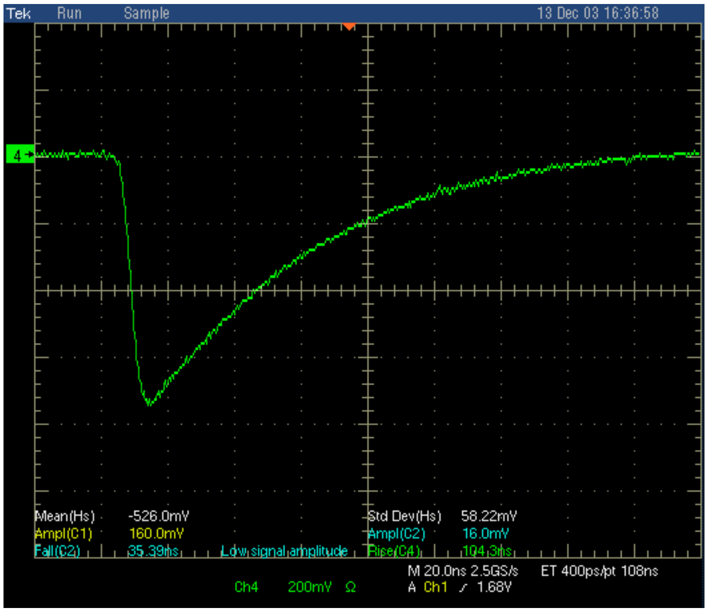


Figure 4.24: The raw waveform of our preamplifier. The horizontal scale is 20ns/div, and the vertical scale is 200mV/div.

range without the effect from the limited slew rate of our amplifier. The gain of the amplifier is, however small for the slower input signal.

The measured result is shown in the figure 4.25. We obtain the dynamic range of 0 to -5V. We fit the data in the region of $0 < Q < 4.5{\times}10^7$[electron] with a linear function in the upper graph, and obtain the relation between input charge Q[electron] and output voltage V_o[V] as:

$$V_o = -4.1 \times 10^{-8} Q \tag{4.14}$$

In order to check the linearity, we also plot the difference between measured output voltage and the calculated output voltage with above linear relation, in the figure 4.25 (the lower graph). We obtain the good linearity in the region of $0 < Q < 4.5{\times}10^7$[electron] ($0 > V_o > -1.8$[V]). In this region, the deviation from the *linear* is 2.1mV in RMS. This value corresponds to $\sim$10bit linearity. The linearity gets worse in the region of $-1.8 < V_o$[V], but is not so disastrous.

Since the gain with slow($\sim$40ns) signal is $\sim$40% smaller than that with fast($\sim$5ns) signal, we have to rewrite the equation 4.14 in order to apply the result to the real HPD signal, as follows:

$$V_o = -6.8 \times 10^{-8} Q \tag{4.15}$$

where we calculate the gain using the original value in the equation 4.14 and the ratio of 40% as $4.1 \times 10^{-8}/(1 - 0.4) = 6.8 \times 10^{-8}$. The dynamic range, which corresponds to $0 > V_o > -5$[V], is about 740p.e.(photo-electron) equivalent, and the good linearity region, which corresponds to $0 > V_o > -1.8$[V], is about 260p.e. equivalent. Here we have assumed 10^5 as the gain of HPD.

The slew rate

Since the slew rate can limit the linearity and the dynamic range, we also measure it. Keeping the leading edge of the input step pulses constant, we have measured the relation between the amplitude and the rise time of the output signal. We have set two values as the leading edge, $\sim$2.5ns(fast) and 6ns(slow).

Figure 4.26 shows the result. We write two lines in the figure 4.26, $T_r = 4.2{\times}V_o$ and $T_r = 6{\times}V_o$ where T_r and V_o are rise time in nanoseconds and output voltage in volts, respectively. $T_r = 4.2{\times}V_o$

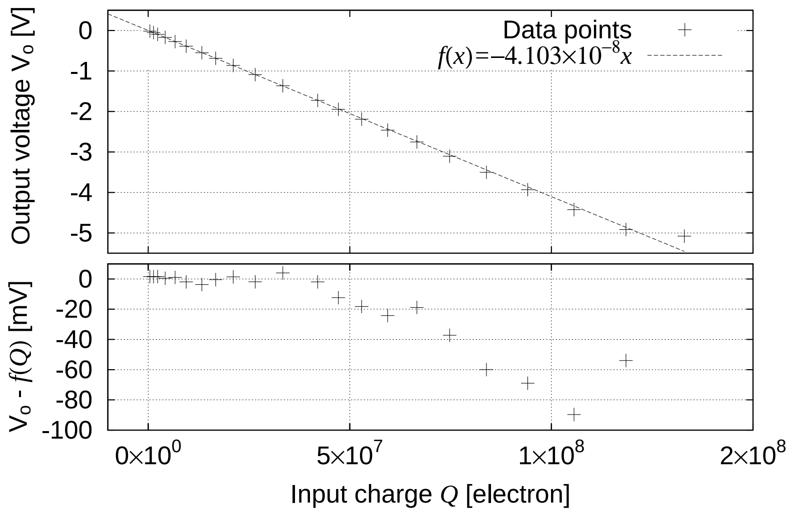


Figure 4.25: The measured dynamic range and the linearity of our amplifier.

is a roughly fitted line to the data points at which rise time is proportional to input voltage.[ix] This means that the inclination corresponds to the slew rate of our amplifier. From this line, the slew rate is $1/4.2[\mathrm{V/ns}] = 240[\mathrm{V}/\mu\mathrm{s}]$.

We roughly write the line $T_r = 6\times V_o$ so that the rise time is constant, independent of output voltage, in the region of $T_r \gtrsim 6\times V_o$. Ideally, the rise time should not depend on amplitude, since linearity gets worse if rise time varies depending on signal amplitude. The timing resolution also deteriorates in the case. Thus, we have to use the preamplifier in the region where the rise time is independent of amplitude. To avoid the variation of the rise time, we have to use the amplifier in the region $T_r > 6\times V_o$. The signal width of HPD is about 7ns and thus a separate signal height should be less than $\sim$1.2V, which corresponds to $\sim$180p.e. Note that the slew rate is independent of the dynamic range, e.g., two sequential 180p.e. signals with a 10ns interval are still acceptable.

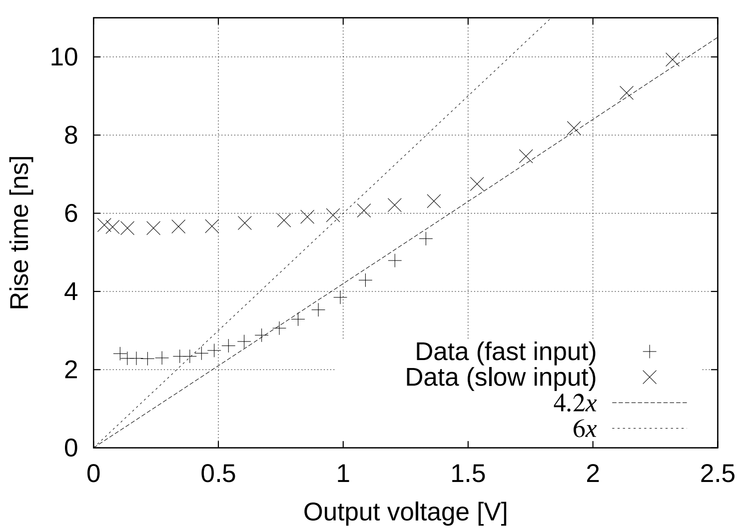


Figure 4.26: The measured relation of rise time v.s. the amplitude of the output signal.

[ix] We have used the data points of $V_o > 1$[V] ($V_o > 1.6$[V]) for the fast (slow) test pulse input.

On the improvement of the dynamic range and the slew rate

We have checked which part of the preamplifier limits the dynamic range and the slew rate, and we have figured out that the cause of the dynamic range limitation is the bias point of Q27 in the figure 4.19, and the cause of the slew rate limitation is the emitter followers, Q21 and Q22 in the figure. We can optimize the bias currents of these nodes, and thereby we can improve the dynamic range and the slew rate further.

Summary

We summarize here the performance of our amplifier.

- The noise for $C_D = 70$pF is $\sim 3,000$[electron] and it is good enough[x].
- The rise time is $\lesssim$ 4ns and it meets our requirement.
- The dynamic range is $\sim$ 700p.e. and it meets our requirement.
- The slew rate is about 240[V/μs].
- We expect the dynamic range and the slew rate to improve by the further optimization of bias currents.

4.3 Waveform sampler

As described in section 4.1, waveform sampling with fast speed ($\sim$ 1[GHz]) is essential to achieve the high timing resolution and the short dead time. This section is to show our strategy to realize the fast waveform sampling. We have two choices in sampling devices: Flush ADC (FADC) and Analog Memory Cell (AMC). First, we describe the principle of AMC. Then we discuss about the merits and the demerits of FADC and AMC. Finally, we show our R&D status of AMC.

4.3.1 Principle of AMC

The simplest description of our AMC's principle is *the analog version of Time Memory Cell* (TMC[35]). The schematic drawing of the AMC is shown in figure 4.27. The mechanism used in the analog memory part is called "Switched Capacitor Array(SCA[36])." Whereas SCA usually needs an external clock pulse to turn on/off the switch in each stage(cell) of the analog memory, our AMC needs only one trigger to sample a waveform with all stages(cells), using gate delays like TMC.

Each stage of the AMC consists of an analog switch, a cell capacitor and a delay buffer. At first, all analog switches are on. If the trigger is issued, the switches are turned off. The timings of turning off are delayed by the delay buffers. When an analog switch of a stage is turned off, voltage of analog input at the timing is *memorized* by the cell capacitor of the stage. Consequently, with triggered time T_t and the delay time of delay buffer t_d, the first cell capacitor memorizes voltage of analog input of the time T_t, the second cell capacitor memorizes that of the time $T_t + t_d$, ... and the nth cell capacitor memorizes that of the time $T_t + nt_d$. As a result, we can sample the waveform of analog input by frequency of $1/t_d$ and length of Nt_d where N is the number of the stages. The sampled waveform will be read out by using the analog multiplexer and probably undergo AD conversion. In our plan, t_d will be set $0.5 \sim 1$[ns] and N will be $512 \sim 1024$.

The essence of the device is that one can execute fast ($\sim$ 1[GHz]) waveform sampling without a fast clock. In other words, one can make the speed of AD conversion much slower than that of waveform sampling. For example, one can sample a waveform in 1GHz speed with "AMC + 10MHz FADC" configuration. Then cost and power dissipation per channel will be reduced compared with the simple solution: waveform sampling with 1GHz FADC.

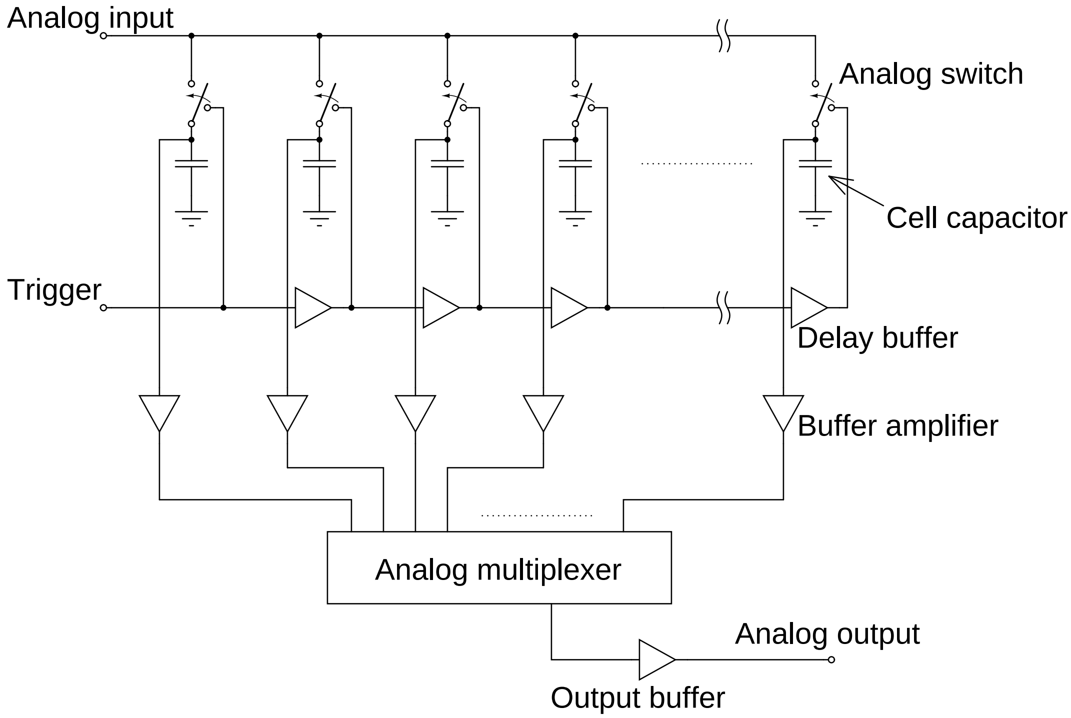


Figure 4.27: Schematic view of AMC.

Table 4.7: The comparison table of FADC and AMC

	AMC (+10MHz FADC)	FADC
speed	1∼2 GHz	0.5∼1 GHz
power/channel	$\lesssim$0.5W	1∼2W
cost/channel	$\sim 10^4$yen	$5 \times 10^4 \sim 10^5$yen
depth(# of stages)	512∼1024	$\gtrsim 10^4$
linearity	$\gtrsim$ 10 bit	∼ 8 bit

4.3.2 Comparison of FADC and AMC

Comparison of FADC and AMC is summarized in the table 4.7.

The merits of AMC are its low power dissipation and cost per channel. Also, since it is made as ASIC, it can be included in the larger system ASIC with other functionalities like preamplifier and/or ADC, for example. The demerit is its limited depth, though we can get rid of the demerit to some extent by connecting multiple AMCs, in daisy chain pattern.

The greatest merit of FADC is that it is already available. Also, its depth is much larger than that of AMC. The demerits are its large power dissipation and cost per channel. Its implementation density of channels is also limited. Though FADC with better linearity than the 8 bit is also available, it has larger cost and power dissipation[xi] than that with 8 bit linearity.

4.3.3 The R&D of AMC

We have developed AMC ASIC(Application Specific Integrated Circuit). We made the first prototype of AMC ASIC with 16stages(cells). The picture and the circuit pattern of our prototype are shown

[x]Note that S/N is expected to improve with digital filtering in the later stage.

[xi] In order to achieve 10 bit linearity, for example, the cost and the power dissipation become 2 ∼ 4 times larger than that with 8 bit linearity.

in the figures 4.28 and 4.29. We have two types of the prototype: AMC, and AMC with an analog multiplexer(AMCM). In this section, we report the result of our evaluation of the prototypes and then show the plan for the next version AMC ASIC.

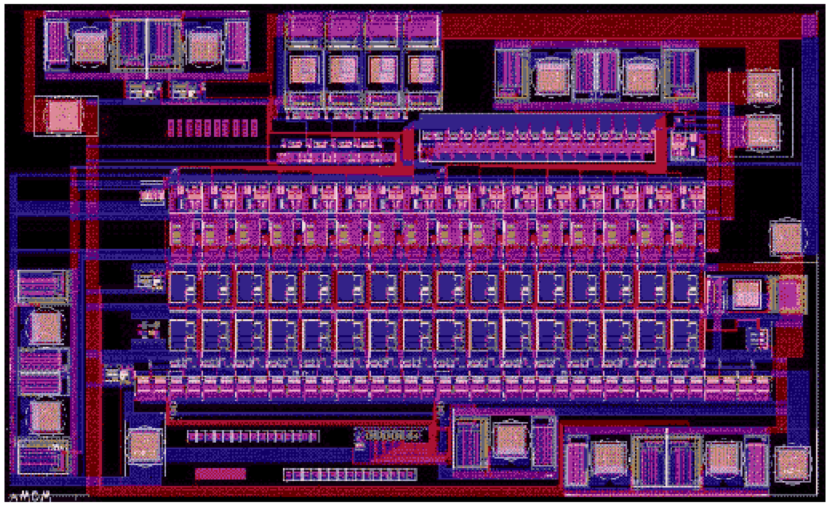


Figure 4.28: The circuit pattern of the prototype AMC ASIC.

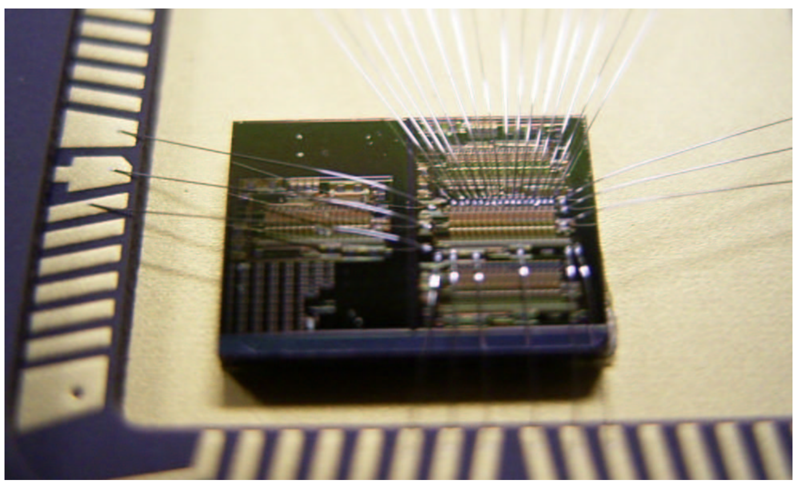


Figure 4.29: The prototype AMC ASIC chip.

Basic functionality

First, we check the basic functionality of AMC and AMCM: if they can hold analog input. We input some signals: leading edges and trailing edges of step pulses, for example. Then we confirm that AMC and AMCM can hold the input signal properly. The example is shown in the figure 4.30.

We also confirm that the prototypes can keep up with the fast input signals, which means that the bandwidths of them are wide enough.

Stage by stage (cell by cell) dispersion

Putting in the constant voltage from analog input, we sample the level input using AMC and AMCM. Then we check the uniformity of the output level from each stage(cell).

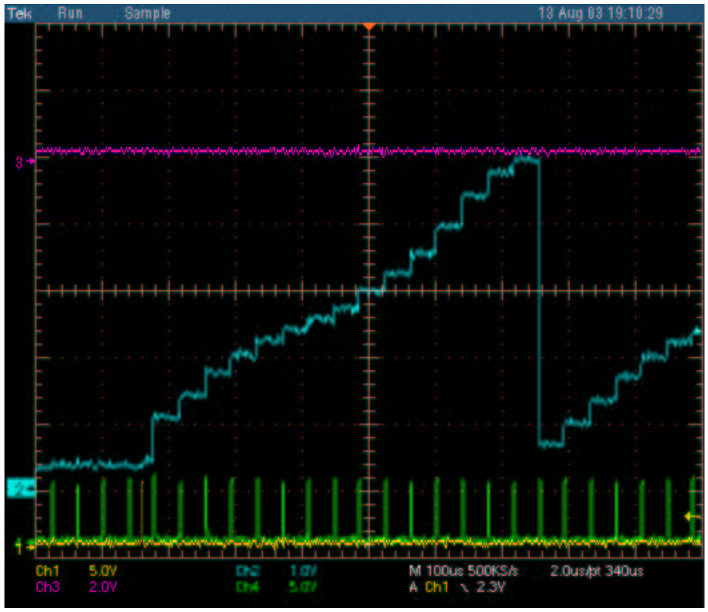


Figure 4.30: The leading edge waveform of a step pulse sampled using AMCM. The blue line is multiplexed output of AMCM and green line is the clock for the analog multiplexer. One can see blue line has 16steps, which correspond to multiplexed analog output for 16stages(cells) of AMCM. The horizontal scale is 100μs/div and the vertical scale is 1V/div for the blue line.

The measured result for one of the AMC samples is shown in the figure 4.31. The dispersion is ∼400mV, and is much worse than our requirement. The all other samples of AMC and AMCM also have the similar behavior. The pattern of the dispersion varies depending on the height of input level and hence it relates to the deterioration of linearity, as discussed later. We also check the dispersion without turning off the switches in each stage. In this case, the dispersion is small enough (∼5mV).

Linearity

We evaluate the linearity of each stage in AMC and AMCM. Putting in a level voltage from analog An input, we measure the level of output, varying the level of input. One of the results for AMC is shown in the figure 4.32. It does not satisfy our goal of the 8∼10bit equivalent linearity. As mentioned above, we consider the deterioration of linearity relates to the stage by stage dispersion because the pattern of deterioration varies stage by stage.

We also check the linearity when without turning off the switches in each stage, in the same way as we have done for the dispersion measurement. The linearities are 10bit equivalent and are good enough for both AMC and AMCM. Thus, we have confirmed the linearities of buffer amplifiers are satisfactory.

Delay interval

We measure the delay interval of delay buffers (see figure 4.27).

We measure a delay for a specific stage of AMC. Figure 4.33 schematically shows the method to measure it. *Delay A* is the delay of the specific stage, which we want to measure. It is a sum of delays yielded by the delay buffers before the stage. *Delay B* is the delay between the edge of analog input and the trigger, which we can control. We input analog signal and trigger as in the left side of the figure 4.33. The point sampled by the stage is determined by *delay A* and *delay B*. *Delay A* is constant for the specific stage. The sampled point moves along the analog input signal as *delay B* changes. We measure the output signal level of the stage changing *delay B*, and plot the levels as a function of *delay B*, which corresponds to the waveform of input, as in the right side of the figure 4.33. The edge of the sampled

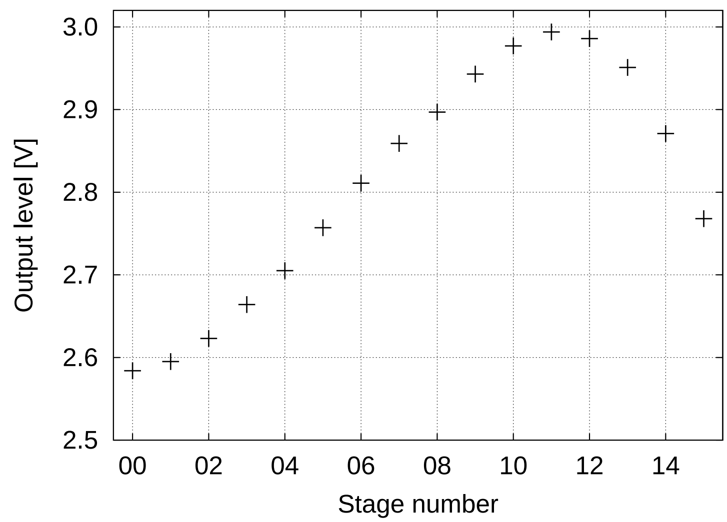


Figure 4.31: A measurement result of stage by stage dispersion.

