Dynamic Holography Using Ferroelectric Liquid Crystal on Silicon Spatial Light Modulators

A dissertation submitted for the degree of
Doctor of Philosophy
at the University of Cambridge

Kim Leong Tan

Wolfson College

February 1999

Declaration

This dissertation contains the results of research undertaken by the author between October 1995 and December 1998 at the Engineering Department of Cambridge University. No part of this dissertation is the result of work done in collaboration with others, except where explicitly described or cited within the text. The contents have not been submitted, in whole or in part, for any other University degree or diploma.

Kimleong

Kim Leong Tan
Cambridge, February 1999

Keywords

CMOS VLSI silicon backplane, computer generated holograms, diffraction efficiency, dynamic holography, ferroelectric liquid crystals, free-space holographic switch, multi-level phase holograms, spatial light modulators.

Summary

Optical networking is expected to evolve from the current point-to-point photonic transport networks. Optical wavelength- and space-division interconnects allow for a fuller exploitation of the intrinsic bandwidths of optical fibres as *the* transmission medium. The use of costly electronic Time Division Multiplexing (TDM) switches in high bit-rate telecommunications systems can also be reduced/eliminated. These networks require dynamic and transparent Wavelength Division Multiplexing (WDM) filters and Optical Cross-Connects (OXC). Dynamic holograms recorded on reconfigurable ferroelectric liquid crystal (FLC) spatial light modulators (SLM) have the potential to meet these requirements.

The research outlined in this dissertation relates to an original theoretical analysis, validated by numerical simulations and experimental measurements, of the performance of these routing holograms when used in 4*f* coherent optical systems. The principal analytical results demonstrated are *the dependence of the hologram replay crosstalk on the clipping of the Gaussian beam by the hologram aperture* and *the predictability of the locations and intensities of the hologram replay peaks.* The level of sidelobe ripple noise and the effects of phase quantisation, arrangement of phase elements, spatial quantisation and SLM pixel dead-space are quantified within the valid regime of paraxial diffraction theory. The understanding of routing hologram replay has been exploited to devise a novel deterministic hologram generation algorithm and to determine the scalability of these switches as constrained by the switch path, temporal loss variations and optical bandwidths. It can be also used to design hologram configurations for a large free-space optical matrix switch.

The first silicon backplane for driving a holographic SLM was designed and commercially fabricated during the course of this research. The emphasis of the backplane design and the SLM assembly was to obtain good optical modulation for coherent applications. The backplane contains a binary and a quaternary modulator array having 540 × 1 pixels. The silicon backplane and all assembled SLMs were fully tested and characterised. The binary-phase SLM, with a high quality post-processed mirror array, was to be used in a collaborative free-space 1 × 8 optical switch demonstration. Several optical experiments were performed using a fixed intensity grating on glass and a reconfigurable binary-phase SLM with a view of verifying the crosstalk isolation and insertion loss aspects of routing hologram analyses. These experiments simulating *infinite* switch configurations gave > 50 dB crosstalk figures and insertion losses to within 3 dB of the theoretical values. Two hologram refresh schemes for maintaining a holographic interconnect over a long period were also evaluated.

Acknowledgements

The work reported here owes much to the unique environment in the Photonics and Sensors Group. For that, I would like to thank Bill Crossland with the help of Caryn Wilkinson, for leading the group and providing the resources and interactions with fellow researchers from elsewhere in my chosen field of research. Many thanks are also due to Bill for sorting out my maintenance support. My gratitude is also extended to Mike Robinson at Sharp Labs. Europe for starting the ball rolling by suggesting work on analogue SLMs. I would like to thank Robert Mears, my supervisor at Cambridge, for his occasional but critical appraisals of my work and for the tremendous help in the preparation of papers and this dissertation.

I would especially like to thank Tim for his help in SLM processing/fabrications and proof reading the draft dissertation; Steve for fruitful interactions on hologram analysis and replay; Maura for liquid crystal alignment work; Ilias for help in mounting the fibre on the quarter disc and optical system aligning; Ping Chiek for varied assistance throughout the two and a half years as well as encouragement when I have been frustrated and dispirited; James Collington for demonstrating the VLSI layout using Mentor Graphics; Ian Underwood and David Holburn for contributing to the chip design review; Mike Hands for loaning the stepper motor stages; Mike Bradley for helping with C-programming; Mike Parker for discussions on quaternary hologram analysis; Sazzad Nasir at Wolfson college for numerous mid-night discussions on the wonderful world of physics and mathematics research; Tony Davey for his critical opinions on liquid crystals; Eddie, Sabesan, Cheng-Huan, Karsten, Anna, Huan, Jeon, Clarence, Terence, Faheem, Louis, Niall, Adam, Brian, Mark and many others for friendship, help, comments, advice, etc.; Overseas Research Studentship, Cambridge Commonwealth Trust and Thomas Swan Co. Ltd. for providing financial assistance and making my existence at Cambridge possible!

The support by the workshop technicians, Russell, Steve, Adrian and Mick is greatly appreciated. The Roses project with its industrial and university partners has provided many opportunities for interactions and the funding for the chip development. My sincere thanks to them all and especially to CRL for bonding the SLMs.

Lastly, I am indebted to my wife, Sian, for her never ending love and support and in putting up with the inevitable maniacal work routines during the final stages of my Ph.D.

Author's note: the image on the cover page is the fibre-scan intensity pattern of a grating replay, see §8.5.1.

To Sian

and

our families

Contents

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

CHAPTER 9

APPENDIX A

APPENDIX B

APPENDIX C

APPENDIX D

List of figures

List of tables

CHAPTER 1

Introduction

The strengths of optics and electronics complement each other. Whereas optics is associated with 2-D parallelism and non-interference of crossed paths, electronics is renowned for high speed logic operations, signal buffering and physical wiring for interconnections. The use of electronic chips is prevalent in computing and data processing. Optics on the other hand has found applications in high-bandwidth telecommunication links using fibres and optical storage in the forms of volume holograms and optical disks. In the overlapping domain of telecommunications switching, electronic switches are the established solution through digitisation and time-division-multiplexing techniques. Optics is still seeking a role to complete the terminal to terminal transmissions, switching/routing and amplification in the optical medium.

The adoption of optical transmission and amplification owes much to the invention of optical fibre [1] and the erbium-doped fibre amplifier [2]. A similar technological break through has not occurred for optical switching. Earlier attempts at implementing optical switching were mostly based on configuring 2-D matrices of directional couplers linked by electro-optic [3] or semiconductor [4] waveguides. Free-space techniques have also been investigated [5] largely due to the potential of utilising the full 3-D switch volume. More recently the advent of wavelength-division multiplexing [6] has added a fourth dimension: wavelength routing.

It is not clear whether an optical switch should include space-routing, wavelength-routing or a combination of the two. Wavelength routing has held the upper hand with the development of key WDM components such as fibre Bragg gratings (FBG) [7], long period fibre gratings (LPG) [8], arrayed-waveguide gratings (AWG) [9] and the ultra-wide band optical amplifier covering the conventional wavelength band as well as the new 1570-1600 nm long-wavelength 'L'-band [10]. WDM point-to-point transmission promises to deliver one hundred 100-GHz spaced channels at up to 10-Gb/s rates [11]. The attraction of multiplying the capacity of fibre networks by merely upgrading to WDM switches is ultimately limited by nonlinearities in the fibre and WDM components (e.g., four-wave mixing, self-phase and cross-phase modulation, etc. [12]) due to high aggregated optical power and by chromatic and polarisation-mode dispersions.

Not only is the scope of WDM networking limited by fibre nonlinearities and dispersion management/compensation, these WDM links have to be inter-connected and network functionality extended by using WDM/optical add-drop multiplexers (OADM) and optical cross-connects (OXC) to enable the current photonic-layer transport to evolve into full optical networking [13]. Dynamically reconfigurable phase-only computer-generated holograms may have important applications in such all-optical networks. These "*routing holograms*" are encoded onto a pixellated spatial light modulator (SLM) and used as the active element in beam steering free-space [14,15] and wavelength [16] switches. For beam steering purposes, blazed phase-only holograms are often required to produce a single replay peak. It is this unique characteristic that has led many to adopt a replay peak efficiency expression given by Dammann [17]. The analysis of *stepped phase functions* is readily applicable when the routing hologram is a 1-D grating written onto an SLM without pixel dead-space and is illuminated by an infinite-expanse plane wave.

In practice, pixel dead-space exists regardless of the SLM technology and arbitrary 2-D routing patterns other than gratings are often desired. Furthermore, the fibre-mode approximated as a Gaussian distribution is more appropriate as the input to a 4*f* coherent optical configuration for beam-steering applications. In order to construct a measure of insertion loss and crosstalk isolation performance of such a holographic switch, an analysis of the replay plane in terms of the peak intensities and locations and sources of crosstalk is needed. Understanding the effects of input illumination, inherent hologram characteristics and SLM imperfection is essential, given the stringent requirements of transparent optical networks [18].

Some aspects of these theoretical performance figures for dynamic hologram replay can be verified experimentally. The prevalent SLM technology, liquid crystal on silicon, is typically designed for displays. Scattering and absorption reduces the efficiency as well as presenting crosstalk problems in switching light. In order to enable a reliable evaluation of hologram replay properties as well as the capabilities of such a holographic switch, an optimised SLM has to be fabricated. The first stage of the SLM assembly is getting functional silicon backplanes, where the optical quality of SLM pixels, the size of pixel dead-space and the size of the pixel array are the main considerations. In addition, more than two phase levels are required to restore some of the binary hologram symmetry losses. A silicon backplane comprising a binary and a quaternary array was designed and fabricated within the scope of this Ph.D. research, specifically to address these issues.

The research described in this dissertation seeks to unfold the 'mysteries' concerning the performance figures of dynamic holography using pixellated SLMs. Several research methodologies including theoretical analysis, numerical simulations with realistic examples and the design, fabrication, characterisation and experimentation of demonstrator devices for theory verifications were carried out in the course of this Ph.D. The dissertation is divided into nine chapters and is briefly outlined below:

CHAPTER 2: An introduction of several SLM technologies for coherent optical processing and optical routing/switching is given, followed by a discussion of the application of ferroelectric liquid crystal on silicon SLMs in moderate speed optical network management traffic routing.

CHAPTER 3: This chapter describes the effects of the input illumination profile. The replay beam profile is dependent on the clipping of the collimated Gaussian by the finite expanse of SLM/hologram aperture. The main figure of merit for such a switch will be the efficiency of the desired replay beam coupled into a single mode fibre. The sidelobes arising from this clipping ought to be considered for calculating the crosstalk power of the holographic switch fabric. The overlap integral of the replay beam profile with the weakly-guiding fibre mode [19] and its Gaussian approximation, having lateral and angular offsets [20] is evaluated numerically for a blank hologram replay.

Analytic expressions profiling the dependence of coupling intensity on only one of the two offsets are presented. The exact on-axis coupling intensity expression taking into account the power clipping and the Gaussian fibre mode matching is also given. Inasmuch as the true fibre mode is described by Bessel functions and the lateral offset replay/coupling intensity expressions are asymptotic approximations, numerical simulation results by employing fast Fourier transforms (*fft* also FFT) with both Gaussian and Bessel function fibre modes are compared with analytic expressions. Having treated the effects of an apertured Gaussian illumination separately, the phase-only hologram can be analysed as an infinitely repeated base hologram which suits the *fft* routines.

CHAPTER 4: In this chapter, phase and spatial quantisation effects are decoupled. A convenient fractional hologram representation and a general skip-rotate rule are introduced to locate each peak of the "*numerical replay*" along with its intensity as a result of the availability of a finite number of phase states and the arrangement of these phase states in a hologram. The pixel periodicity and finite pixel clear aperture give sampling and interference effects which result in scaled multiple replay replications. As Fourier holograms are being replayed, paraxial scalar wave propagation with respect to the pixel size (and the hologram period) and thickness must hold true for the wavelength of operation. Given the typical SLM pixel sizes and focal lengths of the transform lens, the replay plane description should be accurate for a small number of replay replications close to the optical axis. For optical correlation applications, an analysis of the effects of SLM dead-space has been given by Gianino *et al.*[21]. The same principle in treating pixellation and dead-space is applicable for optical beam steering. Within the paraxial regime, expressions governing the replay intensities in any replay replication for a hologram having any number of phase levels are presented. Results from the numerical hologram replay taking into account phase and spatial quantisation are presented.

CHAPTER 5: The understanding of the distribution of replay peaks, their intensities and the beam profile is used for evaluating the insertion loss and crosstalk isolation figures. This is also used to predict some device and operational parameters in order to meet a set of performance specifications for holographic replay. All-optical interconnects have stringent polarisation dependence, path loss, crosstalk, reconfiguration speed, switch matrix size and optical bandwidth requirements. Understanding the performance figures of a dynamic hologram replay helps in the evaluations of the prospects of these switches. Most important of all, the method of shift-rotate is applied to generating deterministic multi-level phase-only holograms to replay a single routing peak. This type of hologram is probably useful and adequate for routing purposes.

CHAPTER 6: An analysis of the mechanics of polarisation rotation using planar switching liquid crystals in obtaining phase modulation is given. The limitations of a four-level phase modulator is described and ways of achieving ideal four-level switching using double pass cells are presented. The architectural issues of optimised backplane designs for holographic SLMs are discussed and choices made based on the obtainable processing steps from the standard CMOS fabrication. In particular, the design of the binary backplane allows for the deposition and patterning of a third layer metal to form the pixels. The quaternary pixel design utilises as few post-foundry steps as possible to expose the polysilicon modulating layer. The quest for higher voltage LC driving and smaller pixel pitch is assessed by discussing the circuit requirements of a moderate and a high voltage CMOS process. This is not to say future silicon designs should not lean towards the high voltage processing, requiring more complicated circuitry for drain protection.

CHAPTER 7: The full-custom mask design of a silicon backplane for an FLC/SLM to be used in holographic applications is presented. The specifications for the design were given by a collaborative project. The circuit design and layout was performed using Mentor Graphics computer-aided-design (CAD) tools. Details concerning the implementation of electronic functionality are described.

CHAPTER 8: This chapter deals with the characterisation of and the experimentation on the holographic SLMs. Tests of electronic functionality and yield results of the fabricated devices are given. The design and fabrication of the silicon backplane has produced electrically functioning devices. The processing steps required to turn these silicon backplanes into SLMs are briefly described. Optical inspection results, using both the probe station and the cross-polarising microscope then follow.

Finally, several proof-of-principle holographic replay experiments using both a fixed intensity grating and binary phase holograms are described. The results from these experiments are compared to crosstalk and insertion loss figures detailed in the theoretical analyses. These measurements were made using a holographic optical switch configuration. Hence, the results are also indicative of the capabilities of the dynamic holographic switches. The suitability of the binary SLM design approach can be assessed by the final processing yield and the quality of the optical replay. A reliable gauge of the optical quality is to compare the first order efficiency of the fixed intensity grating having a cold-deposited electrode array with that of the dynamically reconfigurable binary SLM possessing liquid crystal scattering, dead-space loss, non-ideal switching angle and poor retardation efficiency. Temporal optical losses in updating FLC holograms meant for 'static' interconnect are quantified by simulation and experiments.

CHAPTER 9: Conclusions are drawn from the analyses and experimentation of the dynamic holography using FLC/Si SLMs described in this dissertation. Several areas for future work are presented to progress the understanding and the use of dynamic holography in conjunction with pixellated SLMs.

CHAPTER 2

Spatial light modulators: active devices for optical processing

2.1 Introduction

In this chapter, spatial light modulators (SLMs) for coherent optical processing and optical routing are briefly reviewed. The purpose is to show the diversity of materials used, modulation/detection mechanisms, addressing methods, device configurations, capabilities and limitations. An excellent review article on SLMs is given by Neff *et al.* [22]. Incoherent optical processors, such as displays, do not feature in this review. Finally, the most suitable SLM application together with its enabling components are chosen in terms of prospects and feasibility within the scope of this Ph.D. research.

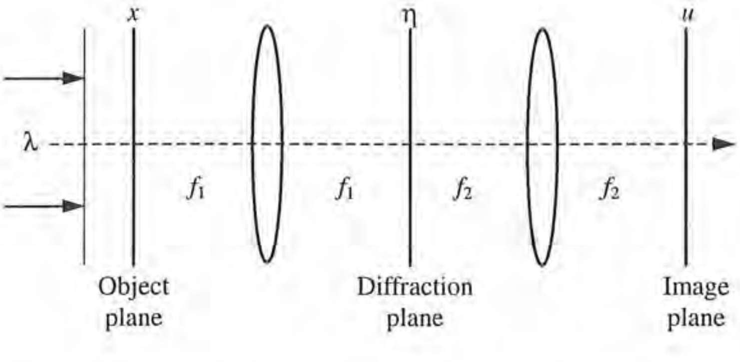


Figure 2.1: Basic arrangement for coherent optical filtering.

A transistor in an electronic processing system is the active element providing gain and switching functions. Likewise, an SLM in an optical processing system controls the light paths and sometimes adds gain to the signals. A typical *4f* optical set-up for coherent processing is shown in Figure 2.1. The three 'co-ordinate planes' are separated by the focal lengths, f_1 and f_2, of the two Fourier transform (FT) lenses. The object plane contains the input spatial function, e.g., a transparency. The diffraction plane is where a filter is located. Correlation or convolution patterns are formed on the image plane depending on whether the filter is a conjugate FT of the object function or a FT of any other function.

It is obvious that SLMs are required to present the spatial functions of the object and diffraction plane. These functions are reconfigurable at high speed and are ideally encoded with high spatial resolution. However, an SLM with a low spatial resolution is still useful in configuring the routing patterns of a relatively small number of input/output optical channels.

The basic structure of an SLM cell is as shown in Figure 2.2. SLMs can be electrically (EASLM) or optically addressed (OASLM). Therefore the 'input write' can be an electron beam, a row-column matrix or a laser beam. In the case of EASLMs, there is no need for the optically opaque layer. The charge generating layer refers to the column-row overlap in matrix addressed EASLMs. Both transmissive and reflective devices can be made. Reflective readout is necessary for OASLMs. However, EASLMs require reflective readout if the charge generating layer is opaque to the wavelength of operation, as is the case with silicon based active matrix devices.

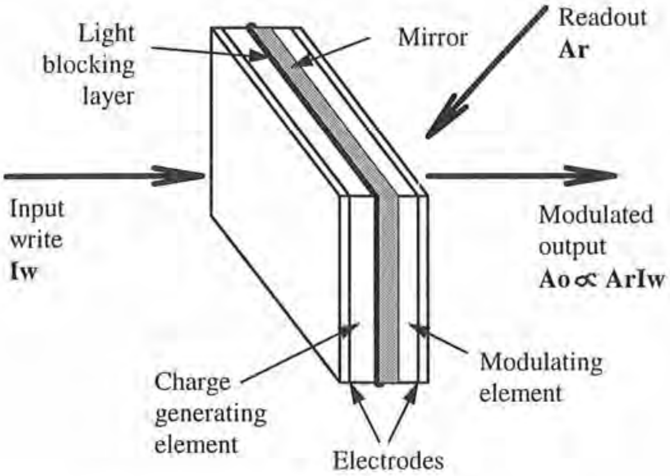


Figure 2.2: General structure of an SLM.

2.2 SLMs for use in coherent optical processing systems

SLMs that are considered suitable for optical processing must be capable of high resolution and good optical quality readout. Other secondary criteria such as speed, contrast ratio, dynamic range, linearity and memory effects are also important. Therefore, SLMs that have their origins as pixellated displays and storage devices may not be suitable. This section includes SLMs that are not commonly pixellated.

2.2.1 Nematic liquid crystal (NLC) modulators

The liquid crystalline state exists between the highly ordered solid crystal and the isotropic liquid phase. Conventionally, there are three types of thermotropic liquid crystals: nematic, chiral nematic (cholesteric) and smectics. These organic mixtures are often associated with the ease of optic axis alignment and a large birefringence (difference of the two perpendicular ordinary and extraordinary refractive indices). There are several methods of configuring non-pixellated liquid crystal modulators such as field, twisted nematic, hybrid field, guest-host, voltage-controlled birefringence and scattering effects.

A successful liquid crystal device is Liquid Crystal Light Valve (LCLV) by Hughes [23] employing the nematic hybrid field effect. In a nematic liquid crystal cell, the liquid crystal molecules are aligned parallel to the grooves on an alignment layer (e.g., polyimide or nylon). If the alignment is homogeneous, this gives rise to field and voltage-controlled birefringence effects. Upon applying a voltage, positive dielectric-anisotropy materials tend to align parallel to the field and negative dielectric-anisotropy materials tend to tilt to an angle, hence giving controllable birefringence.

In a twisted nematic configuration, the cell alignment results in the liquid crystal layers tracing a helix or twist [24]. The pitch is modified by an electric field applied across the cell. However, if the cell thickness is such that the twist results in a 90° rotation without an applied electric field, an off-state is obtained with a parallel polariser/analyser combination. With an applied electric field, the hybrid field effect relies on a mixture of polarisation rotation and optical birefringence to produce elliptical outputs. For LCLV, the write light generates electrons in the Cadmium Sulphide, CdS, photodetector which modify the field locally and therefore spatially modulate the read light phase. The response times are tens of milliseconds due to the slow switching of nematic liquid crystals.

2.2.2 Photorefractive crystals

Photorefractive crystals are non-fragile and robust electro-optic materials. Some examples of photorefractive crystals are bismuth silicon oxide (BSO), bismuth germanium oxide (BGO), strontium barium niobate (SBN), barium tionate, iron doped lithium niobate (LNB) and deuterated potassium-dihydrogen-phosphate (DKDP). Only DKDP can be both electrically and optically addressed. All others have the advantage of transmissive mode optical detection and modulation in a single material. The speed of response is typically fast. However the required hundreds of volts and extreme working temperature make these crystals less attractive for low power optical computing/processing uses. These crystals typically give a high contrast ratio at high spatial resolutions. Devices made from these materials exploit the linear electro-optic effect (Pockels Effect). LNB is the most common crystal for waveguiding applications such as directional couplers and external cavity modulation of continuous-wave (CW) lasers.

2.3 Smart-pixel arrays for free-space optical interconnects

A so-called 'smart pixel' can be integrated with a combination of light modulator, photodiode, receiver circuits, modulator driver, processing electronics and electrical inputs/outputs. Therefore, the integration platform has to be based on a semiconductor process. In the following, self electro-optic effect device (SEED) and electro-absorption (EA) modulators constructed from multiple quantum wells (MQW) which are capable of high speed response and/or functionality are reviewed in the context of optical transmission/switching.

2.3.1 SEED modulators

Quantum wells (QW) are made up of alternate thin layers (~ 5 → 10 nm) of two different semiconductors such as GaAs and $Ga_xAl_{1-x}As$. The absorption spectrum of a multiple QW (MQW) stack shows a series of steps due to the confinement of electron-hole pairs (excitons) within the layers. The path refractive index of the ~ 850 nm optical beam at normal-incidence is modified in association with an absorption. This quantum confined Stark effect (QCSE) [25] can be used for phase modulation. Amplitude modulation is achieved by means of electrically altering the absorption spectra. MQW-structure based SLMs are attractive because they are fast, can be pixellated and are compatible with semiconductor processing technologies (e.g., molecular beam epitaxy). However, the contrast ratio is low (2→6:1). Two important types of optical bistable (OB) modulator have been fabricated by placing the MQW structure in a Fabry-Perot (FP) cavity [26] and in the depletion region of a *p-i-n* photodetector (Self-electro-optic-Effect Device, SEED [27]).

There are a variety of SEED configurations depending on the biasing and driving of the *p-i-n* photodetector (Figure 2.3) [28]. The simplest SEED uses a resistor or a diode for biasing. However, the problems with critical biasing and input/output isolation of a two-terminal driving have compelled progress to using a three-terminal transistor or a similar SEED with a control light input for biasing (symmetric S-SEED). The response of S-SEED modulators can be very fast (e.g., 33 ps [29]), limited only by the optical power of the control input.

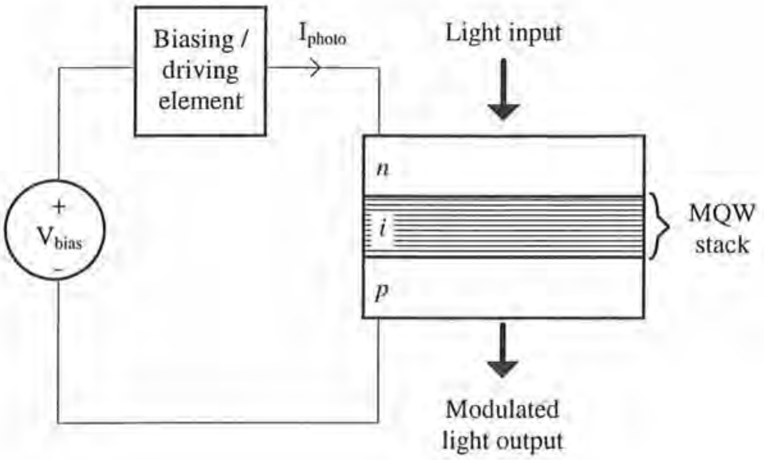


Figure 2.3: Generalised SEED device.

While S-SEED devices have been extensively used in opto-electronic/optical logic operations [30] and in free-space multi-stage-interconnection networks (MIN) [31], its use in general smart pixel arrays is hindered by the lack of electronic processing capabilities. Arrays of smart pixels comprising a SEED modulator, a field-effect-transistor (fet or FET) driver and other GaAs electronic elements have therefore become prominent [32]. Nonetheless further advance in SEED devices has taken the route of flip-chip bonding SEED devices with CMOS silicon VLSI chips [33] followed by substrate removal to access the optical modulators. Several hybrid-optoelectronic switches involving photo-detection and regeneration have been demonstrated using CMOS-SEEDs [34].

2.3.2 Electro-absorption (EA) modulators

High-speed (up to ~ 60 GHz, 3 dB electrical bandwidth) and low chirp (i.e., change in optical frequency) modulation which are essential in digital optical transmission and switching systems are obtainable from EA modulators. The modulator uses the undoped region of a *p-i-n* photodetector for optical absorption. If the active region is constructed as a non-MQW quasi-bulk structure, the optical absorption coefficient is changed by the applied reverse bias field through the Franz-Keldysh (FK) effect. With stacked compressive/tensile strained QWs in the active region, excitons provide enhanced modulation through the QCSE. The steepness of the absorption spectrum at the band-edge is increased and the modulation can take place at room temperature with lower fields. Unlike the SEED devices, the optical beam is incident parallel to the relatively long (~ 200 μm) active region so as to realise > 30 dB optical power change, as shown in Figure 2.4 [35]. The modulation is polarisation independent and is relatively wavelength-insensitive (large optical bandwidth > 10 nm).

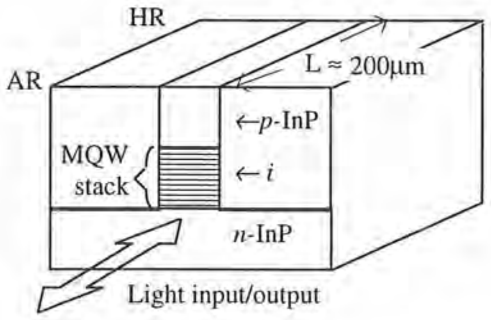


Figure 2.4: Schematic view of an electro-absorption modulator.

2.4 SLMs for use in free-space optical switches

Where an SLM is not fast enough to modulate digital data streams at several Gb/s, nor is it adequate for providing a high spatial-bandwidth product (SBWP-product of spatial resolution and achievable size), the SLM can still be useful in matrix-vector multiplier [36] based optical systems involving discrete arrays in the input and output planes. The main requirements for these applications are low power, compact electrical and optical interfaces and low device costs.

Excluding free-space optical switches that use smart-pixel arrays [37], free-space optical switches that use low-pixel-functionality SLMs can be divided into three architectures. The first is the matrix-matrix multiplier version of the original matrix-vector switches connecting an input single-mode fibre array with an output multimode fibre array [38]. Using the 2-D numerical aperture of the optical system enables the replication of the 2-D input array on a shutter SLM. By switching the through channels to the focusing micro-lens array, multi-cast of several inputs to a large core output fibre is established as shown in Figure 2.5(a). The main drawbacks are the use of polariser/analyser intensity modulation and replication loss proportional to the number of channels. A polarisation-sensitive 64:64 matrix switch employing this architecture has been demonstrated under the Optically Connected Parallel Machine (OCPM) with 29 dB insertion loss and 15 dB crosstalk isolation [39]. Nonetheless there are additional tilted-coupling losses if small-core and low numerical-aperture single-mode fibres are to be connected in the output stage as a consequence of Liouville's theorem [40].

The second design seeks to avoid the replication loss by using a holographic fan-out stage as shown in Figure 2.5(b) [15]. A micro-lens array collimates the input channels to a sub-hologram array. Each sub-hologram steers its respective beam to the desired output fibre port. The same hologram pattern diffracts the light from any input channels to a particular output port due to the shift-invariant property of holograms using a single Fourier transform lens. Nonetheless, the coupling of light into the output fibres which are not square-law intensity detectors entails an intrinsic loss due to angular tilt [20]. The additional loss in this switch architecture has been shown to scale with N where N is the number of input/output channels [41].

The final architecture tries to correct the optical wave-front tilts at the output stage by an active holographic fan-in as shown in Figure 2.5(c). This switch design is equivalent to the two-stage fixed holographic interconnects with an added ability to reconfigure the routing patterns [42]. It is thought that a 32:32 optical matrix switch connecting two single-mode fibre arrays is feasible with low loss (< 20 dB) and high crosstalk isolation (> 40 dB) [43].

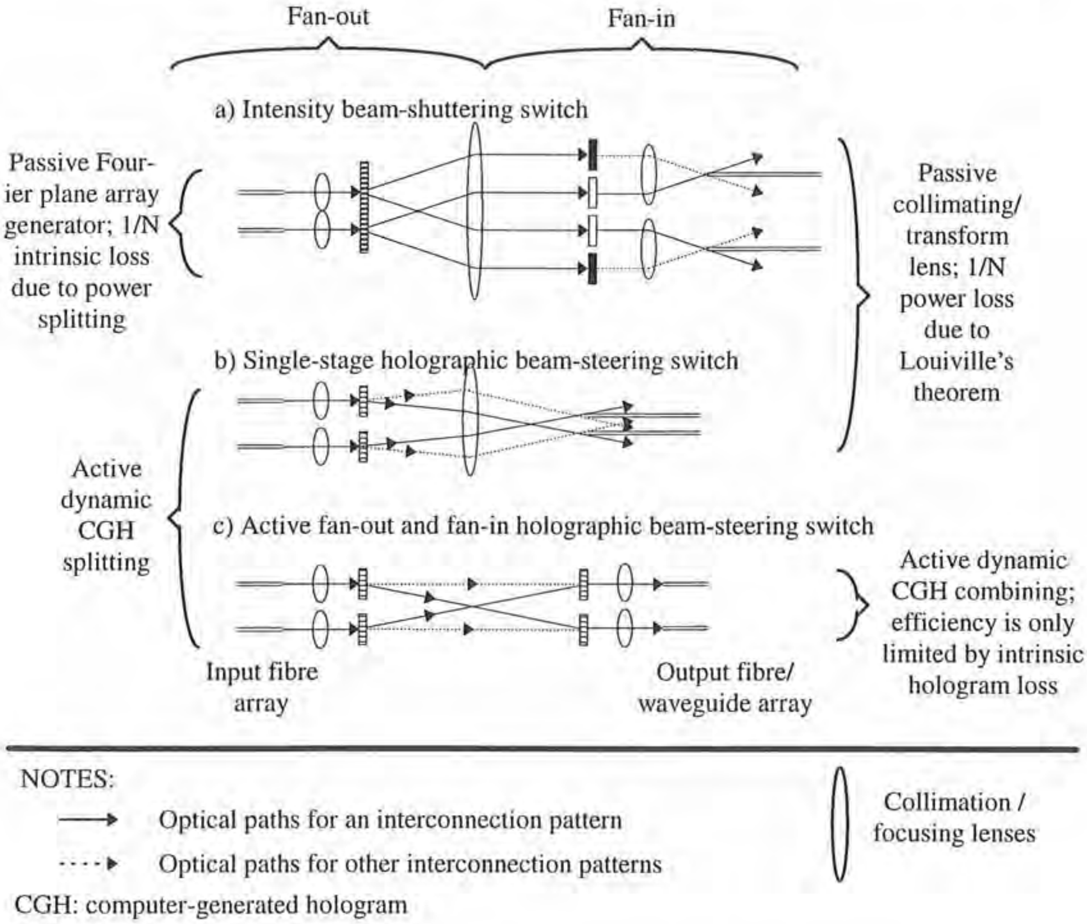


Figure 2.5: Single-mode fibre to fibre free-space switch techniques.

Free-space holographic matrix switches are potentially comparable with waveguide-based switches in terms of insertion loss and crosstalk figure and could surpass waveguide-based switches in scalability and control aspects. Waveguide-based matrix switches constructed from multiple stages of directional couplers or digital optical switches (DOS) (electro-optic and thermo-optic X-junction) [44, 4, 45] and 3-dB couplers with single [46] or double Mach-Zehnder optical gates have been demonstrated with 5 ~ 6 dB insertion loss and 55 ~ 60 dB crosstalk isolation for 8 × 8 [47] and 16 × 16 [48] matrix switch sizes.

In the following subsections, binary magneto-optic and grey-scale Si-PLZT electro-optic intensity as well as opto-mechanical and ferroelectric liquid crystal (FLC) phase modulators suitable for one/two-stage free-space optical switches are reviewed. In addition, a Digital Micro-mirror Device (DMD) used in deflecting light paths is also described.

2.4.1 Opto-mechanical modulators

Modulation of light by mechanically changing the modulation variables can be effected on a pixellated or a continuous plate/film. The Membrane Light Modulator (MLM) (both optically and electrically addressed) [49] and the Charge-coupled Device (CCD) addressed Deformable-membrane Device (precursor to the hinged free-rotating Digital-Micro-mirror Device, DMD) [50] are the more successful examples of this family. These devices modify the phase of the readout light due to an optical path length change caused by the membrane deformation. The force required for membrane deformations can be an electrostatic field or the charge distribution in an optical/electrical addressed SLM.

2.4.2 Magneto-optic modulators

Magneto-optic SLMs are available commercially by the name of SIGHT MOD [51]. They are binary SLMs operating on Faraday effect. The magnetisation direction of a bismuth-doped yttrium iron garnet film is determined by the electrical control inputs. The polarisation of the readout beam is thus modulated accordingly. These bistable SLMs have been used to record binary phase holograms for reconfigurable interconnects [52]. They are capable of fast response (microsecond) with a high contrast (1000:1) as well as realisable in large matrix arrays (256 × 256 or higher). The main disadvantages are the low light throughput (~ 5%) and the high current required to effect switching.

2.4.3 Si-PLZT modulators

Photoconductor materials, such as Si-PLZT (silicon-lead/lanthanum zirconate/titanate) which are compatible with semiconductor processing, offer good prospects for spatial light modulation [53]. The SLM is pixellated with an array of silicon photodetectors. The applied electric field acts on the polarisation of the readout beam by a transverse effect. This device is capable of MHz frame rates at moderate operating voltages (10→50V). Although the array size is typically small, this modulation mechanism gives a large dynamic range with a high contrast ratio (10→1000:1).

2.4.4 Ferroelectric liquid crystal (FLC) modulators

Since the 80s, there have been many devices made using surface-stabilised chiral smectic C (Sc*) liquid crystals. Due to surface stabilisation [54], there are only two orientations which the molecular directors can occupy; both occurring in the plane of device as shown in Figure 2.6. Each director orientation is associated with a spontaneous polarisation which is normal to the plane of director and layer-normal. This dipole moment can be switched by the polarity of an applied field. The bistability feature enables passive matrix electrical addressing. The size of matrix can be made arbitrarily large. For display applications, passive matrix addressing with large achievable display size coupled with microsecond response times can be very attractive. It is even more so with the τ-v_{min} schemes where the faster DC-balance mono-pulse addressing enables video rate line scans. The crosstalk due to column voltages being applied to the pixels in any non-selected row is minimal [55].

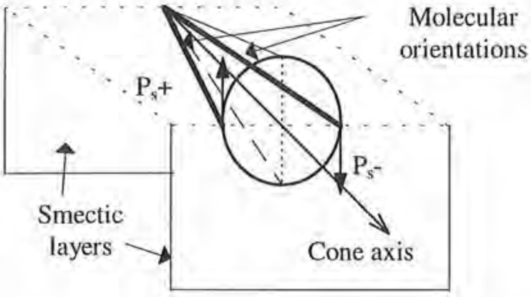


Figure 2.6: Binary (bistable) ferroelectric liquid crystal orientations.

Ferroelectric liquid crystals in passive matrix glass SLMs have been used to encode binary phase-only holograms [56, 15]. However, for holographic purposes, unlike display, large pixel dimensions produce a small replay field. The integration of liquid crystal modulators with silicon backplane circuitry has resulted in much higher space-bandwidth products as well as pixel level processing [57].^ Binary holographic elements offer a poor efficiency as well as undesired replay peaks due to the inverse symmetry. At least three or more phase levels are needed for improved diffraction efficiency and breaking the replay symmetry. There is not much diffraction efficiency gain in using holograms with a large number of phase levels. For example, the 'theoretical' diffraction efficiencies are 81% and 95% for four- and eight-level holograms, respectively (see §4.3.2 for more practical values). A general scheme for obtaining an arbitrary number of phase levels by means of cascading multiple SLMs has been presented [58]. This does not appear attractive due to the difficulty in

^ Silicon backplane devices have been used to record dynamic binary phase holograms [87].

aligning the two devices pixel to pixel. Similarly, for filtering applications requiring asymmetric replay, a pseudo-quaternary hologram obtained by combining a [0,π] modulator with a randomised [0,π/2] fixed grating [59] is easier to implement than cascading two corresponding dynamic holograms on SLMs [60].

2.4.5 Digital-micro-mirror-device (DMD)

Amongst the various SLM technologies, the Digital-micro-mirror-device (DMD) is arguably the most attractive one in terms of pixel density, cost, speed and contrast ratio. These SLMs were initially developed by Texas Instruments targeting projection-based and miniature displays [61]. The DMD consists of a 2-D array of small tilting mirrors (~ 16 × 16 μm^2) each hinged over the top of a static pixel storage element. Each DMD pixel mirror is pre-tilted by an electrostatic field. A 0 or 1 pixel state modulates the tilt such that the incident light is either directed at an opaque region or back into the optical system at a high speed (< 20 μs). These SLMs can be suitable for OCPM-type switches while also eliminating the polarisation dependence.

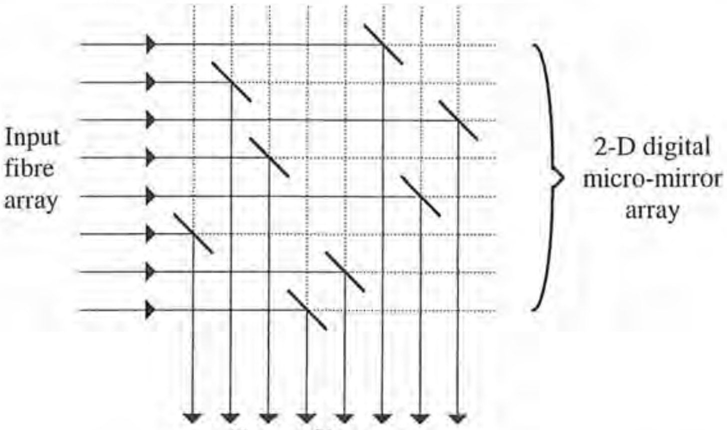


Figure 2.7: $N \times N$ free-space optical switching using DMDs.

Recently, micro-machining techniques have been used to produce scratch-driven flipping micro-mirror arrays [62]. An *N*:*N* pixel array is able to form a deflecting matrix to interconnect two *N*-port fibre arrays as shown in Figure 2.7. Each pixel mirror is either raised to 90° to the substrate and 45° to the light path to deflect the incident light beam to an output port or lowered to allow for a through path. Due to the large mechanical movements, a 1 ms response time is typical. The scalability limit of such a switch is not so much an issue of the associated loss but the integration of ball lenses, fibre connectors, etc. in one dimension as well as the reliability and precision of the mechanical moving parts. Further progress in micro-electro-mechanical systems (MEMS) may overcome some of these limitations [63].

2.5 Conclusions

Real-time optical processing requires too many logic operations at too high speeds compared to the capabilities of current spatial light modulator or arrayed optical-source (micro-lasers and VCSEL-vertical cavity surface emitting lasers) technologies. Hence coherent filtering applications are likely to continue to be performed off-line using high resolution modulators such as photorefractive (electro-optic) crystals.

At low operating voltages, few modulation techniques are able to produce high response speeds. A notable exception is the stacked multiple-quantum well modulator. In the symmetrical self-electro-optic device (S-SEED) configuration, these arrayed modulators are capable of switching at the bit rate of fast packet data (< 1 ns). In order to achieve low cost and dense driving circuitry, these field effect transistor (fet or FET) drivers have been replaced by complementary-metal-oxide-silicon (CMOS) very-large-scale-integrated (VLSI) circuits. Thus far, yield has been the main set-back for integrating the two platforms via the flip-chip bonding technique. In the long term, a solution can be expected from the CMOS fabrication industry owing to the enormous potential benefits in vertical integration. In the electro-absorption configuration, MQW modulators are already replacing lithium niobate electro-optic crystals in high speed optical transmission modulation (up to several Gb/s).

An attractive application for the rest of moderate speed and resolution SLMs lies in transparent optical switches used in optical/wavelength-division-multiplexing(WDM) path restoration, reconfiguration and management where bit-rate response is not essential. Transparency to the signal modulation formats (analogue, amplitude/frequency/phase or digital intensity modulation at any bit rate) and carrier wavelengths (WDM channels) requires the switch not to perform any electro-optical conversion and *v.v.* In this application domain, switches based on *holographic beam steering* rather than *intensity beam shuttering* appear to be more scalable with input/output port numbers.

Ferroelectric liquid crystal on silicon seems to be a suitable technology to record binary phase holograms for polarisation-insensitive beam steering switches. If offers a means of creating low-cost and intrinsically reliable arrays of 10^6 pixels or more, with moderate response speeds of order tens of microseconds. These should meet the needs of network reconfiguration traffic.

CHAPTER 3

Analysis of fibre-to-fibre 1:N switch coupling efficiency: the replay of hologram aperture

3.1 Introduction

The output field of a 4f holographic system is a convolution of the discrete hologram replay and the Fourier transform of the hologram illumination. The collimated beam illuminating the hologram of an optical fibre switch is often approximated as a Gaussian distribution. Starting from this assumption, the replay field and coupling intensity expressions, for a blank hologram aperture clipping the hologram illumination, are derived. A holographic switch contains hologram patterns encoded onto an SLM with inherent physical defects. Having considered the "*aperture replay*", the holograms can be treated as infinitely extending/ repeating, as given by the analysis in §4.

The analytic expressions are compared with numerical simulations using the Gaussian and Bessel function descriptions of the weakly-guiding input fibre mode. The collimated beam (*fft* of the input beam) is truncated at the hologram plane. Two aspects of the beam clipping effects are of interest: *efficiency* of the coupling of replay power into a single mode fibre with small lateral or angular offsets and *crosstalk* power received at a fibre port due to the sidelobes of the replay beam (large lateral offsets).

3.2 Weakly-guiding approximation for the fundamental (HE_{11}) fibre mode

The normalised field solution of the scalar wave equation (valid only for the weakly-guiding assumption) for an isotropic step-index, circular-core single mode fibre is [64],

$$E_x = \begin{cases} J_0(UR) & R \le 1 \\ \dfrac{J_0(U)}{K_0(W)} K_0(WR) & R > 1 \end{cases} \qquad R = \frac{r}{a}, \tag{3.1}$$

where J_0 and K_0 are Bessel functions of the first kind and second kind of order zero, respectively; U and W are the fibre parameters in the core and cladding, respectively; r is the radius from the fibre axis, a is the fibre core radius and R is the normalised radial co-ordinate. The values of U and W satisfy the following condition, $V^2 = U^2 + W^2$, where V is the normalised frequency parameter of the fibre. The values of V must be < 2.405 for single mode operation. Typically, V values of around 2 are chosen ($V = 2$ for this analysis) in order to achieve a well-confined mode so as to reduce bending loss [65]. The parameter U is the solution of the eigenvalue equations for continuity of the fibre mode field and its derivative at the fibre core-cladding boundary [64]:

$$\frac{U}{n_{co}} \frac{J_1(U)}{J_0(U)} = \frac{W}{n_{cl}} \frac{K_1(W)}{K_0(W)}, \tag{3.2}$$

where n_{co} and n_{cl} are the refractive indices in the fibre core and cladding, respectively. This transcendental equation has to be solved numerically. Assuming n_{co} of 1.6 and a numerical aperture (N.A.) of 0.12 for the single mode fibre, the values obtained are $U = 1.5294$ and $W = 1.2887$. These parameter values are consistent with tabulated solutions [66]. The fibre core radius, a, for these conditions is 4.1 μm at $\lambda = 1.55$ μm. The optical power carried by the fundamental mode with normalised radius and unit on-axis intensity is,

$$P_o = 2\pi \left\{ \int_0^1 J_0^2(UR) R dR + \frac{J_0^2(U)}{K_0^2(W)} \int_1^\infty K_0^2(WR) R dR \right\},$$

$$P_o = \pi J_0^2(U) \left(\frac{J_1^2(U)}{J_0^2(U)} + \frac{K_1^2(W)}{K_0^2(W)} \right) = 2.3793. \tag{3.3}$$

The Gaussian field distribution with unity power is defined as,

$$G(w/a\,; R) = \frac{1}{w/a} \sqrt{\frac{2}{\pi}} \exp\left(-\left(\frac{R}{w/a} \right)^2 \right), \tag{3.4}$$

where w is the Gaussian radius at 1/e on-axis field amplitude. The two well-known parameters for Gaussian approximation of the fundamental fibre mode are:

i) Marcuse's Gaussian parameter [20] where, $\frac{w_{mar}}{a} = 0.65 + \frac{1.619}{V^{1.5}} + \frac{2.789}{V^{6}}$ and (3.5)

ii) Snyder's Gaussian parameter [67] where, $\frac{w_{sny}}{a} = \frac{1.}{\sqrt{\ln(V)}}$. (3.6)

The power integrals of the Bessel function and Gaussian distribution descriptions of the fibre mode have to be equal. Thus, the Gaussian field distribution with unity power has to be scaled by the factor, $\sqrt{P_o}$. Using the Gaussian parameter by Marcuse gives too low on-beam-axis field magnitude and that of Snyder's parameter gives too high on-beam-axis field magnitude. These parameters for Gaussian approximation are either too broad or too narrow. It was decided to proceed by taking the average of these two parameters to obtain a value of $w_{ave} = 1.2343a$ or 5.06 μm as the Gaussian radius.

The problem in approximating the fundamental fibre mode as a Gaussian field distribution is that the Gaussian field magnitude does not fall rapidly enough for R between 1 and 1.5 and the true fibre field is underestimated for $R > 1.5$. However, even the transverse field solutions of the fibre mode in terms of Bessel functions are only roughly true for a weakly guiding fibre. It was thought that a Gaussian approximation need not have to fit this transverse field exactly. It will be shown in §3.5.2 and §3.5.3 that approximations based on this Gaussian parameter give fairly small errors of coupling intensities, for cases where the output fibres are aligned with either lateral offsets or angular tilts.

3.3 Analysis of the replay field profile using Gaussian fibre mode description

A holographic switch to route the input fibre signal to an array of output fibres is configured as a 4*f* coherent optical system as shown in Figure 3.1. The optical field confined within a single-mode waveguide is approximated as a Gaussian distribution, as alluded in §3.2. In the figure, w is the Gaussian beam radius, '$_i$', '$_h$' and '$_r$' are subscripts denoting *input*, *hologram* and *replay* (or *output*) planes and γ is the ratio of the hologram (or SLM) aperture to the Gaussian radius of the collimated beam at the *hologram* plane.

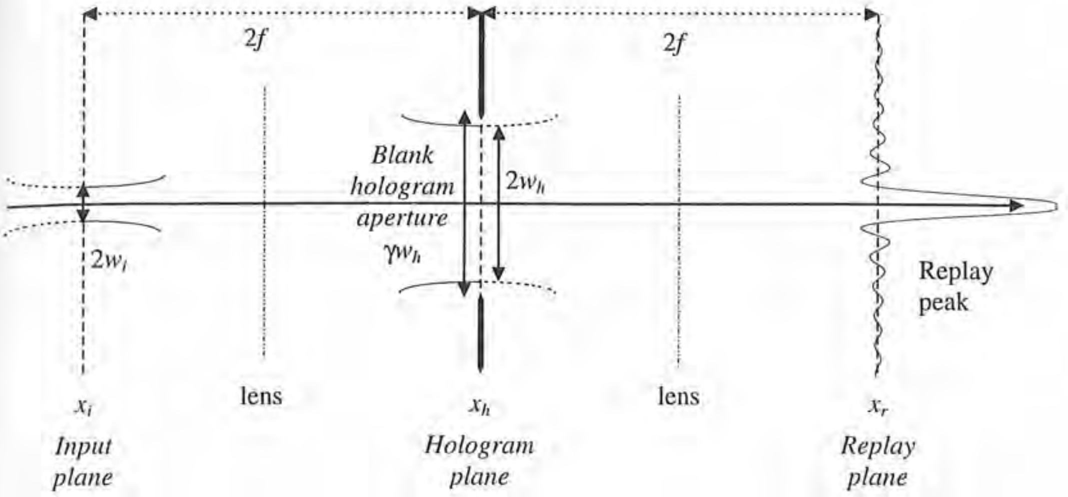


Figure 3.1: The replay of a blank hologram using a 4*f* coherent optical configuration.

The replay field, $R(x_r,y_r)$ of a square-dimension hologram is given by:

$$R(x_r, y_r) = \mathrm{F}\left\{G(w_h; x_h, y_h) \times rect\left(\frac{x_h}{\gamma w_h}, \frac{y_h}{\gamma w_h}\right) \times H(x_h, y_h)\right\}, \tag{3.7}$$

where $G(w_h;x_h,y_h)$ is the Gaussian beam at the hologram plane; $rect(x_h/\gamma w_h, y_h/\gamma w_h)$ is the aperture function that has a value of 1 for $|x_h|$ and $|y_h| \leq \gamma w_h/2$, 0 else where; γ being the ratio of hologram aperture to the spot-radius of the illuminating Gaussian beam, w_h; F denotes the optical Fourier transform, as defined by Yariv [68]. The normalised unit intensity Gaussian field at the hologram plane is as below:

$$G(w_h; x_h, y_h) = \frac{1}{w_h}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{x_h^2 + y_h^2}{w_h^2}\right), \tag{3.8}$$

where $w_h = f\lambda/(\pi w_i)$; f is the focal length of the collimating lens, λ is the illuminating wavelength and w_i is the Gaussian radius at the *input* plane.

In investigating the effect of the clipping ratio, γ, on the replay peak profile, a blank hologram, i.e., $H(x_h,y_h) = 1$ can be assumed. A complete field distribution of the illuminated hologram requires knowledge of the characteristics of the encoding SLM, the calculated hologram, the size of the hologram aperture and the input illumination profile. Assuming the illumination is a truncated Gaussian beam, the asymptotic approximation of the replay of a blank hologram can be shown to be [Appendix A]:

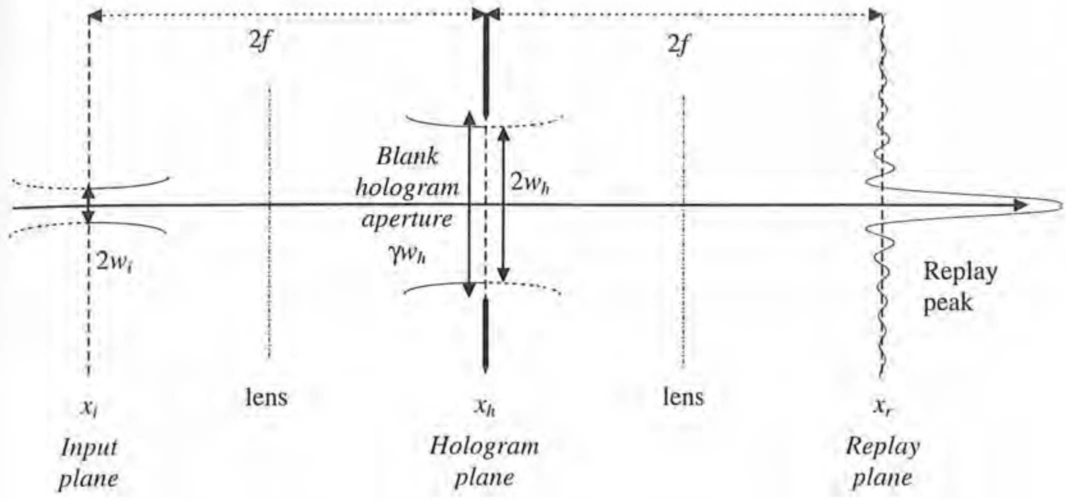


Figure 3.1: The replay of a blank hologram using a 4*f* coherent optical configuration.

The replay field, $R(x_r,y_r)$ of a square-dimension hologram is given by:

$$R(x_r, y_r) = \mathrm{F}\left\{G(w_h; x_h, y_h) \times rect\left(\frac{x_h}{\gamma w_h}, \frac{y_h}{\gamma w_h}\right) \times H(x_h, y_h)\right\}, \tag{3.7}$$

where $G(w_h;x_h,y_h)$ is the Gaussian beam at the hologram plane; $rect(x_h/\gamma w_h, y_h/\gamma w_h)$ is the aperture function that has a value of 1 for $|x_h|$ and $|y_h| \leq \gamma w_h/2$, 0 else where; γ being the ratio of hologram aperture to the spot-radius of the illuminating Gaussian beam, w_h; F denotes the optical Fourier transform, as defined by Yariv [68]. The normalised unit intensity Gaussian field at the hologram plane is as below:

$$G(w_h; x_h, y_h) = \frac{1}{w_h}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{x_h^{\,2} + y_h^{\,2}}{w_h^{\,2}}\right), \tag{3.8}$$

where $w_h = f\lambda/(\pi w_i)$; f is the focal length of the collimating lens, λ is the illuminating wavelength and w_i is the Gaussian radius at the *input* plane.

In investigating the effect of the clipping ratio, γ, on the replay peak profile, a blank hologram, i.e., $H(x_h,y_h) = 1$ can be assumed. A complete field distribution of the illuminated hologram requires knowledge of the characteristics of the encoding SLM, the calculated hologram, the size of the hologram aperture and the input illumination profile. Assuming the illumination is a truncated Gaussian beam, the asymptotic approximation of the replay of a blank hologram can be shown to be [Appendix A]:

$$R(x_r, y_r) = \frac{1}{w_r}\sqrt{\frac{2}{\pi}} \times R(x_r) \times R(y_r), \tag{3.9}$$

$$R(x_r) \sim \exp\left(-\frac{x_r^2}{w_r^2}\right) - \frac{\exp\left(-\frac{\gamma^2}{4}\right)}{\sqrt{\pi}\left(\frac{\gamma^2}{4} + \frac{x_r^2}{w_r^2}\right)}\left[\frac{\gamma}{2}\cos\left(\frac{\gamma x_r}{w_r}\right) - \frac{x_r}{w_r}\sin\left(\frac{\gamma x_r}{w_r}\right)\right], \tag{3.10}$$

where γ is the ratio of the SLM aperture to the Gaussian radius in both x_h and y_h directions (equal truncation) and both shoulders of the Gaussian beam are clipped symmetrically.

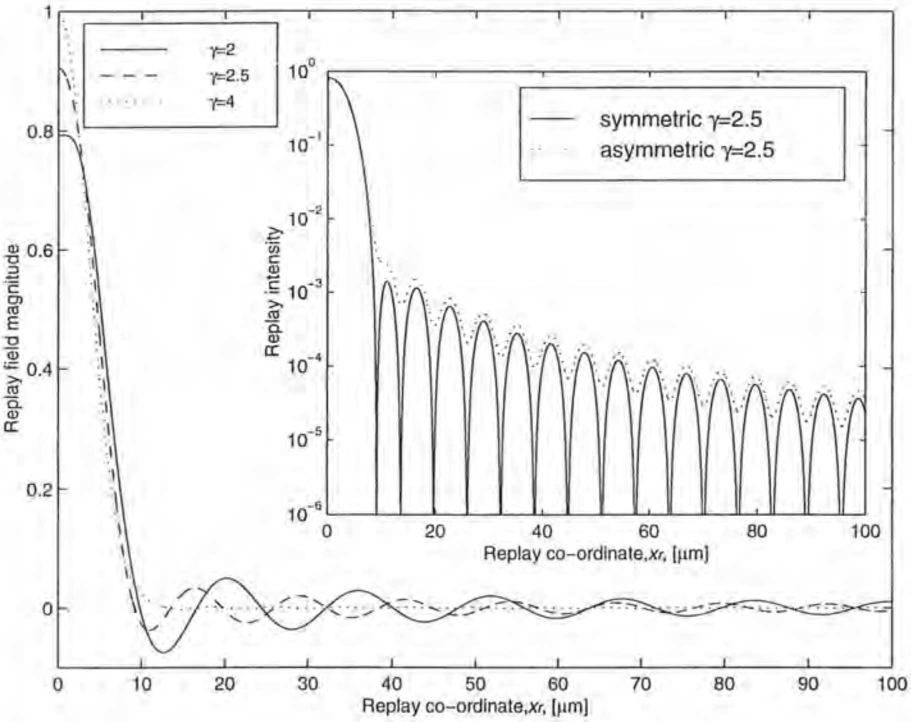


Figure 3.2: Asymptotic approximation of the replay field of a blank hologram illuminated by a Gaussian beam with $w_r = 5.06$ μm. The inset contains the replay field for an asymmetrically clipped beam.

Eq. (3.10) is plotted in Figure 3.2 for $\gamma = 2$, 2.5 and 4 (solid, dashed and dotted lines, respectively). Note the rippling nature of the sidelobe field. This is due to the replay field profile being the convolution result of the transform of a square aperture and the transform of a Gaussian field distribution. For the case of $\gamma = 4$, the clipping of the collimated Gaussian beam has negligible effect on the replay field profile (i.e., nearly a Gaussian replay beam). On the other hand, a clipping ratio of 1 will result in a near *sinc*-like replay field profile. However, eq. (3.10) is not valid for $\gamma >> 2$, as the argument of the asymptotic approximation will tend towards π/2.

For cases where γ is between 2 and 4, sidelobe rippling diminishes with increasing γ. Decreasing γ not only clips the input power by blocking part of the collimated beam, it also causes more power loss due to diffraction effects which steers optical power to sidelobes. It is important to point out that symmetrical truncation is not a general case. When both shoulders of the Gaussian beam are not clipped similarly, field cancellation will not be as exact as those depicted in Figure 3.2. The replay field approximation is generally complex and the troughs of its intensity plot will not be as sharp. This replay intensity is illustrated by the inset with a 0.1γ lateral shift of a $\gamma = 2.5$ aperture. In essence the approximate replay is the superposition of a symmetrical and real field due to the $\gamma' = 2$ sub-aperture and a complex field due to the extended 0.5 aperture. The complex field causes the "*smearing*" effect.

3.4 Analysis of the coupling intensity profile using Gaussian fibre mode description

If a large photo-detector is used, the sidelobes of the replay field obviously may cause severe crosstalk. But, we are only interested in the field amplitude of these sidelobes being coupled into other fibre ports as well as the on-beam-axis coupling efficiency into the desired fibre port with/without angular tilt. In order to analyse these effects, the overlap-integral of the replay field profile and fibre mode displaced laterally and having an angular tilt has to be solved, i.e., determining how closely the replay field profile resembles a Gaussian field.

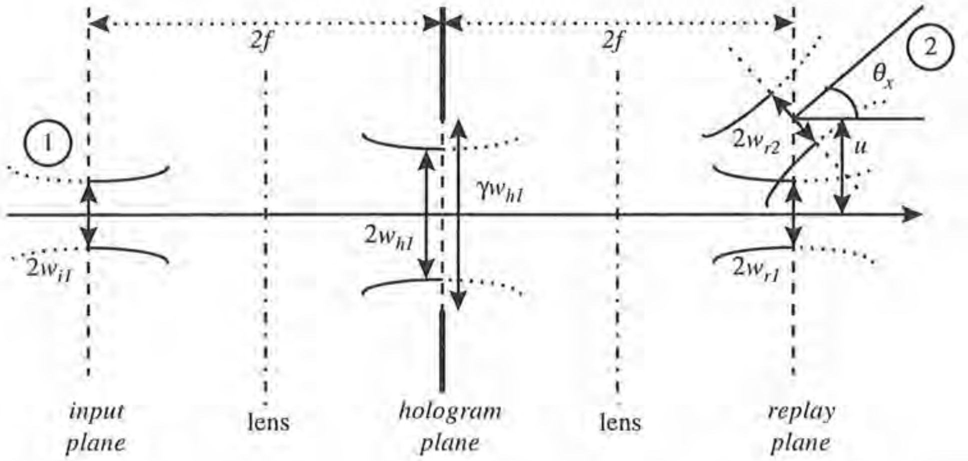


Figure 3.3: Power coupling of the blank hologram replay and the weakly-guiding fibre mode with an angular-tilt and a lateral-shift in a 4*f* holographic routing architecture.

Figure 3.3 illustrates the output power coupling of a 4*f* holographic replay system; subscripts '$_1$' and '$_2$' denote the Gaussian field distributions, as an approximate description of the true transverse fibre mode fields of the fibres located at the *input* and *replay* planes, each in the absence of the other. The output Gaussian fibre mode has a spot width w_{r2}, whereas the desired transformed Gaussian has a spot width w_{r1} (without any clipping/vignetting).

The coupling intensity into a fibre using the overlap-integral is thus[*]:

$$\left|R_c(u,v;\theta_x,\theta_y)\right|^2 = \frac{\left|\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} R(x_r,y_r)G_2(w_{r2};x_r-u,y_r-v;\theta_x,\theta_y)dx_r dy_r\right|^2}{\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left|R(x_r,y_r)\right|^2 dx_r dy_r \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left|G_2(w_{r2};x_r-u,y_r-v;\theta_x,\theta_y)\right|^2 dx_r dy_r}, \tag{3.11}$$

The replay field $R(x_r,y_r)$ is generally not known analytically but it is the optical Fourier transform of the Gaussian beam at the hologram plane:

$$G_1(w_{h1};x_h,y_h;0,0) = \frac{1}{w_{h1}}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{x_h^2+y_h^2}{w_{h1}^2}\right). \tag{3.12}$$

G_1 is the transform of the input Gaussian transverse field distribution. G_2 is the Gaussian transverse field distribution of the output fibre at the replay plane having a unity power:

$$G_2(w_{r2};x_r-u,y_r-v;\theta_x,\theta_y) = \frac{1}{w_{r2}}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{(x_r-u)^2+(y_r-v)^2}{w_{r2}^2} + jk\left(\theta_x(x_r-u)+\theta_y(y_r-v)\right)\right). \tag{3.13}$$

The output fibre is placed at an arbitrary co-ordinate, (u,v), from the replay beam axis ($x_r = 0$, $y_r = 0$), to simulate both the coupling of the desired main peak and sidelobes. The fibre axis is also tilted from the optical axis by (θ_x,θ_y) on the x-z and y-z planes, respectively. Using small angle approximation, $sin(\theta_{x,y}) = \theta_{x,y}$, the tilted fibre can be assumed to add an additional phase to the exponent of the Gaussian distribution [69], $k\theta_x x_r$ and $k\theta_y y_r$, where k is the wavenumber. The angular tilts have to be small in order that small angle approximation is valid, but the lateral offsets can assume fairly large values.

Although Figure 3.3 implicitly assumes the replay of a blank hologram, the analysis is valid in that the steering function of the hologram can be considered elsewhere. The normalising factors in the denominator of eq. (3.11) will be both unity, if power loss due to the blocked illumination at the hologram plane is included (coupling efficiencies as a percentage of the highest coupling efficiency when there is no truncation at the hologram plane). We are

---

[*] The complex conjugate of the replay field, $R^*(x_r,y_r)$, should be used in the overlap-integral for a (non-blank) hologram replay where the replay field may have a phase dependence.

interested in the coupling intensity of the replay peak into a single mode fibre at the replay plane, hence wavefront curvature effects of Gaussian beams have been neglected (i.e., both input and replay planes are the focal planes of the collimating and transform lenses).

