# USC–SIPI REPORT #258

## Optoelectronic Sorting Networks

by

**Yibin Lu**

**March 1994**

Signal and Image Processing Institute
**UNIVERSITY OF SOUTHERN CALIFORNIA**
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 404
Los Angeles, CA 90089-2564 U.S.A.

# Dedication

To my parents,
for their constant love and support.

To my wife and my sons,
who fulfill this journey with joy.

# Acknowledgments

First and foremost, I would like to express my deep gratitude to my advisor, Professor Alexander A. Sawchuk, for his guidance and support throughout this work. It was truly an invaluable and wonderful experience for me to work with him both intellectually and personally.

I would also like to thank my dissertation committee: Dr. P. Daniel Dapkus, Dr. B. Keith Jenkins, Dr. Alan Willner, and Dr. C-C Jay Kuo. Their effort and advice have made this work better.

I am graceful to the faculty, staff and friend in Signal and Image Processing Institute (SIPI) for their support during my years at USC. Special thanks go to my officemates and friends, Jeng-Feng Lin, Wei-Feng Hsu and his wife Syying Lee. We share many valuable discussions and great game-time. The friendship and interesting conversations from Dr. Charles Kuznia, Mr. Kuang-Yu Li, Mr. Ching-Chu Huang, and Mr. Chun-Hsiung Chuang are acknowledged.

Last but not least, I would like to express deep love to my parents and my sisters and brother for the love they have given me. Finally, my thanks and love go to my wife Shu-Wen Chang. Her continuous love, care, and support make my PhD dream come true.

# Abstract

This dissertation presents an optoelectronic implementation of multistage sorting networks. Volume (3-D) sorting networks which utilize the high spatial bandwidth and parallelism of optics are discussed. Designs for the diffractive optical elements (DOEs) used for the interconnections and optoelectronic compare-and-exchange (C&E) modules used for the dynamic sorting nodes are given.

The first part of this work describes the network properties of various conventional 2-D sorting networks and the development of their corresponding 3-D structures. A merge procedure is proposed for network expansion in 3-D space. A mathematical framework is established, and a detailed analysis of the network performance is given. The decision of choosing a specific network structure depends on two factors: the cost of optical elements and switching elements; and the complexity and reliability of the network.

Designs of optical interconnection systems for one-to-one linear mappings and 2-D folded shuffles are presented in the second part of this work. Our study of the characteristics of Gaussian microbeams propagating in microscaled diffractive optical elements provides a useful method for analyzing the packing limitations and light efficiency of each network

structure. In our work, short and long distance linear mapping interconnections are accomplished by using microlens arrays. Various approaches are proposed to implement 2-D folded shuffles. Phase-only blazed gratings and off-axis microlenses are used for space-variant designs, and a demagnifying algorithm is proposed for space-semivariant designs. Mathematically modeled error functions are developed for performance analysis.

The last part of this dissertation describes systematic design methods for the optoelectronic 2 input/ 2 output multifunctional C&E modules. We choose smart pixel optoelectronic devices having both optical and electronic functionality for implementing the logic functions. The particular device technologies used are L-SEEDs, FET-SEEDs, and OEICs. We consider the packing density limitations due to three main factors: optical diffraction, pixel size, and power dissipation. Based on state-of-the-art technology, the analysis shows that the power dissipation of the smart pixel devices currently limits the number of modules that can be integrated in a single chip.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Motivation

With the growing number of areas in which computers are being used, there is an ever-increasing demand for more information processing power than today's machines can deliver. Extremely fast electronic computers and communication networks are being sought for many applications to process enormous quantities of data in reasonable amounts of time. However, it is becoming apparent that it will very soon be impossible to achieve significant increase in speed by simply using faster electronic devices, as was done in the past three decades. This is due, on one hand, to the fact that with today's superfast VLSI circuits more time is needed for data to travel between two devices than it takes for it to be processed by either of them. On the other hand, the reduction of distance between devices through very large scale integration is quickly reaching a limit beyond which the reliability and speed of circuit elements decrease [1].

An alternative solution to achieve very high processing capacity is to use parallel system architectures. Here, different processing units are locally or globally interconnected and they process the information simultaneously.

Unfortunately, potential packaging and interconnection bottlenecks of electronic VLSI circuits prevent the development of high density parallel systems. As a result, extensive research efforts have been directed towards free space optoelectronic technologies which utilize the inherent parallelism of optics to provide solutions to the interconnection bottleneck problem. The goal of this research is to explore the potential of utilizing current optoelectronic technologies to construct a special-purpose, 2-D parallel computing architecture, named a sorting network, to transform it into an equivalent 3-D volume architecture, and to evaluate its performance in various ways.

## 1.2 Background

### 1.2.1 Definition

The chosen problem of this research is based on the concept of sorting. Let us begin by giving a formal definition of sorting.

**Definition 1.1** A set of elements are said to satisfy a *linear order* "<" if and only if

1. for any two elements $a$ and $b$, either $a < b$, $a = b$, or $b < a$.
2. for any three elements $a$, $b$, and $c$, if $a < b$ and $b < c$, then $a < c$.

The linear order "<" is usually read "precedes."