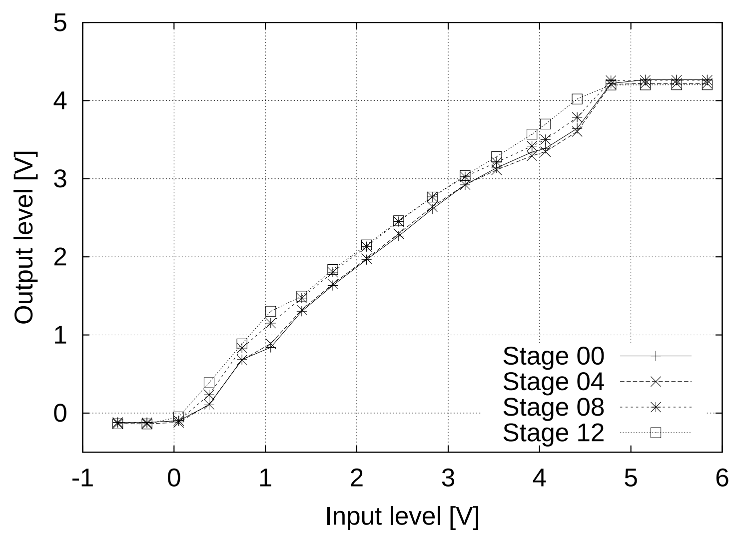


Figure 4.32: A measurement result of linearity.

waveform in the right side of the figure 4.33 corresponds to *delay A*[xii], which we want to measure.

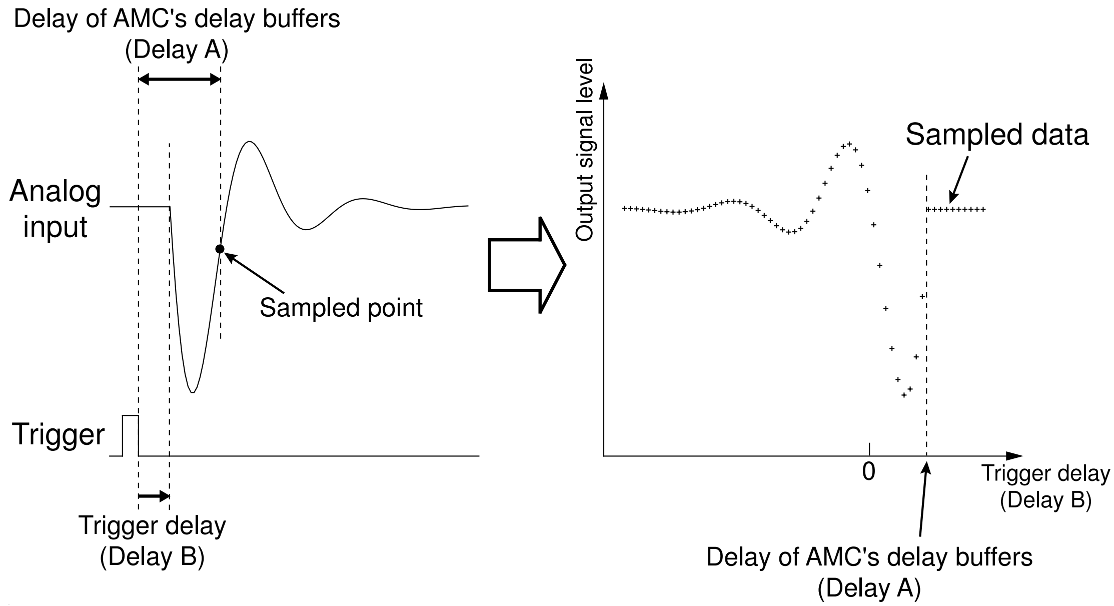


Figure 4.33: The method to measure the delay of a stage.

We measure the buffer delays(*delay A*) of all stages in this way. We calculate the positions of the edges by fitting the part of waveforms with a parabolic function. The result for one of the AMC samples is shown in the figure 4.34. Here we fit the data points with a straight line. The inclination of the line corresponds to the delay interval yielded par one delay buffer. Though the data points have a common offset, it does not matter since only the inclination of the fitted line is essential.

By the inclination, we conclude the delay yielded by one delay buffer is ∼500ps. The dispersion of the delays is smaller than the measurement precision(∼100ps)[xiii], since the reduced chi-square, χ^2/n.d.f., of the fitting is 1.6.

Speed of readout(for AMCM)

The speed of readout is limited by the speed of the analog multiplexer. We measure the settling time of analog multiplexer, which determines the speed. The settling time is ∼100ns and hence the maximum readout speed is a few MHz.

Summary and plan

The following items are satisfactory.

- The basic functionality.
- The delay interval (both the length and the dispersion).

The following items need further improvements. First two items are highly correlated.

- The stage by stage dispersion.
- The linearity.
- The readout speed.

[xii]Imagine what happens when *delay B* equals to *delay A* in the left side of the figure 4.33.
[xiii] The dispersion is measured more precisely to be ∼ 50ps using a TMC ASIC with the same delay buffer circuit.

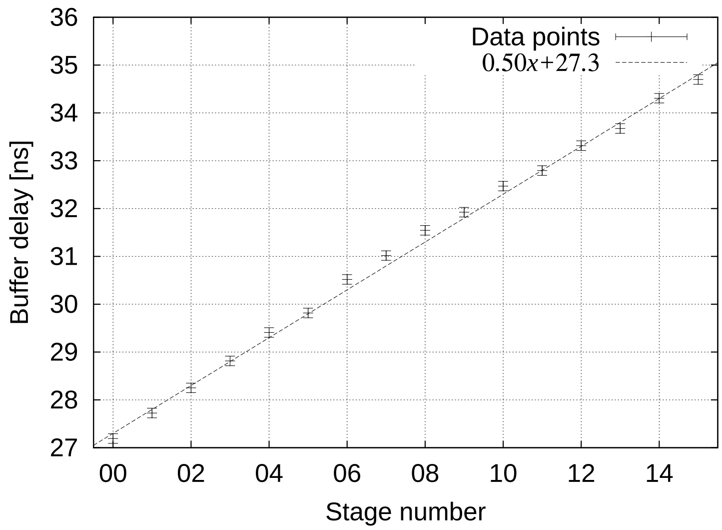


Figure 4.34: The measurement result of delay yielded by delay buffers.

We consider the bad linearity is due to parasitic capacitances in the switch of each cell. The effect of the capacitances seems to get larger since the buffer amplifiers are voltage followers.

We are planning the next version of AMC ASIC based on the evaluation above. The modifications from the evaluated version are listed in the followings. We will change the buffer amplifiers from the voltage followers to inverting amplifiers, which have low input impedance, and thereby we expect the linearity and the readout speed will be improved.

- Increase the number of stages to 512∼1024.
- Implement functionality to change the delay interval(0.5ns∼1ns).
- Decrease the stage by stage dispersion(<50mV) and improve the linearity(8bit∼10bit).
- Improve the readout speed to be 10MHz.

4.4 Digital signal processing

In this section, we describe how we process the waveform data sampled by the waveform sampler.

Here we note our use of a word "signal." In this section, We use the word "signal" in two meanings. One is as an antonym of *noise*. In this use, we use the word like "*signal*," that is, use it as an uncountable word and write in *italic* characters. Another is as an input or output waveform of digital filters. In this use, we use the word like "a signal," as a countable word, or more clearly "a waveform." A signal may not consist of *signal*, but consist of *noise* only.

4.4.1 Design concept

We divide the digital signal processing into two subprocesses: 1.filtering a signal waveform using digital filters and 2.deduce the amplitude(Q)[xiv] and the timing(T) of the signal.

In this section, first we introduce the principle of *digital filters* and some concepts often used in digital filtering. Then we mention about the deduction of Q and T, and finally, we mention a little about the implementation of filters.

[xiv] We call the amplitude "Q" since it corresponds to the amount of input charge.

Roles and merits of digital filters

As described in the section 4.1, digital filters in our readout system replace the shaping amplifier in general readout systems. Thus, the roles of digital filters are 1.to maximize the signal to noise ratio and 2.to avoid signals to pile up.

These functionalities can also be implemented using analog filters, since both analog filters and digital filters are linear transformations from an input signal waveform to an output signal waveform, as shown in the figure 4.35.[xv] The analog filters are cheap[xvi] and fast. On the other hand, digital filters are vastly superior in the level of performance that can be achieved, that is, we can make, with digital filters, the filters that cannot be realized with the analog filters. Another merit of the digital filters is that we can easily change the property of the filter, e.g., the bandwidth, by changing the filter kernel, which we mention in the following. The demerits are that the digital filters can take significant calculation time, can consume larger power, and can be more expensive. We can remove the demerits, however, with the FPGA technology, where the remarkable progress has been accomplished in recent years.

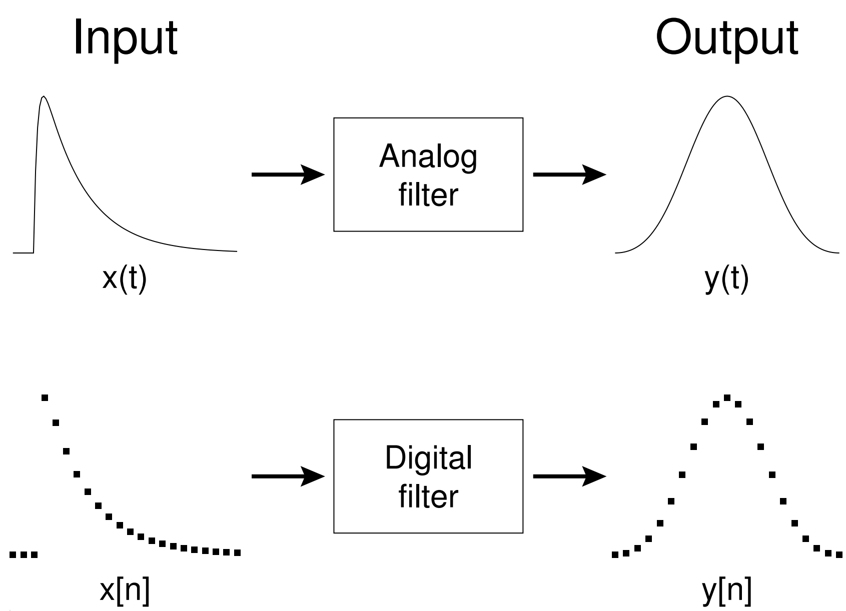


Figure 4.35: A schematic picture of analog and digital filters.

Note that the primary reason we adopt the waveform sampler in our readout system is not for adopting digital filters but for high rate tolerance(see section 4.1). We cannot avoid, therefore, the waveform sampling even if we adopt analog filters. In other words, we cannot utilize one of the largest merits of analog filters that we do not need waveform samplers, even when we adopt analog filters.

Principle of digital filters

We adopt FIR(Finite Impulse Response) filters. In FIR filters, an output signal $y[n]$ for an input signal $x[n]$ can be calculated as:

$$y[i] = \sum_{j=0}^{N-1} x[i+j] \times h[N-j-1] \tag{4.16}$$

[xv]Though there is a difference that the waveforms are continuous for analog filters while they are discrete for digital filters, it is not essential.

[xvi]Because they do not need waveform samplers and FPGA(or DSP).

where $h[n]$ is a series that defines the property of the filter and N is the length of $h[n]$. We call $h[n]$ *filter kernel*. We can symbolically rewrite the equation 4.16 as follows, using the symbol of convolution $\otimes$:

$$y[n] = (h \otimes x)[n] \tag{4.17}$$

Then we apply Discrete Fourier Transformation(DFT) to the equation 4.17. By DFT, a signal $x[n]$ is changed to a complex spectrum $X[k]$. The input to DFT, $x[n]$, is said to be in the *time domain*, while the output of DFT, $X[k]$, is said to be in the *frequency domain*[xvii]. Following the standard notation, we represent time domain signals using lower-case letters, such as $x[n]$, $y[n]$, and $z[n]$, and corresponding frequency domain signals (spectrums) using corresponding upper-case letters, such as $X[k]$, $Y[k]$, and $Z[k]$. $x[n]$ and $X[k]$ are related by DFT and inverse DFT as:

$$X[l] \equiv \sum_{j=0}^{N-1} x[j] e^{-2\pi i \cdot jk/N} \tag{4.18}$$

$$x[j] = \frac{1}{N} \sum_{l=-(N-1)/2}^{(N-1)/2} X[l] e^{2\pi i \cdot jk/N} \tag{4.19}$$

$$\Big(\quad X[-l] = X[N-l] \quad \Big), \tag{4.20}$$

where N is the length of the signal $x[n]$. Though we assume N is odd here and after, nothing changes essentially when N is even. By Convolution Theorem, we can write the result of applying DFT to the equation 4.17 as:

$$Y[k] = H[k] \times Z[k] \tag{4.21}$$

where the convolution in the equation 4.17 is replaced by multiplication. Although the length of $h[n]$ is usually longer than that of $x[n]$, we can easily adjust the length of $h[n]$ by padding it with the required number of zeros.

We show an example of the effect of the digital filter in the figure 4.36. The bandwidth of the kernel $H[k]$ is directly related to the bandwidth of the filtering. We call the frequency region where $|H[k]|$ is large *passband*, and the region where $|H[k]|$ is small *stopband*. The frequency components of $x[n]$ in the *passband* pass through the filter, while the frequency components of $x[n]$ in the *stopband* are rejected by the filter.

We can make high-pass, low-pass, band-pass, band-reject, or more complex filters by changing $H[k]$[xviii]. The transformation from $x[n]$ to $y[n]$ described by the equation 4.16 is regarded as *filtering* in this way. Note that the complex phase of $H[k]$, i.e., $\arg(H[k])$, also plays important roles, as in the Matched filter, though we have only paid attention to $|H[k]|$ in the description above.

Shaping time and the signal to noise ratio

As in the figure 4.37, *signal* spectra are generally large in the small frequency region, while *noise* spectra are constant assuming white noise. We have the similar condition of *signal* and *noise* spectra. Under the condition like this, the *signal* to *noise* ratio gets larger when we pick out the low frequency region by the filtering. That is, the longer shaping time is the better to increase the *signal* to *noise* ratio.

If the shaping time of the filter is long, however, the output waveforms of *signal* get longer in the time domain, and the pileups of the output waveforms become more frequent to occur. Hence, we have to reject the low-frequency region and pick out the high-frequency region, in order to avoid the pileups. That is, the shorter shaping time is the better to avoid the pileups, or to reduce the dead time.

Consequently, our two requirements for the filtering, 1) maximize the *signal* to *noise* ratio, and 2) minimize the possibility that the pileup occurs, are contrary.

Design concept of our filter

To come to terms with the contradictory requirements, we adopt the strategy to apply two different filters in parallel to one input signal. We call the two filters the *slow filter* and the *fast filter*. The *slow*

[xvii] Note that n in $x[n]$ represents the axis of *time*, and k in $X[k]$ corresponds to *frequency*, here.
[xviii] Note that we can make the filter kernel $h[n]$ that corresponds to arbitrary $H[k]$, by using inverse Fourier transformation.

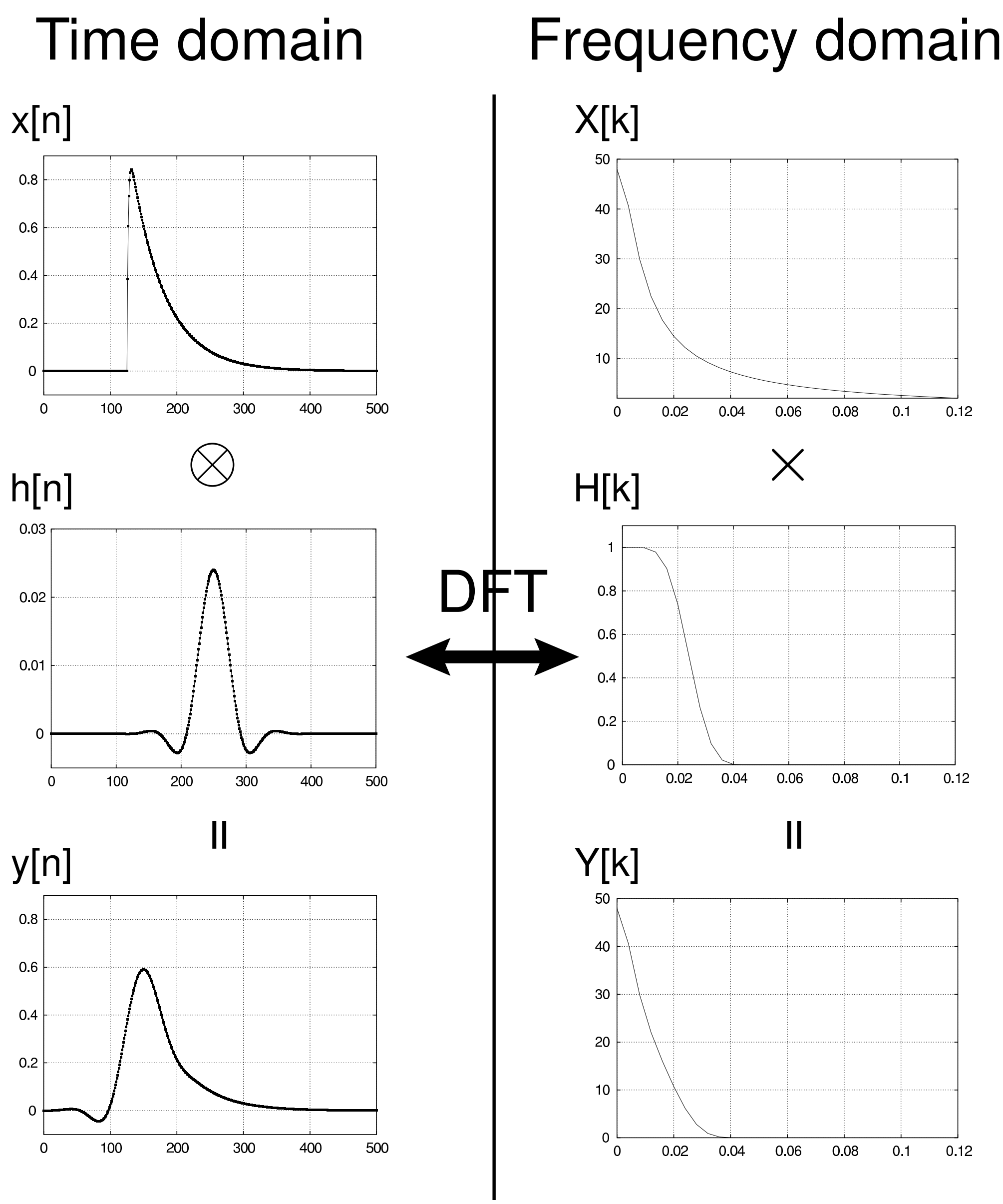


Figure 4.36: An example of the effect of the digital filter in the time domain and frequency domain.

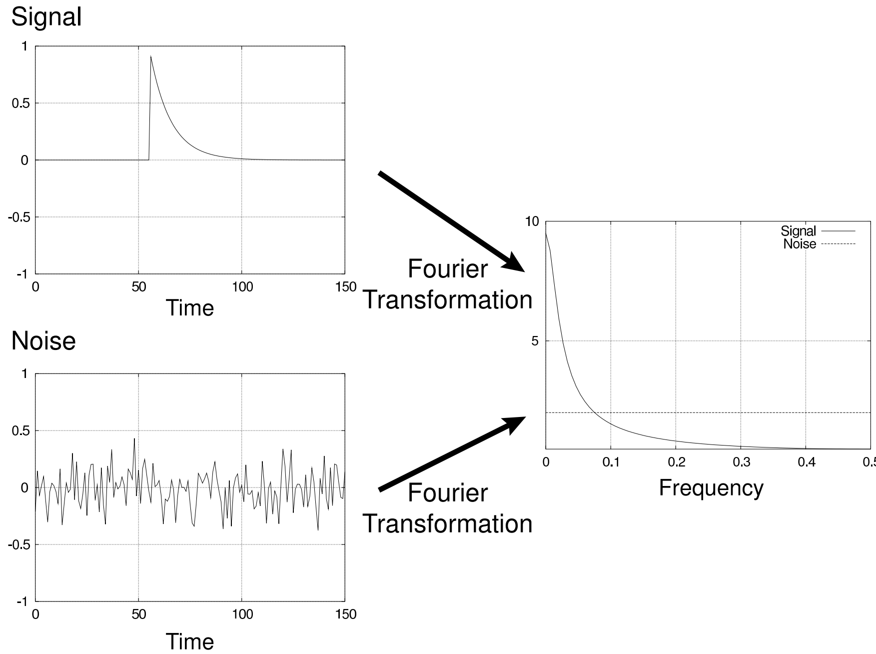


Figure 4.37: Comparison of signal and noise spectra.

filter is the filter optimized to maximize the *signal* to *noise* ratio. The *fast filter* is the filter optimized to minimize the possibility that the pile up occurs.

Our strategy is schematically shown in the figure 4.38. First we check the output of *fast filter* and count the number of peaks that exceed a specific threshold. We adopt the output of the *fast filter* when the number is more than one. We adopt the output of the *slow filter* when the number is one or zero.

The reason we consider this strategy effective is as follows. We have simulated the *two-neutrino event* using the simulation code of the K2K water Cerenkov detector in the near side[xix]. The size of the water tank is $\sim 10\text{m} \times 10\text{m} \times 10\text{m}$, and the PMTs of 50cm diameter are placed with 70cm intervals on the wall of the tank. We have counted the number of PMTs that detect Cerenkov lights from both of the two neutrinos and the number of PMTs that detect Cerenkov light only from one of the two. The result is that the number of the former is comparable with number of the latter. We consider our strategy effective in the condition like this where the numbers of adoption of the *slow filter* and the *fast filter* are comparable.

Deduction of Q and T

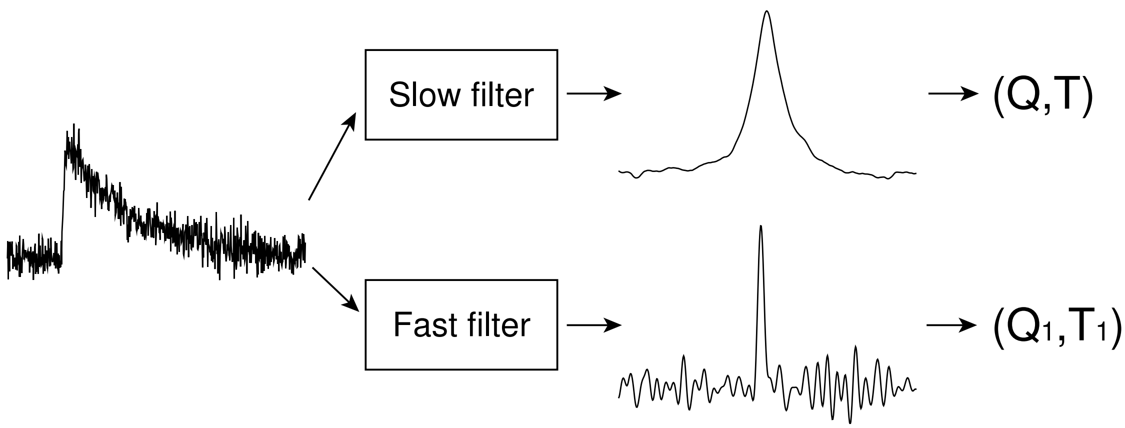
The method to deduce Q(amplitude of a signal) and T(timing of a signal) is very simple: to search for the peak(s) in the output from the digital filters. The peak height corresponds to Q and the peak position (time) corresponds to T. The result is one set of (Q, T) or nothing when we adopt the *slow filter*. The result is multiple set of (Q, T), like $(\mathrm{Q}_1, \mathrm{T}_1)$, $(\mathrm{Q}_2, \mathrm{T}_2) \cdots$, when we adopt the *fast filter*.