It is difficult to solve the integral of the replay field and a Gaussian fibre mode with a lateral offset and angular tilt as in eq. (3.11). Instead, using Parseval's theorem, the integrated power can be obtained from the overlap of the field distributions at the transform plane. Therefore, the overlap-integral of the *optical Fourier transforms* of the *laterally-shifted & angularly-tilted output Gaussian fibre mode* and the *non-shifted, non-tilted input Gaussian fibre mode* is evaluated at the hologram plane over the truncation aperture. The coupling intensity into a fibre, for an equal and symmetric truncation, $|R_c(u,v;\theta_x,\theta_y)|^2$ is thus:

$$\begin{aligned}|R_c(u,v;\theta_x,\theta_y)|^2 = \Bigg| \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Bigg\{ \Bigg[\int_{-\frac{\gamma}{2}w_{h1}}^{\frac{\gamma}{2}w_{h1}} \int_{-\frac{\gamma}{2}w_{h1}}^{\frac{\gamma}{2}w_{h1}} \frac{G_1(w_{h1};x_h,y_h;0,0)}{f\lambda} \exp\left(\frac{jk}{f}(x_r x_h + y_r y_h)\right) dx_h dy_h \Bigg]. \\ \times G_2(w_{r2};x_r-u,y_r-v;\theta_x,\theta_y) \Bigg\} dx_r dy_r \Bigg|^2\end{aligned} \tag{3.14}$$

Interchanging the order of integration,

$$\begin{aligned}|R_c(u,v;\theta_x,\theta_y)|^2 = \Bigg| \int_{-\frac{\gamma}{2}w_{h1}}^{\frac{\gamma}{2}w_{h1}} \int_{-\frac{\gamma}{2}w_{h1}}^{\frac{\gamma}{2}w_{h1}} \Bigg\{ \Bigg[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{G_2(w_{r2};x_r-u,y_r-v;\theta_x,\theta_y)}{f\lambda} \exp\left(\frac{jk}{f}(x_r x_h + y_r y_h)\right) dx_r dy_r \Bigg], \\ \times G_1(w_{h1};x_h,y_h;0,0) \Bigg\} dx_h dy_h \Bigg|^2\end{aligned} \tag{3.15}$$

$$\begin{aligned}|R_c(u,v;\theta_x,\theta_y)|^2 = \Bigg| \int_{-\frac{\gamma}{2}w_{h1}}^{\frac{\gamma}{2}w_{h1}} \int_{-\frac{\gamma}{2}w_{h1}}^{\frac{\gamma}{2}w_{h1}} \frac{1}{w_{h1}w_{h2}} \frac{2}{\pi} \exp\left(-\frac{(x_h+f\theta_x)^2}{w_{h2}^2} - \frac{x_h^2}{w_{h1}^2} + \frac{jkux_h}{f}\right). \\ \times \exp\left(-\frac{(y_h+f\theta_y)^2}{w_{h2}^2} - \frac{y_h^2}{w_{h1}^2} + \frac{jkvy_h}{f}\right) dx_h dy_h \Bigg|^2\end{aligned} \tag{3.16}$$

Eq. (3.16) is the expanded form of the overlap integral at the hologram plane:

$$\left|R_c(u,v;\theta_x,\theta_y)\right|^2 = \frac{\left|\iint\limits_{\substack{hologram\\plane}} G_1\left(w_{h1};x_h,y_h;0,0\right)G'_2\left(w_{h2};x_h+f\theta_x,y_h+f\theta_y;\frac{u}{f},\frac{v}{f}\right)dx_h dy_h\right|^2}{\iint\limits_{\substack{input\\plane}}\left|G_1\left(w_{i1};x_i,y_i;0,0\right)\right|^2 dx_i dy_i \iint\limits_{\substack{replay\\plane}}\left|G_2\left(w_{r2};x_r-u,y_r-v;\theta_x,\theta_y\right)\right|^2 dx_r dy_r} \text{,where} \quad (3.17)$$

$$G'_2(w_{h2};x_h+f\theta_x,y_h+f\theta_y;\frac{u}{f},\frac{v}{f}) = \frac{1}{w_{h2}}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{(x_h+f\theta_x)^2}{w_{h2}{}^2}+\frac{jkux_h}{f}\right)\exp\left(-\frac{(y_h+f\theta_y)^2}{w_{h2}{}^2}+\frac{jkvy_h}{f}\right). \quad (3.18)$$

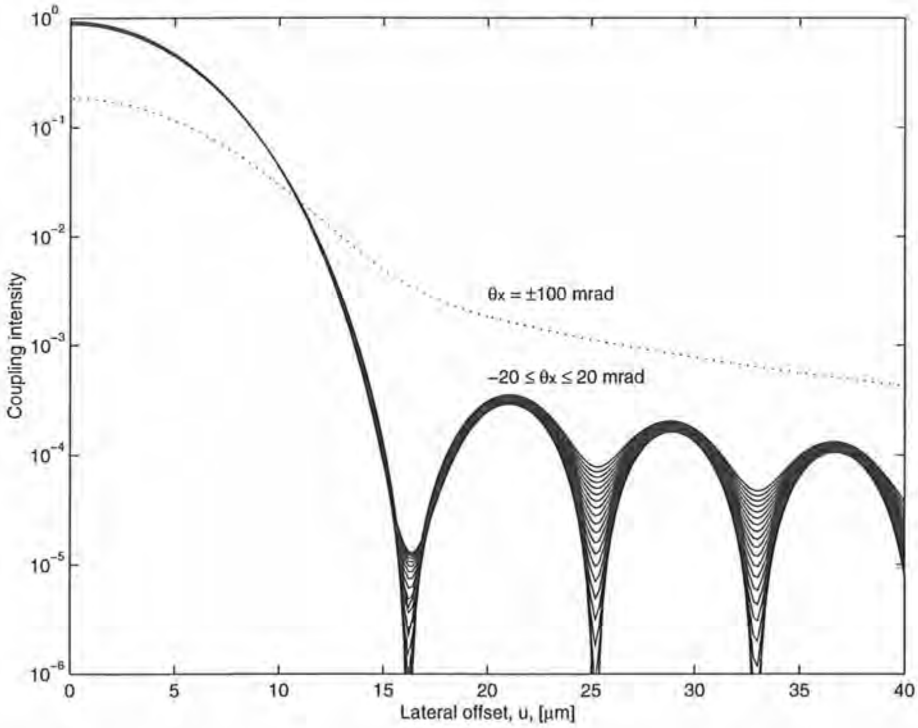


Figure 3.4: Numerical integration of the coupling intensity into an output fibre with 1D lateral offsets and angular tilts.

It is intuitively expected that the lateral shifts and phase due to the angular tilts of the output fibre have been transformed to additional phase shifts and beam axis offsets, respectively. Analytical solution of the integral in eq. (3.16) is impossible. The numerical integration results for the 1D case are shown in Figure 3.4. The integration has been performed by assuming a symmetric $\gamma = 2$ truncation and $f = 25$ mm for $\lambda = 1.55$ µm Gaussian beams. The coupling intensity profile with lateral offsets is similar to the replay intensity profile. However, with the inclusion of angular tilts in the output fibre, 'blurring' effects are obtained. The troughs of the coupling intensity profile are smoothed out and the general sidelobe level is raised. For large angular tilts, e.g., ±100 mrad, the coupling intensity profile is broadened along with a reduced on-beam-axis launch power and levelled-off sidelobes.

Analytically, there are two cases to be considered; the first when it involves only lateral shifts (i.e., $\theta_x = 0$ & $\theta_y = 0$); and the other when it involves only angular tilt (i.e., $u = 0$ & $v = 0$) alignment errors. These two coupling conditions are examined in the following subsections.

3.4.1 Coupling intensity for a single mode fibre with lateral offsets (u,v)

Eq. (3.16) becomes simpler when $\theta_x = 0$ & $\theta_y = 0$. It is the familiar integral of a truncated Gaussian field. The x and y integrals are separable. Hence the analysis in appendix A is applicable. For the x integral, the complex arguments for the error function are:

$$z_1 = \frac{\gamma\sqrt{w_{h1}^2 + w_{h2}^2}}{2w_{h1}} - \frac{ju}{\sqrt{w_{h1}^2 + w_{h2}^2}}, \qquad z_2 = \frac{\gamma\sqrt{w_{h1}^2 + w_{h2}^2}}{2w_{h1}} + \frac{ju}{\sqrt{w_{h1}^2 + w_{h2}^2}}. \tag{3.19}$$

This yields the coupling intensity into a single-mode fibre positioned at (u,v) from the axis of the replay peak as below:

$$\left|R_c(u,v;0,0)\right|^2 \approx \left(\frac{2w_{r1}w_{r2}}{w_{r1}^2 + w_{r2}^2}\right)^2 \times \left|R_c(u;0,0)\right|^2 \times \left|R_c(v;0,0)\right|^2, \text{ where} \tag{3.20}$$

$$\begin{aligned}\left|R_c(u;0,0)\right|^2 \sim \Bigg|&\exp\left(-\frac{u^2}{w_{r1}^2 + w_{r2}^2}\right) - \frac{\exp\left(-\frac{\gamma^2(w_{r1}^2 + w_{r2}^2)}{4w_{r1}^2}\right)}{\sqrt{\pi}\left(\frac{\gamma^2(w_{r1}^2 + w_{r2}^2)}{4w_{r1}^2} + \frac{u^2}{w_{r1}^2 + w_{r2}^2}\right)}. \\ &\times\left[\frac{\gamma\sqrt{w_{r1}^2 + w_{r2}^2}}{2w_{r1}}\cos\left(\frac{\gamma u}{w_{r1}}\right) - \frac{u}{\sqrt{w_{r1}^2 + w_{r2}^2}}\sin\left(\frac{\gamma u}{w_{r1}}\right)\right]\Bigg|^2\end{aligned} \tag{3.21}$$

For identical input/output single mode fibres (i.e., $w_r = w_{r1,2}$) further simplification yields,

$$\left|R_c(u;0,0)\right|^2 \sim \left|\exp\left(-\frac{u^2}{2w_r^2}\right) - \frac{\exp\left(-\frac{\gamma^2}{2}\right)}{\sqrt{\pi}\left(\frac{\gamma^2}{2} + \frac{u^2}{2w_r^2}\right)}\left[\frac{\gamma}{\sqrt{2}}\cos\left(\frac{\gamma u}{w_r}\right) - \frac{u}{\sqrt{2}w_r}\sin\left(\frac{\gamma u}{w_r}\right)\right]\right|^2. \tag{3.22}$$

Eq. (3.22) is the asymptotic representation of the replay field that is coupled into a single mode fibre located at (u,v) from the axis of the replay peak. It agrees with Marcuse's [20] coupling efficiency expression for a perfect Gaussian replay when $\gamma \rightarrow \infty$. At this limit, the two x-, y-orthogonal components become radially symmetric. The validity of eq. (3.22) can be obtained by referring to the argument requirement expression of the asymptotic approximation, which means γ cannot be >> $\sqrt{2}$. The same condition, $|tan^{-1}(u/\gamma w_r)| < \pi/2$ must apply and thus, u cannot take indefinitely large numbers.

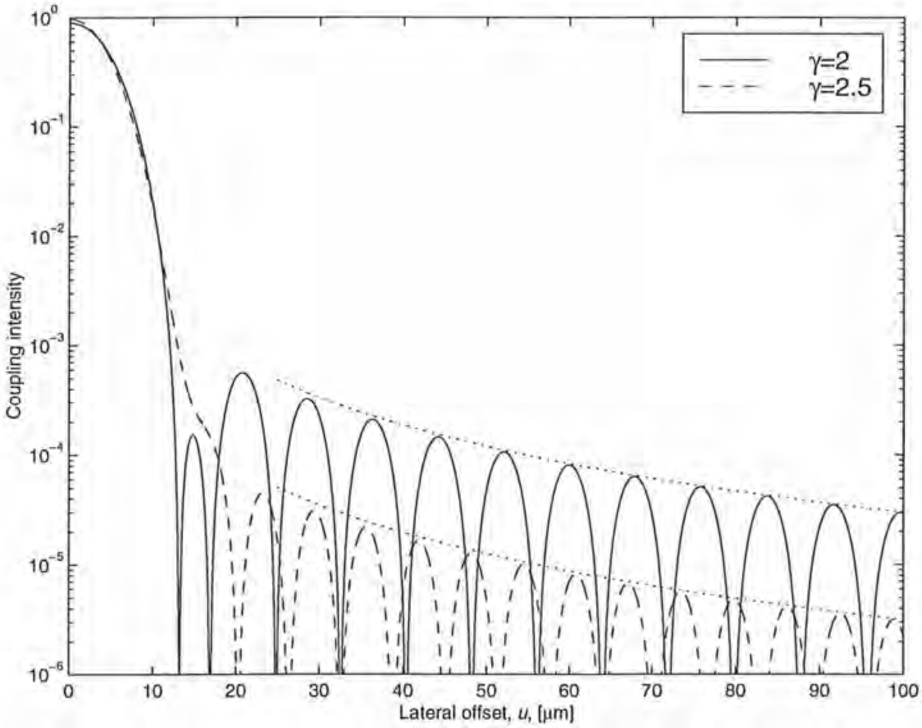


Figure 3.5: 1-D coupling intensity profile using asymptotic approximation for $w_r = 5.06$ μm.

The vignetting of Gaussian tails by a square/rectangular hologram aperture causes ripples in the coupling intensity profile, described by an approximation with two separable x-, y-orthogonal components and a prefactor. The prefactor is the defocusing term due to the mismatch of input and output fibre mode radii. Eq. (3.22) represents the 1-D contribution when the input and output fibre mode radii are similar. It contains the truncation ratio γ, Gaussian radius and offset co-ordinate as variables. Figure 3.5 plots the coupling intensity according to eq. (3.22) for $\gamma = 2$ and 2.5. Note that the coupling intensity is not much reduced after the initial *exponential-like* fall. The troughs in these plots are rather sharp but will be less so for a hologram illumination truncated asymmetrically. A fibre port with its axis located at some large distances away (within the paraxial limit) from the main replay peak is likely to couple in the peak of the sidelobe power, which is approximately $(2/\pi)(w_r/u)^2\exp(-\gamma^2)$. This is shown by the dotted lines in Figure 3.5.

The coupling of the replay field into a fibre is not a 1-D phenomenon. The total coupling intensity has contributions from both orthogonal components. The 2-D coupling intensity is given by a simple product of the u and v terms. Thus, we can expect the coupling intensity into a fibre port located at off-transverse-axes locations of the main replay peak to have roughly half, on the logarithmic scale, the on-transverse-axis coupling power provided a minimum distance criterion is met (e.g., from -30 dB to -60 dB). The minimum distance criterion is the separation between the beam axis of the main peak and the first minimum in the coupling intensity profile. With $\gamma = 2$ and 2.5, the minimum distance corresponds to a spacing of $2 \rightarrow 3$ 'coarse' replay resolutions where a coarse replay resolution, $\Delta r = \pi w_r/\gamma$ (coarse as only N addressable points for an N-pixel SLM is considered).

Further approximation of the peak sidelobe coupling intensity for off-transverse-axes locations yields:

$$\left|R_c(u,v;0,0)\right|^2 \approx \left(\frac{2}{\pi}\right)^2\left(\frac{w_r^{\,2}}{uv}\right)^2 \exp\left(-\left(\gamma_x^{\,2}+\gamma_y^{\,2}\right)\right). \tag{3.23}$$

Offsets $|u|$ and $|v|$ are $> 5a$, where a is the fibre core radius and γ_x and γ_y are the truncation ratios in x_h and y_h directions, respectively. Where the offset is aligned to a transverse-axis direction (e.g., y), further approximation gives,

$$\left|R_c(u=0,v;0,0)\right|^2 \approx \left(\frac{2}{\pi}\right)\left(\frac{w_r}{v}\right)^2\left(erf\left(\frac{\gamma_x}{\sqrt{2}}\right)\right)^2 \exp\left(-\gamma_y^{\,2}\right), \tag{3.24}$$

where $|v| > 5a$ and $u = 0$. The contribution from the x-component at $u = 0$ has been taken to be an *erf* function which is exact, rather than the value of $|R_c(u;0,0)|^2$ at $u = 0$. A quadrant of the 2-D coupling intensity profile, according to eq. (3.22) is shown in Figure 3.6 for $\gamma_x = \gamma_y = 2$. Eqs. (3.23) and (3.24) are plotted as dashed and dotted lines. The square or rectangular clipping of the Gaussian beam at the hologram plane gives rise to a significant sidelobe power along both transverse-axis directions. However, off-transverse-axis locations would typically receive a crosstalk power of -70 to -60 dB. This agrees qualitatively with the numerical results for an optical system without aberrations by Churin *et al.* [70]. Before aberrations dominate for $\gamma > 3$, the crosstalk level at off-transverse-axis locations is about twice as low in dB terms compared to on-transverse-axis locations.

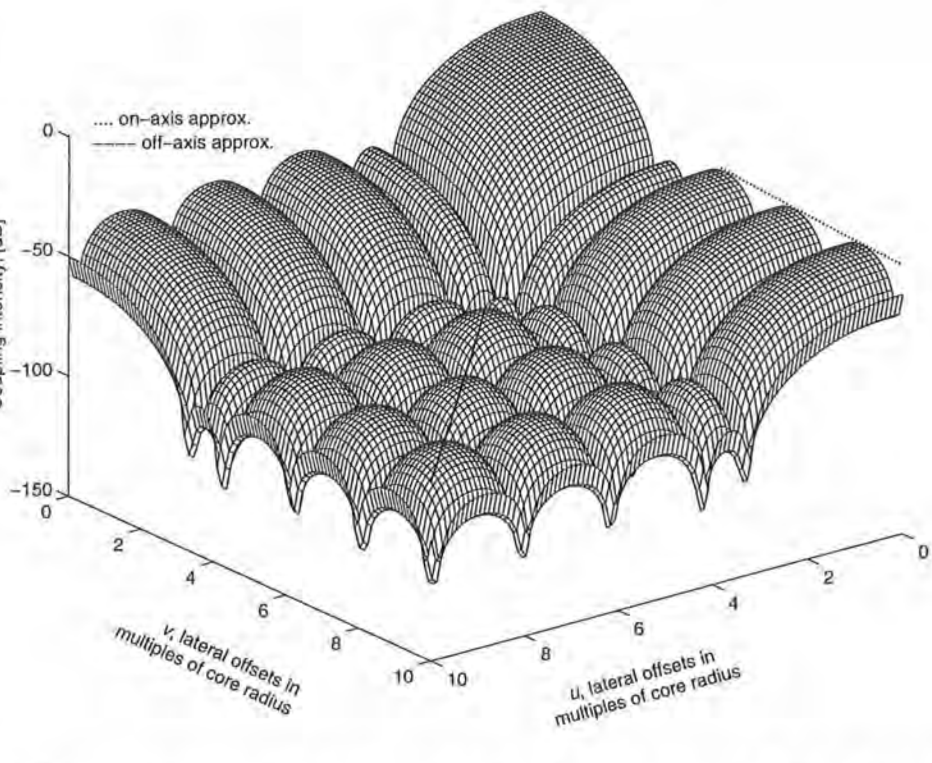


Figure 3.6: Coupling intensity for an output fibre with two-dimensional lateral offsets (u,v) for a symmetric truncation with ratio $\gamma = 2$ and similar input/output fibre modes.

The asymptotic approximation has a well-defined error expression. The error expression of including only the first n terms in the asymptotic expansion, for a symmetric truncation and similar input/output fibres, is:

$$Error(u) < \left| \frac{\exp\left(-\frac{u^2}{2w_r^2}\right)\exp\left(-z_1^2\right)\Gamma(n+\frac{1}{2})}{2\pi|z_1|^{2n+1}\cos(\phi_{z1})} + \frac{\exp\left(-\frac{u^2}{2w_r^2}\right)\exp\left(-z_2^2\right)\Gamma(n+\frac{1}{2})}{2\pi|z_2|^{2n+1}\cos(\phi_{z2})} \right|. \quad (3.25)$$

Since $z_1 = z_2^*$, thus $|z_1| = |z_2|$ and $\cos(\phi_{z1}) = \cos(\phi_{z2})$, the absolute error expression becomes:

$$Error(u) < \left\{\exp\left(-\frac{\gamma^2}{2}\right)\Gamma(n+\frac{1}{2})\left|\cos\left(\frac{\gamma u}{w_r}\right)\right|\right\} \Big/ \left\{\pi\left(\frac{\gamma^2}{2}+\frac{u^2}{2w_r^2}\right)^{2n}\left(\frac{\gamma}{\sqrt{2}}\right)\right\}. \quad (3.26)$$

The errors of coupling intensity expression are illustrated graphically in Figure 3.7. These errors are $< 10^{-3}$ and hence are within an acceptable range.

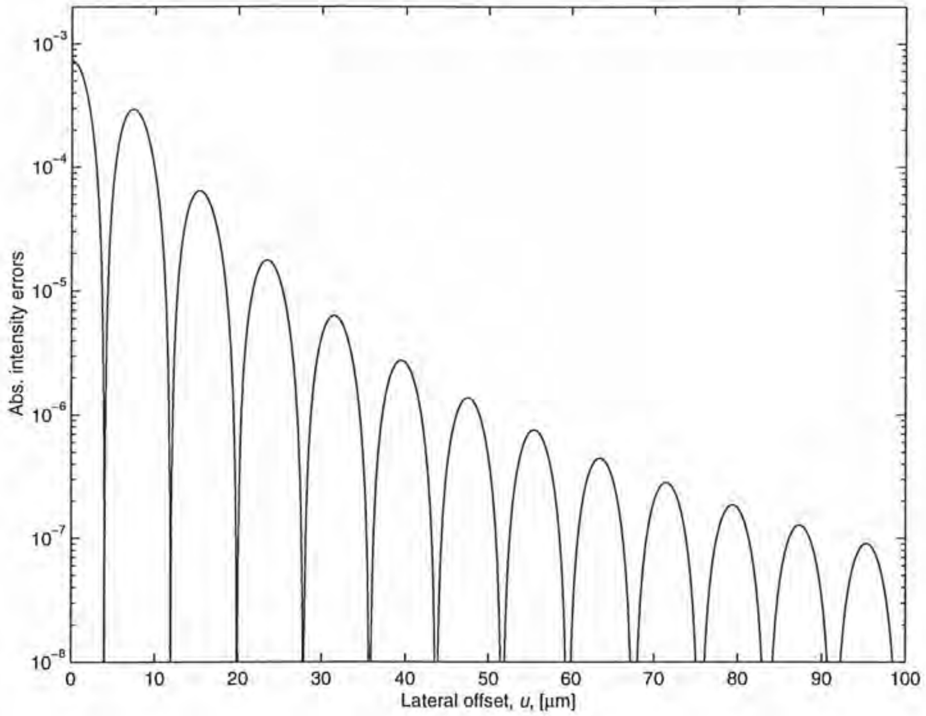


Figure 3.7: Absolute intensity errors incurred by including only one term in the asymptotic series expansion for $\gamma = 2$ and $w_r = 5.06$ μm.

3.4.2 *Coupling intensity for a single mode fibre with angular tilts (θ_x, θ_y)*

For the case where only angular tilts are involved in the coupling of the replay field into a single mode fibre, (u,v) in eq. (3.16) is (0,0). Thus the coupling power into a fibre, separating the x and y integrals, is,

$$\left|R_c(0,0;\theta_x,\theta_y)\right|^2 = \left(\frac{2w_{r1}w_{r2}}{w_{r1}^2 + w_{r2}^2}\right)^2 \times \left|R_c(0,0;\theta_x)\right|^2 \times \left|R_c(0,0;\theta_y)\right|^2 \text{ where,}$$

$$\left|R_c(0,0;\theta_x)\right|^2 = \left|\exp\left(-\frac{(k\theta_x w_{r1}w_{r2}/2)^2}{(w_{r1}^2 + w_{r2}^2)}\right) \times \frac{1}{2}\left[erf\left(\frac{\gamma\sqrt{(w_{r1}^2 + w_{r2}^2)}}{2w_{r1}} + \frac{k\theta_x w_{r2}^2}{2\sqrt{(w_{r1}^2 + w_{r2}^2)}}\right) + erf\left(\frac{\gamma\sqrt{(w_{r1}^2 + w_{r2}^2)}}{2w_{r1}} - \frac{k\theta_x w_{r2}^2}{2\sqrt{(w_{r1}^2 + w_{r2}^2)}}\right)\right]\right|^2 . \quad (3.27)$$

Erf is the error function definition [71] which takes real-only arguments. When the Gaussian spot widths of the input and output fibres are identical (i.e., $w_{r1} = w_{r2} = w_r$), eq. (3.27) yields,

$$\left|R_c(0,0;\theta_x)\right|^2 = \left|\exp\left(-\frac{1}{2}\left(\frac{k\theta_x w_r}{2}\right)^2\right)\times\frac{1}{2}\left[erf\left(\frac{\gamma}{\sqrt{2}}+\frac{k\theta_x w_r}{2\sqrt{2}}\right)+erf\left(\frac{\gamma}{\sqrt{2}}-\frac{k\theta_x w_r}{2\sqrt{2}}\right)\right]\right|^2 . \quad (3.28)$$

Figure 3.8 plots eq. (3.28) for $\gamma = 2$, 2.5 and 4. Marcuse's expression (i.e., the exponential term in eq. (3.28) [20]) scaled by the on-beam-axis coupling value is plotted as a 'dash-dot' line for $\gamma = 2$. The apparent errors of using only the exponential term point to the necessity of evaluating the two *erf* functions for each tilt value. The fall in coupling power as the output fibre tilt increases is rapid. The 3 dB coupling points are approximately ±80 mrad. The coupling intensity into an output fibre aligned with an angular tilt for large γ values can assume only the exponential term. In this case, the replay beam is Gaussian-like.

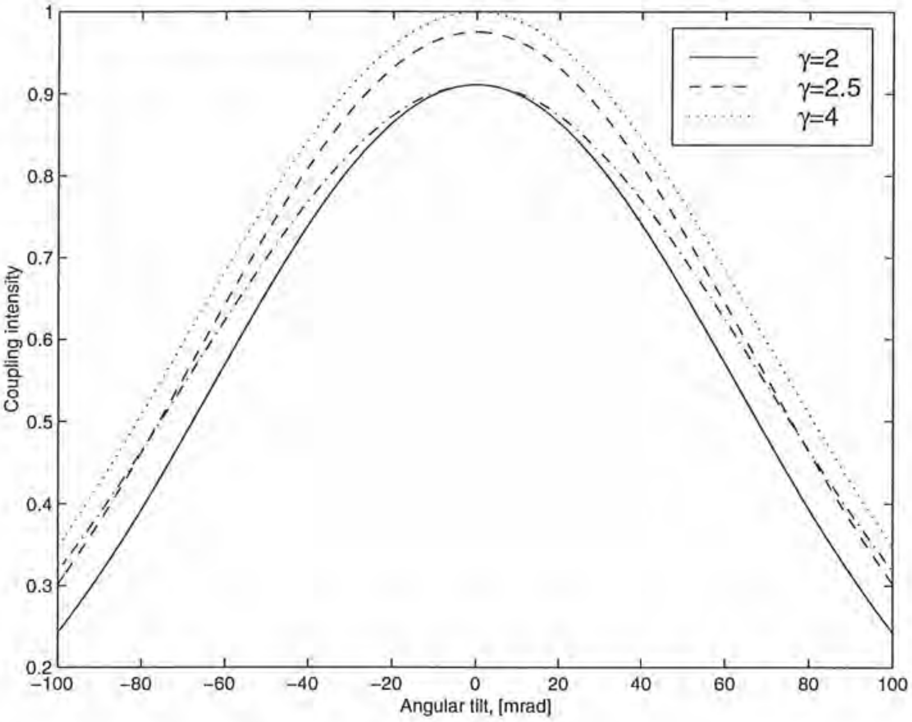


Figure 3.8: Coupling intensity for tilted fibres.

When the '1-D output fibre' is not tilted with respect to the optic axis (i.e., $\theta_x = 0$), the on-beam-axis coupling intensity given by eq. (3.28) is $(erf(\gamma/\sqrt{2}))^2$. This fraction of the coupling of power within the truncation window into an output fibre differs from the on-beam-axis coupling intensity of eq. (3.22) when $u = 0$. The reason is that, eq. (3.22) is only an approximation and is most accurate for large u values, whereas eq. (3.28) is an exact analytic expression, although only numerical evaluation of the error function is possible. Note that, the small angle approximation is only valid for θ_x, θ_y << 1 (e.g., 0.1 rad). The coupling intensity plot for an output fibre with 2-D angular tilts, (θ_x,θ_y) is shown in Figure 3.9 for a symmetric truncation ratio $\gamma = 2$ and similar input/output fibres.

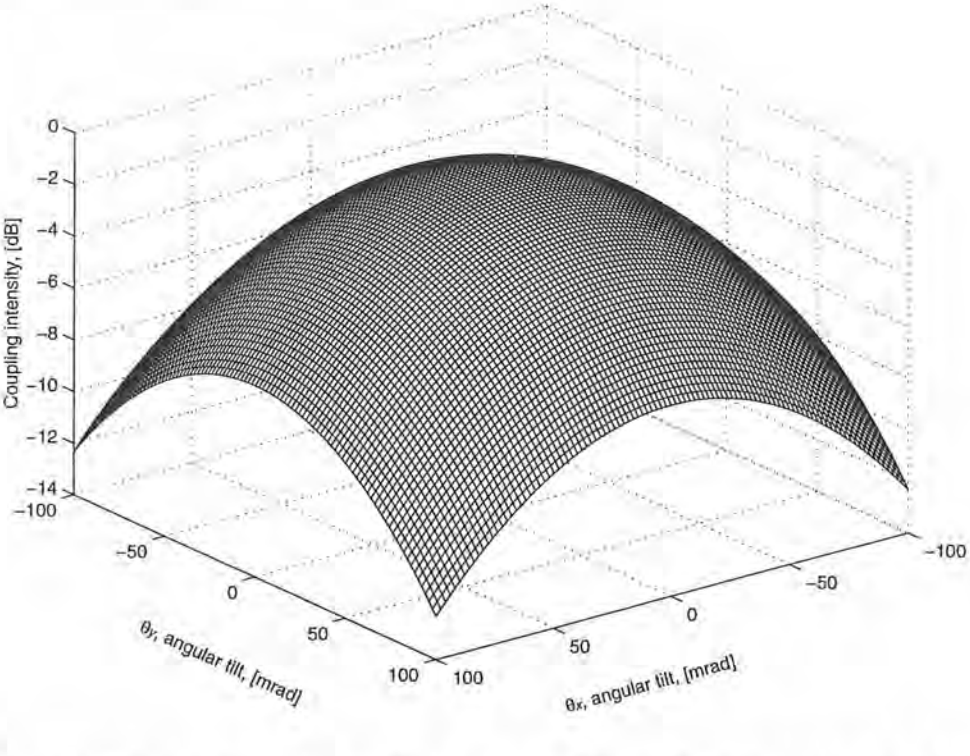


Figure 3.9: Coupling intensity profile for an output fibre with 2-D angular tilts, (θ_x,θ_y)

3.5 Numerical simulations of the replay field and coupling intensity profiles

The coupling of replay power into an output fibre is a 2-D phenomenon. All expressions involved other than the Bessel function fibre mode have two orthogonal components. Hence, by assuming the 2-D mathematical representations of the optical fields and setting y truncation ratio to be ∞ and its lateral offset and angular tilt to be 0, manageable and accurate simulations of *'a 1-D slice of the 2-D coupling intensity'* can be performed. As in the analysis of §3.4, the coupling intensities with lateral offsets and angular tilts are simulated separately. The Gaussian radii for the input and output fibres, w_{r1} and w_{r2} are assumed equal and assigned w_r. This takes the average of Marcuse's and Snyder's Gaussian parameters. The core radius of fibre, a, is 4.1 μm.

The primary care in performing numerical *fft* is to ensure adequate sampling of the functions represented by weighted deltas. In Figure 3.10(a) the solid line shows the Bessel function transverse field of a fibre mode exiting the cleaved end of a single-mode fibre at the *input* plane. Its Gaussian approximation is plotted as the dotted line. The solid line plot in Figure 3.10(b) represents the magnitude of either field distribution at the exit pupil of the *hologram* plane after an *fft* operation. Note that, there is no $f\lambda$ (focal length × wavelength) value assumed in the *fft* algorithm; the size of the transformed beam depends entirely on the aperture of the input vector, D_i, which in this case is $1000a$ at $\Delta_i = 0.02a$ *steps*. By taking an analogy from the digital and optical Fourier transforms,

$$\Delta_h = \frac{1}{D_i},\ w_h = \frac{f\lambda}{\pi w_r} \text{ and } N_{w_h} \times \Delta_h = \frac{w_h}{a^2}, \tag{3.29}$$

where Δ_h is the hologram plane resolution; N_{w_h} is the number of points used to represent the transformed Gaussian distribution from 0 to its spot radius w_h. Normalisation of w_h by a^2 is needed for consistency of the dimensions. The number of points for the spot width of the transformed Gaussian is thus,

$$N_{w_h} = \frac{D_i}{\pi w_r} \times \frac{f\lambda}{a^2}. \tag{3.30}$$

The co-ordinate axis for the plots in (b) has been multiplied by Δ_i and then normalised by $f\lambda/a$. Hence the spot width of the numerically transformed Gaussian corresponds to $5.16a$ or > 500 points within the spot diameter. This is an adequate representation of the fields.

The replay field is the *fft* of this transformed field truncated at $\pm w_h$. Consequently the truncation ratio $\gamma = 2$ for these simulations. This represents the worst case accuracy of eq. (3.22). The solid line plot in Figure 3.10(c) shows the magnitude of the replay field, which is the *fft* of the field shown by the solid line at the *hologram* plane. The dashed line plot in (c) represents the field magnitude of the fibre mode with a tilted axis. Its *fft* is shown as the dashed line at the *hologram* plane in (b). This is done so that the coupling intensity for an output fibre mode with angular tilts can be evaluated as a numerical convolution at the exit pupil of the *hologram* plane. The evaluation of the coupling intensity for a shifted output fibre mode is also a numerical convolution process. This is performed at the *replay* plane between the replay field and the fibre mode with a lateral shift (dotted line plot in (c)).

It has been assumed that a tilted fibre mode adds an additional phase to the field expression, regardless of whether a Bessel function or a Gaussian distribution is used. Simulations using both the Bessel function and Gaussian fibre mode descriptions were conducted. The results are compared with analytic expressions assuming a Gaussian distribution for single mode fibres.

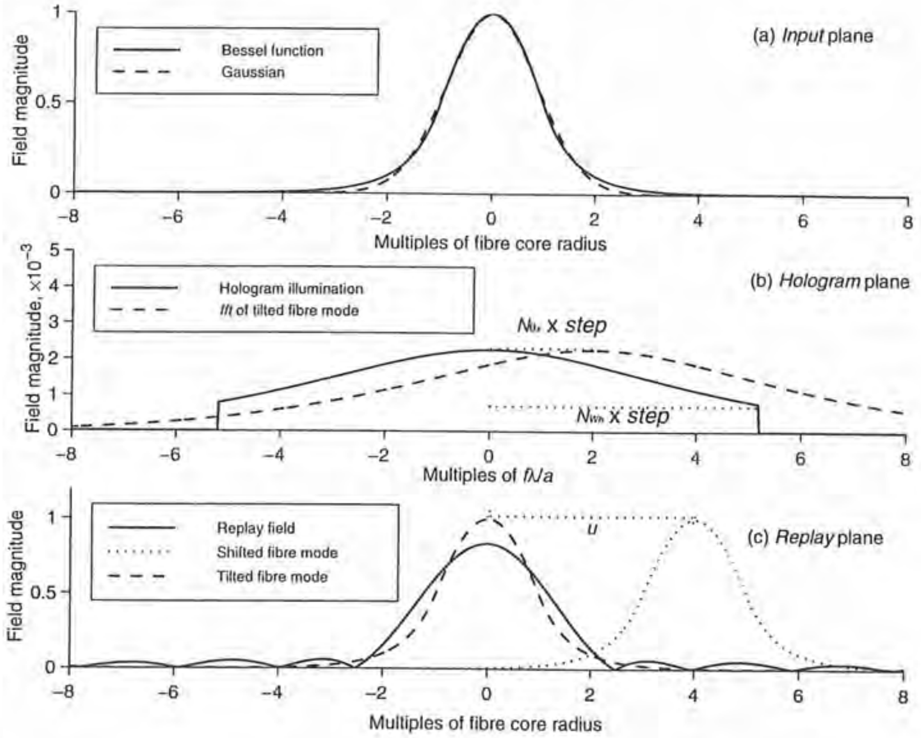


Figure 3.10: The field profiles at the (a) *input*, (b) *hologram* and (c) *replay* planes for numerical simulations.

3.5.1 *Approximate replay peak descriptions*

The normalised (i.e., Gaussian field at the input plane has been normalised by the square root of total fibre mode power) replay field, or the optical Fourier transform of the Gaussian field within the truncation aperture, can be expressed asymptotically as,

$$R(x_r,y_r) \approx \sqrt{P_o \Big/ \frac{\pi}{2}\left(\frac{w_r}{a}\right)^2} \times R(x_r) \times R(y_r), \tag{3.31}$$

where $R(x_r)$ is given by eq. (3.10). This assumes a symmetrical truncation in x and y directions. Making the x-direction truncation ratio 2 and the y-direction ∞, the variation of the replay field magnitude in the x-direction according to eq. (3.31) is plotted as the dotted line in Figure 3.11. The solid and dashed lines in the same figure show the numerical results for the replay field using the Bessel function and Gaussian input fibre modes, respectively. The transform of input fibre mode is truncated at $\pm w_{hl}$ in the x direction and not truncated in the y direction. An *fft* of this field gives the replay field profile as shown.

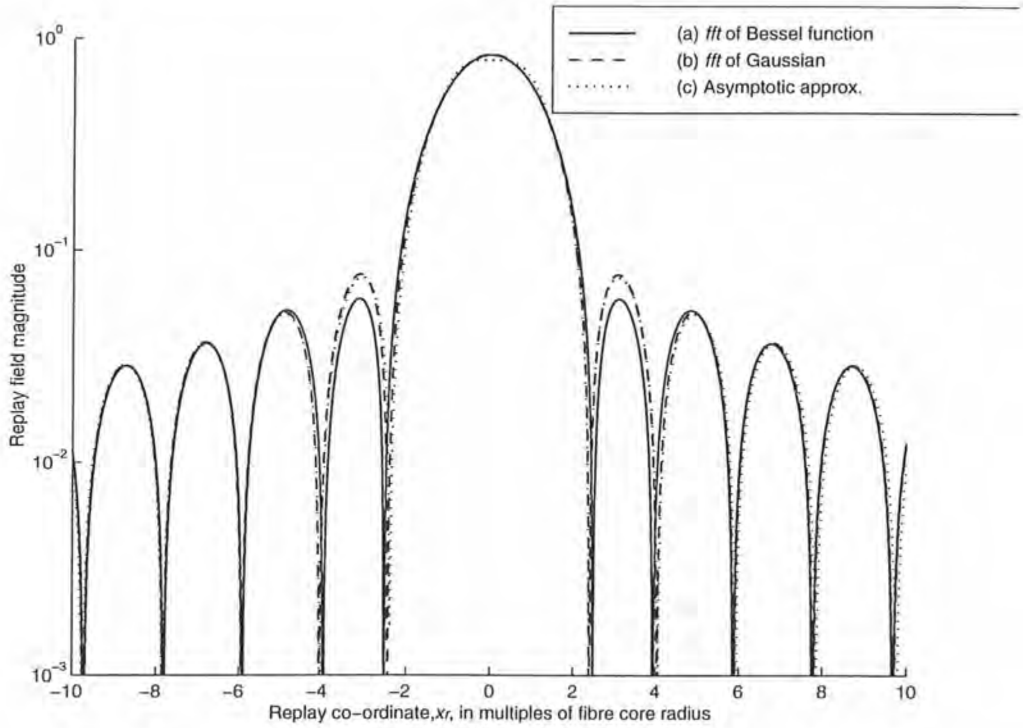


Figure 3.11: Approximate replay field magnitude using analytic expression and numerical *fft* of the blank hologram field at the exit pupil.

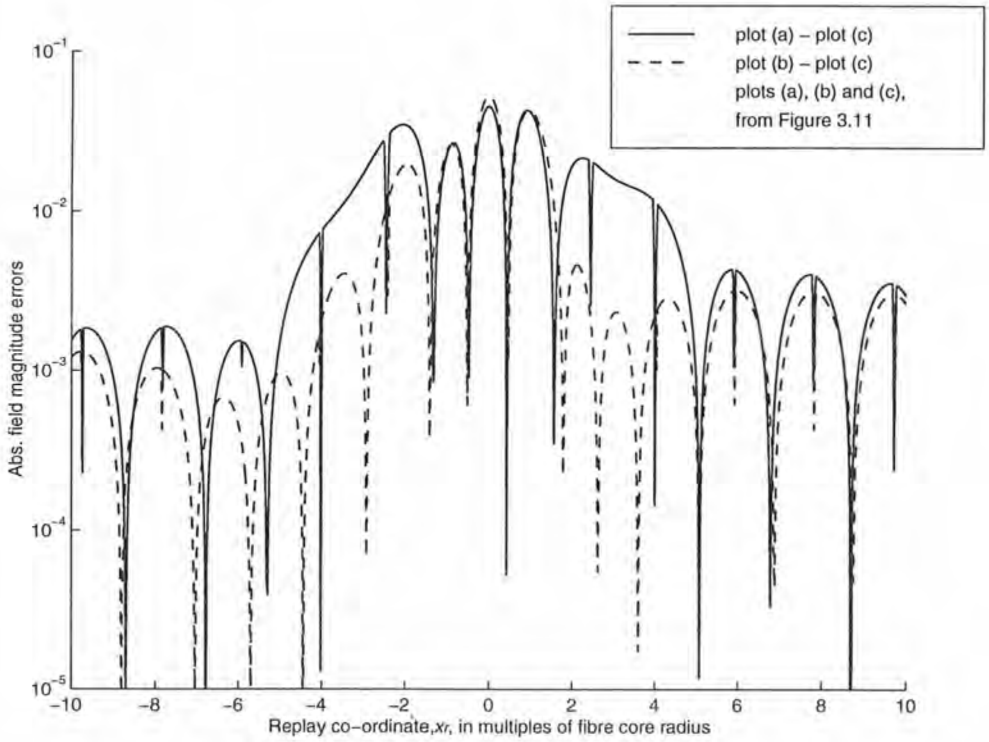


Figure 3.12: Absolute errors of the replay field magnitude by using asymptotic approximation compared to numerical *fft*.

The errors of using an asymptotic approximation based on Gaussian fibre mode field, instead of the *fft* of the fibre mode fields are shown in Figure 3.12. Solid and dashed lines plot the errors of using the asymptotic replay field approximation as compared to the *fft* of the Bessel function and Gaussian input fibre modes, respectively. For $|x_r| > 5a$, the errors are $< 0.2\%$. As expected, the asymptotic expression is most accurate for large arguments. The condition of the expansion dictates that $\gamma << 2$ gives erroneous results. Also, x_r cannot assume indefinitely large numbers. The numerical results for the replay field are not symmetrical. The field description at the hologram plane as a result of performing an *fft* on the input field is Hermitian. Hence, its field distribution at the replay plane using another *fft* is asymmetric.

Asymptotic approximation for the replay of a truncated fibre mode field at the *hologram* plane is fairly accurate, especially for large displacements from the beam axis. However, it is more important to describe how this power is received by a finite size detector. For an optical switch, the '*detector*' acts as a fibre-mode filter. This is simulated using the derived expressions eqs. (3.22) and (3.28) in the following two subsections.

3.5.2 Approximate coupling intensity for an output fibre with a 1-D lateral offset, u

The coupling intensity expressions of eqs. (3.22) and (3.28) are compared with numerical convolution results in order to evaluate their accuracy. For the case of coupling with lateral offsets, the numerical convolution of a '*1-D slice of 2-D fields*' performs the following expression:

$$\left|R_c(u;0,0)\right|^2 = \left|F_{fm}(w_r;x_r,0;0,0) * sinc(\frac{\gamma}{w_r}x_r,0) * F_{fm}(w_r;x_r,0;0,0)\right|^2, \tag{3.32}$$

where F_{fm} is the fibre mode field described as a Bessel function or Gaussian distribution with an equivalent Gaussian spot width w_r and $*$ is the numerical convolution operator. The analytical approximation (dotted line) and numerical convolution (solid line for Bessel function and dashed line for Gaussian fibre modes) results are plotted in Figure 3.13 for $\gamma = 2$. Figure 3.14 shows the absolute value of the difference between the numerical (solid line for Bessel function and dashed line for Gaussian fibre modes) and analytic approximation results in Figure 3.13. All parameters are as shown in Figure 3.13. The on-beam-axis numerical results have been normalised to $(erf(\gamma/\sqrt{2}))^2$ (i.e., the overlap of two Gaussian distributions evaluated over the aperture).

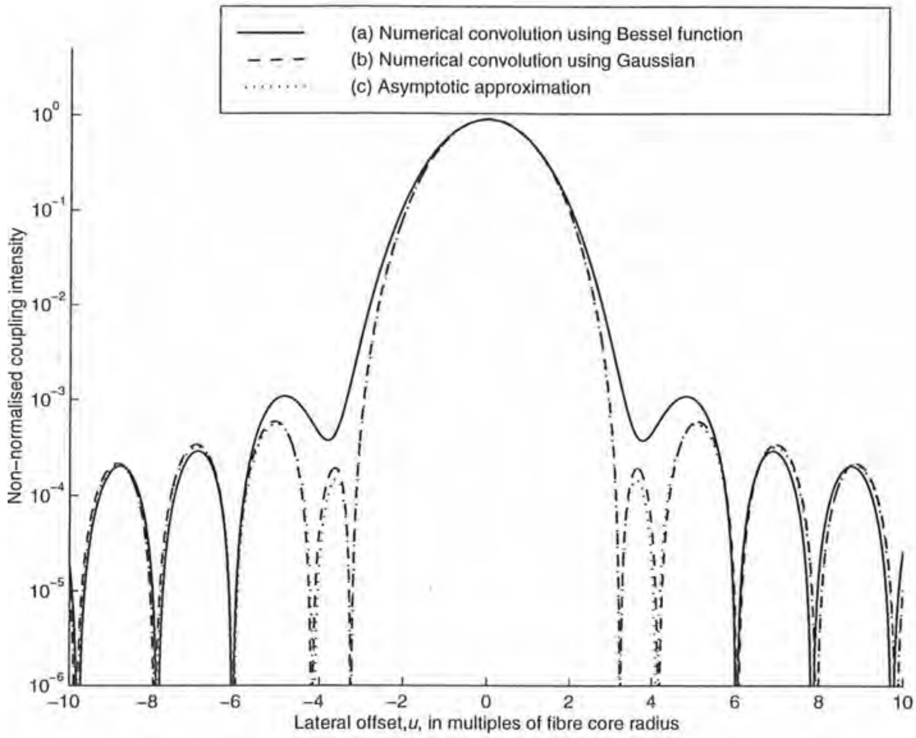


Figure 3.13: Coupling intensity for an output fibre with a lateral offset u, Gaussian beam radius $w_r = 5.06$ μm and truncation ratio $\gamma = 2$.

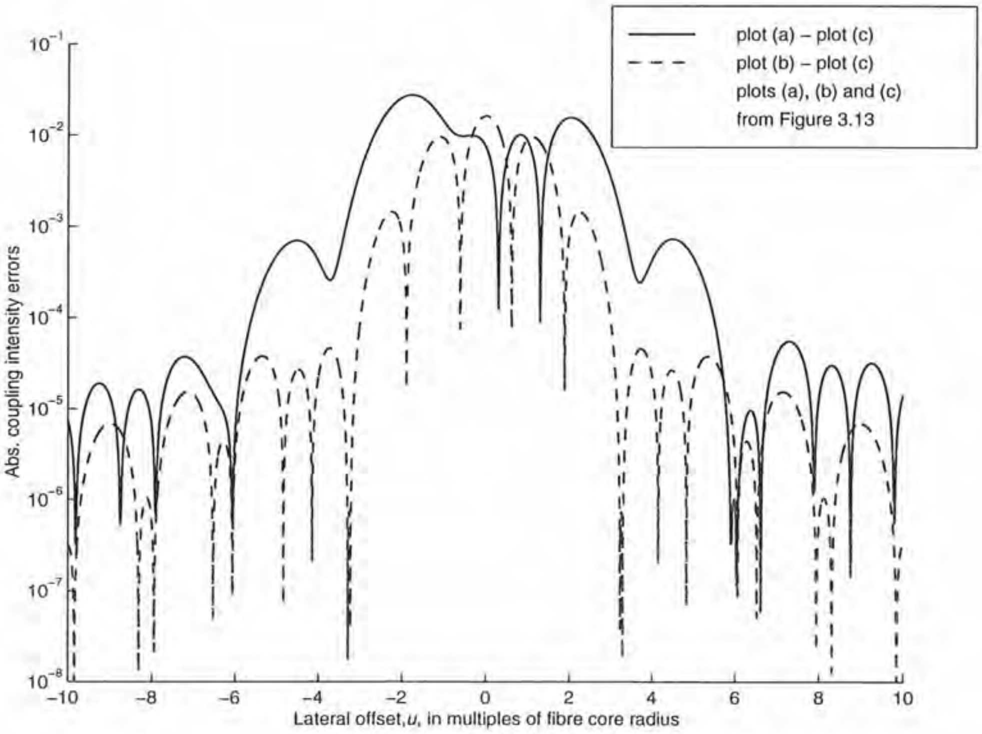


Figure 3.14: Coupling intensity errors using the asymptotic approximation expression as compared to numerical results.

The approximation fits the numerical results fairly well for $|u| > 6a$. In these regions, the errors are ~ 10^{-5} regardless of whether the fibre mode is described as a Bessel function or Gaussian distribution. Between $2a$ and $6a$ lateral offset, the Gaussian distribution underestimates the actual fibre field. Consequently, the coupling intensity given by the asymptotic approximation is lower. As γ increases, this underestimate becomes worse. Nevertheless, the theoretical coupling intensity for small offsets will be more accurate.

3.5.3 Coupling intensity for an output fibre with a 1-D angular tilt, θ_x

An additional phase factor of the fibre mode field at the *replay* plane correlates to a lateral shift at the *hologram* plane. Again, taking an analogy from digital and optical Fourier transforms, the lateral shift in the 'imaginary reverse' optical replay is related to the digital hologram plane grid by,

$$N_{\theta_x} \times \Delta_h = \frac{-f\theta_x}{a^2}, \tag{3.33}$$

where $N\theta_x$ is the number of grid points shifted for θ_x angular tilt. After normalisation so as to create the same co-ordinate axis as the *fft* of the input beams (see eq. (3.30)), N_{θ_x} becomes dependent on wavelength, λ, and the size of the input vector, D_i, such that,

$$N_{\theta_x} = \frac{-D_i\theta_x}{\lambda} \times \frac{f\lambda}{a^2}, \tag{3.34}$$

Hence, the coupling efficiency of the output fibre with an angular tilt can now be evaluated using numerical convolution of the *truncated* and the *shifted* fibre mode fields. The angular tilts of the output fibre are obtained from eq. (3.34). This simulation technique avoids the use of numerical integration of the overlapped fibre mode fields at the replay plane.

The tilted coupling intensity expression, eq. (3.28) is plotted as the dotted line in Figure 3.15. It almost overlaps all the points of the plot generated by the numerical convolution of the transformed fibre modes from the *input* and *replay* planes at the exit pupil of the hologram aperture. The numerical convolution results using the Bessel function and Gaussian fibre modes are shown by the solid and dashed lines respectively. The dash-dot line plots the corresponding expression by Marcuse, for the coupling efficiency of two fibres aligned with angular tilts. This expression has been scaled by $(erf(\gamma/\sqrt{2}))^2$ due to the truncation window. This is the coupling value when there is no tilt between the input and output fibres.

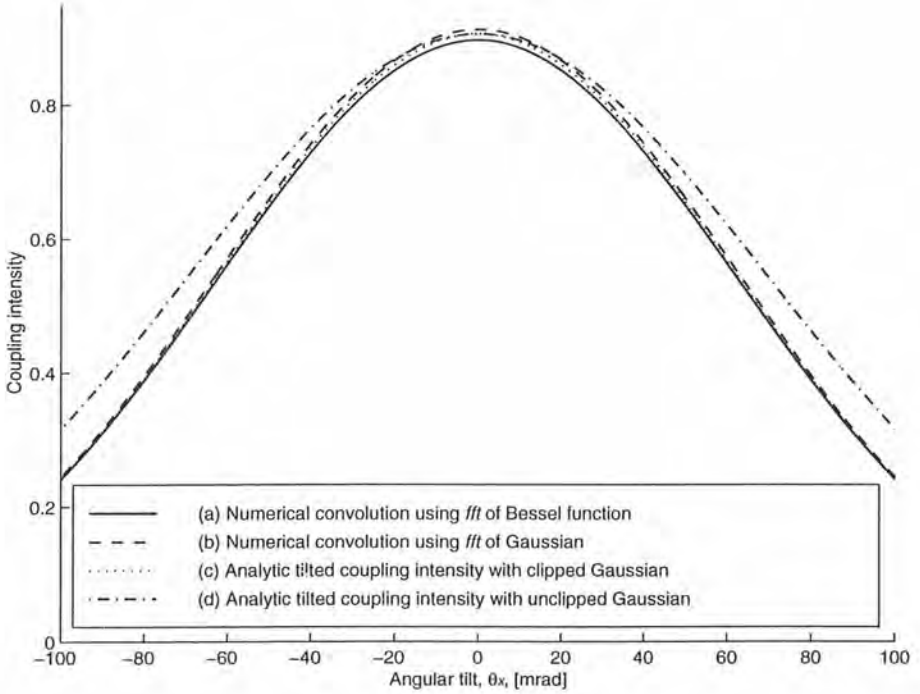


Figure 3.15: Coupling intensity for an output fibre with an angular tilt θ_x for Gaussian beam radius $w_r = 5.06$ μm and truncation ratio $\gamma = 2$.

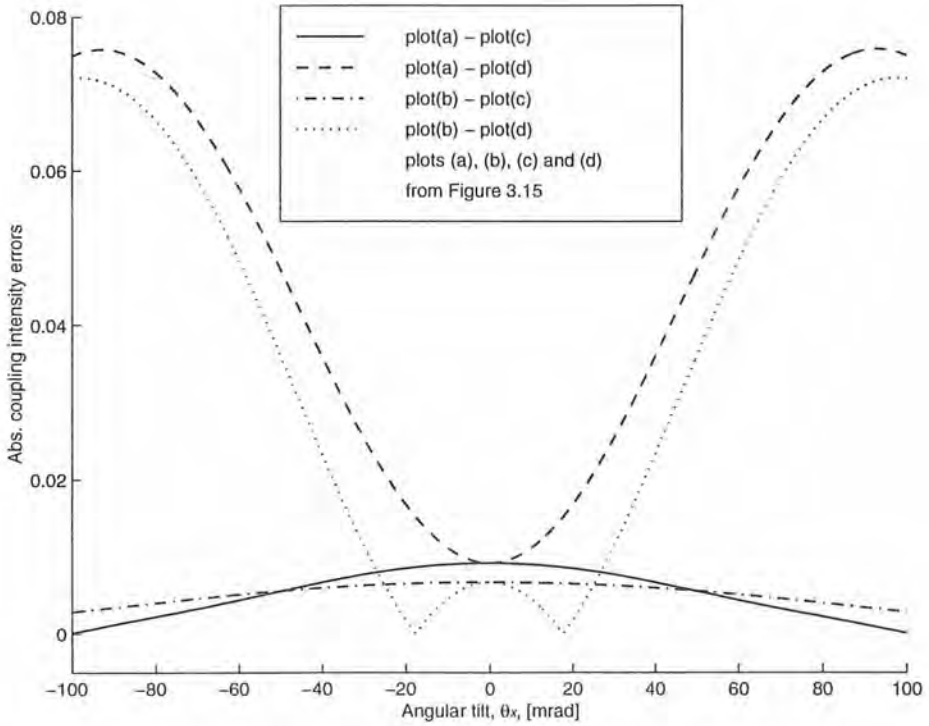


Figure 3.16: Coupling intensity errors using the *erf* and scaled Marcuse's expressions as compared to numerical results.

Figure 3.16 shows the coupling intensity errors due to the use of analytic expressions as compared to numerical convolution results for tilted coupling. The errors of using the *erf* and scaled Marcuse's expressions are plotted as solid and dashed lines for the Bessel function fibre mode. These corresponding errors are plotted as dash-dot and dotted lines for the Gaussian fibre mode. In contrast to the asymptotic approximation above, the exact *erf* expression for the tilted coupling intensity is only valid for small angles. Validity of the small angle approximation is restricted to ±0.1 radian. Within this range, the Gaussian fibre mode description agrees to within 1% of the values obtained by numerically evaluating the overlap integral using the Bessel function fibre mode description. The scaled version of Marcuse's expression has rather large errors for $\gamma = 2$.

3.6 Conclusions

The analytic expressions for power coupling with lateral offsets (eq. (3.22)) and angular tilts (eq. (3.28)) are analogous to those derived by Marcuse [20]. However these are for the case where the hologram illumination is truncated, with γ as the ratio of the truncation window and the Gaussian spot width at the *hologram* plane. For a 1:*N* holographic splitting stage, eq. (3.22) is applicable in calculating crosstalk power received by non-selected output fibre ports. The on-beam-axis coupling efficiency can be evaluated using the tilted coupling expression with the angular tilts being (0,0). For a single-stage *N*:*N* holographic space switch, the fan-in stage involves a tilted fibre launch. Hence, the coupling efficiency can be estimated by using eq. (3.28). However, attempting to evaluate crosstalk power with tilted wavefront using both equations will probably be erroneous. Unless the two expressions are used for small lateral shifts and angular tilts, the two results are generally not additive [72]. The truncation is normally asymmetric. This has a pronounced effect on the lateral-shift coupling intensity in that the troughs are no longer as sharp. Crosstalk power rises whereas on-beam-axis coupling efficiency falls.

The approximate replay and coupling equations are compared with numerical *fft* replay and coupling intensities using Bessel function and Gaussian descriptions of the fundamental fibre mode. It is shown that the approximate expression for coupling efficiency with a small lateral offset is not very accurate when the Bessel function fibre mode is used. However, the crosstalk values predicted by the expression have errors $\sim 10^{-5}$. In addition, the coupling efficiency for a small angular tilt is accurate to within 1% of the numerical results using the Bessel function description of the fundamental fibre mode. Numerical simulations using the Gaussian fibre mode return smaller replay and coupling intensity errors.

Having determined the accuracy of these analytic expressions, the replay and coupling intensity of a blank hologram at a small/large lateral offset or small angular tilt can be calculated with ease. From crosstalk considerations, there is a need to construct an output fibre array, which neatly avoids the placing of fibres along two transverse directions of an existing fibre port. The transverse axes are parallel to the edges of the square/rectangular hologram aperture. The crosstalk level in this case can be as low as -70 to -60 dB for a truncation ratio of 2. A full crosstalk analysis will need to take account of aberrations, scattering and Fresnel reflections.

CHAPTER 4

Theory of the replay of routing holograms written onto a programmable SLM

4.1 Introduction

In the context of optical reconstruction of thin holograms, phase and spatial quantisation effects determining the replay intensity and location of each replay order are treated separately. The phase quantisation effect is concerned with the number of phase levels available for recording holograms and the hologram patterns (or the distribution of phase elements within each hologram unit, or "base hologram"). An equation describing the intensity of all the peaks in an *aliased* digital replay is presented in §4.2.2. The locations of these replay peaks are to be found by using a skip-rotate rule, in order that the intensity expression is applicable. These intensities are exactly as given by a discrete Fourier transform (e.g., fast Fourier transform, *fft* algorithm [73]) of *any* hologram.

Recording the holograms on a dynamically reconfigurable SLM entails a further drop in replay intensity as a result of pixellation and dead-space as well as phase errors. The *sampling* of the optical field by the pixels having a finite size and a non-unity modulating to non-modulating area ratio is appended to the expression describing replay peak intensities. This optical replay can be simulated by over-sampling each encoding point of the digital hologram. Dynamic electro-optic modulation schemes including the fast planar-switching liquid crystal phase modulation are of interest here [74]. Utilising the birefringence effect for phase modulation will likely encounter a lack of switching angle. In fact, there are always phase errors for each desired phase level even for a fixed surface-relief diffractive element etched onto fused-quartz [75]. The expression is then modified to include the effect of *inadequate phase modulation* (*phase-mismatch*) for a subset of holograms where each phase level is present in equal measure.

This chapter is only concerned with the replay of holograms which are repeated infinitely such that the Fourier series is applicable. The replay peaks from such a "*numerical*" or "*digital*" hologram replay have been treated as Dirac's delta functions which work best for the numerical examples shown. The input beam profile and its clipping by the hologram aperture dictate the replay beam shape, as given in §3. This is most important for crosstalk consideration although it also gives additional insertion loss. Throughout this analysis, the paraxial optical reconstruction region (an area close to the optical axis at the replay plane where scalar diffraction theory is valid) is divided into subsections of $f\lambda/d$ in size centred at the optical axis, where f is the focal length of the transform lens, λ is the wavelength of illumination and d is the pixel pitch. These subsections of equal sizes are a result of the *pixel periodicity* of the pixellated SLM and are called "*replications*". Within each replication, the distribution of replay peaks or "*diffraction orders*" is identical except that their intensities are scaled by a *sinc* squared roll-off arising from the finite pixel width (i.e., pixel clear aperture or pixel transmittance function). The locations and relative intensities of these diffraction orders are entirely caused by phase quantisation and the distribution of these "*phase elements*" within a base hologram. The structures of the hologram and replay planes are as shown in Figure 4.1(a).

4.2 Phase quantisation and the distribution of quantised phase elements

Routing holograms are dynamically reconfigurable holograms that are used in holographic free-space routing applications. A single spot replay from these holograms is sufficient for the simplest free-space switch configuration. A simulated phase-only hologram (or computer-generated hologram, CGH) of a fixed number of phase levels can either be *noise* or *efficiency* optimised. An efficiency-optimised hologram can be obtained by the inverse *fft*. This produces only one solution of the phase distribution over a 1-D or 2-D hologram plane. All phase- or spatially-shifted variations of this distribution are eigen-sets of the original hologram. Noise-optimised holograms, such as those generated by Gerchberg-Saxton (GS) [76], direct binary search (DBS) [77], simulated annealing (SA) [78], error diffusion (ED) [79], etc., algorithms retain the fundamental characteristics of their efficiency-optimised counterpart. Hence, it is useful to analyse the replay of a general hologram based on the underlying spatial frequency properties of an efficiency-optimised type.

For a multi-level hologram (with m quantised "*phase states*" or "*phase levels*"), the n^{th}-order replay intensity, is often assumed to be given by Dammann's expression [17]:

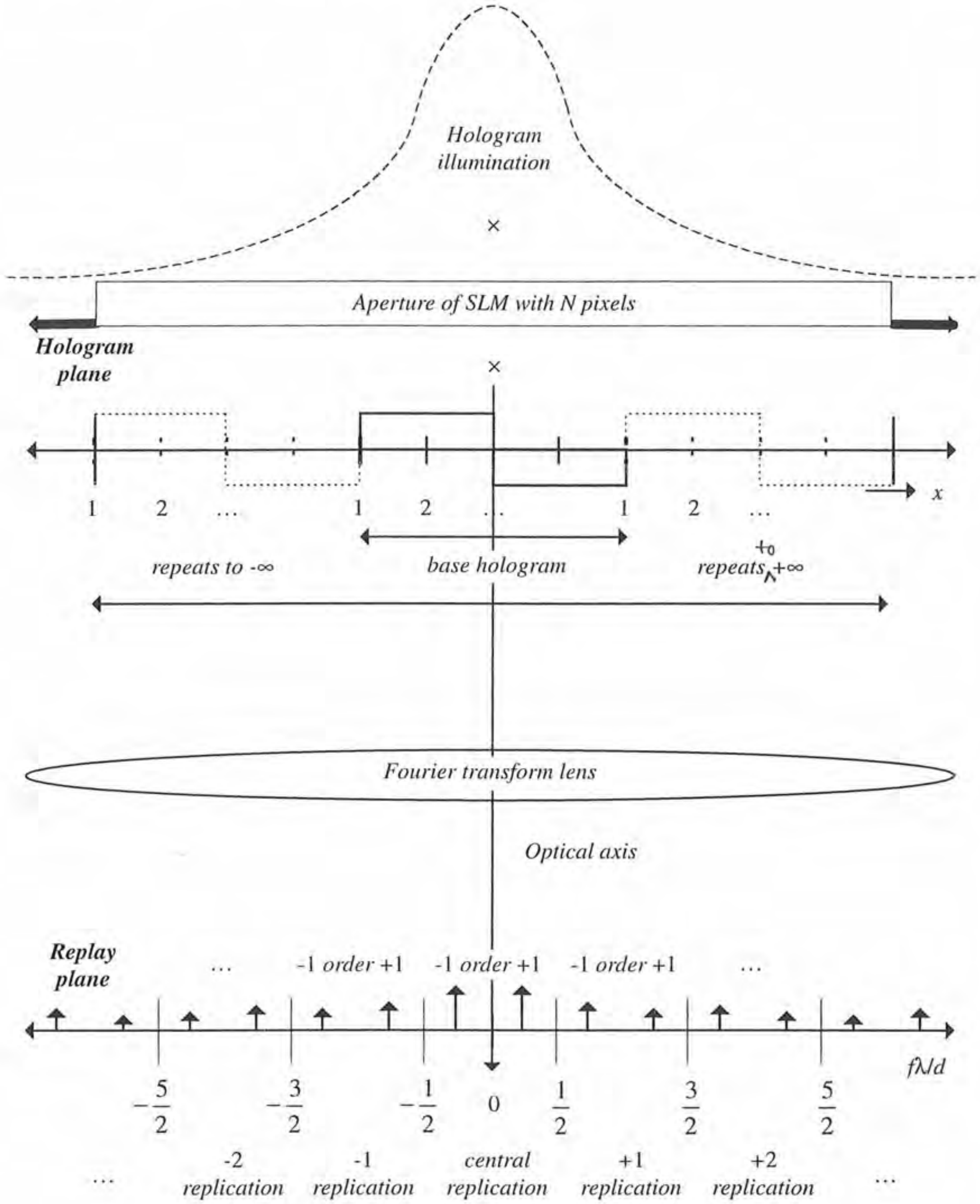


Figure 4.1(a): The division of SLM pixel array into a number of base hologram repeats and the grouping of diffraction orders into each replication at the replay plane.

$$\eta_{m;n} = sinc^2\left(\frac{n\pi}{m}\right), \tag{4.1}$$

where $sinc(x) = sin(x)/x$ and $n = gm + 1$ with g being an integer.

This expression is only valid for a 1-D *step-phase grating* written onto an SLM without dead-space which produces a single Dirac's delta-function main peak and its higher order replay peaks. The use of 1-D *step-phase grating* is too restrictive for routing holograms. Typically, a large number of holograms for dynamically reconfigurable routing applications are generated with a small number of phase levels. These are then written onto 2-D pixellated SLMs with a limited spatial bandwidth product (SBWP or equivalently number of pixels) and other processing limitations such as dead-space and phase uniformity of each modulating element.

