Electronic design issues in high-bandwidth parallel optical interfaces to VLSI circuits

Mark Graham Forbes

Submitted for the degree of Doctor of Philosophy

Heriot-Watt University Department of Physics

March 1999

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that the copyright rests with its author and that no information derived from it may be published without the prior written consent of the author or the University (as may be appropriate).

Table of contents

| Acknowledgements | vii |
|--|------|
| Abstract | viii |
| List of publications | ix |
| | |
| Chapter 1: Introduction | 1 |
| 1.1 Scope and overall research contribution | 1 |
| 1.2 Motivation | 2 |
| 1.2.1 The interconnect problem | 2 |
| 1.2.2 Capabilities and limitations of electrical interconnects | 4 |
| 1.2.3 Advantages of optical interconnects | 9 |
| 1.2.4 Summary | 10 |
| 1.3 Thesis organisation and specific research contributions | 10 |
| | |
| Chapter 2: Optical interface technology | 12 |
| 2.1 Introduction | 12 |
| 2.2 Monolithic integration | 13 |
| 2.3 Hybrid integration technology | 15 |
| 2.3.1 Bonding technologies | 15 |
| 2.3.2 Regrowth technologies | 17 |
| 2.4 Optoelectronic devices | |
| 2.4.1 MQW modulators | |
| 2.4.2 VCSELs | 25 |
| 2.5 Optical packaging | 27 |
| 2.5.1 Fibre-ribbon interconnects | 27 |
| 2.5.2 Free-space interconnects | |
| 2.5.3 Example of a system using free-space interconnects | 29 |
| 2.6 Conclusion | 32 |
| | |
| Chapter 3: Design of a free-space optoelectronic crossbar | |
| 3.1 Introduction | |
| 3.2 System description | |
| 3.2.1 Overview of system architecture | |
| 3.2.2 Optical implementation | |
| 3.3 Implementation details | |
| 3.3.1 Behavioural description | |

| 3.3.2 Logical implementation | 41 |
|---|----|
| 3.3.3 Timing conventions | 43 |
| 3.3.4 Physical implementation of the switching chip | 44 |
| 3.4 Other system components | 49 |
| 3.4.1 InGaAs detectors and modulators | 49 |
| 3.4.2 Optical power budget | |
| 3.4.3 Optical components | 53 |
| 3.5 Progress towards experimental realisation | 54 |
| 3.6 Conclusion | 56 |
| 3.7 Appendix: scheme for handshaking with input queues | 57 |
| Chapter 4: Design trade-offs in photoreceivers for terabit/s scale optical interfaces | 58 |
| 4.1 Introduction | 58 |
| 4.2 Review of receiver design | 59 |
| 4.2.1 Structure of a conventional telecommunications receiver | 59 |
| 4.2.2 Smart-pixel implementation | 60 |
| 4.2.3 Synchronous sense-amplifier receivers | 63 |
| 4.3 Front-end small signal performance | 64 |
| 4.3.1 Small-signal analysis | 64 |
| 4.3.2 Effect of the feedback resistor | 66 |
| 4.3.3 Effect of the front-end amplifier | 68 |
| 4.3.4 Damping factor | 70 |
| 4.3.5 Calculations for example parameters | 71 |
| 4.3.6 Effect of photodiode capacitance | 73 |
| 4.4 Noise limits on receiver sensitivity | 73 |
| 4.5 Post-amplifier gain limits on sensitivity | 76 |
| 4.5.1 Introduction | 76 |
| 4.5.2 Small signal analysis | 78 |
| 4.5.3 Design variables | 79 |
| 4.5.4 Design trade-offs | 79 |
| 4.5.5 Simulation | 80 |
| 4.5.6 Multistage amplifier designs | 81 |
| 4.5.7 Performance together with front-end | 82 |
| 4.6 MOSFET mismatch limits on receiver sensitivity | |
| 4.6.1 Introduction | 84 |
| 4.6.2 Physical origin of offsets | 86 |
| 4.6.3 An analytical model for the offset limit on receiver sensitivity | 90 |
| | |

| 4.6.4 Application of the offset model to a 0.7 μm process | |
|---|-----|
| 4.6.5 Possible solutions to the offset problem | |
| 4.7 Conclusion | 98 |
| 4.8 Appendix: modelling of feedback resistor | 99 |
| 4.9 Appendix: behaviour of decision stage | 100 |
| | |
| Chapter 5: Case study: a single-beam receiver for SPOEC | 102 |
| 5.1 Introduction | 102 |
| 5.2 System requirements | 102 |
| 5.3 Detailed design considerations | 104 |
| 5.3.1 Circuit schematic | 104 |
| 5.3.2 Summary of simulated performance | |
| 5.3.3 Power supply distribution | 105 |
| 5.3.4 Matching considerations | 106 |
| 5.3.5 Feedback resistor | 107 |
| 5.3.6 Disable mechanism | 108 |
| 5.3.7 Circuit layout | 108 |
| 5.4 Experimental performance | 111 |
| 5.4.1 Test structures | 111 |
| 5.4.2 Results | 113 |
| 5.5 Conclusion | 117 |
| | |
| Chapter 6: Scaling of receiver performance in advanced CMOS technology | 119 |
| 6.1 Introduction | 119 |
| 6.2 Scaling of basic transistor characteristics | |
| 6.2.1 Ideal scaling | 120 |
| 6.2.2 Non-ideal scaling | 121 |
| 6.2.3 Scaling of offset voltage | 126 |
| 6.2.4 Scaling of capacitance | 128 |
| 6.2.5 Scaling of inverter gain | 129 |
| 6.3 Impact on receiver performance | |
| 6.3.1 General approach | 130 |
| 6.3.2 Scaling of front-end dimensions | 131 |
| 6.3.3 Trends in switching energy | 132 |
| 6.3.4 Trends in power consumption | 134 |
| 6.3.5 Changes to the scaled design to exploit the reduced power consumption | 135 |
| 6.3.6 Limit on scaling due to power supply distribution | |

| 6.3.7 Limitations of this analysis | 138 |
|---|-----|
| 6.3.8 Comparison with other studies | 139 |
| 6.4 Conclusions | 140 |
| 6.5 Appendix: Mobility degradation model | 141 |
| Chapter 7: Transconductance-transimpedance post-amplifiers for smart-pixel receivers. | 142 |
| 7.1 Introduction | 142 |
| 7.2 Description of circuit technique | 143 |
| 7.3 Small-signal analysis of the transconductance-transimpedance circuit | 144 |
| 7.4 Detailed comparison with low-gain voltage amplifier | 148 |
| 7.4.1 Introduction | 148 |
| 7.4.2 Methodology | 149 |
| 7.4.3 Results and discussion | 151 |
| 7.4.4 Small-signal results | 155 |
| 7.4.5 Conclusion | 156 |
| 7.5 Offset performance of the transconductance-transimpedance cascade | 156 |
| 7.6 Application to a differential receiver circuit | 158 |
| 7.6.1 Introduction | 158 |
| 7.6.2 Application requirements | 158 |
| 7.6.3 General design approach | 159 |
| 7.6.4 Detailed design | 161 |
| 7.6.5 Simulation results | 167 |
| 7.6.6 Experimental results | 168 |
| 7.6.7 Evaluation of the design | 177 |
| 7.7 Conclusion | 178 |
| Chapter 8: Electrical crosstalk in large photoreceiver arrays | 179 |
| 8.1 Introduction | 179 |
| 8.2 Analysis of supply sensitivity | 180 |
| 8.2.1 Basic approach | 180 |
| 8.2.2 Front-end | 181 |
| 8.2.3 Second-stage | 185 |
| 8.2.4 Decision stage | 186 |
| 8.2.5 Digital stage | 186 |
| 8.2.6 Optimum partitioning of power supplies | 187 |
| 8.3 Estimation of voltage noise | 187 |
| 8.3.1 Introduction | 187 |

| 8.3.2 Equivalent circuit of a receiver cell | |
|---|--------|
| 8.3.3 Modelling of receiver array power supply network | 191 |
| 8.3.4 Modelling of package impedance | 194 |
| 8.3.5 Combining on-chip admittance matrix and off-chip impedance matrix | |
| 8.3.6 Estimation of IR drop | 201 |
| 8.4 Numerical estimates of supply crosstalk | 202 |
| 8.4.1 Introduction | 202 |
| 8.4.2 Noise immunity of receiver circuit | 202 |
| 8.4.3 Simulation results | 203 |
| 8.5 Discussion | 208 |
| 8.6 Conclusions and further work | 212 |
| 8.7 Appendix: Norton equivalent of a distributed line with a distributed current sour | rce214 |
| 8.8 Appendix: calculation of package impedance matrix | 216 |
| 8.9 Appendix: effect of a single admittance term on an impedance matrix | 218 |
| | |
| Chapter 9: Conclusions | 221 |
| 9.1 Introduction | 221 |
| 9.2 Design of a prototype terabit/s scale system | 221 |
| 9.3 Investigation of receiver design issues | |
| 9.4 Future work | 223 |
| 9.5 Closing remarks | 224 |
| | |
| Appendix A: Detailed simulations of SPOEC data receiver | 225 |
| Appendix B: Detailed simulations of SPOEC clock receiver | |
| | |

| References |
|------------|
|------------|

Acknowledgements

I've had the good fortune to carry out this work in a friendly research group made up of people eager to share their wide ranging expertise.

Particular thanks are due to my supervisor, Professor Andy Walker, for keeping me on track, for getting me to ask myself the right questions and for not letting me lose sight of what research is all about. I really appreciate having had the opportunity to define my own path of work. Dr. Frank Tooley also provided key guidance at certain points.

The research of which this work forms a part could only have been carried out as a team effort, and I'd like to thank all my colleagues on the SPOEC project, especially those with whom I worked most closely: at Heriot-Watt, Marc Desmulliez and Stuart Fancey and at Supélec Paris, Alain Gauthier and Philippe Benabes.

The opportunity for constant informal discussion, over coffee and elsewhere, was and is one of the strengths of the Heriot-Watt research group, and this work would have been impossible without this environment. Discussions with Julian Dines, David Neilson, James Gourlay, Tsung-Yi Yang and Alex Fritze on subjects ranging from the details of computer architecture to the philosophy behind what we were doing were invaluable. Doug Baillie's assistance in the lab is also acknowledged, and none of the experimental work would have been possible without the technical genius of George Smith. Thanks also to Lucy Wilkinson, Simon Prince and Neil Ross.

Finally, I want to thank my flatmates, in particular Yeong, Zak, Sandra and Ian, for putting up with me over the last three and a bit years and for reminding me of the existence of a world outside the Physics department.

Abstract

This thesis investigates electronic design issues in systems using parallel optical interconnections to provide terabit/s scale I/O to VLSI circuits. It focuses on the design of arrays of transimpedance photoreceiver circuits in the context of optically interconnected digital switching systems.

The trade-offs in photoreceiver design are discussed. Transistor mismatch is shown to limit receiver sensitivity. Trends in performance in future CMOS technology are projected. By the 0.1 µm generation, the analysis forecasts a major improvement in electrical power consumption but, due to transistor mismatch, only a limited improvement in optical switching energy.

The design of the receiver subsystem for a prototype optoelectronic-VLSI switching system, implemented in a hybrid 0.6 µm CMOS-InGaAs MQW modulator technology, is used to provide several case studies. The system is designed to implement a 1 Tbit/s optical interface using 4096 optical channels operating at 250 Mbit/s. Two receivers for this system are described. Experimental results from prototypes tested with electrical inputs verify that the designs meet the DC sensitivity requirements of the system. One of the designs applies the transconductance-transimpedance circuit technique to this application area for the first time. A detailed study of the merits of this approach in high bit-rate post-amplifiers shows that its high gain-bandwidth product can improve sensitivity at some cost in power consumption and layout area.

A method for analysing electrical crosstalk through the power supply network in twodimensional receiver arrays is presented. A case study shows that crosstalk has a major impact on the performance of single-ended receivers suggesting that electrically differential designs may be advantageous.

The work concludes that, although certain challenges remain, electronic design issues are unlikely to prevent parallel optical interconnects from meeting the I/O requirements of future VLSI circuits.

List of publications

The work described in this thesis has been reported in the following journal articles and conference proceedings:

Journal articles

- FORBES, M.G., WALKER, A.C.: 'Wideband transconductance-transimpedance postamplifier for large photoreceiver arrays', *Electron. Lett.*, 19 March 1988, **34** (6), pp. 589-590
- WALKER, A.C., DESMULLIEZ, M.P.Y., FORBES, M.G., FANCEY, S.J., BULLER, G.S., TAGHIZADEH, M.R., DINES, J.A.B., STANLEY, C.R., PENNELLI, G., HORAN, P., BYRNE, D., HEGARTY, J., EITEL, S., GAUGGEL, H.-P., GULDEN, K.-H., GAUTHIER, A., BENABES, P., GUTZWILLER, J.L., GOETZ, M.: 'Design and construction of an optoelectronic crossbar containing a terabit/s free-space optical interconnect', accepted for *IEEE J. Selected Topics in Quant. Electron.*, March/April 1999, Special issue on Smart Photonic Components, Interconnects and Processing (Invited Paper)
- WALKER, A.C., YANG, T.-Y., GOURLAY, J., DINES, J.A.B., FORBES, M.G., PRINCE, S.M., BAILLIE, D.A., NEILSON, D.T., WILLIAMS, R., WILKINSON, L.C., SMITH, G.R., DESMULLIEZ, M.P.Y., BULLER, G.S., TAGHIZADEH, M.R., WADDIE, A., UNDERWOOD, I., STANLEY, C.R., POTTIER, F., VOGELE, B., SIBBETT, W.: 'Optoelectronic systems based on InGaAs-complementary-metal-oxide-semiconductor smart-pixel arrays and free-space interconnects', *Appl. Opt.*, 10 May 1998, **37** (14), pp. 2822-2830

Conference proceedings

- FORBES, M.G.: 'Electronic design issues in optically interconnected systems', 1-4 September 1997, Rank Prize Funds Mini-Symposium on Devices and Systems for Optical Interconnects and Data Links, Grasmere
- FORBES, M.G., WALKER, A.C., POTTIER, F, VOGELE, B., STANLEY, C.R.: 'Uniformity measurements on large InGaAs-AlGaAs multiple quantum well modulator arrays', *Spatial Light Modulators 97*, Technical Digest (Optical Society of America, Washington DC, 1997), pp. 123-124
- 3. WALKER, A.C., DESMULLIEZ, M.P.Y., FORBES, M.G., FANCEY, S.J., BULLER, G.S., TAGHIZADEH, M.R., DINES, J.A.B., STANLEY, C.R., PENNELLI, G., BOYD, A., PEARSON, J.L., HORAN, P., BYRNE, D., HEGARTY, J., EITEL, S., GAUGGEL, H.-P., GULDEN, K.-H., GAUTHIER, A., BENABES, P., GUTZWILLER, J.L., GOETZ, M.: 'An

optoelectronic crossbar switch as a demonstrator test-bed for terabit/s i/o', accepted for *Optics in Computing 99*, Technical Digest (Optical Society of America, USA, 1999)

- FANCEY, S.J., FORBES, M.G., DINES, J.A.B., YANG, T., WILKINSON, L.C., DESMULLIEZ, M.P.Y., WALKER, A.C., MARSH, J.H., STANLEY, C.R: 'Progress in optoelectronic systems based on InGaAs/CMOS smart-pixel arrays and free-space optical interconnects', *Technical digest of Optoelectronic Integration and Switching*, Glasgow, UK, IEE, pp.11.1-11.6, (November 13, 1997)
- YANG, T., BAILLIE, D.A., GOURLAY, J., FORBES, M.G., DINES, J.A.B., WALKER, A.C., POTTIER, F., STANLEY, C.R.: 'Experimental investigations of MQW InGaAs on CMOS smart pixels for optically connected computers', technical digest of Quantum Electronics (QE-13), Cardiff, UK, IOP, p. 89, (September 8-11, 1997).
- FANCEY, S.J., FORBES, M.G., TAGHIZADEH, M.R., DINES, J.A.B., BULLER, G.S., WALKER, A.C., DESMULLIEZ, M.P.Y., PENNELLI, G., MARSH, J.H., STANLEY, C.R., HORAN, P., BYRNE, D., HEGARTY, J., EITEL, S., GULDEN, K.-H., GAUTHIER, A., BENABES, P., GOETZ, M.: 'A free-space optoelectronic crossbar interconnect with Terabit/s communication to silicon electronics', *Conf. on Lasers and Electro-Optics Europe* 1998 – Technical Digest, Glasgow, UK, 14-18 September 1998, p. 50
- GOURLAY, J., YANG, T.-Y., DINES, J.A.B., FORBES, M.G., WADDIE, A.J., WALKER, A.C., VASS, D.G., UNDERWOOD, I., STANLEY, C.R., SIBBETT, W.: 'An optoelectronic sorter system', *Conf. on Lasers and Electro-Optics Europe 1998 – Technical Digest*, Glasgow, UK, 14-18 September 1998, p. 51
- WALKER, A.C., TOOLEY, F.A.P., TAGHIZADEH, M.R., DESMULLIEZ, M.P.Y., BULLER, G.S., NEILSON, D.T., PRINCE, S.M., BAILLIE, D.A., DINES, J.A.B., WILKINSON, L.C., FORBES, M.G., SNOWDON, J.F., WHERRET, B.S., STANLEY, C.R., POTTIER, F., VASS, D.G., UNDERWOOD, I., WILLIAMS, R., SIBBETT, W., DUNN, M.H.: 'Development of an optoelectronic parallel data sorter based on CMOS/InGaAs smart pixel arrays', *Optics in Computing 1997*, Technical Digest (Optical Society of America, Washington DC, 1997), pp. 149-151

Chapter 1

Introduction

1.1 Scope and overall research contribution

This thesis considers the problem of how to communicate data between VLSI integrated circuits over distances of several centimetres at overall data rates in the terabit/s region. In particular, it considers the application of optical communication technology to this problem.

The idea of using optical techniques to address the chip-to-chip interconnection problem has been around for a long time [1]. However, it is only in the last few years that technology with a realistic promise of eventual commercial application has emerged. Progress can be attributed to a shift away from trying to develop custom VLSI technologies with in-built optoelectronic capability towards developing techniques to allow parallel arrays of separately fabricated optoelectronic devices to be tightly integrated with standard foundry VLSI electronics. Parallel optical interfaces can be conceived consisting of arrays of optoelectronic devices providing of the order of one thousand optical channels each running at speeds around 1 Gbit/s and hence offering an overall capacity of 1 Tbit/s to a single integrated circuit. The technology has now developed to the point that it is possible to contemplate its use in commercial systems within a time-frame of 5-10 years, but it remains the case that a complete system of a realistic scale based on the technology has yet to be demonstrated. In the absence of such a demonstration, the risk attached to the considerable investment required to move the technology from the research labs into commercial application may be unacceptably high.

The specific focus of this work is on aspects of the design of the electronic circuitry that forms the interface between the digital VLSI electronics and the optoelectronics devices, in particular the analogue receiver circuits that convert photocurrent produced by the optoelectronic detectors into digital logic signals. While the essential structure of these receiver circuits is the same as that of conventional telecommunications receivers, the need to integrate between several hundred and a few thousand receivers onto the same chip in close proximity with digital logic in order to satisfy the bandwidth requirement creates a new set of design problems. A qualitative change in the nature of the receiver circuit is caused by the severe constraints on performance parameters such as power consumption and layout area that result from the large number of circuits in the array. Although several other authors have carried out work in this area, the complete understanding of the electronic design issues in large two-dimensional receiver arrays that is a prerequisite for the design of a fully operation demonstrator system is yet to be achieved. This work makes significant progress towards this goal through a number of design studies in the area covering, in particular, issues of transistor matching, electrical

crosstalk, predicted performance in future silicon technology and new receiver design structures for this application.

Although the design studies that make up this work are generic, they draw on the experience gained from the design, as part of larger team, of an experimental system that aims to demonstrate a parallel optoelectronic interconnect to a VLSI circuit with a bandwidth approaching 1 Tbit/s. Case studies from the design of this system are used to illustrate the general discussion throughout. Design of the demonstrator is complete and it is hoped to start initial experimental tests in early 1999. By attempting to design and build a system on the scale of the eventual application, it has been possible to identify problems in the underlying technology requiring further study before commercial exploitation can proceed. Even though complete operation to specification of this system is not expected, the problems identified in the design process make it more likely that successful operation of future systems will be achieved. The contribution made to the design effort on this project is an important part of the work reported in this thesis.

The remainder of this chapter justifies in more detail the motivation behind studying this topic and highlights the specific contributions to the area made by this work.

1.2 Motivation

1.2.1 The interconnect problem

In order to support the continuing increase in processing capability of integrated circuits and the overall improvement in the performance of digital systems that this has produced, a commensurate improvement in the capacity of the interconnects is required throughout the hierarchy of the system.

Since its beginnings, the semiconductor industry has sustained an exponential rate of progress as predicted by Moore in 1965 [2]. There is a significant effort within the industry to sustain the trend: the Semiconductor Industry Association (SIA) periodically produce a 'Technology Roadmap' [3] to identify the technological developments required over a 15 year time-scale to sustain exponential growth. It seems likely that the scaling down of transistor dimensions that is responsible for this trend will continue until at least the 0.05 μ m generation which is expected to occur by 2012. Although below 0.1 μ m, fundamental physics will limit further reduction of transistor dimensions, it is likely that even beyond this point there will still be commercial pressure to sustain the improvement in overall system performance by other means.

A large, digital system that is too complex to fit on a single chip must be partitioned into several modules. At different levels of the system, these modules may consist of individual

2

chips, multichip modules (MCMs) or entire printed circuit boards. In general, as the processing capability of each module increases, the capacity of the interconnect that connects the modules must also increase. This relationship is captured empirically by one formulation of Rent's rule which states that the bandwidth requirement of a module with processing capability C increases in proportion to C^{α} where α is an exponent between 0.5 and 0.75 [4]. Assuming that this relationship continues to hold and that the processing capability of each module continues to increase, there must be a requirement for an interconnect technology with a higher capacity.

However, there is an increasing gap between the off-chip bandwidth available using electrical interconnect techniques and the processing capability of integrated circuits. Even today, off-chip bandwidth is considered by the electronics community to be the limiting factor in the design of many systems [4]. It is not expected to grow at a rate that keeps pace with the increase in bandwidth requirement implied by the higher gate count and operating frequency together with Rent's rule. The 1997 SIA Roadmap indicates that no currently known packaging technology will be capable of meeting the year 2012 I/O requirements for high-performance integrated circuits¹ within the packaging cost targets of about \$125 per package in 1997 US dollars [3][5]. Whilst this is not an indication that packaging such chips using electronic interconnect techniques is a physical impossibility, it does suggest that alternative technologies for providing interfaces to ICs will become increasingly attractive. In identifying the six major "Grand Challenges" facing the semiconductor industry to sustain the same rate of growth beyond the 0.1 µm generation, the roadmap highlights off-chip communication in the gigahertz frequency range as "perhaps even greater than the challenge of on-chip performance at this frequency" [6].

It is possible to contemplate improvements in computer architecture that might reduce the system interconnect bandwidth requirement, but these would require that significant fractions of the available processing capacity be devoted to compensating for the limitations of the interconnect. For example, although historically on-chip caches have been used in CPUs primarily to reduce latency, they also serve to reduce the off-chip bandwidth [7][8]. A higher performance interconnect would free up these resources to do useful processing.

There are a number of applications that might require interconnect bandwidths in the terabit/s region. Terabit/s Internet routers are under active consideration for next-generation switching systems [9]. Indeed, a commercial switch that scales to a throughput of 5.6 Tbit/s has already been announced [10][11]. Massively parallel computer systems are another example [12] although the market for such systems is currently rather small and may not be sufficient in

¹ High-performance microprocessors are expected to require 512-bit data buses operating at 1.5 Gbit/s.

itself to justify commercial development of a completely new technology. Thus, although the improvements in silicon technology within the next ten years may allow many low-cost consumer systems to be integrated onto a single chip, and thus eliminate the need for a packaging hierarchy, it seems likely that there will remain a class of large digital systems that requires partitioning between several modules and thus an improved interconnect technology. It is this class of system that the work in this thesis targets.

1.2.2 Capabilities and limitations of electrical interconnects

This section discusses the physical origins of the limitations of conventional electrical interconnects and reviews how the interconnect problem might be addressed electronically, before going on in the next section to look at how optical interconnects can overcome some of these limitations. The issues in the design of high-performance electronic signalling systems are reviewed in [13] and an extensive bibliography of this field is available in references [14] and [15].

Frequency dependent loss

The main physical limitation on the use of electrical signalling over long distances is frequency dependent loss due to the skin-effect and dielectric absorption.

The attenuation due to the skin effect increases in proportion to $f^{1/2}$ above a certain critical frequency. This gives rise to a so-called 'aspect ratio' limit on the bandwidth of an electrical interconnect [16] [17] relating the maximum total capacity of an electrical interconnect B_{MAX} to the overall cross-section A and the length L:

$$B_{MAX} = B_0 \frac{A}{L^2} \tag{1.1}$$

The constant of proportionality B_0 is related to the resistivity of copper interconnects and is only weakly dependent on the particular fabrication technology; it is about 10^{15} bit s⁻¹ for typical MCM technologies² and 10^{16} bit s⁻¹ for on-chip lines.

This limit is scale-invariant and so applies equally to board-to-board interconnects as to connections on an MCM to the extent that B_0 is independent of technology. Also, for a fixed cross-section, the limit is independent of whether the interconnect is made up of many, slow wires or a few fast wires up to the point where other effects start to limit performance.

² Calculated by Miller and Ozaktas from experimental loss data assuming loss is dominated by skin-effect related mechanisms.

The aspect ratio limit is part of the reason why fibre-optics have replaced coaxial cables in telecommunications networks. The bandwidth requirement of digital systems has already reached the point where, in high aspect-ratio long haul telecommunications links, the limit of equation (1.1) is exceeded. As the bandwidth requirement continues to increase, the limit will become important in systems with aspect ratios more typical of those found in self-contained digital systems.

The attenuation due to dielectric absorption increases in proportion to frequency leading to an upper limit on operating speed that is inversely proportional to distance. It is independent of conductor cross-section and is not scale-invariant. For a 1 Gbit/s interconnect, it would in itself limit the length to 1 m for a stripline in a standard FR4 epoxy fibre-glass PCB interconnect and maybe 10 m in a good low-loss material such as PTFE. However, it does not limit the overall bandwidth of an interconnect over a certain distance in the same way as the skin effect because a higher overall bandwidth could be obtained by using more conductors within the same cross-section.

Figure 1-1 compares the limits on interconnect length imposed by the skin-effect and dielectric loss as a function of bit-rate and conductor cross-section.



Figure 1-1: Comparison of the bit-rate limit imposed by skin-effect and dielectric losses for a single-channel data link as a function of conductor cross-section

Impedance discontinuities

Propagation of electrical signals over appreciable distances requires the use of transmission lines; there are a number of practical difficulties in constructing an ideal transmission line with a uniform characteristic impedance. Impedance discontinuities created by package pins, vias, connectors and gaps in the signal return plane create unwanted reflections which contribute to the total noise on the digital link. This becomes an increasing problem at high signalling rates because the physical size of discontinuities that are important becomes smaller.

Manufacturing tolerances in the transmission line medium and termination resistors also contribute towards signal reflections.

Crosstalk

Electrical interconnects are susceptible to various forms of crosstalk but this is not a fundamental limit and can be controlled with careful design. Sources of crosstalk include mutual inductance and capacitance with adjacent signal lines and the finite impedance of a common signal return in a connector or a package pin. The former source can be completely controlled by enforcing rules on minimum signal separation. Both problems can be substantially reduced by employing differential signalling.

Power consumption

The power dissipation in termination resistors is often cited as a major disadvantage of electronic signalling but is a much less serious problem with low-voltage signalling technologies specifically designed for high-speed operation than with conventional CMOS signalling which has very poor performance in high-speed signalling applications. For example, a 5V swing across a parallel terminated 50 Ω transmission line³ produces a power dissipation of 250 mW; in contrast, for a 400 mV differential swing across a 100 Ω termination (as used in the LVDS standard [18]), the power dissipation is only 1.6 mW.

Nevertheless, the overall power consumption of a long distance electrical interconnect is typically an order of magnitude higher than in an optical equivalent using highly integrated optoelectronic components. Reasons for this include the larger parasitic capacitance of electrical interconnects (of the order of 1 pF compared to 100 fF for a small optoelectronic device, constrained in part by the need to provide adequate electrostatic discharge protection) and the need to employ complex synchronisation circuitry to compensate for the skew in long-distance electrical interconnects.

For example, a state-of-the-art commercial 2.5 Gbit/s bidirectional transceiver core [19] has a typical power consumption of 200 mW per channel and allows a total chip I/O bandwidth of 160 Gbit/s for a 32-channel system. In contrast, 1 Gbit/s optical transceivers have been demonstrated with a power consumption of the order of 10 mW [20].

6

Electromagnetic compatibility

Meeting regulatory requirements on electromagnetic emissions is a challenge in large digital systems and can increase the overall cost of the system. This problem is again substantially reduced by the use of low-voltage differential signalling.

Example problem

As an example of the limitations imposed by aspect ratio and the skin-effect, consider the problem of constructing an electrical backplane for a rack-mounted electronic system. Assume that the backplane has a horizontal dimension of 40 cm and a vertical dimension of 10 cm and consists of a multilayer printed circuit board. The signals use thin copper striplines, of width w and with a ground layer between each signal layer⁴, configured as point-to-point differential links⁵. The lines are assumed to run parallel and the longest link is assumed to extend across the entire backplane.

A typical construction [15] for such a link would be a 150 μ m wide track with a pitch of 0.5 mm for each track for an overall pitch of 1 mm for the differential pair with a dielectric thickness h of about 0.4 mm. This gives a 50 Ω characteristic impedance. The pitch is set by crosstalk requirements rather than fabrication constraints. Assuming that current flows uniformly within one skin depth of each surface in the stripline and in an equal area in the return path through the ground plane, the maximum bit-rate per line is:

$$B_{MAX} = \frac{8 \sigma Z_0^2}{\pi \mu_0} \frac{w^2}{L^2} \left(\frac{\alpha}{20 \log_{10} e}\right)^2$$
(1.2)

where $\sigma = 5.80 \times 10^7 \,\Omega^{-1} \text{m}^{-1}$ is the conductivity of copper, Z_0 is the characteristic impedance of the line (which is a function of w / h), α is the maximum tolerable loss in dB at a frequency $B_{MAX} / 2$ (taken to be 2 dB [13]) and μ_0 is the permeability of free space. This gives a data rate limited to 2 Gbit/s by frequency dependent loss due to the skin effect alone. Within the 10 cm vertical dimension of the backplane, there is room for 100 signals per layer giving an absolute

 $^{^{}_3}$ Parallel termination of 100 Ω resistor to +5V and 100 Ω resistor to 0V assumed.

⁴ It might be possible to place two layers of conductors between ground planes to reduce the number of layers, provided the separation between the two signal layers is sufficient to avoid crosstalk (in this backplane application, all conductors are assumed to be running in the same direction).

⁵ The advantages of differential links for high-speed signalling are outlined below. Point-topoint links are much easier to implement at high line rates than multidrop buses because they have fewer impedance discontinuities.

maximum capacity of 200 Gbit/s per signal layer. In practice, it would be difficult to achieve this routing density because of the area occupied by vias in the board and the fact that the required pattern of interconnections is likely to be more complicated than a set of parallel lines.

Now consider the effect of varying the dielectric thickness h by a factor K. If the thickness is decreased, then the track width must be decreased to keep the characteristic impedance the same. The minimum pitch required to achieve a given level of crosstalk is proportional to h and so the pitch can also be decreased by a factor K. There is an overall increase in the number of signals K but a reduction in the maximum rate of each line by a factor K^2 and thus an overall reduction in the capacity of the backplane by a factor K reflecting the increase in the aspect ratio. Conversely if the dielectric thickness goes up by a factor of K, pitch must be increased to keep the crosstalk at the same level; the capacity of the backplane goes up by a factor of K with K times fewer links running at K^2 times the speed. Since other factors such as impedance discontinuities due to the backplane connectors will limit the maximum speed of an individual line, there is clearly a limit to which the capacity of a single layer can be increased by this means.

To provide a significant increase in the backplane capacity, it is necessary to add more layers, which will increase the cost of the electronic system and may impose a practical limit on the capacity that can be achieved using standard electrical interconnects. The number of layers required to implement a terabit/s capacity backplane (5 signal plus 5 power) is not prohibitive but for capacities much higher than this, the limit starts to become important. Because of the scale invariance of (1.1), the above calculation applies to any planar interconnect with a length to overall width ratio of 4:1. The limit may differ slightly from technology to technology because of differences in the maximum number of layers that can be fabricated economically. For example, it is currently practical to fabricate printed circuit boards with 50 layers [21][22] whereas MCM technology is restricted to between 5 and 8 [23][24]. On the other hand, the physically smaller size of the impedance discontinuities in multichip modules will increase the maximum possible bit-rate per line and thus allow an overall increase in capacity per layer through the use of proportionately thicker dielectric spacing.

This calculation says nothing about the other practical difficulties of building an electrical backplane with this capacity, in particular the problems created by the impedance discontinuities of the backplane connectors.

Electronic solutions to this problem are being considered but will tend to increase the overall cost of a system through, for example, increased power consumption. One promising technique is to employ line equalisation to increase the amount of tolerable loss in (1.2). Laboratory implementations have demonstrated that 10 dB of loss can be equalised in a 4 Gbit/s link,

8

which increases the bandwidth limit by a factor of 25. The power consumption in current technology is around 100 mW per transceiver [13].

1.2.3 Advantages of optical interconnects

This section reviews the motivation behind the exploration of optical interconnects as an alternative to electrical interconnects. The physical advantages of optical connections are discussed in more detail by Miller [25].

The narrowband nature of optical signals makes it relatively simple to construct high-quality, uniform transmission lines that operate at high data rates. An optical signal is modulated onto a very high frequency (~ 10^{14} Hz) carrier and thus the signal can be treated as narrowband for the purposes of matching the impedance of the transmission line (in this context, free space) to the load (in this context, the detector material) using quarter-wave sections of transmission line (in this context, a dielectric antireflection coating) in a way directly analogous with the termination of narrowband microwave signals. Unlike the resistive termination required to achieve a broadband match to a baseband digital signal, there is no power dissipation associated with this termination.

The narrowband nature also eliminates frequency dependent loss: there is no equivalent to the aspect ratio limit for optical interconnects. Dispersion limits the distance over which signals can propagate in optical fibres, leading to a fixed bandwidth-distance product (1 GHz km for typical multimode fibre [26]), but this limit is not important over the distances encountered in self-contained digital systems.

Various forms of optical crosstalk exist in parallel optical data links but, as with electrical signals, can be controlled by separating the channels by sufficient distance. Optical signals do not in themselves suffer from signal-return crosstalk, although the transceiver electronics at both ends of the link are still susceptible to the problem. Optical signals do not produce and are immune from electromagnetic interference.

Parallel optical data links also have lower skew than parallel electrical data links. In free-space optical data links, the intrinsic skew is zero whilst in fibre data links, the skew resulting from manufacturing variations can be an order of magnitude better than achievable in coaxial cable (8 ps m⁻¹ for typical fibre ribbon compared to 40-50 ps m⁻¹ for micro-coaxial cable [27]). This can eliminate the need to resynchronise each channel independently in a medium-distance parallel optical interconnect and lead to an overall reduction in power consumption as discussed in Section 1.2.2.

1.2.4 Summary

In the short-term, relatively simple improvements to established electrical signalling techniques, such as the widespread adoption of low-voltage differential signalling, are sufficient to overcome many of the disadvantages of current CMOS signalling practice.

In the longer term, the aspect ratio limit on conventional electrical interconnections will make it increasingly expensive to sustain the growth in overall system performance in the absence of revolutionary approaches to the interconnect problem. It is important that, at this point in time, a range of such approaches are considered so that, by the time the interconnect problem is faced by commercial systems, at least one is sufficiently mature for adoption by industry. Equalised electrical signalling, new architectures that reduce the requirement for global interconnect and optical interconnections are all possible solutions. The goal of this work is to contribute towards the development of the latter.

1.3 Thesis organisation and specific research contributions

The organisation of the thesis is as follows.

Chapter 2 begins by reviewing the different technologies that can be used to form highbandwidth optoelectronic interfaces to VLSI electronics; it discusses the current state-of-the-art and describes in more detail the specific optoelectronic device technology, based on surfacenormal InGaAs multiple-quantum-well diodes, that is used as the focus for this work. It describes some experimental characterisation of these devices, carried out by the author, which demonstrates for the first time that devices in this particular fabrication process can be modulated at speeds that are useful for high-bandwidth parallel optical interconnects.

Chapter 3 describes the design of a prototype switching system, based on this technology, that includes an optoelectronic interface with a target capacity of 1 Tbit/s. This system is the focus of a large research effort to which more than 25 people have contributed. Its aim is to act as a vehicle for investigating some of the technologies required for high-bandwidth optical interfaces; although, as mentioned above, experimental tests of the complete system are not due to start until early 1999, it has already partially fulfilled this goal in the design phase of the project by identifying the main issues that must be considered when a system based on this kind of technology is built.

The chapter concentrates primarily on the system architecture and high-level digital design. Some of the general electronic design issues that were raised in the course of this project are highlighted. The author made significant contributions to these areas of the project, although the detailed digital design was carried out by Philippe Benabes and Alain Gauthier of Supélec, Paris. The author was solely responsible for the design of the analogue receiver subsystem and,

10

by adopting electrically differential design techniques for some of the electronic receiver circuitry and modulator drive circuits, contributed towards the development, by the University of Glasgow, of a new and more manufacturable optoelectronic device fabrication process.

The remainder of the thesis deals specifically with the design of receiver circuits for this application area and was carried out entirely by the author.

Chapter 4 is a general study of the design trade-offs in such circuits. It contrasts the design optimisation with that in conventional single-channel optical receiver circuits and includes the first quantitative analysis of the role of MOS transistor mismatch in large receiver arrays, updating related work by Novotny on FET-SEED technology [166]. Chapter 5 illustrates this discussion with a case study from the switching system described in Chapter 3, and includes experimental results from an electrical test implementation of a receiver circuit in $0.6 \mu m$ technology.

Chapter 6 extends the design study of Chapter 4 by looking at how receiver performance can be expected to change in future silicon technology. While there has been previous work in this area [167][168], it has not included the effects of transistor offset, which the analysis shows to be the main factor limiting circuit performance. This result suggests that slightly different approaches may be required in future to alleviate the offset problem and discusses some of the possibilities.

Chapter 7 describes the first application of the transconductance-transimpedance circuit technique to receivers in this application area. Its general advantages for high-data rate receivers are discussed and illustrated with another case study from the prototype switching system including experimental results.

Chapter 8 presents a method for analysing electrical crosstalk resulting from simultaneous switching noise in large receiver arrays. It introduces a simple analytical technique for dealing with the distributed nature of the receivers in two-dimensional arrays. It highlights some significant shortcomings of the simple receiver circuits used to date in this application area in terms of immunity to crosstalk and suggests some alternative approaches.

Finally, Chapter 9 concludes by summarising the main results and identifying areas requiring further study.

Chapter 2

Optical interface technology

2.1 Introduction

In this chapter, the technologies required to implement high-bandwidth optical interfaces to VLSI electronics are reviewed.

The technology for implementing single-channel optical links is very mature. The use of single-channel optical links in long-haul telecommunications is well established. More recently, optical data links have been used in a number of mass-market applications including local-area-networking. Fibre Channel [28] and Gigabit Ethernet [29] are two example standards that provide low-cost data links at data rates of around 1 Gbit/s using a single fibre.

Single-fibre terabit/s optical data links have been demonstrated in the laboratory using wavelength-division multiplexing [30]. However, these demonstrations are aimed at long-haul telecommunications; the technology is not suitable as it stands for constructing highly-integrated interfaces to VLSI electronics: it uses a large number of individual optoelectronic devices and fibre components to combine tens of independent data sources that are generated using special purpose electronics operating at \geq 20 Gbit/s.

It is more likely that terabit/s scale optical interfaces to individual VLSI circuits will be implemented using more channels operating at lower speeds that are compatible with mainstream electronics (perhaps somewhere between 256 channels \times 4 Gbit /s and 1024 channels \times 1 Gbit/s). The technology required to support optical links with this many channels is still at the research stage. There is a requirement for large arrays of optoelectronic devices, a transmission medium that can carry multiple channels (such as a fibre-ribbon or a free-space imaging system) and new packaging techniques to interface the optoelectronic devices to the transmission medium. In particular, two-dimensional arrays of optoelectronic devices and associated packaging technologies may be required to provide sufficient channels.

Integration of the optoelectronic devices with mainstream VLSI electronics is almost certainly required to make these multiple-channel optical links economically feasible. Monolithic optoelectronic integrated circuits (OEICs) comprising, for example, a detector and receiver in a special-purpose fabrication process, have been demonstrated to provide high-performance for single-channel links. However, multiple-channel links require a VLSI process to implement arrays of interface circuits: receivers to convert photocurrents into

12

digital logic levels, and transmitters to convert digital logic levels to the signals required to drive modulators or emitters. A capability to include complex digital functionality such as buffering and routing is also desirable.

A huge investment would be required to develop, from scratch, a high-yield, state-of-the-art VLSI process with a full optoelectronic capability. The market for optical interconnect technology is simply not large enough to justify this development even if it were technically feasible. A more realistic approach, commonly referred to as 'optoelectronic VLSI', is to leverage off existing investment in mainstream VLSI by implementing the electronic functionality in a conventional, state-of-the-art fabrication process and developing techniques for integrating separately fabricated optoelectronic devices with the VLSI electronics. This approach also allows the performance of the optoelectronic devices to be optimised separately from the electronics. Silicon CMOS is a strong candidate for the base VLSI technology because of its low-cost and widespread use, but other technologies with a VLSI capability that are available at moderate cost, such as GaAs MESFETs, could also be considered.

This chapter begins by reviewing monolithic and hybrid techniques for accomplishing this integration. The characteristics of the most important optoelectronic device types are discussed and some experimental characterisation of the particular device technology used in this work is described. The chapter concludes by comparing the fibre-ribbon and free-space approaches to implementing multiple-channel optical data links.

Several detailed reviews of optical interface technology have been published by other authors [31][32][33].

2.2 Monolithic integration

It is possible to fabricate detector structures as part of the standard processing sequence of certain VLSI technologies. Although the performance of these detectors is, in general, not as good as optimised devices in custom processes, there are obvious cost advantages in using a standard foundry.

Unfortunately, the responsivity of high-speed photodiodes fabricated in silicon CMOS processes at the standard short-distance data-link wavelength of 850 nm is poor. The physical origin of the poor performance is the long absorption length at this wavelength in silicon. Carriers are generated deep in the silicon substrate and take a long time to reach the *pn* junction, which is relatively shallow in a modern CMOS process, resulting in a detector with a slow response. Improved performance can be obtained at shorter wavelengths. Techniques have been used to improve the speed of response at the expense of responsivity.

Ayadi [34] discusses the design of detectors in a standard CMOS process in detail and describes a 180 Mbit/s receiver circuit [35]. Woodward et. al. [36] report a 1 Gbit/s receiver using a photodiode with a responsivity of between 0.01 A/W and 0.04 A/W at 850 nm in a standard CMOS process; Kuchta et. al. [37] report a 0.07 A / W detector in a standard BiCMOS process with an intrinsic bandwidth of 700 MHz. Kuijk [38] describes a low-capacitance (0.012 fF / μ m²) detector structure implemented in standard CMOS with a responsivity of 0.05 A/W; the low capacitance of this structure might compensate in part for the low responsivity (see chapter 4). The ideal responsivity at 850 nm is 0.68 A/W and so these detectors provide external quantum efficiencies of between 5% and 10%. The low responsivity of these devices compared to optimised *pin* diodes means that higher power optical sources would be needed to achieve the same data rate. Whether these monolithic detectors are suitable for use in large arrays will depend on whether the cost saving resulting from the use of a standard, monolithic process is enough to offset the increased cost of the more powerful optical sources.

Recently, GaAs MESFET technology has emerged as an alternative to silicon for highspeed circuits with medium circuit complexity [39][40]. This technology may offer a suitable platform for implementing high capacity optical data links and has been used in several experimental systems [41][42]. One advantage of this technology is that it allows easy fabrication of interdigitated metal-semiconductor-metal detectors, or MSMs, as part of the standard fabrication process. A single additional mask step can be used to improve the responsivity of the detectors. Compared to *pin* diodes, MSM detectors have the advantage of a low capacitance per unit area (in the range 0.01-0.06 fF / μ m² depending on the design [43][44]) but have a lower responsivity due to the shadowing effect of the metal fingers (typically in the range 0.2-0.35 A/W [45][46][100]). A theoretical comparison has shown [47] that, when the effect of photodiode capacitance on receiver performance is considered, MSMs can offer better overall performance than *pin* diodes. 8×8 arrays of MSM detectors have been reported [48]; because of the extremely simple device structure, it can be anticipated that very large arrays could be fabricated with high yield.

Integrated GaAs MSMs have been used commercially to fabricate a low-cost 1 Gbit/s integrated receiver for local area network applications [49].

The performance of GaAs MSMs is easily good enough for optical interconnect applications, but it is likely that GaAs MESFET based VLSI electronics will remain a relatively expensive, niche technology with lower levels of integration than a state-of-the-art silicon CMOS process.

In any case, standard monolithic technologies can only offer a solution to half of the optical interconnect problem because the device structures required to implement the optical output

14

devices – whether they be modulators or lasers – require dedicated processing steps. Indeed, the silicon material system is fundamentally unsuitable for emission based output devices because of its indirect bandgap. Efforts to monolithically integrate modulator devices with GaAs MESFET technology by Bell Labs [50], although successful in producing devices for prototype optical interconnect systems [65], were abandoned in favour of hybrid integration techniques.

2.3 Hybrid integration technology

In the absence of a completely monolithic solution to the problem of fabricating optoelectronic device arrays, some form of hybrid integration technology must be considered. Two basic approaches have been followed:

- fabricating the optoelectronic devices separately and then bonding them to foundry VLSI chips or wafers;
- taking foundry wafers and using them as a starting point for growth of optoelectronic material and fabrication of optoelectronic devices.

2.3.1 Bonding technologies

Conventional wire-bonding has been used for parallel optical data link applications [41][103] but is limited in its applicability to chips with a relatively small numbers of channels. Typically this would involve wire bonding both optoelectronic devices and electronic interface circuits to an intermediate substrate such as a ceramic multi-chip-module. This scheme is relatively simple for linear arrays of devices, but in two-dimensional arrays, interconnect traces are required to route the device connections to the periphery of the optoelectronic die for bonding. A key problem with this approach is the effect of the parasitics associated with the wire bond pads and interconnect traces. It will be shown in chapter 4 that the low parasitics associated with more advanced packaging techniques are not so much a useful advantage but rather an essential requirement for the feasibility of very large receiver arrays because of the strong relation between receiver input capacitance and power consumption.

A more promising technique for hybrid integration is flip-chip bonding. This is a derivative of IBM's C4 (controlled-collapse chip connections) technology [51] that was developed in the 1960s and is widely used in high-density electronic packaging today. It was first used to attach optoelectronic devices more recently [52][53]. Figure 2-1 shows schematically an optoelectronic device (a *pin* diode) attached to a CMOS circuit using the flip-chip attachment technique. The electrical connections to the two terminals of the photodiode are formed by spherical bumps of solder.





The diode is referred to as a 'surface-normal' optoelectronic device because the light is incident from above. Such devices allow the construction of two-dimensional arrays of devices, in contrast to, for example, edge emitting lasers. In this geometry, the optical signal has to pass through the substrate. In certain material systems (strain-balanced InGaAs on GaAs between 980 nm and 1064 nm, and InGaAs on InP at 1.3 μ m and 1.55 μ m), the substrate is transparent and the device can be used as it stands. In other material systems (GaAs at 850 nm), the substrate is absorbing and must be removed [54].

The flip-chip solder connections are formed as follows. The starting point for the process is two wafers covered in passivation: one silicon wafer fabricated in a standard foundry and one wafer of optoelectronic devices. The wafers contain metal pads at the location of the flip-chip connections. Holes are opened in the passivation layer above the metal pads, and a secondary pad of solder-wettable material is deposited over the opening¹. On one wafer, solder is then evaporated over the wettable pad and melted to form spheres of solder (Figure 2-2) bounded by the non-wettable passivation layer. After dicing into individual chips, the optoelectronic chip is turned upside down and aligned approximately with the silicon chip. The solder is then melted again and the two chips pulled by surface tension into accurate alignment. This process requires three extra lithographic steps for each chip. A typical dimension for the wettable pad is 15-20 µm.

¹ The wettable pad typically consists of a three layer stack of metals such as chromium, copper and gold. A thin outer layer of gold protects the copper from oxidation and is dissolved by the solder. Copper is the wettable layer itself. Chromium forms a barrier layer to prevent a reaction between the copper and the underlying aluminium metallisation.

The parasitic capacitance and partial inductance of the flip-chip connection are very small [55]. The flip-chip hybridisation technique is suitable for a very large numbers of connections. The failure rate of solder bumps is of the order of 1 in 50 000 [56] which is typically much better than the optoelectronic device yield. Other techniques for attaching optoelectronic devices have been used and are reviewed in [31].





Figure 2-2: Flip-chip integration technology. The left hand picture shows an InGaAs modulator array with reflowed solder bumps on top of the optoelectronic devices prior to flip-chip attachment. The diameter of the solder pad is 20 µm. The right-hand

picture shows the device array after flip-chip assembly with a foundry 1 µm CMOS circuit packaged into a standard chip carrier. The inner cavity dimension is 7.5 mm.

2.3.2 Regrowth technologies

The second approach, growing an optoelectronic substrate on top of a wafer from a VLSI foundry, is being investigated by a number of groups.

This technique has the general problem that the growth and fabrication steps used to form the optoelectronic devices must not degrade the performance of the underlying electronics. In particular, processing at temperatures above \sim 470°C degrades the interconnect metallisation and must be avoided [57].

Nevertheless, attempts to grow GaAs-InGaP LEDs [57] and MQW modulators [58] using low-temperature molecular beam epitaxy on top of a VLSI GaAs MESFET process have been reasonably successful. Attempts to grow III-V material on silicon [59] have additional problems due to the difference in lattice constant between the two materials [60].

This technology is still at the research stage in contrast to the flip-chip bonding technology which is in commercial use.

2.4 Optoelectronic devices

Optoelectronic device technology is sufficiently well developed to fabricate device arrays of the size required for terabit/s scale interfaces with reasonable yield, although some improvements in output device technology are desirable.

The main candidates for a detector technology are *pin* photodiodes and monolithically integrated MSM detectors. MSM detectors have already been discussed and are primarily a candidate for integration with GaAs MESFET circuits. Surface normal *pin* photodiodes are ideal for flip-chip integration with silicon. Responsivity can be comparable with discrete devices and small-diameter devices with low capacitance can easily be fabricated. The device diameter is limited by the optomechanical packaging requirements rather than fabrication limits.

The main candidates for output devices are multiple-quantum-well (MQW) modulators and vertical cavity surface emitting lasers (VCSELs).

2.4.1 MQW modulators

MQW modulators are the more mature technology. Zero-defect arrays containing more than 4000 modulators have been reported [61] with an average device failure rate of 1 in 1000. Spatial light modulator arrays with 65536 devices have been constructed with device failure levels of 1 in 5000 [62]. Thus, fabrication of device arrays with of the order of 1000 devices with commercially acceptable yields are practical.

MQW modulators are formed by a series of quantum wells contained within the intrinsic region of a reverse biased *pin* junction. Application of an electric field across the wells changes the absorption coefficient of the well material according to the quantum confined Stark effect [63][64] and results in modulation of an external beam incident on the device (Figure 2-3). A mirror at one end of the *pin* stack reflects the incident beam which passes through the intrinsic region twice (Figure 2-1). Electrically, the devices are essentially just small capacitors and can be driven using a standard CMOS inverter.

The external beams are typically generated by a single, high-power laser in combination with a diffractive array generator element [65]. The need to bring in an external beam and route the reflected beam through a different path in the optical system complicates the optomechanical packaging of modulators compared to emitters.





Figure 2-3: Absorption of an InGaAs/AlGaAs MQW modulator as a function of wavelength and reverse bias voltage. Calculated from transmission measurements in Wilkinson [68] assuming a mirror reflectivity of 1 and neglecting absorption in the substrate and buffer layers. Notice how, at a fixed wavelength in the range 1040-1080 nm, altering the bias voltage produces a change in absorption. Increasing the voltage shifts the band-edge to longer wavelengths and broadens the exciton peak.

Two principal material systems have been used to fabricate large arrays of surface-normal modulator devices: GaAs/AlGaAs devices² for 850 nm and InGaAs/AlGaAs for 980 nm-1064 nm. GaAs/AlGaAs devices offer better modulation depths and are simpler to grow but require substrate removal because of the opaque substrate. Historically, an important reason for pursuing the InGaAs material system was the availability of high power (~1W) solid-state laser sources at 1047 nm and 1064 nm (Nd:YLF and Nd:YAG) but improvements in diode-laser technology [66] have made this advantage less important. Substrate removal is not perceived to be a major disadvantage and, in the future, the GaAs/AlGaAs material system is likely to be the preferred choice. Table 2-1 compares the performance of example devices in the two material systems. The reflectivity change ΔR is the fraction of the incident power which translates into useful modulation which is a useful figure of merit for modulator performance. The output data beams typically have relatively low contrast ratios of about 2:1 – typically the modulated beam might contain 30% of the incident power in one logic state and 60% of the incident power in the other logic state.

² The notation GaAs/AlGaAs denotes a device with GaAs wells and AlGaAs barriers.

| material system | ΔR | operational temperature | operating wavelength |
|-----------------|------------|----------------------------|----------------------|
| | (peak) | range ($\Delta R > 0.2$) | |
| GaAs/AlGaAs | 0.46 | 27°C | 852 nm |
| InGaAs/AlGaAs | 0.33 | 27°C | 1068 nm |

Table 2-1: Comparison of modulator performance in GaAs/AlGaAs and InGaAs/AlGaAs material systems. GaAs/AlGaAs data is from [67]; InGaAs/AlGaAs data is calculated from experimental measurements in Wilkinson [68] (sample B491).

A temperature coefficient of 0.35 nm / K is assumed for the InGaAs devices [69].

A drawback of modulator devices is that they are inherently sensitive to the wavelength of the excitonic absorption feature which is affected by temperature and manufacturing tolerances. The read laser wavelength can easily be stabilised (for example, by using an external-cavity with a semiconductor laser). The temperature dependence is a more serious problem. The temperature coefficient of the band edge is approximately 0.3 nm K⁻¹ in both material systems [70][69]. Table 2-1 gives an approximate indication of the useful wavelength range defined arbitrarily as the temperature range over which a ΔR of better than 0.2 is maintained. Note that the disparity in performance between the two material systems with a realistic tolerance on temperature is less; the broader exciton peak of the InGaAs material system that is the cause of the poorer modulation performance under optimum conditions also makes it less sensitive to temperature changes. As it stands, this temperature range is insufficient for operation in a standard electronics environment; however, active regulation of chip temperature can be used [70]. The useful temperature range can be extended by adjusting the reverse bias on the modulators in response to temperature changes on the chip³ [71] which, for a particular design of GaAs/AlGaAs device, extends the wavelength range for a 2:1 contrast to 17 nm or 60°C [72].

As part of this study, a 64×32 array of InGaAs/AlGaAs MQW modulators, fabricated by the University of Glasgow, was characterised experimentally [73]. A portion of the array is shown in Figure 2-4. The intrinsic region consisted of 95 periods of $In_{0.22}Ga_{0.78}As$ -Al_{0.15}Ga_{0.85}As; the details of the device design are described in Wilkinson [68]⁴. The active area of the array was 2.9×2.9 mm². Each diode had a 20 µm diameter optical window. Each of the 32 rows of the array consisted of 32 differential pairs of diodes with the centre node tied together to allow all devices in the row to be driven in tandem (Figure 2-5). Because the

³ The detector bias also has to be adjusted if the detectors are fabricated in the same wafer as the modulators to maintain the resonant enhancement of absorption coefficient.

⁴ Devices fabricated in wafer B492.

fabrication process had been designed to produce devices suitable for flip-chip integration, a complicated mounting scheme was required to allow conventional wire-bonding to the devices (Figure 2-6). The scheme involved supporting the die round its periphery in a machined brass holder and accessing the devices optically through a hole in the back of a circuit board.

tht TDT.

Figure 2-4: Portion of a 64 \times 32 array of InGaAs MQW modulators on 45 μm / 90 μm

pitch. Each circle is a MQW diode.



Figure 2-5: Configuration of modulator diodes on 64 × 32 test array



Figure 2-6: Mounted InGaAs MQW modulator test array. Optical access is through a hole in the back of the circuit board and the substrate of the device. Thirty-two 50 Ω microstrip signal lines, spaced to avoid crosstalk, are visible, with bias voltages supplied to the top and bottom of the chip.

Large signal modulation tests (Figure 2-7) demonstrated a contrast ratio of about 2:1 for a 5 V swing which is consistent with predictions from transmission spectrum measurements on the same material [68] and measurements of large signal modulation on flip-chipped devices [74].



Figure 2-7: Large signal modulation of an array with a 2:1 contrast ratio (5V swing /

5V prebias at 1056 nm) The oscilloscope trace shows the signal reflected from one modulator device monitored through a fibre-coupled photodiode. The zero-level on the trace corresponds to the position of the horizontal graticule. The photograph shows the circular devices illuminated by an array of spots and viewed, through the substrate. The two diodes in the centre of the photograph have a reverse bias of 10 V and absorb more of the incident light. The horizontal pitch is 45 μm.

Tests on this array revealed yield problems with the InGaAs fabrication process (Figure 2-8). The device yield was approximately 90%. Notice the non-planar interconnect required to implement the diode-pair structure; polyimide bridges were used to overcome the change in height at the vertical mesa side-wall. The polyimide bridges are believed to be the main cause of the low yield. They were completely eliminated in a revised process, devised by the University of Glasgow, which used sloping mesas and was previously shown in Figure 2-1. The system described in the next chapter deliberately avoided the use of structures requiring a non-planar electrical interconnect to side-step these problems.



fabrication process that used polyimide bridges

The intrinsic response speed of MQW modulators is very fast and in practice the speed of operation is limited by the time required for the drive electronics to charge the device capacitance. Intrinsic switching times as low as 33 ps have been measured [75] with large signal modulation to at least 2 Gbit/s limited by test equipment [76]. High-speed tests of the device array just described [77] confirmed intrinsic operation to at least 500 Mbit/s (Figure 2-9). In these tests, a single row of devices was driven by a terminated 50 Ω transmission line. The combined rise time of the signal source, detector and oscilloscope (estimated from instrument specifications to be 1.2 ns) should have been just sufficient for observation of an open eye at 622 Mbit/s; the discrepancy with the experimental results may in part be due to the difficulty that was encountered in obtaining stable triggering of the oscilloscope at this speed. The RC rise-time due to the source impedance of the terminated transmission line and the capacitance of a row of devices was estimated to be about 200 ps and should not have limited performance. Nevertheless, the experiment demonstrates that the individual MQW devices in this fabrication process are capable of operating at the speed required to implement terabit/s scale optical interfaces.



Figure 2-9: Eye diagrams obtained by direct modulation of a test 64 × 32 MQW modulator array

Low-voltage operation of future generation CMOS technology may result in a reduction in the reflectivity change possible with MQW modulators but a number of solutions have been proposed including new driver circuits [78], stacked devices that are electrically in parallel and optically in series [79] and Fabry-Pérot enhanced devices [80][62].

The *pin* diode that forms the modulator device also makes a good detector. For example, *pin* diodes in a GaAs/AlGaAs modulator process have been measured to have a responsivity of about 0.5 A/W (corresponding to an external quantum efficiency of about 75%) and a device capacitance of about 0.11 fF μ m⁻² [81]. The device capacitance is somewhat higher than a *pin* diode optimised for detection; a relatively narrow intrinsic region is required to provide an electric field that is high enough to produce a significant shift in the band edge for good modulation with the drive voltages available from a modern CMOS process. The relatively high responsivity is obtained in part due to an excitonic resonance in the absorption coefficient close to the band edge, which is a consequence of the quantum well structure of the device.

Wafers containing arrays of *pin* detectors and modulators have been used to construct several large scale experimental systems based on free space optical interconnects [77][82][83] and are available commercially as a 'research-grade technology' on a 'best-effort' basis from Lucent Technologies [84].

2.4.2 VCSELs

VCSELs are a more recent technology and, unlike modulators, are not quite yet sufficiently well developed for the construction of terabit/s scale systems. However, the emergence of high-volume commercial applications for single VCSELs such as local area networking can be expected to push the development of the technology and, in the long-term, the simpler optical packaging requirements of an emitter device may make the VCSEL the output device of choice.

The fabrication of VCSELs is reviewed in reference [85]. VCSELs can be classified according to the technique used to define the extent of the laser cavity. Ion-implantation and selective-oxidation are two of the techniques used.

Ion-implanted VCSELs are ready for commercial application in small arrays but improvements in electrical-to-optical power conversion efficiency are desirable to avoid power dissipation problems in large arrays. Honeywell have reported devices with typical wall-plug efficiency and output power of 11% and 2 mW at 10 mA which is representative of the performance of this class of VCSEL. They report a device yield of 99.8% across 3" wafers (1 in 500 devices outside specification) with excellent long-term reliability [86].

