# CMOS Image Sensors

**Online at: https://doi.org/10.1088/978-0-7503-3235-4**

# CMOS Image Sensors

**Konstantin D Stefanov**
*Centre for Electronic Imaging, The Open University, Milton Keynes, UK*

**IOP** Publishing, Bristol, UK

© IOP Publishing Ltd 2022

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher, or as expressly permitted by law or under terms agreed with the appropriate rights organization. Multiple copying is permitted in accordance with the terms of licences issued by the Copyright Licensing Agency, the Copyright Clearance Centre and other reproduction rights organizations.

Permission to make use of IOP Publishing content other than as set out above may be sought at permissions@ioppublishing.org.

Konstantin D Stefanov has asserted their right to be identified as the author of this work in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

ISBN 978-0-7503-3235-4 (ebook)
ISBN 978-0-7503-3233-0 (print)
ISBN 978-0-7503-3236-1 (myPrint)
ISBN 978-0-7503-3234-7 (mobi)

DOI 10.1088/978-0-7503-3235-4

Version: 20221101

IOP ebooks

British Library Cataloguing-in-Publication Data: A catalogue record for this book is available from the British Library.

Published by IOP Publishing, wholly owned by The Institute of Physics, London

IOP Publishing, No.2 The Distillery, Glassfields, Avon Street, Bristol, BS2 0GR, UK

US Office: IOP Publishing, Inc., 190 North Independence Mall West, Suite 601, Philadelphia, PA 19106, USA

*To David Burt, from whom I have learnt many of the things in this book,*
*and to my family, for their support and encouragement.*

# Contents

**Preface** xi

**Acknowledgement** xiii

**Author biography** xiv

**List of frequently used abbreviations** xv

**Table of common symbols and units** xvi

**1 The fundamentals** 1-1

1.1 Introduction—what is an image sensor and what does it do? 1-1
1.2 Charge generation 1-2
1.2.1 Photoeffect 1-2
1.2.2 Ionisation 1-7
1.3 Charge collection 1-9
1.3.1 Carrier lifetime 1-10
1.3.2 Recombination 1-12
1.3.3 Drift 1-14
1.3.4 Diffusion 1-17
1.4 Charge transfer 1-19
1.5 Charge conversion 1-20
1.6 *pn* junction 1-22
1.6.1 *pn* junction in equilibrium 1-22
1.6.2 *pn* junction under reverse bias 1-26
1.6.3 Charge collection 1-29
1.6.4 Junction capacitance 1-32
1.7 MOS capacitor 1-33
1.7.1 Depletion 1-33
1.7.2 Gate capacitance 1-37
1.8 MOS transistor 1-38
1.8.1 Structure 1-38
1.8.2 MOSFET characteristics 1-40
1.8.3 Output resistance and body effect 1-44
1.8.4 Transistor threshold 1-47
1.8.5 Analogue switch 1-51
1.8.6 MOSFET capacitor 1-53
1.9 Source follower 1-54
1.9.1 Gain 1-54

1.9.2 Input capacitance 1-58
Chapter summary 1-60
References 1-61

**2 CMOS pixel architectures 2-1**

2.1 History and technology 2-1
2.2 Photodiode APS 2-2
2.2.1 Structure 2-2
2.2.2 Operation 2-5
2.2.3 Performance 2-9
2.3 Pinned photodiode (4T) 2-11
2.3.1 Structure 2-11
2.3.2 Operation 2-15
2.3.3 Charge storage and full well capacity 2-17
2.3.4 Charge transfer 2-22
2.3.5 Image lag 2-27
2.3.6 Transistor sharing 2-33
2.4 Other PPD-based pixels 2-33
2.4.1 Global reset (5T) 2-33
2.4.2 In-pixel signal storage 2-35
2.4.3 High dynamic range 2-37
2.5 Hybrid and 3D image sensors 2-41
Chapter summary 2-43
References 2-44

**3 Advanced image sensor topics 3-1**

3.1 Photocurrent 3-1
3.2 Dark current 3-6
3.2.1 Sources of dark current 3-6
3.2.2 Depletion dark current 3-9
3.2.3 Diffusion dark current 3-12
3.2.4 Surface dark current 3-15
3.2.5 Dark current suppression by pinning 3-16
3.2.6 Temperature for dark current doubling 3-17
3.3 Reflective barrier 3-18
3.4 Back-side illumination 3-21
3.4.1 Front and back-side illumination 3-21
3.4.2 Back-side interface 3-24

3.4.3 BSI technologies 3-28
3.5 Depletion depth and potential gradients 3-29
3.5.1 Depletion depth as a 3D effect 3-29
3.5.2 Potential gradients in PPDs 3-31
3.6 Punch-through 3-31
3.7 Field-induced junctions 3-35
Chapter summary 3-36
References 3-37

**4 Noise and readout techniques 4-1**
4.1 Noise in image sensors 4-1
4.1.1 Thermal and reset noise 4-1
4.1.2 Shot noise 4-6
4.1.3 $1/f$ and random telegraph noise 4-8
4.1.4 MOSFET noise 4-10
4.1.5 Source follower noise 4-12
4.2 Correlated double sampling 4-14
4.2.1 Reset noise suppression 4-14
4.2.2 Double sampling 4-15
4.2.3 Dual slope integrator 4-19
4.2.4 Optimal signal processing 4-22
4.2.5 Digital CDS and multiple sampling 4-24
4.2.6 Column-level noise 4-28
4.2.7 MOSFET optimisation 4-32
Chapter summary 4-33
References 4-34

**5 Characterisation 5-1**
5.1 Introduction 5-1
5.2 Readout modes 5-2
5.3 Principles of EO characterisation 5-4
5.4 Photoresponse, non-uniformity and nonlinearity 5-7
5.5 Photon transfer curve 5-17
5.5.1 Principles 5-17
5.5.2 Frame differencing 5-21
5.5.3 System gain, CVF and noise 5-24
5.5.4 Nonlinear PTC 5-27
5.5.5 PTC from dark current 5-30

5.5.6 Practical tips for the PTC 5-31
5.5.7 The PTC as a diagnostic tool 5-34
5.6 X-ray calibration 5-36
5.7 Full well capacity and dynamic range 5-39
5.8 Dark current and DSNU 5-41
5.9 Noise measurement 5-44
5.10 Image lag 5-48
5.11 Quantum efficiency 5-50
5.11.1 Principles 5-50
5.11.2 Pain–Hancock method 5-54
5.11.3 Modulation transfer function 5-57
5.12 Electrical transfer function 5-62
Chapter summary 5-65
References 5-66

**6 Electronics 6-1**

6.1 On-chip electronics 6-1
6.1.1 Architecture 6-1
6.1.2 Column buffers 6-2
6.1.3 Column amplifiers 6-3
6.1.4 CDS circuits 6-7
6.1.5 Row drivers 6-12
6.1.6 Pixel addressing 6-14
6.1.7 Analogue switches and multiplexers 6-17
6.1.8 Output amplifier 6-19
6.2 Off-chip electronics 6-21
6.2.1 General requirements 6-21
6.2.2 Signal amplifiers 6-23
6.2.3 Power supplies 6-29
6.2.4 Bias circuits 6-30
6.2.5 Noise measurements 6-32
Chapter summary 6-34
References 6-35

# Preface

Image sensors are fascinating devices that straddle the boundary between semiconductor physics and electronic engineering. Nowadays, they are used in almost everything, come in bewildering varieties, and are mostly made with complementary metal-oxide semiconductor (CMOS) technology, like the billions of other integrated circuits (IC) manufactured every year. And they look good! They are among the handful of IC types which can be easily seen through their transparent glass cover.

To understand how they work, we need to know a fair bit of semiconductor physics, especially when dealing with high performance imagers designed for science applications. However, this is not always enough; there are a lot of electronic circuits inside a CMOS image sensor (CIS), and even the simple ones can show subtle behaviour and throw up surprises. Without claiming to cover everything, this book strives to cover both the semiconductor physics and the essential electronics found inside a CMOS image sensor.

A relatively small number of concepts from semiconductor physics form the backbone of image sensors' operation—depletion, drift, diffusion, recombination, charge conversion, to name a few. Knowing them well provides the foundations to understand practically all image sensors and helps with the more complex structures that exist or are yet to be invented.

It is impossible to imagine doing any serious work into image sensors without semiconductor technology CAD (TCAD). Very often it is the only way to 'see' what is happening inside, and this book offers many examples of device simulations. A successful TCAD simulation is a good result, but does not guarantee that something will work. However, if something doesn't work in TCAD, it's probably not going to work in silicon either.

To make full use of this book, some basic knowledge of electronics is essential. Knowing what an amplifier is, being familiar with gain, bandwidth, and noise, can be very advantageous. Freely available electronic simulation tools, such as SPICE, are great for designing and verifying the performance of various circuits. Throughout my career, my hobby in electronics has helped me enormously when working with image sensors. Building my first electronic circuit at around 14 years of age, I was fascinated but I only had a faint understanding of how it worked. In a few years, electronics gradually started to make more sense, and after learning semiconductor physics at university it was clear to me that this is what I wanted to do.

I have often found that many important 'bread and butter' topics in image sensors and their operation are difficult to find in books and papers, and sometimes are not there at all. Some of those I have only been able to find out in discussions with more experienced colleagues who have been longer in the field. This book is an attempt to put some of this 'unofficial knowledge', some of which could be simply due to my ignorance, in one place.

This book is intended to be used as a tutorial and has many examples, taken mostly from practice. Solved examples are essential for proper understanding of the theory and bring 'life' to the formulas. They also help with appreciating the

parameters in real-world applications. Very often, a good enough grasp of the phenomena can be obtained from relatively simple formulas. They may not be super-accurate, but can give a decent approximate answer and can serve as a 'sanity check' for more detailed results derived from TCAD and SPICE.

Some of the contents of this book are derived from a practical course on CMOS image sensor operation and characterisation techniques that we deliver at the Open University to students and staff.

Chapter 1 covers the fundamentals of image sensors, starting with the photoeffect and charge generation, and including charge collection and transfer, drift and diffusion, recombination, and carrier lifetime. The chapter also describes the fundamental building blocks of CIS—diodes, MOS capacitors and transistors, and the basic MOSFET circuits—source followers and analogue switches.

Chapter 2 deals with 3T, 4T and other CMOS pixel architectures. Most of the material in this chapter is dedicated to the pinned photodiode (PPD) because of its importance for image sensor technology. This includes the operating principles of the PPD, doping profiles, charge transfer, full well capacity and image lag. Other PPD-based designs, such as the 5T, pixels with charge domain signal storage, and several high dynamic range pixels are also covered, as well as hybrid and 3D-intgerated sensors.

In chapter 3, some of the more specialised subjects in CIS performance are discussed, such as the collection of photogenerated signal in *pn* junctions, the sources of dark current, reflective barriers, the backside interface and its effect on the quantum efficiency in backside-illuminated sensors, potential gradients and punch-through.

Chapter 4 is dedicated to the different sources of noise in electronics components and MOSFETs, the readout techniques used for CIS and their noise performance. The two main correlated double sampling (CDS) methods, based on the double sampling and the dual slope integrator are discussed in detail and their noise performance is compared. In addition, the chapter deals with digital CDS, noise in the column readout in CIS, and MOSET noise optimisation.

Chapter 5 begins with the principles of electro-optical characterisation and readout modes in CIS. It describes the measurement methods for obtaining the most important sensor parameters: photoresponse, linearity, system gain, readout noise, dynamic range, full well capacity, dark current, image lag, quantum efficiency and modulation transfer function. Special attention is paid to the photon transfer curve (PTC) because of the wealth of information it provides, and many experimental tips are given.

Chapter 6 covers the on-chip and off-chip electronics used to control and readout the pixels and the sensor. On-chip amplifiers, correlated double sampling circuits, buffers, switches, drivers and logic are some of the circuits described here. Off-chip electronics providing power, bias and amplification is shown with practical examples and performance calculations.

I hope that this book will be useful to all who are using, characterising, or designing CMOS image sensors, beginners and experts alike.

# Acknowledgement

The stimulating research environment at the Centre for Electronic Imaging (CEI) at the Open University is one of the main reasons for the existence of this book. Many of the topics and ideas presented here came up from discussions with my colleagues, students, industrial collaborators, and external partners, and by the challenges and the difficult questions they often had.

I am grateful for the inspiration and the knowledge I have gained thanks to David Burt, Andrew Holland, Chris Damerell, Ray Bell, Jérôme Pratlong, Paul Jerram, Doug Jordan, Neil Murray, Pete Turner, Dave Barry, David Hall, Matthew Soman, Julian Heymes, Martin Prest, James Ivory, Chiaki Crews, Nathan Bush, Steve Bowring and Giulio Villani.

Finally, I would like to thank IOP Publishing, and in particular John Navas for overseeing the production and for making this book possible.

Konstantin Stefanov
August 2022

# Author biography

## Konstantin D Stefanov

Konstantin D Stefanov was born in Rousse, Bulgaria, and has received his MSc in applied physics from Sofia University 'St. Kliment Ohridski'. He received his PhD degree in physics from Saga University, Japan, in 2001. As a research scientist at the Rutherford Appleton Laboratory in Oxfordshire, UK, Dr Stefanov has worked on the development of CCD and CMOS sensors for particle physics. Since 2012 he has been working at the Centre for Electronic Imaging at the Open University, Milton Keynes, UK, where he is developing CMOS image sensors for scientific and space applications. His research interests are in the areas of physics, technology, and design of CMOS image sensors for science applications, semiconductor device simulations, device characterisation, radiation damage effects, detector electronics and data acquisition systems. He has published over 80 research papers, has co-written two book chapters and holds several patents on CMOS image sensors.

# List of frequently used abbreviations

| | |
|---|---|
| AC | Alternating current |
| ADC | Analogue-to-digital converter |
| ADU | Analogue-to-digital unit |
| APS | Active pixel sensor |
| BSI | Back-side illumination |
| CCD | Charge coupled device |
| CDS | Correlated double sampling |
| CG | Conversion gain |
| CIS | CMOS image sensor |
| CMOS | Complementary metal-oxide-semiconductor |
| CVF | Charge-to-voltage conversion factor |
| DC | Direct current |
| DN | Digital number |
| DR | Dynamic range |
| DSNU | Dark signal non-uniformity |
| DUT | Device under test |
| EO | Electro-optical |
| ENC | Equivalent noise charge |
| ETF | Electrical transfer function |
| FPN | Fixed pattern noise |
| FSI | Front-side illumination |
| FWC | Full well capacity |
| HDR | High dynamic range |
| IR | Infrared |
| LED | Light emitting diode |
| MOS | Metal-oxide-semiconductor |
| MOSFET | Metal-oxide-semiconductor field effect transistor |
| MTF | Modulation transfer function |
| MVC | Mean-variance curve |
| NIR | Near infrared |
| NMOS | N-channel MOSFET |
| PD | Photodiode |
| PMOS | P-channel MOSFET |
| PPD | Pinned photodiode |
| PRNU | Photo response non-uniformity |
| PTC | Photon transfer curve |
| RMS | Root mean square |
| RTN | Random telegraph noise |
| RTS | Random telegraph signal |
| SF | Source follower |
| SNR | Signal-to-noise ratio |
| QE | Quantum efficiency |
| UV | Ultraviolet |

# Table of common symbols and units

| Symbol | Description | Value/units |
|---|---|---|
| $\alpha$ | Photon absorption coefficient | $cm^{-1}$ |
| $B$ | Signal bandwidth | Hz (Hertz) |
| $B_n$ | Noise power bandwidth | Hz |
| $C$ | Capacitance | F (Farad) |
| $C_{ox}$ | Area oxide capacitance | $F\ cm^{-2}$ |
| $C_{GS}$ | MOSFET gate-source capacitance | F |
| $D_n$, $D_p$ | Diffusion coefficient for electrons or holes | $cm^2\ s^{-1}$ |
| $e_n$ | Voltage noise density | $V/\sqrt{Hz}$ |
| $e_{nw}$ | Voltage white noise density | $V/\sqrt{Hz}$ |
| $e_{nf}$ | Voltage $1/f$ noise density | V (Volts) |
| $E$ | Electric field | $V\ cm^{-1}$ |
| $E_a$ | Activation energy | eV |
| $E_c$ | Conduction band energy | eV |
| $E_g$ | Bandgap energy | eV |
| $E_v$ | Valence band energy | eV |
| $E_i$ | Intrinsic Fermi level | eV |
| $E_F$ | Fermi level | eV |
| $E_t$ | Trap energy | eV |
| $E_e$ | Irradiance | $W\ cm^{-2}$ |
| $E_{ph}$ | Photon energy | eV |
| $E_w$ | Ionisation energy | eV |
| $\varepsilon_0$ | Dielectric permittivity of vacuum | $8.85 \times 10^{-14}\ F\ cm^{-1}$ |
| $\varepsilon_{Si}$ | Relative dielectric permittivity of silicon | 11.9 |
| $f$ | Frequency | Hz |
| $f_{nc}$ | $1/f$ noise corner frequency | Hz |
| $f_c$ | Cut-off frequency | Hz |
| $\Phi$ | Photon flux | $cm^{-2}\ s^{-1}$ |
| $\varphi_T$ | Thermal potential, $kT/q$ | V (Volts) |
| $G$ | Carrier generation rate | $cm^{-3}\ s^{-1}$ |
| $G_c$ | Conversion gain | $\mu V/e^-$ |
| $G_{SF}$ | Source follower gain | – |
| $g_m$ | MOSFET gate transconductance | $A\ V^{-1}$ |
| $h$ | Planck's constant | $6.62 \times 10^{-34}\ J\ s$ |
| $I$ | Current | A (Amperes) |
| $i_n$ | Current noise density | $A/\sqrt{Hz}$ |
| $J$ | Current density | $A\ cm^{-2}$ |

| | | |
|---|---|---|
| $k$ | Boltzmann constant | $1.38 \times 10^{-23}$ J K$^{-1}$ |
| $K$ | System gain | $e^-$/ADU |
| $L_n$, $L_p$ | Diffusion length for electrons or holes | cm |
| $\lambda$ | Photon wavelength | nm |
| $\mu_n$, $\mu_p$ | Mobility for electrons or holes | cm$^2$V$^{-1}$s$^{-1}$ |
| $n$ | Electron concentration | cm$^{-3}$ |
| $n_p$ | Electron concentration in $p$-type semiconductor | cm$^{-3}$ |
| $n_{p0}$ | Electron concentration in $p$-type semiconductor in equilibrium | cm$^{-3}$ |
| $N_\mathrm{A}$ | Acceptor concentration | cm$^{-3}$ |
| $N_\mathrm{D}$ | Donor concentration | cm$^{-3}$ |
| $N_\mathrm{ss}$ | Surface trap energy density | cm$^{-2}$eV$^{-1}$ |
| $N_\mathrm{st}$ | Surface trap density | cm$^{-2}$ |
| $N_\mathrm{t}$ | Trap concentration | cm$^{-3}$ |
| $p$ | Hole concentration | cm$^{-3}$ |
| $p_n$ | Hole concentration in $n$-type semiconductor | cm$^{-3}$ |
| $p_{n0}$ | Hole concentration in $n$-type semiconductor in equilibrium | cm$^{-3}$ |
| $P_\mathrm{ph}$ | Optical power | W (Watts) |
| $Q$ | Charge | C (Coulomb) |
| $q$ | Elementary charge | $1.6 \times 10^{-19}$ C |
| $R$ | Resistance | Ω (Ohm) |
| $\sigma$, $\sigma_n$, $\sigma_p$ | Trap capture cross section for electrons or holes | cm$^{-2}$ |
| $S_\mathrm{ph}$ | Photosensitivity | A W$^{-1}$ |
| $S_n$ | Surface recombination velocity for electrons | cm s$^{-1}$ |
| $t$ | Time | s (seconds) |
| $\tau_n$ | Electron lifetime | s |
| $T$ | Temperature | K, °C |
| $U$ | Recombination rate | cm$^{-3}$ s$^{-1}$ |
| $U_\mathrm{s}$ | Surface recombination rate | cm$^{-2}$ s$^{-1}$ |
| $v_\mathrm{th}$ | Electron or hole thermal velocity | cm s$^{-1}$ |
| $v_\mathrm{d}$ | Drift velocity | cm s$^{-1}$ |
| $V_\mathrm{GS}$ | MOSFET gate-source voltage | V |
| $V_\mathrm{T}$ | MOSFET threshold voltage | V |

**IOP** Publishing



**Konstantin D Stefanov**

# Chapter 1

# The fundamentals

## 1.1 Introduction—what is an image sensor and what does it do?

An image sensor has the job to convert an image, consisting of photons emitted or reflected by an object, into an electronic signal. To register a photon, it must be *absorbed* by the sensor and converted to an electric signal. For most sensors, this happens using the internal photoelectric effect. Once the electric signal is registered and processed, we obtain an electronic image, a representation of the incoming photons. The photon energy can span from the far-infrared up to x-rays.

Image sensors come in a bewildering variety of types, shapes and sizes. Most commonly, image sensors are segmented into individual sensitive elements, called pixels. Each pixel is intended to register photons independently of its neighbours, however, in most real-world sensors there is some electrical and optical crosstalk. An array of pixels forms a 2-D image sensor, which is by far the dominant type, and a line of pixels is a linear array, or 1-D image sensor. A simple photodiode has no pixel structure and can therefore be called a 0-D image sensor.

Typically, we expose the image sensor to photons for a certain time, called integration time. During that time, each pixel absorbs photons and registers an electric signal. The job of the image sensor, broadly speaking, is to determine *how many photons have been received in each pixel during the integration time*. Ideally, every photon should be registered separately, and some image sensors can do exactly that. More often, however, the pure photon-induced signal is mixed with the intrinsic noise of the sensor, and we can find out only the *average* number of registered photons. A linear photoresponse is often desired in an image sensor, so that the electrical output signal is proportional to the number of registered photons. Most sensors have small nonlinearity in their response not exceeding a few percent, but some types have intentionally much stronger nonlinearity, for example with logarithmic response, helping to achieve wide dynamic range.

The output signal from a sensor can be a continuous current proportional to the illumination, and this is the way the photodiodes are normally operated. However,

doi:10.1088/978-0-7503-3235-4ch1



© IOP Publishing Ltd 2022

in low light conditions the photogenerated currents can be extremely small and be counted in few electrons *per second*. Such currents are impossible to measure directly, therefore the method we use is to *integrate* the charge over a certain time on a collection element because this creates a signal that is much easier to measure. Most image sensors, including the ones described in this book, are of the integrating type.

## 1.2 Charge generation

### 1.2.1 Photoeffect

For a semiconductor image sensor to detect light, the photons impinging on it must interact with the sensitive regions of the sensor and be converted to an electric charge, which is then collected and recorded.

The dominant process of photon conversion in image sensors is the internal photoeffect. An incident photon with sufficient energy can liberate a valence electron from an atom, which becomes a free electron in the conduction band and can move about in the crystal lattice. At the same time, the missing valence electron becomes a hole, which is also mobile, as shown in figure 1.1. In this way, electrons and holes are created in pairs, and the photon is absorbed and disappears. The minimum photon energy for the photoeffect to occur is the bandgap of the semiconductor $E_g$. The excess photon energy above this threshold is dissipated as crystal vibrations or by generating secondary electron–hole pairs as the primary pair dissipates its kinetic energy.

Photons with energy lower than the bandgap $E_g$ are not able to create electron–hole pairs, and since there is no other mechanism for photons to lose energy, silicon appears transparent at their wavelength. The photon energy $E_{ph}$ is given by:

$$E_{ph} = \frac{hc}{\lambda} \tag{1.1}$$

where $h$ is Planck's constant, $c$ is the velocity of light and $\lambda$ is the photon's wavelength. Equation (1.1) can be more conveniently written as (1.2) where the photon energy $E_{ph}$ is in electron-volts (1 eV = $1.6 \times 10^{-19}$ J) and the wavelength $\lambda$ is in nanometres:

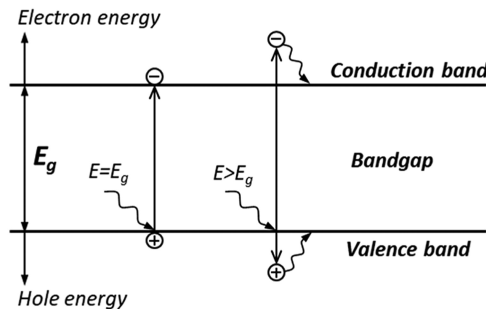


**Figure 1.1.** Photoeffect in semiconductors.

$$E_{ph} = \frac{1240}{\lambda} \tag{1.2}$$

Using (1.2) we can calculate that a photon with energy equal to the bandgap of silicon ($E_{ph} = E_g = 1.12$ eV at 300 K) has a wavelength of 1107 nm, and this is commonly referred to as the cut-off wavelength. The bandgap increases slightly at lower temperatures [2] leading to shorter cut-off wavelength and weaker absorption in the near-IR band.

In silicon, as an indirect bandgap semiconductor, the excitation of a valence electron into the conduction band requires that lattice vibrations (phonons) are involved to obey both energy and momentum conservation laws [2]. As the photon energy increases, the amount of momentum transfer must increase too, and the energy required to generate one electron–hole pair gradually increases. Because of this, a photon with energy equal to double the silicon bandgap (2.24 eV, or 554 nm) still generates one electron–hole pair, and not two.

The band structure of silicon has a direct bandgap of 3.1 eV [3] (400 nm) as well, allowing an electron–hole pair to be created directly, without the assistance of a phonon. However, it is thanks to its indirect bandgap that silicon is sensitive to visible light; if it only had the 3.1 eV direct bandgap it would have been sensitive only to wavelengths shorter than 400 nm and unusable for mainstream imaging.

A single electron–hole pair, corresponding to internal quantum yield of unity, is created up to a photon energy equal to three times the bandgap (3.36 eV, or 369 nm) [4]. Above this energy two e–h pairs begin to be created, but at a very low rate. Multiple e–h pair creation becomes significant only when the photon energy exceeds 4 eV (310 nm) [5, 6].

As the photon energy increases further, the ionisation energy $E_w$ needed for the creation of one e–h pair reaches a peak of approximately 4.5 eV [4, 5]. For $E_{ph} > 10$ eV the pair creation energy $E_w$ levels off to around 3.65 eV, as shown in figure 1.2.

An incoming beam of photons with flux $\Phi_0$ (number of photons per unit area per second) and energy higher than the bandgap is gradually absorbed in the

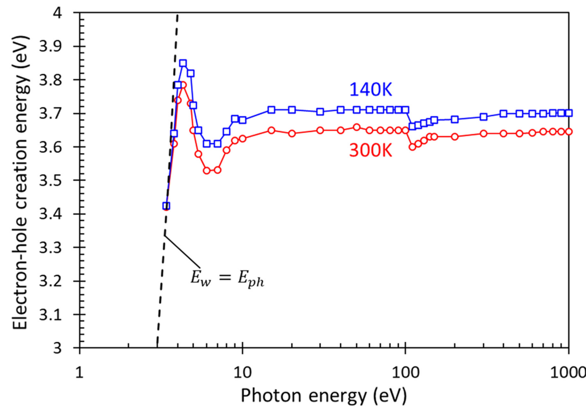


**Figure 1.2.** Electron–hole pair creation energy in silicon at 140 and 300 K (data from [1]). The dashed line marks $E_w = E_{ph}$.