4.2.1 Fractional representation for the replay of routing holograms

The numerical replay field of $N \times N$ grid points for an m-level hologram is a series of delta functions:

$$\hat{h}(x_d, y_d) = \sum_n \sum_{s=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} A_{m;n}\delta\Big(x_d - \big(\langle n\xi + sN\rangle\big), y_d - \big(\langle n\zeta + tN\rangle\big)\Big), \tag{4.2}$$

where (x_d, y_d) is the discrete co-ordinate at the simulated replay plane and (ξ,ζ) is the co-ordinate of the first order replay peak. $A_{m;n}$ is the simulated amplitude of the n^{th}-order replay given that the hologram has m phase levels. This follows directly from the property of the *fft* whereby an N-point hologram is transformed to an N-point replay. Depending on the hologram pattern one or more numerical replay points will have non-zero intensity. The non-zero intensity terms are related to "*modulo-N*" functions, $\langle n\xi + sN\rangle$ and $\langle n\zeta + tN\rangle$, such that:

$$-\frac{N}{2} \le \big(\langle n\xi + sN\rangle, \langle n\zeta + tN\rangle\big) \le \frac{N}{2} - 1, \tag{4.3}$$

$n = gm + 1$ and (s,t) can take any values, as long as the elements of $\big(\langle n\xi + sN\rangle, \langle n\zeta + tN\rangle\big)$ set are distinct. If (σ,τ) are now assigned as the fractional co-ordinate of the replay field, (i.e., $\sigma = \xi/N$, $\tau = \zeta/N$), criterion (4.3) becomes:

$$-\frac{1}{2} \le (\langle n\sigma + s\rangle, \langle n\tau + t\rangle) \le \frac{1}{2} - \frac{1}{N}, \tag{4.4}$$

and $\langle n\sigma + s\rangle$ and $\langle n\tau + t\rangle$ become "*modulo-1*" functions. The main peak of the numerical replay grid is located at (σ,τ) where $\sigma = x/x_o$ and $\tau = y/y_o$; x and y are integers and x_o and y_o are positive integers. It becomes apparent that a routing hologram of $N \times N$ size contains multiple "*base holograms*" routing to "(σ,τ) *replay fraction*". The (σ,τ) base hologram requires only $x_o \times y_o$ sample points for a complete representation of its form, its replay intensities and the diffraction orders that characterise it. See §5.3 for details of hologram generation.

Henceforth, the properties of a routing hologram are analysed by considering its base hologram having a (σ,τ) first order replay fraction. The optical replay of this hologram yields the first diffraction order at $(\sigma f\lambda/d, \tau f\lambda/d)$ co-ordinate from the optical axis. The fractional representation eliminates the need to know about the physical parameters of the holographic replay system without compromising any aspect of the hologram properties. It is assumed that x_o, y_o and m can be any positive odd or even integer and x_o and y_o may/may not be a multiple of m. The grating is a special case where $x = 1$ and $y = 1$ and x_o and y_o are multiples of m.

4.2.2 The replay intensities of grating holograms

In a numerical *fft* grating replay where only a single sampling point for each pixel state is used, an *aliased* numerical replay of the actual hologram with finite pixel width will always result, consistent with Nyquist's sampling criterion. For example, any integer multiples of binary $\{0\ \pi\}$ grating period will give a replay of 1 unit intensity at the first numerical replay point, corresponding to -1/2 normalised spatial frequency as illustrated in Figure 4.1. Due to the use of two sample points per grating period, the highest frequency component that can be replayed by the *fft* is that of the fundamental frequency. If the fractional co-ordinate of the first order peak is now -1/8 with the use of binary $\{0\ 0\ 0\ 0\ \pi\ \pi\ \pi\ \pi\}$ grating period, the four times over-sampling gives up to four times the fundamental frequency being replayed in the same output replay resulting in 4 peaks appearing at ±1 and ±3 orders. These two grating functions and their replays as well as that of $\sigma = \pm 1/4$ and $\pm 1/16$ are shown in Figure 4.1.

Figure 4.1 depicts four grating functions and their replays in which Dammann's efficiency expression ought to be applicable. However, the higher orders exceeding ±1/2 of a simulated hologram replay should not be thought to be 'spilling over' to the higher order replay replications as each pixel is only sampled by a single point in the numerical *fft* and no ±

spatial frequencies < -1/2 or ≥1/2 are possible by the numerical replay. Instead, these higher orders overflowing the frequency co-ordinate at one end must be folded back into the same numerical replay grid through the opposite end. It is possible that these contributions from the higher order replications overlap exactly with the main peak or other higher order noise peaks produced in the central replication.

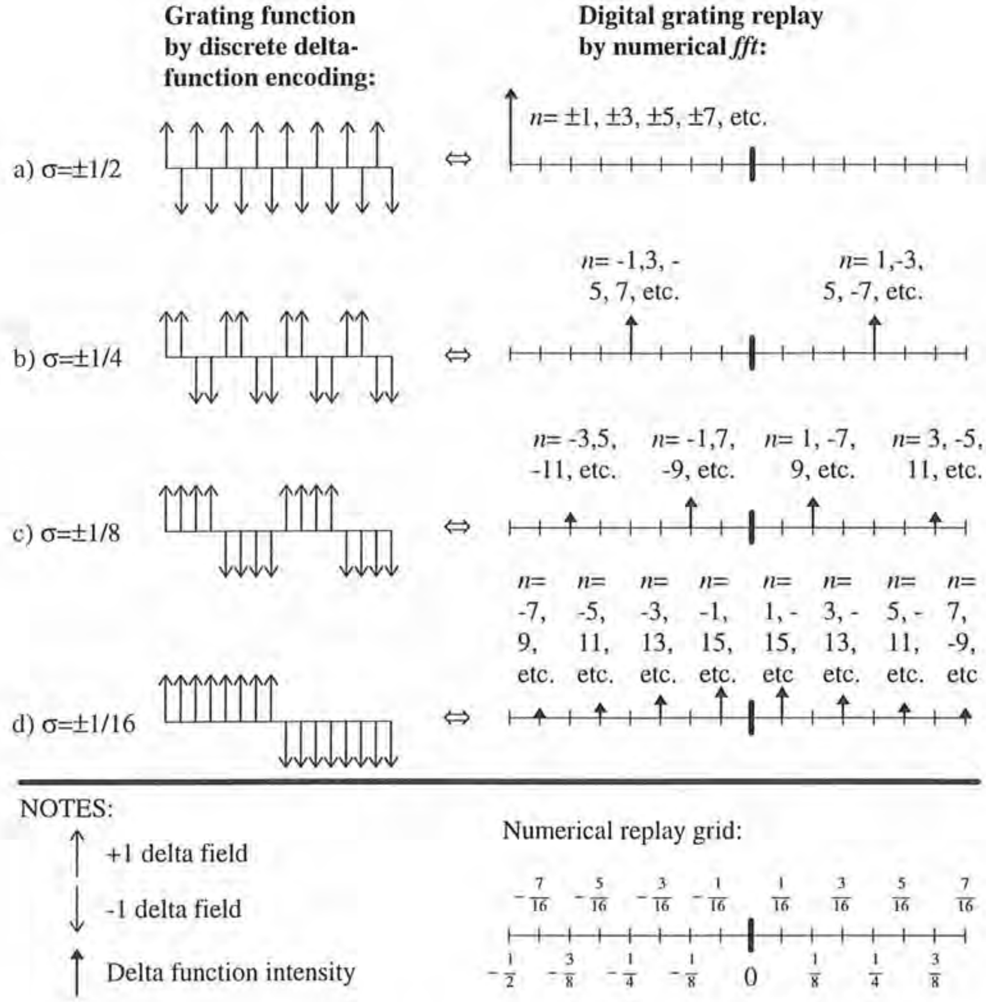


Figure 4.1: Higher orders overlap in the numerical replay grid of binary gratings.

For a general grating period, x_o ($\sigma = 1/x_o$, x_o = a multiple of m), the peaks that appear within the numerical replay grid, due to phase quantisation and the distribution of m phase-states to create a grating of x_o phase-elements, have an aggregated intensity:

$$\eta^{\sigma}_{m;n} = \sum_{g=-\infty}^{\infty} sinc^2\left(\frac{n\pi}{m} + g\frac{x_o\pi}{m}\right), \tag{4.5}$$

where n denotes the replay order and g is an integer. The summation gives the intensity due to phase quantisation, $\eta^{\sigma}_{m;n}$:

$$\eta^{\sigma}_{m;n} = sinc^2\left(\frac{n\pi}{m}\right)\sum_{g=-\infty}^{\infty}\left(\frac{n}{n+gx_o}\right)^2 . \tag{4.6}$$

Utilising the residue method, the convergent series sums to

$$\eta^{\sigma}_{m;n} = \frac{sinc^2\left(\frac{n\pi}{m}\right)}{sinc^2\left(\frac{n\pi}{x_o}\right)} . \tag{4.7}$$

Eq. (4.7) gives an accurate description of the intensity of each frequency component present in the numerical replay grid. In the actual grating encoding using rectangular/square aperture pixels and optical grating replay by the use of a transform lens, each harmonic frequency component will be replayed with an appropriate scaling by the transform of the pixel aperture function. If the pixel aperture is rectangular/square without dead-space, the reverse process of the summation in eq. (4.5) takes place. The spatial quantisation term of a grating, which is considered as an infinite repetition of an m-level phase ramp, effectively cancels out the denominator of eq. (4.7) to yield Dammann's expression. Decoupling the phase quantisation term from the overall efficiency expression is important because many 2-D hologram patterns in addition to 1-D gratings are often required; each has to be generated using a limited number of phase levels. Besides, pixel dead-space results in an intensity scaling term due to the spatial quantisation (see §4.3), which does not cancel out the denominator of eq. (4.7).

For $\sigma = \pm 1/2$, $\pm 1/4$, $\pm 1/8$ and $\pm 1/16$ gratings shown in Figure 4.1, the first order intensities are:

$$\begin{aligned} \eta^{\sigma=1/2}_{m=2;n=\pm1} &= \frac{sinc^2\left(\frac{\pi}{2}\right)}{sinc^2\left(\frac{\pi}{2}\right)} = 100\% \quad , \quad \eta^{\sigma=1/8}_{m=2;n=\pm1} = \frac{sinc^2\left(\frac{\pi}{2}\right)}{sinc^2\left(\frac{\pi}{8}\right)} = 42.68\% \\ \eta^{\sigma=1/4}_{m=2;n=\pm1} &= \frac{sinc^2\left(\frac{\pi}{2}\right)}{sinc^2\left(\frac{\pi}{4}\right)} = 50\% \text{ and } \eta^{\sigma=1/16}_{m=2;n=\pm1} = \frac{sinc^2\left(\frac{\pi}{24}\right)}{sinc^2\left(\frac{\pi}{16}\right)} = 41.05\% \end{aligned} , \tag{4.8}$$

respectively. The number of peaks appearing within the N numerical replay points is x_o/m, bearing in mind that this has a minimum of 1 regardless of m values. As an example, optimisation of a four-level phase hologram to route to -1/2 fractional co-ordinate will always produce a binary grating.

Although x_o is 2 and m is 4 for this case, it should really be considered as $m' = 2$, because that is the underlying hologram property for this routing. For a grating which consists of as many encoding points as there are phase levels (e.g. $\sigma = 1/4$, $m = 4$; $\sigma = 1/5, m = 5$, etc.), its numerical replay will always contain a single peak of 100% which means phase quantisation produces 100% efficiency ($x_o = m$ in eq. (4.7)).

Extending the analysis to 2-D gratings, the least common multiple of x_o and y_o is used in eq. (4.7), to enhance the intensity of each overlapped location within the numerical replay grid. The equivalent expression relating to phase quantisation for a 2-D grating is:

$$\eta_{m:n}^{\sigma,\tau} = \frac{sinc^2\left(\frac{n\pi}{m}\right)}{sinc^2\left(\frac{n\pi}{x_o y_o}\right)}, \tag{4.9}$$

where $x_o y_o$ takes the least common multiple (*lcm*) of x_o and y_o, or $x_o y_o = lcm(x_o, y_o)$.

There are $x_o y_o/m$ peaks within $x_o \times y_o$ points in the numerical replay grid. In addition, the location of each replay order can be predicted, by 'folding' higher orders with fractional locations $< -1/2$ or $\geq 1/2$ back to the numerical replay grid. The numerical replay of a routing hologram (grating) can be represented mathematically by (for *even* $x_o y_o$),

$$\hat{h}(x_r, y_r) = \sum_{n=-\frac{x_o y_o}{2}+1}^{n=\frac{x_o y_o}{2}-m+1} \sum_{s\to-\infty}^{\to\infty} \sum_{t\to-\infty}^{\to\infty} \frac{sinc\left(\frac{n\pi}{m}\right)}{sinc\left(\frac{n\pi}{x_o y_o}\right)} \delta\left(x_r - \left(\langle s + n\sigma \rangle\right), y_r - \left(\langle t + n\tau \rangle\right)\right), \tag{4.10}$$

where (s,t) can, but does not necessarily take all values $\to \pm\infty$, and their largest values depend on x_o, y_o and the sign of the fractional co-ordinate of the main peak; (x_r, y_r) is the fractional co-ordinate of the replay replication; $n = gm + 1$, g being an integer. It turns out that $s = 0$ and $t = 0$ for the numerical replay of all grating patterns. It is important that 2-D Fourier (continually-shifted) phase-only holograms (gratings) rather than crossed holograms (gratings) are being replayed. The crossed phase-only holograms (gratings) can be considered to be producing independent routings in x- and y-directions giving $(x_o/m) \times (y_o/m)$ number of peaks appearing in the numerical replay grid each with an intensity of:

$$\eta_{m:n}^{\sigma,\tau} = \frac{sinc^2\left(\frac{n\pi}{m}\right)}{sinc^2\left(\frac{n\pi}{x_o}\right)} \frac{sinc^2\left(\frac{n\pi}{m}\right)}{sinc^2\left(\frac{n\pi}{y_o}\right)}. \tag{4.11}$$

4.2.3 The replay locations of general holograms

In the above, a method which considers the phase quantisation effect separately from spatial quantisation on the intensity replay peaks has been shown to be consistent with Dammann's stepped-phase grating analysis. In addition to the dead-space consideration, the principle is most useful for the replay of a general phase-only hologram (i.e. $\sigma = x/x_o$ and $\tau = y/y_o$) where x_o and y_o are not necessarily a multiple of m and x and y are not necessarily 1. At the digital stage, the numerical replay field of a 1-D general phase-only hologram of $\sigma = x/x_o$ simple-fraction co-ordinate contains the same fundamental properties of a $1/x_o$ phase-only hologram. The relative intensities of all the replay peaks are shifted according to which harmonic of the basic $1/x_o$ hologram is excited. Identifying the location of each order (or equivalently the order number at each non-zero locations) enables the intensities to be calculated by eq. (4.7).

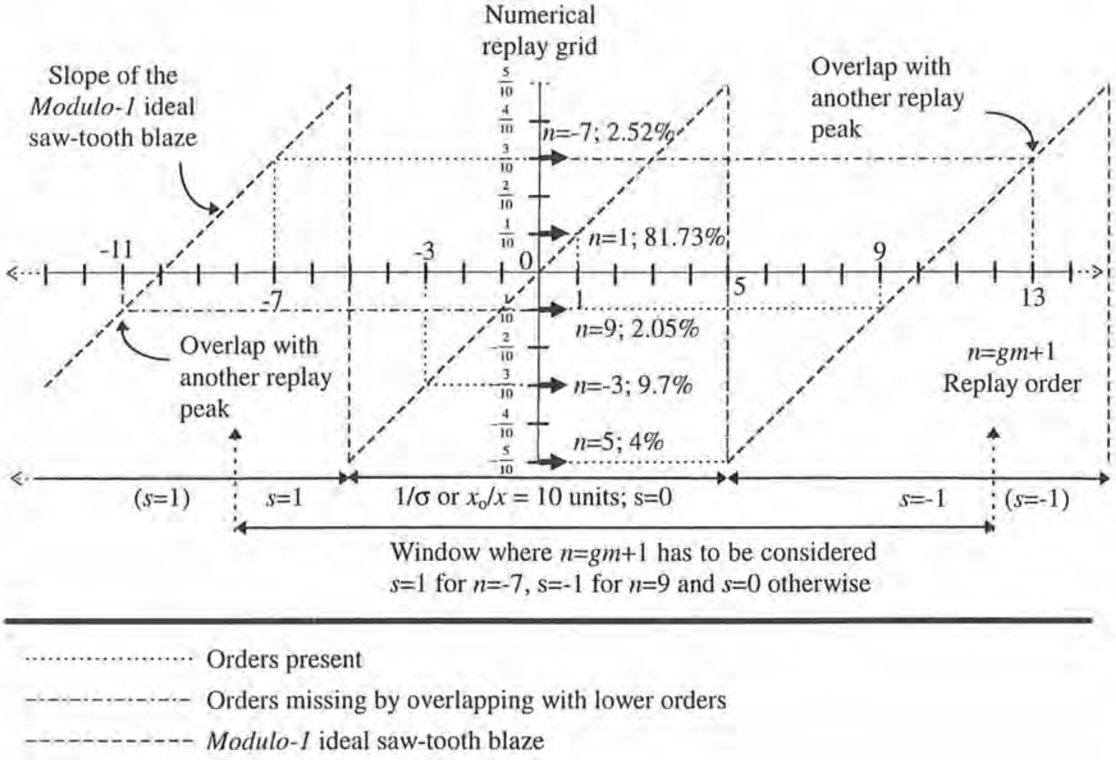


Figure 4.2: *Modulo-1* skip-rotate rule used to locate higher orders of $\sigma = 1/10$ quaternary hologram replay fraction.

The numerical replay field of a 1-D phase-only hologram is given by a series of delta functions:

$$\hat{h}(x_r) = \sum_{n} \sum_{s \to -\infty}^{\to \infty} \frac{sinc\left(\frac{n\pi}{m}\right)}{sinc\left(\frac{n\pi}{lcm(m,x_o)}\right)} \delta\left(x_r - \left(\langle s + n\sigma \rangle\right)\right). \tag{4.12}$$

The *modulo-1* skip-rotate rule, in order to determine the 1-D numerical replay according to eq. (4.12), is illustrated in Figure 4.2 for σ = 1/10 and $m = 4$. It has been assumed that phase-matching condition is satisfied (i.e. there is $2\pi(m\text{-}1)/m$ phase depth between the lowest and the highest phase levels, see §4.4). With this routing example, m and x_o are both even numbers but x_o is not a multiple of m. The key to obtain the correct number of n orders appearing in the numerical replay field is to draw an ideal saw-tooth blaze along the n axis passing through $n = 0$ and the centre of the numerical replay grid with a slope of x. With these conditions, each period of the ideal saw-tooth blaze is drawn over the length of 1/σ along the n-axis.

The peaks that do appear in the numerical replay grid have decreasing intensities in accordance with Dammann's criterion $n = gm + 1$ where $g = 0,-1,+1,-2,+2,-3,+3\ldots$. Each integer g is taken successively until a high order begins to overlap with one of the existing replay peaks (i.e., only the lowest $lcm(m,x_o)/m$ peaks at ± frequencies need to be considered). At this point all replay orders have been located and the aggregated intensity of each peak as a result of the *modulo-1* skip-rotate rule is given by eq. (4.7) with x_o replaced by $lcm(m,x_o)$. That is to say the overlap of a higher order peak with an existing replay peak occurs only after every $lcm(m,x_o)$ order separation. It may seem an irony that a higher order with a large n value can be located at a shorter distance from the centre of replication than a higher order of a smaller n value. The 5 peaks shown in Figure 4.2 are actually $n = \{-5,-3,-1,1,3\}$. However, by assuming $n = \{5,-3,9,1,-7\}$, their intensities are correctly predicted using eq. (4.7). It also gives the intensity of an n order obtained by this consideration diminishes as $|n|$ increases.

Another example of the use of the skip-rotate rule is shown in Figure 4.3 where σ = 1/5. Noting that a higher order peak is located at the centre of the replay replication. This is not surprising since the hologram does not average out to 0 DC. In fact, for an even m and a simple routing fraction with an odd denominator x_o, there is always a higher order peak at 0 spatial frequency point so that the rest of x_o-1 peaks can be evenly distributed across the positive/negative halves of the numerical replay grid. A summary of the number of peaks obtained from a numerical replay of a general 1-D Fourier phase-only hologram is shown in Table 4.1.

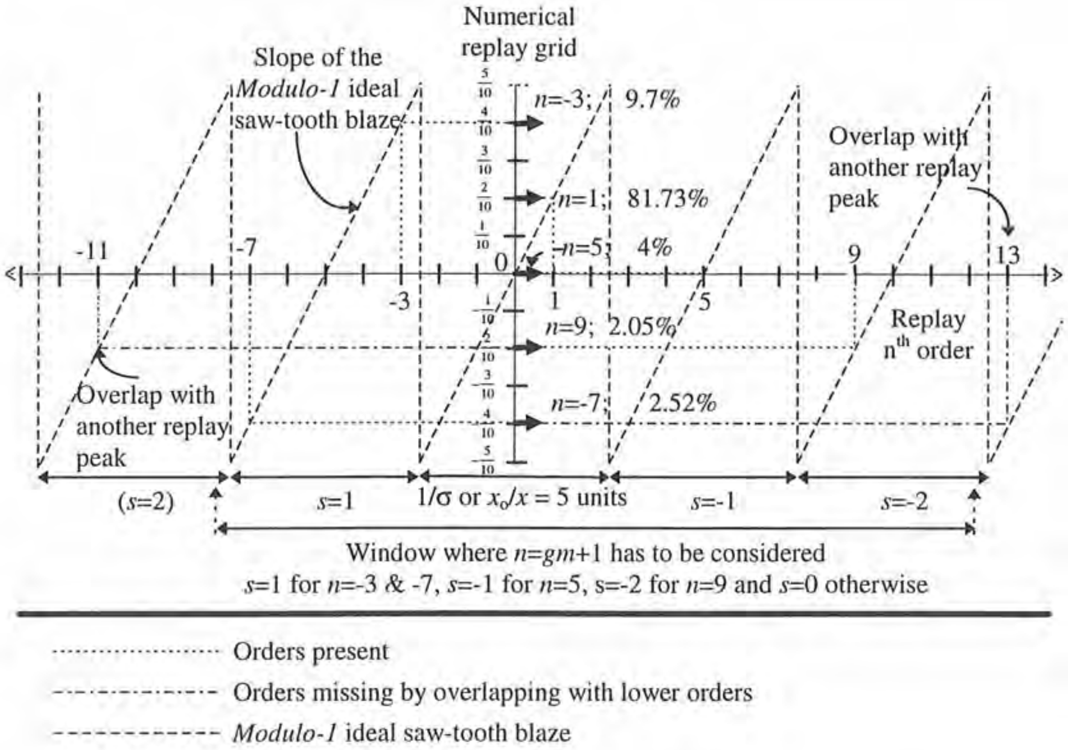


Figure 4.3: *Modulo-1* skip-rotate rule used to locate higher orders of $\sigma = 1/5$ quaternary hologram replay fraction.

Phase-matched hologram, $\phi = 2\pi(m-1)/m$ (see §4.4) replay fraction $(\sigma,\tau) = (x/x_o, y/y_o)$ and $x_oy_o = lcm(x_o,y_o)$		
$x_oy_o =$ multiples of m	$x_oy_o \neq$ multiples of m	$x_oy_o \neq$ multiples of m
$lcm(m, x_oy_o) = x_oy_o$ no. of peaks = x_oy_o/m $n = \Sigma\,(gm+1)$, $g = 0, -1, +1, -2, +2 \ldots$ etc. for the first x_oy_o/m peaks. $\{m, x_oy_o\} \in \{\{\text{even,even}\}, \{\text{odd,even}\}, \{\text{odd,odd}\}\}$	$lcm(m, x_oy_o) = m\, x_oy_o$ no. of peaks = x_oy_o $n = \Sigma\,(gm+1))$ $\{m, x_oy_o\} \in \{\{\text{even,odd}\}, \{\text{odd,even}\}, \{\text{odd,odd}\}\}$	$lcm(m, x_oy_o) = m\, x_oy_o/2$ no. of peaks = $x_oy_o/2$ $n = \Sigma\,(gm+1)$ $\{m, x_oy_o\} \in \{\text{even,even}\}$

Table 4.1: Summary of the number of peaks appearing in the numerical replay field for any combination of m, x_o and y_o. For a 1-D hologram, $y_o = 1$ and $x_oy_o = x_o$.

4.3 The effect of pixellation and dead-space (spatial quantisation)

Dynamically reconfigurable routing holograms are written onto a pixellated SLM. The SLM contains a clear aperture and some dead space. A finite pixel aperture introduces sampling effects and the replay field consists of multiple replay replications. Provided that the transmission characteristic of the dead-space is constant for all the pixels, the fraction of light routed by dead-space will not change the replay pattern of the numerical hologram replay. The pixel pitch has been normalised to 1 and its geometry is taken to be a square. Again this is to free up the dependence on the physical parameters as was in §4.2. The ratio of clear aperture to pixel pitch is ρ in both x and y directions, as shown in Figure 4.4; α and β are the amplitudes of complex transmittance in the in the pixel and dead-space, respectively.

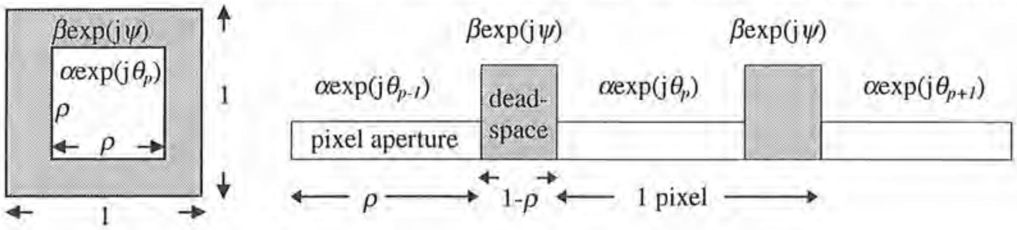


Figure 4.4: Pixel dimensions, transmittance and a 1-D cross-section of pixels

The hologram written onto an SLM is the convolution of the pixel clear aperture, $rect(x_h/\rho, y_h/\rho)$ and the calculated transmittance at each point. The calculated transmittance function to fill the N-point SLM dimension in 1-D contains a number of "*hologram repeats*" (i.e. repetition of a base hologram). The 1-D case is shown in Figure 4.5.

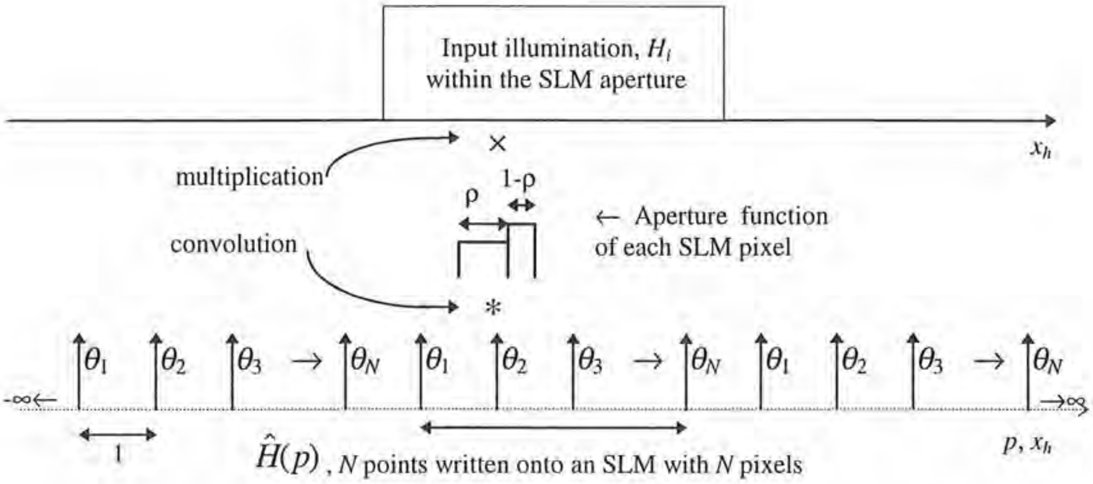


Figure 4.5: Real holograms depicted by the 1-D convolution of each calculated hologram point with the pixel transmittance and multiplied by the finite hologram illumination.

The hologram function is then multiplied by the apertured profile of the hologram illumination which is the Fourier transform of the beam profile at the input plane of the *4f* coherent optical system and hence the notation $H_i(x_h,y_h)$. Mathematically, the 2-D case is represented as the following:

$$\begin{aligned} H(x_h, y_h) = \Bigg\{ & \left[\sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \delta(x_h - p, y_h - q) \times \hat{H}(p,q) \right] * \alpha \, rect\left(\frac{x_h}{\rho}, \frac{y_h}{\rho} \right) + \\ & \left[\sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \delta(x_h - p, y_h - q) \beta \exp(j\psi) \right] * \left[1 - rect\left(\frac{x_h}{\rho}, \frac{y_h}{\rho} \right) \right] \Bigg\} \times H_i(x_h, y_h) \end{aligned} . \quad (4.13)$$

In eq. (4.13), (p,q) denotes the summation variables of the sampling 'pulse train' and $*$ represents the convolution operator. If the hologram illumination is a plane-wave and the hologram aperture is very large, the hologram replay can be assumed to consist of many delta functions, located within each set of (p,q) replay replication.

The validity of this treatment breaks down as soon as scalar wave propagation in the Fraunhofer regime fails. Within each replay replication system, there are higher order replay peaks which can be located by applying the *modulo-1* skip-rotate rule. The hologram illumination is typically a fibre mode in free-space, truncated by the finite expanse of the hologram aperture. This was considered in the previous chapter, §3. The replay field is the optical Fourier transform of eq. (4.13) to yield,

$$\begin{aligned} h(x_r, y_r) = \Bigg\{ & \sum_{p,q} \sum_{s,t} \sum_{n} A_{m;n} \, \alpha \, \rho^{2} sinc\big(\rho\pi(p + \langle s + n\sigma \rangle)\big) sinc\big(\rho\pi(q + \langle t + n\tau \rangle)\big) + \\ & \beta\left(1 - \rho^2\right) \exp(j\psi) - \beta \exp(j\psi) {\textstyle\sum_p'} {\textstyle\sum_q'} \rho^2 sinc(\rho p\pi) sinc(\rho q\pi) \Bigg\} * h_i(x_r, y_r) \end{aligned} . \quad (4.14)$$

$|A_{m;n}|^2$ is equivalent to $\eta_{m;n}^{\sigma,\tau}$; the primes on the second summation indicate that p and q can take any integer values, but not zero simultaneously. Assuming that the transform of input profile is a delta function, further simplification yields,

i) replay intensity of non-zero orders:

$$\left| h(x_r, y_r)_{m;n\neq0} \right|^2 = \sum_{n} \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \sum_{s\to-\infty}^{\to\infty} \sum_{t\to-\infty}^{\to\infty} \eta_{m;p;n\neq0}^{\sigma,\tau;p,q} , \quad (4.15)$$

where $\eta_{m:\rho:n\neq 0}^{\sigma,\tau;p,q} = \eta_{m:n\neq 0}^{\sigma,\tau} \times \eta_{\rho:n\neq 0}^{\sigma,\tau;p,q}$ and

$$\eta_{\rho:n\neq 0}^{\sigma,\tau;p,q} = \alpha^2 \rho^4 sinc^2\big(\rho\pi(p+\langle s+n\sigma\rangle)\big) sinc^2\big(\rho\pi(q+\langle t+n\tau\rangle)\big), \tag{4.16}$$

and $\eta_{m:n}^{\sigma,\tau} = \dfrac{sinc^2\left(\frac{n\pi}{m}\right)}{sinc^2\left(\frac{n\pi}{lcm(m,x_0,y_0)}\right)}$ as given by eq. (4.9) and,

ii) intensity of zero orders in all replay replications:

$$\left|h(x_r,y_r)_{m;n=0}\right|^2 = \left|\beta\left(1-\rho^2\right)\exp(j\psi)\right|^2 + \textstyle\sum_p^{'}\sum_q^{'}\left|-\beta\exp(j\psi)\right|^2 \rho^4 sinc^2(\rho p\pi) sinc^2(\rho q\pi). \tag{4.17}$$

It has been assumed that there is no DC-average intensity from pixels areas (i.e. phase-matching applies and each phase state is present with equal measure). If there is DC from the replay of a phase-matched hologram, the zero order term due to dead-space has to be added with the amplitude contribution from pixel areas, forming a grating which routes to 0 and integer multiples of *f*λ*/d*-sized replay replications. Eqs. (4.15), (4.16) and (4.17) take into consideration the *spatial* (pixellation and dead-space) and *phase quantisation* effects for the case of constant transmittance dead-space. A simple interpretation of eqs. (4.15), (4.16) and (4.17) is that the intensity of any replay peak appearing in the numerical replay grid will be subjected to a *sinc* squared scaling resulting in peaks appearing in higher order replications. A unique property of *sinc* squared that diffracts light to higher order replications is:

$$\sum_{p=-\infty}^{\infty} \rho^2 sinc^2\big(\rho\pi(p+x)\big) = \rho, \tag{4.18}$$

where $0 \leq \rho \leq 1$ and $|x| \leq 1/2$. The sum of all replay peak intensities corresponds to the SLM fill-factor provided that the condition for $|x|$ in eq. (4.4) is satisfied. Hence, knowing the intensity and location of the main replay peak of a simulated hologram is sufficient to predict the whole discrete hologram replay intensity values and locations by applying eq. (4.16).

4.3.1 Numerical simulations of 1-D hologram replay

In a system implementation, only the diffraction orders that fall within the central replication are of any use. The replay intensities within the central replication can be described by the expression below,

$$\eta_{m;p;n\neq 0}^{\sigma;\tau;p=0,q=0} = \frac{sinc^2\left(\frac{n\pi}{m}\right)}{sinc^2\left(\frac{n\pi}{lcm(m,x_o,y_o)}\right)} \sum_{s,t} \alpha^2 \rho^4 sinc^2\left(\rho\pi(\langle s+n\sigma\rangle)\right) sinc^2\left(\rho\pi(\langle t+n\tau\rangle)\right). \qquad (4.19)$$

Notice that the central first order intensity of a real hologram replay will be lower than Dammann's expression predicts. Dammann's expression includes both phase and spatial quantisation and it is valid as long as a 1-D grating ($\sigma = 1/x_o$) is written on an SLM without any pixel dead-space.

This "*composite sinc squared*" scaling for various replay orders across all replay replications can be shown to give accurate intensity values by numerical simulations. The 1-D version of eq. (4.19) is applied to the replay of a phase-matched (ϕ = 270°, see §4.4) four-level (m = 4) phase hologram with its first order peak of $\sigma = 3/8$. The pixel transmittance is assumed unity; the dead-space is opaque and comprises ~ 10% of a pixel. Each point of the calculated base hologram is represented by 32 samples with 3 samples made zero transmittance. There exists a separate *composite sinc squared* scaling for each replay order n, as shown in Figure 4·6. 'o' marks the replay peaks within the central replication and 'x' is the intensity of the central zero order.

Due to pixel periodicity, the replay of a hologram is replicated. This is then scaled by the single pixel aperture term, as shown in Figure 4.7. The delta plots are obtained by replacing n in eq. (4.7) with $n + g \times lcm(m,x_o)$. Note the remarkable similarity with the diffraction of an N-slit diffraction grating [80] when the hologram phase-quantisation is treated separately.

In Figure 4.6, the first order peak intensity has been reduced to 47.22% from the phase-quantised value of 85.36% due to spatial quantisation. However, the third order peak intensity at -1/8 is only reduced slightly, from 14.64% to 11.53%. In addition, the first order peak belonging to -1 replication system is very pronounced at 21.22%. This, and all other peaks at multiples of integers from the first order peak in the central replication originate from the phase-quantised-only first order replay peak. The fill-factor ratio has an effect on how the peaks located at some distance away from the central zero order are scaled by spatial quantisation as depicted in Figure 4.8 for three fill-factors: ρ = 0.8, 0.9 and 1. The scaling for replay peaks located far-away from the optical axis (zero order in the central replication) is very severe. Therefore, it is advantageous to choose replay locations close to the central zero order, e.g., within ± 0.15 fractional co-ordinate.

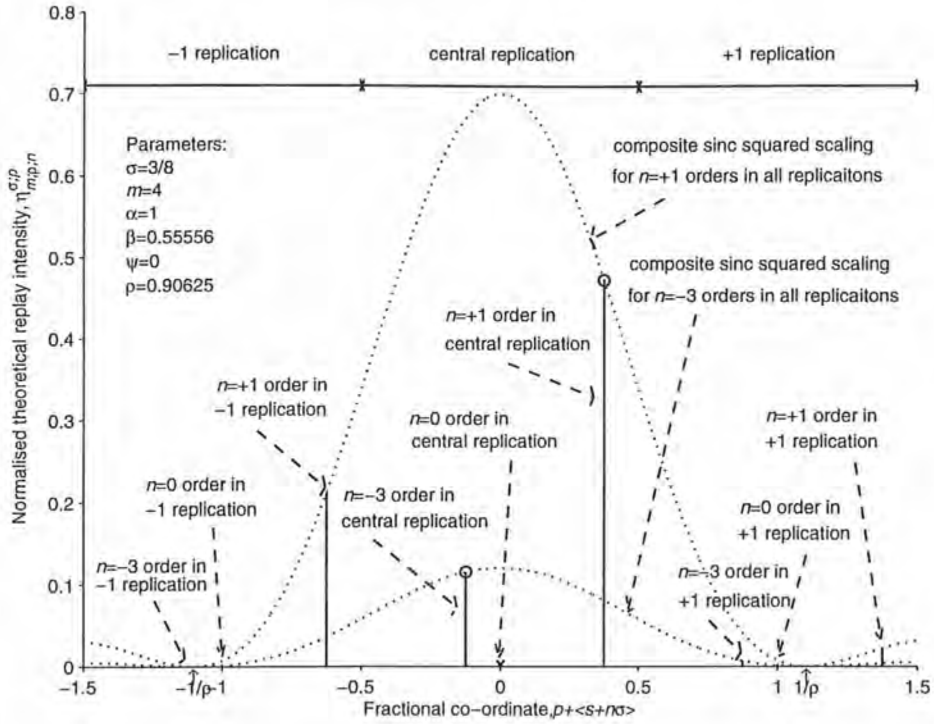


Figure 4.6: Composite effects of spatial and phase quantisation on the intensity of a phase-matched 3/8 quaternary replay fraction.

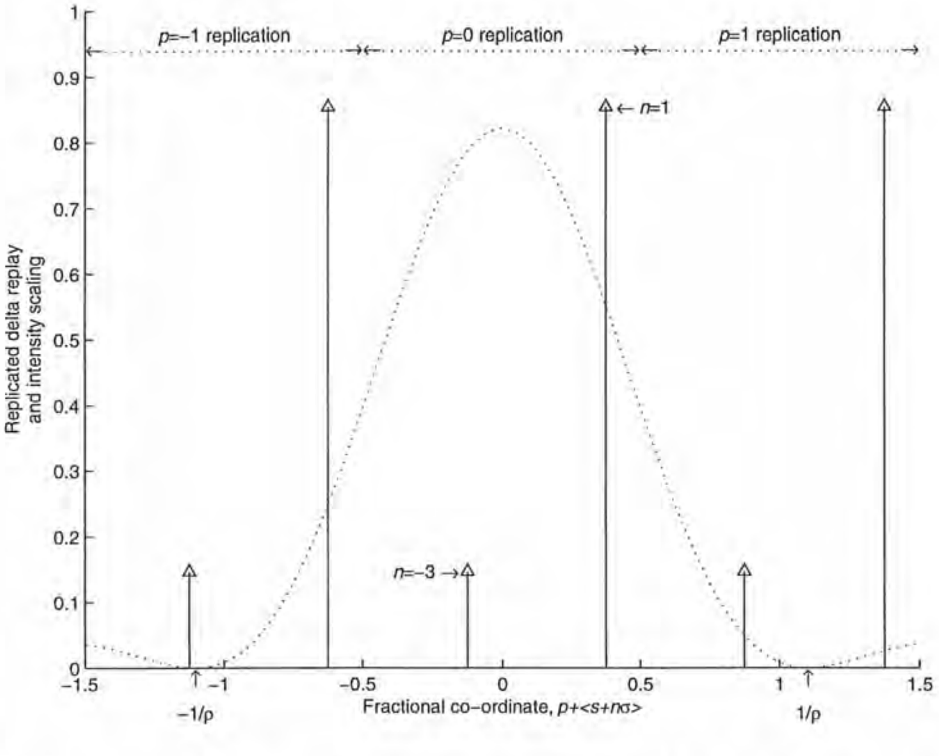


Figure 4.7: Separating the hologram term (delta function plot) and the single pixel aperture term (dotted line).

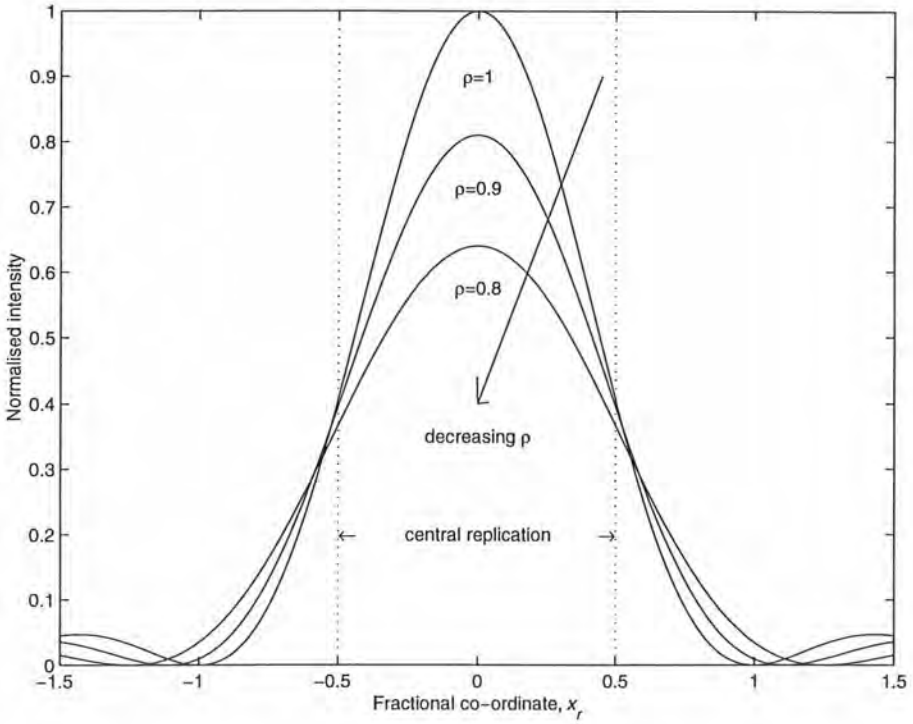


Figure 4.8: *Sinc* squared scaling due to 1-D spatial quantisation.

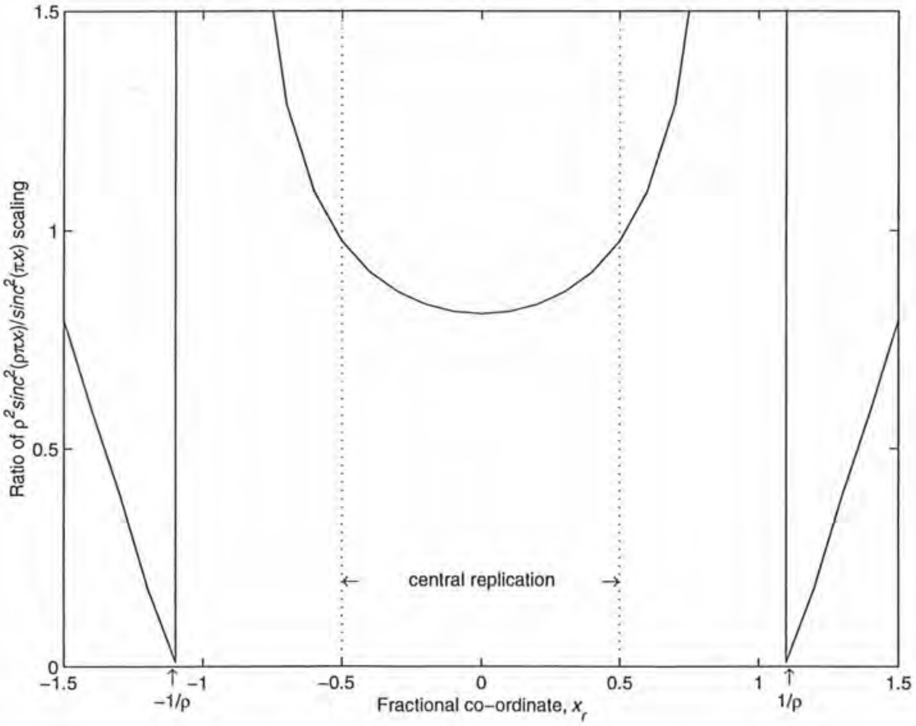


Figure 4.9: The ratio of *sinc* squared scaling for ρ = 0.9 and 1.0.

When the fill-factor is reduced due to the presence of dead-space, the central region (where the spatial quantisation effect has a near uniform scaling) becomes larger, at the expense of reduced replay intensity. The difference in pixellation scaling for the case with and without dead-space is shown in Figure 4.9. For an SLM without dead-space ($\rho = 1$) the replicated replay peaks in ± (0.5→1.0) normalised replay co-ordinate are much reduced compared to those in the central replication. When the fill-factor becomes non-unity, the replay peaks in these regions are boosted.

4.3.2 The upper-bound of the replay efficiency of routing holograms

The most important characteristic of a discrete hologram replay is the predictable replay intensity of each (σ,τ) base hologram having taken into account phase and spatial quantisation effects. Each base hologram produces a first order peak at (σ,τ) fractional co-ordinate of the central replication. The 2-D expression predicting the first order ($n = 1$) replay intensity of an m- level (σ,τ) base hologram is

$$\eta_{m;\rho;n=1}^{\sigma,\tau;p=0,q=0} = \frac{sinc^2\left(\frac{\pi}{m}\right)}{sinc^2\left(\frac{\pi}{lcm(m,x_0,y_0)}\right)}\alpha^2\rho^4 sinc^2(\rho\sigma\pi)sinc^2(\rho\tau\pi). \tag{4.20}$$

The first order efficiency of 1-D holograms is plotted against the main replay fraction σ in Figure 4.10 with these assumptions: $\alpha = 1$, $\rho = 0.9$ and $m = 2$, 4 and 8. Note that within the window of ± 0.15 normalised co-ordinate, the binary, quaternary and octonary holograms have approximately 33, 66 and 77% (or -4.8, -1.8 and -1.1 dB) first order diffraction efficiencies, respectively. If the first order peak locations are chosen close to the edges of the central replay replication (e.g. ±1/2), there is an additional 3.1 dB loss due to spatial scaling with $\rho = 0.9$. The sole effect of spatial quantisation effect with fill-factor $\rho = 0.9$ is plotted by the dotted line.

For all first order replay locations, the replay intensities of the quaternary hologram are twice those of the binary hologram replay except for $\sigma = \pm 1/2$ binary gratings. Similarly, the replay intensities of the octonary hologram are higher than those of the quaternary hologram by $sinc^2(\pi/8)/sinc^2(\pi/4)$ except for the binary and quaternary gratings ($\sigma = \pm 1/2$ and $\pm 1/4$).

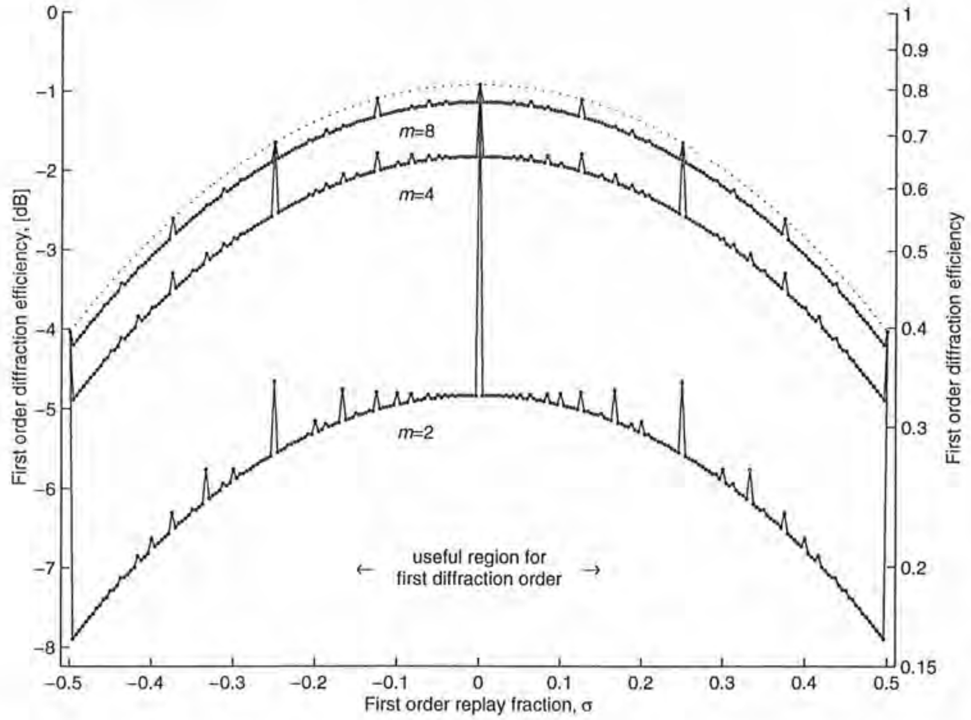


Figure 4.10: The intensity of the first order replay peaks within the central replication.

4.4 Inadequate phase modulation (phase mismatch)

In the above, a phase-matched hologram has been assumed, in which the difference between the highest and the lowest phase modulation is $2\pi(m\text{-}1)/m$. When the phase difference is less than ideal, wave theory has to be employed to sum the intensity at each replay order. For the replay of a 1-D grating, this has been shown to be [81,82]:

$$\eta_{m:n}^{\phi} = \sum_{n=-\infty}^{\infty} sinc^2\left(\frac{n\pi}{m}\right) \times \left| \sum_{g=1}^{g=m} \frac{\exp\left(j(\phi'-2n\pi)\cdot\frac{g}{m}\right)}{m} \right|^2 , \tag{4.21}$$

where m is the number of phase levels and n is the order number. ϕ' denotes the largest phase difference with 2π being the ideal case. Considering the overlapping nature of the phase quantisation and taking ϕ with an ideal value of $2\pi(m\text{-}1)/m$, the efficiency expression becomes $\eta_{m:n}^{\sigma,\tau} \times \eta_{m:n}^{\phi}$, where the intensity due to phase quantisation, $\eta_{m:n}^{\sigma,\tau}$, is given by eq. (4.9). The inadequate phase modulation term is:

$$\eta^{\phi}_{m;n} = \left| \exp\left(j\left(\frac{\phi}{2} - \frac{m-1}{m} n\pi \right)\right) \frac{sinc\left(\frac{\phi}{2}\frac{m}{m-1} - n\pi\right)}{sinc\left(\left(\frac{\phi}{2}\frac{m}{m-1} - n\pi\right)/m\right)} \right|^2, \tag{4.22}$$

where $n = \pm(gm + 1)$, $n = \pm(gm + 2)$, $n = \pm(gm + 3)$, … $n = \pm(gm + m/2)$, g being an integer. An implicit assumption of the summation in eq. (4.21) is that there are equal numbers of each phase present in the hologram. Hence, its simplification in eq. (4.23) must only be valid for x_o and y_o being multiples of m.

With the inclusion of this effect, a subset of hologram replay where the denominators of the x- and y-fraction are both multiples of an even number of phase levels, m, becomes:

$$\begin{aligned} \widehat{h}(x_r, y_r) = \sum_{n=-\left(\frac{x_o y_o}{2} - \frac{m}{2}\right)}^{\frac{x_o y_o}{2} - \frac{m}{2}} \sum_{s,t} & \left\{ \frac{sinc\left(\frac{n\pi}{m}\right)}{sinc\left(\frac{n\pi}{x_o y_o}\right)} \exp\left(j\left(\frac{\phi}{2} - \frac{m-1}{m} n\pi \right)\right) \right. \\ & \left. \times \frac{sinc\left(\frac{\phi}{2}\cdot\frac{m}{m-1} - n\pi\right)}{sinc\left(\left(\frac{\phi}{2}\cdot\frac{m}{m-1} - n\pi\right)/m\right)} \delta\left(x_r - (\langle s + n\sigma \rangle), y_r - (\langle t + n\tau \rangle)\right) \right\} \end{aligned} \tag{4.23}$$

For the other subset of m and $x_o y_o$ combinations where $x_o y_o$ is not a multiple of m, the number of replay peaks is the same for both phase-matched and phase-mismatched cases, i.e., $x_o y_o$. Summing the contribution of each phase element to the replay of n order cannot be simplified readily. The replay for each order is best evaluated using a discrete Fourier transform.

4.4.1 The zero orders for phase-mismatched holograms

Taking spatial quantisation into account, eq. (4.23) yields the intensity of zero orders across multiple replications of the discrete hologram replay,

$$\begin{aligned} \left| h(x_r, y_r)_{m;n=0} \right|^2 = & \left| \alpha\rho^2 \exp\left(j\frac{\phi}{2}\right) \frac{sinc\left(\frac{\phi}{2}\frac{m}{m-1}\right)}{sinc\left(\frac{\phi}{2}\frac{1}{m-1}\right)} + \beta\left(1 - \rho^2\right)\exp(j\psi) \right|^2 \\ & + \sum_p{}' \sum_q{}' \left| \alpha \exp\left(j\frac{\phi}{2}\right) \frac{sinc\left(\frac{\phi}{2}\frac{m}{m-1}\right)}{sinc\left(\frac{\phi}{2}\frac{1}{m-1}\right)} - \beta\exp(j\psi) \right|^2 \rho^4 sinc^2(\rho p\pi) sinc^2(\rho q\pi) \end{aligned} \tag{4.24}$$

Note that the DC average $\phi/2$ phase modulation of the pixel areas forms a complex amplitude grating with the constant dead-space phase modulation of ψ. It is likely that the amplitudes of pixel aperture and dead-space transmittance are not the same. This grating has twice the spatial frequency of the $\{0,\pi\}$ grating. Hence, its replay peaks are located at integer multiples along the two normalised fractional co-ordinates. In addition, there is a DC term due to the non-π relative phase of pixel average and dead-space phase modulation as well as a result of the "*area mismatch*" between pixel and dead-space.

An example is given in Figure 4.11(a) with $m = 2$, $\alpha = 0.8$, $\beta = 1$, $\rho = 0.8125$, $\sigma = 3/8$ and $\psi = 234°$. The variations of the first and zero order intensities in the central replication with binary phase depth are shown by the solid and dotted lines, respectively. As expected, the maximum of the central first order replay peak occurs at π binary modulation. However, the minimum of the central zero order depends largely on the constant phase of the dead-space (ψ) and α, and to a lesser extent β and ρ for a typical SLM ($\rho > 0.5$). This relationship may be important in characterising the phase response of an SLM using continuous modulation. The maximum ratio of the first to zero order intensities may not correspond to π phase modulation.

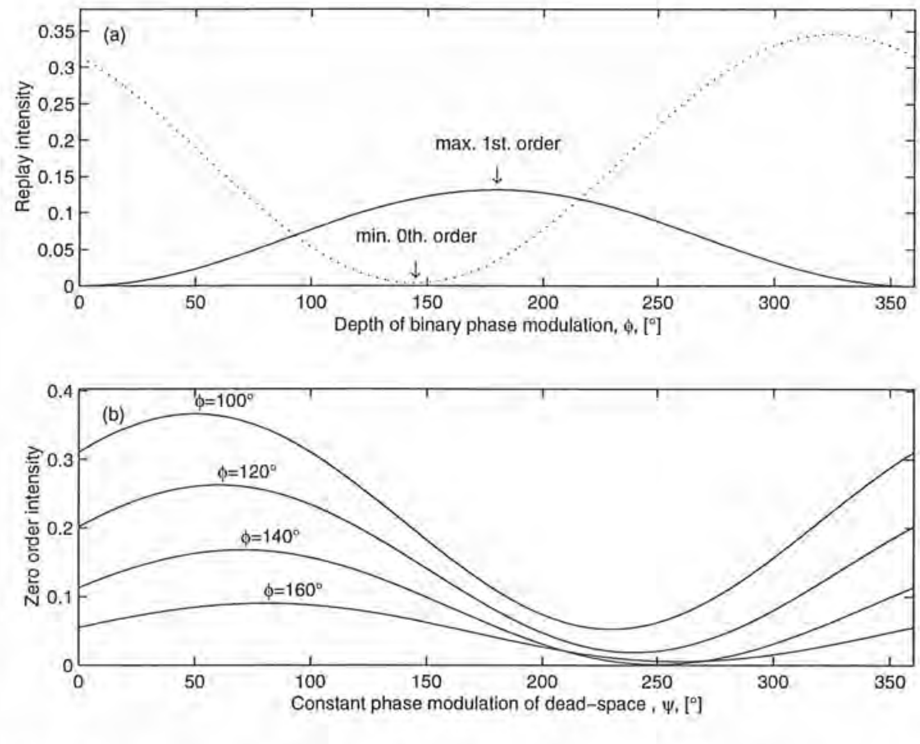


Figure 4.11: (a) The intensities of the central first and zero order replay peaks.
(b) The intensities of the central zero order for phase-mismatched binary holograms.

In the best case (i.e. $\psi = \phi/2 + \pi$), most of the light from the dead-space and the average of pixels would be routed to ±1 order and its multiples (i.e. the zero orders of each non-central replications). This gives the lowest replay intensity for the central zero order, as depicted in Figure 4.11(b). The four plots correspond to $\phi = 100°$, $120°$, $140°$ and $160°$ of a phase-mismatched binary hologram. All other parameters are as in Figure 4.11(a).

The intensity of the central zero order is often higher than the first order replay. Reducing the intensity of the central zero order can decrease the crosstalk power due to the coupling of sidelobes. An opaque dead-space does not help to diffract the non-modulated light away from the optical axis. But a transmissive dead-space can be used to boost/reduce the fraction of light going to the central zero order depending on the amplitude and phase of each component. If the dead-space modulation cannot be used to an advantage, its indirect influence on the intensity of the central order should be considered especially where an optical switch requires very high crosstalk isolation. In the worst case where the average of the pixel and dead-space phase modulation is the same, crosstalk level may be limited by the evanescent coupling of the central zero order.

4.4.2 The replay of non-zero orders for phase-mismatched holograms

The replay of non-zero orders can be expressed as:

$$\left|h(x_r, y_r)_{m;n\neq 0}\right|^2 = \sum_{n} \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \sum_{s\to-\infty}^{\to\infty} \sum_{t\to-\infty}^{\to\infty} \eta_{m;n\neq 0}^{\sigma,\tau;\phi;p,q} , \tag{4.25}$$

where $\eta_{m;p;n\neq 0}^{\sigma,\tau;\phi;p,q} = \eta_{m;n\neq 0}^{\sigma,\tau} \times \eta_{m;n\neq 0}^{\phi} \times \eta_{p;n\neq 0}^{\sigma,\tau;p,q}$; $\eta_{m;n\neq 0}^{\sigma,\tau}$, $\eta_{m;n\neq 0}^{\phi}$ and $\eta_{m;n\neq 0}^{\sigma,\tau;p,q}$ are eqs. (4.9), (4.23) and (4.16), respectively. The intensity expressed by eq. (4.25) has taken into account phase quantisation using m levels, the largest phase difference between the top and bottom phase levels (ϕ), pixellation and dead-space with fill-factor ρ and the 2-D distribution of the m phase states to route light to the main replay peak with fractional co-ordinate (σ,τ).

The use of the 1-D version of eq. (4.25) is shown in Figure 4.12 with exaggerated phase depth for the clarity of displaying the intensities of higher diffraction orders. Dotted lines are the composite scalings (phase/spatial quantisation and phase mismatch); 'o' marks the replay peak locations within the central replay replication; dashed line is the zero order scaling for zero orders located outside the central replication; and 'x' is the zero order intensity in the central replication. The effect of phase-mismatch severely reduces the first order intensity and causes inadequate field cancellation of the replay of higher orders. Hence, noise peaks creating potential crosstalk spread all over the replay plane.

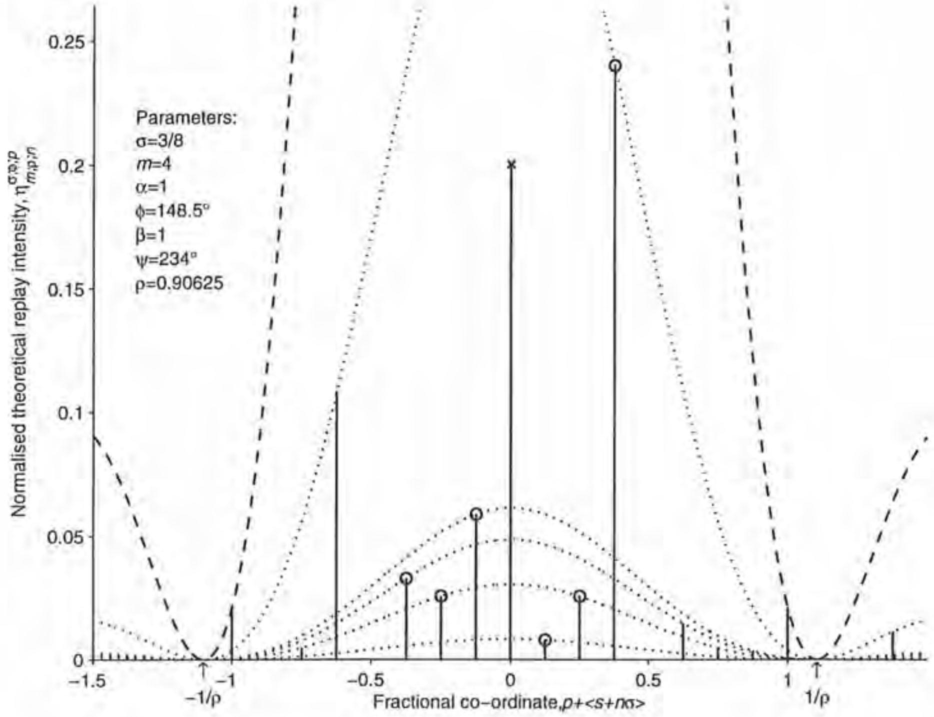


Figure 4.12: Replay intensity for a phase-mismatched quaternary replay fraction σ = 3/8.

If x_oy_o (i.e., $lcm(x_o, y_o)$) is a multiple of m, the location and intensity of all the replay peaks in the central replication can be determined using,

$$\left|h(x_r, y_r)_{p,q=0}\right|^2 = \sum_{s,t} \eta^{\sigma,\tau}_{m:n\neq0}\eta^{\phi}_{m:n\neq0}\,\alpha^2\,\rho^4 sinc^2\left(\rho\pi(\langle s+n\sigma\rangle)\right)sinc^2\left(\rho\pi(\langle t+n\tau\rangle)\right). \tag{4.26}$$

For this subset of all combinations of m and x_oy_o, the number of peaks appearing in the central replay replication for a phase-mismatched hologram is (m-1) times that of a phase-matched hologram. In fact, all diffraction orders (e.g., n = ... -3,-2,-1,0,1,2,3 ...) can be considered present and higher orders which are multiples of m give zero intensity due to $\eta^{\sigma,\tau}_{m:n}$. Nonetheless, a binary hologram has the same number of replay peaks regardless of the depth of phase modulation.

4.5 Conclusions

In order to meet the stringent crosstalk requirements of transparent optical networks, non-iterative-search single-peak holograms with a predictable distribution of replay peaks and intensities are essential. In this chapter, phase-only holograms are described according to the fraction of the central replay replication where their first diffraction order resides. A modified form of Dammann's efficiency expression is used to predict the replay intensity of a base hologram with a limited number of phase levels. The distribution of phase elements within the base hologram determines which harmonic frequency, if not the fundamental (i.e., corresponding to a grating replay), is excited. The locations of each diffraction order can be deduced by applying the *modulo-1* skip-rotate rule.

The underlying pixel periodicity of each SLM dictates the interval of the replay replications. The ratio of pixel clear aperture to pixel pitch determines the scaling of these replications. For the simplest physical modulation characteristics of pixels where the dead-space has a constant transmittance function, scaling expressions have been derived and found accurate by simulated optical hologram replays. Within a small window of approximately 30% of the central replay replication, the theoretical upper boundsof the first order replay efficiencies are 33, 66 and 77% for binary, quaternary and 8-level 1-D holograms assuming a 90% linear pixel fill-factor. The efficiency falls rapidly for holograms with a higher first order spatial frequency and for holograms performing a 2-D routing.

The theoretical limit is reduced by phase mismatch present in any real hologram recording device. For a limited number of cases where there is an equal measure of each phase state composing the base hologram, the scaling expression with this additional factor was found to be correct by simulation. The intensity of the central zero order which may potentially limit the crosstalk performance of an optical switch by way of sidelobe coupling was also investigated.

CHAPTER 5

Applications of coupling intensity and discrete hologram replay descriptions

5.1 Introduction

The understanding of coupling intensity profile and discrete hologram replay as detailed in the previous two chapters has two important applications. By combining the analyses of coupling intensity and discrete hologram replay, the requirements for a desired crosstalk level of a 1:*N* switch can be explored. The largest denominator of the replay fractions and the minimum number of hologram "*repeats*" for a given SLM SBWP in order to achieve this crosstalk figures are given. In addition, using the knowledge of the first order efficiency variation, the required crosstalk isolation and the physical dimensions of various SLM, optical and holographic parameters as well as the scalability and insertion loss limits of a large 1:*N* free-space holographic switch are analysed.

The *modulo-1* shift rule introduced in §4.2.3 is directly applicable in generating 1-D and 2-D multi-level phase-only routing holograms. The optical reconstruction of these holograms produces a single peak replay, along with its higher order peaks and sidelobes between peaks.

5.2 The design limitations of a large 1:*N* holographic switch

The complete replay field profile is the convolution of the replay field due to a *blank* hologram replay and the *discrete* hologram replay comprising a series of weighted delta functions. If the hologram replay is adequately resolved, this gives, within each replication, pronounced replay "*peaks*" and "*sidelobes*" corresponding to the non-zero and zero delta function locations of the discrete hologram replay. Obviously some of the peaks and sidelobes will be indistinguishable if the intensities of those discrete replay orders are very low. For a hologram replay which is not adequately resolved, the sidelobes can be completely missing.

