Lecture 7

• Optical receivers
  – p–i–n diodes
  – Avalanche diodes
  – Receiver design
  – Receiver noise
    • Shot noise
    • Thermal noise
  – Signal-to-noise ratio
Optical receivers

• The purpose of a traditional receiver for OOK is:
  – Convert the optical signal into an electrical signal
  – Recover the data by:
    • Doing clock recovery
    • Performing decisions on the obtained signal

• In state-of-the-art coherent receivers, additional functionality is performed in digital signal processing (DSP)
  – Electronic dispersion compensation (EDC)
  – Adaptive equalization
  – Phase synchronization

• This lecture is about OOK systems
  – Necessary to know about...
  – ...and still common
Photodetectors

• The most critical component is the photodetector
  – Converts the optical signal to an electrical current

• We want these components to have:
  – High sensitivity
  – Fast response time
  – Low noise
  – High reliability
  – Size compatible with fibers

• This means that semiconductor materials are exclusively used
  – Photons are absorbed and generate electron–hole (e–h) pairs
  – This produces a photo-current.

• Basic requirement: The detector material bandgap energy \( E_g \) < the photon energy \( hv \)
Photodiodes (4.1.1)

- The photocurrent is proportional to the optical power \( I_p = R_d P_{in} \)
- The constant \( R_d \) is the **responsivity**

\[
R_d = \eta \frac{q}{h\nu} = \eta \frac{\lambda}{1.24} \text{ [A/W]} \text{ with } \lambda \text{ in } \mu\text{m}
\]

- \( \eta \) = the **quantum efficiency** = the number of e–h pairs per incident photon
- Ideally \( \eta = 1 \)
- \( R_d \) increases with \( \lambda \) until \( h\nu = E_g \)
  - \( R_d \to 0 \) when the photon energy becomes too low
- Si or GaAs can be used for short wavelengths (\( \lambda < 900 \text{ nm} \))
- InGaAs is most common at 1.3 and 1.55 \( \mu\text{m} \)
- Most communication systems use reverse-biased p–n junctions (photodiodes) of two main types:
  - p–i–n photodiodes
  - Avalanche photodiodes (APD)
p–i–n diodes (4.2.2)

absorption of photons ⇒ e–h pair generation ⇒ carrier drift due to built-in and applied field ⇒ induced current in the external circuit

- p–i–n diode: p–n junction with an intrinsic (un-doped) layer
- Response time is limited by the transit time through the i-region
  \[ \tau_{tr} = \frac{W}{v_s} \]
- Responsivity increases with \( W \) ⇒ a trade-off between responsivity and speed
- High speed (~50 GHz) diodes with \( \eta \) close to unity are available
p–i–n diodes, performance

p–n diodes are limited by diffusion (absorption outside the depletion region).

In a p–i–n diode, the depletion region is wide (intrinsic, undoped).

- p–i–n bandwidth limitations:
  - Parasitic capacitance
    - Reduces the speed of voltage changes
  - Transit time
    - Takes time to collect the carriers
- The *dark current* should be low
  - Current without input signal
    - Due to stray light and thermal generation of carriers
Examples of p–i–n diodes

Schematic picture of a p–i–n diode
Green is anti-reflection coating

• Important parameters are:
  – Bandwidth
  – Sensitivity
  – Responsivity
  – Polarization dependence
  • No dependence is preferred

p–i–n diodes without and with “pigtail”
Avalanche photodiodes (APDs) (4.2.3)

- An APD is a p–i–n diode with an extra layer next to the i-region
  - Gives gain through \textit{impact ionization} and amplifies the signal
  - The responsivity can be $> \frac{q}{h\nu}$

- The responsivity of an APD is
  - $M$ is the \textit{multiplication factor}

- The increased responsivity comes at the expense of
  - Enhanced noise
  - Reduced bandwidth

$$R_{\text{APD}} = M \frac{\eta q}{h \nu} = MR_d$$
APD multiplication factor

- The multiplication factor $M$ depends on the geometry of the APD, the electric field etc
- The frequency dependence is

$$M(\omega) = \frac{M(0)}{\sqrt{1+[\omega \tau_e M(0)]^2}}$$

- $\tau_e$ is the effective transit time for the multiplication process
- A trade-off between multiplication and bandwidth
- Si-APDs have very good performance
  - $M > 100$, high bandwidth, relatively low noise
  - Very useful for systems operating near 0.8 $\mu$m
- InGaAs-APDs can be used at 1.3 and 1.55 $\mu$m
  - Suffer from smaller multiplication and bandwidth, and higher noise
Receiver design (4.3)