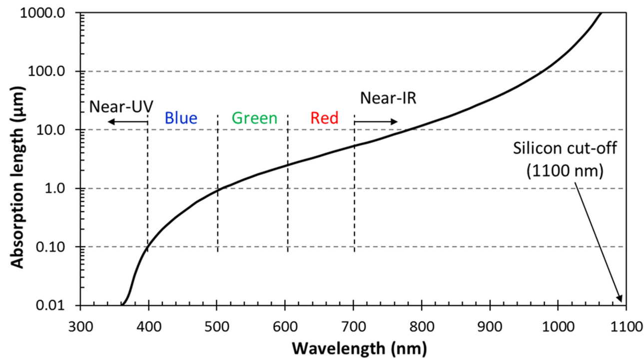


**Figure 1.3.** Photon absorption in silicon for low energy photons from near-UV to near-IR at 300 K. The data is from [7].

semiconductor. Ignoring any reflections, the flux $\Phi(x)$ at a distance $x$ away from the illuminated surface is given by the Beer–Lambert law:

$$\Phi(x) = \Phi_0 e^{-\alpha x} \tag{1.3}$$

where $\alpha$ is the absorption coefficient, typically measured in units of $\text{cm}^{-1}$. At distance $x_0 = 1/\alpha$ the incoming photon flux is attenuated by $1/e$, which means that 63% of the light has been absorbed. The distance $x_0$ is called absorption length and is often more practical to use than the absorption coefficient because it allows straightforward comparison with the dimensions used in image sensors.

The absorption length depends strongly on the wavelength of light and changes by a factor of 50 between the lower end (400 nm) and the top end (700 nm) of the visible light range[1], as shown in figure 1.3 and table 1.1. As the bandgap increases at low temperatures the absorption length increases too, especially at near-IR wavelengths [7]. It is worth noting that photon absorption does not depend on the doping concentration or the free carrier concentration (either electrons or holes) in silicon for most practical cases.

Very often we would like to know what the silicon thickness should be to achieve certain level of photon absorption.

---

**Example 1.1.** Calculate the silicon thickness for 95% photon absorption for light with 400, 700 and 900 nm wavelength.

**Solution:** 95% absorption means that only 5% of the light is left. From formula (1.3) we have $e^{-\alpha x} = 0.05$ and therefore $x = -\ln(0.05)/\alpha = 3.0/\alpha$. From table 1.1 we get $1/\alpha = x_0 = 0.105$ µm for 400 nm wavelength, therefore the thickness is $x = 0.31$ µm. For 700 nm we have $1/\alpha = 5.263$ and 15.8 µm silicon thickness; and for 900 nm $x = 97.9$ µm.

---

[1] In the CIE (The International Commission on Illumination) luminous efficiency functions [8] the wavelength range of visibility is 380–700 nm.

**Table 1.1.** Absorption length in silicon for light wavelengths from 300 to 1100 nm at 300 K. Data from [7].

| Wavelength (nm) | Absorption length (μm) |
|---|---|
| 300 | 0.006 |
| 350 | 0.010 |
| 400 | 0.105 |
| 450 | 0.392 |
| 500 | 0.901 |
| 550 | 1.565 |
| 600 | 2.415 |
| 650 | 3.559 |
| 700 | 5.263 |
| 750 | 7.692 |
| 800 | 11.77 |
| 850 | 18.69 |
| 900 | 32.68 |
| 950 | 63.69 |
| 1000 | 156.3 |
| 1050 | 613.5 |
| 1100 | 2857 |

The light attenuation with distance for the wavelengths used in example 1.1 is plotted in figure 1.4 and illustrates the huge differences in silicon thickness required for the same absorption. This example shows the extremes of the visible range; in practice silicon thickness of 5 μm is often sufficient for visible light imagers because it allows acceptable absorption in the red end of the spectrum around 600–650 nm.

The fact that in the visible wavelength range each photon creates one electron–hole pair can be used to calculate the total light-generated charge in a volume of silicon. This charge, if collected, is the electrical output of the image sensor. Knowing the incident optical power $P_{ph}$ (measured in watts) and the photon energy we can calculate the number of e–h pairs $\Delta N_{e-h}$ generated per unit time $\Delta t$ based on the energy conservation law, simply as this:

$$\frac{\Delta N_{e-h}}{\Delta t} = \frac{P_{ph}}{E_{ph}} \tag{1.4}$$

As we can see the number of generated e–h pairs is inversely proportional to the photon energy, therefore lower energy photons, corresponding to near-IR and red light generate more carriers at the same optical power. While this is true, e–h pair generation requires that the photons are absorbed; for those long wavelengths the silicon must be very thick to ensure full absorption as figure 1.4 tells us.

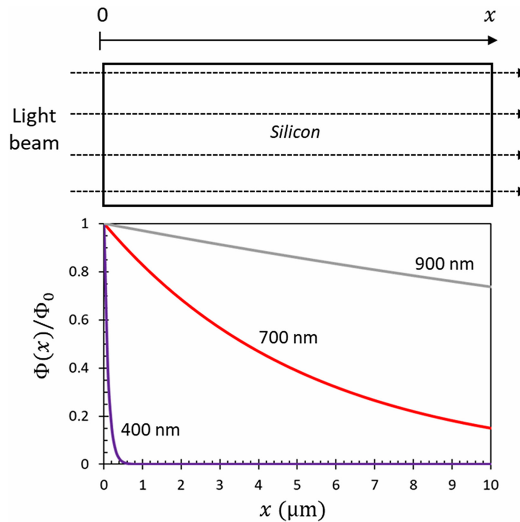


**Figure 1.4.** Photon absorption in silicon for violet ($\lambda = 400$ nm), red ($\lambda = 700$ nm) and for 900 nm light in the near-infrared.

---

**Example 1.2.** Calculate the number of e–h pairs generated per second by red light ($\lambda = 650$ nm) with an irradiance (power per unit area) of $E_e = 1$ W m$^{-2}$ in a square pixel with a size $a = 10$ μm. Assume that the light is fully absorbed in the pixel's volume.

**Solution:** From equation (1.2) we find that for $\lambda = 650$ nm the photon energy is $E_{ph} = 1.91$ eV. The energy deposited in the pixel per second is the irradiance multiplied by the pixel area $a^2$. The number of generated e–h pairs per second is the energy deposited in the pixel per second, divided by the energy to create one e–h pair:

$$\frac{\Delta N_{e-h}}{\Delta t} = \frac{E_e a^2}{q E_{ph}} = \frac{1 \times 10 \times 10^{-6} \times 10 \times 10^{-6}}{1.6 \times 10^{-19} \times 1.91} = 3.27 \times 10^8 \text{ s}^{-1}$$

Here we have multiplied $E_{ph}$ by the elementary charge $q$ to convert the photon energy from eV to Joules. Irradiance of 1 W m$^{-2}$ at 650 nm corresponds to approximately 68.3 lux, or a dimly lit room. This example shows that even meagre illumination manages to create a third of a billion e–h pairs every second in a tiny 10 μm pixel.

---

If all the e–h pairs were collected, a steady state *photocurrent* $I_{ph}$ will flow; it is given by multiplying formula (1.4) by the elementary charge to convert the number or e–h pairs per second to coulombs per second, which is current:

$$I_{ph} = \frac{q \Delta N_{e-h}}{\Delta t} \tag{1.5}$$

**Example 1.3.** Calculate the photocurrent flowing in the pixel in example 1.2.
**Solution:** Multiplying the answer by the elementary charge gives:

$$I_{ph} = \frac{q\Delta N_{e-h}}{\Delta t} = 1.6 \times 10^{-19} \times 3.27 \times 10^{8} = 5.235 \times 10^{-11}\text{A} = 52.4 \text{ pA}$$

Very often the sensitivity of a semiconductor material is given as the photocurrent per watt of incident light energy. To compare devices irrespective of their structure or pixel size the current can be expressed as *current density* $J_{ph}$, i.e. amperes per unit area. If we use the optical power per unit area $E_e$ (irradiance), equation (1.5) can be rewritten by dividing both sides by the device (or pixel) area, and using (1.4) we get

$$J_{ph} = \frac{qE_e}{E_{ph}}. \tag{1.6}$$

From (1.6) we arrive at the astonishingly simple expression (1.7) for the sensitivity $S_{ph}$ in terms of detector current density per unit of optical irradiance. $S_{ph}$ is measured in units of ampere per watt (A $W^{-1}$) because the area cancels from both $J_{ph}$ and $E_e$.

$$S_{ph} = \frac{J_{ph}}{E_e} = \frac{q}{E_{ph}} \tag{1.7}$$

Equation (1.7) gives the theoretical maximum photosensitivity of an ideal image sensor having perfect light absorption, without any light losses due to reflection, and of course with complete charge collection. The ratio $E_{ph}/q$ gives the photon energy in units of eV. From here we can calculate that the theoretical maximum photosensitivity for $\lambda = 650$ nm (as in example 1.2) is $S_{ph} = 1/1.91 = 0.52$ A $W^{-1}$. Using (1.2), expression (1.7) can also be written as $S_{ph} = \lambda/1240$, where $\lambda$ is in nanometres.

### 1.2.2 Ionisation

Silicon makes an excellent sensor material not just for the near-IR, visible and UV light, but also for much more energetic photons, such as x-rays. The absorption length covering photon energies from 1.2 eV to 10 keV in figure 1.5 shows what happens at the higher end of silicon's usable range: similarly to the near-IR end, silicon becomes transparent beyond photon energy of about 10 keV. The absorption above 100 eV shows discontinuous absorption edges at the energy levels of the L-shell (≈100 eV) and the K-shell (1839 eV) of the silicon atom. As the photon energy exceeds the binding energy of a shell, the electrons occupying it can be excited and the photon absorption sharply goes up. This corresponds to a stepwise *decrease* in the absorption length, most clearly seen at the K-edge.

For photons with $E_{ph} > \approx 50$ eV the ionisation energy is $E_w = 3.65$ eV at 300 K [11] and is nearly constant. This allows us to calculate the number of electron–hole pairs $N_{e-h}$ generated by x-rays and gamma-rays using this simple expression:

$$N_{e-h} = \frac{E_{ph}}{E_w} \tag{1.8}$$

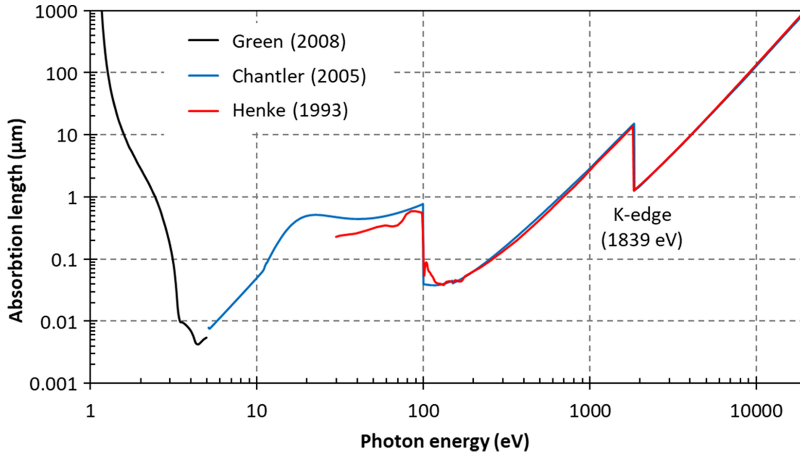


**Figure 1.5.** Photon absorption due to photoeffect in silicon for a wider energy range. Data for <5 eV from [7], 5 – 20000 eV from [9], and >30 eV from [10].

The ionisation energy is temperature dependant; this is due to the reduction of the bandgap as the temperature increases [1, 12]. X-rays in the range 1–10 keV, emitted by radioactive sources and x-ray fluorescence from various materials are particularly useful for sensor characterisation. They are widely used for calibration due to the well-known x-ray energies and the amount of charge created in silicon by them. Also, the initial charge cloud created by low energy x-rays is very compact [13, 14] and this allows the charge to be considered a point source.

One very popular calibration source is the $^{55}$Fe isotope which decays via electron capture to manganese ($^{55}$Mn) with a half-life of 2.737 years. $^{55}$Mn emits characteristic K-shell x-rays with energies 5.89 keV (Mn-K$_\alpha$) and 6.49 keV (Mn-K$_\beta$), with probabilities of 24.4% and 2.9%, correspondingly [15]. Figure 1.6 shows an example of a $^{55}$Fe spectrum obtained by a CMOS image sensor.

The absorption length for 5.9 keV photons in silicon is approximately 28 μm [10]. This length is much larger than the depth of the active silicon in the typical optical sensor, therefore only a small fraction of the incoming x-rays is absorbed and converted to charge.

---

**Example 1.4.** Calculate the number of electron–hole pairs generated by the dominant Mn-K$_\alpha$ x-ray, and the current in a pixel receiving one hundred Mn-K$_\alpha$ x-rays per second, assuming that all the charge is collected.

**Solution:** Using formula (1.8) the number of generated electron–hole pairs per x-ray is

$$N_{\text{e-h}} = \frac{E_{\text{K}\alpha}}{E_{\text{w}}} = \frac{5890}{3.65} = 1614.$$

The current for 100 x-rays per second is calculated using (1.5):

$$I_{\text{ph}} = \frac{qN_{\text{e-h}}}{t} \times 100 = \frac{1.6 \times 10^{-19} \times 1614}{1} \times 100 = 25.8 \text{ fA}$$

---

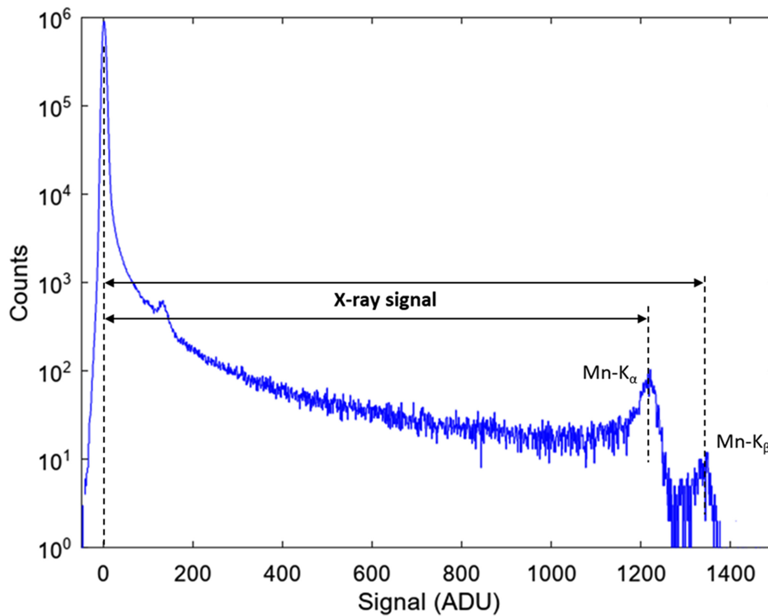


**Figure 1.6.** $^{55}$Fe spectrum obtained by a CMOS image sensor. The peak at 0 ADU is due to pixels without x-ray signal, and the continuum leading to the x-ray peaks is caused by charge collected by more than one pixel.

At much higher photon energies two other mechanisms overtake the photoeffect and begin to play an increasing role—Compton effect and electron–positron pair creation [16]. Besides the detection of optical and x-ray photons, silicon is widely used for the detection of high energy, ionising particles. Traversing the material, charged particles lose energy due to several mechanisms, and most of that energy loss is due to ionisation. The ionisation loss is very high at low particle energies but decreases and flattens off at higher energies in a logarithmic dependence [16]. Particles with energies at and above the plateau, which usually lies at hundreds of MeV, are called minimum ionising particles (MIP).

The average number of the created electron–hole pairs is calculated by the most probable energy loss divided by the ionisation energy as in (1.8). The energy loss due to ionisation has large statistical fluctuations. For particles which lose only a small part of their energy in the material (i.e. the material is 'thin') the energy loss is described by the Landau distribution [17]. In a couple of microns of silicon there is a significant probability that a traversing high energy particle will cause no ionisation at all [18]. At the same time, the probability that the energy loss can far exceed the most probable value is also significant, due to the long tail in the Landau distribution.

The most probable number of e–h pairs increases with the thickness, as shown in figure 1.7, and above 10 μm the energy loss begins to approach the Landau distribution.

## 1.3 Charge collection

In the previous section we looked at how photons generate free charge, consisting of electrons and holes, and how the amount of charge depends on the photon energy. The question now is how to collect this charge and measure it.

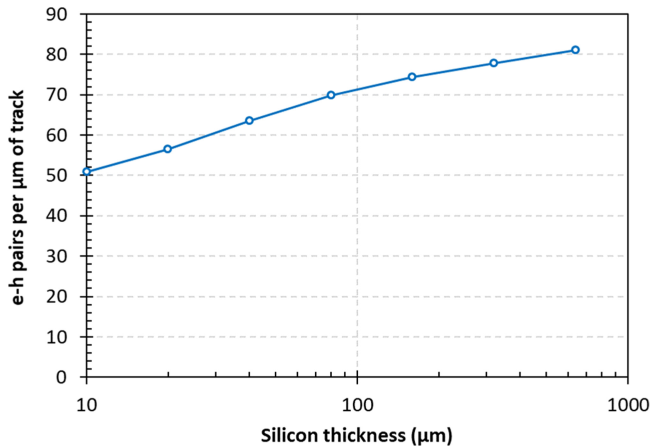


**Figure 1.7.** Most probable number of electron–hole pairs per micrometre of track for high energy charged particles, data from [17].

Left on its own, the charge will simply diffuse out, never to be seen again. Diffusion is fundamental in nature and always occurs when there is a difference in carrier concentration. The charge generated in pixels receiving more illumination will diffuse towards pixels receiving less illumination, until we get nearly uniform charge 'blob' everywhere. Obviously, this is not what we want to happen in an image sensor.

We need a charge collecting element—something that is electrically attractive to either electrons or holes (but obviously cannot be attractive to both). To make the charge move in a particular direction for collection we need to create an *electric field*; within it the charge experiences an electrostatic force and begins to accelerate in a direction opposite to the field (for electrons) and along the field (for holes). This charge movement in the presence of electric field is called *drift* and is the primary mechanism for charge collection. During drift the charge continues to diffuse due to its concentration gradient, regardless of the presence of any electric field; this is unavoidable but not always undesirable.

### 1.3.1 Carrier lifetime

An important point in image sensor operation, which often goes without much mention, is that the generated e–h pairs must survive, i.e. not recombine or get trapped, for sufficiently long time so that they can be collected. The characteristic describing the 'life duration' of electrons and holes is called *carrier lifetime* [19] and is widely used in semiconductor physics. What 'sufficiently long' means in practice will be explored in the next two sections. As a rough indicator the charge collection time rarely exceeds a few hundreds of nanoseconds, and this is how long the carriers must survive. Carrier lifetime can be many orders of magnitude longer, especially in high quality epitaxial silicon.

Whenever electrons and holes are created, for example by illumination with light, the excess carrier concentration will decay back to equilibrium after the source of e–h pair generation is turned off. In silicon the dominant physical mechanism for the decay is trap-assisted recombination. Direct e–h recombination occurs too, but at a much smaller rate. The rate of decay towards the equilibrium concentration, expressed as the change of carrier concentration per unit time $U = \Delta n/\Delta t$, is called *recombination rate*. The simplest possible mathematical description of this process is to assume that the recombination rate $U$ is proportional to the excess carrier concentration. For example, if the electron concentration in $p$-type silicon is $n_p$ and the equilibrium concentration is $n_{p0}$, the recombination rate can be written as $U \propto (n_p - n_{p0})$. Since proportionality is assumed, we need a proportionality constant in units of seconds, so that the recombination rate is measured in units of carrier concentration over time (cm$^{-3}$ s$^{-1}$ in semiconductor physics). This constant is called the carrier lifetime ($\tau_n$ for electrons, $\tau_p$ for holes) and can be thought of as the characteristic time over which the carrier concentration decreases. Now, the recombination rate can be written in its familiar form:

$$U = -\frac{dn_p}{dt} = \frac{n_p - n_{p0}}{\tau_n} \tag{1.9}$$

We have added a negative sign in (1.9) because due to recombination the carrier concentration decreases ($dn_p/dt < 0$) when $n_p - n_{p0} > 0$, and $U > 0$. Because $n_{p0}$ is constant, we can write that

$$\frac{d(n_p - n_{p0})}{dt} = -\frac{n_p - n_{p0}}{\tau_n} \tag{1.10}$$

and after separating the variables we can integrate both sides:

$$\int \frac{d(n_p - n_{p0})}{n_p - n_{p0}} = -\frac{1}{\tau_n}\int dt \tag{1.11}$$

$$\ln(n_p - n_{p0}) + \text{const} = -\frac{t}{\tau_n} \tag{1.12}$$

The final equation can be written by using the initial conditions: at $t = 0$ the excess electron concentration is $n_p(0) - n_{p0}$ and at $t \to \infty$ naturally $n_p - n_{p0} = 0$, with only the equilibrium concentration $n_{p0}$ left. Therefore, the constant in equation (1.12) must equal $-\ln\left(n_p(0) - n_{p0}\right)$ and we arrive at the time dependence of $n_p$:

$$n_p(t) = n_{p0} + (n_p(0) - n_{p0})e^{-\frac{t}{\tau_n}} \tag{1.13}$$

Equation (1.13) tells us that the electron lifetime $\tau_n$ is the characteristic time over which the excess carrier concentration decreases by $1/e$, i.e. only 37% of the excess carriers remain. After three times the lifetime only 5% of the initial charge will be left.

### 1.3.2 Recombination

This is a good place to answer an important question: why don't the electrons and the holes recombine immediately after they are generated? After all, they are created together and close to each other, and it would be reasonable to expect that they should recombine at high rate. Fortunately, such direct (band-to-band) recombination in silicon is very rare because it is an indirect bandgap semiconductor. Carrier lifetime controlled by band-to-band recombination is very long; in high purity silicon the electron and hole lifetimes can be many milliseconds. This long lifetime allows the charge to diffuse out a long distance from the place it was generated without recombining, unless it is quickly collected with the help of an electric field.

Figure 1.8 shows two direct, and very rare band-to-band recombination mechanisms: radiative with the emission of a photon, and Auger recombination where the excess energy is transferred to another carrier, such as a hole in highly doped *p*-type silicon.

The third one, trap-assisted recombination via a mid-band trap, is by far the dominant mechanism in silicon, described by the Shockley–Read–Hall (SRH) theory [20]. Traps are produced by imperfections or impurities in the crystal lattice, which introduce energy levels deep in the bandgap. Traps take part in both capture and emission of carriers and are also called generation-recombination centres.

The recombination rate from a trap with concentration $N_t$ and energy level $E_t$ above the valence band is given by

$$U = \frac{\sigma_n \sigma_p v_{\mathrm{th}}(pn - n_i^2)N_t}{\sigma_n\left[n + n_i \exp\left(\frac{E_t - E_i}{kT}\right)\right] + \sigma_p\left[p + n_i \exp\left(-\frac{E_t - E_i}{kT}\right)\right]} \tag{1.14}$$

Here $E_i$ is the intrinsic Fermi level (approximately the mid-band energy level);
$k$ is the Boltzmann's constant;
$T$ is the absolute temperature;
$v_{\mathrm{th}}$ is the carrier thermal velocity;
$n_i$ is the intrinsic carrier concentration;
$n$ and $p$ are the electron and hole concentrations, respectively;
$\sigma_n$ and $\sigma_p$ are the electron and hole capture cross-sections, respectively.

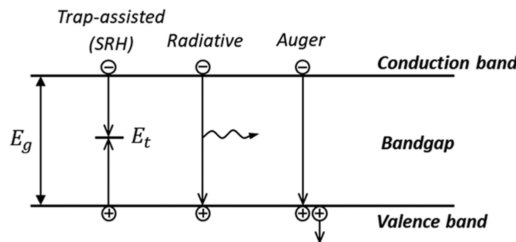


**Figure 1.8.** Recombination mechanisms.

The maximum recombination rate occurs when the denominator is at its minimum, achieved when $E_t = E_i$. This indicates that mid-band traps are the most effective in limiting the carrier lifetime.

In a $p$-type semiconductor we have $p \gg n_i$ and $p \gg n$, and if the two capture cross-sections are similar ($\sigma_n \cong \sigma_p$), the second term in the denominator of (1.14) becomes much larger than the first. Therefore, we can write that

$$U \approx \frac{\sigma_n v_{\text{th}}(pn - n_i^2)N_t}{p} \tag{1.15}$$

Introducing a small excess concentration of electron–hole pairs does not change significantly the equilibrium hole concentration $p_{p0}$ in $p$-type silicon. Using that $p \approx p_{p0}$ and $n_i^2 = p_{p0} n_{n0}$ [21], equation (1.15) can be written as

$$U \approx \frac{\sigma_n v_{th}(p_{p0} n_p - p_{p0} n_{n0})N_t}{p_{p0}} = \sigma_n v_{\text{th}} N_t (n_p - n_{n0}) \tag{1.16}$$

Comparing with (1.9) we see that the term multiplying the concentration difference is the inverse of the electron lifetime (also called recombination lifetime)

$$\tau_n = \frac{1}{\sigma_n v_{\text{th}} N_t} \tag{1.17}$$

Typical capture cross-sections are in the range $10^{-16}$ to $10^{-14}$ cm$^2$.

---

**Example 1.5.** Calculate the electron lifetime due to a mid-band trap ($E_t = E_i$) with $\sigma_n = 10^{-15}$ cm$^2$ and concentration $10^{12}$ cm$^{-3}$, using that the electron thermal velocity is $v_{\text{th}} = 1.4 \times 10^7$ cm s$^{-1}$.
**Solution:** Using (1.17) we get

$$\tau_n = \frac{1}{10^{-15} \times 1.4 \times 10^7 \times 10^{12}} = 71.4\ \mu\text{s}$$

To put this into perspective, trap concentration of $10^{12}$ cm$^{-3}$ corresponds to an average of one trap per cubic micrometre, or one trap per 50 billion Si atoms.

---

Measurements show that in high quality, lightly doped ($<10^{15}$ cm$^{-3}$) silicon the minority carrier lifetime can be tens of milliseconds [22]. Considering the calculation in example 1.5, this implies that the trap density responsible in the SRH model must be less than $10^{10}$ cm$^{-3}$, or one trap per 5 trillion atoms. As the dopant concentration increases above $10^{16}$ cm$^{-3}$ the lifetime begins to decrease and this is taken into account as concentration-dependent SRH lifetime [23].

At high carrier density, such as along a dense ionisation track or in solar cells, Auger and band-to-band radiative recombination can begin to limit the lifetime even in low-doped silicon. Auger recombination involves a direct recombination between

an electron and a hole, with the excess energy transferred to another electron or hole. The Auger lifetime for high excess carrier concentration in a lightly doped semiconductor is [22]:

$$\tau_{\mathrm{Auger}} = \frac{1}{C_{\mathrm{a}}\Delta n^2} \tag{1.18}$$

Here $C_{\mathrm{a}}$ is the ambipolar Auger coefficient ($1.66 \times 10^{-30}$ cm$^6$ s$^{-1}$ in silicon [24]) and $\Delta n = n_p - n_{n0}$ or $\Delta n = p_n - p_{p0}$ is the excess carrier concentration.

The lifetime due to band-to-band radiative recombination is given by

$$\tau_{\mathrm{rad}} = \frac{1}{B\Delta n} \tag{1.19}$$

where $B = 4.7 \times 10^{-15}$ cm$^3$ s$^{-1}$ is the radiative coefficient in silicon at 300 K [25]. It is easy to see that the Auger and radiative lifetimes are much larger than $\tau_{\mathrm{SRH}}$ and can become comparable to the SRH lifetime only at excess concentration above $10^{16}$–$10^{17}$ cm$^{-3}$. In a typical image sensor, the excess carrier concentration due to illumination rarely exceeds $10^8$–$10^{10}$ cm$^{-3}$, therefore the direct recombination mechanisms have negligible influence.

The total lifetime $\tau_{\mathrm{tot}}$ can be calculated from Matthiessen's rule for the three recombination processes as in [19]

$$\frac{1}{\tau_{\mathrm{tot}}} = \frac{1}{\tau_{\mathrm{SRH}}} + \frac{1}{\tau_{\mathrm{Auger}}} + \frac{1}{\tau_{\mathrm{rad}}} \tag{1.20}$$

Equation (1.20) is analogous to the one used to calculate the resistance of parallel resistors; physically it means that the different recombination mechanisms work independently and in parallel.

### 1.3.3 Drift

We are going to consider a hypothetical collection element without specifying what it is and how it is made; then in the following section we will talk about two real charge collection elements—the *pn* junction and the MOS capacitor.