Implementation

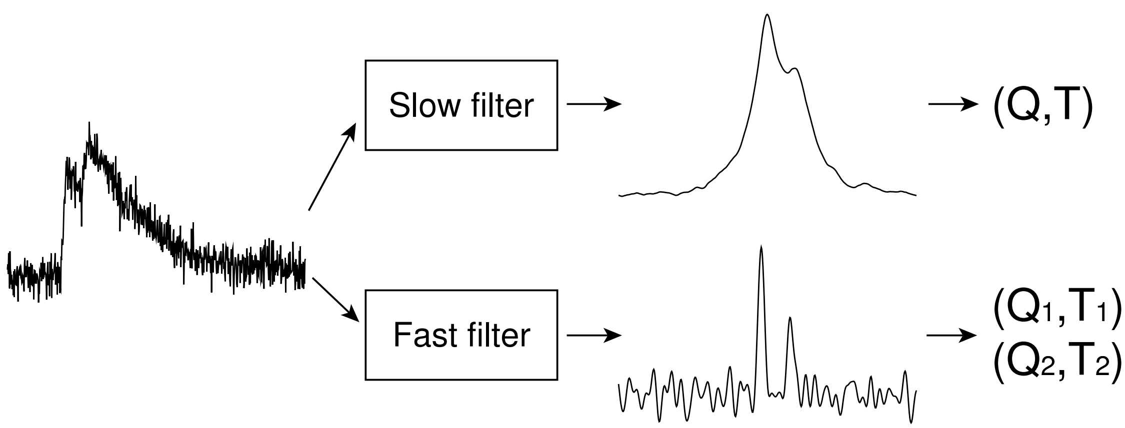
We have not yet tested to implement the processing into FPGA or DSP in practice. We expect, however, the implementation is not so difficult, since we can implement the processing described above with very simple operations: addition, multiplication and maximum search[xx].

[xix] We would like to express our sincere gratitude to K2K collaboration for their great kindness.

[xx] When the lengths of filter kernels get larger, the convolution method using FFT(Fast Fourier Transformation) is faster than the simple method in the equation 4.16. FFT is also implemented with addition and multiplication, and the processes are still implemented with the three simple operations.



(a) The case to adopt the output from the slow filter.



(b) The case to adopt the output from the fast filter.

Figure 4.38: The strategy of DSP.

4.4.2 Assumed signal and noise

We introduce here the *signal* shape and the *noise* spectrum we assume, in this section, to generate filters and to evaluate the performance of the digital signal processing. Note that our algorithms of designing filters itself are general and do not depend on a specific *signal* shape and *noise* spectrum.

We assume the *noise* spectrum in the figure 4.21. For the *signal* shape, we use $x(t)$ defined as follows:

$$x(t) \equiv A \int_0^{\infty} dt' x_{\mathrm{raw}}(t-t') \exp(-t'/\tau) \tag{4.22}$$

This is the convolution of $x_{\mathrm{raw}}(t-t')$ and an exponential, where $x_{\mathrm{raw}}(t-t')$ and the exponential correspond to the raw signal from HPD and the impulse response of the preamplifier, respectively. As $x_{\mathrm{raw}}(t)$, we use the function shape shown in the figure 4.39, taking account of the raw signal waveform of HPD HY0010, which is shown in the figure 3.4(b). We set τ of the exponential to be 80ns, which corresponds to the preamplifier's feedback settings in the section 4.2.4: $C_f = 2$pF and $R_f = 20$kΩ. We set A so that the maximum amplitude of $x(t)$ is 4mV, which corresponds to the single photon equivalent signal assuming the HPD gain of 5×10^4 and the preamplifier setting in the section 4.2.4. Figure 4.40 shows the *signal* and *noise* waveforms generated from the assumptions above. The sampling rate is 2GHz and the bandwidth is set to be full(1GHz). S/N of the raw signal is ~ 5, being different from that in the section 4.2.4, where the bandwidth is limited by oscilloscope and $S/N > 10$.

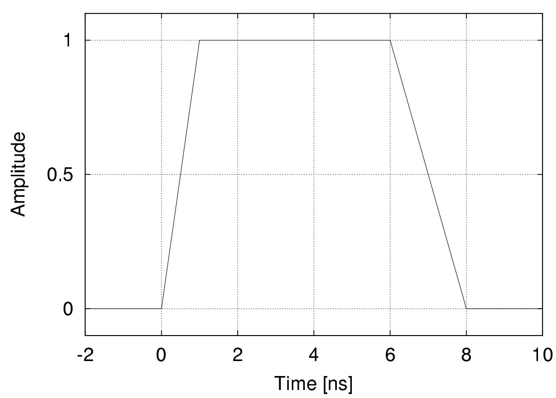


Figure 4.39: The assumed signal shape of $x_{\mathrm{raw}}(t)$.

The *signal* shape and the *noise* spectrum we adopt here are based on the measurement in the previous sections: the measurements of HPD raw signal shape, preamplifier feedback settings, and preamplifier's input transistor. We also take account of the measured gain of HPD in the chapter5. Consequently, the evaluations in the following are not only the evaluation of the digital filters but also the performance expectation of our total system of (HPD) + (readout system) for an input of a single photon.

4.4.3 Filtering algorithms

Slow filter

The aim of this filter is to get the maximum *signal* to *noise* ratio, which is expected to results in the best Q and T resolution. The detail of the algorithms we use in this filter is in the appendix B.1.

As the slow filter, we use the optimal matched filter, whose filter kernel is defined by the equation B.24 or B.35. It gives the best signal to noise ratio. We show the kernel of the slow filter in the figure 4.41, which we have generated with the assumed signal and noise spectrum described in the section 4.4.2. The kernel in the frequency domain shows that the kernel highly enhances the low frequency region.

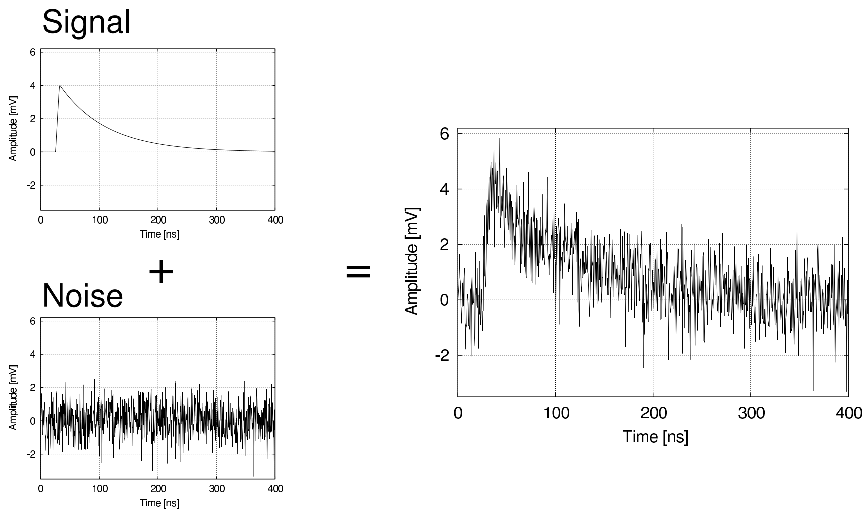
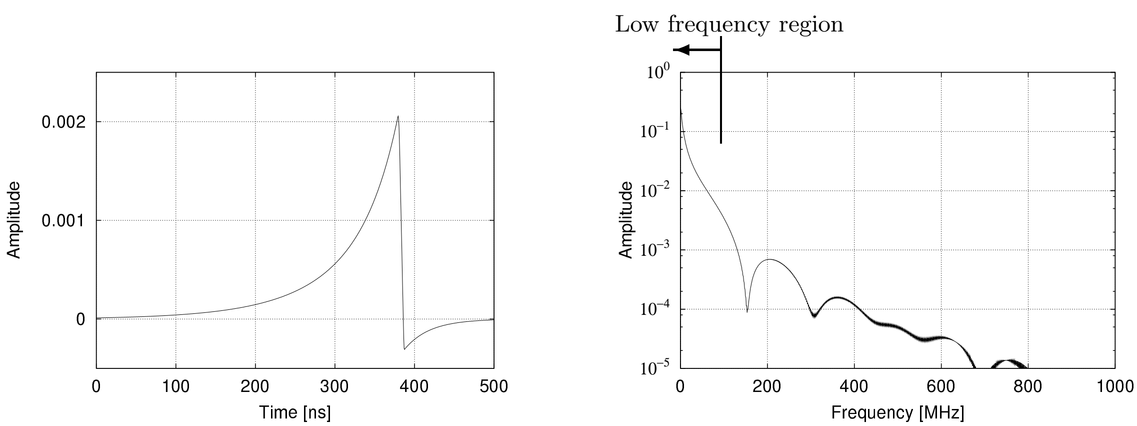


Figure 4.40: The noise and signal of simulation setting.



(a) In the time domain.

(b) In the frequency domain.

Figure 4.41: The kernel of the slow filter.

Fast filter

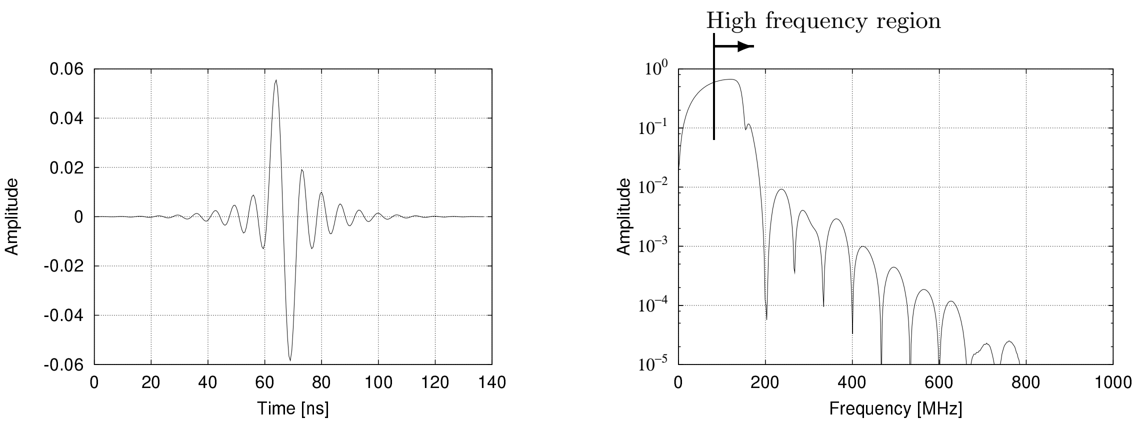
The aim of this filter is to minimize dead time, or to avoid the pileups. The short dead time is essential for the realization of high rate tolerance. We describe the detail of the algorithms we use in the fast filter in the appendix B.2.

We use the combination of the Deconvolution filter and the Wiener optimal filter, since the combination gives the optimal solution for our purpose. We select the Blackman Window shape as the desired output waveform $y_{des}[n]$ of the filter. The function $BW_\tau(t)$ which represents the Blackman Window is as follows.

$$BW_\tau \equiv \begin{cases} 0.42 - 0.5\cos(2\pi t/\tau) + 0.08\cos(4\pi t/\tau) & (0 < t < \tau) \\ 0 & (\text{otherwise}) \end{cases} \tag{4.23}$$

Figure B.1 shows the function shape of "$BW_{10\text{ns}}(t - 5\text{ns})$" for example. FWHM(Full Width Half Maximum) of the Blackman Window BW_τ is $\sim 0.4\tau$. Whereas the smaller τ is better to reduce the dead time, the smaller τ results in the smaller S/N. Since we expect the S/N gets considerably worse when we set τ so that 0.4τ is shorter than the raw signal of the HPD, we set $\tau = 15$ taking account of the raw *signal* waveform of HPD HY0010 shown in the figure 3.4(b). We will mention the relation between S/N and the τ, later. We set the normalization factor c_n of the Wiener filter to be unity, as a consequence of the discussion in the appendix B.2, where $c_n \sim 1$ is considered optimal.

We show the kernel of the fast filter in the figure 4.42, which we generated with the assumed signal and noise spectrum described in the section 4.4.2. Compared with the slow filter kernel in the figure 4.42, the high frequency region ($\sim$ 100MHz) is enhanced.



(a) In the time domain.

(b) In the frequency domain.

Figure 4.42: The kernel of the fast filter.

4.4.4 Processing after the filtering

Deduction of Q and T

We deduce the amplitude of the signal Q and the timing of the signal T, by searching for the maximum of an output signal waveform from the filters.

We search for one global maximum in the output of the slow filter. On the other hand, we search for the multiple local maximums that exceeds a specific threshold in the output of the fast filter, in practical implementation. In this thesis, however, we search for only one maximum in the output of the fast filter, too.

To avoid the error due to discreteness

If we simply adopt the peak search method, the timing resolution of digital signal processing is determined not only by the performance of filters, or, the *signal* to *noise* ratio. When the sampling rate is comparable with the timing resolution, the *discreteness* of the digital signal also deteriorates the resolution.

As an example, here we assume that the sampling rate is 1GHz and the intrinsic timing resolution is zero, i.e., ideally good. Then, the resultant timing T is 1ns when the real timing is 1.2ns, and T is 3ns when the real timing is 2.7ns, because of the discreteness. We can consider the difference between the timing T and the real timing as errors due to the discreteness. The error is −0.2ns in the former case, and is +0.3ns in the latter case. It is clear that the error distributes uniformly from −0.5ns to +0.5ns, where the 0.5ns is determined by the sampling rate 1GHz. In general, the error due to the discreteness is written as follows:

$$\frac{1}{F_s \cdot \sqrt{12}} \tag{4.24}$$

where F_s is the sampling rate. This error is mostly the same as the error called *quantization error*. Since the errors of this type appear very much like random noise, we have to pay attention to the error.

We can avoid the error, however, by interpolating the filter kernels. As shown in the figure 4.43, we interpolate the original kernel $h_0[n]$ and generate kernels $h_1[n], h_2[n] \cdots h_{M-1}[n]$. Then we replace the equation of the convolution, the equation 4.16, with the following equation.

$$y[i_1 + i_2/M] = \sum_{j=0}^{N-1} x[i_1 + j] \times h_{i_2}[N - j - 1] \quad (i_1 \in \mathbf{Z},\ i_2 = 0, 1, 2, \cdots M-1) \tag{4.25}$$

In the original convolution, the point next to $y[0]$ is, of course, $y[1]$. On the other hand, using the convolution with the interpolated kernel, we can make the point next to $y[0]$ to be $y[0 + 1/M]$. By this prescription, we get the output waveform with the effective sampling rate that is M times larger than the original one, and hence we can effectively make the sampling rate much larger than the timing resolution. Thereby, we can avoid the deterioration of the timing resolution due to discreteness.

Then the issue is *how to interpolate the kernel*. When we make the kernel using analytic function, like the Windowed-Sinc filter kernel, we do not have to think of the issue. In our case, however, we have to think of the issue, since we make the filter kernels using the original signal, i.e., $X_s[k]$, in the equations B.24, B.40 and B.53.

First, we simply attacked the issue using linear interpolation, and it did not work well. We filtered input signals contaminated with noise using the equation 4.25 and the linear interpolated kernel. Then the peak positions of the output signals are often in the places at which $i_2 = 0$ in the equation 4.25.

Then we have attacked the issue with the method using Fourier transformation, and it has worked well. When the sampling rate in the time domain gets M times larger, the corresponding spectrum in the frequency domain gets M times longer. Thus, we can make the interpolated kernel as shown in the figure 4.44[xxi]: 1) we transform the original kernel in the time domain, $h_{\mathrm{orig}}[n]$, into that in the frequency domain, $H_{\mathrm{orig}}[k]$, 2) we extend $H_{\mathrm{orig}}[k]$ toward the high frequency region by padding it with zeros and make $H_{\mathrm{inter}}[k]$, and 3) we transform the $H_{\mathrm{inter}}[k]$ back into the interpolated kernel in the time domain, $h_{\mathrm{inter}}[n]$. We evaluate the performance of our filters using this prescription in the section 4.4.5.

Since our goal of the timing resolution is $\sim$ 1ns and the error due to the discreteness is 0.3 $\sim$ 0.15ns for the sampling rate of 1 $\sim$ 2GHz, the error can be negligible in practical implementation. In this thesis, however, we adopt the kernel interpolation to evaluate the intrinsic performance of our system.

4.4.5 Performances

In this section, we show performances of our digital processing. We evaluate the performances assuming the noise spectrum and signal shape described in the section 4.4.2. We also assume the sampling rate of 2GHz.

[xxi] We can also use similar interpolation method when we use the FFT convolution.

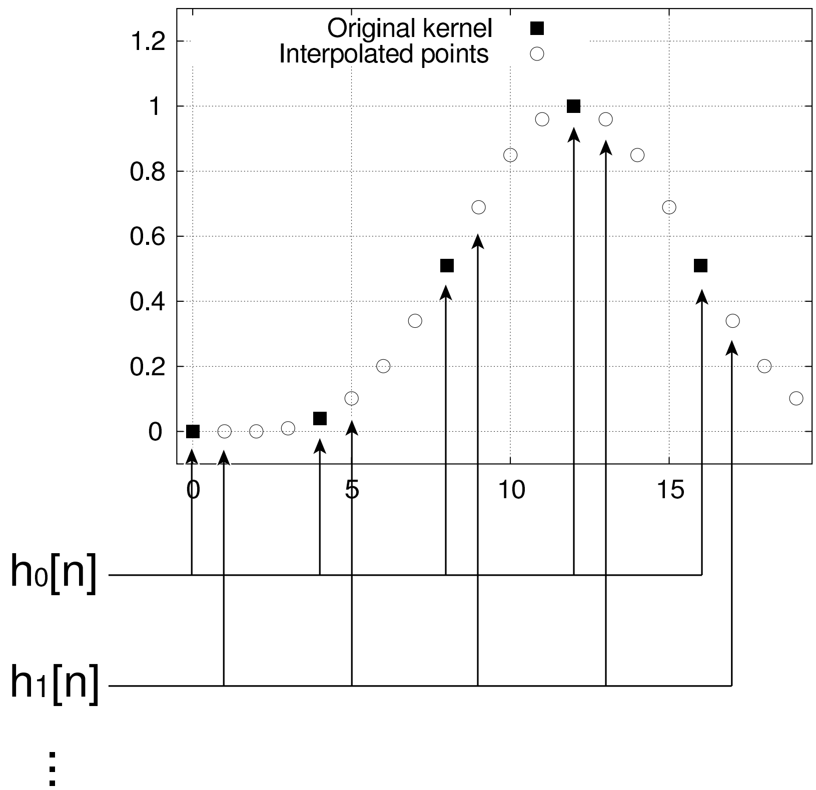


Figure 4.43: An example of generation of kernels $h_1[n], h_2[n], \cdots$, from $h_0[n]$, by interpolation.

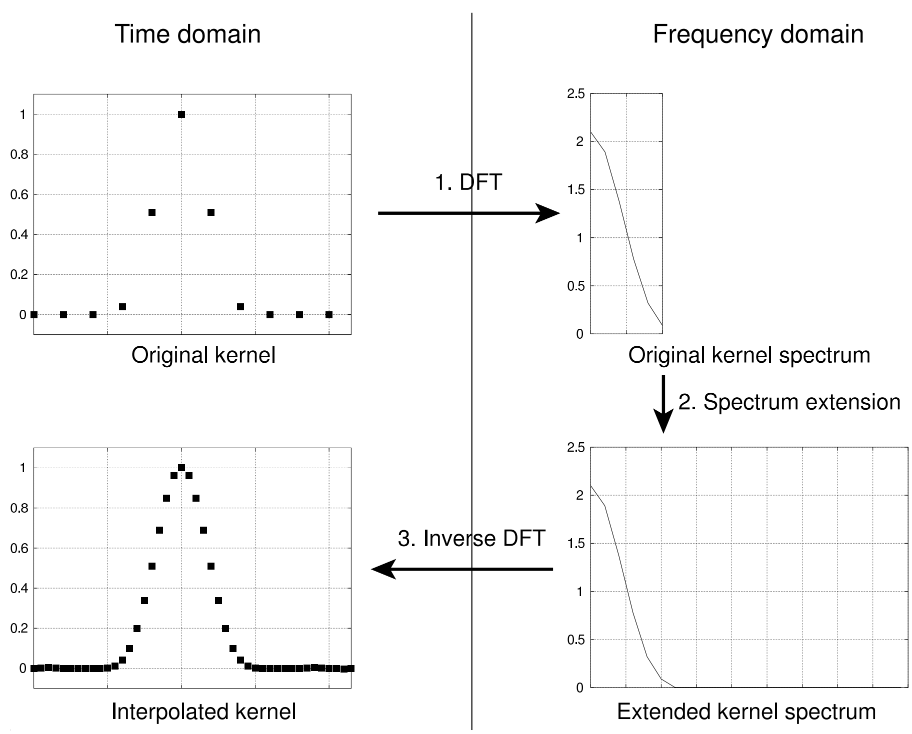


Figure 4.44: The method of interpolation using Fourier transformation.

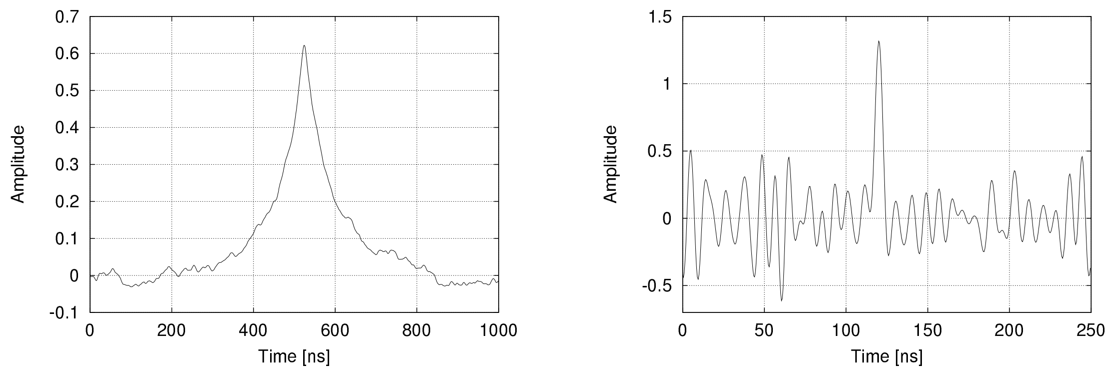
Q resolution

Here, we show the result of evaluation on the resolution of Q, the amplitude, of our slow filter and fast filter. That is, we show the evaluation of the *signal* to *noise* ratio, S/N.

First, we filter the ideal input signal waveform without noise, and get the amplitude of the output waveform. The amplitude is "S," the numerator of S/N. Second, we filter the noise waveform with the length of 10,000 sample (equivalent to 5μs) and calculate the root mean square (RMS) of the output waveform[xxii]. This RMS is "N," the denominator of S/N. Then, we calculate the S/N for our filters with the assumed signal and noise input.

For the slow filter, the signal amplitude after the filtering is 0.610 and the RMS of noise waveforms is 0.0173 and hence the S/N is 35. The noise and the signal after the filtering is shown together in the figure 4.45(a). Note that although we show them together in the figure, we have evaluated them separately.