Selectively-oxidised VCSELs promise better efficiency but have not yet demonstrated the same levels of yield and reliability as ion-implanted devices. Individual devices with efficiencies of the order of 50% have been reported at both 850 nm [87] and 960 nm [88] at output powers of around 1 mW.

VCSELs are also capable of operation at high data rates, although the laser turn-on delay limits operation unless the VCSELs are biased above threshold. Biased modulation to 10 Gbit/s has been reported [89] and bias-free modulation demonstrated at 2.5 Gbit/s [90]. Bias-free modulation would be advantageous in large arrays because it simplifies the driver electronics and reduces the dependence of the optical output power in the low state on variations in threshold current with age, temperature or position within an array.

Unlike modulators, VCSELs are capable of operating over a wide temperature range. However, the threshold current is a function of temperature which can lead to variations in output power and thermal crosstalk. Thermal crosstalk is an outstanding issue in large arrays, but can in principle be overcome by using a DC balanced line code.

The majority of VCSEL arrays that have been developed have been designed for conventional periphery wire bonding. A number of groups have demonstrated 8×8 arrays with good uniformity and 100% yield [91][92] at both 850 nm and 980 nm. 64×1 linear arrays have also been produced [93].

Further improvements are required in VCSEL integration techniques. Some recent attempts to flip-chip VCSELs to CMOS are summarised in Table 2-2. Attempts to integrate VCSEL arrays which do not require substrate removal have been reasonably successful but problems remain with devices that require substrate removal at 850 nm. Other integration techniques have also been considered for attaching VCSELs to silicon circuits [94][95].

| Group | bonded to | wavelength | speed | yield | size |
|------------------------|--------------|-----------------------|------------|---------|---------------|
| Lucent [96] | foundry CMOS | 970 nm | 1 Gbit/s | unknown | 2×10 |
| Army Research Lab [97] | foundry CMOS | 950 nm | slow | 253/256 | 16 	imes 16 |
| MIT [98] | foundry GaAs | 850 nm | not tested | 30-50% | 8×8 |
| NTT [99] | Al N | 850 nm ⁽¹⁾ | 2.6 GHz | 100% | 8×8 |

⁽¹⁾ AlGaAs substrate – no substrate removal required

Table 2-2: Recent work on flip-chip integration of VCSEL arrays

Unlike a MQW modulator, a VCSEL cannot behave as an effective detector; in a VCSEL based system, a second type of device must be integrated to provide both input and output capability. One approach is to fabricate MSM detectors on the same substrate as the VCSELs [95][100]. A second approach is to use separately fabricated *pin* photodiodes. Although the simultaneous flip-chip integration of *pin* diodes and VCSELs to a single
CMOS circuit has yet to be achieved, it has been demonstrated that it is possible to perform successively bonding of two different arrays of optoelectronic devices onto the same circuit by using different solders for each flip-chip operation [101].

VCSEL arrays have been used in a number of experimental and commercial parallel data link applications [102][103][41].

2.5 Optical packaging

Two main approaches have been used to form the connection between source and detector in multiple-channel optical data links: fibre-ribbons and free-space optics. This section briefly reviews the physical construction of optical links based on these two approaches and explains the motivation behind considering the free-space approach.

2.5.1 Fibre-ribbon interconnects

Parallel fibre interconnects operate on the same principle as single-channel fibre links but require more complex packaging schemes and have so far only been used for a small number of channels. The widest fibre-ribbon data link reported to date uses 40 fibres [104].

A number of techniques for interfacing one-dimensional arrays of fibres to detectors have been devised. For example, one technique [41] aligns the fibres by inserting them into Vshaped grooves etched in a silicon substrate and polishes the ends of the fibres at 45° to couple the light into surface-normal optoelectronic devices (Figure 2-10).



(a) side view

(b) cross-section of fibre-array block

Figure 2-10: OptoElectronic Technology Consortium (OETC) fibre packaging technology [41]

There has been less progress in the fabrication of two-dimensional fibre arrays. In part, this is due to limited demand. Several techniques have been used to interface fibre arrays to free-space optical systems [105][106][107][108]. Small, rectangular two-dimensional fibre-arrays (8×2) have been used to couple fibres directly to devices [109].

Fibre-ribbons are appropriate for communication over medium distances of say 0.1-100 m. Experimental and commercial fibre ribbon data links are extensively reviewed in reference [33].

An important strength of the fibre-ribbon approach to optical interconnects is that, once the ribbon has been accurately aligned with and terminated to the optoelectronic circuit, the waveguiding of the light by the fibre ensures that the link is maintained even in the presence of vibration or environmental disturbances. In this sense, the fibre link can be treated simply as a high density, low-EMI, high-speed alternative to an electrical ribbon cable which is, after all, nothing more than an electrical waveguide; the similarity in usage to existing technology is likely to reduce the barrier to market acceptance of the unfamiliar optical technology.

2.5.2 Free-space interconnects

Free-space optical interconnects use the imaging properties of lenses rather than the guided wave technique used by fibres to control the propagation of the optical signals. The operation of a free-space optical interconnect is shown schematically in Figure 2-11 for an emitter based link. The optical system of two lenses images the object plane (a two-dimensional array of emitters) onto the image plane (a two-dimensional array of detectors). The first lens collimates the beams produced by the emitters so that they can propagate over an appreciable distance with minimal divergence. The second lens focuses the beams onto the detector array.





The length of a free-space link of this form is typically four times the focal length of the lens; it is commonly referred to as a '4-f system'. The focal length of a typical bulk-optics lens used in this application would be in the region 10-100 mm; thus, free-space optics is potentially applicable for interconnects over distances of the order of 10 cm. Systems using micro-optics instead of bulk-optics could achieve interconnects over shorter distances. These distances are comparable with the requirements of board-to-board interconnects inside large computer systems.

Maintaining optical alignment is much more difficult in a free-space interconnect. Alignment of the beams must be maintained over the entire link. Nevertheless, experimental systems have shown that, with careful optomechanical design, it is possible to construct systems with long-term mechanical stability [110]. Closed-loop active control of optical alignment has also been considered [111]. Optomechanical design issues are reviewed in [74].

These alignment issues make free-space optical interconnects inherently more complicated and more expensive than fibre-ribbon optical interconnects. For data links with only a few channels, it is likely that fibre-ribbons will remain the optical interconnect technology of choice. As the number of channels required by the application increases, the cost differential between fibre and free-space will decrease and a point may be reached, speculatively at around 100 channels, where free-space optics becomes an attractive alternative. Speculating also that the upper limit on channel rate for low-cost fibre is of the order of 10 Gbit/s, the break even point for free space comes when the interconnect bandwidth approaches 1 Tbit/s.

2.5.3 Example of a system using free-space interconnects

A photograph of an experimental free-space optical system [77][112][113] is shown in Figure 2-12.This system was developed as part of the SCIOS project⁵. It was designed to investigate free-space interconnects based on flip-chip integrated MQW modulator/CMOS technology in the context of a sorting application. The system contains two flip-chip bonded CMOS/InGaAs MQW circuits, one mounted on each circuit board. Each circuit contains 1024 differential modulator pairs and 1024 differential detector pairs that have been designed to run at a channel rate of 100 Mbit/s. A diode-pumped Nd:YLF laser provides 1 W of optical power at 1047 nm. Multi-element lenses in cylindrical barrels provide the interconnect between the two chips. Two CCD cameras are used to view the devices during the alignment of the optics. Beamsplitter cubes introduce the beam from the Nd:YLF laser (as well as additional illumination for alignment purposes) into the optical system.

The author was responsible for the design of the electrical and thermal packaging of the hybrid chips in this system, including the layout of the printed circuit boards. The packaging scheme was in part based on the scheme used in the previous generation SCIOS system [74] and the scheme used by Lucent Technology for packaging of FET-SEED devices [70]. Several of the packaging requirements were particular to chips forming part of a free-space optical interconnect: a 'cavity-up' style of packaging had to be used to allow mechanical access of the objective lens to the chip which had a front working distance of only 3.5 mm; the temperature had to be regulated to approximately 55°C to obtain good modulation

⁵ SCIOS stands for the Scottish Collaborative Initiative in Optoelectronic Science and is an EPSRC funded project involving Heriot-Watt University, University of Glasgow, University of Edinburgh and University of St. Andrews.

performance at the design wavelength. At the same time, the packaging requirements of a high-frequency electronic integrated circuit had to be met including high-speed electrical signalling (separate 16-bit wide 100 Mbit/s input and output buses plus clock signals) and effective high-frequency power supply decoupling.

Closed-loop control of the chip temperature was achieved using a thermoelectric cooler and a thermistor. The high power dissipation of 10 W and 4 W for the two circuits, together with the low maximum permitted junction temperature created a relatively challenging thermal management problem. Commercially available thermally enhanced chip carriers were unsuitable because they used a cavity-down geometry. In production volumes, a custom ceramic carrier would provide the best technical solution but was prohibitively expensive for a one-off prototype. Instead, the chips were mounted directly on a copper heat-spreader, which was inserted through a hole machined in a circuit board; the chips were then bonded directly to selectively patterned nickel-gold bond pads on the board. A thermoelectric cooler and heat sink were clamped to the back of the heat-spreader. This chip-on-board packaging technique provides many of the benefits of a custom carrier at a lower cost and has excellent high-frequency performance. The circuit boards used a standard four layer epoxy-fibre glass construction with only a small cost premium for the nickel-gold bond pads. Bonded electrical test versions of the chips are shown in Figure 2-13. Chip-on-board mounting techniques for free-space optical packaging were independently developed at McGill University [114] at around the same time.



Figure 2-12: An experimental free-space optical system



(a) shift register circuit (10 W power dissipation, 16-bit wide input and output buses)



(b) sorting node circuit (4 W power dissipation)

Figure 2-13: Chip-on-board packaging of electrical test versions of hybrid CMOS/InGaAs circuits. The smaller surface-mount components are bypass capacitors; the larger ones are terminating resistors. Notice the holes for mounting the

heat-sink and the optomechanical mount.

2.6 Conclusion

This chapter has reviewed the technology available for implementing multiple-channel optical interfaces to mainstream VLSI electronics.

It is primarily intended to place in context the work that follows in later chapters. The experimental characterisation of surface-normal MQW modulators in the InGaAs/AlGaAs material system is of significance primarily as an illustration of the performance available from this class of device; the results on high-speed modulation to 500 Mbit/s are also important insofar as they are the first experimental verification that the devices in this fabrication process can be modulated at speeds that are useful for optical interconnects. The use of chip-on-board packaging for low-cost, high-performance prototyping of hybrid optoelectronic circuits is also of note.

The integration and device technology has been shown to be relatively mature. Neither area presents a barrier to the use of the technology in a commercial application: hybrid CMOS-MQW modulator devices could be used to construct a commercial terabit/s scale system today. Nevertheless, some improvement in the flip-chip integration of VCSELs is desirable to permit the use of an optomechanically simpler emitter-based link.

The optical packaging technology for links comprising several hundred channels is less mature: fibre-ribbon links of this scale are yet to be investigated and, whilst prototype freespace systems of this scale have been constructed in the laboratory, progress in achieving straightforward initial system alignment and in maintaining this alignment under standard environmental conditions with low-cost optomechanics is required for the systems to be commercially feasible.

The remainder of this thesis examines issues in the design of the electronics, in particular the photoreceiver circuits, which are specific to systems employing the technology discussed in this chapter. The need to integrate a large number of receivers onto a single chip imposes several design constraints on power consumption, layout area and crosstalk which are less relevant to single channel receivers. Whilst this problem is much less fundamental than that of cost-effective optomechanical packaging, nevertheless there is not yet a complete understanding of the issues that affect receiver design in large arrays; this thesis makes progress in addressing some of the outstanding questions in this area.

Although much of the work has been carried out in the context of a free-space system, which is described in the next chapter, the majority is equally applicable to fibre data links with a similar number of channels.

Chapter 3

Design of a free-space optoelectronic crossbar

3.1 Introduction

The previous chapter has reviewed the optoelectronic device and hybrid integration technology that can be used to provide high-bandwidth optical interfaces to VLSI circuits.

In this chapter, the design of an experimental system that uses this technology is described. This system is a 250 Mbit/s 62×64 packet-switched crossbar which contains, internal to the crossbar, an optical interface with a capacity of 1 Tbit/s.

A switch was selected as the focus for the system because switching is typical of the applications which might require very high bandwidth interfaces. A switch performs only limited processing: its main function is to route data. Its cost can be dominated by the cost of the interconnect rather than that of the processing electronics. Switching is therefore one of the most promising applications for optical interconnect technology.

However, the system described in this chapter does not attempt to meet the needs of a specific switching application; rather, it had the more general aim of developing the technologies associated with high-bandwidth optical interconnects. In the course of the system design, advances have been made in understanding the issues that affect receiver design in large arrays, in the design and fabrication of arrays of both VCSELs and InGaAs MQW modulators and detectors and in free-space optomechanical design and packaging. In particular, the core of the work presented in this thesis has been a direct or indirect result of the design process.

The main component of the system, a hybrid CMOS-InGaAs MQW circuit that implements the crossbar functionality, is an example of the specific class of optoelectronic VLSI circuits generally referred to as 'smart-pixel arrays' [115]. Like any optoelectronic VLSI circuit, a smart-pixel array consists of an array of optoelectronic devices and some digital electronics. The characteristic that distinguishes a smart-pixel array is that the digital electronics is physically interleaved with the optoelectronic devices – the circuit consists of a regular array of 'pixels' which contain a detector and/or optoelectronic output device, some interface electronics and some logic. Commonly, each pixel is physically small (50-200 µm) and has

limited functionality¹. For circuits having a layout which maps naturally onto a twodimensional array and only requires local routing, the smart-pixel approach in theory permits operation at high clock frequencies by reducing the importance of RC interconnect delays, which scale as the square of the length of the wire. The optical interface covers the majority of the chip and requires optics that can cope with large fields. This approach can be contrasted with the 'photonic-interface-module' approach [148][149] in which the optoelectronic interface is located in a separate area of the chip and connected to the digital electronics using on-chip, non-local wiring. The latter permits a smaller optical field but is subject to the RC delays and real-estate requirements of the non-local interconnect.

This chapter, as well as giving an overview of the system to set the work on receiver design in later chapters in context, also provides an insight into some of the design issues which arise in implementing large-scale optoelectronic VLSI circuits, in particular those based on the smart-pixel approach. Although many groups are developing optical interconnect technology, only a few [116][158][148][161][164] are involved in constructing systems of a realistic scale in terms of number of channels and overall interconnect capacity. The experience of participating in the design of this system and of the SCIOS sorting demonstrator described in Chapter 2 has borne out the belief that many of these issues do not become apparent until the detailed work required to actually construct a system is carried out.

The chapter begins by giving an overview of the system architecture and its optical implementation. The architectural and physical design of the switching circuit are then described in more detail. Aspects of the performance of other system components which impact on the design of the silicon interface circuitry are briefly considered. The chapter concludes by describing progress towards realising the design experimentally and evaluating its significance in relation to other recently reported systems.

¹ The term 'smart-pixel' is potentially misleading because it has connotations of display applications. Spatial-light-modulators for display and analogue imaging processing applications are also referred to as 'smart-pixels'. However, the term is commonly applied in its more general sense to systems such as this one that have nothing to do with displays.

The switching system described in this chapter is the focus of the SPOEC research project²; several others have contributed to the work presented in this chapter. The author was responsible for the design of the receiver and transmitter interface electronics, had a major input to the high-level system design and specification and proposed the modulator drive scheme described in Section 3.4.1 as a means of simplifying the InGaAs fabrication process. The detailed digital design and layout was performed by Philippe Benabes and Alain Gauthier of Supélec Paris.

At the time of writing, the system is yet to be assembled and tested in the laboratory. Much work remains before the project goal of an experimental demonstration of a terabit/s interface is achieved. However, in terms of the understanding of the component technologies that it has provided, the design work described herein has already been of value in its own right. If, in addition, the system assembly proceeds according to plan, the work will also have significance in its contribution towards achieving the wider research objective of the project: demonstrating the feasibility of high bandwidth optical interfaces to VLSI electronics.

3.2 System description

3.2.1 Overview of system architecture

The SPOEC system implements a packet-switched crossbar between N input ports and M output ports. Each of the M output ports uses a separate N:1 multiplexer to select the data from one of the N input channels. Each input channel is thus fanned out by a factor of M between the input port and the switch and fanned in by a factor of N at the output port. This scheme is illustrated in Figure 3-1 for a 4×4 crossbar. The overall throughput of the switch is N times the channel rate; however, internal to the switch, this method of implementing a crossbar requires an interconnect with a capacity of NM times the channel rate, although obviously not all the data in this interconnect is independent.

² SPOEC stands for "Smart-pixel optoelectronic connections" and is part of the "Optoelectronic Interconnects for Integrated Circuits (OPTO)" cluster of projects funded by the European Commission through the Advanced Research Initiative in Microelectronics (ESPRIT MEL-ARI) programme. The institutions involved in the SPOEC project are Heriot-Watt University, Supélec Paris, Supélec Metz, University of Glasgow, CSEM Zürich and Trinity College Dublin.



Figure 3-1: 4 × 4 crossbar implemented using fan-out and fan-in

The SPOEC system was originally intended to have N=M=64; however, two input channels were allocated for clock distribution and the final system is actually a 62×64 crossbar.

The system uses optics to implement the high-bandwidth internal interconnect. The very high internal bandwidth required by this architecture makes it attractive to use as a test vehicle for the interconnect technology. Test equipment that is capable of generating high speed data channels is expensive; this architecture allows a relatively modest amount of hardware (62 input channels) to be used to fully exercise an interconnect with a much larger capacity $(62 \times 64 = 3968 \text{ channels})$. The fact that not all the channels are independent does not in any way ease the requirements on the interconnect technology. The target channel rate of the system was 250 Mbit/s giving an aggregate bandwidth at the optoelectronic interface of approximately 1 Tbit/s. This is in the region where free-space optical interconnect might arguably become preferable to an electrical interconnect over moderate distances as discussed in chapter 1. The primary merit of this system is in its ability to act as a test vehicle in this way; no claim is made that the architecture in its own right is a good way to implement a large crossbar.

The output multiplexers are configured on a packet-by-packet basis using address information encoded in packet headers in the input data stream. The output ports also contain logic to arbitrate between inputs that simultaneously request access to the same output port.

3.2.2 Optical implementation

The internal interconnect of the crossbar is implemented using a free-space optical link that employs VCSELs as the transmitter technology, free-space bulk-optics as the interconnect medium and InGaAs MQW pin diodes as the detector devices. The switching functionality is implemented in a custom CMOS circuit and integrated to the optoelectronic devices using the flip-chip integration technique described in chapter 2. Optical outputs from the switching circuit are provided by operating the InGaAs pin diodes as QCSE modulators.



The optical layout of the system is shown in Figure 3-2.

Figure 3-2: SPOEC system layout (after [117])

The inputs to the system are provided by an 8×8 array of 960 nm top-emitting VCSELs. The main optical interconnect between the VCSEL array and the switching chip is a 4-f system consisting of the hybrid lens formed by microlens array 1 and bulk-lens 1, diffractiveoptical-element DOE-1, polarising beamsplitters PBS-A and PBS-B and bulk-lens 2. For the input beams at 960 nm, the two beamsplitters are designed to operate as polarisation independent mirrors. The diffractive optical element [118] at the Fourier plane [119] accomplishes the required 1 to 64 fan-out optically³, replicating the two-dimensional array of input beams into a larger 8×8 array of 64-beam groups (Figure 3-3). The main switching chip thus contains a total of 4096 detectors.



(a) input 8x8 VCSEL array(b) switching chip detector arrayFigure 3-3: Organisation of the optoelectronic devices on the switching chip

Each of the 64 smaller arrays corresponds to one of the output ports of the switch. The electronics associated with a single detector is referred to as a "pixel"; each array of 64 pixels is referred to as a "super-pixel". Each pixel contains the analogue interface circuits that convert the detected photocurrent into a standard CMOS logic signal while each super-pixel contains the digital circuits implementing the routing and arbitration logic. Two of the VCSELs supply a differential clock signal to the super-pixel electronics via a dedicated clock receiver circuit. The receiver circuits for the data and clock channels are discussed in detail in chapters 5 and 7 respectively.

³ The fan-out exploits the ability of a lens to perform the two-dimensional Fourier transform [119]. The VCSEL input plane is Fourier transformed by the combination of the microlens array and lens 1; the result is produced in the plane of the diffractive optical element. The DOE has a phase-profile that implements a two-dimensional frequency domain filter; the phase profile or 'transfer-function' of the DOE is an approximation to the Fourier Transform of the desired impulse response which in this case is an 8×8 array of delta functions with the same pitch as the super-pixel; multiplication by the transfer function in the Fourier plane results in a convolution with the impulse response in the plane of the switching chip. Lens 2 completes the 4-f system by transforming the frequency domain representation back to the spatial domain.

The output of each port is converted back to an optical signal using a differential pair of modulator diodes located within each super-pixel. The 64 output channels are relayed to the third chip in the system, which is again a hybrid CMOS-InGaAs device. This second interconnect operates at 1047 nm. The optical source is a Nd:YLF laser. Unlike the VCSEL array, it has a well controlled polarisation: a standard polarisation routing technique (see, for example, references [120][121]) is used to control the path taken by the beam. At 1047 nm, the beamsplitters are designed to reflect light of s-polarisation and transmit light of p-polarisation. The correct optical path through the system is obtained by using two passes through a quarter-wave plate to change the p-polarised input beam to s-polarisation, and a single-pass through a half-wave plate to change the light reflected by PBS-A back to p-polarisation. The output chip converts the optical signals back to electronic form for testing purposes.

The physical dimensions of the optomechanical system are approximately 30 cm \times 20 cm \times 15 cm.

3.3 Implementation details

3.3.1 Behavioural description

This section describes in more detail the high-level architecture of the switch and its operation.

The input data streams are bit serial and consist of a sequence of packets. The packet format is shown in Figure 3-4. Each packet consists of two sections: a header and a payload. The header encodes a six-bit destination address A < 5:0> and a flag F to indicate whether or not a packet is present.





The switch operates in two phases: an address phase and a data phase. During the address phase, each output port examines the addresses of all packets arriving at its inputs, determines which packets have a destination address that matches its own address and, if there is contention between several inputs, makes a decision on which one to route. During the data phase, the data from the selected input is routed to the output.

The duration of the data phase and hence the length of the packet is controlled by a global synchronisation signal *SYNC*!, which is provided as an electrical input to the switching chip.

During the address phase, the system requires the full capacity of the optical interconnect. For the system to operate correctly, all of the data receivers in the array must operate simultaneously without error for the duration of the header. However, during the data phase, only the data receiver corresponding to the selected input channel needs to operate. The power consumption of the system can be reduced by disabling the unused input channels. The modifications to the data receiver design that were required to implement this feature are discussed in Chapter 5. The header gap and cell gap in the timing diagram are required to allow the analogue receiver circuits to stabilise after being enabled or disabled.

Packets from inputs that are denied access to the output port are not buffered by the switch. The system therefore relies on buffering at the input to the switch to provide a reliable transport. To provide the same switching performance as a classic input-queued switch [122][123], information on the results of the arbitration process must be communicated back to the input queues so that successfully transmitted packets can be removed from the queue. A simple scheme that could be used to implement this function within the duration of a single packet was devised (see Appendix 3.7). However, for simplicity, this feature was not implemented; thus, the switching fabric provides an unreliable transport. A higher level protocol must be used to accomplish the handshaking function of detecting and re-transmitting dropped packets.

With random traffic, an input-queued switch saturates as the arrival probability at the input ports approaches 0.586 [122]; thus, the system would only be suitable for use in a telecommunications or local-area-network type application if the line rate was some fraction of the channel rate used internally inside the switch. However, the architecture would provide reasonable performance as an interconnect for a multiprocessor system running an application where the communication pattern is regular and the probability of contention is low. A number of other groups [124][125][126][158][127][128][129] have investigated optoelectronic switch designs that employ similar technology but are based on the Growable Packet Switch architecture [130][131] and would be more suitable for building large telecommunications switches.

40

3.3.2 Logical implementation

This section briefly describes the structure of the digital electronics used to implement the functionality just described. More details about the digital electronics that implement the contention resolution function can be found in references [132][133] and [134]. A block diagram of each output port is shown in Figure 3-5.

The address comparison is performed bit-serially on all input channels in parallel. The comparator is implemented using a single D-type flip-flop. At the end of the address phase, each address comparator produces a signal MATCH < n > that indicates whether or not the packet arriving on its associated input port has an address which matches the address of the output port. The arbitration logic uses this information to select one of the inputs for routing to the output. The priority of input channels that simultaneously request routing to the same output port is varied as a function of time by means of a priority generation state machine. For simplicity, the state-machine used was a simple binary counter and provides round-robin access to the output port. Better switch performance could be obtained using a pseudo-random priority sequence; this could be easily implemented using a linear-feedback shift-register with approximately the same amount of logic. Both schemes give all input channels equal priority averaged over a large number of packets.

The arbitration logic and multiplexer is implemented using a 6-level tree of two-way arbitration blocks and 2:1 multiplexers. Part of this structure is shown in Figure 3-6. Each arbiter receives two *MATCH* signals from the previous level of the tree that indicate whether any of the channels above this point in the tree have a packet destined for the output node. The address comparator blocks form the leaf nodes of the tree and provide the *MATCH* inputs to the first level of arbiter blocks.

Each arbiter sets the local 2:1multiplexer select signal according to which, if any, of the two inputs have a packet available; in the case where both inputs have a packet available, an input is selected based on the *PRIORITY*<*k*> signal which is shared amongst all the arbiters at the *k*th level of the tree. If either input contains a valid packet, a *MATCH* signal is passed on to the next level of the tree.



Figure 3-5: Structure of output port logic

The example shown in Figure 3-6 illustrates the case where inputs IN0, IN1, IN4 and IN6 have packets with addresses matching the output node. At the first level of the tree, IN1 is chosen over IN0 because *PRIORITY*<0> is 1; at the second level of the tree, IN6 is chosen over IN4 because *PRIORITY*<1> is 1; at the third level of the tree, IN1 is selected to provide the final output because *PRIORITY*<2> is 0.

The logic to implement the arbitration and multiplexer tree is fully asynchronous and can be implemented with low fan-out gates; it is therefore potentially extremely fast. The six levels of the multiplexer add propagation delay to the input signal but do not limit the data rate that may be routed. The propagation delay at all output ports should be approximately the same. Even if synchronous routing were employed, the tree structure can be easily pipelined by adding latches after every two or three levels of the tree.

Note that this scheme does not explicitly generate the multiplexer select signals SELECT < 0.5>; rather, the signals are distributed throughout the tree.



Figure 3-6: Detailed implementation of multiplexer arbitration logic (first 3 levels of tree shown for first 8 inputs)

3.3.3 Timing conventions

The input channels are assumed to be fully synchronous at both the bit and packet level. This assumption requires that some means of synchronising the input sources is available. For the purposes of experimental operation of the system, this synchronisation is achieved by manually adjusting the delay on each input channel using a variable delay control on the test equipment. In a practical system, achieving synchronisation between boards at a clock frequency of 250 MHz is more difficult, but a number of suitable techniques exist [135][136][137]. The effort required to implement system-level clock distribution constitutes a major disadvantage of this architecture as a practical means of implementing a crossbar; however, it does not detract from its strengths for its intended purpose of demonstrating a terabit/s scale optical interconnect.

Clock distribution within the main switching chip is achieved by using two of the VCSEL channels to carry a differential optical clock signal. The optical fan-out used to replicate the input data channels also accomplishes the fan-out of the clock. Although the optical fan-out itself is intrinsically free of skew, some skew is introduced by the clock receiver circuits used to interface the optical signals to the electronics. The worst-case skew between two super-pixels is estimated to be about \pm 300 ps (see chapter 7). However, the super-pixels operate independently and the only consequence of this skew is a lengthening of the cycle time – the chip does not use any delay-critical global signalling.

Electrical clock distribution at 250 MHz over a 14.2 mm \times 15.6 mm chip is a non-trivial problem, but would have been far from impossible with careful design. No suggestion that chip-level optical clock distribution can be justified over an electronic solution in its own

right is intended. The primary motivation behind using optical clock distribution in this system was to shift design effort away from the well-understood problem of electrical clock distribution to another with more relevance to the general research area of the project. The clock receiver design is of interest in its own right and, in addition, the problem of supplying an optical clock in parallel with a wide bus of optical data channels is of wider interest.

For example, consider the switching nodes in the interconnection network of a large multiprocessor system, in which each processor has an independent clock domain, and in which each logical channel in the interconnection network is implemented as a parallel bus of several physical channels. Each switching node must synchronise with data from a number of independent sources. One approach to this might be to recover a clock from one of the physical data channels and use this to retime the data. However, if the bus is wide, then the overhead of including a clock in parallel with the data is low and the need for clock recovery can be avoided by using this clock to retime the data.

The solution to the clock receiver design problem gives some insight into how to implement a receiver front-end capable of detecting a return-to-zero clock signal with the same optical power per photodiode but twice the signal bandwidth requirement as the non-return-to-zero data channels in the same link. This is discussed in more detail in chapter 7. However, the design does not fully address the example retiming problem just described; specifically, the issue of maintaining correct phase alignment of the clock channel relative to the data is not considered. In addition, there are several other approaches which might avoid the need for this extra bandwidth such as the use of double-edge triggered flip-flops [138] and frequencydoubling phase-locked-loops.

3.3.4 Physical implementation of the switching chip

The switching circuit was implemented in a 0.6 μ m CMOS technology with two levels of metal. The chip had overall dimensions of 14.2 mm × 15.6 mm and contained approximately 580 000 transistors. It was packaged in a 256-pin ceramic pin-grid array carrier; most of the pins were required for the analogue and digital supplies. The detailed digital design and physical layout of the chip was carried out by Supélec Paris.

The floorplan of a super-pixel is shown in Figure 3-7. The super-pixel consists of rows of standard-cell logic interleaved with full-custom analogue receiver circuits and flip-chip pads. A combination of manual and automatic cell placement and routing was used. The pixel pitch was 149.5 µm and the super-pixel pitch was 1614.6 µm.

The structure of the contention resolution circuits is suitable for layout within a pixellated array. Each address decoder is associated with a single pixel. The first two levels of the multiplexer tree can be split into sixteen independent groups, each consisting of four input channels, three multiplexers and three arbiters. These groups contain 75% of the gates in the entire tree. Much of the wiring is local to these groups; wiring global to the super-pixel is required only for the lower levels of the multiplexer tree, the bit-serial address signal, the priority signals, the clocks and some test signals. The grouping of the logic into sixteen half-rows of four inputs is clearly visible in the floorplan.



Figure 3-7: Floorplan of a super-pixel

The integration of analogue and digital circuitry on the same chip means that substratecoupling of digital noise is a potential problem [139][140]. However, the use of a process with a lightly doped p-type epitaxial layer on a degenerately doped bulk will have helped to improve the isolation between nearby circuits [141]. A number of general precautions were taken to minimise the problem including the use of guard rings in the analogue layout to provide low resistance substrate contacts. However, no quantitative analysis of the problem was attempted because of the difficulty of analysing the problem by hand. Software tools suitable for tackling this analysis have recently become available [142][143].

The power supply distribution network created one of the main layout difficulties. Separate power supply rails were required for the analogue and digital circuits to prevent switching transients in the digital supply current from upsetting the operation of the sensitive analogue circuitry. A separate power supply was also used for the receiver front-end to control the electrical crosstalk between receivers (see chapter 8). The scheme used is illustrated in Figure 3-8 which shows the detail of the two metal layers for an example pixel. The horizontal routing layer is metal 1 and the vertical routing layer is metal 2. The analogue power supply lines run vertically; two pairs of the front-end supply rails (wide) and one pair of the decision stage supply rails (narrow) are visible. The digital power supply lines run horizontally at the top and the bottom of the two rows of standard cells. The flip-chip pads are in metal 2. Bypass capacitors formed by thin-oxide capacitors are placed underneath the front-end analogue power supplies in both rows of standard cells. It can be seen that, for a two-metal process, the need to route multiple, separate power supplies leads to inefficient area usage in the routing channel and the standard cell rows, although some of the unused area underneath the analogue power supply rails in the upper standard cell row is recovered by a bypass capacitor.





In retrospect, this may not have been the best power supply distribution scheme. An alternative strategy would have been to replace the lower standard cell row with a row dedicated exclusively to the analogue receiver circuits and to run the analogue power supply rails horizontally. The bypass capacitor in the upper row could then be exchanged with the digital circuitry in the lower row. The most important advantage of this alternative is that it would make it easier for the layout of the analogue and digital circuitry to proceed independently. In addition, a more compact layout of the analogue cells would be possible because they would no longer be constrained to fit exactly within the height of a digital standard cell and could omit the horizontal digital power supply rails.

To a certain extent, the difficulty in achieving an efficient power-supply distribution layout is a consequence of the limited number of metallisation layers in this process. Dedicated power-distribution layers for the analogue and/or digital supplies, routinely available in state-of-the-art silicon processes with three or more levels of metallisation, would significantly reduce the overall size of the chip and hence relax the demands on the performance of the optics.

The requirement to achieve a regular pixel pitch with the same spacing in both x- and ydirections (to obtain a match to a VCSEL array with a standard pitch) also reduced the efficiency of the layout that could be achieved with a standard cell approach. The number and width of the cells determined the horizontal pitch; the vertical pitch was constrained to be the same even though not all the space in the routing channel was required. Designing a custom set of standard cells with a more appropriate aspect ratio may have permitted a more compact layout, but would have required significantly more design effort. An alternative approach, which was not explored in any detail, would have been to use a transistor-level layout synthesis tool [144] to automatically generate a layout with the required aspect ratio from the schematics. At the very least, this technique would allow different options for the aspect ratio to be evaluated very quickly to determine a good choice before proceeding with the custom layout of a set of standard cells. The extent to which this would improve the overall density depends on how much optimisation has already been carried out on the existing standard cell library.

Previous smart-pixel circuits have tended to use full-custom layout techniques [125][145]. On balance, the improvement in design productivity offered by the standard-cell approach outweighed its disadvantages in terms of layout density in the context of this project; it is doubtful whether a circuit design of this complexity could have been realised within the resource constraints of the project using a full-custom approach.

Alternative approaches to automating the design of smart-pixel circuits are currently being investigated by some groups. The use of an upper layer of metallisation to route the optical input signals to separately located receiver circuits, with free placement of digital logic underneath the flip-chip pads, has been proposed and implemented in experimental designs [146]; enhancements to existing CAD placement tools are being developed to automatically place receiver circuits close to the flip-chip pads amongst the other digital cells [147]. An outstanding problem with this approach is the susceptibility of the sensitive analogue inputs connected to the flip-chip pads to electrical crosstalk from the underlying digital circuits. Deliberately degrading the sensitivity of the receivers might be one way to provide an adequate noise margin.

Another practical difficulty that was encountered in the design process was the interdependence between the layout of the silicon and the design of the optics. The area occupied by the silicon circuitry determined the detector pitch and hence the optical field size and magnification. This made it difficult, but not impossible, for design work to proceed in parallel. This interdependence would seem to be a general feature of systems using the smart-pixel approach in which the logic is localised to the detector positions.

The interdependence could have been avoided by specifying conservative values for the detector pitches at an early stage in the design process. The disadvantage of this method is inefficient utilisation of both the silicon area and the optical field. Indeed, in this particular system, the fact that both the optical field of view (17.5 mm diagonal) and the dimensions of the silicon circuit were close to their practical limits made this approach unworkable.

This interdependence in the design process is something that would add cost to any system that uses the smart-pixel approach, either by requiring custom design of optics and optoelectronic devices for every system or, perhaps more realistically, by inefficient utilisation of silicon through the adoption of a standard pitch.

An alternative approach for using optical interconnect technology is the 'photonic interface module' [148] [149] in which a densely packed array containing only transceivers and wiring is used to provide the optical interface, and the digital logic is located in a separate area of the chip and designed using standard CAD tools. A chip based on this approach [148] contained 504 receivers in a 1 mm \times 1 mm area in 0.8 µm technology. The advantage of the technique is that a self-contained design with a standard pitch can be designed once using a full-custom, carefully optimised layout. Once proven experimentally, such a block could, in principle, be reused with minimal effort by people without specific expertise in optical interconnect technology. The technique also permits dense packing of the optoelectronic devices thus significantly reducing the cost of the optical packaging, arguably the most important barrier to the commercial adoption of free-space optics. Nevertheless, the approach has certain drawbacks. The on-chip interconnects between the interior of the transceiver array and the digital logic have a power consumption and layout area overhead. It is also possible that the RC delays in these electrical interconnects might ultimately limit the overall capacity of the optical interface: more work is required to establish how well the 'photonic interface module' approach scales to aggregate bandwidths of several terabits/s.

3.4 Other system components

In this section, the main characteristics of other system components which impact on the design of the interface electronics are briefly discussed.

3.4.1 InGaAs detectors and modulators

Although the InGaAs modulator/detector device structure was essentially the same as that of the devices described and characterised in chapter 2, a significant effort was made by the

group at the University of Glasgow to simplify the fabrication process to overcome the problems with InGaAs yield that had previously been encountered.

Changes to the interface circuits used to implement a two-beam modulator driver and a twobeam receiver, proposed by the author, helped to achieve this process simplification by eliminating the requirement for two levels of tracking.

The new modulator drive scheme is shown in Figure 3-9.



(a) conventional differential modulator drive scheme(b) modulator drive scheme with a single bias voltage

Figure 3-9: New modulator driver with a single bias rail

In the new scheme, separate digital driver circuits are used to drive the two p-type contacts with the true and complementary data. Compared to the conventional approach of using two series-connected modulator diodes with a single driver [150], this approach permits the use of a single, common modulator bias voltage. A secondary benefit is that it results in first-order cancellation of the switching noise on the modulator bias rail, although the switching transient on the silicon modulator driver supply remains in both schemes.

A disadvantage of this approach is that the modulator driver transistors must be sized to sink the total photocurrent through each diode, which is larger than the difference in photocurrent between the two; consequently, the power consumption of the new scheme is higher. In this system, the inverter that drives the modulator devices used minimum length transistors with widths of 24.8 μ m and 44.8 μ m for the NMOS and PMOS transistors respectively. The size was determined by the relaxed DC specification for the maximum photocurrent (500 μ A) rather than by the dynamic requirement to drive the modulator capacitance. The power consumption with random data at 250 Mbit/s was simulated to be 2 mW. The electrically differential design that was chosen for the clock receiver also permitted the use of a single bias voltage for the detectors. In contrast, the most commonly used two-beam smart-pixel receiver design (see chapter 4) uses a series-connected pair of photodiodes configured similarly to Figure 3-9 (a) and consequently requires two bias voltages. The detector bias voltage used for the clock receiver circuit was shared with the data receiver circuits; however, the modulator and detector bias voltages were separated to avoid electrical crosstalk.





Figure 3-10: Scanning-electron-micrograph of InGaAs devices fabricated with the new InGaAs process. The left-hand picture shows four detectors and a modulator after the mesa-definition etch; the right-hand picture shows a completed modulator with isolation trench and lower modulator bias line (Courtesy of University of Glasgow).

With these modifications to the interface electronics, the InGaAs processing sequence can be reduced to four mask steps: combined mirror/upper p+ contact deposition; mesa definition etch; lower n+ contact deposition and lift-off; and trench isolation. In contrast, the previous process required six steps: mesa definition; mesa isolation; p+ contact deposition; n+ contact deposition; via layer deposition and etch; and mirror metallisation. Photographs of prototype devices fabricated by the University of Glasgow using the new process are shown in Figure 3-10. Glasgow have recently proposed a refinement to this process that requires only three mask steps [151].

The common detector bias voltage takes the form of a plane of metal that is broken only by the modulator bias line which runs vertically from the top to the middle of the chip (Figure 3-11). This provides a low-inductance and low-resistance bias to the detectors which is important for controlling crosstalk. It is also tolerant to open-circuit failures. Isolated detector bias planes were used for each half-column of four super-pixels to localise short-circuits to the detector bias to a single section of the chip.



modulator bias rail

Figure 3-11: InGaAs supply voltage layout for one super-pixel

3.4.2 Optical power budget

The sensitivity required of the receivers in the switching chip was determined by the VCSEL output power together with the optical losses in the link. The VCSELs were specified to provide 1 mW of optical power. The optical loss in the main arm of the system is estimated to be 3.5 dB excluding the 18 dB loss resulting from the 1:64 fan-out. The overall responsivity of the detector is estimated to be 0.54 A/W based on an ideal responsivity at 960 nm of 0.77 A/W and an estimate of the external quantum efficiency of 70%. These figures predict a detected photocurrent of 3.8 μ A per photodiode; the specification on the receiver sensitivity was 3.5 μ A. The measured VCSEL output power exceeded the specification by 25% which provides additional margin.

The channel data rate was determined by the detected photocurrent and the photodiode capacitance.

The detected photocurrent affects the channel rate because of the direct trade-off in the design of the receiver between speed and sensitivity. This trade-off will be discussed in detail in chapter 4.

The photodiode capacitance is determined by the detector diameter. In general, this is in turn determined by the resolution of the lenses in the link and the alignment tolerance. However, because a diffractive element is used in this system, the wavelength tolerance on the VCSELs was also important. At the edge of the field, the VCSEL divergence together with the performance of the optical system gave a spot-size of 20 μ m. The VCSEL wavelength variation came from both intrinsic device-to-device variation and thermal effects. For the device arrays that will be used in the final system, the wavelength variations were 0.25 nm and 0.3 nm respectively, which translates into a maximum position error of 5 μ m at the edge

of the field. The intrinsic variation is the variation in emission wavelength across a single array. Best uniformity was obtained by selecting chips from near centre of the VCSEL wafer.

After budgeting for lens manufacturing tolerances and alignment, a detector diameter of $35 \,\mu\text{m}$ was chosen. This gives a photodiode capacitance of 95 fF based on a MQW thickness of 1.2 μ m [152] and a dielectric constant of 13.3 [153⁴].

3.4.3 Optical components

The overall system performance was strongly determined by the optical throughput of the first branch of the system between the VCSEL array and the switching chip. The second branch, between the switching chip and the output chip, used a high-power optical source; there was ample optical power available. Consequently, the first branch of the system was optimised at the expense of the second.

Optimisation of the performance of the first branch had a particular impact on the design of the beamsplitters. Achieving good performance as both a polarisation independent reflector at 960 nm and a polarisation dependent beamsplitter at 1047 nm was difficult because of the small separation in wavelength. However, by sacrificing performance at the less critical wavelength of 1047 nm, adequate performance could be obtained.

A consequence of this trade-off was that the reflectivity and transmittance of the beamsplitter at 1047 nm was a strong function of the angle of incidence in the Fourier plane and hence of the horizontal co-ordinate of the modulator in the switching chip plane. The two modulators of a differential pair are separated by a considerable distance (half the super-pixel pitch or about $800 \ \mu\text{m}$) – simply orientating the switching chip as drawn in Figure 3-11 would result in a reduction in contrast due to the beamsplitter non-uniformity. However, by rotating the circuit by 90° on the optical mount, this reduction in contrast can be eliminated.

The lens performance was also optimised for operation at 960 nm.

The optical design is described in detail elsewhere [154][155][156]. The author's contributions to the optical design were the proposals to include a half-wave plate in the system to permit an identical design for the two beamsplitters and the decision to rotate the modulator pair.

⁴ linear interpolation between values for GaAs (12.91) and InAs (15.15) for $In_{0.185}Ga_{0.815}As$ wells assuming that the barriers have a similar dielectric constant

3.5 Progress towards experimental realisation

The design of the system that has been described in this chapter is mostly complete. In this section, the progress to date towards realising this design experimentally is summarised and the programme of characterisation that is planned for the assembled system is reviewed.

The main switching circuit has been designed and fabricated. The wafer delivered by the commercial foundry is currently undergoing post-processing to prepare it for flip-chip bonding. Electrical testing of the final circuit will begin once the post-processing has been completed. Preliminary electrical tests of earlier prototype receiver circuits are reported in chapters 5 and 7 and have demonstrated that the receiver designs are functionally correct, although the speed of operation appears to have been limited to around 100 MHz by the structures included for testing. Simulations predict that, under typical process conditions, the circuits will operate to around 200 Mbit/s which comes close to meeting the initial design target. However, there are indications that electrical crosstalk may start to degrade performance when the entire array is operated simultaneously; this will be discussed in detail in chapter 8. The design of the output chip is in progress.

The InGaAs detector array has been designed by the University of Glasgow and is currently being fabricated. A prototype device array using the same fabrication process has already been demonstrated.

The input VCSEL array has been fabricated and experimentally tested by CSEM and has been found to meet or exceed the system requirements. Preliminary crosstalk measurements have been performed with a small number of active channels; more comprehensive crosstalk measurements are planned once the VCSEL array is incorporated into the final system.

The optical design is complete and the assembly of the optical system is underway. Preliminary testing of the system will be carried out using a discrete fibre-coupled photodiode in the output-chip plane.

Although a considerable number of steps still need to be completed successfully to achieve full system operation, an important characteristic of the system is that, in general, partial failures of individual components are unlikely to cause a catastrophic failure of the entire system. Many of the critical components such as the receiver designs have already been validated at least in part experimentally. The likelihood of achieving at least partial operation of the system is therefore high, although the only real test of this is experiment.

The system has been designed with testability in mind; a detailed experimental investigation of the performance of the optical interface is planned. The output chip will provide a

capability to select any subset of four output channels for simultaneous observation on an oscilloscope without realigning the optical system. It will also provide a capability to simultaneously measure the bit-error-rate on any four output channels.

The main switching chip also contains circuitry for test purposes. Scan path flip-flops [157] are used for the main components of the design. This allows the multiplexer settings to be configured electrically. It also permits the results of the address comparison to be acquired and analysed after the end of the address phase; this will determine whether or not simultaneous error-free operation of the entire receiver array has been achieved for the duration of the address phase and thus allow the peak capacity of the interconnect to be measured. This test capability will allow the effect of electrical crosstalk through the power supply distribution network of the switching circuit to be monitored as the number of active inputs applied to the system is increased.

The sustained capacity of the interconnect will be more difficult to verify. The main factor that is expected to limit operation of the entire array is the electrical crosstalk through the analogue power supply network. The amount of crosstalk is determined by the number of receivers that are enabled and simultaneously switching. Subject to power consumption constraints, it will be possible to override the amplifier disable circuitry during the data phase (using an external control signal) to achieve continuous operation of the entire receiver array. Under this worst-case crosstalk condition, a test of whether each output channel in turn can operate error-free for a sustained period of time will be possible. Provided there is no realignment of the optics throughout this test, it represents a fairly convincing demonstration of sustained operation of the entire interconnect.

One limitation is that with the test equipment currently available, only partial operation of the input array (24 out of 62 data channels) will be possible. Nevertheless, after fan-out, this represents approximately 1500 simultaneously active optical inputs, which exceeds previous attempts to measure simultaneous operation of high-speed receiver arrays.

Table 3-1 lists some recent free-space optoelectronic VLSI circuits. It can be seen that a number of very high capacity optical interfaces have been constructed, and that the system described in this chapter is one of the most ambitious in terms of number of channels, aggregate data rate and optical field dimensions.

Of particular note is the Lucent Technologies ATM distribution network [158] which has demonstrated operation of individual channels at 622 Mbit/s on a chip with 1024 optical inputs and 1024 optical outputs to give a "potential" capacity of greater than 1 Tbit/s. However, the Lucent system did not permit simultaneous operation of the full array because a separate laser was required for each input. Indeed, there has been very little quantitative investigation of the effects of simultaneous operation of large numbers of channels on biterror rates in very large arrays. The most detailed study of crosstalk effects has been on a 50 channel array [159] and has shown a 2.5 dB penalty for simultaneous operation of the entire array. It is hoped that the system described in this chapter will permit a more detailed characterisation of these effects. Another difference from the Lucent system is that the SPOEC circuit contains more complex digital logic. The Lucent circuit comprises sixty-four externally configured 16×16 crossbars that do not carry out any processing of the packet header information. The process of designing the SPOEC system has provided additional insight into how to approach the integration of complex digital logic with smart-pixel receivers. Nevertheless, unlike the SPOEC system, the Lucent circuit is based on a high-performance switch architecture that is suitable for use in telecommunications applications.

| system | design targets | | experimentally demonstrated | | |
|------------------------------|-----------------------|-----------------|--------------------------------|-------------------------------|--|
| | number of channels | channel rate | optical field diagonal | simultane- ous channels | bit-rate |
| Lucent crosstalk test [159] | 50 | - | 2 mm | 50 | 311 Mbit/s (1) |
| Lucent/UNC SRAM [160] | 96 | - | 2 mm | 96 | 125 Mbit/s $^{\scriptscriptstyle (2)}$ |
| McGill backplane[161][162] | 256 | 300 Mbit/s | 9 mm (3) | in progress | |
| UNC/Lucent SRAM [148] | 512 | 100 Mbit/s | 1.4 mm | ? | ? |
| Lucent ATM switch [163] | 512 | 622 Mbit/s | 3.2 mm | 1 | 790 Mbit/s |
| Lucent ATM switch [158] | 1024 | 622 Mbit/s | 7.5 mm | 16 | 625 Mbit/s |
| SCIOS sorter [116][145] | 1024 | 100 Mbit/s | 8 mm | in progress | |
| OptiComp [164][165] | 1024 | 500 Mbit/s | 23 mm | in progress | |
| SPOEC (this work) | 3968 | 250 Mbit/s | 18 mm | in progress | |
| Lucent ATM switch [126][125] | 4096 | 208 Mbit/s | 8 mm | 896 | 160 Mbit/s ⁽⁴⁾ |

⁽¹⁾ BER testing with 2.5 dB sensitivity penalty for array operation

⁽²⁾ single bit burst test

⁽³⁾ microlens optics with clustered detectors

⁽⁴⁾ qualitative test of parallel operation only. Number of channels limited by fan-out grating.

Table 3-1: Comparison of recent large scale free-space optics receiver arrays

3.6 Conclusion

In this chapter, an experimental demonstrator system based on free-space optical interconnect technology has been described; its intended purpose as a vehicle for improving the understanding of the design issues in systems using high-bandwidth optical interconnects

has been emphasised. The chapter has discussed some of these issues. In particular, the difficulties that have been encountered in using a standard-cell layout approach in a 'smart-pixel' design have been highlighted and some of the alternatives proposed in the literature have been reviewed.

In the remainder of this thesis, issues relating specifically to the design of the analogue receiver circuits are considered in more detail. Specific case studies from the system described in this chapter are used throughout to illustrate the discussion: chapter 5 considers the design of the data receiver circuit; chapter 7 introduces a new approach to smart-pixel post-amplifier design that was used in the clock receiver circuit; chapter 8 discusses the problem of electrical crosstalk that was highlighted by the design of the switching chip.

The next chapter begins the discussion by considering the general trade-offs in the design of receiver circuits in this application area.

3.7 Appendix: scheme for handshaking with input queues

This appendix describes a possible scheme for communicating back the results of the arbitration from the switching fabric to the 62 input ports.

A total of 62 bits of information must be determined by the switching fabric in a given slot, corresponding to whether the packet presented by each of the 62 inputs has been routed to an output. After the arbitration process, this information is distributed throughout the output ports of the switching chip. This information can be collected over a period of 62 clock cycles as follows. The *k*th cycle is used to determine whether a packet has been routed from input *k*. Each of the 64 output ports locally determines (from the multiplexer select signals) the source address of the packet, if any, that has been routed. During the *k*th cycle, the output node asserts its input to a chip-wide 64-input OR gate if this address is equal to *k*. The output of this gate determines whether any of the output ports have routed the packet from the *k*th input; it is broadcasts electrically to the input ports. Each input port examines this broadcast signal during its own cycle to determine whether the packet that it sent has been accepted.

The chip-wide 64-input OR gate could be implemented using, for example, an 8-input wiredor bus for each row of super-pixels together with an 8-input OR gate at the edge of the chip.

A limitation of this technique is that it sets a limit on the minimum packet duration of 62 clock cycles. Because of the need for chip-wide signalling, the clock used for this process might be some fraction of the channel rate implying a minimum packet length of perhaps 128-256 bits.

Chapter 4

Design trade-offs in photoreceivers for terabit/s scale optical interfaces

4.1 Introduction

The first part of this thesis has given an overview of the technology that forms the basis of terabit/s scale optical interfaces to VLSI electronics, including an example of the type of system that might utilise this technology.

The remainder of this thesis, which forms the core of the research work presented, focuses on the design of photoreceiver circuits for this application; in particular, it deals with aspects of photoreceiver design that are specifically relevant to arrays which are of a scale suitable for implementing terabit/s optical interfaces to VLSI electronics, and which contain many hundreds of channels.

Such photoreceivers, commonly referred to as "smart-pixel" receivers, are similar in structure to the single-channel photoreceivers used in long-haul telecommunications systems, but have very different performance requirements. In telecommunications receivers, sensitivity is of primary importance and can be optimised at the expense of other factors such as power consumption. In contrast, smart-pixel receivers must meet tight constraints on power consumption and layout area to permit a high level of integration (of the order of 1000 receivers in an area of 1 cm²) creating a qualitatively different design problem.

This chapter begins the study by discussing in detail the design trade-offs in smart-pixel receiver circuits. The structure of a smart-pixel receiver and how it compares to a conventional telecommunications receiver is first reviewed. The main issues affecting receiver performance – front-end small signal characteristics , post-amplifier gain, front-end noise and inter-stage offsets are then discussed in turn and their relative importance assessed. The ways in which the design variables can be traded off against the basic characteristics of speed, sensitivity and power consumption are considered.

The analysis in this chapter builds on the work of several authors in the specific field of smart-pixel receiver design [166][167][168][174][187] and in the more general area of telecommunications receivers. Previous work on smart-pixel receivers has highlighted the key differences in performance requirements compared to telecommunications receivers and considered the trade-offs between factors such as power consumption, bit-rate and

sensitivity. To an extent, the analysis in this chapter is merely an alternative approach to the same problem with a slightly more detailed investigation of the influence of some of the circuit design variables such as transistor dimensions. Similarly, the noise analysis is essentially a review of existing theory from the telecommunications field [192].

However, the main contribution of this chapter is to extend earlier studies of the smart-pixel receiver design problem to include a quantitative assessment of the effect of MOSFET mismatch on the sensitivity of DC coupled receivers. The discussion of the other aspects of receiver performance is included primarily to allow an assessment of the relative importance of mismatch, and to form the basis of the study of the scaling of receiver performance in advanced CMOS technology in Chapter 6.

The analysis of receiver performance in this chapter is specific to large arrays only insofar as it considers designs with constrained power consumption and layout area; it still considers the performance of a single receiver circuit in isolation. Later chapters will look at issues arising from the incorporation of these designs into large arrays, through a case study of the receiver design for the SPOEC system in Chapter 5 and an analysis of power supply crosstalk in Chapter 8.

4.2 Review of receiver design

In this section, the architecture of a photoreceiver is reviewed with specific emphasis on the features that distinguish a smart-pixel photoreceiver from a conventional telecommunications photoreceiver.



4.2.1 Structure of a conventional telecommunications receiver

Figure 4-1: Structure of a conventional telecommunications receiver

Figure 4-1 shows the basic structure of a receiver of the type found in a long-haul or medium-distance telecommunications link. The overall purpose of the circuit is to convert an

optical signal, which may have been attenuated and corrupted by noise and dispersion in the optical link, into a clean, high level electronic signal that can be used as an input to logic circuitry or to regenerate the optical signal.

The front-end converts the input photocurrent to a low-level voltage. The main design objective in the front-end is to minimise the electronic noise added to the optical signal. The output of the front-end is then amplified using a post-amplifier; a decision stage produces a regenerated logic signal which is often re-timed with a clock extracted from the data stream. A decision threshold equal to the mean optical input power is extracted from the input data stream using a low-pass filter.

A transimpedance front-end is normally preferred over a high-impedance front-end [190][191] because it offers a combination of high dynamic range and low noise; high-impedance front-ends can, in theory, give superior noise performance but are more difficult to realise in practice. The design of integrated single-channel transimpedance front-ends is a mature field; an excellent review is given by Williams [169][198].

The post-amplifier is either a linear amplifier with an automatic gain control [170][171] or a non-linear limiting amplifier [172].

4.2.2 Smart-pixel implementation

Smart pixel receivers have a structure which is a simplified form of Figure 4-1. The differences are required to reduce circuit complexity, in particular power consumption and layout area. The front-end and post-amplifier remain but typically use somewhat smaller transistors, simpler biasing schemes and fewer gain stages. The clock recovery feature is omitted; this is made possible either using asynchronous routing of the data without retiming or by using a global system clock to retime the data (there has been some work looking at very low power clock recovery techniques [173]). The circuitry to extract the optimum decision threshold is also omitted; consequently, the system response extends down to DC and there is no requirement to code or scramble the data (this issue is discussed in more detail in Section 4.6).

Another major difference is that it is convenient to use a two-beam (differential) optical data link (Figure 4-2). Historically, this was motivated by a requirement to receive signals from low-contrast multiple-quantum well modulator devices of the type described in Chapter 2; however, differential encoding of the optical signal has many advantages even for highcontrast optical data. In particular, the optical signal carries its own reference level which reduces the sensitivity penalty incurred by omitting the circuitry to determine the optimum

60

decision threshold. The increase in complexity in the optical system arising from a doubling of the number of optical signals in a free space optical system is not large, although it would be more significant in a fibre based interconnect. Nevertheless, single ended implementations are still possible. Comparison of designs at similar speed indicates a penalty of between 5 dB and 8 dB in total optical power for choosing to use a single-ended design [174].



Figure 4-2: Two-beam differential encoding of a 2:1 contrast optical signal

Table 4-1 compares the performance of a typical two-beam smart-pixel optical receiver, illustrated in Figure 4-3 [174], with an example commercial SONET receiver [196]. Major differences in the smart-pixel receiver are the much lower power consumption and area of the smart-pixel receiver and the significantly lower sensitivity. It is only possible to integrate thousands of receivers into a single chip if the power consumption of each receiver is of the order of a few milliwatts.



Figure 4-3: Lucent smart-pixel receiver (transistor W/L shown in µm) (after [197])

| parameter | SDH / SONET | smart-pixel |
|----------------------------------|-------------------|-------------------------|
| data rate | 622 Mbit/s | 622 Mbit/s |
| technology | bipolar | CMOS |
| front-end power consumption | 85 mW | 8 mW |
| limiting amplifier power | 59 mW | |
| consumption | | |
| clock recovery power consumption | 155 mW | - |
| sensitivity @ 850 nm | -31 dBm | -18 dBm |
| dynamic range | 25 dB | >16 dB |
| front-end feedback resistance | 6 kΩ | 15 kΩ |
| input referred rms current noise | 55 nA | not relevant |
| area | two die + several | $45 	imes 25 \ \mu m^2$ |
| | external filter | |
| | components | |
| number of optical beams | 1 | 2 |
| switching energy | 1.3 fJ | 28 fJ per beam |



The transimpedance gain stage can be implemented using a number of circuit topologies. Simple gain stages based on CMOS, NMOS and PMOS inverters are shown in Figure 4-4. Many variants are possible: multiple stages [170][230]; the use of a diode connected load transistor to reduce and stabilise the circuit gain [168][230]; addition of a source follower to reduce the influence of capacitance on the output node [198]; common-gate amplifiers and other current-mode structures [175][176][177]; a bipolar gain stage with shunt-series feedback (e.g. [178]). Many of the more complex variants are not suitable for smart-pixel implementation. The work in this chapter concentrates on analysing the simple structure with a single-pole transfer function.



Figure 4-4: Example gain stages suitable for use in smart-pixel receivers

Woodward [174] has recently reviewed approaches to smart-pixel receiver design.

The analysis of design trade-offs in this chapter is based on the circuit structure of Figure 4-3 with a complementary gain stage. The complementary inverter gain stage has been found to
be particularly simple to use in a smart-pixel environment because it is fully self-biasing (see Chapter 5). However, the analysis can easily be adapted for the other gain stages shown in Figure 4-4. This design implements the feedback resistor of the transimpedance front-end with an ohmic region transistor.

4.2.3 Synchronous sense-amplifier receivers

It is worth mentioning in passing that there exists a class of smart-pixel receiver that is completely distinct from the class using a transimpedance front-end. These receivers are called synchronous sense-amplifiers; they are based on the same circuit techniques used in memory sense amplifiers.

The key features of this class of receiver are low power consumption and a requirement for an external clock signal which is synchronous to the data signal.

The circuits are all based on some form of regenerative, bistable latch. The clock is required to reset the latch into a metastable state at the start of each cycle. The stable state into which the latch resolves is determined by a differential input signal. The positive feedback of the latch allows the input signal to be restored to a full digital logic level in a single stage.

The requirement for a clock is the main disadvantage of this class of receiver. A clock cannot be extracted from the data stream if it is not amplified with a linear amplifier first. However, the overhead of supplying a clock channel in parallel with a wide data bus is low. The high bandwidth optical interfaces for which this class of receiver is intended will contain many hundreds of channels but the number of independent sources of data is likely to be much fewer in most cases.

Sense amplifiers with optical inputs were first used in [179] and independently developed by a number of authors [180][181][182][183][184]. Two distinct approaches have been taken: a charge-sense amplifier and a current-sense amplifier.

The charge-sense amplifier operates by integrating the input photocurrent on a capacitor during one clock phase and sensing the integrated voltage using a regenerative latch during a second clock phase. This can potentially offer very high sensitivity at low clock speeds because of the integration of the input signal which has fundamental noise advantages in the same way as a high-impedance receiver [185]; however, it requires complicated timing signals which make it difficult to apply at high speeds.