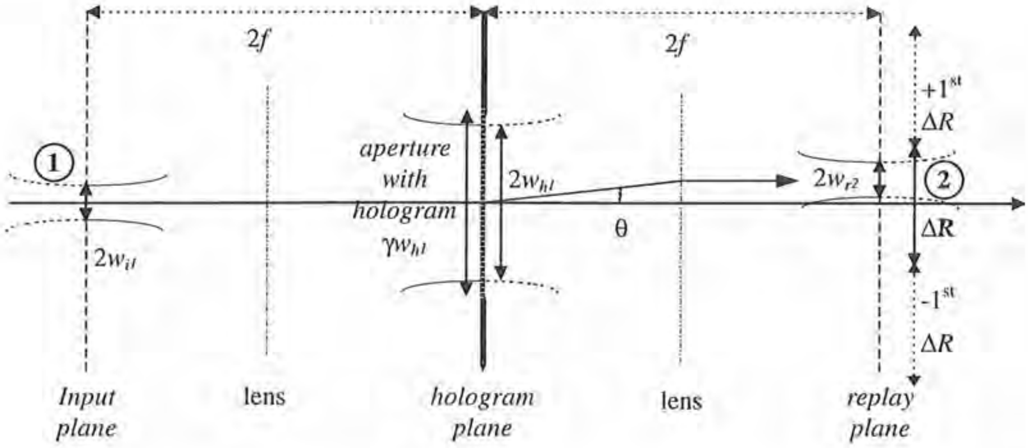


Figure 5.1: Coupling of holographic replay power into output fibres.

The aim of this section is to explore the limits of a 1:N switch in terms of the smallest holographic "*replay resolution*" and its associated on-beam-axis coupling efficiency and off-beam-axis crosstalk level using a highly specified SLM. The number of hologram "*repeats*" on the SLM as a function of the required crosstalk is also given. The smallest replay resolution refers to the separation between two replay peaks for a given hologram which contains *multiple repeats of a base hologram.* This is done by considering the combined effects of the analysis of the replay field and coupling intensity profile of a blank hologram and the analysis of discrete delta function replay. Figure 5.1 depicts the 4f configuration of a 1:N holographic switch. Further assumptions are based on figures achievable using a Silicon backplane liquid crystal SLM as given in the following:

i) A Silicon chip typically has a 15 mm maximum linear dimension (with current technology) and the modulator aperture, D_h is assumed 12 mm. The areas outside the SLM aperture is assumed opaque. The whole modulator array is to be used for a 1-to-N holographic switch.

ii) The size of the central replay replication, ΔR is $f\lambda/d$ where f is the focal length of the transform lens, λ is the operating wavelength at 1.55 μm and d is the SLM pixel pitch. For a given λ and d, ΔR is only dependent on f.

iii) The largest steering angle, θ, for a useful holographic connection has to be smaller than the diffraction-limited operation of the transform lens at < 5°.

5.2.1 On-beam-axis coupling efficiency and off-beam-axis crosstalk power

The input power of Gaussian beam transmitted through the SLM aperture having a *symmetrical* (the aperture clips both shoulders of Gaussian field equally) and *equal* x- and y-direction aperture size to Gaussian beam radius of γ is $(erf(\gamma/\sqrt{2}))^2$. However, evaluating the overlap-integral at the hologram plane for the coupling intensity of an output fibre coinciding with the replay beam axis (whatever the angular tilt caused by the hologram phase modulation is exactly matched by the offset of the fibre from the optical axis) entails a further drop in coupling power as given by $(erf(\gamma/\sqrt{2}))^4$. This is due to the diffraction effect of the truncation aperture which results in a main lobe and many sidelobes replay for each replay order obtained from the discrete analysis. The reduced coupling efficiency has taken into consideration the mismatch between the fibre mode and the broadened beam. The trade-off is between choosing a large γ value(so as to increase on-beam-axis coupling intensity and reducing crosstalk due to sidelobes), and the size of central replay replication.

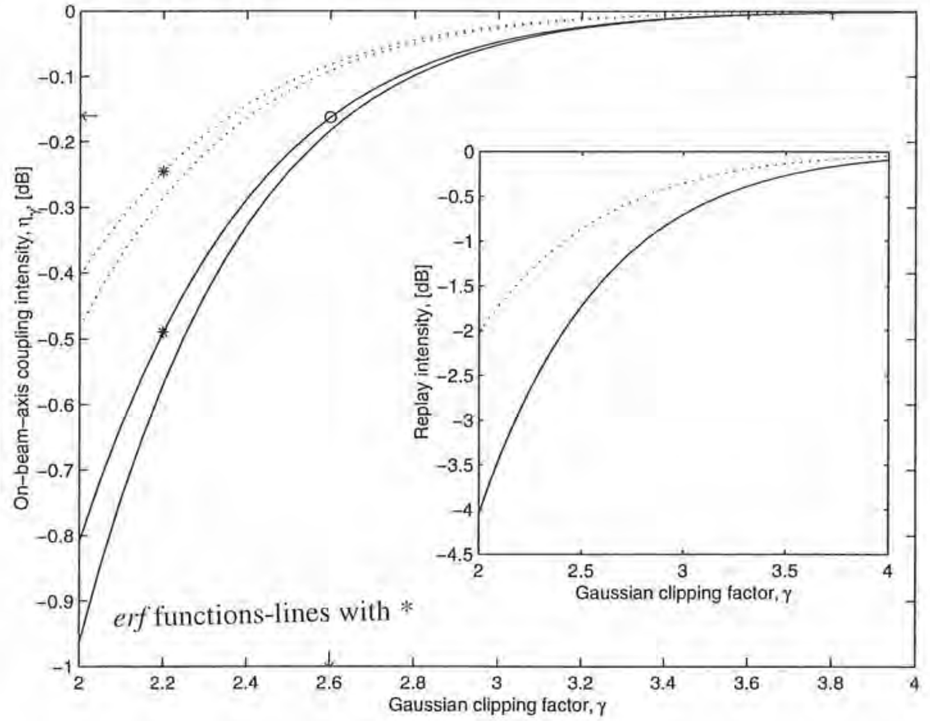


Figure 5.2: The dependence of coupling and replay intensity on the truncation ratio, γ.

Figure 5.2 depicts the on-beam-axis coupling intensity using the exact numerical solution (*erf* functions-lines with *) and the asymptotic approximation expression. Solid lines denote 2-D symmetrical and equal truncation and dotted lines denote the collimated beam is truncated in one direction only. It could be deduced that a γ value of greater than 2.6 is necessary in order not to incur > 0.2 dB coupling power penalty. Inset shows the corresponding replay intensity using the asymptotic approximation. However it is a less useful figure of merit for a fibre-to-fibre switch because the values cannot be measured experimentally.

Where a fibre port does not coincide with any higher order locations of the replay of all other holograms, the crosstalk noise floor is likely to be the maximum level of the sidelobe power. For a clipping ratio of > 2.6, the noise floor level is approximately -45 dB and -90 dB for a fibre with ($u = \pm 30$ μm, $v = 0$) and ($u = \pm 30$ μm, $v = \pm 30$ μm) lateral offsets[*] from the replay beam axis, respectively. This is shown in Figure 5.3. The dotted line denotes on-transverse axis offsets and solid line denotes offsets from both transverse axes. Dashed-dot and dashed lines plot the coupling intensity according to the asymptotic approximation at 30 μm along one and away from both transverse axes, respectively.

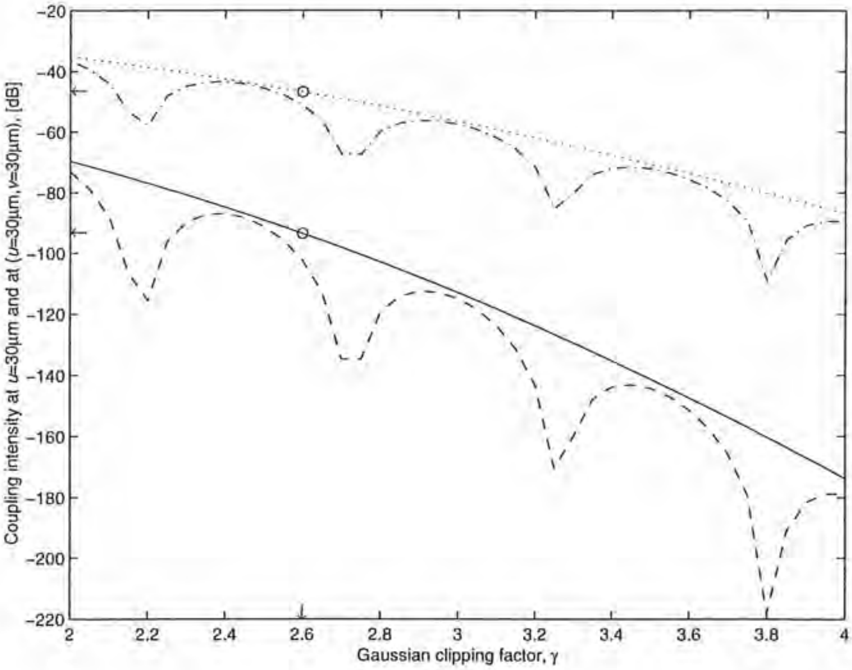


Figure 5.3: Crosstalk level at 30 μm offset due to the replay of a single beam.

5.2.2 The number of hologram repeats for a 1:N holographic switch

Considering an actual 1-D hologram replay with non-ideal phase encoding, there will be at most x_o replay peaks in the central replication. A further approximation in the convolution of the coupling field profile with discrete hologram replay field assumes there are sufficient replications of the base hologram displayed on the SLM. If $\gamma = 1$ the illuminating Gaussian beam can be approximated as a plane wave. In this case, there should be at least 2× replication of the base hologram in order to resolve the replay peaks adequately. From the earlier analysis, γ has to be > 2.6 to maximise the on-beam-axis coupling efficiency and

---

[*] At offsets greater than 25 μm, the crosstalk power is sufficiently low, see §3.5.2. A 30 μm offset was chosen such that the off-beam-axis crosstalk was calculated at around the peak of a sidelobe.

reduce the off-beam-axis crosstalk power. Scaling γ to large values does not necessarily give rise to proportional crosstalk reduction according to eq. (3.22).

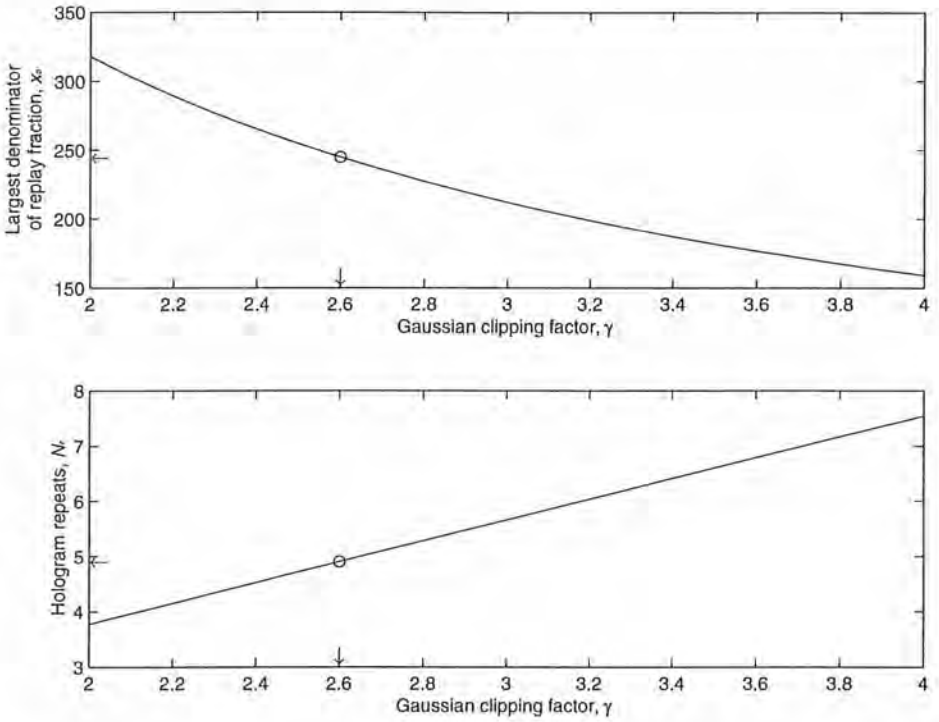


Figure 5.4: (a) The largest x_o of a *base hologram* to adequately resolve the replay. (b) The corresponding *hologram repeats* for $N = 1200$.

For $1 < \gamma < 4$ where neither plane wave nor Gaussian illumination assumption is appropriate, the coupling intensity for a fibre located at ~ 30 μm along the transverse axes of a replay beam will couple in approximately -20γ dB of crosstalk power. For a given γ value, increasing the denominator of the replay fraction will ultimately limit the crosstalk level to the noise level of higher order peaks instead of the noise floor due to the sidelobes. Hence for the crosstalk level to be limited by the sidelobe power, the general rule is to have approximately 2γ replication of the base hologram on the SLM. This limits the largest denominator of the replay fraction to approximately $N/(2\gamma)$. Taking the Gaussian radius as 5.06 μm at 1.55 μm wavelength, it was found that a 30 μm separation between two neighbouring replay orders was necessary to adequately resolve the relay. Assuming this "*replay resolution*" together with a highly-specified SLM of 1200 pixels at 10 μm pixel pitch, the largest denominator and the minimum hologram repeat in order to meet the 30μm replay resolution are plotted in Figure 5.4(a) and (b).

The numerical convolution of the fibre coupling field approximation and the discrete hologram replay field was performed for various replay fractions and clipping factor, γ. A section of the coupling intensity profile close to the main replay peak is illustrated in Figure 5.5 and Figure 5.6 for 90% phase-matched 1/20 and 9/200 base holograms with $\gamma = 6.15$. The corresponding number of *hologram repeats* are 60 and 6. In both figures, dashed line plots the coupling intensity profile multiplied by the replay intensity of the main peak and shifted to the fractional co-ordinate of the main peak. The dotted line is the *sinc* squared scaling of the main peak multiplied by the on-beam-axis coupling intensity. Clearly, the 6× hologram repeat is not sufficient to resolve the replay of $\sigma = 9/200$ at $\gamma = 6.15$. However, choosing $\gamma = 3.07$ for a 2γ *hologram repeat* adequately resolves the replay peaks as shown in Figure 5.7. The increase in ΔR frees up more locations to place other fibre ports, where the crosstalk power is < -60 dB.

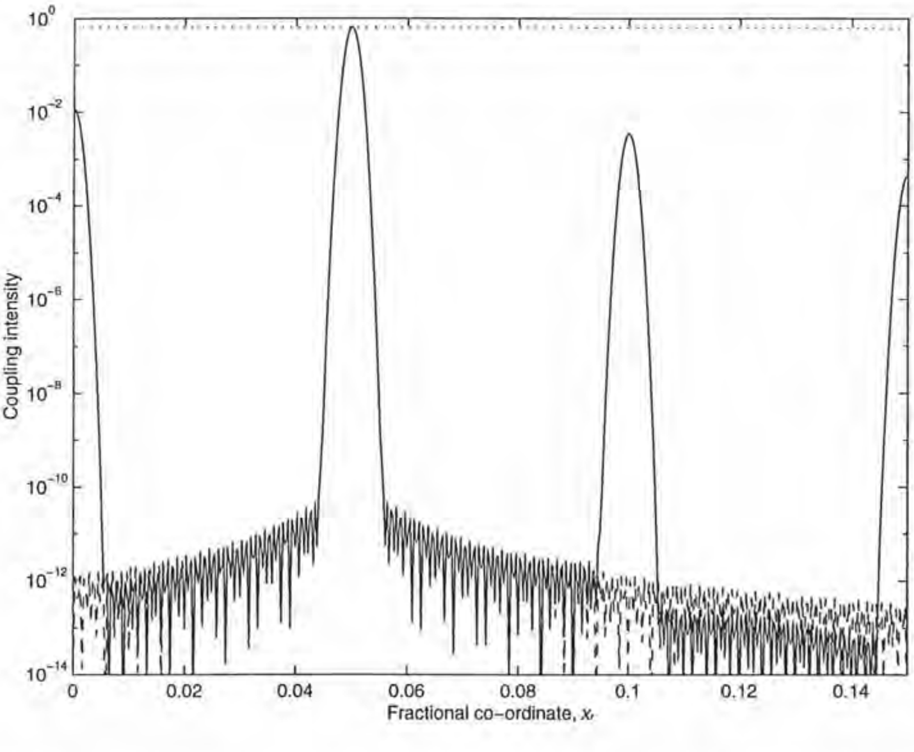


Figure 5.5: The coupling intensity profile at the replay plane for $\gamma = 6.15$ and $\sigma = 1/20$.

Attempting to obtain the benefit of crosstalk reduction beyond -60 dB by choosing $\gamma >> 3$ is both unrealistic and wasteful of the SLM SBWP. Aberrations, scattering and stray reflections in the optical system for large pupil size and temporal phase modulation shift will severely distort the calculated crosstalk level. The optimum value for γ is between 2.6 and 3 which gives -45 ~ -60 dB crosstalk at transverse axes locations so long as these locations are not in close proximity of higher order replay peaks. For $2.6 > \gamma > 3$ the separation between two replay peaks has to be > 30 ~ 40 μm in order that all replay peaks are sufficiently resolved.

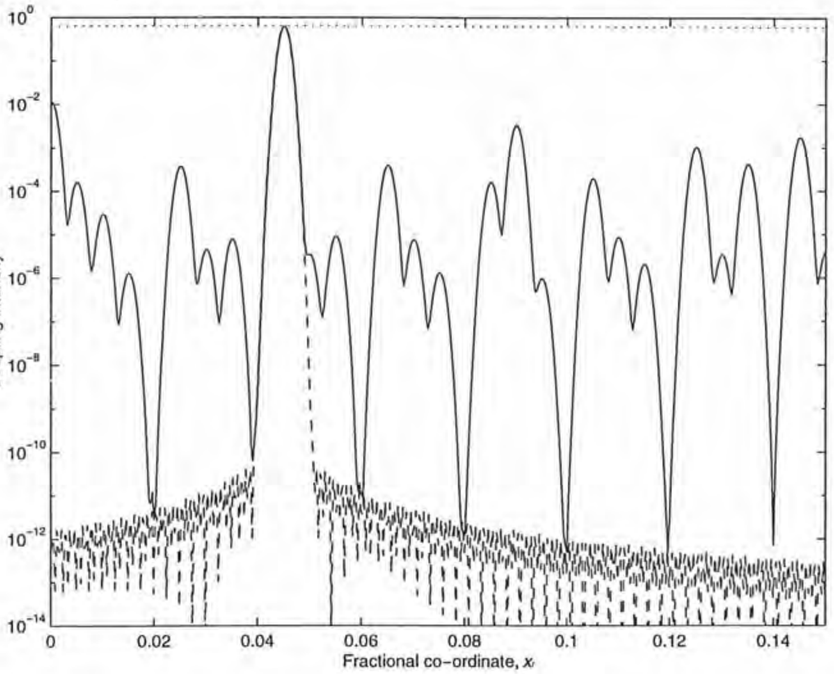


Figure 5.6: The coupling intensity profile at the replay plane for $\gamma = 6.15$ and $\sigma = 9/200$.

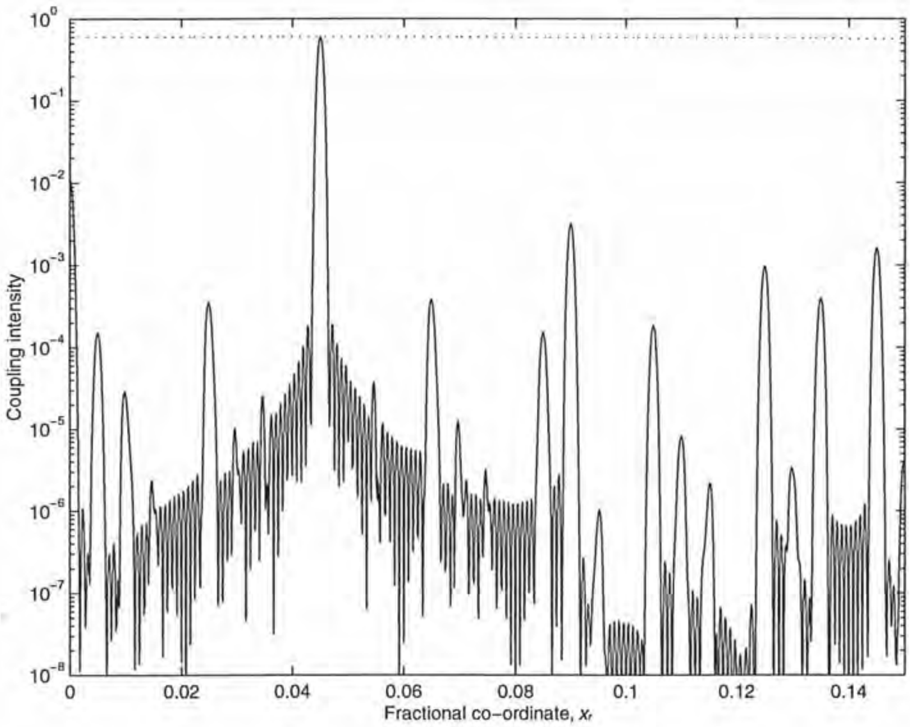


Figure 5.7: The coupling intensity profile at the replay plane for $\gamma = 3.07$ and $\sigma = 9/200$.

For an optical switch with a smaller matrix size, a 40 dB crosstalk isolation may be adequate. In this case, it is not detrimental to the crosstalk level if the number of peaks occurring in the output plane is increased such that the maximum denominator of the replay fraction is half the number of pixels (i.e., $x_o = N/2$). The large number of peaks appearing will result in unresolved replay peaks. These higher order peaks are weak and the unresolved replay does not increase the -40 dB crosstalk level significantly for the case of d = 10 μm and D_h = 12 mm.

The 2-D hologram replay with a finite SLM aperture follows the same treatment as the 1-D case. The combination of (x_o,y_o) arbitrary denominators for the replay fraction at worst gives rise to $x_o \times y_o$ peaks. If the criterion for the number of hologram repeats, 2γ in each direction is satisfied, N^2 points within the central replay replication will resolve the replay adequately such that the crosstalk is limited to no greater than -20γ dB along the transverse axis of a beam by the sidelobes. Again, for a target crosstalk isolation of ~ -40 dB, the minimum number of base hologram repeats is 2 in each direction.

5.2.3 Prospective 1:N switch using a highly specified SLM

A Silicon backplane liquid crystal SLM with 20 μm pixel pitch (d) is easily achievable. For display applications pixel pitch < 10 μm has been demonstrated [83]. However, other device criteria such as the flatness of modulating layer and the pixel dead-space are more critical for the performance of holographic switches compared to displays. Based on the current SLM development capability, a 1 μm dead-space is a feasible target. In Figure 5.8(a) the largest steering angle corresponding to replaying the largest spatial frequency is plotted against SLM pixel pitch, d, at λ = 1.55 μm, $\sin(\theta) = \lambda/(2d)$.

Assuming that the entire replay field is to be utilised and the tolerance to wavelength shift (optical bandwidth) is neglected, catalogue lenses which typically have < 5° diffraction limited operation will not be suitable to replay holograms written onto SLMs with d smaller than 10 μm. Besides, the power loss due to dead-space will be intolerable given that dead-space size is severely constrained by the pixel circuit construction or the photolithography resolution for depositing a planarised mirror. The fraction of power loss due to dead-space is approximately, $1-\rho$ for 1-D and $1-\rho^2$ for 2-D SLMs where ρ is the 1-D fill-factor. This loss is shown in Figure 5.8(b) for 1-D ('dotted') and 2-D ('solid') SLMs. A state of the art SLM in the near future will probably have a 10 μm pixel pitch and a 1 μm dead-space. In Figure 5.8(b), it is assumed that dead-space is 1μm for any pixel pitches. A loss of 10% is incurred for d = 10 μm.

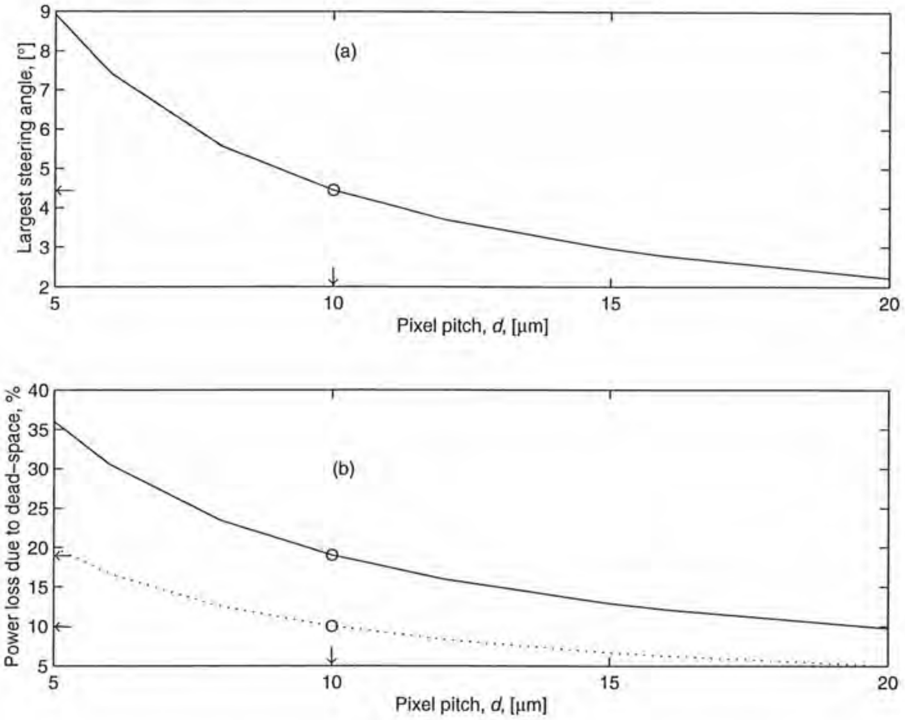


Figure 5.8: (a) The largest beam steering angle using $\lambda = 1.55$ μm.

(b) The fraction of power loss due to a 1 μm dead-space.

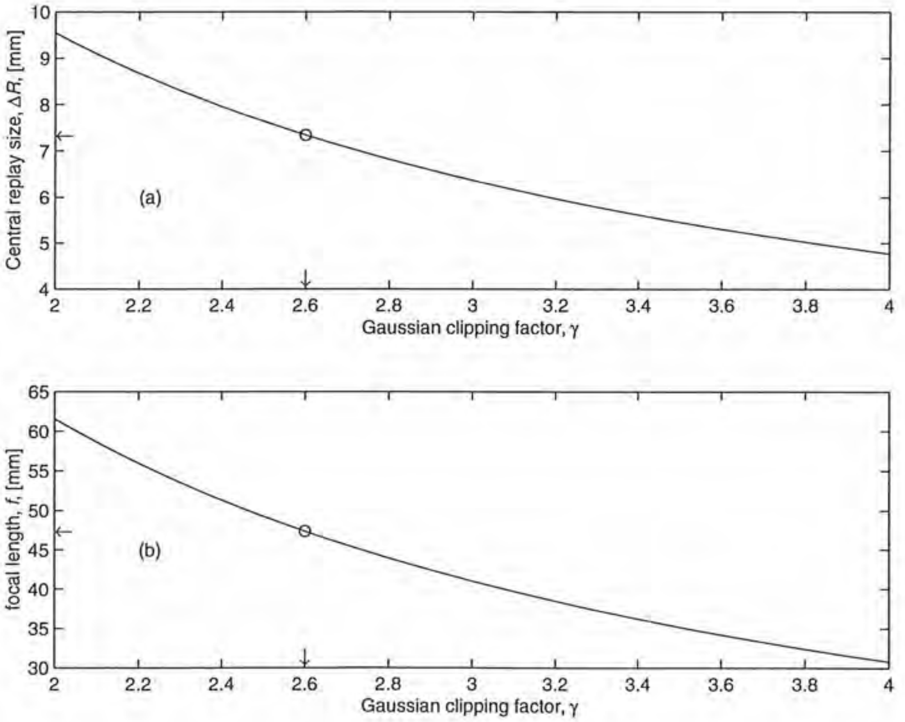


Figure 5.9: (a) The size of the central replay replication for $d = 10$ μm and $\lambda = 1.55$ μm.

(b) The corresponding focal length in order to achieve the γ values above.

The size of the central replay replication, $\Delta R = f\lambda/d$, is dependent on both γ and d. For $d = 10$ μm, this is illustrated in Figure 5.9(a). Of interest is the size of the central replay replication corresponding to $\gamma > 2.6$ which has a maximum of 7.34 mm for $\gamma = 2.6$. It is obvious that a large γ value offers the best on-beam-axis coupling intensity and sidelobe crosstalk at the expense of reducing the size of the central replay replication. The corresponding focal length f reduction with increasing γ is plotted in Figure 5.9(b). The overall system length of $4f$ is reduced accordingly with increasing γ.

The prospective SLM SBWP of 1200 in 1-D and pixel pitch of 10 μm give adequate replay size and manageable system length for $\gamma = 2.6$ Gaussian clipping factor. Some SLM technologies (e.g. FLC/Si) will probably not work properly if the electro-optic modulation layer is effected by < 10 μm pixellation. Assuming that a maximum of 1 dB variation in path loss (i.e., hologram replay efficiency) is permitted, the fraction of ΔR that can be used for locating the output waveguide ports spaced at 30 μm, κ, is ~ 0.6. Likewise, the maximum permitted power loss due to chromatic dispersions of a broadband source by the diffractive hologram is assumed to be 1 dB. The optical bandwidth is given by,

$$\Delta\lambda = 2\times\frac{2d}{\kappa}\times\frac{w_o}{f}\times\sqrt{\frac{\ln(10)\times\delta}{10}}, \tag{5.1}$$

where d, κ, f, w_o and δ are the pixel pitch, useful replay size ratio, focal length, Gaussian spot radius and the additional power loss, respectively. The optical bandwidth for the parameters listed in Table 5.1 is 3.4 nm.

For applications requiring a small optical bandwidth, this 1:N free-space optical switch design can accommodate ~ 146 waveguide ports spaced at 30 μm apart with < 10 dB insertion loss and > 40 dB crosstalk isolation (provided high orders are avoided). These best-case estimates also allow for additional losses due to non-ideal wave-plate configurations and FLC switch angle in the case of a holographic switch using binary FLC phase holograms. For applications requiring > 30 nm optical bandwidths, the largest number of output waveguides spaced at 30 μm apart is 16. This assumes the three switch design parameters, $\{d, \kappa, f\}$ can be conveniently chosen and $\kappa < 0.6$. Combinations such as {20 μm, 0.25, 25 mm} and {15 μm, 0.31, 15 mm} meet the 30 nm optical-bandwidth requirement, although a full design will need to take account of the bandwidth narrowing due to optical components.

1:N switch parameter	Symbol and value	1:N switch parameter	Symbol and value
No. of SLM pixels	$N_{SLM} = 1200$	Focal length	$f \approx 47$ mm
Pixel pitch	$d = 10$ μm	Central replay size	$\Delta R = 7.34$ mm
Pixel dead-space	$ds = 1$ μm	Useful replay-size ratio	$\kappa = 0.6$
1-D fill-factor	$\rho = 0.9$	Efficiency variation	$\Delta\eta = 1$ dB
Gaussian clip-ratio	$\gamma = 2.6$	waveguide spacing	$\Delta s = 30$ μm
Clipped coupling	$\eta_\gamma = -0.2$ dB	No. of waveguides	$N_{wg} = 146$
Centre wavelength	$\lambda = 1.55$ μm	Chromatic dispersion loss	$\delta = 1$ dB
Gaussian radius	$w_o = 5$ μm	Optical bandwidth	$\Delta\lambda = 3.4$ nm
Retardation efficiency ($\phi = \pi \pm 10\%$)	$\eta_\phi = -0.1$ dB	Phase efficiency (FLC switching angle, $\theta_o = 60°$)	$\eta_{\theta_o} = -1.25$ dB
Expected switch efficiency: $\eta = \eta_{hg} + \eta_\phi + \eta_{\theta_o} + \eta_\gamma - \Delta\eta - \delta = -8.35$ dB; where η_{hg} is the highest 1-D binary phase-only hologram efficiency including phase and spatial quantisation with $\rho = 0.9$; $\eta_{hg} \approx 33\%$ or -4.8 dB (see §4.3.2).			

Table 5.1: 1:N switch parameters and insertion loss components.

5.3 Deterministic routing hologram generation

Routing holograms are often required to steer the highest possible fraction of the input optical power to a single spot at the replay plane. Hence, it is useful to devise techniques that calculate the hologram patterns rapidly. Inasmuch as each of the patterns deflects the light only to one arbitrary location, a deterministic technique based on Fourier theory is not beyond the reach of a typical matrix manipulation software package.

Given that there are at most m phase levels available for hologram encoding, the description of a general routing hologram by the normalised replay fraction ($\sigma = x/x_o, \tau = y/y_o$) encapsulates both the generation procedures and the replay characteristics. A 1-D base hologram contains x_o or y_o phase elements whereas a $x_o \times y_o$ matrix defines an arbitrary 2-D replay fraction for any x_o, y_o or a combination of the two. As evident from the analysis in §4, the fractional representation of a hologram replay peak is intuitive and follows directly from the inverse co-ordinate relationship between the hologram plane and the transform plane. For example, the relay spot of a grating with x_o period will be located at $1/x_o$ away from the optical axis, along the normalised replay co-ordinate. This notation frees up the dependence on the pixel pitch of the encoding device, *d*, the wavelength of illumination, λ, and the focal length of the transform lens, *f*. Thus, this representation is most suited to the numerical hologram generation and replay evaluation of holograms. The actual hologram will contain several *repeats* of the base hologram. For the grating above, the actual optical replay will produce the first order peak at $[1/x_o]\times f\lambda/d$ as well as other higher order noise peaks.

The symbols used in calculating base holograms have the following meaning:

i) m the maximum number of *phase levels/states* available.

ii) x_o the denominator or the number of *phase elements* required for the base hologram providing the *x*-direction routing.

iii) x the numerator of the replay fraction for which the x^{th} harmonic (or the fundamental if $x = \pm 1$) is to be excited, $|x| \leq x_o/2$.

iv) Λ_x the grating period (length) of the intermediate $1/\Lambda_x$ *grating replay fraction*, where, Λ_x is the least common multiple of m and x_o.

v) y_o, y and Λ_y are the corresponding descriptions of the *y*-direction routing component.

5.3.1 1-D hologram generation by choosing a combination of x_o phase elements

The key to deterministically generate single-spot-replay holograms lies in creating 1-D grating functions and re-orientating the order of these phase elements to form the required base holograms. It is intuitive that writing a repeated linear array of {1 1 j j -1 -1 -j -j} using $m = 4$ will replay the first order peak at 1/8 of the central replay replication. {1 1 j j -1 -1 -j -j} is hence the base hologram (grating) corresponding to 1/8 replay fraction. The 1/4 (or 2/8) base hologram is obtained by taking every 2nd phase element of the above grating, i.e., {1 j -1 -j}. Similarly the 3/8 base hologram is the combination of {1st, 4th, 7th, 2nd. 5th, 8th, 3rd and 6th} phase elements from the 1/8 grating giving {1 j -j 1 -1 -j j -1} base hologram.

Each base hologram contains phase elements in the grating function pointed to by an index array which is returned by *modulo*(x_o,x) function. This skip and rotate rule takes the first index and every subsequent index *skips* along at x interval for a total of x_o entries. Indices that exceed x_o are *rotated* to within $1{:}x_o$ range by deducting multiples of x_o. Figure 5.10 shows the skip and rotate rule being used to generate base holograms for the examples above as well as 1/2 replay fraction.

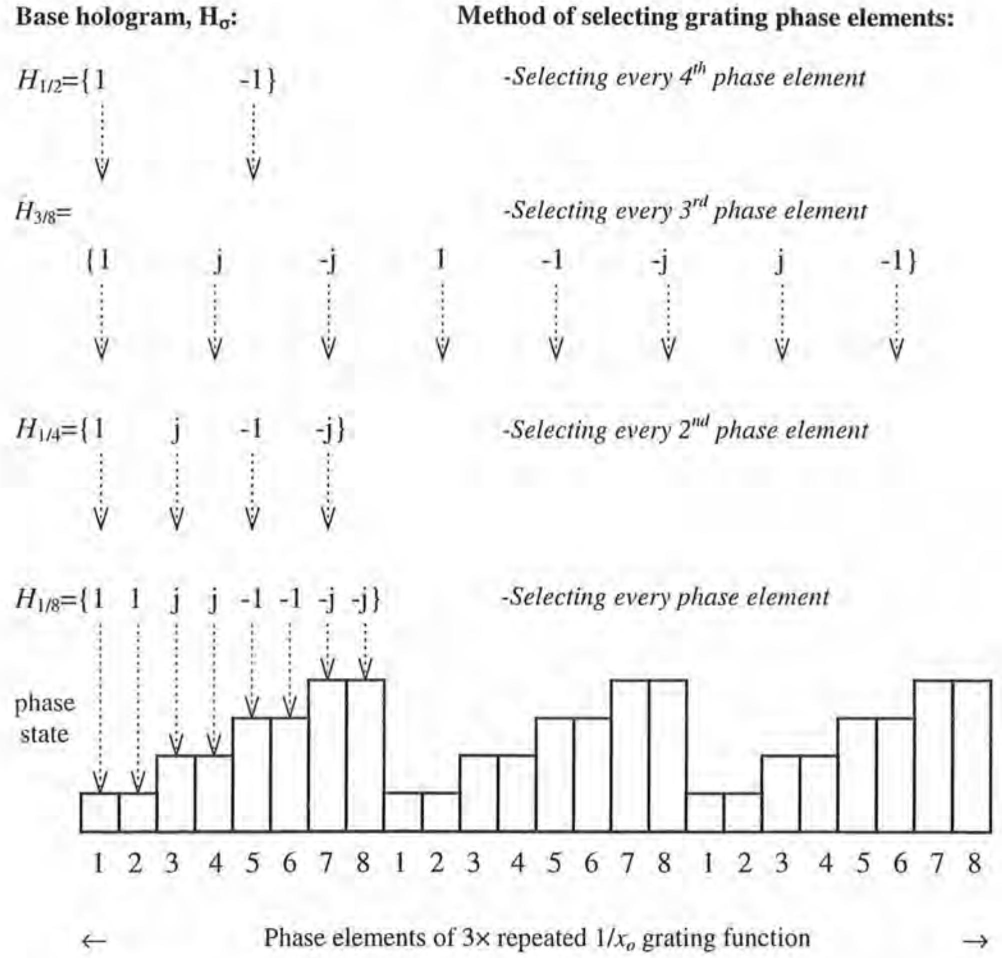


Figure 5.10: *Modulo-x_o* skip and rotate rule used for generating 1-D holograms.

For eight phase elements with four different states and given the circular symmetry, there are eight different combinations to be obtained. In addition to 1/8, 2/8, 3/8 and 4/8, other replay fractions are 5/8, 6/8, 7/8 and 8/8. However in order to adhere to the definition of σ, these fractions are translated to -3/8, -2/8,-1/8 and 0/8. Hence for negative fractions, *modulo*(x_o,x) works by either reversing the phase change direction of the intermediate grating function or by converting negative harmonics x to positive integers by adding with x_o. Using a typical software package capable of matrix operations, rearranging the grating phase elements to form the desired base hologram can be carried out easily. One thing to note is that an *fft* algorithm treats phase in the negative way. Thus the four examples shown in Figure 5.10 will be numerically transformed to -1/2, -3/8, -1/4 and -1/8, respectively. This inversion property does not complicate *modulo* function but it has to be considered consistently.

The idea illustrated in Figure 5.10 can be simpler in practice. The intermediate grating function will only have x_o phase elements where x_o is the denominator of a simple fraction x/x_o. For replay fractions with denominators which are not a multiple of m, the least common multiple (*lcm*) operation will give a grating period of Λ_x. The input arguments for *modulo* function are Λ_x and $x \times \Lambda_x/x_o$. As an example, if $m = 4$ and $\sigma = 1/5$, Λ_x will be 20. The intermediate grating will consist of five elements of each phase state arranged in increasing or decreasing order. The five harmonics have *skip* steps of $[-2,-1,0,1,2] \times 4$. There are cases where only a subset of all the phase states will be used in base holograms generated, e.g., $m = 4$ and $\sigma = \pm 1/2$.

5.3.2 Offsetting the 1-D combinations to provide 2-D routing patterns

There are two options one can choose to obtain 2-D routing holograms. One method is to take x and y orthogonal 1-D base holograms and form a 2-D crossed hologram. The first order replay efficiency is poorer as given by eq. (4.11). For example, crossed binary holograms will produce four symmetrical peaks within the central replication region. For each peak resulted from the horizontal routing, the vertical routing component further splits these peaks to form a series of peaks along the vertical axis. The two crossed components diffract light orthogonally in both directions resulting in rows and columns of crowded noise peaks.

A better way of implementing 2-D steering will be to form continuously-shifted Fourier holograms. This makes use of the properties of Fourier plane holograms whereby all spatially-shifted or phase-shifted holograms produce the same replay intensity pattern. Hence, each row or column alone does a horizontal or a vertical beam steering; the combination of the two steers the beam off the two transverse axes. The Fourier and crossed holograms are equivalent if at least one routing fraction is $1/m$.

In order to generate continuously-shifted base holograms, an intermediate 1-D grating that can be used for generating 1-D base holograms for both routing components is needed. A simple way to produce this m-level grating is to choose the length of the grating as $\Lambda_{xy} = lcm(m,x_o,y_o)$. The indices for both routing components are returned by *modulo*-Λ_{xy} function, having adjusted the skip intervals. The algorithm can start with either components (e.g. x-direction routing). For each 1-D row of the final 2-D base hologram, the corresponding y-direction index element is used to offset the x-direction index array. The grating phase elements at locations pointed to by the offset index array form the base hologram pattern for that row. An iteration through all the rows returns the 2-D base hologram. This is depicted by the flowchart in Figure 5.11. Alternatively, one can start with the index array for y-direction routing and iterate column wise. The choice of which direction to start with is arbitrary. The intermediate $1/\Lambda_{xy}$ grating has to be repeated twice to rotate overflow indices back to the same range.

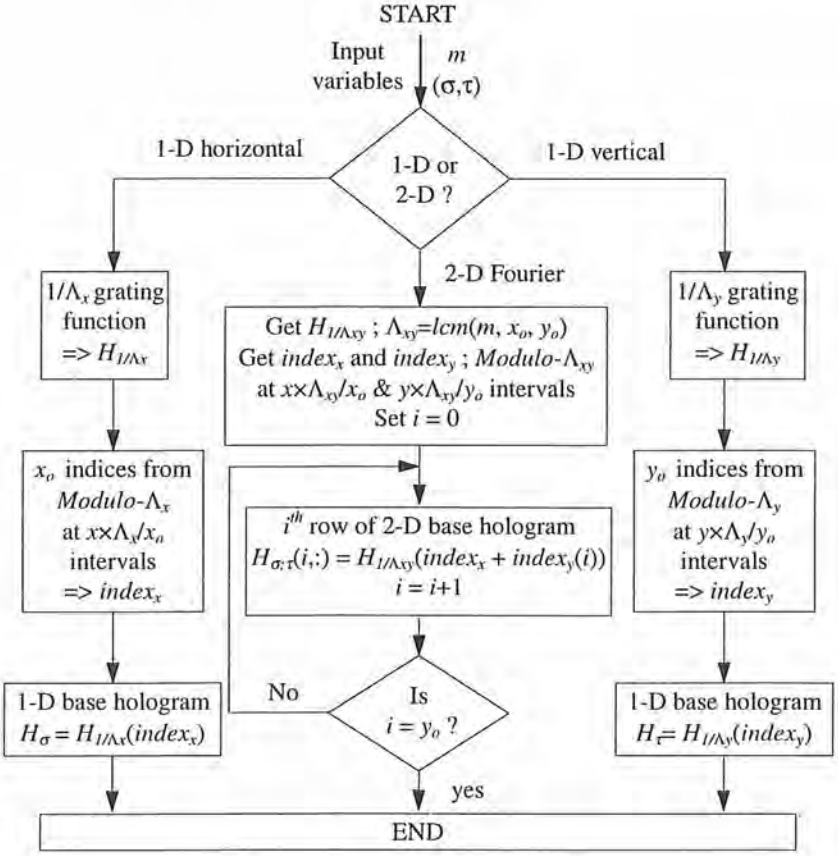


Figure 5.11: Flowchart of the algorithm for generating determinstic multi-level phase-only routing holograms.

5.3.3 The advantages of the skip-rotate hologram generation technique

The ability to calculate and display hologram on the fly is an asset in optimising the performance of an optical switch. Using the *skip* and *rotate* method, there is no need to perform any inverse Fourier transform. No iteration is needed for 1-D patterns and only x_o or y_o calculation loops have to be performed for 2-D cases. For a typical requirement of $m \leq 16$ and x_o and $y_o \leq 256$, each base hologram generation takes < 1s. As with *fft* algorithm, the fastest computation will involve m, x_o and y_o being powers of two.

The hologram synthesis algorithm and the fractional representation allows for fine tuning a replay fraction, and not just constrained by the SLM pixel resolutions [84]. This is only possible due to the use of Gaussian illumination. For certain optimisation algorithms, maximum efficiency base holograms with regular higher order noise peaks are required [85]. Optimisation reduces the noise intensity of the selected replay orders by deflecting these components to high spatial frequencies. As evident from the first order efficiency plot of each replay fraction in §4.3.2 and optical bandwidth consideration, the useful region within the central replay replication may be small. Hence, selective noise optimisation process does not reduce the variety of switch configurations.

5.4 Conclusions

The Gaussian clipping factor, γ, is an important parameter in designing a holographic switch fabric. A ratio of at least 2.6 is required in order to reduce clipping loss to < 0.2 dB. The corresponding on- and off-transverse-axis crosstalk levels due to sidelobe coupling are excellent at ~ -45 and ~ -90 dB, respectively, before aberration effects are considered. The approximate on-transverse-axis crosstalk power of -20γ dB using 2γ hologram repeats is verified by simulations. However, a more realistic target of 40 dB crosstalk isolation requires just 2× replications of the base hologram.

At 1.55 μm operation, the minimum peak-to-peak separation is ~ 30 μm. Using an achievable SLM specification of d = 10 μm pixel pitch and N_{SLM} = 1200 pixels, the size of the central replay replication is ~ 7 mm. This will accommodate a large number of output ports. The system length, power loss due to dead-space and the largest diffraction angle are all within acceptable ranges. A 1:*N* holographic switch using such an SLM is capable of providing ~ 146 waveguide ports at 30 μm pitch with < 10 dB insertion loss and crosstalk figures approximately at the sidelobe levels. However, the highly diffractive switch has a very small optical bandwidth of 3.4 nm. If the full 30 nm bandwidth of a conventional erbium-doped fibre amplifier is to be utilised, only 16 waveguide ports are possible. These switch sizes are limited by the 1-D routing of the holographic SLM. A substantial increase in the number of output ports is likely by using 2-D SLMs albeit at a higher loss due to spatial quantisation and dead-space.

The analysis of discrete routing hologram replay gives sufficient insights to enable the synthesis of this type of hologram by applying a set of 'skip and rotate' rules. This deterministic generation routine is capable of producing non-regular-grating holograms with large numbers of phase levels and large dimensions at high speed. These holograms can be directly usable in single-port holographic interconnects or as the base hologram for further optimisation.

CHAPTER 6

FLC on CMOS/VLSI Si spatial light modulators for holographic applications

6.1 Introduction

Diffractive optical elements (DOE) based on etched silica/quartz using electron beam lithography (EBL) are well-developed and capable of implementing various filter functions [86]. The main drawback of such techniques is the inability to reconfigure them. Hence dynamically reconfigurable holographic optical elements (HOE) have many applications domains such as optical interconnects, beam steering/deflection, adaptive optics and dynamic phase masks. In particular, phase-only holograms do not have any intrinsic loss.

One way to realise phase-only holograms is to encode the patterns onto a pixellated ferroelectric liquid crystal (FLC) on silicon SLM [87]. The ferroelectric liquid crystals used for binary holograms can be switched at speeds of order (tens of μs) [54]. The Fourier replay loss due to the inherent symmetry of binary holograms can be restored by utilising more than two phase levels. It is often thought that continuous analogue grey-scale planar-switching liquid crystal (LC) modulation gives very high hologram replay efficiency. The mechanics and the expressions governing the diffractive power of phase modulation by way of polarisation rotation are given in §6.2. The author's published results [88] inferring the limited benefits of using multi-level holograms recorded by planar-switching LCs are presented in §6.2.2.

The rest of the chapter focuses on the generic issues of designing a silicon backplane SLM for displaying holograms. Although multi-level phase-only holograms have been relegated to specialist applications, the capability of displaying four grey levels was to be designed into the demonstrator chip. Hence, several design options concerning the data conversion for analogue drivers and the SLM construction to obtain ideal quaternary phases are explored.

The principle design aim is to produce a linear silicon modulator array with good optical quality pixels. However, a CMOS process is optimised to implement electrical functionality with unrivalled circuit complexity and density. Therefore, attempting to make good optical devices necessitates a post-processing of the silicon backplane pixels. The quest for higher driving voltages is critical for dynamic holography due to the requirements of large tilt FLCs. A suitable silicon process to implement the backplane circuitry was chosen after considering the circuit and layout requirements of a moderate (11V) and a high (45V) voltage process.

6.2 Polarisation rotation as a means of phase modulation

The speed advantage of the polarisation rotation technique [56] over the changing of optical path lengths [82] by varying the refractive index is enormous. Binary phase-only holograms written using the FLC electro-optic effect at best give a replay of -4 dB first order power if a half-wave plate configuration and a 90° switching angle are available and the pixels are free from any defects. Some of the losses can be restored by having more phase levels. In order to achieve multi-level phase and fast switching, electroclinic and deformed-helix ferroelectric families of the chiral smectic liquid crystals (CSLC) are the promising candidates for multi-level holograms.

6.2.1 Multi-level modulation

Multi-level polarisation rotation, as in the binary case, assumes that the CSLC is an electro-optic medium possessing uniaxial optic axis. Ideally, the multi-level CSLC modulator is configured as a half-wave retarder. The maximum switching angle of the planar switching CSLC, θ_o, is divided into m evenly-spaced states.

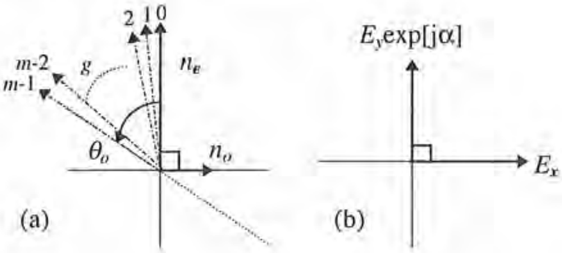


Figure 6.1: (a) m switched states of a CSLC cell and (b) two linear orthogonal polarisation components of the illuminating beam.

Figure 6.1(a) shows m CSLC molecular director positions, corresponding to $g = 0, 1, 2 \ldots m-1$ switched states. Each switched state has $g\theta_o/(m-1)$ angular difference from the extreme state at $g = 0$. The ordinary and extra-ordinary refractive indices are assigned n_o and n_e respectively, and the x and y axes are arbitrarily taken to be parallel to n_o and n_e. The input polarisation is resolved into the x and y linear orthogonal components, assigned E_x and $E_y e^{j\alpha}$ as shown in Figure 6.1(b); α being the phase difference between the x and y components of the incident beam.

Using matrix notation, the matrices for the input polarisation $\underline{E}_{in}$, anticlockwise β rotation $\mathbf{R}_\beta$ and the Jones matrix for π and φ retardation plates, $\mathbf{J}_\pi$ and $\mathbf{J}_\varphi$, are:

$$\underline{E}_{in} = \begin{bmatrix} E_x \\ E_y e^{j\alpha} \end{bmatrix}, \ \mathbf{R}_\beta = \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix}, \ \mathbf{J}_\varphi = \begin{bmatrix} e^{-j\varphi} & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \mathbf{J}_\pi = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{6.1}$$

respectively, where the relative retardance of the ordinary and extra-ordinary rays is $\varphi = k\Delta nT$; k is the wavenumber in free space, Δn is the birefrigence of the CSLC ($\Delta n = n_e - n_o$) and T is the thickness of the CSLC electro-optic medium for a full optical path. Assuming that the retardation by the CSLC is π (i.e., $T = \lambda/2\Delta n$ in transmission mode), the polarisation of the light beam exiting a modulating pixel switched to state g resolved ‖ and ⊥ to the state m-1 director (i.e., θ_o from the y-axis), $\underline{E}_{out}$ is as follows:

$$\underline{E}_{out} = \left[\mathbf{R}_{\theta_o - \beta}\right]\left[\mathbf{J}_\pi\right]\left[\mathbf{R}_\beta\right]\left[\underline{E}_{in}\right], \tag{6.2}$$

where β refers to the switched angle $g\theta_o/(m-1)$ for g^{th} state. It can be shown that the DC undiffracted light is independent of the amplitudes and the relative phase of the two linear orthogonal components (E_x, E_y and α) (Appendix B). In this case, the DC power fraction is,

$$DC^2_{\varphi=\pi;m} = \frac{sinc^2\left(\dfrac{m\theta_o}{m-1}\right)}{sinc^2\left(\dfrac{\theta_o}{m-1}\right)}. \tag{6.3}$$

If however, the retardance of the CSLC cell is not π, the DC term is still not affected by the input polarisation although *retardation efficiency* loss increases the DC undiffracted light to,

$$DC_{\varphi;m}^2 = \cos^2\left(\frac{\varphi}{2}\right) + \sin^2\left(\frac{\varphi}{2}\right) \times \frac{sinc^2\left(\frac{m\theta_o}{m-1}\right)}{sinc^2\left(\frac{\theta_o}{m-1}\right)}. \tag{6.4}$$

The imperfect retardation and the lack of switching angle causes power loss to the zero order but both are independent of the input polarisation. A LC cell with a non-π retardation due to fabrication tolerances and broadband sources does not give rise to critical power loss. Similarly, the dynamic fluid nature of liquid crystals can only result in averaged macroscopic director orientations. For ±5% of binary phase mismatch and non-π wave retardation, the excess loss to the zero order is < 2%.

An example of multi-level polarisation rotation is shown in Figure 6.2 for $m = 4$. In general the *polarisation shape* of light exiting from a fibre is elliptical and it could vary with time if uncontrolled. Polarisation control is impractical in a fibre network where each incoming signal could pass through numerous optical components and be subjected to different and time-varying fibre transmission characteristics. Given that the incoming signals can assume any polarisation shape before being phase modulated by a polarisation rotation device, the output from each of the modulated states can be different as shown in Figure 6.2.

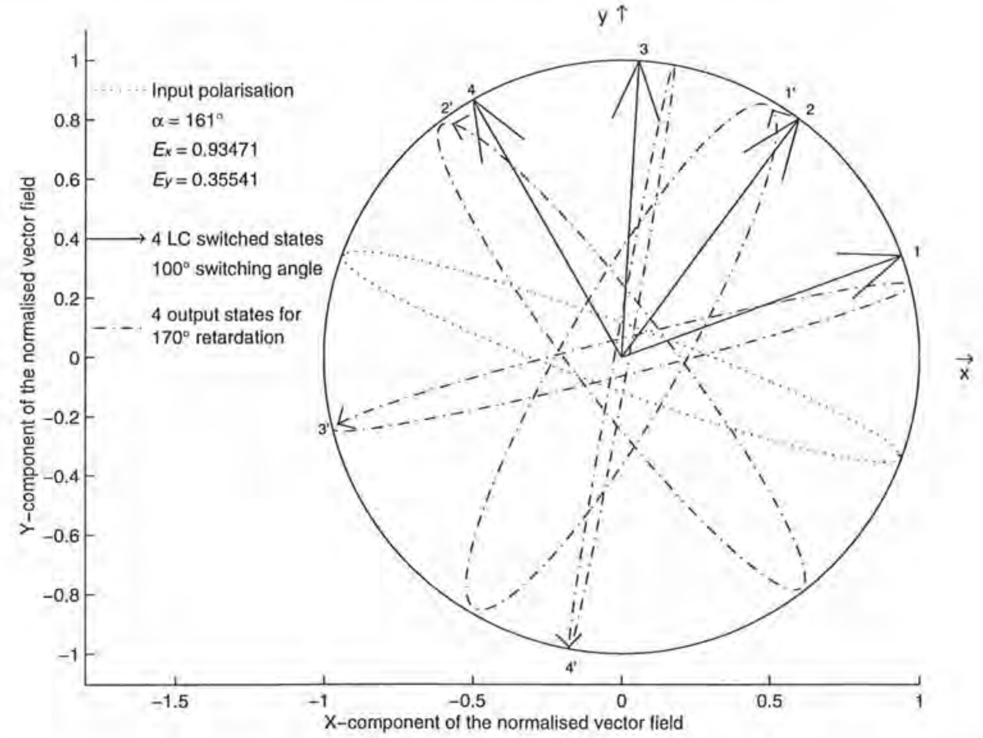
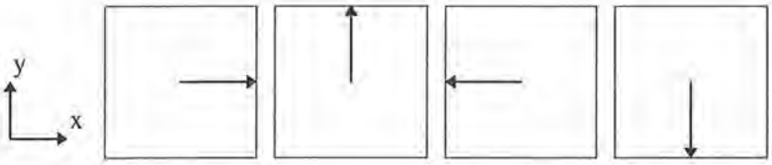


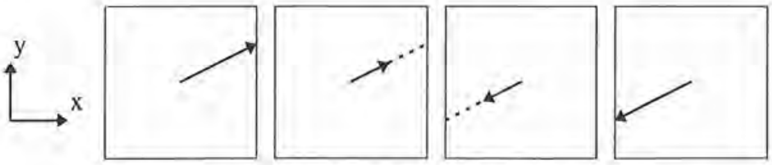
Figure 6.2: Four-level polarisation rotation given an input polarisation and four CSLC states.

The four full-length arrows denote the long optic axes of the four differently switched pixels and their corresponding output states for a given TEM_{00} input polarisation (dotted line) are plotted in dash-dot lines. All four little arrows on the output polarisation loci are the corresponding polarisation vectors at a given time exiting their respective pixels. It is the coherent interference of these four vector fields that gives rise to a hologram replay. For a given distribution of these four vectors, the instantaneous states of all vectors will produce a replay field and *the incoherent sum of the replay over one optical cycle gives the apparent replay power.*

Phase modulation by polarisation rotation produces a 1D or 2D array of polarisation vectors that point to a range of directions on the transverse plane. All instantaneous vectors are equal in length at any stage of the optical cycle if the wave retardation is exactly π, i.e., similar polarisation shape for any switch state. Whether all the vectors are of equal length or not, they all interfere to produce a replay field. Each vector can be resolved into its constituent x- and y- components whose replay add up incoherently. Therefore the proportion of replay power being deflected to the ± orders would depend on the input polarisation.



CSLC: Phase modulation by polarisation rotation



DOE: Phase modulation by optical path length change

Figure 6.3: Four-level phase modulation using optically active and isotropic media.

In contrast to phase modulation by passing the light through an optically active medium, DOEs produce different path lengths and hence the incoming polarisation is irrelevant (within the paraxial regime). The interference effect of the array of polarisation vectors all pointing in the same direction but with different lengths (real part of the fields) results in a replay without any inverse orders. This crucial distinction is shown in Figure 6.3 where only the real parts of the instantaneous fields are shown for four pixels.

Determining the dependence of the switching angle and the wave plate retardation on the DC power gives an indication of the remaining useful power fraction for diffraction. Although the input polarisation state does not affect the DC level, it does determine the distribution of the diffracted power between the positive and the inverse hologram replay components. The effects of E_x, E_y, α and φ on the diffracted orders are coupled. For a polarisation rotation device, an input beam is best decomposed into its constituent clockwise and anticlockwise circular components [89]. The input polarisation parameters determine its ellipticity and hence the distribution of ± modulation whereas the non-π retardance term φ simply results in additional loss without changing the ratio of the ± modulation terms. The replay intensity expression for a positive n^{th} order is given by:

$$AC^2_{\varphi;m;n} = E_{n+}\left(E_x, E_y, \alpha\right) \times \sin^2\left(\frac{\varphi}{2}\right) \times \frac{sinc^2\left(\frac{m\theta_o}{m-1} - n\pi\right)}{sinc^2\left(\left(\frac{m\theta_o}{m-1} - n\pi\right)\Big/m\right)}, \tag{6.5}$$

where $E_{n+}(E_x, E_y, \alpha)$ is the ellipticity term for positive n^{th} orders. It can vary from 0 to 1 for purely clockwise to anticlockwise circularly polarised light illuminating the hologram, designed to produce an asymmetric replay pattern. The corresponding ellipticity term for negative n^{th} orders is $(1 - E_{n+}(E_x, E_y, \alpha))$. Other than the obvious symmetric replay using binary phase, certain applications such as a Fresnel lens also require symmetric hologram encoding. In this case, multi-level phase can be more efficient as some higher orders can be eliminated. In addition, for non-telecommunications applications such as correlation, reliable circularly polarised light sources can be obtained and an asymmetric replay is a plus point.

6.2.2 Binary modulation

The case of binary phase modulation using FLCs has been demonstrated without any control of the incident polarisation and retardation of the electro-optic medium [90]. The two polarisation outputs for any arbitrary input, FLC switching angle and thickness are shown in Figure 6.4. *Unless it is a π retardation wave plate, the output polarisation shapes cannot be maintained and are different for the two states.* This means that the instantaneous lengths and angular orientations for the two vectors are different and varying through a complete optical cycle. The temporal ratio of DC to AC power fractions changes accordingly. *However, unless the input polarisation shape changes very rapidly and always results in the worst case DC loss, the replay should contain the average AC and DC over any complete polarisation cycle.* The DC loss is due to both the lack of switching angle ($\theta_o \neq \pi/2$) and the non-ideal retardation ($\varphi \neq \pi$) such that,

$$DC^2_{\varphi;m=2} = \cos^2\left(\frac{\varphi}{2}\right)+\sin^2\left(\frac{\varphi}{2}\right)\times\cos^2(\theta_o) \text{ or equivalently } \cos^2(\theta_o)+\sin^2(\theta_o)\times\cos^2\left(\frac{\varphi}{2}\right). \tag{6.6}$$

A binary hologram is inherently symmetric hence it produces equal ± orders regardless of the input polarisation shape. The theoretical ±1 order efficiency for a normalised replay fraction $\sigma = x/x_o$ before considering dead-space and pixellation loss is thus,

$$AC^2_{\varphi;m=2;n=\pm1} \propto sinc^2\left(\frac{\pi}{2}\right)\times\sin^2\left(\frac{\varphi}{2}\right)\times\sin^2(\theta_o). \tag{6.7}$$

The corresponding expression for a four-level polarisation rotation system, assuming a *pure circularly-polarised* input beam is,

$$AC^2_{\varphi;m=2;n=1} \propto sinc^2\left(\frac{\pi}{4}\right)\times\sin^2\left(\frac{\varphi}{2}\right)\times\sin^2\left(\frac{2}{3}\theta_o\right)\cos^2\left(\frac{\theta_o}{3}-\frac{\pi}{4}\right). \tag{6.8}$$

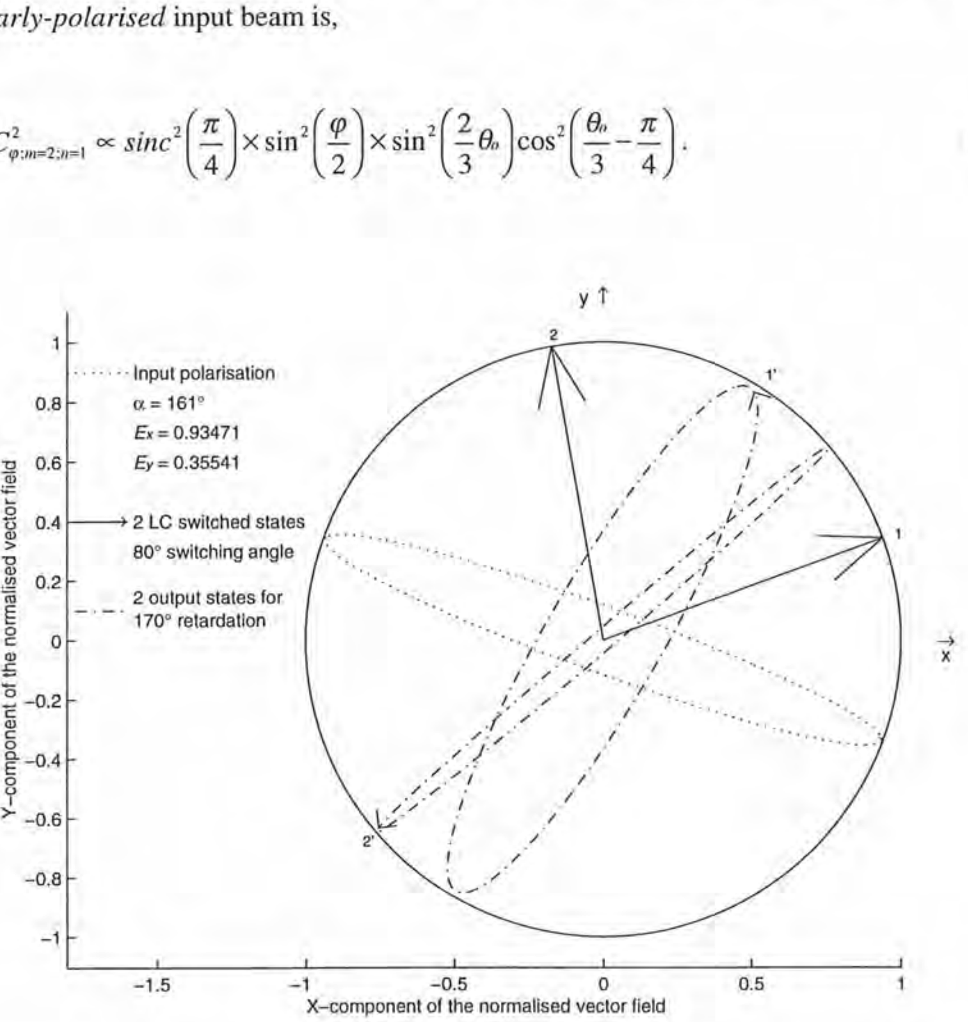


Figure 6.4: Binary polarisation rotation given an input polarisation and two FLC states.