For the fast filter, the signal amplitude after the filtering is 1.0 and the RMS of noise waveforms is 0.174, and hence the S/N is 5.8. The noise and the signal after the filtering is shown together in the figure 4.45(b). We have also measured the S/N for different $y_{des}[n]$, by changing the τ of the Blackman Window $BW_\tau(t)$, which is defined in the equation 4.23. We show the result in the figure 4.46. The S/N gets worse than 5 when we set τ less than 15ns.



(a) The output of the slow filter.

(b) The output of the fast filter.

Figure 4.45: The signal and noise in the output of the filters.

Table 4.8: The signal to noise ratio of the output waveforms.

Filter type	Signal S	Noise N	S/N
Slow filter	0.610	0.0173	35
Fast filter	1.0	0.174	5.8

T resolution

To evaluate the timing resolution of the filters, we generate 10,000 waveforms with the assumed *signal* shape and the *noise* spectrum. We filter the waveforms using the slow and fast filters with the interpolated kernels. The interpolated kernel has $\times 5$ effective sampling rate, that is, the sampling rate of the output waveforms is effectively 10GHz. Then, we search the maximum in each output waveform and determine the timing T.

[xxii]Note that we use the *valid* output length of filters, that is, the length of $N_{\text{sig}} - 2N_{\text{kernel}}$ where N_{kernel} is the length of the filter kernel and N_{sig} is the length of the input noise, which is 10000 in our case.

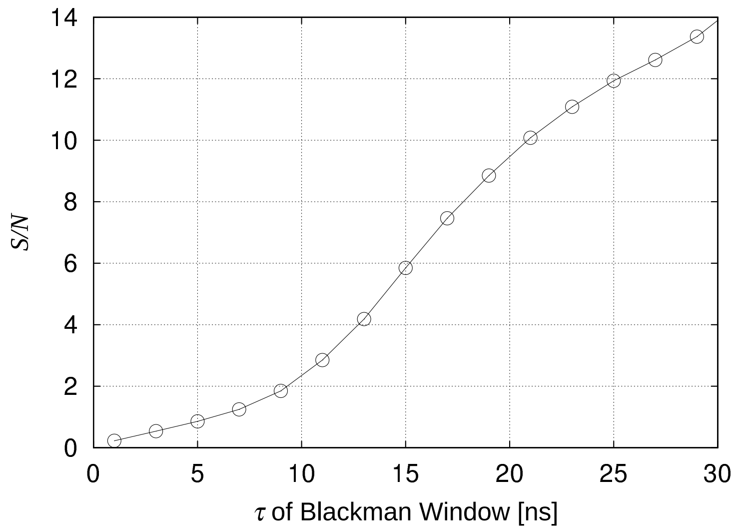


Figure 4.46: The dependence of Blackman Window's τ on S/N.

Figure 4.47 shows the distributions of the resultant T of the slow and fast filters. The RMS, which we define as *resolution*, is $\sim$ 0.4ns for the slow filter, and $\sim$ 0.9ns for the fast filter. We also fit the distribution of the slow (fast) filter with Gaussian (double Gaussian) and calculated the width, or standard deviation (σ) of the Gaussian. We define the width (σ) of the double Gaussian as:

$$\sigma \equiv \sqrt{\frac{A_1\sigma_1 \cdot {\sigma_1}^2 + A_2\sigma_2 \cdot {\sigma_2}^2}{A_1\sigma_1 + A_2\sigma_2}} \tag{4.26}$$

where the two Gaussian functions, $g_i(x)$, are defined as:

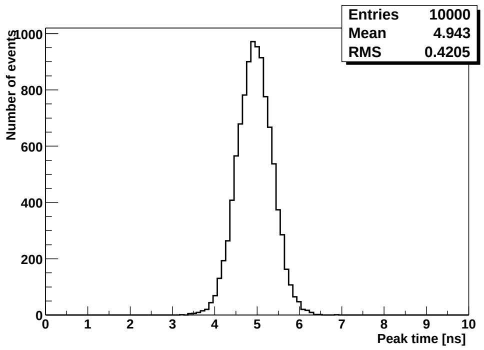
$$g_i(x) \equiv A_i \exp\left(-\frac{x^2}{2{\sigma_i}^2}\right) \ . \tag{4.27}$$

χ^2/n.d.f. of each fitting is ~ 1. The resultant widths are consistent with the root mean squares of the distributions.

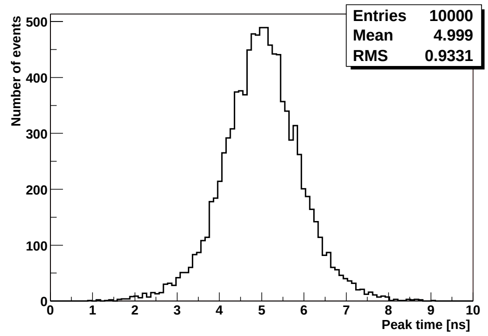
Rate tolerance (for the fast filter)

If we use the Deconvolution filter only, the Deconvolution filter perfectly recovers the desired output waveform, which is the Blackman window with τ = 15ns here. Consequently, the dead time of the output is guaranteed to be exactly 7.5ns. We use, however, the combination of the Deconvolution filter and the Wiener optimal filter. In this case, the output signal shape deviates from the desired waveform, the Blackman window, and we have to confirm that the dead time is really short as we intend.

Figure 4.48 shows the waveform that we get, by filtering the signal waveform without noise with the fast filter. The deviation from the original Blackman window is small enough that the dead time is still $\sim$ 7.5ns for the fast filter.



(a) The result of the slow filter.



(b) The result of the fast filter.

Figure 4.47: The T distribution.

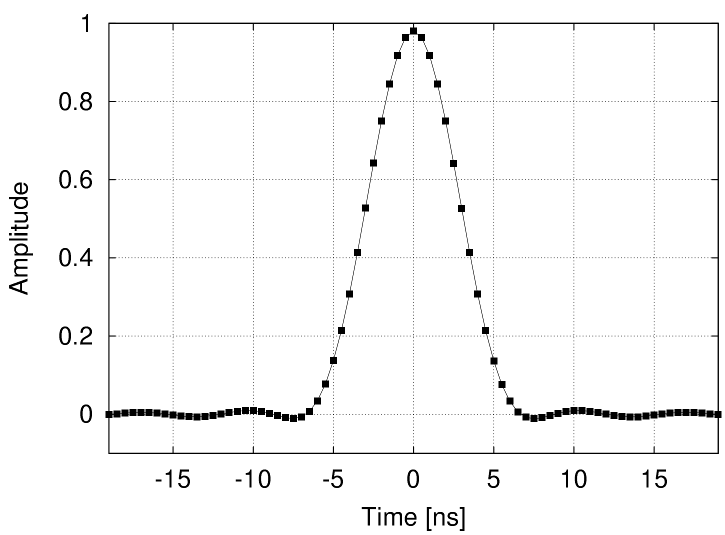


Figure 4.48: The signal shape of the fast filter's output.

Chapter 5

The connection test of HPD and readout

In this chapter, we report the result of the test using 5 inch prototype HPD, which we evaluated in the section 3.2, the preamplifier mentioned in the section 4.2, and the digital filter mentioned in the section 4.4.

5.1 Setup

Figure 5.1 shows the schematic drawing of our test setup, and figure 5.2 shows the pictures of the setup. We use the HPD HY0009 and HY0010 that we evaluate in the section 3.2. We place the light collimator to avoid the signal timing jitter due to the deviation of photoelectron orbits, since the 5 inch prototypes are not optimized for the optics of the photoelectrons and it is meaningless to measure the jitter due to the deviation. The applied HV is -8.5kV for both HPDs, and the applied bias on AD is 155V (350V) for HY0009 (HY0010). The preamplifier in the setup is the one described in the section 4.2. As a waveform sampler, we have used the oscilloscope HP Infinium 54825A with 2GHz sampling rate mode, since the AMC described in the section 4.3.3 is still under development. The analog bandwidth of the oscilloscope is 500MHz. We acquire the data taken with the oscilloscope into a PC. Then we apply the digital signal processing described in the section 4.4 to the offline data.

We want to emphasize that the setup described above is close to our goal described in the section 4.1. The signal from HPD goes through the same path in principle, although there are differences between the final system and the setup here: we use an oscilloscope instead of AMC, and PC instead of FPGA or DSP. Therefore, with this setup we can perform not only the evaluation of our current system, but also the feasibility test of the final system.

5.2 Data before analisys

The noise spectrum

We have measured the noise spectrum of the preamplifier connected with the HPDs. After acquiring 2,000 noise waveforms with 1000ns length for each waveform, we transform each waveform using DFT into a spectrum in the frequency domain, $Y_n[k]$, and average the 2,000 spectra. Then we get the nose power spectrum $S_n(f)$ using the following equation,

$$S_n(F_s \cdot \frac{l}{N}) = \frac{2}{N \cdot F_s}\left|Y_n[l]\right|^2 \quad (l = 0, 1, 2, \cdots, N/2) \tag{5.1}$$

where F_s, N, and $Y_n[l]$ are the sampling rate, the length of one waveform, and the averaged noise spectrum. Here, the length of waveforms is counted in the number of sample point, and the sampling rate F_s is 2GHz.

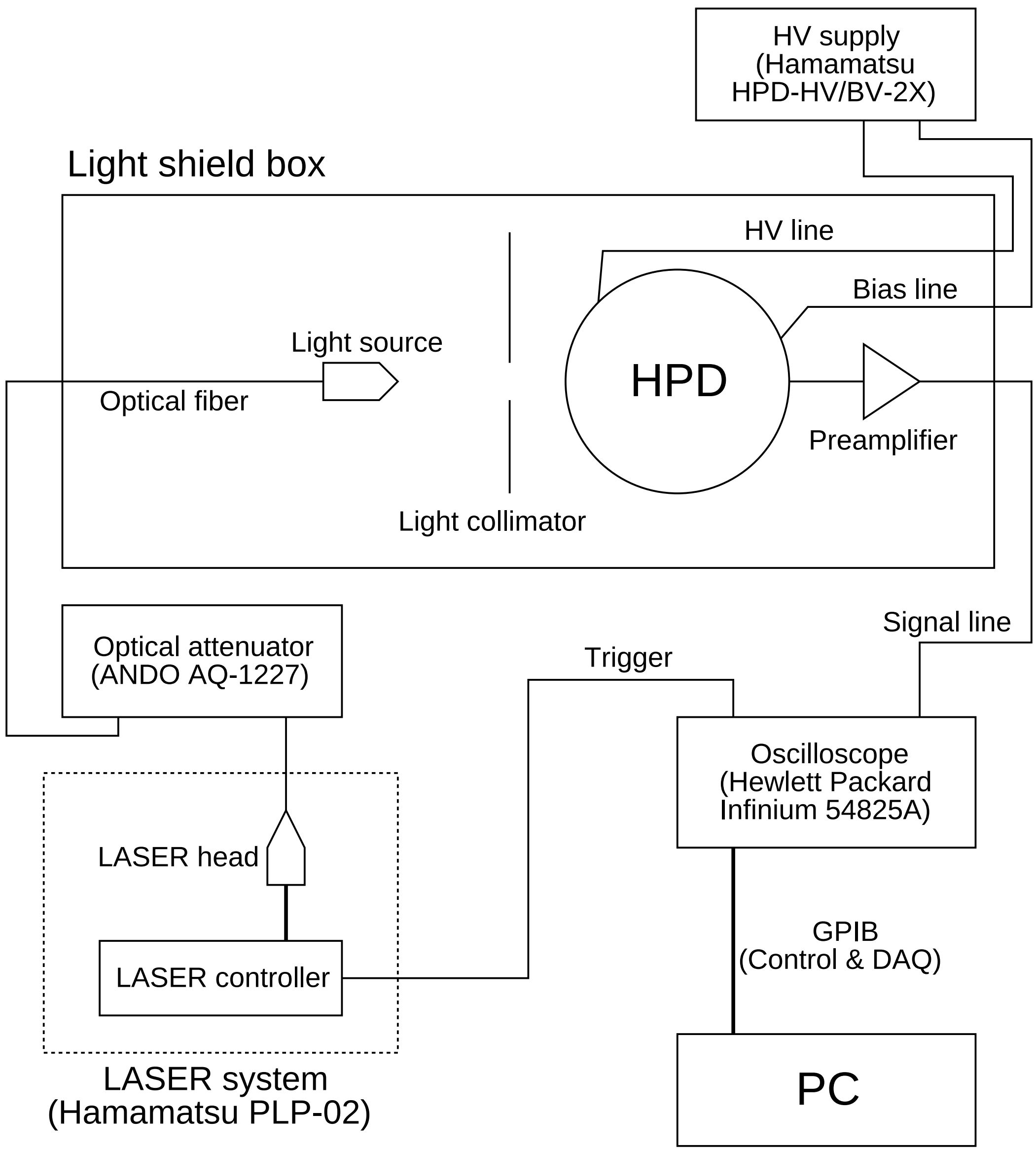
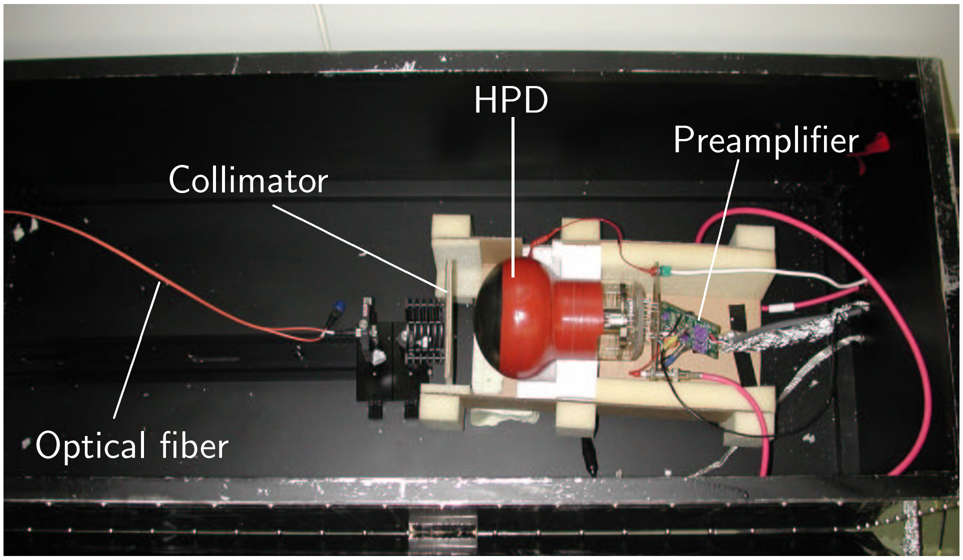
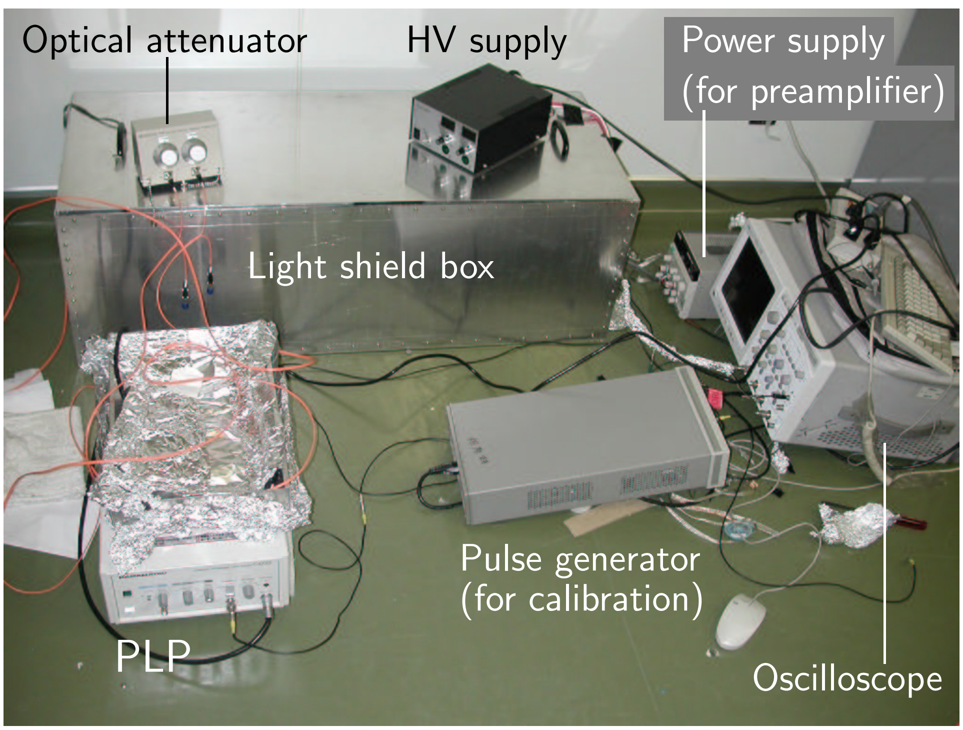


Figure 5.1: A schematic drawing of the connection test setup.



(a) A picture of the light shield box's inside.



(b) A picture of the whole setup. The PLP is placed far from the light shield box in actual measurement, in order to avoid the noise from the laser head.

Figure 5.2: Pictures of connection test setup.

We show the result of the measurements with HY0009 and HY0010 in the figure 5.3, together with the noise spectrum without HPD ("Preamplifier only" line), where we set the capacitor of 50pF emulating the detector capacitance. In both spectra of HY0009 and HY0010, there are large spikes in the frequency region of $\sim$ 200MHz. We consider the spikes are due to external noise sources, since, in the frequency region without the spikes, the noise powers are almost the same as the noise power of the measurement without HPD.

The primary source of the external noise is from the electromagnetically dirty environment where we did the measurements. The structure of the 5 inch prototype HPD is not optimized to avoid such noise, and its structure has a long interval of $\sim$ 10cm between the AD and the preamplifier. Thus, the structure picks up the external noise. On the other hand, the 13 inch prototype has more optimal structure, and the interval is reduced to be a few cm. We therefore expect the noise problem will be less severe for the 13inch prototype HPD.

Note that, due to the buffer part of our amplifier circuit in the figure 4.19, the gain of the amplifier is $2 \sim 3$ times larger than the gain without the buffer. Thus, to compare the noise spectra in the figure 5.3 with that in the figure 4.21, we should scale down the spectra in the figure 5.3 by $2^2 \sim 3^2$ times.

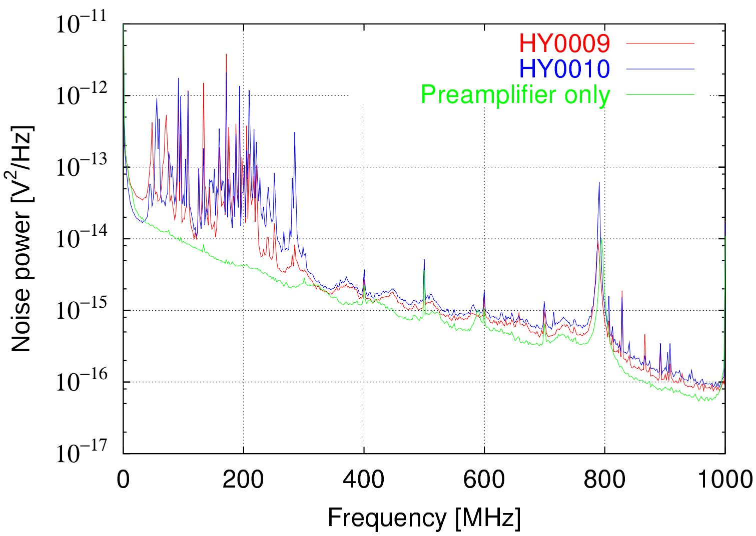


Figure 5.3: The noise power spectrum of each setting.

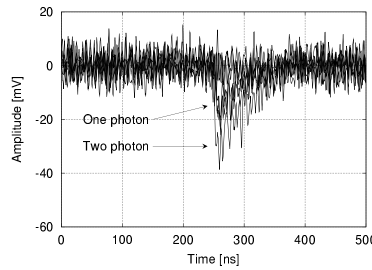
The shapes of waveforms

We measure the signal shape of HPD through the preamplifier, and filter them using the slow and fast digital filters described in the section 4.4

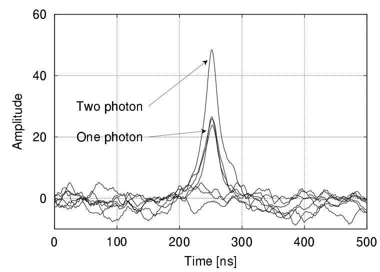
Figures 5.4(a) and 5.5(a) show the raw signal shape of HY0009 and HY0010, respectively. In the figures, we can see the waveforms which correspond to one and two photon signals. Although the noise appears considerably large, we can still extract the meaningful information because of the localization of the noise power spectra.

Figures 5.4(b) and 5.5(b) show waveforms after filtering with the slow filters. We have generated the slow filters using the algorithm described in the section 4.4.3 and assuming the measured signal and noise spectra. The waveforms in the figures 5.4(b) and 5.5(b) exactly correspond to the waveforms in the figures 5.4(a) and 5.5(a). Here the signal to noise ratio is highly enhanced by the slow filters.

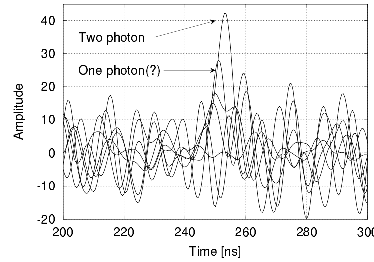
Figures 5.4(c) and 5.5(c) show waveforms after filtering with the fast filters. Note that the scales of horizontal axises are different from the other figures. We have also generated the fast filters using the algorithm described in the section 4.4.3 and assuming the measured signal and noise spectra. The waveforms correspond to that of 5.4(a) and 5.5(a), too. The signal widths are narrow, though the signal to noise ratio is small. It is particularly small for the waveforms of HY0009. This is because we have decided the width of the output signal to be optimal for the raw signal shape of HY0010 as



(a) Raw signal

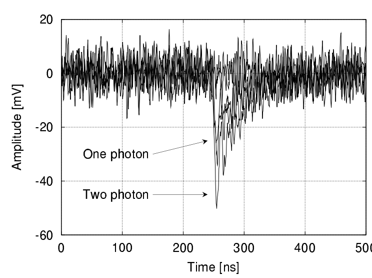


(b) After slow filter

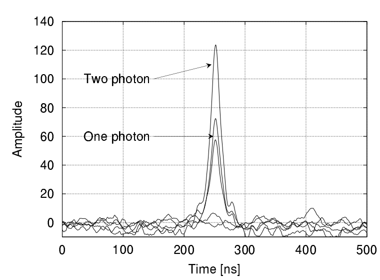


(c) After fast filter

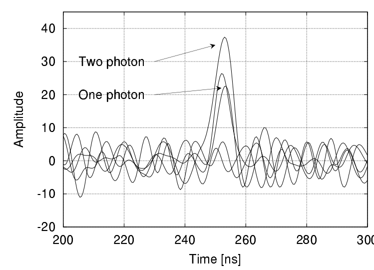
Figure 5.4: Waveforms of HY0009.



(a) Raw signal



(b) After slow filter



(c) After fast filter

Figure 5.5: Waveforms of HY0010.

described in the section 4.4.3. The width is probably too narrow for HY0009, whose raw signal is slower than that of HY0010. Thus, we do not mention anymore the result from HY0009 with the fast filter.