**Definition 1.2** Given a sequence $S = \{x_1, x_2, \ldots, x_n\}$ of $n$ items on which a linear order is defined, the purpose of *sorting* is to arrange the elements of $S$ into a new sequence $S' = \{x'_1, x'_2, \ldots, x'_n\}$ such that $x'_i < x'_{i+1}$ for $i = 1, 2, \ldots, n-1$.

To get an intuitive understanding of this definition, $S$ can be thought of as a sequence of names to be arranged in alphabetical order. Another example is a sequence of numbers to be arranged in non decreasing order.

### 1.2.2 Parallel Sorting Architectures

For both practical and theoretical reasons, sorting is one of the most common tasks in general-purpose computation. It was estimated that one fourth of all computer time was devoted to sorting [2]. For example, databases and expert systems often sort the elements of a data structure to simplify searching and the addition of new elements. Furthermore, data manipulation operations like projection, set union, and intersection can be directly implemented by modified sorting algorithms [3]. With the tremendous demand for faster computing speed, and the sharp decline in the price of processing elements, the era of parallel computing systems was born. In the case of sorting we distinguish between two general approaches: multipurpose parallel structures and special-purpose parallel architectures [4].

Multipurpose parallel architectures are destined for computing systems with a broad range of applications. These are usually classified into one of two main categories: single instruction stream multiple data stream

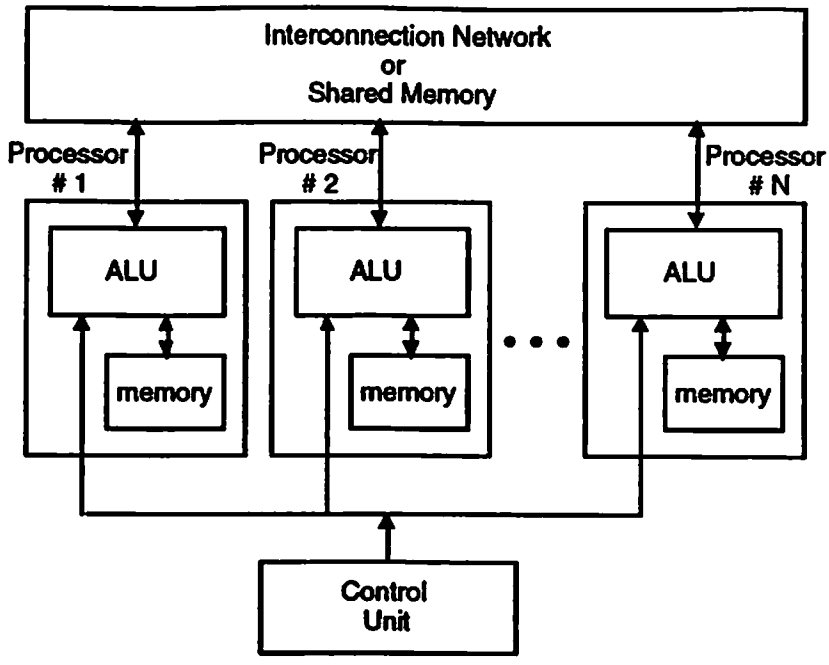



Figure 1.1: SIMD computing system.

(SIMD) computing systems and multiple instruction stream multiple data stream (MIMD) computing system.

An SIMD computing system consists of a number of processors operating under the control of a single instruction stream provided by a central control unit. Figure 1.1 shows the structure of an SIMD computing system with the input and output units omitted. In order to sort a set of data, each processor loads programs and different data into its local memory. During a given time unit, a selected number of processors are active and execute the same instructions such as comparison, addition, or register shifting in a synchronous manner, each on a different data set. In order to be able to exchange data, the processors either communicate through an interconnection network or share a common memory.

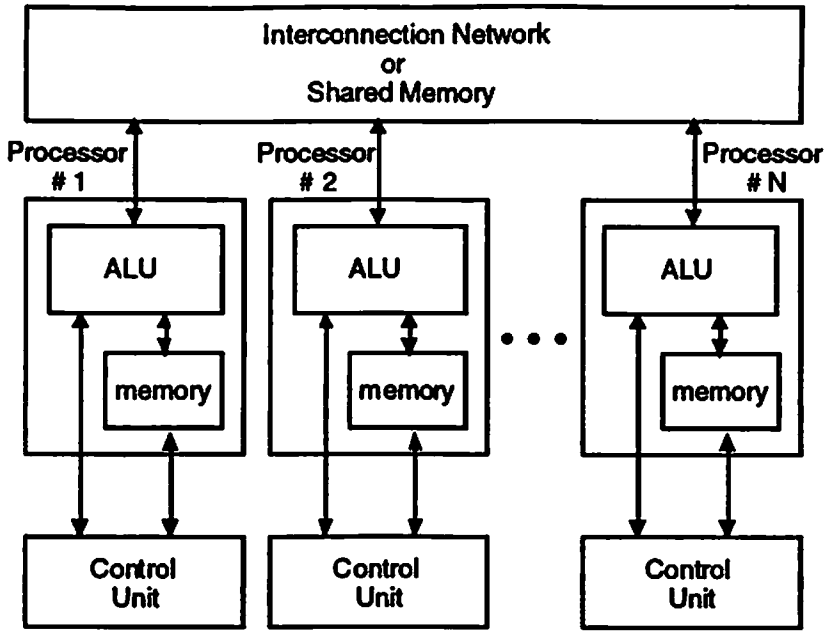



Figure 1.2: MIMD computing system.