In DOEs the phase matching condition can be made near perfect and hence it is commonly perceived that a four-level phase hologram gives twice the first order efficiency of the corresponding binary hologram. However a reconfigurable four-level phase hologram made from CSLC is restricted by the switching angle. A comparison of the respective first order efficiencies of binary and quaternary CSLC holograms (for π retardation) is shown in Figure 6.5 [88]. Intuitively, the ideal switching angles for the binary and quaternary systems are 90° and 135°, respectively. Unless the CSLC switching angle is > 67.5°, there is no benefit in increasing the number of phase levels from two to four. In typical cases, CSLC switching angles are < 90°. Hence, the binary replay is the favoured type based on insertion loss criterion alone. In addition, its insensitivity [91] to input polarisation is a clear advantage for telecommunications applications. All that remains is whether a four-level system will give better crosstalk isolation figures.

Assuming that high SBWP SLMs are easily available, an optimum choice of fibre arrays avoids the coupling of the -1 replay orders. This eliminates the SNR advantage of using quaternary holograms. The use of fibres for spatial filtering is effective and attractive. In any case, optimum fibre placing has to be implemented even for ideal 135° switching CSLC quaternary holograms with polarisation control, because the slightest imperfections in any of the four states or input polarisation will give rise to cross-modulation peaks.

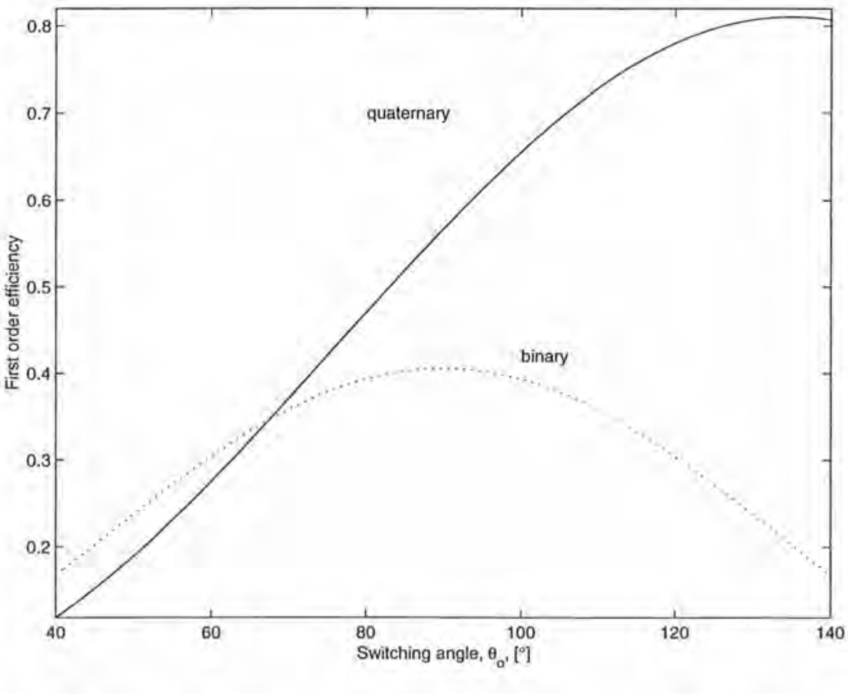


Figure 6.5: The first order efficiencies of binary and quaternary phase-only holograms.

6.3 Fast four-level phase-only modulation by polarisation rotation

Continuous phase modulation of up to ~ 360° using a single pixel has been shown [92], however, a moderate number of phase levels is most attractive and presents the least technical hurdles to realise as an array of modulating elements. Even though SLMs capable of providing four-level modulation are not useful for fibre-to-fibre routing/switching, there exists applications such as correlation and symmetric multi-level encoding where a four-level system might be desirable.

6.3.1 Analysis of enhanced switching using a double-pass configuration

An obvious way to overcome the lack of switching angle is to cascade two or more stages [93], however, pixel-to-pixel alignment problems are a major restriction. Double-pass LC cell configurations with or without a resonant cavity have been widely used to enhance intensity and phase modulation [94]. The challenge is to design a silicon backplane that is suitable for double passes in an integrated manner, as opposed to a single-pixel sandwiched construction.

The double-pass configuration contains a fixed quarter-wave retarder adjacent to a mirror. This provides a symmetrical fold-back optical system as shown in Figure 6.6. Using the matrices in eq. (6.1) with $\varphi = \pi/2$ and the mirror matrix, $\mathbf{R}_m$, the output field becomes,

$$\underline{E}_{out} = [\mathbf{R}_m]\times\left[[\mathbf{J}_{\pi/2}][\mathbf{R}_{-\theta}][\mathbf{J}_{\pi}][\mathbf{R}_{\theta}]\right]^{t}\times\left[[\mathbf{J}_{\pi/2}][\mathbf{R}_{-\theta}][\mathbf{J}_{\pi}][\mathbf{R}_{\theta}]\right]\times[\underline{E}_{in}] \text{ where } \mathbf{R}_m = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (6.9)$$

The superscript t indicates the ordinary matrix transpose operation. The transpose gives the Jones matrix under reverse propagation [95]. $\mathbf{R}_m$ is the co-ordinate axis reversal term in order to maintain a right-handed system, which in this case has been chosen to be the x-axis. Eq. (6.9) simplifies to yield,

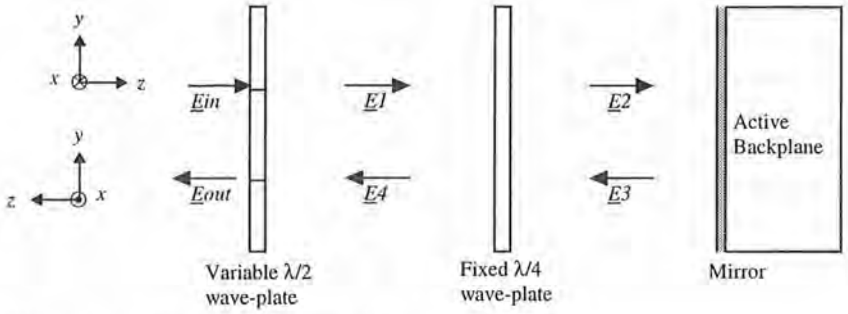
$$\underline{E}_{out} = \begin{bmatrix} \cos(4\theta) & \sin(4\theta) \\ -\sin(4\theta) & \cos(4\theta) \end{bmatrix}\times[\underline{E}_{in}]. \quad (6.10)$$

With circular input polarisation, $\underline{E}_{in} = \begin{bmatrix} 1 \\ \pm j \end{bmatrix}$, the output polarisation vectors are,

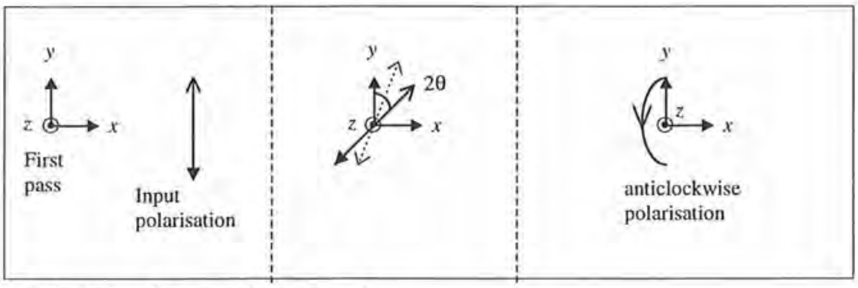
$$\underline{E}_{out} = \exp[\pm j4\theta]\times\begin{bmatrix} 1 \\ \pm j \end{bmatrix}. \quad (6.11)$$

In order to obtain four equally spaced phase modulation values, θ has to be 67.5°. This requirement is well within the capability of DHF LC mixtures. However, polarisation rotation using DHF is associated with a voltage dependent modulation of the average effective refractive indices, which causes deviations from the ideal λ/2 wave plate configuration. It has been shown that an additional loss is incurred if the λ/2 retardation wave plate in the double pass system departs from the ideal condition [92].

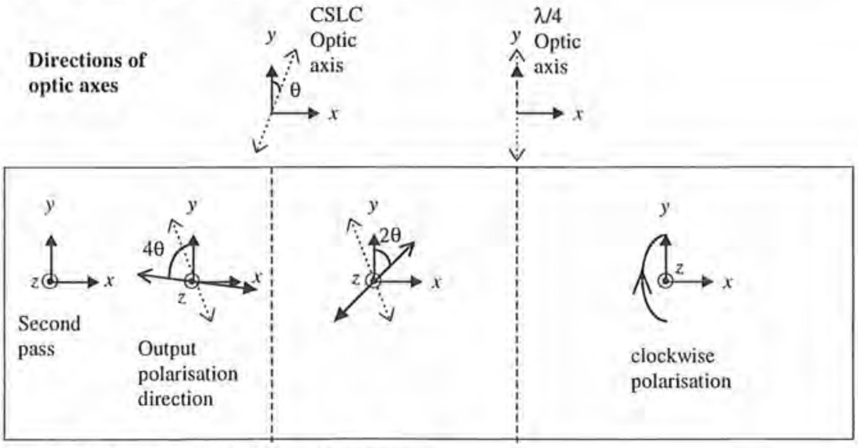
The analysis above is valid for any relative orientation of the instantaneous input polarisation, the variable electro-optic switched states and the optic axis of the fixed λ/4 wave plate. The input polarisation has to be considered in terms of the two orthogonal clockwise and anticlockwise circularly polarised modes, in which case, absolute orientation differences do not matter.



Cross-sectional view of a double-pass device



Polarisation diagrams of the forward pass



Polarisation diagrams of the reverse pass

Figure 6.6: Double-pass holographic optical element structure.

6.3.2 Double-pass devices

In Figure 6.6, it is assumed that the pixel array configured as a λ/2 retarder is driven via the fixed λ/4 wave plate. The voltage drop due to the potential divider between fixed and variable electro-optic media is acceptable provided that the thickness of the fixed λ/4 wave retarder does not stop the switching of the CSLC-filled pixels. A technology to produce these thin-film optical elements is based on polymerising and fusing reactive monomers [96]. Liquid crystals mixtures containing siloxane side chains can be polymerised into plastic-like wave plates. This device construction does not require vias to connect to the modulating pads at the expense of reduced effective electric fields across the CSLC cells. A cross-sectional view of this construction is shown in Figure 6.7 assuming $\lambda = 633$ nm and $\Delta n = 0.1$.

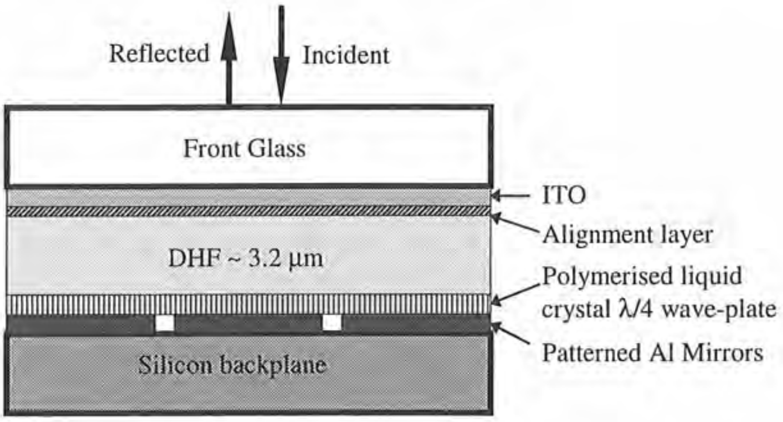


Figure 6.7: A silicon backplane SLM integrated with a thin-film λ/4 wave plate.

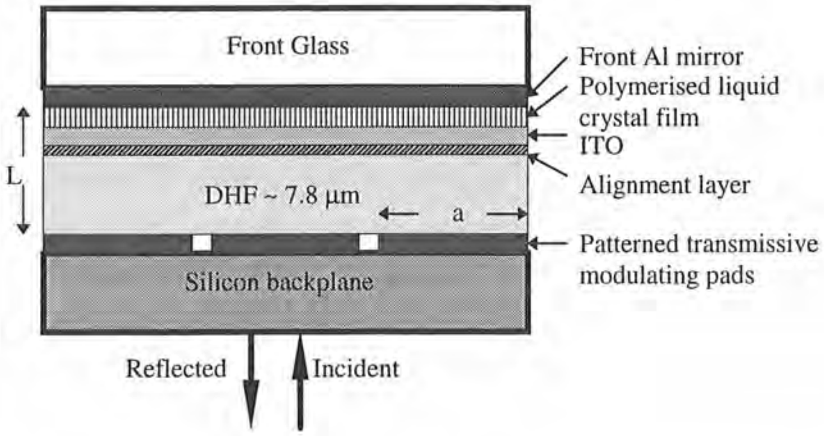


Figure 6.8: A double-pass silicon backplane SLM with fused reactive monomers on glass.

The silicon substrate is transparent to wavelengths > 1 μm, although there are high Fresnel reflection losses at the Si/air interface due to severe refractive index mismatch. If the double-pass quaternary SLM is to be configured for the communications wavelength windows of 1.3 and 1.55 μm, transmissive operation through the silicon backplane can be realised. In this case, the metal modulating pads on the silicon backplane must be replaced by IR transmissive media such as highly-doped diffusion regions or polysilicon electrodes. The thin-film λ/4 wave plate can be coated onto the front glass above a layer of metal reflector. The mirror layer serves as the front electrode contact. If the insulating effect of the λ/4 wave plate prevents LC switching, an electrode (ITO) can be deposited on top of the glass/mirror/wave plate construction as shown in Figure 6.8. The input beam illuminates the back of the silicon substrate, passes the variable electro-optic medium, the fixed λ/4 wave plate, the switch-able retarder again and then exits the SLM having performed a 4θ polarisation rotation.

Another way to realise the double-pass cell is to use a rigid solid crystal (e.g., quartz, mica or plastic) wave plate as the λ/4 retarder and the front glass of the CSLC cells as depicted in Figure 6.9. This device construction is no different from a conventional LC/Si SLM cell, except that the front glass is replaced by an 'optically active' glass coated with a metal layer on the outer facet.

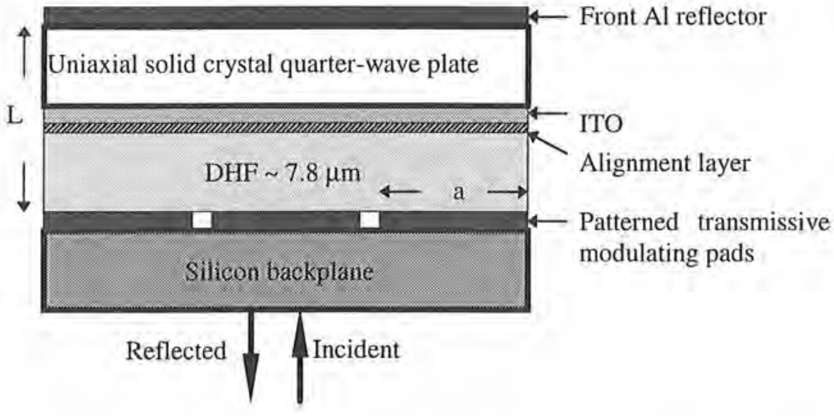


Figure 6.9: A double-pass silicon backplane SLM with a rigid solid crystal λ/4 wave plate.

One possible difficulty of this cell construction lies in the thickness of the λ/4 wave plate. In a double-pass system, the optical field modulated by a backplane pixel in the first pass must be incident on the same pixel on its return path, as shown in Figure 6.10. The largest far-field diffraction angle after the first pass for a given pixel pitch, d, is $\theta \approx \lambda/2d$. In order that the return pass hits the same pixel with an aperture, a, the dept, L, of the λ/4 wave plate should be $\leq da/\lambda$. For an SLM with a 20 μm pixel pitch and a 2 μm pixel gap operating at 1.55 μm, the maximum thickness for the λ/4 wave plate is 230 μm. In practice, the short distance ensures that near-field diffraction dominates and the thickness requirement of the λ/4 wave plate is probably not as stringent.

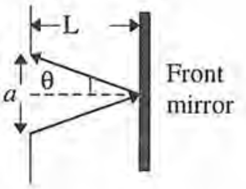


Figure 6.10: Approximation of the diffraction within a single pixel aperture.

6.4 Processing digital data for analogue devices

Analogue electro-optic modulation requires grey-scale voltages to be applied at each pixel. For nearly all processing and storage purposes, a digital format is preferred. Therefore a frame of hologram data, calculated off-line or on the fly, has to be presented to digital-to-analogue converters (DACs). The location of the D/A conversion as well as the conversion technique is heavily influenced by the intended application of the SLM. The SLMs considered here are those based on very-large-scale-integrated (VLSI) circuits realised using complementary-metal-oxide-silicon (CMOS) semiconductor processing and assembled with a LC electro-optic layer.

Conventional DACs use R-2R thin film resistor networks to weight different bits of a digital word. This is both spatially bulky and incompatible with CMOS transistor technology. With CMOS VLSI, the switched-capacitor approach is preferred due to the higher conversion accuracy. There are also DACs that function by summing mirrored current sources. The size factor has made this technique attractive [97], however, it requires a current-to-voltage converter. This is too large to be implemented on the pixel real estate of a silicon backplane SLM.

Another approach is to digitally derive the analogue modulation [98]. The essence of this scheme is to sum the AND logic products of various binary bit-weighted pulse densities with the input data bits. This technique is meant for RMS responding nematic liquid crystals. The DACs located at pixel-level simply modify the effective frequency of pulses applied to the pixels. Two disadvantages arise if this design is adopted for a FLC/Si SLM. First it requires several separate pulse densities and data electrodes to each pixel. Secondly, RMS weighting of the digital input does not work for smectic liquid crystals. Smectic liquid crystals are fast switching DC-responding mixtures.

In the following sub-sections, three strategies regarding the location and type of DACs are explored and their suitability for a holographic chip explained.

6.4.1 Global DACs

In conventional active-matrix twisted nematic (AM-TN) or super-twist nematic (STN) LCDs, the data for a display are read from a frame store and converted to analogue voltages by means of one or several parallel DACs. The analogue voltages are latched into the LCD driver by sample-hold circuitry and fed to the switching/storage elements of the pixel array column-by-column and row-by-row (in a sequential or interlaced manner). A general layout of the high-level functionality blocks is shown in Figure 6.11 [113].

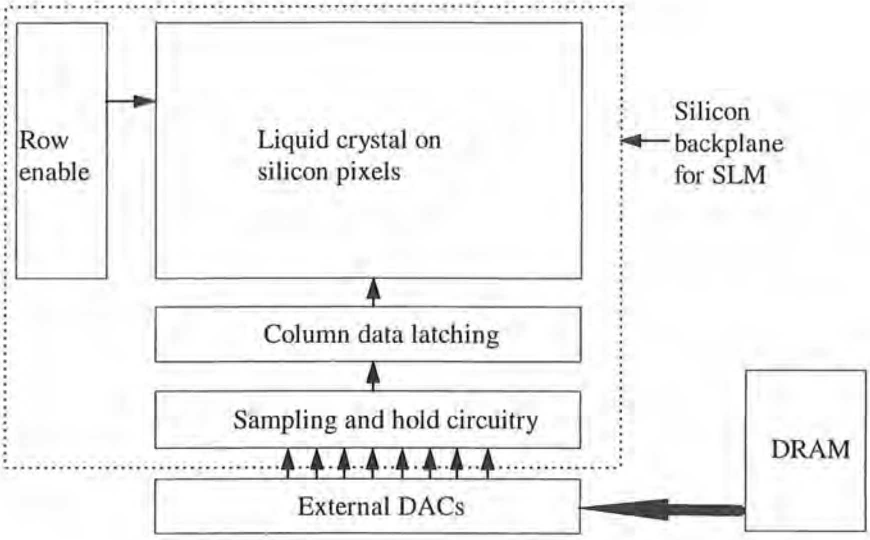


Figure 6.11: A general layout of an SLM driven from global DACs.

This arrangement is adequate for the low switching speed of polysilicon or amorphous silicon thin-film-transistors (TFT) memory cells on glass. Also, the TFT transistor technology is unable to realise high complexity electronic functionality. At video frame rates, the determining step in these systems is the LC switching time. Hence, several global DACs with a short settling time (e.g., 6 ns for an 8-bit DAC) will achieve < 32 μs video line rate.

For silicon backplane SLMs, an opportunity arises for the inclusion of D/A conversion functionality on chip and addressing the pixel array using a fast scan scheme. It is generally not possible to implement full grey-scale (3 colours × 8 bits) conversion on the SLM chip due to silicon real-estate constraints. For micro-display applications, analogue data will be generated off chip, even for high frame rate 3D displays with angular multiplexing (if true analogue driving rather than spatial/temporal dithering is preferred). SLMs fabricated on non-CMOS/VLSI chips will require the conversion functionality to be implemented on silicon electronics [33] using specialised D/A chips. For holographic applications, several phase levels are required at high frame rates. Hence, using external DACs is not appropriate.

6.4.2 Column-select DACs

If the number of digital bits is traded off for a smaller DAC, it is possible to integrate the converters at a column level. With four grey levels the two-bit inputs are used to select one of V_0, V_1, V_2 or V_3 voltages to be latched to the DRAM at each pixel. The low number of grey scales makes this design less useful as a display or amplitude modulator. However, it is sufficient for multi-level holograms. V_0 and V_3 are the lowest and the highest backplane voltages, (e.g., ground and the supply voltage to the transistor drains, VDD). V_1 and V_2 have to be calibrated experimentally to compensate for the non-linear switching of the electro-optic medium.

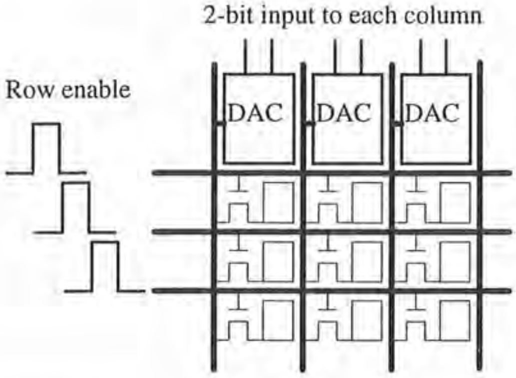


Figure 6.12: Column-select D/A conversion for CMOS/VLSI SLMs.

The layout for a silicon backplane with column-select D/A conversion is shown in Figure 6.12. The row-enable circuitry is simply a shift register. The column data latches are digital demultiplexers and static memory cells. The conversion of the data for an entire row (e.g., 512 columns) takes place in parallel. The speed advantage is considerable, in addition to the simple circuitry for digital data input (sample and hold circuit not required).

The column-select DAC can be realised in many ways with different transistor counts. One way to implement the column-select is by employing address-decoder type structure as shown in Figure 6.13. D1 and D0 data inputs determine which of the NAND gates is activated to connect the column electrode to one of the supply rails via a switching transistor. The voltage is then reflected at the pixel mirror via a DRAM transistor at the enabled row.

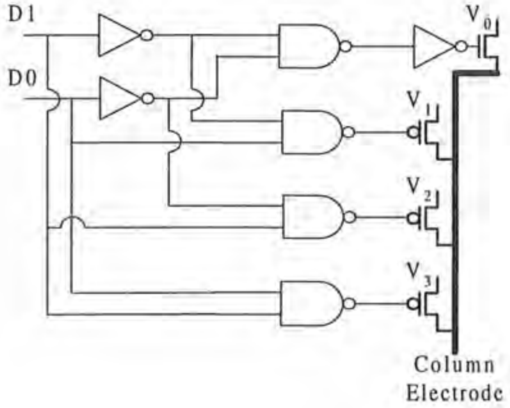


Figure 6.13: Combinational logic circuits for a column-select DAC.

6.4.3 Pixel-level DACs[*]

The main drawbacks of both the global and column-select D/A conversion are the voltage droop after LC switching and the inadequate charge stored to switch LC with a high spontaneous polarisation due to the use of DRAM storage at the pixels. The DRAM transistor is typically made of n-type metal-oxide-silicon field effect transistor (MOSFET) for a higher carrier mobility. When the polysilicon gate is switched on by a positive pulse, a conducting channel is formed linking the drain to the source of the MOSFET as depicted in Figure 6.14. In order to increase the storage capacitance at each pixel, an additional diffusion area is included in parallel with the metal pad.

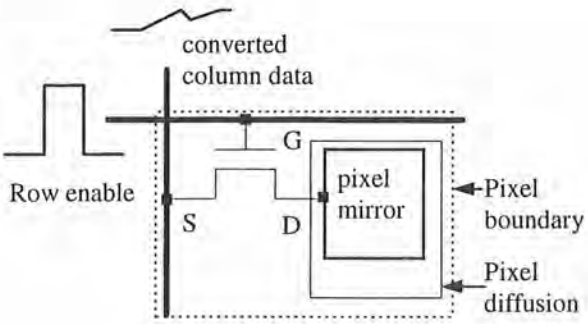


Figure 6.14: Pixel layout using a single DRAM transistor.

When pulling the drain of the transistor to VDD, capacitive charging of the modulating pad (C_{pad}) and liquid crystal dielectric (C_{LCd}) media occurs via the resistive path of the driving

---

[*] The analysis of pixel-level storage circuitry is applicable for both binary and multi-level SLMs.

stage and the transistor channel as shown in Figure 6.15. The drains of the DRAMs are isolated from the supply sources once the addressing is complete. The mechanism of LC switching results in a minimum energy state. The sharing of the charges initially stored on the pixel dielectric capacitance (C_{pad} & C_{LCd}) with the 'charge compensation' capacit‸-ance associated with LC switching, C_{LCs} gives rise to a charge equilibrium. Typically, the front electrode voltage (VFE) is fixed at either GND, VDD/2 or VDD for a given period of pixel capacitance charging and LC switching.

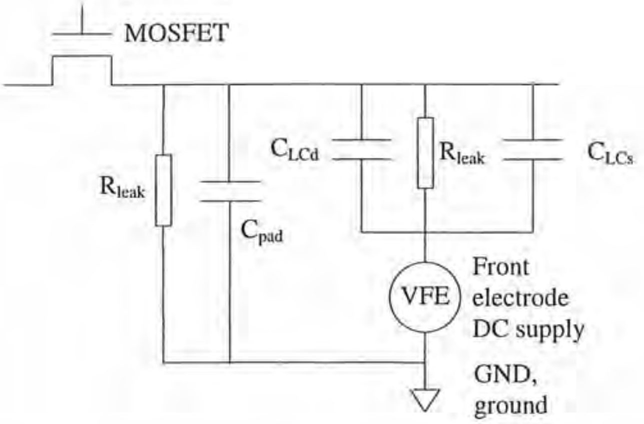


Figure 6.15: Equivalent circuit components of a LC/DRAM pixel.

The voltage droop and inadequate switching cause perturbation in the LC switched state, which in turn appears as ripples in the optical link. One way to reduce this effect is to build static-random-access-memory (SRAM) cells at pixels. A 1-bit SRAM requires six transistors (Figure 6.16). For a moderate increase in transistor count (20 transistors), a 2-bit SRAM cell with supply-select combination logic circuits can be implemented at each pixel [99].

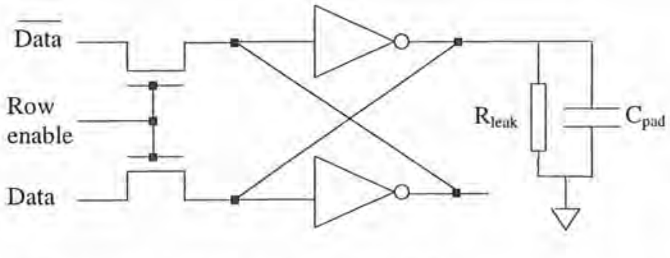


Figure 6.16: Six-transistor SRAM memory element.

If a static memory cell is used to store the state of a pixel, the charging of dielectric capacitance and switching of the LC can draw current from the supply rails continuously. The current dynamics of this process, assuming small signal equivalent circuit, consist of a sharp current peak over several ns (charging the dielectric media) and a broader current pulse over tens of μs (slower compensation of the LC switching/rotation). The equivalent relative permittivity (ε_r) of electro-optic layer is ~ 5. However, the switching dynamics are dominated by the viscosity parameter of the LC mixture; a simple RC time constant is not applicable.

6.4.4 Choice of DAC location for the demonstrator chip

The multi-level pixel array of the demonstrator chip is meant for displaying holograms. Thus, a modest scale of four grey levels is adequate. Given the low number of grey levels intended, the data conversion bottleneck can be eliminated readily by building the DACs on chip. For the first demonstration of multi-level phase modulation, only a linear array is aimed for. Hence, both column-select and pixel-level DACs are equivalent and can be hidden out of view of the optical illumination.

6.5 Pixel design for coherent optical phase modulation

The thrust of designing a silicon backplane for assembly into holographic SLMs lies in obtaining high quality optical modulation. Therefore, the optical interface (i.e., pixels) has to be the flattest possible and its construction has to enable other intricate SLM assembly processes. The pixel designs for the binary and quaternary arrays are discussed separately.

6.5.1 Enhanced optical reflectors for the binary SLM

In seeking $\lambda/10$ surface roughness for coherent optical applications, the VLSI metal layers must not be used for direct optical interface. The sintering/annealing processes at the silicon foundry in order to produce high quality electrical interconnects result in hillocks [100]. Silicon processing utilises conformal depositions. Hence, if the underlying surface topology is flat, subsequent layer depositions could return ~ 200 Å RMS roughness [114].

With regard to designing the backplane for high optical quality 1D pixel arrays, the task has been made simpler by not having to planarise the driver circuit areas. These drivers could be hidden out of view of optical illumination. The choices of pixel design are between transmissive pixels using doped regions or reflective pixels to be deposited at low temperature after the silicon fabrication. The former is ruled out because of the difficulty in strapping the diffusions with substrate or well taps, back reflections at air/Si interface and half the operating electrical field in transmission mode. In addition, exposing the electrodes by removing the oxides/nitride insulation/protective layers could prove difficult.

The reduced field is especially critical considering that the binary devices are to be used at 1.5 μm IR. At this wavelength window, the combined effect of reduced LC birefringence, Δn

and longer wavelength requires the LC cell to be ~ 3× thicker than that for visible illumination. Operating the SLM in reflection mode would halve this thickness penalty. Thus the SLM pixels have to be made of metal reflectors. The pixel array area should be left with only bare-silicon and oxide/nitride layers.

6.5.2 Transmissive pixels for quaternary SLM

The quaternary pixel design has limited choices. A transmissive operation must be enabled, in order that any double-switching SLM construction can be attempted. There is no hope of altering the silicon fabrication sequence or processes. Hence, any insulating layer protecting the transmissive modulating layer must be removed to obtain the full voltage drop. The two intrinsic silicon fabrication layers which are transmissive at 1.3 and 1.5 μm IR are diffusion and polysilicon. Diffusion has to be strapped to the Si-substrate or well to minimise latch-up risks. Polysilicon electrodes can be used as electrical interconnects with its grain size (400 Å) posing no problems for coherent optics. The details of exposing the polysilicon electrodes, which comes before the doping of diffusions are given in §7.8.2.

6.6 Semiconductor fabrication process

There are numerous design options to implement digital CMOS circuits. The trade-offs are between the design/fabrication-time and the design-complexity/layout-control. Automated design methodologies using PLDs[*] and FPGAs[†] (Programmable Logic Devices and Field Programmable Gate Arrays) drastically reduce the design/fabrication time but they do not allow the users to manipulate the layout and access the voltage modulators. Hence, a *full-custom CMOS mask design* is the only option for an SLM backplane where the size and location of the optical interface has to be optimised. The ASIC (Application Specific Integrated Circuit) chip requires a fabrication process at a silicon foundry.

Although numerous silicon foundries (e.g., AMS [101], Alcatel Mietec [102], Newport [103], TSMC [104], etc.) offer independent engineering runs, Austria Mikro Systeme (AMS) was chosen for it has suitable processes to meet the requirements of ROSES[‡] devices and it is familiar with the application of such opto-electronic chips having previously processed two other batches of devices. Typically there are complementary-metal-oxide-silicon (CMOS), dual-diffusion MOS (DMOS) and bipolar-CMOS (BiCMOS) processes to realise VLSI

---

[*] Such as PALs® (Programmable Array Logic) from AMD (Advance Micro Devices, Inc.).

[†] Such as XILINX Configurable Logic Block® (CLB).

[‡] The ROSES project under which these devices were developed is described in §7.

circuits. BiCMOS processes enable both NPN and PNP bipolar and CMOS transistors to be integrated on a common substrate. It was decided to avoid the extra complexity of a BiCMOS process. As the proposed ROSES binary driver circuits were required to implement only logical functions without static current flow, they could be implemented using simpler CMOS or DMOS technologies. The quaternary drivers have an additional digital voltage-select to perform the D/A conversion. Hence, only CMOS and DMOS processes are suitable for ROSES chip. In the following the two processes are compared.

6.6.1 CBH 10V 2-μm CMOS process

CBH is a high voltage variant of a conventional CMOS process. Its positive supply voltage can be as high as 11V. This p-substrate process uses two metal interconnects and two polycrystalline silicon (polysilicon) layers to obtain thin-film sandwiched capacitors. The 2-μm feature size enables moderate integration of logic drivers without resorting to costly low dimension processes with stringent design rules. Input/output signals and internal logic manipulation can be executed at the standard 5V level whereas the final driving of pixel electrodes can be shifted to a higher level of 11V.

Other than the high integration density and speed, CMOS circuits are preferred for low DC power dissipation. N-channel and p-channel MOSFETs or *n-fet* and *p-fet* (also p-FET and n-FET) transistors typically operate as enhancement devices where there is no conduction between the *drain* and *source* diffusions at zero *gate*-source voltage (Vgs = 0). When the transistor gate bias is greater than the threshold voltage (Vtn (+ve) and Vtp (-ve) for n-fets and p-fets respectively), a conducting channel is obtained by the inversion of the semiconductor type beneath the gate. These devices operate in three regions:

a) *cut-off* (sub-threshold) where drain-source current, Ids ≈ 0.

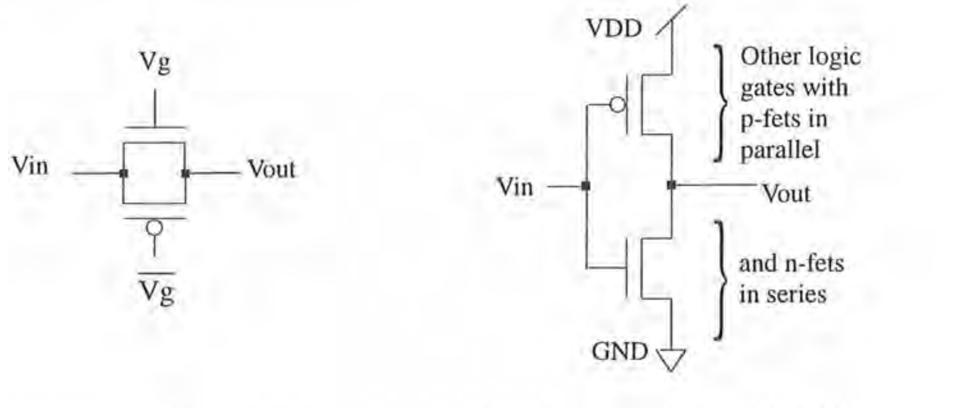
b) *linear* (non-saturation) where Ids is dependent on Vgs and drain-source voltage, Vds, and

c) *pinch-off* (saturation) where the inverted channel is terminated before the drain edge; Ids is nearly constant for a fixed Vgs.

Digital logic circuits rely on MOS transistors being in the cut-off or linear regions. In the following, five useful digital circuit elements for building SLM drivers are described.

i) CMOS switch: n-fet and p-fet transistor pass gate (transmission gate)

A MOS transistor will remain in conduction provided |Vgs| > |Vt(n,p)|. Hence an n-fet cannot pass logic 1 signals with full pull-up (gate voltage, Vg = Vin = 5V, Vout = Vin-Vtn) whereas a p-fet cannot pass logic 0 signals with full pull-down (Vg = Vin = 0V, Vout = |Vtp|). By connecting an n-fet and a p-fet transistors in parallel as shown in Figure 6.17(a), a transmission gate capable of passing both logic levels without any degradation is obtained. The switch requires complementary gate signals. It is useful in 'opening' and 'shutting' signal propagating paths at various junctures.



a) CMOS switch (transmission gate) b) CMOS inverter and other 'NAND' logic gates

Figure 6.17: Transistor-level schematics of a complementary n-fet/p-fet switch (transmission gate) and a CMOS inverter.

ii) CMOS inverter and logic gates

Logic circuits including the most common Boolean logic gates, uni-input inverter and dual-input AND2, OR2, NOR2, XOR2 and XNOR2 can be realised by combinations of n-fets and p-fets: either n-fets in series and p-fets in parallel (NAND-design) or p-fets in series and n-fets in parallel (NOR-design). NAND-design is faster due to the use of higher carrier mobility n-fets in series. An example of a CMOS inverter is shown in Figure 6.17 (b), with possible extension to a multiple-input NAND logic gate.

iii) Dynamic shift register

A shift register latches in logic signals and ripples them through its memory units in response to clock signals. A straight forward dynamic shift register uses CMOS inverters and transmission gates to control and regenerate the logic signals. A single-bit dynamic shift-register is shown in Figure 6.18. The two non-overlapping clock signals, CK0 and CK1, control the first and second transmission gates, respectively. There is no through path from the input to the output at any time. When the transmission gate before each inverter is enabled, the inverter input is driven by the output stage of the previous inverter. When the path through the transmission gate is in high impedance, the input to the following inverter is sustained by the charges on the drain capacitor. Because of this, the dynamic shift register cannot operate below certain clock rates at which the capacitive charges would have decayed before the next shift.

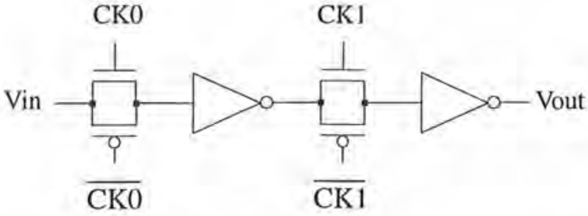


Figure 6.18: 1-bit dynamic shift register with two non-overlapping clock signals.

iv) Static shift register

When the dynamic shift register unit is added with positive feedback by means of cross-coupled inverters, the internal logic state is held static. A transmission gate in the feedback path isolates input stage of the static memory unit when it is being updated.

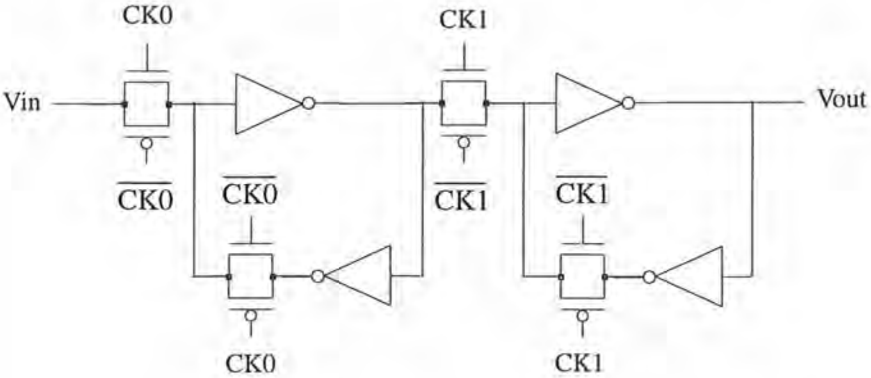


Figure 6.19: 1-bit static shift register using positive feedback.

v) Level shifting of logic signals

A straight forward level shifter consisting of two cross-coupled p-fets and two n-fets requiring complementary inputs arranged in two parallel branches is as shown in Figure 6.20. Regardless of whether the input is 0 or VDD (drain power supply), one of the two n-fets will be cut-off (turned off) whereas the other will be conductive (turned on). The conductive nfet has its drain pull ed-down to 0 V, thus turning on the p-fet in the opposite branch which has its drain pulled-up to VDDH (the shifted supply). The p-fet in the same branch as the conducting n-fet is cut-off because it has both its drain and gate voltages at VDDH. Therefore, the diagonal p-fet and n-fet remain either on or off together. There is no DC path from VDDH to ground at steady states. The classic CMOS push-pull effect is being utilised here. Power dissipation is only incurred when the logic input changes.

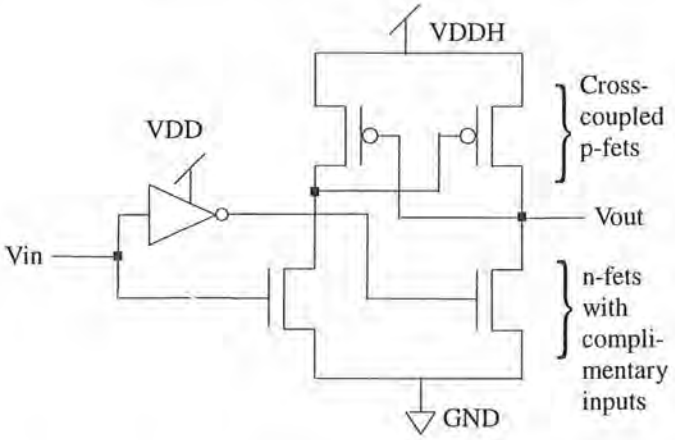


Figure 6.20: Level shifting from [0,VDD] to [0,VDDH] using cross-coupled p-fets.

Using the five digital circuit elements and CBH 10V process, complete VLSI drivers can be implemented on a silicon backplane. The core functionality of data latching and storing is fabricated on the standard 5V logic circuits whereas the actual LC driving voltage is stepped up to 11V. The availability of ±5.5 V (assuming the front electrode of the SLM is fixed at mid rail) is adequate for a wide range of LC mixtures and applications. However, there are two severe voltage limitations for holographic SLMs:

i) At 1.5 μm IR, the combined effect of longer wavelength and decreased LC birefringence, Δn at IR requires 3× thicker LC cell and

ii) Most high-tilt LC mixtures (e.g., Chisso CS2005) are slow at moderate fields.

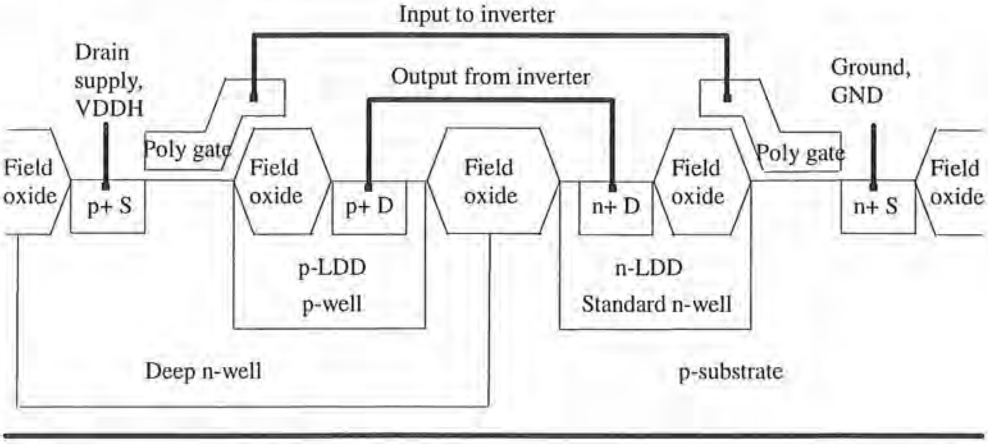
These can be overcome by making higher fields available as examined in §6.6.2. Once a LC cell has been switched, there should be little current drawn at static state. Therefore, very high voltage driving will not give rise to power dissipation problems.

6.6.2 CBY 50V 2-μm DMOS process

A DMOS process uses lightly-doped-drain (LDD) transistors to prevent *drain punch-through* at high bias fields. An example of the LDD structure used in asymmetric n-fet/p-fet inverter is shown in Figure 6.21. LDD protection has two main features:

i) Hot-electron degradation: For a fixed channel length, the electric field at the drain of an n-fet in saturation can be so high that electrons are accelerated with enough energy to dislodge holes in the *drain*. These are then swept to the negatively biased substrate and appear as a substrate current [105]. This effect can change transistor parameters such as threshold voltage, sub-threshold current, etc. and is termed *hot electron degradation*. Having an LDD diffusion and a spacer oxide inevitably lengthens the channel and presents a non-straight-line path for the drifting electrons. This structure slows the electron drift velocity and prevents hot electron effects. The LDD function is particularly important for n-channel MOSFETs where the mobility of electrons is ~ 2× the hole mobility of p-channel MOSFETs. Hot hole degradation is normally not a problem.

ii) Drain punch-through: Regardless of the steady input state, one of the two transistors making up the inverter has a reverse-biased *drain-bulk* by the full VDDH drop. That itself can cause *drain punch-through* by extending the *p-n* junction depletion into the source. Hence, the channel conduction takes place irrespective of the gate bias. The decreased LDD doping density increases the junction breakdown voltage. The maximum permissible VDDH value is primarily a function of the difference between LDD, the bulk impurity densities and the polysilicon gate biasing. Any slight inversion in the channel due to Vgs reduces the breakdown voltage significantly. That is why the drain of p-fets needs LDD protection even though the hot-hole effect is not common.



S & D: source and drain of FETs; n+ & p+: highly-doped diffusions; n- & p-: lightly-doped diffusions

Figure 6.21: A CBY DMOS inverter with asymmetric LDD transistors.

LDD structures will also be required for the *sources* (symmetric transistors) if they are not tied to the well or substrate voltages (or Vsb > 10V). The layout of a CBZ DMOS inverter was drawn and found to be ~ 4× larger than the corresponding inverter using CBH design rules. Although all logic operations of the chip will be performed using standard CMOS transistors, using DMOS transistors merely for the level-shifting and buffering would have increased the driver size substantially.

Two important parameters of MOSFETs are the drain-source voltage, Vds and the gate-source voltage, Vgs. For the CBH process, Vds must not exceed 12 V for both n-fets and p-fets. Clearly, VDDH must not be scaled up without protective structures at the drains. The straight forward level shifting seems to work for shifting standard 5V logic signals to 45V given that the maximum |Vds| for high voltage n-fet and p-fet of CBY process are 50 and 45V, respectively. If the four transistors of the level shifter are implemented as DMOS, the gate bulk (i.e., n-well or p-substrate) voltage, Vgb of the cross-coupled p-fet that is tied to the branch with 0V intermediate voltage will exceed the maximum rating of 25 V. This rating is limited by the gate-oxide break-down and electron tunnelling.

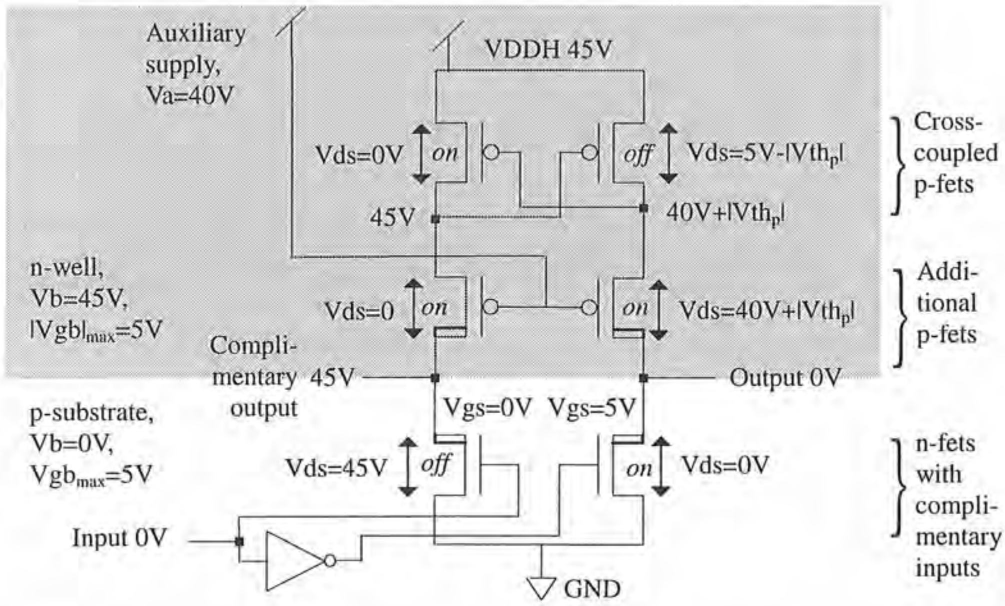


Figure 6.22: Asymmetric-high-voltage (AHV) level shifter showing the transistor states for 0V input.

A level shifter design using six transistors is presented in Figure 6.22. Two p-fets are added and biased by an auxiliary supply, Va, ~ 5V below VDDH. This prevents the source of these p-fets from dropping below Va + |Vth$_p$|. Hence, the gates of the cross-coupled p-fets will not be pulled > 25V below the n-well bias (= VDDH). Figure 6.22 shows the intermediate voltages and Vds across all six transistors at steady state for a 0V input. The voltages in the two branches are reversed if a 5V input is given. Although CBY permits the n-well to be biased up to 50V, the maximum p-fet |Vdb| of 45V does not allow a [0,50V] output.

The auxiliary voltage supply can be lowered to half of VDDH. The rating of Vgb (< 25V) will not be exceeded. An SLM backplane fabricated using symmetric additional p-fets in the symmetric high voltage (SHV) level shifters has been demonstrated [106]. However, the design in Figure 6.22 for a high voltage level shifter has the following advantages:

i) The auxiliary supply of VDDH-5 can be generated internally by cascading a series of diode-connected MOSFETs,

ii) *Low voltage* p-fets can be used for the two cross-coupled p-fets, as $|Vds| < 5.5$V (asymmetric high voltage p-fets required for SHV),

iii) asymmetric *high voltage* p-fet for the two additional p-fets, as the $5.5 < |Vds| < 45$V and $|Vdb| > 10$V but $|Vsb| < 10$V (symmetric high voltage p-fets required for SHV),

iv) asymmetric n-fets for the two n-fets, as $0 < Vds < 45$ and $Vdb > 10$V but $Vsb = 0$V; this can utilise the deep n-well for LDD drain and the p-well for the substrate (as offered by CBY process of AMS).

There is no loss in speed due to the circuit design of AHV. The reduced Vgs will turn on one of the cross-coupled p-fets more slowly but this is offset by a quicker turn-off of the cross-coupled p-fet in the opposite branch. The resultant time delay in shutting the DC path of the level shifter branch having a logic 1 input is similar.

A straight forward high voltage inverter working at [0,45V] is not feasible due to gate-oxide break-down limitations. Buffering will require separate control of the n-gate and p-gate. The CBY AHV level shifter will occupy more than 4× the area of a CBH level shifter. Coupled with the pre-commercial run nature of process, it was decided that the Roses silicon backplane would be designed for the 10V 2-μm CBH process.

6.7 Conclusions

Multi-level phase-only holograms are thought to be capable of improving the insertion loss and the crosstalk of binary hologram replay recorded on FLC/Si SLMs. Both of these are found to be insignificant by placing the practical constraints of fast switching analogue LCs into the expressions describing the diffractive power of multi-polarisation states. In addition, using CSLCs to encode these phase elements is a case of lossy binary phase modulation whereby more light can be deflected to one of the ± orders depending on the composition of the input polarisation. Nevertheless there are instances where the input polarisation can be controlled and an asymmetric replay is required. Double-pass optical configurations are needed to obtain ideal four-phase modulation.

In addition to an analysis of dynamic CSLC holograms, the architectural issues for the Roses silicon backplane design have been identified and considered in this chapter. There are two linear arrays of voltage modulators designed onto the silicon. Both the binary and quaternary arrays are designed for high quality optical modulation targeting holographic applications. For an analogue modulator, it is possible to implement analogue drivers by locating the DACs at the external, column-level or pixel-level of the chip. However, for the linear array with four modulation levels, the digital data are converted to analogue values simultaneously for all the pixels on chip. Polysilicon has been identified as the suitable VLSI layer to implement IR transmissive modulating elements. For the binary array, the key design approach has been to leave an array of tabs for connecting to high quality third layer metalisation pixels to be deposited after the silicon fabrication.

The CBH 10V 2-μm CMOS process from AMS has been chosen to implement the VLSI driver circuits and pixel layouts. The logic operations of the digital drivers are implemented using transmission gates, inverters and shift registers at the standard 5V supply. The voltages for driving pixel are stepped up to 11V using level shifter circuits. A dual-drain MOS CBY 50V process is probably able to meet the unusual requirement of small MOSFET dimensions and high voltage operation of an SLM backplane. However, this is not suitable for this particular design exercise due to the larger layout area for level shifters and an unknown process commission date.

CHAPTER 7

Circuit design and layout of the Roses chip

7.1 Introduction

Since the laboratory demonstration of polarisation insensitive free-space optical switching using reconfigurable FLC holograms [90], there has been tremendous interest in evaluating the true capabilities and robustness of this technique. Whereas pixel density and frame rate are the only critical requirements for general-purpose/display SLMs, an SLM designed for coherent optical applications has to incorporate a range of desirable features such as:

i) high optical quality metal reflectors (flat and highly reflective),

ii) pixels with a high fill-factor and a fine pixel pitch,

iii) large pixel array and long pixels to reduce the Gaussian beam clipping,

iv) drive electronics capable of high frame rates and

v) high electric fields to enable large electro-optic modulation.

Although a glass SLM can provide high quality optical modulation, a silicon based SLM with integrated drivers is preferred for circuit density and speed. In line with these objectives, a silicon modulator was designed and fabricated under the Roses* collaborative project.

---

* "Reconfigurable Optical Switches for Aerospace and Telecommunications Systems (Roses)", is a collaborative DTI/EPSRC (Department of Trade and Industry / Engineering and Physical Sciences Research Council) Link Photonics project between British Aerospace (BAe), Central Research Laboratories (CRL), Cambridge University Engineering Department (CUED), King's College London (KCL), Northern Telecom (Nortel) and Thomas Swan & Co. Ltd. (TS) and is managed by TS.

CHAPTER 7

Circuit design and layout of the Roses chip

7.1 Introduction

Since the laboratory demonstration of polarisation insensitive free-space optical switching using reconfigurable FLC holograms [90], there has been tremendous interest in evaluating the true capabilities and robustness of this technique. Whereas pixel density and frame rate are the only critical requirements for general-purpose/display SLMs, an SLM designed for coherent optical applications has to incorporate a range of desirable features such as:

i) high optical quality metal reflectors (flat and highly reflective),

ii) pixels with a high fill-factor and a fine pixel pitch,

iii) large pixel array and long pixels to reduce the Gaussian beam clipping,

iv) drive electronics capable of high frame rates and

v) high electric fields to enable large electro-optic modulation.

Although a glass SLM can provide high quality optical modulation, a silicon based SLM with integrated drivers is preferred for circuit density and speed. In line with these objectives, a silicon modulator was designed and fabricated under the Roses* collaborative project.

---

* "Reconfigurable Optical Switches for Aerospace and Telecommunications Systems (Roses)", is a collaborative DTI/EPSRC (Department of Trade and Industry / Engineering and Physical Sciences Research Council) Link Photonics project between British Aerospace (BAe), Central Research Laboratories (CRL), Cambridge University Engineering Department (CUED), King's College London (KCL), Northern Telecom (Nortel) and Thomas Swan & Co. Ltd. (TS) and is managed by TS.

The silicon design of the Roses SLM was conducted using Mentor Graphics layout tools. The design process is highly iterative and evolves around *digital circuit design* using Design Architect®, *VLSI layout* using IC Station® and *analogue simulation* of the implemented VLSI circuits with extracted parasitic capacitance values using AccuSimII®. The inter-relations between these three design sub-processes are encapsulated in the top left of Figure 7.1, which also shows the complete sequence in silicon backplane design/fabrication/post-processing and SLM assembly/testing.

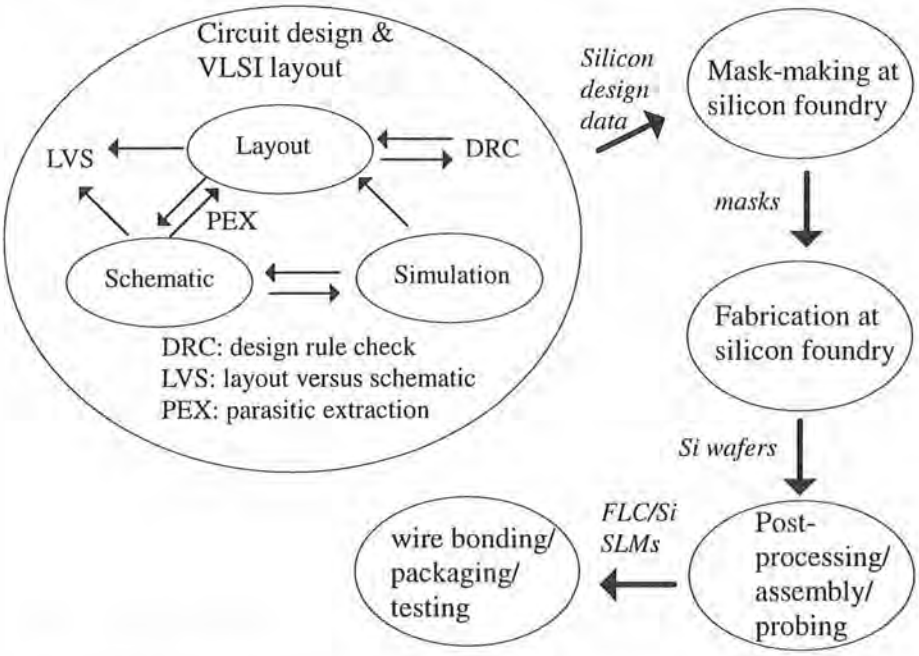


Figure 7.1: The design and fabrication of silicon backplane FLC/SLMs.

Several *generic design issues* have been dealt with in the previous chapter taking into account the optics and SLM applications. In this chapter, the *chip design* for implementation using CBH 2-μm 10V silicon fabrication process is described. It is aimed at meeting a set of specifications. The 'black-box' approach is the most suitable considering that the chip has to implement a set of electronic functions with several geometrical constraints in the VLSI layout of the chip imposed by both the chip size and the intended applications. The circuit design for pixel drivers is typical for SLM backplanes, rather, the laying out of VLSI structures in order to meet the area constraints forms the thrust of this design exercise.

7.2 Specifications for the binary and quaternary modulators

The primary aim of Roses chip design was to provide a Silicon backplane for a *binary* FLC/SLM. The linear SLM was intended for one-dimensional holographic beam steering applications. Due to the 1-D nature of the modulating array, the driver circuitry can be hidden from the optical illumination. This also enables a backplane design where the pixel array was left with a flat silicon area without any VLSI structures. Metal tabs were provided for making contacts with the cold-evaporated/sputtered aluminium (Al) mirrors during post-processing. This design was adopted to produce optically flat (~ λ/10 as there are only the insulating oxide and the protective nitride layers above the bare silicon substrate) and highly reflective pixel mirrors by avoiding the use of CMOS metalisation layer for the optical interface. Absorption and scattering losses were expected to be reduced. In addition to the binary array, a four-level voltage modulator was also designed onto the same chip.

7.2.1 The reflective binary array

A summary of the specifications for the 1-D binary SLM/switch which are dependent on the silicon design, post-processing and optical design, respectively, are as follows:

a) Those which are solely/largely determined by the silicon design:

- 20 μm pixel pitch,
- 6 mm long pixels and
- at least 320 pixels (> 6.4 mm array width),
- stacked VLSI-layer pillars to support a minimum of ~ 2 μm LC cell thickness,
- on chip temperature sensors to monitor thermal drifts,
- moderate frame time of 200 ~ 250 μs (4k to 5k frame/s) and
- asynchronous blanking to high and low functionality.

The figures for pixel pitch, length and number are not cemented at the silicon design stage. They are the upper bounds to be realised at post-processing. The actual figures depend on the design of layer-3 metalisation mask and the success of mirror deposition.

b) Those which are dependent on post-processing capabilities:

- 2 μm pixel-gap giving a 90% linear fill-factor,
- high reflectivity mirrors to reduce absorption loss and
- λ/10 (155 nm) optically flat mirrors sputtered at low temperature to reduce scattering loss.

c) Those to be met by the optical system design:

- reflective operation at 1.3 and 1.55 μm IR telecommunications wavelength windows,
- 1 dB optical bandwidth of ~ 30 nm determined by the choice of output port locations, hologram sets, lens focal length and focal length tolerance,
- 10 dB insertion loss for a 1:8 switch and
- 20 dB worst case crosstalk isolation.

7.2.2 The transmissive quaternary array

In addition to the binary modulator array, a 1-D quaternary array capable of four-level voltage modulation was included in the chip. The D/A conversion of 2-bit input data for all pixels is to be performed simultaneously on chip. The quaternary array ideally has the same pixel pitch and number as the binary array. However, the length of each pixel would depend on the size of available silicon real-estate. The quaternary driver array should have approximately similar electronic frame rates as the binary driver array. Most importantly, the construction of the quaternary pixel array should allow for through-silicon modulation at Si-transparent wavelengths.

7.3 Overall floor-plan

The overall floor-plan of the full-custom designed Roses chip is shown in Figure 7.2. The binary array of pixels and drivers are located at the top section of the die whereas the quaternary array occupies the remaining area of the lower section. A gap of ~ 310 μm has been allocated between the two pixel arrays. A circuit-free 1-mm ring around the periphery of the two pixel arrays has been allocated for printing the glue line. VLSI pillars made from stacked POLY1/POLY2/MET1/MET2 (polysilicon layer 1&2 and metal layer 1&2) are distributed evenly within the glue lane. These structures prevent the front-electrode from shorting to the backplane circuitry as well as setting a minimum LC cell gap. Data from the binary and quaternary pixel drivers are routed to the pixels by 6-μm wide MET2 tracks across the glue lane. These tracks are protected by the nitride passivation layer. There is a large n-type hard guard ring surrounding each pixel array. Substrate grounding for both pixel arrays is hence achieved.

The binary pixel array is driven by two rows of interlaced drivers. These drivers take their input from two serial links. The quaternary driver array occupies an area larger than the binary driver array. In order to space the quaternary pixels at the same pitch as the binary pixels, three staggered rows of quaternary drivers were used. Each quaternary driver element contains a 2-bit digital to analogue converter (DAC). The control circuitry before each driver array buffers the clock, transfer, asynchronous set/reset and data signals. The two clock signals are only buffered once for each driver row.

The left edge of the die contains 37 bond pads. The pad layouts are standard family cells from AMS measuring 215 × 540 μm^2. The more speculative driver design for the quaternary array has greater risk of failure. In order to reduce its cascading effect on the operation of the binary array, all positive supplies to the two arrays are kept separate. The binary drivers require standard 5V and high voltage (up to 11V) supplies whereas the quaternary drivers require three positive rails in addition to the standard 5V supply. There are two sets of control signals, one set each for the binary and the quaternary arrays due to the separate power lines. The use of serial links, two for binary data and three for quaternary data, reduces the total pin-out count.

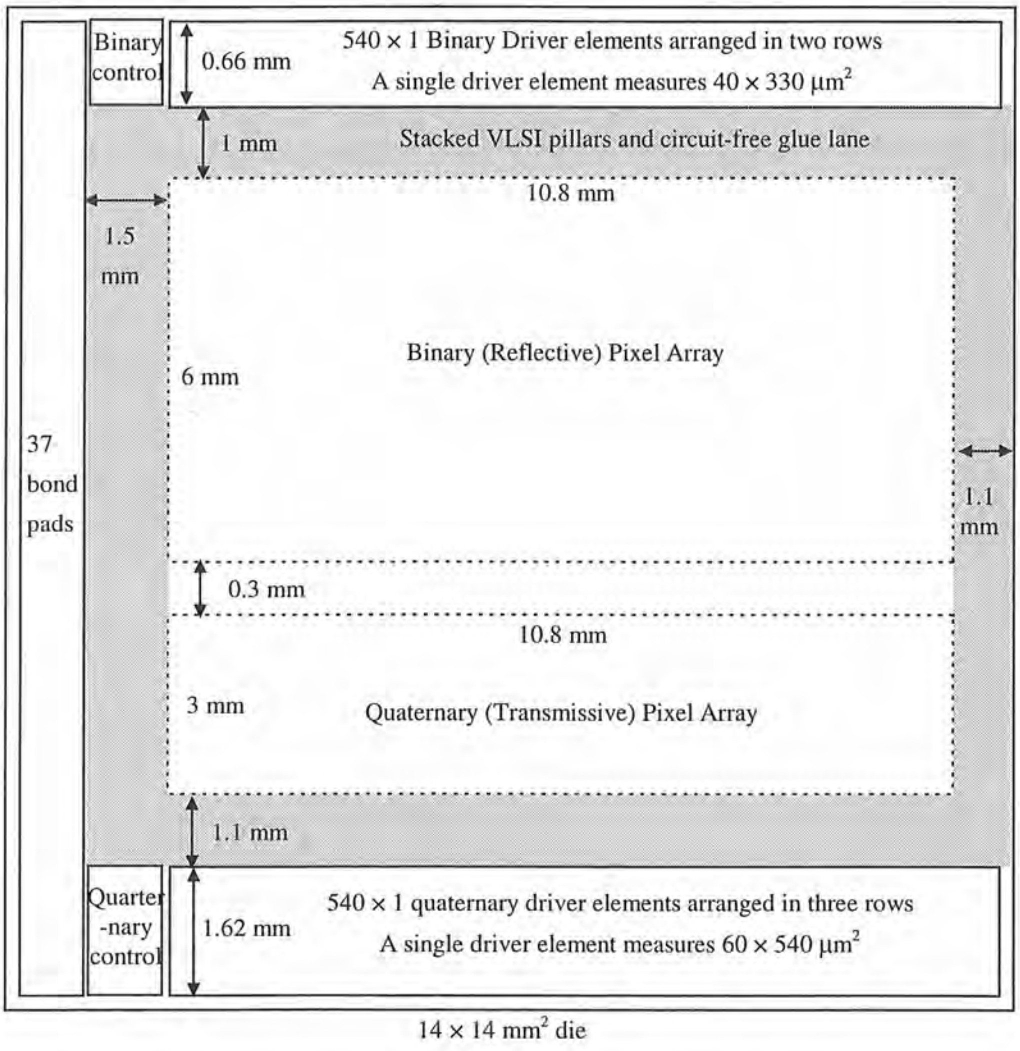


Figure 7.2: The silicon die floor-plan of the Roses chip.