Figure 1.9 shows our hypothetical collection element. From the surface down to depth $d$ there is a constant electric field $E$, and below $d$ the field is zero. This may look artificial but is not far off from reality.

In this structure only the electrons are collected, and the holes are discarded never to be seen again, as is typical for most image sensors. Electrons are preferred because they move much faster than the holes, resulting in shorter collection times. Holes are allowed to diffuse until they reach the backside substrate electrode, or they recombine after travelling a long distance away from the charge collection element.

Within a region having an electric field $E$ electrons experience a force $F = -qE$ and begin to accelerate. Holes experience the same force but with the opposite sign and move in the other direction. As mentioned previously, this movement under the influence of an electric field is called *drift*. In semiconductors it is experimentally

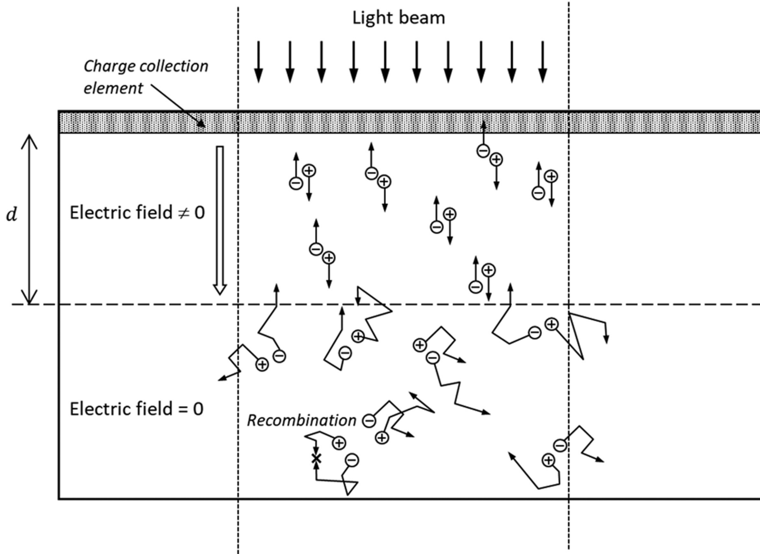


**Figure 1.9.** Charge collection of electrons created by a light beam and experiencing drift and diffusion.

observed that at low electric fields (<~$10^4$ Vcm$^{-1}$) the charge carriers acquire drift velocity $v_d$ proportional to the electric field $E$:

$$v_d = \mu E \tag{1.21}$$

The proportionality factor $\mu$ is called *mobility* and is not a constant—it decreases as the temperature and the doping concentration increase [26]. The electron mobility $\mu_n$ at low fields is about three times higher than the hole mobility $\mu_p$; the values in low-doped silicon at 300 K are $\mu_n$ = 1400 cm$^2$ V$^{-1}$ s$^{-1}$ and $\mu_p$ = 470 cm$^2$ V$^{-1}$ s$^{-1}$ [26]. The higher electron mobility and velocity is one of the main reasons why we prefer collecting and transferring electrons, rather than holes.

The drift velocity decreases below the value calculated by (1.21) at higher electric fields (above $10^4$ V cm$^{-1}$), and stops increasing altogether (i.e. saturates) at $E > 10^5$ V cm$^{-1}$. The saturation drift velocity $v_d^{sat}$ for both electrons and holes is approximately $10^7$ cm s$^{-1}$.

Drift velocity is directional and determined by the applied electric field; it is also superimposed on the thermal carrier velocity caused by their random movement in the crystal lattice. The thermal velocity for electrons is given by

$$v_{th} = \sqrt{\frac{3kT}{m_0^*}} \tag{1.22}$$

where $k$ is the Boltzmann constant, $T$ is the absolute temperature and $m_0^* = 0.26m_0$ is the effective electron mass [2]. At 300 K formula (1.22) gives $v_{th} = 2.3 \times 10^7$ cm s$^{-1}$. This is similar to the experimentally observed saturation velocity $v_d^{sat}$, and is an indication that the simple proportionality in formula (1.21) is valid only when $v_d \ll v_{th}$.

---

**Example 1.6.** Calculate the electron drift velocity for $E = 1000$ V cm$^{-1}$ and electron mobility $\mu_n = 1400$ cm$^2$ V$^{-1}$ s$^{-1}$ (value for Si at 300 K and for low doping concentration), and compare it to the thermal velocity $v_{\text{th}}$.
**Solution:** From formula (1.21) we get

$$v_{\text{d}} = \mu_n E = 1400 \times 1000 = 1.4 \times 10^6 \text{ s}^{-1}$$

This velocity is about 20 times lower than the random thermal electron velocity.

---

Knowing the drift velocity allows us to calculate the time it takes to collect the charge. For simplicity we can consider that $v_{\text{d}}$ is much smaller than the saturation velocity. The travel distance $x$ is simply the velocity multiplied by the time:

$$x = v_{\text{d}} t = \mu_n E t \tag{1.23}$$

The charge collection time $t$ is the device thickness divided by the drift velocity, and substituting the drift velocity from (1.21) we arrive at:

$$t = \frac{d}{v_{\text{d}}} \cong \frac{d}{\mu_n E} = \frac{d^2}{\mu_n V} \tag{1.24}$$

Here we have used that the electric field is the applied voltage $V$ divided by the thickness $d$. This is an approximate and simple, but very useful formula; we will refine it further in the following sections.

---

**Example 1.7.** Calculate the charge collection time in silicon with thickness $d = 5$ μm and applied voltage across it $V = 1$ V. The electron mobility is $\mu_n = 1400$ cm$^2$ V$^{-1}$ s$^{-1}$. Compare with the charge collection time under velocity saturation ($v_{\text{d}}^{\text{sat}} \cong 10^7$ s$^{-1}$).
**Solution:** First, we need to see how the electron drift velocity compares to the saturation velocity. Using (1.21)

$$v_{\text{d}} = \mu_n E = \mu_n \frac{V}{d} = 1400 \times \frac{1}{5 \times 10^{-4}} = 2.8 \times 10^6 \text{ s}^{-1}$$

we see that it is about a factor of 3 lower than $v_{\text{d}}^{\text{sat}}$, therefore formula (1.24) can be used and gives:

$$t = \frac{(5 \times 10^{-4})^2}{1400 \times 1} = 0.18 \text{ ns}$$

The collection time under velocity saturation is calculated as

$$t = \frac{d}{v_{\text{d}}^{\text{sat}}} = \frac{5 \times 10^{-4}}{10^7} = 0.05 \text{ ns}$$

---

This example shows that for the typical sensor thicknesses the charge collection time by drift is very short and may become an issue only if very fast operation is required. However, in a much thicker sensor, made so that it can have higher absorption at near-IR wavelengths, the charge collection time could be substantially longer due to the quadratic dependence on the device thickness.

### 1.3.4 Diffusion

Charge generated in a field-free semiconductor diffuses out from the point at which it is created and can travel large distances. It can reach a region with an electric field, where it is swept away, or it can recombine with the assistance of bulk or surface traps.

If we generate a point-like sphere of electron–hole pairs, they will expand stochastically in a cloud described by the Gaussian distribution. After time $t$, the RMS cloud radius $r_n$ for electrons and $r_p$ for holes in one dimension is given by:

$$r_n = \sqrt{2D_n t} \tag{1.25}$$

$$r_p = \sqrt{2D_p t} \tag{1.26}$$

where $D_n$ is the diffusion coefficient for electrons and $D_p$ for holes, measured in $cm^2\ s^{-1}$. The diffusion coefficients are connected to the mobility via the Einstein relationship:

$$D_n = \mu_n \frac{kT}{q} \tag{1.27}$$

$$D_p = \mu_p \frac{kT}{q} \tag{1.28}$$

The diffusion radii are the standard deviations of the charge density spread; from the properties of the Gaussian distribution, we know that 95% of the charge is contained within radius $2r_n$ for electrons and $2r_p$ for holes.

Figure 1.10 shows the diffusion spread with time of a point-like charge generated at $t = 0$ without any recombination. The electron density is described by a Gaussian with standard deviation given by (1.25), with the peak moving by a distance $x$ determined by (1.23) when the electric field is not zero.

Electrons spread $\sqrt{3}$ times faster than holes due to their diffusion coefficient being three times higher; the values can be calculated from (1.29) and (1.30) and are $D_n = 36\ cm^2\ s^{-1}$, $D_p = 12\ cm^2\ s^{-1}$ at 300 K.

During diffusion the charge carrier concentration decreases because they spread out; at the same time their concentration decreases also because they are subjected to various recombination processes, with their combined influence reflected in the carrier lifetime. The longest distance the carriers can travel is naturally limited by their lifetime and is called *diffusion length*. It is an important parameter in semiconductors and enters numerous formulas describing image sensor operation. The diffusion length $L_n$ for electrons is defined by:

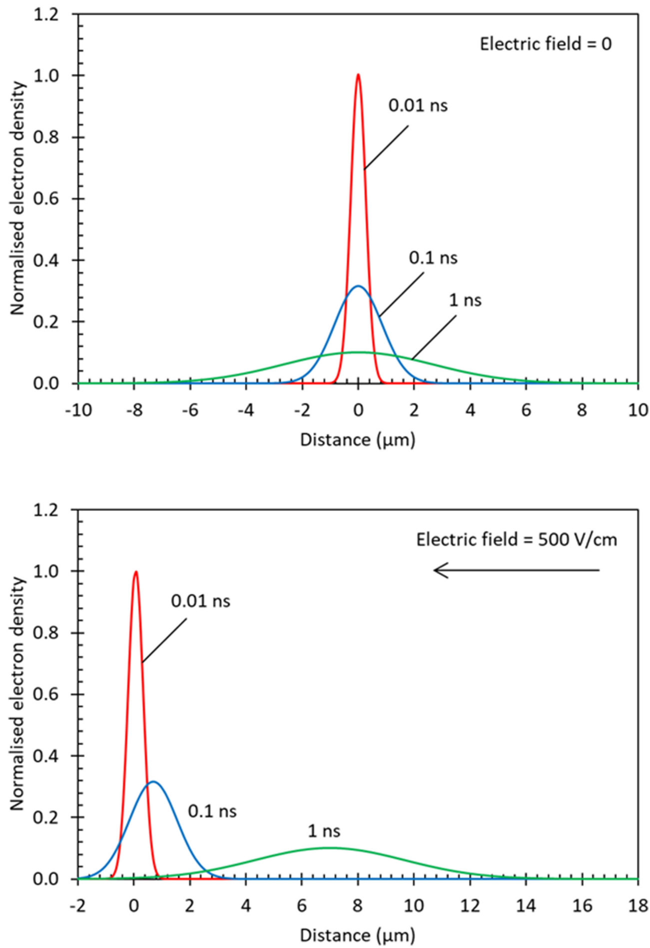


**Figure 1.10.** Electron spread only due to diffusion (at zero electric field) and due to drift and diffusion with electric field = 500 V cm$^{-1}$ (after [2], p 55).

$$L_n = \sqrt{D_n \tau_n} \tag{1.29}$$

and a similar expression can be written for hole diffusion length $L_p$. The diffusion length is usually much larger than the typical pixel sizes due to the long carrier lifetime.

---

**Example 1.8.** Calculate the diffusion length for electrons in silicon with electron lifetime $\tau_n = 1$ ms (fairly typical for low-doped, high quality epitaxial silicon). The diffusion coefficient is $D_n = 36$ cm$^2$ s$^{-1}$.

**Solution:**

$$L_n = \sqrt{36 \times 10^{-3}} = 1897\ \mu\text{m}$$

This is a *really long* distance; some image sensors are physically smaller than this!

---

Since charge travels so well on its own, can we collect it using only diffusion? The short answer is 'No'—collection entirely by diffusion can lead to significant losses because the charge is not going to stay in the confines of the pixel where it was generated. Diffusion is isotropic, and charge moves in all directions. To capture it, we need the collection element to surround the charge on all sides, or a special structure that forces the charge to go in one predominant direction.

Drift is a far better choice for charge collection because it is directional towards the source of electric field. It is also much faster and minimises the chance of charge loss. The distance travelled under drift (1.23) is proportional to time, while the diffusion radius (1.25) increases much slower as a square root. Also, electrons move three times faster than holes in electric field and not only $\sqrt{3}$ times faster as in diffusion.

---

**Example 1.9.** Calculate the electron collection time if the charge moves only by diffusion for the values in example 1.7, assuming that the charge is forced to travel in one direction, and there is no charge loss.
**Solution:** Using equation (1.25) we can calculate the time it takes the electrons to travel the longest distance, equal to the depth of the sensor $r_n = d = 5\ \mu$m:

$$t = \frac{r_n^2}{2D_n} = \frac{(5 \times 10^{-4})^2}{2 \times 36} = 3.5 \text{ ns}$$

Because of the stochastic nature of the diffusion this time can be considered as a time constant; three time constants would suffice for 95% charge collection, and equals to 10.5 ns. This time is two orders of magnitude longer than in charge collection by drift.

---

## 1.4 Charge transfer

After the charge is collected it needs to be converted to a voltage or a current that can be measured using electronic circuits. This could happen at the collection element itself, but very often the tasks of collection and conversion are physically separated for good reasons. The charge needs to travel to a dedicated place where it is converted to an electrical signal; therefore, the task is to perform an efficient *charge transfer*.

Figure 1.11 shows schematically three cases of charge transfer. Figure 1.11(a) corresponds to CMOS image sensors using pinned photodiode as a collection element and is probably the most widely used. Here the photogenerated electrons are directed to the collection element by an electric field. The conversion element is shielded from direct charge collection and receives only the charge transferred to it. This is usually accomplished by either an optical shield over the conversion element, or electrically—by steering the electrons away from it.

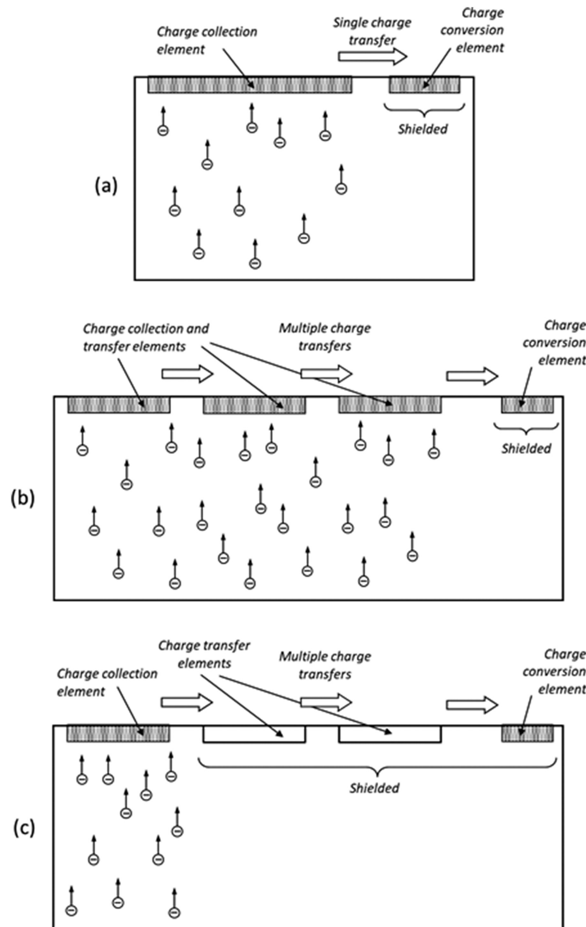


**Figure 1.11.** Charge transfer in image sensors: (a) single transfer; (b) multiple transfers with charge collection elements capable of charge transfer; (c) multiple transfer with dedicated transfer elements which do not collect charge.

Figures 1.11(b) and 1.11(c) show examples where the charge travels larger distances, and the transfer is done is many steps. The transfer elements can either be the same as the charge collection elements, as in full frame CCDs, or they can be dedicated to charge transport only, as in interline transfer CCDs.

## 1.5 Charge conversion

Charge conversion is a crucial step that occurs at a *sense node*, where the charge is converted to an electrical signal—most frequently a voltage. This can happen *destructively*—meaning that after the conversion the charge cannot be recovered or returned to its original state. This is the typical conversion in image sensors.

The conversion can also occur *non-destructively*, which is when the charge remains unaltered and intact, and the same charge can be converted multiple times.

A circuit allowing non-destructive conversion would normally couple capacitively to the charge without getting in physical contact with it.

Why is destructive conversion preferred? Because it is more sensitive and a larger electrical signal can be obtained for the same charge, it is usually sufficient to convert the signal into voltage only once. Multiple signal measurements requiring non-destructive conversions are used only in some specialised sensors where they offer some advantages, such as charge measurement during collection, or for noise reduction.

Figure 1.12 shows four methods of charge conversion: (a)–(c) are destructive because there is an electrode attached to the collection element; once the electrons reach it, they exit the device through that external connection and there no easy way of reversing this process.

In figure 1.12(a) the photogenerated current is continuously flowing out of the device through a resistor; the voltage drop across it (i.e. the difference between the bias voltage and the output) is the measure of the photocurrent.

Figure 1.12(b) shows the typical charge conversion circuit in the vast majority of image sensors. A transistor act as an electronic switch to momentarily connect the output to a stable bias voltage, and after that it is disconnected. The output terminal has an equivalent capacitance $C$ to ground, and when the charge $Q$ reaches the output a voltage step is produced:

$$\Delta V = \frac{Q}{C} \tag{1.30}$$

This voltage step is proportional to the photogenerated charge and is the output signal. Figure 1.12(c) shows a more sophisticated version which uses a Charge

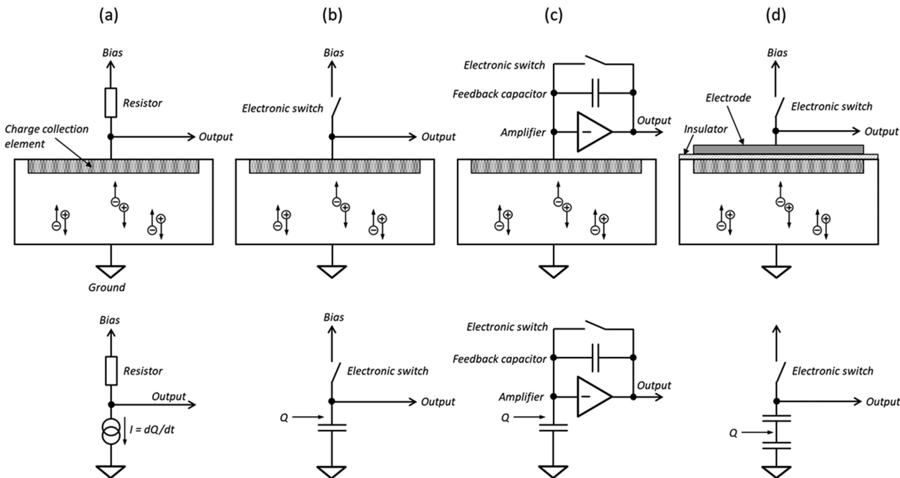


**Figure 1.12.** Charge conversion types: (a) continuous current on a resistor; (b) on the effective capacitance of the charge collection element; (c) using a charge-sensitive amplifier; (d) non-destructive by electrostatic induction.

Sensitive Amplifier (CSA). Due to the negative feedback the CSA 'moves' the incoming charge to the feedback capacitor and the voltage step (1.30) appears across it, while the input stays at nearly constant voltage. The output signal can be large because the feedback capacitor can be made very small.

Figure 1.12(d) illustrates the fourth method of charge conversion discussed here, where the charge is sensed non-destructively by capacitive coupling. The output is the top electrode of a floating capacitor, separated by an insulator from the structure underneath. The collection element has capacitance to both the output and to ground, so the equivalent circuit has two capacitors connected in series. After the switch is disconnected, the collected charge couples by electrostatic induction to the output electrode and produces a voltage change across the effective capacitance. This method does not interfere with the collected charge because there is no connection to it.

The common feature between the charge conversion methods in figures 1.12(b) and (c) is that the charge is converted to voltage on a capacitance. This may not be a physical capacitor as the feedback capacitor in figure 1.12(c), but an effective capacitance to ground formed by the structure of the charge collection element.

The conversion to voltage is characterised by the *conversion gain* $G_c$, expressed as the voltage change at the output per one collected electron. Using (1.30), the conversion gain is written as:

$$G_c = \frac{q}{C} \tag{1.31}$$

where $q$ is the elementary charge and $C$ is the *conversion capacitance*. The term charge to voltage factor (CVF) is also frequently used as a synonym for the conversion gain.

## 1.6 *pn* junction

The *pn* junction is arguably the most important element in image sensor technology. It can be used for both charge collection and charge-to-voltage conversion and is used in practically all image sensors. Many excellent books describing the *pn* junction have been written (for example, [2] and [21]) and we are going to spend a great deal on it too, due to its importance to image sensors.

### 1.6.1 *pn* junction in equilibrium

The *pn* junction is one of the types of charge collection elements with internal electric field, discussed in the previous section. To form a *pn* junction a donor dopant is implanted into *p*-type silicon (or the other way around), and then the device is annealed to activate the implant[2]. The free electrons from the *n*-side start diffusing into the *p*-type silicon and the free holes from the *p*-side into the *n*-type silicon; this occurs naturally because of the concentration differences. Donor and acceptor atoms on both sides of the junction are left without their electrons and holes to keep

[2] Nobody is forming junctions by bringing *p*- and *n*-type silicon bars together.

them neutral, and this creates a region of fixed *space charge*—positively charged donor atoms on the *n*-side and negatively charged acceptor atoms on the *p*-side. This fixed charge creates a built-in electric field which counteracts the diffusion of both majority carriers. The same field, however, is accelerating the minority holes in the *n*-side and the minority electrons in the *p*-side towards the opposite sides of the junction. As the space charge region grows, the electric field increases, slowing down the diffusion until an equilibrium is reached. The two opposing currents balance each other for both electrons and holes, the net flow becomes zero and the width of the space charge region settles to $x_n$ on the *n*-side to $x_p$ on the *p*-side as shown in figure 1.13.

The concentration of majority electrons and holes in the space charge region is nearly zero due to the electric field which forces them away. The space charge region is depleted from majority carriers and this is why it is more often called the *depletion region*. The depletion behaves essentially as an insulator, separating the electrically neutral (and conducting due to the large carrier concentration) *n* and *p* regions on both sides.

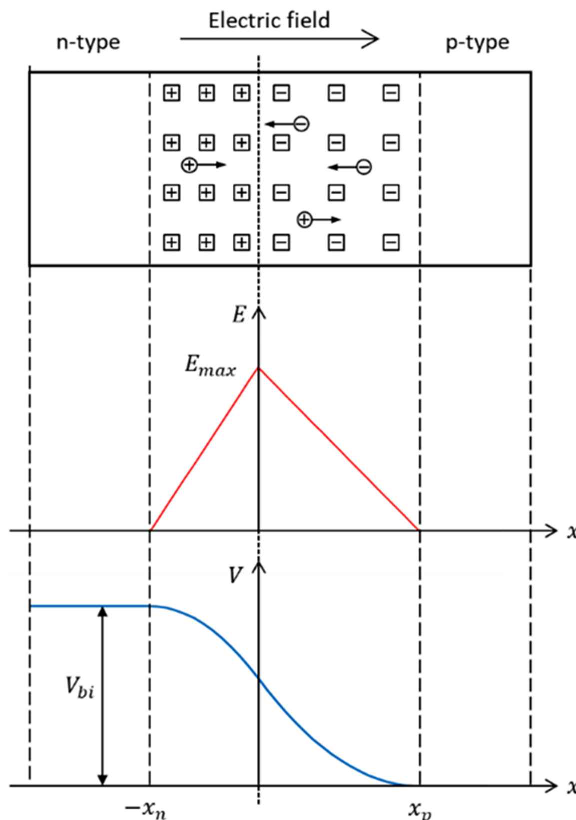


**Figure 1.13.** Structure, electric field and potential in a *pn* junction in equilibrium.

It is this electric field that makes the *pn* junction so useful in image sensors. Electrons and holes generated by the photoeffect get separated upon entering the depletion region and create current, which can be measured. Also, only a negligible current will flow in the absence of light due to the depletion behaving as an insulator. This is what was described in the previous sections when we discussed the idealised charge collection.

In equilibrium the whole of the *pn* junction must remain electrically neutral regardless of the fixed space charge. If we consider an *abrupt junction*, where the doping concentrations $N_D$ (in the *n*-side) and $N_A$ (in the *p*-side) are uniform and change in a stepwise fashion at the junction at $x = 0$, the condition of electrical neutrality can be written as

$$N_D x_n = N_A x_p \tag{1.32}$$

Equation (1.32) mathematically means that in the space charge region the number of donor and acceptor atoms per unit area is equal. Despite its idealistic appearance, the abrupt junction is a very good approximation to real *pn* junctions found in image sensors.

From the condition of thermal equilibrium, it follows that the Fermi level through the device is constant. Therefore, the valence and the conduction bands must bend so that the Fermi level stays flat, creating the built-in potential $V_{bi}$ in figure 1.13. Because the electric field is the gradient of the potential distribution, this is just another way to describe the presence of electric field in the space charge region. The built-in voltage can be calculated from the uniformity of the Fermi level and is [2]

$$V_{bi} = \frac{kT}{q} \ln\left(\frac{N_A N_D}{n_i^2}\right) \tag{1.33}$$

where $n_i$ is the intrinsic carrier concentration.

Integrating the Poisson equation for both sides [2], we can calculate the electric field on the *n*-side ($-x_n \leqslant x \leqslant 0$)

$$E(x) = \frac{qN_D}{\varepsilon_0 \varepsilon_{Si}}(x + x_n) \tag{1.34}$$

and on the *p*-side ($0 \leqslant x \leqslant x_p$)

$$E(x) = \frac{qN_A}{\varepsilon_0 \varepsilon_{Si}}(x_p - x) \tag{1.35}$$

Here $\varepsilon_0$ is the dielectric permittivity of vacuum and $\varepsilon_{Si}$ is the dielectric constant of silicon. The electric field has a linear dependence on distance because of the constant doping concentrations in the abrupt junction approximation. We can see in figure 1.13 that the electric field has a characteristic triangular shape and decreases linearly from its maximum $E_{max}$ at $x = 0$, where the formulas (1.34) and (1.35) join up.

$$E_{\max} = \frac{qN_A x_p}{\varepsilon_0 \varepsilon_{Si}} = \frac{qN_D x_n}{\varepsilon_0 \varepsilon_{Si}} \tag{1.36}$$

The field becomes zero in the neutral silicon at $x \leqslant -x_n$ and $x \geqslant x_p$.

Furthermore, we can calculate the potential distribution in the junction by integrating (1.34) and (1.35), using the boundary conditions $V_n(-x_n) = V_{bi}$ and $V_p(x_p) = 0$. This gives the expected quadratic dependence on distance, since we are integrating a linearly changing electric field:

$$V_n(x) = V_{bi} - \frac{qN_D}{2\varepsilon_0 \varepsilon_{Si}}(x + x_n)^2 \tag{1.37}$$

$$V_p(x) = \frac{qN_A}{2\varepsilon_0 \varepsilon_{Si}}(x - x_p)^2 \tag{1.38}$$

Here we must clarify an important point about the built-in potential—it does not appear across the terminals of the junction. Anybody who has measured the voltage across a *pn* diode with a voltmeter will testify that the voltage is zero, unless the diode is illuminated. Obviously the *pn* junction is not a battery! The reason we do not see the built-in voltage is the contact potential between the silicon and the electrodes used to connect it to the outside world. The contact potentials between the metal electrodes on the *n* and the *p*-side precisely cancel the built-in voltage, and there is no potential difference between the two external electrodes. If that were not the case, a shorted diode would generate a continuous current through itself, which obviously does not happen.