In an MIMD computing system, processors posses independent instruction cycles and operate asynchronously. Figure 1.2 shows the structure of an MIMD computing system with the input and output units omitted.

Special-purpose parallel architectures are designed with a particular problem in mind. They result in parallel computing systems well suited for solving that problem, but cannot in general be used for other purposes. Sorting networks fall into this class. They consist of a number of special processors arranged in a certain way and connected to each other through communication lines. Such a "constrained" type of sorting scheme is especially interesting because of its applications and its rich underlying theory. In addition to be used for parallel computing, sorting networks also

have important applications in communications. In particular, parallel processor architectures can be interconnected by pipelined sorting networks serving as message passing systems [5]. Similarly, telecommunication packet switches can be based on sorting networks [6].

### 1.2.3 Interconnection Schemes

Considering an SIMD computing system, most of the processing procedures rely on the interconnection network to perform permutations of data among the processors. There are three most commonly studied interconnection architectures: *Mesh*, *Perfect Shuffle*, and *Hypercube*. We will compare the performance of all three architectures in the case of parallel computing and sorting of $N$ data elements, where $N = m \times m = 2^n$. In this thesis, unless indicated otherwise, all logarithms are assumed to be base 2 and are denoted by $\log(\bullet)$. Thus, we have

$$\log N = n. \tag{1.1}$$

Due to the regularity of data movements within most parallel computing algorithms, the most useful permutations have been studied and a family of permutations called Frequently Used Permutations (FUPs) have been identified [7]. A group of FUPs have been chosen as test permutations and a comparison is made on the number of cycles required for each interconnection architecture to perform the operations [8]. The results are shown in Table 1.1. The Perfect Shuffle and Hypercube all

perform the FUPs in $O(\log N)$, except that the Perfect Shuffle architecture needs only one cycle to perform the perfect shuffle permutation. The Mesh requires at least $O(N)$ cycles to perform any of the FUPs.

| FUP | Mesh | Perfect Shuffle | Hypercube |
|---|---|---|---|
| cyclic shift of $d$ units | $d$ | $2\log N - 1$ | $\log N$ |
| perfect shuffle | $N-1$ | 1 | $\log N$ |
| bit reverse | $2(\sqrt{N}-1)^2$ | $3\log N/2$ | $\log N$ |

Table 1.1: Performance comparison of three types of interconnection architectures based on the number of cycles to accomplish a group of FUPs.

In the case of parallel sorting operations, processors need to perform the compare-exchange instruction and interconnection networks provide the routing instruction, respectively. The performance of Mesh, Perfect Shuffle, and Hypercube in an SIMD sorting machine has been studied [9, 10, 11] and we summarize it in Table 1.2. The Perfect Shuffle and Hypercube all perform the sorting operation in $O(\log^2 N)$. The Mesh requires $O(\sqrt{N})$ cycles to perform the same sorting operation.

| Sorting | Mesh | Perfect Shuffle | Hypercube |
|---|---|---|---|
| route steps | $O(\sqrt{N})$ | $O(\log^2 N)$ | $O(\log^2 N)$ |
| compare-exchange steps | $O(\sqrt{N})$ | $O(\log^2 N)$ | $O(\log^2 N)$ |
| fan-out | 4 | 1 | $\log N$ |
| interconnection | space-invariant | space-variant | space-variant |

Table 1.2: Performance comparison of three types of interconnection architectures based on the number of cycles to accomplish a sorting operation.

Among all three types of interconnection architectures, the Perfect Shuffle has fast data permutation and routing abilities. Another advantage is that it has a single fan-out structure which avoids the problem of communication conflict between processing units caused by a multiple fan-out structure. Thus, we favor the Perfect Shuffle in our system design.

### 1.2.4 Multistage Sorting Networks

Sorting networks can be described by their construction and data format. A 2-D multistage sorting network, which can be implemented by electronic VLSI technology, has its signal channels and switching units arranged in 1-D arrays (node stages) and the link patterns arranged in planes (link stages) as shown in Fig. 1.3. A 3-D multistage sorting network, on the other hand, occupies a physical volume in space, and has its channels and switching units rearranged in 2-D arrays which are interconnected by link patterns in 3-D space as shown in Fig. 1.4. This structure intrinsically favors an optical implementation.

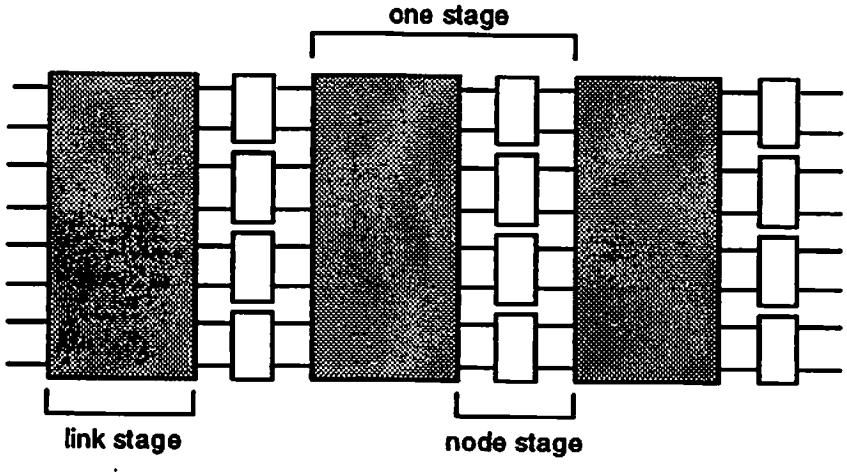



Figure 1.3: 2-D multistage sorting network.