7.4 Features and simulated performance of the Roses chip

The top-level Roses chip features are listed in Table 7.1 [107]. The entries regarding pixel and array sizes as well as the maximum clock rates are useful for optical systems designers. For an operator of the backplane (or assembled SLMs), the data formats and main interface requirements are contained in Table 7.2 [108]. The operation of either array is fully represented in Figure 7.3. In essence, a frame of data consisting of two (binary) or three (quaternary) parallel streams are clocked into the shift registers by the two non-overlapping clock signals, CK0 and CK1. At the end of each frame time (270 and 360 clock periods for binary and quaternary arrays, respectively), a global transfer (TR) signal latches all data bits from the shift registers to the static memory cells.

Si-backplane features	Binary (reflective) array	Quaternary (transmissive) array
Die size	14 × 14 mm^2 data-extrema	
Silicon Process	AMS CBH 10V 2-μm CMOS process with p-substrate, double-METAL and double-POLY	
Total pin-out	37 bond pads of 100 × 100 μm^2 at 375 μm pitch	
Array dimension	540 × 1	540 × 1
Pixel pitch, *d*	20 μm	20 μm
Pixel size	18[*] μm × 6 mm	18 μm × 3 mm
Pixel-array size	10.8 × 6 mm^2	10.8 × 3 mm^2
Pixel addressing	SRAM direct-drive	SRAM direct-drive
Data inputs	2 serial links	3 serial links
Driver array	270 × 2 each 40 × 330 μm^2	180 × 3 each 60 × 540 μm^2
Minimum clock period (=min. data bit duration)	2-phase non-overlapping 20 ns (50 MHz cycle)	2-phase non-overlapping 20 ns (50 MHz cycle)
Max. rate; min. time	185k frames/s; 5.4 μs	139k frames/s; 7.2 μs

Table 7.1: Summary of designed features for Roses chip.

---

[*] The actual pixel size is dependent on post-processing results.

Si-backplane elements	Binary Array Operation	Quaternary Array Operation
Driver circuitry	Serial to parallel shift registers and SRAMs	Serial/parallel shift registers, SRAMs and 2-bit DACs
No. of shift registers/SRAMs	270 × 2 (E and O rows)	360 × 3 (A, B & C rows)
Data input names	BDE and BDO	QDA, QDB and QDC
Control signal names	BCK0, BCK1 and BTR	QCK0, QCK1 and QTR
No. of data periods /frame	270	360
Data format convention	Data for last pixel latched first (presented to input pad). Alternate two 1/2-frames: **E** and **O** from the first pixel (see Figure 7.4)	Data for last pixel latched first Interleaved three 1/3-frames: **A**, **B** and **C** from the first pixel (see Figure 7.5); for each pixel state, LSB is latched first
Power input names	GND, VDD and VDDH	GND, VDD & VDDH[1 2 3]/3
Active-low asynchronous reset/set signal names	NRST/NSET	NRST/NSET
Driver output voltages	GND and VDDH	GND and VDDH[1 2 3]/3

Table 7.2: Operational parameters for Roses silicon backplane.

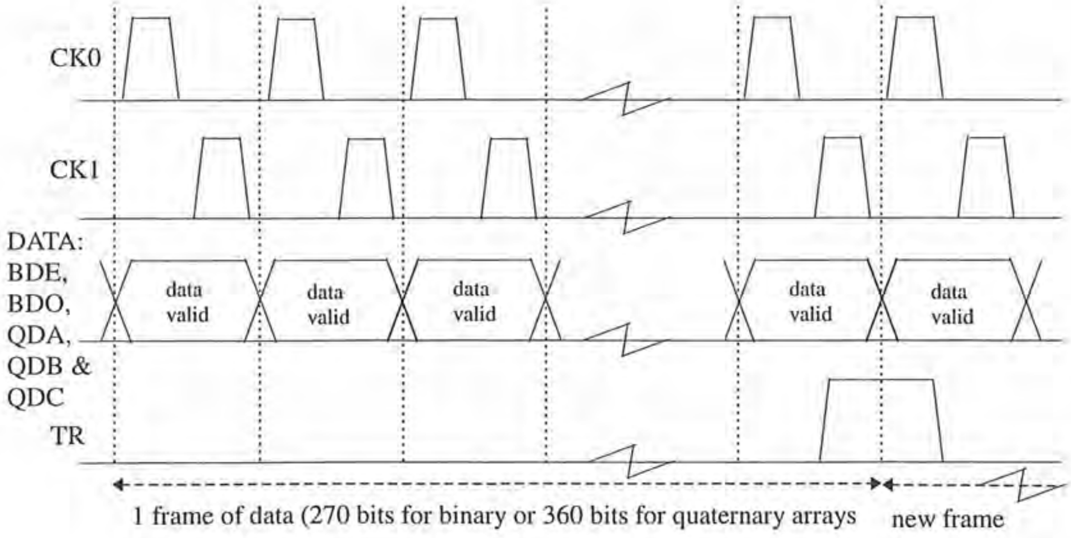


Figure 7.3: Relative timing diagram of clock signals and data inputs.

For the binary array, the voltage level of logic 1 signals is stepped up to 11V before driving the pixels. For the quaternary array, data from two successive memory units are combined to select one of the four power rails which then drives a single pixel. All voltage conversions for the binary and quaternary arrays occur simultaneously. During the LC/pixel driving, the TR signal disconnects the shift register memories from the static memory units before the next clock cycle. This enables the latching of the next data frame while LC switching is still taking place.

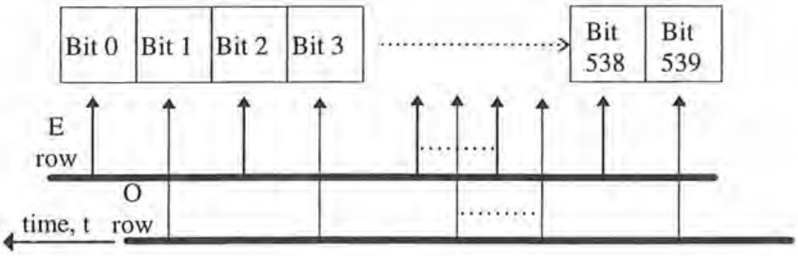
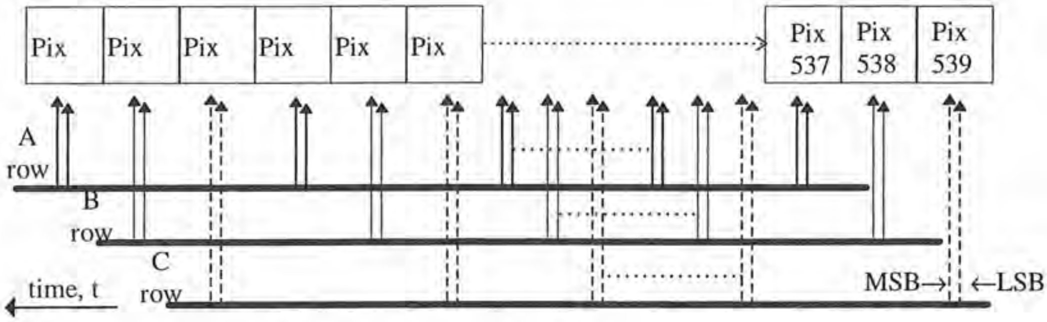


Figure 7.4: Two alternating 1/2-frame data format for binary array operation.



MSB: most significant bit, LSB: least significant bit

Figure 7.5: Three 1/3-frame data format for quaternary array operation.

The PSPICE simulation results concerning delay, rise and fall times of the output stages of the drivers under the typical mean and the worst case speed scenarios are given in Table 7.3. The 200 ns response time is much quicker than the several microseconds required to load a frame of data. The response time of the dynamic shift registers at 20 ns is conservative. This value sets the 50 MHz maximum data line and clock rates that the arrays can be operated on. However, even operating the data latching circuitry at 1/10 of their capabilities would still meet the intended 200 µs SLM reconfiguration rate comfortably. At 5 MHz clock rates, the liquid crystal would remain fully switched for 3/4 of the reconfiguration interval. The DC power was found to be negligible at this clock rate.

Simulation domains	Binary Array Operation	Quaternary Array Operation
Delay, τ_d (TR to output)	~ 30 ns	~ 30 ns
Delay, τ_d (NRST/NSET-output)	~ 40 ns	~ 40 ns
rise & fall times, τ_r & τ_f for *typical mean* (tm) case and *worst case speed* (ws)	~ 80 ns (tm) ~ 110 ns (ws)	~ 35 - 80 ns depending on transition step (tm); ~ 40% more (ws)
Current peak for tm and *worse case power* (wp) (single element)	~ 1 mA (tm) ~ 1.45 mA (wp) (VDDH)	~ 1.1; 0.73; 0.73 mA (tm) ~ 1.6; 1; 1 mA (wp) (VDDH;VDDH2/3;VDDH1/3)
Current pulse width	~ 100 ns (estimate) (VDDH)	~ 100 ns (estimate) (VDDH; VDDH2/3; VDDH1/3)
RMS Power dissipation per driver at 5MHz clock rate	~ 10 nW (tm) ~ 40% more for wp	~ 12.5 nW (tm) ~ 40% more for wp
RMS Power dissipation for SLM drivers + control @ 5MHz	~ 5.5 μW (tm) ~ 40% more for wp	~ 7 μW (tm) ~ 40% more for wp

Table 7.3: Simulation results of Roses binary and quaternary drivers.

7.5 Circuit design for binary modulation

7.5.1 Binary driver schematics and functionality

The schematic diagram of a binary driver element is shown in Figure 7.6. In each row of binary drivers, two non-overlapping clock signals (CK0 & CK1) and their inverse (NCK0 & NCK1) shift serial data inputs (DBE or BDO) into a 270-bit dynamic shift register. Each bit of the dynamic shift register has two transmission gates and two inverters. During this time, the global transfer pulse is held low to separate the data of the current frame rippling through the shift register and the data of the previous frame stored in the static shift register. At the end of the data shifting, the global transfer pulse (TR) updates the states of the 270-bit static shift register. These standard logic levels [0,5V or GND, VDD] are buffered before being converted to [0,11V or GND, VDDH] logic levels. The output of the last stage buffer is reflected at pixels. The second set of circuits drive the interlaced sets of pixels in parallel with the first set of drivers. The simulation results of the operation of four binary drivers are given in Figure 7.7. These results show a response time of < 100 ns.

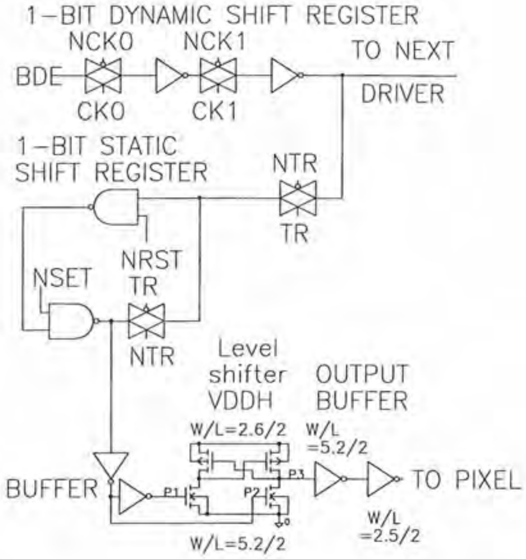


Figure 7.6: Transistor and logic-level schematics of a binary driver element.

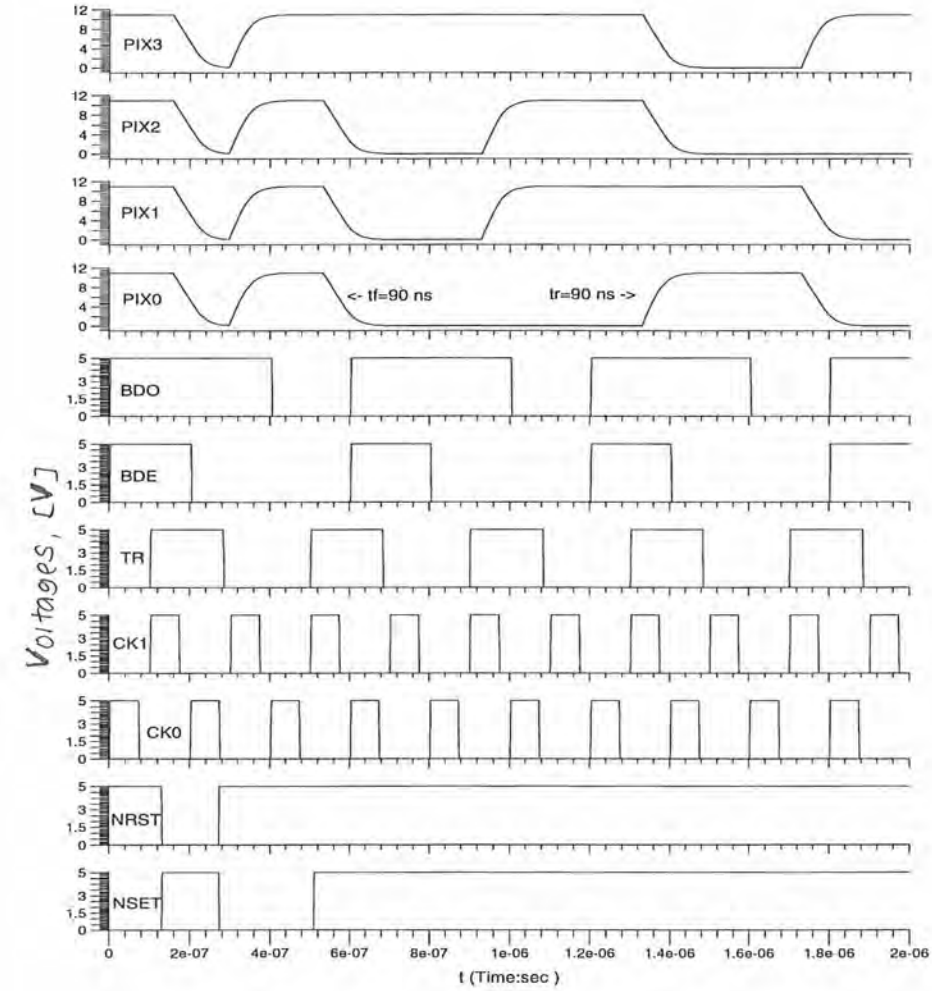


Figure 7.7: Simulation results of the binary driver functionality.

7.5.2 Asynchronous global blanking

An effective DC-balance FLC/SLM addressing scheme may require the use of frame blanking. Therefore, although cross-coupled inverters are sufficient to implement a static memory unit, they have been replaced by 2-input NAND gates to enable asynchronous global set and reset. These active low signals (NSET & NRST) override the contents of all static memories. The output of all binary drivers are either pulled-up or pushed-down depending on whether NSET or NRST signal goes low. This happens regardless of the state of the transfer pulse (TR). However, unless NSET or NRST stays low for the entire duration, having TR active can corrupt the blanking high/low as the outputs of the dynamic shift registers are not isolated. If both NSET and NRST are active, the binary array is blanked high.

7.5.3 Standard geometry MOSFET design

One of the key considerations of the pixel driver design has been the peak power requirement during the update of a hologram frame. In Figure 7.6, all the p-fets and n-fets have channel width/length (W/L) ratios of 5.2/2 and 2.6/2, respectively, unless shown otherwise. The standard geometry design for n-fet (W/L = 2.6/2) is just slightly wider than the minimum 2.4/2 allowed for in CBH process. The ~ 1 ns transit delay of a standard geometry inverter is at least two orders faster than the binary driver requirements.

7.5.4 Current-limiting in sizing level shifter transistors

In contrast to a typical inverter or NAND gate, the short-circuit current grounding through the serial n-fet and p-fet of a cross-coupled level shifter during switching is significant. The level shifter requires two complementary inputs (P1 & P2) at standard logic levels [0,5V]. The output P3 reflects P1 logic at a higher voltage level, VDDH. When a L(ow) to H(igh) transition occurs at P1 or P2, the corresponding p-fet in the same branch is not immediately turned off. Therefore, the path from VDDH to ground remains in a low impedance state until the p-fet in the opposite branch is fully conducting. The duration of this short-circuit current depends on the speed at which the n-fets are switched on, whereas its peak value is limited by the channel on-state impedance of the p-fets. It was found from simulation that an acceptable short-circuit current spike of ~ 650 μA is obtained by using W/L ratios of 2.6/2 and 5.2/2 for the p-fets and n-fets in both branches.

7.5.5 Speed versus current-limiting trade-off in sizing buffering transistors

The trade-off in non-standard geometry transistor sizing is the maximum current spike and the driver speed response. The W/L ratios of the output buffer transistors were obtained after many layout and simulation iterations. The peak current problem lies in driving *huge* pixels.

In the simulation, metal-2 (MET2) electrodes of 18 μm and 6 mm have a parasitic capacitance of ~ 2.8 pF. In addition, charging of the *liquid crystal dielectric* layer would have to be considered. However, the *charge compensation* effect due to the ferroelectric dipole rotation need not be included because this slow process does not cause high current-spikes. For simulation purposes, a minimum cell thickness of 1.9 μm and a relative dielectric constant, $\varepsilon_r = 5$, for the LC cells were assumed. This gave 2 pF additional capacitive loading at each pixel.

With 4.8 pF of pixel capacitance (modulator electrode + liquid crystal dielectric), the use of standard geometry transistors for the output buffering (5.2/2 and 2.6/2 for p-fet and n-fet) limits the current pulse to 1.3 mA for each driver and gives a response time of < 100 ns. For the purpose of equating rise and fall times, the n-fet of the second inverter in the output buffer has been reduced to 2.4/2. The design of output 'buffers' essentially limits the current spikes rather than boosting the current capabilities.

7.6 Layout of the binary backplane

Each driver element has 32 transistors and takes up an area of $40 \times 330\ \mu\text{m}^2$. Without having to fan in driver outputs using the metal-3 (MET3) mask at post-processing, two staggered rows of binary drivers are needed to meet the 20 μm pixel pitch.

7.6.1 Standard geometry MOSFET layout

The design rules of CBH process dictate the minimum channel width and length of 2.4 μm and 2 μm, respectively. In order to roughly equate the rise and fall times of an inverter output, the channel width of a p-fet has to be twice that of an n-fet due to lower hole mobility. Having considered the minimum diffusion layer extension of a diffusion contact (1.6 μm) and the size of the contact (2 μm), the W/L ratio for a standard geometry p-fet was chosen to be 5.2/2. At W = 5.2 μm, the layout of a p-fet is shorter as it does not require the notching of the diffusion layer that overlaps with the polysilicon layer-1 (POLY1) gate. Thus, the standard geometry n-fet has a W/L ratio of 2.6/2. These W/L ratios for the diffusion layer and minimum geometry for all other layers have been used in all MOSFETs except those used in the level shifter and the last stage buffer.

7.6.2 Protecting supply lines from peak current effects

The worst case current spike would occur at setting and resetting. If the whole row of 270 drivers were to be blanked high, the peak current pulse drawn from the VDDH line by the p-fets in the level shifters and the charging of output capacitance would have been 350 mA (1.3

mA × 270). The duty ratio of switching current (~ 100 ns) is very small compared to the frame time. Hence, the 1 mA RMS/µm-width current density rating for MET2 would not be exceeded. The standard rule-of-thumb is to limit the AC current density rating to 10× that of its RMS rating. Hence, a 50 µm wide VDDH MET2 track has been provided to withstand the 350 mA current spike. The corresponding ground line for discharge current is also carried by a 50 µm MET2 line. All other VDD and GND MET2 tracks are at least 5.2 µm wide. The current spike drawn from each power track has been simulated to be within the capacity provided.

7.6.3 Increasing the decoupling-capacitance of power lines

In the layout of the binary circuitry, there are multiple power rails extending from the left to right edges of the binary driver array. These power rails overlap with substrate/well diffusions throughout. Abutting a substrate tap with the source of an n-fet or a well tap with the source of a p-fet was avoided. Instead, the whole of the power line was tapped to substrate/well periodically where there was no metal interconnect. This increases the decoupling capacitance of the VDD and VDDH lines to ~ 1 nF and ~ 0.1 nF, respectively. A high decoupling capacitance helps to stabilise the power rails, which reduces the risk of latch-ups from parasitic thyristors. The intrinsic capacitance due to power supply rails was augmented by 100 nF external capacitors mounted externally to the IC.

7.6.4 Routing driver outputs to the pixel array

Owing to the use of two staggered driver rows, the outputs of the even row drivers have to be routed through the gaps between the odd row drivers before reaching the pixel array. This vertical interconnect is made from a 4-µm wide MET1 track. An increase of 1 µm from the minimum 3 µm MET1 width will improve fabrication yield for the ~ 330-µm long tracks. From the end of the O-row drivers, the outputs of all binary drivers are extended by 1 mm through the glue lane to an array of *pixel tabs* using 6-µm wide MET2 tracks. For a *thin and long* metal track in a circuit-free area, the metal etching process can sometimes result in broken lines. This 'lonely line' problem can be reduced by having ≥ 5 µm wide metal lines.

7.6.5 Pixel tabs to contact MET3 pixels

The cold-deposited Al mirrors must be contacted to the output of the binary drivers. There are two options to implement these contacts. The first is to etch the nitride protection layer from the whole of the pixel array including the end section of the 6-µm data tracks. It is probable that the poor step coverage of the MET2 edge by a thin MET3 (several hundred nm to 1 µm) would result in voids and broken MET2/MET3 interconnects. The other option is to leave the passivation layer as it is and cut several nitride-free windows (PAD layer) to provide the electrical connection. The deposited MET3 over the MET2 step would be supported by the nitride layer underneath. Besides, having the scratch-resistant nitride protection layer minimises handling damage to the area designated as the pixel array. The final layout of these pixel tabs is shown in Figure 7.8. Three 10 × 10 μm^2 PADs were provided to maximise the area of contact for each pixel. The complete backplane has a 540 × 3 PAD array. This layout was agreed with AMS as the PAD size and MET2 extension did not meet the standard design rules.

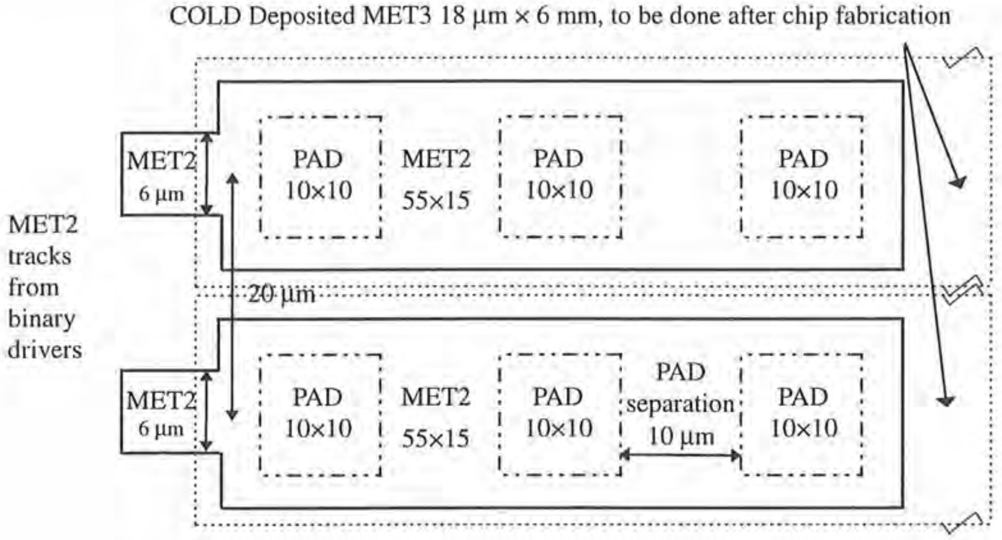


Figure 7.8: Pixel tabs contacting two post-processed binary pixels.

7.7 Circuit design for four-level modulation

7.7.1 Quaternary driver schematics and functionality

The transistor and logic-level schematics of a quaternary driver element is shown in Figure 7.9. Each driver element contains a two-bit shift register, 2-bit static memory cells, a 4-level logic decoder and four voltage-select MOSFETs. The functionality of the quaternary driver schematic up to the buffer of the static register is the same as the binary driver design. The outputs of the two data bits stored in the static memory cells and their complementary outputs are used to generate four enabling-signals by way of four NAND gates (address-decoder-style). The four active-low signals (D00, D01, D10 and D11) are non-overlapping. The active triggering signal digitally selects one of four power rails [GND, VDDH1/3, VDDH2/3 and VDDH] to represent one of the four different levels [D00, D01, D10 and D11] at the pixels. The simulation results for all twelve possible output transitions of the converter are shown in Figure 7.10. The settling time for any transition is < 100 ns.

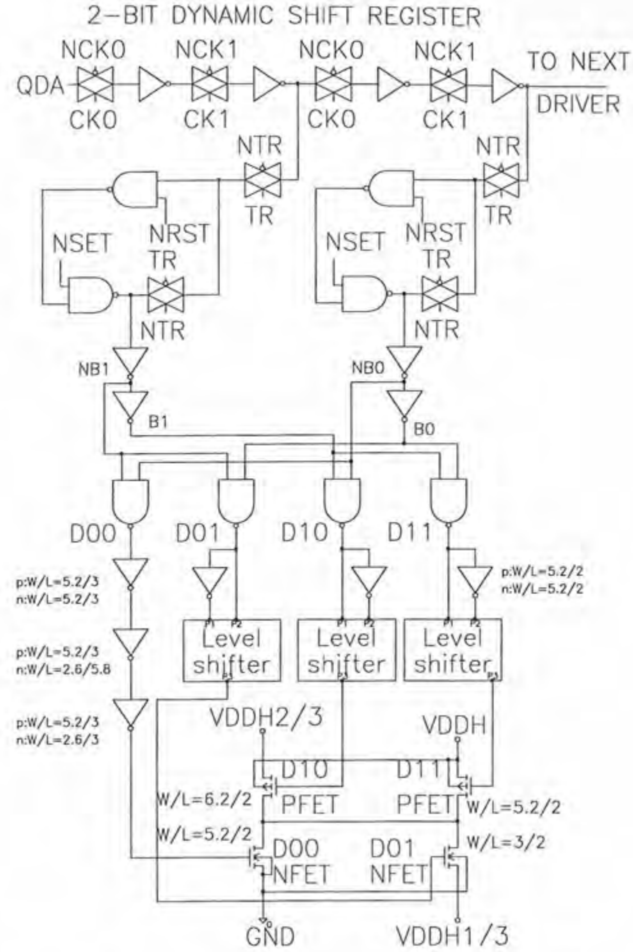


Figure 7.9: Transistor and logic-level schematics of a quaternary driver element.

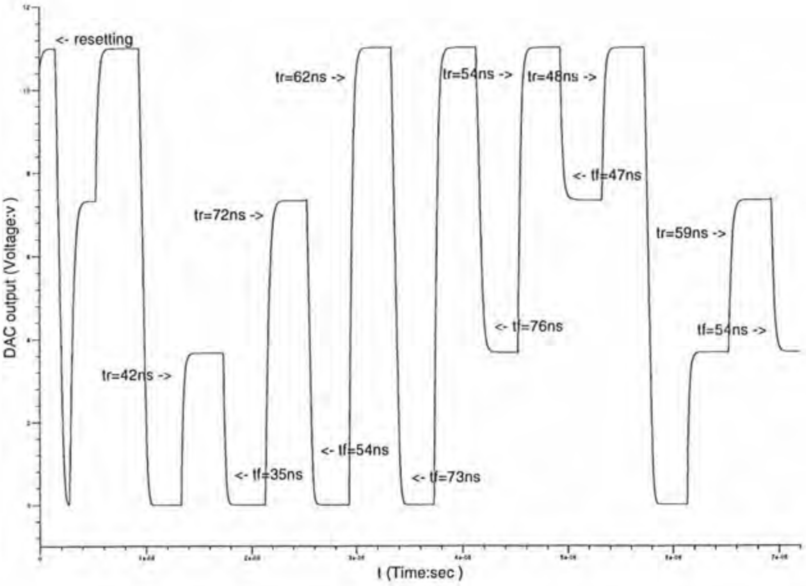


Figure 7.10: Simulation results of the D/A conversion of the quaternary drivers.

7.7.2 Standard geometry MOSFET design

All quaternary driver circuit elements up to the four-NAND-gate decoder are constructed using 5.2/2 p-fets and 2.6/2 n-fets for the same reasons as the binary driver design. In order to reduce the 'short-circuit' current passing through the select transistors, the design was aimed at providing a quick turn-off and a slow turn-on of any select transistor. This was achieved by using 5.2/2 p-fet and n-fet in the inverter stage following the decoder which gives short fall and long rise times. The three level-shifters in each quaternary driver have transistors with the same dimension and pin naming as the level shifter of the binary driver.

7.7.3 Converting drive voltages by digital selection of power rails

D00 and D11 of the four-level modulation require an n-fet and a p-fet, respectively. For selecting D11, the decoder output has to be converted to [GND, VDDH]. When the D11 select p-fet switches, the magnitude of the initial gate-source voltage (Vgs) is always VDDH, regardless of which level it switches from. For selecting D00, the decoder output of [GND, VDD] is passed through three inverters to equate delays and to obtain an active-high signal. Vgs in this case is always VDD, regardless of the previous state at the driver output.

Both n-fets and p-fets can be used for D01 and D10 select transistors. However, these levels are switched from a previous level which can be *higher* or *lower*. The Vgs value determines the speed of switching and the maximum current peak. Having performed the simulation, an n-fet and a p-fet have been chosen as the select transistors for D01 and D10, respectively. In both cases, the input voltage to the select transistor is shifted to [GND, VDDH]. The voltage for selecting D01 level is an active high input. The Vgs value is VDDH2/3 for switching from D10 and D11 but this increases to VDDH (*source-drain* interchanged) for switching from D00 to D01. The case for switching to the D10 level is the opposite for switching to the D01 level. The Vgs value is VDDH (*source-drain* interchanged) for switching the output from D11 to D10 and VDDH2/3 for switching from the other two lower voltage levels.

7.7.4 Current limiting by sizing select transistors

The output liquid crystal dielectric capacitance is small due to increased cell thickness (~ 4.7 μm and $\varepsilon_r = 4$ for DHF 9848) for transmissive operation. Although the length of quaternary pixels is half that of the binary pixels, the pixel pad capacitance is higher than that for binary pixels due to the use of POLY1 electrodes on field oxides. The total extracted pixel pad and LC dielectric capacitance used in simulations was 3.8 ~ 4.2 pF.

The W/L ratio of the D11 select p-fet is 5.2/2. This gives a maximum current pulse of 1.1 mA and a maximum rise time of ~ 62 ns. For the D00 select n-fet, a wider 5.2/2 channel was used because Vgs is roughly half that of the D11 select p-fet. The peak current pulse was 0.55 mA and a maximum 0-90% rise time of ~ 80 ns was obtained in simulations. The trade-off in optimising the channel width of the D01 and D10 select transistors was between limiting the peak current when Vgs = VDDH and increasing the speed when Vgs = VDDH2/3. The D01 select n-fet has W/L ratio of 3/2 whereas the D10 select p-fet has 6.2/2. In both cases, the slowest response time was ~ 80 ns and the maximum current peak was ~ 0.73 mA.

7.8 Layout of the quaternary backplane

The quaternary driver array contains twice as many dynamic shift register and static memory units (1080 bits) as the binary driver array. In addition, each quaternary driver element has a built-in DAC as well as requiring more supply lines. The 92-transistor quaternary driver element layout occupies an area of 60 × 540 μm^2. Thus these driver elements were arranged in three rows in order to meet the 20 μm pixel pitch specification.

In the layout of the quaternary drivers, similar considerations were given to *standard geometry MOSFET layout, protecting supply lines from peak current effects, increasing the decoupling-capacitance of power lines* and *routing driver outputs to the pixel array*. These considerations are not elaborated again here. The final layout adopted after many design iterations are described below. In addition, much attention was paid to the layout design of the quaternary pixels.

7.8.1 Layout of the quaternary drivers

In the layout of a quaternary driver element, the widths of MET2 tracks providing VDDH, VDDH2/3, VDDH1/2 and GND to the four select transistors are 50, 25, 25 and 25 μm, respectively. In addition, a 36-μm MET2 line provides GND to the sources of the n-fets in the D10 and D11 level shifters.

The vertical data routing of output signals from the *middle* and *bottom* driver rows through the *top* driver row is done by 3.2 μm MET1 lines running in parallel with a gap of 1.4 μm. The length of these two parallel tracks is > 500 μm. In the middle driver row, the output signals from the bottom driver row are carried by 5 μm MET1 lines. All the outputs are extended to the pixels by 6 μm MET2 tracks over a glue lane of 1.1 mm.

7.8.2 Layout of the quaternary pixel array

The layout of transmissive pixels is aimed at producing exposed electrodes. Either diffusion regions or doped polysilicon lines could be used for 1.55 μm transmissive operation. POLY1 layer is chosen for its lower resistivity, without the need for substrate/well strapping. The grain size of 200 nm should not cause optical beam scattering. The main challenge of laying out these pixels has been the removal of as many VLSI layers as possible at the silicon foundry. AMS agreed to remove the inter-layer dielectric (ILD) oxides by allowing for design-rule-violating CONTACT and VIA windows over the entire pixel array as shown in Figure 7.11.

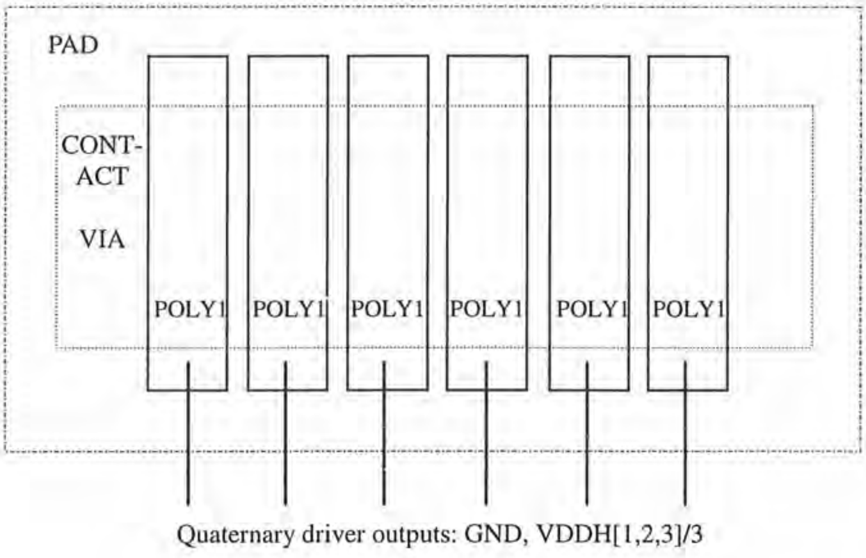


Figure 7.11: A section of the quaternary pixel array showing 6 POLY1 electrodes each measuring 18 μm × 3 mm.

However, POLY1 electrodes can only be prevented from etch damage by the subsequent etching processes (i.e., MET1 and VIA) if the CONTACT hole is filled with MET1 layer. Similarly the MET1 layer within the VIA hole has to be topped up with MET2 layer in order to protect it from MET2 and PAD etch impairments. This is because in the conventional VLSI processing, each VLSI layer (other than diffusions and wells doping) is first deposited, photoresist patterned and developed before the unwanted areas of VLSI layer are etched. A metal *lift-off* reverses this process sequence by first depositing the photoresist layer on the semiconductor wafer, patterning and developing the photoresist before depositing the metal layer into the areas with the photoresist removed.

However, even with the two metal layers protecting the POLY1 pixel array, their removal at post-processing is straight forward. The risk associated with this design was that the plasma dry-etch of the ILD oxides for a large area may not be uniform. There could be residual oxides on the modulators. The pixel design for 4-level modulation had been greatly helped by the design and process engineering support of AMS. They were of the opinion that over-etch resulting in broken pixels should not be the case.

7.9 Buffering of control signals

In both the binary and quaternary backplane designs, large four-stage graded buffers are provided for CK0, CK1 and TR input signals. These buffers are distributed to each row of drivers. The last two stages of the buffers for each driver row use 166.4/2 p-fets and 83.2/2 n-fets. The smearing of global setting and resetting pulses, NRST and NSET is less critical. Hence, two-stage graded buffers having 83.2/2 p-fets and 41.6/2 n-fets for the last stage are utilised for these signals. Data inputs to the binary array, DBE and DBO and to the quaternary array, DQA, DQB and DBC are also lightly buffered. These inputs are obtained from TTL-compatible digital pads.

The capacitance values for CK0, NCK0, CK1, NCK1, TR, NTR, NRST and NSET buffer elements are obtained by parasitic extraction of the actual layout. The difference in extracted capacitance from the layouts with 16 and eight drivers per row is scaled to the total number of drivers in a row. No attempt was made to extract the capacitance for each buffer element from the complete chip layout. The capacitance values range from 4 pF to 6 pF for clocking signals and 8 ~ 10 pF for transfer signals. Setting and resetting signals have very high capacitive loads for the quaternary array (14 ~ 22 pF) and moderate for the binary array (10 ~ 14 pF).

7.10 Bonding pads

It is likely that the 0.5A AC current limit for a single pair of power-ground pads would be exceeded at setting and resetting. Two VDDH pads have been provided for each array. There is only one VDD pad included for each array because the current consumption of the standard logic circuits in 540 pixel drivers and control blocks would not be higher than 0.5A. For VDDH1/3 and VDDH2/3 power supplies of the quaternary array, only one pad is provided for each power rail. In both arrays, three ground pads are implemented. There are no separate supplies [GND, VDD] to power the I/O pad buffers.

Other than the power pads, all I/O pad layouts are also standard cells from AMS. The periphery-limited digital input pad used was IB35P. The pad layout has a buffer and a pull-up p-fet. When the inputs are in the high impedance state (e.g., broken connector), the internal chip logic circuits are fed with 5V logic 1 inputs. The active low blanking signals are disabled unless deliberately taken low. The digital test outputs are obtained from OB33P pads which have an LDD n-fet constructed from the n-well layer. The analogue outputs from the quaternary dummy driver and the first quaternary pixel are connected to IOA4P analogue pads.

7.11 Functionality tests

There is no attempt to test the yield of the semiconductor fabrication process. Instead, additional circuitry is added onto the binary and quaternary drivers in order to verify that the intended binary and quaternary functionality is implemented correctly by the VLSI structures on-chip. The complete block schematic of the Roses chip is shown in Figure 7.12. The signal names on the left hand side and right hand side (except temperature sensor outputs) of the diagram are the test inputs and test outputs, respectively. The functionality tests focus on data latching/shifting and storing of a data frame in the static memory cells before being converted into suitable drive voltages. These tests are to be carried out during probing of an unprocessed wafer or an un-packaged die. Thus, all the pads for test outputs (except DRV_A0) are not wire-bonded.

7.11.1 Test of the dynamic shift register data latching and shifting

This test is concerned with detecting at the last element of each driver row ([E, O, A, B, C]) tokens that have been injected into the first element. The outputs at the last shift register elements, SREG_[E, O, A, B, C], are fed to digital output pads. If an injected token does not come out of a particular driver row, it could mean that the design is faulty or that the shift action is broken by a < 100% fabrication yield in that row. Although the data latches are implemented using dynamic shift registers, the capacitive charge storage should allow for a data interval of several milliseconds during chip probing.

7.11.2 Test of frame update, level shifting and D/A conversion

There are two dummy pixel drivers, one each for O-row of the binary and C-row of the quaternary circuitry. If the global transfer signal, TR and the voltage conversion circuitry have worked correctly, the stored pixel voltage level will be reflected at the output of these dummy drivers. The binary dummy driver output, DRV_O, is buffered before being connected to a digital output pad. The quaternary dummy driver output, DRV_C, which can assume ground or one of the three supply rails, is linked to an analogue output pad. In addition, the output of the first quaternary pixel, DRV_A0, is also channelled to an analogue output pad. The available driving voltages at the far end of the resistive POLY1 pixels can thus be determined. This test signal pad is to be wire-bonded out.

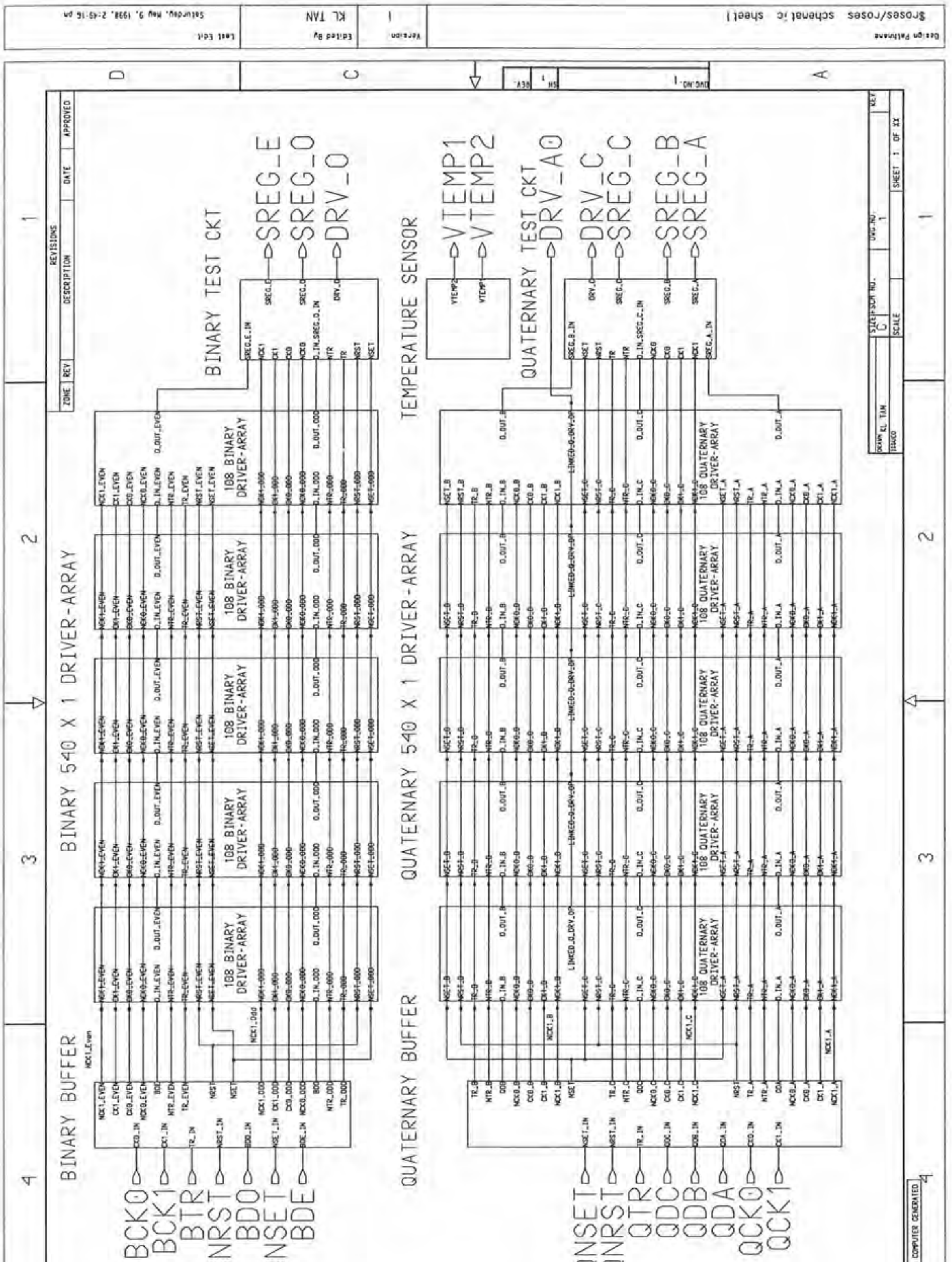


Figure 7.12: The complete block schematics of the Roses chip including test circuit blocks.

7.12 Conclusions

The important considerations in designing the circuitry and layout for the binary and quaternary Roses SLM drivers have been detailed in this chapter. The descriptions here do not include the details in the design iterations. The adopted end designs are reported to guide interested readers undertaking a similar work. The finalised layout accomplishes the electrical and geometrical aspects of the design specifications. Both the binary and the quaternary pixels are of 540 × 1 dimensions with pixel pitch of 20 μm. The backplane design aims to reduce the number of bonding pads and current requirements while still meeting the 200-μs frame time. The design was subsequently reviewed and tested for DRC errors before being fabricated at AMS. The following chapter will detail the probing and testing procedures of SLMs assembled from the Roses backplane.

CHAPTER 8

Characterisation of fabricated Roses devices

8.1 Introduction

This chapter describes all the important steps in realising SLMs from fabricated Si backplane devices and the use of these SLMs in several proof-of-principle holographic replay evaluations. Unlike the usual Si backplane design for displays, one of the Roses devices requires extensive post-processing in order to achieve enhanced mirror quality and the other requires the removal of the protective metal coating. Initial tests were conducted separately on the binary and the quaternary arrays of a whole wafer to ascertain that the circuit functionality was correctly implemented. Several wafers were then processed, diced and SLMs assembled from functioning die[*].

Un-packaged and packaged SLMs were inspected for optical response. A flexible software routine, albeit with a slow hardware interface was used to display a variety of patterns on the pixel arrays. The enhanced optical replay using these 'holographic chips': the binary modulator with good quality mirrors and the quaternary modulator, were demonstrated experimentally using typical optical set-ups. Similar experiments using a fixed glass grating were also conducted to enable the comparison of results with those obtained from the binary SLM. Temporal losses associated with hologram refresh schemes were also evaluated.

8.2 Initial tests of an unprocessed wafer

The aim of the first test is to ensure that functioning devices were fabricated by the silicon foundry. To this end, evidence of simple token shifting by the shift registers, level-shifting

---

[*] Silicon backplane post-processing and SLM assembly were performed by T.D. Wilkinson.

and digital to analogue conversion is required, before any further processing needs doing. Each device on the unprocessed wafer was also tested for short-circuits (S/C) between any two supply lines (GND, VDD, VDDH, VDDH2/3 and VDDH1/3).

8.2.1 Binary array

The binary array consists of two banks of drivers each with 270 elements. The two rows are termed 'O' and 'E'. At the end the 270-element serial shift register of each row, a buffer is used to feed the shift register output to a digital output pad. In addition, the 'O' row has a dummy pixel driver at position 270th. Hence, by latching in a single high-pulse token (BDE and BDO) for every two frames and having the transfer pulse, TR rising to H(igh) at clock period 270(Figure 8.1), the output of the 'O' row dummy driver (DR_O) will toggle H and L(ow) alternately (Figure 8.2). With the shift register outputs triggering a monostable chip with a duration (RC time constant in Figure 8.2) of just under the frame time, alternating token output (SR_E and SR_O) can be viewed easily.

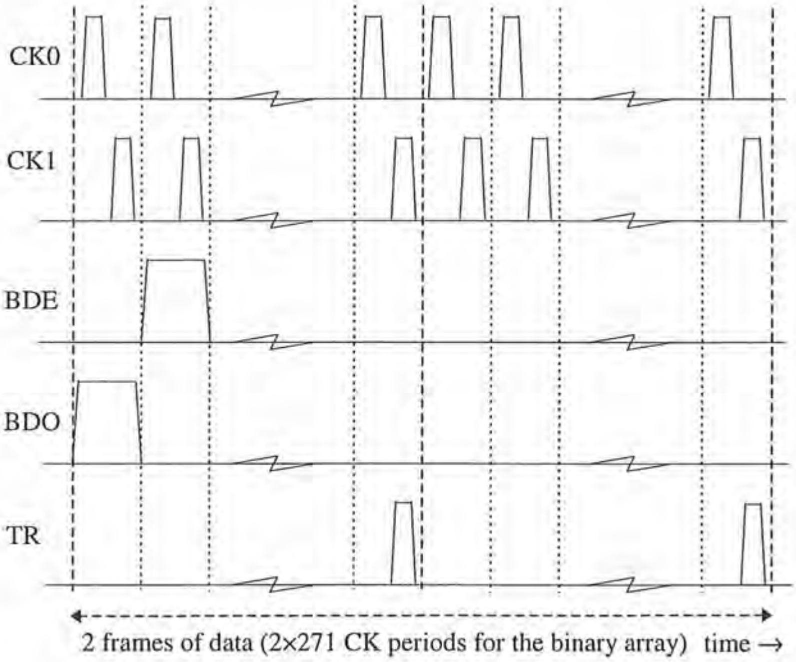


Figure 8.1: Control and data signals for probing the binary drivers.

This simple test procedure can only verify whether a binary device is shorted, any of the two shift registers is not functioning or the level shifting and buffering of each data bit is not carried out correctly. Yield failures of a section of a device cannot be determined. The quiescent current readings for VDD and VDDH are low at ~ 30 and ~ 17 mA, respectively. An upper limit of 100 mA has been set such that any device drawing the limit current is deemed short-circuited (S/C). Figure 8.3 gives a summary of the probing results. 19 binary devices from 26 complete-die passed circuit functionality (shift-register, level-shifter and buffer) tests, giving a silicon fabrication yield of 73%. It is also evident that the failed devices are located at around the edge of the 4" silicon wafer (the binary device sits at the top half of each die).

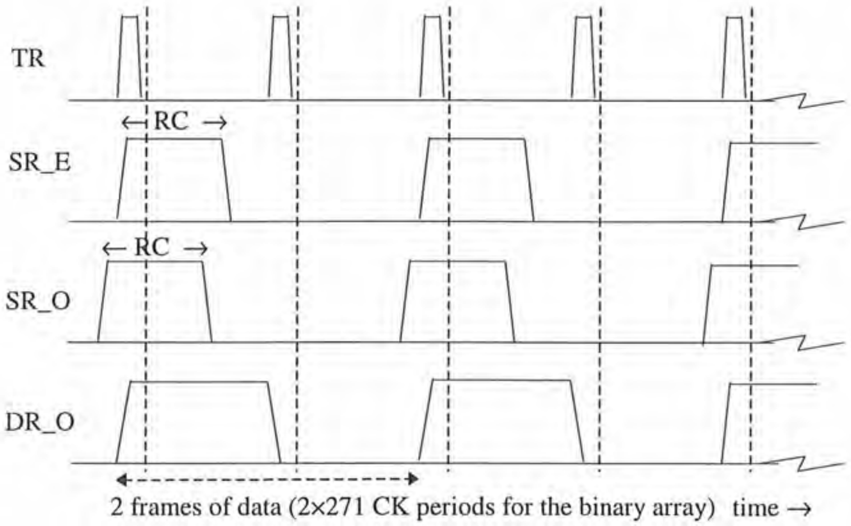


Figure 8.2: Expected output signals for binary devices with the correct electrical functionality.

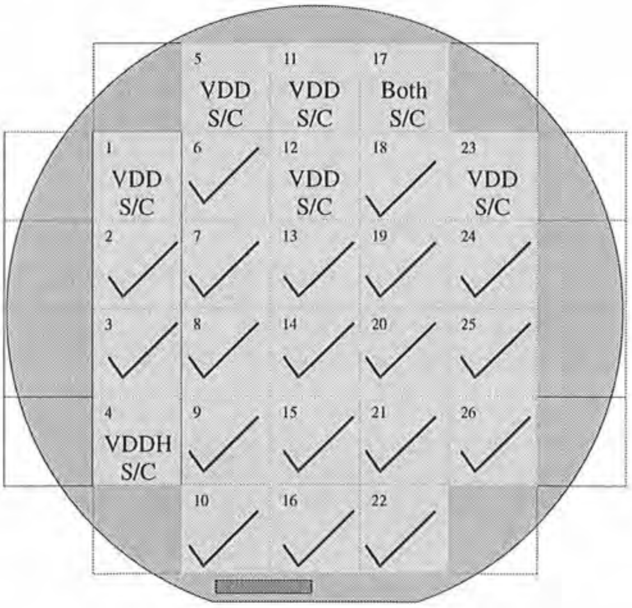


Figure 8.3: Results of binary device probing.

8.2.2 Quaternary array

The probing of the quaternary circuit functionality is very similar to that of the binary array. In addition to quiescent current tests, the QDC data input of the device is fed with waveforms corresponding to states '11', '10', '01' and '00' over four frame periods. The 540 drivers of the quaternary device are arranged in three rows, A, B and C. The last element of each shift register is buffered and connected to probing pads, SR_A, SR_B and SR_C. In addition, there is a dummy pixel driver at position 180 of C-row. Since each driver element requires two bits of data input, each frame duration corresponds to 362 clock periods.

The input waveforms for probing the quaternary device are as shown in Figure 8.4. During the second clock period, all data inputs (i.e., QDA, QDB and QDC) have to be the same due to the globally connected output stage of all drivers on the unprocessed wafer. The example in Figure 8.4 shows test sequences being fed to all three data inputs. Other probing pad outputs are as shown in Figure 8.5 if the three shift register outputs are first fed to monostables with a *RC* time constant. The results of functionality and short-circuit tests on all 26 quaternary devices are as shown in Figure 8.6. 20 of the 26 complete die were functioning, giving a silicon fabrication yield of 77%. Again, devices which failed were located near the edge of the wafer.

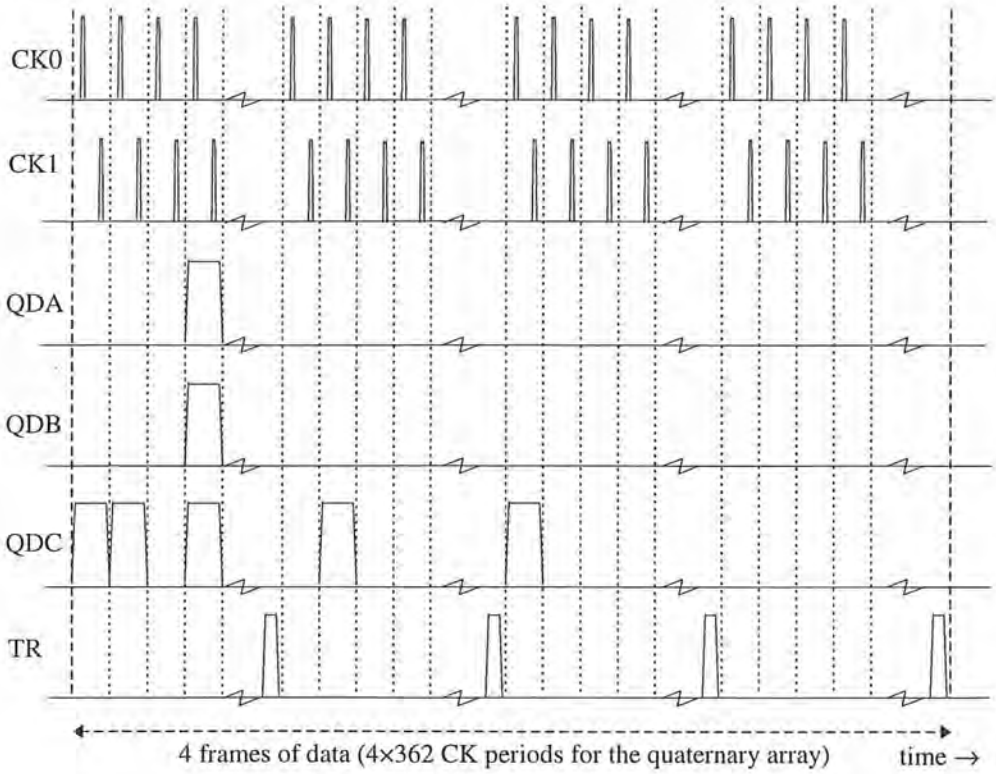


Figure 8.4: Control and data signals for probing the quaternary drivers.

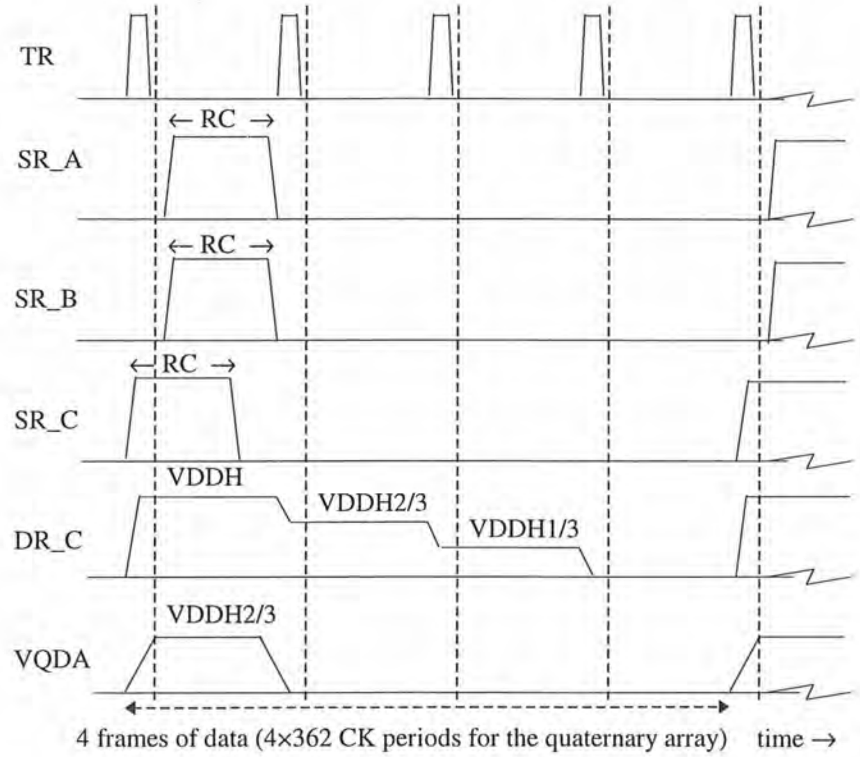


Figure 8.5: Expected output signals for electrically functioning quaternary devices.

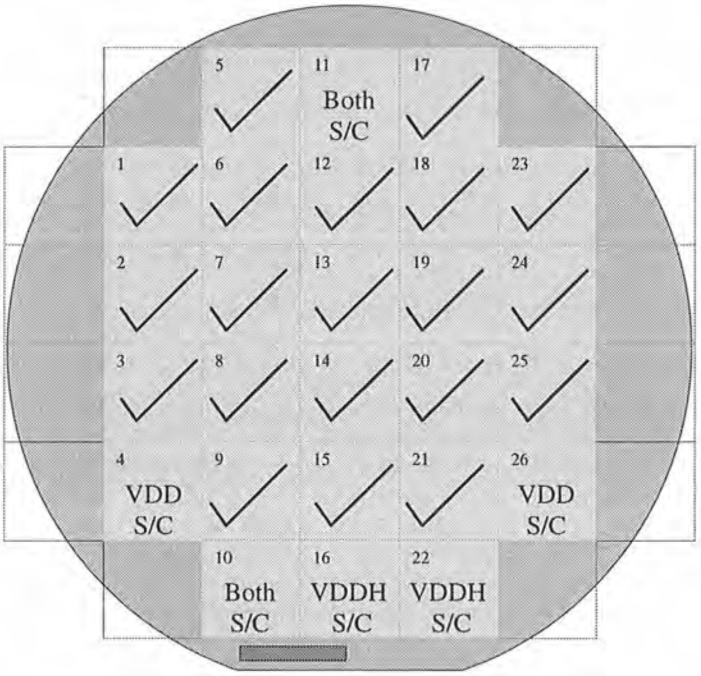


Figure 8.6: Results of quaternary device probing.

8.3 Processing silicon backplane devices and SLM assembly

Other than the sacrificial wafer #1, all other wafers from the Roses collaborative project were not subjected to the probing sequence detailed in §8.2 on a wafer scale. At this stage, the circuit design had been verified to be correct and the silicon fabrication yield was deemed acceptable. All assembled devices (including those from wafer #1) were treated with the processing steps to be described below, depending on whether the binary or the quaternary circuitry was to be used. These steps can be divided into pre-assembly processing and assembly of the LC filled SLMs.

8.3.1 Optical quality mirror deposition

Binary devices were returned by the silicon foundry with only nitride and oxide layers on silicon in the area designated for the binary pixel array. The intention was to coat the wafer with a thin but highly reflective Al third layer metal. A pre-deposition oxygen-plasma etch was required to 'cleanse' ~ 40 Å of the entire wafer. This should remove the layer of native oxides within each nitride-cut window and ensures good contact with layer 3 metallisation. A wafer-scale mask was prepared which was to be used to pattern the photo-resist. A dry/wet etch of the wafer with developed resist would result in isolated Al pixel mirrors with an overlap to their driver interconnects. The sputtering of the Al layer and the dry/wet etch of the pixel gaps can be controlled to obtain the required metal thickness and dead-space. Step-coverage at the edges of the underlying MET2 driver interconnects, electrical connectivity to the MET2 driver interconnects within nitride-cut windows and inter-pixel isolation will determine the yield of binary pixel array processing.

8.3.2 Al-etch of protective mirror coating

The extra processing step required before the assembly of quaternary devices was to remove the MET1/MET2 protective coating within the pixel array. Al wet-etch would naturally stop at Si, polysilicon or oxide layer. A negative photo-resist was patterned at die-scale and developed before the exposed metal layer was removed by aluminium etch solutions.

8.3.3 Assembly of silicon backplane SLMs

After the processed silicon backplane wafers were diced (for binary devices) or the diced wafers were processed (for quaternary devices), probable functioning devices using the yield charts from the initial electrical tests could now be assembled into SLMs. This is best done while the devices remain in the same clean room to minimise contamination risks. The assembly procedures involve spinning on and rubbing a thin 6-6 nylon layer, baking this alignment layer, printing a glue seal, attaching the ITO-coated front glass, filling the cell by capillary force at isotropic phase, cooling the LC to Sm C* phase, sealing the filling hole and connecting the front electrode to its driver line. The alignment layer has to be baked for several hours to ensure near-permanent alignment features/directions. This is especially critical if the SLMs are to be treated with cooling/heating cycles and AC fields for LC alignment purposes.

8.4 Tests of assembled SLMs

The assembled SLMs are then probed for silicon fabrication defects and additional damages introduced by the processing and assembly procedures. Power lines, inter-pixel shorts and driver-pixel non-contact are the most common electrical defects. However, shorting of the backplane mirrors and the front electrode has been virtually eliminated by the use of VLSI pillars on chip. The optical modulation characteristics of each SLM including interference fringes due to non-uniform LC cell gap across the pixel array, reduced pixel array due to areas filled with glue or air bubbles, damaged LC patches due to localised defects and LC alignment quality, can also be inspected. These tests had to be carried out in an iterative loop while optimum processing and assembly steps were still being worked out. In many ways, these tests are extremely useful before the SLMs are bonded and packaged onto carrier boards.

8.4.1 Initial optical inspection using the probe-station

After the processing and assembly of binary/four-level silicon backplane devices, a filtering process has to be implemented in order to select electrically and optically functioning SLMs for bonding and packaging. This was carried out using the probe station equipped with a microscope. Fully assembled SLMs were driven by the probing needles and the optical modulation was investigated using a sheet-polariser.

The probe-card used in device electrical functionality tests has the complete interface capability to drive both binary and quaternary SLMs. Hence, frames of images comprising of binary or four-level stripes could be displayed on the pixel arrays before the I/O pads were bonded. Several frames with successive shifting of the stripes forming an approximately 1 Hz animation are the best for direct viewing. When the linear sheet polariser is appropriately oriented with respect to one of the extreme switched states, the stripes corresponding to this switched state will show up as the brightest or the darkest.

The connection to the front electrode is not critical to successful observation of contrasts. A frame typically contains equal numbers of each of the switched states, hence, there are an equal number of areas where current has to be drawn from or returned to the front electrode in order to compensate for the polarisation rotation. For both the binary and quaternary array operation, the unconnected front electrode should have approximately VDDH/2 volts. Even if the number of pixels for each of the states is different, a charge equilibrium for the front electrode would still be obtained and partial switching is not a problem. Conversely, driving the front electrode with a bipolar AC square wave and floating the backplane pixels will also show contrast although charge equilibrium on the isolated pixels is not possible. The backplane pixels lie above the silicon substrate and tend to be grounded when they are not driven.

Table 8.1 shows the summary of both electrical and optical test results for SLMs made from devices on wafer #1; "*b*" or "*q*" at the end of each die number denotes whether the device was used for binary or quaternary operation in these tests. The current readings for the VDD and VDDH power rails are approximately 30 and 17 mA. These are largely unaffected by the choice of frame rates from the slow interface board or the patterns being displayed unless there are numerous inter-pixel shorts.

Wafer # / Die #	Liquid crystal	Electrical functionality	Optical modulation
1/2b, 1/7b, 1/13b & 1/19b	CDDR8	working	3 groups of pixels showed contrast
1/22b	unfilled	3 groups of pixels connected to external probes, but no signals detectable.	
1/17q	unfilled	Voltage probe dropped onto polysilicon fingers but no signals detectable.	
1/9q	unfilled	VDD short-circuit	
1/20q	unfilled	probe pads blocked by the front glass	
1/15q	CDDR8	working	binary stripes showing contrast
1/14q	DHF 10817	working	no contrast observed
1/21q	DHF 9848	working	4 grey levels shown
1/6q	DHF 9848	VDDH pads damaged during dicing	
1/3q, 1/18q, 1/24q & 1/25q	DHF 9848	working	quaternary operation observed

Table 8.1: Probe station test results of SLMs made from wafer #1.

The initial deposition of metal reflectors on wafer #1 was not a success. Sintering process at the foundry has resulted in a fairly thick layer of native oxides on MET2. Therefore, most of the deposited MET3 fingers did not form any contact with the MET2 pad underneath. It was suggested that a plasma etch has to precede MET3 deposition if no seed material (e.g., titanium) that can break through the oxides is used.

The triple contact/via/pad openings of the quaternary polysilicon pixel did not function as expected. Dropping a probe tip onto the top surface of the pixels did not produce any readings. In addition, there were visible fringes across the array due to uneven etches. However, the quality of optical modulation cannot be judged by observing the stripes using a single sheet-polariser on the probe station. Imaging or at least viewing under a cross-polariser/analyser set-up has to be carried out.