The current-sense amplifier also uses a regenerative latch but with a differential current input; it does not integrate the input photocurrent. Experimental results for a two-beam

current-sense optical receiver have demonstrated, for example, operation at 320 Mbit/s with a peak photocurrent per beam of 5 μ A corresponding to a switching energy of 100 fJ, and a power consumption of 1-2 mW [186]; this may be compared with the performance of the transimpedance design in Table 4-1.

Hybrid approaches using a transimpedance front-end followed by a clocked sense amplifier with a global clock have also been proposed [187][188].

This class of photoreceiver is explicitly outside the scope of this chapter.

4.3 Front-end small signal performance

This section begins the discussion of the receiver design trade-offs by considering a smallsignal model of the front-end, and examining the influence of the feedback resistor, the frontend gain stage and the photodiode capacitance on the overall circuit performance.

4.3.1 Small-signal analysis

Figure 4-5 shows a small signal equivalent circuit of the transimpedance front-end.



Figure 4-5: Small signal model of a transimpedance front-end

This model can be applied to gain stages containing NMOS and/or PMOS gain transistors; g_m is the total transconductance of the gain transistor(s) and g_{ds} is the output conductance of the gain transistor(s) and bias transistor if any. The open loop gain of the stage is $A = g_m/g_{ds}$.

 C_{IN} is the capacitance of the input node; it includes the photodiode capacitance and the gatesource capacitance of the input transistor. In smart-pixel circuits, the photodiode capacitance is the main contributor to this capacitance.

 C_{F} is the feedback capacitance between the input and output nodes. It includes the gate-drain capacitance of the input transistor and any parasitics.

 C_L is the capacitance at the output node of the front-end. It includes the input capacitance of the post-amplifier, the drain junction capacitance of the front-end, and the source-gate capacitance and junction capacitance of the feedback transistor. The Miller approximation is used to include the effect of the post-amplifier feedback capacitance in C_L^{-1} .

Table 4-2 shows typical values of these parameters, and how they relate to transistor dimensions, for a complementary gain stage in a specific 0.6 μ m process with a 5 V supply. Identical widths are assumed for the NMOS and PMOS transistors. W_{FRONT-END} and W_{POST-AMP} denote the widths of the NMOS transistors in the front-end and post-amplifier gain stages. W_{FB} and L_{FB} denote the width and length of the feedback transistor.

As the transistors used are typically quite small, routing capacitance is also significant. The table includes parasitic values extracted from an actual layout with $W_{\text{FRONT-END}} = 10 \ \mu\text{m}$ and $W_{\text{POST-AMP}} = 5 \ \mu\text{m}$.

| | terms proportional to transistor dimensions | | | | |
|-----------------|---|------------------------|-----------------------|-----------------------------------|-------------------------------|
| parameter | component | front-end | post- amplifier | feedback transistor | constant terms and parasitics |
| | multiplier | W _{FRONT-END} | W _{POST-AMP} | $W_{_{FB}} 	imes L_{_{FB}}$ | |
| C _{IN} | | 2.0 fF / µm | | $1.4 \; fF / \mu m^2$ | 100 fF photodiode |
| | | | | | 20 fF parasitics |
| $C_{_{\rm F}}$ | | 0.7 fF / µm | | -0.5 fF / μm^2 | 2 fF |
| C _L | | 0.4 fF / µm | 2.0 fF / µm | $1.4~\mathrm{fF}/\mu\mathrm{m}^2$ | 9 fF |
| g_{m} | | 190 µS / µm | | | |
| А | | | | | 16.0 |

Table 4-2: Typical parameter values for a smart-pixel transimpedance receiver

Assuming that A >> 1, C_{IN} >> C_F and $g_m R_F$ >>1, simple nodal analysis of the circuit gives a transimpedance gain $Z_T(s)$ of:

$$Z_{T}(s) = \frac{v_{out}}{i_{in}} = -\frac{R_{F}\left(1 - s\frac{C_{F}}{g_{m}}\right)}{1 + s\tau + \frac{1}{\omega_{0}^{2}}s^{2}}$$
(4.1)

¹ Note, however, that the load capacitance is primarily important in determining the response of the front-end close to the cut-off frequency through its influence on ω_0 ; because the postamplifier voltage gain can have significant phase shift at these frequencies, the reactive component of the Miller load can be somewhat less than (A+1) C_{gp.}

where $s = j\omega$ is the complex frequency variable and where

$$\tau = R_F (C_F + \frac{C_{IN}}{A+1}) + \frac{C_{IN}}{g_m} + \frac{C_L}{g_m}$$

$$\frac{1}{\omega_0^2} = \frac{R_F C_{IN} (C_L + C_F)}{g_m}$$
(4.2)

The zero at $\omega = g_m / C_F$ is usually at a high frequency and does not have a big effect on performance. The circuit therefore has a second order response with damping factor $\zeta = \omega_0 \tau / 2$. The requirement for a response in the time domain with acceptable overshoot requires that ζ is above a certain minimum. For $\zeta > 1 / \sqrt{2}$, the step response of a second order system settles to within 5% of its final value after a settling time of approximately²

$$T_s = 3\tau \tag{4.3}$$

 T_s is used to estimate the minimum bit-period of a particular front-end. It is assumed for the moment that the post-amplifier has sufficient bandwidth that it does not degrade the rise-time. The bit-rate B is then given by:

$$B = \frac{1}{3\tau} \tag{4.4}$$

Note that the model of Figure 4-5 starts to break down at frequencies above about ω_{T} / 8 due to non-quasi-static behaviour [189] where ω_{T} is the transit frequency of the gain transistor. For the example parameters, the limit of validity is about 2 GHz. The feedback transistor typically uses a channel length of several times minimum to realise a large resistance and so can have a much lower transit frequency; the equivalent circuit used for the feedback transistor is therefore a first order correction to quasi-static behaviour as discussed in Appendix 4.8 but can still be accommodated within the circuit model of Figure 4-5. It predicts the two-port y-parameters of the feedback transistor to better than 5% up to 1.5 ω_{T}

4.3.2 Effect of the feedback resistor

For a fixed gain stage, the receiver bit-rate can be traded off against sensitivity by varying the feedback resistor. Large values of R_F will give high sensitivity but lower speed and vice-versa. It is useful to define the **switching energy** of a receiver as a measure of the

² This expression is accurate to about 20% for the stated range.

transimpedance-bandwidth product. Following the definition in [167], the switching energy of a two-beam receiver is defined to be the peak optical energy in each beam per bit³:

$$E_{OPT} = \frac{P_{PEAK}}{B} \tag{4.5}$$

Let V_{MIN} be the minimum peak-to-peak voltage at the post-amplifier input that will produce a valid logic signal at the receiver output. Assume that V_{MIN} is determined by the properties of the second-stage amplifier only. This assumption is valid for designs that are limited by the gain of the post-amplifier and is a reasonable approximation for designs that are limited by DC offset (see Section 4.6). It is not valid for designs limited by the front-end noise, but Section 4.4 will show that smart-pixel receivers are not noise-limited. In the first instance, V_{MIN} is assumed to be independent of bit-rate, although this assumption will be modified when the behaviour of the post-amplifier is examined.

The peak photocurrent per photodiode for a high contrast optical signal required to produce an output swing of V_{MIN} is $I_{MIN} = R_F V_{MIN} / 2$ where the factor of two is due to the fact that there are two input beams. If the responsivity of the photodiode is S A/W, then the switching energy is:

$$E_{OPT} = \frac{1}{2S} \frac{V_{MIN}}{R_F} T_{BIT}$$

$$\tag{4.6}$$

By eliminating $R_{_{\rm F}}$ from the expression for τ , an alternative form can be derived:

$$E_{OPT} = \frac{3}{2} \left(C_F + \frac{C_{IN}}{A+1} \right) \frac{1}{1 - B / B_0} \frac{V_{MIN}}{S}$$
(4.7)

where

$$B_0 = \frac{g_m}{3(C_{IN} + C_L)}$$
(4.8)

This equation predicts that the switching energy is constant for large values of R_F (low bitrates) but increases asymptotically as the bit-rate B approaches B_0 . Because the switching energy does not vary very much over a range of bit-rates, the low-frequency limit provides a

³ Definitions varying by a factor of two in either direction have also been used in the literature, depending on whether the peak or average optical energy is used and whether the energy per beam or total energy in two beams is calculated.

useful guide to the amount of optical power required to implement an optical interface with a given total bandwidth.





Figure 4-6: Variation in switching energy with bit-rate for a fixed amplifier

4.3.3 Effect of the front-end amplifier

The design variables of the front-end gain stage can be adjusted to influence the trade-off between speed and sensitivity through the value of B_0 . Increasing B_0 shifts the position of the asymptote in the graph of switching energy to a higher bit-rate. In order to achieve a switching energy close to the low-frequency limit, B_0 must be chosen to be greater than about 5B. Since $C_L << C_{IN}$, B_0 is primarily determined by the transconductance of the input stage and the photodiode capacitance.

The transconductance is directly proportional to the width of the front-end transistors. The supply current is also proportional to the width; there is thus a reasonably direct trade-off between power consumption and maximum operating speed for low switching energy.

To a lesser extent, the transconductance can also be adjusted by choosing the transistor drive voltage V_{gs} - V_T . In long-channel transistors, the transconductance is directly proportional to V_{gs} - V_T . However, in short-channel devices, saturation of the carrier drift velocity at high

electric fields causes the transconductance to reach a maximum value at relatively low drive voltages; the influence of this variable is therefore less strong. This is particularly true in the complementary inverter configuration where the NMOS and PMOS transistors are biased with relatively large drive voltages.

Two design variables affect the drive voltage in the complementary inverter: the choice of power supply voltage and the ratio of the width of the PMOS transistor to the width of the NMOS transistor.

The effect of the power supply voltage is illustrated in Table 4-3 for the 0.6 μ m process: increasing the power supply voltage from 3.3 V to 5.0 V gives only a 20% increase in g_m for a substantial penalty in supply current and power consumption, suggesting that the lower power supply voltage would be preferred. However, in the design of a large chip, issues other than the figures of merit in Table 4-3 influence the choice of power supply voltage. These practical issues will be discussed in Chapter 5.

| V _{DD} | g _m | I | P/mW | $g_{\rm m}^{}/I_{\rm supply}^{}$ | g _m /P |
|-----------------|----------------|---------|--------|----------------------------------|----------------------|
| 3.3 V | 1.51 mS | 0.35 mA | 1.2 mW | 4.3 V ⁻¹ | 1.3 V ⁻² |
| 5.0 V | 1.86 mS | 1.02 mA | 5.0 mW | 1.8 V ⁻¹ | 0.4 V^{-2} |

Table 4-3: Effect of supply voltage on g_m for Wn=Wp=10µm, L=0.6µm

The PMOS / NMOS width ratio does not have a strong influence on the transconductance. Figure 4-7 shows the transconductance and total transistor width as a function of the inverter ratio at fixed supply current using a 5 V supply. An inverter ratio much less than one (small PMOS, large NMOS, low operating point) gives a small improvement in the transconductance per unit current but has other disadvantages that make it undesirable in practice. For example, assuming that subsequent stages in the receiver are designed to have the same operating point as the front-end, a low ratio gives rise to highly asymmetric large-signal rise and fall times. It can also be seen from Figure 4-7 that, for reasonable ratios in the range 1 to 3, the total transistor width required to obtain a given transconductance is only weakly dependent on the exact ratio. Consequently, the term proportional to the transistor feedback capacitance C_F in the expression for the switching energy (4.7) is to a first order independent of the ratio.

In summary, for the complementary inverter topology, the drive voltage does not allow for much control over the transconductance.



Figure 4-7: Effect of inverter ratio on transconductance and total width at fixed current

4.3.4 Damping factor

The expression for switching energy (4.7) is valid provided the circuit has an acceptable damping factor ($\zeta > 1 / \sqrt{2}$). This may limit the range of bit-rates at which a given gain circuit may be made to operate by adjusting R_F . As the bit-rate is varied by adjusting the feedback resistor, the damping factor varies as:

$$\zeta = \frac{\zeta_{MIN}}{2} \frac{1}{\sqrt{(B/B_0)(1 - B/B_0)}}$$
(4.9)

where

$$\zeta_{MIN}^{2} = (1 + \frac{C_{L}}{C_{IN}}) \frac{\frac{C_{IN}}{A+1} + C_{F}}{C_{L} + C_{F}} \approx \frac{\frac{C_{IN}}{A+1} + C_{F}}{C_{L} + C_{F}}$$
(4.10)

is the minimum value of ζ . The form of this function is plotted in Figure 4-8.



Figure 4-8: Variation in damping factor with B/B₀

As similar values of B / B_0 will be used, independent of bit-rate, to ensure a good trade-off between power consumption and switching energy, ζ_{MIN} , as determined by equation (4.10), indicates the influence of the various parameters on the damping factor.

Notice, in particular, the importance of minimising the load capacitance. If the post-amplifier input transistor scales in proportion to the front-end, the load capacitance will scale in proportion to the bit-rate; the damping factor constraint thus limits the maximum speed of operation. This limit can be overcome by reducing the front-end gain (e.g. by using diode connected load transistors inside the feedback loop) but at the expense of switching energy.

Damping also limits the extent to which the switching energy may be improved by using longer-channel, higher gain transistors in the front-end.

4.3.5 Calculations for example parameters

In this section, the expressions developed in the previous sections and the example parameters in Table 4-2 are used to investigate the performance trade-offs numerically.

Figure 4-9 shows the effect on the switching energy of adjusting the front-end width. In this example, $W_{POST-AMP}$ has been taken equal to $W_{FRONT-END}$ and a post-amplifier gain of 5 has been assumed in the estimation of the Miller load capacitance. A responsivity of 0.5 A/W and a V_{MIN} of 200 mV were assumed.



Figure 4-9: Effect of front-end width on switching energy at different bit-rates

For simplicity, a fixed size of feedback transistor of $1.2 \ \mu m \times 2.0 \ \mu m$ is used, independent of the required resistor value; it is assumed that the resistance can be adjusted by varying the gate voltage. For an NMOS transistor with the example dimensions, the resistance can be adjusted between 7 k Ω and 70 k Ω for drive voltages between 200 mV and 2V. The resistor value required to achieve the necessary rise time varied between 12 k Ω and 120 k Ω depending on bit-rate and so, in a practical implementation, the transistor would have to be slightly longer at lower-speeds. However, the extra capacitance would be less important at these lower speeds. The transit frequency of the feedback transistor varied between 200 MHz and 2 GHz depending on bit-rate and in all cases was sufficient to ensure the accuracy of the resistor model within the bandwidth of the circuit.

It can be seen that for each data rate, there is a broad minimum in switching energy for frontend widths greater than a particular value. This is a sensible region in which to design the receiver: in this region, the switching energy should be insensitive to small variations in process parameters. Increasing the width beyond this optimum slightly degrades the switching energy due to the increase in the front-end feedback capacitance.

This optimal width increases with data rate. For example, at 200 Mbit/s, the model predicts a minimum switching energy of 8 fJ at 2.5 μ m, whereas at 1 Gbit/s, the minimum is 15 fJ at 8 μ m. The switching energy is thus relatively insensitive to bit rate, but the power consumption of a good design varies quite strongly. For reference, the power consumption of the front-end only for W_{FRONT-END} = 5 μ m under typical process conditions is 2.6 mW.

The damping factor calculated for these parameters was greater than 0.7 at bit-rates up to 400 Mbit/s but fell to around 0.6 at 1 Gbit/s. It was possible to improve the damping factor to 0.7 by sizing the front-end to be twice as wide as the post-amplifier. However, $\zeta = 1 / \sqrt{2}$ is not a hard limit, and it is possible to operate a receiver with lower values of ζ and hence longer settling times than equation (4.3) if increased pattern dependent jitter can be tolerated.

4.3.6 Effect of photodiode capacitance

An important implication of the results in Section 4.3.3 is the very strong influence that the photodiode capacitance has on the performance of the receiver circuit. In order to achieve a switching energy close to the low frequency limit, the input stage must be sized in proportion to the photodiode capacitance. Choosing B_0 equal to about 5 B can provide a good starting point for a design for a given photodiode capacitance. Thus, the power consumption of the front-end is approximately proportional to the photodiode capacitance. The low frequency limit of the switching energy is also proportional to the photodiode capacitance.

Receivers with performance characteristics that are compatible with requirements of large arrays in terms of power consumption and switching energy have used 100 fF photodiodes with diameters of around 30 μ m. The development of cost-effective optomechanical packaging that can focus light onto a detector of around this size is therefore important for large receiver arrays to be practical.

4.4 Noise limits on receiver sensitivity

The sensitivity of stand-alone receiver circuits is usually limited by noise. In this section, the fundamental noise limits of a transimpedance front-end are reviewed and a comparison made between the optimisation of a front-end for low noise and the optimisation of a design for the smart-pixel environment.

Better noise performance is theoretically possible using an integrating front-end and equalisation in the post-amplifier [190][191]. However, this technique is rarely used in high speed designs and is not practical for data-link applications; it is not considered here.

The integrated output noise $\langle i_{in}^2 \rangle$ of a transimpedance front-end referred back to the input is [192]

$$\left\langle i_{in}^{2} \right\rangle = \left(2q(I_{GATE} + I_{DARK}) + \frac{4kT}{R_{F}} \right) \int_{0}^{\infty} \frac{\left| Z_{T}(f) \right|^{2}}{\left| Z_{T}(0) \right|^{2}} df + 4kT \Gamma \frac{(2\pi C_{IN})^{2}}{g_{m}} \int_{0}^{\infty} f^{2} \frac{\left| Z_{T}(f) \right|^{2}}{\left| Z_{T}(0) \right|^{2}} df \quad (4.11)$$

 Γ is a numerical factor describing the channel thermal noise which is 2/3 for long channel devices but significantly higher for short-channel devices under certain bias conditions [193][194][195]. I_{GATE} is the transistor gate leakage current and I_{DARK} is the detector dark current; both these noise sources are usually negligible for MOSFET transistors⁴. Low-frequency flicker noise has also been neglected.

These integrals are evaluated in [192] for the transfer impedance $Z_{T}(f)$ from equation (4.1) to give:

$$\left\langle i_{in}^{2} \right\rangle = \frac{kT}{\tau} \left(\frac{1}{R_{F}} + \Gamma \frac{C_{IN}^{2} \omega_{0}^{2}}{g_{m}} \right) = \frac{kT}{\tau R_{F}} \left(1 + \Gamma \frac{C_{IN}}{C_{L} + C_{F}} \right)$$
(4.12)

The first and second terms represent the thermal noise in the feedback resistor and the transistor channel respectively. The noise power can be related to a sensitivity for a particular bit error rate by treating it as a Gaussian noise source with the same power. Defining I_{MIN} as the peak photocurrent per photodiode for a high contrast optical signal as before,

$$I_{MIN} = Q_{\sqrt{\left\langle i_{in}^{2} \right\rangle}}$$
(4.13)

where Q is the solution of

$$BER = \frac{1}{2} erfc(\frac{Q}{\sqrt{2}}) \tag{4.14}$$

-1/2

and is 6.00 for a bit error rate of 10^{-9} .

Eliminating R_F from equations (4.2), (4.5) and (4.12) as before gives an expression for the switching energy:

$$E_{OPT} = \frac{3Q}{S} \frac{1}{\sqrt{1 - B/B_0}} \left[kT \left(C_F + \frac{C_{IN}}{A+1} \right) \left(1 + \Gamma \frac{C_{IN}}{C_L + C_F} \right) \right]^{1/2}$$
(4.15)

This expression is quite similar to (4.7). Both expressions contain the factor $1 / (1 - B/B_0)$ raised to a power: 1 in the minimum output signal limited case and 0.5 in the noise limited

⁴ The shot noise in the photocurrent is also neglected. The magnitude of this is often much less than the amplifier noise at the sensitivity limit of the amplifier for high contrast optical data. For example, the amplifier noise quoted below of 55nA is equivalent to the shot noise in a dc photocurrent of 18μ A in a noise equivalent bandwidth of 500 MHz which is much greater than the minimum detectable photocurrent. However, it might become a relevant source of noise in very low contrast optical data.

case. The optimum gain-transistor width from a noise perspective is therefore lower than in the minimum output signal case because the factor falls off more rapidly as B_0 is increased above B. Nonetheless, the general principle of sizing the gain transistor to make B_0 significantly greater than B is still valid. Sizing the transistor wider than this optimum will reduce the noise spectral density but move the second order pole, which determines the effective upper limit of the second integral, to a higher frequency, offseting the reduction in spectral density. If there are higher order poles in the transfer function, or if the postamplifier can be used to band-limit the noise (as is quite commonly the case), then the preceding analysis ceases to be valid and the noise at the input to the decision stage can be reduced by further increasing the width of the front-end gain transistor. However, the expression still gives an upper-bound on the noise.

Figure 4-10 plots the noise limited switching energy for the example parameters, assuming Q=6 and, optimistically, $\Gamma = 2/3$. Notice from the graph that a smart-pixel receiver in which the front-end has been sized to lie on the flat portion of the switching energy curve will also be near optimum from a noise point of view for a given amplifier configuration.



Figure 4-10: Upper bound on noise-limited switching energy as a function of front-end width

This result is perhaps surprising – one might expect that the restrictions on size and power consumption of the smart-pixel environment would result in degraded noise performance. The result can be rationalised in light of the small photodiode capacitance resulting from

hybrid integration. The comparison would be different if the photodiode capacitance was significantly larger than the 100 fF assumed here since the front-end has to be scaled in proportion to the photodiode capacitance to achieve good switching energy. In any case, all the result implies is that the noise performance of the front-end is quite good - it does not say anything about the performance which is achievable in the post-amplifier and decision stage of the receiver.

For comparison, a commercial bipolar 622 Mbit/s transimpedance front-end specified for a 300 fF photodiode capacitance [196] has a worst-case input referred rms noise current of 55 nA compared with a prediction of a typical value of 33 nA for a 5 μ m front-end at the same speed with a 100 fF photodiode capacitor. This indicates that the noise performance of the smart-pixel design is respectable.

The importance of the thermal noise in a smart-pixel circuit can be assessed by writing an expression for the voltage signal swing at the output of the first stage and comparing it with typical values of V_{MIN} for the second stage amplifier.

$$V_{OUT} = 2Q\sqrt{1 - B/B_0} \left[kT \left(1 + \Gamma \frac{C_{IN}}{C_L + C_F} \right) / \left(C_F + \frac{C_{IN}}{A + 1} \right) \right]^{1/2}$$
(4.16)

For a design with B_0 significantly larger than B, this function is only weakly dependent on the bit-rate and the exact sizing of the front-end; C_{IN} and C_L are the main influences on its value. For designs satisfying this condition using the example parameters, the value was between 9mV and 14mV. The next two sections will show that this is less than the minimum signal requirement imposed by post-amplifier gain and DC offsets.

4.5 Post-amplifier gain limits on sensitivity

4.5.1 Introduction

The analysis so far has expressed the sensitivity of the receiver in terms of the minimum input signal V_{MIN} required at the output of the front-end to produce a valid logic level at the output of the decision stage. The trade-offs in the design of the front-end at constant V_{MIN} have been discussed.

However, the value of V_{MIN} is determined by the design of the post-amplifier and decision stage. In this section, the parameter V_{MIN} is related to the design variables of the post-amplifier circuit and the results used to discuss the trade-off between power consumption, switching energy and bit-rate in the photoreceiver as a whole.

The analysis in this section considers the optimisation of the design in the absence of random DC offsets. Section 4.6 considers how the conclusions of this section are modified by the presence of DC offsets and discusses the circumstances in which the results of this section remain valid.

The analysis is based on the two-beam DC coupled post-amplifier circuit described in Section 4.2.2.



Figure 4-11: Single stage DC coupled post-amplifier circuit together with front-end

The post-amplifier (Mn2/Mp2/Mn3/Mp3) is designed to amplify the small-signal voltage at the output of the front-end (Mn1/Mp1) to the level required to produce non-linear thresholding behaviour in the decision stage. Typically, the post-amplifier has a relatively low voltage gain. Transistors Mn2/Mp2 are matched to the front-end transistors Mn1/Mp1 so that the operating point of the two stages is the same; in a two-beam receiver, a symmetrical photocurrent swing at the input produces a symmetrical voltage swing about this operating point at A and B, the amplitude at B being larger than the amplitude at A by the gain of the second stage.

By definition, the post-amplifier is a small-signal, linear amplifier: if the signal amplitude at the output of the front-end was sufficient to produce non-linear behaviour, then no post-amplifier would be required and the decision stage could be connected directly to the front-end. Consequently, the small-signal bandwidth must be adequate to pass the data signal in order to avoid pattern dependent effects. The purpose of diode-connected transistors Mn3/Mp3 is to reduce the gain and hence improve the bandwidth of the linear amplifier from that of an unloaded inverter. They also help to stabilise the gain against process variation and interstage offsets and to improve the linearity of the stage [197] [198].

The decision stage (Mn4/Mp4) is essentially a very simple limiting amplifier. Unlike the post-amplifier, it is non-linear and it is not necessary for the small-signal bandwidth to be

enough to pass the signal: the distortion introduced by the non-linear gain reduces the rise and fall time from the small-signal value. The output of the stage is clamped by the power supply rails and is a valid digital logic signal. This output would typically be used to drive a small digital inverter.

For simplicity, the minimum input signal to the decision stage, $V_{DECISION}$, is assumed to be set by the requirement to create a fully restored DC output level after a single inverter stage. This approach is consistent with other studies of smart-pixel receivers [168]. The input signal must also be large enough to ensure adequate dynamic response. Simulations (see Appendix 4.9) suggest that the input signal required to satisfy this constraint is smaller than the signal required to satisfy the DC constraint. $V_{DECISION}$ as defined here is therefore independent of bit rate and, consequently, V_{MIN} is controlled by the gain of the second stage $A_{POST-AMP}$. For the 0.6 µm process considered so far with a 5 V supply, $V_{DECISION}$ is about 800 mV.

Because adequate dynamic response can be obtained with a smaller input signal, it is possible to use a two-stage decision circuit consisting of a non-linear limiting stage that produces a large but not fully restored swing, followed by a second stage which generates a completely valid digital logic level. This approach has been found to provide better overall sensitivity. In some respects it is similar to adding an additional stage to the post-amplifier, but has a smaller cost in power consumption. However, for simplicity, this possibility is not considered in the analysis which follows. Allowing this design option would modify the results in two ways: V_{DECISION} would be reduced and would become somewhat dependent on bit-rate. For example, the input signal required to produce an output edge time equal to the input edge time reduces from 420 mV for 1 ns edges to 140 mV for 4 ns edges.

The effective load capacitance of the decision stage is defined to be $\mathbf{C}_{\text{\tiny DECISION}}$

4.5.2 Small signal analysis

Figure 4-12 shows a small signal model of the post-amplifier circuit.



Figure 4-12: Small signal model of post-amplifier circuit

The small signal voltage gain is:

$$A_{POST-AMP} = \frac{g_{m2}}{g_{m3} + g_{ds2} + g_{ds3} + s C_{LOAD}} \approx \frac{g_{m2} / g_{m3}}{1 + s \frac{C_{LOAD}}{g_{m3}}}$$
(4.17)

where

$$C_{LOAD} = C_{DS2} + C_{GS3} + C_{DS3} + C_{DECISION} + C_{PARASITIC}$$
(4.18)

is the total load capacitance.

4.5.3 Design variables

The design variables of the amplifier are the width of transconductor Mn2/Mp2 and the width of the diode connected load Mn3/Mp3, defined as $W_{POST-AMP}$ and W_{LOAD} respectively. It is assumed for simplicity that the length of the diode-connected load is identical to the length of Mn2/Mp2 in order to avoid a systematic offset between the operating point of the frontend and the operating point of the post-amplifier under process variation in the channel length.

4.5.4 Design trade-offs

The speed of the post-amplifier will be discussed in terms of a required small signal bandwidth of the stage ω_{POST_AMP} . This will be related to the overall bit-rate of the receiver when the interaction with the front-end is discussed.

For a fixed capacitive load, the width of the load transistor is determined by ω_{POST_AMP} which is approximately

$$\omega_{POST-AMP} = \frac{g_{m3}}{C_{LOAD}} \tag{4.19}$$

The gain of the stage can then be adjusted by varying $W_{POST-AMP}$. To a first approximation, the gain of the stage is given by the ratio of $W_{POST-AMP}$ to W_{LOAD} , although the finite output conductance of Mn2/Mp2 limits the maximum gain that can be achieved to that of the unloaded inverter.

It follows that there is a clear trade-off between the post-amplifier power consumption and the gain of the second stage (and hence to sensitivity of the photoreceiver as a whole). The power consumption is proportional to $W_{POST-AMP} + W_{LOAD}$ which is in turn approximately proportional to $(A_{POST-AMP}+1)$.

There is also a fairly direct trade-off between sensitivity and speed. For a fixed $W_{POST-AMP}$, the gain-bandwidth product of the amplifier is approximately.

$$GBW = \frac{g_{m2}}{C_{LOAD}} \tag{4.20}$$

although the load capacitance is weakly dependent on the gain because of the capacitance of the load transistor.

4.5.5 Simulation

These predictions were validated using HSpice simulations. A decision stage with $W_N = W_P = 4 \mu m$ was loaded with two identical stages. The maximum gain achievable by varying the width of the load transistor under constraint of a minimum 3 dB bandwidth was determined using a small-signal analysis. Each node was loaded with a parasitic capacitance of 5 fF. The post amplifier input was biased about its operating point and driven with a voltage source.

Figure 4-13 shows the results which confirm that the power consumption can be traded off against gain up to a point, although there is a diminishing return for gains larger than about 5. The gain-bandwidth trade-off is only evident at higher bandwidths. This is because there is a minimum allowed transistor width of 0.8 μ m which placed a minimum on g_{m3} under a constraint of fixed length. At this width, the bandwidth of the stage is still quite high.



Figure 4-13: Trade-off between gain and power consumption in a single-stage postamplifier

4.5.6 Multistage amplifier designs

A more effective way to trade power consumption for gain beyond this point of diminishing returns is to employ a multistage post-amplifier. A two-stage amplifier is shown in Figure 4-14. Identical first and second stages were used, although this is not necessarily optimal because the load on the two stages is not the same.



Figure 4-14: Two stage post-amplifier design

The results of similar simulations performed on this design are shown in Figure 4-15. In this circuit, $\omega_{POST AMP}$ was defined as the overall bandwidth of the cascade.



two-stage post-amplfier



It can be seen that, at low bandwidths, this circuit can deliver much higher gains compared to a single stage design of the same power consumption. For example, a two stage design with a width of 4 μ m has a small-signal gain of 12.6 for bandwidths up to 400 MHz compared to a gain of 4.8 for a single-stage design of the same power consumption and a width of 8 μ m.

However, at higher overall bandwidths closer to the gain-bandwidth product of a single stage, this technique seems to be less useful because the gain per stage must be quite low to achieve the necessary overall bandwidth. Increasing the number of stages further can help; a

multistage cascade of low-gain amplifiers is a common technique for wide-band amplifiers [199][200].

An additional advantage of this technique is that it reduces the loading on the front-end both directly, by reducing the width of the post-amplifier input transistors, and indirectly, by reducing the gain of the first post-amplifier stage and hence reducing the contribution to the load capacitance of the Miller multiplied gate-drain capacitance.

4.5.7 Performance together with front-end

The performance trade-offs in the design of the post-amplifier have now been considered. To establish the trade-offs in the receiver as a whole, the interactions between the design variables of the two components must be considered.

The main interaction is through the load capacitance presented by the post-amplifier. It has been seen that the load capacitance does not have a strong effect on B_0 or on the switching energy at fixed V_{MIN} . However, it does affect the damping factor ζ ; this imposes an upper limit on the effective load capacitance.

The input capacitance of the post-amplifier is proportional to $W_{POST-AMP}$. However, in practice, the width of the second stage is not a completely free parameter because of layout requirements for low systematic offset . It is desirable to have the ratio $W_{PRONT-END}$: $W_{POST-AMP}$ expressable as a ratio of small integers so that they can be constructed from paralleled unit transistors. Ratios of 1:1 and 2:1 have layouts that are particularly simple (and hence have low parasitic capacitance) and, for the example process parameters, have been found to ensure acceptable damping with 2:1 preferred at higher bit-rates.

Assuming that the ratio remains fixed, the power consumption of the post-amplifier and the front-end cease to be independent and it is possible to treat the front-end width as a direct measure of the power consumption of the photoreceiver. It is then useful to weight the switching energy graph of Figure 4-9 by $V_{\text{DECISION}} / A_{\text{POST-AMP}}$ to provide an indication of the overall trade-off between power consumption and sensitivity as a function of bit-rate in the absence of DC offset.

The overall bandwidth of the front-end / post-amplifier combination is less than that of the component stages. The frequency response of the two stages has been approximated by taking the rms sum of the rise times. Where possible, the second stage has been designed to

have the same bandwidth as the front-end⁵ such that the overall bandwidth in MHz is half the bit-rate in Mbit/s. In low bit-rate cases where a post-amplifier with a minimum width load transistor has more bandwidth than is required to satisfy the equal bandwidth condition, the bandwidth of the front-end has been chosen to give the required overall bandwidth.

A one- or two-stage post-amplifier has been selected to provide the highest gain for a particular post-amplifier power consumption, comparing a one-stage design with width $W_{FRONT-END}$ with a two-stage design with width $W_{FRONT-END} / 2$.



WFRONT-END / µm

Figure 4-16: Trade-off between front-end width and switching energy including the bitrate dependence of the second stage gain

The results of this calculation are plotted in Figure 4-16. As the width is increased, there is initially a very rapid reduction in switching energy due to the front-end requirement that $B \ll B_0$. As the width increases further, the additional gain available from the post-amplifier causes the switching energy to continue to decrease; consequently, there is no optimum width as there is if a V_{MIN} independent of width is assumed.

It is also useful to estimate typical values of $V_{_{MIN}}$. If an arbitrary upper limit on the total postamplifier width of 10 µm is assumed, then gains of between 3 and 18 are possible depending on speed and power consumption requirements. This corresponds to values of $V_{_{MIN}}$ between

⁵ In general, a cascade of two stages each having a fixed gain-bandwidth product and the same basic shape of frequency response will have a maximum overall gain for a given overall bandwidth when both stages are designed to have the same bandwidth.

40 mV and 250 mV. In many cases, these figures are such that that the overall performance of the photoreceiver is gain limited [167]. More precisely, the figure of merit

$$\frac{gain \times bandwidth}{power} \tag{4.21}$$

of this post-amplifier stage is such that, within a given constraint on power consumption, it is not possible to implement enough gain to amplify the front-end signal to the detection threshold of the decision stage. Chapter 7 looks at how this limit can be addressed by using a post-amplifier topology with a fundamentally higher gain-bandwidth product.

The next section considers how the results of this section are modified in the presence of random offsets.

4.6 MOSFET mismatch limits on receiver sensitivity

4.6.1 Introduction

In this section, the impact of random DC offsets arising from transistor mismatch on the sensitivity of smart-pixel receivers is quantified and compared with the limits due to post-amplifier gain and thermal noise that have already been discussed.

A major difference between the majority of smart-pixel receiver circuits that have been implemented to date and long-haul telecommunication and optical data link receivers is that the smart-pixel circuits have a frequency response that extends down to DC.

This characteristic has certain advantages in the context of smart-pixel systems. No capacitors are required to implement a low-frequency cut-off; very simple and compact DC coupled designs with modest power consumption exist. A DC coupled channel can also pass an arbitrary sequence of non-return-to-zero data (which has a frequency spectrum that extends down to DC) which is convenient for any smart-pixel system that has to process (rather than just route) a data sequence and is also spectrally efficient. Other coding schemes such as Manchester or 8B / 10B coding [201] are either spectrally inefficient or require relatively complex coding/decoding circuits which add power consumption and layout area.

In systems with a modest number of channels, these benefits are less important than some of the benefits obtained from removing the DC component of the signal: the DC offsets between amplifier stages can be eliminated; low-frequency flicker noise can be filtered out; an optimum decision threshold independent of input optical power level (Figure 4-17) can be

achieved even in a single-beam receiver (although this can also be achieved without requiring line coding by using a two-beam receiver or by other techniques [202][203]).



(a) optimum decision threshold

(b) non-optimum decision threshold

Figure 4-17: Pulse width distortion caused by non-optimum decision threshold in a DC coupled receiver

The presence of a DC offset creates several problems:

- it sets an absolute minimum on the detectable signal
- it introduces pulse-width distortion into the signal which effectively increases the minimum acceptable bit period
- it makes it difficult to implement multi-stage high gain post-amplifiers because the DC offset is amplified and could shift later stages out of the linear region of operation

The importance of DC offsets in optical receivers is well recognised [200] [204] and in general it is solved in stand-alone receivers by blocking or attenuating the DC component of the signal. Novotny [166] has previously identified the importance of offsets as a primary limit on smart-pixel receiver sensitivity in the context of FET-SEED technology and has investigated them experimentally by measuring the distribution of offsets across smart-pixel arrays. The issue of high DC offsets in GaAs / MSM optical receiver arrays has also been highlighted [204]. A recent study on CMOS smart-pixel receiver optimisation [168] also mentioned that DC offsets limited post-amplifier gain but did not include the effect quantitatively. The study presented here quantifies, for the first time, the importance of DC offsets in smart-pixel applications by applying models from the literature on MOSFET mismatch as a function of transistor parameters.

After a brief review of the physical origins of offset voltages and the standard models used to characterise transistor mismatch, an expression is derived for the minimum detectable signal in a smart-pixel receiver in the presence of offsets in terms of the physical parameters of the transistors. Parameters from a specific technology are then used to quantitatively assess the importance of offsets.

The studies of the noise and gain limits of smart-pixel receivers have been based on a particular 0.6 µm technology. To make a good comparison with these limits, it would be ideal to estimate the offset limits in the same technology. Unfortunately, information on the matching characteristics of transistors in this technology was not available at the time of writing. Instead, the offset limits are investigated using a 0.7 µm process from Alcatel-Mietec [214] in which transistor matching has been extensively characterised in open literature.

4.6.2 Physical origin of offsets

Random offsets are a direct consequence of transistor mismatch. The physical parameters of a group of identically designed transistors can be expected to show a distribution about their nominal value. The mismatch ΔP in a parameter P is formally defined as the difference between the value of the parameter for two identically designed devices and is a random variable with mean zero and variance σ_{P}^{2} , commonly with a normal distribution.

Mismatch can be caused by the stochastic nature of the various steps of the manufacturing process, which will affect two transistors independently of their relative position, and also by systematic variations in process parameters across a die due to, for example, die attachment stress [205] or gradients in oxide thickness.

The model

$$\sigma^{2}(\Delta P) = \frac{A_{p}^{2}}{WL} + S_{p}^{2}D^{2}$$
(4.22)

has been shown to model a wide class of mismatch sources [206]. Here, W and L are the width and length of the device, A_p and S_p are process dependent constants characterising mismatch due to stochastic processes and gradients respectively, and D is the separation of the transistors.

From this equation, several comments can be made about the relevance of mismatch in smart-pixel photoreceivers. Firstly, because devices are typically chosen to be small in order to reduce power consumption and layout area, random mismatch in general would be expected to be more important since the variance in the mismatch is inversely proportional to the device area, in contrast to precision analogue circuits where large devices can deliberately be used. Secondly, the device separation is small, so the second term should be unimportant.

The current in a MOS transistor in saturation can be modelled for the purposes of matching by the equation:

$$I = \frac{\beta}{2} \frac{(V_{GS} - V_T)^2}{(1 + \theta \ (V_{GS} - V_T))}$$
(4.23)

where θ is introduced to model velocity saturation and series resistance in the source. Mismatch parameters A_{vT} , A_{β} and A_{θ} can be defined.

$$\Delta V_T^2 = \frac{A_{VT}^2}{WL}; \qquad \frac{(\Delta\beta)^2}{\beta^2} = \frac{A_{\beta}^2}{WL}; \qquad \frac{(\Delta\theta)^2}{\theta^2} = \frac{A_{\Theta}^2}{WL}$$
(4.24)

Several terms contribute towards the threshold voltage of the transistor, but the dominant cause of mismatch has been shown to be due to the voltage required to deplete the substrate underneath the gate [207], which is related to the charge in the channel depletion region and the gate capacitance.

$$V_{DEPLETION} = \frac{Q_{SUBSTRATE}}{C_{OXIDE}}$$
(4.25)

The charge in the channel depletion region is in turn related to the number of dopant atoms it contains. The threshold voltage mismatch arises from the variance in the number of dopant atoms, which is a random variable with a Poisson distribution.

The physical origin of mismatch in the current factor β and the mobility reduction factor θ is variations in carrier mobility [208] [209].

Figure 4-18 and Figure 4-19 show the values of the constants A_{vT} and A_{β} from a selection of process reported in open literature. The oxide thickness is used to characterise the technology. Notice that A_{vT} scales linearly with oxide thickness but that A_{β} shows no clear trend with technology.



Figure 4-18: Matching constant A_{vT} for a range of technologies (after [209])



Figure 4-19: Matching constant A_B for a range of technologies (after [209])

Recent work [209][210][211] has shown that in sub-micron transistors, this model is inadequate for predicting the threshold voltage mismatch. In devices with short channel lengths, it underestimates the mismatch because the depth of the substrate depletion region controlled by the gate varies along the channel due to the extension of the source and drain depletion regions underneath the channel (Figure 4-20). In devices with narrow widths, it overestimates the mismatch because fringing fields at the edge of the gate can deplete an additional volume of charge which does not form part of the channel. These effects are essentially the same effects which require complex models for predicting the threshold voltage of sub-micron transistors [212]. This has been modelled in [209] by adding additional terms to (4.22) which, in a slightly modified form, can be written:

$$\sigma^{2}(\Delta V_{T}) = \frac{A_{VT}^{2}}{WL} \left(1 + \frac{L_{VT}}{L} - \frac{W_{VT}}{W}\right) + S_{P}^{2}D^{2}$$
(4.26)

where the L_{vT} and W_{vT} parameters characterise the short and narrow channel effects. However, the study showed that the simple model for the current factor mismatch remained valid.



Figure 4-20: Diagram of a MOS transistor showing the physical channel length (which is the drawn length corrected for etching effects) and the lateral diffusion of the source and drain underneath the gate. Dimensions are approximately to scale.

Table 4-4 shows the mismatch parameters for the Alcatel-Mietec 0.7 µm process characterised using this extended threshold mismatch model in [209]. Notice that the short channel effects cause a 40% increase in threshold voltage mismatch above that predicted by the standard model.

Note also that the figures for S_{vT} and S_{β} confirm that the spatially dependent mismatch component is indeed negligible in a smart-pixel application. Consequently, there is no motivation for using common centroid layout techniques to improve matching in this context.

| parameter | NMOS | PMOS |
|-------------------------------|-------------------------|--------------------------|
| A _{vt} | 11 mV μm | 22 mV μm |
| Αβ | 1.9 % μm | 2.8 % μm |
| L _{vt} | 0.67 μm | - |
| $W_{_{ m VT}}$ | 0.40 µm | - |
| $\mathbf{S}_{_{\mathrm{VT}}}$ | 0.6 mV mm ⁻¹ | 0.1 mV mm^{-1} |
| S_{β} | 0.2 % mm ⁻¹ | 0.3 % mm ⁻¹ |

| Table 4-4: Mismatch | parameters for | · Alcatel-Mietec 0.7 | µm process |
|---------------------|----------------|----------------------|------------|
|---------------------|----------------|----------------------|------------|

4.6.3 An analytical model for the offset limit on receiver sensitivity

Offset voltage of a simple transimpedance receiver

This section derives an expression for the standard deviation in the offset voltage of a transimpedance receiver with a complementary gain stage using a simple extension of a standard procedure [268].

The first step is to derive an expression for the variance in the operating point of an inverter.

Let the mean operating point of the inverter be $V_{_{OP}}$ the mean values of the transistor parameters are $V_{_{T}}$, β , θ and the operating current with the parameters at their mean value be I; suppose that, as a result of mismatch, these values vary by $\delta V_{_{OP}}$, $\delta V_{_{T}}$, $\delta\beta$, $\delta\theta$ and δI .

Assuming the variation in each parameter is small compared to its mean value, a Taylor expansion for the current in the NMOS transistor about the nominal operating point can be written using the model equation (4.23).

$$\delta I = \frac{\partial I}{\partial V_{OP}} \delta V_{OP} + \frac{\partial I}{\partial V_{Tn}} \delta V_{Tn} + \frac{\partial I}{\partial \beta_n} \delta \beta_n + \frac{\partial I}{\partial \theta_n} \delta \theta_n$$

$$= g_{mn} \left(\delta V_{OP} - \delta V_{Tn} \right) + \frac{I \delta \beta_n}{\beta_n} - \frac{I}{1 + \theta \left(V_{OP} - V_{Tn} \right)} \left(V_{OP} - V_{TN} \right) \delta \theta$$
(4.27)

A similar expression can be written for the PMOS transistor. The change in current in both transistors must be the same; by equating the two expressions for δI , an expression for $\delta V_{_{OP}}$ can be obtained.

The parameters δP are random variables with a variance $\sigma_{\delta_P}^2$ which is related to the variance in the mismatch parameter $\Delta P = (P_1 - P_2)$ by $\sigma^2(\Delta P) = 2\sigma_{\delta_P}^2$ since δP for different transistors is statistically independent.

Studies have shown [209] that $\delta\theta$ and $\delta\beta$ are almost entirely positively correlated and so:

$$\frac{\delta\theta}{\theta} = \frac{\delta\beta}{\beta} \tag{4.28}$$

Bringing all these results together, the variance in the operating point $\sigma_{_{VOP}}^{^{2}}$ is:

$$2\sigma_{V_{OP}}^{2} = \frac{g_{mp}^{2}\sigma^{2}(\Delta V_{T_{P}}) + g_{mn}^{2}\sigma^{2}(\Delta V_{T_{n}})}{(g_{mn} + g_{mp})^{2}} + \left(\frac{I}{g_{mn} + g_{mp}}\right)^{2} \left[\frac{\left(\frac{1}{1 + \theta_{n}(V_{OP} - V_{T_{n}})}\right)^{2}\frac{\sigma^{2}(\Delta\beta_{n})}{\beta_{n}^{2}}}{+\left(\frac{1}{1 + \theta_{P}(V_{DD} - V_{OP} - |V_{T_{P}}|)}\right)^{2}\frac{\sigma^{2}(\Delta\beta_{P})}{\beta_{P}^{2}}} \right]$$
(4.29)

The study of mismatch in the Alcatel-Mietec process found that a simpler current mismatch model was adequate to predict the matching in that technology. The simplified model for the current mismatch at constant gate source voltage was:

$$\frac{\delta I}{I} = \frac{2 \ \delta V_T}{V_{GS} - V_T} + \frac{\delta \beta}{\beta}$$
(4.30)

where the θ mismatch term has been dropped⁶ and a long channel approximation for g_m has been used. The mismatch parameters have been fitted to this model and so it must be used to calculate the current mismatch.

The variance in the operating point is then:

$$2\sigma_{V_{OP}}^{2} = \frac{\sigma_{In}^{2} + \sigma_{Ip}^{2}}{(g_{mn} + g_{mp})^{2}}$$
(4.31)

where the transconductance values in equation (4.31) should be calculated using a full transistor current model such as equation (4.23) that includes the effects of velocity saturation.

⁶ Note that although the velocity saturation term is not required in the analysis in this chapter, it is included in the derivation here as it is required in the investigation of the scaling of receiver performance in advanced CMOS processes in Chapter 6.

Now consider a receiver consisting of a front-end, a single post-amplifier with a gain A and a decision stage. It is easily shown that the offset voltage, defined as the nominal voltage required at the output of the first stage to produce zero output, is

$$V_{OFFSET} = -\delta V_{FRONT-END} + (1 + \frac{1}{A}) \delta V_{POST-AMP} - \frac{\delta V_{DECISION}}{A}$$
(4.32)

and hence

$$\sigma_{VOFFSET}^{2} = \sigma_{FRONT-END}^{2} + (1 + \frac{1}{A})^{2} \sigma_{POST-AMP}^{2} + \frac{\sigma_{DECISION}^{2}}{A^{2}}$$
(4.33)

In the simple case when the gain of the second stage is high, the mismatch in the postamplifier load transistors is neglected, and the second stage is the same size as the front-end, this can be approximated as $\sigma_{voFFSET}^2 = 2\sigma_{voP}^2$. This is the expression that is used to estimate the offset voltage. Note, however, that this approach may under-estimate the offset voltage because the gain of the post-amplifiers used in smart-pixel circuits can be quite low, and the transistors in the decision stage may be smaller than those in the front-end and postamplifier.

Relationship between circuit yield and offset voltage

If the standard deviation of the offset voltage is $\sigma_{voreset}$ then the probability that, in any one particular circuit, the offset is within $\pm k \sigma_{voreset}$, assuming a normal distribution, is:

$$P_{WORKING} = erf\left(\frac{k}{\sqrt{2}}\right) \tag{4.34}$$

The probability that, in a group of N circuits, all of the circuits are within specification is P_{WORKING} .

Figure 4-21 plots the value of k as a function of the required failure rate. It can be seen that, in large arrays, the worst-case offset is substantially higher than in a single channel receiver. For example, to achieve a failure rate of 1 in every 100 chips, a 4096 channel smart-pixel array must be designed to accommodate a maximum offset of $\pm 4.7 \sigma_{voFFSET}$ compared with $\pm 2.6 \sigma_{voFFSET}$ for a single channel circuit. Similar issues occur in the design of sense amplifiers in large DRAM arrays [213] and in high resolution D/A converters.

Because it would be very difficult and hence expensive to test the receivers in smart-pixel arrays for compliance with an offset specification before assembling them into a proper

optical package, it is likely that commercial devices would have to be designed with a conservative failure rate.



Figure 4-21: Relation between number of channels in smart pixel array and number of offset standard deviations

Minimum detectable signal in the presence of an offset voltage

 V_{MIN} has been defined as the peak-to-peak signal amplitude at the post-amplifier input required to produce a rail-to-rail output swing at the output of the decision stage. Equivalently, a rail-to-rail swing will be produced when the input signal exceeds its switching point by V_{MIN} / 2. It is assumed that this remains true when there is an offset voltage present. However, in this case, the actual switching point may be anywhere within a band of $\pm V_{\text{OFESET}}$ centred about the nominal switching point (Figure 4-22).



Figure 4-22: Minimum signal required with offset

The minimum input signal to produce a full swing output is therefore

$$V_{MIN} = V_{MIN} + 2 V_{OFFSET}$$
(4.35)

However, a slightly larger signal may be required to achieve acceptable pulse-widthdistortion. This is illustrated in the simulated eye diagrams in Figure 4-23.With no offset, an input signal of $V_{\text{MIN}} = 100 \text{ mV}$ is required to produce a good quality eye; with a 50 mV offset, although the minimum input voltage of 200 mV predicted by equation (4.35) is just sufficient to produce a subjectively acceptable eye, there is more pulse-width-distortion.



Figure 4-23: Simulated effect of offset on minimum detectable signal and pulse width distortion ($W_{POST-AMP} = 5 \ \mu m$; $W_{LOAD} = 0.8 \ \mu m$; small signal gain = 4.4; period = 1.5 ns). The eye diagrams were obtained by driving the input to the post-amplifier with a pseudo-random bit sequence voltage source and observing the output of the post-amplifier after a single stage of buffering.

4.6.4 Application of the offset model to a 0.7 µm process

Method

As discussed above, the offset limits are investigated using a 0.7 μ m process with a nominal oxide thickness of 17 nm from Alcatel-Mietec [214] in the absence of matching information on the 0.6 μ m process. The matching parameters for the 0.7 μ m process were given in Section 4.6.2.

Lack of information and some peculiarities of this process still require that some estimates be made in order to calculate the offset voltage. The published information on this process includes NMOS model parameters (Table 4-5) for equation (4.23). However, this information is not available for the PMOS transistors. Also, the PMOS transistors that were characterised for matching were special purpose "low-threshold" transistors requiring an additional processing step⁷ and a minimum channel length of 1.2 μ m. No short channel effects were observed because of this restriction on channel length. It is assumed that the standard PMOS transistors in this process would have the same matching parameters as the low threshold

 $^{^{7}}$ An additional threshold adjust implant to reduce the threshold voltage from -1.0 V to - 0.7 V.

devices $^{\rm s}$ and would display short channel characteristics with the same values of $L_{_{\rm VT}}$ and $W_{_{\rm VT}}$ as the NMOS transistors.

A front-end with $W_{PMOS} = W_{NMOS}$ and $L_{PMOS} = L_{NMOS} = 0.7 \,\mu\text{m}$ is assumed. The transconductance of the PMOS transistor is estimated to be half that of the NMOS transistor and the operating point is estimated to be 1.75V based on the operating point of an inverter in the 0.6 μm technology with the same W_{PMOS} : W_{NMOS} ratio.

| β | $550 \mu\text{A/V}^2$ |
|----------------|-----------------------|
| V _T | 0.743 V |
| θ | 0.560 V ⁻¹ |

Table 4-5: Model parameters for Alcatel-Mietec 4µm / 0.7µm NMOS transistor

The offset voltage was calculated as a function of $W_{FRONT-END}$ assuming a 1024 channel receiver array with a required yield of 0.99 (k = 4.5) for designs having a $W_{FRONT-END}$: $W_{POST-AMP}$ ratio of 1:1 and 2:1. Figure 4-24 shows the results of the calculation.

In this technology, about 75% of the offset voltage could be attributed to the threshold voltage mismatch term.

⁸ In practice, the extra threshold adjust implant would degrade the matching; however, the additional implant dose is a factor of 3 lower than the implant dose used on both $low-V_{T}$ and normal PMOS transistors so the difference in matching should be quite small.



Figure 4-24: Calculation of the worst-case offset voltage as a function of transistor width in the Alcatel 0.7µm process

Discussion

The results in Figure 4-24 show that the offset voltage is quite large for inverters constructed from minimum length transistors. To an extent, this is due to the large number of standard deviations that must be considered in a smart-pixel array with many elements. These values are also significantly larger than the estimate made in the smart-pixel design study in [168].

In particular, the minimum signal imposed by the offset voltage, which is twice the offset plotted in the graph, is significantly larger than the estimates of the thermal noise in Section 4.4. This indicates that, in a DC coupled receiver, thermal noise does not set an important limit on receiver performance.

The results also suggest that the offset may be more important in determining the sensitivity than the post-amplifier gain limit under many circumstances. Direct comparison with the 0.6 μ m process post-amplifier gain limits is difficult because of the use of a different process for the analysis; however, the worst-case offset voltage varies between ±20 mV and ±70 mV for reasonably dimensioned transistors and it has been seen that the gain-limited V_{MIN} ranged between 40 mV and 250 mV depending on bit-rate and power consumption requirements. The offset limit is more important in low bit-rate circuits where it is possible to achieve high post-amplifier gain and/or sufficient bandwidth with narrow transistors.

Taking offset voltage into account in the design process leads to a slightly different design approach. Firstly, it limits the amount of gain that can be implemented in the post-amplifier to between 10 and 20. Also, if a design is offset limited, then a higher overall sensitivity can be attained by putting more gain/less bandwidth into the front-end and less gain/more bandwidth into the post-amplifier stage; this contrasts with the post-amplifier gain-bandwidth limited case where it is better to make the bandwidths of the two stages the same as discussed in Section 4.5

4.6.5 Possible solutions to the offset problem

In this section, some possible solutions to the offset problem are briefly considered.

Bipolar transistors tend to have lower offsets than MOSFETs because they do not have a threshold voltage. There is a case for investigating whether they might be more suitable for constructing large, DC coupled receiver arrays although there are many other differences to consider, not least of which is the additional cost of a BiCMOS process.

Regenerative latches can in principle be compensated for offset during the reset phase of the latch. This approach is used in sense amplifier arrays in DRAM chips [213] and in comparators [215]. It may be possible to adapt these techniques to work with the sense amplifier class of photoreceiver described briefly in Section 4.2.3, although no specific implementations have been proposed to date.

Offsets can be eliminated completely if it becomes feasible to implement a receiver with a lower frequency cut-off in large arrays. The difficulty in implementing the lower frequency cut-off comes from the area required to implement the filter capacitor, and, to a lesser extent, from the additional power consumption necessary in a differential decision stage. Coding circuitry is not necessarily prohibitive: for example, if the smart-pixel chips are only used to implement a switching fabric, then it is sufficient to code and decode the data at the inputs and outputs of the switching fabric and not at all intermediate stages.

A simple calculation allows the area of the filter capacitor to be estimated. The lower frequency cut-off causes droop in the signal amplitude when there is a long sequence of ones or zeros in the data; this in turn leads to pattern dependent jitter [196]. In a first approximation, the maximum allowed lower cut-off frequency is determined by the maximum run length of the coding scheme, the data rate, the amount of pattern dependent jitter that is acceptable and the edge time of the data signal. In the case of 8B/10B coding, the maximum run length is 5 [201]; assuming a data rate of 1 Gbit/s, and a maximum permissible droop of 5% of the signal amplitude, then the lower frequency cut-off must be

97

1.6 MHz. A resistance of 100 k Ω , which could be readily implemented using an ohmic region MOSFET of a few microns in length, would require a 1 pF capacitor⁹. The capacitor value is non-critical and it can be implemented using the gate oxide; the capacitor would occupy 20 × 20 µm² in 0.6 µm technology scaling to 5 × 5 µm² in a hypothetical future 0.05 µm process with a 1 nm equivalent gate oxide thickness [237]. A possible approach is described in Chapter 6.

At lower data rates in 0.6 µm technology, the capacitor area is unacceptable for use in large arrays; however, it appears that it will become increasingly feasible to implement the low frequency cut-off as channel data rates reach 1 Gbit/s and as the oxide capacitance per unit area increases. This is a good argument for pursuing power efficient implementations of such receivers; such an approach may allow smart-pixel receivers to approach noise limited sensitivity.

Example implementations of receivers with a low-frequency cut-off can be found in [200][216][217][230].

4.7 Conclusion

In this chapter, the basic structure of a smart-pixel receiver and how it compares to a conventional telecommunications receiver has been reviewed. The design variables that can be used to achieve a trade-off between sensitivity, power consumption and speed have been identified and the analysis of their influence provides a starting point for a design procedure for smart-pixel receiver circuits.

In particular, the linear relation between photodiode capacitance and both power consumption and switching energy has been highlighted.

The relative importance of thermal noise, post-amplifier gain and mismatch-related offset on sensitivity has been discussed. In marked contrast to conventional telecommunications receivers, noise is not a limiting factor: firstly, because the low photodiode capacitance resulting from hybrid integration provides good noise performance even at low power dissipation; secondly, because the DC coupled nature of the design imposed by layout area constraints and the limited number of post-amplifier gain stages imposed by power consumption limitations make it difficult to design a post-amplifier that can detect signals at

⁹ The capacitor would have to be larger by the factor by which the offset was attenuated if the RC filter forms part of a feedback loop.
the noise limit. It is the properties of the post-amplifier, not those of the front-end, that limit the performance of smart-pixel receivers.

The analysis in the chapter has been based on a simple small-signal model of the receiver circuit and has considered the performance of an individual receiver circuit in isolation without consideration of its environment and without consideration of practical issues such as tolerance to process variation and realisation of resistors in a digital CMOS technology. In the next chapter, a specific case study of a receiver design for the SPOEC system described in Chapter 3 is used as a focus for the discussion of these issues. Chapter 6 will consider how the conclusions of this chapter are modified by the changes in receiver performance that occur as a result of advances in silicon technology.

4.8 Appendix: modelling of feedback resistor

This section justifies the equivalent circuit used for the feedback transistor and contains comments on the correctness of the BSIM simulation capacitor models for modelling the feedback transistor.

The feedback transistor operates in the ohmic region and has the same electrical properties as a distributed RC line.

A distributed RC line can be modelled by two-port Y parameters. The analytic solution for a distributed line is [218]:

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} \sqrt{s} \coth \sqrt{s} & -\sqrt{s} \operatorname{csch} \sqrt{s} \\ -\sqrt{s} \operatorname{csch} \sqrt{s} & \sqrt{s} \coth \sqrt{s} \end{bmatrix}$$
(4.36)

where the complex frequency variable s has been normalised to $\omega_T = 1 / R C$.

By forming a first order Taylor expansion about s=0, the following approximation to the twoport parameters is obtained:

$$[\mathbf{y}] \approx \begin{bmatrix} 1+s/3 & -1+s/6 \\ -1+s/6 & 1+s/3 \end{bmatrix}$$
(4.37)

This approximation predicts the real and imaginary parts of all y parameters to better than 5% up to a frequency of about 1.5 ω_r .

The first order approximation to [y] can be realised with the equivalent circuit in Figure 4-25.



Figure 4-25: First order equivalent circuit of distributed RC line

Note the negative drain-source capacitance used to model the imaginary component of y_{12} and y_{21} . This will tend to cause more overshoot than the simple pi-section model without the source-drain capacitor which has Y parameters:

$$\begin{bmatrix} \mathbf{y} \end{bmatrix} = \begin{bmatrix} 1+s/2 & -1 \\ -1 & 1+s/2 \end{bmatrix}$$
(4.38)

Figure 4-25 is equivalent to the Ward MOSFET capacitance model [219] as implemented in the BSIM 3v3 model [212] with the 40/60 drain-source charge partitioning option selected (XPART=0). This is a first order approximation to non-quasi-static behaviour [220][221]. The default 0/100 charge partitioning option (XPART=1) has y parameters:

$$\begin{bmatrix} \mathbf{y} \end{bmatrix}_{0-100} = \begin{bmatrix} 1 + \frac{3}{4} s & -1 - s/4 \\ -1 - s/4 & 1 + \frac{3}{4} s \end{bmatrix}$$
(4.39)

and produces incorrect physical behaviour in the context of a transimpedance amplifier circuit. Notice that the source-drain 'capacitance' here is positive and will cause simulations to underestimate the overshoot and underestimate the bandwidth.