We also have to check the shape of the waveform after the fast filter, as described in the section 4.4.5. Figure 5.6 shows the typical shape of the waveform of HY0010 after the fast filtering. We have made the waveform using the averaged HY0010 signal shape and the fast filter we actually use to process data. The dead time is still about 7.5ns and we achieve the short dead time with the filter.

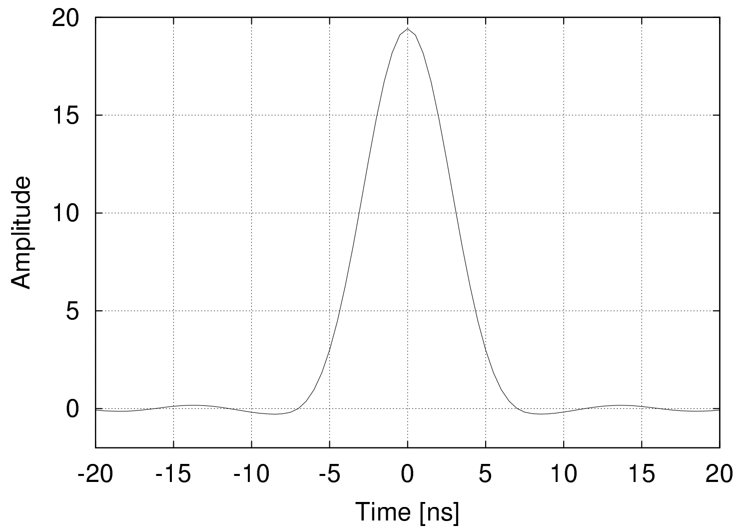


Figure 5.6: The typical shape of the waveform of HY0010 after the fast filtering.

5.3 Performance evaluation

Q distribution

We calculate the raw signal amplitude(Q_{raw}) by the following process. *Raw* means *before calibration*, here. With filter kernels, we filter the typical signal shape, which we get by averaging signal waveforms. We define the maximum position of the filtered typical waveform as T_{max}, which corresponds to the timing of the input signal. Then we define Q_{raw} for each event as[i]

$$Q_{raw}[\text{count}] \equiv y[T_{max}] \tag{5.2}$$

$$= (h \otimes x)[T_{max}] \tag{5.3}$$

where $x[n]$, $h[n]$ and $y[n]$ are the raw signal waveform of the event before filtering, the filter kernel, and the signal waveform after filtering, respectively.

Then we calibrate the Q_{raw}[count] and get the calibrated amplitude Q, where the unit is [electron], as follows. We put the test pulse into the "Test pulse input" terminal of the preamplifier, which one can see in the figure 4.19. To reproduce the same detector capacitance condition, we connect the preamplifier to HPDs and we apply the same AD bias as we do in the measurement of photons. In addition, to reproduce the signal shape from HPDs, we set the leading edge length of the test pulse to be the same as the width of the raw signal of HPDs shown in the figure 3.4: $\sim$ 10ns for HY0009 and $\sim$ 6ns for HY0010, in the meaning of 10%-90% time. Then we acquire the output from preamplifier using the oscilloscope and PC, in the same way as the other measurements, and process with the same filters. We define the maximum of the filtered calibration waveform as Q_{cal}[count]. Note that we have three different types of data sets, HY0009 with slow filter, HY0010 with slow filter and HY0010 with fast

[i] Here we define the unit of Q_{raw} as "count," since Q_{raw} is something like ADC count.

filter, Accordingly, we measure Q_{cal} for each data set. On the other hand, the input signal Q_{in}[electron] is calculated as

$$Q_{in}[\text{electron}] = \frac{C_{in}[\text{F}] \times A_{in}[\text{V}]}{1.6 \times 10^{-19}[\text{C/electron}]} \tag{5.4}$$

where C_{in} is the capacitance connected to the "Test pulse input" terminal in the figure 4.19, which is 2pF here, and A_{in} is the amplitude of the test pulse. Using Q_{cal}[count] and Q_{in}[electron], we calculate the calibration constant c_{cal}[electron/count] as follows.

$$c_{cal}[\text{electron/count}] = \frac{Q_{in}[\text{electron}]}{Q_{cal}[\text{count}]} \tag{5.5}$$

We then calculate the calibrated signal amplitude Q[electron] for each raw signal amplitude Q_{raw}[count] as follows.

$$Q[\text{electron}] = c_{cal}[\text{electron/count}] \times Q_{raw}[\text{count}] \tag{5.6}$$

Figures 5.7, 5.8 and 5.9 show the distributions of the calibrated signal amplitudes, Q[electron], for each data set with vertical axises in linear scale and log scale. The peaks at 0 electron correspond to noise, and we define this peak as 0-th peak. The peaks next to the 0-th peak are from the single photon contribution. We define the peak as the first peak. In the same manner, one can clearly see up to the fourth peak. We can also observe the fifth and more peaks increasing the magnitude of input light. The peak of HY0010 with slow filter is the most narrow and that of HY0010 with fast filter is the broadest.

We fit each peak with a Gaussian, or the sum of a Gaussian and a first order polynomial. We consider the mean and the standard deviation of the fitted Gaussian as the position and the width of each peak, respectively. The result of the fitting is shown in the figure 5.10, where the error bars correspond to the widths of the peaks.

We fit the peak position, Q_{peak}, with the function of peak number as:

$$Q_{\text{peak}} = G \times n_p + Q_{\text{offset}} \tag{5.7}$$

where n_p is the peak number, $n_p = 0$ for 0-th peak, $n_p = 1$ for 1st peak, ..., and so on. G and Q_{offset} are free parameters and are determined by the fitting. G corresponds to the gain of HPD, and is shown in the table 5.1. The fitted lines are also shown in the figure 5.10. One can see the line and data points are in consistent. As an indicator of the linearity, we calculate the RMS of residuals (σ_{residual}) as:

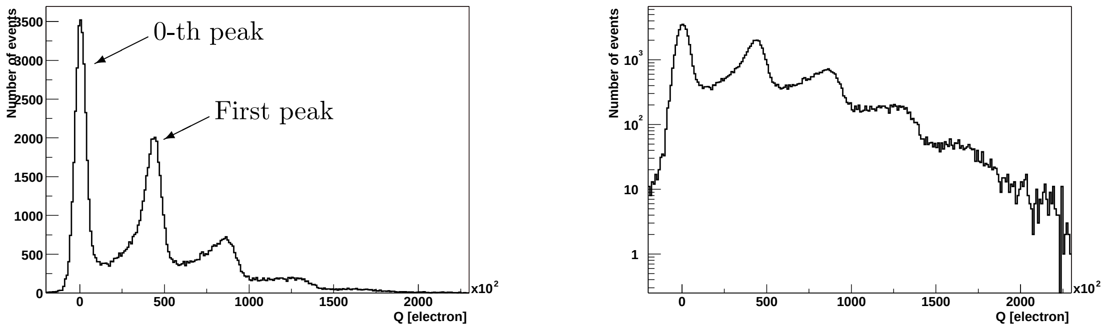
$$\sigma_{\text{residual}} \equiv \sqrt{\frac{\sum_{n_p=0}^{4} \left[Q_{\text{dat}}(n_p) - f(n_p)\right]^2}{\text{n.d.f.}}}, \tag{5.8}$$

where $Q_{\text{dat}}(n_p)$ and $f(n_p)$ are the data point for the n-th peak and the fitted function, respectively. Here the n.d.f. (number of degree of freedom) equals 3, since the number of data points is 5 and we use a first order polynomial as the function to be fitted. Table 5.1 shows the resultant σ_{residual}. They are small enough compared with the noises, σ_n, in the table, which we calculate in the following. Thus, the gain linearity for each data set is good enough.

The gain of HY0010 with the fast filter is larger than that of HY0010 with the slow filter. We consider the difference is from the different uncertainties in the calculation of the calibration constants. Whereas the slow filter works like an integrator and is sensitive to the height of input signals, the fast filter works like a differentiator and is sensitive to the shape of the leading edge of input signals. On the other hand, though we have modeled the test pulse after the signal from HPD, the detailed shape can deviate from the real HPD signal. The deviation of the test pulse shape results in the deviation of the leading edge shape of the preamplifier output, which is the input to the digital filters. Consequently, the error can be larger in the calibration of the fast filtered data, than the slow filter. Thus, we conclude, adopting the values of the slow filter, the gain of HY0009 and that of HY0010 are 4.2×10^4 and 4.4×10^4, respectively.

We also acquire the data without the input signal, that is, the noise data. We analyze them in the same way as the signal data and calculate the root mean square (RMS), σ_n, of the noise data for each data set. Considering the gain G as the signal height for single photon input and the noise RMS σ_n as the noise, we calculate the signal to noise ratio S/N as follows.

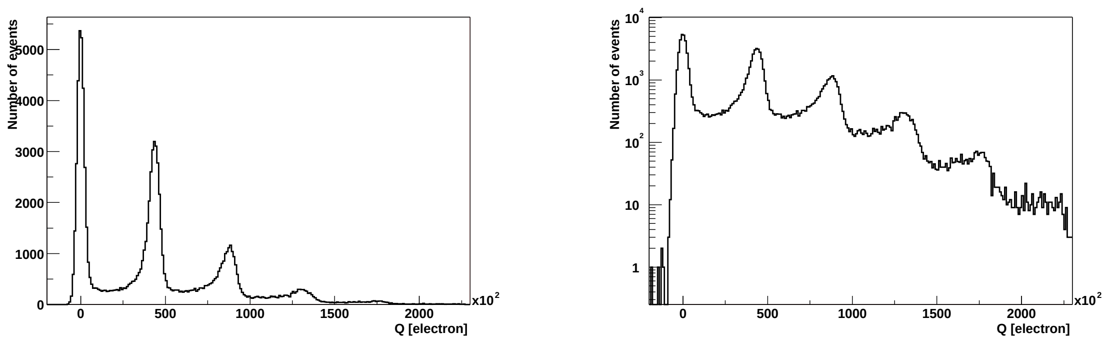
$$S/N = \frac{G}{\sigma_n} \tag{5.9}$$



(a) In linear scale.

(b) In log scale.

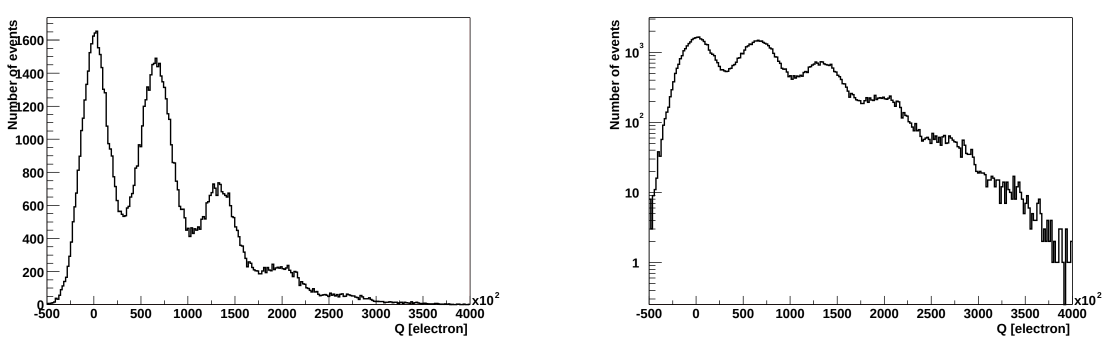
Figure 5.7: The distributions of Q(signal amplitude) for HY0009 with slow filter.



(a) In linear scale.

(b) In log scale.

Figure 5.8: The distributions of Q(signal amplitude) for HY0010 with slow filter.



(a) In linear scale.

(b) In log scale.

Figure 5.9: The distributions of Q(signal amplitude): HY0010 with fast filter.

The gain, noise, and the signal to noise ratio for each data set are shown in the table 5.1, together with the $\sigma_{\mathrm{residual}}$. The noises are consistent with the width of the 0-th peaks in the figures 5.7, 5.8 and 5.9, in accuracies of 5% (20%) for HY0010 with the slow filter data set (HY0009 with the slow filter and HY0010 with the fast data sets).

The signal to noise ratio for HY0010 with the fast filter is consistent with the estimation in the section 4.4.5, where $S/N = 5.8$. Although the S/N with the slow filter here is more different from the estimation ($S/N = 35$), the measured S/N is still good enough and the difference is allowable. It can be different, since there are differences between the measurement condition in the real measurement and the assumptions in the section 4.4.5, the detector capacitance is 30pF for HY0010 and 70pF in the section 4.4.5, and there are extra noises in real measurement. The estimated noise of preamplifier can also differ from the real one.

The measured performances are satisfactory as to the signal to noise ratio.

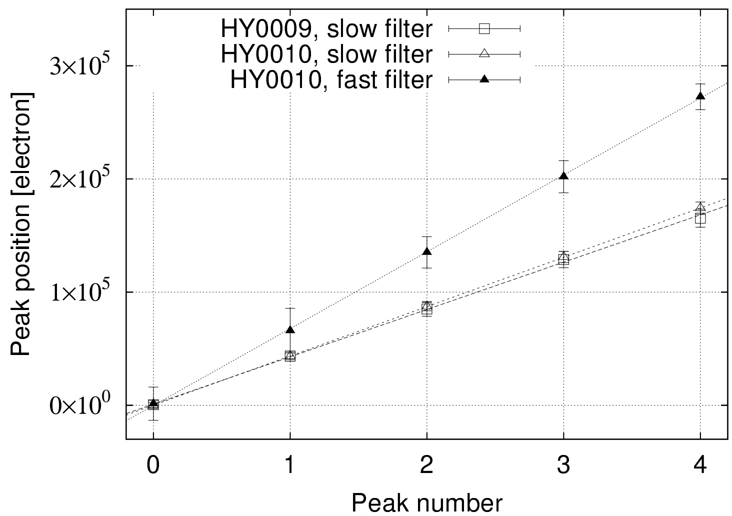


Figure 5.10: The plot of the peak positions Q_{peak} v.s. the peak numbers n_p.

Table 5.1: The gain, noise and the signal to noise ratio for each data set.

Data set	gain G	noise σ_n[electron]	S/N	$\sigma_{\mathrm{residual}}$[electron]
HY0009 with slow filter	4.2×10^4	3.0×10^3	14	2.2×10^3
HY0010 with slow filter	4.4×10^4	2.1×10^3	21	2.2×10^2
HY0010 with fast filter	6.8×10^4	1.1×10^4	6.4	1.8×10^3

T resolution

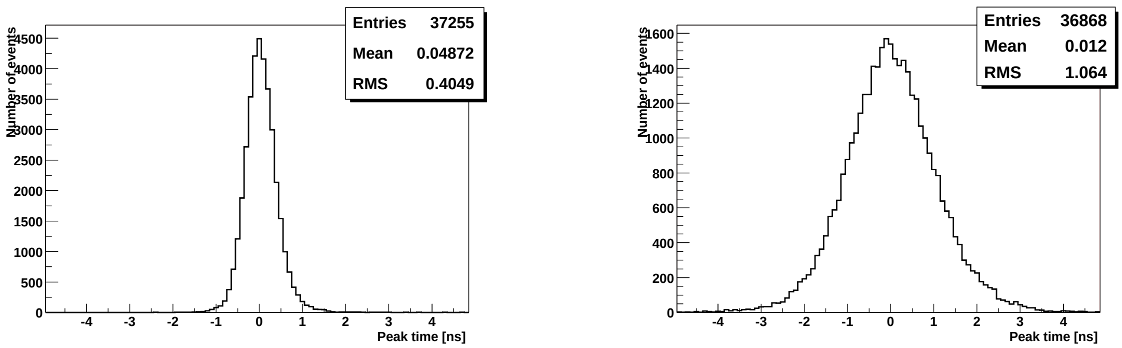
We divide each event into the first, second, third and forth peaks and calculate the timing resolution for each group.

We classify, into n-th peak, the event with the amplitude of Q that meets the following equation

$$G \times (n - 0.5) + Q_{\mathrm{offset}} < Q < G \times (n + 0.5) + Q_{\mathrm{offset}} \tag{5.10}$$

where Q is the amplitude of the event, G and Q_{offset} are gain and offset that we calculate above. After the classification, we calculate the timing T of each event with the method described in the section 4.4.4.

Figure 5.11 shows the obtained T distributions for the first peak (one photon peak) of HY0010 data. Here we ignore underflow and overflow events ($|T| > 5$ns). Such events are considered to be from spike noise with external origins. The contribution of such events are smaller than 1% (0.1%) for HY0009



(a) The result of slow filter.

(b) The result of fast filter.

Figure 5.11: The timing resolution of HY0010 for the first peak.

(HY0010). The RMSs in the figure 5.11 are mostly consistent with the timing resolutions estimated in the section 4.4.5, which resolutions are shown in the figure 4.47, though the RMS in the figure 5.11(b) is a little larger than that in the figure 4.47(b).

We fit each distribution with a double Gaussian and calculate the width (σ) of each distribution using the equation 4.26. In order to avoid the distributions' tail effect, we define the width (or σ) as the timing resolution, not the RMS of each distribution. We calculate the timing resolutions for all peaks of all data sets. The χ^2/n.d.f. for all fittings are from ~ 1 to ~ 6 and hence the fittings are appropriate. Figure 5.12 shows the result. We also check the T distributions for the events with $Q < 0$ for each data set, and confirm that there is no peaking background.

The timing resolution is good enough as a whole, since our goal is a resolution of 1ns.

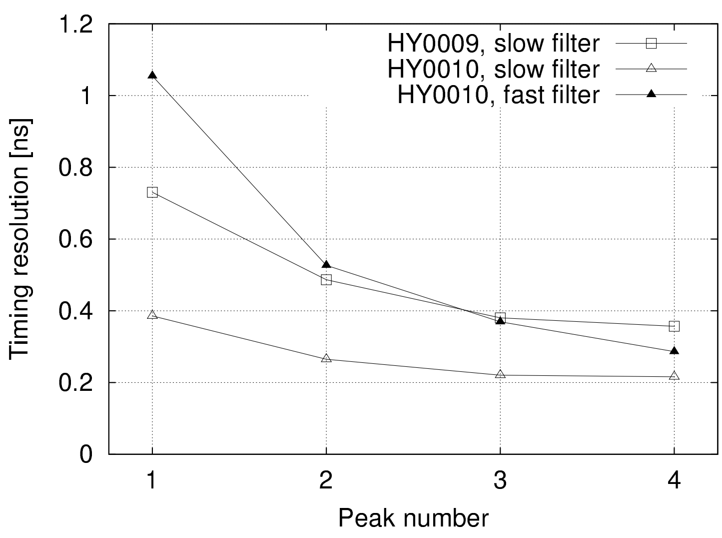


Figure 5.12: The relation between timing resolutions and signal heights.

5.4 The measurement of 13inch HPD prototype

We also have several 13 inch prototype HPDs, whose picture is shown in the figure 5.13. We carry out some basic measurements using one of the prototypes, LHP0022, in HPK by cooperating with HPK staffs.

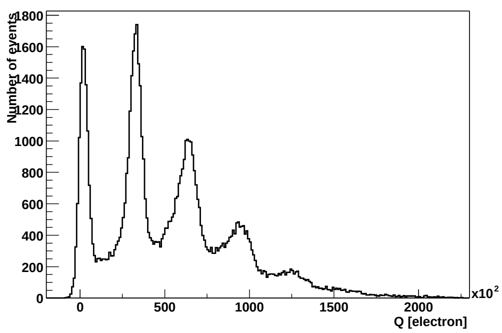
The measurement setup for LHP0022 is different from that for 5inch prototypes. We use the different preamplifier, though it has the similar configuration as we developed. The oscilloscope we use is Hewlett Packard 54522A with a 2GHz sampling rate and a 500MHz analog bandwidth. The other part of the setup, the light source and the method to acquire the data from the oscilloscope, is mostly the same as that for the 5inch prototypes. The applied HV is +8.5kV and the bias on AD is 360V. Another important difference is that we have not done the calibration with the real detector, but with a capacitor of 70pF, which emulates the detector capacitance. This can result in an error factor of the calibration.

Using the acquired signal, noise, and calibration data, we make the slow filter, apply it to the signal data, and calculate the signal amplitude Q using the calibration data. We repeat all processes used for the 5 inch prototype measurement in the same manner. The resultant Q distribution is shown in the figure 5.13. The single photon peak can be clearly seen in the figure. We calculate the gain and the noise in the same way as we do for the 5 inch prototypes. Table 5.2 shows the results of the calculation and the signal to noise ratio calculated from them, together with the RMS of the residuals ($\sigma_{\mathrm{residual}}$) of the fitting.

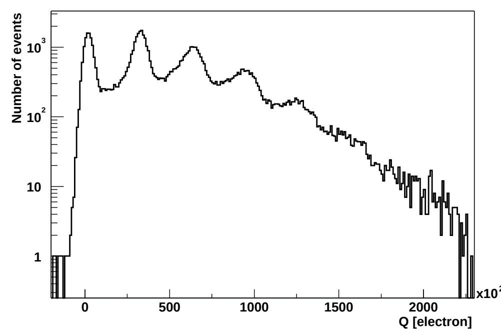
Figure 5.13: The picture of 13inch prototype HPD.

Table 5.2: The gain, noise and the signal to noise ratio for the data of 13 inch prototype LHP0022.

Data set	gain G	noise σ_n[electron]	S/N	$\sigma_{\mathrm{residual}}$[electron]
HLP0022	3.1×10^4	2.4×10^3	12.7	2.4×10^2



(a) In linear scale.



(b) In log scale.

Figure 5.14: Distributions of Q(signal amplitude) for LHP0022 with slow filter.

Chapter 6

Summary and future prospects

6.1 Summary

We have developed the 5 inch prototype HPD and its readout system. The performances of the prototype HPD are satisfactory, with a signal speed of ∼6ns and a multiplication gain of $\sim 4\times10^4$. The development of the readout system has also been proceeded successfully. We have developed the preamplifier with the wide bandwidth and the good noise performance. The preamplifier's rise time is ∼4ns, and its noise is ∼3,000[electron] for 70pF detector capacitance. Its dynamic range is ∼700p.e. equivalent and it meets our requirement. As for the waveform sampler, the development of prototype Analog Memory Cell (AMC) has been done and we have confirmed the essential functions of the AMC work well. We have also found several items to be improved, and the next version designing is ongoing based on the evaluation. The algorithms of the Digital Signal Processing (DSP) have also been studied. We have adopted the strategy to use two types of digital filters: the slow filter, which is optimized for the signal to noise ratio (S/N) and timing resolution (σ_t), and the fast filter, which is optimized to minimize the dead time (T_{dead}) of the photo sensor system. Using simulated signal and noise with realistic spectra, we have checked the performances of the DSP. It shows good performances: $S/N = 35$ and $\sigma_t = 0.4$[ns] with the slow filter, and $S/N = 5.8$, $\sigma_t = 0.9$ and $T_{\mathrm{dead}} = 7.5$[ns] with the fast filter, for a single photon equivalent input.

We have evaluated the performance of the present 5 inch prototype HPD and the prototype readout system. The timing resolution is ∼ 0.4ns (∼ 1ns) with the slow (fast) filter. The dead time of the fast filter is ∼7.5ns. The performances are satisfactory. We have also acquired the data for the 13 inch prototype HPD and have shown the preliminary result. We have detected the single photon even with the 13 inch HPD.