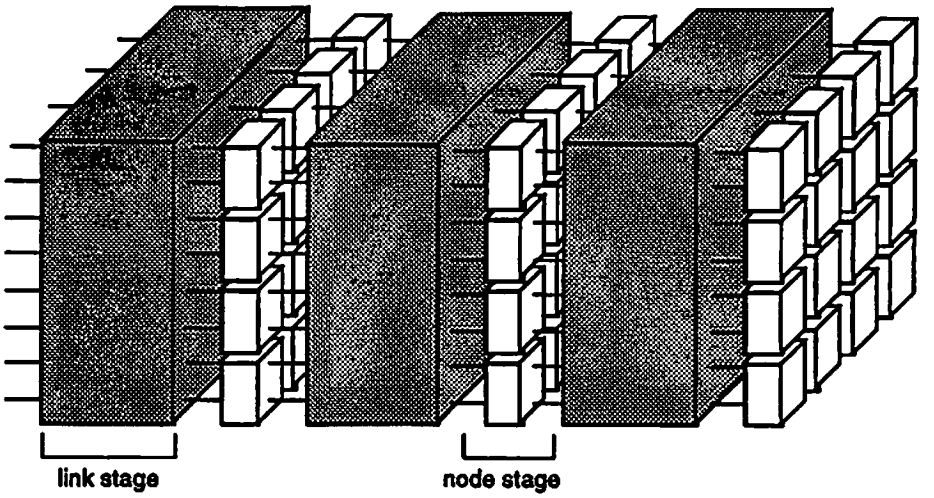



Figure 1.4: 3-D multistage sorting network.

Multistage networks with different interconnection structures, such as the shuffle-exchange network, Banyan network, etc., can be used as a multistage sorting network. The perfect shuffle interconnection network is of most interest because of its relatively regular and simple structure.

There are two kinds of perfect shuffles: 1-D and 2-D perfect shuffles. In the 1-D perfect shuffle, channels are arranged in a 1-D array and are shuffled in a plane. As shown in Fig. 1.5, The $N$ channels (assuming $N$ is even) are divided into two halves, then these two halves are interleaved.

For the 2-D perfect shuffle, channels are arranged in a plane and are shuffled in free space. There are many variations of 2-D perfect shuffle, two of which are the 2-D separable shuffle and the 2-D folded shuffle. The algorithm for a 2-D separable shuffle is performed in two steps as shown in Fig. 1.6. The two steps are to perform 1-D perfect shuffles on the rows and

the columns independently. However, in the 2-D folded shuffle, rows and columns are related in the way that the 2-D plane is a rastered version of a long 1-D channel array as shown in Fig. 1.7. The purpose of 2-D folded shuffle is to perform the 1-D perfect shuffle more efficiently.

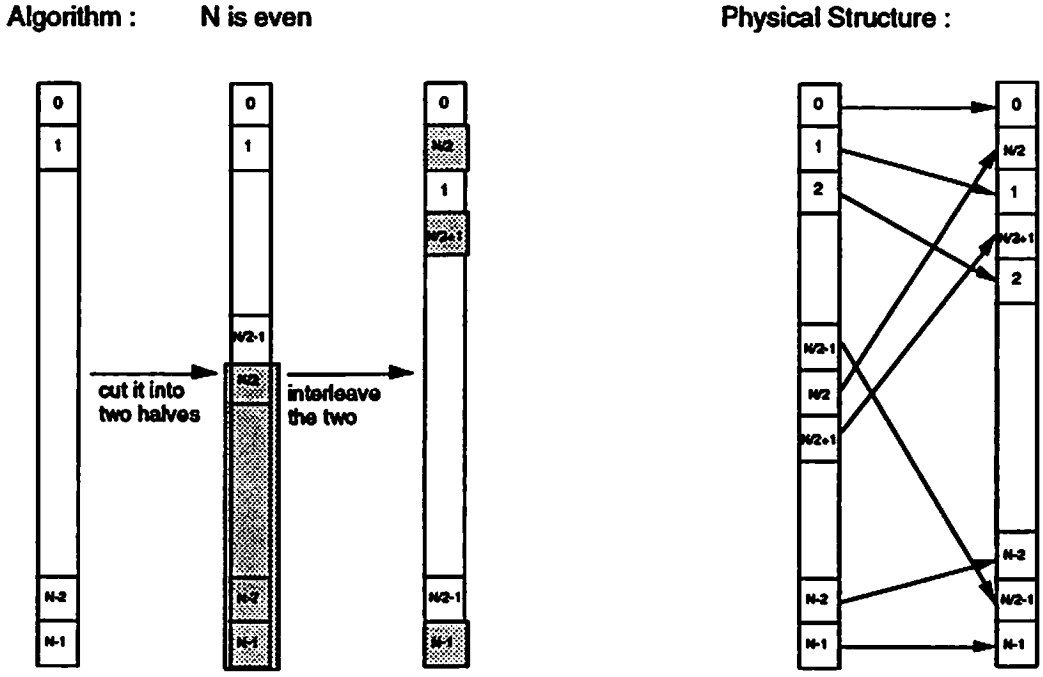



Figure 1.5: 1-D perfect shuffle.