4.9 Appendix: behaviour of decision stage

The behaviour of the decision stage is defined by the minimum signal swing at the input required to produce a fully restored digital signal at the output with an edge which is short compared with the bit period.

The behaviour of a simple inverter as a decision stage was investigated using transient simulations in the example $0.6 \mu m$ process considered in this chapter.



Figure 4-26: Decision stage used to estimate $V_{\mbox{\tiny DECISION}}$

The inverter was driven by a voltage source with a hyperbolic tangent pulse shape symmetrical about the operating point of the inverter. The edge time was defined between the 10% and 90% of the wave-form.

The inverter was loaded with two identical stages and a parasitic capacitance of 5 fF.

Rising and falling edges of amplitudes between 0 and 2V and rise times between 0.5 ns and 5.0 ns were applied. The output rise and fall time was measured between 1V and 5V.

The results showed that a minimum input signal in the range 700-900 mV was required to produce an output edge which is shorter than the input edge. The minimum input signal increased slightly at higher data rates, but only by about 10%. For the purposes of the comparison of switching energy at different bit rates, it is assumed that the minimum input signal required at the input of the decision stage is independent of bit-rate to a first approximation.

Chapter 5

Case study: a single-beam receiver for SPOEC

5.1 Introduction

This chapter describes the detailed design of the single-beam receiver circuit used for the data channels in the main switching chip of the SPOEC system outlined in Chapter 2.

As well as providing a concrete illustration of the general design trade-offs discussed in the previous chapter, the circuit highlights certain practical issues that constrain the design of smart-pixel receivers, including process variation and the resistance of the chip power-supply distribution network.

The chapter reports experimental results from an electrical test version of the receiver design that verifies correct DC operation of the circuit. A version of the circuit designed for optical inputs has been fabricated as part of the main switching chip and is undergoing flip-chip assembly at the time of writing.

The chapter begins by reviewing the main system requirements of the circuit. It then discusses the detailed implementation and experimental characterisation, and evaluates the design in light of the results.

5.2 System requirements

The switching chip contains approximately 4000 data receiver circuits; power consumption is therefore a significant concern. The target power consumption on the worst-case process corner was set at 5 mW per receiver to give a peak power consumption of 20 W during the header phase. From experience in the design of the cooling system for similar circuits¹, this power consumption is close to the maximum that can be extracted from a chip of this size without using forced-air cooling and while keeping the junction temperature below 55°C to maintain good MQW modulator device performance. Thus a facility to power-down receivers that are not in use was required.

The sensitivity of the receiver is set by the optical output power of the VCSELs together with the budget for power loss in the optical path between the VCSELs and the switching chip. Initial estimates indicated that a peak photocurrent of around 5 μ A would be available; additional losses discovered in the course of the design reduced this figure to 3.5 μ A.

¹ The SCIOS demonstrator system briefly described in Chapter 2.

The required dynamic range was small. The optical losses in the system were well controlled and the uniformity of VCSEL output power across an array is high. The final laser array displayed a spread in output power at fixed operating current of between 1.1 mW and 1.4 mW [222]. The circuit was specified for a dynamic range of about 2. Although a commercial system would have to allow for a higher spread in the VCSEL output power, the optical losses in a free space optical system would remain well controlled; it can generally be expected that smartpixel receivers do not require high dynamic range.

A single-ended receiver was a requirement given the number of channels and the limit on the number of diodes that could be fabricated in the InGaAs process.

Individual receivers had to be distributed across the area of the chip and surrounded by blocks of digital logic. This tightly constrained the layout area and led to a DC coupled design based on self-contained complementary inverter gain stages. NMOS/PMOS gain stages were not used because they require bias circuitry; sharing this circuitry between receivers would have been difficult because of the need to distribute a noise-sensitive bias voltage to the current-source load. Including the bias circuitry in each receiver cell would add a significant layout area and power consumption overhead. The bias signal was also found to require substantial decoupling capacitance to prevent noise from the post-amplifier from disturbing the front-end. A possible compromise may have been to group the receivers into blocks of four in which there was no digital circuitry.

The possibility of using a synchronous sense amplifier in the receiver was not investigated in any detail, primarily because the optoelectronic designs proposed to date are more suitable for two-beam operation and also work better with return-to-zero data [173]. It has been proposed [167] that asynchronous receivers can enable operation at a speed beyond the normal rated speed of a particular silicon technology; one of the goals of the design was to investigate the extent to which this is true in a practical system. The system architecture in principle allows the data section of the packet to operate at a higher bit-rate than the header.

The initial target data rate for the receiver was 250 Mbit/s. The circuit was originally designed for operation with a 20 μ m diameter photodiode with a capacitance of 50 fF; however, changes in the specifications of the optical system meant that the final photodiode had a diameter of 35 μ m and a capacitance of 95 fF. In the end, this specification proved too demanding within the other constraints on the design; the final circuit is expected to operate between 150 Mbit/s and 200 Mbit/s.

5.3 Detailed design considerations

5.3.1 Circuit schematic

A schematic of the receiver circuit is shown in Figure 5-1. The circuit was implemented in the 0.6 µm digital 5 V CMOS process used for the design study in Chapter 4.



Figure 5-1: Schematic of SPOEC data receiver (final version)

The design is similar to those discussed in Chapter 4. The front-end Mn1/Mp1 is followed by a single stage post-amplifier Mn2/Mp2 with gain broadening transistors Mn3/Mp3. The narrow width of Mn2/Mp2 means that acceptable second stage gain cannot be achieved using broadening transistors with a matched channel length. This gives rise to a spread in systematic offset voltage between the first and second stage due to process variation of about 15 mV. The gain of the second stage is about 5.

The operating points of the first two stages are nominally equal (about 1.9 V) but the operating point of the decision stage is designed to be somewhat higher (about 2.5 V) to introduce the systematic offset required for single-beam operation (Figure 5-2). The single ended design used in earlier work [231] introduced the fixed offset between the first and second stages. Introducing the offset after an additional stage of gain allows the input referred offset to be set more precisely for a given absolute voltage tolerance on the operating point of an inverter, but

leads to slightly reduced dynamic range. It also allows identical transistors to be used for the first two stages, which should improve matching.



Figure 5-2: Illustration of switching thresholds in receiver

The decision circuit consists of two inverter stages as discussed in Section 4.5.7.

The length of the transistors in the second inverter of the decision stage is made quite long in order to slow down the switching transients on the power supply (see Chapter 8).

The output of the amplifier is loaded by a single digital inverter connected to the digital supplies (not shown) which drives the standard cell logic.

5.3.2 Summary of simulated performance

Detailed simulation results are given in Appendix A. In summary, the DC switching point of the receiver occurred at an input photocurrent of 2.2 μ A with an additional 0.4 μ A penalty to allow for random offset. Dynamic performance was acceptable up to somewhere between 150 Mbit/s and 200 Mbit/s with 95 fF photodiode capacitance.

5.3.3 Power supply distribution

Two separate analogue power supplies were used to address problems with power supply noise: the first supplied the front-end and post-amplifier (AVDD! / AGND!) and the second supplied the decision stage (ATHVDD! / ATHGND!). Both supplies were separate from the digital supply. The analogue power supply rails were fed vertically across the chip from the top edge to the centre and from the bottom edge to the centre. Each half-column of 32 receivers had its own set of analogue rails. The rationale behind the power supply strategy is discussed in more detail in Chapter 8.

The need to keep the IR voltage drop across the analogue power supply rails to an acceptable level limited the power-supply current of the circuit. This limit was more fundamental than the power consumption constraint; whilst disabling amplifiers during part of the cycle helps to reduce the average power consumption, it does not reduce the peak voltage drop.

The process used to implement the design had only two levels of metal. The final design used metal 2 widths of 15.4 µm for AVDD! / AGND! and metal 2 widths of 4.9 µm for ATHVDD! / ATHGND!; thus, the analogue power supply rails occupied about 27% of the horizontal extent of a pixel. The estimated voltage drops are shown in Table 5-6. Although the voltage drops on the worst case process corner are only borderline acceptable, increasing the width much further would have led to an unacceptably large chip. A contingency has been provided to selectively disable the amplifiers in each row of super-pixels in case the voltage drop proves problematic during experimental operation of the system.

| process | AVDD! | half-column | drop in power | power consumption |
|------------|-------------|-------------|----------------|----------------------|
| corner | current per | resistance | supply voltage | (analogue + digital) |
| | receiver | | | |
| typical | 0.34 mA | 19 Ω | 210 mV | 2.8 mW |
| worst case | 0.54 mA | 35 Ω | 600 mV | 4.1 mW |

Table 5-6: Estimated DC voltage drop in SPOEC switching chip

To an extent, this problem is an artefact of the use of a two-level metal process. Availability of an additional level of metal that could be dedicated to power supply distribution would have significantly eased the problem.

In retrospect, it may have been better to use a 3.3 V power supply for the design, because the g_m/I figure of merit is more than a factor of two better (Table 4-3); a circuit with the same transconductance could have been designed with say half the current which would have produced a smaller relative voltage drop along the supply rail.

5.3.4 Matching considerations

An acceptable supply voltage drop required a low supply current and thus narrow transistors; minimum-length transistors would give poor matching because of the small transistor area. Transistors slightly longer than minimum (L=0.9 μ m) were selected to increase the area of the transistors and to reduce short channel effects on the matching. This also helped to improve the process sensitivity of the operating point and the supply current but increased the overshoot in the step-response because of the higher gain of the long-channel transistors.

Matching parameters for this process were not available. The threshold-voltage mismatch term was estimated by extrapolating the trend in Figure 4-18 to the oxide thickness of the process. Neglecting the current-factor mismatch term, equation (4.29) gives $\sigma_{vofFSET}$ of 6.6 mV. Allowing for ±4.7 σ to ensure that all 4000 receivers are within specification gives an offset voltage of ±31 mV. This neglects any short channel effects in the threshold-voltage mismatch.

5.3.5 Feedback resistor

A MOS transistor was used to implement a non-linear feedback resistor. The choice of an NMOS rather than a PMOS transistor ensures that the rise-time of the circuit is not degraded at high input currents but has the disadvantage of a lower sheet resistance.

A major difficulty in the design of this circuit was implementing a feedback resistor which was not sensitive to process variation. There were two main reasons for this problem:

- the requirement to keep parasitic capacitance low meant that the dimensions of the transistor were small and therefore not well controlled.
- the operating point of the front-end inverter varied significantly with process (± 250 mV worst-zero to worst-one) and therefore using a fixed bias voltage such as the power supply rail did not give good control of the resistance.

Using a fixed gate voltage of 5 V gives a feedback resistance which varies between 23 k Ω and 73 k Ω between fast and slow process corners.

Following the approach of others [197][223], this problem was addressed by using an external bias voltage to control the gate resistance. In an experimental system of this kind, in which there is some uncertainty about whether the system components will meet their specified performance, a facility to influence the speed/sensitivity tradeoff during actual operation is useful in its own right.

Using an operating point below mid-rail gives a wider range of adjustment of the gate-source overdrive $(V_{GS}-V_T)$ of the feedback transistor. This is quite important because the threshold voltage of the feedback transistor is significantly increased by the body effect. The nominal bias voltage of 4.0 V gives a gate-source overdrive of 0.66 V. The nominal overdrive voltage cannot be too small to allow for some variation in the operating point of the front-end across the chip due to voltage drops in the power supply. Providing sufficient supply head-room for adjustment of the feedback resistor on the slow process corner was one of the main reasons for the choice of a 5 V supply.

Whilst the approach of using of an external bias signal does provide a means of compensating against process variation, it is not without disadvantages. The main problem is that the signal has to be distributed globally across the chip; it is therefore susceptible to pickup of noise from the digital parts of the circuit, particularly since the node is high impedance. Fortunately, the circuit is not particularly sensitive to small amounts of noise on this node, and careful routing of the bias signal should in theory allow this problem to be addressed. However, requiring that an analogue signal be routed close to digital logic throughout the majority of a chip makes the task of designing and verifying a circuit significantly harder, especially when, as in this project,

different people are working on the analogue and digital sections of the design. It remains to be seen the extent to which noise coupling onto this signal proves to be a problem experimentally.

An arguably better approach is to set the bias voltage on the gate of the feedback transistor in a way which compensates for process variation [198]. A technique for doing this was implemented in the clock receiver described in Chapter 7. However, in the context of this receiver circuit, the overhead associated with this approach was deemed to be unacceptable.

A simpler approach might be to utilise the high-resistance polysilicon resistors available as an option in many digital processes. Typically, these offer a sheet resistance of about $1 \text{ k}\Omega$ / square and an absolute tolerance of about $\pm 25\%$ for a narrow resistor. These would be particularly suited to high speed designs with a relatively low feedback resistance.

5.3.6 Disable mechanism

A requirement of the design was a facility to power down the amplifier. During the header phase of the packet, the outputs of all of the receivers are processed in order to perform address recognition and arbitration. However, during the data phase, only the outputs from receivers on input channels that have been routed to one of the outputs of the switch are actually used; it is possible to power down the other receivers in order to reduce the power consumption of the system.

The disable circuitry (Figure 5-1) was designed to have minimal impact on the performance of the circuit in normal operation. All the additional transistors connect to non-critical nodes in the circuit. In normal operation, Mn6 and Mn7 are off and Mp7 operates as a pass transistor.

When *disable* is asserted, transistor Mn6 shunts the photocurrent to the analogue ground, pulling the input node low. The bias voltage is disconnected from the gate of the feedback transistor by opening Mp7; Mn7 pulls the gate of the feedback transistor low to prevent current flow through the feedback transistor from *y* to *optin*. An additional transistor (Mp6) is required to turn off the bias current in the gain broadening transistors.

Lucent have previously used power down techniques to reduce static power consumption in an experimental smart-pixel crossbar system [224] using a slightly different circuit approach where a NAND gate is used in place of an inverter as the gain element in the front-end. The technique used here has less impact on the operation of the front-end.

The circuit takes about 25 ns to stabilise after being enabled (see Appendix A).

5.3.7 Circuit layout

The circuit was laid out to fit within a 38 μ m high standard cell row (Figure 5-3) to make it compatible with standard-cell place-and-route. 8 μ m of this height had to be reserved for the

horizontal digital power supply rails. To fit the circuit within this height required aggressive packing of the transistors.

Care was taken to use a symmetrical layout for the transistors for which matching was critical (Mn1/Mn2,Mp1/Mp2). These transistors were surrounded by guard rings. However, common centroid layout was not used in the final layout; as explained in Chapter 4, this only helps to reduce mismatch caused by process gradients which are extremely small over the area of the receiver circuit. Each transistor was constructed from a pair of unit transistors with a common-drain contact to minimise drain junction capacitance.

Noise on the feedback-resistor bias signal was controlled by minimising the on-chip resistance. A separate bias pin was used for each quadrant of the chip; a 3 μ m wide vertical metal 2 track at both sides of each super-pixel column was used for global distribution.

The signal was routed far away from any 5 V digital logic signals and a limited signal shield [225] was provided within a super-pixel using the digital power and ground supplies².

² Voltage noise on these lines should be in antiphase; capacitive coupling should cancel to first order.



Figure 5-3: Layout of data receiver circuit. Digital power supply rails run left to right in metal 1; analogue power supplies run vertically. The flip-chip pad is located to the right of the analogue power supply rails. The circles on the flip-chip pad represent the overglass cut and wettable metal deposition. The large transistors underneath the wide analogue power supply rails are used as decoupling capacitors.

5.4 Experimental performance

5.4.1 Test structures

A version of the receiver design described in the previous section has been fabricated by an external foundry and tested experimentally. This circuit was intended to permit electrical testing of the receiver circuit in advance of the availability of photodiodes suitable for flip-chip integration. No optical tests have been performed to date.

Meaningful electrical high-speed characterisation of smart-pixel receiver circuits is difficult because the circuits are designed to operate with very low input capacitance photodiodes. To obtain accurate results, the parasitic capacitance of the input test fixture must be small and reasonable well controlled.

Standard 50 Ω high-speed probes are unsuitable because the source impedance must be much higher than the input impedance of the receiver. Current probes, consisting of a 50 Ω transmission line with a high-value resistor in series with the probe tip, could be used. However, the maximum frequency of operation is limited by the relatively high resistor value required to provide a good current source to a smart-pixel receiver, which typically has a higher input impedance than a conventional receiver because of the low photodiode capacitance, and the self-capacitance of the series resistor (typically 10-50 fF).

The approach used, which was only partially successful, was to include a test structure that mimicked the electrical characteristics of a photodiode at the input of the photoreceiver. The test structure consisted of a metal2-metal1-poly sandwich capacitor with a nominal capacitance of 50 fF (matching the capacitance of the originally specified photodiode diameter of 25 μ m) and a simple voltage-controlled current-source with a high input-impedance to allow the circuit to be driven from a 50 Ω line. The current-source was implemented with a PMOS transistor. The input voltage was also connected to a nominally identical transistor which was used to calibrate the output current as a function of input voltage.

These test structures can be seen bottom-left in the photomicrograph of the receiver circuit shown in Figure 5-4: the sandwich capacitor is to the left of the voltage controlled current source. A schematic is shown in Figure 5-5.



Figure 5-4: Photomicrograph of test chip implementation of data receiver



Figure 5-5: Schematic of data receiver test structure

Unfortunately, experimental tests revealed a design flaw in this test structure that limited highspeed operation: the transconductance of the PMOS transistor rolls off at frequencies close to its transit frequency. Simulations had failed to predict this behaviour because quasi-static transistor models, which assume that the drain current responds instantaneously to changes in the terminal voltages, had been used [226][227]. The room-temperature cut-off frequency calculated from the basic process parameters using a second-order non-quasi-static longchannel transistor model [228] is about 40 MHz.

The output of the receiver circuit is buffered by a single digital inverter connected to a library pad cell that drives a discrete 1 k Ω resistor in series with a terminated 50 Ω transmission line to form a low-parasitic passive probe [229].

There were some minor differences between the test chip implementation of the receiver and the final design: the test chip used a slightly smaller feedback resistor (W/L=1.0 μ m / 4.0 μ m) and a decision stage with a slightly lower operating point (Mn4 had a width of 1.8 μ m). The

changes were made to account for the revised optical power budget as discussed in Section 5.2. The test chip design also used a common power supply for all three analogue stages.

The test circuit was packaged in a 48-pin dual-in-line (DIL) package.

5.4.2 Results

DC transfer characteristics

The experimental DC transfer characteristic of the circuit (Figure 5-6) agreed reasonably well with the simulation results presented in Appendix A. The DC sensitivity at the nominal feedback-resistor bias voltage of 4 V was about 2 μ A.



Figure 5-6: Experimental DC transfer characteristic of data receiver circuit as a function of feedback resistor bias voltage

Dynamic performance

Only limited tests of the dynamic performance of the receiver were possible because of the limited bandwidth of the test structure. Nevertheless, the results give a lower bound on the performance that can be expected from the design.

Figure 5-7 and Figure 5-8 show eye diagrams obtained by applying pseudo-random bit sequences to the circuit. Clean eye diagrams were produced at 50 Mbit/s at the target sensitivity of 5 μ A. Operation up to a speed of 100 Mbit/s was possible at this sensitivity but with significant eye-closure due to pattern-dependent-jitter. The maximum speed is consistent with the calculated cut-off frequency of the test structure.



Figure 5-7: Eye diagrams for the data receiver. A rising edge in the output waveform corresponds to a rising edge in the test-transistor gate voltage.





Slightly faster operation was obtained by operating the circuit with a higher input current and a lower feedback resistance. This can also be explained in terms of the behaviour of the test structure: to produce the higher input current, it is necessary to increase the gate-source drive voltage on the input transistor, and the transit frequency of the input transistor increases in proportion to this voltage. The feedback resistor must be adjusted to maintain an optimum decision threshold at the higher input current. The results do not in themselves provide conclusive evidence that this is a correct explanation; indeed, the faster operation could also be attributed to the increase in front-end bandwidth with the smaller feedback resistance. However, results in Chapter 7 from another circuit employing the same test structure support the first explanation.

Inadequate buffering of the receiver output in the test circuit may also have contributed to the slow performance. The parasitic routing capacitance between the buffer and the external output cell was not allowed for in the test-chip design; the calculated edge time at the input to the library pad driver is about 4 ns which is enough to produce some pattern-dependent-jitter above 150 Mbit/s.

Dynamic range

The circuit had limited dynamic range. Experimental eye diagrams (Figure 5-8), taken at a fixed data rate of 50 Mbit/s and a fixed feedback bias of 4.25 V, illustrate the problem. There is only a small amount of pattern-dependent-jitter in the eye diagrams, indicating that the test structure is adequate for this bit-rate. As the input current is increased above threshold (4.24 μ A), the eye opening increases at first; the maximum eye opening is obtained when the eye is centred, indicating that the decision threshold is equal to half the peak input current. As the input current is increased further, the eye opening reduces; the dynamic range, defined in terms of the input current required to reduce the eye opening at the 20% and 80% levels to less than half the bit-period, was 3.7 (5.7 dB).

Although dynamic range is not a critical requirement for smart-pixel systems and the measured dynamic range is adequate for this application, it is questionable whether this figure is good enough to ensure robust, long-term operation in a real-world system. The root cause of the poor dynamic range – the fixed decision threshold of the receiver – is, however, a fundamental characteristic of a single-beam receiver with a response down to DC. This is a significant argument against employing a single-beam data link in a smart-pixel system. Several alternative solutions to improve the dynamic range do exist: a diode-connected transistor in parallel with the front-end feedback transistor can be used as a clamp to limit the voltage swing at the output of the front-end to the (body-effected) threshold voltage of the clamp transistor [231]; the overall bandwidth of the receiver can be increased to reduce pulse-width-distortion at

116

high power levels at the cost of reduced sensitivity; attenuating the DC component of the signal or, equivalently, adaptively setting the decision level based on the mean operating current [230] would also help but would add significant complexity to the design.

5.5 Conclusion

This chapter has presented a detailed case-study of a practical receiver design of the type analysed in the previous chapter. Several issues in this circuit made it difficult to achieve in practice the performance figures predicted by the theory of the previous chapter. These include the large process variation in the front-end feedback resistance and the difficulty of obtaining good matching in a very low-power design with minimum-length transistors.

The experimental results from a prototype circuit with an electrical input confirm that the receiver design meets the DC requirements of the target application. The experimental DC sensitivity is about 2 μ A peak input photocurrent. A limitation of the test circuitry made it impossible to verify high-speed performance in full. Nonetheless, results confirm that operation to at least 100 Mbit/s with an input current of 5 μ A is possible. Simulations of the slightly modified final design suggest that operation to somewhere between 150 Mbit/s and 200 Mbit/s at the target sensitivity of 3.5 μ A can be expected.

The failure to meet the initial design target can be attributed to the constraint on power-supply current imposed by the two-level metal process and the high sensitivity required by the high fan-out system architecture. In terms of the analysis of the previous chapter, the supply current limited the value of B_0 that could be achieved using a 5 V supply to about 600 Mbit/s under typical process conditions and thus it was not possible to achieve a switching energy close to the low-frequency limit at a bit-rate of 250 Mbit/s. At this bit-rate, the receiver operates close to the steep part of the graph in Figure 4-6 of switching energy against B. A 3.3 V supply would have given a higher value of B_0 for the same supply current and might have given better overall performance. An interesting conclusion of this work is that the limit on receiver performance in arrays of this size (in terms of number of channels) is strongly determined by the root cause of the limit on the power supply current – the properties of the power distribution layer. This should be borne in mind in the selection of the silicon process in future projects.

In this system, the choice of an asynchronous receiver did not allow the data channel to operate at a higher speed than the digital logic because of these additional constraints on the receiver design. For this particular application, a synchronous receiver may have provided better overall performance. Asynchronous receivers are expected to be more competitive in architectures that use low fan-out or point-to-point links and have more relaxed sensitivity requirements.

117

Nevertheless, the switching energy³ of the receiver is comparable with other similar circuits. The simulated value with a 0.5 A/W photodiode is 47 fJ at 150 Mbit/s with a power consumption of 2.8 mW compared to the experimental switching energy of 41 fJ at 550 Mbit/s with a power consumption of 5 mW of the single-beam smart-pixel receiver described in reference [231].

A version of the design with optical inputs has been included in the SPOEC switching chip. Although the analysis of chapter 8 suggests that electrical crosstalk is likely to limit the performance of this circuit, the results in this chapter have demonstrated that the sensitivity of the circuit is sufficient to allow an experimental investigation of simultaneous operation of a large array of smart-pixel receivers.

³ Defined for a single-beam receiver as the peak optical input energy per bit.

Chapter 6

Scaling of receiver performance in advanced CMOS technology

6.1 Introduction

There will not be a commercial requirement for optoelectronic interfaces of the kind considered in this thesis for a number of years, by which time silicon technology can be expected to have advanced significantly beyond the 0.6 µm technology used as the basis for this study. Indeed, even today, state-of-the-art commercial silicon processes have reached 0.18 µm [232] with production 0.10 µm technology scheduled for 2001 [233]. It is the performance of receiver circuits in these more advanced technologies that will determine the commercial feasibility of smart-pixel circuits; it is valuable to make projections about this performance, both as a guide to research in other areas which are impacted by aspects of receiver performance (for example, by giving guidelines on the amount of optical power required from optoelectronic devices to implement a data link) and as a means of identifying aspects of receiver design requiring further research attention.

There have been several attempts to look at the scaling of receiver performance. Williams [198] considered the scaling of noise limited stand-alone receivers and showed that, under an implicit assumption that the post-amplifiers limited the noise bandwidth, the noise limited sensitivity improves as the square root of the reduction in channel length. Krishnamoorthy and Miller [167] have looked at scaling of smart-pixel receiver circuits due to gain and noise limits based on empirical data on transistor performance and predict improvements in speed, switching energy and power consumption but make assumptions about aggressive scaling of photodiode diameters to 5-7 μ m which may be incompatible with cost-effective packaging. Van Blerkom [168] has also looked at receiver performance in 0.8 μ m, 0.6 μ m and 0.1 μ m technologies but does not draw general conclusions on scaling trends.

The emphasis of this study is on predicting general trends in various performance figures with a view to identifying how the structure of receivers might evolve, rather than on making precise quantitative predictions of performance. A key feature is the inclusion of matching in the analysis.

The study is approached in two steps: first, it predicts the way in which the basic transistor characteristics will scale using a combination of theoretical analysis based on MOSFET scaling theory [234] and empirical data. Then, it considers how the performance characteristics of a receiver circuit implemented in an existing technology will change when all transistor

119

| Parameters changing by definition | description | scaling |
|-----------------------------------|----------------------------------|--------------|
| W | transistor width | 1/α |
| L | channel length | $1/\alpha$ |
| T _{ox} | oxide thickness | $1/\alpha$ |
| V _{DD} | supply voltage | $1/\alpha$ |
| $\mathbf{V}_{_{\mathrm{T}}}$ | threshold voltage ¹ | $1/\alpha$ |
| Parameters changing as a | | |
| result | | |
| $V_{_{ m GS}}$ - $V_{_{ m T}}$ | gate-source drive ² | $1/\alpha$ |
| $g_{_{ m m}}$ | transconductance | 1 |
| $C_{_{ m GS}}$ | gate-source capacitance | $1/\alpha$ |
| $\mathbf{f}_{_{\mathrm{T}}}$ | transit frequency | α |
| Ι | supply current | $1/\alpha$ |
| Р | power consumption | $1/\alpha^2$ |
| I / g_m | I per unit g_m figure of merit | $1/\alpha$ |

Table 6-1: Predictions of first-order scaling theory for transistor parameters

dimensions and operating voltages are changed according to these scaling rules based on the expressions describing the receiver performance derived in Chapter 4.

6.2 Scaling of basic transistor characteristics

6.2.1 Ideal scaling

Classical constant-field MOSFET scaling theory [234] predicts that if all the linear dimensions and supply voltages of a transistor are scaled down by a factor α , where $\alpha > 1$, and all doping densities are increased by a factor α , then the electric fields within the device are unchanged and so the basic form of the transistor characteristic is unaltered.

The predictions for the constant-field scaling of basic transistor parameters based on a first order transistor model are shown in Table 6-1.

This chapter uses the example 0.6 μ m technology from Chapter 4 as a reference ($\alpha = 1$).

 $^{^{1}}$ V_T does not scale exactly as 1 / α under strict constant-field scaling when the bulk-source voltage is zero but can be made to do so by a threshold adjust implant; this is taken as part of the definition of ideal scaling.

² By definition, V_{DD} and V_T scale as $1 / \alpha$; in an inverter biased at mid-rail, V_{GS} - V_T also scales as $1 / \alpha$.

In sub-micron processes, transistors are not accurately modelled by a first order model. The validity of the scaling model can be extended by considering velocity saturation effects. The current in strong-inversion is described by

$$I_{DS} = \frac{\mu_{EFF} W_{EFF} C_{OX}}{2n L_{EFF}} \frac{(V_{GS} - V_T)^2}{1 + \theta (V_{GS} - V_T)}$$
(6.1)

where velocity saturation effects are described by a parameter θ given by

$$\theta = \frac{\mu_{EFF}}{2n\upsilon_{SAT}L_{EFF}} \tag{6.2}$$

and where v_{sAT} is the carrier saturation velocity (1 × 10⁵ ms⁻¹ for electrons and holes), μ_{EFF} is the effective carrier mobility, W_{EFF} and L_{EFF} are the effective channel width and length respectively, and n is the bulk charge effect factor which is typically close to 1 [212][235]. The effects of finite source resistance, which is also important in sub-micron transistors, can also be incorporated in the expression for θ but is not considered here.

Simple algebraic manipulation of expressions for g_m , I and I / g_m derived from (6.1) and (6.2) shows that, although velocity saturation causes these parameters to be less than they would be in the low-field case, exactly the same scaling factors apply provided the mobility does not change with technology.

6.2.2 Non-ideal scaling

In early CMOS generations, constant-voltage rather than constant-field scaling was applied [236]. However, SIA projections for technology scaling [237] appear to follow constant-field scaling closely in terms of oxide thickness and power supply voltage. Nevertheless, there are reasons why a deviation from the ideal scaling can be expected as the fundamental physical limits of transistor performance are approached. This section considers the influence of two such effects: the non-scaling of the thermal voltage kT / q and the mobility degradation due to the electric field perpendicular to the channel [245].

The influence of these two effects is quantified by calculating I, g_m and I / g_m for different technologies using the basic process parameters of oxide thickness, threshold voltage, power supply voltage and effective channel length together with the current model of (6.1) and (6.2).

Calculations were performed using data for a number of actual processes [238][239][240][241] [242] and using figures from the SIA roadmap³ [237].

The scaling analysis is applied to the transistors in inverter-based gain stages of the simple receivers of the type discussed in Chapter 4. Specifically, it is applied to complementary gain stages, biased at mid-rail, in which the scaling of V_{GS} - V_T is determined by the scaling of the power supply voltage. The advantages of the self-biased complementary-inverter gain stage in smart-pixel receivers has already been highlighted in Chapter 5. Nevertheless, the analysis is equally applicable to an amplifier using a single type of transistor as the transconductance element in which V_{GS} - V_T is made to follow the same scaling as in the inverter by appropriate choice of bias current.

Non-scaling of kT / q

The fact that the thermal voltage kT / q is independent of technology limits the extent to which both the threshold voltage V_T and the gate-source drive voltage V_{GS} - V_T can be made to scale. The origin of these limitations is the subthreshold current of the transistor which, below or close to threshold, depends exponentially on the ratio of V_{GS} - V_T to the thermal voltage [243].

A gate-source drive voltage of at least 200 mV is required for the strong-inversion current, modelled by equation (6.1), to dominate over the subthreshold current [244]. Operation in moderate-inversion, in which both modes of conduction are important, could be considered⁴; however, this analysis assumes that strong-inversion operation is required and thus $V_{GS}-V_{T}$ cannot be scaled below 200 mV. Table 6-2 indicates how ideal scaling is modified by the constraint that $V_{GS}-V_{T}$ is held constant in both the low-lateral-field (no velocity saturation, $\theta (V_{GS}-V_{T}) << 1$) and high-lateral-field (velocity saturated, $\theta (V_{GS}-V_{T}) >> 1$) limits. Transistors in sub-micron processes are more closely approximated by the high-lateral-field limit. Scaling

³ The effective channel length was assumed to be equal to the "isolated line (MPU gate) dimension" from the Roadmap Overall Technology Characteristic Table (Table 1) which is typically 20-30% smaller than the "technology generation" (for example gates are 0.2 μ m in the 0.25 μ m technology generation). The power supply voltage used was the upper limit of the range for analogue supply voltage (Table 15).

⁴ Operation in moderate inversion at lower drive voltages is possible provided a constant current bias is used to achieve temperature insensitive operation (this would preclude the use of a complementary inverter). The consequence would be that the current model used here would no longer apply and, under scaling for constant transconductance, this would allow further improvement in I (but at a slightly slower rate) but a less rapid improvement or even a reduction in the transit frequency.

at constant V_{GS} - V_T will tend to shift the behaviour even closer towards this limit. From the table, the main consequences of the lower limit on V_{GS} - V_T for ideal scaling are an end to the improvement or even a slight degradation of I and I / g_m but a small additional improvement in g_m and f_T .

| parameter | ideal scaling | constant V_{GS} - V_{T} | | |
|---------------------------|---------------|--|------------------------|--|
| | | high-field | low-field | |
| V_{GS} - V_{T} | 1/α | 1 | 1 | |
| g _m | 1 | 1 | α | |
| Ι | 1/α | 1 | α | |
| Р | 1/α | 1 | α | |
| \mathbf{f}_{T} | α | α | α^2 | |
| I / g _m | 1/α | \mathbf{V}_{GS} - \mathbf{V}_{T} | $(V_{GS} - V_{T}) / 2$ | |

Table 6-2: Modifications to ideal-scaling by holding V_{cs} - V_{T} constant

A threshold voltage of at least 0.4 V is required to keep the power-consumption due to subthreshold leakage in low-activity digital transistors at an acceptable level. For this reason, the threshold voltage is expected to remain at this value for technologies below 0.25 μ m [245]. Subthreshold power-consumption is less critical in high activity digital transistors and analogue transistors designed to operate in saturation; some processes now offer special low-V_T transistors for these applications.

The consequence of the fixed threshold voltage is that the value of V_{GS} - V_T in a complementary gain stage is slightly lower than that predicted by ideal scaling. g_m and I are thus slightly lower than ideal scaling would predict and the g_m / I figure of merit is slightly better. However, the generation at which the strong-inversion limit is reached occurs sooner. The fixed threshold voltage obviously has no impact on NMOS-only gain stages because V_{GS} - V_T can be chosen arbitrarily by appropriate biasing.

These two problems together mean that the complementary inverter topology becomes impractical for power supply voltages below about 1.2 V. This is the target analogue supply voltage for the 0.1 μ m generation (0.07 μ m gate). Beyond this generation, it is assumed here that, in order to maintain the drive voltage at 200 mV, either low-V_T transistors are employed, or the power supply voltage is not reduced below 1.2 V. Neither of these assumptions at all modifies the scaling of the transistor parameters but the higher power supply voltage would lead to a proportionately higher power consumption. Neither assumption is necessary for a gain stage based on a single transistor type in which the drive voltage is scaled in the same way.

Mobility degradation

The effective carrier mobility μ_{EFF} is expected to degrade as technology is scaled due to an increase in the electric field perpendicular to the channel. The current equation (6.1) indicates that both g_m and I decrease in proportion to μ_{EFF} when velocity saturation is not present. However, in the high lateral-field (velocity saturated) limit where θ (V_{GS} - V_T) >> 1, the current equation reduces to:

$$I_{DS} = W_{EFF} \upsilon_{SAT} C_{OX} \left(V_{GS} - V_T \right)$$
(6.3)

and thus the current and transconductance are independent of the mobility. The actual influence of mobility degradation will be somewhere between these two extremes. This is quantified using an established model for mobility degradation (Appendix 6.5).

As a result of additional mobility degradation, the calculations indicate that the transconductance of the scaled transistor in 0.1 μ m technology is a fraction 0.6 below the value predicted when mobility degradation is neglected. The g_m / I figure of merit is almost unaffected. The practical consequence of this is that, when scaling at constant transconductance, the transistor has to be made slightly wider than would otherwise be required leading to a less than ideal improvement in capacitance. The degradation is less if a smaller threshold voltage is assumed and so the estimate is conservative.

Results and discussion

The scaling of these parameters is plotted in Figure 6-1 for NMOS transistors. PMOS transistors show essentially the same trends.

Although the two non-ideal effects have a clearly noticeable influence on transistor performance, the influence is sufficiently small that, at least until the 0.1 μ m (0.07 μ m gate) technology generation, they can be considered as corrections to ideal scaling rather than factors that completely invalidate it. Indeed, the deviation from ideal scaling is not that much larger than the scatter about the ideal for actual processes.

The transconductance of the scaled transistor is slightly reduced, reaching about 50% of the value of the reference 0.6 μ m transistor by the 0.05 μ m (0.035 μ m gate) generation. Mobility degradation is the more important effect; the additional mobility degradation introduced between the 0.6 μ m generation and the 0.05 μ m generation accounts for about 75% of the total reduction. The practical consequence of the lower transconductance of the directly scaled transistor is that, in order to maintain the same transconductance, the transistor has to be sized somewhat wider than in the ideal case and hence the gate capacitance is about twice as large.

scaling of transistor parameters



Figure 6-1: Non-ideal effects on scaling of I, g_m and I / g_m . Ideal scaling relative to a reference 0.6 µm technology is shown with a dashed line. Points marked with a diamond are calculated using parameters from the 1997 SIA roadmap. Other points are calculated using parameters from actual processes. Note that all three graphs contain the same number of decades on the vertical scale; the vertical distance from the ideal scaling line is a measure of the factor by which the non-ideal scaling differs from the ideal scaling. Values of I and g_m are for a transistor with $W_{EFF} = L_{EFF}$. Note that $L_{EFF} = 0.7-0.8 \times SIA$ technology generation.

As discussed above, the I / g_m figure of merit improves slightly faster than in the ideal case beyond the first technology generation at which the threshold voltage is held constant but then levels off once the strong-inversion limit is reached so that, by the 0.05 µm generation, the figure becomes about twice as bad as that predicted by ideal scaling. The calculations indicate that mobility degradation by itself has very little effect on this figure-of-merit.

6.2.3 Scaling of offset voltage

This section predicts the future trend in the offset-voltage of the inverter gain stage.

Scaling of threshold voltage mismatch

Historically, A_{vT} has shown a clear linear relationship with technology (Figure 4-18). Extrapolation of this trend predicts that, due to the finite intercept of the graph, the standard deviation in the threshold voltage of a scaled transistor is somewhat larger.

Predictions of the scaling of the mismatch have been made by Mizuno [207] based on a theoretical model of the mismatch. The threshold voltage of a MOS transistor, neglecting the influence of fixed oxide charge, is given by:

$$V_T = \phi + 2\psi_B + \frac{Q_B}{C_{OX}} \tag{6.4}$$

where ϕ is the gate work function, 2 ψ_B is the surface potential at the onset of strong inversion and Q_B is the bulk depletion charge. The work function and the inversion terms are only logarithmically dependent on the doping density; for processes which use a n-type polysilicon gate for NMOS transistors and a p-type polysilicon gate for PMOS transistors (for example, the process described in reference [238]) the terms approximately cancel [234]. Their influence is neglected in the analysis which follows. Q_B is a Poisson random variable with standard deviation proportional to $Q_B^{1/2}$ (assuming that all the charges contributing towards Q_B have the same sign).

In the case of a uniformly doped channel,

$$Q_B = \sqrt{2q\varepsilon_{Si}N_A 2\psi_B} \tag{6.5}$$

and the standard deviation in the threshold voltage is then [207]:

$$\sigma_{VT} \propto \frac{t_{OXIDE}}{\sqrt{WL}} N_A^{-1/4}$$
(6.6)

Under constant field scaling, this expression increases as $\alpha^{1/4}$.

However, as pointed out by Mizuno, this model is overly simplistic because practical MOSFET transistors do not use a uniform channel and because the model ignores the requirement for threshold voltage adjustment; an implant is used to increase the doping concentration at the surface.

From equation (6.4), it can be seen that, in the constant V_T scenario considered here, the total bulk charge must increase in proportion to α . In this scenario, σ_{VT} would increase as $\alpha^{1/2}$ in technologies beyond the first to employ $V_T = 0.4$ V. The predicted value for the threshold voltage mismatch of a transistor with width ten times the minimum length based on this model are shown in Table 6-3 based on the equation

$$\sigma(\Delta V_T) = \sqrt{\frac{2 q t_{OXIDE} V_T}{WL \varepsilon_{OXIDE}}}$$
(6.7)

which can be derived from (6.4).

| technology | oxide thickness | V _T | $A_{_{VT}} / mV\mu m$ | $\sigma \left(\Delta V_{_{T}} \right) / mV$ |
|------------|-----------------|----------------|-----------------------|--|
| 0.7 μm | 17 nm | 0.743 V | 10.8 | 4.9 |
| 0.6 µm | 12.5 nm | 0.8 V | 9.6 | 5.1 |
| 0.25 μm | 4.5 nm | 0.4 V | 4.1 | 5.2 |
| 0.1 µm | 1.75 nm | 0.4 V | 2.5 | 8.1 |
| 0.05 µm | 1 nm | 0.4 V | 1.9 | 12.2 |

Table 6-3: Estimated scaling of threshold mismatch standard deviation for a mediumsized analogue transistor with W=10 L ignoring work function and inversion terms andassuming V_T control is achieved using implantation

Some limited experimental support for the validity of this model is provided by agreement between the prediction for the Alcatel 0.7 μ m technology and the experimental value of 11 mV μ m. Based on this model, the contribution of the threshold voltage mismatch to the offset voltage is predicted to get worse by a factor of between 2 and 3 between 0.6 μ m and 0.05 μ m.

An $\alpha^{1/2}$ trend has also been proposed in [246] based on three-dimensional simulation of transistors.

Note that the mismatch problem is considered to be sufficiently serious to impact on digital circuits and it is therefore possible that technological solutions to the problem may be developed. Possibilities that have been proposed include the use of a retrograde channel [245], which in principle can be used to eliminate threshold voltage mismatch completely, and the accomplishment of threshold voltage adjustment by control of the gate work function [240] instead of an increase in doping concentration. Indeed, other factors make it likely that a

retrograde channel structure will be used below 0.2 μ m [245] which would make this analysis of V_T mismatch overly simplistic. Indeed, the SIA roadmap specifies a target "V_T 3 σ variation (± mV) (For minimum L device)" reducing from 60 mV in 0.25 μ m technology to 40 mV in 0.05 μ m technology, although below 0.1 μ m no currently known technique is capable of meeting the target. If it is reasonably assumed that the width of the transistor to which this entry refers also scales with technology, then, assuming the SIA targets can be met, there would be a small reduction in the offset voltage of a scaled analogue circuit.

Consequently, the discussion in this section should only be interpreted as providing a qualitative indication that the threshold voltage mismatch will not improve dramatically and may get slightly worse.

It can be seen from the expression for the offset voltage (4.29) in terms of $\sigma(\Delta V_T)$ that, since the ratio of g_{mn} and g_{mp} can be expected to stay constant, the contribution of the threshold voltage mismatch term to the photoreceiver offset voltage scales in the same was as $\sigma(\Delta V_T)$.

Scaling of current factor mismatch

Simple algebraic manipulation of (4.29) shows that the contribution to the offset voltage of the current mismatch term is proportional to

$$\sqrt{\sigma^2 \upsilon_{OP,CURRENT}} \propto \frac{V_{GS} - V_T}{2 + \theta \left(V_{GS} - V_T\right)} \frac{A_{\beta}}{\sqrt{WL}}$$
(6.8)

The historical trend is that A_{β} has remained relatively unchanged with technology. Assuming that this trend continues, the contribution to the offset is unchanged under ideal scaling. The current mismatch term is less important than the V_{T} mismatch for the designs examined in the 0.7 µm technology in Chapter 4; given that there are indications that the V_{T} mismatch may get worse, it is reasonable to conclude that the scaling of the offset voltage will be determined primarily by the scaling of the V_{T} mismatch.

6.2.4 Scaling of capacitance

Although the intrinsic gate-source capacitance follows ideal scaling almost by definition, a number of other capacitances are large enough to influence the performance of smart-pixel receiver circuits; this section verifies that none of these capacitances scales any worse than $1 / \alpha$.

Gate-drain overlap capacitance

Empirical data [247] has shown that the gate-drain overlap capacitance per unit width has remained more or less constant down to 0.25µm and thus the gate-drain overlap capacitance

scales as $1 / \alpha$. This trend can be explained based on an analytical model of overlap capacitance presented in [248]. The overlap capacitance per unit width will remain constant in spite of the decrease in oxide thickness provided the lateral diffusion of the drain under the gate (as shown in Figure 4-20) also scales as $1 / \alpha$.

Interconnect and parasitic capacitance

The historical trend has been for the interconnect capacitance per unit length to remain constant [245] which is consistent with equal scaling of the line-width and the inter-layer dielectric thickness [249][250]. Scaling of interconnect width is expected to continue; the SIA roadmap defines the target minimum metal dimension to be equal to the technology generation. The length of the routing required can be expected to reduce as technology shrinks, driven by the requirement to increase logic density. If a linear trend in the length of the interconnect is assumed, the parasitic interconnect capacitance scales in line with the gate capacitance and so its relative importance remains the same. Improvements in interconnect technology, such as low-permittivity dielectrics, introduced to alleviate RC delays in long-distance on-chip interconnects, can be expected to result in an additional reduction in interconnect capacitance.

The drain junction capacitance also makes a significant contribution to the parasitic capacitance. The width of the junction scales as $1 / \alpha$ by definition; the lateral extent must also scale down to keep the series resistance of the source junction low. Under constant field scaling, the junction capacitance per unit area and the sidewall capacitance per unit length can be expected to increase somewhere between $1 / \alpha^{1/2}$ and $1 / \alpha$ based on the standard equation for the capacitance of a *pn* junction [251]. On this basis, a scaling in the junction capacitance between $1 / \alpha$ and $1 / \alpha^{3/2}$ is predicted. This trend is approximately supported by empirical data until the 0.25 µm technology generation where there is evidence of a dramatic drop in the area and sidewall capacitance suggesting a qualitative change in process technology [247]. Other innovations such as the introduction of silicon-on-insulator technology to high volume manufacturing, recently announced by IBM [252], may significantly reduce this component. Although the exact scaling is unclear because of uncertainty about the length of the drain region, this evidence generally indicates that the drain junction capacitance will not increase in relative importance to the gate capacitance and may even decrease slightly.

The overall scaling of load capacitance in the transimpedance circuit is thus not significantly modified by inclusion of parasitics.

6.2.5 Scaling of inverter gain

The dependence of V_{T} on drain bias, or drain-induced barrier-lowering, becomes more pronounced in short-channel transistors and is expected to lead to a reduction in the inverter

129

gain A. This effect must be suppressed, even in digital circuits, in order to obtain a well defined threshold voltage. An estimate of the gain can be obtained from the digital power supply voltage and a typical specification for the V_{T} shift due to short-channel effects in a modern CMOS process. In Intel's 0.25 µm 1.8 V technology, the specification is 0.12 V shift in V_{T} over the entire output voltage swing [253]. Assuming that the output conductance is dominated by drain-induced barrier-lowering, this indicates an average gain of about 15. If the specification remains the same then, in processes designed for a lower supply voltage, the gain will be reduced, maybe by a factor of two by the 0.05 µm generation. However, it is difficult to make accurate predictions without a detailed understanding of the device physics and two scenarios are considered in which A remains unchanged and in which A falls by a factor of 2.

6.3 Impact on receiver performance

6.3.1 General approach

Having established how the basic transistor characteristics can be expected to scale, the implications of this scaling on receiver performance are now considered. The class of receiver circuit investigated is the two-beam receiver discussed in Chapter 4.

The general approach is to start with a reference design in a current technology, for which the transistor parameters are known, then to scale this design according to the rules in Section 6.2.

How this scaled design must be modified in order to achieve a chosen scaling of the bit-rate is then considered. Two different bit-rate scaling scenarios are investigated: scaling the design at constant bit-rate and scaling the design such that the bit-rate increases by α . The increase in operating speed of digital electronics and the simplification of the optical packaging that arises from using as few channels as possible to implement a given overall capacity would both favour the scaled bit-rate scenario.

The analysis assumes the validity of a small-signal model for the front-end and post-amplifier.

A fixed photodiode capacitance is assumed. It is primarily determined by the detector diameter which is currently limited by the capabilities of optomechanical packaging technology and not by device fabrication limits. Any improvement in photodiode capacitance that did occur would give rise to a further reduction in switching energy and front-end power consumption as discussed in Chapter 4.

As a reference circuit, a 500 Mbit/s receiver in 0.6 µm technology is used. The data rate was chosen to be somewhat faster than typical chip-level clock rates in this technology generation, again in the expectation that relatively high data rates will be favoured to minimise the complexity of the optics for a given overall capacity. Implicit in this is the assumption that

130

some form of demultiplexing or asynchronous routing is used. However, the data rate is sufficiently low that a relatively simple front-end can perform well. In the scaled bit-rate scenario, the chosen reference speed scales to 3 Gbit/s by the 0.10 μ m technology generation; this is in line with the SIA target of 2 Gbit/s for high-performance off-chip electrical connections.

Estimated characteristics of the reference design are shown in Table 6-4. These were calculated using the same techniques and parameters used in the investigation of smart-pixel receiver design trade-offs in Section 4.3. A two-stage post-amplifier design was chosen because it offered a higher gain for the selected bandwidth than a single-stage design of comparable power consumption. The offset voltage was estimated using the value of A_{vT} calculated in Table 6-3 for a 0.6 µm technology and allowing 40% extra to account for short-channel effects (consistent with the penalty in the Alcatel-Mietec 0.7 µm process). Current factor mismatch was neglected. A damping factor of greater than 0.7 confirms that the design has an acceptable step-response.

| front-end width | 6 μm |
|--|----------------------------|
| number of stages in post-amplifier | 2 |
| post-amplifier width (per stage) | 3 μm |
| post-amplifier load width | 0.8 μm |
| power consumption | 7 mW |
| front-end bandwidth | 400 MHz |
| post-amplifier bandwidth | 400 MHz |
| estimated bit-rate | 500 Mbit/s |
| post-amplifier small signal gain | 7.9 |
| V _{DECISION} | 800 mV |
| estimated V _{MIN} based on gain limit | 100 mV |
| offset voltage ($\pm 4.7 \sigma$) | $\pm 31 \text{ mV}$ |
| switching energy (adjusted for offset) | 18 fJ peak energy per beam |
| photodiode capacitance | 50 fF per diode |
| damping factor ζ | 0.71 |

| Table 6-4: | Characteristics | of a P | reference | receiver | design | in 0.6 | μm | technol | ogy |
|-------------------|-----------------|--------|-----------|----------|--------|--------|----|---------|--------------|
| | | | | | | | | | - O V |

6.3.2 Scaling of front-end dimensions

The maximum bit-rate for which the front-end can achieve a low switching energy is determined by parameter B_a , which was defined in Chapter 4 as:

$$B_0 = \frac{g_m}{3(C_{IN} + C_L)}$$
(6.9)

Since the fixed detector capacitance typically dominates the overall capacitance in this expression, B_0 is largely determined by the transconductance of the front-end. In the constant bit-rate scenario, the front-end must be scaled for constant transconductance; to a first approximation, this corresponds to the directly scaled design. In the scaled bit-rate scenario, B_0

must be increased by a factor of α ; this can be accomplished by leaving the width of the frontend transistors unchanged instead of scaling them by $1/\alpha$. In both cases, non-ideal scaling effects require a proportionately wider transistor.

6.3.3 Trends in switching energy

The switching is determined by equation (4.7) which is repeated for ease of reference:

$$E_{OPT} = \frac{3}{2} \left(C_F + \frac{C_{IN}}{A+1} \right) \frac{1}{1 - B / B_0} \frac{V_{MIN}}{S}$$
(6.10)

This expression has a factor determined by the front-end and a factor determined by the postamplifier and decision stage.

Front-end

The factor determined by the front-end is the effective input capacitance $C_F + C_{IN} / (A+1)$. In the reference design, the feedback capacitance accounts for about 40% of the total effective capacitance; faster designs in the same technology have a proportionately higher contribution from C_F .

The contribution from the larger term, $C_{IN} / (A+1)$, will not improve with technology in either bit-rate scenario because the input capacitance will remain dominated by the detector. Depending on the extent to which short-channel effects result in a reduction in gain, this term could remain unchanged or get worse by a factor of two.

In the constant bit-rate scaling scenario, C_F decreases due to the reduction in gate-source overlap capacitance and the general reduction in parasitic capacitance in the scaled front-end; the relative importance of the C_F in the receiver performance will therefore decline. However, even if C_F was completely eliminated, there would only be a modest improvement in the switching energy.

In the scaled bit-rate scenario, only the routing capacitance contribution to C_{F} will reduce. This only accounts for about 15% of the total effective capacitance. The additional transistor width required to compensate for non-ideal scaling effects will increase the contribution from the gate-drain overlap capacitance; at worst, this produces a 40% increase in switching energy.

Decision stage and post-amplifier

The factor determined by the decision stage and post-amplifier is the minimum signal required at the output of the front-end $V_{_{MIN}}$.

In the absence of DC offsets, this is in turn determined by the gain of the post-amplifier and V_{decision} .

The directly scaled post-amplifier circuit has the same gain (set by a ratio of transistor widths in a gain-broadened design) and a bandwidth that is a factor of α higher. Thus the directly scaled design corresponds to the scaled bit-rate scenario. The reduction in f_{T} arising from mobility degradation requires that the gain be slightly reduced in order to maintain the same improvement in bandwidth.

In the constant bit-rate scenario, the scaled post-amplifier has more than enough bandwidth to pass the signal; to a limited extent, the additional bandwidth may be traded for increased gain. This may be achieved by adjusting the size of the post-amplifier load transistor or even removing the load transistors altogether⁵. Neither of these changes have a significant impact on the scaling of the power consumption. The increase in gain is limited by the unloaded gain of an inverter.

 V_{DECISION} can be expected to scale down due to the reduction in power supply. It was previously defined as the voltage width of the transition region of the inverter transfer characteristic. Using a first order transistor model, the extent of this region, as defined by the unity slope points, predicts a value of $V_{\text{DD}} / 4 - V_{\text{T}} / 2$ for a symmetrical inverter [254]; similar approximations have been used to predict the scaling by other authors [168]. The estimate of V_{DECISION} obtained using this simple formula with a 1.2 V supply and a 0.4 V threshold voltage (corresponding to the 0.1 µm technology generation) gives a value of 100 mV for V_{DECISION} .

However, the non-scaling of the offset voltage severely limits the extent to which this benefit can be exploited. Although the value of V_{DECISION} and the post-amplifier gain indicate that a small signal of about 10 mV centred around the switching point of the post-amplifier could be amplified to give a logic signal, it is evident that it is not possible to amplify signals that are smaller than the offset voltage without also filtering the DC component of the signal.

The limit on switching energy is then largely determined by how the offset voltage scales. In the constant bit-rate scenario, the offset voltage scales according to the discussion in Section 6.2.3. In the scaled bit-rate scenario, the proportionately wider front-end will tend to slightly reduce the offset voltage; however, the dominant contribution will still come from the smaller post-amplifier input transistor and at most this effect gives a 20% reduction in the offset compared to the directly scaled design.

Unfortunately, the exact scaling of the offset voltage is one of the more uncertain quantities in this analysis. If it is optimistically assumed that it remains unchanged at about \pm 30 mV then

 $^{^{5}}$ It is interesting to note that Lucent 0.35 μ m receivers do not use gain broadening

there is a lower limit on V_{MIN} of about 70-100 mV. At this level, there is little benefit in using a post-amplifier at all; indeed, it would be difficult to keep the post-amplifier within its region of linear operation with such a combination of low supply voltage, non-scaled threshold voltage and high offset voltage. Eliminating the post-amplifier stage or significantly reducing its gain would allow the front-end bandwidth to be reduced slightly; this might give a further small improvement in switching energy. Taking V_{MIN} as 85 mV, assuming a front-end bandwidth just enough to pass the signal and assuming that the inverter gain does not get any worse gives a switching energy of about 5 fJ peak optical energy per beam. Short-channel effects on the gain might degrade this to 10 fJ. This is better than the 0.6 µm value of 18 fJ but still somewhat short of the noise limit. If the DC offset could be removed, it would be possible to amplify signals close to the thermal noise limit predicted by (4.16) using these simple post-amplifier structures.

6.3.4 Trends in power consumption

One area in which a significant improvement can be anticipated is in power consumption. In both scaling scenarios, the reduction in power supply voltage contributes to this. In the fixed bit-rate case, there is also a reduction of $1/\alpha$ in the supply current. In the scaled bit-rate case, the front-end current remains unchanged, but the post-amplifier current reduces in line with $1/\alpha$; the overall power consumption then becomes dominated by the front-end.

A useful metric of the power consumption is the power consumption per terabit/s aggregate data input capacity. In the constant bit-rate scenario, the bit-rate remains unchanged but the supply current drops by α ; in the scaled bit-rate scenario, the bit rate increases by the α but the front-end supply current remains unchanged. Hence, to a first approximation this metric is independent of which bit-rate scaling scenario applies. However, since the post-amplifier power consumption in the reference design accounts for about half the total power consumption, the metric may be as much as a factor of two better in the scaled bit-rate scenario.

For the reference design, the metric is 14 W/(Tbit/s). A 4.2× reduction comes directly from the scaling of the supply voltage from 5 V to 1.2 V. Up to the 0.10 μ m technology generation, there is an improvement of 6× in the supply current or channel bit-rate due to ideal-scaling with a further 1.6× benefit in supply current coming from the corrections to the first order scaling model. This improves the metric to 0.3 W/(Tbit/s) which is very small compared with the estimated power consumption of typical digital chips in this technology generation. However, a further improvement in this metric beyond 0.10 μ m will only be obtained if the power supply voltage can be successfully reduced below 1.2 V because of the end to the improvement in I/g_m.

134
Even if this estimate is optimistic, it would appear that the electrical power dissipation in large receiver arrays in technology that is actually capable of exploiting the bandwidth is not a major concern, in sharp contrast to the experimental test bed systems implemented in today's silicon technology.

This estimated value of this metric is more or less consistent with some of the designs investigated by Van Blerkom [168] where for example, a 3 Gbit/s receiver in 0.1 μ m technology is estimated to have a power consumption of 600 μ W giving a metric of 0.2 W / (Tbit/s).

6.3.5 Changes to the scaled design to exploit the reduced power consumption

An important implication of the predicted dramatic fall in power consumption of a basic receiver design in future technology is that there may be more scope for improving the sensitivity or dynamic range by using more complex receiver circuits. In this section, several such options are discussed.

Constant offset scaling

Because of the increasing importance of offset voltage, it is valuable to consider an alternative scaling scenario in which the transistors are scaled for constant offset voltage. In the worst-case offset voltage scenario, where the offset voltage scales as $\alpha^{1/2}$, this could be achieved by holding the width of the post-amplifier input transistor constant and by adopting the scaled bit-rate scenario for the front-end. The supply current per channel then remains more or less unchanged; however, the power consumption per terabit/s metric still improves because of the increase in channel rate. This is an additional reason to favour the scaled bit-rate scenario.

In the optimistic scenario for offset voltage scaling where it remains unchanged for the directly scaled design, a constant width scaling would lead to an improvement in offset voltage of $\alpha^{1/2}$ and a commensurate improvement in switching energy.

Increase post-amplifier gain

In principle, extra power consumption allows the use of additional gain stages in the postamplifier to be considered. We have already seen that increasing the post-amplifier gain is not useful unless a lower frequency cut-off is implemented to compensate for the offset voltage. However, in Section 4.6.5 it was shown that the area requirement of the capacitor required to implement a lower frequency cut-off becomes compatible with smart-pixel requirements in future technologies.

135

Amplification of high-frequency signals that are smaller than the offset voltage in low-voltage circuits requires some care. If the filtering is simply implemented at the input to the next stage, then the offset (which is larger than the signal of interest in this scenario) will also be amplified by the post-amplifier and may shift the signal outside the common-mode input range of the following stage. Filtering prior to the post-amplifier does not eliminate the offset-voltage of the post-amplifier itself. The solution is to implement the low-frequency cut-off by using negative feedback to subtract the low-frequency component of the post-amplifier output signal from the current flowing into the post-amplifier load.

A possible implementation is shown in Figure 6-2. The output of the post-amplifier Mn2/Mp2 is filtered by the RC network and converted to a current signal using Mn3/Mp3. The current signal is fed back into the post-amplifier transistors; this effectively shifts the operating point of Mn2/Mp2 to compensate for the offset voltage between the first and second stages.