6.2 Future prospects

As described above, the development of the 13 inch prototype HPD is steadily going on. The completion versions of the prototype will be available soon. The development of AMC is also proceeding. The next version with 512∼1024 stages is under planning, and we will submit it soon. As to the DSP, we are planning to implement it with FPGA. We have another choice to use PC based system,[37, 38] and the choice is also under review.

Acknowledgement

I would like to express my great thanks to my supervisor H. Aihara. I am especially thankful to M. Iwasaki.

I greatly acknowledge M. Tanaka and the other members of the KEK electronics group. I would like to express my appreciation to following people for their help: Y. Fujita (KEK online group). M. Shiozawa and Y. Itow (ICRR). Y. Kawai, M. Suyama and the other members of the electron tube center (ETC) of Hamamatsu Photonics K.K. (HPK).

I also thank H. Kawai, K. Tanabe, R. Ishida, K. Ito, Y. Nakahama, H. Nakayama, T. Higuchi, J. Tanaka, M. Yokoyama, T. Nakadaira, T. Tomura, N. Uozaki and Y. Yamashita.

Appendix A

The calculations on the electronics

A.1 The noise of electronics

Here we express the current and voltage in upper case, like I_x or V_x and the noise of them in lower case, like i_x or v_x. The definitions of the noise i and v are

$$i^2 = S_i(f)\Delta f, \quad v^2 = S_v(f)\Delta f \tag{A.1}$$

where $S_i(f)$ and $S_v(f)$ are the *noise spectral density* or the *noise power spectrum* and Δf is the bandwidth. $S_i(f)$ and $S_v(f)$ have the dimension of [A^2/Hz] and [V^2/Hz]. The multiplication by Δf has to be replaced by integral if the bandwidth is not small enough compared with the change of the $S(f)$ in that region.

A.1.1 The noise of the bipolar junction transistor (BJT)

The noise of Bipolar Junction Transistor (BJT) is described here. In BJT, the current flow of the collector current I_c and the base current I_b can be modeled as random processes. Consequently, I_c and I_b shows *full shot noise*, that is, the shot noise i_{cshot} and i_{bshot} can be written as:

$$i_{cshot}^2 = 2eI_c\Delta f, \quad i_{bshot}^2 = 2eI_b\Delta f \tag{A.2}$$

where e is the electron charge (1.6×10^{-19}C). BJT has a physical resistor at base, which is called base spreading resistor $R_{bb'}$, and the resistor is the source of thermal noise. We can write the noise as:

$$v_{bthermal}^2 = 4kTR_{bb'}\Delta f \tag{A.3}$$

where k is the Boltzmann constant and T is the temperature of the transistor.

BJT also has the flicker noise and the burst noise and they have been found empirically to be represented by the current noise of I_b. Here we take them together approximately and call it $1/f$ noise. The $1/f$ noise can be written as:

$$i_{b1/f}^2 = K\frac{I_b^a}{f^b}\Delta f \tag{A.4}$$

where K is the coefficient for a particular device, and a and b are constant and are close to 1.

Taking the above noise sources together and ignoring the other noise sources (they are negligible, in practice), the noise can be written as follows:

$$\begin{aligned} v_b^2 &= 4kTR_{bb'}\Delta f && \text{(A.5)} \\ i_c^2 &= 2eI_c\Delta f && \text{(A.6)} \\ i_b^2 &= 2eI_b\Delta f + K\frac{I_b^a}{f^b}\Delta f && \text{(A.7)} \end{aligned}$$

We can simply sum up squared noise since they are from independent physical processes. The schematic drawing of this noise model of BJT is shown in the figure A.1. Note that the base resistance $R_{bb'}$ is written as the resistance outside the transistor.

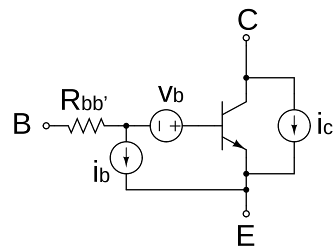


Figure A.1: The noise model of BJT.

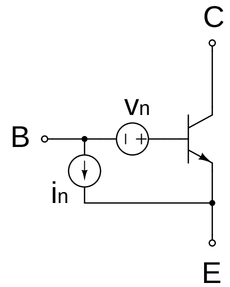


Figure A.2: The noise model of BJT with noise sources transfered to input.

Then we calculate the input equivalent noises, which are schematically shown in the figure A.2, from the noises above. We can write the input equivalent noise i_n and v_n as:

$$v_n^2 = 4kTR_{bb'}\Delta f + \frac{2eI_c}{{g_m}^2}\Delta f + \left(2eI_b + K\frac{I_b^a}{f^b}\right){R_{bb'}}^2\Delta f \tag{A.8}$$

$$i_n^2 = 2eI_b\Delta f + K\frac{I_b^a}{f^b}\Delta f \tag{A.9}$$

where g_m is the transistor's transconductance from base to collector.

As far as we concern about a negative feedback amplifier, the bandwidth of the amplifier is always smaller than the transistor's gain bandwidth f_T. Thus, we do not have to consider about the contribution from i_c to i_n. Note that there is correlation among the third term of v_n^2 and the first and second terms of i_n^2. Thus, we cannot simply sum up the noise contributions from v_n^2 and i_n^2 when the contributions are comparable. Also, note that the base spreading resistance $R_{bb'}$ is now included inside the transistor.

A.1.2 The propagation of the noise in a negative feedbacked amplifier

The input device, which is mostly a transistor, usually determines the noise characteristic of an amplifier. We calculate the propagation of the noise of the input device in a negative feedback amplifier. The input equivalent current (voltage) noise is written as i_n (v_n), here.

We define the signal current as I_{sig} and output voltage as V_{out}. Then the contribution of I_{sig}, i_n and v_n to V_{out} are written as:

$$V_{\mathrm{out}} = K_0 I_{\mathrm{sig}} + K_1 i_n + K_2 v_n \tag{A.10}$$

where K_i are determined by characteristics of amplifiers. Note that K_i depend on the frequency. Then the input equivalent noise i_{eq} is written as follows:

$$i_{\mathrm{eq}} = \left|\frac{K_1}{K_0}\right|^2 {v_n}^2 + \left|\frac{K_2}{K_0}\right|^2 {i_n}^2 \tag{A.11}$$

Assuming certain values of i_n and v_n, we only have to calculate K_i. We use the schematic model of the negative feedbacked amplifier written in the figure A.3, with open loop gain A, detector impedance Z_D and feedback impedance Z_f.

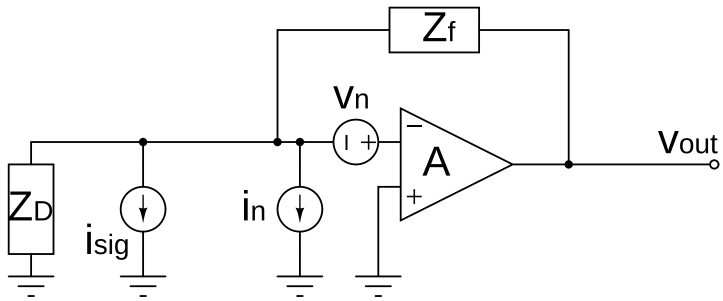


Figure A.3: The amplifier model for calculating the noise propagation.

From this figure, one can easily see that I_{sig} and i_n contribute to v_{out} in the same way and hence:

$$K_0 = K_2 \tag{A.12}$$

As for the current through Z_D and Z_f, using Kirchhoff's rule, we get the following equations:

$$i_D + i_f = i_n + I_{\text{sig}} \tag{A.13}$$
$$-i_D Z_D = V_{\text{out}} - i_f Z_f \tag{A.14}$$

where I_D and i_f are current through Z_D and Z_f, respectively. Using the open loop gain of the amplifier, we have the relation:

$$-A\left(-i_D Z_D + v_n\right) = v_{\text{out}} \tag{A.15}$$

Making three equations together, we get:

$$v_{\text{out}}\left(1 + \frac{Z_D + Z_f}{AZ_D}\right) = -I_{\text{sig}} Z_f - i_n Z_f + \frac{Z_D + Z_f}{Z_D} v_n \tag{A.16}$$

thus, ignoring the term of order $1/A$, we get:

$$K_0 = -Z_f \tag{A.17}$$
$$K_1 = \frac{Z_D + Z_f}{Z_D} \tag{A.18}$$

Consequently, the input equivalent noise is written using the noise of the input device, i_n and v_n, as:

$$i_{\text{eq}}{}^2 = \left|\frac{1}{Z_D} + \frac{1}{Z_f}\right|^2 v_n{}^2 + i_n{}^2 \tag{A.19}$$

The noise in output v_{out} is written as:

$$v_{\text{out}}{}^2 = \left|\frac{Z_D + Z_f}{Z_D}\right|^2 v_n{}^2 + \left|Z_f\right|^2 i_n{}^2 \tag{A.20}$$

A.2 Estimation of base spreading resistance using S parameter

We can estimate the base spreading resistance, $R_{bb'}$, of a transistor, using the S parameter of the transistor. The method is described here.

The process of the estimation is summarized as follows:

1. Convert S parameter, i.e., **S** matrix, to admittance matrix **Y**.
2. Deduce $R_{bb'}$ from **Y**, since $R_{bb'}$ is one of the parameters that determine **Y**.

For transistors, we have two ports and hence **S** and **Y** are 2×2 matrices. Usually, port 1 and 2 corresponds to base and collector port respectively, with common emitter configuration. The relation between **S** and **Y** can be written as:

$$\mathbf{Y} = (1 - \mathbf{S}) \cdot (1 + \mathbf{S})^{-1} / Z_0 \tag{A.21}$$

where commonly $Z_0 = 50[\Omega]$ is used.[i] The definition of **Y** is:

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \mathbf{Y} \cdot \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \tag{A.22}$$

where V_1, V_2, I_1, and I_2 are the voltage applied to the base, the voltage applied to the collector, the base current, and the collector current, respectively. Note that V_i and I_i are not DC voltage and current but AC, since here **S** and **Y** are defined in the frequency domain. We can relate **Y** and transistor parameters as

$$\mathbf{Y} = \begin{pmatrix} R_{bb'} + Z_\pi & 0 \\ \frac{Z_\pi}{Z_\pi + R_{bb'}} g_m & r_o \end{pmatrix} \tag{A.23}$$

$$\left(Z_\pi = \frac{r_\pi}{1 + 2\pi j f r_\pi C_\pi} \right) \tag{A.24}$$

using the simple transistor model shown in the figure A.4. Y_{11} is useful to estimate $R_{bb'}$.

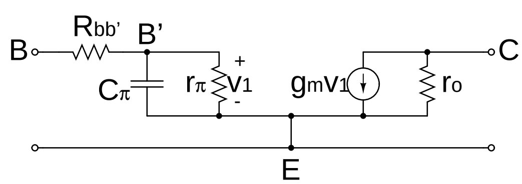


Figure A.4: The simple model for BJT.

$$\mathrm{Re}\left(\frac{1}{Y_{11}}\right) = R_{bb'} + \frac{r_\pi}{1 + (2\pi j f)^2 {\tau_\pi}^2} \tag{A.25}$$

$$\mathrm{Im}\left(\frac{1}{Y_{11}}\right) = -\frac{2\pi j f \tau_\pi r_\pi}{1 + (2\pi j f)^2 {\tau_\pi}^2} \tag{A.26}$$

$$(\ \tau_\pi = r_\pi C_\pi\ ) \tag{A.27}$$

The real part of $1/Y_{11}$ asymptotically converges to $R_{bb'}$ as frequency increases, and hence we can use it to deduce $R_{bb'}$. The imaginary part asymptotically converges to 0, and thus we can use it to check if the transistor is described well using the model in figure A.4. The estimation is valid only when the behavior of the transistor follows the model. Figures A.5 and A.6 show examples of *good* and *bad* $1/Y_{11}$ function shapes. That is, figure A.5 matches with the model in the figure A.4, while figure A.6 does not.

[i] Z_0 is not scalar value but a matrix in general.

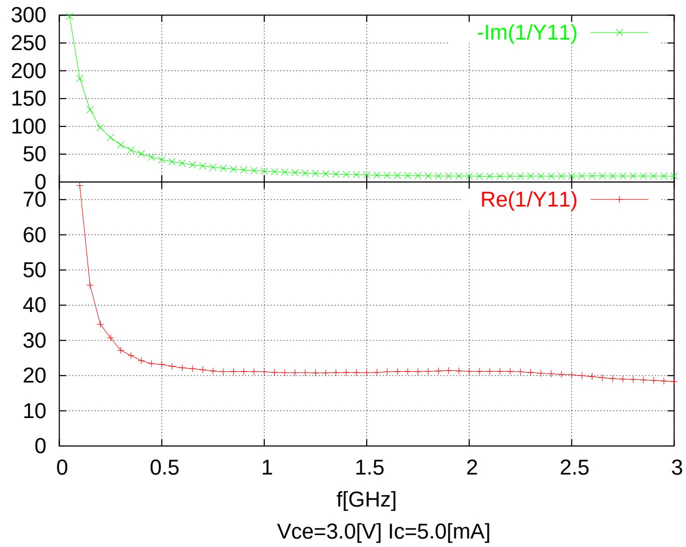


Figure A.5: The example that matches with the simple model: 2SC5086

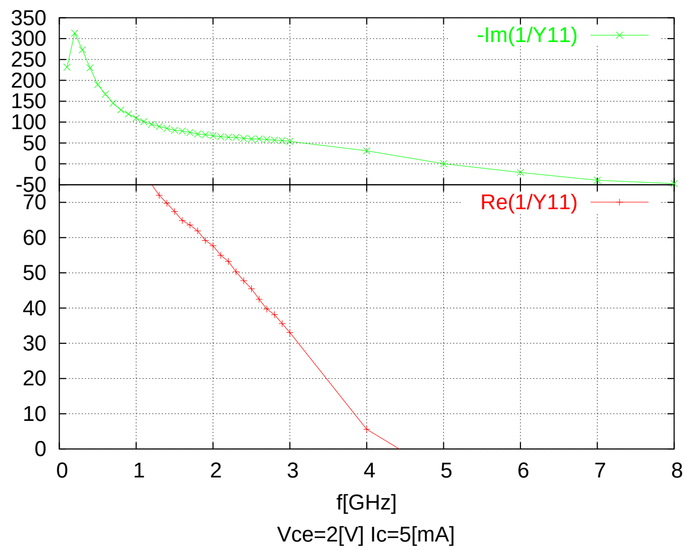


Figure A.6: The example that does not match with the simple model: 2SC5667

Appendix B

The algorithms for the digital filter

B.1 The algorithms for the slow filter

B.1.1 Matched filter

First, we have to define the *signal to noise ratio*. We define $x_s[n]$ and $x_n[n]$ as signal waveforms of *signal (without noise)* and *noise* before filtering, and $h[n]$ is a filter kernel. By definition, output waveforms of signal and noise after filtering, $y_s[n]$ and $y_n[n]$, are described as follows.

$$y_s[n] = (h \otimes x_s)[n] \tag{B.1}$$
$$y_n[n] = (h \otimes x_n)[n] \tag{B.2}$$

Then the noise σ_n is described as follows.

$$\sigma_n \equiv \sqrt{\frac{1}{N-1}\sum_j y_n[j]^2} \tag{B.3}$$
$$= \sqrt{\frac{2}{N(N-1)}\sum_l |Y_n[l]|^2} \tag{B.4}$$
$$= \sqrt{\frac{2}{N(N-1)}\sum_l |H[l]|^2 \, |X_n[l]|^2} \tag{B.5}$$

We used Parseval's Relation in the second equality. The peak value P of the waveform of signal $y_s[n]$ is described as follows.

$$P \equiv \max_j y_s[j] \tag{B.6}$$
$$= \frac{2}{N}\max_j\left(\sum_l Y_s[l]e^{2\pi i\cdot jl/N}\right) \quad (Y[-l] \equiv Y[N-l]) \tag{B.7}$$

We can take j of the maximum point as $j = 0$ without losing generality, by shifting the origin of time. Then, the equation B.7 can be rewritten as follows.

$$P = \frac{2}{N}\sum_l Y_s[l] \tag{B.8}$$
$$= \frac{2}{N}\sum_l H[l]\; X_s[l] \tag{B.9}$$

Note that P is always real number. Then we get our definition of the signal to noise ratio, S/N, as follows.

$$
\begin{aligned}
S/N &\equiv \frac{P}{\sigma_n} && \text{(B.10)} \\
&= \sqrt{2\left(1-\frac{1}{N}\right)}\,\frac{\sum_l H[l]\;X_s[l]}{\sqrt{\sum_l |H[l]|^2\,|X_n[l]|^2}} && \text{(B.11)}
\end{aligned}
$$

Consequently, we have replaced optimization of S/N with that of $\mathcal{F}$, which is a functional of a filter kernel in the frequency domain, $H[k]$ for given $X_s[k]$ and $X_n[k]$.

$$
\mathcal{F}[H] \equiv \frac{\sum_l H[l]\;X_s[l]}{\sqrt{\sum_l |H[l]|^2\,|X_n[l]|^2}} \tag{B.12}
$$

First, we optimize $\mathcal{F}$ by changing the phase of $H[k]$, keeping $|H[k]|$ constant. Here, the denominator of $\mathcal{F}[H]$ in the equation B.12 is obviously constant, being independent of the phases of $H[k]$. Since P is a real number, the upper limit of the numerator can be written as follows.

$$
\sum_l H[l]\;X_s[l]\;\leq \sum_l |H[l]|\;\;|X_s[l]| \tag{B.13}
$$

The equality is realized in the following condition.

$$
\arg\left(H[l]\right) = \arg\left(X_s[l]^*\right) \quad (\text{for } \forall l) \tag{B.14}
$$

This is the condition to get optimal $\mathcal{F}[H]$ when we keep $|H[k]|$ constant.

Then we optimize $\mathcal{F}$ by changing the absolute value of $H[k]$, keeping $\arg\left(H[k]\right)$ to satisfy the equation B.14. For simplicity, we assume *white noise*, here. Since the normalization of $X_s[k]$ does not influence the optimization of $\mathcal{F}[H]$, we can write the assumption as follows.

$$
X_n[l] = 1 \tag{B.15}
$$

Also, $\mathcal{F}[H]$ does not depend on the normalization of $H[k]$ and hence we took the normalization of $H[k]$ as follows.

$$
\sum_l |H[l]|^2 = 1 \tag{B.16}
$$

Using the equations B.12, B.14, B.15 and B.16, we rewrite $\mathcal{F}[H]$ as follows.

$$
\mathcal{F}[H] = \sum_l |H[l]|\;\;|X_s[l]| \tag{B.17}
$$

It is easy to find the condition on $H[k]$ to maximize the right-hand side the equation B.17. We can consider the right-hand side of the equation B.17 as an inner product of vectors $|H[k]|$ and $|X_s[k]|$ where k is the index of the vectors. $|H[k]|$ is a norm vector because of the normalization in the equation B.16. Then the condition to maximize $\mathcal{F}[H]$ in the equation B.17 is obvious: make $|H[k]|$ to be in parallel with $|X_s[k]|$, or

$$
|H[l]| = a\,|X_s[l]| \quad (\text{for } \forall l) \tag{B.18}
$$

where a is a constant.

The equation B.14 and B.18 are the condition to get the optimal signal to noise ratio. Note that although once we have assumed the normalization written in the equation B.16, the result is not dependent on it. Thus, we can take $a = 1$ in the equation B.18. Then, we can put the equation B.14 and B.18 together as follows.

$$
H[l] = X_s[l]^* \quad (\text{for } \forall l) \tag{B.19}
$$

This is the definition of the filter kernel of the *Matched filter*.

Here we briefly describe why the filter is called the *Matched* filter. By a property of Fourier transformation, the complex conjugate in the frequency domain corresponds to the time reversal in the time domain. Thus, the filter kernel $h[n]$ that corresponds to $H[k]$ in the equation B.19, is the *time reversal* of the waveform of signal, $x_s[n]$. Since *filtering* is the convolution of $h[n]$ and $x_s[n]$ as written in the equation 4.16, "filtering with the inversion of the waveform of *signal* $x_s[n]$" is equivalent to "taking an inner product with the waveform of *signal* $x_s[n]$." In other words, the filtering with Matched filter is to calculate how the input signal matches with the waveform of *signal (without noise)*, $x_s[n]$. It is intuitive that processing like this enhances *signal* while suppressing *noise*.

The Matched filter, which is defined by the equation B.19, is therefore the optimal filter when the noise spectrum is white.

B.1.2 Optimal matched filter

In the deduction of the Matched filter, we have assumed *white noise*. Here we consider about the optimal filter without such an assumption. Note that we have not needed the white noise assumption to deduce the equation B.14. Thus, we only have to get the condition on $|H[k]|$ without the white noise assumption.

With the condition of the equation B.14, we rewrite the equation B.12 as follows.

$$\mathcal{F}[H] = \frac{\sum_l |H[l]|\ |X_s[l]|}{\sqrt{\sum_l |H[l]|^2 |X_n[l]|^2}} \tag{B.20}$$

Our task is to maximize $\mathcal{F}[H]$ for given $|X_s[k]|$ and $|X_n[k]|$ by changing $|H[k]|$. Since $\mathcal{F}[H]$ does not depend on the normalization of $|H[k]|$, we can take following normalization.

$$\sum_l |H[l]|^2 |X_n[l]|^2 = C \tag{B.21}$$

Then we can replace the problem of optimization of $\mathcal{F}[H]$ with the maximization of $\mathcal{F}'[H]$ defined in the following

$$\mathcal{F}'[H] = \sum_{l=-N/2}^{N/2} |H[l]|\ |X_s[l]| \tag{B.22}$$

with the restraint condition of the equation B.21. This problem can easily be solved using the method of Lagrange's Multipliers, as

$$|H[l]| = \lambda \frac{|X_s[l]|}{|X_n[l]|^2} \quad (\text{for } \forall l) \tag{B.23}$$

where λ is a Lagrange's undetermined multiplier. Since normalization of $|H[k]|$ is arbitrary, we can determine C so that λ equals to 1.

Putting the equation B.14 and B.23 together, we got $H[k]$ to maximize the signal to noise ratio as follows.

$$H[l] = \frac{X_s[l]^*}{|X_n[l]|^2} \quad (\text{for } \forall l) \tag{B.24}$$

Note that the equation B.24 is equivalent to the equation B.19 if we assume white noise. This result suits our anticipation.

We call the filter defined by the equation B.24 the *optimal matched filter.*

B.1.3 On timing resolution

Though it is plausible to consider the filter with the best S/N is the filter with the best timing resolution, it is not so obvious that we cannot take the anticipation as true without any confirmation. Thus, we have considered how the S/N is related to the timing resolution.

To attack this problem, we have assumed the following two assumptions.

1. Using a parabolic function, we can describe, the neighborhood (compared with the length of timing resolution) of the peak of the waveform of signal (without noise) after the filtering.

2. Using a linear function, we can describe the waveform of noise in narrow regions compared with the length of timing resolution.

These assumptions are realized if the bandwidth of the filter is narrow compared with the timing resolution. For example, when the timing resolution is 1ns, the assumptions are realized if the bandwidth of the filter satisfies $B \ll 1\text{GHz}$, where B is the bandwidth. Our situation is close to this condition, while expected timing resolution is $\sim$ 1ns and we expect the bandwidth of our filter to be smaller than the sampling rate, 1GHz $\sim$ 2GHz, at least.