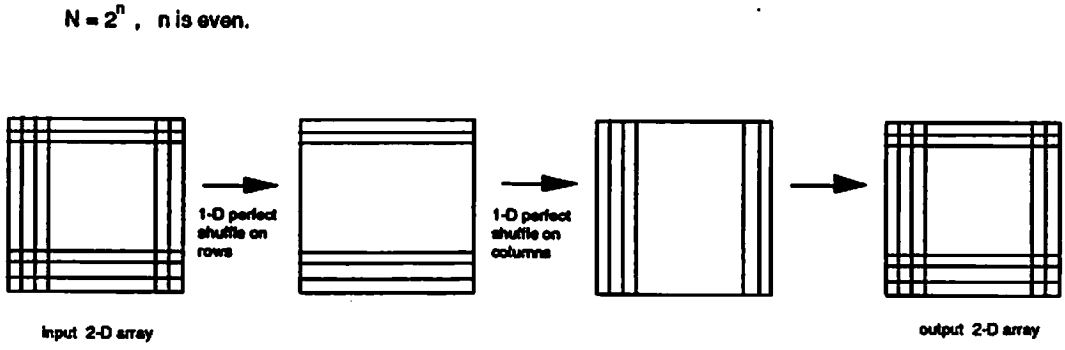



Figure 1.6: 2-D separable shuffle.

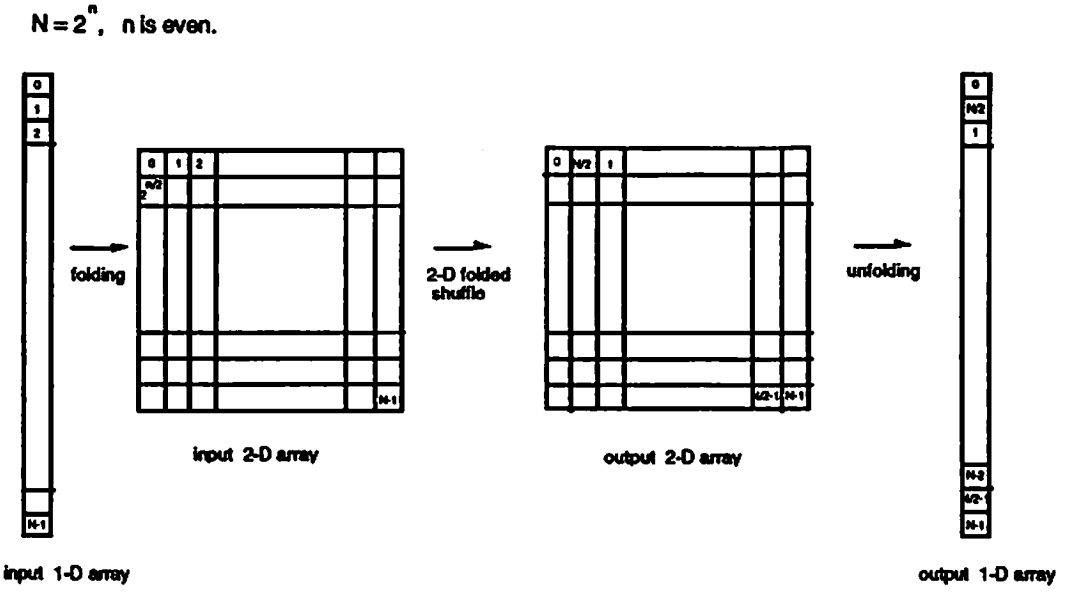



Figure 1.7: 2-D folded shuffle.

The special processors in the node stages of a multistage sorting network are dynamic switching units called compare-and-exchange (C&E) modules. There are three basic functions of a C&E module: bypass, maxsort, and minsort. These C&E modules and their input-output relations are shown in Fig. 1.8. The bypass function performs a simple

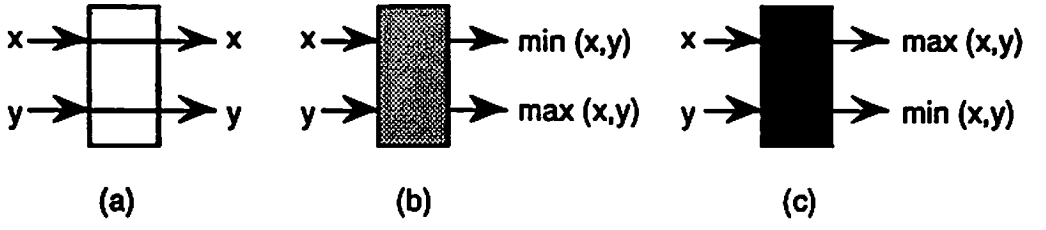



Figure 1.8: Three basic functions of a C&E module. (a) bypass function. (b) minsort function. (c) maxsort function.

straight-through connection. The maxsort function sorts two inputs so that the input whose label is maximum emerges from the port labeled max($x$,$y$) (shown as the top port in Fig. 1.8(c)). The minsort function sorts two inputs so that the input whose label is minimum emerges from the port labeled min($x$,$y$) (shown as the top port in Fig. 1.8(b)).