The total width of the depleted region $W = x_n + x_p$ can be found from (1.37) and (1.38) by using that $V_n(0) = V_p(0)$, which gives

$$V_{bi} - \frac{qN_D}{2\varepsilon_0 \varepsilon_{Si}} x_n^2 = \frac{qN_A}{2\varepsilon_0 \varepsilon_{Si}} x_p^2 \tag{1.39}$$

From here, using (1.32) we can write

$$V_{bi} = \frac{q}{2\varepsilon_0 \varepsilon_{Si}} \frac{(N_A + N_D)N_D}{N_A} {x_n}^2 = \frac{q}{2\varepsilon_0 \varepsilon_{Si}} \frac{(N_A + N_D)N_A}{N_D} {x_p}^2 \tag{1.40}$$

and finally, we get

$$x_n = \sqrt{\frac{2\varepsilon_0 \varepsilon_{Si}}{q} \frac{N_A V_{bi}}{(N_A + N_D)N_D}} \tag{1.41}$$

$$x_p = \sqrt{\frac{2\varepsilon_0 \varepsilon_{Si}}{q} \frac{N_D V_{bi}}{(N_A + N_D)N_A}} \tag{1.42}$$

The depletion width $W$ is the sum of (1.41) and (1.42)

$$W = \sqrt{\frac{2\varepsilon_0\varepsilon_{Si}}{q}\left(\frac{N_A + N_D}{N_A N_D}\right)V_{bi}} \tag{1.43}$$

Most of the *pn* junctions used in image sensors are one-sided (also called asymmetric), i.e. one of the dopants has much higher concentration than the other. Typical examples are $n^+p$ junctions used as photodiodes and for charge conversion.

In a one-sided $n^+p$ junction $N_D \gg N_A$, which leads to almost all of the depletion being on the *p*-side because $x_p \gg x_n$. From (1.42) and (1.43) we see that $x_p \cong W$ and the expression for the depletion width simplifies to

$$W = \sqrt{\frac{2\varepsilon_0\varepsilon_{Si}}{qN_A}V_{bi}} \tag{1.44}$$

Why are we using one-sided junctions? If one dopant has significantly higher concentration than the other, for example by a factor of 100, the depletion width is determined entirely by the low-doped side, and only its concentration has to be precisely controlled. The built-in voltage depends on both dopant concentrations, but thanks to the logarithmic dependence in (1.33) it has much weaker effect on the depletion width.

### 1.6.2 *pn* junction under reverse bias

The *pn* junction is very useful even without any voltage applied to it. Whether its terminals are shorted or left floating, the internal electric field is still there and can separate photogenerated electron–hole pairs. However, the depletion width in equilibrium can be quite small, unless very low doping concentrations are used. Increasing the depletion is often desired as it reduces the size of the field-free regions and the extent of charge diffusion. This can happen when a reverse bias is applied across the junction with the same polarity as the built-in voltage, i.e. positive on the *n*-side (cathode) relative to the *p*-side (anode).

When a reverse bias $V_r$ is applied, the total voltage across the junction is $V_{bi} + V_r$ and the depletion width in a $n^+p$ junction becomes

$$W = \sqrt{\frac{2\varepsilon_0\varepsilon_{Si}}{qN_A}(V_{bi} + V_r)} \tag{1.45}$$

Due to the asymmetric doping, the depletion width is almost entirely contained in the *p*-side of the junction, and so is the electric field.

---

**Example 1.10.** Calculate the depletion depth of a silicon one-sided $n^+p$ junction with $N_A = 6.7 \times 10^{14}$ cm$^{-3}$ (resistivity $\rho = 20$ Ωcm) and $N_D = 10^{18}$ cm$^{-3}$ for $V_r = 0$ and $V_r = 5$ V and 300 K. Use that $n_i = 1.45 \times 10^{10}$ cm$^{-3}$, $\varepsilon_0 = 8.85 \times 10^{-14}$ F cm$^{-1}$ and $\varepsilon_{Si} = 11.9$. Also calculate $E_{max}$ and $x_n$ at $V_r = 5$ V.
**Solution:** First, from (1.33) we calculate $V_{bi}$:

$$V_{\rm bi} = \frac{kT}{q} \ln\left(\frac{N_{\rm A}N_{\rm D}}{n_i^2}\right) = \frac{1.38 \times 10^{-23} \times 300}{1.6 \times 10^{-19}} \ln\left(\frac{6.7 \times 10^{14} \times 10^{18}}{2.1 \times 10^{20}}\right) = 0.74 \text{ V}$$

Next, from (1.45) we calculate the depletion widths

$$W(0{\rm V}) = \sqrt{\frac{2 \times 8.85 \times 10^{-14} \times 11.9 \times 0.74}{1.6 \times 10^{-19} \times 6.7 \times 10^{14}}} = 1.21\ \mu\text{m}$$

$$W(5{\rm V}) = \sqrt{\frac{2 \times 8.85 \times 10^{-14} \times 11.9 \times (0.74 + 5)}{1.6 \times 10^{-19} \times 6.7 \times 10^{14}}} = 3.36\ \mu\text{m}$$

The maximum electric field and $x_n$ can be calculated from (1.36) using that $x_p \cong W$:

$$E_{\max} = \frac{qN_{\rm A}x_p}{\varepsilon_0\varepsilon_{\rm Si}} = \frac{qN_{\rm A}W}{\varepsilon_0\varepsilon_{\rm Si}} = \frac{1.6 \times 10^{-19} \times 6.7 \times 10^{14} \times 3.36 \times 10^{-4}}{8.85 \times 10^{-14} \times 11.9} = 34\,200 \text{ Vcm}^{-1}$$

$$x_n = \frac{\varepsilon_0\varepsilon_{\rm Si}E_{\max}}{qN_{\rm D}} = \frac{8.85 \times 10^{-14} \times 11.9 \times 34\,200}{1.6 \times 10^{-19} \times 10^{18}} = 2.3 \text{ nm}$$

---

This example shows how much smaller the depletion is on the *n*-side in a one-sided $n^+p$ junction compared to the *p*-side; in this case $x_n$ is entirely negligible. The characteristic right-triangular shape of the electric field and its drop to zero at the edge of the depletion are important features with implications to charge collection.

Figure 1.14 shows the calculated potential and electric field from example 1.10 using (1.34)–(1.38), and a simulation using commercial TCAD software [27].

It is useful to compare the analytic formulae with a finite element device simulation using the same parameters; this provides a necessary cross-check even if both cannot be made exactly the same. The matching in figure 1.14 is good, considering the approximations in the analytical calculation and the discrete structure of the simulation model.

The approximation of ideal abrupt *pn* junction assumes that the space charge region and the majority carrier concentration have infinitely sharp edges. This is of course an idealisation and is the reason why the electrical field falls to zero at both ends of the depletion. In practice infinitely sharp edges do not exist and the majority carrier concentration changes smoothly.

The finite element analysis (FEA) TCAD simulation solves the Poisson equation without this assumption and reveals the fine details at the edge of the depletion. Plotted on a logarithmic scale, the electric field in figure 1.15 continues to be above zero for about a micron beyond the calculated edge of the space charge region. This field is small but can make a sizeable effect on the charge collection time due to drift being much faster than diffusion, as discussed in section 1.3.3.

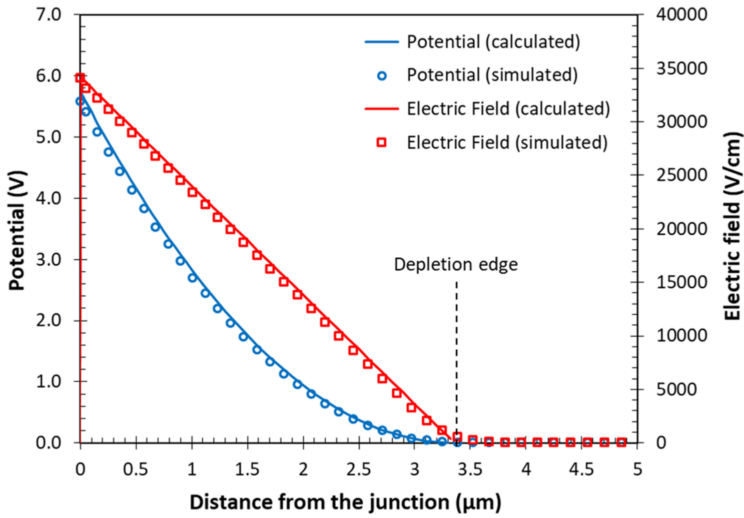


**Figure 1.14.** Potential and electric field calculated and simulated in example 1.10 for $V_r = 5$ V.

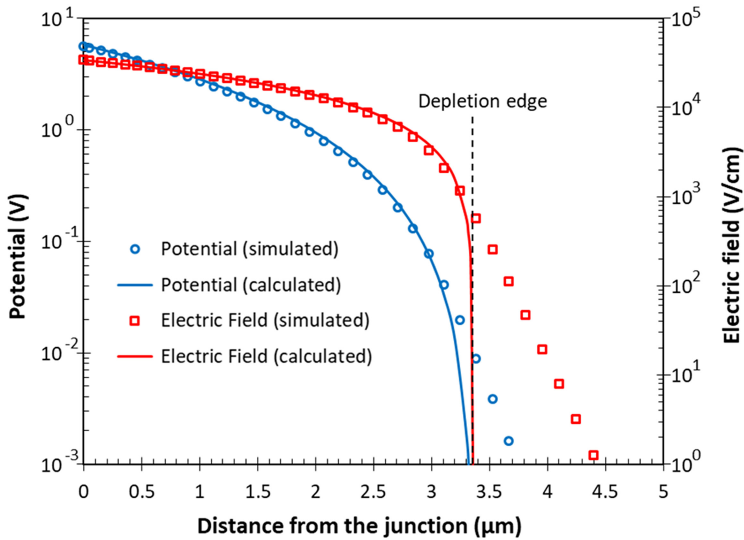


**Figure 1.15.** A comparison between the electric field calculated using the abrupt $n^+p$ junction approximation in example 1.10 for $V_r = 5$ V, and an FEA device simulation which does not make this approximation.

The calculations in this section used the abrupt one-sided *pn* junction model, which is a good approximation to most practical devices. This gives the familiar square root dependence (1.45) of the depletion width on the reverse bias. A different doping profile would produce a different voltage dependence, with the linearly graded junction [2] the most notable example, giving $W \propto (V_{bi} + V_r)^{1/3}$.

### 1.6.3 Charge collection

Due to its electric field, the *pn* junction collects electrons on the cathode, biased or not. Looking back at figure 1.9, showing our then hypothetical charge collection element, this is exactly what the *pn* junction does.

Without a reverse bias, the depletion is usually much smaller than the depth of the device. In this case some of the charge will diffuse before collection, as shown in figure 1.16(a). To minimise charge spread the field-free region must be eliminated, so that *full depletion* is achieved. This is shown in figure 1.16(b), and the mathematical condition is that the depletion width becomes equal to the device thickness. Applying that $W = d$ to (1.45) we can calculate the reverse bias $V_{\mathrm{fd}}$ at which full depletion is achieved:

$$d = \sqrt{\frac{2\varepsilon_0\varepsilon_{\mathrm{Si}}}{qN_{\mathrm{A}}}(V_{\mathrm{bi}} + V_{\mathrm{fd}})} \tag{1.46}$$

which gives

$$V_{\mathrm{fd}} = \frac{qN_{\mathrm{A}}d^2}{2\varepsilon_0\varepsilon_{\mathrm{Si}}} - V_{\mathrm{bi}} \tag{1.47}$$

Equation (1.47) tells us that the full depletion voltage is proportional to the dopant concentration, and this is why it is easier to reach in high resistivity, low-doped semiconductors.

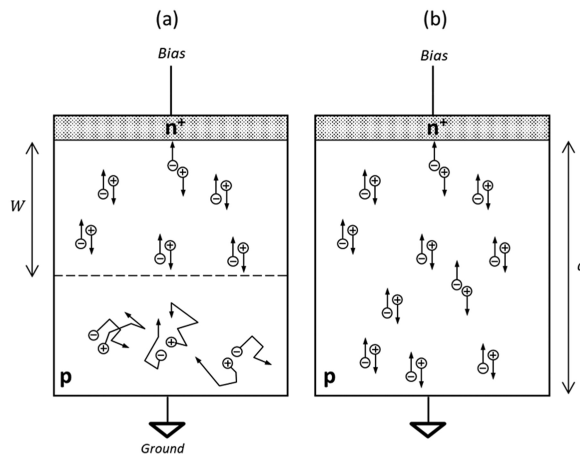


**Figure 1.16.** Charge collection in partially depleted (a); and in a fully depleted *pn* junction (b).

---

**Example 1.11.** Calculate the full depletion voltage of the $n^+p$ junction in example 1.10 for a device thickness $d = 5$ μm.

**Solution:** Using (1.47) and the previously calculated $V_{\mathrm{bi}}$

$$V_{\text{fd}} = \frac{1.6 \times 10^{-19} \times 6.7 \times 10^{14} \times (5 \times 10^{-4})^2}{2 \times 8.85 \times 10^{-14} \times 11.9} - 0.74 = 11.98 \text{ V}$$

For very thick devices the full depletion voltage is large and we can ignore $V_{\text{bi}}$ because $V_{\text{fd}} \gg V_{\text{bi}}$.

We can now refine the treatment of the charge collection time in (1.24) for an abrupt one-sided *pn* junction. Figure 1.17 shows the familiar triangular shape of the electric field, which in full depletion is described by (1.35) with $x_p = W = d$.

$$E(x) = \frac{qN_{\text{A}}}{\varepsilon_0 \varepsilon_{\text{Si}}}(d - x) \tag{1.48}$$

To avoid the electric field dropping to near zero at the back of the device, the reverse bias can be increased above that required for full depletion so that the junction becomes *over-depleted.* This creates an additional electric field $E_{\text{od}} = V_{\text{od}}/d$, where $V_{\text{od}}$ is the voltage top-up above $V_{\text{fd}}$. Adding $E_{\text{od}}$ to $E(x)$, and using (1.47) with $V_{\text{fd}} \gg V_{\text{bi}}$ the electric field becomes

$$E(x) = \frac{2V_{\text{fd}}}{d^2}(d - x) + \frac{V_{\text{od}}}{d} \tag{1.49}$$

Due to the linear dependence of the electric field the electron velocity $v_n$ will change linearly too, depending on the position in the device according to $v_n(x) = \mu_n E(x)$. The travel time can be obtained by solving equation (1.50) as in [28]

$$v_n(x) \equiv \frac{dx}{dt} = -\mu_n \left[ \frac{2V_{\text{fd}}}{d^2}(d - x) + \frac{V_{\text{od}}}{d} \right] \tag{1.50}$$

Here the minus sign takes into account that the electrons move opposite to the direction of the $x$-axis, and we ignore velocity saturation. The solution for charge generated at coordinate $x_0 < d$ is:

$$t = \frac{d^2}{2\mu_n V_{\text{fd}}} \ln \left[ \frac{1}{1 - \left( \frac{2V_{\text{fd}}}{2V_{\text{fd}} + V_{\text{od}}} \right) \frac{x_0}{d}} \right] \tag{1.51}$$

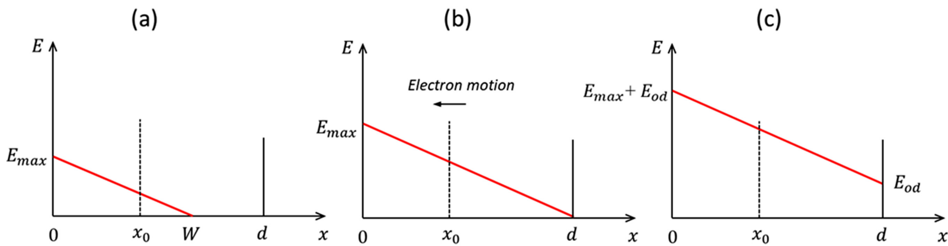


**Figure 1.17.** Electric field in a partially depleted (a), fully depleted (b) and over-depleted (c) abrupt *pn* junction.

If $V_{od} = 0$ (1.51) simplifies to

$$t = \frac{d^2}{2\mu_n V_{fd}} \ln\left(\frac{d}{d - x_0}\right) \tag{1.52}$$

and shows that when charge is generated near the back of the device ($x_0 \approx d$) the charge collection time tends to infinity. This is because we are considering only charge drift and the electric field at the back of the device for $V_{od} = 0$ is zero, therefore an electron will stay there forever and will not be collected. Fortunately, diffusion is ever-present and takes care of this 'unphysical' situation—the charge will diffuse to regions with non-zero field will be quickly swept away.

The maximum charge collection time $t_{max}$ for over-depleted junction ($V_{od} > 0$) can be obtained from (1.51) for $x_0 = d$

$$t_{max} = \frac{d^2}{2\mu_n V_{fd}} \ln\left(\frac{2V_{fd} + V_{od}}{V_{od}}\right) \tag{1.53}$$

Comparing (1.53) to (1.24), the differences are a factor of two in the denominator and the logarithmic term, which can become significant for low over-depletion voltages. The two formulas give the same result when $V_{od} = 2V_{fd}/(e^2 - 1) = 0.31\,V_{fd}$.

---

**Example 1.12.** Calculate the maximum charge collection time for the *pn* junction in example 1.11 for $V_{od} = 1$ V, ignoring $V_{bi}$.
**Solution:** Using formula (1.53) we get

$$t_{max} = \frac{(5 \times 10^{-4})^2}{2 \times 1400 \times 11.98} \ln\left(\frac{2 \times 11.98 + 1}{1}\right) = 24 \text{ ps}$$

However, this is wrong because we have ignored velocity saturation in the derivation of the formula. A quick check using (1.49) for $x = 0$ gives $v_n = 6.7 \times 10^7$ cm s$^{-1}$, well above the saturation velocity. Therefore, this calculation gives too short charge collection time; a better estimation as in example 1.7 would give around 50 ps.

---

You may ask the question: do we need a specialised semiconductor structure (e.g. *pn* junction or a MOS capacitor) to generate an electric field? Can we just apply some voltage across silicon to collect the charge? Let's investigate this, picking a square pixel with size $a = 10$ μm in silicon substrate $d = 5$ μm thick. Applying voltage $V$ across the pixel will force current to flow because silicon has finite resistance. This current will look like photogenerated signal and must be reduced as much as possible; therefore, we have to choose intrinsic (i.e. undoped and pure) silicon, which has the highest resistivity $\rho = 230$ kΩcm at room temperature. If we now apply one volt across the pixel, the current can be calculated from the Ohm's law:

$$I = \frac{V}{R} = \frac{V}{\rho d / a^2} = \frac{1}{230 \times 10^3 \times 5 \times 10^{-4} / (10 \times 10^{-4})^2} = 8.7 \text{ nA}$$

Comparing with examples 1.3 and 1.4 we see that this current is many orders of magnitude higher than the photogenerated current; obviously a pixel made like this will be a very poor image sensor. It may have a chance in very bright illumination conditions, or when the sensor is cooled down (to increase the resistivity and reduce the current), but not as a normal image sensor we are all used to.

### 1.6.4 Junction capacitance

Besides for charge collection, the *pn* junction can be used for charge-to-voltage conversion as in the diagrams shown in figures 1.12(a)–(c). Two of the circuits rely on the conversion of the photocurrent on an external resistor (figure 1.12(a)), or on an external capacitor (figure 1.12(c)). The diagram in figure 1.12(b) shows the most popular use, where the junction capacitance itself is used for the conversion.

As mentioned before, the depletion region is an insulator separating the conducting *p* and *n*-type field-free regions. This is exactly the situation in a parallel plate capacitor with a distance between the electrodes $W$ and capacitance $C = \varepsilon_0\varepsilon_{\mathrm{Si}}A/W$, where $A$ is the electrode area. Using (1.45) the capacitance of the *pn* junction is

$$C = \frac{\varepsilon_0\varepsilon_{\mathrm{Si}}A}{W} = A\sqrt{\frac{\varepsilon_0\varepsilon_{\mathrm{Si}}qN_{\mathrm{A}}}{2(V_{\mathrm{bi}} + V_{\mathrm{r}})}} \tag{1.54}$$

---

**Example 1.13.** Calculate the capacitance of the *pn* junction with an area $A = 25\ \mu\mathrm{m}^2$ with the parameters given in example 1.10 for $V_{\mathrm{r}} = 1$ V.
**Solution:** From (1.54)

$$C = 25 \times 10^{-8} \times \sqrt{\frac{8.85 \times 10^{-14} \times 11.9 \times 1.6 \times 10^{-19} \times 6.7 \times 10^{14}}{2 \times (0.74 + 1)}} = 1.42\ \mathrm{fF}$$

---

After the junction has been biased to $V_{\mathrm{r}}$ and left floating by disconnecting the switch, photogenerated electrons will collect at the cathode and reduce its potential according to formula (1.30). However, because the junction capacitance depends on the applied voltage, it is not an ideal parallel plate capacitor. As electrons are collected, the voltage on the junction goes down, and the capacitance goes up. The change of the capacitance is nonlinear, and its dependence on the reverse bias can be found by differentiating (1.54) by $V_{\mathrm{r}}$:

$$\frac{dC}{dV_{\mathrm{r}}} = -\frac{A}{2}\sqrt{\frac{\varepsilon_0\varepsilon_{\mathrm{Si}}qN_{\mathrm{A}}}{2(V_{\mathrm{bi}} + V_{\mathrm{r}})^3}} = -\frac{C}{2(V_{\mathrm{bi}} + V_{\mathrm{r}})} \tag{1.55}$$

The nonlinear capacitance means that the conversion from charge to voltage will be nonlinear too, which is not what we normally want from an image sensor. As figure 1.18 shows, the change of capacitance can exceed 20% when the voltage across the junction changes by 1 V. This change is large but can be counteracted by connecting the junction in parallel with a larger, linear capacitance, so that the nonlinearity is much reduced. In image sensors the role of this additional capacitance is performed by the readout circuitry, as well as by actual capacitors.

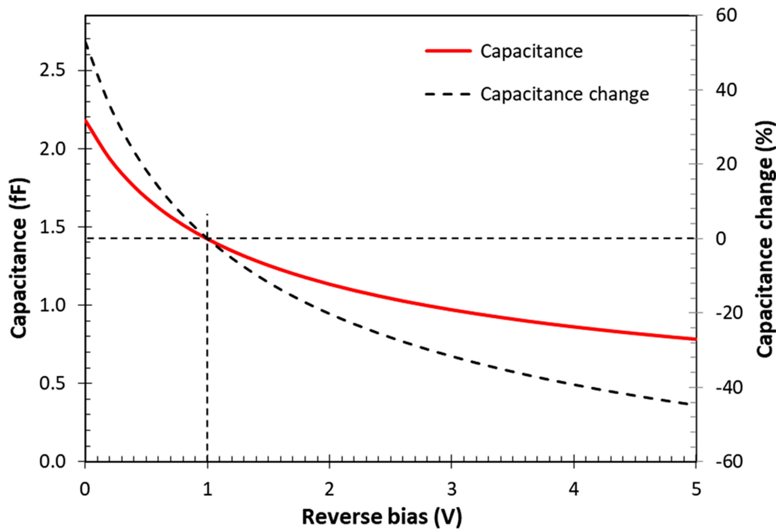


**Figure 1.18.** Junction capacitance and the change of the capacitance relative to $V_r = 1$ V for the *pn* junction in example 1.13.

## 1.7 MOS capacitor

### 1.7.1 Depletion

The MOS capacitor is a structure that combines a metal electrode (usually called a 'gate') deposited on top of an insulator (typically $SiO_2$), which is grown on silicon, as shown in figure 1.19. Normally the gate is not made of metal but of heavily doped, highly conductive polycrystalline silicon because this greatly improves the quality of the Si–$SiO_2$ interface and device yield. The gate was made of metal in the early days of semiconductor technology and hence the term 'MOS' was coined back then, but it remains in use to this day.

The MOS capacitor is the basic building block of the CCD and also of certain type of CIS using photogates instead of a photodiode. Photogate-based CIS find applications where fast or multiple charge transfer is needed, such as in some time-of-flight (ToF) sensors.

Similarly to the *pn* junction, the MOS capacitor can be used to create an electric field and depletion region suitable for charge collection. We will consider the example in figure 1.19 with a $n^+$-doped polysilicon gate and a *p*-type substrate. Intuitively, a positive gate potential with respect to the substrate should force the holes away from the Si–$SiO_2$ interface and create a depletion region with depth $W$. In the band diagram in figure 1.19 this is shown as downward bending of the conduction and valence bands. Downward bending indicates increased potential near the interface relative to the neutral bulk, which is connected to ground.

In an electrically neutral semiconductor, the difference between the Fermi level $E_F$ and the mid-band (intrinsic) Fermi level $E_i$ is the same everywhere, including at the

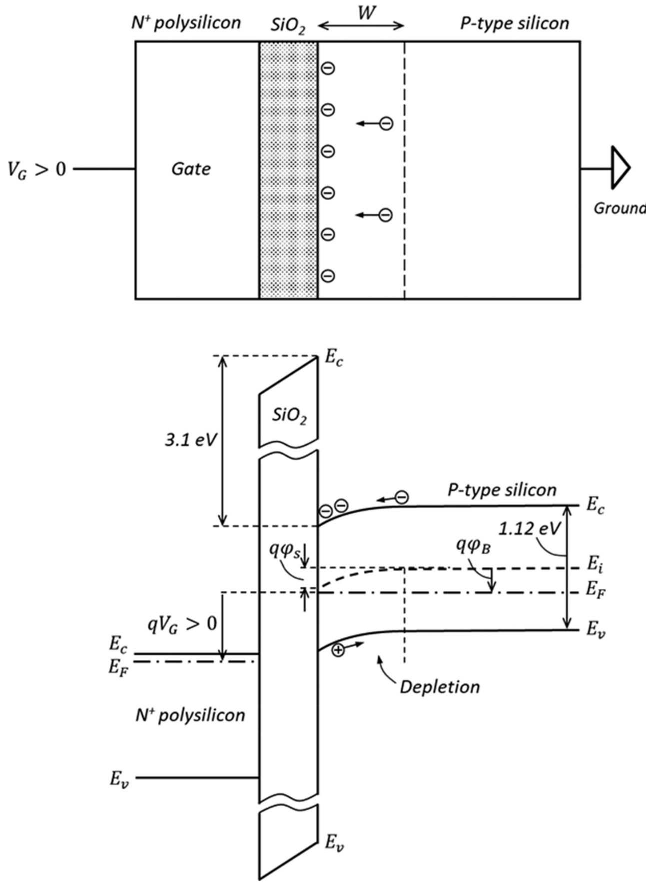


**Figure 1.19.** Energy-band diagram of a MOS capacitor in depletion.

Si–$SiO_2$ interface. The difference $E_i - E_F$ in the bulk of a neutral $p$-type semiconductor can be expressed as:

$$q\varphi_B = E_i - E_F = kT\ln\left(\frac{N_A}{n_i}\right) \tag{1.56}$$

When a gate voltage $V_G$ is applied so that the bands bend downwards, the surface potential $\varphi_s$ increases, and the surface hole concentration decreases according to

$$p = N_A \exp\left(-\frac{q\varphi_s}{kT}\right) \tag{1.57}$$

When $\varphi_B > \varphi_s > 0$ the semiconductor is depleted because the hole concentration at the surface is lower than in the bulk. When $\varphi_s = \varphi_B$ the surface hole concentration becomes the intrinsic concentration $n_i$, as can be verified by using equations (1.56) and (1.57). If the gate voltage is increased even further, so that $\varphi_s > \varphi_B$, the surface becomes inverted because there are more electrons than holes. Strong inversion

happens when $\varphi_s > 2\varphi_B$ and is the condition used in enhancement mode MOS transistors.