Wafer # / Die #	Liquid crystal	Electrical functionality	Optical modulation
2/10b	CDDR8	working	supplies shorted when driven
2/16b, 2/19b 2/21b, 2/22b & 2/26b	CDDR8	working	Contrast observed on most pixels; some pixels are dead/broken, potentially useful devices for experiments
2/9b & 2/15b	CS2005	working	Very few pixels switched
2/3b & 2/25b	CS2005	working	Contrast observed on most pixels, worth bonding out to carriers
2x9b & 2x3b	CS2005	working	most pixels showed contrast, worth bonding out to carriers
2x6b, 2x5b & 2x1b	CS2005	working	few pixels showed switching
2x8b	CS2005	not probed	LC looks well aligned
2x4b	CS2005	working	VDDH shorted when driven: pixels shorted at many places
2x2b & 2x7b	CS2005	VDD shorted	--

Table 8.2: Probe station test results of SLMs made from wafer #2.

The vacuum deposition of Al pixel array without pre-etching native oxides within the passivation windows did not produce any useful binary devices. Subsequently, the post-processing of binary arrays from wafer #2 involved pre-etching and sputtering with a ~ 0.5 μm thick Al layer before pixel gaps were etched*. This produced a much improved yield of Al modulating electrodes. However, problems with the step coverage of MET2 tracks by the Al electrodes due to uneven etching resulted in some unconnected (dead) pixels and the yield varies from die to die. Devices from wafer #2 had a pixel gap larger than 2 μm. Even so, there were still cross-linked pixels. The results of tests conducted using the probe station are shown in Table 8.2. In summary, mirror deposition with prior native oxide removal can be used to layout high quality optical reflectors, but the technique can be improved by dry etching the pixels.

8.4.2 SLM interface and carrier design†

The simple driver board for test purposes is not adequate when the SLMs are to be integrated into an optical switch system. Using the parallel port of a computer, each data period requires several microseconds and hence the slow interface produced a frame time of ~ 10 ms. A high speed 1D SLM driver board was developed by Central Research Laboratories (CRL) as a reduced feature version of the 2D SLM interface board. A photograph of the driver board for binary SLMs with the opto-mechanics of the demonstrator switch is shown in Figure 8.7. The basic clock for the board runs at 5 MHz and the fastest frame time is ~ 54 μs.

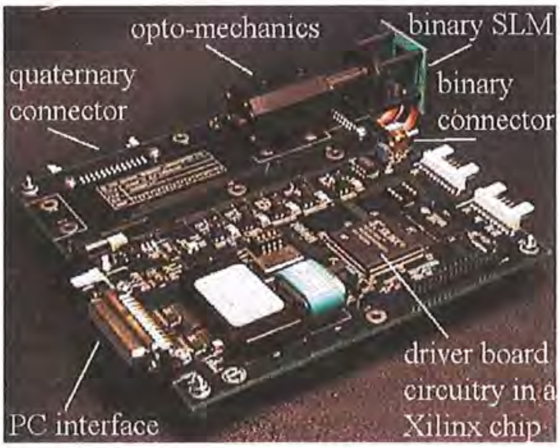


Figure 8.7: Photograph of the 1D SLM driver board for the 1 × 8 demonstrator switch.

---

* Deposition was done at GEC Marconi Gt. Baddow by M. Worboys and S. Radcliffe.

† The design and fabrication of SLM carriers and driver boards were performed by CRL.

The carriers for binary and quaternary SLMs are based on a single-layer substrate. The binary device has been designed to be integrated into the 1D demonstrator switch and hence it has a copper-clad aluminium substrate whereas the substrate for the quaternary devices is a PCB material. The silicon backplane SLM is glued onto the substrate using Araldite Rapid epoxy at room temperature to avoid damaging the LC. Gold bond wires are used to connect bond sites on the carrier to bond pads on the chip. Simple routing tracks link gold-clad bond sites and a row of solder pads. Decoupling surface-mount capacitors are included for power rail stability. A ribbon cable with connector ensures easy interfacing with the SLM driver board. The photograph in Figure 8.8 shows a binary SLM fully bonded onto its carrier.

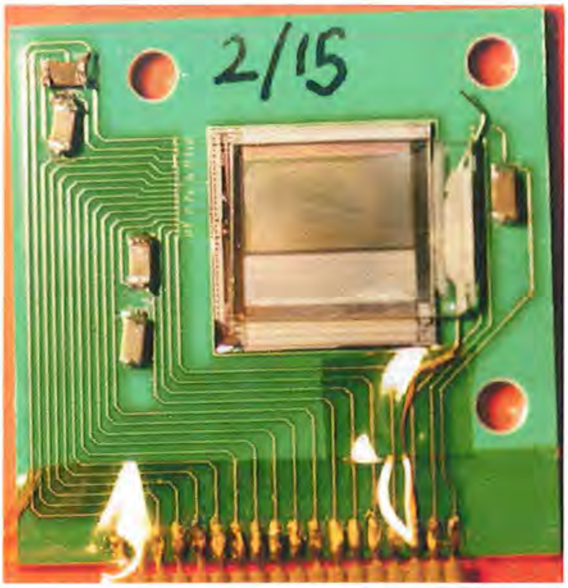


Figure 8.8: Photograph of a bonded binary SLM.

8.4.3 Imaging optical modulating using a polarising microscope

After the initial optical inspection using the probe station, electrically and optically functioning devices were selected for bonding and packaging onto carriers. The bonded SLM could be driven without the probe-card. Hence, liquid crystal alignment procedures could be implemented by grounding the backplane pixels while simultaneously applying AC fields to the front electrode*. High voltage (up to 10Vrms/µm) and low frequency AC square waves (2 ~ 3 kHz) are applied to the LC cell via the simple interface board for probing/testing while the cell is cooled from the isotropic phase [109].

---

* AC-field LC alignments for both binary and quaternary SLMs were performed by M.M. Redmond.

Several quaternary SLMs filled with DHF 9848 and 10817 were inspected under the polarising microscope. The quality of the quaternary SLMs with on-chip DAC and transmissive pixels was largely degraded by the pixel surface roughness, pixel integrity, inadequate removal of the oxide layer above polysilicon pixels and the difficulty in aligning LC on silicon/polysilicon. Figure 8.9 shows the intensity modulation of device 1/14q with aligned LC under the polarising microscope.

Figure 8.9: Four grey-level optical modulation using Roses 1/14q quaternary array.

Wafer # / Die #	Currents	Exposed data tracks	Contrasts within pixel array
1/14q	OK	Switching observed	Contrasts observed after LC alignment using AC low frequency electric field treatment, ~ 3/4 of each pixel length was broken.
1/21q	Shorted	--	Device damaged during bonding process
1/24q	Shorted	--	Device damaged during bonding process
1/3q, 1/15q, 1/18q & 1/25q	OK	Switching observed	Suspected loss of birefringence due to excessive heating while gluing the SLM to its carrier. These weak areas were damaged by localised defects when high amplitude AC fields were applied during LC aligning.

Table 8.3: Test results of the bonded quaternary SLMs.

The defects with all bonded quaternary SLMs are detailed in Table 8.3. None of these SLMs was suitable for imaging or holographic evaluations. The experience with bonding and gluing these SLMs to carriers has helped to improve yields with later batches of devices (including binary SLMs).

Wafer # / Die #	$\sigma = 1/2$ pattern	$\sigma < 1/2$ pattern	Remarks
2/22b	OK	OK	Best SLM, an area of non-birefringent and unconnected pixels at far edges
2/21b	OK	OK	Not as good, with two groups of cross-linked pixels at the centre of the array
2/16b	high currents	currents ↑ with hologram periods	Many areas of cross-linked pixels
2/10b	high currents	high currents	A failure.
2/25b	OK	OK	By far the best binary SLM so far, < 20 backplane pixels do not switch independently or unconnected, however, the front electrode was shorted to Si-substrate by the silver-loaded glue
2/3b & 2x3b	OK	OK	Areas of non-working pixels scattered across array
2x9b	high currents	OK	--
2/9b & 2/15b	OK	OK	Too few pixels contacted to drivers

Table 8.4: Test results of the bonded binary SLMs

The optical modulation of binary SLMs was also inspected under the polarising microscope. The binary modulation pattern corresponding to $\sigma = 1/40$ hologram replay fraction is shown in Figure 8.10. With suitable magnification, individual pixels can be imaged. Thus, the success of post-processing pixellation can be visually inspected. In addition, LC characterisation including tilt angle and response time measurements were conducted[*]. Results for the LC CDDR8 using device 2/22b were reported in full elsewhere [110]. Remarks on the bonded binary SLMs are given in Table 8.4[†].

Figure 8.10: Photograph of 1/40 binary base hologram modulation pattern.

8.5 Holographic SLM demonstrator

Having conducted the theoretical routing hologram analysis and synthesis, designed, fabricated, tested, assembled and packaged potentially high optical quality linear SLMs, an opportunity to put the theory to test emerged. The main motivation was to verify that the first order replay efficiency of any routing hologram fraction and the relative intensity distribution of all significant replay orders could be predicted with a high degree of accuracy. These results were obtained by mapping the far field replay of binary holograms using a scanning output fibre. Additionally, these experiments evaluated the Roses binary array with enhanced optical reflectors as a device for hologram displays.

8.5.1 Replay field mapping using intensity modulation

The analysis in §3 emphasises the critical dependence of the crosstalk figure on the clipping of the Gaussian beam by the hologram aperture. In order to eliminate device and material imperfections in dynamic holograms (i.e., SLM dead-space, phase-mismatch of LC and

---

[*] These measurements were undertaken solely by M.M. Redmond as part of a Roses work package.

[†] Several 'fully' functioning binary SLMs have since been assembled from dry etched pixel arrays. They have less than 10 of the 540 pixels not working independently or not contacted. These SLMs are currently being integrated into the Roses demonstrator switch. Preliminary results have shown an improvement of several dB in the switch insertion loss.

temporal ripples associated with LC switching requirements), a static binary intensity grating in a 4*f* coherent optical system (as in Figure 8.11) was used to produce a replay field. The replay field was 'mapped' by a single mode fibre (SMF) to simulate the coupling of crosstalk power in a fibre-to-fibre switch. The radiation modes launched into the SMF were stripped by winding the fibre several times around a 16 mm diameter hose. Using a SMF as a detector allows the large dynamic range of optical sensors/power meters to be exploited. The fibre was mounted on motorised stages and stepped along at sub-micron resolutions. The main drawback of this technique is the length of time required for data collection. However, this is well justified given the absence of intensity averaging by CCD camera pixels or photodiodes.

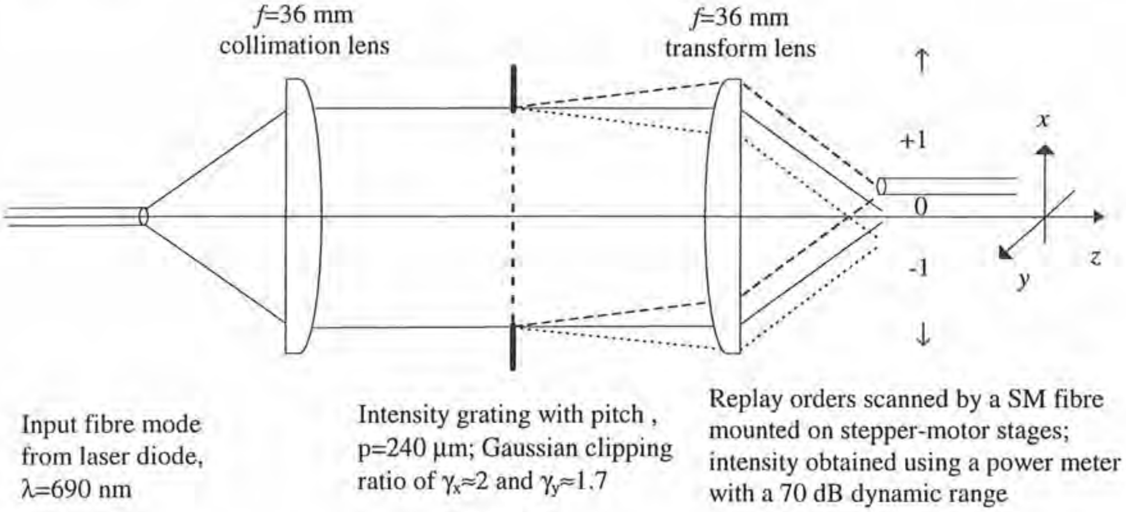


Figure 8.11: Optical set-up for transmissive grating replay measurements.

The fixed grating has vertical pixel/dead-space sizes of 120/120 μm forming a 10.8 × 3 mm^2 rectangle[*]. For the practical wavelength window of 633-1550 nm, the spot radius of the collimated Gaussian beam is roughly proportional to focal length *f* only ($w \approx 0.1f$). In order not to severely truncate the collimated Gaussian beam in the *y*-direction and present flexibility in the *x*-direction clipping ratio of up to 3, a pair of cemented achromat lenses with a nominal focal length of 36 mm were chosen. Variable size masks were prepared for a range of truncation factors in *x*-direction.

---

[*] The fixed intensity grating corresponds to σ=1/2 base hologram encoded onto the Roses SLM. The fabrication was undertaken by T.D. Wilkinson.

The 690 nm source used was not a distributed feedback laser. Hence, multi-moded behaviour which could cause smearing effects in the replay field was expected. The laser diode was cooled by passing a 1 A current at 1 V through an attached Peltier device. This laser source was tested to have 2% power stability and a measured aggregated power of 9 mW at 25°C. These experiments were conducted at ~ 17 °C, hence 15 mW (11.8 dBm) output power was expected. Its spectrum is shown in Figure 8.12.

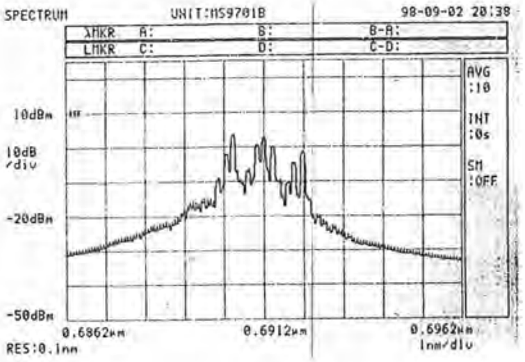


Figure 8.12: Spectrum of 690 nm laser diode source.

A 2-D image of the intensity launched into the fibre detector is illustrated in Figure 8.13. The figure shows the three main peaks, 0, -1 and +1 orders. The noise floor was limited to -70 dBm by the lower bound of the optical sensor/power meter. The peaks are Gaussian-like and sidelobes in the transverse *x*- and *x*-directions are rather pronounced. Line plots of the *x*- and *y*- directions are shown by the solid lines in Figure 8.14. The truncation ratios in *x*- and *y*-directions are approximately 2 and 1.7 respectively. The fibre used for detection had a cut-off wavelength of ~ 610 nm. Assuming a numerical aperture (N.A.) of 0.1 for the detection fibre, the spot radius was thus $w_{r2} = 2.77$ μm. The output of the laser diode should be well-confined within the input fibre. Thus, it has a high N.A. and normalised velocity parameter, V. Taking N.A. = 0.12 and V = 2 and averaging Marcuse's and Snyder's mode size parameters (eqs. 3.5 and 3.6) for a 4*f* replay without clipping, the spot radius of the replay peak would be w_{r1} = 2.25 μm, if beam broadening by the lenses and hologram aperture is negligible. Using these spot radii and the respective truncation ratios in the approximation of the solution to the overlap-integral in eqs. (3.20) and (3.22), the predicted replay field profiles for a single peak by assuming a Gaussian fibre mode are depicted by the dashed lines in Figure 8.14. There is a good match around the main peaks and the sidelobes. The measured 1/e^2 diameters of the coupling intensity in the *x*- and *y*-directions are both 8 μm c.f. expected values of 7.4 and 7.9 μm, respectively.

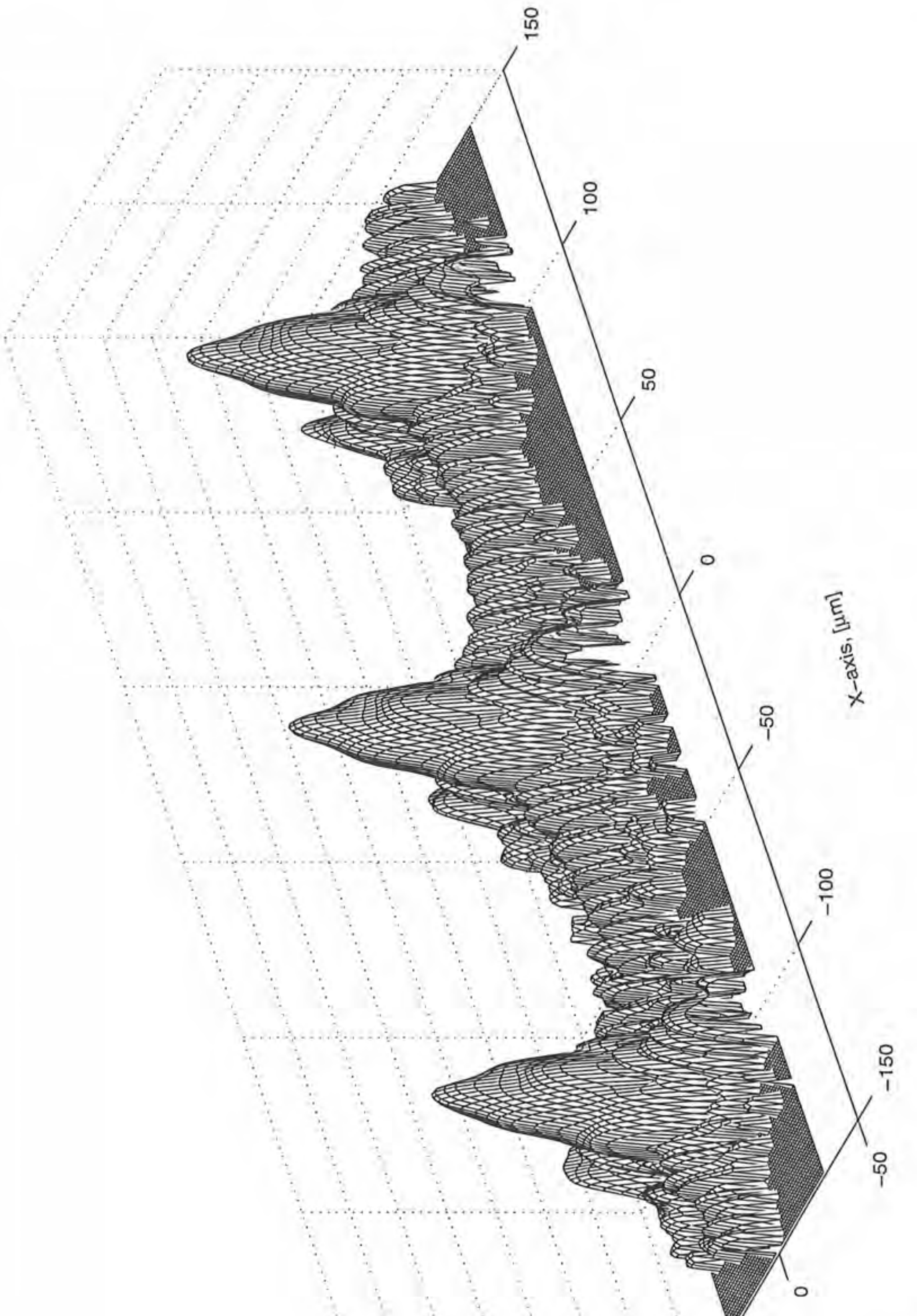


Figure 8.13: 2-D coupling intensity profile of the zero and ±1 orders of the fixed intensity grating replay.

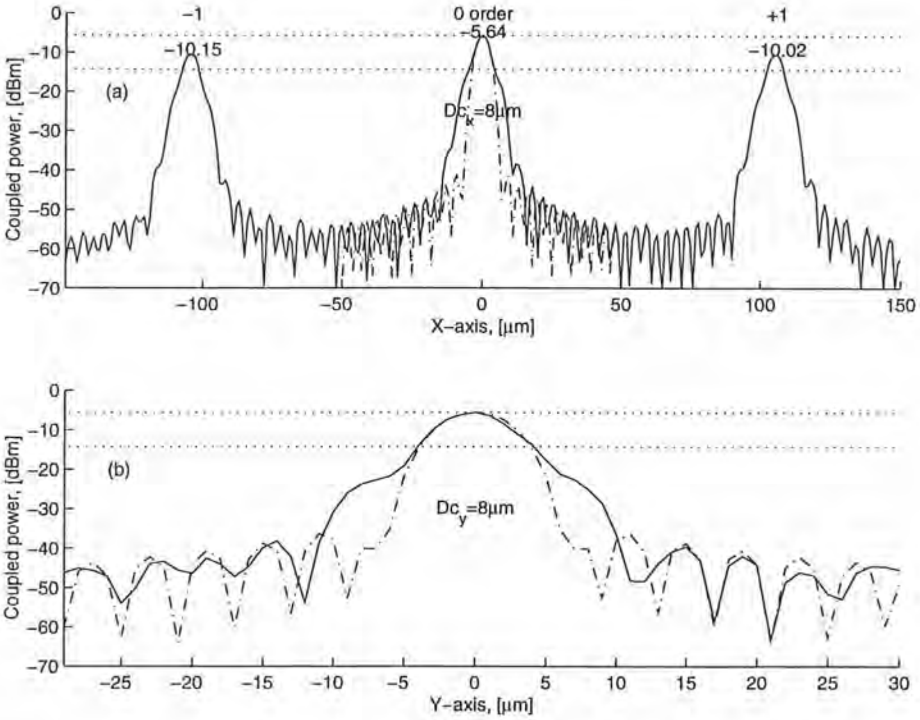


Figure 8.14: Line scans along and orthogonal to the dispersion plane (grating line).

The true fibre mode is a Bessel function K_o within the cladding. Hence, the steep shoulders of the main peaks are broadened. The collimation and transform lenses used in these experiments have an *f/#* of 2.4. Thus, there is not much aberration control which will also contribute to beam broadening. The coupling intensity profile would be improved if a laser source with a better spectral purity were used, although smearing by chromatic dispersions has not been noticeable judging from the clear side-lobe ripples and identical beam profile within the central 70 × 300 μm^2 area.

Under ideal beam collimation and laser spectrum, a 50:50 pixel/dead-space intensity grating will produce the zero and ±1 orders at -6 and -10 dB of the input power, respectively. The GaAs optical sensor used was calibrated at 750 nm IR and gave a -5 dB responsitivity at 690 nm. The power loss due to the hologram aperturing at $\gamma_x = 2$ and $\gamma_y = 1.7$ is 1.2 dB according to $[erf(\gamma_x/\sqrt{2}) \times erf(\gamma_y\sqrt{2})]^2$. Hence, the measured zero and ±1 diffraction intensity was ~ 5.6 dB less than the theoretical values. This may be due to imperfect lens transmission, aberrations, back reflection and fibre coupling. A single line scan approximately along the grating line direction covering the first 14 diffraction orders and the zero spot of the intensity grating is shown in Figure 8.15. It is apparent that the sidelobes will set the crosstalk isolation limits and their 'envelope' level is determined by the Gaussian clipping factor.

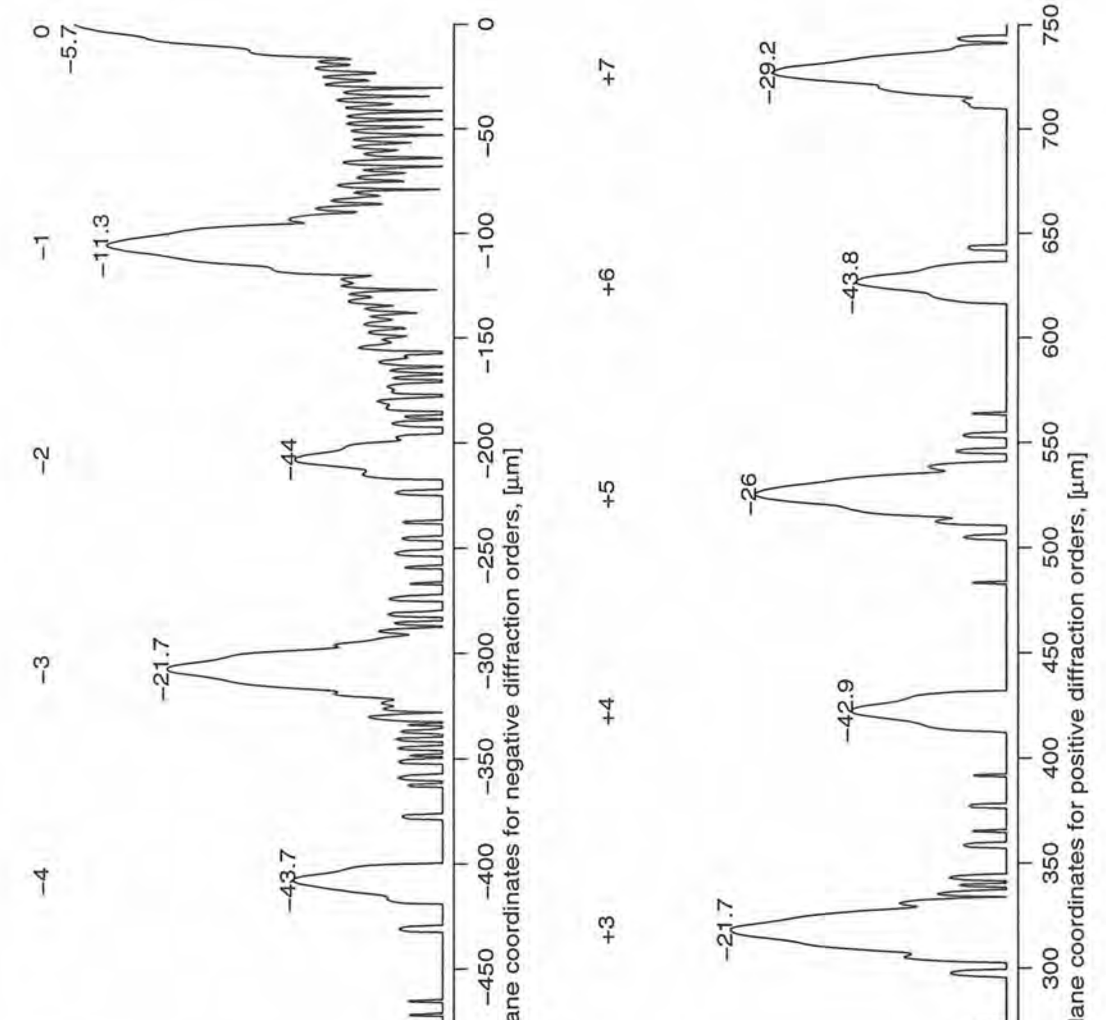


Figure 8.15: Line scan across 14 + zero orders of the grating replay.

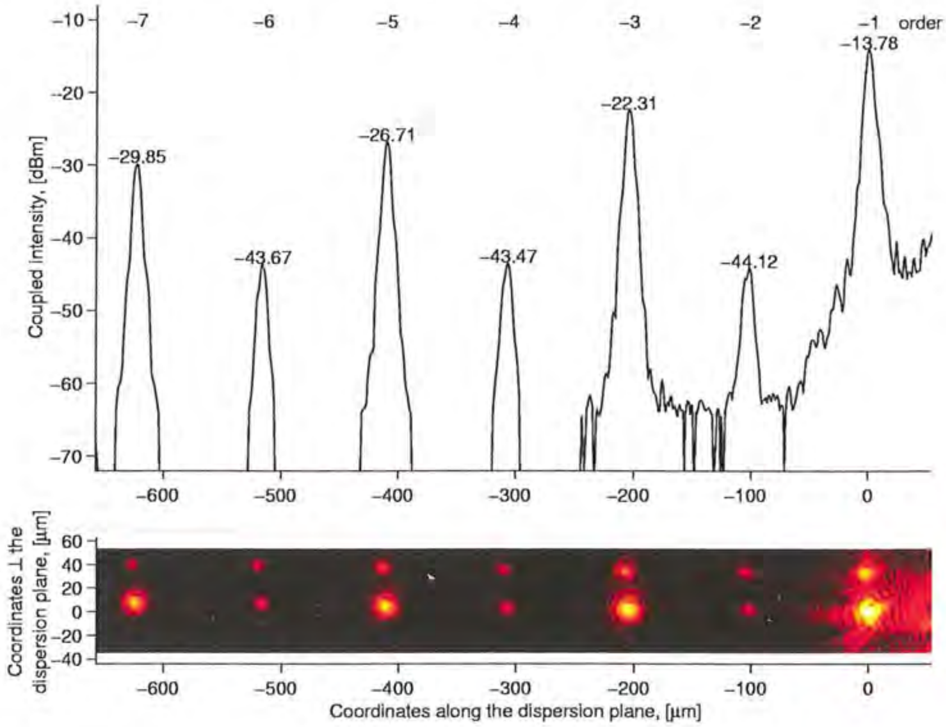


Figure 8.16: Line scan along the grating line and 2-D image of the fixed grating in reflection.

The same glass grating was also used to evaluate the optical set-up for reflective hologram replay (see §8.5.2 for optical set-up). A line scan across the dispersion plane encompassing -5 to -1 orders is shown in the top section of Figure 8.16. Two points are apparent:

i) The set-up was adequate for taking hologram replay in reflection and

ii) The Al pixels were of good optical quality as the diffraction peaks were approximately as intense as the diffracted peaks in transmission.

The -1 order at ~ 100 μm from the zero order was too close to be scanned by the output fibre without distorting the illumination. Hence, the beam profile of the first order was distorted and considerably less power was measured at an on-beam-axis location. Also evident from Figure 8.15 is the multiple paths in the collimation and far-field Fourier transform using the reflection set-up.

Using the transmittance function shown in Figure 8.17 with the theoretical analysis in §4.3 produces theoretical replays that agree with the transmissive and reflective measurements. Figure 8.18 (a) and (b) show the theoretical central 15 replay orders using the fixed glass pattern as a transmissive and reflective intensity grating, respectively. The two plots are to be compared with experimental measurements in Figure 8.15 and Figure 8.16. The loss mechanisms are not accounted for here. Rather, the relative intensities between replay orders are within 1 ~ 2 dB of expected values. Error bars are not shown in the experimental results as they are small (~ 0.1 dB) and ten data points were collected for averaging. The theoretical plot shows that all even orders will be present in the far field replay if the pixel/dead-space ratio is not 1.

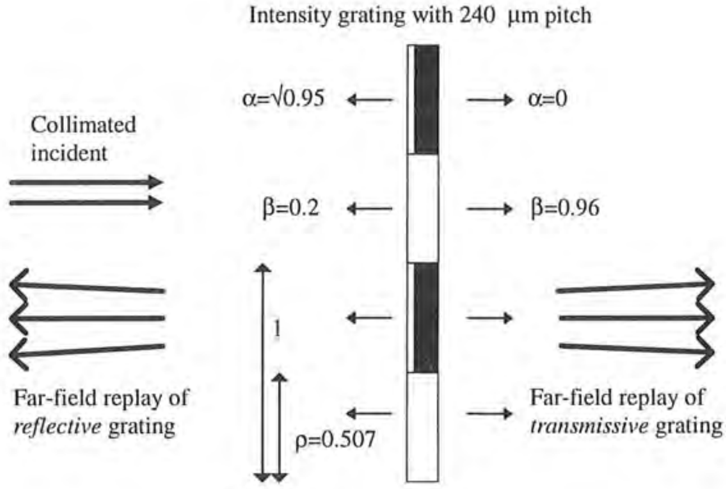


Figure 8.17: Estimated transmittance function of the fixed grating in reflection and transmisison.

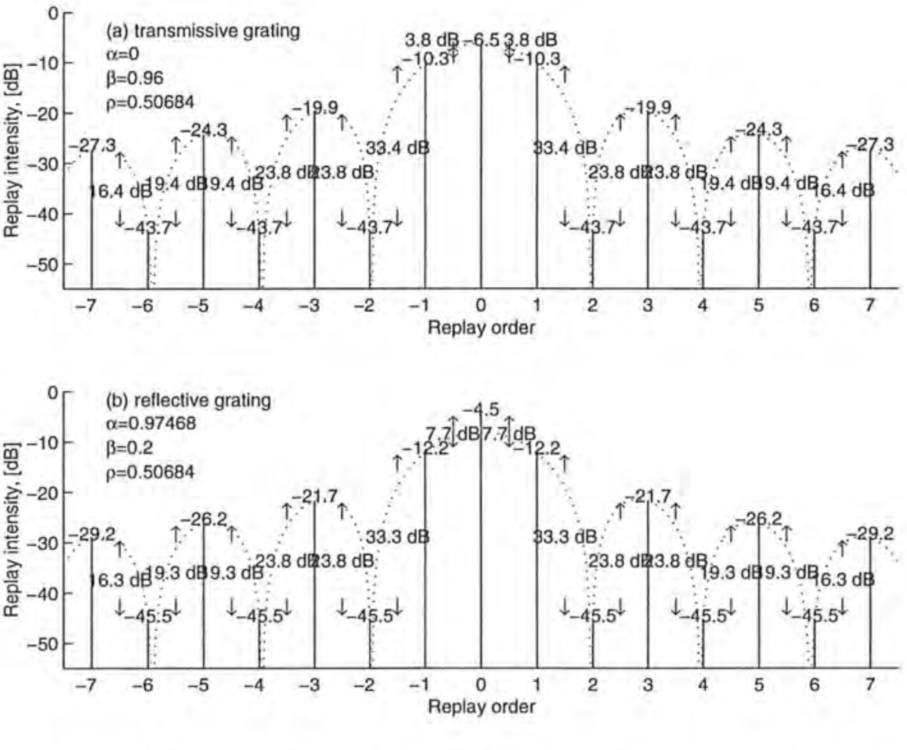


Figure 8.18: Theoretical grating replay in transmission and reflection modes.

8.5.2 Reflective binary-phase holographic operation[*]

The challenge was to measure the far field replay of holograms encoded using a FLC on silicon backplane SLM, in a way that would ensure normal incidence on the electro-optic medium and without the use of a non-polarising beam splitter cube. As the SLM has to be used in reflection mode, a single lens to collimate the fibre launch and perform the far-field transform of the hologram will suffice. Based on the pixel pitch, d = 20 μm, focal length, f = 36 mm and wavelength, λ = 690 nm, a straight forward $2f$ collimation/transform of the optical field emitted from the cleaved end of a standard SMF fibre would result in the fibre cladding (125 μm) and the bare-fibre mount posing severe constraints on the measurable size of the replay plane. At this wavelength, the Gaussian spot diameter is ~ 5 μm and there is considerable space saving if the focus spot is relayed to the $2f$ hologram collimation/replay system by a standard $4f$ system of identical lenses. The scanning fibre was glued onto a V-groove etched on the curved facet of a 15 mm-radius quarter disc. This ensures normal incidence and normal detection without the output fibre blocking the incoming beam. The complete $6f$ optical set-up is as shown in Figure 8.19.

The disadvantage of this set-up is that the first ~ 100 μm including the zero order at the replay plane is inaccessible. This would rule out any measurements of the first order intensity for replay fractions smaller than 1/12. In addition, the lack of zero-order intensity data makes the evaluation of FLC/SLM as a holographic device harder.

The 2/21b SLM used in this experiment was filled with CDDR8 [111]. The switching time and tilt angle of this FLC mixture are 500 μs and ~ 34° [110], respectively. The SLM had ~ 220 (4.4 mm) continuous pixels with ~ 2/3 of pixel length (4.5 mm) working on the far-left section of the array. Internal pixel shorts result in eight pairs of pixel switching together depending on whether the two driven states are similar or opposite. In the worst case, a hologram with replay fraction 1/2 is displayed such that every neighbouring pixel is of the opposite state. Each of the eight pairs could be pulled low or high depending on which electrical state the localised shorts favour. Other hologram patterns would require more than two pixels of the same optical states. Hence, the additional loss due to these pixels could not be quantified for a general case.

---

[*] Fibre mounting on the quarter disc and optical set-up/alignment were assisted by I. Manolis

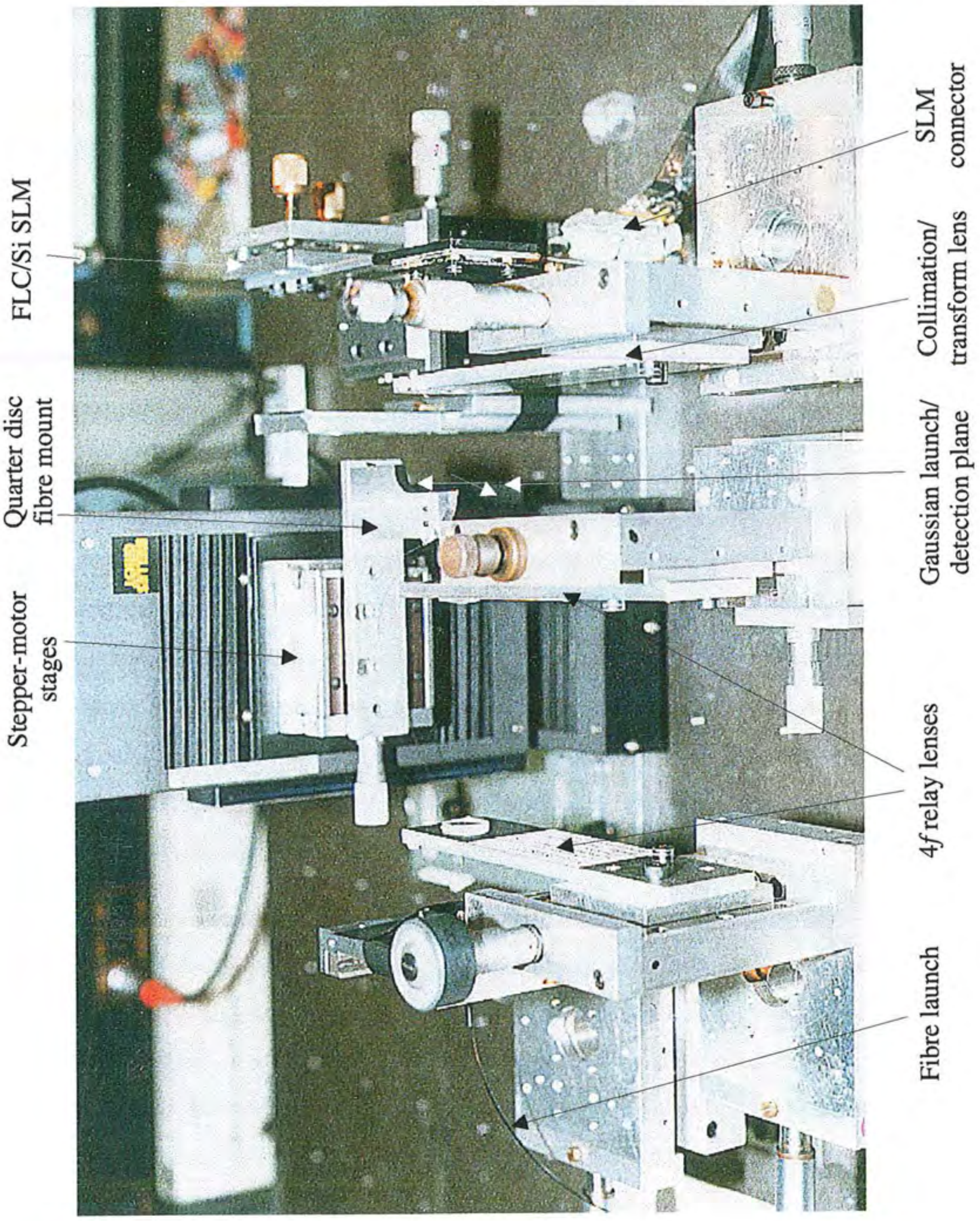


Figure 8.19: Photograph of the experimental set-up for measuring reflective hologram replays.

Due to the slowness of the fibre-scanning set-up to collect optical power corresponding to a positional co-ordinate, a safer approach at driving the SLM was adopted. The SLM was updated with a new frame and then its inverse frame at 100 Hz. The frame update was limited by the slow interface. The data were collected at ~ 0.4 s intervals using three sample points for averaging. This was not synchronised with hologram updates, hence it was possible that a particular sample point could have been collected while the LC was being refreshed and not fully switched. Running the SLM at 100 Hz with a 2 kHz CDDR8 material did minimise these erroneous data collections.

The aims of these experiments were

i) to verify the first order replay intensity versus its replay fraction (see §4.3.2) and

ii) to verify the relative peak intensities for several higher orders of a hologram (especially non-grating holograms) across several replay replications (see §4.3.1).

In order to achieve any realistic comparison with diffraction theory, SLM dead-space determination is of the utmost importance. With the SLM being blanked high and low at ~ 1 kHz so as to minimise LC scattering, several diffraction orders due to dead-space and pixel reflectance were mapped using the fibre-scan set-up. With the above-mentioned λ, f and d, the separation between two successive diffraction orders is ~ 1.24 mm. With this separation, only the first three negative orders are within the diffraction-limited region (±5°) of the transform lens. The beam broadening due to chromatic dispersions was corrected for and ~ 82% fill-factor was found to scale the distribution of intensity peaks well.

Other parameters were taken from theoretical analyses (see §3.4, §4.2 and §4.3) and parallel measurements by Roses partners (e.g., θ_o). All factors affecting the measured intensity values that could be quantified are given in Table 8.5. Some of the loss components listed are at best a guide. For example, it was impossible to know the exact size of the hologram illuminated, the spot radius of the heavily multi-moded light source and its collimated beam size. Others like Δn and Al pixel reflectance at the operating wavelength, LC cell thickness, were not calibrated for, instead typical/expected values were assumed.

Power/loss Component	Loss expression & assumptions	Power/loss value
A) Hologram-dependent losses		
I) ***phase quantisation*** - number of phase levels - required routing fraction	$sinc^2(\pi/2)/sinc^2(\pi/x_o)$ for $\sigma = x/x_o$, first order *digital* diffraction efficiency for binary holograms of replay fraction σ	i) $\sigma = 1/2$: 0 dB ii) $\sigma = 1/4$: -3 dB iii) $\sigma = 1/8$: -3.7 dB

II) ***pixellation and dead-space*** - SLM pixel isolation	$\rho^2 \mathit{sinc}^2(\rho\pi\sigma)$, ρ = fill factor @ 82%	i) $\sigma = 1/2$: -4.27 dB ii) $\sigma = 1/4$: -2.33 dB ii) $\sigma = 1/8$: -1.87 dB
B) Hologram-independent losses		
III) ***retardation efficiency*** -non-$\lambda/2$ LC cell configuration	$\sin^2(\pi\Delta n 2T/\lambda)$, $\Delta n = 0.145$ [91] @ λ = 690 nm and T = LC cell thickness @ 1.95 μm	-5.4 dB
IV) ***phase efficiency of electro-optic modulation*** - non-$\pi/2$ FLC switching angle	$\sin^2(\theta_o)$, θ_o = FLC switch angle @ 68°[110]	-0.66 dB
V) ***Gaussian beam aperturing***	$[\mathit{erf}(\gamma_x/\sqrt{2}) \times \mathit{erf}(\gamma_y/\sqrt{2})]^2$ @ $\gamma_x = 1.26$, $\gamma_y = 1.29$	-3.96 dB
VI) ***sundry losses*** in the optical system including lens transmission and back reflections	3 MgF_2 coated lenses with 6 reflecting surfaces each giving 1.5% loss [112], SLM front glass with 2 × 4% loss and 95% reflectance of Al SLM pixels.	1 dB
VII) ***fibre coupling loss*** -losses unaccounted for	Replay beam/fibre axial tilts, strains on the output fibre glued onto quarter disc, chromatic dispersions of source, lens aberrations and system misalignment.	~ 2.3 dB
Total hologram-independent losses		13.32
C) input power: The λ = 690 nm laser diode has 15 mW output at 17 °C. The optical power meter has a -5 dB responsitivity at 750 nm calibration.		6.8 dBm
D) Expected power	After taking into account the major loss components in the above	i) σ=1/2: -10.8 dBm ii) σ=1/4: -11.9 dBm iii)σ=1/8: -12.0 dBm
E) Detected power	Actual measured first order intensity values	i) σ=1/2: -10.8 dBm ii) σ=1/4: -11.9 dBm iii)σ=1/8: -13.6 dBm

Table 8.5: Losses in using Roses 2/21b SLM for binary hologram encoding.

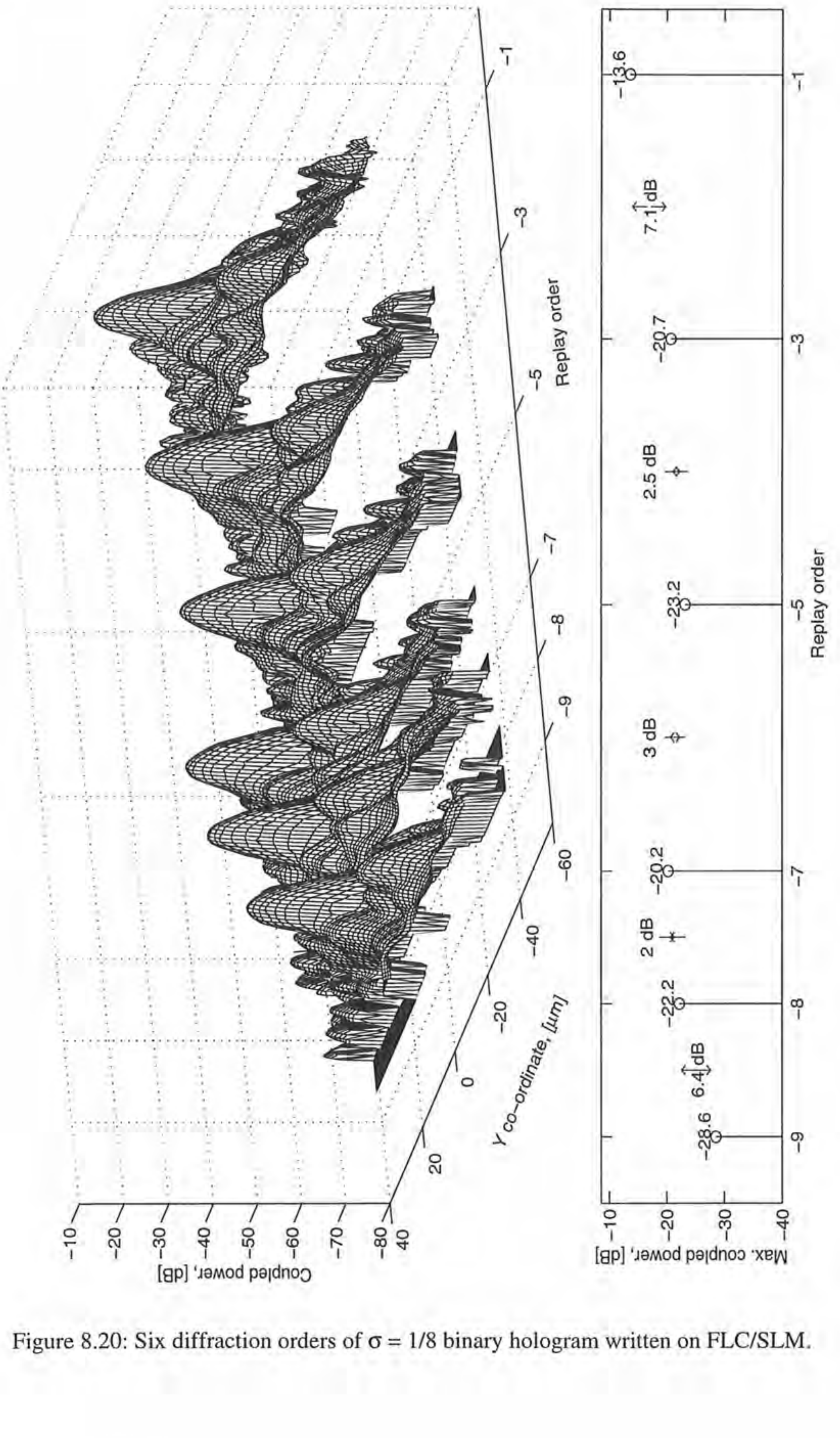


Figure 8.20: Six diffraction orders of σ = 1/8 binary hologram written on FLC/SLM.

The intensity values given in **"Detected power"** were the highest intensity from scanning a $80 \times 80\ \mu m^2$ at 1 μm resolution corresponding to that replay fraction. These values should be very close to the on-beam-axis intensities without any lateral offsets. An example of replay beam profile is shown in Figure 8.20. It encompasses six consecutive non-zero-intensity orders (at 1/8, 3/8, 5/8, 7/8, 8/8 and 9/8 fractional co-ordinates) of 1/8 hologram replay fraction.

Without any curve-fitting, the diameter of both the x- and y-direction coupling intensity profiles is ~ 10 μm for all orders except -1 which is ~ 11 μm. From the analysis in §3.4, there could be 32% beam broadening due to severe Gaussian clipping and hence 9.4 μm was expected. The main point is, these measurements were not an order of magnitude deviations from expectations! From the results of §4.3.2, gratings with longer length should return poorer first order efficiency. The first order intensity of $\sigma = 1/8$ was distorted due to the output fibre partially obstructing the launch light. Multiple reflections in a $2f$ optical configuration may degrade the high crosstalk isolations inherent in holographic free-space systems with single-mode fibre or waveguide output filtering.

8.6 Drive schemes issues for holographic applications

In order to prevent electro-chemical degradation of FLCs, multi-field and multi-time-slot addressing schemes are often employed for displays based SLMs. These schemes utilise the slow response of our eyes to average out an intensity image for continuous viewing. SLMs for coherent optical applications especially those carrying high data-rate traffic would fare miserably with averaging schemes. However, as far as writing holograms on SLMs is concerned, the symmetry features due to phase and spatial eigen sets can be exploited. In the simplest case, the true and the inverse frames of a binary hologram both give identical far-field intensity patterns [115].

DC balance addressing of holograms can be sped up due to these inherent symmetries of a known hologram frame. However, addressing these linear-independent groups rather than row-by-row updates is not an urgency considering LC rather than the drive electronics is the bottleneck in readying a frame of hologram.

In these subsections, simulation and experimental results of hologram refreshing in order to maintain a fixed optical link are presented. These are most relevant to the current application areas of FLC/Si holograms such as WDM and tributary traffic management. For these applications, the moderate reconfiguration rate of the optical interconnection is not critical, rather maintaining the interconnection for very long duration and small polarisation dependence are the key requirements.

8.6.1 Non-DC balanced refreshing to maintain holograms for long periods

Bistable FLCs will help to alleviate the long term electro-chemical degradation problems of LC, however, no one is certain that leaving it in one state for a long period of time will not cause image sticking when the opposite state is addressed. One of the solutions is to periodically flip a group of pixels to their inverse states and flip them back after the completion of switching.

A numerical iteration has been performed with 65 sets of 128 × 128 binary phase-only holograms. The replay of these holograms represents the routing to one of the 65 output ports, arranged non-symmetrically about the central grid point [88]. In each simulation, 1, 2, 4, 8, 16, 32, 64 and 128 pixels in a continuous section of a row have been updated to the inverse states and the first order replay intensity calculated at intermediate times, assuming a linear phase-time response and plane-wave input.

As soon as switching is complete, the same section is re-written to its original states. This ensures that the updating would not result in two sub-frames which belong to two orthogonal hologram sets giving the same intensity replay pattern with a phase difference. Using the *time-delayed* inverse-frame refreshing scheme, a loss as large as 6 dB will be incurred when the frame has been updated by a quarter. The *global* inverse-frame updating would have resulted in a blank frame half way between the two switched states crossing over.

Re-writing the original states in each section helps in DC balancing the pixels. Between a section being completely updated and when it is next addressed, the simulation assumes perfect bistability without any fields applied. Boosting the voltages periodically is one technique which has been investigated for restoring the desired optical modulations.

The simulation results obtained indicate that an optical ripple loss of up to 0.14 dB is incurred if a section as large as a whole row were to be updated to its opposite states. This is valid only for 128 × 128 holograms. For 32 × 32 holograms, updating a single row to its inverse states results in an optical ripple loss of 0.54 dB. Figure 8.21 shows the temporal intensity plots of two row updates for various fractional-row updates. The average value determined from all 65 holograms is used to extrapolate the optical ripple as a function of the ratio between the number of pixels in an updated section and the 2D hologram SBWP. This is illustrated in Figure 8.22

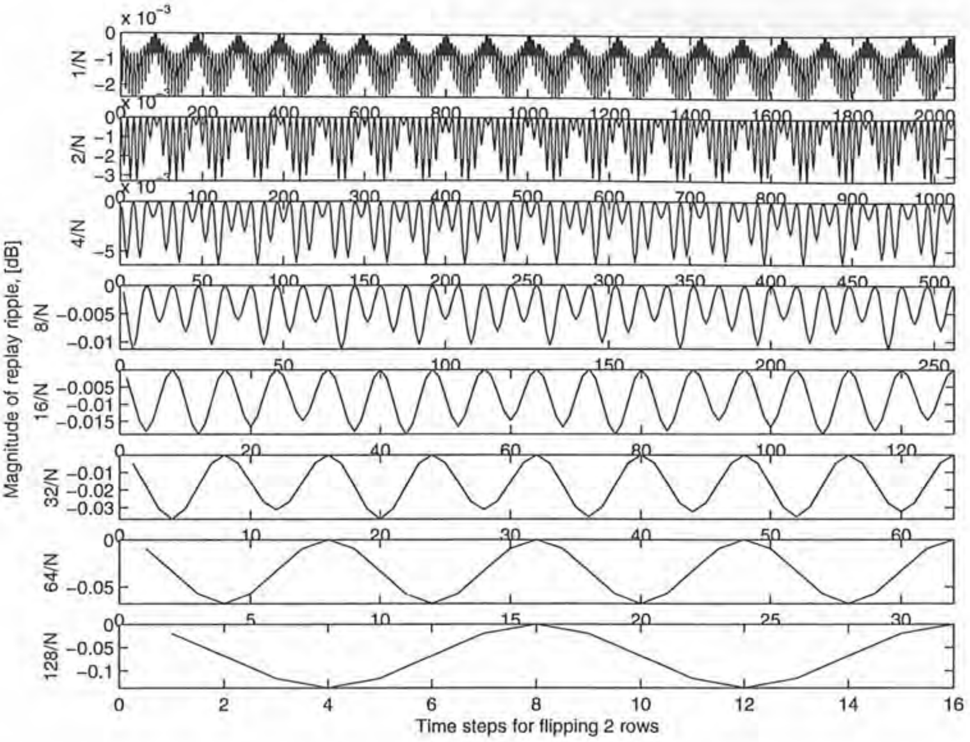


Figure 8.21: Simulated temporal optical response of a 128 × 128 binary hologram with plane-wave input.

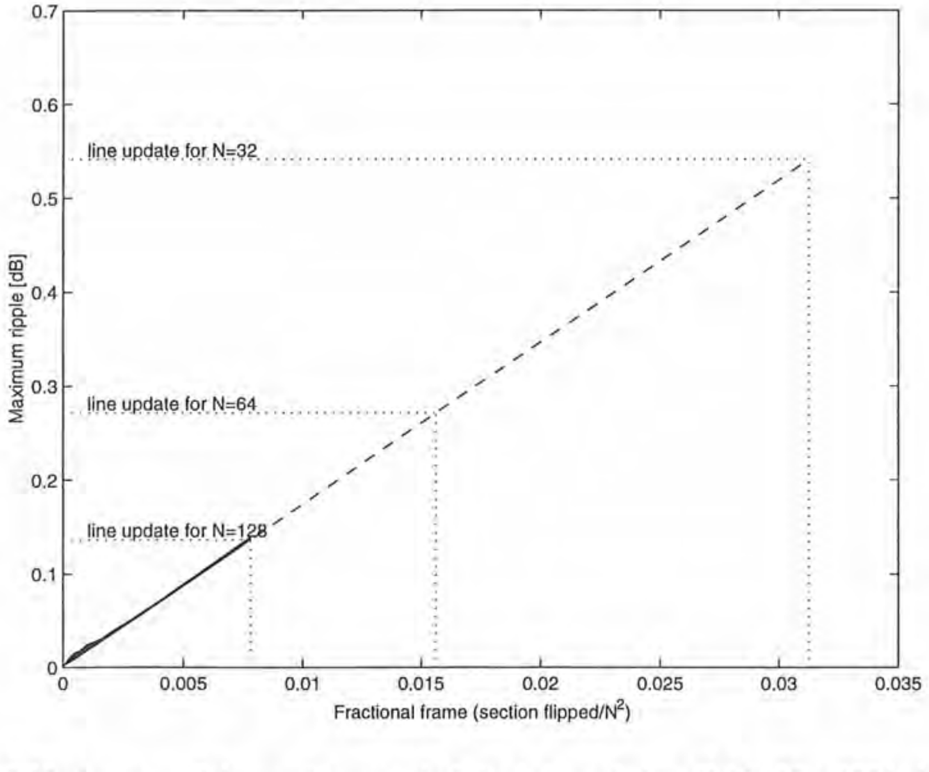


Figure 8.22: Average of the maximum optical ripple in updating all 65 128 × 128 holograms and extrapolated values for other fractional updates.

There was no attempt to vigorously select the best set of orthogonal sections to be updated. Line updating probably represents the simplest of all updating schemes. When the SLM is not illuminated uniformly, updating a hologram frame with row-by-row flipping will not give the same optical ripple noise. The simulation results of various fractional-row flipping and back for a Gaussian distribution input is shown in Figure 8.23 for a 32 × 32 binary phase-only hologram. Obviously, the ripple noise increases when a section of hologram close to the Gaussian beam axis is updated. The optical response becomes more complicated when dead-space, phase-mismatch and non-$\lambda/2$ wave retardation are considered.

An experiment of 'flipping and flipping back' a single pixel of a linear binary phase-only hologram was conducted. The 1-D case represents the simplest evaluation of the optical loss when a fraction of the hologram frame is in the 'wrong' state. The hologram corresponds to $\sigma = 1/2$ (1 pixel on and 1 pixel off) and was recorded onto a Roses SLM with fill-factor ~ 82% [see 8.5.2 for more details of the experimental set-up and device]. Approximately ten sample points were collected within each flip interval of ~ 324 ms. The results, as shown in Figure 8.24, indicate a maximum of 0.2 dB power loss for 1/220 of a frame updated in flip & flip-back manner. This ripple size is ~ 30% worse than the 4/(32 × 32) 2-D simulation results in Figure 8.23, possibly due to a more severe truncation and cross-linked/defective pixels.

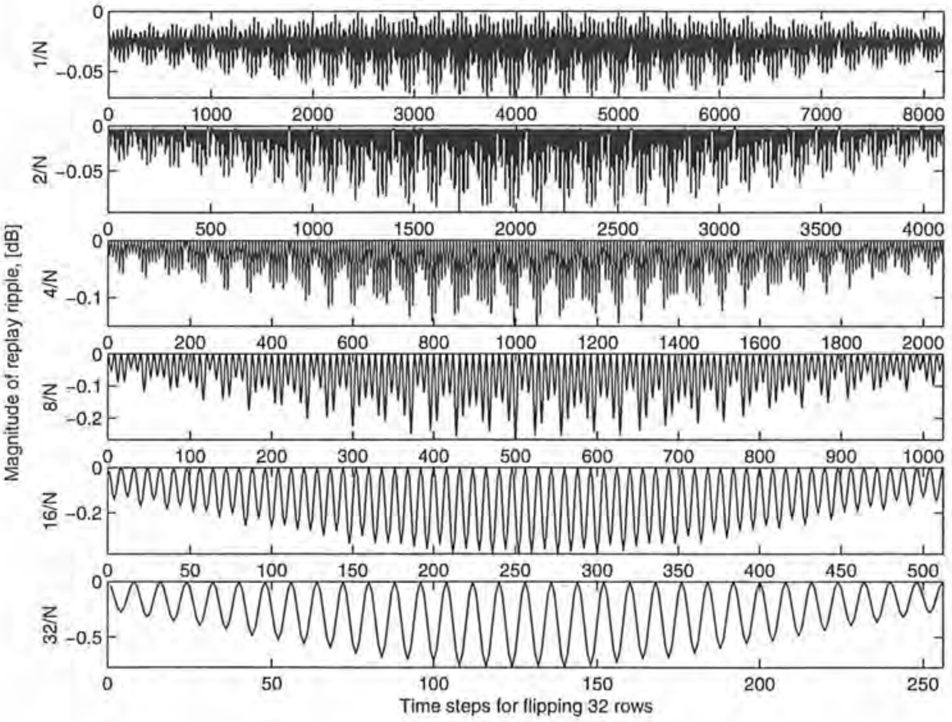


Figure 8.23: Simulated temporal optical ripples for 32 × 32 binary holograms illuminated with a Gaussian input truncated at $\gamma = 2$.

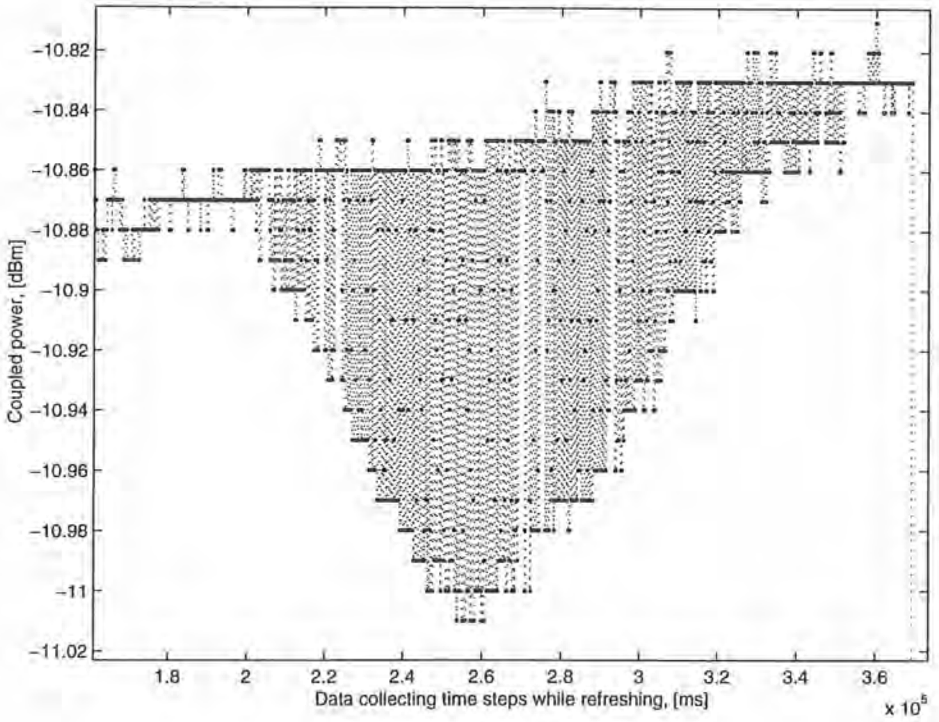


Figure 8.24: Experimental optical ripple noise for $\sigma = 1/2$ binary phase-only hologram.

8.6.2 DC-balance refreshing to maintain holograms for long periods

If the bistability of FLCs is poor, *flip & flip-back* scheme may not be suitable. Instead, a scrolling method whereby the 2D or 1D hologram is continuously being shifted by several rows or columns in the horizontal or vertical directions. This uses the spatial eigen modes of a base hologram. For example, for a hologram having a period of 20 pixels will require 20 spatial eigen modes. The storage for 20 sub-frames may be required unless the lateral shift is implemented as internal chip logic. This is particularly easy for a 1D SLM where there is already a shift-register on chip acting as the memory for the current frame. The updating of all the eigen sets is illustrated using a four pixel grating in Figure 8.25. The implementation on a 2D SLM will be difficult unless a global update can be carried out. Under typical matrix addressing schemes, each row has to be updated sequentially. Hence, both the phase mismatches due to row-scrolling and the integrity of the column routing pattern can produce large ripples. The previous method of flipping each row momentarily seems attractive here.

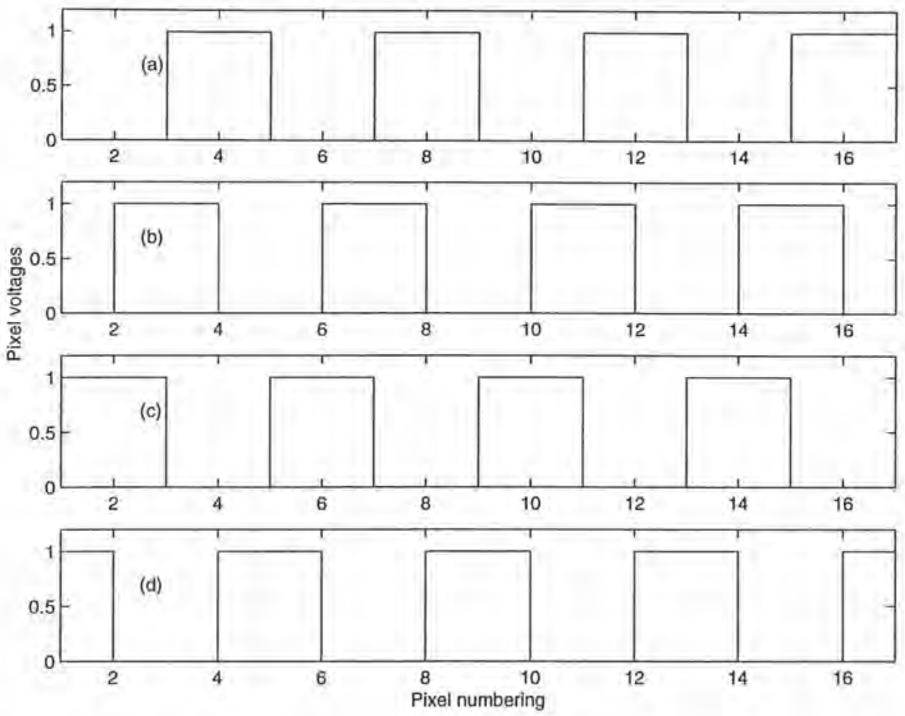


Figure 8.25: Scrolling of 1D hologram to attain DC-balanced refreshing.