## 1.3 Objectives

With the increasing demand for solving computation-intensive problems, a volume, dynamic interconnection network that interconnects a 2-D array of channels or processing elements to another is desirable [12]. The objective of this research has been to investigate and study the practicality of a 3-D optoelectronic multistage sorting network utilizing diffractive optics and OEIC technologies. It explores and discusses 3-D structures for various types of planar sorting networks, the network implementation, and the design of optical interconnection elements and dynamic switching units. Both theoretical analysis and practical considerations are included in this work.

## 1.4 Thesis Organization

The body of the thesis has two main parts. The first part consists of Chapter 3, dealing with network architectures and implementations from a system point of view. The second part in Chapter 4, 5, 6, and 7 is devoted to technological issues, design and performance analysis. Figure 1.9 shows the organization of this thesis.

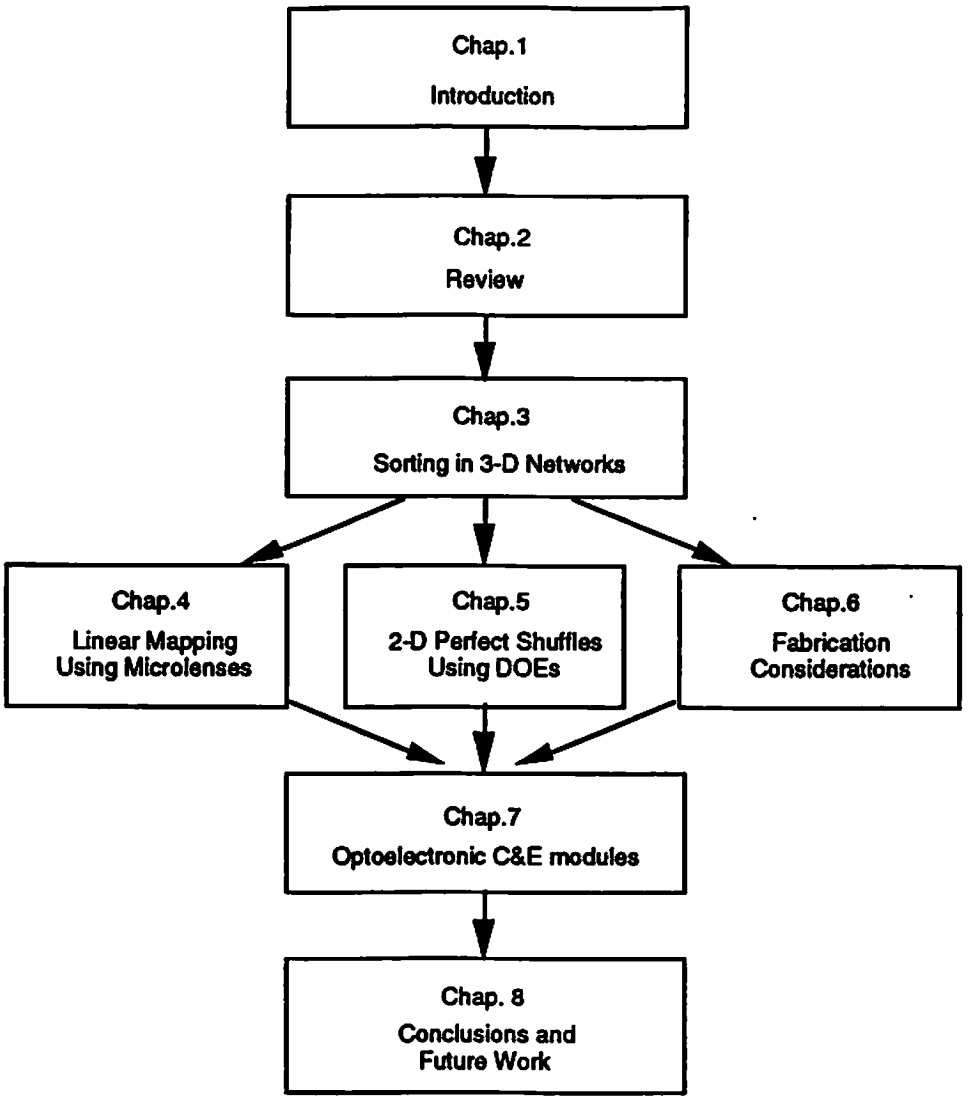



Figure 1.9: Thesis organization

Chapter 2 reviews the work of sorting algorithms on a variety of sorting networks. It also describes inherent problems of electronic and optical technologies for realizing a sorting network. The rest of Chapter 2 gives previous work done on optical interconnections and optoelectronic C&E modules.

Chapter 3 discusses three different types of multistage sorting networks and the relationship between their 2-D and 3-D structures. A merge procedure is presented for network expansion in 3-D space. The mathematical framework is established in this discussion. Toward the end of the chapter we give a comparison of these networks and summarize the results.

In Chapter 4 optical system designs for one-to-one linear mapping interconnection using microlenses are presented. The characteristics of Gaussian microbeams in on-axis diffractive microlens systems are first studied. Optical system structures for short and long distance interconnections are analyzed later on. In the rest of the chapter we focus on the packing limitations and light efficiency of each structure.