Another way to think about this circuit is that for high frequency signals, the post-amplifier looks like an unloaded inverter with a fairly high gain, but for low frequency signals, the post-amplifier looks like a gain-broadened amplifier with gain determined by the ratio of the transconductance of Mn2/Mp2 to Mn3/Mp3 which can be designed to be say 1 so that the offset is not amplified. This gain cannot be made very small because the output conductance of Mn3/Mp3 will load down the gain for signals within the pass-band. If the low-frequency gain is say 1, then the high-frequency gain will be half the gain of the unloaded inverter.

Note that because this design uses the capacitor in a feedback loop, the lower frequency cut-off is higher than 1/RC by the factor by which the input-referred offset is attenuated, requiring a proportionately larger capacitor to achieve a given cut-off frequency.



Figure 6-2: A post-amplifier with a filter to attenuate the DC offset

In the scaled bit-rate scenario, the power-consumption is dominated by the front-end and so a modest increase in the power consumption of the post-amplifier need not result in a large increase in the overall power-consumption of the receiver.

Methods for increasing the gain of the post-amplifier at some cost in power consumption have already been discussed in Section 4.5. Such modifications together with the implementation of a low-frequency cut-off should, in principle, allow noise limited performance to be obtained.

Increase front-end gain

There is some limited scope for increasing the front-end gain, and hence improving the switching energy, as a consequence of the fact that the front-end transconductance scales in proportion to the bit-rate but the front-end load capacitance decreases in proportion to α . There is thus an overall improvement of α in the ratio of the open-loop to the closed-loop bandwidth and hence the maximum acceptable gain. However, under constant-offset scaling, the ratio does not improve.

The increase in gain could be achieved, for example, by using non-minimum length transistors in the front-end.

Ease optoelectronic packaging constraints

Since the power consumption, switching energy and speed of the scaled design are easily good enough to implement a terabit/s scale optical interface, it may be worthwhile to use some of the extra power consumption to accommodate larger detectors. This would ease the optomechanical packaging problem, which is by far the greatest obstacle to the commercial feasibility of smart-pixel systems.

The detector capacitance scales as the square of the diameter, so a factor of four in front-end power consumption could be traded for an increase in detector diameter from about 25 μ m to 50 μ m with the additional penalty of a 4 × increase in optical switching energy.

6.3.6 Limit on scaling due to power supply distribution

The SPOEC system demonstrated that one of the factors limiting the performance in a large array of receivers is the DC voltage along the power supply rails. In this section, we briefly analyse the extent to which this will limit modifications to the scaled design along the lines discussed in the previous section.

Consider a square array with a regular pitch. Let there be N channels in the array. Assume that the power supply rails are fed separately from the top and the bottom of the array and the analogue power and ground rails together occupy k entire layers where k might be around 0.7 if a dedicated distribution layer is used. If k remains fixed, then the number of squares of resistance is determined by the number of receivers in a column and not by the physical dimension of the array. Let R_{super} be the sheet resistance of the power distribution layer and let

the current per receiver be I. One can easily show that the maximum number of receivers N_{MAX} that can be used while maintaining a DC voltage drop of less than ΔV is given by:

$$N_{MAX} = 2 \frac{k}{R_{SHEET}} \frac{\Delta V}{I}$$
(6.11)

If we take as a reference $\Delta V = 500 \text{ mV}$ (for a 5 V supply), k = 0.7, R_{SHEET} = 0.04 Ω / square and I = 1.4 mA (based on the reference receiver design) then N_{MAX} is 12 500. This gives an upper bound on the aggregate bandwidth of 6 Tbit/s which is independent of the channel rate to the extent that the supply current is proportional to the channel rate. In practice, the maximum bandwidth would be perhaps a factor of two lower than this in order to allow for process tolerance in the supply current and sheet resistance.

Assuming that the relative tolerance on the power-supply voltage remains the same in more advanced technology, the constant bit-rate scaling scenario leaves $N_{_{MAX}}$ unchanged whilst the scaled bit-rate scenario causes a reduction in $N_{_{MAX}}$ with a scaling that is slightly slower than $1 / \alpha$. This means that there is little improvement in this upper limit on the aggregate bandwidth.

However, this limit is not fundamental. The sheet resistance can be reduced by using thicker layers of metal – Intel already employ 1.90 μ m thick metal in the top level of their 0.25 μ m technology [255]. A shift to copper interconnect will reduce the metal resistivity from 0.028 Ω μ m to 0.017 Ω μ m. These changes alone give a 4× reduction in sheet resistance over the value used above. It is also possible to use a two layer gridded power distribution scheme which gives a further factor of four benefit.

Thus, for aggregate bandwidths of the order of several terabit/s, it does not appear that this is a major limit on the performance of large receiver arrays, provided that a dedicated power distribution layer is used. There is plenty of margin for considering the modifications discussed in the previous section.

6.3.7 Limitations of this analysis

This analysis is intended only to provide a general guide to trends in receiver performance and has been based on a very simple small signal analysis. There are a number of specific limitations that must be borne in mind in applying the results.

The analysis is based on the complementary inverter gain stage; the feasibility of this structure with very low supply-voltages (especially below 1.2 V) is questionable because of the very small input-voltage range over which linear operation is obtained. It is acceptable for the frontend (which is self-biasing and has a low voltage swing at the input) but it is not obvious that it will offer acceptable dynamic range in the later stages of the receiver. The increased dynamic range of differential amplifiers may make them more attractive. The analysis will not successfully predict how the performance of, say, a 0.6 μ m differential amplifier design will scale – in current technology, differential amplifiers typically use gate-source drive voltages in the range 200 - 400 mV which is much smaller than that in an inverter which is biased at midrail. However, in low-voltage technologies, in which the inverter drive voltage scales to within this range, the values of g_m and I / g_m will be comparable for the two topologies. Thus the predictions of performance can be expected to be approximately correct for a differential amplifier structure, even though they were obtained by considering a scaled inverter design.

The analysis has not considered power-supply crosstalk. The discussion of this issue in Chapter 8 suggests that it may be an important limiting factor on receiver performance in large arrays and it may prevent thermal-noise limited performance from being obtained even if a low-frequency cut-off can be implemented.

A third limitation of this analysis is that it has not looked at the scaling of the thermal noise performance, in particular how the parameter Γ scales. There is some work on this topic in the literature (see the references in Section 4.4).

The argument that noise limited performance is attainable depends on the feasibility of correctly biasing the analogue circuitry in the linear region, which has only been examined superficially in the case of the complementary inverter topology and not at all in other structures such as differential amplifiers.

The general problem of receiver design with very low supply voltages is a subject worthy of more detailed investigation. In principle, it is possible to investigate this in current technology; however, the higher parasitic capacitance and limited f_{T} at low drive voltages would make it difficult to design high-speed circuits.

The uncertain information about the scaling of the threshold voltage mismatch also means that the estimate of the switching energy must be treated with caution.

6.3.8 Comparison with other studies

Krishnamoorthy and Miller [167] project scaling to 1fJ peak optical energy per beam by the $0.10 \mu m$ technology generation based on a data rate of 1 Gbit/s. This is close to their analysis of the noise limit of 0.4 fJ. Although the approach of their analysis is slightly different from here, and some discrepancy is therefore expected, their estimate nevertheless seems to be optimistic compared to the 5-10 fJ limit estimated here. This is in part due to the assumption that the photodiode capacitance scales down to 20 fF and the omission of offsets from their analysis.

6.4 Conclusions

This section has examined the scaling of small signal transistor and interconnect parameters in advanced CMOS technology including some corrections to the standard first-order theory. These parameters have been used to analyse the scaling of the performance of a common smart-pixel transimpedance receiver circuit. A dramatic reduction in electrical power consumption per terabit/s to about 0.3 W/Tbit/s by the 0.1 µm technology generation is predicted but beyond this point, little further improvement is expected.

A possible realisation of a 1 Tbit/s interface in 0.1 µm technology might be 256 channels running at 4 Gbit/s. A scaled DC coupled two-beam receiver design operating close to this speed is projected to have a switching energy of between 5 fJ and 10 fJ per beam limited by transistor offset which is only a modest improvement over 18 fJ for a 500 Mbit/s receiver in 0.6 µm technology.

These improvements are possible without any reduction in photodiode diameter from a typical current value of $25 \ \mu m$.

Designs with a low-frequency cut-off seem at first sight to be feasible and in principle allow thermal noise limited performance to be obtained. Although there is a case for examining receiver designs with a low-frequency cut-off as a means to improve switching energy, it is by no means a necessity that receivers with a response down to DC be abandoned; a switching energy of the order of 5 fJ per beam is perfectly compatible with very high data-rate point-topoint optical links over short distances. For example, a two-beam optical link, implemented using VCSELs with a projected performance of 500 μ W output power and a data rate of 2 Gbit/s [256], has a peak energy per beam of 250 fJ.

The switching energy corresponds to a total optical power per chip after losses of only 5 mW for high contrast data which is easily compatible with the capabilities of optoelectronic device technology. The number of channels required to implement a 1 Tbit/s interface is not at all aggressive and could conceivably be accomplished with either free-space or fibre based interconnect systems.

None of the receiver design issues examined in this chapter appear to indicate that implementing aggregate data rates much higher than a terabit/s using large receiver arrays is difficult. Because transistor performance supports the scaled bit-rate scenario without any problems, it seems likely that the capability of free-space optics to offer several thousand optical inputs can only be fully exploited in interfaces with a capacity much larger than 1 Tbit/s.

6.5 Appendix: Mobility degradation model

This appendix describes the mobility model used to study mobility degradation due to the electrical field perpendicular to the channel based on a description in [212]. References to the original work on which the model is based can be found therein.

The model is an empirical one designed to take into account a variety of scattering mechanisms.

$$\mu_{EFF} = \frac{\mu_0}{1 + \left(\frac{E_{EFF}}{E_0}\right)^{\nu}}$$
(6.12)

 E_0 , μ_0 and ν are constants (Table 6-5). E_{EFF} is a so-called effective electric field which represents the average electric field experienced by carriers in the channel which is process dependent and can be empirically described by:

$$E_{EFF} = \frac{V_{GS} + V_T}{6T_{OX}} \tag{6.13}$$

A plot of this function matches the experimental data presented in [245] where the problem of mobility degradation is discussed.

| parameter | electron | hole |
|-----------------------|----------|------|
| $\mu_0 / cm^2 V^{-1}$ | 670 | 160 |
| $E_0 / MVcm^{-1}$ | 0.67 | 0.7 |
| ν | 1.6 | 1.0 |

 Table 6-5: Parameters for mobility degradation model

Chapter 7

Transconductance-transimpedance post-amplifiers for smart-pixel receivers

7.1 Introduction

In previous chapters, it has been shown that the post-amplifier can be the component that limits the overall performance of a smart-pixel receiver circuit. This is especially true in a circuit that approaches the speed limit of a particular technology; in such cases, the gain that can be obtained from a simple, single-stage voltage amplifier is limited by the gain-bandwidth product of the technology. Whilst operation at such speeds is still possible using a voltage-gain amplifier, it can only be achieved at the cost of a switching energy greater than the DC offset limit.

Nevertheless, in a systems context, there is an arguable advantage in operating the optical channels at as high a data rate as is permitted by the electronics. For a given aggregate data rate, this minimises the number of optical channels, and hence the complexity and cost of the optomechanical packaging of the chip.

In this chapter, the application of the transconductance-transimpedance circuit technique to smart-pixel post-amplifiers is introduced as a method for increasing the gain-bandwidth product of the post-amplifier and thus allowing lower switching energies at speeds closer to the technology limit.

The transconductance-transimpedance circuit technique, originally due to Cherry and Hooper [257], is commonly used in wideband amplifier design. It has been used, for example, in wideband operational amplifiers [258], RF front-ends [259] and hard-disk read-head preamplifiers [260]. In the context of stand-alone optical receiver circuits, it has been used in post-amplifier design in both bipolar [261] and MOS [262][263][264] technology and also in front-ends [262][265]. The same principle has also been used to increase the speed of a cascade of digital inverters [266].

The significance of the work described in this chapter is that it demonstrates that implementations of the circuit technique exist that have performance characteristics which are compatible with the special requirements of smart-pixel circuits and that such implementations have potential to offer better performance than conventional wideband post-amplifier designs for smart-pixel applications based on low-gain voltage stages. The structure of the chapter is as follows. After a brief description of the general principle of the transconductance-transimpedance circuit technique, a simple implementation of the approach suitable for use in smart-pixel post-amplifiers is presented. The advantage of the circuit technique over the gain-broadened voltage amplifier designs described in Chapter 4 is demonstrated in small-signal terms. A more thorough comparison with the gain-broadened design is made by optimising circuits with the same specification using both circuit approaches and comparing their performance using large-signal transient simulations. The chapter concludes by describing the application of the circuit technique to the differential clock receiver in the SPOEC system and describes preliminary experimental results.

7.2 Description of circuit technique

The transconductance-transimpedance circuit technique uses a cascade of a transconductance stage and a transimpedance stage. The signal is converted from a voltage at the input to a current at the intermediate node and back to a voltage at the output node of the circuit. This is shown in Figure 7-1(a).



(a) transresistance load

(b) resistive load



The advantage of this technique at high frequencies can be explained qualitatively in terms of the impedance levels at the nodes of the circuit.

First, consider an unloaded CMOS inverter. The small-signal output impedance of the inverter is relatively high. When used to drive a high-impedance load such as the gate of a second inverter, a low-frequency pole occurs. Only if the inverter is used to drive a low-impedance load can high bandwidth operation be obtained. The gain-broadened amplifier described in Chapter 4 meets this condition using a resistive load formed by diode-connected transistors: the output impedance of the inverter is high compared with the load impedance and the CMOS inverter can be considered as a transconductance element (Figure 7-1 (b)). The condition sets an upper limit on the load resistance and hence the voltage gain of the stage.

However, the same impedance condition can be met at the output of the first stage but with a larger resistor by connecting the load resistor as a shunt-feedback element around a voltage gain stage to form a transimpedance amplifier (Figure 7-1 (a)). The feedback reduces the input-

impedance by the loop-gain of the transimpedance stage; a higher feedback resistor can be used for the same bandwidth. This is directly analogous to the use of a transimpedance front-end in preference to a low-impedance front-end to provide lower thermal noise (larger resistor) for the same bandwidth.

The maximum voltage gain that can be obtained from a single stage at low speeds is also higher because a larger load resistor can be used before the finite output impedance of the transconductance element starts to limit the gain.

7.3 Small-signal analysis of the transconductance-transimpedance circuit

This argument is now formalised using a small-signal analysis of a simple CMOS implementation of the transconductance-transimpedance concept.

Figure 7-2 shows the circuit. This can be directly coupled to a front-end of the form studied in Chapter 4. The transconductance element and transimpedance gain element are based on a complementary inverter.

Implementing a suitable feedback resistor in a standard digital CMOS process is not straightforward. The voltage swings at the output of the transimpedance stage in this circuit are large signals; unlike in the front-end, a simple ohmic region MOSFET is unsuitable because of the strong non-linear increase in resistance with drain-source voltage which reduces the speed of operation of the circuit. A well resistor provides sufficient linearity for this application but occupies significantly more layout area; nevertheless, the sheet resistance and junction capacitance are compatible with reasonably high-speed circuits; the use of a well resistor is considered in Section 7.4. A high-resistance polysilicon resistor process option would be attractive for implementing this circuit.



Figure 7-2: Simple CMOS implementation of the transconductance-transimpedance cascade

A small-signal model of this circuit is shown in Figure 7-3. The decision stage has been replaced with an equivalent load capacitance.



Figure 7-3: Small-signal model of the transconductance-transimpedance cascade

Simple nodal analysis of this circuit gives a transfer function of the form:

$$H(s) = A_0 \frac{1 - s\tau_Z}{1 + s\tau + s^2 \frac{1}{\omega_0^2}}$$
(7.1)

where the zero is at a high frequency (g_m / C_F) and does not affect circuit performance. The full expressions for A_0 , τ and ω_0 are:

$$A_{0} = g_{m1} \frac{R_{F} - \frac{1}{g_{m2}}}{1 + \frac{g_{ds1}}{g_{m2}} + \frac{g_{ds2}}{g_{m2}}}$$
(7.2)

$$\tau = \frac{C_x + C_L + R_F (C_L g_{ds1} + C_F (g_{ds1} + g_{ds2} + g_{m2}) + C_x g_{ds2})}{g_{m2} (1 + \frac{g_{ds1}}{g_{m2}} + \frac{g_{ds2}}{g_{m2}} + R_F \frac{g_{ds2}}{g_{m2}} g_{ds1})}$$
(7.3)

and

$$\frac{1}{\omega_0^2} = \frac{R_F (C_L C_X + C_L C_F + C_F C_X)}{g_{m2} (1 + \frac{g_{ds1}}{g_{m2}} + \frac{g_{ds2}}{g_{m2}} + R_F \frac{g_{ds2}}{g_{m2}} g_{ds1})}$$
(7.4)

As in the analysis of the front-end, these expressions are simplified by assuming that the unloaded gain of an inverter is much larger than 1 so that $g_{m2} >> g_{ds1}$ and $g_{m2} >> g_{ds2}$, and that the feedback resistor R_F is sufficiently high that it does not load down the transimpedance stage so that $g_{m2} R_F >> 1$. It is further assumed that the low frequency output impedance of the transconductance stage is much greater than the input impedance of the transimpedance stage so that the circuit operates in the current mode at the internal node. This corresponds to the assumption $1/g_{ds1} >> R_F / A_2$. For the purposes of factoring the expression for ω_0 only, it is assumed somewhat more crudely that $C_X >> C_F$ which is equivalent to assuming that the gate-source capacitance of a transistor is much larger than the gate-drain overlap capacitance.

The expressions for A_0 , τ and ω_0 then reduce to:

$$A_0 = g_{m1} R_F \tag{7.5}$$

$$\tau = \frac{C_X + C_L}{g_{m2}} + C_F R_F + \frac{C_X}{A_2} R_F + \frac{C_L}{A_1} \frac{g_{m1}}{g_{m2}} R_F$$
(7.6)

and

$$\frac{1}{\omega_0^2} = \frac{R_F C_X (C_F + C_L)}{g_{m2}}$$
(7.7)

The 3 dB bandwidth is estimated using the dominant-pole approximation in which $\omega_{_{3dB}} = 1 / \tau$. Provided there is no peaking in the amplitude response, this is a conservative estimate of the 3dB bandwidth; in the extreme case of a maximally flat response, it underestimates $\omega_{_{3dB}}$ by a factor $\sqrt{2}$.

The first term in the expression for τ is close to the unity gain frequency of the device. Since the circuit technique is intended for high-speed applications, this term will be large enough to affect the performance; however, to a first approximation it is assumed that τ is dominated by the terms proportional to R_F. The gain-bandwidth product is then:

$$GBW = \frac{g_{m1}}{C_F + \frac{C_X}{A_2} + \frac{C_L}{A_1} \frac{g_{m1}}{g_{m2}}}$$
(7.8)

In comparison, the result for the gain-broadened design from Chapter 4 in the notation used in this chapter is:

$$GBW = \frac{g_{m1}}{C_L} \tag{7.9}$$

This demonstrates the gain-bandwidth advantage of the transconductance-transimpedance cascade over a single-stage voltage-gain amplifier: assuming all the inverters are the same size, the capacitance in the denominator in equation (7.8) is smaller than in equation (7.9). The feedback capacitance C_F is usually smaller than the gate-source capacitance; the contribution of the gate capacitances, C_x and C_L to the total capacitance in the denominator is reduced by the unloaded voltage gain A_1 and A_2 of the inverters. The advantage is significant but not dramatic; for example, C_{GS} is about 3 C_F for a minimum length transistor in the 0.6 µm technology examined in Chapter 4.

Although the GBW of the transconductance-transimpedance cascade is higher than a voltagegain stage, an important difference between the two circuit topologies is that, in the case of the transconductance-transimpedance cascade, the second-order nature of the transfer function implies that there is an upper limit on the frequency to which the trade-off may be applied by varying the feedback resistor. To obtain an acceptably damped small-signal transient response, a minimum damping factor ζ of $1 / \sqrt{2}$ is required. This corresponds to a maximally flat amplitude response. Considering only the terms in (7.6) proportional to R_F , the maximum value of ω_0 is

$$\omega_{MAX} = \frac{C_F + \frac{C_L}{A_1} \frac{g_{m1}}{g_{m2}} + \frac{C_X}{A_2}}{\sqrt{2}(C_F + C_L)} \frac{g_{m2}}{C_X}$$
(7.10)

and at this frequency, the 3 dB bandwidth is equal to ω_{MAX} . The second factor, $\omega_{T} = g_{m2} / C_{X}$ is the unity gain frequency of the process. To give a feel for the value of this expression, assume identically sized stages, take $C_{X}=C_{L}=3C_{F}$, and assume a large gain; equation (7.10) indicates that ω_{MAX} is about 0.2 ω_{T} . The limit is slightly higher when the gain is finite. For these parameters, the benefit in gain-bandwidth of the transconductance-transimpedance design, given by the ratio of equation (7.8) to equation (7.9), is a factor of three.

In principle, it is possible to increase ω_{MAX} at the expense of k by reducing the gain A_2 , for example by including broadening transistors in the voltage gain stage of the transimpedance amplifier. However, for signals that are large enough to produce a degree of limiting in the

transimpedance stage, the requirement for low overshoot is less important. The implementation described in Section 7.6 uses a diode clamp in parallel with the feedback resistor to enhance this effect.

An additional advantage of the transconductance-transimpedance approach is that the high voltage-gain output node is isolated from the input node due to the current-mode signal representation at the internal node. As in a cascoded voltage gain stage, the gate-drain capacitance of the post-amplifier input transistor is not multiplied by the gain of the amplifier resulting in less loading of the transimpedance input stage which can help to reduce overshoot.

In summary, this section has shown that, based on a small-signal analysis, the transconductance-transimpedance cascade can, in principle, give a small (up to $3\times$) benefit in gain-bandwidth.

7.4 Detailed comparison with low-gain voltage amplifier

7.4.1 Introduction

The analysis in the previous section has shown that there should be a theoretical advantage from the transconductance-transimpedance circuit topology. However, there are a number of reasons why it might be difficult to translate this theoretical advantage into practical benefits.

The most important is the difference in process sensitivity of the two approaches.

The gain-broadened voltage amplifier has a well controlled gain determined by a ratio of transistor widths; its bandwidth tracks the process unity gain frequency and hence the speed of operation of the other circuits on the chip.

In contrast, in a transconductance-transimpedance amplifier, the tolerance in the gain and the bandwidth depend on how closely the transconductance and the feedback resistance track. If they are realised with different types of physical device, then they will not track and a large process spread in gain and bandwidth will occur. The use of MOS transistors for the feedback resistor can give better control; the gain is then insensitive to certain parameter variations such as oxide thickness. However, if the feedback transistor is biased with a fixed voltage, then the resistance will remain sensitive to the poorly controlled operating-point voltage of the transimpedance stage. More complex biasing circuits can adaptively bias a feedback transistor to track the transconductance [267] or, alternatively, can track the transconductance to the sheet resistance of the device type used to implement the feedback resistor; however, these capabilities are not offered by the simple implementation of Figure 7-2. In addition, any MOS implementation of the resistor must still overcome the non-linearity problem.

There are also a number of tricks that can be used to improve the gain-bandwidth of postamplifiers based on voltage-gain stages. For example, the transconductance transistors in a stage can be sized larger than the input transistors of the subsequent stage. This is often called "inverse-tapering". Also, provided the frequency of operation is not too close to the unity gain frequency, a significant improvement in overall gain-bandwidth can be obtained by cascading multiple low-gain voltage-amplifier stages. If inverse-tapering is also applied to this cascade, then a further improvement in gain-bandwidth can be obtained.

This section investigates the extent to which the theoretical gain-bandwidth advantage can be translated into practical benefits by optimising designs based on both circuit topologies for a particular specification and comparing their performance.

A summary of the analysis in this section has been published in [268].

7.4.2 Methodology

The study is based on a two-beam receiver specified to operate at a speed of 1 Gbit/s in the 5 V digital 0.6 µm technology described in Chapter 4. A speed that was relatively high for this technology was chosen in the expectation that the transconductance-transimpedance cascade would demonstrate an advantage at high speeds of operation. The circuit was designed to operate at this speed under worst-case variations in process parameters. A well resistor was used for the feedback element.

Three different circuit topologies were considered for the post-amplifier circuit: voltage-gain circuits with both one and two linear stages and a transconductance-transimpedance cascade.

A simple front-end, common to each circuit, was employed (Figure 7-4). The front-end gain transistors and feedback resistor were sized to support 1 Gbit/s operation over all process corners. A capacitance of 53 fF per diode [269] was assumed. For simplicity, a design using the positive power supply as a fixed bias voltage for the front-end feedback transistor was used. This is not optimal. However, achieving the best possible switching energy from the front-end is not particularly important in comparing the post-amplifier designs, provided a common front-end is used.



Figure 7-4: Front-end used for comparison of voltage-gain and transconductancetransimpedance designs

The decision stages were loaded with identical three-stage digital buffers that had been tapered to drive a MQW modulator pair to form a simple optical repeater circuit as in reference [270]. The inverter widths (NMOS/PMOS) were $2.8 \mu m / 5.6 \mu m$, $5.6 \mu m / 11.2 \mu m$ and $12 \mu m / 18 \mu m$; all lengths were 0.6 μm .

An objective figure-of-merit was required to compare the designs. Since large-signal nonlinearity is one of the practical difficulties in implementing the transconductancetransimpedance approach, the small-signal bandwidth is not an adequate metric for comparing the circuits. Instead, the designs were compared using large-signal transient simulations. The figure-of-merit used was obtained by analysing eye diagrams at the modulator output. The eyeopening was defined as the interval within the eye during which the channel could be guaranteed to have a valid logic level. The logic thresholds were defined at 20% and 80% of the full 5 V output swing.

For each basic circuit topology, the transistor dimensions were manually optimised to achieve the lowest input current for a 500 ps eye opening. Based on experience in previous designs, the channel lengths of all analogue transistors were chosen slightly longer than the 0.6 μ m minimum (0.7 μ m for PMOS / 0.8 μ m for NMOS) to achieve a reasonable degree of process control. To reduce the number of variables in the optimisation, the NMOS width and PMOS width were taken to be the same.

In the gain broadened designs, the variables considered in the optimisation were the widths of the gain and load transistors. These were allowed to vary independently, but the gain transistor was restricted to $1\times$, $1/2 \times$ and $1/3 \times$ the front-end width to permit layout with unit transistors.

In the transconductance-transimpedance design, the variables considered in the optimisation were the widths of the transconductance stage and the transimpedance stage and the dimensions of the feedback transistor.

In both designs, the width of the decision stage was an additional variable.

Both process variation and DC offsets were taken into account in the optimisation. Each design was simulated with offsets of 0 mV and ± 25 mV on typical, worst power (fast n/p), worst speed (slow n/p), worst zero (slow n/fast p) and worst one (fast n/slow p) process corners; for each case, simulations were also performed for low-resistance, typical-resistance and high-resistance well resistor process corners. The figure-of-merit used to guide the optimisation was the worst-case eye opening over all these simulations.

During the optimisation, the circuit was simulated with two input patterns: a single one in a field of zeros and a single zero in a field of ones. These are the two worst-case bit sequences if the circuit bandwidth is the limiting factor on circuit performance; however, they do not properly characterise the reduction in eye opening caused by overshoot and resultant jitter. To give a better measure of actual performance, the final optimised designs were characterised by stimulating the input with one period of a maximal length pseudo-random bit sequence [271] with pattern length 2⁹-1. The eye opening was calculated from the transient simulation output using a script written in the AWK text processing language.

Simulations were performed using BSIM 3v2 transistor models¹ in HSpice. The well-resistor model was a JFET level 1 model which included parameters to model the distributed capacitance using a single π -section lumped approximation. To reduce simulation time during the optimisation, routing capacitance was neglected.

Physical layouts of the optimised circuits were produced to compare the real-estate requirements of the two approaches.

7.4.3 Results and discussion

The performance of the optimised designs is summarised in Table 7-1 and the transistor dimensions are shown in Figure 7-5. The sensitivity excludes the additional penalty due to stochastic noise which was estimated to be $\pm 0.5 \ \mu$ A based on HSpice noise simulations. The

¹ The default MOS capacitance model was used; the problem with this capacitance model discussed in the Appendix to Chapter 4 had not been discovered at the time of the study. However, the use of this capacitance model should only have affected the performance of the front-end feedback transistor and does not invalidate the comparison of the two post-amplifier designs. A thermal noise model that was accurate in the ohmic region was used.

power consumption includes the front-end (5.5 mW) but excludes the digital buffer used to drive the MQW modulator.

The transconductance-transimpedance design demonstrated almost a factor of two improvement in sensitivity over the most sensitive single-stage voltage-gain design. It was also slightly more sensitive than the two-stage voltage-gain design when assessed in terms of the optimisation criterion (the sensitivity under worst-case process conditions) yet still managed to achieve this performance with a power consumption comparable to the single-stage voltagegain design. Two-stage designs of comparable power consumption that were assessed were less sensitive.

| | sensitivity (I _{PEAK} / µA) | | | | typical | |
|-----------------------------|--------------------------------------|--------|-----------------|--------|-------------|--|
| | worst-case process | | typical process | | power | |
| | 0 mV ±25 mV | | 0 mV | ±25 mV | dissipation | |
| | offset | offset | offset | offset | / mW | |
| voltage gain (single-stage) | 11.0 | 21.5 | 6.7 | 12.0 | 9.8 | |
| voltage gain (two-stage) | 5.9 | 15.0 | 2.2 | 9.2 | 12.3 | |
| transconductance- | 5.9 | 11.1 | 2.5 | 7.2 | 9.7 | |
| transimpedance | | | | | | |

 Table 7-1: Sensitivity of optimised circuits excluding noise penalty at 1Gbit/s (peak

 photocurrent per beam)



(c) transconductance-transimpedance

Figure 7-5: Schematics of optimised post-amplifiers. W/L sizes in µm are indicated (u).



Figure 7-6: Comparison of eye diagrams for two-stage voltage-gain and transconductance-transimpedance designs (typical process, 0 mV offset, ±3 μA photocurrent, 2⁶-1 bit PRBS)

The power consumption of all the designs is somewhat higher than those considered in Chapter 4. This is in part a reflection of the higher speed of the design, and in part a consequence of ignoring the power consumption during the optimisation. The transconductance-

transimpedance design intrinsically contains two amplifiers operating in the linear region and, at lower speeds, it is clear that single-stage voltage-gain designs will tend to have lower power consumption. In the higher speed circuits where the gain-bandwidth advantage is more useful, the power consumption penalty of using two linear stages is offset by the fact that a single stage design must use larger transistors to provide acceptable gain.

Some general remarks can be made about the dimensions of the transistors in the optimised designs (Figure 7-5). In all cases, sizing the post-amplifier input transistors smaller than the front-end helped to reduce overshoot. However, in the voltage-gain designs, the additional gain produced by a slightly larger post-amplifier input transistor was worthwhile. The use of inverse-tapering in the two-stage design is also evident.

Eye diagrams for the two-stage and transconductance-transimpedance design are compared in Figure 7-6 at slightly above the sensitivity threshold under typical conditions. Notice that the bandwidth of the optimised two-stage design is not quite sufficient to produce a full voltage swing at the post-amplifier output; this translates into pattern dependent jitter at the modulator output after the decision stage and modulator driver circuit. In contrast, the transconductance-transimpedance design has sufficient bandwidth resulting in a relatively clean eye diagram.

7.4.4 Small-signal results

Small-signal analysis was also performed on the optimised post-amplifiers (Table 7-2). These results must be treated with caution because of the large signal nature of the design. Nevertheless, the transconductance-transimpedance design has a small-signal *bandwidth* similar to that of the single-stage voltage amplifier, and a small-signal *gain* comparable to that of the two-stage design. The gain-bandwidth product is better than that of the other designs by a factor of about three.

| | gain | | | 3 dB bandwith / GHz | | |
|-------------------|------|------|------|---------------------|------|------|
| | min. | typ. | max. | min. | typ. | max. |
| single-stage | 2.8 | 3.4 | 3.6 | 0.8 | 1.3 | 2.9 |
| two-stage | 6.6 | 9.1 | 10.2 | 0.2 | 0.4 | 1.1 |
| transconductance- | 6.1 | 10.8 | 22.0 | 0.7 | 1.2 | 2.2 |
| transimpedance | | | | | | |

Table 7-2: Process tolerance of post-amplifier small-signal gain and bandwidth

Under all process conditions, the small-signal amplitude response was flat indicating that the overshoot is acceptable.

The small-signal results also illustrate the process sensitivity of the two designs. The gain of the single-stage amplifier is indeed well controlled. The overall gain of the two-stage design is

less well controlled because the relative tolerance is squared by cascading two voltage stages. As expected, the transconductance--transimpedance design has poor worst-case tolerance on both gain and bandwidth; despite this, it produces better overall worst-case performance.

7.4.5 Conclusion

An important disadvantage of the transconductance-transimpedance design in this application is the layout area required by the well resistor. The receiver occupied $47 \times 29 \ \mu\text{m}^2$ versus $25 \times 29 \ \mu\text{m}^2$ for the single-stage voltage gain design. This is still compatible with including several hundred receivers on a chip; however, it would not be the preferred design in lower speed applications where sufficient gain-bandwidth is available from the more compact, voltage-gain approach.

The study has been based on one particular optimisation approach and specification. It is not, therefore, sufficient to demonstrate a universal advantage. However, the results give clear evidence that the design technique can be of value in smart-pixel applications, and, for the particular set of assumptions made here, it offers a small but distinct advantage in performance at some cost in layout area. This is achieved in spite of poor process sensitivity; it is conceivable that an even greater advantage could be obtained with a more complex variant of the design that could compensate for the this variation. As it stands, it is questionable whether the benefit is sufficient to warrant the additional design and verification effort required in a circuit that is highly sensitive to parameter variation.

It must also be pointed out that, since performing this study, the n-well sheet resistance specification in this particular process has been substantially reduced. Consequently, the layout area and parasitic capacitance associated with the feedback resistor have increased and so the design may be less advantageous. This does not invalidate the design concept, but does suggest that its general applicability may depend on the availability of controlled analogue resistor process modules.

Nevertheless, the study indicates that the transconductance-transimpedance cascade is worth evaluating as a possible design option in systems using high bit-rate channels and validates the theoretical argument given in the preceding section.

7.5 Offset performance of the transconductance-transimpedance cascade

In the previous section, it was shown that the gain of a typical transconductancetransimpedance cascade is comparable with that of a two-stage voltage amplifier. It could be argued that that the transconductance-transimpedance cascade is in fact a two-stage amplifier; in certain respects, this is undoubtedly the case. However, this section demonstrates that in one important respect – input referred offset – its behaviour is characteristic of a single-stage voltage amplifier and in that sense can be considered as a single stage.

The input-referred offset of a two-stage voltage-gain post-amplifier has already been analysed. Because of the relatively low gain per stage used in a wideband, multistage voltage amplifier, the offset in the second stage can make a significant contribution to the input referred offset. This is particularly true if inverse tapering is used and the second stage transistors are small.

To determine the input referred offset of the transconductance-transimpedance design, consider the small-signal equivalent circuits of the two gain stages formed about their own operating point (Figure 7-7) with the input biased at the operating point of the first stage.



Figure 7-7: Small-signal circuit used to evaluate offset

Performing nodal analysis on this circuit gives the equations:

$$-(v_{OP1} - v_{OP1})g_{m1} + (v_{OP1} - v_X)g_{ds1} + (v_{OUT} - v_X)/R_F = 0$$

$$(v_X - v_{OUT})/R_F - (v_X - v_{OP2})g_{m2} + (v_{OP2} - v_{OUT})g_{ds2} = 0$$
(7.11)

Solving these equations and making the same assumptions used in the derivation of the transfer function about the relative output and input impedance of the transconductance and transimpedance stages respectively gives:

$$\upsilon_{OUT} \approx g_{ds1} R_F (\upsilon_{OP2} - \upsilon_{OP1}) + \upsilon_{OP2}$$
(7.12)

and thus the input referred offset due to the transimpedance transistors is approximately:

$$\upsilon_{OFFSET} = \frac{\upsilon_{OP2} - \upsilon_{OP1}}{A_1} \tag{7.13}$$

Since the unloaded gain of the transconductance stage is high, the input offset is not sensitive to an offset voltage between the transconductance and transimpedance gain stages. This is a reflection of the fact that the signal is represented as a current at the internal node. The total input referred offset is thus characteristic of a single-stage voltage amplifier with a large gain equal to the overall voltage gain of the cascade. Two benefits of this characteristic are highlighted. It permits the use of a different amplifier topology for the transimpedance stage with only a small penalty in systematic offset. For example, an NMOS only inverter could be used for the front-end and the transconductance stage with a low gate-source drive voltage to provide a high transconductance per unit current; simultaneously, a CMOS inverter could be used for the transimpedance stage and decision stage to provide symmetrical rise and fall times. It also addresses the problem that was discussed in Chapter 6 of maintaining correct biasing of later stages in a high-gain multistage amplifier in future-generation low-voltage circuits: the final transimpedance stage will automatically bias about its own operating point. A method for correctly biasing the transconductance stage is still required, but this is easier because of the lower signal amplitudes at this point in the circuit.

7.6 Application to a differential receiver circuit

7.6.1 Introduction

In this section, the use of the transconductance-transimpedance circuit technique in a real application is described. A prototype circuit has been fabricated in a 0.6 µm digital process and characterised successfully at low frequency. However, a limitation of the circuitry included to test the high-frequency performance has prevented the speed advantage of the design technique from being verified.

The design presented in this section also illustrates how the circuit technique may be applied to an electrically differential receiver circuit and shows one possible method for using a MOSFET feedback transistor to provide better process control of the amplifier gain.

After a brief review of the application requirements and a discussion of the general motivation for the adoption of an electrically differential design, the detailed circuit design of the postamplifier and front-end used in this receiver are discussed. A summary of the performance predicted by simulations is presented. Experimental results from the prototype circuit are then discussed. The section concludes with an evaluation of the design based on the simulated and experimental performance.

7.6.2 Application requirements

The technique has been applied to the clock receiver in the SPOEC routing chip. In this section, the performance requirements of the clock receiver are briefly reviewed.

The optical clock signal was provided by the same VCSEL array used for the data channels. The system optical power budget predicts an optical power of about 3.5 μ A per photodiode. As in the data receiver, the photodiode capacitance in the final system was about 95 fF.

The target clock frequency for the receiver was 250 MHz. The timing of the digital circuits required a clock signal consisting of short bursts of pulses interspersed with long runs of zeros. This timing scheme requires a flat receiver response down to low frequencies and, essentially, the dynamic performance required of the receiver is similar to a 500 Mbit/s data receiver. The clock signal must also have low duty-cycle distortion, must be free of glitches and must have low jitter. It is therefore desirable that the receiver design has good immunity to digital noise.

The demands on power consumption and layout area were less severe for the clock receiver than for the data receiver. Whereas the routing chip contained approximately 4000 data receivers, only 64 clock receivers were required. This provides the extra design freedom necessary to improve upon the dynamic performance of the data receiver. Nevertheless, it was decided to contain the circuit layout within a digital standard cell. It is true in general that the circuitry used to detect an optical clock signal in parallel with a wide bus of data signals can afford to use more power and layout area. The final design occupied an area of 255 μ m × 38 μ m. In retrospect, the height of a single standard cell was too small to permit an effective layout of the receiver and many of the circuit nodes had significantly higher parasitic capacitance than might have been obtained in an equivalent square layout area. A square layout area could easily have been achieved without hindering the digital layout by stacking two standard cell rows and using the height of both cells to contain the clock receiver layout.

To avoid the need for a second feedback resistor control signal, it was decided to take advantage of some of the extra design freedom to implement a resistor biasing scheme that could automatically compensate for certain process variations.

7.6.3 General design approach

The two requirements of a generally robust design together with twice the equivalent data rate led to the selection of a two-beam, electrically differential design for the clock receiver.

It has previously been seen that, in general, a single-beam DC-coupled receiver cannot perform as well as a two-beam receiver. In particular, the post-amplifier design is limited by the requirement to provide a built-in threshold. Although the switching energy of the data receiver front-end is somewhat less than optimal because of the tight constraints on power consumption, achieving at least a factor of two improvement in switching energy as well as further improvements in pulse-width distortion and noise immunity was not considered to be a realistic target with a DC-coupled single ended receiver. It has already been highlighted that a lowerfrequency cut-off was not an option because of the burst nature of the clock signal. Had the clock been continuous, a single-ended design which took advantage of the extra layout area to provide a low-frequency cut-off may have been able to deliver the required performance. A two-beam approach was therefore adopted.

159

The majority of two-beam smart-pixel receivers to date have used two photodiodes connected in series; the designs that have been discussed so far are of this type. This is convenient with low-contrast input data because the subtraction of the two complementary optical signals is performed directly on the photocurrents; the common-mode rejection is limited only by the matching of the optical inputs and the photodiodes which can be very good. The DC operating voltages in the receiver circuit are independent of the common-mode photocurrent. However, series connected photodiodes require additional InGaAs fabrication steps as was discussed in Chapter 3.

The two-beam clock receiver circuit was therefore implemented using an electrically differential receiver. The electrically differential receiver was implemented using two single-ended front-ends followed by a post-amplifier with a differential input.

Table 7-3 makes a simple comparison of some of the properties of a single-beam, two-beam single-ended and two-beam differential receivers for a fixed size of photodiode and the same front-end bandwidth. It has been assumed that the front-end has been sized in proportion to the total photodiode capacitance to achieve a switching energy close to the limit set by the capacitance $C_F+C_{IN}/(A+1)$ as discussed in Chapter 4, that the photodiode capacitance dominates the input capacitance and that the output load capacitance is such that the damping factor of the receiver is acceptable.

| number of | electrical | input | photodiode | feedback | output | front-end | signal |
|-------------|--------------|---------|-------------|------------|--------|------------|---------------------|
| beams | structure | current | capacitance | resistance | swing | transistor | offset |
| | | | per branch | per branch | | width | voltage |
| single-beam | single-ended | Ι | С | R | IR | W | σ |
| two-beam | single-ended | 2I | 2C | R/2 | IR | 2W | $1/\sqrt{2} \sigma$ |
| two-beam | differential | 2I | С | R | 2IR | W | $\sqrt{2} \sigma$ |

Table 7-3: Comparison of the properties of single-beam and two-beam receivers

Notice that the two-beam single-ended design must use a smaller feedback resistor because of the extra photodiode capacitance on the input. Also notice that the power consumption of both types of two-beam front-ends is the same but that the differential mode offset voltage of the electrically differential design is twice that of the single-ended design, although the signal-to-offset ratio is the same. It is obviously possible to trade off extra power consumption for the same offset performance. Nevertheless, for designs of the same power consumption, this comparison suggests that, all things being equal with the post-amplifier, the electrically differential approach should provide comparable overall sensitivity in an offset limited design and better overall sensitivity in a gain limited design.

However, a major disadvantage of the electrically differentially receiver design is that the common-mode photocurrent is amplified by the front-end; a large common-mode photocurrent

could potentially shift the post-amplifier out of its linear region of operation. This is not a problem in this particular application because only limited dynamic range is required and the sources used are high-contrast. However, it might be a problem in modulator based interconnects.

Increased dynamic range could be obtained using a common-mode feedback circuit, possibly in conjunction with a fully differential front-end. However, a common-mode feedback circuit involves a considerable increase in complexity.

Electrically differential receivers have previously been used in optical interconnect applications [204][272].

An electrically differential front-end also provides the improved immunity to common-mode electrical noise required of the clock channel. Sources of common-mode noise include power supply switching noise and substrate voltage fluctuations. The good noise rejection of an electrically differential design is a useful feature of any receiver in a mixed-signal system where the analogue circuits are located in close proximity to the digital circuits; it is particularly important in a receiver designed for a clock signal because of the requirement for a glitch free signal with low jitter. The two-beam electrically single-ended receiver design has poor immunity to common-mode voltage noise.

A secondary benefit of the use of a two-beam receiver is that the two VCSELs used for the data channel may be biased slightly above threshold to reduce turn-on delay jitter. However, in practice it has proved possible to operate the VCSELs at the required speed in this system without such a pre-bias.

7.6.4 Detailed design

Overall structure

The overall structure of the receiver is shown in Figure 7-8.



Figure 7-8: Simplified clock receiver schematic. NMOS and PMOS lengths are 0.8 μm and 0.7 μm respectively unless stated otherwise

The photocurrents from the two detectors are converted into a differential voltage signal using two separate transimpedance front-ends. A differential transconductance amplifier converts the differential voltage into two currents, one of which is mirrored and subtracted from the other at the internal node X. The difference current is converted to a near full swing output voltage using a simple single-ended transimpedance stage. A final analogue inverter is used to fully restore the signal to digital levels.

Thus, as well as providing a high gain-bandwidth product, the transconductancetransimpedance approach provides a convenient way to implement a differential to singleended conversion. The transconductance stage is biased directly by the common-mode output voltage of the frontend. The low operating point of the front-end makes it difficult to interface to a conventional CMOS operational transconductance amplifier (OTA) with a constant current bias at the common source terminal. In general, the conventional OTA would be preferred because it has lower common-mode gain and has a transconductance that is independent of the common-mode input voltage. However, for a high-contrast input signal, the common mode input signal is a quarter of the differential mode input signal; the higher common-mode gain is not a problem.

Front-end amplifier

The front-end used a gain block consisting of a cascode and a source follower. The original motivation for choosing a cascode gain stage was to improve sensitivity by increasing the gain compared with a simple inverter design so that a larger feedback resistor could be used. However, the higher gain resulted in an amplifier that was less stable and it was necessary to compensate the amplifier with additional feedback capacitance. This is a form of feedback-zero compensation [273]; this form of compensation is useful because it achieves the required reduction in closed-loop bandwidth by changing the feedback network but without significantly affecting the frequency response of the forward transfer function.

The compensation capacitor was realised with a metal2-metal1-poly sandwich capacitor placed underneath the flip-chip pad; the inner metal1 plate with the lower parasitic capacitance to ground was used for the front-end output.

The net benefit in performance of the cascode amplifier over a conventional inverter design after it had been adequately compensated was marginal. In retrospect, the significant increase in design complexity and layout area was probably not justifiable.

Front-end feedback resistor biasing

The front-end circuit illustrates a method for biasing the feedback resistor to compensate for variations in threshold voltage and operating point. The biasing scheme is based on a technique used in CMOS operational amplifiers to improve the process control of the resistor used to implement pole-zero compensation [274].



Figure 7-9: Schematic of clock receiver front-end with transistor widths in microns. NMOS and PMOS lengths are 0.8 μm and 0.7 μm respectively unless stated otherwise

The small-signal resistance of an ohmic-region MOSFET is given by:

$$R = \frac{1}{KP\frac{W}{L}(V_{G} - V_{S} - V_{T})}$$
(7.14)

where KP is the process gain factor and where the threshold voltage is itself a function of the source voltage through the body effect. One of the main problems with using a fixed bias voltage such as a power supply rail to control the feedback transistor is the uncertainty in the feedback transistor's threshold voltage and source voltage, which is determined by the operating point of the transimpedance amplifier. These two variables are positively correlated which makes matters worse.

The uncertainty in $V_{g}-V_{s}-V_{T}$ can be overcome by deriving V_{g} from a second transistor of the same type that is biased to have the same source voltage. To a first approximation, the threshold voltages of the feedback transistor and the reference transistor, including the increase due to the body effect, will be the same; the bias voltage will automatically adapt to

compensate for process variation in V_s and V_T . The feedback resistance will still vary due to the process gain factor KP and the effective width.

A scheme for achieving this is illustrated in the front-end schematic (Figure 7-9). Transistor Mn2 is the reference transistor used to generate the gate voltage for the feedback transistor. The current in the biasing transistors Mn2 and Mn3 is ratioed to the current through the cascode stage using the current mirror Mp1/Mp2/Mp3. Transistor Mn3 is sized to have the same current density as the input transistor of the cascode stage. Neglecting the finite output impedance of the transistors, the gate of Mn3, and hence the source of Mn2, is at the same potential as the gate of Mn5 and hence the source of the feedback transistor. This achieves the required compensation as explained above. Mn2 is sized to achieve the required gate-source drive voltage for the feedback transistor.

The example biasing scheme gives a nominal resistance of 78 k Ω at 50°C with a 5 V supply. The sensitivity of the resistance to worst-case variations in process, voltage and temperature is shown in Table 7-4. The process tolerance is a significant improvement over a fixed bias. Nevertheless, a simultaneous combination of worst-case variation in all three areas still leads to a significant spread in the resistor value.

| variation | minimum | maximum | |
|-------------|------------|------------|--|
| process | -20% fast | +27% slow | |
| supply | -13% 5.5 V | +17% 4.5 V | |
| temperature | -27% 0°C | +29% 70°C | |

Table 7-4: Process sensitivity of front-end feedback resistor relative to 5 V and 50°C

Similar schemes have been proposed in the literature. In particular, Williams suggests connecting the source of the reference transistor directly to the output of the transimpedance stage [198] which also improves the linearity of the feedback transistor.

MOSFET feedback resistor

The feedback element in the transimpedance stage of the post-amplifier used an ohmic MOSFET with a fixed bias voltage.

To overcome the non-linearity at large voltage swings, a clamp transistor was connected in parallel with the feedback resistor. No clamp is required for output voltage swings below the operating point of the transimpedance stage; the resistance of the feedback resistor decreases as the output voltage falls. However, when the output voltage swings above the operating point of the stage, the resistance of the feedback resistor starts to increase. Once the output voltage swing exceeds the threshold voltage of the clamp transistor², the clamp transistor turns on, significantly reducing the incremental feedback resistance and providing a sharp limit in the output swing. The transient response is illustrated in Figure 7-10 for worst-case data patterns with a 2 ns pulse time. Compared to the circuit without the clamp, the fall time is reduced and the pattern dependent jitter eliminated.



Figure 7-10: Comparison of post-amplifier operation with and without clamp present with 400 mV differential input swing

A similar technique has been used to increase the dynamic range of front-end circuits [270].

Because an NMOS transistor is used for both the transconductance element and the feedback resistor, the post-amplifier gain is much better controlled with respect to process variation and temperature than in the circuit considered in Section 7.4 where a well resistor was used for the feedback element.

The nominal small-signal gain of the post-amplifier is 5.2 at 50°C for a typical process. Over the five standard process corners, the gain varies from 4.4 to 7.4, a spread of about $\pm 25\%$.

² Note that the threshold voltage of the clamp transistor is increased by the body effect.

Although this is still not as tightly controlled as in a single-stage gain-broadened amplifier, it is sufficient for a practical design. In contrast, the transconductance varies by $\pm 50\%$.

As a result of temperature variation between 0°C and 70°C, the gain is controlled to within about $\pm 5\%$, despite a 35% variation in transconductance over this temperature range. The drop in transconductance at high temperature is primarily due to a drop in mobility which affects the feedback transistor and the transconductance element equally.

The process spread in bandwidth is significantly larger, varying from 0.3 GHz on a slow process to 1 GHz on a fast process. The bandwidth is determined by two poorly controlled parameters: the feedback resistance and the gate-drain overlap capacitance (variation \pm 50% and \pm 30% respectively). Better control of the bandwidth would require a biasing scheme that could control both R_F and g_m .

Although some improvement in the control of the bandwidth is desirable, it is no worse than in the single-stage gain-broadened design considered in the previous section. This design illustrates that manufacturable designs using the transconductance-transimpedance approach are possible.

7.6.5 Simulation results

Detailed simulation results are presented in Appendix B. In summary, the receiver had a DC sensitivity that was primarily limited by offset to about $\pm 1.7 \,\mu$ A.

The dynamic performance is illustrated in the simulated 500 Mbit/s eye diagram in Figure 7-11.



Figure 7-11: Eye diagram at the output of the clock receiver decision stage at 500 Mbit/s / 3.5 μ A input current. Eye diagrams for ± 4.7 $\sigma_{voffset}$ are overlaid with the typical eye.

A reasonably open eye is obtained in a typical receiver in the array; however, the offset voltage causes clock skew of about \pm 300 ps between worst-case super-pixels in the array. In the SPOEC system, the super-pixels are functionally independent and the only consequence of this skew is an increase in the minimum cycle time. However, this result indicates a fundamental

problem in the more general use of DC coupled receiver circuits for low-skew optical clock distribution. The skew could be reduced by operating at higher input currents or by using a receiver with a lower frequency cut-off.

The typical power consumption is 8.3 mW from a 5 V supply at 500 Mbit/s.

7.6.6 Experimental results

A version of the receiver design described in the previous section has been fabricated by an external foundry and tested experimentally. In this section, the results of the experimental characterisation are presented.

There were a number of wiring errors on the test-chip; these were corrected using an ion-beam milling facility. Apart from a very small imbalance in the input capacitance of the high-speed test circuit, these corrections should not have affected the circuit performance.

Test structures

Two instances of the clock receiver circuit were included on the prototype circuit for test purposes, one for DC characterisation and one for high-speed characterisation.

Current inputs to the inverting and non-inverting terminals of the circuit were derived from two external voltage signals using a test structure similar to that used to test the data receiver circuit (Section 5.4.1). The layout of the test structure is shown in Figure 7-12. The width of the test transistor was scaled from the 5 μ m value used in the data receiver test structure to 3 μ m in proportion to the target sensitivity. The length of the transistor was unchanged and so the limit on high-frequency operation is the same. The nominal value of the capacitor used to mimic the photodiode capacitance was again 50 fF.



Figure 7-12: Photomicrograph of the prototype clock receiver circuit

DC voltages were monitored with a single output pin using a simple on-chip 8-input analogue multiplexer implemented with complementary transmission gates. The multiplexer select inputs were provided by a 3-bit shift register; the clock and data input to this shift register were provided by external pins.

To permit high-speed testing of the clock receiver circuit, the output of the receiver was used to clock an on-chip four-state Gray code counter (Figure 7-13). The two state variables could be monitored via external pads.

A buffered copy of the clock receiver output was also provided as an external pin to permit measurement of eye diagrams. However, the I/O pad driver used in this chip contained circuitry to reduce short-circuit current during logic level transitions; this circuitry appeared to limit the speed of operation³.



Figure 7-13: Clock receiver test structure

Differences in the prototype design

The overall structure of the prototype design was similar to the final design described in Section 7.6.4. However, there were two significant differences in the implementation of the post-amplifier.

The transimpedance stage used a well resistor as a feedback element without a diode clamp. The well resistor was sized to give a nominal resistance of $18 \text{ k}\Omega$; however, a change in the foundry specification of the n-well sheet resistance resulted in an actual resistance value of about 7 k Ω . Consequently, the post-amplifier gain in the prototype circuit is significantly below the design value. In the final design, the increase in junction capacitance for the same nominal resistance that was caused by the reduced sheet resistance made a diode-clamped ohmic MOSFET the device of choice.

The transconductance stage used a cascode current mirror and was followed by a transimpedance stage with a lower operating point. The initial reason for using a cascode

³ Note that the I/O cells used in the data receiver chip were lower current drivers which did not contain this circuitry and seemed to be capable of operating at higher data rates

current mirror was to reduce the systematic offset between the transconductance stage and the transimpedance stage. However, more detailed simulation showed that the high-frequency common-mode rejection of the cascode current mirror was poor; a simple current mirror in conjunction with a transimpedance stage with a mid-rail operating point was used in the final design.

DC characterisation

The measured DC transfer characteristics illustrated in Figure 7-14 are approximately in line with simulations. A DC photocurrent swing of approximately 1 μ A on each photodiode is required to produce a fully restored logic level at the output of the decision stage; this is well within specification, despite the smaller-than-designed feedback resistor in the post-amplifier.

The differential offset voltage in this circuit resulting from the front-end only, determined from the inverting and non-inverting output voltages when the input transistor is biased well below threshold is 15 mV. Assuming that there is no systematic mismatch between the front-ends, one can state with 95% confidence that the standard deviation in the contribution of the front-end to the offset voltage is greater than 7 mV. A larger sample set is required to estimate an upper bound on the offset voltage. Nevertheless, the lower bound is large enough to support the argument that offset voltages are important in smart-pixel receivers.


(c) decision stage

Figure 7-14: DC transfer characteristics of the prototype clock receiver circuit

High frequency characterisation

The high-frequency measurements on the buffered output of the receiver demonstrated experimentally that the speed of operation was indeed limited by the transit frequency of the input test transistor (Figure 5-5) and, to a lesser extent, by the bandwidth of the current mirror in the post-amplifier transconductance stage.

The evidence that the transit frequency was the limiting factor comes from the asymmetry in the jitter on rising and falling transitions on the input voltage: the amount of pattern-dependent-jitter was greater in the falling edge. When the input voltage is high, the transit frequency of the PMOS transistor that forms the test structure is lower and so its output current takes longer (more bit-periods) to reach its equilibrium value. The current into the front-end at the onset of the falling edge therefore depends on the number of preceding bit-periods in which the input voltage has been high. In contrast, when the input voltage is low, the transit frequency of the PMOS transistor is higher and the output current can reach its equilibrium value within fewer bit periods.

The jitter produced by a falling edge in the input voltage is illustrated in the oscilloscope trace in Figure 7-15 (a) which shows the response of the circuit to a falling edge on one of the differential inputs (VCLK+) while a DC bias is maintained on the other differential input (VCLK-) to maintain the correct decision threshold⁴. The trace shows the response for two data patterns which should give the early and late extremes for the falling transition: a long sequence of ones followed by a zero (latest falling edge) and a long sequence of zeros followed by a single one and a zero (earliest falling edge). These two edges are not coincident indicating that the bandwidth of the circuit is not sufficient to allow the circuit voltages and currents to reach equilibrium values within a bit period.

Increasing the gate-source overdrive voltage by reducing the high-level input voltage, while holding the low-level input voltage constant was also found to significantly reduce the jitter in the falling edge even though this also reduces the differential input current swing (Figure 7-15 (b)⁵). This is consistent with the increase in transit frequency in the input high state. In itself, this trend could also be qualitatively explained in terms of an increase in front-end

⁴ By operating the circuit in single-ended mode and applying the signal to the branch of the circuit that does not pass through the current mirror, the bandwidth limit of the input transistor and front-end can be examined in isolation.

⁵ The DC bias was optimised separately for each value of the high-level input voltage such that the rising edge crossed the falling edge in the centre of the eye diagram. However, changes in the bias only had a small effect on the amount of pattern-dependent-jitter in the falling edge.

bandwidth produced by the reduction in the front-end feedback resistance at higher mean input currents shown in the DC characteristics. However, this explanation is unlikely because a significant reduction in jitter was produced by changes in the high input voltage that only produced a small shift in the mean input current. For example, reducing the high input voltage from 3.9 V to 3.8 V only slightly increased the mean common-mode photocurrent from $1.33 \mu\text{A}$ to $1.36 \mu\text{A}$ but reduced the jitter by approximately 10%.

These results provide reasonably conclusive evidence that the transit frequency of the input test transistor, and not the receiver itself, is the main performance limiting factor of the circuit.

However, the results also show that the jitter is greater when the signal is applied to the inverting clock input (Figure 7-15 (c) and (d)). The only asymmetry in the inverting branch is that the signal passes through the cascode current-mirror in the post-amplifier transconductance stage. Thus, the bandwidth of the current-mirror must also be a limiting factor; at frequencies well above the cut-off of the mirror, as much as half of the differential signal will be attenuated at in the normal differential mode of operation. However, as has already been discussed, this bandwidth limit was identified in more detailed simulation of the initial design and corrected in the design included on the final switching chip. The simulated mirror bandwidth in the revised design is about 650 MHz, well in excess of the requirement.

The common-mode rejection inherent in a differential circuit made it possible to partially overcome the bandwidth limit of the input test transistors by biasing the two input transistors well above threshold to increase the transit frequency. Figure 7-16 shows eye diagrams obtained from the buffered output of the clock receiver operated in differential mode with a current swing on each input of 5 μ A. Relatively clean eye diagram obtained up to 100 Mbit/s with closure at 125 Mbit/s when the input transistor is prebiased with a voltage of 3.0 V in the high state. The pattern dependent trajectory in the rising edge is due to the I/O pad driver and is the primary factor limiting the eye diagram. The same characteristic shape of rising edge was shown by transistor level simulations of the I/O pad drivers, although the rise time of the edge shown in the simulations was shorter. Without the common-mode bias (Figure 7-16 (d)), there is pattern-dependent-jitter in the falling edge and, to a much lesser degree, the rising edge. In a perfectly balanced differential circuit, one would not expect any asymmetry between the rising and falling edges. The asymmetry can again be attributed to the limited bandwidth of the cascoded current mirror. Because the signal in the inverting branch is significantly attenuated by the mirror at high-frequencies, the jitter is determined primarily by the signal on the noninverting input. The falling edge in the output corresponds to the falling edge in the noninverting input voltage, which is the slower of the two transitions due to the lower transit frequency in the high voltage state as explained above.



(d) pulse on VCLK-, 3.4 V high input voltage

Figure 7-15: Origin of frequency limitation on clock receiver circuit. The traces show the voltage waveform at the output of the clock receiver for the two input data patterns that produce the latest and earliest falling edge with a signal applied to only one of the differential inputs. The output voltage is non-inverting with respect to a voltage swing on VCLK+ and inverting with respect to a voltage swing on VCLK-. The pulse frequency is 80 MHz.