If the assumptions are true, we can write the signal and noise waveform after the filtering as

$$y_s[j] = a \cdot j^2 + C_1 \tag{B.25}$$

$$y_n[j] = b \cdot j + C_2 \tag{B.26}$$

where a and b are constant. We have defined the peak position of signal as the origin of time. Note that b differs event by event, while a is really *constant*, in the statistical meaning. If the signal $y_s[n]$ is smeared with the noise $y_n[n]$, the waveform we measure is $y_s[n] + y_n[n]$. The noise then shifts the position of peak from $j = 0$ to $j = -b/2a$. Thus the timing resolution σ_t is defined as follows.

$$\sigma_t \equiv \sqrt{\left\langle \left(\frac{b}{2a}\right)^2 \right\rangle} = \frac{\sqrt{\langle b^2 \rangle}}{2|a|} \tag{B.27}$$

Then we think of how to relate a and b to the spectra of signal and noise. Using the equation of inverse DFT 4.19, we can relate a and b with the spectra of signal $Y_s[k]$ and noise $Y_n[k]$ as follows.

$$a = \frac{y_s[1] + y_s[-1] - 2y_s[0]}{2} \tag{B.28}$$

$$= -\frac{4}{N} \sum_l Y_s[l] \, \sin^2\left(\frac{\pi l}{N}\right) \tag{B.29}$$

$$b = y_n[1] - y_n[0] \tag{B.30}$$

$$= \frac{2i}{N} \sum_l Y_n[l] \, e^{\pi i l/N} \sin\left(\frac{\pi l}{N}\right) \tag{B.31}$$

Since b is real, $\arg(Y_n[l])$ is random, and there is no correlation between $\arg(Y_n[l_1])$ and $\arg(Y_n[l_2])$ for $l_1 \neq l_2$, we can calculate $\langle b^2 \rangle$ as follows.

$$\langle b^2 \rangle = \frac{4}{N^2} \sum_l |Y_n[l]|^2 \sin^2\left(\frac{\pi l}{N}\right) \tag{B.32}$$

Using the equations B.27, B.29 and B.32, we rewrite the timing resolution σ_t as

$$\sigma_t = \frac{1}{4} \frac{\sqrt{\sum_l |Y_n[l]|^2 \sin^2\left(\frac{\pi l}{N}\right)}}{\sum_l Y_s[l] \sin^2\left(\frac{\pi l}{N}\right)} \tag{B.33}$$

$$= \frac{1}{4} \frac{\sqrt{\sum_l |H[l]|^2 \, |X_n[l]|^2 \sin^2\left(\frac{\pi l}{N}\right)}}{\sum_l H[l] X_s[l] \sin^2\left(\frac{\pi l}{N}\right)} \tag{B.34}$$

Our task is to minimize σ_t, or to maximize $1/\sigma_t$. The $1/\sigma_t$ has a similar form to the $\mathcal{F}[H]$ in the equation B.12. The essential difference is that $X_s[l]$ and $|X_n[l]|$ in the equation B.12 are replaced with $X_s[l] \sin^2(\pi l/N)$ and $|X_n[l] \sin(\pi l/N)|$ in the equation B.34, respectively. Thus, we can do the same argument as we calculated the $H[k]$ that maximize $\mathcal{F}[H]$. We can then calculate the $H[k]$ that maximize $1/\sigma_t$, applying the replacement to the equation B.24. The result is as follows.

$$H[l] = \frac{X_s[l]^* \sin^2(\pi l/N)}{|X_n[l]|^2 \sin^2(\pi l/N)} = \frac{X_s[l]^*}{|X_n[l]|^2} \quad (\text{for } \forall l) \tag{B.35}$$

This result is the same as the equation B.24. It suits our first anticipation that we probably get the best timing resolution when we maximize the signal to noise ratio.

Note that we assumed the two assumptions to calculate the equation B.35. The assumptions are only approximately true in practice. The condition to get the best signal to noise ratio may therefore not be the same as that to get *the best* timing resolution. However, we have confirmed, at least, that the condition to get the best signal to noise ratio leads to *better* timing resolution.

B.2 The algorithms for the fast filter

B.2.1 Deconvolution filter

The Deconvolution filter is a filter to remove the effect of unwanted convolution. For example, an output waveform of a preamplifier connected to a HPD, $x_{\mathrm{out}}[n]$, is a convolution of the impulse response of the preamplifier $h_p[n]$ and a raw signal waveform from HPD $x_{\mathrm{HPD}}[n]$.

$$x_{\mathrm{out}}[n] = (h_p \otimes x_{\mathrm{HPD}})[n] \tag{B.36}$$

Using DFT, we can replace the convolution with multiplication.

$$X_{\mathrm{out}}[l] = H_p[l] \times X_{\mathrm{HPD}}[l] \tag{B.37}$$

Then we think of the filter kernel $h_{\mathrm{Deconv}}[n]$ to recover $x_{\mathrm{HPD}}[n]$ from $x_{\mathrm{out}}[n]$. From the equation B.37, we can easily calculate the DFT result of $h_{\mathrm{Deconv}}[n]$, $H_{\mathrm{Deconv}}[k]$, as follows.

$$H_{\mathrm{Deconv}}[l] = \frac{1}{H_p[l]} \tag{B.38}$$

Then the following equation is obviously realized.

$$X_{\mathrm{HPD}}[l] = H_{\mathrm{Deconv}}[l] \times X_{\mathrm{out}}[l] \tag{B.39}$$

By the filter defined in the equation B.38, the signal $X_{\mathrm{out}}[k]$ or $x_{\mathrm{out}}[n]$ recovered the original shape: $X_{\mathrm{HPD}}[k]$ or $x_{\mathrm{HPD}}[n]$. This is the basic idea of the Deconvolution filter.

As an application of the basic idea, we can make a filter $h_D[n]$ that generates an arbitrary desired signal output $y_{des}[n]$ from an arbitrary input signal $x_s[n]$. The filter kernel in the frequency domain $H_D[k]$ is easily calculated using the input and output signals in the frequency domain, $X_s[k]$ and $Y_{des}[k]$ as follows.

$$H_D[l] = \frac{Y_{des}[l]}{X_s[l]} \tag{B.40}$$

The filters of this type are also called the Deconvolution filter in a large meaning.

We can use the Deconvolution filter, using a narrow waveform as the $y_{des}[n]$, to narrower the signal shape from the preamplifier. The Blackman Window shown in the figure B.1 is an example of the narrow waveform.

B.2.2 Wiener optimal filter

The discussion above has been done without noise. When there is no noise, we can perfectly get the same signal waveform as $y_{des}[n]$. On the other hand, we cannot do that perfectly from a signal waveform contaminated with random noise. Then the issue is that from a contaminated input signal $x_{s+n}[n]$, how to make the signal *closest* to the desired signal shape $y_{des}[n]$. Here $x_{s+n}[n]$ is defined using the waveform of signal $x_s[n]$ and noise $x_n[n]$ as follows.

$$x_{s+n}[n] = x_s[n] + x_n[n] \tag{B.41}$$

The issue is to make the output $y_{s+n}[n]$ to be closest to the desired shape of waveform $y_{des}[n]$, using a filter kernel $h[n]$. Here the $y_{s+n}[n]$ is defined as:

$$y_{s+n}[n] \equiv (h \otimes x_{s+n})[n]\ . \tag{B.42}$$

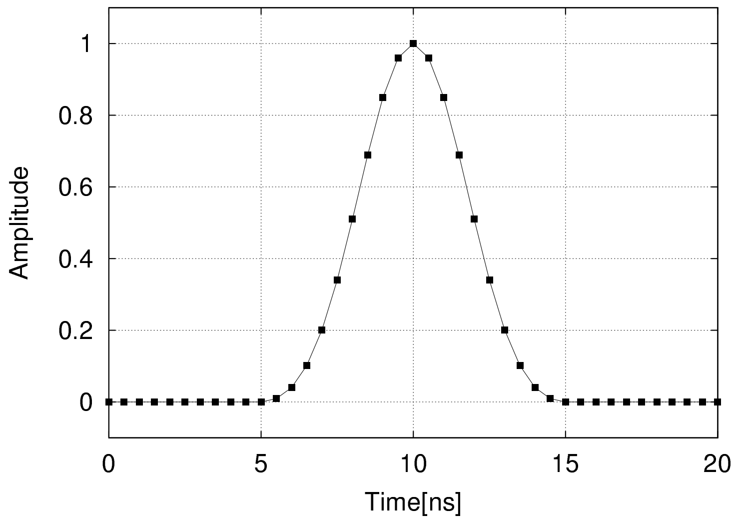


Figure B.1: A signal waveform of the Blackman Window with the width of $\sim$ 4ns.

First, we define the meaning of the word *close*. Our definition of *close* is that in *the least-square sense*, that is, $y_{s+n}[n]$ and $y_{des}[n]$ is close if the following sum is small.

$$\sum_j \left| y_{s+n}[j] - y_{des}[j] \right|^2 \tag{B.43}$$

Second, we divide the filter kernel $h[n]$ into two parts, $h_D[n]$ and $h_W[n]$. $h_D[n]$ is the Deconvolution filter part and $h_W[n]$ is the Wiener filter part. $h_D[n]$ is defined by the following equation.

$$y_{des}[n] = (h_D \otimes x_s)[n] \tag{B.44}$$

Note that $y_{des}[n]$ is given. The definition of $h_W[n]$ is as follows.

$$h[n] = (h_W \otimes h_D)[n] \tag{B.45}$$

The definitions may be more clear in the frequency domain.

$$H_D[l] \equiv \frac{Y_{des}[l]}{X_s[l]} \tag{B.46}$$

$$H_W[l] \equiv \frac{H[l]}{H_D[l]} \tag{B.47}$$

Then we rewrite the equation B.43 using the Perseval's Relation as follows.

$$\sum_j \left| y_{s+n}[j] - y_{des}[j] \right|^2 = \frac{2}{N} \sum_l \left| Y_{s+n}[l] - Y_{des}[l] \right|^2 \tag{B.48}$$

$$= \frac{2}{N} \sum_l \left| H_D[l]\, H_W[l]\; (X_s[l] + X_n[l]) - H_D[l]\, X_s[l] \right|^2 \tag{B.49}$$

Then the issue is to minimize the functional $\mathcal{F}[H_W]$ defined as

$$\mathcal{F}[H_W] \equiv \sum_l \left| H_D[l]\, H_W[l]\; (X_s[l] + X_n[l]) - H_D[l]\, X_s[l] \right|^2 \tag{B.50}$$

for given $X_s[k]$, $X_n[k]$ and $H_D[k]$[i].

[i] $H_D[k]$ is given since $Y_{des}[k]$ and $X_s[k]$ are given.

Since there is no correlation between $X_s[k]$ and $X_n[k]$, we rewrite the equation B.50 as follows.

$$\mathcal{F}[H_W] = \sum_l \left|H_D[l]\right|^2 \left|\left(H_W[l]-1\right)X_s[l] + H_W[l]\,X_n[l]\right|^2 \tag{B.51}$$

$$= \sum_l \left|H_D[l]\right|^2 \left\{\left|H_W[l]-1\right|^2 |X_s[l]|^2 + |H_W[l]|^2\,|X_n[l]|^2\right\} \tag{B.52}$$

Differentiating $\mathcal{F}[H_W]$ with respect to H_W, and setting the result equal to zero, gives the following equation.

$$H_W[l] = \frac{\left|X_s[l]\right|^2}{\left|X_s[l]\right|^2 + \left|X_n[l]\right|^2} \tag{B.53}$$

This is the definition of the Wiener optimal filter.

B.2.3 On the normalization of the Wiener optimal filter

Here we have to pay attention to the fact that the normalization of $X_s[l]$ and $X_n[l]$ matters in the equation B.53. This situation is different from the filter such as the equation B.24. Speaking more strictly, not the absolute values but the ratio of the normalization of $X_s[k]$ and $X_n[k]$ matters. By tracing back the argument above, we can find out that it is the length of the summation in the equation B.43 that determines the normalization of $X_s[k]$ and $X_n[k]$. Thus, for the signal processing for repetitive signals, e.g., audio processing or part of image processing, we do not have to pay attention to the normalization issue. This is because the ratio of the normalization does not change when we vary the length of the summation. On the other hand, in cases in which signal is not repetitive, like ours, we have to pay attention to the normalization issue.

Since there is no correlation between signal and noise, we rewrite the equation B.43 as follows.

$$\sum_{j=0}^{N-1}\left|y_{s+n}[j] - y_{des}[j]\right|^2 = \sum_{j=0}^{N-1}\left|(h\otimes x_s)[j] + (h\otimes x_n)[j] - y_{des}[j]\right|^2 \tag{B.54}$$

$$= \sum_{j=0}^{N-1}\left|(h\otimes x_s)[j] - y_{des}[j]\right|^2 + \sum_{j=0}^{N-1}\left|(h\otimes x_n)[j]\right|^2 \tag{B.55}$$

The normalization is determined by the summation length of the first and second term of the equation B.55. A simple consideration in the following leads to the conclusion that the normalization simply determined by the equation B.55 is not appropriate, or to be more strict, the equation *does not* determine the normalization at all. Suppose if the N in the equation gets larger and larger. The second term (noise term) of the equation gets larger in proportion to N, while the first term stays constant since the signal is localized in a specific region. Thus, if N gets infinitely large, $X_n[k]$ gets much larger than $X_s[k]$ and $H_W[l] \sim 0$ for all l. This behavior is not the desired one.

The fundamental idea of the Wiener optimal filter is to decrease noise with the trade off that $(h\otimes x_s)[n]$ deviates from $y_{des}[n]$, and to minimize the sum of the deviation and the influence of noise. Here, we have to take care about the fact that the deviation of $(h\otimes x_s)[n]$ from $y_{des}[n]$ is localized in the region where $y_{des}[n]$ is non-zero. Thus, we have to evaluate the magnitude of the noise in the region of the length $\sim N_{des}$. It is therefore plausible for us to replace the equation B.55 as follows

$$\sum_{j=0}^{N-1}\left|(h\otimes x_s)[j] - y_{des}[j]\right|^2 + c_n\frac{N_{des}}{N}\sum_{j=0}^{N-1}\left|(h\otimes x_n)[j]\right|^2 \tag{B.56}$$

where N_{des} is the length of the non-zero region of $y_{des}[n]$ and c_n is a normalization constant with size of ~ 1. Then the definition of the Wiener optimal filter $H_W[k]$ in the equation B.53 also changes as follows.

$$H_W[l] = \frac{\left|X_s[l]\right|^2}{\left|X_s[l]\right|^2 + c_n^2\left|X_n'[l]\right|^2} \tag{B.57}$$

$$\left(X_n'[l] \equiv \frac{N_{sig}}{N}X_n[l]\right) \tag{B.58}$$

We have made the "Deconvolution+Wiener optimal" filter with some c_n and have studied the difference due to c_n. The result is shown in the figure B.2. We have put two sequential signals with amplitudes of one p.e. equivalent and three p.e. equivalent, with an interval of 10ns. When c_n is 0.1, the filter is over optimized to reproduce $y_{des}[n]$ and noise is not efficiently suppressed. When c_n is 10, the filter is over optimized to suppress noise and the output signal does not reproduce $y_{des}[n]$ well. The behavior suits our consideration above.

Note that this result may be partly due to the simple waveform of signal, which exponentially decays in the most regions. When the waveform of signal is more complex, it may be better to replace N_{des} with the length of input signal. Because, $(h \otimes x_s)[n]$ may deviate from $y_{des}[n]$ also in the region where $y_{des}[j]$ is zero but $x_s[j]$ is complex shaped. We have the simple shaped input signal, anyway, and it is suitable for us to use N_{des}, which the result in the figure B.2 supports.

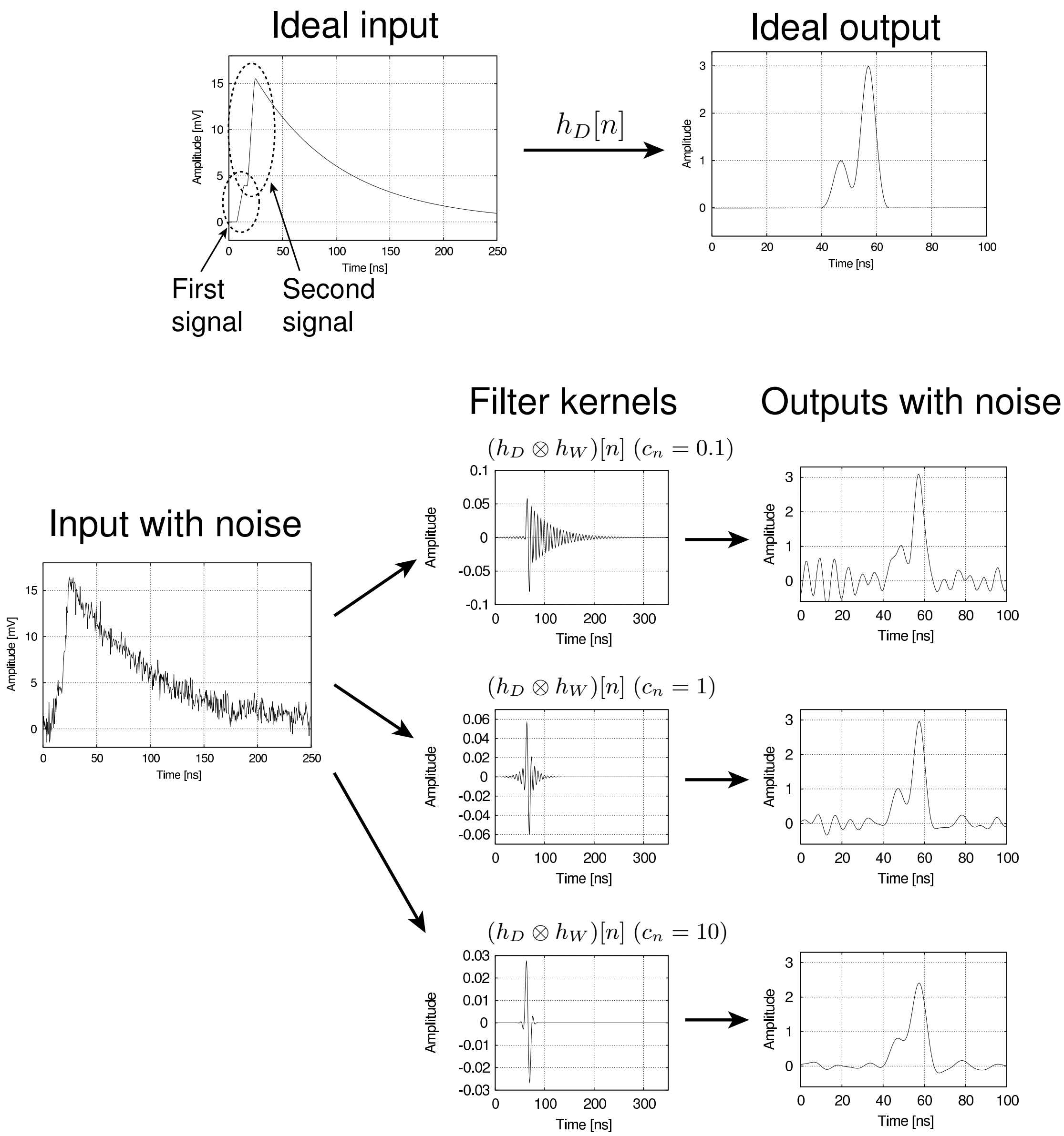


Figure B.2: The Deconvolution filter + Wiener filter with various normalization factor c_n.

Appendix C

Avalanche Diode

C.1 Avalanche multipllication

If an electron or hole is created in, or moved into, a high-field region inside a semiconductor, it may be accelerated strongly enough between collisions to obtain sufficient energy for the creation of an electron-hole pair. An avalanche may thereafter develop.

A single primary electron or hole is assumed to be generated at a position, and the charge carrier is then accelerated by an electric field applied on the semiconductor. At low field strengths, the gain (or the acceleration) in kinetic energy of the charge carriers between collisions is too small to create a secondary electron-hole pair. At collision the kinetic energy of electrons and holes will simply be transferred to the crystal lattice before they are accelerated in the electric field again, and hence no multiplication occurs. (Figure C.1(a))

In contrast, at high field strengths, the energy gain between collisions may be high enough to allow the creation of an electron-hole pair. Part of the kinetic energy is used for the creation of the additional electron-hole pair, and part goes into lattice vibration (or phonon). After the collision, there are two electrons and one hole, as shown in the figure C.1(b). Each of the carriers will be accelerated by the field again, creating with a certain probability further electron-hole pairs in subsequent collisions.

Here we will estimate the gain of the avalanche multiplication. We assume primary electrons are incident at an edge of a depletion region ($x = 0$), which is usually formed in a reverse biased P-N junction, and the primary electrons are accelerated by the electric field toward another edge of the depletion region ($x = L$). In part of the depletion region, the field may be high enough so that multiplication process can occur. Introducing the field-dependent multiplication coefficient α_n (α_p) for electrons (holes), we write the change of electron current $I_n(x)$ and holes current $I_p(x)$ in a thin region dx:

$$dI_n = -dI_p = \alpha_n(x)\, I_n(x)\, dx + \alpha_p(x)\, I_p(x)\, dx \tag{C.1}$$

or

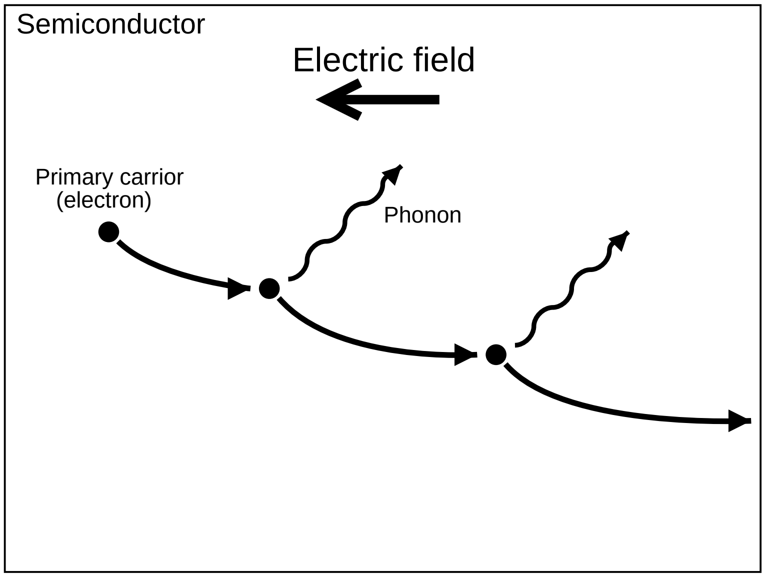
$$\frac{dI_n(x)}{dx} = [\alpha_n(x) - \alpha_p(x)]\, I_n(x) + \alpha_p(x) I \quad (I \equiv I_n + I_p) \tag{C.2}$$

where $\alpha_n(x)$ and $\alpha_p(x)$ are determined by the electric field, temperature, doping concentration, etc., of the point x. Note that the total current $I = (I_p + I_n)$ is constant at steady state, and thus $dI_p = -dI_n$.