The depletion region in figure 1.19 looks very similar to the *p*-side of an abrupt *pn* junction (figure 1.14) if we take the voltage at the junction as $\varphi_s$. Therefore, the depletion depth can be calculated with equation (1.45)

$$W = \sqrt{\frac{2\varepsilon_0\varepsilon_{Si}\varphi_s}{qN_A}} \tag{1.58}$$

and from it, the surface potential is

$$\varphi_s = \frac{qN_A W^2}{2\varepsilon_0\varepsilon_{Si}} \tag{1.59}$$

The potential in the depletion region away from the surface (which is at $x = 0$) is quadratic as in equation (1.38) describing the *pn* junction:

$$\varphi(x) = \varphi_s\left(1 - \frac{x}{W}\right)^2 \tag{1.60}$$

Since the substrate is at ground, we can write that the gate voltage is the sum of the voltage across the oxide $V_{ox}$ and the surface potential, ignoring for the moment the flat-band voltage offset (described in the next section):

$$V_G = V_{ox} + \varphi_s \tag{1.61}$$

The oxide voltage is equal to the space charge in the depleted region divided by the oxide capacitance per unit area $C_{ox}$:

$$V_{ox} = \frac{qN_A W}{C_{ox}} \tag{1.62}$$

Substituting (1.59) and (1.62) into (1.61), we get the equation (1.63)

$$V_G = \frac{qN_A W}{C_{ox}} + \frac{qN_A W^2}{2\varepsilon_0\varepsilon_{Si}} \tag{1.63}$$

which can be solved to give the depletion depth as

$$W = -\frac{\varepsilon_0\varepsilon_{Si}}{C_{ox}} + \sqrt{\left(\frac{\varepsilon_0\varepsilon_{Si}}{C_{ox}}\right)^2 + \frac{2\varepsilon_0\varepsilon_{Si}}{qN_A}V_G} \tag{1.64}$$

We see that the depletion depth changes as the square root of the gate voltage, similarly to the reversed-biased *pn* junction.

When $\varphi_s > 2\varphi_B$ and in the absence of a ready source of electrons, as in figure 1.19, inversion takes some time to take hold because the electrons needed to populate the interface have to be generated thermally ([26], chapter 5). This allows the surface to be taken far beyond $\varphi_s > 2\varphi_B$ into *deep depletion* (figure 1.20) without surface inversion, so that photogenerated electrons can be collected. This condition is what

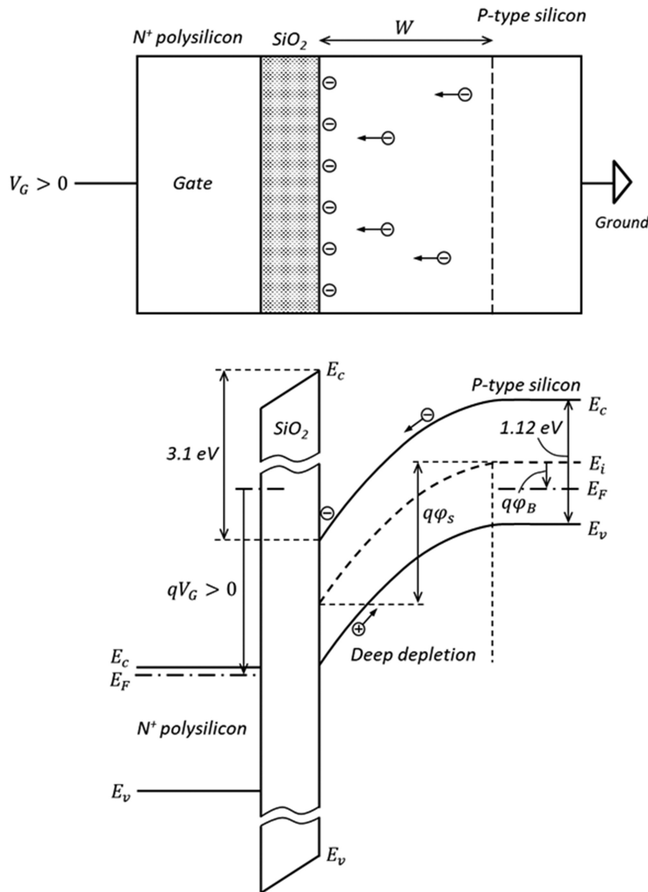


**Figure 1.20.** Energy-band diagram of a MOS capacitor in deep depletion.

makes image sensors using photogates work. If inversion were to be established immediately as in MOSFETs, the photogenerated charge would be swamped by the many more electrically induced carriers.

Once collected, the signal electrons have nowhere else to go because they cannot cross over the oxide, therefore signal detection via current is not an option. However, the charge can be measured electrostatically (non-destructively) as in figure 1.12(d), or it can be transferred to a separate charge conversion element. In CCDs, for example, the charge is transported over large distances to a reverse-biased *pn* junction for detection.

Regardless of how the signal is measured, it must be cleared from the MOS capacitor before the next signal is collected. Unlike a *pn* junction the photogenerated charge cannot be drained away simply by re-connecting the bias since the MOS capacitor has no conductive path. The possibilities for clearing the charge are to either: (a) transfer it in a controlled manner; (b) drain it by using an additional

structure next to the MOS capacitor, or (c) to recombine the signal electrons by flooding the interface with holes.

### 1.7.2 Gate capacitance

Similarly to the *pn* junction, the depletion depth in the MOS capacitor is nonlinear due to the square root in (1.64). This makes the MOS capacitance voltage dependent too. The difference is that the total gate capacitance is determined by the oxide and the depletion capacitances connected in series, as illustrated in figure 1.21:

$$\frac{1}{C_\mathrm{G}} = \frac{1}{C_\mathrm{ox}} + \frac{1}{C_\mathrm{dep}} \tag{1.65}$$

The oxide capacitance per unit area is given by

$$C_\mathrm{ox} = \frac{\varepsilon_0 \varepsilon_\mathrm{ox}}{t_\mathrm{ox}} \tag{1.66}$$

where $\varepsilon_\mathrm{ox}$ is the dielectric permittivity of $SiO_2$ and $t_\mathrm{ox}$ is the gate oxide thickness. The depletion capacitance per unit area can be determined by using the formula for the parallel plate capacitor and (1.64):

$$C_\mathrm{dep} = \frac{\varepsilon_0 \varepsilon_\mathrm{Si}}{W} = \frac{1}{-\frac{1}{C_\mathrm{ox}} + \sqrt{\frac{1}{C_\mathrm{ox}^2} + \frac{2V_\mathrm{G}}{\varepsilon_0 \varepsilon_\mathrm{Si} q N_\mathrm{A}}}} \tag{1.67}$$

Due to the thin gate oxide $C_\mathrm{ox}$ is usually very large, and in deep depletion $V_\mathrm{G}$ is large too, so that $C_\mathrm{ox} \gg C_\mathrm{dep}$ and $1/C_\mathrm{ox}$ becomes negligible in (1.67). In this condition the gate capacitance can be approximated with

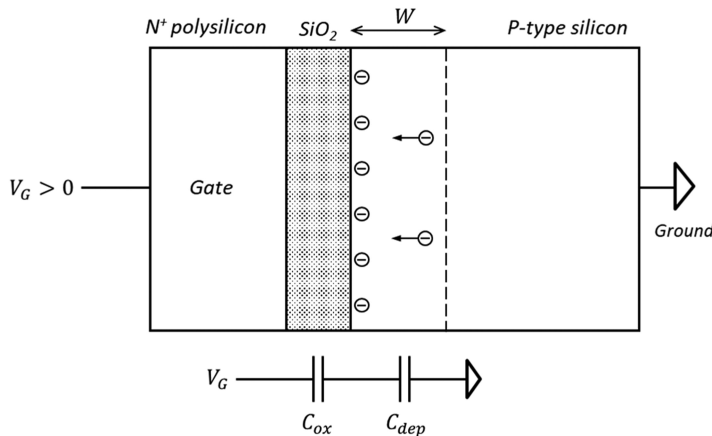


**Figure 1.21.** MOS gate capacitance.

$$C_G \approx \sqrt{\frac{\varepsilon_0 \varepsilon_{Si} q N_A}{2V_G}} \tag{1.68}$$

which is almost identical to the capacitance of the *pn* junction in (1.54).

---

**Example 1.14.** Calculate the oxide capacitance per square centimetre and square micron of 7 nm thick $SiO_2$, using that $\varepsilon_{ox} = 3.9$. Also calculate the gate capacitance in deep depletion for $N_A = 6.7 \times 10^{14}$ cm$^{-3}$ (resistivity $\rho = 20$ Ωcm) and $V_G = 3.0$ V.
**Solution:** From (1.66)

$$C_{ox} = \frac{3.9 \times 8.85 \times 10^{-14}}{7 \times 10^{-7}} = 4.93 \times 10^{-7}\ \mathrm{F\ cm^{-2}} = 4.93\ \mathrm{fF}\ \mu\mathrm{m}^{-2}$$

Next, using (1.68)

$$C_G = \sqrt{\frac{8.85 \times 10^{-14} \times 11.9 \times 1.6 \times 10^{-19} \times 6.7 \times 10^{14}}{2 \times 3.0}} = 4.34 \times 10^{-9}\ \mathrm{F\ cm^{-2}}$$
$$= 0.04\ \mathrm{fF}\ \mu\mathrm{m}^{-2}$$

The gate capacitance is about two orders of magnitude smaller than the oxide capacitance, which should not come as a surprise because the depletion depth in silicon is 2.4 μm (as can be verified from (1.64)), compared to just 0.007 μm oxide thickness. In this case, ignoring $C_{ox}$ in the approximate formula (1.68) is fully justified.

---

The gate capacitance in deep depletion is dominated by the depletion depth in silicon because the oxide capacitance is normally much larger. However, if the depletion is to shrink to zero, the silicon under the oxide becomes conductive and acts as an electrode to the gate oxide. This makes the gate capacitance equal to the very large $C_{ox}$, which is the upper limit. The way to reach this condition is to operate the MOS capacitor in *accumulation* by biasing the gate sufficiently negative relative to substrate, so that holes gather at the Si–$SiO_2$ interface. Another way to increase the capacitance to $C_{ox}$ is to invert the surface and populate it with electrons, which act as the second electrode. This is used in capacitors based on MOSFETs and is described in the following section.

## 1.8 MOS transistor

### 1.8.1 Structure

MOS field effect transistors (MOSFETs) are the staple of microelectronics and are the fundamental building block of nearly every integrated circuit. In image sensors MOSFETs are used as buffers for signals with high output impedance, such as the photogenerated voltage in a photodiode, and as amplifiers, current sources and sinks, active loads, switches and capacitors.

MOSFETs work using the *field effect*, describing the strong change of conductivity in a semiconductor (or another material such as graphene) under the influence of an electric field. In *p*-type semiconductor the electric field induced by the gate–substrate voltage can create a thin inversion layer of electrons at the Si–$SiO_2$ interface. Devices operating like this are called surface channel transistors, shown in figure 1.22(a), and comprise the vast majority of MOSFETs. They are also called enhancement mode, or 'normally off' transistors, because at zero voltage on the gate the transistor does not conduct.

Another type of MOSFET, called buried channel, or depletion mode transistor, is shown in figure 1.22(b). In the buried channel *n*-MOSFET a *n*-type dopant is implanted in the channel, which becomes conducting so that the transistor is normally on without a voltage applied to the gate. To turn it off, a *negative gate voltage* with respect to the source must be applied. This depletes the channel of electrons until the conduction is cut off.

Described in simple terms, when the gate–source voltage $V_{GS}$ is above the threshold voltage $V_T$ the transistor is conducting current, and when below threshold the transistor is off, representing an infinite resistance. For MOSFETs in digital circuits this description is very accurate—the transistors are either 'fully on' or 'fully off', with no intermediate state. Analogue circuits use precise ways to smoothly control the drain current so that the MOSFETs work as amplifiers of small signals, or as current sources and sinks.

The source and the drain of the transistors shown in figure 1.22 are symmetrical and interchangeable. The source and the drain are *pn* junctions in their own right, and so is the whole of the buried channel in figure 1.22(b). They are always reverse biased with respect to the substrate, and in common with any *pn* junction can be photosensitive.

MOSFETs are essential for image sensor operation—the main reason is their extremely high input impedance. As we saw already, the currents involved in imaging can be very low; as an example, the dark current though a pixel could be below 1 electron per second at room temperature—this is $1.6 \times 10^{-19}$ A (or 160 zepto amps), a phenomenally low current. Only a good insulator such as the $SiO_2$ used as a gate oxide in MOSFETs has a chance of low enough leakage that is not disturbing this tiny current. For comparison, a very good discrete *pn* junction (e.g. a low

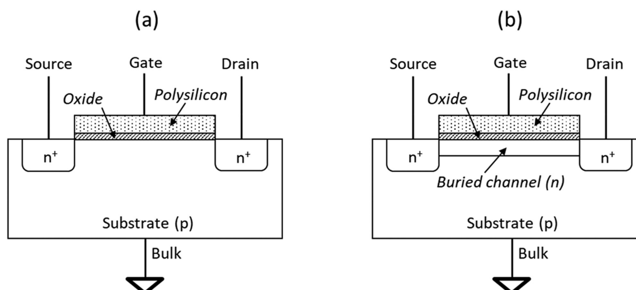


**Figure 1.22.** Surface (a) and buried channel (b) *n*-channel MOSFETs.

leakage diode or the gate of a JFET) has reverse current of around 1 pA at room temperature.

Both transistors in figure 1.22 can sit in *p*-wells in order to have good control over certain MOSFET parameters such as the threshold. This also allows the substrate to have different, usually lower doping concentration, which is needed to achieve the desired depletion depth and charge collection.

### 1.8.2 MOSFET characteristics

There are plenty of excellent books describing how MOS transistors work [26, 29]. Here we are just going to give a short summary with the most relevance to image sensors.

N-channel MOSFETs (*n*-MOSFETs) are made on *p*-type substrate and are used in the pixel of most CIS which collect electrons as photogenerated signal. The classic long channel model gives a good approximation for the characteristics of the MOSFET. In the active region, when the drain–source voltage is small, i.e. $0 < V_{DS} < (V_{GS} - V_T)$, the drain current in a transistor with channel length $L$ and width $W$ is given by

$$I_D = \frac{\mu_n C_{ox} W}{L}\left[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}\right] \tag{1.69}$$

For surface channel transistors the electron mobility $\mu_n$ is much smaller than the bulk mobility used to describe charge collection because of increased electron scattering at the Si–$SiO_2$ interface ([26], p 203). The active region is called a 'linear regime' because the drain current depends linearly on $V_{GS}$. Also, when $V_{DS} \ll (V_{GS} - V_T)$ the second term in the brackets in (1.69) can be ignored and the drain current becomes approximately linearly dependent on the drain–source voltage:

$$I_D \approx \frac{\mu_n C_{ox} W}{L}(V_{GS} - V_T)V_{DS} \tag{1.70}$$

In the linear regime the MOSFET behaves as a resistor with voltage-controlled resistance given by:

$$R_{lin} = \frac{V_{DS}}{I_D} = \frac{L}{\mu_n C_{ox} W(V_{GS} - V_T)} \tag{1.71}$$

This is used in analogue switches which operate at very low voltage drop across the drain–source $V_{DS} \ll (V_{GS} - V_T)$ in their 'on' state, when $V_{GS} > V_T$.

For fixed $V_{GS}$ the drain current described by (1.69) initially increases with the drain–source voltage, but as $V_{DS}$ continues to increase, the second term in the brackets becomes larger and the drain current increase slows down. Eventually the drain current stops increasing with $V_{DS}$, or in other words, *saturates*. Mathematically, saturation is expressed as $\partial I_D/\partial V_{DS} = 0$, and by differentiating (1.69) we can see that this happens when $V_{DS} = V_{GS} - V_T$. For $V_{DS} \geqslant V_{GS} - V_T$ the drain current in saturation depends quadratically on the gate–source voltage as

$$I_D = \frac{\mu_n C_{ox} W}{2L}(V_{GS} - V_T)^2 \tag{1.72}$$

Equation (1.72) describes a *voltage-controlled current source* because the drain current does not depend on the drain voltage, but only on the input $V_{GS}$. A measure of the dependence of the drain current on the gate–source voltage is the gate transconductance $g_m$, a very important device parameter:

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \frac{\mu_n C_{ox} W}{L}(V_{GS} - V_T) \tag{1.73}$$

Using (1.72), the gate transconductance can be expressed also as

$$g_m = \sqrt{\frac{2\mu_n C_{ox} W}{L} I_D} \tag{1.74}$$

The transconductance describes how a small change $v_{GS}$ of the gate–source voltage forces a drain current change $i_D = g_m v_{GS}$. This is provided that its DC gate–source voltage $V_{GS}$ is above threshold, the MOSFET is biased in saturation, and the drain current is not limited by the supply. The changes $v_{GS}$ and $i_D$ can be thought of as small AC signals on top of the larger $V_{GS}$ and $I_D$, correspondingly. If the drain current passes through the load resistor $R_L$ as in figure 1.23(a), $i_D$ will induce a change $v_o$ of the drain voltage equal to

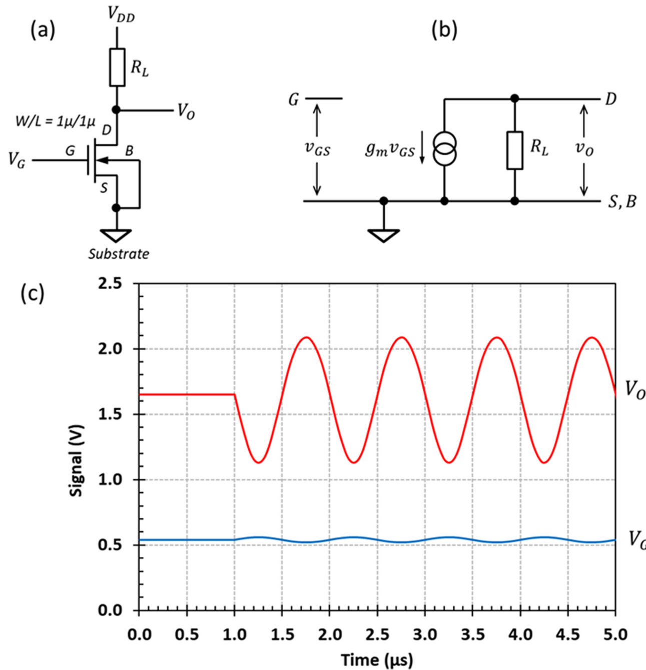


**Figure 1.23.** Common source MOSFET amplifier with resistive load with the substrate (bulk) connected to the source (a) and its equivalent schematic (b). SPICE simulation with $R_L = 1\text{M}\Omega$, $V_{DD} = 3.3$ V and 40 $\text{mV}_{pp}$, 1 MHz input sinewave signal starting from 1 µs onwards, showing 954 $\text{mV}_{pp}$ at the output (c).

$$v_o = i_D R_L = -g_m v_{GS} R_L \tag{1.75}$$

The circuit is *inverting* because increasing $V_G$ makes the drain current increase too, which in turn makes the output voltage $V_O$ decrease due to the larger voltage across the load resistor. Since the top end of $R_L$ is at the supply $V_{DD}$, which is AC ground (figure 1.23(b)), the AC gain of this circuit is the change of the output voltage divided by the change of the input voltage:

$$G = \frac{v_o}{v_{GS}} = -g_m R_L \tag{1.76}$$

The gain given by (1.76) applies only to AC signals well below the bandwidth of the circuit and should not be confused with the ratio of the DC voltages at the drain and the gate; the DC voltage ratio is a completely different matter.

The gain of this simple single-transistor circuit can be substantial, as figure 1.23(c) demonstrates. Here the input AC signal is superimposed on a DC gate voltage of 0.54 V, needed to bias the transistor in saturation. This DC bias is chosen so that the static drain voltage is at half the supply (1.65 V), thus maximising the swing at the output.

---

**Example 1.15.** Calculate the gate transconductance of an $n$-MOSFET with $W = L = 1$ µm, $\mu_n = 400$ cm$^2$ V$^{-1}$ s$^{-1}$, $C_{ox} = 4.93$ fF µm$^{-2}$ (7 nm thick gate oxide), $V_T = 0.4$ V and $V_{GS} = 0.5$ V. Also, calculate the $g_m$ of the transistor in figure 1.23.
**Solution:** Substituting the parameters in formula (1.73) we get:

$$g_m = \frac{400 \times 4.93 \times 10^{-15} \times 10^8 \times 10^{-4}}{10^{-4}}(0.5 - 0.4) = 19.7\ \mu\text{A V}^{-1}$$

To calculate the transconductance of the transistor in figure 1.23 we can use (1.76) after calculating the gain from the given input and output AC voltages:

$$g_m = \frac{|G|}{R_L} = \frac{954/40}{10^6} = \frac{23.9}{10^6} = 23.9\ \mu\text{A V}^{-1}$$

We can also see that the DC current through the MOSFET is $I_D = (V_{DD} - V_o)/R_L = 1.65$ µA, and the DC voltage ratio $V_O/V_G = 1.65/0.54 = 3.1$ has nothing to do with the AC gain, which is 945 mV/40 mV = 23.9.

---

The achievable gain depends not only on the product $g_m R_L$ but also on the supply voltage. Increasing $g_m$ makes the drain current increase too, according to (1.74). With the supply fixed, the load resistance $R_L$ cannot be made arbitrarily high because the voltage drop across it would 'eat up' into the available voltage for the MOSFET. A way to get around this limitation is to use another MOSFET as an active load instead of a resistor [29], which is the preferred method in IC design.

The formulas describing MOSFET operation assume that no current flows when the gate–source voltage is below threshold. In practice when $V_{GS} < V_T$ the drain current is very small but not zero, and the MOSFET operates in *subthreshold* (also known as weak inversion) mode. The drain current is caused by diffusion of electrons from the source because their concentration is higher than at the drain [2]. The subthreshold current can be described by [29]:

$$I_D = I_{D0}\frac{W}{L}\exp\left(\frac{V_{GS}}{n\varphi_T}\right) \tag{1.77}$$

where $I_{D0}$ is a technology parameter of the order of $10^{-12}$ A called off-state leakage, $\varphi_T = kT/q$ is the thermal potential and $n$ is the subthreshold slope factor. Formula (1.77) is valid for $V_{DS} \gg \varphi_T$.

In subthreshold mode the drain current depends exponentially on the gate–source voltage, rather than quadratically as in saturation (1.72). The gate transconductance from (1.77) is

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \frac{I_{D0}}{n\varphi_T}\frac{W}{L}\exp\left(\frac{V_{GS}}{n\varphi_T}\right) = \frac{I_D}{n\varphi_T} \tag{1.78}$$

The much higher rate of change of the drain current would appear to make the transconductance high, but the drain current in subthreshold is in the nanoamp range, so in practice $g_m$ is much lower than in saturation. This is one reason why MOSFETs operating in subthreshold mode are rarely used as amplifiers. However, subthreshold operation occurs in many MOSFET circuits where the source is driving a high impedance load, such as the sense node connected to the reset transistor in image sensors.

TCAD models such as the one in figure 1.24 are used to extract the transistor characteristics, which after parameterisation are converted to SPICE models for more convenient and faster simulations. The simulated input characteristic in

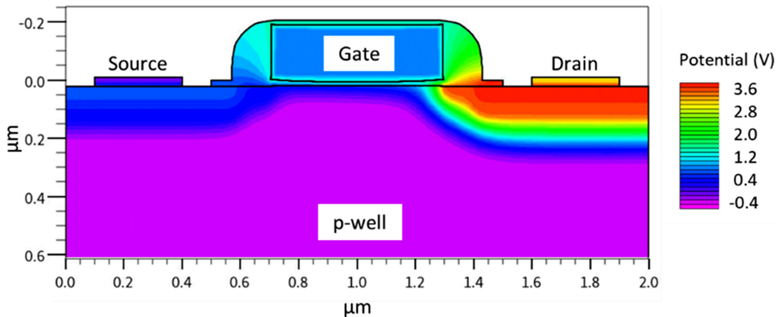


**Figure 1.24.** 2D TCAD model of an *n*-MOSFET with $L = 0.6$ µm, $W = 1.0$ µm, $t_{ox} = 12$ nm and uniform *p*-well with boron doping of $2 \times 10^{17}$ cm$^{-3}$. The substrate (bulk) and $V_S$ are at ground, and $V_G = 1.0$ V, $V_D = 3.3$ V. The structure is symmetrical, and the source and the drain are interchangeable.

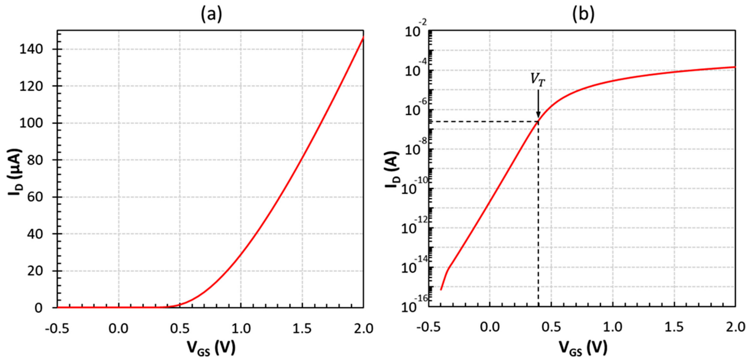


**Figure 1.25.** TCAD simulation of the input characteristic of the *n*-MOSFET in figure 1.24 for $V_{DS}$ = 3.3 V plotted on a linear scale (a), and on a semi-log scale (b).

figure 1.25 shows that the threshold of this transistor, defined as the gate–source voltage at which the drain current is 100 nA $\times W/L$ [26], is approximately 0.4 V.

Below threshold, the exponential dependence (1.77) holds for nearly 8 orders of magnitude change of the drain current (figure 1.25(b)). The off-state leakage at $V_{GS} = 0$ V is around 20 pA, which may be low enough for many applications, but is enormous in the context of image sensors—125 million electrons per second. The transistor turns off completely only when its gate–source voltage is negative, and the drain current becomes too low to be reliably simulated when $V_{GS} < -0.5$ V. Such low off-state currents are indeed needed in image sensors, where the pixel dark current can be measured in few electrons per second at room temperature.

### 1.8.3 Output resistance and body effect

Formula (1.72) tells us that in saturation the drain current does not depend on the drain–source voltage, i.e. the MOSFET behaves as an ideal current source. This is not exactly what happens in practice; as the drain–source voltage increases, the effective channel length decreases through a mechanism known as channel modulation [26] and the drain current slightly increases. The effect is easily seen in the output transistor characteristics, such as those in figure 1.26, where the drain current continues to increase with $V_{DS}$ after saturation is reached. In the absence of channel modulation, the drain current would be 'flat' for high $V_{DS}$ and the output impedance would be infinity.

This effect can be approximated with a resistor connected in parallel with the channel, between the source and the drain. Mathematically this is expressed by multiplying the drain current (1.72) with a term containing the drain–source voltage and the channel modulation parameter $\lambda$:

$$I_D = \frac{\mu_n C_{ox} W}{2L}(V_{GS} - V_T)^2(1 + \lambda V_{DS}) \quad (1.79)$$

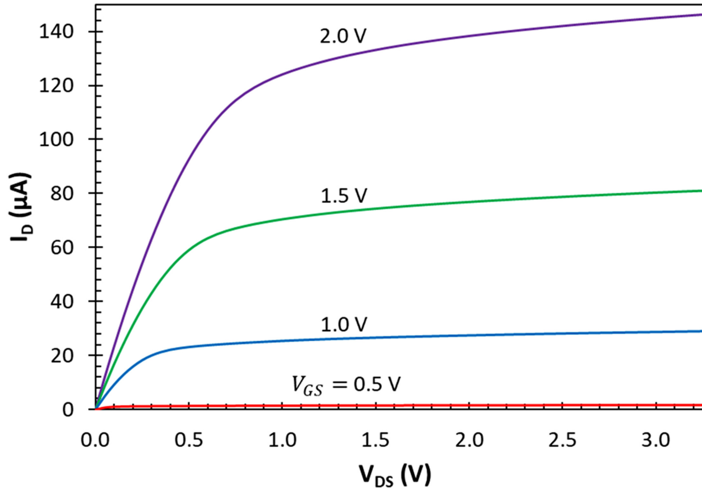


**Figure 1.26.** Output characteristics of the transistor in figure 1.24.

The change of the drain current caused by the drain–source voltage is called output conductance[3] $g_{\mathrm{ds}}$ and is defined as

$$g_{\mathrm{ds}} = \frac{\partial I_{\mathrm{D}}}{\partial V_{\mathrm{DS}}} = \frac{\mu_n C_{\mathrm{ox}} W}{2L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2 \lambda = \frac{\lambda I_{\mathrm{D}}}{1 + \lambda V_{\mathrm{DS}}} \tag{1.80}$$

Normally the parameter $\lambda$ is small, and as a rule of thumb $g_{\mathrm{ds}}$ is about a hundred times smaller than the gate transconductance $g_{\mathrm{m}}$ [29]. The output conductance (1.80) for small $\lambda$ simplifies to

$$g_{\mathrm{ds}} = \frac{\lambda I_{\mathrm{D}}}{1 + \lambda V_{\mathrm{DS}}} \cong \lambda I_{\mathrm{D}} \tag{1.81}$$

The output conductance $g_{\mathrm{ds}}$ can be substituted by the output resistance[4] $r_{\mathrm{ds}}$ in the MOSFET model in figure 1.27. The output resistance is simply the inverse of the output conductance:

$$r_{\mathrm{ds}} = \frac{1}{g_{\mathrm{ds}}} = \frac{1}{\lambda I_{\mathrm{D}}} \tag{1.82}$$

A much stronger effect on the drain current is caused by the source–substrate voltage through modulation of the transistor threshold. This is called body effect and is important for many circuits such as source followers, amplifiers and analogue switches.