Chapter 5 studies the properties of a 2-D perfect shuffle and implements it using diffractive optical elements (DOEs). Phase-only blazed gratings and off-axis microlenses are used for space-variant designs, and a demagnifying algorithm is proposed for space-semivariant designs. The analysis of packing density and light efficiency is given at the end of each section.

In Chapter 6, we systematically analyze the fabrication problems of the multi-level DOEs discussed in Chapter 4 and 5. Mathematical proofs are given through the discussion.

Chapter 7 is devoted to the design of optoelectronic 2 input/2 output compare-and-exchange modules. The chapter starts with the logic design of C&E modules based on truth table minimization, finite state machines, and latching logic. The device technologies we choose to implement the logic circuits are logic self-electro-optic effect devices (L-SEEDs), field-effect transistor-self-electooptic effect devices (FET-SEEDs), and optoelectronic integrated circuits (OEICs) with built-in detectors and laser sources. Concepts and details are given through each design. The rest of the chapter deals with the packing density limitations considering three main factors: optical limitations due to diffraction, pixel size, and power dissipation. The performance analysis is summarized at the end of this chapter.

Chapter 8 gives a brief conclusion of this thesis. A few suggestions for the future research is given.

## 1.5 Contributions

In this research, we systematically study, design, and analyze 3-D optoelectronic sorting networks. The contributions of this work are summarized as follows:

- The investigation of various types of structured sorting networks.
- The exploration of corresponding 3-D sorting network structures.
- The development of merging algorithms for the 3-D network expansion.
- The analysis of the reflected-symmetric property of 2-D folded shuffle.

- Novel designs of optical interconnection architectures for the 2-D folded shuffle based on diffractive optical elements.
- Generic designs of computer-generated diffractive gratings for the 2-D folded shuffle.
- The analysis of packing limits based on the characteristics of the light sources and diffractive optical elements.
- Systematic analysis of the efficiency loss of multi-level DOEs caused by lithographic fabrication problems.
- The design of 2 input/ 2 output optoelectronic compare-and-exchange modules based on different optoelectronic devices.
- The analysis of the performance of C&E modules based on the characteristics of the devices and the complexity of the circuitry designs.

# Chapter 2

# Previous Work - A Review

It is known that serial sorting algorithms require at least $O(N \log N)$ temporal complexity to sort $N$ inputs [2]. Due to the need for rapidly sorting large quantities of information, parallel sorting algorithms which can offer enhanced performance have been studied since at least the 1960's. In 1968, Batcher [13] discovered the bitonic sorting network which has sublinear ($O(\log^2 N)$) temporal complexity to sort $N$ inputs. The bitonic sorting network can be pipelined in a multistage architecture that requires global interconnection patterns and dynamic switching units. Both electronic technology and optical devices have been considered for realizing these kinds of sorting networks. This includes:

- Electronics: very large scale integrated circuits (VLSI)
- Optics: lenses, prisms, holograms, etc.

As pointed out by Goodman et. al. [14], VLSI networks suffer from crosstalk, clock-skew, pin-in/pin-out limitations, and delay problems when the interconnections are spatially dense or long distance. Other problems, including high frequency parasitic coupling between circuit elements,

excessive power dissipation, and impedance mismatch, limit the performance of electronically implemented networks. These limitations restrict the utilization of electronics in volume sorting networks. On the other hand, optics has useful properties such as high bandwidth, low crosstalk and inherent parallelism. Thus, optics is ideally suited for providing the interconnections needed in a volume sorting network.

Unfortunately, due to physical and technical constraints, a high speed, power efficient and easily controllable purely optical C&E module has been very difficult to build. A new group of devices called optoelectronic devices have been developed by researchers in order to utilize the capabilities of both optics and electronics for the dynamic switch elements. These devices include spatial light modulators (SLM), liquid crystal light valves (LCLV), self electro-optic effect devices (SEEDs) based integrated circuits, vertical-cavity surface-emitting lasers (VCSELs) based integrated circuits, and optoelectronic integrated circuits (OEIC) [56, 85]. In the remainder of this section, we will review previous work on optical interconnections and optoelectronic C&E modules.

## 2.1 Optical Interconnections

It has been recognized that optical interconnections can provide the global connections needed between the stages of a interconnection network [14, 15]. For example, each bitonic interstage connection pattern can be emulated by a number of stages of optical perfect shuffle interconnections with global, space-variant communications [10]. Optical elements such as

simple lenses, mirrors, and prisms have be used to implement free-space perfect shuffles [16, 17, 18]. However, as the array size and density of new optoelectronic devices in the interconnection networks have increased, and more complex and irregular non-local interconnections are required, it has become more obvious that conventional deflective and reflective optical elements are unable to meet the requirements. New optical elements, such as holograms and micro-optical components, offering unique wavefront transformation capabilities have been proposed to solve these problems.

Diffractive optical elements (DOEs) are phase-only micro-optical components. Some early DOEs were made by Dammann [19] who used computer-designed surface-relief structures and integrated-circuit etching methods to transfer a binary surface-relief pattern to a dielectric or metallic substrate. These versatile devices can be much thinner and lighter than conventional ones because they need to induce changes of only a fraction of a wavelength in the shape of an optical wavefront. Another advantage of DOEs over conventional bulk optics from a manufacturing point of view is the possibility of monolithic integration of different elements such as micro-lenses and gratings as well as aberration correction optics [20].