M 2.00ms Ch2 J

Intensity

Infinite Persistence

Variable Persistence 70.00 s

Readout Options

20mV

Graticule Fail



Telk 21008 25.0G5/s ET 8.480M Acqs

Persistence: 10.00 s

InstaVu Style

Vectors

Infinite Persistence

Variable Persistence 70.00 s

Readout

InstaVu Style

Vectors

Infinite Persistence

Variable Persistence 70.00 s

Readout Options

20m/

Graticule Fail

20mV

Graticule Fail

(b) 75 Mbit/s – with common-mode bias

122

8

301 St. 0m/52

Settings Mode «Display» <InstaVa

(d) 75 Mbit/s – without common-mode bias

Figure 7-16: Clock receiver eye diagrams with and without a common-mode voltage bias on the test input transistors

The operation of the clock receiver with an input signal similar to the input signal that is required to operate the SPOEC switching chip was also tested. A repetitive sequence consisting of 8 return-to-zero pulses followed by 24 zeros was applied. At frequencies up to a certain maximum, the four-state counter produced a stable divided output as shown in Figure 7-17. Slightly above this frequency, the counter occasionally slipped one state, resulting in an unstable waveform on the oscilloscope display. Stable operation for ten seconds was the criterion used to define an upper limit on the frequency of operation; however, there was a rapid transition between a completely stable waveform and an unstable waveform as the frequency was increased.



Figure 7-17: Divided clock output with 100 MHz RZ input pulses. Upper trace is the counter output; lower trace is a trigger signal.

The increase in the maximum operating frequency with a prebias on the input transistor provides further evidence that the test transistor is responsible for the limit in high-frequency performance. Without prebiasing the input transistor (input voltage swing between 3.90 V and 2.25 V), the maximum pulse frequency for stable operation was 65 MHz; this increased to 95 MHz with an input voltage swing between 3.00 V and 1.93 V. In both cases, the input current swing on each input was 5.0 μ A with a input high current of 0.05 μ A and 1.75 μ A respectively. Marginally faster operation to 105 MHz was possible by using an inverted input pattern with 8 return-to-one pulses followed by 24 ones.

The sensitivity of the circuit with prebias at the fixed operating frequency of 100 MHz was $1.6 \,\mu\text{A}$ current swing on each input. This result was obtained with an input voltage swing between 2.28 V and 2.66 V corresponding to a current swing between 3.0 μ A and 4.6 μ A. At larger gate-source drive voltages, it seemed that the benefit gained from the increase in the transit frequency of the test transistor was offset by the reduction in the transconductance of the second-stage input transistors at higher common-mode input currents. This sensitivity, if it can be obtained without prebias in the final implementation of the circuit with optical inputs, is well within the system specification of 3.5 μ A.

7.6.7 Evaluation of the design

Whilst certain details of the front-end, such as the biasing scheme for the feedback resistor, are of some wider relevance, this design is of value primarily as an illustration of the practical use of the transconductance-transimpedance post-amplifier structure in a smart-pixel application.

It demonstrates that it is possible to implement a post-amplifier with a gain that is reasonably well controlled with respect to process variation using only standard transistors while still achieving a high gain-bandwidth product. This is an advance over the design considered in Section 7.4.

It also demonstrates that the technique can be applied to an electrically differential receiver circuit. However, the circuit, as it stands, does not allow one of the main benefits of a differential receiver circuit to be realised: it is not possible to split the front-end and post-amplifier power supplies in order to reduce switching noise because of the low operating point of the front-end. Simple variants of the design using, for example, a conventional CMOS operational transconductance amplifier for the transconductance stage would provide this benefit.

Alternative approaches do exist. No comparison has been made against the more conventional technique of using a low-gain source-coupled differential amplifier with a resistive load [275]. This structure has an inherently limited output voltage swing and can be used to construct a limiting amplifier with a much more non-linear gain characteristic than a simple CMOS inverter for low input voltage swings. It was not evaluated for this system.

The design provides further evidence that the high gain-bandwidth product of this circuit structure is useful. However, the analysis of Chapter 6 has already shown that, in advanced CMOS technologies, offset voltages may be the primary factor limiting circuit performance, rather than post-amplifier gain-bandwidth product. To fully exploit the benefits of the circuit technique, it will be necessary to adapt the designs presented in this chapter to provide a low-frequency cut-off.

177

In more specific terms, the simulations of the design suggest that it will come close to fulfilling the requirements of the SPOEC system. Although limitations of the test circuitry have prevented experimental verification of the design at the target operating frequency, all tests that have been possible have produced results that can be explained in terms of factors which will not affect the final design. Although a number of modifications have been made in the final circuit⁶, it is reasonable to anticipate that the final design will perform at least as well as the prototype.

7.7 Conclusion

In this chapter, the transconductance-transimpedance circuit approach has been applied for the first time to post-amplifiers in smart-pixel receiver circuits.

The theoretical gain-bandwidth advantage that provides the motivation for considering the circuit technique has been reviewed. A simple implementation using a well resistor, suitable for smart-pixel circuits, has been proposed. A detailed comparison against a conventional wideband low-gain voltage amplifier structure, taking into account practical considerations such as process variation, has shown that the technique can indeed provide improved sensitivity in high-speed circuits at some cost in power consumption and layout area. However, process variation reduces the performance advantage in this simple design.

A modified structure which gives performance that is less sensitive to process variation has been applied in the two-beam, electrically differential clock receiver for the SPOEC switching chip. Preliminary experimental results from a prototype implementation in 0.6 μ m CMOS, tested with electrical inputs, have verified that the design operates correctly with a 100 MHz RZ clock input with a peak differential input current swing of 1.6 μ A per photodiode, although the high-frequency performance has been degraded by limitations of the circuitry used to provide the electrical current inputs. A revised version of the circuit designed for optical inputs has been fabricated and, at the time of writing, testing is awaiting flip-chip assembly of the optoelectronic devices. Simulations indicate that the revised design will operate at 250 MHz with a photocurrent of 3.5 μ A.

⁶ The increase in photodiode capacitance is not expected to have a big impact on performance; the high gain of the cascode stage means that the bandwidth is largely determined by the feedback capacitance.

Chapter 8

Electrical crosstalk in large photoreceiver arrays

8.1 Introduction

Previous chapters have considered how the problem of designing a photoreceiver circuit is altered by the constraints imposed on power consumption and layout area by the need to integrate a large number of circuits on a single chip. However, the performance of the resulting receiver circuits has been considered in isolation.

A second major difference in the design of large arrays, as opposed to single receivers, is the possibility of interaction between the receiver circuits. Experimental evidence [276] has shown that simultaneous operation of receivers in an array significantly degrades performance; for example, Woodward et. al. report a 2.5 dB reduction in the sensitivity of an individual channel in an array of 50 two-beam smart-pixel receivers when all channels in the array are active in comparison to the sensitivity when the channel is tested in isolation. The shared power supply network has been proposed as a mechanism for this degradation. It has already been noted in chapter 3 that demonstration of simultaneous operation of a terabit/s scale optical interface is a significant outstanding research goal, and a full understanding of crosstalk is essential if robust operation of arrays of this scale is to be achieved.

This chapter makes progress towards this goal by analysing the problem of crosstalk between receivers in large arrays as a result of a shared power supply network. The problem is tackled in two stages: first, the sensitivity of a simple smart-pixel photoreceiver circuit to voltage noise on its power supplies is analysed and discussed in relation to the main receiver design variables such as photodiode capacitance; then, a method for estimating the supply voltage noise from the magnitude of the switching transients generated by other receiver circuits and a simplified model of the power supply network impedance is presented. Finally, using the results of these two sections, an order of magnitude estimate of the input referred noise resulting from crosstalk in an example circuit loosely modelled on the SPOEC switching chip is made. This calculation is not intended to provide accurate quantitative predictions of the crosstalk in the SPOEC system; rather, it is intended to illustrate a method by which crosstalk as a guide to how best to improve the crosstalk immunity of existing receiver designs.

The analysis confirms that crosstalk is indeed an extremely important issue in the design of large receiver arrays; some of the techniques that might be used to control the problem are discussed. In particular, the measures adopted in the SPOEC system to partially control the crosstalk are described. These measures are not in themselves expected to be sufficient to

completely eliminate crosstalk from the final system: a detailed understanding of the crosstalk problem was not achieved until well into the system design process by which time important system specifications that could have been adjusted to allow for crosstalk, in particular the receiver sensitivity, had been fixed. The measures were retrofitted to an existing receiver design and, in future, it would be more appropriate to redesign the circuit taking into account the analysis presented in this chapter from the outset.

Whilst this chapter falls somewhat short of providing a complete analysis of and solution to the crosstalk problem, it nevertheless identifies the main issues that must be considered and is a useful starting point for future receiver array designs.

8.2 Analysis of supply sensitivity

8.2.1 Basic approach

The sensitivity to supply noise is analysed by considering the signal coupled onto a 'quiet' receiver in which the input photocurrent is not changing. The signal is expressed in terms of an input referred photocurrent, which is calculated by dividing the signal produced by the supply voltage noise at the input to the decision stage by the overall DC transimpedance gain of the receiver. The problem of power-supply noise induced jitter is not considered although it is also an important issue.

The receiver circuit considered is typical of smart-pixel designs. It consists of a front-end, a single-stage post-amplifier and a decision stage, each based on a complementary inverter (possibly including gain-broadening transistors). The output of the decision stage drives a gate connected to the digital supply. This model is appropriate for the SPOEC data receiver described in chapter 5 and the Lucent smart-pixel receivers described in chapter 4.

For generality, we initially assume that each of the four stages uses a separate power supply and subsequently discuss the consequences of using a common supply for subsets of these stages. A separate detector bias is also assumed; this is consistent with the approach adopted in most hybrid CMOS/QCSE modulator smart-pixel circuits to date, in which a separate bias voltage is required to allow the detector absorption peak to be tuned to the wavelength of the incident light, although the analysis shows that this approach may not be ideal from a crosstalk point of view.

To simplify the analysis, the detector bias terminal is grounded and the noise expressed in terms of the differential-mode and common-mode voltage on the pairs of power supply terminals. It is possible to analyse the supply sensitivity in terms of any set of linearly independent voltages. However, differential-mode and common-mode voltages with respect to the detector bias are an appropriate choice for a reasonably symmetric complementary inverter

180

front-end because, as the analysis will show, the sensitivity to common-mode noise is much greater than the sensitivity to differential-mode noise. This choice of variables also has several advantages when the impedance of the power supply network is analysed.



equivalent input noise = [Yn][Vn] = [Yn][In][Z]

Figure 8-1: Pictorial overview of approach to analysing power supply noise

Figure 8-1 illustrates the overall approach to the power supply noise analysis. The switching transients [I] generated by the active receivers through the impedance [Z] of the supply network produce a voltage noise $[V_{a}]$ on the supply terminals. In this section, the vector of noise transfer admittance [Y], which relates the noise on the 8 independent supply voltage variables to the input referred noise current, is derived.

8.2.2 Front-end

The analysis method for sensitivity to power supply noise is similar to the standard technique for calculating the effects of random noise [277].

The circuit with a noise voltage v_{N} applied to the power supply terminals can be rewritten in terms of a circuit without any noise voltage applied with two noise generator current sources $i_{N1} = g_{N1}v_N$ and $i_{N2} = g_{N2}v_N$ added (Figure 8-2). g_{N1} and g_{N2} represent the current coupled into the output and input nodes of the amplifier.



Figure 8-2: Small-signal model used to analyse supply crosstalk

In the case of common-mode voltage noise, we have that:

$$g_{N1CM} = g_m + g_{ds} + sC_L$$

$$g_{N2CM} = sC'_{IN}$$
(8.1)

where $C'_{IN} = C_{IN} - C_{PHOTO}$ is the input capacitance not associated with the photodiode. In the case of differential voltage noise, we have that:

$$g_{N1DM} = \frac{1}{2}(g_{mp} - g_{mn} + s(C_{Lp} - C_{Ln})) = g_X + sC_X$$

$$g_{N2DM} = \frac{1}{2}s(C_{INp} - C_{INn}) = sC_Y$$
(8.2)

where the total transconductance, load capacitance and input capacitance have been expressed in terms of the components associated with the positive and negative supply:

$$g_{m} = g_{mp} + g_{mn}$$

$$C_{L} = C_{Lp} + C_{Ln}$$

$$C'_{IN} = C_{INp} + C_{INn}$$
(8.3)

and where g_x , C_x and C_y have been defined as short-hand to represent half the difference between these components.

The generator current sources are related to the output from the post-amplifier using the zparameters of the front-end circuit in the absence of supply noise:

$$v_{OUT} = z_{21}i_{N1} + z_{22}i_{N2} \tag{8.4}$$

where

$$z_{21} = \frac{1 - g_m R_F}{g_m + g_{ds}} \frac{1}{P(s)} \approx \frac{-R_F}{P(s)}$$

$$z_{22} = \frac{1}{g_m + g_{ds}} \frac{1 + sR_F(C_F + C_{IN})}{P(s)}$$
(8.5)

and P(s) represents the poles of the front-end transfer function normalised to P(0)=1. Straightforward simplification with reasonable approximations gives the coupled signal referred back to the input:

$$i_{CM} = \frac{1}{R_F} \frac{1 + sR_F C_{PHOTO}}{P(s)} H_{POST-AMP}(s)$$

$$i_{DM} = \frac{1}{R_F} \frac{g_X}{g_m + g_{ds}} \frac{1 + s(R_F C_{IN} + C_X / g_X)}{P(s)} H_{POST-AMP}(s)$$
(8.6)

where the current into the amplifier is defined as positive and where $H_{POST-AMP}(s)$ is the transfer function of the post-amplifier normalised to $H_{POST-AMP}(0)=1$.

At low frequencies, the rejection of differential-mode voltage noise is high but common-mode noise couples one-to-one onto the receiver output. The differential-mode noise is proportional to the ratio of the difference between the NMOS and PMOS transconductances to the overall transconductance, which is close to zero for a nominally symmetric inverter.

However, the high-frequency noise rejection is a more serious problem. There is a zero in the noise transfer admittance at a frequency of $\omega = 1 / R_F C_{PHOTO}$. Above this frequency, the noise transfer admittance increases linearly with frequency until the cut-off in the combined frequency response of the front-end and post-amplifier is reached. The noise in this frequency range is equivalent to that produced by direct capacitive coupling through a capacitor of value C_{PHOTO} in the common-mode case and a small fraction of C_{PHOTO} in the differential-mode case. The qualitative behaviour of the noise transfer admittance as a function of frequency is shown in Figure 8-3, assuming that the overall photoreceiver has a second-order transfer function. It can be seen that, at the worst case frequency slightly below the cut-off of the receiver, the noise transfer is approximately a factor of A+1 higher than the DC value where A is the voltage-gain of the front-end amplifier (A+1 is the factor by which the input time constant formed by the photodiode and the feedback resistor is reduced by the negative feedback of the front-end amplifier). In practice, the factor is somewhat less than this because of the contribution of other time constants to the overall response.



Figure 8-3: Qualitative behaviour of front-end noise transfer admittance

The worst-case high-frequency noise transfer admittance can be used to make some general statements about how the overall noise susceptibility is related to the receiver design variables. In the simple case where the photodiode capacitance dominates the frequency response, an upper bound on the signal-to-noise ratio in terms of the noise voltage v_N on the front-end supply is given by:

$$SNR = \frac{V_{MIN}}{V_N} \frac{1}{A+1}k$$
(8.7)

where k is, for the common-mode and differential-mode cases respectively, $k_{CM} = 1$ and $k_{_{DM}} = g_x / (g_{_{m}} + g_{_{ds}})$ and where $V_{_{MIN}}$ is the minimum signal required at the output of the front-end as defined in chapter 4. Thus the susceptibility of this simple receiver design to power supply noise is largely determined by the front-end gain and $V_{_{MIN}}$ which are both directly related to switching energy at a fixed photodiode capacitance. The photodiode capacitance does not in itself affect the signal-to-noise ratio; although the equivalent input current noise produced by a given amount of supply noise increases in proportion to the photodiode capacitance, the minimum input current signal required to produce a signal V_{MIN} at the front-end output in the absence of crosstalk increases by the same factor. Recall from chapter 4 that typical values of V_{MIN} and A were around 200 mV and 20; thus, as a first-approximation, the common-mode voltage noise relative to the photodiode bias at the amplifier cut-off frequency must be less than about 10 mV. This is a demanding constraint in a noisy digital environment. The constraint on the differential-mode voltage noise is less severe because a symmetric inverter design can be used to achieve a low value of k_{DM} . Independent process variation of the NMOS and PMOS transistors will limit the value that can be achieved; sample simulations give a spread of -0.03 to 0.06 for a nominally symmetric inverter. The nominal value of k_{DM} in the SPOEC data receiver, which did not use a perfectly symmetrical inverter, was 0.2.

8.2.3 Second-stage

The same technique can be used to analyse the supply sensitivity of the post-amplifier. The small signal model is shown in Figure 8-4.



Figure 8-4: Small-signal model of post-amplifier used to analyse supply crosstalk

The common-mode and differential-mode noise generators are:

$$g_{N3CM} = g_{m2} + g_{L2} + sC_{L2}$$

$$g_{N3DM} = \frac{1}{2}(g_{m2p} - g_{m2n}) + \frac{1}{2}s(C_{L22p} - C_{L22n}) = g_{X2} + sC_{X2}$$
(8.8)

and the input referred noise currents are:

$$\frac{i_{CM}}{v_{CM}} = -\frac{1}{R_F} \left(1 + \frac{1}{|A_2|}\right) H_{POST-AMP}(s) \left(1 + s \frac{C_{L22}}{g_{m2} + g_{L2}}\right)$$

$$\frac{i_{DM}}{v_{DM}} = -\frac{1}{R_F} \left(1 + \frac{1}{|A_2|}\right) H_{POST-AMP}(s) \frac{g_{X2}}{g_{m2}} \left(1 + s \frac{C_{X2}}{g_{X2}}\right)$$
(8.9)

where the total load capacitance C_{L2} comprises a capacitance to the post-amplifier supplies C_{L22} and a capacitance to the decision stage supplies C_{L23} $A_2 = -g_{m2} / g_{L2}$ is the DC gain of the post-amplifier. In both cases, the zero is (slightly) above the cut-off of the post-amplifier and the noise coupling is thus determined primarily by the DC shift in operating point.

Low-frequency common-mode voltage noise couples approximately one-to-one onto the postamplifier input signal. If the post-amplifier supply is shared with the front-end, then the larger part of the low-frequency common-mode noise transfer admittance cancels with the corresponding term in the front-end to leave an overall low-frequency noise transfer admittance of $-1 / (R_FA_2)$.

The differential-mode noise transfer admittance is again much smaller and will cancel in the same way when a shared supply is used if the inverter ratio is the same¹.

The noise generator associated with the current coupled through the input capacitance of the post-amplifier, which would appear in the same place as i_{N2} in the front-end circuit, has been neglected. The justification for this is that, unless the voltage noise on the post-amplifier supply

¹ In a single-ended receiver, the inverter ratio may differ in order to introduce a fixed threshold. However, the inverter ratio is the same in the SPOEC data receiver as discussed in chapter 5.

is very much greater than the voltage noise on the front-end supply, this contribution to i_{N2} will be smaller than that due to the front-end supply noise.

Overall, the noise transfer admittance for the post-amplifier supply is significantly lower than for the front-end.

8.2.4 Decision stage

The switching characteristic of the decision stage together with the signal swing at its input determines the overall noise margin of the data link. As discussed in chapter 4, the decision stage requires a minimum input signal swing centred about its switching point to produce a fully restored output level with an acceptable edge time. This minimum input signal defines an upper and a lower switching threshold. The noise transfer admittance is then determined by the shift in these switching thresholds, as a function of the differential and common-mode supply noise, referred back to the input.

Assuming the thresholding inverter is followed by a second inverter connected to the same supply to clean up the signal, a common-mode voltage shift produces an equal and opposite equivalent voltage swing at the post-amplifier output. Referred back to the input this gives a noise transfer admittance of:

$$\frac{i_{CM}}{v_{CM}} = \frac{1}{R_F} \frac{1}{|A_2|}$$
(8.10)

which cancels with the low-frequency component of the post-amplifier and front-end terms if all three stages are connected to the same supply.

Differential-mode noise is analysed in the same way as in the post-amplifier. Although the decision stage is a large signal circuit, a small-signal analysis is still appropriate for a quiet receiver. The differential noise rejection is again good for a symmetric inverter and dominated by the low-frequency behaviour. Delay sensitivity is a separate issue and it may be that the jitter performance of the decision stage sets the constraint on maximum allowable differential noise.

8.2.5 Digital stage

The input to the first digital gate could be analysed in exactly the same way as the decision stage. However, the output signal from the decision stage is a rail-to-rail swing digital logic level and, provided the digital power supply network has been adequately designed, the interface should be relatively immune to common-mode noise. The interface between the decision stage and the first digital gate is not considered further in this chapter.

8.2.6 Optimum partitioning of power supplies

In this section, the advantages and disadvantages of using a common power supply for some or all of the receiver stages are discussed.

As shown in the analysis above, an advantage of using a common power supply is a reduced susceptibility to low-frequency common-mode voltage noise. The issue of a DC common-mode shift is particularly relevant at the interface between the front-end and the post-amplifier where the signal amplitude is small. Although the common-mode DC voltage between the two stages should be zero in the ideal case where the resistance of the positive supply connection is identical to the resistance of the negative supply connection, some asymmetry in the resistance is inevitable in practice². This, together with the fact that the current transients in both stages are relatively small, makes it attractive to join the two supplies together. In the remainder of this chapter, it is assumed that this is the case. The common supply is referred to simply as the "front-end" supply from now-on.

Connecting the decision stage to the front-end supply is also possible. However, the transient component of the decision stage supply current is comparable with the static DC component. It is therefore reasonable to anticipate that the high frequency common-mode voltage noise introduced by the decision stage on a shared supply would be comparable to the low frequency component. Since the high-frequency common-mode noise susceptibility of the front-end is much worse than the low-frequency common-mode susceptibility of the front-end/post-amplifier combination, it would seem, on balance, it makes sense to separate the decision stage supply. This is assumed in the analysis that follows. The case is not as clear-cut as that for joining the front-end and the post-amplifier supplies, but the assumption is partially vindicated by the voltage noise estimates in section 8.4.3.

The optimum partitioning of the supplies in different receiver designs, in particular in electrically differential circuits, could well be different.

8.3 Estimation of voltage noise

8.3.1 Introduction

This section discusses a method for calculating the voltage noise on the receiver power supply in an array of receivers from estimates of the transients in the receiver power supply current together with a description of the power supply distribution network. The receivers are assumed

² The DC current of these stages is large and results in a significant DC differential-mode voltage drop across the supplies; consequently, a small relative mismatch in the resistance of the positive and negative supplies will produces a significant common-mode voltage drop.

to be distributed in an array across an area of an integrated circuit with external power supply connections at the edge of the array.

The voltage noise is estimated from the response of the power supply network to the current transient produced when all the receivers switch in the same direction simultaneously. This scenario will produce the largest peak current and, if the voltage transients produced by an input step persist for less than one bit period, it will also give a worst-case estimate of the voltage noise. If the voltage transient persists for several bit periods (for example, if there is a resonance in the power supply network) then the worst-case voltage noise could be somewhat higher than this estimate.

The steps in the analysis are as follows:

- 1. a simplified equivalent circuit of a single receiver cell that captures the main behaviour of the power supply current transient is established;
- the distributed array of receiver circuits is reduced into a Norton equivalent circuit, comprising a set of current sources and a simplified admittance matrix representing the onchip power network, with the package pins as the external connections;
- 3. techniques for modelling the package impedance are discussed;
- the admittance matrix of the Norton equivalent circuit of the on-chip network is incorporated into the impedance matrix of the external package to obtain an overall impedance matrix that allows calculation of the voltage noise at the pin from the Norton current sources;
- 5. the voltage noise at a receiver in the interior of the array is calculated from the voltage noise at the package pin.

8.3.2 Equivalent circuit of a receiver cell

The receiver cell has seven power supply terminals: front-end supplies AVDD/AGND, decision stage supplies ATHVDD/ATHGND, digital supplies DVDD/DGND and detector bias DETBIAS.

The relationship between the terminal voltages and currents of the receiver cell could be completely described by a seven-terminal Norton equivalent circuit consisting of a timedependent current generator at each terminal and a 7×7 indefinite admittance matrix. The current generator functions can be obtained by measuring the current waveform drawn from a perfect voltage source in a transient simulation. To make the analysis of the power supply network analytically tractable, some entries in the indefinite admittance matrix must be neglected. The simple model used in the analysis that follows assumes that the receiver circuit itself consists of perfect current sources and that the only entries in the indefinite admittance matrix come from the explicit decoupling capacitance included in the cell. Whether or not this approximation is sufficiently accurate for quantitative prediction of the power supply noise has not been examined in detail; however, it should be adequate for its intended purpose of providing a qualitative estimate of the importance of crosstalk in receiver arrays. Implicit in this approximation is the assumption that the voltage transient on the receiver power supply is small enough that it does not significantly effect the magnitude of the current transient.

We define currents I_{AGND} , I_{AVDD} , I_{ATHGND} , I_{ATHVDD} , I_{DGND} , I_{DVDD} and $I_{DETBIAS}$ to represent the current flowing into the power supply network as shown in Figure 8-5. The sign of the currents is defined to be positive going into the power supply network and so I_{AVDD} , I_{ATHVDD} and I_{DVDD} would be expected to be negative.



(b) equivalent circuit in terms of transformed currents

Figure 8-5: Definition of receiver power supply currents and receiver equivalent circuit

Rather than analysing the problem in terms of these terminal currents, it is convenient to use a new set of linearly independent current variables that are obtained from a suitable linear combination of the original set. The main disadvantage of using the original terminal currents is that, because current must always flow in a loop, a given physical source of current (such as, for example, the current between the source and the drain of a transistor) will contribute equal and opposite transients to a pair of terminals. In certain parts of the analysis, the voltages produced by these opposite components can cancel (for example, if the mutual inductance between a certain pin and both DVDD and DGND is similar, then the induced voltage on that pin due to a differential-mode transient in the digital supply current will be small). Whilst an analysis in terms of the original current variables still gives the same result for the voltage

transient provided all signs and coupling impedances are correctly taken into account, it does not give much insight into the relative contribution of the different physical sources of current.

The new set of six current variables chosen to describe the receiver circuit correspond to the following distinct physical sources:

- 1. I_{ADM} : the differential-mode current through the front-end and post-amplifier.
- 2. I_{TDM} : the differential-mode current through the decision stage resulting primarily from the short-circuit current as the inverter passes through its switching point.
- 3. I_{DDM}: the differential mode current through the digital supplies (including that due to the logic circuitry outside the receiver).
- 4. I_{DETBIAS} : the photocurrent transient through V_{DETBIAS} .
- 5. I_{AT} : the current flowing from the front-end supply to the decision stage supply that is used to charge up the input capacitance of the decision stage.
- 6. I_{TD} : the current flowing from the decision stage supply to the digital supply that is used to charge up the input capacitance of the first digital gate. This current must produce a full-rail voltage swing with a sharp edge across the input capacitance and thus is relatively large.

The differential-mode currents are defined by:

$$I_{ADM} = \frac{1}{2} (I_{AGND} - I_{AVDD})$$

$$I_{TDM} = \frac{1}{2} (I_{ATHGND} - I_{ATHVDD})$$

$$I_{DDM} = \frac{1}{2} (I_{DGND} - I_{DVDD})$$
(8.11)

and the common-mode currents are defined by:

$$I_{TD} = I_{DGND} + I_{DVDD}$$

$$I_{AT} = I_{ATHGND} + I_{ATHVDD} + I_{TD}$$

$$I_{DETBIAS} = -I_{AGND} - I_{AVDD} - I_{AT}$$
(8.12)

where Kirchhoff's current law has been applied to Figure 8-5 (a) to relate these currents to the original current variables.

These transformed current variables are particularly useful in conjunction with the transformed differential and common mode voltage variables that were used to analyse supply sensitivity in Section 8.2. Provided the impedance of the positive and negative supplies of each pair is similar, an impedance matrix written in terms of the new voltage and current variables should be approximately diagonal i.e. the main effect of a differential-mode current should be a differential-mode voltage transient on the same supply.

The simulated current transients in response to an input step are shown in Figure 8-6 for the SPOEC data receiver in terms of the transformed current variables. Notice that some current transients are much larger than others: in particular, the transients associated with the decision stage are much larger than those associated with the front-end.

8.3.3 Modelling of receiver array power supply network

The on-chip power supply network is analysed by reducing the distributed array of power supply wires and receiver circuits into a lumped Norton equivalent circuit. This equivalent circuit allows the voltage transient at the edge of the chip to be calculated. The particular network used in the analysis is based on that used in the SPOEC system but is representative of that required by smart-pixel circuits in general.

The form of the on-chip supply network is as follows: the receivers are distributed across a two-dimensional array and the power supply rails are fed from opposite edges of the chip (say top and bottom) and connected to the package by means of wire bonds round the periphery; it is assumed that the use of flip-chip bonding for optoelectronic devices prohibits the use of flip-chip chip connection for the power supply network. The power supply rails are split in the centre of the chip such that the ends of the power supply rails furthest from the external connections are open-circuit. Only the series resistance and the shunt capacitance of the on-chip power supply rails are modelled. The inductance of the on-chip power supply rails is neglected. Explicit decoupling capacitance is included in the model of the receiver circuit as discussed in section 8.3.2



Figure 8-6: Simulated switching transients in SPOEC data receiver in terms of transformed current variables for rising and falling step inputs with an edge time of 1 ns and an amplitude of 5 μ A

For simplicity, the power supplies are assumed to be electrically isolated; in particular, the finite admittance of the substrate is neglected. This would be appropriate for a twin-well

process with a high-resistivity substrate. Further analysis is required to determine how appropriate this model is to a process that uses a lightly-doped epitaxial layer on top of a degenerately doped bulk³.

The detector bias voltage is assumed to be provided by a dedicated plane on the optoelectronic chip, similar to that used in the SPOEC detector arrays, with negligible impedance. Extension of the analysis to include a finite distributed resistance is straightforward.

The first step in the analysis is to reduce the distributed circuit of N receivers and decoupling capacitance in a single column to a single seven-port network.

The Norton equivalent circuit of a column of receivers could be obtained by direct simulation of a linear array of receiver cells in conjunction with a full model of the power supply network but an approximate analytic treatment of the problem gives more insight into the effect of design variables on the problem. The discrete linear array of receiver cells is approximated by a uniformly distributed transmission line. The detailed analysis of the problem is presented in Appendix 8.7. The model of the circuit at the edge of the chip is shown in Figure 8-7. The effect of the decoupling capacitance and the resistance of the power supply rails is to low-pass filter the differential mode current transient on the three supplies with a transfer function H(s). It also adds an admittance between the terminals of the three power supply pairs that can be approximated by a resistor in series with a capacitor. The resistor has value 2 R / 3 where R is the series resistance of a single power supply rail between the edge and the centre of the chip and the capacitor has value of C where C is the total decoupling capacitance of the analogue front-end, thresholding stage and digital stage respectively. The transfer function H(s) is given approximately by:

$$H(s) = \frac{1}{1 + s2/3CR}$$
(8.13)

³ The layout of the SPOEC data receiver used separate n-wells for the front-end, decision stage and digital logic and isolated the front-end p-substrate by surrounding it with an n-well. However, the p-substrate contacts of the decision stage transistors are near to some p-substrate contacts of some digital transistors and, consequently, the isolation between DGND and ATHGND may not be that high.



Figure 8-7: Norton equivalent of receiver array.

The common-mode currents $I_{\rm \scriptscriptstyle DETBIAS}, I_{\rm \scriptscriptstyle AT}$ and $I_{\rm \scriptscriptstyle TD}$ are not filtered.

Having derived a lumped equivalent circuit of a single column of receivers, it is straightforward to combine several columns connected in parallel to the same pin by multiplying the current source and the admittance by the number of pins, assuming that the impedance of the power supply tracks at the edge of the array is negligible.

This analytic treatment is useful for design purposes but for final verification of the power supply noise in a particular design, it would be desirable to perform a full transient simulation of a single column of receivers that includes a full description of the power supply network including effects that are more complicated to treat analytically such as substrate coupling.

8.3.4 Modelling of package impedance

Modelling technique

The package impedance is included by calculating its full impedance matrix.

The most important contribution to the package impedance comes from the inductance of the package pins. Both self- and mutual inductance are important in this problem: not only does the magnitude of the current transients on the different power supplies vary over several orders of magnitude (and so a transient on say a digital supply could still upset the sensitive analogue supply even if the coupling between the supplies is weak) but there is also substantial current flow between supplies (described by $I_{DETBIAS}$, I_{AT} and I_{TD}). It does not therefore make much sense to talk about the 'inductance' of an individual power supply because the overall loop inductance depends on the path taken by the current and hence inherently depends on the mutual inductance between pins⁴.

⁴ If the package contains a ground plane that is separated from the power supply rails by a height that is small in comparison to the pitch of the connections such that the coupling to the ground-plane is much stronger than the coupling to nearby pins then it would make sense to talk about the inductance of individual pins.

The package must therefore be modelled by a full N×N inductance matrix $[L_{ij}]$ relating the N terminal voltages $[V_i]$ to the N terminal currents $[I_i]$ according to:

$$\begin{bmatrix} V_i \end{bmatrix} = s \begin{bmatrix} L_{ij} \end{bmatrix} \begin{bmatrix} I_j \end{bmatrix}$$
(8.14)

and where the mutual inductance L_{ij} between pins i and j is related to the self-inductance L_{ii} and L_{ii} of the two pins by a coupling coefficient K_{ii} that is between -1 and 1.

$$L_{ij} = K_{ij} \sqrt{L_{ii} L_{jj}} \tag{8.15}$$

The inductance matrix can be obtained using both numerical [278][279][280][281] and experimental [282] techniques.

As explained in Section 8.3.2, it is possible to analyse the problem using the package inductance matrix in equation (8.14) directly. However, to provide more physical insight into the origin of the voltage noise on the power supplies, the equation is rewritten in terms of the transformed currents and voltages and a new inductance matrix that is derived from the original package inductance. The procedure for transforming the inductance matrix is outlined in Appendix 8.8.

The actual package used for the SPOEC switching chip has not been modelled in detail. Instead, the crosstalk calculations in this chapter use the inductance matrix of a different package, modelled by Damon [283], with a pin assignment that is loosely based on that used in the actual SPOEC circuit.

The use of an inductance matrix of a different package cannot be expected to provide accurate numerical estimates of the crosstalk in this particular system and would be unacceptable in the design of a new system. Nevertheless, the package is sufficiently representative of the system for it to provide a meaningful order of magnitude estimate of the scale of the crosstalk problem. It should be noted that the requirement to use a full inductance matrix in the modelling of the packaging impedance was not appreciated at the time of the initial circuit design; the calculations in this chapter were performed after the design had been completed.

Description of SPOEC packaging scheme

The SPOEC circuit was packaged in a 256-pin cavity up multilayer ceramic pin-grid array (PGA) package without internal ground planes. The cavity size was 16.0 mm \times 16.0 mm. The choice of package was constrained by the die size and the requirement for a cavity-up geometry imposed by the optomechanics. The overall dimensions of the carrier were 50.8 mm \times 50.8 mm. The estimated height between the plane of the package bond pads and the printed circuit board is 6 mm. The bond pads were arranged in two tiers of 32 on each side of the cavity.

Analogue power supply pins were located at the top and bottom edges of the package. Each half-column of four super-pixels had two pairs of AVDD/AGND pins, one DETBIAS pin, one pair of DVDD/DGND pins and half a pair of ATHVDD/ATHGND pins. One period of the pattern of on-chip bond pads, which repeated after every two columns of super-pixels, is shown in Figure 8-8. Two pairs of AVDD/AGND supplies were required per super-pixel to achieve an acceptable DC voltage drop across the tracks at the edge of the chip connecting the columns of receivers to the pin. The left- and right-hand sides of the package were used for the digital logic signals and additional digital power supply pins.

The location of the power supply pins were chosen in an attempt to minimise coupling on to the quiet supplies. Where possible, adjacent pads were used for the positive and negative terminals of each supply in order to reduce the loop inductance for differential-mode current and to minimise mutual inductive coupling of differential-mode current onto other pins⁵. The detector bias pad was located as far as possible from the digital supply pairs and surrounded by the two relatively quiet AVDD/AGND supply pairs.



Figure 8-8: Sequence of on-chip bond pads used for the analogue power supplies in the SPOEC system

Description of carrier used for inductance model

The carrier modelled in [283] and assumed in this analysis is a 68-pin ceramic quad-flat-pack (Figure 8-9). The nine package pins located at the centre of one side of the carrier are assigned to one half-column of four super-pixels in the sequence: AVDD, AGND, DETBIAS, AVDD, AGND, ATHVDD, ATHGND, DVDD, DGND. This assignment provides the same number of pins for each supply as in the actual SPOEC system except that a complete pair is used for the decision stages so that, for simplicity, the analysis can be done with one half-column of super-pixels rather than two.

⁵ Because the carrier is a pin-grid array with pins in the four outside rows of the grid, the use of adjacent on-chip bond pads for supply pairs did not always result in adjacent pins in the matrix, but did so more often than not. The ATHVDD and ATHGND on-chip bond pads are not adjacent for this reason.

A detailed mechanical description of this package was available in the form of an input file for the inductance calculation program FastHenry [283][284]; this file was used to obtain the inductance matrix. The inductance of the bond-wires was neglected. The self-inductance per pin ranged from 8.7 nH for a centre pin to 10.7 nH for a corner pin. The coupling coefficients as a function of separation are shown in Table 8-1. For comparison, the self-inductance of a 208-pin PGA, similar in construction to the 256-pin PGA used in the final system, was specified by the manufacturer to be between 4.4 nH and 13.8 nH depending on pin location [285]. The self-inductances of the pins in the 68-pin carrier are thus of the same order as those in the actual carrier used in SPOEC but, because of the completely different geometry of the PGA, it is unclear without detailed modelling how much similarity there will be in the coupling coefficients.

The inductance matrix of the package was calculated without an external ground plane; in a practical situation where the package is mounted on a printed circuit board, the coupling coefficients may be significantly smaller, particularly for distant pins, especially since the lead pitch is comparable to the mounted height of the package. This might affect some of the conclusions about the most important sources of crosstalk; it would therefore be worth repeating this study including the effects of a ground plane.



Figure 8-9: Diagram of 68-pin chip carrier used to estimate a typical inductance matrix (reprinted from Damon [283] with permission). The package has dimensions
25 mm × 25 mm, a cavity of 7.8 mm × 7.8 mm, a lead pitch of 1.27 mm and a mounted height of 2.2 mm.

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1.00 | 0.57 | 0.40 | 0.31 | 0.25 | 0.21 | 0.18 | 0.15 | 0.13 | 0.08 | 0.07 | 0.06 | 0.05 | 0.04 | 0.03 | 0.02 | 0.01 | 0.00 |

 Table 8-1: Table of coupling coefficients for centre pin as a function of separation for the example 68-pin package. Values for other pins were very similar.

Transformed impedance matrix

| L/nH | I_{ADM} | I_{DETBIAS} | $I_{_{TDM}}$ | I_{AT} | $I_{\rm DDM}$ | $I_{_{TD}}$ |
|------------------------------|-----------|----------------------|--------------|----------|---------------|-------------|
| V _{ADM} | -3.779 | 015 | .321 | 603 | .072 | .326 |
| V_{ACM} | 015 | -2.848 | 438 | -1.278 | 181 | 567 |
| $V_{\scriptscriptstyle TDM}$ | .321 | 438 | -7.653 | 900 | .717 | 875 |
| V_{TCM} | .587 | -1.570 | .462 | 4.042 | 936 | -2.682 |
| V_{DDM} | .072 | 181 | .717 | .755 | -8.324 | -1.151 |
| V_{DCM} | .261 | -1.003 | 1.337 | 1.928 | .216 | 3.731 |

The inductance matrix in terms of the transformed voltage and current variables is shown in Table 8-2. V_{detBIAS} is used as the reference node.

Table 8-2: Inductance matrix of 68-pin carrier in terms of transformed I and V variables

The fact that the largest element in each row and column is in the diagonal indicates that, to an extent, it is possible to associate each of the six transformed voltage variables with a different current variable. However, the off-diagonal elements are too large to be ignored because of the fact that some current transients are significantly larger than others. On the whole, the mutual inductance relating to differential-mode supply currents is relatively small as a result of the assignment of positive and negative terminals of each supply to adjacent pins. The off-diagonal elements associated with the common-mode voltages are stronger, implying that currents from a number of physical sources contribute towards the common-mode voltage noise.

The self inductance associated with the differential-mode front-end supply is about half that of the decision stage and digital logic due to the use of four pins instead of two.

8.3.5 Combining on-chip admittance matrix and off-chip impedance matrix

The overall impedance of the power supply network is determined by the combination of the admittance of the on-chip decoupling capacitance and the impedance of the package.

In a single supply system, the on-chip decoupling capacitance forms an LC filter with the loop inductance of the external power supply. The high frequency current flows through the on-chip capacitance and the low-frequency current flows through the external pin.

Similar behaviour is obtained in a system such as this one with multiple power supplies. The admittance matrix of the on-chip network described in Section 8.3.4, written in terms of the transformed current and voltage variables, is diagonal and has only three non-zero elements $Y_{VADM-IADM} Y_{VTDM-ITDM}$ and $Y_{VDDM-IDDM}$ corresponding to the admittance between the three pairs of

power supply terminals⁶. It is shown in Appendix 8.9 that when an admittance Y is added to the diagonal element Z_{kk} of the impedance matrix, then, provided the coupling between the different supplies is weak, the admittance has an effect only on the elements in the row k and/or column k of the impedance matrix and that these are filtered by a transfer function:

$$H(s) = \frac{1}{1 + YZ_{kk}}$$
(8.16)

This result allows the overall impedance matrix to be obtained.

The physical interpretation of this result is that capacitance added between two supply pins:

- reduces the differential-mode high frequency voltage noise being induced on that supply by the differential-mode current transient on the same supply;
- 2. reduces the differential-mode high frequency voltage noise being induced on other supplies by this differential-mode current transient;
- 3. reduces the differential-mode high frequency voltage noise being induced on the decoupled supply by current transients on other supplies.

Thus, for example, to ensure low differential-mode voltage noise on the front-end supply, it is necessary to adequately decouple both the analogue supply itself (to filter the current transient from the front-end supply) and also to decouple the digital supply (to filter the mutually induced voltage from the digital current transient).

Note, however, that decoupling capacitance does nothing to reduce the common-mode voltage noise produced by common-mode currents.

The overall filtering effect of the distributed decoupling capacitance on the differential-mode voltage noise produced by a differential-mode current transient on the same supply is given by:

$$V_{DM} = \frac{sL_{kk}}{1 + \frac{2}{3}sCR + s^2 L_{kk}C} I_{DM}$$
(8.17)

The fact that the decoupling capacitance is distributed underneath a slightly lossy power supply network helps to filter the power supply transient and, in particular, damps the resonance formed by the self-inductance of the package and the decoupling capacitance

⁶ Because of the sign conventions chosen in the definitions of the transformed voltages and currents, the entry Y in the on-chip admittance matrix is negative; the negative sign cancels with the negative sign of the differential-mode diagonal elements of the package inductance matrix in (8.16) so that the coefficients of the powers of s in the numerator and denominator are all positive.

Standard design procedures exist for the sizing of on-chip bypass capacitance to meet a specific differential-mode voltage noise specification [286][287]. A reduction in noise can be obtained if it is possible to add sufficient decoupling capacitance to bring the resonant frequency

$$\omega_0 = \frac{1}{\sqrt{L_{kk}C}} \tag{8.18}$$

below the highest frequency component of the current transient. Approximately, this condition requires that:

$$C > \frac{t_{EDGE}^{2}}{L_{kk}}$$
(8.19)

where t_{EDGE} is a measure of the edge time of the current transient.

If this condition can be satisfied, then the high frequency content is absorbed by the on-chip capacitors and only the low-frequency content passes through the external pin. Following the approach in [286], the on-chip capacitance must then be sized to control the high-frequency ripple and the step response of the LC resonant circuit. The step response of a lightly damped LC resonant circuit to a current transient with an edge time much faster than the time constant of the LC is given by [286]:

$$V = I_{AVG} \sqrt{\frac{L_{kk}}{C}} \sin \omega_0 t \tag{8.20}$$

and thus the voltage noise can be reduced by reducing the pin inductance or increasing the decoupling capacitance. If the condition is not satisfied, then the voltage transient is given by L_{kk} di / dt.

However, the total capacitance required to satisfy (8.19) requires significant layout area. For example, if the edge time of the current transient is 2 ns and the loop inductance is 4 nH then approximately 1 nF is required per pin to give significant benefit. The values used in the SPOEC front-end power supply network give some idea of how much filtering can be achieved in a typical smart-pixel array. Each pixel contained 2 pF of gate-oxide decoupling capacitance occupying a layout area of approximately 1500 μ m² which represents a 70% overhead on the basic layout area of the receiver⁷. This gives a total capacitance of about 0.5 nF per pin which is only just enough to provide a benefit. The cut-off frequency of the RC network is at 220 MHz

 $^{^{7}}$ A channel length of 10.2 µm was used to provide low enough series resistance for effective decoupling up to a frequency of about 600 MHz. The area required by the source and drain contacts of the transistor represent a significant overhead.

which is still slightly above the natural LC frequency of 110 MHz and so the network is slightly underdamped ($\zeta = 0.4$).

8.3.6 Estimation of IR drop

The discussion of the receiver current transients (section 8.3.2) together with the analysis of the overall power supply impedance (section 8.3.5) allow the voltage transient at the edge of the chip to be estimated, giving the noise seen by a receiver at the edge of the chip. However, receivers at the centre of the array will see an additional voltage drop due to the resistance of the on-chip power supply network.

It is shown in Appendix 8.7 that the differential-mode voltage V_{CENTRE} seen by an inner receiver is related to the voltage at the pin V_{PIN} by:

$$V_{CENTRE} = V_{PIN} \frac{1}{1 + sCR} - IR \frac{1}{1 + sCR5/6}$$
(8.21)

Physically, this indicates that the inner receiver sees a low-pass filtered version of the voltage at the pin, tending to reduce the crosstalk, but sees additional voltage noise due to the low-pass filtered resistive voltage drop across the power supply rails. Conservatively, it can be assumed that the worst-case receiver sees both the IR voltage drop and the unfiltered voltage noise at the pin.

The voltage drop between the package pin and the centre of the chip due to the common-mode currents are easily calculated from the receiver equivalent circuit (Figure 8-5) and the model of the on-chip power supply network by considering the current sources one at a time with the other current sources set to zero. Table 8-3 summarises the impedance that relates the voltage drop to the receiver currents.

| | I_{ADM} | $\mathbf{I}_{\mathrm{DETBIAS}}$ | $\mathbf{I}_{_{TDM}}$ | \mathbf{I}_{AT} | I_{DDM} | $I_{\rm TD}$ |
|------------------------------|-----------------------------|---------------------------------|-----------------------------|----------------------------|-----------------------------|-------------------|
| V_{ADM} | $-\frac{R_A}{1+sC_AR_A5/6}$ | - | - | - | - | - |
| $V_{\scriptscriptstyle ACM}$ | - | $-R_{A} / 4$ | - | $-R_{A} / 4$ | - | - |
| $V_{_{TDM}}$ | - | - | $-\frac{R_T}{1+sC_TR_T5/6}$ | - | - | - |
| $V_{\scriptscriptstyle TCM}$ | - | - | - | $+R_{_{\rm T}}/4$ | - | $-R_{_{\rm T}}/4$ |
| V_{DDM} | - | - | - | - | $-\frac{R_D}{1+sC_DR_D5/6}$ | - |
| $V_{\rm DCM}$ | - | - | - | - | - | $+R_{\rm D}/4$ |

 Table 8-3: Impedance relating voltage drop between package pin and centre of chip to receiver currents

8.4 Numerical estimates of supply crosstalk

8.4.1 Introduction

This section applies the general method to estimate the crosstalk for the example configuration of a column of 256 receivers outlined in Section 8.3.3 and relates it to the noise immunity of the receiver circuit.

The transient response of the reduced supply network discussed in sections 8.3.5 and 8.3.6 is evaluated numerically using HSpice. The power supply currents from a single receiver instance connected to ideal power supplies provide the stimulus to the network. The reduced form of the power supply network allowed rapid simulation. Although a numerical approach is used to produce the results in this section, the network is simple enough to allow design calculations based on the equations in section 8.3.5.

The input referred current noise is calculated from the supply voltage transients as part of the simulation using equation (8.6) and by modelling the receiver transfer function as a 2^{nd} order Butterworth filter with a 3 dB frequency of 150 MHz and assuming that the frequency content of the noise is well above the low-frequency zero in equation (8.6).

Estimation of the noise produced by the transient on the digital supply requires more information on the details of the digital supply transient and the digital power supply network. However, it is possible to obtain an approximate estimate of the noise by considering the fact that in order to ensure that the digital circuitry operates correctly, the digital power supply network must be designed to limit the L_{DDM} d i_{DDM} / dt voltage drop across the digital power supply rails to some maximum, say 500 mV. The voltage induced on the other supplies will be in proportion to the ratio of the mutual inductance with the digital supply to the self inductance of the digital supply. Since the current flowing through the external digital power supply pin will be below the cut-off frequency of the digital LC filter, it is assumed that this transient is not further filtered by the decoupling capacitance on other supplies.

8.4.2 Noise immunity of receiver circuit

The amount of noise that a receiver can tolerate is determined by the characteristics of the decision stage. Table 8-4 expresses the switching thresholds of the decision stage in the SPOEC data receiver in terms of input referred photocurrent.

The noise margin of a quiet (non-switching) receiver is set by the DC input voltages required to produce a full-swing output but the actual noise margin of the data link is somewhat worse than this because voltage noise will reduce the effective voltage swing and so introduce jitter in the output edge of the decision stage. This has not been studied in detail; however, simulations

showed that an input photocurrent swing of around $\pm 1 \ \mu$ A about the nominal switching threshold of 2.2 μ A was required to achieve a subjectively acceptable additional delay in relation to the delay produced by a full swing photocurrent input. This gives a noise margin of 1.2 μ A and 1.8 μ A in the logic 0 and logic 1 states respectively before allowing for electrical crosstalk.

| | photocurrent / | voltage at decision stage input |
|------------------------------------|----------------|--------------------------------------|
| | μA | relative to switching threshold / mV |
| nominal photocurrent for logic 0 | 0.0 | -700 |
| max current to avoid jitter in 0-1 | ~ 1.2 | -280 |
| transition | | |
| max current for logic 0 (DC) | 2.0 | -100 |
| nominal switching threshold | 2.2 | 0 |
| min current for logic 1 (DC) | 2.4 | +100 |
| min current to avoid jitter in 1-0 | ~ 3.2 | +240 |
| transition | | |
| nominal photocurrent for logic 1 | 5.0 | +550 |
| gross noise margin (logic 0) | 1.2 | 420 |
| gross noise margin (logic 1) | 1.8 | 310 |
| estimated CM front-end noise | -1.4 | |
| estimated DM front-end noise | -0.1 | |
| estimated CM decision stage | -0.9 / -1.7 | 300 |
| noise (0/1) | | |
| net noise margin (logic 0) | -1.2 | |
| net noise margin (logic 1) | -1.4 | |

Table 8-4: Estimated switching thresholds of SPOEC data receiver (typical process conditions)

Several other sources must be included in the noise budget to ensure a reliable link, including random offset voltage, an allowance for process variation in the switching threshold⁸, optical crosstalk and receiver thermal noise. Table 8-4 is only intended to illustrate the scale of the crosstalk in relation to the receiver sensitivity and does not constitute a full calculation of the noise budget. General guidelines for preparing a noise budget are discussed in [288].

8.4.3 Simulation results

Table 8-5 summarises the estimates of the power supply voltage noise produced by all 256 receivers in a half-column of four super-pixels switching in the same direction simultaneously.

⁸ In the context of the SPOEC data receiver, this is assumed to be zeroed out by control of the external bias voltage to the front-end feedback transistor; in practice, the finite tolerance on the bias voltage and process variation across the chip will mean that there is still a contribution.

| | $\mathbf{I}_{_{ADM}}$ | $\mathbf{I}_{\text{detbias}}$ | $I_{_{TDM}}$ | $\mathbf{I}_{_{\mathrm{AT}}}$ | $\mathbf{I}_{\mathrm{DDM}}$ | $I_{_{TD}}$ |
|---|-----------------------|-------------------------------|--------------|-------------------------------|-----------------------------|-------------|
| peak transient current per rx / μA | 3 | 5 | 160 | 16 | | 120 |
| rise time of transient | 2 | 2 | 2 | 2 | 1 | 1 |
| number of receivers | 256 | 256 | 256 | 256 | 256 | 256 |
| front-end | | | | | | |
| DM voltage / mV | 1.4 | - | 6.1 | 1.8 | ~4 | 5.5 |
| DM IR drop / mV | 1.6 | | | | | |
| CM voltage / mV | - | 2 | 23 | 8 | ~11 | 53 |
| CM IR drop / mV | | 0.7 | | 2.2 | | |
| decision-stage | | | | | | |
| DM voltage / mV | | | 260 | | ~40 | 60 |
| DM IR drop / mV | | | 270 | | | |
| CM voltage / mV | | | 20 | 20 | ~60 | 200 |
| CM IR drop / mV | | | | 10 | | 50 |

Table 8-5: Estimates of power supply voltage noise at chip-edge

The supply network parameters used in the simulation were as follows: the resistance of each of the 8 columns of 32 receivers was 17 Ω for each front-end supply rail and 53 Ω for each decision stage supply rail. A decoupling capacitance of 2 pF per receiver for the front-end supply was used. No decoupling capacitance was used for the decision stage. This corresponds approximately to the parameters of the actual SPOEC circuit.

The remainder of this section discusses the results in this table for the front-end and the decision stage supplies.

Front-end

Figure 8-10 shows the estimated voltage transients on the front-end supply at the edge of the chip; Figure 8-11 shows the equivalent input current due to the noise on the front-end supply.

The differential-mode voltage noise on the front-end power supplies is well controlled by the decoupling capacitance, leaving the common-mode voltage noise as the main problem for circuit operation.

The differential-mode voltage noise is dominated by the step response of the LC resonant circuit. The main contributions come from the decision stage and digital supply transients, but the total input referred current noise produced by the differential-mode noise in this case is only 0.1 μ A. The amplitude of the step response is approximately consistent with calculations based on equation (8.20).



Figure 8-10: Contributions to the voltage transient on the front-end supply at the package pin (upper trace: differential-mode; lower-trace: common-mode)



Figure 8-11: Comparison of total input referred current noise produced by commonmode (upper trace) and differential-mode (lower trace) voltage transient on front-end supply for an outer and inner receiver (excluding digital differential-mode current transient)

The input referred current noise produced by the common-mode voltage on the front-end power supply is estimated to be 1.4 μ A, neglecting the contribution from the digital supply (which, assuming an edge time of 1 ns for the induced voltage transient, might add another 1 μ A). This value is enough in itself to just exceed the noise margin of the receiver circuit. The main contributions are again due to mutual inductive coupling from the decision stage differential-mode current I_{TDM} and the current I_{TD} that charges up the input capacitance of the first digital gate.

The additional noise produced by the resistive voltage drop across the front-end power supplies is relatively small in this case.
Decision stage

Figure 8-12 shows the estimated voltage transients on the decision stage supply at the edge of the chip. Although the magnitude of the common-mode and differential-mode voltage noise on the decision stage supplies is comparable, common-mode noise presents a more serious problem to the operation of the decision stage since it appears unattenuated at the input of the stage whereas the rejection of differential-mode noise is quite high.

The dominant source of common-mode voltage noise is the current I_{TD} that charges up the input capacitance of the first digital gate. The total common-mode noise when all receivers switch in the same direction is about 300 mV which produces an equivalent input photocurrent of 1.7 μ A and 0.9 μ A in the high and low states respectively. This is again comparable with the noise margin of the receiver.



Figure 8-12: Contributions to the common-mode (upper-trace) and differential-mode (lower-trace) voltage noise on the decision stage supply



Figure 8-13: Comparison of overall common-mode (upper trace) and differential-mode (lower-trace) voltage noise on decision stage power supply for outer and inner receivers

The main contribution to the differential-mode voltage noise on the decision stage power supply is the differential-mode current I_{TDM} and its main effect will be to introduce some jitter.

Receivers in the interior of the array see a slightly reduced power supply voltage as a result of the IR voltage drop along the power supply (Figure 8-13); this will result in delay variation across the array. The symmetry of the power and ground power supply rails means that the resistive drop does not produce any shift in the common-mode voltage, although in practice some asymmetry is inevitable and an allowance for a small low-frequency common-mode voltage shift should be made in the noise budget.

8.5 Discussion

The results of this case study indicate that, for the particular combination of receiver design and packaging scheme assumed in the calculations, electrical crosstalk is a serious problem. Crosstalk is predicted to prevent simultaneous error-free operation of the entire array. The single-beam nature of the design means that simply increasing the optical power will not solve the problem; the noise margin in the low state is unaffected by the optical power level. Nevertheless, simultaneous operation of a significant fraction of the entire array (perhaps about

20-50%) might be expected and, for practical data patterns in which a mix of transitions in different directions occur, the crosstalk would be expected to be lower.

This calculation does not provide a reliable quantitative prediction of the crosstalk in the SPOEC system because of the differences between the packaging scheme used in the calculation and that used in the actual circuit. The calculation also relies on a number of simplifying assumptions that have not been quantitatively justified and must be treated with a degree of caution. In particular, one of the components that is responsible for much of the noise, the common-mode current I_{TD} between the decision stage and the digital logic, might be significantly influenced by the finite substrate admittance between DGND and ATHGND which has not been included in the calculation.

Nevertheless, the calculation provides definite evidence that electrical crosstalk is not sufficiently small to be ignored in the design of large receiver arrays and must be considered as an integral part of the design of any future system.

In the remainder of this section, alternative design approaches that might be able to achieve robust operation of an array of this scale are discussed. Two basic approaches can be contemplated: reducing the amount of noise generated by the receiver or improving its immunity to noise.

Noise reduction could be achieved by adjusting the design of the circuit or by altering the packaging scheme. In terms of altering the circuit, it is best to focus in the first instance on how to reduce the largest sources of noise which are the differential mode current in the decision stage I_{TDM} and the current I_{TD} that charges up the input capacitance of the first digital gate. The first component could be reduced by the addition of separate decoupling capacitance to the decision stage supply. The penalty for this is a further increase in layout area; this is the main reason why none was included in the SPOEC circuit. There is limited further scope for reducing the second component. The receiver design used in this study took specific measures to minimise I_{TD} : long-channel transistors were used in the output inverter of the decision stage to keep the edge time no faster than necessary to support the target bit-rate, and the input capacitance of the digital stage was minimised by using a small buffer inverter. Differential signalling between the output of the decision stage and the input to the digital logic would result in first order cancellation of the transient.

Significantly improved packaging is in principle possible but is practically difficult to achieve using off-the-shelf packages in prototype volumes. The number of pins required to implement the pin assignment used in this study in an array of 4096 receivers is already at the practical limit (8 columns of 9 pins needs 72 pins per side or 288 pins in total). Addition of a ground plane to the package would help to reduce the mutual inductive coupling; it may even be

possible to retrofit this to the existing packaging scheme by attaching a metal sheet on the top surface of the carrier. The usual means of achieving a low inductance power supply by including both power and ground packages and integrated decoupling capacitors on the carrier is complicated by the fact that both digital and analogue supplies must be accommodated. A custom carrier would address this difficulty but would also be expensive. Whatever packaging approach is adopted, full simulation of the package impedance is necessary.

It is not obvious whether these changes are sufficient to make practical the design of reliable large arrays based on simple single-ended receivers of the type used in the SPOEC system. Detailed and accurate simulation of noise coupling would be required to guarantee a correct circuit in a single fabrication iteration. The simple two-beam receivers of the type discussed in chapter 4 would be one step better because a higher than expected level of crosstalk could always be overcome by increasing the optical power level. Taking advantage of the improvements in silicon process technology discussed in chapter 6 to implement a more complex receiver design with a fundamentally higher noise immunity would seem to be a more attractive approach to take.

A solution must address the sensitivity of both the front-end and the decision stage to commonmode voltage shift relative to the detector bias.

The sensitivity of the front-end can be overcome by connecting the photodiode bias voltage directly to the front-end power supply. All front-end supply noise is then differential-mode and can be controlled by decoupling capacitance. Distributing the connection throughout the receiver arrays using local flip-chip connections to the photodiode bias, rather than locating the bias connections at the edge of the chip as in the SPOEC system, would be preferable to eliminate common-mode voltage noise due to the impedance of the on-chip power network. One problem with the use of a direct connection is the low reverse bias voltage on the photodiode that this scheme would provide in future generation low-voltage CMOS technology; this might reduce the response speed of the photodiode. Instead, the photodiode bias connection could be decoupled to the analogue supply throughout the chip. A second approach to reducing the effect of common-mode voltage noise is to use a two-beam, electrically differential front-end such as that used in the SPOEC clock receiver; in this configuration, common-mode voltage noise on the analogue supply translates predominantly into a common-mode output voltage which could be rejected by a fully-differential post-amplifier.

The sensitivity to common-mode voltage noise between the decision stage and the front-end can also be overcome by using differential signalling between the output of the post-amplifier and the input of the decision stage.

It is clear from this discussion that a fully electrically differential, two-beam receiver would offer a significant improvement in crosstalk performance and is recommended in future systems in the absence of a good reason to adopt a single-beam approach.

However, implementations of single-beam receivers with better immunity to crosstalk are possible. Figure 8-14 illustrates a possible configuration that combines some of the techniques discussed above. A similar circuit is described in [289] but without reference to its crosstalk properties.