[3] The drain–source voltage directly affects the drain current and therefore the term we use is conductance. The gate–source voltage influences the drain current indirectly, and the term transconductance (transfer of conductance) is used. It is also called mutual conductance, which gives the letter m in $g_{\mathrm{m}}$.

[4] Normally we think of the output resistance as a resistor in series with a voltage source. The MOSFET is a current source, therefore its output resistance is in parallel with the channel.

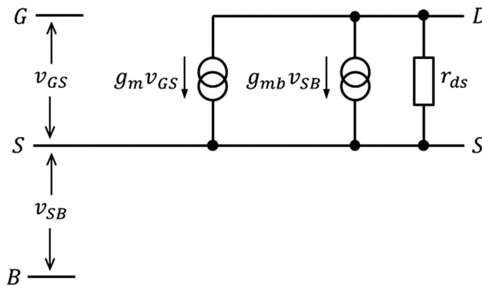


**Figure 1.27.** MOSFET model with output resistance and body effect.

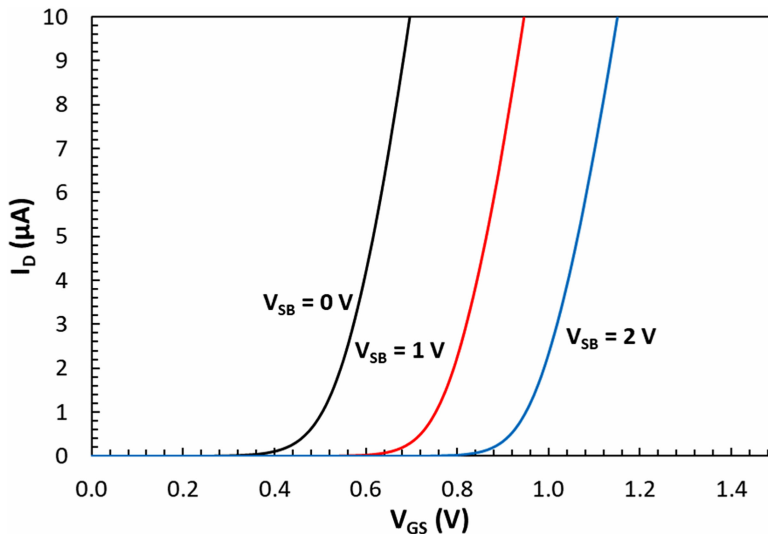


**Figure 1.28.** Body effect in an $n$-MOSFET with $L = 1.0$ μm and $W = 1.0$ μm, manufactured in 180 nm CMOS process.

In normal operation the source–bulk and drain–bulk *pn* junctions are reverse biased. When the source voltage is higher than the bulk ($V_{SB} > 0$), the depletion under the channel grows. This increases the amount of negative space charge, which is coupled capacitively to the inversion layer and induces in it a charge of the opposite polarity. Therefore, the gate voltage forcing channel inversion must increase to maintain the same drain current, which is equivalent to increasing the transistor threshold.

A more detailed description of the body effect is given in the following section, but here we will look at the practical aspects. Figure 1.28 shows the input characteristics of an $n$-MOSFET for three different source–bulk voltages as a parameter. The threshold increases in a sub-linear fashion by over 400 mV for a 2 V change in $V_{SB}$ (the full dependence is given by (1.94)), and more than doubles. This is

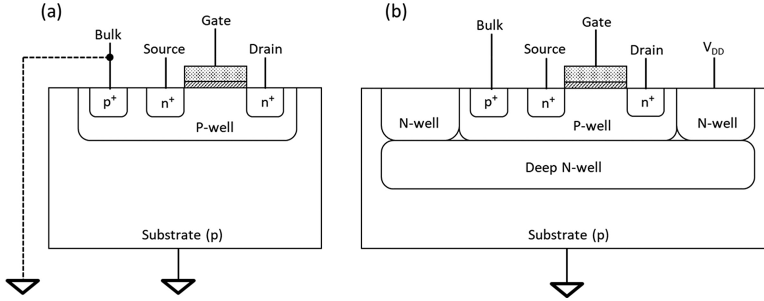


**Figure 1.29.** MOSFET in a *p*-well intrinsically connected to substrate (a) and in a floating *p*-well (hot *p*-well) with deep *n*-well isolation (b).

a substantial change; such large increase in $V_T$ is usually not welcome because it reduces the available voltage range for the signal, since the supply is fixed.

Similarly to the gate transconductance, the body transconductance $g_{mb}$ can be calculated from (1.79) as the rate of change of the drain current caused by $V_{SB}$:

$$g_{mb} = \frac{\partial I_D}{\partial V_{SB}} = \frac{\partial I_D}{\partial V_T}\frac{\partial V_T}{\partial V_{SB}} = -\frac{\mu_n C_{ox} W}{L}(V_{GS} - V_T)(1 + \lambda V_{DS})\frac{\partial V_T}{\partial V_{SB}} \tag{1.83}$$

and using (1.73) with added channel modulation, we can write

$$g_{mb} = -g_m \frac{\partial V_T}{\partial V_{SB}} \tag{1.84}$$

The negative sign in (1.83) is there because $\partial V_T / \partial V_{SB} > 0$ and increasing $V_{SB}$ makes the drain current smaller by increasing the threshold. From the example in figure 1.28 and (1.84) we see that $g_{mb}$ is roughly a factor of 10 smaller than $g_m$, which generally holds as a rule of thumb [29].

The body effect is not unavoidable; if the bulk and the source are at the same potential it can be eliminated. The bulk, which is the *p*-well in *n*-MOSFETs, can either be joined to the substrate intrinsically because they are both *p*-type, as in figure 1.29(a), or are floating, as in figure 1.29(b).

Unless all transistors in the circuit operate with their sources grounded, connecting to sources to the bulk is not an option for figure 1.29(a), therefore the body effect will be there. The floating *p*-well in figure 1.29(b) solves the problem because the bulk can be connected to the source of each transistor and be at a different potential, made possible by the deep *n*-well isolation.

### 1.8.4 Transistor threshold

Normally, transistor thresholds are chosen to be around 0.6–0.8 V. These high thresholds are necessary to reduce the subthreshold leakage in large digital circuits, where the leakage from billions of transistors can add up to unacceptable levels.

High transistor thresholds may not be optimal for analogue circuits. Supply voltages in image sensors are low and transistor thresholds often appear as

undesirable voltage offsets, eating into the precious signal amplitude. Because of this, special transistors with ‘low $V_T$’ or even ‘ultra-low $V_T$’ have been developed and are increasingly being used.

Understanding how the transistor threshold depends on the oxide thickness, substrate doping concentration and bias, and its impact on performance is very important for an image sensor designer.

In an *n*-MOSFET using highly doped *n*-type poly-Si gate at zero gate bias (equilibrium state) the *p*-type body is already in depletion due to the difference in the work functions between the gate and the substrate. Starting with the energy diagram in figure 1.30, we observe that a *negative* voltage has to be applied to the gate so that the flat-band condition is achieved. The voltage $V_{FB}$ is called flat-band voltage and is defined as the difference between the Fermi levels of the gate and the substrate, which in this condition is equal to the difference between the two work functions

$$V_{FB} = \psi_G - \psi_S \tag{1.85}$$

For the structure in figure 1.30 (highly doped *n*-type poly-Si gate over a *p*-type substrate) $V_{FB}$ is negative, at about −0.6 to −0.9 V. From figure 1.30 we can also see that $V_{FB} = \varphi_B - \varphi_G$ and using (1.56) can write

$$V_{FB} = \varphi_B - \varphi_G = -\frac{kT}{q}\ln\left(\frac{N_A}{n_i}\right) - \frac{kT}{q}\ln\left(\frac{N_D}{n_i}\right) \tag{1.86}$$

where $N_A$ and $N_D$ are the dopant concentrations in the substrate and the polysilicon gate, respectively.

The threshold voltage can be calculated from the general equation (1.87) [26], which simply states that the potential drops across the oxide and the substrate balance the gate voltage.

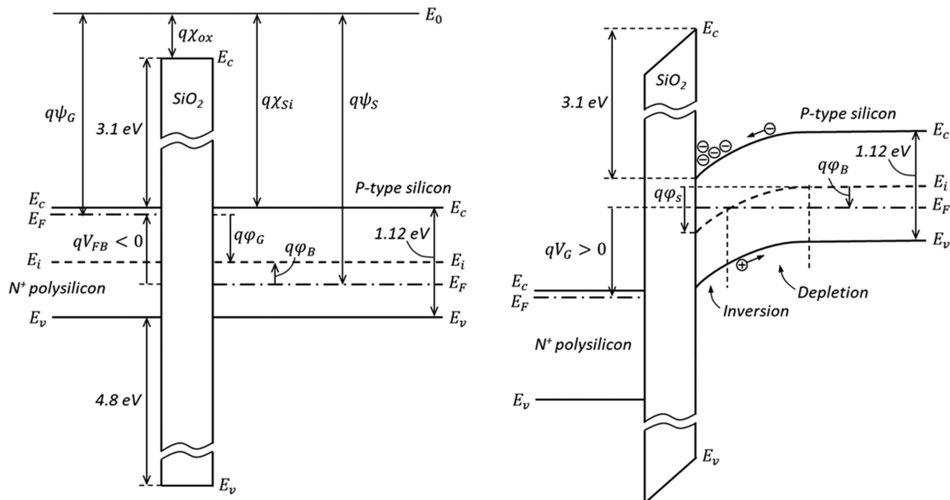


**Figure 1.30.** Band diagram of a MOS capacitor in flat-band condition and in strong inversion.

$$V_G - V_{FB} = \varphi_s + V_{ox} \tag{1.87}$$

In flat-band condition $V_G - V_{FB} = 0$ and $\varphi_s = V_{ox} = 0$. Starting with the flat-band condition, we then calculate the additional gate voltage change to bring the substrate into strong inversion or any other condition.

The voltage across the oxide is given by the space charge in the depleted region per unit area $Q_{b0}$, divided by the gate capacitance:

$$V_{ox} = \frac{Q_{b0}}{C_{ox}} \tag{1.88}$$

Using equation (1.58) for $W$, $Q_{b0}$ is written as:

$$Q_{b0} = qN_A W = qN_A\sqrt{\frac{2\varepsilon_0\varepsilon_{Si}\varphi_s}{qN_A}} = \sqrt{2\varepsilon_0\varepsilon_{Si}qN_A\varphi_s} \tag{1.89}$$

The space charge in a $p$-type substrate is negative due to the negatively charged acceptor atoms. When an additional reverse voltage $V_{SB}$ is applied between the source and the bulk the width of the depletion region expands similarly to a reverse-biased $pn$ junction (1.45) and the space charge per unit area becomes

$$Q_b = \sqrt{2\varepsilon_0\varepsilon_{Si}qN_A(\varphi_s + V_{SB})} \tag{1.90}$$

Strong inversion is achieved when $\varphi_s = 2\varphi_B$. In contrast with the deep depletion shown in figure 1.20, in the $n$-MOSFET strong inversion is achieved very quickly because the highly doped source and drain provide an abundant supply of electrons. Once inversion is reached, the surface potential and the depletion depth stop growing with the gate voltage because the semiconductor below is screened from further potential changes by the thin inversion layer.

The threshold voltage can be calculated from (1.87) using $V_T = V_G$ and $\varphi_s = 2\varphi_B$.

$$V_T = V_{FB} + 2\varphi_B + \frac{Q_b}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} \tag{1.91}$$

Here $Q_{ss}$ is the fixed surface charge density at the oxide, which is normally positive and has density $Q_{ss}/q$ in the range $10^{10}$–$10^{11}$ cm$^{-2}$. Equation (1.91) can be written as

$$V_T = V_{FB} + 2\varphi_B + \frac{Q_{b0}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} + \frac{Q_b - Q_{b0}}{C_{ox}} \tag{1.92}$$

The first four terms do not depend on the substrate voltage and give the threshold voltage for $V_{SB} = 0$:

$$V_{T0} = V_{FB} + 2\varphi_B + \frac{Q_{b0}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} = V_{FB} + 2\varphi_B + \frac{\sqrt{2\varepsilon_0\varepsilon_{Si}qN_A\,|\,2\varphi_B\,|}}{C_{ox}} - \frac{Q_{ss}}{C_{ox}} \tag{1.93}$$

The final expression for the threshold voltage including the dependence on $V_{SB}$ is

$$V_T = V_{T0} + \frac{Q_b - Q_{b0}}{C_{ox}} = V_{T0} + \frac{\sqrt{2\varepsilon_0\varepsilon_{Si}qN_A}}{C_{ox}}\left(\sqrt{|2\varphi_B| + V_{SB}} - \sqrt{|2\varphi_B|}\right) \quad (1.94)$$

It could be very confusing to keep track of the signs of the terms in (1.93), but the following considerations and a careful look at figure 1.30 should help. The gate voltage should bend the bands downwards to reach inversion, therefore the second term $2\varphi_B$ should be positive. Also, the gate voltage should counteract the negative space charge, therefore the third term must also be positive. The fourth term is negative because $Q_{ss}$ reduces the needed voltage to achieve inversion, since $Q_{ss}$ is a positive charge which acts in the same direction.

---

**Example 1.16.** Calculate the threshold voltage of an *n*-MOSFET with the following parameters: $N_A = 2 \times 10^{17}\,\text{cm}^{-3}$, gate doping $N_D = 10^{19}\,\text{cm}^{-3}$, $C_{ox} = 4.93\,\text{fF}\,\mu\text{m}^{-2}$ (7 nm thick gate oxide), $Q_{ss}/q = 10^{10}\ \text{cm}^{-2}$ and 300 K. Use that $n_i = 1.45 \times 10^{10}\ \text{cm}^{-3}$, $\varepsilon_0 = 8.85 \times 10^{-14}\ \text{F cm}^{-1}$ and $\varepsilon_{Si} = 11.9$. Also calculate $V_T$ for $N_A = 1 \times 10^{16}\ \text{cm}^{-3}$.
**Solution:** The first step is to calculate $\varphi_B$, $\varphi_G$ and the flat-band voltage from (1.86)

$$\varphi_B = -\frac{kT}{q}\ln\left(\frac{N_A}{n_i}\right) = -\frac{1.38 \times 10^{-23} \times 300}{1.6 \times 10^{-19}}\ln\left(\frac{2 \times 10^{17}}{1.45 \times 10^{10}}\right) = -0.425\ \text{V}$$

$$\varphi_G = \frac{kT}{q}\ln\left(\frac{N_D}{n_i}\right) = -\frac{1.38 \times 10^{-23} \times 300}{1.6 \times 10^{-19}}\ln\left(\frac{10^{19}}{1.45 \times 10^{10}}\right) = 0.527\ \text{V}$$

$$V_{FB} = \varphi_B - \varphi_G = -0.425 - 0.527 = -0.952\ \text{V}$$

Next, the third term in (1.93) is:

$$\frac{\sqrt{2\varepsilon_0\varepsilon_{Si}qN_A|2\varphi_B|}}{C_{ox}} = \frac{\sqrt{2 \times 8.85 \times 10^{-14} \times 11.9 \times 1.6 \times 10^{-19} \times 2 \times 10^{17} \times 2 \times 0.425}}{4.93 \times 10^{-15} \times 10^{8}} = 0.486\ \text{V}$$

The last term is

$$\frac{Q_{ss}}{C_{ox}} = \frac{10^{10} \times 1.6 \times 10^{-19}}{4.93 \times 10^{-15} \times 10^{8}} = 0.003\ \text{V}$$

Finally, the threshold is:

$$V_{T0} = -0.952 + 2 \times 0.425 + 0.486 - 0.003 = 0.381\ \text{V}$$

Repeating the calculation for $N_A = 1 \times 10^{16}\ \text{cm}^{-3}$ gives $V_{T0} = -0.084$ V.

---

The threshold voltage can be effectively controlled by the dopant concentration of the substrate, and special 'threshold adjust' implants are used for that purpose. This example also shows that the natural threshold of *n*-MOSFETs in higher resistivity

substrates is around 0 V. Such transistors without a threshold adjustment implant are called 'native'.

The threshold volage is temperature dependent via the thermal potential entering $V_{T0}$. For *n*-MOSFETs the temperature coefficient of the threshold is negative and is typically around −2 mV °C$^{-1}$, with a range between −0.5 and −4 mV °C$^{-1}$; for *p*-MOSFETs the coefficient is positive.

### 1.8.5 Analogue switch

MOSFETs working as analogue switches are used to route signals and connect various points together. Their other use is in sample and hold (S&H) circuits to temporarily connect a signal to a capacitor, so that its value can be stored. Figure 1.31 shows a simple S&H circuit built with a MOSFET and a capacitor. With the transistor on, the capacitor charges to the input voltage $V_{in}$ through the resistance of the channel. After the switch turns off, the disconnected capacitor retains its voltage for a very long time (which can be seconds) and this makes it useful as an analogue memory.

To turn the switch on, the gate–source voltage must be above the transistor threshold, but the MOSFET is symmetrical and can work in both directions. The gate–source voltage can be considered as either $V_{GS} = V_G - V_{in}$ or $V_{GS} = V_G - V_{out}$; if one of them is higher than $V_T$ the transistor will turn on. The maximum input voltage this single-transistor analogue switch can handle is $V_G - V_T$, and is limited by the highest voltage that can be applied to the gate, which is typically the supply voltage. The threshold suffers from the body effect.

The source–drain voltage in the 'off' state can be very large, but in the 'on' state it must be low because the MOSFET behaves as a resistor only when $V_{DS} \ll (V_{GS} - V_T)$. When connecting two points with very different potentials, initially the drain current can be large and the MOSFET will be operating in saturation, until the voltages equalise, and the transistor enters linear regime. This is what happens in the switch in figure 1.31—regardless of the initial capacitor voltage (within the operating limits), the end voltage is $V_{in}$.

Analogue switching occurs if at least one of the terminals $V_{in}$ and $V_{out}$ is at high impedance, so that the input and output have a chance to equalise. If both $V_{in}$ and $V_{out}$ are low impedance sources this will not happen, and they will maintain their own voltages—current will flow between them, but we would not call that switching.

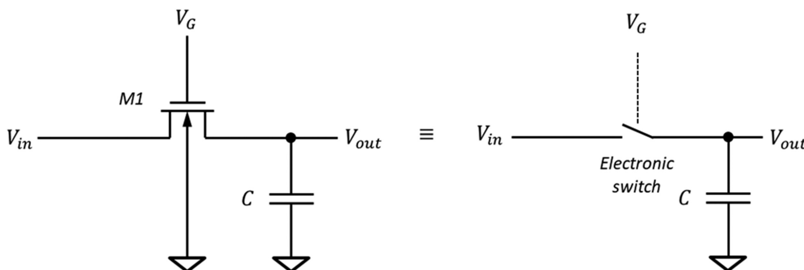


**Figure 1.31.** Analogue sample and hold switch built with an *n*-MOSFET (a) and its equivalent circuit.

---

**Example 1.17.** Calculate the 'on' resistance of $n$-MOSFET with $W = L = 1.0$ μm, $\mu_n = 400\ \text{cm}^2\ \text{V}^{-1}\ \text{s}^{-1}$, $C_{ox} = 4.93\ \text{fF}\ \mu\text{m}^{-2}$ (7 nm thick gate oxide), $V_T = 0.4$ V and $V_{GS} = 1.0$ V. Also, calculate the bandwidth of the circuit in figure 1.31 for $C = 1$ pF.
**Solution:** The drain–source voltage should be much smaller than the overdrive voltage $V_{GS} - V_T$, so that the transistor operates in the linear regime. From equation (1.71) the resistance is:

$$R_{\text{lin}} = \frac{10^{-4}}{400 \times 4.93 \times 10^{-15} \times 10^{8} \times 10^{-4} \times (1.0\text{–}0.4)} = 8.4\ \text{k}\Omega$$

The signal bandwidth is calculated from:

$$BW = \frac{1}{2\pi R_{\text{lin}} C} = \frac{1}{6.28 \times 8400 \times 10^{-12}} = 18.9\ \text{MHz}$$

Despite the MOSFET appearing to have too high a resistance for something to be called an analogue switch, for the usual small load capacitances the response of the circuits like those in figure 1.31 can be quite fast.

---

The simple $n$-MOSFET switch in figure 1.31, and its $p$-MOSFET counterpart, have substantial limitations on their operating voltages. For example, in a circuit supplied with 3.3 V, a typical $n$-MOSFET will not be able to switch voltages higher than about 2.5 V. This is because the threshold is high due to the body effect (see figure 1.28), and the maximum gate voltage is 3.3 V. Also, its 'on' resistance changes wildly as a function of the input voltage, becoming very high as the gate–source voltage approaches threshold according to (1.71).

To solve this problem, the CMOS analogue switch [29] in figure 1.32 is used. It works well for any input voltage within the supply rails and exhibits much smaller change of the 'on' resistance than the single-transistor switch. The CMOS variant has an $n$-MOSFET and a $p$-MOSFET connected in parallel; the $n$-channel transistor

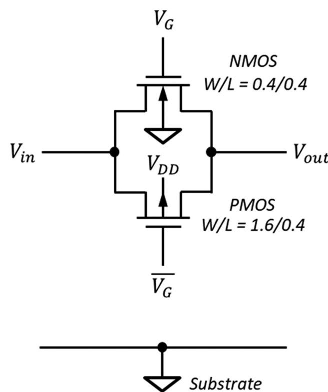


**Figure 1.32.** CMOS analogue switch.

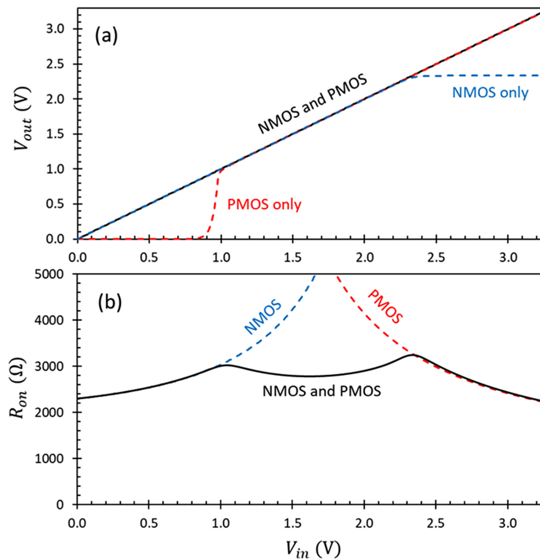


**Figure 1.33.** Output characteristic (a) and resistance (b) of the CMOS analogue switch in figure 1.32 for $V_{DD}$ = 3.3 V, $V_G$ = 3.3 V and $\overline{V_G}$ = 0 V.

takes care of the low input voltages for which $V_{GS}$ is large. To complement this, the *p*-channel transistor works at high inputs because zero gate voltage is applied to turn it on, and its $V_{GS}$ is maximised. The switch control requires complementary gate voltages: $V_G = V_{DD}$ and $\overline{V_G} = 0$ for the 'on' state and the opposite for the 'off' state.

Figure 1.33(a) shows a simulation of the CMOS switch in figure 1.32 for input voltage spanning from substrate potential to the supply. We see that the *n*-MOSFET transistor behaves like a resistor up to an input voltage around 2.4 V (for 3.3 V supply); above that it fails to faithfully follow the input. This is caused by its high threshold—the output can never be higher than the gate voltage minus the threshold $V_G - V_T$, and the threshold is around 0.9 V due to the strong body effect. Similarly, the *p*-MOSFET transistor begins to work only for voltages exceeding about 1 V. By connecting two complementary transistors in parallel we have an excellent switch action across the whole input range, spanning from zero (substrate potential) to the supply $V_{DD}$.

The resistance of the CMOS switch in figure 1.33(b) is the parallel combination of the two transistors: The *n*-MOSFET has low resistance at low input voltages and the *p*-MOSFET at high voltages, so the paralleled resistance does not suffer from large variations. The transistors can be sized appropriately so that the shape of the resistance curve is symmetrical, as done for the switch in figure 1.32.

### 1.8.6 MOSFET capacitor

When operating in inversion or accumulation, the concentration of free carriers in the MOSFET channel is very high. The free carriers act as one half of a parallel plate capacitor, with the other half being the gate. Normally only inversion is used; this is

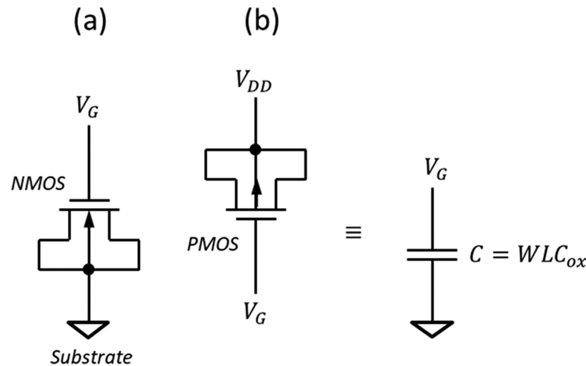


**Figure 1.34.** MOSFETs used as capacitors.

because to bias the channel of a *n*-MOSFET in accumulation the gate must be negative with respect to the bulk, which is possible, but rarely a practical option.

Figure 1.34 shows two ways to create MOSFET capacitors, for *n*- and *p*-channel transistors. For both the bulk, source and drain are connected, and both circuits are equivalent to a capacitor to ground, because the supply $V_{DD}$ is AC ground too. For correct operation in inversion, we need $V_G > V_T$ for the *n*-MOSFET, and $V_G < V_{DD} - V_T$ for the *p*-MOSFET.

---

**Example 1.18.** Calculate the gate capacitance to substrate of a *n*-MOSFET with $W = L = 5.0$ µm, $C_{ox} = 4.93$ fF µm$^{-2}$ (7 nm thick gate oxide), $V_T = 0.4$ V and $V_{GS} = 2.0$ V. Compare with the *pn* junction capacitance in example 1.13.

**Solution:** Because $V_{GS} \gg V_T$ the channel is in inversion, therefore the gate capacitance is

$$C = WLC_{ox} = 5 \times 5 \times 4.93 = 123.3 \text{ fF}$$

This capacitance is about 90 times higher than the *pn* junction with the same area in example 1.13.

---

The gate capacitance of a MOSFET offers the highest capacitance per unit area because it uses very thin oxide measured in just a few nanometres; there is simply no other way to achieve such high specific capacitance. The downside is that the transistor must always be in inversion, and failure to follow this results in much smaller capacitance, paired with nonlinearity.

## 1.9 Source follower

### 1.9.1 Gain

The source follower (SF) is among the simplest, but also the most important circuits in image sensors. The name simply means that the source voltage 'follows' the gate

voltage. Because of the transistor threshold there is a DC offset between the gate and the source and the 'following' applies only to the AC component of the gate voltage.