The experimental verification of this scheme was performed using the same SLM and hologram pattern. Each pixel of the linear array was flipped in turn. When approximately one side of the collimated Gaussian beam was presented with a hologram and the other side with its inverse frame, the optical loss was the maximum. At this stage, the shift corresponds to half of the hologram period and the peak orders of the two half frames are out of phase by π radians. Hence, the interference produced ~ 15 dB lower coupling power. These results are shown in Figure 8.26. Approximately 320 of the 540 pixels are masked, shown by the unchanged power level, corresponding to that of a *static* frame.

The optical replay has much lower ripple magnitude than that of a digital replay assuming perfect binary phase hologram as depicted in Figure 8.27. The cross-linked pixels made the display of an exact inverse frame impossible. As a result the cancellation of replay power was reduced. For all hologram patterns, the scrolling of complete hologram frame involves an intermediate combination of half a true frame and its shifted eigen set. Depending on the hologram period and the number of pixels to be shifted (≥1), this probably gives << π phase difference for the two corresponding replay peaks. Hence, depending on the patterns being displayed, the optical disruption can be minimal.

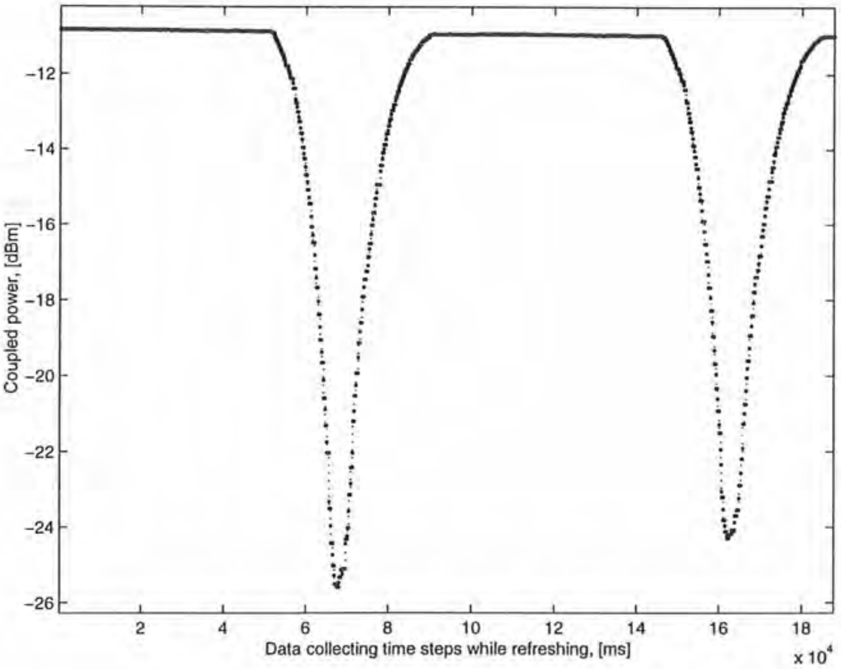


Figure 8.26: Experimental temporal optical ripples with scrolling one pixel of σ = 1/2 binary hologram frame.

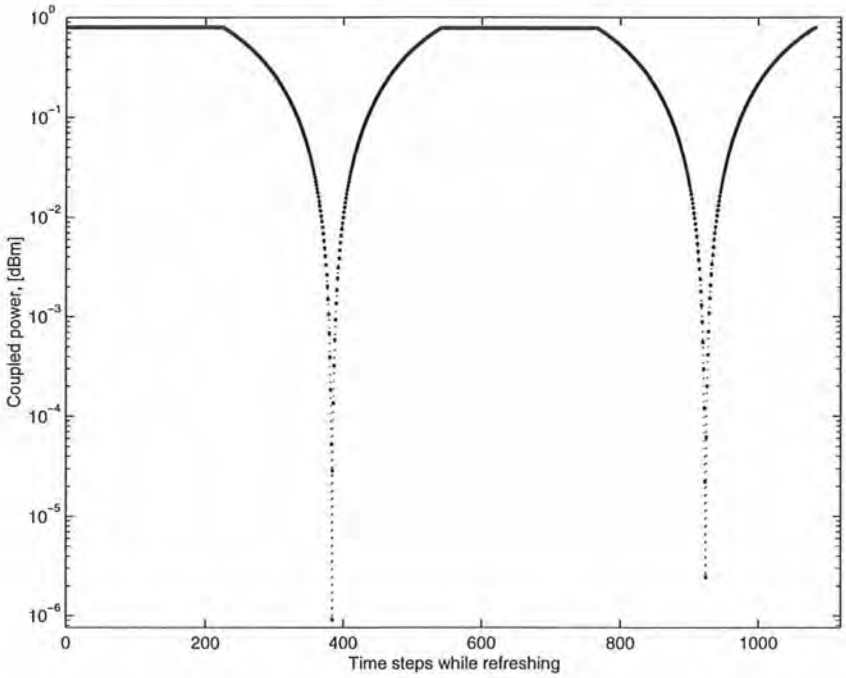


Figure 8.27: Simulated temporal optical ripples with scrolling one pixel of σ = 1/2 binary hologram frame.

8.7 Conclusions

SLMs fabricated on Si backplanes and integrated with FLC electro-optic modulation media were probed and tested for their electrical functionality and optical characteristics. The procedures laid down here are typical for FLC/SLM assembly. To obtain high electrical and optical quality SLMs, not only the backplane has to be designed for high speed interface with intelligent features incorporated, the post-processing/assembly, liquid crystal alignment techniques and driving schemes would have to be improved to reduce failures at each stage. The critical stage seems to be the deposition/etching of a third layer metal as well as the surface quality of the quaternary pixels.

The diffraction intensity measurements were intended to simulate a 1:*N* SM fibre to fibre switch with *infinite* number of output port locations; where the scanning fibre corresponds to the first order co-ordinate, it is the insertion loss of that particular hologram recorded onto a FLC/SLM of known parameters; where it picks up the higher order or the side-lobe power at the replay plane, it gives the likely crosstalk noise present. These measurements have helped to verify the theoretical analyses presented in §3 and §4 in both the insertion loss and crosstalk aspects. Side-lobe crosstalk power of <-50 dB and <-30 dB have been measured using Gaussian illumination with $\gamma \approx 2$ truncation on a fixed grating and $\gamma \approx 1.2$ on a dynamic binary hologram, respectively.

By accounting for the losses in the measurements using imperfect dynamic SLMs, the theoretical routing hologram analyses are proven reasonable. Furthermore, by comparing these losses with a near-ideal fixed grating, scattering by the LC layer and Al reflectors is found to be minimal. In fact, the far field replay of dynamic holograms can be as good as that of fixed gratings if schemes to update and refresh holograms on SLMs, whilst not compromising the electro-chemical integrity of the FLC materials are available. The approach to producing high quality holographic devices is probably correct although some processing techniques have not been perfected yet.

At this stage of development, FLC/Si binary holograms are best suited for switching optical links requiring low reconfiguration rates, and which may remain unchanged for a lengthy period. Flipping & flipping back and lateral shifting schemes in order to refresh a near 'static' binary hologram frame are illustrated with experimental results. With $\sigma = 1/2$ binary hologram displayed, the corresponding optical ripples are 0.2 and 15 dB, respectively. The former scheme which is non-DC balanced and requires FLC bistability gives an additional power loss independent of the hologram pattern. These results have been obtained using a fully assembled and packaged binary SLM from the Roses project wafer run.

CHAPTER 9

Conclusions and further work

9.1 Conclusions

In this dissertation, a full analysis of dynamic phase-only holography used in free-space interconnects with numerical and experimental verifications has been presented. Original results have been described for the following ten areas:

- a skip-rotate rule in locating hologram replay peaks,
- insertion loss due to spatial and phase quantisation effects,
- a deterministic hologram generation algorithm,
- on-beam-axis, lateral and angular offset coupling intensity,
- off-beam-axis crosstalk due to replay sidelobes,
- number of hologram repeats to realise crosstalk benefits,
- limitations of multi-level CSLC holograms,
- Silicon backplane SLMs for holographic applications,
- experimental binary hologram replay, and
- temporal response due to hologram refreshing.

The theme of this dissertation has been the investigation of what a single-peak routing hologram produces given realistic device and system parameters. Taking the routing hologram on its own without any consideration of the optical systems to replay it, the distribution of the replay peaks (first and high orders) is predictable and consistent for a given number of phase levels and replay fraction. Using the *modulo-1* skip-rotate rule in conjunction with fractional hologram representation, all non-zero-intensity replay peaks can be located iteratively. Furthermore, by considering the inadequate resolution of a numerical replay and hence summing of several replay orders at each numerical replay location, each replay peak intensity can be predicted. Incorporating the existing results of finite pixel width and pixel periodicity of a real encoding device (e.g., pixellated SLM) a composite sinc squared expression, giving exactly what the replay intensity of each diffraction order should be, is obtained. The expression agrees with numerical *fft* replay by deliberately increasing the sampling points for each phase element. The results are at least valid for several replay replications within the paraxial diffraction domain. In general, the binary first order diffraction efficiency drops from about 33% to 17% when the normalised spatial frequency is increased from ≈ 0 to $\approx \pm 1/2$. This assumes that the routing holograms are recorded on a 1-D SLM having perfect modulation characteristics except pixel fill-factor (90%).

A nearly similar skip-rotate rule, *modulo-Λ*, where Λ is the effective grating pitch for 1-D or 2-D routing holograms, together with an algorithm to perform the rule have been implemented to generate these one-peak routing holograms. Of course, the peak intensities and locations are predictable. The algorithm only requires the number of phase levels and the replay fraction as inputs. Although this algorithm is still limited to a single-peak hologram generation, it is adequate for single-port optical routing. This technique is attractive due to its speed (typically < 1s) and known (though uncontrollable) high order noise peaks. Nonetheless, optical routing could rely on the spatial filtering of fibres to reject these noise peaks.

With a realistic view of a typical 4*f* holographic system for optical switching, the infinitely extending Gaussian tails (at least theoretically) from the collimated beam could not fit within the hologram aperture. Therefore, there is a trade-off between the clipping of these tails and the diffraction effects from the aperture. Using a single parameter, γ where γ is the ratio of the aperture to Gaussian radius, an asymptotic approximation of the replay peak profile (i.e., the convolution of Gaussian distribution and sinc aperture replay) is obtained. Yet, the more useful description of aperture edge diffraction effects is the coupling intensity into the output fibres. This can be found in a similar way with consideration of the output fibre offsets and tilts. The on-beam-axis coupling power is given by the fourth power of an *erf* function due to the fibre mode distortion and power clipping. For $\gamma > 2$, the coupling loss is less than 1 dB. Off-beam-axis crosstalk is predominantly due to the oscillating nature of the sinc function. The peak crosstalk power is ~ -20γ dB within the valid region of the asymptotic approximation (e.g., 25 μm to 100 μm). Hence, a 40 dB crosstalk isolation is achievable by having $2w$ of the collimated Gaussian beam within the aperture. Between 40 and 80 dB, the number of hologram repeats in order to adequately resolve the replay is 2γ.

Typically achievable four-phase CSLC holograms have been found to have inferior insertion loss and no better crosstalk isolation figures compared to binary FLC holograms. As well as being polarisation sensitive, four-level hologram requires the switching angle of the LCs to be larger than 67.5° to have any throughput advantage. With this switching angle, the additional loss in the binary hologram loss is 0.7 dB. Hence, it is not worth increasing the complexity of the digital binary addressing of SLMs to enable four-level modulation unless an asymmetric replay is desired.

A silicon backplane has been designed with optimum-quality pixels both to meet the requirements of a demonstrator switch (undertaken by other project partners) and experimental verification of the hologram analysis. The 540 × 1 binary array has to be patterned with cold-sputtered pixels. The SLM fabrication approach is novel and its success as an optical modulator is critically dependent on yield issues of pixel gap, pixel cross-links and electrical contacts with the silicon drivers. The project has turned out several CDDR 8 filled binary SLMs with ~ 3.6 μm dead-space for a designed 20 μm pixel pitch. Better SLMs are expected with a new batch backplanes processed using plasma-enhanced dry etching.

High quality dynamic hologram replay has been demonstrated in a $2f$ switch configuration. The insertion loss of the 1/2 replay fraction is ~ 17.6 dB and about 1 dB poorer for 1/4 and 1/8 replay fractions. The results were obtained with an SLM having a much thicker LC cell for $\lambda = 690$ nm and only 220 functioning pixels at ~ 4.5 mm ($\gamma = 1.25$) usable pixel length due to LC scattering. It is almost certain that binary holograms encoded on optimised SLMs will give within 3 dB of the theoretical ~ 5 dB insertion loss for the central 20% routing fractions. The crosstalk power due to sidelobes was measured at ~ -30 dB. A similar experiment using a fixed intensity grating returned greater than 50 dB crosstalk isolation. The coupling intensity profile also agrees with theory, although not sidelobe by sidelobe. The key point is *dynamically* reconfigurable holograms produce as good first order efficiencies as the fixed intensity grating, even though they contain LC scattering, imperfect switch angle and wave-plate configuration, cross-linked pixels, poor fill-factor and severe Gaussian beam truncation.

The non-DC balanced hologram refreshing scheme consisting of flipping and flipping back a small percentage of pixels gives 0.2 dB ripple loss. The corresponding DC-balance refreshing by shifting pixels so as to produce the inverse hologram incurs more than 15 dB loss for the worst case base hologram (i.e., $\sigma = 1/2$). These figures broadly agree with simulations.

9.2 Further work

Although two key research fronts of single-mode fibre to fibre interconnects using dynamic holograms have been dealt with in this thesis, i.e., theoretical performance of hologram replay and holographic device development, there exist several areas for further analysis, device design/optimisation and proof-of-principles experiments to be conducted.

9.2.1 Single discrete description for numerical hologram replay

The intensities of the corresponding non-grating replay orders in all replay replications are scaled by a 'composite-sinc-squared' expression: eq. **(4.15)** . Each diffraction order within a replay replication necessitates a scaling expression. It is possible to build on or modify the *modulo-1* skip-rotate rule to arrive at a single expression that scales all diffraction orders within each replay replications and across all replay replications. The end expression is still a composite term and is analogous to the expression scaling the interference peaks of an N-slit intensity grating [80].

9.2.2 Single continuous description of optical hologram replay

For a phase grating having a uniform illumination, the continuous replay field can be represented by a composite-sinc expression with a variable first order location [82]. It would have been helpful to derive an analytic approximation of the replay or coupling intensity profile of a phase grating (maybe too difficult for a general phase-only hologram) illuminated with a truncated Gaussian beam. The asymptotic replay/coupling intensity approximations are only valid over a short lateral offset range. The use of these expressions in numerical convolutions with the Dirac delta function replay is limited. The numerical replay of a grating unlike a general hologram is scaled by a single composite-sinc squared expression. Hence, a further simplification of the convolution results seems more likely.

9.2.3 Hologram synthesis for multiple replay fractions

It is generally not possible to combine, by an arithmetic function, two or more single-peak holograms in order to obtain a single hologram replaying several first order peaks. One approach will be to analyse multi-peak phase holograms and determine if there is only one unique solution to each first-order combination. If there is, the reverse process of locating the replay peaks would be capable of generating deterministic multi-peak holograms.

9.2.4 Silicon design for a low dimension DMOS process

It is inevitable that lower dimension CMOS processes with reduced operating voltages have to be utilised for future SLM backplane fabrication; SLM applications for silicon chips do not command a 'noticeable' market share and hence have to ride along the micro-electronic trends. DMOS processes hold the promise of integrating standard logic operations and high voltage LC driving on a single chip. To this end, the design of conventional level shifters and buffers suitable for array replication has to be looked into.

9.2.5 Hologram efficiency measurements

An immediate expansion of this work has to be the measurement of the first order diffraction efficiencies for a large number of replay fractions. These should be conducted using a state-of-the-art binary SLM with maximum-efficiency routing holograms. The results could verify/nullify the theoretical upper limit and set the record straight regarding binary phase hologram efficiencies (certainly not 40.5%!).

9.2.6 Experiments using moderate and high number of phase levels

This work has also produced an electrically functioning 4-level modulator with parallel on-chip D/A conversion. The optical quality of the polysilicon seems to be a problem. Within the time scale of this research, there has not been a single quaternary SLM that is adequate for imaging and far-field replay experiments. If some devices become available, light transmission through the polysilicon layer, scattering and replay profile of 4-level modulation should be measured. Furthermore, using an SLM capable of displaying a large number of phases, the replay field profile can be imaged to ascertain whether any SLM technology is capable of replaying m sets of m-level phase holograms having only one replay peak within each replay replication. This grating-like replay makes available the largest area for other routing ports.

Bibliography

[1] K.C. Kao and G.A. Hockham, "Dielectric fibre surface waveguides for optical frequencies", *IEE Proc. Journal*, **133**(7), pp. 1151-1158, 1966.

[2] R.J. Mears, L. Reekie, I.M. Jauncey and D.N. Payne, "Low-noise Erbium-doped fibre amplifier operating at 1.54 μm", *Electron. Lett.*, **23**(19), pp. 1026-1028, 1987.

[3] A. Matoba, H. Okayama, R. Shibuya and T. Ishida, "Low-drive-voltage 8 × 8 Ti:LiNbO$_3$ switch with simplified tree structure", *Electron. Lett.*, **25**(2), pp. 165-166, 1989.

[4] K. Hamamoto, T. Anan, K. Komatsu, M. Sugimoto and I. Mito, "First 8 × 8 semiconductor optical matrix switches using GaAs/AlGaAs electro-optic guided-wave directional couplers", *Electron. Lett.*, **28**(5), pp. 441-443, 1992.

[5] F.B. McCormick, T.J. Cloonan, A.L. Lentine, J.M. Sasian, R.L. Morrison, M.G. Beckman, S.L. Walker, M.J. Wojcik, S.J. Hinterlong, R.J. Crisci, R.A. Novotny And H.S. Hinton, "Five-stage free-space optical switching network with field-effect transistor self-electro-optic-effect-device smart-pixel arrays", *Appl. Opt.*, **33**(8), pp. 1601-1618, 1994.

[6] C.A. Brackett, "Dense wavelength division multiplexing networks: principles and applications", *IEEE J. Selected Areas in Communications.*, **8**(6), pp. 948-964, 1990.

[7] R. Kashyap, J.R. Armitage, R. Wyatt, S.T. Davey and D.L. Williams, "All-fibre narrowband reflection gratings at 1500 nm", *Electron. Lett.*, **26**(11), pp. 730-732, 1990.

[8] A.M. Vengsarkar, P.J. Lemaire, J.B. Judkins, V. Bhatia, T. Erdogan and J.E. Sipe, "Long-period fiber gratings as band-rejection filters", *J. Lightwave Technol.*, **14**(1), pp. 58-65, 1996.

[9] C. Dragone, "An N × N optical multiplexer using a planar arrangement of two star couplers", *IEEE Photon. Technol. Lett.*, **3**(5), pp. 812-815, 1991.

[10] M. Yamada, H. Ono, T. Kanamori, S. Sudo and Y. Ohishi, "Broadband and gain-flattened amplifier composed of a 1.55 mu m- band and a 1.58 mu m-band Er^{3+}-doped fibre amplifier in a parallel configuration", *Electron. Lett.*, **33**(8), pp. 710-711, 1997.

[11] A.K. Srivastava, Y. Sun, J.W. Sulhoff, C. Wolf, M. Zirngibl, R. Monnard, A.R. Chraplyvy, A.A. Abramov, R.P. Espindola, T.A. Strasser, J.R. Pedrazzani, A.M. Vengsarkar, J.L. Zyskind, J. Zhou, D.A. Ferrand, P.F. Wysocki, J.B. Judkins, S.W. Granlund, Y.P. Li, "1 Tb/s transmission of 100 WDM 10 Gb/s channels over 400 km of TrueWave™ fibre", in *OSA Technical digest Series, Optical Fibre Communications (OFC'98)*, paper PD-10, San Jose, USA, February 1998.

[12] A.R. Chraplyvy, "Limitations on lightwave communications imposed by optical fibre nonlinearities", *J. Lightwave Technol.*, **8**, pp. 1548-1557, 1990.

[13] N. Nagatsu and K. Oguchi, "Standardisation activities for optical transport network architectures and systems", in *Technical Digest of Optoelectronics and Communications Conf. (OECC'98)*, pp. 390-391, Chiba, Japan, July 1998.

[14] D.C. O'Brien, R.J. Mears, "Computer generated holograms optimised for illumination with partially coherent light using a silicon backplane spatial light modulator as the recording device", *Proc. SPIE*, **1505**, pp. 32-37, 1991.

[15] D.C. O'Brien, R.J. Mears, T.D. Wilkinson and W.A. Crossland, "Dynamic holographic interconnects that use ferroelectric liquid-crystal spatial light modulators", *Appl. Opt.*, **33**(14), pp. 2795-2803, 1994.

[16] S.T. Warr, M.C. Parker and R.J. Mears, "Optically transparent digitally tunable wavelength filter", *Electron. Lett.*, **31**(2), pp. 129-130, 1995.

[17] H. Dammann, "Blazed synthetic phase-only holograms", *Optik*, **31**, pp. 95-104, 1970.

[18] E.L. Goldstein and L. Eskildsen, "Scaling limitations in transparent optical networks due to low-level crosstalk", *IEEE Photon. Technol. Lett.*, **7**(1), pp. 93-94, 1995.

[19] J.A. Arnaud, *Beam and fiber optics*, Academic Press, New York, 1976.

[20] D. Marcuse, "Loss analysis of single-mode fiber splices", *Bell Systems Technical Journal*, **56**(5), pp. 703-718, 1977.

[21] P.D. Gianino and C.L. Woods, "General treatment of spatial light modulator dead-zone effects on optical correlation. I. Computer simulations and II. Mathematical analysis", *Appl. Opt.*, **32**(32), pp. 6527: 6535 and 6536:6541, 1993.

[22] J.A. Neff, R.A. Athale and S.H. Lee, "Two-dimensional spatial light modulators: a tutorial", *Proc. of IEEE*, **78**(5), pp. 826-855, 1990.

[23] J. Grinberg, A.D. Jacobson, W.P. Bleha, L. Miller, L. Fraas, D. Boswell and G. Myer, "A new real-time noncoherent to coherent light image converter: the hybrid field effect liquid crystal light valve", *Opt. Eng.*, **14**, pp. 217-225, 1975.

[24] F. Leenhouts and M. Schadt, "Optics of twisted nematic and supertwisted nematic liquid-crystal displays", *J. Appl. Phys.*, **60**(9), pp. 3275-3281, 1986.

[25] S. Schmitt-Rink, S.D.S. Chemla and D.A.B. Miller, "Linear and nonlinear optical properties of semiconductor quantum wells", *Advances in Phys.*, **38**, p.89-108, 1989.

[26] N. Peyghambarian and H.M. Gibbs, "Optical bistability for optical signal processing and computing", *Opt. Eng.*, **24**(1), p. 68-73, 1985.

[27] D.A.B. Miller, D.S. Chemla, T.C. Damen, A.C. Gossard, W. Wiegmann, T.H. Wood and C.A. Burrus, "Novel hybrid optically bistable: the quantum well self-electro-optic effect device", *Appl. Phys. Lett.*, **45**(1), pp. 13-15, 1984.

[28] A.L. Lentine and D.A.B. Miller, "Evolution of the SEED technology: bistable logic gates to optoelectronic smart pixels", *IEEE J. Quantum Elect.*, **29**(2), pp. 655-669, 1993.

[29] G.D. Boyd, A.M. Fox and D.A.B. Miller, "33 ps optical switching of symmetric self-electro-optic-effect devices", *Appl. Phys. Lett.*, **57**, pp. 1843-1845, 1990.

[30] L.M. Loh and J.L. LoCicero, "Subnanosecond sampling all-optical analogue-to-digital converter using symmetric self-electro-optic effect devices", *Opt. Eng.*, Vol.35, No.2, pp. 457-466, 1996.

[31] F.B. McCormick, F.A.P. Tooley, T.J. Cloonan, J.L. Brubaker, A.L. Lentine, R.L.

Morrison, S.J. Hinterlong, M.J. Herron, S.L. Waker and J.M. Sasian, "S-SEED-based photonic switching network demonstration", *OSA Proc. on Photonics Switching*, **8**, pp. 48-55, 1991.

[32] D.A.B. Miller, M.D. Freuer, T.Y. Chang, S.C. Shunk, J.E. Henry, D.J. Burrows and D.S. Chemla, "Field-effect transistor self-electro-optic effect device: integrated photodiode, quantum well modulator and transistor", *IEEE Photon. Technol. Lett.*, **1**(3), pp. 62-64, 1989.

[33] K.W. Goossen, J.A. Walker, L.A. Dasaro, S.P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D.D. Bacon, D. Dahringer, L.M.F. Chirovsky, A.L. Lentine and D.A.B. Miller, "GaAs MQW modulators integrated with silicon CMOS", *IEEE Photon. Technol. Lett.*, **7**(4), pp. 360-362, 1995.

[34] A.L. Lentine, K.W. Goossen, J.A. Walker, J.E. Cunningham, W.Y. Jan, T.K. Woodward, A.V. Krishnamoorthy, B.J. Tseng, S.P. Hui, R.E. Leibenguth, L.M.F. Chirovsky, R.A. Novotny, D.B. Buchholz and R.L. Morrison, "Optoelectronic VLSI switching chip with greater than 1 Tbit/s potential optical I/O bandwidth", *Electron. Lett.*, **33**(10), pp. 894-895, 1997.

[35] K. Wakita, I. Kotaka, O. Mitomi, H. Asai, Y. Kawamura and M. Naganuma, "High-speed InGaAlAs/InAlAs multiple quantum well optical modulators", *J. Lightwave Technol.*, **8**(7), pp. 1027-1032, 1990.

[36] A.R. Dias, R.F. Kalman, J.W. Goodman and A.A. Sawchuck, "Fibre optic crossbar switch with broadcast capability ", *Opt. Eng.*, **27**(1), pp. 955-960, 1988.

[37] H.S. Hinton, T.J. Cloonan, F.B. Mccormick, A.L. Lentine and F.A.P. Tooley, "Free-space digital optical-systems", *Proc. of IEEE*, **82**(11), pp. 1632-1649, 1994.

[38] A.G. Kirk, H. Imam, K. Bird and T.J. Hall, "The design and fabrication of computer generated holographic fan-out elements for a matrix-matrix interconnection scheme", *Proc. SPIE*, **1574**, pp. 121-132, 1991.

[39] H.J. White, G.M. Proudley, C. Stace, N.A. Brownjohn, R. Dawkins, A.C. Walker, M.R. Taghizadeh, C.P. Barrett, D.T. Neilson, W.A. Crossland, J.R. Brocklehurst and M.N. Birch, "The OCPM demonstrator system", *Opt. Soc. Amer. 1995 tech. Dig. Ser., Photonics in Switching*, **12**, pp. 152-154, 1995.

[40] D. Marcuse, *Light transmission optics*, Van Nostrand Rheinhold, New York, 1972.

[41] S.T. Warr, *Free-space switching for optical fibre networks-Ph.D. Thesis*, Cambridge University, pp. 83-86, July 1996.

[42] B. Robertson, E.J. Restall, M.R. Taghizadeh and A.C. Walker, "Space-variant holographic optical elements in dichromated gelatin", *Appl. Opt.*, **30**(17), pp. 2368-2375, 1991.

[43] S.T. Warr, "Generic design for *N* to *N* switches specified", *Roses PM 2.4 report*, Oct. 1997.

[44] A. Neyer, "Electrooptic X-switch using single-mode Ti-LiNbO$_3$ channel waveguides", *Electron. Lett.*, **19**(14), pp. 553-554,1983.

[45] N. Keil, H.H. Yao, C. Zawadzki and B. Strebel, "4 × 4 polymer thermooptic directional coupler switch at 1.55 μm", *Electron. Lett*, **30**(8), pp. 639-640, 1994.

[46] T. Nishi, T. Yamamoto, S. Kuroyanagi, "A polarization-controlled free-space photonic switch based on a pi-loss switch", *IEEE Photon. Technol. Lett.*, **5**(9), pp. 1104-1106, 1993.

[47] T. Goh, A. Himeno, M. Okuno, H. Takahashi, K. Hattori, "High-extinction ratio and low-loss silica-based 8 × 8 thermooptic matrix switch", *IEEE Photon. Technol. Lett.*, **10**(3), pp. 358-360, 1998.

[48] T. Goh, N. Yasu, K. Hattori, A. Himeno, M. Okuno, Y. Ohmori, "Low-loss and high-extinction-ratio silica-based strictly nonblocking 16 × 16 thermooptic matrix switch", *IEEE Photon. Technol. Lett.*, **10**(6), pp. 810-812, 1998.

[49] K. Preston Jr., "The membrane light modulator and its applications", *Opt. Acta*, **16**(5), p. 579, 1969.

[50] D.R. Pape and L.J. Hornbeck, "Characteristics of the deformable mirror device for optical information processing", *Opt. Eng.*, **22**(6), p. 675-681, 1983.

[51] W. E. Ross, D. Psaltis and R. H. Anderson, "Two-dimensional magneto-optic spatial light modulator for signal processing", *Opt. Eng.*, **22**(4), pp. 485-491, 1983.

[52] J.A. Davis, S.W. Connely, G.W. Bach, R.A. Lilly and D.M. Cottrell, "Programmable optical interconnections with large fan-out capability using the magneto-optic spatial

light-modulator", *Opt. Lett.*, **14**(1), pp. 102-104, 1989.

[53] S.H. Lee, S.C. Esener, M.A. Title and T.J. Drabik, "Two-dimensional silicon/PLZT spatial light modulators: Design Consideration and Technology", *Opt. Eng.*, **25**(2), pp. 250-260, 1986.

[54] N.A. Clark and S.T. Lagerwall, "Submicrosecond bistable electrto-optic switching in liquid crystals", *Appl. Phys. Lett.*, **36**(11), pp. 899-901, 1980.

[55] P.W.H. Surguy, P.J. Ayliffe, M.J. Birch, M.F. Bone, I. Coulson, W.A. Crossland, J.R. Hughes, P.W. Ross, F.C. Saunders and M.J. Towler, "The JOERS/ALVEY ferroelectric multiplexing scheme", *Ferroelectrics*, **122**, pp. 63-79, 1991.

[56] S.E Broomfield, M.A.A. Neil, E.G.S. Paige and G.G. Yang, "Programmable binary phase-only optical device based on ferroelectric liquid crystal SLM", *Electron. Lett.*, **28**, pp. 26-28, 1992.

[57] K.M. Johnson, D.J. McKnight and I. Underwood, "Smart spatial light modulators using liquid crystals on silicon", *IEEE J. Quantum Elect.*, **29**(2), pp. 699-714, 1993.

[58] S.E. Broomfield, M.A.A. Neil and E.G.S. Page, "Programmable multiple-level phase modulation that uses ferroelectric liquid-crystal spatial light modulators", *Appl. Opt.*, **34**(29), pp. 6652-6665, 1995.

[59] T.D. Wilkinson, D.C. O'Brien and R.J. Mears, "Dynamic asymmetric binary holograms using a ferroelectric liquid crystal spatial light modulator", *Opt. Comm.*, **109**, 222-226, 1994.

[60] S.E. Broomfield, M.A.A. Neil and E.G.S. Paige, "Four-level, phase only, spatial light modulator", *Electron. Lett.*, 29, pp. 1661-1663, 1995

[61] J.M. Florence and L.A. Yoder, "Display system architectures for digital micro-mirror device (DMD) based projectors", *Proc. SPIE*, **2650**, pp. 193-208, 1996.

[62] L.Y. Lin, E.L. Goldstein and R.W. Tkach, "Free-space micromachined optical switches with submillisecond switching time for large-scale optical crossconnects", *IEEE Photon. Technol. Lett.*, **10**(4), pp. 525-527, 1998.

[63] M.C. Wu, "Micromachining for optical and optoelectronic systems", *Proc. of IEEE*, **85**, pp. 1833-1856, 1997.

[64] D. Gloge, "Weakly Guiding Fibres", *Appl. Opt.*, **10**(10), pp. 2252-2258, 1971.

[65] G. Cancellieri, *Single-mode Optical Fibres*, Pergamon Press, pp. 115-118, 1991.

[66] A. W. Snyder and J. D. Love, *Optical Waveguide Theory*, Chapman and Hall, pp. 314-315, 1983.

[67] A.W. Snyder and R.A. Sammut, "Fundamental modes of graded optical fibres", *J. Opt. Soc. Am.*, **12**, pp. 1663-1671, 1979.

[68] A. Yariv, *Optical Electronics*, 4th. Ed., Saunders College Publishing, pp. 705-707.

[69] W. B. Joyce and B. C. Deloach, "Alignment of Gaussian beams", *Appl. Opt.*, **23**(23), pp. 4187-4196, 1984.

[70] E.G. Churin, P. Bayvel, J.E. Midwinter and A.M. Hill, "The influence of aperture size and shape on crosstalk level in free-space grating multiplexers for WDM networks", *IEEE Photon. Technol. Lett.*, **8**(10), pp. 1337-1339, 1996.

[71] I.S. Gradshteyn and I.M. Ryzhik, *Tables of Integrals, Series and Products*, 5th Ed., Academic Press, London, p. 938, 1994.

[72] L. Jeunhomme, *Single-mode fibre optics*, Dekker, New York, pp. 207-209, p. 256, 1983.

[73] E.O. Brigham, *The Fast Fourier Transform*, Prentice-Hall, Englewood Cliffs, NJ, 1974.

[74] N. Collings, W.A. Crossland, P.J. Ayliffe, D.G., Vass and I. Underwood, "Evolutionary development of advanced liquid crystal spatial light modulator", *Appl. Opt.*, **28**, pp. 4740:4747, 1989.

[75] J.A. Cox, "Diffractive efficiency of binary optical elements", *Computer and Optically Formed Holographic Optics*, *Proc. SPIE*, **1211**, pp. 116-124, 1990.

[76] R.W. Gerchberg and W.O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures", *Optik*, **35**(2), pp. 237-246, 1972.

[77] M.A. Seldowitz, J.P. Allebach and D.W. Sweeney, "Synthesis of digital holograms by direct binary search", *Appl. Opt.*, **26**(14), pp. 2788-2798, 1987.

[78] M.R. Feldman and C.C. Guest, *Opt. Lett.*, "Iterative encoding of high-efficiency

holograms for generation of spot arrays", **14**(10), pp. 479-481, 1989.

[79] S. Weissbach, F. Wyrowski and O. Bryngdahl, "Digital phase holograms: coding and quantisation with an error diffusion concept", *Opt. Comm.*, **72**(1), 37-41, 1989.

[80] E.G. Steward, *Fourier optics: an introduction*, 2nd. Ed., Ellis Horwood Ltd., 1987.

[81] H. Dammann, "Spectral characteristics of stepped-phase gratings", *Optik*, **53**, pp. 409-417, 1979.

[82] J. Amako and T. Sonehara, "Hologram using an electrically controlled birefringent liquid-crystal spatial light modulator", *Appl. Opt.*, **30**(32), pp. 4622-4628, 1991.

[83] I. Underwood, "Liquid crystal over silicon spatial light modulators - principles, practice and prospects", *OSA Trends in Optics and Photonics*, Geoffrey Burdge and Sadik C. Esener, eds. (Optical Society of America, Washington DC), *Spatial Light Modulators*, **14**, pp. 76-88, 1997.

[84] S.T. Warr, "Interim report on the first design iteration of a massively parallel matrix switch", *Roses PD 2.1 report*, Jun. 1997.

[85] T. Parker, M.J. Holmes and V. Handerek, "Hologram definition", *Roses PM 3.3 report*, Feb. 1998.

[86] T.K. Gaylord and M.G. Moharam, "Analysis and application of diffraction by gratings", *Proc. IEEE*, **73**(5), pp. 894-937, 1985.

[87] J. Gourlay, S. Samus, P.W. McOwan, D.G. Vass, I. Underwood and M.R. Worboys, "Real-time binary phase holograms on a reflective ferroelectric liquid crystal spatial light modulator", *Opt. Lett.*, **33**, pp. 8251-8254, 1994.

[88] K.L. Tan, W.A. Crossland and R.J. Mears, "A comparison of the efficiency and crosstalk of quaternary and binary phase-only holograms based on ferroelectric liquid crystals (FLC)", *Ferroelectrics*, **213**, pp. 233-240, 1998.

[89] M.C. Parker, *Dynamic holograms for wavelength division multiplexing-Ph.D. Thesis*, Cambridge University, pp. 16-20, November 1996.

[90] S.T. Warr and R.J. Mears, "Polarisation insensitive operation of ferroelectric liquid crossbar devices", *Electron. Lett.*, **31**, pp. 714-716, 1995.

[91] P. Berthele, B. Fracasso and J.L. de Bougrenet de la Tocnaye, "Design and

characterisation of a liquid-crystal spatial light modulator for a polarisation-insensitive optical space switch", *Appl. Opt.*, **37**(23), pp. 5461-5468, 1998.

[92] J.E. Stockley, G.D. Sharp, S.A. Serati and K.M. Johnson, "Analogue optical phase modulator based on chiral smectic and polymer cholesteric liquid crystals", *Opt. Lett.*, **20**(23), pp. 2441-2443, 1995.

[93] M.O. Freeman, T.A. Brown and D.M. Walba, "Quantised complex ferroelectric liquid crystal spatial light modulators", *Appl. Opt.*, **31**(20), pp. 3917-3929, 1992.

[94] G.D. Sharp and K.M. Johnson, "High-speed analogue complex-amplitude liquid crystal-light modulator", *Opt. Lett.*, **19**(16), pp. 1228-1230, 1994.

[95] J.W. Goodman, *Introduction to Fourier optics*, 2nd. Ed., Appendix C, pp. 415-420, McGraw-Hill, Singapore, 1996.

[96] D.J. Broer, "Molecular architectures in thin plastic films by in-situ photopolymerisation of reactive liquid crystals", in *SID 95 Digest*, pp. 165-168, 1995.

[97] P.H. Saul, D.W. Howard and C.J. Greenwood, "VLSI process compatible 8-bit bi-CMOS DAC", *IEE Proc.*, **132**, Parts E & I, no. 2, pp. 99-101,1985.

[98] A. Thomsen, R.G. Lindquist, J.H. Kulick, P.J. Nasiatka, G.P. Nordin and S.T. Kowel, "A pixel scale digital to analogue converter array for liquid crystal on VLSI displays", *IEEE Trans. Circuits and Systems I*, **42**(9), pp. 545-548, 1995.

[99] J.A. Breslin, J.K. Low and I. Underwood, "Smart pixel with four-level amplitude or phase modulation", in *Digest 1998 IEEE/LEOS Summer Topical Meetings on smart pixels*, ThB4, pp. 33-34, July 1998.

[100] A. O'Hara, J.R. Hannah, I. Underwood, D.G. Vass and R.J. Holwill, "Mirror quality and efficiency improvements of reflective spatial light modulators by the use of dielectric coatings and chemical-mechanical polishing", *Appl. Opt.*, **32**(28), pp. 5549-5555, 1993.

[101] Austria Mikro System Internationale, http//www.ams.co.at.

[102] Alcatel Mietec, http://www.dspvalley.com/members/mietec/1n_mietec.htm.

[103] Newport Wafer Fabrication Limited, http//www.nwl-eur.co.uk.

[104] Taiwan Semiconductor Manufacturing Company Limited, http//www.tsmc.com.tw.

[105] N.H.E. Weste and K. Eshraghian, *Principles of CMOS VLSI design*, 2nd. Ed., Addison-Wesley, pp. 57-58, 1993.

[106] C.C. Mao, D.J. McKnight and K.M. Johnson, "High-speed liquid-crystal-on-silicon spatial light modulators using high-voltage circuitry", *Opt. Lett.*, **20**(3), pp. 342-344, 1995.

[107] K.L. Tan, "Finalised SLM layout including specification of AMS process", *Roses PM 5.1 report*, Oct. 1997.

[108] K.L. Tan, "Specifications of SLM pin-outs and interface requirements", *Roses PM 5.3 report*, Jan. 1998.

[109] J.W. Goodby, *et al.*, *Ferroelectric liquid crystals: principles, properties and applications*, Gordon and Breach Science Publishers, pp. 230-239, 1991.

[110] M.M. Redmond, "Alignment and preliminary LC assessment of binary SLM W2/D22", *Roses-CUED 8th quarterly report*, pp. 6-7, 1998.

[111] CDDR materials developed in the Link Molecular Electronics Programme, Project No. IED 2/435/30/001.

[112] "Optical Coatings", *Melles Griot 1997-98 catalogue*, p. 2, 1997.

[113] D.J. McKnight and K.M. Johnson, "Analogue distorted helix ferroelectric-liquid crystal-on-silicon spatial light modulator", *Opt. Lett.*, **20**(5), pp. 513-515, 1995.

[114] Private Communication with C. Buckland, process engineer at Austria Systeme Internationale.

[115] D.C. O'Brien, D.J. McKnight and A. Fedor, "A holographically routed optical crossbar using a ferroelectric liquid-crystal over silicon spatial light modulator", *Ferroelectrics*, **181**, pp. 79-86, 1996.

APPENDIX A

Derivation of the replay field approximation

The replay field, $R(x_r,y_r)$ of a square-dimension hologram is given by:

$$R(x_r, y_r) = \mathrm{F}\left\{ G(w_h; x_h, y_h) \times rect\left(\frac{x_h}{\gamma w_h}, \frac{y_h}{\gamma w_h}\right) \times H(x_h, y_h) \right\}, \tag{A.1}$$

where $G(w_h;x_h,y_h)$ is the Gaussian beam at the hologram plane; aperture function, $rect\left(\frac{x_h}{\gamma w_h}, \frac{y_h}{\gamma w_h}\right)$ has a value of 1 for $|x_h|$ and $|y_h| \le \gamma w_h/2$, 0 elsewhere; γ is the ratio of the hologram aperture to the width of the illuminating Gaussian beam, w_h; F denotes the optical Fourier transform, as defined by Yariv[*]. The normalised unit intensity Gaussian field at the hologram plane is as below:

$$G(w_h; x_h, y_h) = \frac{1}{w_h}\sqrt{\frac{2}{\pi}} \exp\left(-\frac{x_h^2 + y_h^2}{w_h^2} \right). \tag{A.2}$$

Temporally assuming that the Fourier transform of a hologram, $H(x_h,y_h)$ is a 100% Dirac's delta function and there are no pixellation and dead-space effects[†],

---

[*] A. Yariv, *Optical Electronics*, 4th Ed., Saunders College Publishing, pp. 705-707.

[†] The optical replay field is a convolution of the discrete hologram replay and the transform of the hologram illumination.

$$\mathrm{F}\{H(x_h,y_h)\} = \delta(x_r,y_r), \tag{A.3}$$

where (x_h,y_h) and (x_r,y_r) are the hologram and replay plane co-ordinates, respectively. In this case, the replay field is given by the optical Fourier transform of the product of the Gaussian field and the square hologram aperture:

$$R(x_r, y_r) = \mathrm{F}\left\{G(w_h; x_h, y_h) \times rect\left(\frac{x_h}{\gamma w_h}, \frac{y_h}{\gamma w_h}\right)\right\}. \tag{A.4}$$

Assuming that the Gaussian field truncation is symmetrical and the field outside the aperture window is blocked, the replay field is thus:

$$R(x_r, y_r) = \frac{1}{f\lambda} \int_{-\frac{\gamma}{2}w_h}^{\frac{\gamma}{2}w_h} \int_{-\frac{\gamma}{2}w_h}^{\frac{\gamma}{2}w_h} G(w_h; x_h, y_h) \exp\left(\frac{ik}{f}(x_r x_h + y_r y_h)\right) dx_h dy_h. \tag{A.5}$$

Substituting Gaussian field distribution (A.2) into (A.5),

$$R(x_r, y_r) = \frac{1}{f\lambda}\frac{1}{w_h}\sqrt{\frac{2}{\pi}} \int_{-\frac{\gamma}{2}w_h}^{\frac{\gamma}{2}w_h} \exp\left(-\left(\frac{x_h^2}{w_h^2} - \frac{ik(x_r x_h)}{f}\right)\right) dx_h \int_{-\frac{\gamma}{2}w_h}^{\frac{\gamma}{2}w_h} \exp\left(-\left(\frac{y_h^2}{w_h^2} - \frac{ik(y_r y_h)}{f}\right)\right) dy_h. \tag{A.6}$$

Simplifying (A.6), assigning $w_r = \dfrac{f\lambda}{\pi w_h}$ and changing the variable of integration to t gives,

$$R(x_r, y_r) = \frac{1}{4}\frac{1}{w_r}\sqrt{\frac{2}{\pi}} \left[\exp\left(-\frac{x_r^2}{w_r^2}\right)\frac{2}{\sqrt{\pi}} \int_{-\frac{\gamma}{2}-\frac{ix_r}{w_r}}^{\frac{\gamma}{2}-\frac{ix_r}{w_r}} \exp\left(-t^2\right)dt\right]\left[\exp\left(-\frac{y_r^2}{w_r^2}\right)\frac{2}{\sqrt{\pi}} \int_{-\frac{\gamma}{2}-\frac{iy_r}{w_r}}^{\frac{\gamma}{2}-\frac{iy_r}{w_r}} \exp\left(-t^2\right)dt\right]. \tag{A.7}$$

These are the two orthogonal integrals of a Gaussian field with complex limits. Separating the integrals:

$$R(x_r, y_r) = \frac{1}{w_r}\sqrt{\frac{\pi}{2}} \times \frac{1}{2}\exp\left(-\frac{x_r^2}{w_r^2}\right)F(x_r) \times \frac{1}{2}\exp\left(-\frac{y_r^2}{w_r^2}\right)F(y_r), \text{ and hence} \tag{A.8}$$

$$R(x_r, y_r) = \frac{1}{w_r}\sqrt{\frac{\pi}{2}} \times R(x_r) \times R(y_r), \text{ where} \tag{A.9}$$

$$R(x_r) = \frac{1}{2}\exp\left(-\frac{x_r^{\,2}}{w_r^{\,2}}\right)F(x_r)\,. \tag{A.10}$$

The $F(x_r)$ function, with complex limits, is defined as:

$$F(x_r) = \frac{2}{\sqrt{\pi}}\int_{-\frac{\gamma}{2}-\frac{ix_r}{w_r}}^{\frac{\gamma}{2}-\frac{ix_r}{w_r}} \exp\left(-t^2\right)dt\,,$$

$$F(x_r) = \frac{2}{\sqrt{\pi}}\int_{0}^{\frac{\gamma}{2}-\frac{ix_r}{w_r}} \exp\left(-t^2\right)dt + \frac{2}{\sqrt{\pi}}\int_{0}^{\frac{\gamma}{2}+\frac{ix_r}{w_r}} \exp\left(-t^2\right)dt\,, \text{ and}$$

$$F(x_r) = \Phi\left(\frac{\gamma}{2}-\frac{ix_r}{w_r}\right) + \Phi\left(\frac{\gamma}{2}+\frac{ix_r}{w_r}\right). \tag{A.11}$$

The integrals are error functions (Φ or *erf*) with complex limits. An asymptotic expansion[‡] of $\Phi(z)$ for large arguments is:

$$\Phi(z) \sim 1 + \frac{\exp(-z^2)}{\pi}\sum_{m=1}^{n-1}\frac{(-1)^m\Gamma(m-1/2)}{z^{2m-1}} + \frac{\exp(-z^2)}{\pi}R_n\,, \tag{A.12}$$

where $\Gamma(\alpha)$ is the Factorial (Gamma) function and R_n is the remainder.

As the asymptotic expansion is of alternating signs, the absolute value of its remainder satisfies the following inequalities:

$$|R_n| < \frac{\Gamma(n-1/2)}{|z|^{2n-1}\cos(\phi_z)} \text{ and } \phi_z^{\,2} < \left(\frac{\pi}{2}\right)^2 . \tag{A.13}$$

Letting $z_1 = \frac{\gamma}{2} - \frac{ix_r}{wr}$ and $z_2 = \frac{\gamma}{2} + \frac{ix_r}{wr}$, (A.14)

[‡] I.S. Gradshteyn and I.M. Ryzhik, *Tables of Integrals, Series and Products*, 5th Ed., Academic Press, London, pp. 939-940, 1994.

where both z_1 and z_2 satisfy the criteria for the complex argument, taking only the first term of the asymptotic expansion and using $\Gamma(1/2) = \sqrt{\pi}$ gives,

$$\Phi(z_1) + \Phi(z_2) \sim 1 + \frac{-\exp\left(-\left(\frac{\gamma}{2} - \frac{ix_r}{w_r}\right)^2\right)}{\sqrt{\pi}\left(\frac{\gamma}{2} - \frac{ix_r}{w_r}\right)} + 1 + \frac{-\exp\left(-\left(\frac{\gamma}{2} + \frac{ix_r}{w_r}\right)^2\right)}{\sqrt{\pi}\left(\frac{\gamma}{2} + \frac{ix_r}{w_r}\right)}. \tag{A.15}$$

Simplifying and substituting eq. (A.15) into (A.10) yields:

$$R(x_r) \sim \exp\left(-\frac{x_r^2}{w_r^2}\right) - \frac{\exp\left(-\frac{\gamma^2}{4}\right)}{\sqrt{\pi}\left(\frac{\gamma^2}{4} + \frac{x_r^2}{w_r^2}\right)}\left[\frac{\gamma}{2}\cos\left(\frac{\gamma x_r}{w_r}\right) - \frac{x_r}{w_r}\sin\left(\frac{\gamma x_r}{w_r}\right)\right]. \tag{A.16}$$

This replay field profile is the convolution result of the transform of a square aperture and the transform of a Gaussian field distribution. Eq. (A.16) is quoted as eq. (3.10) in §3.3. It is probably not valid for $\gamma << 2$, as the argument of the error function will tend towards $\pi/2$.

APPENDIX B

DC undiffracted light for multi-level polarisation rotation

The modulated polarisation vector, resolved ∥ and ⊥ to the state *m-1* director (i.e., θ_o from the *y*-axis), E_{out} is as below (see §6.2.1):

$$\underline{E}_{out} = \left[\mathbf{R}_{\theta_o-\beta}\right]\left[\mathbf{J}_\pi\right]\left[\mathbf{R}_\beta\right]\left[\underline{E}_{in}\right]. \tag{B.1}$$

Here, β refers to the switched angle, i.e., $g\theta_o/(m-1)$ for *g* state and for this state, E_{out} (assigned E_g) can be defined as:

$$\underline{E}_g = \begin{bmatrix} \cos\left(\dfrac{\theta_o(m-1-g)}{m-1}\right) & \sin\left(\dfrac{\theta_o(m-1-g)}{m-1}\right) \\ -\sin\left(\dfrac{\theta_o(m-1-g)}{m-1}\right) & \cos\left(\dfrac{\theta_o(m-1-g)}{m-1}\right) \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\left(\dfrac{g\theta_o}{m-1}\right) & \sin\left(\dfrac{g\theta_o}{m-1}\right) \\ -\sin\left(\dfrac{g\theta_o}{m-1}\right) & \cos\left(\dfrac{g\theta_o}{m-1}\right) \end{bmatrix} \begin{bmatrix} E_x \\ E_y e^{j\alpha} \end{bmatrix},$$

$$= \begin{bmatrix} -\cos\left(\dfrac{\theta_o(m-1-2g)}{m-1}\right) & \sin\left(\dfrac{\theta_o(m-1-2g)}{m-1}\right) \\ \sin\left(\dfrac{\theta_o(m-1-2g)}{m-1}\right) & \cos\left(\dfrac{\theta_o(m-1-2g)}{m-1}\right) \end{bmatrix} \begin{bmatrix} E_x \\ E_y e^{j\alpha} \end{bmatrix}. \tag{B.2}$$

From symmetry consideration, there will be a state, *m-1-g*, whereby the two rotational matrices are in the reverse order of eq. (B.2):

$$\underline{E}_{m-1-g}=\begin{bmatrix}\cos\left(\frac{g\theta_o}{m-1}\right) & \sin\left(\frac{g\theta_o}{m-1}\right)\\ -\sin\left(\frac{g\theta_o}{m-1}\right) & \cos\left(\frac{g\theta_o}{m-1}\right)\end{bmatrix}\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\cos\left(\frac{\theta_o(m-1-g)}{m-1}\right) & \sin\left(\frac{\theta_o(m-1-g)}{m-1}\right)\\ -\sin\left(\frac{\theta_o(m-1-g)}{m-1}\right) & \cos\left(\frac{\theta_o(m-1-g)}{m-1}\right)\end{bmatrix}\begin{bmatrix}Ex\\ Eye^{j\alpha}\end{bmatrix},$$

$$=\begin{bmatrix}-\cos\left(\frac{\theta_o(m-1-2g)}{m-1}\right) & -\sin\left(\frac{\theta_o(m-1-2g)}{m-1}\right)\\ -\sin\left(\frac{\theta_o(m-1-2g)}{m-1}\right) & \cos\left(\frac{\theta_o(m-1-2g)}{m-1}\right)\end{bmatrix}\begin{bmatrix}E_x\\ E_ye^{j\alpha}\end{bmatrix}. \tag{B.3}$$

The DC undiffracted intensity is the square of the sum of all levels (assuming the hologram consists of an equal number of each state), hence,

$$DC^2=\left|\frac{1}{m}\left(\underline{E}_0+\underline{E}_1+\ldots\underline{E}_g\ldots+\underline{E}_{m-2}+\underline{E}_{m-1}\right)\right|^2. \tag{B.4}$$

Due to the symmetric nature of the m equally spaced phase states, $E_g + E_{m-1-g}$ (i.e., eq. (B.2) + eq. (B.3)) gives a simple diagonal matrix:

$$\underline{E}_g+E_{m-1-g}=\begin{bmatrix}-2\cos\left(\frac{\theta_o(m-1-2g)}{m-1}\right) & 0\\ 0 & 2\cos\left(\frac{\theta_o(m-1-2g)}{m-1}\right)\end{bmatrix}\begin{bmatrix}E_x\\ E_ye^{j\alpha}\end{bmatrix}. \tag{B.5}$$

For an even m, there will be $m/2$ pairs of such output polarisation vectors. For an odd m, there is an additional state corresponding to $g = (m-1)/2$, at the bisector of θ_o. The polarisation vector of this state, exiting the CSLC cell will be:

$$\underline{E}_{(m-1)/2}=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}E_x\\ E_ye^{j\alpha}\end{bmatrix}. \tag{B.6}$$

Therefore, for an even number of phase states, the DC power is:

$$DC^2_{m=even}=\frac{1}{m^2}\left|\sum_{g=0}^{\frac{m-2}{2}}\left(\underline{E}_g+\underline{E}_{m-1-g}\right)\right|^2. \tag{B.7}$$

For the case of an odd number of phase states, the DC. power is:

$$DC^2_{m-odd} = \frac{1}{m^2}\left|\left[\sum_{g=0}^{\frac{m-3}{2}}\left(E_g + E_{m-1-g}\right)\right] + E_{(m-1)/2}\right|^2 . \tag{B.8}$$

Eqs. (B.5) and (B.6) are substituted into eqs. (B.7) and (B.8) to yield the DC undiffracted intensity normalised by the input intensity $|E_{in}|^2$ as below:

$$DC^2_{m=even} = \frac{4}{m^2}\left(\sum_{g=0}^{\frac{m-2}{2}}\cos\left(\frac{\theta_o(m-1-2g)}{m-1}\right)\right)^2 \text{, and}$$

$$DC^2_{m=odd} = \frac{1}{m^2}\left(\left(2\sum_{g=0}^{\frac{m-3}{2}}\cos\left(\frac{\theta_o(m-1-2g)}{m-1}\right)\right)+1\right)^2 . \tag{B.9}$$

Warr[§] has shown analytically that the DC undiffracted power loss of a binary ($m = 2$) hologram is,

$$DC^2_{m=2} = \cos^2(\theta_o) . \tag{B.10}$$

This is readily obtained from eq. (B.9). Also, the two lowest multi-level holograms, $m = 3$ and 4, will have DC undiffracted power as the following:

$$DC^2_{m=3} = \frac{1}{9}\left(2\cos(\theta_o)+1\right)^2 \text{ , and}$$

$$DC^2_{m=4} = \cos^2\left(\frac{\theta_o}{3}\right)\cos^2\left(\frac{2\theta_o}{3}\right). \tag{B.11}$$

Now that the input polarisation states have been shown to have no effects on the DC non-

---

[§] S.T. Warr and R.J. Mears, "Polarisation insensitive operation of ferroelectric liquid crossbar devices", *Electron. Lett.*, **31**, pp. 714-716, 1995.

diffracted light, it may be worthwhile to derive a general expression for the DC light fraction, instead of using the cumbersome eq. (B.9). If we represent all the possible phase states of a m-level hologram by phasors of $\left[e^{j0} \; e^{j\frac{2\theta_o}{(m-1)}} \; \ldots e^{j2\theta_o} \right]$, the DC component is then the geometric average of the phase distributions. Thus,

$DC_m^2 = \dfrac{1}{m^2}\left|\sum_{g=0}^{m-1} e^{\frac{j2g\theta_o}{(m-1)}}\right|^2$ and taking the formula for geometric series yields,

$$DC_m^2 = \frac{1}{m^2}\left|\frac{(1-e^{\frac{j2m\theta_o}{(m-1)}})}{(1-e^{\frac{j2\theta_o}{(m-1)}})}\right|^2 \text{ and } DC_m^2 = \frac{sinc^2(\dfrac{m\theta_o}{m-1})}{sinc^2(\dfrac{\theta_o}{m-1})}. \tag{B.12}$$

For cases where m is a multiple of 2, eq. (B.12) simplifies readily to a product of cosines, e.g.,

$$m = 2,\ DC_{m=2}^2 = \cos^2(\theta_o),$$

$$m = 4,\ DC_{m=4}^2 = \cos^2(\frac{1}{3}\theta_o)\cos^2(\frac{2}{3}\theta_o),$$

$$m = 8,\ DC_{m=8}^2 = \cos^2(\frac{1}{7}\theta_o)\cos^2(\frac{2}{7}\theta_o)\cos^2(\frac{4}{7}\theta_o), \text{ and}$$

$$m = 2^\kappa,\ DC_{m=2^\kappa}^2 = \prod_{i=0}^{\kappa-1}\cos^2(\frac{2^i}{m-1}\theta_o). \tag{B.13}$$

The fraction of non-diffracted light is independent of the input polarisation state when the CSLC cell is configured as a half-wave plate. For a hologram which has 2^κ phase levels (κ an integer), this fraction is $\prod_{i=0}^{\kappa-1}\cos^2(\frac{2^i}{m-1}\theta_o)$, where each phase level is present with an equal measure.

APPENDIX C

Associated publications

Journals:

K.L. Tan, W.A. Crossland and R.J. Mears, "A comparison of the efficiency and crosstalk of quaternary and binary phase-only holograms based on ferroelectric liquid crystals (FLC)", *Ferroelectrics*, **213**, pp. 233-240, 1998.

Conferences:

K.L. Tan, W.A. Crossland and R.J. Mears, "A comparison of the efficiency and crosstalk performance of quaternary and binary phase-only holograms based on ferroelectric liquid crystals (FLC)", in *6th International Conference on Ferroelectric Liquid Crystals (FLC'97)*, E.N.S.T. de Bretagne, Brest, France, July 1997.

K.L. Tan, W.A. Crossland and R.J. Mears, "Limitations of the crosstalk performance of a linear 1:*N* holographic switch due to spatial light modulator (SLM) dead-zones", in *3rd IEEE Malaysia International Conference on Communications (MICC'97)*, Kuala Lumpur, Malaysia, November 1997.

K.L. Tan, T.D. Wilkinson, W.A. Crossland and R.J. Mears, "The design and fabrication of a silicon backplane spatial light modulator for silicon-transparent 1.5 μm IR quaternary operation", in *3rd Optoelectronics and Communications Conference (OECC'98)*, Chiba, Japan, July 1998.

APPENDIX D

Glossary

D.1 Abbreviations and acronyms

1-D	One-Dimension/al	fet (FET)	Field Effect Transistor
2-D	Two-Dimension/al	*fft* (FFT)	Fast Fourier Transform
AWG	Arrayed-Waveguide Grating	FLC	Ferroelectric Liquid Crystal
CCD	Charge-Coupled Device	FP	Fabry-Perot
CGH	Computer-Generated Hologram	ITO	Indium Tin Oxide
CMOS	Complementary Metal-Oxide-Si	*lcm*	Least Common Multiple
CSLC	Chiral Smectic Liquid Crystal	LPG	Long-Period fibre Grating
dB (dBm)	Decibels (dB relative to 1 mW)	LSB	Least Significant Bit
DMD	Digital-Micro-mirror Device	MIN	Multi-stage Interconnect Network
DMOS	Dual-drain MOS (*q.v.*)	MQW	Multiple Quantum Well
DOE	Diffractive Optical Element	MSB	Most Significant Bit
EA	Electro-Absorption (modulator)	OASLM	Optically Addressed SLM (*q.v.*)
EASLM	Electrically Addressed SLM (*q.v.*)	OB	Optical Bistable
EDFA	Erbium-Doped Fibre Amplifier	PLC	Planar Lightwave Circuit
FBG	Fibre Bragg Grating	QCSE	Quantum-Confined Stark Effect

SA	Simulated Annealing
SBWP	Spatial BandWidth Product
SEED	Self-Electro-optic-Effect Device
Si	Silicon
SLM	Spatial Light Modulator
TEM	Transverse ElectroMagnetic
VCSEL	Vertical-Cavity Surface Emitting Laser
VLSI	Very Large Scale Integrated/ion
WDM	Wavelength Division Multiplexing

D.2 Holographic terminology

aperture replay

The *numerical* or *optical replay* (*q.v.*) of a blank aperture at the *hologram plane* considering only the effect of Gaussian input illumination.

base hologram, fractional hologram or hologram unit

The basic building unit of a *routing hologram* (*q.v.*) that requires only $x_o \times y_o$ elements to fully encode the hologram function and describes its complete replay field for a given *replay fraction* (*q.v.*).

composite sinc squared

The overall combined intensity roll-off across all *replay replications* (*q.v.*) due to the orthogonal intensity scaling of *spatial* and *phase quantisation* (*q.v.*).

diffraction order

Replay peak (*q.v.*) due to the steering function of an encoded *routing hologram* (*q.v.*) in the order of diminishing intensity values (except 0th order) within a *replay replication* (*q.v.*). A +1 order is always denoted to the peak with the highest intensity, regardless of the sign of its location from the optical axis or its spatial frequency being higher than some higher order harmonics. Such a notation enables the use of a simple expression for the intensity of each peak even though it makes little sense physically.

Gaussian clipping

The clipping of an infinitely extending (theoretically) collimated Gaussian beam by the finite expanse of the SLM/*sub-hologram* (*q.v.*) aperture

grating (function)

A special case of 1-D *routing holograms*, x/x_o, where $x = 1$ and x_o = multiples of the number of *phase states* (*q.v.*), *m*.

hologram repeats

The number of repeated *base holograms* (*q.v.*) within a *routing hologram* (*q.v.*) that is required to provide a value of replay resolution and a crosstalk isolation figure due to *replay sidelobes* (*q.v.*).

input, hologram and replay planes

The three idealised flat regions corresponding to the focal planes of the collimating and transform lenses in a coherent 4*f* optical configuration where the input light sources, the reconfigurable holograms and the output detectors (could be fibres) are located, respectively.

modulo-1

A set of skip and rotate rules in order to locate all the *replay peaks* (*q.v.*) and their associated intensity values.

modulo-Λ

A set of skip and rotate rules in order to generate a maximum efficiency, single peak *base hologram* (*q.v.*) where Λ is the least common multiple of the number of *phase states* (*q.v.*) and x_o (or y_o).

numerical replay or digital replay

Replaying a CGH using a digital computer by assuming an infinitely repeating hologram function. This results in a series of Dirac's delta functions representing a sampled replay field.

optical replay

Replaying a CGH encoded onto an SLM with device imperfections and an input illumination profile having a finite expanse, resulting in a continuous optical field at the replay plane.

phase element

Each element of a phase-only *base hologram* (*q.v.*) which may contain multiple *phase states* (*q.v.*), may not contain all *phase states* and may/may not contain each *phase state* in equal measures (e.g., 7 *phase elements* (*q.v.*) are required for 1/7 *replay fraction* (*q.v.*) regardless of the number of *phase states* available).

phase quantisation

The twin effects of having a *finite number of phase states* (*q.v.*) for hologram encoding and the *distribution of phase elements* to achieve a normalised *routing fraction* (*q.v.*).

phase state or phase level

Each of the m possible phase values of a multi-level phase modulating device available to encode *routing holograms* (*q.v.*).

replay fraction or routing fraction

The description of the first order of a *routing hologram* (*q.v.*) and therefore a *base hologram* (*q.v.*) replay in terms of a simple fraction along the normalised (by $f\lambda/d$) replay co-ordinates in the x- and y-directions, i.e., ($\sigma = x/x_o, \tau = y/y_o$) where x and y are the harmonic orders and x_o and y_o are the number of *phase elements* (*q.v.*), in the x- and y-directions, respectively; and $|\sigma|$ & $|\tau| \leq 1/2$.

replay peak

The pronounced spot at the replay plane which can be identified as either the fundamental spatial frequency term or one of the higher order harmonics in an *optical replay* (*q.v.*).

replay replication

Subsections of $f\lambda/d$ in size centred at the optical axis, where f is the focal length of the transform lens, λ is the wavelength of illumination and d is the pixel pitch, as a result of the pixel periodicity of the pixellated SLMs.

replay resolution

The separation between two *replay peaks* (*q.v.*) for a given *routing hologram* (*q.v.*) which contains multiple *repeats* of a *base hologram* (*q.v.*).

replay sidelobes

The weak spots at the replay plane surrounding each *replay peak* (*q.v.*) as a result of the diffraction effects at the aperture edges. This occurs in a typical configuration of an *optical replay* (*q.v.*) or in a *numerical replay* (*q.v.*) with a finite support region for the holograms (i.e., zero padded holograms).

routing hologram

A dynamically reconfigurable phase-only computer-generated hologram for beam-steering applications; it is often required to produce only one *replay peak* (*q.v.*). It contains several base holograms and fills up the allocated SBWP of the encoding device.

spatial quantisation

The twin effects of *pixel periodicity* and *finite pixel width* which result in multiple *replay replications* (*q.v.*) and an intensity roll-off across these replications.

sub-hologram

An element of an array of *routing holograms* (*q.v.*) for *N*:*N* optical cross-connects.