Figure 8-14: An alternative single-beam receiver configuration with improved supply noise immunity

The most important feature of this configuration is that the front-end uses only a single type of transistor for the transconductance element and uses a photodiode with its bias connected directly to the same supply. The receiver is then only sensitive to differential-mode noise between AVDD and AGND which can be decoupled using on-chip capacitance. Capacitance is also required between the local ground and the current-source bias to reduce the high-frequency impedance of this node to ground. The transconductance of the front-end supply noise generator consists of only the output conductance of the NMOS bias transistor which is relatively small.

Note that to implement this scheme with a photodiode array using an n-type shared contact, it is necessary to use a PMOS based inverter for the front-end, which will lead to a slightly reduced switching energy compared to a design based on an NMOS inverter because of the larger feedback capacitance for the same bit-rate B_0 (see the discussion in chapter 4).

This front-end design would in itself allow a two-beam, electrically differential receiver circuit with good crosstalk immunity to be implemented. However, it is possible to adapt it to a single-beam receiver circuit as shown in Figure 8-14 by generating a reference voltage using a dummy

photoreceiver circuit shared amongst a small number of receivers. It might be desirable to include a dummy photodiode at the input of this reference circuit to match the high-frequency output impedance of the reference circuit with the active receivers so that common-mode noise current into the output of both the reference and an active receiver is mostly translated into a common-mode output voltage⁹ and is therefore rejected by the differential post-amplifier. The fixed decision threshold required in a single-ended receiver could be introduced in the post-amplifier or by using slightly different transistor dimensions in the reference front-end.

This scheme has not been analysed in any detail and may not be the best way to approach the problem, but is at least a starting point for future designs. It appears to suggest that crosstalk is not an insurmountable problem provided it is taken fully into account from the start of the design process.

8.6 Conclusions and further work

This chapter has presented a method for the analysis of crosstalk between receiver circuits in two-dimensional arrays arising from the finite impedance of the power supply network. The analysis depends on the simplifying assumption that, within a receiver cell, the different power supply connections are electrically isolated with the exception of decoupling capacitance between the power/ ground pair of a given supply. The method allows the effects of the current transients arising from different physical sources to be considered separately by transforming the receiver terminal currents and voltages into differential- and common-mode equivalents. The distributed nature of the on-chip power supply network and a full model of the package impedance are included in the analysis.

A case study applying the method to a receiver array, loosely based on the SPOEC switching chip, shows that power supply crosstalk is sufficiently important in large receiver arrays that it cannot be ignored in the design process. It also provides support for the explanation by others [276] of experimental degradation of receiver performance, during simultaneous operation of large arrays, in terms of power supply crosstalk. Detailed modelling of packaging and budgeting for electrical crosstalk is believed to be essential in the design of future systems.

Common-mode voltage noise between the detector bias, front-end supply and decision stage supply appears to be the most serious problem. Some alternative design approaches which address the largest sources of noise are discussed. In particular, direct connection of the

⁹ The load capacitance of the reference circuit is not matched to the active channels because it is shared amongst several channels and so the cancellation is not complete; however, the photodiode capacitance is the main contribution in the zero of the front-end output impedance.

photodiode bias to the front-end supply and the use of electrically differential design techniques appear, at first sight, to offer the best solutions to the crosstalk problem.

Further work is required to provide a complete analysis of crosstalk. Specifically, the validity of the assumption of isolated supplies requires investigation, in particular in relation to the admittance introduced by a common substrate. An analysis of jitter and some validation of the predictions of the simplified model against more detailed simulations and experimental results is desirable.

Although the case study is not expected to provide accurate quantitative predictions of crosstalk in the SPOEC system, it does suggest that crosstalk will prevent simultaneous error-free operation of the entire switching chip. Nevertheless, experimental results on electrical crosstalk from the system will provide a limited test of the theoretical predictions against experiment. The analysis presented in this chapter was not fully developed until after the system design had been completed, and only some of the results could be fed into the final circuit design. Although the system may not achieve its nominal objective of constructing a fully operational 64×64 crossbar, it can be expected to be a success in terms of its primary objective of investigating and improving the understanding of the design issues in large optoelectronic VLSI systems.

8.7 Appendix: Norton equivalent of a distributed line with a distributed current source

In this appendix, the Norton equivalent for a distributed line, which extends from the inner end at x=0 to the outer end at x=L and is open-circuit at the inner end, with impedance z per unit length, admittance y per unit length and load current j per unit length is derived. An expression relating the voltage at the inner end to the voltage at the outer end is also established.

Consider the section of the line between x- Δx and x+ Δx (Figure 8-15).



Figure 8-15: A section of a distributed transmission line

Apply Kirchhoff's Current Law to the node at x.

$$\frac{V(x-\Delta x)-V(x)}{z\Delta x} + \frac{V(x+\Delta x)-V(x)}{z\Delta x} - V(x)y\Delta x - j\Delta x = 0$$
(8.22)

Expanding V(x $\pm \Delta x$) as a Taylor series about x and taking the limit as $\Delta x \rightarrow 0$ gives the differential equation:

$$V''(x) - V(x)yz = jz$$
(8.23)

Define the characteristic admittance y_0 and the propagation constant γ of the line as:

$$y_0 = \sqrt{\frac{y}{z}}; \ \gamma = \sqrt{yz}$$
 (8.24)

The general solution of the differential equation is:

$$V(x) = C_1 \cosh \gamma x + C_2 \sinh \gamma x - \frac{j}{y}$$
(8.25)

where C_1 and C_2 are determined by boundary conditions at the ends of the line at x=0 and x=L. If the boundary condition is specified in terms of a current I_{IN} into the line at x = L then:

$$V'(L) = I_{IN} z \tag{8.26}$$

To derive the Norton equivalent current source, calculate the current into a short-circuit load at x = L with the line open circuit at x = 0. The boundary conditions are:

$$V'(0) = 0; V(L) = 0$$
 (8.27)

from which the particular solution

$$V(x) = \frac{j}{\gamma} \left(\frac{\cosh \gamma x}{\cosh \gamma L} - 1\right)$$
(8.28)

is derived. The current at x = L is then:

$$I_{IN} = \frac{j}{\gamma} \tanh \gamma \ L \tag{8.29}$$

The one-port admittance is derived in a similar fashion with the current source j set to zero to give:

$$Y_{IN} = y_0 \tanh \gamma \, L \tag{8.30}$$

In the particular case of a distributed RC line where y = sc and z=r where r and c are the resistance and capacitance per unit length, the previous two equations become:

$$I_{INRC} = \frac{j}{\sqrt{scr}} \tanh \sqrt{scr} L \approx I \frac{1}{1 + sCR/3}$$

$$Y_{INRC} = \sqrt{\frac{sc}{r}} \tanh \sqrt{scr} L \approx \frac{sC}{1 + sCR/3}$$
(8.31)

where R, C and I are the total resistance, capacitance and current of the line. A Padé expansion [290] of the two expressions has been used to match the full expression to a single-pole low-pass filter and a series combination of a resistor and a capacitor. In this approximation, the time constant of the filter is CR / 3 and the input admittance consists of a capacitor C in series with a resistor R / 3. The Elmore rise time [291] of the network is 1.0 RC.

The voltage at the inner end is obtained from the voltage at the outer end by solving with the boundary conditions:

$$V'(0) = 0; V(L) = V$$
 (8.32)

which gives:

$$V(x) = (V + \frac{j}{y})\frac{\cosh\gamma x}{\cosh\gamma L} - \frac{j}{y}$$
(8.33)

The voltage at the inner end of the line at x = 0 is then:

$$V(0) = (V + \frac{j}{y})\operatorname{sech} \gamma \operatorname{L} - \frac{j}{y}$$
(8.34)

In the case of an RC line, the far-end voltage is given approximately by:

$$V(0) = V \frac{1}{1 + sCR/2} - \frac{IR}{2} \frac{1}{1 + sCR5/12}$$
(8.35)

where the terms proportional to V and I have been separately matched to a first order low-pass filter function using a Padé expansion.

Equations (8.31) and (8.35) have been derived for an asymmetric unbalanced transmission line in which the impedance of the ground return is zero. The balanced transmission line of interest in analysing the power supply network (Figure 8-16) can be reduced to an unbalanced case by applying symmetry arguments; the same equations can be used with R replaced with 2 $R_{PER-WIRE}$ where $R_{PER-WIRE}$ is the total series resistance of one wire of the balanced transmission line.



Figure 8-16: First-order Norton equivalent of a balanced transmission line with a series resistance of R per wire. H(s) is the transfer function given by (8.31).

8.8 Appendix: calculation of package impedance matrix in terms of transformed current and voltage variables

This appendix describes the calculation of the transformed impedance matrix (which relates the transformed differential- and common-mode voltage variables to the transformed current variables I_{ADM} , I_{TDM} , I_{DDM} , $I_{DETBIAS}$, I_{AT} and I_{TD}) from the original package impedance matrix produced by numerical simulation or experimental measurement (which relates the original pin voltages and pin currents).

The starting point is the inductance matrix for an entire package. In this case, this is calculated numerically using the FastHenry program [279] and consists of a 35×35 matrix for one half of the package. The nine package pins located at the centre of one side of the carrier are considered and are assigned in the sequence AVDD, AGND, DETBIAS, AVDD, AGND, ATHVDD, ATHGND, DVDD, DGND as discussed in Section 8.3.4. All other pins are assumed to be open circuit; thus the 9×9 matrix relating the voltages and currents of these nine pins in isolation consists simply of the elements of the original matrix associated with these pins.

The first step is to reduce the 9×9 matrix to a 7×7 matrix by making the two AVDD and two AGND pins electrically equivalent. The procedure for doing this is discussed in terms of making pins 1 and 2 electrically equivalent; the generalisation is obvious.

The goal is to rewrite the matrix in terms of the total current through the two pins $I_1' = I_1 + I_2$ and a second independent current variable $I_2' = I_1 - I_2$. The terminal voltages are given by:

$$V_{i} = \sum_{j=1}^{N} Z_{ij} I_{j} = Z_{i1} I_{1} + Z_{i2} I_{2} + \sum_{j=3}^{N} Z_{ij} I_{j}$$

$$= Z'_{i1} (I_{1} + I_{2}) + Z'_{i2} (I_{1} - I_{2}) + \sum_{j=3}^{N} Z'_{ij} I_{j}$$
(8.36)

By equating these two expressions we can relate the elements of the new matrix Z_{i1} and Z_{i2} to the elements of the original matrix Z_{i1} and Z_{i2} by:

$$Z'_{i1} = \frac{Z_{i1} + Z_{i2}}{2}; Z'_{i2} = \frac{Z_{i1} - Z_{i2}}{2}; Z'_{ij} = Z_{ij}$$
(8.37)

The voltages at terminals 1 and 2 are given by:

$$V_1 = \sum_{j=1}^{N} Z'_{1j} I'_j$$
(8.38)

$$V_2 = \sum_{j=1}^{N} Z'_{2j} I'_j$$
(8.39)

The terminals are electrically equivalent and so $V_1 = V_2$. Subtracting equation (8.38) from (8.39) gives:

$$0 = \sum_{j=1}^{N} (Z'_{2j} - Z'_{1j})I'_{j}$$
(8.40)

The variable I_2' can then be eliminated using a pivoting step:

$$V_{i} = \sum_{j=1}^{N} Z'_{ij} I'_{j}$$

= $\sum_{j=1}^{N} Z'_{ij} I'_{j} - \frac{Z'_{ij}}{(Z'_{2j} - Z'_{1j})} \sum_{j=1}^{N} (Z'_{2j} - Z'_{1j}) I'_{j}$ (8.41)
= $\sum_{j=1...N, j \neq 2} I'_{j} Z''_{ij}$

This technique can be applied to obtain a new impedance matrix $[Z''_{ij}]$ that relates the voltages at each supply to the total current through each supply. The terminal V_{DETBIAS} , which is used as a reference node, can also be eliminated using a pivoting step.

The next step is to transform this matrix into new current and voltage variables.

To do this, express the transformed currents and voltages as an arbitrary linear combination of the original current and voltage variables using matrix notation.

$$\mathbf{I}' = \mathbf{C}_{\mathbf{I}}\mathbf{I}$$

$$\mathbf{V}' = \mathbf{C}_{\mathbf{V}}\mathbf{V}$$
(8.42)

where C_{I} and C_{v} are matrices of coefficients.

The matrix equation relating the original variables

$$\mathbf{V} = \mathbf{Z}\mathbf{I} \tag{8.43}$$

can be rewritten as

$$\mathbf{V}' = (\mathbf{C}_{\mathbf{V}} \mathbf{L} \mathbf{C}_{\mathbf{I}}^{-1}) \mathbf{I}'$$
(8.44)

and thus the new impedance matrix is:

$$\mathbf{Z}' = \mathbf{C}_{\mathbf{V}} \mathbf{L} \mathbf{C}_{\mathbf{I}}^{-1} \tag{8.45}$$

8.9 Appendix: effect of a single admittance term on an impedance matrix Theorem

Suppose an admittance term Y is added to the diagonal element Y_{kk} of the admittance matrix $[Y_{ji}]$ that is derived from an impedance matrix $[Z_{ij}]$ by calculating the inverse where the voltages are indexed by subscript i and the currents are indexed by subscript j. Then the new impedance matrix $[Z_{ij}']$ can be obtained from the original impedance matrix as follows:

• For entries in row k and/or column k, the new element Z_{ij} is given by:

$$Z'_{ij} = \frac{Z_{ij}}{1 + YZ_{kk}}$$
(8.46)

• For entries not in row k and column k, the new element is given by:

$$Z'_{ij} = Z_{ij} \left(1 - \frac{Y Z_{ik} Z_{kj}}{(Y Z_{kk} + 1) Z_{ij}} \right)$$
(8.47)

Proof

The result is proven for diagonal element Y_{11} . The equation obtained from the first row of the admittance matrix is:

$$I_1 = (Y + Y_{11}) V_1 + Y_{12}V_2 + \dots + Y_{1N}V_N$$
(8.48)

which rearranges to:

$$I_1 - YV_1 = Y_{11}V_1 + Y_{12}V_2 + \dots + Y_{1N}V_N$$
(8.49)

Voltage V_1 can be expressed in terms of the elements of the first row of the new impedance matrix $[Z'_{ij}]$:

$$V_1 = \sum_{j=1}^{N} Z'_{1j} I_j$$
(8.50)

Substitute into equation (8.49):

$$\begin{bmatrix} I_1 - Y \sum_{j=1}^N Z'_{1j} I_j \\ \vdots \\ I_j \\ \vdots \\ I_N \end{bmatrix} = \begin{bmatrix} Y_{ji} \end{bmatrix} \begin{bmatrix} V_i \end{bmatrix}$$
(8.51)

Premultiplying by $[Y_{ij}]^{-1}$ which is equal to $[Z_{ij}]$ gives:

$$\begin{bmatrix} V_i \end{bmatrix} = \begin{bmatrix} Z_{ij} \end{bmatrix} \begin{bmatrix} I_1 - Y \sum_{j=1}^{N} Z'_{1j} I_j \\ \vdots \\ I_j \\ \vdots \\ I_N \end{bmatrix}$$

$$= \begin{bmatrix} Z_{ij} \end{bmatrix} \begin{bmatrix} I - \begin{bmatrix} Y Z'_{11} \cdots Y Z'_{1j} \cdots Y Z'_{1N} \\ 0 & 0 & 0 \cdots & 0 \\ 0 & \ddots \\ \vdots & \ddots \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_1 \\ \vdots \\ I_j \\ \vdots \\ I_N \end{bmatrix}$$
(8.52)

where **I** is the identity matrix. The elements of the new impedance matrix can be obtained by matrix multiplication of the first two square matrices:

$$Z'_{ij} = Z_{ij} - Z_{i1} Y Z'_{1j}$$
(8.53)

Now consider the elements of the first row of $[Z'_{ij}]$.

$$Z'_{1j} = Z_{1j} - Z_{11} Y Z'_{1j}$$
(8.54)

$$Z'_{1j}(1+YZ_{11}) = Z_{1j}$$
(8.55)

$$Z_{1j}' = \frac{Z_{1j}}{1 + YZ_{11}} \tag{8.56}$$

Substitution of (8.56) back into (8.53) gives the desired result after applying some further simplification in the special case of the first column.

The more general case of Y_{kk} can be reduced to this case by exchanging row k and column k with row 1 and column 1 in the matrix and exchanging the labels on V_k/V_1 and I_k/I_1 .

In the specific case where the admittance is capacitive and the impedance is inductive, then equation (8.46) forms an LC low-pass filter. Equation (8.47) becomes:

$$\frac{Z'_{ij}}{Z_{ij}} = \frac{1 + s^2 L_{kk} C (1 - \frac{K_{ik} K_{kj}}{K_{ij}})}{1 + s^2 L_{kk} C}$$
(8.57)

and thus, if the coupling from the decoupled supply to the two supplies whose mutual- or selfinductance is of interest is weak then the term in the impedance matrix is almost unchanged. This is physically sensible.

Chapter 9

Conclusions

9.1 Introduction

This work has considered the problem of photoreceiver design in optoelectronic-VLSI circuits. The work carried out has significance in two areas: firstly, it has made a major contribution to the design of a prototype experimental system, employing CMOS-InGaAs MQW technology, conceived as a vehicle for demonstrating the feasibility of terabit/s scale optical interfaces to VLSI circuits. Secondly, it included a detailed investigation of the key issues in designing receivers for this application, building on the experience gained in the design of the prototype system.

9.2 Design of a prototype terabit/s scale system

An overview of the prototype system has been presented. The system has been designed as part of a collaborative research project and has significant contributions from several research groups. Particular emphasis has been placed on areas in which the author was directly involved: the high-level system architecture and the detailed design of the two photoreceiver circuits required by the system.

Although experimental tests on the prototype system are yet to be performed, the completion of a full system design has already highlighted a number of important issues in optoelectronic-VLSI systems: the importance of adequate power-supply distribution, the closely coupled nature of the optical and electronic design in a full-custom optoelectronic system and the constraints on the physical design of the digital logic imposed by the regularity of the smart-pixel layout. Arguably, these issues could only have been identified by carrying out the full design process on a system of a realistic scale.

Successful operation of the system, if achieved, would obviously be of much greater significance as very few systems of this scale have been tested in the laboratory to date. Preliminary tests on individual components suggest that there is a good chance that it will at least be possible to perform a detailed experimental characterisation of the system as a result of the design work. However, it seems that power supply crosstalk may prevent error-free operation, and it is possible that problems as yet unknown may be highlighted in the course of the experimental test. Nonetheless, irrespective of whether successful operation is achieved, the results obtained will make it easier to construct systems of this type in the future.

As a first step towards achieving experimental operation of the system, results from prototype receiver circuits, designed to use electrical inputs, have been reported. Although high-

frequency testing of the circuits was limited by the test circuitry, the results verify that both receiver designs meet the DC sensitivity requirements of the system.

One of the receiver designs used in this system applies the transconductance-transimpedance circuit technique to smart-pixel receiver post-amplifiers for the first time. A more detailed study of this design technique has shown that the gain-bandwidth advantage of this topology can provide improved sensitivity in high-speed smart-pixel receiver circuits at some cost in power consumption and layout area.

9.3 Investigation of receiver design issues

A detailed study of the design trade-offs in simple smart-pixel receiver circuits in a 0.6 µm technology has been made. Although the low-parasitic hybrid integration technologies used by optoelectronic-VLSI circuits allow low-noise receiver front-ends to be designed with low power-consumption, other factors peculiar to the smart-pixel environment limit performance. In particular, a quantitative analysis has shown that, in many cases, the DC offsets introduced by transistor mismatch limit the minimum detectable signal, although in high-speed circuits in this technology the post-amplifier gain available with reasonable power consumption can also be important.

However, the study made of the impact of anticipated improvements in CMOS technology on receiver performance has shown that power consumption becomes less of a problem in advanced CMOS processes. The study suggests that the most convenient way for receiver performance to evolve is if the bit-rate per channel scales in line with the technology. It predicts that by the 0.1 µm technology generation it should be possible to implement a 1 Tbit/s optical interface with 256 optical channels running at 4 Gbit/s with a power dissipation of about 0.3 W, although some modifications to receiver structure may be required to achieve operation with low-voltage power supplies. These performance improvements are possible even if there is no reduction in photodiode diameter from the value used in prototype smart-pixel systems today. Only a limited improvement in switching energy, to around 5-10 fJ, can be expected in receivers with a response down to DC because of transistor mismatch. However, the increase in capacitance density in advanced CMOS technology will make large arrays of receivers with a low-frequency cut-off a practical possibility. In principle, this would allow noise-limited switching energies to be obtained.

The most significant outstanding problem for smart-pixel receivers is power supply crosstalk. A technique for analysing this effect in large receiver arrays that accounts for the impedance of the chip package and a distributed on-chip power supply network has been presented. An example calculation using this technique has shown that crosstalk is a major issue in the design of large receiver arrays and must be taken into account from the outset in the design of future

systems. The simple, electrically single-ended receivers that have been used in prototype optoelectronic-VLSI systems to date are particularly susceptible to power-supply crosstalk. It seems likely that more complex, electrically differential designs will be required to achieve robust operation of large arrays. However, the improvements in transistor performance outlined above will provide the flexibility to implement these more complex designs.

It is interesting that, as smart-pixel receiver designs become more complex in order to combat the problems of DC offsets, power supply crosstalk and low-voltage operation, the differences in circuit structure between smart-pixel receivers and conventional telecommunications receivers may become less distinct.

9.4 Future work

Topics in which further work is required can be identified in the two areas in which the work has made contributions: demonstration of prototype optoelectronic-VLSI systems and receiver design.

Ultimately, it will be necessary to apply the technology to a real-world problem to show that it can deliver an overall performance improvement over electronic solutions, but the cost of constructing and fully populating such a system suggests that this will only be possible with industrial investment. Future research efforts need to identify and address the perceived concerns of industry about the technology.

Since crosstalk is likely to prevent complete error-free operation of the SPOEC system, further demonstration of the feasibility of terabit/s scale optical interfaces in research labs is arguably required before industry will be prepared to make the investment in the final development work required to introduce the technology into commercial products. There is a case for designing an optoelectronic-VLSI circuit, with the specific purpose of providing a comprehensive verification that long-term error-free operation of optical interfaces of this scale is possible, that takes on board the lessons learnt in the design of the SPOEC system.

In the more specific area of receiver design, there is scope for a more detailed study of the problem of power supply crosstalk that removes some of the simplifying assumptions made in the analysis presented in this work.

The 'photonics-interface-module' approach to implementing high-bandwidth optical interfaces, discussed in Chapter 3, deserves further investigation because of its potential to provide a reusable design block that hides the detail of the optical interface from the electronic designer, thus allowing the optical and electronic design to proceed independently. In particular, work is required to determine whether or not the approach scales to capacities of several terabits/s.

An investigation of techniques for achieving low-voltage operation of receiver circuits would be worthwhile and, since low-voltage operation will present problems for all classes of analogue circuits in future generation CMOS technology, would be of wider interest to the electronic design community.

9.5 Closing remarks

Hybrid optoelectronic-VLSI technology has the potential to deliver vast bandwidths to integrated circuits that satisfy the interconnect requirements of digital systems for the foreseeable future. The question of whether it can do so more economically than purely electronic alternatives is, however, still open. Two-dimensional free-space optical interconnects in particular have the potential to support bandwidths of several terabit/s to single integrated circuits. By exploring the electronic design issues in a prototype free-space VLSI system, this work has shown that, while certain concerns in the area of receiver design require further attention, the performance of the electronic interface circuits does not present an obstacle to optoelectronic-VLSI technology fulfilling its potential. It is likely to be other factors, such as the cost of the optomechanical packaging, that will determine the degree to which optoelectronic-VLSI achieves commercial success.

Appendix A

Detailed simulations of SPOEC data receiver

A.1 Introduction

This appendix contains detailed simulation results for the SPOEC data receiver circuit which are intended to document the expected performance of the circuit.

All simulations were performed in HSpice using BSIM3v2 transistor models.

The multiple quantum well modulators in the system have been designed for operation at a temperature of 50°C and will be controlled by a Peltier element. This was used as the nominal temperature in all simulations.

The final optical design of the system requires a photodiode diameter of 35 μ m which gives a photodiode capacitance of 95 fF based on a MQW thickness of 1.16 μ m and a dielectric constant of 13. This value was used in all simulations.

| Parameter | nominal value |
|------------------------|---------------|
| power supply voltage | 5 V |
| temperature | 50°C |
| photodiode capacitance | 95 fF |
| input photocurrent | 5 μΑ |
| feedback bias voltage | 4.0V |
| | |

All simulations include the parasitic routing capacitance extracted by the layout software.

Table A-1: Nominal parameters used in simulations



Figure A-1: Front-end/post-amplifier/decision output voltage vs. input photocurrent plotted for feedback resistor bias voltages between 3.5 V and 5.0 V..

Notice the relatively sharp transition in the decision stage output and how the switching threshold can be adjusted by varying the bias voltage.

A.2.1 Effect of offset on DC transfer characteristic



Figure A-2: Effect of offset on DC transfer characteristic

The graph shows the spread in performance which can be expected across the chip. Most of the receivers will fall within the $\pm 2\sigma$ range (inner curves); the worst case receivers can be expected around $\pm 4.7\sigma$ (outer curves). Note, however, that, as discussed in the main text, the estimate of σ is not very accurate.

A.2.2 Process sensitivity

The set of graphs contained in this section illustrate the sensitivity of the design to process variation. The performance of the circuit is plotted for the foundry process corners tm (typical mean), wp (worst power=fast n/fast p), ws (worst-speed=slow n/slow p), wz (worst-zero=slow n/fast p) and wo (worst-one=fast n/slow p).

Notice from Figure A-4 and Figure A-5 that the use of a fixed bias voltage (such as the power supply voltage) would lead to a very poor tolerance to process variation. This sensitivity can be attributed to the fact that the operating point of the inverter is not well controlled.

Figure A-4 illustrates that, independent of process, it should be possible to operate the circuit with a bandwidth in the region 120 MHz to 150 MHz with a DC sensitivity of between 1 μ A and 3 μ A by an appropriate choice of tuning voltage, which is consistent with the target specification of the system.

Another point to note is that although it will be possible to set the external bias voltage precisely, the effective value of the bias voltage will vary across the chip due to the shift in the inverter operating point caused by the voltage drop across the analogue power supply rails. Allowance must be made for perhaps a ± 100 mV tolerance on the effective bias voltage.

These graphs do not account for power supply or temperature variation. Fine tuning the power supply voltage is another method that can be used to tweak the performance of the circuit.

The DC sensitivity in these graphs is defined as the current required to produce a voltage at the output of the decision stage equal to 50% of the power supply voltage. This is a reasonable measure because of the narrowness of the transition region of the DC transfer characteristic.





Figure A-3: Speed/sensitivity tradeoff -- process sweep



Figure A-4: Bandwidth vs. bias voltage -- process sweep



Figure A-5: DC sensitivity vs. bias voltage -- process sweep

A.3 Small signal analysis



Figure A-6: Small signal transfer function (front-end and post-amplifier)

Simulation shows the typical AC transfer function of the front-end. Notice that there is some peaking in the amplitude response (which translates into overshoot in the step response). The front-end transimpedance is $67.9 \text{ k}\Omega$ and the small-signal gain of the second stage is 5.2.

The small signal analysis was performed with a DC input photocurrent of 0 μ A which gives the minimum small signal bandwidth.

A.4 Transient response

A.4.1 Step response

The step response of the system provides a more useful measure of the dynamic performance of the circuit because it allows for large signal effects and overshoot. Figure A-7 illustrates a typical step response.



| Figure A-7 | : Typical step | response of front-end |
|------------|----------------|-----------------------|
|------------|----------------|-----------------------|

| front end output | value | notes |
|-----------------------|--------|------------------------|
| rise time | 2.0 ns | 10% / 90% |
| fall time | 1.8 ns | 10% / 90% |
| overshoot | 14% | |
| settling time | 5.1 ns | settling to within 10% |
| post-amplifier output | | |
| rise time | 2.4 ns | 10% / 90% |
| fall time | 2.9 ns | 10% / 90% |
| overshoot | 11% | |
| settling time | 5.5 ns | settling to within 10% |

| Table A-2: | Step | response | charact | teristics |
|------------|------|----------|---------|-----------|
|------------|------|----------|---------|-----------|

A.4.2 Eye diagrams



Figure A-8: Simulated eye diagrams (200 Mbit/s)

Simulated eye diagrams were used to account for the non-linear behaviour of the amplifier. The eye diagrams were created by stimulating the circuit with a 2^6 -1 maximal length pseudo-random bit sequence generated by a linear feedback shift register. C code was written to simulate the LFSR and generate the HSpice stimulus files. The eye diagrams do not include the effects of noise.

A.4.3 Effect of offset voltage on eye diagrams



Figure A-9: Effect of offset voltage on transient response (200 Mbit/s)

Notice that the distribution of switching thresholds creates an uncertainty in the edge position of about 1 ns which is effectively deducted from the cycle time of the system.

A.4.4 Behaviour of power down circuitry

The verification of the settling behaviour of the data receiver after a change in the state of the disable signal was performed using a full RC model of the chip-level bias net. A separate node in the RC network was provided for each row of 4 receiver circuits. Although simpler lumped approximations are no doubt valid, the brute force approach of a full RC model avoids the need to consider how detailed a model is required and the simulation time is still within reasonable bounds (about one hour).

Typical values for metal sheet resistance and capacitance were used in the simulation.

The simulations included an allowance of 15 nH for the inductance of the external bias pin.

At the start of the header phase of a packet, the chip changes from a state where most of the amplifiers are turned off to a state where all of the amplifiers are turned on. This behaviour was simulated by starting with all the receivers turned off, then negating the disable signal and observing the output of the front-end and decision stage of an amplifier located close to the external pin and far away from the external pin. This behaviour is shown in Figure A-10. It can be seen that it takes approximately 25 ns for the output of the front-end to stabilise. This requires that a dead-time of several bits be inserted between the end of one packet and the start of the header of the next.

At the end of the header phase of a packet, most of the amplifiers are turned off. This behaviour was simulated by starting with all the receivers turned on then asserting the disable signal on all bar one of the amplifiers at the end of each half column. The output of the amplifier that stays enabled was observed (Figure A-11). It can be seen the amplifier that remains turned on is not severely affected by the other amplifiers tuning off and has become completely stable within about 15 ns. Again, this might require a gap of one or two bits to be left between the end of the header and the start of the payload.

The data receiver is more sensitive to noise on the bias line when the photocurrent is high; consequently, all simulations were performed with a DC input photocurrent of $5 \,\mu$ A.

This analysis does not include any effects due to power supply transients.



Figure A-10: Enable settling behaviour of data receiver circuit



Figure A-11: Disable settling behaviour of data receiver circuit

A.4.5 Process variation - transient characteristics

The effects of process, offset, supply and temperature variation on the transient performance of the circuit were checked by applying 1-bit long high-to-low and low-to-high pulses to the circuit and measuring the eye opening in the final output of the receiver circuit.

The eye opening was defined to be the interval of time during which the signal output was valid irrespective of whether it was a logic 0 or a logic 1. Assuming that the phase of the clock can be adjusted, this defines the minimum bit period in conjunction with the set-up and hold time of the latch used to retime the data.



Figure A-12: Definition of eye opening

Separate bias voltages were used for each process corner but the bias voltage was not adjusted to compensate for power supply, offset or temperature variations.

Dynamic range was checked by simulating with a photocurrent between 3.5 μ A and 8 μ A. Bit periods of between 4.0 ns (250 Mbit/s) and 10 ns (100 Mbit/s) were used. Offsets of \pm 31 mV corresponding to 4.7 σ values were used.

Performance was tested at supply voltages of 4.5 V and 5.0 V.

Note that the measurement of the eye opening was performed independently for the different offset extremes. The overall eye opening obtained when the offset extremes are overlaid as in Figure A-9 may be slightly smaller if the pulse width distortion is asymmetric for positive and negative offsets.

Because the temperature of this system must be controlled within a fairly tight window to ensure correct operation of the modulators, correct dynamic response at the temperature extremes is not particularly important, although functional correctness is required to allow testing of the optical inputs at room temperature. Note that the same bias voltage was used for all temperature checks; if further adjustment of the bias voltage to compensate for temperature was permitted, somewhat better performance could be expected.

The results in Table A-3 show that for a typical process, robust operation can be expected up to around 200 Mbit/s with performance OK over all process corners at 100 Mbit/s.

| | | 50 degrees | | | | 0 degrees | | | 70 degrees | | | | |
|--------|--------------|------------|---------|---------|----------|-----------|---------|---------|------------|---------|---------|---------|----------|
| | cycle time | 4.00 ns | 5.00 ns | 6.67 ns | 10.00 ns | 4.00 ns | 5.00 ns | 6.67 ns | 10.00 ns | 4.00 ns | 5.00 ns | 6.67 ns | 10.00 ns |
| corner | photocurrent | | | | | | | | | | | | |
| tm | 3.50 µA | - | 2.25 | OK | OK | - | - | 0.76 | 4.68 | 1.23 | OK | OK | OK |
| tm | 5.00 μΑ | 1.79 | OK | OK | OK | OK | OK | OK | OK | 1.49 | OK | OK | OK |
| tm | 6.50 μΑ | 1.42 | 2.45 | OK | OK | OK | OK | OK | OK | 1.14 | 2.19 | OK | OK |
| tm | 8.00 μΑ | 1.46 | 2.46 | OK | OK | OK | OK | OK | OK | 1.21 | 2.23 | OK | OK |
| wp | 3.50 µA | 1.31 | 2.41 | OK | OK | - | - | - | - | 1.59 | OK | OK | OK |
| wp | 5.00 μΑ | OK | OK | OK | OK | OK | OK | OK | OK | OK | OK | OK | OK |
| wp | 6.50 μΑ | OK | OK | OK | OK | OK | OK | OK | OK | OK | OK | OK | OK |
| wp | 8.00 μΑ | OK | OK | OK | OK | OK | OK | OK | OK | OK | OK | OK | OK |
| ws | 3.50 µA | - | 1.81 | OK | OK | - | - | 3.33 | OK | - | - | OK | OK |
| ws | 5.00 μΑ | - | - | 2.59 | OK | - | 1.89 | OK | OK | - | - | 2.21 | OK |
| ws | 6.50 µA | - | - | 2.16 | OK | - | 1.43 | 3.11 | OK | - | - | 1.78 | OK |
| ws | 8.00 μΑ | - | - | 2.28 | OK | - | 1.48 | 3.17 | OK | - | - | 1.94 | OK |
| WZ | 3.50 µA | - | - | 2.77 | OK | - | - | - | - | - | 1.52 | OK | OK |
| wz | 5.00 μΑ | OK | OK | OK | OK | 1.58 | OK | OK | OK | OK | OK | OK | OK |
| wz | 6.50 µA | 1.86 | OK | OK | OK | OK | OK | OK | OK | 1.58 | OK | OK | OK |
| wz | 8.00 μΑ | 1.83 | OK | OK | OK | OK | OK | OK | OK | 1.58 | OK | OK | OK |
| wo | 3.50 µA | 1.53 | OK | OK | OK | OK | OK | OK | OK | 1.18 | 2.35 | OK | OK |
| wo | 5.00 μΑ | 0.70 | 1.78 | OK | OK | 1.51 | OK | OK | OK | - | 1.50 | 3.17 | OK |
| wo | 6.50 µA | 0.35 | 1.50 | 3.17 | OK | 1.21 | 2.21 | OK | OK | - | 1.22 | 2.89 | OK |
| wo | 8.00 μΑ | 0.48 | 1.64 | 3.31 | OK | 1.30 | 2.31 | OK | OK | - | 1.37 | 3.04 | OK |

Table A-3: Eye opening under process/voltage/temperature/offset variation

OK: : eye opening > bit period / 2

- : no eye opening – signal not valid at this speed

x ns : eye opening in ns

Appendix **B**

Detailed simulations of SPOEC clock receiver

B.1 Introduction

This appendix contains detailed simulation results for the SPOEC clock receiver circuit that are intended to document the expected performance of the circuit. It follows the same format as Appendix A describing the performance of the data receiver circuit.

The simulations include DC transfer characteristics, small-signal AC analyses of both the frontend and the post-amplifier and large-signal transient simulations of the receiver. The effect of DC offsets on the receiver is also discussed.

All simulations were performed in HSpice using BSIM3v2 transistor models with extracted parasitic routing capacitance included. Transient and small-signal simulations used the 40/60 drain-source charge partitioning capacitance model (XPART=0).

| Parameter | nominal value |
|------------------------|---------------|
| power supply voltage | 5 V |
| temperature | 50°C |
| photodiode capacitance | 95 fF |
| input photocurrent | 3.5 µA |

Table B-1: Nominal parameters used in simulations

B.2 DC transfer characteristic



Figure B-1: Clock receiver DC transfer characteristic

Figure B-1 shows the DC transfer characteristic of the receiver. The net photocurrent is the difference between the photocurrent in the non-inverting photodiode and the photocurrent in

the inverting photodiode. For positive values of the net photocurrent, the photocurrent is carried by the non-inverting photodiode whilst the photocurrent in the non-inverting photodiode is zero, and vice versa for negative values of the net photocurrent. Thus, the graphs are an indication of the voltage swing for a particular value of the peak photocurrent and do not represent the differential mode DC transfer function.

Notice that the post-amplifier becomes strongly non-linear for input photocurrent swings larger than $\pm 2 \ \mu A$.

Also notice that, at large input currents, the front-end output voltage falls below the threshold voltage of the post-amplifier input transistors (nominally 0.79 V) which will result in full-rail swings at some of the internal nodes of the post-amplifier. The operation of the circuit has not been verified in this region of operation and this may limit the dynamic range of the receiver.

B.3 Effect of offset on DC transfer characteristic



Figure B-2: Effect of offset on DC transfer characteristic at $\pm 2 \sigma$ and $\pm 4.7 \sigma$ limits

The standard deviation in the offset voltage was calculated as the rms combination of the standard deviation of the offset sources listed in Table B-2 for both the front-end and post-amplifier giving eight offset terms in total. The current I and transconductances g_{mn} and g_{mp} are both defined per branch of the differential circuit.

In the front-end, the input transistors are the pair of NMOS transistors connected to the photodiode inputs and the load transistors are the pair of PMOS transistors providing the bias current. Similarly, in the post-amplifier, the input transistor mismatch is the mismatch between the two NMOS transistors and the load transistor mismatch is between the two PMOS transistors that form the current mirror. The cascode transistor and the transistors in the source
follower do not make a first-order contribution to the offset voltage. The threshold voltage offset parameters were extrapolated from Figures 4-18 and 4-19 and values of $A_{\beta_N}=2$ %µm and $A_{\beta_P}=3$ %µm assumed. The calculated $\sigma_{voFFSET}$ is 10.3 mV; using ± 4.7 σ limits to achieve the same yield as the data receivers gives an estimated worst-case offset voltage of 48.4 mV between the first and second stages.

| description | expression |
|--|---|
| input transistor V_{T} mismatch | $\sigma(\Delta V_{TN})$ |
| load transistor V_{T} mismatch | $\frac{g_{mp}}{g_{mn}}\sigma(\Delta V_{TP})$ |
| input transistor current factor mismatch | $\frac{I}{g_{mn}}\sigma(\frac{\Delta\beta_{\rm N}}{\beta_{\rm N}})$ |
| load transistor current factor mismatch | $\frac{I}{g_{mn}}\sigma(\frac{\Delta\beta_{P}}{\beta_{P}})$ |

Table B-2: Mismatch terms in clock receiver circuit

The simulations show that offset is the primarily limit on the low-speed sensitivity of the circuit, incurring a penalty of approximately $\pm 1 \ \mu A$.

B.4 Small signal analysis

B.4.1 Front-end



Figure B-3: Front-end small-signal transimpedance and group delay as a function of bias photocurrent

The simulation shows the typical AC transfer function of the front-end. The small-signal transimpedance of the front-end is 76 k Ω with no common-mode photocurrent reducing to 51 k Ω with a common-mode photocurrent bias of 2.5 μ A.

The 3dB bandwidths are 141 MHz and 250 MHz respectively.

B.4.2 Post-amplifier



Figure B-4: Post-amplifier differential and common mode voltage gain (simulated with a common-mode photocurrent of 2.5 µA)

The post-amplifier has a nominal gain of 5.2 and a small-signal bandwidth of 622 MHz. There is significant peaking in the response, but we shall see that this does not degrade the transient response.

The common-mode gain is low at low-frequencies but degrades at higher frequencies. The primary effect of this will be to degrade the immunity to high-frequency power supply noise.

B.5 Transient response

B.5.1 Step response



Figure B-5: Clock receiver step response: differential-mode front-end output and singleended post-amplifier output

The step response gives a much better indication of the speed of operation of the circuit than the small-signal transfer function because of the large signal effects.

Note the limiting action of the clamp in the post-amplifier output waveform which removes some of the overshoot in the front-end signal, despite the peaking in the small-signal response of the post-amplifier.

The circuit was designed using a capacitance model that did not accurately model distributed RC effects in the feedback transistor; this accounts for the large front-end overshoot indicated in the table on the slow process corner with the more accurate capacitance model. The overshoot predicted by the original model determined the choice of compensation capacitor (4% overshoot with 2.0 ns settling time).

| front end output | slow | typical | fast | notes |
|-----------------------|--------|---------|--------|------------------------|
| rise time | 1.7 ns | 1.3 ns | 1.1 ns | 10% / 90% |
| fall time | 1.7 ns | 1.3 ns | 1.2 ns | 10% / 90% |
| overshoot | 16 % | 9% | - | |
| settling time | 4.2 ns | 1.3 ns | 1.1 ns | settling to within 10% |
| post-amplifier output | | | | |
| rise time | 1.8 ns | 1.3 ns | 1.1 ns | 10% / 90% |
| fall time | 1.2 ns | 1.0 ns | 0.9 ns | 10% / 90% |
| overshoot | 2% | 2% | - | |
| settling time | 1.8 ns | 1.3 ns | 1.1 ns | settling to within 10% |

Table B-3: Step response characteristics

B.5.2 Eye diagrams



Figure B-6: Simulated eye diagrams at 400 Mbit/s

The clean eye diagrams at 400 Mbit/s indicate that operation with a 200 MHz burst clock signal is feasible. The simulated eye diagram at 500 Mbit/s is still open but shows some signs of pattern-dependent jitter.

B.5.3 Effect of offset voltage on eye diagrams





As in the data receiver, the effect of the offset voltage is to introduce pulse width distortion in the output waveform. In the context of a clock circuit, this in turn results in a clock skew of approximately ± 300 ps between different super-pixels in the chip.

The skew is reduced at higher input photocurrents.

B.5.4 Process variation – transient characteristics

The effects of process, offset, supply and temperature variation on the transient performance of the clock receiver were checked by applying 1-bit long high-to-low and low-to-high pulses to the receiver and measuring the eye opening in the final output of the circuit. This is the same technique used for the data receiver circuit and is described in more detail in Appendix A.

Dynamic range was checked by simulating with a photocurrent between 3.5 μ A and 8 μ A per photodiode. Pulse widths of between 2 ns and 5 ns were used corresponding to clock frequencies between 250 MHz and 100 MHz. The ±4.7 σ limits on the differential offset voltage were used.

Table B-4 shows the eye opening at the nominal operating temperature of 50°C and worst-case values over the full temperature range 0°C to 70°C. Satisfactory operation is achieved up to at least 250 MHz under typical conditions; operation to 150 MHz is possible over the full process and temperature range.

| | | 50 degrees | | | 0-70 degrees | | | | |
|--------|--------------------|------------|---------|---------|--------------|---------|---------|---------|---------|
| | t _{PULSE} | 2.00 ns | 2.50 ns | 3.33 ns | 5.00 ns | 2.00 ns | 2.50 ns | 3.33 ns | 5.00 ns |
| corner | Ι/μA | | | | | | | | |
| tm | 3.50 | OK | OK | OK | OK | 0.93 | OK | OK | OK |
| tm | 5.00 | OK | OK | OK | OK | OK | OK | OK | OK |
| tm | 6.50 | OK | OK | OK | OK | OK | OK | OK | OK |
| tm | 8.00 | OK | OK | OK | OK | OK | OK | OK | OK |
| wp | 3.50 | 0.89 | OK | OK | OK | 0.61 | 1.11 | OK | OK |
| wp | 5.00 | OK | OK | OK | OK | OK | OK | OK | OK |
| wp | 6.50 | OK | OK | OK | OK | OK | OK | OK | OK |
| wp | 8.00 | OK | OK | OK | OK | OK | OK | OK | OK |
| ws | 3.50 | - | 0.96 | OK | OK | - | 0.86 | OK | OK |
| ws | 5.00 | 0.36 | 1.07 | OK | OK | - | 0.95 | OK | OK |
| ws | 6.50 | 0.31 | 1.14 | OK | OK | - | 1.02 | OK | OK |
| WS | 8.00 | 0.20 | 1.22 | OK | OK | - | 1.09 | OK | OK |
| WZ | 3.50 | 0.71 | OK | OK | OK | 0.60 | 1.21 | OK | OK |
| WZ | 5.00 | 0.84 | OK | OK | OK | 0.76 | OK | OK | OK |
| WZ | 6.50 | 0.87 | OK | OK | OK | 0.78 | OK | OK | OK |
| WZ | 8.00 | 0.89 | OK | OK | OK | 0.79 | OK | OK | OK |
| wo | 3.50 | 0.99 | OK | OK | OK | 0.96 | OK | OK | OK |
| wo | 5.00 | OK | OK | OK | OK | OK | OK | OK | OK |
| wo | 6.50 | OK | OK | OK | OK | OK | OK | OK | OK |
| wo | 8.00 | OK | OK | OK | OK | OK | OK | OK | OK |

There is no indication of dynamic range problems.

| Table B-4: Pulse-width checks of | n clock receiver a | at nominal temperature aı | nd over full |
|----------------------------------|--------------------|---------------------------|--------------|
|----------------------------------|--------------------|---------------------------|--------------|

temperature range

- OK: : eye opening > bit period / 2
- : no eye opening signal not valid at this speed
- x ns : eye opening in ns

References

[1] GOODMAN, J.W., LEONBERGER, F.J., KUNG, S.-Y., ATHALE, R.A.: 'Optical

interconnections for VLSI systems', Proc. IEEE, July 1984, 72 (7), pp. 850-866

[2] MOORE, G.E.: 'Cramming more components onto integrated circuits', *Electron.*, 19

April 1965, pp. 114-117, reprinted in Proc. IEEE., January 1998, 86 (1), pp. 82-85

[3] SEMICONDUCTOR INDUSTRY ASSOCIATION: 'The national technology roadmap for semiconductors: technology needs', 1997 Edition

[4] DALLY, W.J., POULTON, J.W.: 'Digital systems engineering', 1998, Cambridge University Press, pp. 19-20

[5] CHANG, C.S., OSCILOWSKLI, A., BRACKEN, R.C.: 'Future challenges in electronics packaging', *IEEE Circuits and Devices*, March 1998, **14** (2), pp. 45-54

[6] SEMICONDUCTOR INDUSTRY ASSOCIATION: *ibid.*, p.11

[7] SMITH, B.: 'Alternatives and imperatives for optical interconnect in high performance computers', in *Optics in Computing*, Technical Digest (Optical Society of America,

Washington DC, 1997), p.28, oral presentation

[8] DINES, J.A.B.: private communication

[9] 'Next generation IP switches and routers', *IEEE J. Selected Areas in Comm.*, forthcoming issue, 1999

[10] AVICI SYSTEMS INC.: 'The world of terabit/s switch/router technology', White paper, 1998, http://www.avici.com/

[11] AVICI SYSTEMS INC.: 'AVICI TSR Optical Terabit switch router', data sheet, 1998
[12] STUNKEL, C.B.: 'Commercial MPP networks: time for optics', Proc. 4th Int. Conf. on Massively Parallel Processing using Optical Interconnections - MPPOI 97, Montreal, Canada, 1997, p.90-95

```
[13] DALLY, W.J., LEE, M.-J.E., AN, F.-T., POULTON, J., TELL, S.: 'High-performance
electrical signalling', Proc. 5<sup>th</sup> Int. Conf. on Massively Parallel Processing using Optical
Interconnections - MPPOI 98, Las Vegas, USA, 1998, see
```

http://www.cs.unc.edu/~fast_links/pubs.html

[14] TELL, S.: 'Selected bibliography on high-speed signalling',

http://www.cs.unc.edu/~fast_links/bib.html

[15] DALLY, W.J., POULTON, J.W.: ibid., Chapters 7-12

[16] MILLER, D.A.B., OZAKTAS, H.M.: 'Limit to the bit-rate capacity of electrical interconnects from the aspect ratio of system architecture', *J. Parallel Distributed Computing*, 1997, **41** (1), pp. 42-52

[17] SMITH, B.: 'Interconnection networks for shared memory parallel computers', *Proc.* 2nd *Int. Conf. on Massively Parallel Processing using Optical Interconnects*, 1995, IEEE Press, pp. 255-256

[18] IEEE: '1596.3-1996 – IEEE standard for low-voltage differential signals (LVDS) for Scalable Coherent Interface (SCI)', 1996

[19] TEXAS INSTRUMENTS INC.: 'Texas Instruments delivers world's first 2.5 Gbps transceiver core for CMOS ASICs', 5 October 1998,

http://www.ti.com/sc/docs/news/1998/98080a.htm

[20] WOODWARD, T.K., KRISHNAMOORTHY, A.K., LENTINE, A.L., GOOSEN, K.W., WALKER, J.A., CUNNINGHAM, J.E., YAN, W.Y., D'ASARO, L.A., CHIROVSKY,

L.M.F., HUI, S.P., TSENG, B., KOSSIVES, D., DAHRINGER, D., LEIBENGUTH, R.E.: '1-Gb/s two-beam transimpedance smart-pixel optical receivers made from hybrid GaAs MQW modulators bonded to 0.8 µm silicon CMOS', *IEEE Photonics Technology Lett.*, March 1996, **8** (3), pp. 422-424

[21] ALTRON INC.: 'Multilayer PCB product information', http://www.altron.com/, now owned by Sanmina Corporation, http://www.sanmina.com/

[22] CHEN, R.T., WU, L., TANG, S., DUBINOVSKY, M., QI, J., SCHOW, C.L.,

CAMPBELL, C.L., WICKMAN, R., PICOR, B., HIBBS-BRENNER, M., BRISTOW, J.,

LIU, Y.S., RATTAN, S., NODDINGS, C.: 'Si CMOS process compatible guided-wave multi-Gbit/s optical clock distribution system for Cray T-90 supercomputer', Proc. 4th Int. Conf. on Massively Parallel Processing using Optical Interconnections - MPPOI 97,

Montreal, Canada, 1997, p.10-24

[23] MICROMODULE SYSTEMS INC .: 'MCM technology information',

http://www.mosis.org/New/Technical/MCM/mms.html

[24] IBM MICROELECTRONICS: 'IBM Multichip module on laminate', product data sheet, Available http://www.chips.ibm.com/products/interconnect/

[25] MILLER, D.A.B.: 'Physical reasons for optical interconnection', *Int. J. Optoelectronics*, May-June 1997, **11** (3), pp. 155-168

[26] SENIOR, J.: 'Optical fiber communications: principles and practice', 2nd edition, 1992,Prentice-Hall, p.106

[27] LEVI, A.F.J.: 'High-performance optoelectronic physical layers in systems', Proc. 4th
 Int. Conf. on Massively Parallel Processing using Optical Interconnections - MPPOI 97,
 Montreal, Canada, 1997, p.2

[28] FibreChannel Association web site: http://www.fibrechannel.com/

[29] GIGABIT ETHERNET ALLIANCE: 'Gigabit ethernet – accelerating the standard for speed', white-paper, 1998, http://www.gigabit-ethernet.org/technology/whitepapers/
[30] WILLNER, A.E.: 'Mining the optical bandwidth for a terabit per second', *IEEE Spectrum*, April 1997, **34** (4), pp. 32-41

[31] MILLER, D.A.B.: 'Dense two-dimensional integration of optoelectronics and electronics for interconnections', *SPIE Symposium on Photonics West Optoelectronics* 98, San Jose, 24-30 January 1998

[32] KRISHNAMOORTHY, A.V., GOOSSEN, K.W.: 'Progress in optoelectronic-VLSI smart-pixel technology based on GaAs/AlGaAs MQW modulators', *Int. J. Optoelectronics*, May-June 1997, **11** (3), pp. 181-198

[33] TOOLEY, F.A.P.: 'Challenges in optically interconnecting electronics', *IEEE J.* Selected Topics in Quant. Electron., April 1996, **2** (1), pp. 3-13

[34] AYADI, K.: 'High speed CMOS photoreceivers and technology for OptoElectronic ICs and data communications', July 1998, PhD thesis, Vrije Universiteit Brussel, Belgium
[35] AYADI, K., KUIJK, M., HEREMANS, P., BICKEL, G., BORGHS, G., VOUNCKX, R.: 'A monolithic optoelectronic receiver in standard 0.7 μm CMOS operating at 180 MHz and 176-fJ light input energy', *IEEE Photonics Tech. Lett.*, January 1997, 9 (1), pp. 88-90
[36] WOODWARD, T.K., KRISHNAMOORTHY, A.V.: '1 Gbit/s CMOS photoreceiver with integrated detector operating at 850 nm', *Electron. Lett.*, 11 June 1998, 34 (12), pp. 1252-1253

[37] KUCHTA, D.M., AINSPAN, H.A., CANORA, F.J., SCHNEIDER, R.P.Jr: 'Performance of fiber-optic data links using 670-nm CW VCSELs and a monolithic Si photodetector and CMOS preamplifier', *IBM J. Research and Development*, January-March 1995, **39** (1-2), pp. 63-72

[38]KUIJK, M., COPPEE, D., VOUNCKX, R.: 'Spatially modulated light detector in CMOS with sense-amplifier receiver operating at 180 Mb/s for optical data link applications and parallel optical interconnects between chips', *IEEE J. Selected Topics in Quantum Electron.*, November-December 1998, **4** (6), pp. 1040-1045

[39] VITESSE SEMICONDUCTOR INC.: http://www.vitesse.com/

[40] TRIQUINT SEMICONDUCTOR INC .: 'TriQuint Foundry',

http://www.triquint.com/Foundry

[41] WONG, Y.M., MUEHLNER, D.J., FAUDSKAR, C.C., BUCHHOLZ, D.B.,

FISHTEYN, M., BRANDNER, J.L., PARZYGNAT, W.J., MORGAN, R.A., MULLALLY,

T., LEIBENGUTH, R.E., GUTH, G.D., FOCHT, M.W., GLOGOVSKY, K.G., ZILKO, J.L.,

GATES, J.V., ANTHONY, P.J., TYRONE, B.H.Jr., IRELAND, T.J., LEWIS, D.H.Jr., SMITH, D.F., NATI, S.F., LEWIS, D.K., ROGERS, D.L., AISPAIN, H.A., GOWDA, S.M., WALKER, S.G., KWARK, Y.H., BATES, R.J.S., KUCHTA, D.M., CROW, J.D. :

'Technology development of a high-density 32-channel 16-Gb/s optical data link for optical interconnection applications for the optoelectronic technology consortium (OETC)', *J*.

Lightwave Technology, June 1995, 13 (6), pp. 995-1016

[42] FONSTAD, C.: 'OPTOCHIP project guide', 1996,

http://web.mit.edu/afs/athena.mit.edu/user/f/o/fonstad/optochip/opto.home.html

OPTOCHIP was a multi-project fabrication run in the Vitesse HGaAsIII process containing optoelectronic circuits from a number of US universities.