Figure 1.35 shows two basic SF circuits using a resistor as a load; they are useful to explain how the SF works but rarely used in practice because resistors take up too much space in integrated circuits.

We can derive the voltage gain of the simple SF in figure 1.35(b). The transistor operates in saturation and the relationship between the DC voltages (marked with capital letters) for $V_{\mathrm{GS}} > V_{\mathrm{T}}$ is:

$$V_{\mathrm{O}} \approx V_{\mathrm{G}} - V_{\mathrm{T}} \tag{1.95}$$

For small signal analysis we use the AC components of the output voltage $v_{\mathrm{O}}$, drain current $i_{\mathrm{D}}$ and the gate–source voltage $v_{\mathrm{GS}}$, marked with lower case letters.

The definition of the transconductance (1.73) applied to small signals states that a small change of the gate–source voltage $v_{\mathrm{GS}}$ around a DC bias point induces a small change in the drain current $i_{\mathrm{D}}$, and we can write

$$i_{\mathrm{D}} = g_{\mathrm{m}} v_{\mathrm{GS}} \tag{1.96}$$

The small signal schematic in figure 1.36 shows that the drain current flows through the load resistor $R_{\mathrm{L}}$ since the drain is connected to the supply voltage $V_{\mathrm{DD}}$, which is at AC ground. For simplicity, the load resistance here includes the

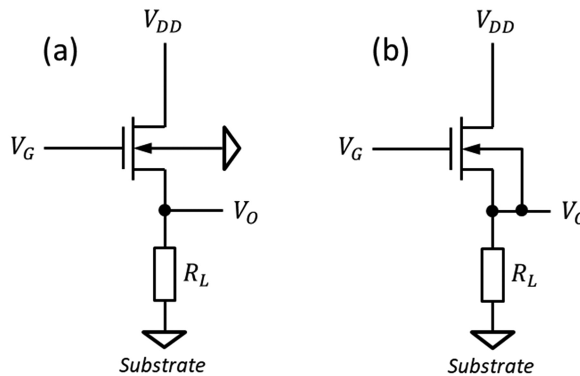


**Figure 1.35.** Source follower with resistor load and body connected to the substrate(a); and with the body connected to the source to avoid the body effect (b).

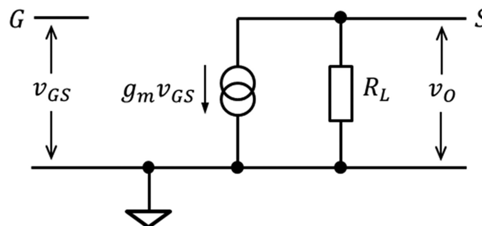


**Figure 1.36.** Small signal equivalent schematic of a source follower without body effect.

transistor's output resistance $r_{ds}$, which similarly to (1.82), can be replaced by the transconductance $g_L = 1/R_L$.

The output voltage $v_O$ of the SF is the drain current (1.96) multiplied by the load resistance $R_L$:

$$v_O = i_D R_L = g_m v_{GS} R_L \tag{1.97}$$

Using that $v_{GS} = v_G - v_O$ and substituting in (1.97) we get

$$v_O = \left( \frac{g_m R_L}{1 + g_m R_L} \right) v_G \tag{1.98}$$

The term in brackets is the gain of the source follower $G_{SF}$:

$$G_{SF} = \frac{v_O}{v_G} = \frac{g_m R_L}{1 + g_m R_L} = \frac{g_m}{g_m + g_L} \tag{1.99}$$

In the absence of body effect, the gain $G_{SF}$ given by (1.99) is very close to one and is typically above 0.9. With body effect the gain is

$$G_{SF} = \frac{g_m}{g_m + g_{mb} + g_L} \tag{1.100}$$

Because usually $g_{mb} \approx 0.1 g_m$ the theoretical maximum gain with body effect is around 0.9 (if $R_L$ is very large, $g_L \approx 0$), but in practice is around 0.8.

Figure 1.37 shows the DC transfer characteristic of a source follower with resistive load, which can be used to determine the gain. Formulas (1.99) and (1.100) give the gain for small input signals when the source follower is biased in saturation and should not be confused with the ratio between the DC voltages at the source and the gate. When the input voltage is close to $V_T$ the ratio of the DC source and gate

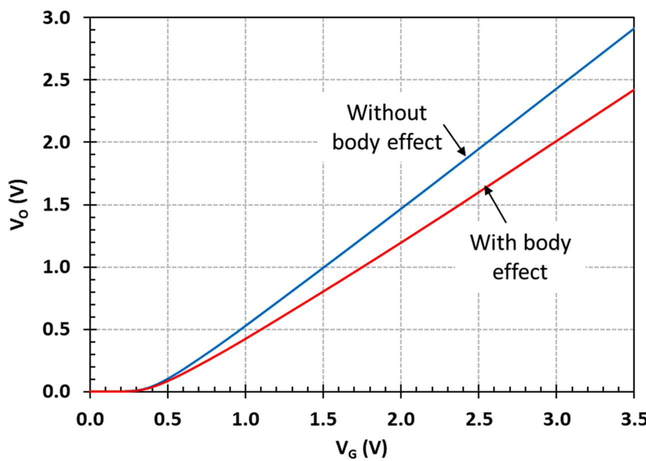


**Figure 1.37.** DC transfer characteristics of a source follower with $W/L = 1\mu/1\mu$, $R_L = 1$ MΩ and $V_{DD} = 3.3$ V, with and without body effect, corresponding to figure 1.35.

voltages can be very far from the AC small signal gain, but it does approach it for high input voltages. This is because the output voltage stays at zero until the input exceeds the transistor threshold.

---

**Example 1.19.** Calculate the approximate SF gain from the data in figure 1.37.
**Solution:** The gain is defined as the change of the output voltage over a small change in the input voltage. Since the output voltage is linear for $V_G > 0.6$ V, we can use any input voltage difference above that, for example between $V_G = 2.0$ V and $V_G = 3.0$ V. Without body effect $G_{SF} \approx (2.4 - 1.5)/1.0 = 0.9$; with body effect $G_{SF} \approx (2.0 - 1.2)/1.0 = 0.8$. A proper fit to the numerical data gives $G_{SF} = 0.94$ and $G_{SF} = 0.77$, correspondingly.

---

The gain of the source follower $G_{SF}$ is always less than one. For the typical source follower operating currents of few microamps the load resistor $R_L$ should be in the mega-ohm range. It is not practical or even possible to include such a large resistor on a chip; instead, an active load is used virtually everywhere. Figure 1.38(a) shows a source follower M1 working with an active current load. Because the current is flowing into the load it is also called a current sink.

The simplest possible current load is built with just one MOSFET, as shown in figure 1.38(b). Since in saturation the MOSFET behaves as a voltage-controlled current source according to (1.72), when an appropriate voltage $V_{bias}$ is applied to the gate of M2 the drain current $I_{bias}$ is nearly constant and independent of the output voltage $V_O$. In practice $I_{bias}$ depends slightly on $V_O$ due to the output resistance of M2.

The current load can be taken a step further in figure 1.38(c) with a current mirror consisting of the transistors M2 and M3. An external current $I_{in}$ is supplied to M3, which is connected as a MOS diode. This creates the voltage drop $V_{bias}$ across M3, which in turn supplies the gate of M2. What this does is to generate the voltage $V_{bias}$ locally to the source follower, using a current input instead of a voltage input. If transistors M2 and M3 were the same, the same gate–source voltage $V_{bias}$ ensures that their drain currents would be the same too. When $I_{bias}$ is just a few microamps,

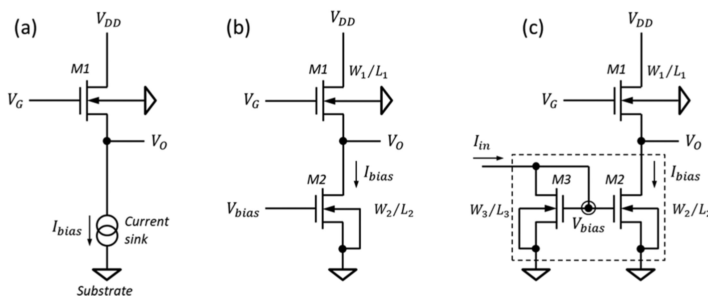


**Figure 1.38.** Source follower using an active current load: (a) symbolic representation; (b) with externally biased load transistor; (c) with current mirror.

the voltage $V_{\text{bias}}$ is very close to the transistor threshold and is typically around 0.4–0.6 V.

The current $I_{\text{bias}}$ depends on the channel width and length of M2 and M3 as [29]

$$I_{\text{bias}} = I_{\text{in}}\left(\frac{W_2}{W_3}\right)\left(\frac{L_3}{L_2}\right) \tag{1.101}$$

Equation (1.101) tells us that the bias current $I_{\text{bias}}$ 'mirrors' the input current $I_{\text{in}}$ through a proportionality constant, determined by the width and the length of the two transistors. Normally $L_2 = L_3$ and this allows the current ratio to be chosen simply as the width ratio of M2 over M3. The transistor M3 can generate the bias for many current load transistors and can be shared among them.

Source followers with an active load have better performance and higher gain than the resistor-loaded followers due to the high dynamic resistance of the load transistor. They are almost invariably used in all CMOS image sensors and will pop up frequently in the following chapters.

### 1.9.2 Input capacitance

The input capacitance of the source follower adds to the sense node capacitance, therefore, it is part of the charge-to-voltage conversion. Electrically, the input capacitance consists of two parts, as shown in figure 1.39(a):

(a) $C_{\text{d}}$ 'detection capacitance' is all the capacitance between the gate and the substrate, with the substrate being at AC ground. The gate–drain capacitance is also connected between gate and AC ground, and for simplicity is included in $C_{\text{d}}$.

(b) $C_{\text{GS}}$ is the capacitance between the gate and source.

We would like to know the effective gate capacitance to substrate $C_{\text{in}}$. We can find it by introducing a small charge $Q$ at the input, which changes the gate voltage by $v_{\text{G}} = Q/C_{\text{in}}$. The change of the output voltage, including the source follower gain, is $v_{\text{o}} = QG_{\text{SF}}/C_{\text{in}}$, therefore we need to find an expression for $v_{\text{o}}$.

The charge $Q$ is shared between the two capacitances, so that from conservation of change we have

$$Q = C_{\text{d}}v_{\text{G}} + C_{\text{GS}}v_{\text{GS}} \tag{1.102}$$

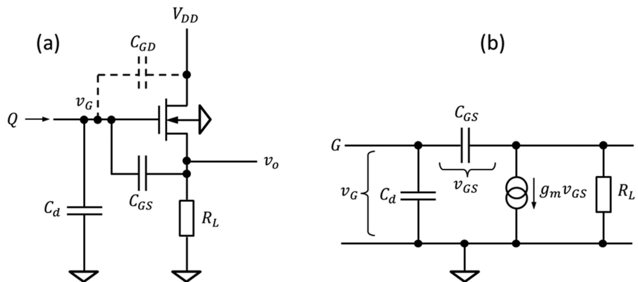


**Figure 1.39.** Source follower driven by a high impedance charge input (a) and its simulation schematic (b).

Using that $v_{GS} = v_G - v_o$ and $v_o = G_{SF}v_G$ from (1.99), equation (1.102) can be written as

$$Q = C_d\frac{v_o}{G_{SF}} + C_{GS}\left(\frac{v_o}{G_{SF}} - v_o\right) \tag{1.103}$$

From here, solving for $v_o$ we get

$$v_o = \frac{QG_{SF}}{C_d + C_{GS}(1 - G_{SF})} \tag{1.104}$$

From equation (1.104) we can conclude that the effective input capacitance of the source follower is

$$C_{in} = C_d + C_{GS}(1 - G_{SF}) \tag{1.105}$$

It is not surprising that $C_d$ appears unaltered in the input capacitance, because it is connected to the substrate. On the other hand, $C_{GS}$ is significantly attenuated; the voltage change across $C_{GS}$ is much smaller than the gate voltage because the source closely follows the gate. A reduced voltage change means that the capacitance is reduced too; in the extreme case when the source follows the gate exactly ($G_{SF} = 1$) there would be no voltage change across $C_{GS}$, therefore no current would flow through it. This is the same as if $C_{GS}$ is not connected at all, i.e. $C_{GS} = 0$. For the typical SF gains of 0.8–0.9, $C_{GS}$ is reduced by a substantial factor.

Figure 1.39 gives a schematic view of the capacitances as they appear electrically, but we can also consider their physical location in figure 1.40. There are three physical gate capacitances—to the source, the drain and the channel.

The gate–source and gate–drain edge capacitances are identical for the symmetrical transistor in figure 1.40 and are given by

$$C_{GSe} = C_{GDe} = C_e W \tag{1.106}$$

Here $C_e$ is the edge capacitance per unit length of gate overlap and can be found from the process specifications for the used CMOS technology.

The gate area capacitance $C_{Ga}$ is between the gate and the conducting channel. In a source follower configuration, the potential along the channel is superimposed on the source potential, therefore $C_{Ga}$ is effectively connected between the gate and the source, the same as $C_{GSe}$. $C_{Ga}$ is [30]

$$C_{Ga} = \frac{2}{3}C_{ox}WL \tag{1.107}$$

and is smaller than the expected $C_{ox}WL$ due to the potential distribution in the channel.

Therefore, substituting $C_{GDe}$ for $C_d$ and $(C_{GSe} + C_{Ga})$ for $C_{GS}$ in (1.105) the input capacitance of the source follower becomes

$$C_{in} = C_{GDe} + (C_{GSe} + C_{Ga})(1 - G_{SF}) \tag{1.108}$$

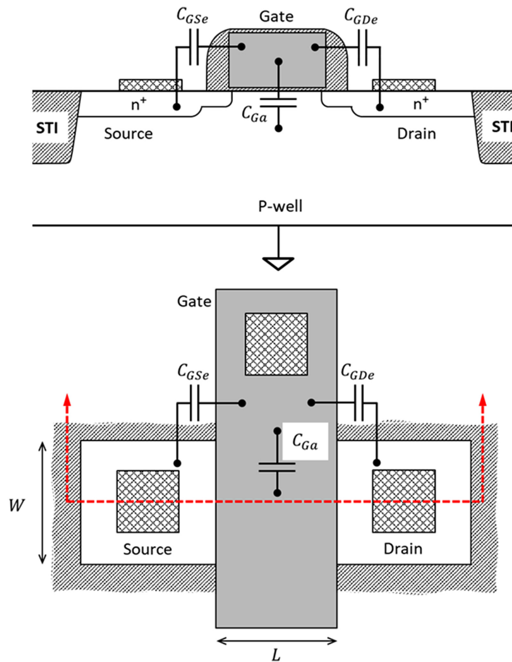


**Figure 1.40.** Location of the gate capacitances in a MOSFET. The shallow trench insulation (STI) oxide surrounds the MOSFET on all four sides.

Finally, using the transistor geometry, the capacitance can be written as:

$$C_{\text{in}} = C_{\text{e}}W + \left(C_{\text{e}}W + \frac{2}{3}C_{\text{ox}}WL\right)(1 - G_{\text{SF}}) \tag{1.109}$$

Usually $C_{\text{Ga}}$ is much larger than the edge capacitances due to the large $C_{\text{ox}}$, but fortunately it is strongly suppressed by the factor $(1 - G_{\text{SF}})$.

## Chapter summary

1. Photons generate electron–hole pairs in silicon via the internal photoeffect. Energetic charged particles generate electron–hole pairs via ionisation. Charge is generated independently of the doping concentration and the concentration of free carriers in the semiconductor.
2. A single photon with wavelength between 1100 and 369 nm can generate only one electron–hole pair in silicon. Above 1100 nm silicon is transparent and insensitive to light.
3. Photons with wavelengths shorter than 369 nm (3.36 eV) generate more than one electron–hole pair, and for photons with energies above 10 eV the electron–hole creation energy levels off to about 3.65 eV.
4. The absorption length of light in silicon depends strongly on the photon energy, and changes between few nanometres at UV to hundreds of micrometres for near-IR light and soft x-rays.

5. Carrier lifetimes in high quality silicon are in the millisecond range and decrease with increasing trap or doping concentration. Trap-assisted recombination is the dominant mechanism controlling carrier lifetime.
6. Charge is collected by diffusion and drift. Drift occurs under electric field and is the preferred way of charge collection due to its high speed.
7. Charge diffuses regardless of whether there is an electric field or not, but a short collection time under drift does not allow for significant diffusion. In field-free semiconductor the charge diffuses until it is trapped or reaches a region with electric field.
8. The rate of direct electron–hole recombination in silicon is vanishingly small except at very high carrier concentrations. This is why electrons and holes generated in a dense cloud diffuse away before they have a chance to recombine with each other.
9. Charge is converted to electrical signal, usually to voltage, at the capacitance of a sense node.
10. The depletion region established in a *pn* junction acts like an insulator and facilitates charge collection due to the internal electric field. The capacitance of the *pn* junction can be used as a sense node.
11. MOSFETs act as a voltage-controlled current sources. The drain current in saturation has a quadratic dependence on the gate–source voltage. As an amplifier, the MOSFET gain is reduced by the body effect and the output resistance.
12. Source followers are used as buffers due to their very high input impedance and well-defined gain, which is always less than one.

## References

[1] Mazziotta M 2008 Electron–hole pair creation energy and Fano factor temperature dependence in silicon *Nucl. Instrum. Methods Phys. Res.* A **A584** 436–9

[2] Sze S 1981 *Physics of Semiconductor Devices* 2nd edn (New York: Wiley)

[3] Geist J and Wang C 1983 New calculations of the quantum yield of silicon in the near ultraviolet *Phys. Rev.* B **27** 4841–7

[4] Scholze F, Henneken H, Kuschnerus P, Rabus H, Richter M and Ulm G 2000 Determination of the electron–hole pair creation energy for semiconductors from the spectral responsivity of photodiodes *Nucl. Instrum. Methods Phys. Res.* A **439** 208–15

[5] Kübarsepp T, Kärhä P and Ikonen E 2000 Interpolation of the spectral responsivity of silicon photodetectors in the near ultraviolet *Appl. Optics* **39** 9–15

[6] Kuschnerus P, Rabus H, Richter M, Scholze F, Werner L and Ulm G 1998 Characterization of photodiodes as transfer detector standards in the 120 nm to 600 nm spectral range *Metrologia* **35** 355–62

[7] Green M A 2008 Self-consistent optical parameters of intrinsic silicon at 300 K including temperature coefficients *Solar Energy Mater. Solar Cells* **92** 1305–10

[8] Shevell S K (ed) 2003 *The Science of Color* 2nd edn (Oxford: Elsevier)

[9] Chantler C, Olsen K, Dragoset R, Chang J, Kishore A, Kotochigova S and Zucker D 2005 *X-Ray Form Factor, Attenuation and Scattering Tables (version 2.1)* (Gaithersburg, MD: National Institute of Standards and Technology) http://physics.nist.gov/ffast

[10] Henke B, Gullikson E and Davis J 1993 X-ray interactions: photoabsorption, scattering, transmission, and reflection at $E = 50–30{,}000$ eV, $Z = 1–92$ *At. Data Nucl. Data Tables* **54** 181–342
[11] Scholze F, Rabus H and Ulm G 1998 Mean energy required to produce an electron–hole pair in silicon for photons of energies between 50 and 1500 eV *J. Appl. Phys.* **84** 2926–39
[12] Lowe B and Sareen R 2007 A measurement of the electron–hole pair creation energy and the Fano factor in silicon for 5.9 keV X-rays and their temperature dependence in the range 80–270 K *Nucl. Instrum. Methods Phys. Res.* A **576** 367–70
[13] Janesick J 2001 *Scientific Charge-Coupled Devices* (Bellingham, WA: SPIE Press)
[14] Tsunemi H, Hiraga J, Yoshita K, Miyata E and Ohtani M 2000 Comparison of methods of measuring the primary charge-cloud shape produced by an x-ray photon inside the CCD *Nucl. Instrum. Methods Phys. Res.* A **439** 592–600
[15] Zula P *et al* 2020 Commonly used radioactive sources *Prog. Theor. Exp. Phys.* 083C01 https://pdg.lbl.gov/
[16] Zula P *et al* 2020 Passage of particles through matter *Prog. Theor. Exp. Phys.* 083C01 https://pdg.lbl.gov/)
[17] Bichsel H 1988 Straggling in thin silicon detectors *Rev. Modern Phys* **60** 663–99
[18] Wang F, Dong S, Nachman B, Garcia-Sciveres M and Zeng Q 2018 The impact of incorporating shell-corrections to energy loss in silicon *Nucl. Instr. Meth. Phys. Res.* A **899** 1–5
[19] Schroder D K 1997 Carrier lifetimes in silicon *IEEE Trans. Electron Devices* **44** 160–70
[20] Shockley W and Read W T 1952 Statistics of the recombinations of holes and electrons *Phys. Rev.* **87** 835–42
[21] Grove A S 1967 *Physics and Technology of Semiconductor Devices* (New York: Wiley)
[22] Richter A, Glunz S W, Werner F, Schmidt J and Cuevas A 2012 Improved quantitative description of Auger recombination in crystalline silicon *Phys. Rev.* B **86** 165202
[23] Law M, Solley E, Liang M and Burk D 1991 Self-consistent model of minority-carrier lifetime, diffusion length, and mobility *IEEE Electron Dev. Lett.* **12** 401–3
[24] Sinton R A and Swanson R M 1987 Recombination in highly injected silicon *IEEE Trans. Electron Devices* **34** 1380–9
[25] Nguyen H T, Baker-Finch S C and Macdonald D 2014 Temperature dependence of the radiative recombination coefficient in crystalline silicon from spectral photoluminescence *Appl. Phys. Lett.* **104** 112105
[26] Hu C C 2009 *Modern Semiconductor Devices for Integrated Circuits* (Upper Englewood Cliffs, NJ: Pearson)
[27] Silvaco Atlas, Silvaco Inc. https://silvaco.com/
[28] Spieler H 2005 *Semiconductor Detector Systems* (Oxford: Oxford University Press)
[29] Allen P E and Holberg D R 2012 *CMOS Analog Circuit Design* (Oxford: Oxford University Press)
[30] Boukhayma A, Peizerat A and Enz C 2016 Temporal readout noise analysis and reduction techniques for low-light CMOS image sensors *IEEE Trans Electron Devices* **63** 72–8

# Full list of references

## Chapter 1

[1] Mazziotta M 2008 Electron–hole pair creation energy and Fano factor temperature dependence in silicon *Nucl. Instrum. Methods Phys. Res.* A **A584** 436–9
[2] Sze S 1981 *Physics of Semiconductor Devices* 2nd edn (New York: Wiley)
[3] Geist J and Wang C 1983 New calculations of the quantum yield of silicon in the near ultraviolet *Phys. Rev.* B **27** 4841–7
[4] Scholze F, Henneken H, Kuschnerus P, Rabus H, Richter M and Ulm G 2000 Determination of the electron–hole pair creation energy for semiconductors from the spectral responsivity of photodiodes *Nucl. Instrum. Methods Phys. Res.* A **439** 208–15
[5] Kübarsepp T, Kärhä P and Ikonen E 2000 Interpolation of the spectral responsivity of silicon photodetectors in the near ultraviolet *Appl. Optics* **39** 9–15
[6] Kuschnerus P, Rabus H, Richter M, Scholze F, Werner L and Ulm G 1998 Characterization of photodiodes as transfer detector standards in the 120 nm to 600 nm spectral range *Metrologia* **35** 355–62
[7] Green M A 2008 Self-consistent optical parameters of intrinsic silicon at 300 K including temperature coefficients *Solar Energy Mater. Solar Cells* **92** 1305–10
[8] Shevell S K (ed) 2003 *The Science of Color* 2nd edn (Oxford: Elsevier)
[9] Chantler C, Olsen K, Dragoset R, Chang J, Kishore A, Kotochigova S and Zucker D 2005 *X-Ray Form Factor, Attenuation and Scattering Tables (version 2.1)* (Gaithersburg, MD: National Institute of Standards and Technology) http://physics.nist.gov/ffast
[10] Henke B, Gullikson E and Davis J 1993 X-ray interactions: photoabsorption, scattering, transmission, and reflection at $E = 50$–30,000 eV, $Z = 1$–92 *At. Data Nucl. Data Tables* **54** 181–342
[11] Scholze F, Rabus H and Ulm G 1998 Mean energy required to produce an electron–hole pair in silicon for photons of energies between 50 and 1500 eV *J. Appl. Phys.* **84** 2926–39
[12] Lowe B and Sareen R 2007 A measurement of the electron–hole pair creation energy and the Fano factor in silicon for 5.9 keV X-rays and their temperature dependence in the range 80–270 K *Nucl. Instrum. Methods Phys. Res.* A **576** 367–70
[13] Janesick J 2001 *Scientific Charge-Coupled Devices* (Bellingham, WA: SPIE Press)
[14] Tsunemi H, Hiraga J, Yoshita K, Miyata E and Ohtani M 2000 Comparison of methods of measuring the primary charge-cloud shape produced by an x-ray photon inside the CCD *Nucl. Instrum. Methods Phys. Res.* A **439** 592–600
[15] Zula P *et al* 2020 Commonly used radioactive sources *Prog. Theor. Exp. Phys.* 083C01 https://pdg.lbl.gov/
[16] Zula P *et al* 2020 Passage of particles through matter *Prog. Theor. Exp. Phys.* 083C01 https://pdg.lbl.gov/)
[17] Bichsel H 1988 Straggling in thin silicon detectors *Rev. Modern Phys* **60** 663–99
[18] Wang F, Dong S, Nachman B, Garcia-Sciveres M and Zeng Q 2018 The impact of incorporating shell-corrections to energy loss in silicon *Nucl. Instr. Meth. Phys. Res.* A **899** 1–5
[19] Schroder D K 1997 Carrier lifetimes in silicon *IEEE Trans. Electron Devices* **44** 160–70
[20] Shockley W and Read W T 1952 Statistics of the recombinations of holes and electrons *Phys. Rev.* **87** 835–42
[21] Grove A S 1967 *Physics and Technology of Semiconductor Devices* (New York: Wiley)

[22] Richter A, Glunz S W, Werner F, Schmidt J and Cuevas A 2012 Improved quantitative description of Auger recombination in crystalline silicon *Phys. Rev.* B **86** 165202
[23] Law M, Solley E, Liang M and Burk D 1991 Self-consistent model of minority-carrier lifetime, diffusion length, and mobility *IEEE Electron Dev. Lett.* **12** 401–3
[24] Sinton R A and Swanson R M 1987 Recombination in highly injected silicon *IEEE Trans. Electron Devices* **34** 1380–9
[25] Nguyen H T, Baker-Finch S C and Macdonald D 2014 Temperature dependence of the radiative recombination coefficient in crystalline silicon from spectral photoluminescence *Appl. Phys. Lett.* **104** 112105
[26] Hu C C 2009 *Modern Semiconductor Devices for Integrated Circuits* (Upper Englewood Cliffs, NJ: Pearson)
[27] Silvaco Atlas, Silvaco Inc. https://silvaco.com/
[28] Spieler H 2005 *Semiconductor Detector Systems* (Oxford: Oxford University Press)
[29] Allen P E and Holberg D R 2012 *CMOS Analog Circuit Design* (Oxford: Oxford University Press)
[30] Boukhayma A, Peizerat A and Enz C 2016 Temporal readout noise analysis and reduction techniques for low-light CMOS image sensors *IEEE Trans Electron Devices* **63** 72–8