- The digital receiver consists of three parts:
  - Front end (photo-detector, trans-impedance amplifier)
  - Linear channel (amplifier, low-pass filter)
  - Data recovery (clock recovery, decision circuit)
**Receiver front-ends (4.3.1)**

**Transimpedance front-end**

- Simple
- Electrically stable
- Low sensitivity for small $R_L$
- Small bandwidth for high $R_L$

$$\Delta f = \frac{1}{2\pi R_L C_p}$$

- High bandwidth
- High sensitivity
- Potentially unstable

$$\Delta f = \frac{G}{2\pi R_f C_p}$$

Effective input resistance = $R_f/G$
Linear channel (4.3.2)

- The linear channel consists of:
  - A high-gain amplifier with automatic gain control
    - Constant average output voltage irrespective of the input (within limits)
  - A low-pass filter with bandwidth chosen to:
    - Reject noise outside signal bandwidth
    - Avoid introducing inter-symbol-interference (ISI)
- The best situation is when the filter (and not other components) limits the overall bandwidth of the receiver
- The output voltage spectrum is given by $H_{\text{out}}(\omega) = H_T(\omega)H_p(\omega)$
  - $H_p(\omega)$ is the photocurrent spectrum
  - $H_T(\omega)$ is the total transfer function of the front end and the linear channel
- Normally, $H_T(\omega)$ is dominated by the filter transfer function
  - $H_T(\omega) \approx H_f(\omega)$
Data recovery (4.3.3)

- The data-recovery section consists of
  - A *clock-recovery circuit*
    - Extracting a sinusoidal component at $f = B$ to enable proper synchronization of the decision circuit
  - Easily done for an OOK RZ signal with a narrow-band filter
    - The signal contains a delta function at $f = B$
  - More difficult for NRZ
    - No sinusoidal spectral components are present
    - Can use a full-wave rectifier to convert the NRZ signal to RZ containing a delta function at $f = B$
  - A *decision circuit* comparing the input voltage with a threshold at the time obtained from the clock recovery
    - Deciding whether a "1" or a "0" was received
Eye diagrams

• The *eye diagram* is a superposition of all bits on top of each other
  – Looks like an eye
  – Gives a visual way to monitor the receiver performance

• Left: An ideal NRZ eye diagram

• Right: An eye diagram degraded by noise and timing jitter

• A measured RZ eye diagram at 640 Gbit/s
Eye diagram interpretation

Slope indicates sensitivity to timing error, smaller is better.

Signal excursion or wasted power

Amount of distortion at sampling instant, relates to signal SNR

Amount of noise that can be tolerated by the signal, the larger the better.

Amount of distortion, or variation in where zero crossing occurs.

Best time to sample

Opening of the eye, time over which we can successfully sample the waveform
Receiver noise (4.4)

- The detected photo current in the receiver will contain noise.
- There are two fundamental sources of noise:
  - **Shot noise** due to field and charge quantization.
  - **Thermal noise** due to thermal motion of charges.
- The total current, signal + noise, can be written as:
  \[ I(t) = R_d P_{in}(t) + i_s(t) + i_T(t) \]
- In addition, there can also be optical noise in \( P_{in} \):
  - Comes from lasers and optical amplifiers.
  - Will be treated later in the course.
- Remember:
  \[ I_p(t) = R_d P_{in}(t) \]
Shot noise

- Shot noise arises from the particle nature of the photocurrent
  - Current consists of electrons that can only be described statistically
  - Current is not constant but fluctuates
  - Compare with cars on a highway or hails on a roof
- The variance of the shot noise photocurrent is
  $$\sigma_s^2 = \langle i_s^2(t) \rangle = 2 \int_0^{\Delta f} S_s(f) df = 2qI_p \Delta f$$
  - \(\Delta f\) is the effective noise bandwidth of the receiver
  - \(S_s(f)\) is the shot noise two-sided power spectral density (PSD)
- If the detector dark current \(I_d\) cannot be neglected we have
  $$\sigma_s^2 = 2q(I_p + I_d) \Delta f$$
  - Originating from stray light or thermally generated e–h pairs
Thermal noise

- Thermal noise originates from the thermal motion of the electrons
- The two-sided PSD is

\[ S_T(f) = \frac{2hf}{R_L \left[ \exp(hf / k_B T) - 1 \right]} \approx \frac{2k_B T}{R_L} \]