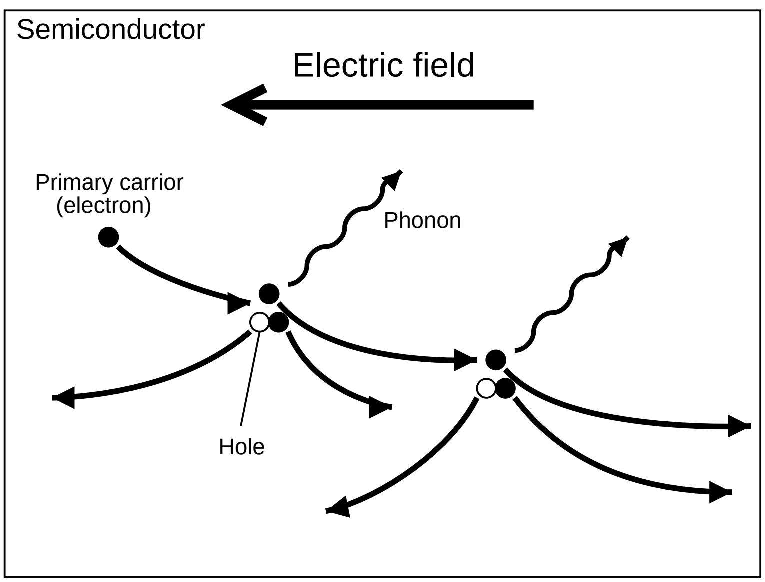
We assume here the region is fully depleted and hence the holes in the region are originating from multiplication processes only. Then the hole current I_p equals zero at the boundary $x = L$ and thus $I_n(L) = I$. Introducing the multiplication gain $M \equiv I_n(L)/I_n(0)$, the solution of the equation C.2 is given by:

$$I_n(x) = I \left\{ \frac{1}{M} + \int_0^x ds\, \alpha_p(s) \exp\left[-\int_0^s dt\, (\alpha_n(t) - \alpha_p(t)) \right] \right\} \Big/ \exp\left[-\int_0^x ds (\alpha_n(s) - \alpha_p(s)) \right] \tag{C.3}$$

where the constant of integration is fixed to be $1/M$ by the boundary condition for $x = 0$. Using the



(a) No multiplication occurs in a low electric field.



(b) Avalanche multiplication occurs in high electric field.

Figure C.1: A schematic drawing of avalanche multiplication.

boundary condition for $x = L$, that is, $I(L) = I$, we get[i]:

$$1 - \frac{1}{M} = \int_0^L ds\, \alpha_n(s) \exp\left[- \int_0^s dt\, (\alpha_n(t) - \alpha_p(t)) \right] \tag{C.4}$$

This is the equation that determines the gain of the avalanche multiplication. We can analytically calculate the integration when α_n and α_p are position independent, as follows:

$$M = \frac{\left(1 - \frac{\alpha_p}{\alpha_n}\right) e^{L(\alpha_n - \alpha_p)}}{1 - \frac{\alpha_p}{\alpha_n} e^{L(\alpha_n - \alpha_p)}} \tag{C.5}$$

C.2 Avalanche Diode

An Avalanche Diode (AD) is a sensor with intrinsic amplification owing to the avalanche multiplication. It is operated at high reverse bias voltage where the avalanche multiplication takes place. As in a PIN diode, the depletion region in an avalanche diode is sensitive to photons and/or charged particles. The electron-hole pairs made by the incident particles are multiplied with the avalanche multiplication.

The avalanche diode for photon detection is often called Avalanche Photo Diode (APD). Since the main use of avalanche diodes is the photon detection, it is often called APD, not AD. We call it AD, however, because we do not detect photons but electrons using the avalanche diode in the HPD.

Avalanche diodes have two operational modes: proportional mode and Geiger mode. The operational mode of an avalanche diode is determined by its material, structure, applied bias voltage, etc. We describe the two operational modes in the following.

C.2.1 Operation in proportional mode

In some materials, such as silicon, germanium, etc., the multiplication coefficient for electrons (α_n) and holes (α_p) are much different. Here we assume $\alpha_n \gg \alpha_p$, which is the case for silicon and germanium. In this case, we can operate the avalanche diode at an electric field strength where only electrons can generate secondary electron-hole pairs, that is, $\alpha_n > 0$ and $\alpha_p \simeq 0$. The amplified signal will be proportional to the primary ionization signal, and hence we call this mode "proportional mode." The ADs in our HPDs are also operated in this mode. The multiplication gain here is calculated from the equation C.4 and $\alpha_p = 0$ as:

$$M = \exp\left[\int_0^L ds\, \alpha_n(s) \right] \tag{C.6}$$

or assuming the position independent multiplication coefficient α_n:

$$M = \exp(L\, \alpha_n) \tag{C.7}$$

Note that the multiplication gain M is finite for finite α_n. Thus, in the proportional mode, where the contribution from α_p is negligible, we can well control the multiplication gain M through α_n, by controlling the applied bias voltage. In this mode, no avalanche breakdown occurs. In practice, however, the α_p becomes significant when the applied bias voltage gets above a specific value, and the avalanche breakdown may occur.

C.2.2 Operation in Geiger mode

At even higher field than in the proportional mode, holes also start to generate secondary electron-hole pairs (or α_p becomes significant) as mentioned above. Secondary electrons generated by holes will again pass through the avalanche region, thereby possibly generating other electron-hole pairs. The avalanche

[i] In this calculation, we use the following equality:

$$\int_0^x ds\, \alpha(s) \exp\left[- \int_0^s dt\, \alpha(t) \right] = 1 - \exp\left[- \int_0^x dt\, \alpha(t) \right]$$

substituting "$\alpha_n(x) - \alpha_p(x)$" for $\alpha(x)$.

process continues again and again. Consequently, the avalanche multiplication gain M becomes very large. This situation, where both α_n and α_p are significant, also occurs for the materials whose α_n and α_p have similar values, as in GaAs, GaP, etc. To be simple, here we assume the equal multiplication coefficient: $\alpha_n = \alpha_p \equiv \alpha$. Then from the equation C.4, we get:

$$M = \frac{1}{1 - \int_0^L ds \, \alpha(s)} \tag{C.8}$$

or assuming the position independent multiplication coefficient:

$$M = \frac{1}{1 - L\,\alpha} \tag{C.9}$$

The two equations above show that M can diverge to infinity, or an avalanche breakdown can occur, even with finite multiplication coefficient α. In a practical device, however, the equations above do not determine the multiplication gain. Instead, the continuing avalanche process is stopped by one of the following three processes: 1) a statistical fluctuation in the process, 2) a sufficiently large drop of the externally supplied voltage, or 3) a space-charge effect that lowers the field strength in a moment. Consequently, the output signal is not proportional to the input signal, because all of the three processes to stop the avalanche are nonlinear. We call this operational mode "Geiger mode," due to the nonlinear multiplication behavior. This operational mode is suit for single photon detection.

Appendix D

The history of water Cerenkov detectors and their photo sensors

D.1 IMB

IMB (Irvine Michigan Brookhaven experiment) was a 8,000 ton ring-imaging water Cerenkov detector. The detector was built 600 meters underground in the Morton salt mine in Ohio, and started data taking in 1982. The water tank of the detector had a size of 18m×17m×22.5m. Around the tank, it had 2048 photomultiplier tubes (PMTs) of 5.5 inch diameter, which was replaced with 8 inch PMTs with wave-shifting plates by two upgrades. It was the first generation water Cerenkov detector experiment, with the KamiokaNDE experiment, and had achieved very fruitful results on neutrino physics:[39] the search for the atmospheric neutrino oscillation[40] and the observation of supernova neutrinos,[41] for example. Also, the proton decay search is one of its most important results.[42]

At first, the experiment (IMB-1) used 5 inch (EMI 8970B) and 8 inch (EMI 8934) PMTs. The experiment consisted of about 2,400 PMTs, where the large part was the 5 inch PMTs. The quantum efficiency (QE) was $\sim$ 25% for the 5 inch PMT with 400nm light. The 5 inch PMTs had timing resolution of $\sim$ 9ns in standard deviation, for single photoelectron input.[43] It was named IMB-3 after two upgrades. IMB-3 was equipped with 8 inch PMTs (Hmamatsu R1408) with wave-shifting plates.[44, 45] The time resolution is 15ns FWHM (equivalent to σ = 6.4ns standard deviation, assuming normal distribution), and the quantum efficiency is $\sim$ 30% with the wave shifting plate.

D.2 KamiokaNDE

KamookaNDE (Kamioka Nucleon Decay Experiment) was located 1,000 meters underground in Kamioka mine in Japan, and started data taking in July 1983. Its detector was an imaging-type water Cerenkov detector with a tank, which contained 3,000 tons of pure water and had about 1,000 photomultiplier tubes (PMTs) attached to the inner surface. The size of the tank was 16.0m in height and 15.6m in diameter. The detection of supernova neutrinos[46] and the long time measurement of solar neutrinos[47] are two of the greatest achievements of the experiment, and are the milestones of the neutrino physics today. The experiment also searched for the proton decay and ruled out the simplest Grand Unified Theories.[48]

The photo sensor of the detector is Hamamatsu 20 inch PMT developed for the experiment.[49] The quantum efficiency is 22% for 400nm light, and the timing resolution is 7ns FWHM (equivalent to $\sigma \sim$ 3ns standard deviation) for single photoelectron input.

D.3 Super Kamiokande

Super-Kamiokande, the successor experiment of the KamiokaNDE experiment, was constructed about 200m away from the KamiokaNDE detector, has started data taking in April 1996, and is still running.

Its detector is the large-sized version of the KamiokaNDE: its tank contains 50,000 tons of pure water, with 11,100 20 inch PMTs equipped around it as photo sensors.[i] The achievements of the experiment cannot be fully listed here. Some of the most important ones are the results on the neutrino oscillation: the observation of the atmospheric neutrino oscillation[1] and the solar neutrino oscillation,[4] and the result of the first long base line neutrino oscillation experiment[6] where Super Kamiokande plays an essential role as the far detector.

The main photo sensor of the detector is Hamamatsu R3600, the upgrade version of KamiokaNDE's 20 inch PMT.[50] Its quantum efficiency is 22% at a wavelength of 390nm. It has good performances of timing resolution ($\sigma \sim 2.2$ns for single-photoelectron), and pulse height distribution (being capable of single-photon peak identification).

D.4 1 Kt detector in K2K

K2K (KEK to Kamioka long-baseline neutrino experiment) is the first accelerator-based long baseline neutrino oscillation experiment. The experiment has been under operation since March 1999. In the experiment, the neutrino (ν_μ) beam is generated at KEK in Tsukuba, with KEK 12GeV PS (proton synchrotron). The detector in the far site, the Super Kamiokande, detects the generated neutrinos and counts the number. The experiment also has detectors in the near site, KEK, to monitor the initial flux of the neutrino beam. 1 Kt (kiloton) detector is one of the detectors in the near site. The experiment has confirmed the ν_μ disappearance, which has been observed with atmospheric neutrino oscillation, for the first time with artificial neutrinos.[6]

The 1 Kt detector is a water Cerenkov detector with a water tank of an 8.6m diameter, 8.6m high cylinder, and 680 PMTs around the tank.[12] The PMTs used are the same ones as Super Kamiokande.

D.5 SNO

SNO (Sudbury Neutrino Observatory) is the experiment with an objective of solving the solar neutrino problem. The experiment has been under operation since November 1999. The detector was built 2,000 meter underground in Creigton mine near Sudbury. It is a heavy-water Cerenkov detector with water tanks containing 1,000 tons of heavy water and 7,000 tons of light water, and it has 10,000 PMTs around the tank. Its unique feature is that by using the heavy water, it can discriminate three different types of neutrino interactions: The Charged Current(CC) interactions, the Neutral Current(NC) interactions, and the Electron Scattering(ES) interactions. By the power of the feature, the SNO experiment has given us a good understanding on the solar neutrino problem.[5]

The photo sensor used in the detector is the Hamamatsu R1408 PMT. The effective photocathode coverage is enhanced to 54% with a 27cm diameter light concentrator mounted, although the original photocathode coverage is 31%. The timing resolution for single-photoelectron is $\sigma = 1.7$ns on the average.[51]

[i] After the accident in November 2001, Super Kamiokande has reduced the number of PMTs to 5,200.

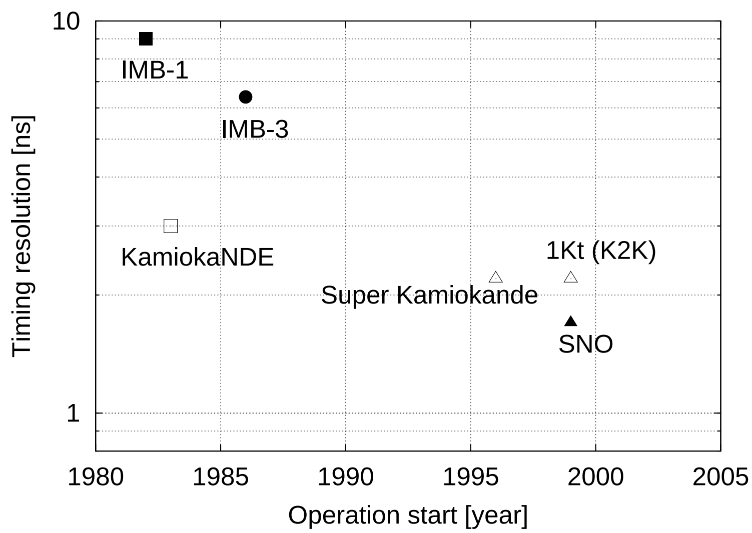


Figure D.1: The change of photo sensors' timing resolution.

Bibliography

[1] **Super-Kamiokande** Collaboration, Y. Fukuda *et al.*, "Evidence for oscillation of atmospheric neutrinos," *Phys. Rev. Lett.* **81** (1998) 1562–1567, `hep-ex/9807003`.

[2] "Letter of Intent, Neutrino oscillation experiment at JHF," January, 2003. `http://www-nu.kek.jp/jhfnu/index_e.html`.

[3] T. Nakaya, "The JHF-Kamioka neutrino experiment," *Nucl. Instrum. Meth.* **A503** (2003) 167–172.

[4] **Super-Kamiokande** Collaboration, S. Fukuda *et al.*, "Determination of solar neutrino oscillation parameters using 1496 days of Super-Kamiokande-I data," *Phys. Lett.* **B539** (2002) 179–187, `hep-ex/0205075`.

[5] **SNO** Collaboration, Q. R. Ahmad *et al.*, "Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory," *Phys. Rev. Lett.* **89** (2002) 011301, `nucl-ex/0204008`.

[6] **K2K** Collaboration, M. H. Ahn *et al.*, "Indications of neutrino oscillation in a 250-km long-baseline experiment," *Phys. Rev. Lett.* **90** (2003) 041801, `hep-ex/0212007`.

[7] **KamLAND** Collaboration, K. Eguchi *et al.*, "First results from KamLAND: Evidence for reactor anti-neutrino disappearance," *Phys. Rev. Lett.* **90** (2003) 021802, `hep-ex/0212021`.

[8] T. Kobayashi, "Conventional neutrino beams at K2K and JHF," *Nucl. Instrum. Meth.* **A503** (2003) 418–420.

[9] T. Kobayashi, "Status of J-PARCν facility," November, 2003. Talk from the 4th International Wrokshop on the Neutrino Beams and Instrumentation (NBI2003). `http://www-ps.kek.jp/nbi2003/presentation/index.html`.

[10] H. Noumi, "Proton beam for K2K," March, 2002. Talk from the 3rd International Wrokshop on the Neutrino Beams and Instrumentation (NBI2002). `http://proj-cngs.web.cern.ch/proj-cngs/2002_workshop/NBI2002_talks/pb_noumi.pdf`.

[11] Y. Itow, "Super-Kamiokande as a Far Detector," May, 2001. Talk from the International Workshop on a Next Generation Long-Baseline Neutrino Oscillation Experiment, JHF-SK Nu. `http://neutrino.kek.jp/jhfnu/workshop2/ohp.html`.

[12] **K2K** Collaboration, S. H. Ahn *et al.*, "Detection of accelerator produced neutrinos at a distance of 250-km," *Phys. Lett.* **B511** (2001) 178–184, `hep-ex/0103001`.

[13] **Super-Kamiokande** Collaboration, M. Shiozawa, "Reconstruction algorithms in the super-kamiokande large water cherenkov detector," *Nucl. Instrum. Meth.* **A433** (1999) 240–246.

[14] Y. Itow *et al.*, "The JHF-Kamioka neutrino project," `hep-ex/0106019`.

[15] K. Nakamura, "Hyper-Kamiokande — A next generation water Cerenkov detector," *Int. J. Mod. Phys.* **A18** (2003) 4053–4063.

[16] M. Shiozawa, "Photo-sensor R&D for 2km water Cherenkov detector and Hyper-Kamiokande," September, 2002. Talk from the International Workshop on Nuclear and Particle Physics at 50-GeV PS (NP02). `http://jnusrv01.kek.jp/jnu/NP02/Sep28/shiozawa.pdf`.

[17] M. Suyama, "Development of a multi-pixel photon sensor with single-photon sensitivity,". KEK-REPORT-2002-16.

[18] F. A. White and J. C. Sheffield, "Reversed-biased P-N junctions in electron tubes," *Proc. IRE* **50** (1962), no. 6, 1523–1524.

[19] R. DeSalvo, "Hybrid photodiode tube,". CLNS-87-92.

[20] T. Gys, "The pixel hybrid photon detectors for the LHCb-RICH project," *Nucl. Instrum. Meth.* **A465** (2000) 240–246.

[21] A. Braem *et al.*, "Development of large area hybrid photodiodes for the LHCb ring imaging Cherenkov detectors," *IEEE Trans. Nucl. Sci.* **46** (1999) 365–370.

[22] P. Cushman *et al.*, "Multi-pixel hybrid photodiode tubes for the CMS hadron calorimeter," *Nucl. Instrum. Meth.* **A387** (1997) 107–112.

[23] M. Suyama *et al.*, "A hybrid photodetector (HPD) with a III-V photocathode," *IEEE Trans. Nucl. Sci.* **45** (1998) 572–575.

[24] N. Kanaya *et al.*, "Test results on hybrid photodiodes," *Nucl. Instrum. Meth.* **A421** (1999) 512–521.

[25] Hamamatsu Photonics K. K., "HPD (Hybrid Photo-Detector) Modules H8236-07/40.". `http://www.hamamatsu.com/`.

[26] S. M. Bradbury, R. Mirzoian, J. Gebauer, E. Feigl, and E. Lorenz, "Test of the new hybrid INTEVAC intensified photocell for the use in air Cerenkov telescopes," *Nucl. Instrum. Meth.* **A387** (1997) 45–49.

[27] R. Mirzoyan, D. Ferenc, and E. Lorenz, "An evaluation of the new compact hybrid photodiodes R7110U-07/40 from Hamamatsu in high-speed light detection mode," *Nucl. Instrum. Meth.* **A442** (2000) 140–145.

[28] M. Akatsu *et al.*, "Time-of-propagation Cherenkov counter for particle identification," *Nucl. Instrum. Meth.* **A440** (2000) 124–135, `physics/9904009`.

[29] S. Matsui *et al.*, "HAPD time-resolution study under single-photon irradiation," *Nucl. Instrum. Meth.* **A463** (2001) 220–226.

[30] N. Kanaya, "HPD for linear colliders," *Nucl. Instrum. Meth.* **A442** (2000) 295–299.

[31] M. Suyama *et al.*, "Development of a multi-pixel photon sensor with single-photon sensitivity," `physics/0312017`.

[32] K. Deiters *et al.*, "Properties of the most recent avalanche photodiodes for the CMS electromagnetic calorimeter," *Nucl. Instrum. Meth.* **A442** (2000) 193–197.

[33] K. Deiters *et al.*, "Properties of the avalanche photodiodes for the CMS electromagnetic calorimeter," *Nucl. Instrum. Meth.* **A453** (2000) 223–226.

[34] M. Tanaka, H. Ikeda, K. Tamai, M. Takemoto, and H. Hayashii, "Development and evaluation of charge sensitive preamplifier for CsI calorimeter in the KEK B factory," *IEEE Trans. Nucl. Sci.* **41** (1994) 1208–1211.

[35] Y. Arai and T. Ohsugi, "TMC: A CMOS time to digital converter VLSI," *IEEE Trans. Nucl. Sci.* **36** (1989) 528–531.

[36] S. A. Kleinfelder, "A 4096 cell switched capacitor analog waveform storage integrated circuit," *IEEE Trans. Nucl. Sci.* **37** (1990) 1230–1236.

[37] T. Higuchi *et al.*, "Development of a PCI based data acquisition platform for high intensity accelerator experiments," *ECONF* **C0303241** (2003) TUGT004, `hep-ex/0305088`.

[38] Y. Igarashi *et al.*, "The data acquisition system based on PMC bus," *ECONF* **C0303241** (2003) TUGP009, `physics/0305137`.

[39] R. Becker-Szendy *et al.*, "Neutrino measurements with the IMB detector," *Nucl. Phys. Proc. Suppl.* **38** (1995) 331–336.

[40] R. Becker-Szendy *et al.*, "A search for muon-neutrino oscillations with the IMB detector," *Phys. Rev. Lett.* **69** (1992) 1010–1013.

[41] R. M. Bionta *et al.*, "Observation of a neutrino burst in coincidence with supernova SN1987A in the Large Magellanic Cloud," *Phys. Rev. Lett.* **58** (1987) 1494.

[42] W. Gajewski *et al.*, "Search for proton decay into e+ + pi0 in IMB-3 detector," *Phys. Rev.* **D42** (1990) 2974–2976.

[43] C. Wuest *et al.*, "The IMB photomultiplier test facility," *Nucl. Instrum. Meth.* **A239** (1985) 467–486.

[44] R. Claus *et al.*, "A wave shifter light collector for a water cerenkov detector," *Nucl. Instrum. Meth.* **A261** (1987) 540–542.

[45] R. Becker-Szendy *et al.*, "IMB-3: A Large water Cerenkov detector for nucleon decay and neutrino interactions," *Nucl. Instrum. Meth.* **A324** (1993) 363–382.

[46] K. S. Hirata *et al.*, "Observation in the Kamiokande-II detector of the neutrino burlst from supernova SN1987A," *Phys. Rev.* **D38** (1988) 448–458.

[47] **Kamiokande** Collaboration, Y. Fukuda *et al.*, "Solar neutrino data covering solar cycle 22," *Phys. Rev. Lett.* **77** (1996) 1683–1686.

[48] **KAMIOKANDE-II** Collaboration, K. S. Hirata *et al.*, "Experimental limits on nucleon lifetime for lepton + meson decay modes," *Phys. Lett.* **B220** (1989) 308.

[49] H. Kume *et al.*, "20 inch diameter photomultiplier," *Nucl. Instrum. Meth.* **205** (1983) 443–449.

[50] A. Suzuki *et al.*, "Improvement of 20-inch diameter photomultiplier tubes," *Nucl. Instrum. Meth.* **A329** (1993) 299–313.

[51] **SNO** Collaboration, J. Boger *et al.*, "The sudbury neutrino observatory," *Nucl. Instrum. Meth.* **A449** (2000) 172–207, `nucl-ex/9910016`.