## Chapter 2

[1] Fossum E 1995 CMOS image sensors: electronic camera on a chip *Proc. Int. Electron Devices Meeting* 17–25
[2] Fossum E 1997 CMOS image sensors: electronic camera-on-a-chip *IEEE Trans. Electron Devices* **44** 1689–98
[3] Noble P 1968 Self-scanned silicon image detector arrays *IEEE Trans. Electron Devices* **15** 202–9
[4] Weckler G 1967 Operation of p-n junction photodetectors in a photon flux integrating mode *IEEE J. Solid-State Circuits* **2** 65–73
[5] Dyck R and Weckler G 1968 Integrated arrays of silicon photodetectors for image sensing *IEEE Trans. Electron Devices* **15** 196–201
[6] Teranishi N, Kohno A, Ishihara Y, Oda E and Arai K 1982 No image lag photodiode structure in the interline CCD image sensor *1982 Int. Electron Devices Meeting (San Francisco, CA, USA)* 324–7
[7] Teranishi N, Kohno A, Ishihara Y, Oda E and Arai K 1984 An interline CCD image sensor with reduced image lag *IEEE Trans. Electron Devices* **31** 1829–33
[8] Lee P, Gee R, Guidash R, Lee T-H and Fossum E 1995 An active pixel sensor fabricated using CMOS/CCD process technology *1995 IEEE Workshop on Charge-Coupled Devices (Dana Point, CA, USA)*
[9] Kagawa Y and Iwamoto H 2019 3D integration technologies for the stacked CMOS image sensors *Int. 3D Systems Integration Conf. (3DIC) (Sendai)*
[10] Allen P E and Holberg D R 2012 *CMOS Analog Circuit Design* (Oxford: Oxford University Press)
[11] Teranishi N 2016 Analysis of subthreshold current reset noise in image sensors *Sensors* **16** 663
[12] Fowler A M and Gatley I 1991 Noise reduction strategy for hybrid IR focal-plane arrays *Proc. of SPIE 1541 (San Diego)*

[13] Mendis S K *et al* 1997 CMOS active pixel image sensors for highly integrated imaging systems *IEEE J. Solid-State Circuits* **32** 187–97

[14] Teranishi N 2016 Effect and limitation of pinned photodiode *IEEE Trans. Electron Devices* **63** 10–5

[15] Fossum E R and Hondongwa D B 2014 A review of the pinned photodiode for CCD and CMOS image sensors *IEEE J. Electron Devices Soc.* **2** 33–43

[16] Sze S 1981 *Physics of Semiconductor Devices* 2nd edn (New York: Wiley)

[17] Kawai S, Mutoh N and Teranishi N 1997 Thermionic-emission-based barrier height analysis for precise estimation of charge handling capacity in CCD registers *IEEE Trans. Electron Devices* **44** 1588–92

[18] Pelamatti A, Goiffon V, Estribeau M, Cervantes P and Magnan P 2013 Estimation and modeling of the full well capacity in pinned photodiode CMOS image sensors *IEEE Electron Device Lett.* **34** 900–2

[19] Pelamatti A *et al* 2015 Temperature dependence and dynamic behavior of full well capacity in pinned photodiode CMOS image sensors *IEEE Trans. Electron Devices* **62** 1200–7

[20] Khan U and Sarkar M 2018 Dynamic capacitance model of a pinned photodiode in CMOS image sensors *IEEE Trans. Electron Devices* **65** 2892–8

[21] Sarkar M, Buttgen B and Theuwissen A J P 2013 Feedforward effect in standard CMOS pinned photodiodes *IEEE Trans. Electron Devices* **60** 1154–61

[22] Bonjour L, Blanc N and Kayal M 2012 Experimental analysis of lag sources in pinned photodiodes *IEEE Electron Device Lett.* **33** 1735–7

[23] Rizzolo S, Goiffon V, Estribeau M, Marcelot O, Martin-Gonthier P and Magnan P 2018 Influence of pixel design on charge transfer performances in CMOS image sensors *IEEE Trans. Electron Devices* **65** 1048–55

[24] Han L, Yao S and Theuwissen A J P 2016 A charge transfer model for CMOS image sensors *IEEE Trans. Electron Devices* **63** 32–41

[25] Xu J, Wang R, Han L and Gao Z 2020 Analysis and modeling of spill back effect in high illumination CMOS image sensors *IEEE Sensors J.* **20** 3024–31

[26] Cao Z, Zhou Y, Li Q, Qin Q, Liu L and Wu N 2013 Design of pixel for high speed CMOS image sensors *Int. Image Sensor Workshop* 7.11

[27] Shin B, Park S and Shin H 2010 The effect of photodiode shape on charge transfer in CMOS image sensors *Solid-State Electron.* **54** 1416–20

[28] Miyauchi K *et al* 2014 Pixel structure with 10 nsec fully charge transfer time for the 20 m frame per second burst CMOS image sensor *Proc. SPIE* 9022

[29] Cao X *et al* 2015 Design and optimisation of large 4T pixel *Int. Image Sensor Workshop*

[30] Janesick J, Andrews J, Tower J, Grygon M, Elliott T, Cheng J, Lesser M and Pinter J 2007 Fundamental performance differences between CMOS and CCD imagers: part II *Proc. of SPIE (San Diego, CA)*

[31] Wang X *et al* 2010 A 2.2M CMOS image sensor for high-speed machine vision applications *Proc. of SPIE 7536 (San Jose)*

[32] Miyauchi K, Mori K, Otaka T, Isozaki T, Yasuda N, Tsai A, Sawai Y, Owada H, Takayanagi I and Nakamura J 2020 A stacked back side-illuminated voltage domain global shutter CMOS image sensor with a 4.0 μm multiple gain readout pixel *MDPI Sensors* **20** 486

[33] Velichko S *et al* 2016 CMOS global shutter charge storage pixels with improved performance *IEEE Trans. Electron Devices* **63** 106–12

[34] Yasutomi K, Itoh S and Kawahito S 2011 A two-stage charge transfer active pixel CMOS image sensor with low-noise global shuttering and a dual-shuttering mode *IEEE Trans. Electron Devices* **58** 740–7

[35] Boukhayma A, Peizerat A and Enz C 2016 A sub-0.5 electron read noise VGA image sensor in a standard CMOS process *IEEE J. Solid-State Circuits* **51** 2180–91

[36] Seo M, Kawahito S, Kagawa K and Yasutomi K 2015 A 0.27e-RMS read noise 220-μV/e-conversion gain reset-gate-less CMOS image sensor with 0.11-μm CIS process *IEEE Electron Device Lett.* **36** 1344–7

[37] Vu P *et al* 2011 Low noise high dynamic range 2.3Mpixel CMOS image sensor capable of 100 Hz frame rate at full HD resolution *Int. Image Sensor Workshop (Hokkaido)*

[38] CMOS sensors (https://fairchildimaging.com/products/scmos-sensors)

[39] Solhusvik J *et al* 2013 A comparison of high dynamic range CIS technologies for automotive applications *Int. Image Sensor Workshop (Snowbird, UT)*

[40] Kabir S, Guidash M, Vogelsang T, Smith C, Schneider A and Endsley J 2017 A small pixel high performance full frame HDR sensor *Int. Image Sensor Workshop (Hiroshima)*

[41] Willassen T *et al* 2015 A 1280×1080 4.2 μm split-diode pixel HDR sensor in 110 nm BSI CMOS process *Int. Image Sensor Workshop (Vaals, The Netherlands)*

[42] Ma C, Liu Y, Li Y, Zhou Q, Wang X and Chang Y 2017 A 4-M pixel high dynamic range, low-noise CMOS image sensor with low-power counting ADC *IEEE Trans. Electron Devices* **64** 3199–205

[43] Tanaka S *et al* 2018 Single exposure type wide dynamic range CMOS image sensor with enhanced NIR sensitivity *ITE Trans. MTA* **6** 195–201

[44] Solhusvik J *et al* 2017 A 1392×976 2.8 μm 120 dB CIS with per-pixel controlled conversion gain *Int. Image Sensor Workshop (Hiroshima)*

[45] Sugawa S, Akahane N, Adachi S, Mori K, Ishiuchi T and Mizobuchi K 2005 A 100 dB dynamic range CMOS image sensor using a lateral overflow integration capacitor *IEEE Int. Solid-State Circuits Conf.*

[46] Akahane N, Sugawa S, Adachi S, Ishiuchi M K T and Mizobuchi K 2006 A sensitivity and linearity improvement of a 100-db dynamic range CMOS image sensor using a lateral overflow integration capacitor *IEEE J. Solid-State Circuits* **41** 851–8

[47] Fujihara Y, Murata M, Nakayama S, Kuroda R and Sugawa S 2021 An over 120 dB single exposure wide dynamic range CMOS image sensor with two-stage lateral overflow integration capacitor *IEEE Trans. Electron Devices* **68** 152–7

[48] Lalanne F, Malinge P, Hérault D and Jamin-Mornet C 2017 A native HDR 115 dB 3.2 μm BSI pixel using electron and hole collection *Int. Image Sensor Workshop (Hiroshima)*

[49] Bai Y *et al* 2012 4K×4K format 10-micron pixel pitch H4RG-10 hybrid CMOS silicon visible focal plane array for space astronomy *Proc. of SPIE 84530M (Amsterdam)*

[50] Garnett J *et al* 2004 2K×2K molecular beam epitaxy HgCdTe detectors for the James Webb Space Telescope NIRCam instrument *Proc. of SPIE 5499 (Glasgow)*

[51] Bai Y *et al* 2008 Teledyne imaging sensors: silicon CMOS imaging technologies for x-ray, UV, visible, and near infrared *Proc. of SPIE 702 102 (Marseille)*

[52] Douence V *et al* 2005 Hybrid image sensor with multiple on-chip frame storage for ultrahigh-speed imaging *Proc. of SPIE 5580 (Alexandria, VA)*

[53] Gove R J 2020 CMOS image sensor technology advances for mobile devices *High Performance Silicon Imaging* ed D Durini 2nd edn (Cambridge: Woodhead Publishing) pp 185–240

[54] Suntharalingam V *et al* 2005 Megapixel CMOS image sensor fabricated in three-dimensional integrated circuit technology *2005 IEEE Int. Digest of Technical Papers. Solid-State Circuits Conf.*

[55] Kwon M *et al* 2020 A low-power 65/14 nm stacked CMOS image sensor *2020 IEEE Int. Symp. on Circuits and Systems (ISCAS)*

[56] Miura T *et al* 2019 A 6.9 μm pixel-pitch 3D stacked global shutter CMOS image sensor with 3M Cu-Cu connections *2019 Int. 3D Systems Integration Conf. (3DIC) (Sendai)*


## Chapter 3


[1] Sze S 1981 *Physics of Semiconductor Devices* 2nd edn (New York: Wiley)

[2] Richter A, Glunz S W, Werner F, Schmidt J and Cuevas A 2012 Improved quantitative description of Auger recombination in crystalline silicon *Phys. Rev.* B **86** 165202

[3] Janesick J, Andrews J, Tower J, Grygon M, Elliott T, Cheng J, Lesser M and Pinter J 2007 Fundamental performance differences between CMOS and CCD imagers: Part II *Proc. of SPIE, 669003 (San Diego)*

[4] MacKenty J 2012 Performance and calibration of the HST wide field camera 3 *Proc. of SPIE, 84421V (Amsterdam)*

[5] Andor Neo 5.5 sCMOS specifications Andor https://andor.oxinst.com/assets/uploads/products/andor/documents/andor-neo-scmos-specifications.pdf)

[6] Pichler P and Point I 2012 *Defects, Impurities, and Their Diffusion in Silicon* (Vienna: Springer)

[7] McGrath D, Tobin S, Goiffon V, Magnan P and Le Roch A 2018 Dark current limiting mechanisms in CMOS image sensors *IS&T Int. Symp. on Electronic Imaging 2018, Image Sensors and Imaging Systems (Burlingame, CA)*

[8] Brunetti A M, Musolino M, Strangio S and Choubey B 2020 Pixel design driven performance improvement in 4T CMOS image sensors: dark current reduction and full-well enhancement *IEEE Trans. Electron Devices* **67** 409–12

[9] Seo M, Kawahito S, Yasutomi K, Kagawa K and Teranishi N 2014 A low dark leakage current high-sensitivity CMOS image sensor with STI-less shared pixel design *IEEE Trans. Electron Devices* **61** 2093–7

[10] Grove A S 1967 *Physics and Technology of Semiconductor Devices* (New York: Wiley)

[11] Janesick J 2001 *Scientific Charge-Coupled Devices* (Bellingham, WA: SPIE Press)

[12] Gradshteyn I S and Ryzhik I M 2007 *Table of Integrals, Series, and Products* 7th edn (Amsterdam: Elsevier)

[13] Teranishi N 2016 Effect and limitation of pinned photodiode *IEEE Trans. Electron Devices* **63** 10–5

[14] Snoeys W 2013 Monolithic pixel detectors for high energy physics *Nucl. Instr. Meth. Phys. Res.* A **731** 125–30

[15] Lahav A, Fenigstein A, Strumb A and Rizzolo S 2020 Backside illuminated (BSI) complementary metal-oxide semiconductor (CMOS) image sensors *High Performance Silicon Imaging* ed D Durini 2nd edn (Cambridge: Woodhead Publishing) pp 95–117

[16] Nikzad S *et al* 2017 High-efficiency UV/optical/NIR detectors for large aperture telescopes and UV explorer missions: development of and field observations with delta-doped arrays *J. Astron. Telesc. Instrum. Syst.* **3** 036002

[17] Green M A 2008 Self-consistent optical parameters of intrinsic silicon at 300 K including temperature coefficients *Solar Energy Mater. Solar Cells* **92** 1305–10

[18] Law M, Solley E, Liang M and Burk D 1991 Self-consistent model of minority-carrier lifetime, diffusion length, and mobility *IEEE Electron. Dev. Lett.* **12** 401–3
[19] Hoenk M *et al* 2009 Delta-doped back-illuminated CMOS imaging arrays: progress and prospects *Proc. of SPIE 74190T (San Diego)*
[20] Hoex B *et al* 2008 On the c-Si surface passivation mechanism by the negative-charge-dielectric $Al_2O_3$ *J. Appl. Phys.*, **104** 113703
[21] Shin B, Park S and Shin H 2010 The effect of photodiode shape on charge transfer in CMOS image sensors *Solid-State Electron.* **54** 1416–20
[22] Cao X *et al* 2015 Design and optimisation of large 4T pixel *Int. Image Sensor Workshop (Vaals, The Netherlands)*
[23] Kalgi A K, Crouwels A, Dierickx B, Verbruggen W and Aken D V 2019 Fast charge transfer in 100µm long PPD pixels *Int. Image Sensor Workshop (Snowbird, UT)*
[24] Miyauchi K *et al* 2014 Pixel structure with 10 nsec fully charge transfer time for the 20M frame per second burst CMOS image sensor *Proc. of SPIE, 902203 (San Francisco, CA)*
[25] Reich R K *et al* 1993 Integrated electronic shutter for back-illuminated charge-coupled devices *IEEE Trans. Electron Devices* **40** 1231–7
[26] Juntunen M A, Heinonen J, Vähänissi V, Repo P, V D and Savin H 2016 Near-unity quantum efficiency of broadband black silicon photodiodes with an induced junction *Nat. Photon.* **10** 777–81
[27] Goiffon V, Cervantes P, Virmontois C, Corbière F, Magnan P and Estribeau M 2011 Generic radiation hardened photodiode layouts for deep submicron CMOS image sensor processes *IEEE Trans. Nucl. Sci.* **58** 3076–84

## Chapter 4

[1] Horowitz P and Hill W 2015 *The Art of Electronics* 3rd edn (New York: Cambridge University Press)
[2] Fano U 1947 Ionization yield of radiations. II. The fluctuations of the number of ions *Phys. Rev.* **71** 26–9
[3] Janesick J 2007 *Photon Transfer DN λ* (Bellingham, WA: SPIE Press)
[4] Caloyannides M 1974 Microcycle spectral estimates of 1/*f* noise in semiconductors *J. Appl. Phys.* **45** 307
[5] Vandamme L K J and Hooge F N 2008 What do we certainly know about 1/*f* noise in MOSTs? *IEEE Trans. Electron Devices* **55** 3070–85
[6] Chang J, Abidi A A and Viswanathan C R 1994 Flicker noise in CMOS transistors from subthreshold to strong inversion at various temperatures *IEEE Trans. Electron Devices* **41** 1965–71
[7] Janesick J, Andrews J and Elliott T 2006 Fundamental performance differences between CMOS and CCD imagers: part I *Proc. SPIE 6276 (Orlando, FL)*
[8] Leyris C, Martinez F, Valenza M, Hoffmann A, Vildeuil J C and Roy F 2006 Impact of random telegraph signal in CMOS image sensors for low-light levels *2006 Proc. of the 32nd European Solid-State Circuits Conf. (Montreux, Switzerland)*
[9] Hopkinson G and Lumb D 1982 Noise reduction techniques for CCD image sensors *J. Phys. E: Sci. Instrum.* **15** 1214–22
[10] Jordan A and Jordan N 1965 Theory of noise in metal oxide semiconductor devices *IEEE Trans. Electron Devices* **12** 148–56

[11] Nemirovsky Y, Corcos D, Brouk I, Nemirovsky A and Chaudhry S 2011 1/*f* Noise in advanced CMOS transistors *IEEE Instrum. Meas. Mag.* **14** 14–22

[12] Yadid-Pecht O, Mansoorian K, Fossum E and Pain B 1997 Optimization of noise and responsivity in CMOS active pixel sensors for detection of ultra low light levels *Proc. SPIE* **3019** 125–36

[13] Burt D 1991 CCD performance limitations: theory and practice *Nucl. Instr. Meth.* A **305** 564–73

[14] Kansy R 1980 Response of a correlated double sampling circuit to 1/*f* noise *IEEE J. Solid-State Circuits* **15** 373–5

[15] Kawahito S 2011 Architectures for low-noise CMOS electronic imaging *Single-Photon Imaging* ed P Seitz and A Theuwissen (Berlin: Springer ) pp 197–217

[16] Janesick J 2001 *Scientific Charge-Coupled Devices* (Bellingham, WA: SPIE Press)

[17] Hegyi D and Burrows A 1980 Optimal sampling of charge coupled devices *Astron. J.* **85** 1421–4

[18] Levanon N and Mozeson E 2004 Matched filter *Radar Signals* (New York: Wiley) pp 20–33

[19] Stefanov K 2015 Digital CDS for image sensors with dominant white and 1/*f* noise *J. Instrum.* **10** 04003

[20] Wey H M and Guggenbühl W 1990 An improved correlated double sampling circuit for low noise charge-coupled devices *IEEE Trans. Circuits Syst.* **37** 1559–65

[21] Kleinpenning T and de Kuijper A 1988 Relation between variance and sample duration of 1/*f* noise signals *J. Appl. Phys.* **63** 43–5

[22] Kawai N and Kawahito S 2005 Effectiveness of a correlated multiple sampling differential averager for reducing 1/*f* noise *IEICE Electron. Express* **2** 379–83

[23] Kawahito S and Seo M-W 2016 Noise reduction effect of multiple-sampling-based signal-readout circuits for ultra-low noise CMOS image sensors *Sensors* **16** 1867

[24] Stefanov K and Murray N 2014 Optimal digital correlated double sampling for CCD signals *IET Electron. Letters* **50** 1022–4

[25] Gach J-L, Darson D, Guillaume C, Goillandeau M, Cavadore C, Balard P, Boissin O and Boulesteix J 2003 A new digital CCD readout technique for ultra-low-noise CCDs *Publ. Astron. Soc. Pacific* **115** 1068–71

[26] Fowler B 2011 Single photon CMOS imaging through noise minimization *Single-Photon Imaging* ed P Seitz and A Theuwissen (Berlin: Springer) pp 159–95

[27] Boukhayma A, Peizerat A and Enz C 2016 Noise reduction techniques and scaling effects towards photon counting CMOS image sensors *Sensors* **16** 514

[28] Boukhayma A, Peizerat A and Enz C 2016 Temporal readout noise analysis and reduction techniques for low-light CMOS image sensors *IEEE Trans. Electron Devices* **63** 72–8

## Chapter 5

[1] EMVA Standard 1288 Standard for Characterization of Image Sensors and Cameras, European Machine Vision Association, 15 March 2021 (www.emva.org)

[2] International Organization for Standardization (https://www.iso.org/home.html)

[3] Jacquot B C, Monacos S P, Hoenk M E, Greer F, Jones T J and Nikzad S 2011 A system and methodologies for absolute quantum efficiency measurements from the vacuum ultra-violet through the near infrared *Rev. Sci. Instrum.* **82** 043102

[4] Spivak A, Belenky A, Fish A and Yadid-Pecht O 2009 Wide-dynamic-range CMOS imagesensors—comparative performance analysis *IEEE Trans. Electron Devices* **56** 2446–61

[5] Fernandez F, Steward B, Gross K and Hawks M 2019 Implementation of a non-linear CMOS and CCD focal plane array model in ASSET *Proceedings of SPIE, 110010C (Baltimore)*
[6] Janesick J 2007 *Photon Transfer DN → λ* (Bellingham, WA: SPIE Press)
[7] Wang F and Theuwissen A J P 2017 Linearity analysis of a CMOS image sensor IS&T Inter. Symp. on Electronic Imaging *(San Francisco)*
[8] Bonjour L, Blanc N and Kayal M 2012 Experimental analysis of lag sources in pinned photodiodes *IEEE Electron Device Lett.* **33** 1735–7
[9] Soman M, Stefanov K, Weatherill D, Holland A, Gow J and Leese M 2015 Non-linear responsivity characterisation of a CMOS active pixel sensor for high resolution imaging of the Jovian system *J. Instrum.* **10** C02012
[10] Berendsen H J C 2011 *A Student's Guide to Data and Error Analysis* (Cambridge: Cambridge University Press)
[11] Levski D, Wäny M and Choubey B 2021 Compensation of signal-dependent readout noise in photon transfer curve characterisation of CMOS image sensors *IEEE Trans. Circuits Syst. II: Express Br.* **68** 102–5
[12] Okura S *et al* 2019 A 2-Mpixel CMOS image sensor with device authentication and encryption key generation based on physically unclonable function *Int. Image Sensor Workshop (Snowbird, UT)*
[13] Janesick J 2001 *Scientific Charge-Coupled Devices* (Bellingham, WA: SPIE Press)
[14] Downing M, Baade D, Sinclaire P, Deiries S and Christen F 2006 CCD riddle: (a) signal vs time: linear; (b) signal vs variance: non-linear *Proc. of the SPIE 627609 (Orlando, FL)*
[15] Stefanov K D 2013 A statistical model for signal-dependent charge sharing in image sensors *IEEE Trans. Electron Devices* **61** 110–5
[16] Pain B and Hancock B R 2003 Accurate estimation of conversion gain and quantum efficiency in CMOS imagers *Proc. of SPIE 5017 (Santa Clara, CA)*
[17] Bohndiek S E *et al* 2008 Comparison of methods for estimating the conversion gain of CMOS active pixel sensors *IEEE Sensors J.* **8** 1734–44
[18] Janesick J, Andrews J and Elliott T 2006 Fundamental performance differences between CMOS and CCD imagers: part I *Proc. of SPIE 6276, 62760M (Orlando, FL)*
[19] Michelot J, de Ipanema Moreira A, Monsinjon P and Caranhac S 2016 Effects of transfer gate spill back in low light high performances CMOS image sensors *Photon Counting, Low Flux and High Dynamic Range Optoelectronic Detectors Workshop (Toulouse)*
[20] Ivory J, Stefanov K D and Holland A D 2020 Mitigating charge spill-back induced image lag with a multi-level transfer gate pulse in PPD image sensors *Proc. of SPIE 11454*
[21] Lowe B and Sareen R 2007 A measurement of the electron–hole pair creation energy and the Fano factor in silicon for 5.9 keV X-rays and their temperature dependence in the range 80–270 K *Nuclear Instrum. Methods Phys. Res.* A **576** 367–70
[22] Widenhorn R, Dunlap J C and B E 2010 Exposure time dependence of dark current in CCD imagers *IEEE Trans. Electron Devices* **57** 581–7
[23] McGrath D, Tobin S, Goiffon V, Magnan P and Le Roch A 2018 Dark current limiting mechanisms in CMOS image sensors IS&T *Int. Symp. on Electronic Imaging 2018, Image Sensors and Imaging Systems (Burlingame, CA)*
[24] Teranishi N 2013 Dark current and white blemish in image sensors *2013 Int. Symp. on VLSI Technology, Systems and Application (VLSI-TSA) (Hsinchu, Taiwan)*
[25] Chao C Y-P *et al* 2019 Identifying the sources of random telegraph noises in pixels of CMOS image sensors *Int. Image Sensor Workshop (Snowbird, UT)*

[26] Kwon H-M *et al* 2013 Effects of high-pressure annealing on random telegraph signal noise characteristic of source follower block in CMOS image sensor *IEEE Electron Device Lett.* **34** 190–2
[27] Gao W, Guidash M, Li N, Ispasoiu R, Ailuri P R, Palaniappan N, Tekleab D and Rahman M 2017 Photodiode barrier induced lag characterization using a new lag versus idle time methodology *Int. Image Sensor Workshop (Hiroshima, Japan )*
[28] Crews C, Soman M, Allanwood E A, Stefanov K, Leese M, Turner P and Holland A 2020 Quantum efficiency of the CIS115 in a radiation environment *Proc. of SPIE, 114540E*
[29] Janesick J, Andrews J, Tower J, Grygon M, Elliott T, Cheng J, Lesser M and Pinter J 2007 Fundamental performance differences between CMOS and CCD imagers: part II *Proc. of SPIE 6690, 669003 (San Diego, CA)*
[30] Boreman G D 2021 Point-, line-, and edge-spread function measurement of MTF *Modulation Transfer Function in Optical and Electro-Optical Systems* 2nd edn (Bellingham, WA: SPIE Press) pp 67–84
[31] Zhang X, Kashti T, Kella D, Frank T, Shaked D, Ulichney R, Fischer M and Allebach J P 2012 Measuring the modulation transfer function of image capture devices: what do the numbers really mean? *Proc. of SPIE 8293 (829307) (Burlingame, CA)*
[32] Janesick J, Pinter J, Potter R, Elliott T, Andrews J, Tower J, Grygon M and Keller D 2010 Fundamental performance differences between CMOS and CCD imagers: part IV *Proc. of SPIE 7742, 77420B (San Diego, CA)*
[33] Estribeau M and Magnan P 2004 Fast MTF measurement of CMOS imagers at the chip level using ISO 12233 slanted-edge methodology *Proc. of SPIE 5570 (Maspalomas)*

## Chapter 6

[1] Fowler B 2011 Single photon CMOS imaging through noise minimization *Single-Photon Imaging* ed P Seitz and A Theuwissen (Berlin: Springer) pp 159–95
[2] Boukhayma A, Peizerat A and Enz C 2016 A sub-0.5 electron read noise VGA image sensor in a standard CMOS process *IEEE J. Solid-State Circuits* **51** 2180–91
[3] Ivory J, Stefanov K D and Holland A D 2020 Mitigating charge spill-back induced image lag with a multi-level transfer gate pulse in PPD image sensors *Proc. of SPIE, 11454*
[4] Kaeslin H 2008 *Digital Integrated Circuit Design, From VLSI Architectures to CMOS Fabrication* (Cambridge: Cambridge University Press)
[5] Wakerly J F 2001 *Digital Design: Principles and Practices* 3rd edn (Englewood Cliffs, NJ: Prentice-Hall)
[6] Allen P E and Holberg D R 2012 *CMOS Analog Circuit Design* (Oxford: Oxford University Press)
[7] Horowitz P and Hill W 2015 *The Art of Electronics* 3rd edn (New York: Cambridge University Press)
[8] Jung W 2005 *Op Amp Applications Handbook* (Oxford: Newnes)
[9] AN-1515 A Comprehensive Study of the Howland Current Pump, Texas Instruments (https://www.ti.com/lit/an/snoa474a/snoa474a.pdf)
[10] Jefferts S R 1989 A very low-noise FET input amplifier *Rev. Sci. Instrum.* **60** 1194–6
[11] Shen S and Yuan J 2020 1/fγ low frequency noise model for buried channel MOSFET *IEEE J. Electron Devices Soc.* **8** 126–33