- \( k_B \) is Boltzmann’s constant
- \( T \) is the temperature
- \( R_L \) is the load resistance

- The noise variance is

\[ \sigma_T^2 = \left\langle i_T^2(t) \right\rangle = 2 \int_0^{\Delta f} S_T(f) df \approx (4k_B T / R_L) \Delta f \]

- In addition, thermal noise is also generated in electrical amplifiers
  - Introduce the **amplifier noise figure** \( F_n \) to obtain

\[ \sigma_T^2 = (4k_B T / R_L) F_n \Delta f \]
Signal-to-noise ratio (SNR)

• The different noise sources are uncorrelated
  – We obtain the total noise power according to
    \[ \sigma^2 = \langle (\Delta I)^2 \rangle = \sigma_s^2 + \sigma_T^2 = 2q(I_p + I_d)\Delta f + (4k_B T / R_L) F_n \Delta f \]

• The *signal-to-noise ratio* (SNR) of an electrical signal is defined as
  \[ \text{SNR} = \frac{\text{average signal power}}{\text{noise power}} = \frac{I_p^2}{\sigma^2} \]

• This definition is for an analog signal
  – This is not the usual meaning of ”SNR” in digital communication theory
    • Instead \( E_b / N_0 \) or \( E_s / N_0 \) is used there
      – \( E_b \) is the energy per bit
      – \( E_s \) is the energy per symbol
      – \( N_0 \) is the noise PSD
Noise in p–i–n receivers (4.4.2)

- For a p–i–n receiver we have

\[
\text{SNR} = \frac{R_d P_{in}^2}{2q(R_d P_{in} + I_d)\Delta f + 4(k_B T / R_L) F_n \Delta f}
\]

- When thermal noise dominates, we have

\[
\text{SNR} = \frac{R_L R_d^2 P_{in}^2}{4k_B T F_n \Delta f} \propto P_{in}^2
\]

- When shot noise dominates, we have

\[
\text{SNR} = \frac{R_d P_{in}}{2q\Delta f} = \frac{\eta P_{in}}{2h \nu_0 \Delta f} \propto P_{in}
\]

- We note:
  - Different scaling with input power in the two limits
  - Thermal noise dominates at low input power
  - Shot noise dominates at high input power
Noise in APD receivers (4.4.3)

- Since $R_{\text{APD}} = MR_d$, the power of the current increases by $M^2$
  - But the noise increases too, so the SNR increase is smaller
- The APD shot noise variance is
  $$\sigma_s^2 = 2qM^2 F_A (R_d P_{\text{in}} + I_d) \Delta f$$
- The *excess noise factor* is
  $$F_A(M) = k_A M + (1 - k_A)(2 - 1/M)$$
  - $1 < F_A < M$ since $0 < k_A < 1$, ($k_A = \alpha_h/\alpha_e$, see (4.2.3))
- The SNR becomes
  $$\text{SNR} = \frac{(MR_d P_{\text{in}})^2}{2qM^2 F_A (R_d P_{\text{in}} + I_d) \Delta f + 4(k_B T / R_L) F_n \Delta f}$$
  - The shot-noise is increased by $M^2 F_A$
Noise in APD receivers

- In the thermal noise limit we have
  \[ \text{SNR} = \frac{R_L R_d M^2 P_{in}^2}{4k_B T F_n \Delta f} \propto M^2 P_{in}^2 \]
  - A factor of $M^2$ higher than for the p–i–n

- In the shot noise limit we have
  \[ \text{SNR} = \frac{R_d P_{in}}{2q F_A \Delta f} = \frac{\eta P_{in}}{2h \nu_0 F_A \Delta f} \propto \frac{P_{in}}{F_A} \]
  - A factor of $F_A$ lower than for the p–i–n diode

The SNR is increased by an APD in the thermal-noise limit
The SNR is decreased by an APD in the shot-noise limit
The APD vs the p–i–n

- The SNR ($\Delta f = 30$ GHz) for a p–i–n receiver and an APD receiver
  - APD is best at low power
  - p–i–n is best at high power
  - $M = 10$ is worse than $M = 5$

- There is an optimum value for $M$

$$M_{\text{opt}} \approx \left[ \frac{4k_B TF_n}{k_A q R_L (R_d P_{\text{in}} + I_d)} \right]^{1/3}$$

- Optimum value depends on $k_A = \alpha_h/\alpha_e$
  - Highest $M_{\text{opt}} \sim 100$ for silicon APD
  - Highest $M_{\text{opt}} \sim 10$ for InGaAs APD