[43] YUANG, R.-H., SHIEH, J.-L., CHYI, J.-I., CHEN, J.-S.: 'Overall performance

improvement in GaAs MSM photodetectors using recessed-cathode structure', IEEE

Photonics Technol. Lett., February 1997, 9 (2), pp. 226-228

[44] MATSUO, S., NAKAHARA, T., KOHAMA, Y., OHISO, Y., FUKUSHIMA, S.,

KUROKAWA, T.: 'Monolithically integrated smart pixel using an MSM-PD, MESFET'S,

and a VCSEL', IEEE J. Selected Topics in Quant. Electron., April 1996, 2 (1), pp. 121-127

[45] CHOI, J., SHEU, B.J., CHEN, O.T.-C.: 'Monolithic GaAs receiver for optical

interconnect systems', IEEE J. Solid-State Circuits, March 1994, 29 (3), pp.328-331

[46] DUTTA, N.K., TU, K.Y., LEVINE, B.F.: 'Optoelectronic integrated receiver', *Electron*. *Lett.*, 3 July 1997, **33** (14), pp.1254-1255

[47] ROGERS, D.L.: 'Integrated optical receivers using MSM detectors', *J. Lightwave Technology*, December 1991, **9** (12), pp. 1635-1638

[48] HAYES, E.M., JURRAT, R., PU, R., SNYDER, R.D., FELD, S.A.: 'Foundary (*sic*) fabricated array of smart pixels integrating MESFETs/MSMs and VCSELs', *Int. J. Optoelectronics*, May-June 1997, **11** (3), pp. 229-237

[49] VITESSE SEMICONDUCTOR CORPORATION: 'VSC7810 – photodetector / transimpedance amplifier family for optical communication', August 1997, product data sheet

[50] D'ASARO, L.A., CHIROVSKY, L.M.F., LASKOWSKI, E.J., PEI, S.-S.,

LEIBENGUTH, R.E., WOODWARD, T.K., FOCHT, M., LENTINE, A.L., ASOM, M.T., GUTH, G., KOPF, R.F., KUO, J.M., PEARTON, S.J., PRZYBYLEK, G.J., REN, F., SMITH, L.E.: 'Batch fabrication and structure of integrated GaAs-AlGaAs field-effect transistor selfelectro-optic devices (FET-SEEDs)', *IEEE Electron Device Lett.*, October 1992, **13** (10), pp. 528-531 [51] MILLER, L.F.: 'Controlled collapse reflow chip joining', *IBM J. Res. Develop.*, May 1969, pp. 239-250

[52] WIELAND, J., MELCHIOR, H., KEARLEY, M.Q., MORRIS, C.M., MOSELEY, A.M., GOODWIN, M.J., GOODFELLOW, R.C.: 'Optical receiver array in silicon bipolar technology with self-aligned, low parasitic III-V detectors for DC-1 Gbit/s parallel links', *Electron. Lett.*, 21 November 1991, **27** (24), pp. 2211-2213

[53] GOODWIN, M.J., MOSELEY, A.J., KEARLY, M.Q., MORRIS, R.C., KIRKBY,
C.J.G., THOMPSON, J., GOODFELLOW, R.C., BENNION, I.: 'Optoelectronic component arrays for optical interconnection of circuits and subsystems', *J. Lightwave Tech.*, December 1991, 9 (12), pp. 1639-1645

[54] GOOSSEN, K.W., WALKER, J.A., DASARO, L.A., HUI, S.P., TSENG, B.,

LEIBENGUTH, R., KOSSIVES, D., BACON, D.D., DAHRINGER, D., CHIROVSKY,

L.M.F., LENTINE, A.L., MILLER, D.A.B.: 'GaAs MQW modulators integrated with silicon CMOS', *IEEE Photonics Tech. Lett.*, April 1995, **7** (4), pp. 360-362

[55] NOVOTNY, R.A., LENTINE, A.L., BUCHHOLZ, D.B., KRISHNAMOORTHY, A.V.:
'Analysis of parasitic front-end capacitance and thermal resistance in hybrid flip-chipbonded GaAs SEED/Si CMOS receivers', in *Optical Computing*, 1995 OSA Technical Digest Series 10, pp. 207-209

[56] STIRK, C.W., NEFF, J.: 'The cost of optical interconnects vs. MCMs', *Optics in Computing*, Technical Digest (Optical Society of America, Washington DC, 1997), pp. 21-23
[57] AHADIAN, J.F., VAIDYANATHAN, P.T., PATTERSON, S.G., ROYTER, Y., MULL, D., PETRICH, G.S., GOODHUE, W.D., PRASAD, S., KOLODZIEJSKI, L.A., FONSTAD, C.G.Jr.: 'Practical OEIC's based on the monolithic integration of GaAs-InGaP LED's with commercial GaAs VLSI electronics', *IEEE J. Quant. Electron.*, July 1998, **34** (7), pp. 1117-1123

[58] WANG, H., LUO, J., SHENOY, K.V., ROYTER, Y., FONSTAD, C.G.Jr., PSALTIS,D.: 'Monolithic integration of SEEDs and VLSI GaAs circuits by epitaxy on electronics',

IEEE Photonics Tech. Lett., May 1997, 9 (5), pp. 607-609

[59] GERARD, B.: 'Esprit Project 22614 - MONOLITH: Monolithic integration of light emitting devices with Si-ICs using conformal epitaxy', 1996,

http://www.cordis.lu/esprit/src/synops.htm#monolith

[60] EUROPEAN COMMISSION: 'Technology Roadmap – optoelectronic interconnects for integrated circuits', Esprit programme Long Term Research Microelectronics advanced research initiative (ESPRIT MEL-ARI), June 1998, ISBN 92-828-3950-8

[61] GOOSEEN, K.W.: 'Optoelectronic VLSI', in *Spatial Light Modulators*, Technical Digest (Optical Society of America, Washington DC, 1997), pp. 2-4

[62] TREZZA, J.A., POWELL, J.S., GARVIN, C.G., KANG, K., STACK, R.D.: 'Creation and application of a very large format high-fill-factor GaAs-on-CMOS binary and gray-scale modulator and emitter arrays', in *Optics in Computing* '98, Pierre Chavel, David A.B. Miller, Hugo Thienpont, Editors, Proc. SPIE **3490** pp. 78-81 (1998)

[63] LENTINE, A.L., MILLER, D.A.B.: 'Evolution of the SEED technology: bistable logic gates to optoelectronic smart pixels', *IEEE J. Quantum Electron.*, February 1993, **29** (2), pp. 655-669

[64] MILLER, D.A.B., CHEMLA, D.S., DAMEN, T.C., GOSSARD, A.C., WIEGMANN,
W., WOOD, T.H., BURRUS, C.A.: 'Electric field dependence of optical absorption near the band gap of quantum-well structures', *Phys. Rev. B*, 15 July 1985, **32** (2), pp. 1043-1060
[65] HINTON, H.S., CLOONAN, T.J., McCORMICK, F.B.Jr., LENTINE, A.L., TOOLEY,
F.A.P.: 'Free-space digital optical systems', *Proc. IEEE*, November 1994, **82** (11), pp. 1632-

1649

[66] SDL INC.: 'SDL TC-30 500 mW tunable external cavity laser diode system', 1998, http://www.sdli.com/

[67] VENDITTI, M.B., KABAL, D., AYLIFFE, M.H., RICHARD, E., PLANT, D.V., TOOLEY, F.A.P., CURRIE, J., SPRINGTHORPE, A.J.: 'Electrical, thermal and optomechanical packaging of large 2D optoelectronic device arrays for free-space optical interconnects', accepted for publication in *J. European Optical Soc.*, 1998
[68] WILKINSON, L.C.: 'Indium gallium arsenide multiple quantum well devices for optically interconnected smart pixels', PhD Thesis, Heriot-Watt University, UK, 1998
[69] BHATTACHARYA, P.: 'Properties of lattice-matched and strained Indium Gallium Arsenide', EMIS Datareview series No. 8 (INSPEC, IEE, London, 1993)
[70] BUCHHOLZ, D.B., LENTINE, A.L., NOVOTNY, R.A.: 'Thermal considerations in the design of optoelectronic device mounts', in *Photonics in Switching*, 1995 OSA Technical Digest Series (Optical Society of America, Washington DC, 1995), **12**, pp. 118-120
[71] NOVOTNY, R.A.: 'Analysis of smart pixel digital logic and optical interconnections', PhD thesis, Heriot-Watt University, UK
[72] GOOSSEN, K.W., CUNNINGHAM, J.E., JAN. W.Y., LEIBENGUTH, R.: 'On the operational and manufacturing tolerances of GaAs-AIAs MQW modulators', *IEEE J.*

Quantum Electron., March 1998, **34** (3), pp. 431-438

[73] FORBES, M.G., WALKER, A.C., POTTIER, F, VOGELE, B., STANLEY, C.R.:

'Uniformity measurements on large InGaAs-AlGaAs multiple quantum well modulator arrays', *Spatial Light Modulators*, Technical Digest (Optical Society of America, Washington DC, 1997), pp. 123-124

[74] BAILLIE, D.A.: 'Free-space optical interconnection of digital electronics', PhD Thesis, Heriot-Watt University, UK, September 1996

[75] BOYD, G.D., FOX, A.M., MILLER, D.A.B., CHIROVSKY, L.M.F., D'ASARO, L.A.,
KUO, J.M., KOPF, R.F., LENTINE, A.L.: '33 ps optical switching of symmetric self-electrooptic effect devices', *Appl. Phys. Lett.*, 29 October 1990, **57** (18), pp. 1843-1845

[76] LENTINE, A.L., CHIROVSKY, L.M.F., D'ASARO, L.A., LASKOWSKI, E.J., PEI, S.-S., FOCHT, M.W., FREUND, J.M., GUTH, G.D., LEIBENGUTH, R.E., SMITH, L.E.,
WOODWARD, T.K.: 'Field-effect-transistor self-electro-optic-effect-device (FET-SEED) electrically addressed differential modulator array', *Appl. Opt.*, 10 May 1994, **33** (14), pp. 2849-2855

[77] WALKER, A.C., YANG, T.-Y., GOURLAY, J., DINES, J.A.B., FORBES, M.G.,
PRINCE, S.M., BAILLIE, D.A., NEILSON, D.T., WILLIAMS, R., WILKINSON, L.C.,
SMITH, G.R., DESMULLIEZ, M.P.Y., BULLER, G.S., TAGHIZADEH, M.R., WADDIE,
A., UNDERWOOD, I., STANLEY, C.R., POTTIER, F., VOGELE, B., SIBBETT, W.:
'Optoelectronic systems based on InGaAs-complementary-metal-oxide-semiconductor
smart-pixel arrays and free-space interconnects', *Appl. Opt.*, 10 May 1998, **37** (14), pp. 2822-2830

[78] WOODWARD, T.K., KRISHNAMOORTHY, A.V., GOOSSEN, K.W., WALKER, J.A., TSENG, B., LOTHIAN, J., HUI, S., LEIBENGUTH, R.: 'Modulator driver circuits for optoelectronic VLSI', *IEEE Photonics Tech. Lett.*, June 1997, **9** (6), pp. 839-841
[79] GOOSSEN, K.W., CUNNINGHAM, J.E., JAN, W.Y.: 'Stacked-diode electroabsorption modulator', *IEEE Photonics Tech. Lett.*, August 1994, **6** (8), pp. 936-938
[80] BYRNE, D., HORAN, P., HEGARTY, J.: 'Optimisation of InGaAs MQW modulator structures operating with 5V or less modulation', in *Optics in Computing* '98, Pierre Chavel, David A.B. Miller, Hugo Thienpont, Editors, Proc. SPIE **3490** pp. 389-392 (1998)
[81] KRISHNAMOORTHY, A.V., MILLER, D.A.B.: 'Scaling optoelectronic-VLSI circuits into the 21st century: a technology roadmap', *IEEE J. Selected Topics in Quantum Electron.*, April 1996, **2** (1), pp. 55-76
[82] LIU, Y., ROBERTSON, B., BOISSET, G., AYLIFFE, M.H., IYER, R., PLANT, D.V.:

⁽⁸²⁾ LIU, Y., ROBERTSON, B., BOISSET, G., AYLIFFE, M.H., IYER, R., PLANT, D.V.: ^{(Design, implementation and characterization of a hybrid optical interconnect for a four-} stage free-space optical backplane demonstrator', *Appl. Opt.*, 10 May 1998, **37** (14), pp. 2895-2914

[83] LENTINE, A.L., GOOSSEN, K.W., WALKER, J.A., CUNNINGHAM, J.E., JAN,

W.Y., WOODWARD, T.K., KRISHNAMOORTHY, A.V., TSENG, B.J., HUI, S.P.,

LEIBENGUTH, R.E., CHIROVSKY, L.M.F., NOVOTNY, R.A., BUCHHOLZ, D.B.,

MORRISON, R.L.: 'Optoelectronic VLSI switching chip with greater than 1Tbit/s potential optical I/O bandwidth', *Electron. Lett.*, 8 May 1997, **33** (10), pp. 894-895

[84] WOODWARD, T.K.: 'Optoelectronic VLSI foundry services from Lucent Technologies', 1998, http://www.bell-labs.com/project/oevlsi/

[85] CHOQUETEE, K.D., HOU, H.Q.: 'Vertical-cavity surface emitting lasers: moving from research to manufacturing', *Proc. IEEE.*, November 1997, **85** (11), pp. 1730-1739

[86] HIBBS-BRENNER, M.K., MORGAN, R.A., WALTERSON, R.A., LEHMAN, J.A.,

KALWEIT, E.L., BOUNNAK, S., MARTA, T., GIESKE, R.: 'Performance, uniformity, and yield of 850-nm VCSEL's deposited by MOVPE', *IEEE Photonics Tech. Lett.*, January 1996, **8** (1), pp. 7-9

[87] JAGER, R., GRABHERR, M., JUNG, C., MICHALZIK, R., REINER, G., WEIGL, B.,
EBELING, K.J.: '57% wallplug efficiency oxide-confined 850 nm VCSELs', *Electron. Lett.*,
13 February 1997, **33** (4), pp. 330-331

[88] LEAR, K.L., CHOQUETTE, K.D., SCHNEIDER, R.P.Jr., KILCOYNE, S.P., GEIB,

K.M.: 'Selectively oxidised vertical cavity surface emitting lasers with 50% power conversion efficiency', *Electron. Lett.*, 2 February 1995, **31** (3), pp. 208-209

[89] FIEDLER, U., REINER, G., SCHNITZER, P., EBELING, K.J.: 'Top-surface emitting laser diodes for 10 Gbit/s data transmission', *IEEE Photonics Tech. Lett.*, June 1996, **8** (6), pp. 746-748

[90] SCHNITZER, P., GRABHERR, M., JAGER, R., JUNG, C., EBELING, K.J.: 'Bias-free
2.5 Gbit/s data transmission using polyimide passivated GaAs VCSELs', *Electron. Lett.*, 19
March 1998, **34** (6), pp. 573-575

[91] KOHAMA, Y., OHISO, Y., FUKUSHIMA, S., KUROKAWA, T.: '8×8 independently addressable vertical-cavity surface emitting laser diode arrays grown by MOCVD', *IEEE Photonics Tech. Lett.*, August 1994, **6** (8), pp. 918-920

[92] SPOEC 24-month progress report, September 1998, pp. 17-19

[93] MORGAN, R.A., ROBINSON, K.C., CHIROVSKY, L.M.F, FOCHT, M.W., GUTH,

G.D., LEIBENGUTH, R.E., GLOGOVSKY, K.G., PRZYBYLEK, G.J., SMITH, L.E.:

'Uniform 64×1 arrays of individually-addressed vertical cavity top surface emitting lasers', *Electron. Lett.*, 1 August 1991, **27** (16), pp. 1400-1402

[94] PU, R., HAYES, E.M., WILMSEN, C.W., CHOQUETTE, K.D., HOU, H.Q., GEIB,K.M.: 'Comparison of techniques for bonding VCSELs directly to ICs', *Optics in Computing* '98, pp. 498-501

[95] MATSUO, S., NAKAHARA, T., TATENO, K., KUROKAWA, T.: 'Novel technology for hybrid integration of photonic and electronic circuits', *IEEE Photonics Tech. Lett.*, November 1996, **8** (11), pp. 214-216

[96] KRISHNAMOORTHY, A.V., CHIROVSKY, L.M.F., HOBSON, W.S., LEIBENGUTH,
R.E., HUI, S.P., ZYDZIK, G.J., GOOSSEN, K.W., WYNN, J.D., TSENG, B.J., LOPATA, J.,
WALKER, J.A., CUNNINGHAM, J.E., D'ASARO, L.A.: 'Vertical cavity surface emitting
lasers flip-chip bonded to CMOS circuits', *Optics in Computing* '98, Post-deadline paper
[97] PAANANEN, D., WASSERBAUER, J., SCOTT, J., BRUSENBACH, P., LEWIS, D.,
SIMONIS, G.: 'A 16×16 vertical cavity surface emitting laser array module for optical
processing and interconnect applications', *Optics in Computing* '97, Post-deadline paper PD-1

[98] PU, R., HAYES, E.M., JURRAT, R., WILMSEN, C.W., CHOQUETTE, K.D., HOU,
H.Q., GEIB, K.M.: 'VCSELs bonded directly to foundry fabricated GaAs smart pixel arrays', *IEEE Photonics Tech. Lett.*, December 1997, **9** (12), pp. 1622-1624

[99] OHISO, Y., TATENO, K., KOHAMA, Y., WAKATSUKI, A., TSUNETSUGU, H.,
KUROKAWA, T.: 'Flip-chip bonded 0.85-µm bottom-emitting vertical-cavity laser array on an AlGaAs substrate', *IEEE Photonics Tech. Lett.*, September 1996, 8 (9), pp. 1115-1117
[100] LIU, Y., HIBBS-BRENNER, M.K., MORGAN, B., NOHAVA, J., WALTERSON, B.,
MARTA, T., BOUNNAK, S., KALWEIT, E., LEHMAN, J., CARLSON, D., WILSON, P.:
'Integrated VCSELs, MSM photodetectors and GaAs MESFETs for low cost optical interconnects', in *Spatial Light Modulators*, Technical Digest (Optical Society of America, Washington DC, 1997), pp. 22-24

[101] GOOSSEN, K.W., TSENG, B., HUI, S.P., WALKER, J.A., LEIBENGUTH, R., CHIROVSKY, L.M.F., KRISHNAMOORTHY, A.: 'Multiple-attachment GaAs-on-Si hybrid optoelectronic/VLSI chips', *Proc. 1996 IEEE/LEOS Summer Topical Meet. on Smart Pixels*, Keystone, Colorado, 5-9 August 1996, pp. 24-25

[102] HAHN, K.H., GIBONEY, K.S., WILSON, R.E., STRAZNICKY, J., WONG, E.G., TAN, M.R., KANESHIRO, R.T., DOLFI, D.W., MUELLER, E.H., PLOTTS, A.E., MURRAY, D.D., MARCHEGIANO, J.E., BOOTH, B.L., SANO, B.J., MADHAVEN, B., RAGHAVEN, B., LEVI, A.F.J.: 'Gigabyte/s data communication with the POLO parallel optical link', *Proc. Electronic Components and Technol. Conf. 1996*, pp. 301-307

[103] GRULA, J.: 'Application note AN1572 – Applying the Optobus[™] I multichannel optical data link to high-performance communication systems: SCI, fibre channel and ATM', Motorola Inc., June 1996

[104] USUI, M., MATSUURA, N., SATO, N., NAKAMURA, M., TANAKA, N., OHKI, A., HIKITA, M., YOSHIMURA, R., TATENO, K., KATSURA, K., ANDO, Y.: '700-Mb/s 40channel parallel optical interconnection module using VCSEL arrays and bare fiber connectors (ParaBIT: Parallel inter-Board optical Interconnection Technology)', *Proc. Lasers and Electro-Optics Society Annual Meeting 1997 – volume 1*, 1997, pp.51-55

[105] KOYABU, K., OHIRA, F., YAMAMOTO, T.: 'Fabrication of two-dimension fiber arrays using microferrules', *IEEE Trans. Components, Packaging and Manufacturing Technol. – Part C*, January 1998, **21** (1), pp. 11-19

[106] BASAVANHALLY, N., BORUTTA, R., CRISCI, R., NIJANDER, C., WATKINS, L.:
'Evolution of fiber arrays for free-space interconnect applications', in *Photonics in Switching*, 1995 OSA Technical Digest Series (Optical Society of America, Washington DC, 1995), **12**, pp. 124-128

[107] CRYAN, C.V.: 'Two-dimensional multimode fibre array for optical interconnects', *Electron. Lett.*, 19 March 1998, **34** (6), pp. 586-587

[108] PROUDLEY, G.M., STACE, C., WHITE, H.J.: 'Fabrication of two-dimensional fiber optic arrays for an optical switch', *Opt. Eng.*, February 1994, **33** (2), pp. 627-635

[109] KOSAKA, H., KAJITA, M., YAMADA, M., SUGIMOTO, Y., KURATA, K.,

TANABE, T., KASUKAWA, Y..: '2D alignment free VCSEL-array module with push/pull fibre connector', *Electron. Lett.*, 10 October 1997, **32** (31), pp. 1991-1992

[110] LENTINE, A.L., REILEY, D.J., NOVOTNY, R.A., MORRISON, R.L., SASIAN, J.M., BECKMAN, M.G., BUCHHOLZ, D.B., HINTERLONG, S.J., CLOONAN, T.J.,

RICHARDS, G.W., McCORMICK, F.B.: 'Asynchronous transfer mode distribution network by use of an optoelectronic VLSI switching chip', *Appl. Opt.*, 10 March 1997, **36** (8), pp. 1804-1814

[111] HIRABAYASHI, K., YAMAMOTO, T., HINO, S., KOHAMA, Y., TATENO, K.:
'Optical beam direction compensating system for board-to-board free space optical interconnection in high-capacity ATM switch', *J. Lightwave Technol.*, May 1997, **15** (5), pp. 874-882 [112] DESMULLIEZ, M.P.Y, TOOLEY, F.A.P., DINES, J.A.B., GRANT, N.L.,

GOODWILL, D.J., BAILLIE, D., WHERRETT, B.S., FOULK, P.W., ASHCROFT, S.,

BLACK, P.: 'Perfect-shuffle interconnected bitonic sorter: optoelectronic design', *Appl. Opt.*, 10 August 1995, **34** (23), pp. 5077-5090

[113] NEILSON, D.T., PRINCE, S.M., BAILLIE, D.A., TOOLEY, F.A.P.: 'Optical design of a 1024-channel free-space sorting demonstrator', *Appl. Opt.*, 10 December 1997, **36** (35), pp. 9243-9252

[114] KABAL, D., AYLIFFE, M.H., BOISSET, G.C., PLANT, D.V., ROLSTON, D.R.,

VENDITTI, M.B.: 'Chip-on-board packaging of a hybrid-SEED smart pixel array', in *Optics in Computing*, Technical Digest (Optical Society of America, Washington DC, 1997), pp. 76-78

[115] JOHNSON, K.M., McKNIGHT, D.J., UNDERWOOD, I.: 'Smart spatial light modulators using liquid crystals on silicon', *IEEE J. Quantum Electron.*, February 1993, 29
(2), pp. 699-714

[116] WALKER, A.C., DESMULLIEZ, M.P.Y., TOOLEY, F.A.P., NEILSON, D.T., DINES, J.A.B., BAILLIE, D.A., PRINCE, S.M., WILKINSON, L.C., TAGHIZADEH, M.R., BLAIR, P., SNOWDON, J.F., WHERRETT, B.S., STANLEY, C., POTTIER, F., UNDERWOOD, I., VASS, D.G., SIBBETT, W., DUNN, M.H.: 'Construction of an optoelectronic bitonic sorter based on CMOS/InGaAs smart pixel technology', Proc. 2nd Int. Conf. on Massively Parallel Processing using Optical Interconnections - MPPOI 95, San Antonio, USA, 1995, pp.180-187

[117] SPOEC 24-month progress report, September 1998, p.45

[118] TAGHIZADEH, M.R., TURUNEN, J.: 'Synthetic diffractive elements for optical interconnection', *Opt. Computing and Processing*, 1992, **2** (4), pp. 221-242

[119] GOODMAN, J.W.: 'Introduction to Fourier optics', 2nd Edition, McGraw-Hill, 1996, pp. 210-214, pp.96-120

[120] HINTON, H.S., CLOONAN, T.J., McCORMICK, F.B., LENTINE, A.L., TOOLEY,

F.A.P.: 'Free-space digital optical systems', *Proc. IEEE*, November 1994, **82** (11), pp. 1632-1649

[121] DESMULLIEZ, M.P.Y., TOOLEY, F.A.P., DINES, J.A.B., GRANT, N.L.,

GOODWILL, D.J., BAILLIE, D., WHERRETT, B.S., FOULK, P.W., ASHCROFT, S.,

BLACK, P.: 'Perfect-shuffle interconnected bitonic sorter: optoelectronic design', *Appl. Opt*, 10 August 1995, **34** (23), pp. 5077-5090

[122] KAROL, M.J., HLUCHYJ, M.G., MORGAN, S.P.: 'Input versus output queuing on a space-division packet switch', *IEEE Trans. Commun.*, December 1987, **35** (12), pp. 1347-1356

[123] HLUCHYJ, M.G., KAROL, M.J.: 'Queuing in high-performance packet switching', *IEEE J. Selected Areas in Communications*, December 1998, **6** (9), pp. 1587-1597

[124] CLOONAN, T.J.: 'A high bit-rate packet switch architecture with advanced electronic packaging and free-space optical interconnects', PhD thesis, Heriot-Watt University, UK, February 1993

[125] LENTINE, A.L., GOOSEN, K.W., WALKER, J.A., CHIROVSKY, L.M.F.,

D'ASARO, A., HUI, S.P., TSENG, B.J., LEIBENGUTH, R.E., CUNNINGHAM, J.E., JAN, W.Y., HUO, J.-M., DAHRINGER, D.W., KOSSIVES, D.P., BACON, D.D., LIVESCU, G., MORRISON, R.L., NOVOTNY, R.A., BUCHHOLZ, D.B.: 'High-speed optoelectronic VLSI switching chip with > 4000 optical I/O based on flip-chip bonding of MQW modulators and detectors to silicon CMOS', *IEEE J. Selected Topics in Quantum Electron.*, April 1996, **2** (1), pp.77-83

[126] LENTINE, A.L., REILEY, D.J., NOVOTNY, R.A., MORRISON, R.L., SASIAN, J.M., BECKMAN, M.G., BUCHHOLZ, D.B., HINTERLONG, S.J., CLOONAN, T.J.,

RICHARDS, G.W., McCORMICK, F.B.: 'Asynchronous transfer mode distribution network by use of an optoelectronic VLSI switching chip', *Appl. Opt*, 10 March 1997, **36** (8), pp. 1804-1814

[127] DAMES, M.P., COLLINGTON, J.R., CROSSLAND, W.A., SCARR, R.W.A.:

'Scalable approach to high performance ATM switching by the interconnection of free-space photonic crossbar modules', *Electron. Lett.*, 4 July 1996, **32** (14), pp. 1313-1314

[128] SCARR, R.W.A., COLLINGTON, J.R., CROSSLAND, W.A., DAMES, M.P.: 'Highly parallel optics in ATM switching networks', *IEE Proc. Optoelectronics*, April 1997, **144** (2), pp. 53-60

[129] DAMES. M.P.: 'Report on preferred POETS ATM switch architecture, including theoretical analysis of expected performance', Cambridge University internal report, April 1995

[130] ENG, K.Y., KAROL, M.J., YEH, Y.-S.: 'A growable packet (ATM) switch architecture: design principles and applications', *IEEE Trans. Commun.*, February 1992, 40
(2), pp. 423-430

[131] KAROL, M.J., I, C.-L.: 'Performance analysis of a growable architecture for broadband packet (ATM) switching', *IEEE Trans. Commun.*, February 1992, **40** (2), pp. 431-439 [132] GAUTHIER, A., BENEBES, P., KIELBASA, R., FORBES, M., DESMULLIEZ, M.:

'Project SPOEC – system specification', Version 3, SPOEC project internal document, 22 January 1997

[133] SPOEC 12-month progress report, September 1997, pp. 56-86

[134] GAUTHIER, A., BENEBES, P., KIELBASA, R., FORBES, M., DESMULLIEZ, M.:'Report on Optical Crossbar Demonstrator System Design Concept', SPOEC internal report (Deliverable 1), 17 March 1997

[135] MOTOROLA, INC.: 'Timing solutions', Data book BR1333 revision 5, 1995[136] DALLY, W.J., POULTON, J.W.: 'Digital systems engineering', 1998, Cambridge University Press, Chapter 9

[137] KIEFER, D.R., SWANSON, V.W.: 'Implementation of optical clock distribution in a supercomputer' in 'Optical Computing', *1995 OSA Technical Digest Series*, **10**, pp. 260-262
[138] AFGHAHI, M., YUAN, J.: 'Double edge triggered D flip-flops for high-speed CMOS circuits', *IEEE J. Solid-state circuits*, August 1991, **26** (8) pp. 1168-1170

[139] SCHMERBECK, T.J.: 'Noise coupling in mixed-signal ASICs', in 'Low-power HF microelectronics: a unified approach', G.A.S.MACHADO (Ed.), Institution of Electrical Engineers, 1996, pp. 373-430

[140] Foundry internal application note, 'Crosstalk in mixed signal systems', July 1996
[141] SU, D.K., LOINAZ, M.J., MASUI, S., WOOLEY, B.A.: 'Experimental results and modelling techniques for substrate noise in mixed-signal integrated circuits', *IEEE J.Solidstate circuits*, April 1993, **28** (4), pp. 420-430

[142] VERGHESE, N.K., ALLSOT, D.J.: 'Computer-aided design considerations for mixedsignal coupling in RF integrated circuits', *IEEE J. Solid-state circuits*, March 1998, **33** (3), pp. 314-323

[143] CADENCE DESIGN SYSTEMS, INC.: 'Analog-artist simulation help – substrate coupling analysis', on-line documentation, Cadence Release 4.4.1, 1997

[144] CADENCE DESIGN SYSTEMS, INC.: 'Virtuoso layout synthesizer help', on-line documentation, Cadence Release 4.4.1, 1997

[145] DINES, J.: 'Optoelectronic computing: interconnects, architectures and a system demonstrator', PhD thesis, Heriot-Watt University, UK, 1998

[146] KRISHNAMOORTHY, A.V., FORD, J.E., GOOSEN, K.W., WALKER, J.A.,LENTINE, A.L., HUI, S.P., TSENG, B., DAHRINGER, D., D'ASARO, L.A., KIAMILEV,F.E., APLIN, G.F., ROZIER, R.G., MILLER, D.A.B.: 'Photonic page buffer based on GaAs

MQW modulators bonded directly over active silicon CMOS circuits', *Appl. Opt.*, 10 May 1996, **35** (14), pp. 2439-2448

[147] ROZIER, R., FARBARIK, R., KIAMILEV, F., EKMAN, J., CHANDRAMANI, P.,
KRISHNAMOORTHY, A.V., OETTEL, R.: 'Automated design of integrated circuits with area-distributed input-output pads', *Appl. Opt.*, 10 September 1998, **37** (26), pp. 6140-6150
[148] KIAMILEV, F.E., ROZIER, R.G., KRISHNAMOORTHY, A.V.: 'Smart pixel IC layout methodology and its application to photonic page buffers', *Int. J. Optoelectron.*, May-June 1997, **11** (3), pp.199-215

[149] KRISHNAMOORTHY, A.V., ROZIER, R.G., FORD, J.E., KIAMILEV, F. : 'CMOS static RAM chip with high-speed optical read and write', *IEEE Photonics Tech. Lett.*, November 1997, **9** (11), pp. 1517-1519

[150] WOODWARD, T.K., KRISHNAMOORTHY, A.K., LENTINE, A.L., GOOSEN,
K.W., WALKER, J.A., CUNNINGHAM, J.E., YAN, W.Y., D'ASARO, L.A., CHIROVSKY,
L.M.F., HUI, S.P., TSENG, B., KOSSIVES, D., DAHRINGER, D., LEIBENGUTH, R.E.: '1-Gb/s two-beam transimpedance smart-pixel optical receivers made from hybrid GaAs MQW
modulators bonded to 0.8 μm silicon CMOS', *IEEE Photonics Technology Lett.*, March
1996, 8 (3), pp. 422-424

[151] SPOEC 24-month progress report, September 1998, pp.40-42

[152] WILKINSON, L.C.: 'Indium Gallium Arsenide quantum well devices for optically interconnected smart-pixels', PhD Thesis, Heriot-Watt University, UK, 1998, Table 5.1 and pp. 116

[153] BHATTACHARYA, P.: 'Properties of lattice-matched and strained Indium Gallium Arsenide', EMIS Datareview series No. 8 (INSPEC, IEE, London, 1993)

[154] FANCEY, S.J., TAGHIZADEH, M.R., BULLER, G.S., DESMULLIEZ, M.P.Y.,

WALKER, A.C.: 'Optical components of the smart-pixel opto-electronic connection (SPOEC) project', *J.European Opt. Soc. A - Pure and Appl. Opt.*, submitted for publication 1998

[155] WALKER, A.C., DESMULLIEZ, M.P.Y., FORBES, M.G., FANCEY, S.J., BULLER, G.S., TAGHIZADEH, M.R., DINES, J.A.B., STANLEY, C.R., PENNELLI, G., HORAN, P., BYRNE, D., HEGARTY, J., EITEL, S., GAUGGEL, H.-P., GULDEN, K.-H., GAUTHIER, A., BENABES, P., GUTZWILLER, J.L., GOETZ, M.: 'Design and construction of an optoelectronic crossbar containing a terabit/s free-space optical interconnect', submitted to *IEEE J. Selected Topics in Quant. Electron.*, Special issue on Smart Photonic Components, Interconnects and Processing, 1999

[156] FANCEY, S.J., BULLER, G.S., TAGIZADEH, M.R., WALKER, A.C.: 'Report on optical design and optomechanics for SPOEC demonstrator', SPOEC internal report (Deliverable 7), March 1998 [157] WESTE, N.H.E., ESHRAGHIAN , K.: 'Principles of CMOS VLSI design - a system perspective', 2nd edition, 1993, Addison-Wesley, chapter 7 [158] LENTINE, A.L., GOOSSEN, K.W., WALKER, J.A., CUNNINGHAM, J.E., JAN, W.Y., WOODWARD, T.K., KRISHNAMOORTHY, A.V., TSENG, B.J., HUI, S.P., LEIBENGUTH, R.E., CHIROVSKY, L.M.F., NOVOTNY, R.A., BUCHHOLZ, D.B., MORRISON, R.L.: 'Optoelectronic VLSI switching chip with greater than 1 Tbit/s potential optical I/O bandwidth', Electron. Lett., 8 May 1997, 33 (10), pp. 894-895 [159] WOODWARD, T.K., LENTINE, A.L., KRISHNAMOORTHY, A.V., GOOSSEN, K.W., WALKER, J.A., CUNNINGHAM, J.E., JAN, W.Y., TSENG, B.T., HUI, S.P., LEIBENGUTH, R.E.: 'Parallel operation of 50 element two-dimensional CMOS smart-pixel receiver array', Electron. Lett., 14 May 1998, 34 (10), pp.936-937 [160] KRISHNAMOORTHY, A.V., ROZIER, R.G., FORD, J.E., KIAMILEV, F.E.: 'CMOS static RAM chip with high-speed optical read and write', IEEE Photon. Tech. Lett., November 1997, 9 (11), pp. 1517-1519 [161] ROBERTSON, B.: 'Design of an optical interconnect for photonic backplane applications', Appl. Opt., 10 May 1998, 37 (14), pp. 2975-2984 [162] KIRK, A.G., BROSSEAU, D.F., LACROIX, F.K., BERNIER, E., AYLIFFE, M.H., ROBERTSON, B., TOOLEY, F.A.P., PLANT, D.V.: 'Design and implementation of twostage optical power supply spot array generator for a modulator-based free-space interconnect', Optics in Computing 1998, pp. 48-51 [163] LENTINE, A.L., GOOSEN, K.W., WALKER, J.A., CUNNINGHAM, J.E., JAN, W.Y., WOODWARD, T.K., KRISHNAMOORTHY, A.V., TSENG, B.J., HUI, S.P., LEIBENGUTH, R.E., CHIROVSKY, L.M.F., NOVOTNY, R.A., BUCHHOLZ, D.B., MORRISON, R.L.: 'High throughput optoelectronic VLSI switching chips', in Spatial Light Modulators, Technical Digest (Optical Society of America, Washington DC, 1997), pp. 8-10 [164] TRAVERS, C.M., HESSENBRUCH, J.M., KIM, J., STONE, R.V., GUILFOYLE: 'CMOS compatible free-space optical interconnects', Optics in Conputing 1998, Proc. SPIE, 3490, pp. 560-563 [165] TRAVERS, C.M., HESSENBRUCH, J.M., KIM, J., STONE, R.V., GUILFOYLE, P.S., KIAMILEV, F.: 'VLSI photonic smart pixel array for I/O system architectures', SPIE

Photonics West Conf., January 1998, http://www.opticomp.com/Publishing.html

[166] NOVOTNY, R.A.: 'Analysis of smart pixel digital logic and optical interconnections', PhD thesis, Heriot-Watt University, UK, 1996

[167] KRISHNAMOORTHY, A.V., MILLER, D.A.B.: 'Scaling optoelectronic-VLSI circuits into the 21st century: a technology roadmap', *IEEE J. Selected Topics in Quantum Electron.*, April 1996, **2** (1), pp.55-76

[168] VAN BLERKOM, D.A., FAN, C., BLUME, M., ESENER, S.C.: 'Transimpedance receiver design optimization for smart pixel arrays', *J. Lightwave Tech.*, January 1998, 16 (1), pp.119-126

[169] WILLIAMS, G.F: 'Active feedback lightwave receivers', *J. Lightwave Tech.*, October 1986, **4** (10), pp.1502-1507

[170] ABIDI,A.A: 'Gigahertz transresistance amplifiers in fine line NMOS', *IEEE J. Solid-state Circuits*, December 1984, **SC-19** (6), pp. 986-994

[171] MOLLER, M., REIN, H.-M., WERNZ, H.: '13 Gbit/s Si-bipolar AGC amplifier IC with high gain and wide dynamic range for optical fiber receivers', *IEEE J. Solid-state circuits*, July 1994, **29** (7), pp. 815-822

[172] REIMANN, R., REIN, H.-M.: 'Bipolar high-gain limiting amplifier IC for optical fiber receivers operating up to 4 Gbit/s', *IEEE J. Solid-state circuits*, August 1987, **22** (4), pp. 504-511

[173] RAZAVI, B.: 'A 2.5 Gbit/s 15 mW clock recovery circuit', *IEEE J. Solid-state circuits*,
 April 1996, **31** (4), pp. 472-480

[174] WOODWARD, T.K., KRISHNAMOORTHY, A.V., LENTINE, A.L., CHIROVSKY,
L.M.F.: 'Optical receivers for optoelectronic VLSI', *IEEE J. Selected Topics in Quantum Electron.*, April 1996, 2 (1), pp.106-116

[175] WILSON, B., DARWAZEH, I.: 'Transimpedance optical preamplifier with a very low input resistance', *Electron. Lett.*, 12 February 1997, **23** (4), pp. 138-139

[176] VANISRI, T., TOUMAZOU, C.: 'Integrated high frequency low-noise current-mode optical transimpedance preamplifiers: theory and practice', *IEEE J. Solid-state circuits*, June 1995, **30** (6), pp. 677-685

[177] TOUMAZOU, C., PARK, S.M.: 'Wideband low noise CMOS transimpedance amplifier for gigahertz operation', *Electron. Lett.*, 20 June 1996, **32** (13), pp. 1194-1196

[178] GRAY, P.R., MEYER, R.G.: 'Analysis and design of analog integrated circuits (3rd
 Edition)', 1993, Wiley, p.638

[179] MIYAKE, K., NAMBA, T., HASHIMOTO, K., SAKAUE, H., MIYAZAKI, S., HORIKE, Y., YOKOYAMA, S., KOYANAGI, M., HIROSE, M.: 'Fabrication and evaluation of three-dimensional optically coupled common memory', *Jpn. J. Appl. Phys. Part 1*, February 1995, **34** (2B), pp.1246-1248

[180] DINES, J.A.B: 'Optoelectronic computing: interconnects, architectures and a system demonstrator', PhD thesis, Heriot-Watt University, UK, 1998, Chapter 4

[181] DINES, J.A.B: 'Smart pixel optoelectronic receiver based on a charge sensitive amplifier design', *IEEE J. Selected Topics in Quantum Electron.*, April 1996, 2 (1), pp. 117-120

[182] AYADI,K., KUIJK,M., HEREMANS,P., BICKEL,G., BORGHS,G., VOUNCKX,R.:

'A monolithic optoelectronic receiver in standard 0.7 μm CMOS operating at 180 MHz and 176-fJ light input energy', *IEEE Photonics Tech. Lett.*, January 1997, **9** (1) pp. 88-90

[183] WOODWARD, T.K., KRISHNAMOORTHY, A.V. GOOSEN, K.W., WALKER, J.A.,

CHIROVSKY, L.M.F., HUI, S.P., TSENG, B., KOSSIVES, D., DAHRINGER, D.,

LEIBENGUTH, R.E., CUNNINGHAM, J.E., JAN, W.Y.: 'Synchronous sense amplifier based optical receivers for smart-pixel applications', Optical computing, Sendai, April 1996, pp.80-81

[184] SHANG, A.Z., TOOLEY, F.A.P: 'A high-sensitivity and low-power smart-pixel receiver', Optical computing, Sendai, April 1996, pp. 96-97

[185] DALLY, W.J., POULTON, J.W.: 'Digital systems engineering', 1998, Cambridge University Press, pp. 540-548

[186] WOODWARD, T.K., KRISHNAMOORTHY, A.V. GOOSEN, K.W., WALKER, J.A., CUNNINGHAM, J.E., JAN, W.Y, CHIROVSKY, L.M.F., HUI, S.P., TSENG, B.,

KOSSIVES, D., DAHRINGER, D., BACON, D., LEIBENGUTH, R.E.: 'Clocked senseamplifier based smart-pixel optical receiver', *IEEE Photonics Tech. Lett.*, August 1996, **8** (8), pp. 1067-1069

[187] SHANG, A.Z.: 'Transceiver arrays for optically interconnected electronic systems',PhD thesis, McGill University, Canada, 1997

[188] TOOLEY, F., SINHA, P., SHANG, A.: 'Time-differential operation of an optical receiver', *Optics in Computing*, 1997, **8**, OSA Technical Digest Series, pp.73-75

[189] BAGHERI, M., TSIVIDIS, Y.: 'A small signal dc-to-high-frequency nonquasistatic model for the four-terminal MOSFET valid in all regions of operation', *IEEE Trans*.

Electron. Devices, November 1985, 32 (11), pp. 2383-2391

[190] PERSONICK, S.D.: 'Receiver design for digital fiber optic communication systems, I and II', *Bell System Technical J.*, July 1973, **52** (6), pp. 843-886

[191] SMITH, R.G., PERSONICK, S.D.: 'Receiver design for optical fiber communication systems', in *Semiconductor Devices for Optical Communications (Topics in Applied Physics v. 39)*, 2nd edition, Kressel, H.G. (ed.), Springer-Verlag, 1982, pp.89-159

[192] MORIKUNI, J.J., DHARCHOUDHURY, A., LEBLEBICI, Y., KANG, S.:

'Improvements to the standard theory for photoreceiver noise', *J. Lightwave Tech.*, July 1994, **12** (4) pp.1174-1183

[193] ABIDI, A.A: 'High-frequency noise measurements on FET's with small dimensions', *IEEE Trans. Electron Devices*, November 1996, **ED-32** (11), pp.1801-1805

[194] JINDAL, R.P.: 'Hot-electron effects on channel thermal noise in fine-line NMOS field-effect transistors', *IEEE Trans. Electron Devices*, September 1986, ED-33 (9), pp.1395-1397
[195] TRIANTIS, D.P., BIRBAS, ALEXIOS, A.N., KONDIS, D.: 'Thermal noise modelling for short-channel MOSFET's', *IEEE Trans. Electron Devices*, November 1996, ED-43 (11), pp.1950-1955

[196] MAXIM INTEGRATED PRODUCTS INC.: 'MAX3664 622Mbps, Ultra-Low-Power
3.3VTransimpedance Preamplifier for SDH/SONET', 'MAX3675 622Mbps, low-power,
3.3V clock recovery and data-retiming IC with limiting amplifier', product datasheets, 1996,
1997, http://www.maxim-ic.com/

[197] WOODWARD, T.K., KRISHNAMOORTHY, A.K., LENTINE, A.L., GOOSEN,
K.W., WALKER, J.A., CUNNINGHAM, J.E., YAN, W.Y., D'ASARO, L.A., CHIROVSKY,
L.M.F., HUI, S.P., TSENG, B., KOSSIVES, D., DAHRINGER, D., LEIBENGUTH, R.E.: '1-Gb/s two-beam transimpedance smart-pixel optical receivers made from hybrid GaAs MQW
modulators bonded to 0.8 μm silicon CMOS', *IEEE Photonics Technology Lett.*, March
1996, 8 (3), pp. 422-424

[198] WILLIAMS, G.F.: 'Lightwave receivers' in *Topics in Lightwave systems*, Li, T. (ed.), Academic, 1991, pp.79-149

[199] GRAY, P.R., MEYER, R.G.: 'Analysis and design of analog integrated circuits (3rd Edition)', 1993, Wiley, p.314

[200] HU, T.H., GRAY, P.R.: 'Monolithic 480 Mb/s parallel AGC/decision/clock-recovery circuit in 1.2 μm CMOS', *IEEE J. Solid-State Circuits*, December 1993, **28** (12) pp.1314-1320

[201] WIDMER, A.X., FRANASZEK, P.A.: 'A DC balanced, partitioned-block 8B/10B transmission code', *IBM J. Res. Develop.*, September 1983, **27** (5), pp.440-451
[202] OTA, Y., SWARTZ, R.G., ARCHER III, V.D., KOROTKY, S.K., BANU, M., DUNLOP, A.E.: 'High-speed, burst-mode packet capable optical receiver and instantaneous

clock recovery for optical bus operation', *J. Lightwave Tech.*, February 1994, **12** (2), pp. 325-331

[203] SU, C., CHEN, L.-K., CHEUNG, K.-W., 'Theory of burst-mode receiver and its applications in optical multiaccess networks', *J. Lightwave Tech.*, April 1997, **15** (4), pp.590-606

[204] LI, C-S., STONE, H.S., KWARK, Y., OLSEN, C.M: 'Fully differential optical interconnections for high-speed digital systems', June 1993, *IEEE Trans. VLSI Systems*, 1 (2), pp. 151-163

[205] BASTOS, J., STEYAERT, M.S.J., PERGOOT, A., SANSEN, W.M.: 'Influence of die attachment on MOS transistor matching', *IEEE Trans. Semiconductor Manufacturing*, May 1997, **10** (2), pp.209-218

[206] PELGROM, M.J.M, DUINMAIJER, A.C.J., WELBERS, A.P.G.: 'Matching properties of MOS transistors', *IEEE J. Solid-state Circuits*, October 1989, 24 (5), pp. 1433-1440
[207] MIZUNO, T., OKAMURA, J-I., TORIUMI, A.: 'Experimental study of threshold voltage fluctuation due to statistical variation of channel dopant number in MOSFETs', *IEEE Trans. Electron Devices*, November 1994, 41 (11), pp.2216-2221

[208] LAKSHMIKUMAR, K.R., HADAWAY, R.A., COPELAND, M.A.: 'Characterization and modeling of mismatch in MOS transistors for precision analog design', *IEEE J. Solid-State Circuits*, December 1986, **21** (6), pp. 1057-1066

[209] BASTOS, J.: 'Characterization of MOS transistor mismatch for analog design', PhD thesis, Katholieke Universiteit Leuven, Belgium, 1998

[210] STEYAERT, M., BASTOS, J., ROOVERS, R., KINGET, P., SANSEN, W.,

[211] BASTOS, J., STEYAERT, M., PERGOOT, A., SANSEN, W.: 'Mismatch

GRAINDOURZE, B., PERGOOT, A., JANSSENS, Er.: 'Threshold voltage mismatch in short-channel MOS transistors', *Electron. Lett.*, 1 September 1994, **30** (18), pp. 1546-1548

characterisation of submicron MOS transistors', *Analog Integrated Circuits and Signal Processing*, 1997, **12** (2), pp.95-106

[212] CHENG, Y., CHAN, M., HUI, K., JENG, M.-C., LIU, Z., HUANG, J., CHEN, K., CHEN, J., TU, R., KO, P.K., HU, C.: 'BSIM3v3 Manual', 1996, Dept. of Electrical Engineering and Computer Sciences, University of California, Berkeley (http://wwwdevice.eecs.berkeley.edu/~bsim3)

[213] KAWAHARA, T., SAKATA, T., ITOH, K., KAWAJIRI, Y., AKIBA, T., KITSUKAWA, G., AOKI, M.: 'A high-speed, small-area, threshold-voltage-mismatch compensation sense amplifier for gigabit-scale DRAM arrays', *IEEE J. Solid-State Circuits*, July 1993, **28** (7), pp. 816-823

[214] Technical details are available from Europractice:

http://www.imec.be:8000/europractice/on-line-docs/prototyping/ti/ti_mtc0u7.html

[215] RAZAVI, B., WOOLEY, B.A.: 'Design techniques for high-speed, high-resolution comparators', *IEEE J. Solid-state circuits*, December 1992, **27** (12), pp.1916-1926

[216] YOON, T., JALALI, B.: '1 Gbit/s fibre channel CMOS transimpedance amplifier', *Electron. Lett.*, 27 March 1997, **33** (7), pp. 588-589

[217] NAGARJAN, R., SHA, W., LI, B., CRAIG, R.: 'Gigabyte/s parallel fiber-optic links based on edge emitting laser diode arrays', *J. Lightwave Tech.*, May 1998, **16** (5), pp. 778-787

[218] METASOFT INC.: 'HSpice user manual', 1997, Chapter 25

[219] OH, S.-Y., WARD, D.E., DUTTON, R.W.: 'Transient analysis of MOS transistors', *IEEE J. Solid-state circuits*, August 1980, **15** (4), pp. 636-643

[220] FOSSUM, J.G., JEONG, H., VEERARAGHAVAN, S.: 'Significance of the channelcharge partition in the transient MOSFET model', *IEEE Trans. Electron Devices*, October 1986, **33** (10), pp. 1621-1623

[221] PAULOS, J.J., ANTONIADIS, D.A.: 'Limitations of quasi-static capacitance models for the MOS transistor', *IEEE Electron Device Lett.*, July 1983, **4** (7), pp. 221-224

[222] SPOEC internal report: 'Supply of 8×8 VCSEL array as input device for demonstrator (Deliverable 4)', April 1998

[223] WILLIAMS, R.: 'SCIOS Memory Chip - EU32x32SR', 1996, University of Edinburgh internal report

[224] LENTINE, A.L., GOOSEN, K.W., WALKER, J.A., CHIROVSKY, L.M.F.,

D'ASARO, A., HUI, S.P., TSENG, B.J., LEIBENGUTH, R.E., CUNNINGHAM, J.E., JAN, W.Y., HUO, J.-M., DAHRINGER, D.W., KOSSIVES, D.P., BACON, D.D., LIVESCU, G., MORRISON, R.L., NOVOTNY, R.A., BUCHHOLZ, D.B.: 'High-speed optoelectronic VLSI switching chip with > 4000 optical I/O based on flip-chip bonding of MQW modulators and detectors to silicon CMOS', *IEEE J. Selected Topics in Quantum Electron.*, April 1996, **2** (1), pp.77-83

[225] JOHNS, D.A., MARTIN, K.: 'Analog integrated circuit design', 1997, Wiley, pp. 115-118

[226] TSIVIDIS, Y.P., SUYAMA, K.: 'MOSFET modelling for analog circuit CAD: problems and prospects', *IEEE J. Solid-state circuits*, March 1994, **29** (3), pp. 210-216

[227] CHAN, M., HUI, K.Y., HU, C., KO, P.K.: 'A robust and physical BSIM3 non-quasistatic transient and AC small-signal model for circuit simulation', *IEEE Trans. Electron Devices*, April 1998, **45** (4), pp. 834-841

[228] PAULOS, J.J., ANTONIADIS, D.A.: 'Limitations of quasi-static capacitance models for the MOS transistor', *IEEE Electron Device Lett.*, July 1993, **4** (7), pp. 221-224

[229] JOHNSON, H.W., GRAHAM, M.: 'High-speed digital design – a handbook of black magic', Prentice-Hall, 1993, p.98

[230] INGELS, M., VAN DER PLAS, G., CROLS, J., STEYAERT, M.: 'A CMOS 18

THz Ω 240 Mbit/s transimpedance amplifier and 155 Mbit/s LED-driver for low cost optical fiber links', *IEEE J. Solid-state circuits*, December 1994, **29** (12), pp. 1552-1559

[231] KRISHNAMOORTHY, A. V., WOODWARD, T. K., GOOSSEN, K. W., WALKER,

J. A., LENTINE, A. L., CHIROVKSY, L. M. F., HUI, S. P., TSENG, B., LEIBENGUTH, R.,

CUNNINGHAM, J. E., JAN, W. Y.: 'Operation of a single-ended 550 Mbit/s, 41fJ, hybrid

CMOS/MQW receiver-transmitter', Electron. Lett., April 1996, 32 (8), pp.764-766

[232] LSI LOGIC INC.: 'G12 CMOS Techology 0.18 μm drawn / 0.13 μm Leff', product information, 1998, http://www.lsilogic.com/products/unit5_2j.html

[233] TEXAS INSTRUMENTS INC.: '0.07-Micron CMOS Technology Ushers in Era of gigahertz DSP and analog performance', News Release, 26 August 1998,

http://www.ti.com/sc/docs/news/1998/98079.htm

[234] DENNARD, R.H., GAENSSLEN, F.H., YU, H.N., RIDEOUT, V.L., BASSOUS, E.,

LEBLANC, A.R.: 'Design of ion-implanted MOSFETs with very small physical

dimensions', IEEE J. Solid-state circuits, October 1974, 9 (5), pp. 256-268

[235] LAKER, K.R., SANSEN, W.M.C: 'Design of analog integrated circuits and systems', 1994, Mc-Graw Hill, Chapter 1

[236] WESTE, N.H.E, ESHRAGHIAN, K.: 'Principles of CMOS VLSI design: a systems perspective', 2nd Edition, Addison-Wesley, 1993, pp. 250-256

[237] Semiconductor Industry Association: 'The national technology roadmap for semiconductors', 1997, Tables 14, 15, 22

[238] KOBURGER III, C.W., CLARK, W.F., ADKISSON, J.W., ADLER, E., BAKEMAN,
P.E., BERGENDAHL, A.S., BOTULA, A.B., CHANG, W., DAVARI, B., GIVENS, J.H.,
HANSEN, H.H., HOLMES, S.J., HORAK, D.V., LAM, C.H., LASKY, J.B., LUCE, S.E.,
MANN, R.W., MILES, G.L., NAKOS, J.S., NOWAK, E.J., SHAHIDI, G., TAUR, Y.,
WHITE, F.R., WORDEMAN, M.R.: 'A half-micron CMOS logic generation', *IBM J. Res. Develop.*, January/March 1995, **39** (1/2), pp.215-228

[239] SHAHIDI, G.G, WARNOCK, J.D., COMFORT, J., FISCHER, S., McFARLAND,

P.A., ACOVIC, A., CHAPPELL, T.I., CHAPPELL, B.A., NING, T.H., ANDERSON, C.J.,

DENNARD, R.H., SUN, J., Y.-C., POLCARI, M.R., DAVARI, B.: 'CMOS scaling in the 0.1

 μ m, 1.X volt regime for high-performance applications', *IBM J. Res. Develop.*,

January/March 1995, 39 (1/2), pp.229-244

[240] TAUR, Y., MII, Y.-J., FRANK, D.J., WONG, H.-S., BUCHANAN, D.A., WIND, S.J., RISHTON, S.A., SAI-HALASZ, G.A., NOWAK, E.J.: 'CMOS scaling into the 21st century:

0.1 µm and beyond', *IBM J. Res. Develop.*, January/March 1995, **39** (1/2), pp.245-260

[241] IBM CORPORATION: 'BlueLogic[™] BiCMOS 4S Technology – a proven 0.8 μm BiCMOS process', June 1998,

http://www.chips.ibm.com/services/foundry/offerings/bicmos/4s/

[242] IBM CORPORARION: 'BlueLogic[™] BiCMOS 5S Technology – a proven 0.5 μm BiCMOS process', June 1998,

http://www.chips.ibm.com/services/foundry/offerings/bicmos/5s/

[243] SZE, S.M.: 'Physics of semiconductor devices (2nd edition)', 1981, Wiley, pp. 446-448
[244] LAKER, K.R., SANSEN, W.M.C: 'Design of analog integrated circuits and systems', McGraw Hill, International Edition, 1994, p.31

[245] TAUR, Y., BUCHANAN, D.A., CHEN, W., FRANK, D.J., ISMAIL, K.E., LO, S-H.,
SAI-HALASZ, G.A., VISWANATHAN, R.G., WANN, H-J.C., WIND, S.J., WONG, H-S.:
'CMOS scaling into the nanometer regime', *Proc. IEEE*, April 1997, **85** (4), pp.486-504
[246] ASAI, S., WADA, Y.: 'Technology challenges for integration near and below 0.1 μm', *Proc. IEEE*., April 1997, **85** (4), pp.505-520

[247] HUANG, Q., PIAZZA, F., ORSATTI, P., OHGURO, T.: 'The impact of scaling down to deep submicron on CMOS RF circuits', *IEEE J. Solid-state Circuits*, July 1998, **33** (7), pp.1023-1036

[248] SHRIVASTAVA, R., FITZPATRICK, K.: 'A simple model for the overlap capacitance of a VLSI MOS Device', *IEEE Trans.Electron Devices*, December 1983, **29** (12), pp.1870-1875

[249] RYAN, J.G., GEFFKEN, R.M., POULIN, N.R., PARASZCZAK, J.R.: 'The evolution of interconnection technology at IBM', *IBM J. Res. Develop*, July 1995, **39** (4), pp.371-381 [250] SARASWAT, K.C., MOHAMADI, F.: 'Effect of scaling of interconnections on the time delay of VLSI circuits', *IEEE J. Solid-state Circuits*, April 1982, **17** (2), pp.275-280 [251] GRAY, P.R., MEYER, R.G.: 'Analysis and design of analog integrated circuits (3rd edition)', 1993, Wiley, p. 6

[252] IBM Press release, 3rd August 1998, http://www.ibm.com/news/1998/08/03.phtml

[253] BOHR, M.T.: 'Technology for advanced high-performance microprocessors', *IEEE Trans. Electron. Devices*, March 1998, **45** (3), pp. 620-625

[254] UYEMURA, J.P.: 'Fundamentals of MOS digital integrated circuits', Addison-Wesley, 1998, pp. 147-148

[255] BOHR, M.T: 'Dense electrical interconnects for ULSI circuits', *ESPRIT MEL-ARI First Annual Workshop*, October 1997, Zurich

[256] FORCHEL, A., MALINVERNI, P. (Editors): 'ESPRIT Microelectronics Advanced Research Initiative: Optoelectronic interconnections for integrated circuits (MEL-ARI OPTO) Technology Roadmap' Revision 1.4, April 1998: Projection for 2007 VCSEL performance

[257] CHERRY, E.M., HOOPER, D.E.: 'The design of wide-band transistor feedback amplifiers', *Proc. IEEE*, February 1963, **110** (2), pp.375-389

[258] GRAY, P.R., MEYER, R.G.: 'Analysis and design of Analog Integrated Circuits' (3rd edition), Wiley, 1993, pp. 594

[259] SHENG, S., LYNN, L., PEROULAS, J., STONE, K., O'DONNELL, I., and

BRODERSEN,R.: 'A low-power CMOS chipset for spread-spectrum communications',

Digest of Technical Papers - IEEE International Solid-State Circuits Conference, 1996, **39**, pp.346-347

[260] PAN, T-W., ABIDI, A.A: 'A wide-band CMOS read amplifier for magnetic data storage systems', *IEEE J. Solid-State Circuits*, June 1992, **27** (6), pp. 863-873

[261] MOLLER, M., REIN, H.-M., WERNZ, H.: '13 Gbit/s Si-bipolar AGC amplifier IC with high gain and wide dynamic range for optical fiber receivers', *IEEE J. Solid-state circuits*, July 1994, **29** (7), pp. 815-822

[262] ABIDI, A.A: 'Gigahertz transresistance amplifiers in fine line NMOS', *IEEE J. Solid-state circuits*, December 1984, **19** (6), pp. 986-994

[263] WANG, C., HUANG, P.-C., HUANG, C.-Y.: 'A fully differential CMOS

transconductance transimpedance wideband amplifier', IEEE Trans. Circuits Syst. II -

Analog and Digital Signal Processing, November 1995, 42 (11), pp. 745-748

[264] HUANG, P.-C., WANG, C., WU, W.C., and WANG, Y.D.: 'Fully differential CMOS transconductance-transimpedance wide-band amplifier', U.S. Patent #5 451 902, 19 September 1995

[265] PETRI, C., ROCCHI, S., VIGNOLI, V.: 'High dynamic CMOS preamplifiers for quantum-well diodes', *Electron. Lett.*, 30 April 1998, **34** (9), pp. 877-878

[266] DALLY, W.J., POULTON, J.W.: 'Digital systems engineering', Cambridge University Press, 1998, pp. 282-285

[267] JOHNS, D.A., MARTIN, K.: 'Analog integrated circuit design', Wiley, 1997, pp. 246-251

[268] FORBES, M.G., WALKER, A.C.: 'Wideband transconductance-transimpedance post-amplifier for large photoreceiver arrays', *Electron. Lett.*, 19 March 1988, 34 (6), pp. 589-590
[269] KRISHNAMOORTHY, A.V., WOODWARD, T.K., NOVOTNY, R.A., GOOSSEN, K.W., WALKER, J.A., LENTINE, A.L., D'ASARO, L.A., HUI, S.P., TSENG, B., LEIBENGUTH, R., KOSSIVES, D., DAHRINGER, D., CHIROVSKY, L.M.F., APLIN, G.F., ROZIER, R.G., KIAMILEV, F.E., and MILLER, D.A.B.: 'Ring oscillators with optical and electrical readout based on hybrid GaAs MQW modulators bonded to 0.8µm silicon
VLSI circuits', *Electron. Lett.*, 26 October 1995, 31 (22) pp.1917-1918

[270] WOODWARD, T.K., KRISHNAMOORTHY, A.V., LENTINE, A.L., GOOSSEN,

K.W., WALKER, J.A., CUNNINGHAM, J.E., JAN, W.Y., D'ASARO, L.A.,

CHIROVSKY,L.M.F., HUI, S.P., TSENG, B., KOSSIVES, D., DAHRINGER, D., and LEIBENGUTH, R.E.: '1-Gb/s two-beam transimpedance smart-pixel optical receivers made from hybrid GaAs MQW modulators bonded to 0.8µm silicon CMOS', *IEEE Photonics Technol. Lett.*, March 1996, **8** (3), pp.422-424

[271] MUTAGI, R.N.: 'Pseudo noise sequences for engineers', *Electron. & Comm. Eng. J.*, April 1996, **8** (2), pp.79-87

[272] NOVOTNY, R.A., WOJCIK, M.J., LENTINE, A.L., CHIROVSKY, L.M.F.,

D'ASARO,L.A., FOCHT, M.W., GUTH, G., GLOGOVSKY, K.G., LEIBENGUTH, R.,

ASOM, M.T., FREUND, J.M.: 'Field effect transistor-self electrooptic effect device (FET-

SEED) differential transimpedance amplifiers for two-dimensional optical data links', J.

Lightwave Tech., April 1995, 13 (4), pp. 606-614

[273] GRAY, P.R., MEYER, R.G.: *ibid.*, pp. 638-642

[274] JOHNS, D.A., MARTIN, K: 'Analog integrated circuit design', Wiley, 1997, pp. 246-253

[275] DALLY, W.J., POULTON, J.W.: ibid., pp. 199-207

[276] WOODWARD, T.K., LENTINE, A.L., KRISHNAMOORTHY, A.V., GOOSSEN,

K.W., WALKER, J.A., CUNNINGHAM, J.E., JAN, W.Y., TSENG, B.T., HUI, S.P.,

LEIBENGUTH, R.E.: 'Parallel operation of 50 element two-dimensional CMOS smart-pixel receiver array', *Electron. Lett.*, 14 May 1998, **34** (10), pp.936-937

[277] GRAY, P.R., MEYER, R.G.: *ibid.*, Chapter 11

[278] ANSOFT, INC.: 'Ansoft SI product family data sheet', 1996

[279] KAMON, M. SILVEIRA, M., SMITHHISLER, C., WHITE, J.: 'FastHenry User's

guide', 11 November 1996, version 3.0, available from ftp://rle-vlsi.mit.edu/pub/fasthenry

[280] KAMON, M.: 'Fast parasitic extraction and simulation of three-dimensional

interconnect via quasistatic analysis', PhD Thesis, 1998, Massachusetts Institute of Technology, USA

[281] RUEHLI, A.E.: 'Survey of computer-aided electrical analysis of integrated circuit interconnections', *IBM J, Res. and Develop.*, November 1979, **23** (6), pp. 626-639

[282] TSAI, C.T., YIP, W.Y.: 'Experimental technique for full package inductance matrix characterization', *IEEE Trans. Components, Packaging and Manufacturing Technol. Part B: Advanced packaging*, May 1996, **19** (2), pp. 338-343

[283] KAMON, K., TSUK, M.J., WHITE, J.: 'FastHenry: a multipole-accelerated 3D inductance extraction program', *IEEE Trans. Microwave Theory and Techniques*, September 1994, **42** (9), pp. 1750-1758

[284] RUEHLI, A.E.: 'Inductance calculations in a complex integrated circuit environment', *IBM J. Res. and Develop.*, September 1972, **16**, pp. 470-481

[285] KYOCERA FINECERAMICS LTD.: private communication relating to 208-pin PGA part number PB-P88639. Test conditions (location of ground plane) unspecified.

[286] DALLY, W.J., POULTON, J.W., *ibid.*, Chapter 5

[287] JOHNSON, H.W., GRAHAM, M.: ibid., Chapter 8

[288] DALLY, W.J., POULTON, J.W.: ibid., Chapter 6

[289] WIELAND, J., MELCHIOR, H., KEARLEY, M.Q., MORRIS, C.M., MOSELEY,

A.M., GOODWIN, M.J., GOODFELLOW, R.C.: 'Optical receiver array in silicon bipolar technology with self-aligned, low parasitic III-V detectors for DC-1 Gbit/s parallel links', *Electron. Lett.*, 21 November 1991, **27** (24), pp. 2211-2213

[290] CHAR, B.W., GEDDES, K.O., GONNET, G.H., LEONG, B.L., MONAGAR, M.B.,

WATT, S.M.: 'Maple V Library Reference Manual (1st edition)', Waterloo Maple Publishing, 1991, pp. 50-51

[291] ELMORE, W.C.: 'The transient response of damped linear networks with particular regard to wide-band amplifiers', *J. Appl. Phys.*, January 1948, **19** (1), pp. 55-63