The growth of integrated-circuit manufacturing technology and the developments in numerical design techniques starting from the 1970s and 1980s has made diffractive optics a real alternative to their conventional counterparts. Design concepts for computer-generated holograms (CGH) for

diffractive elements have existed for decades [21]. The problem with holograms is that, as more and more optical systems operate at low fiber loss, near-infrared wavelengths, volume holographic recording materials become less sensitive and need special exposure procedures. On the other hand, DOEs are fabricated using simpler repeated masking and etching, which makes it possible to fabricate complex binary optical interconnection patterns cheaply and in volume. Thus, diffractive optics has shown its superiority in this area.

Diffractive optics have so far been applied to various research areas, for example, multiple imaging of single objects [19, 22], optical communications [23, 24], laser beam profile shaping [25, 26] and beam combining [27]. The interest in the all-optical digital computers and photonic switching circuit has also led to a very intensive research effort on the design and fabrication of both binary and multi-level DOEs for 2-D beam array generation [28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40]. Those DOEs perform functions such as power and the clock signal illumination, fan-in, fan-out, and optical interconnection between logic element arrays. In this work, we focus on the implementations of free-space interconnection using diffractive optics.

Extensive research has been done in designing diffractive optical elements for free-space interconnections. From the system design point of view, the proposed implementations for optical interconnections can be classified into following two categories:

**Fourier-Plane DOEs:** These elements, first pioneered by Dammann [19], are periodic Fourier-type surface-relief structures which generate a given far-field pattern insensitive to the shape of incident light beam. The design of the binary phase pattern of a Fourier-plane DOEs is a complicated numerical problem. Two main design methods that utilize the freedom to optimize the phase of grating functions were proposed in 1970s: the nonlinear parametric optimization technique; and the iterative Fourier transform algorithm. The latter, which based on the Gerchberg-Saxton phase retrieval algorithm [41], is presently the most computer-time efficient technique. Parameter optimization method, first applied to the design of DOEs by Dammann [19] and later also known as the simulated annealing algorithm [42], can always find a global optimum of the phase grating. It is the most efficient of the general-purpose nonlinear design methods. Fourier-plane DOEs have been mostly used as beam splitters for providing equal-intensity, equal-spacing spot arrays. They are also capable of generating light spot arrays with irregular patterns and density [43,44]. These elements can serve as space-invariant interconnections which have one-to-many or many-to-many communicating capabilities.

The concept for parallel optoelectronic bus-type and star coupler interconnections (many-to-many) utilizing one single space-invariant diffractive optical system was proposed by Killat [23] and Krackhardt [24], respectively. In their work, a Dammann grating was used as a multiple beam splitter providing simultaneous fan-in and fan-out

capabilities. Morrison [45] demonstrated a novel use of a space-invariant DOE design to create a space-variant Banyan interconnection.

The major drawback of Fourier-plane DOEs is their complicated design process needed because of the position of the transitions and the phase depth must be optimized for the desired beam configuration [46].

**Fresnel-Plane DOEs:** These elements are essentially multifacet devices, in which each facet deflects and focuses a single incident beam in a prespecified manner. Theoretically, any one-to-one space-variant interconnection can be performed using multifacet Fresnel-plane DOEs. For example, microlens arrays fabricated by diffractive optics technology were applied to 2-D free-space optical interconnections by McCormick [47]. DOEs for optical perfect shuffle interconnection have been proposed by Song [48]. It was designed by calculation of the interference fringes between diverging waves from four point sources and a reference plane wave. A different approach has been taken by Zhan [49], in which a quadrant-encoded grating and spatial filters were used to perform 2-D optical perfect shuffles.

As pointed out above, Fresnel-plane DOEs are essentially diffraction gratings with a different fringe period and orientation in each facet. In the realization of these multifacet elements, it is usually necessary to fabricate multi-level surface-relief profiles to achieve sufficient light efficiency. Currently, mask alignment error is the bottleneck in

fabrication of such structure. More detailed discussion will be given in later chapters.

## 2.2 Optoelectronic Compare-and-Exchange Modules

A compare-and-exchange module performs the job of rearranging inputs according to their values. Two major operations, comparison and exchange, need to be done in a single module. These two operations can be implemented with a variety of optical technologies [50]. The optimal technology depends on the requirements of the application of interest.

Zhang [51] demonstrated a circuit that performed the compare-and-exchange operation with ZnS interference filters as bistable devices. The ZnS interference filters were used in less common modes of operation including latching and bidirectional logic. In addition they employed polarization multiplexing and filtering to achieve channel isolation, 4-port bidirectional devices, and reduced feedback.

Murdocca [52] proposed an optical design of C&E modules using programmable nonlatching logic arrays made by symmetric self electro-optic effect devices (S-SEEDs). The comparison and exchange functions were designed utilizing arbitrary finite state machines with regularly interconnected arrays of logic gates and masks. All logic gates had fan-in and fan-out of two and functioned as OR or NOR gates according to the operation of S-SEED devices.

Mao [53] has shown the possibility of applying bistable light emitting diode (BILED) circuits for C&E modules. The threshold characteristics of the photodetector were employed in the C&E modules. The resistors and transistors were needed in the circuit to bias the LEDs such that the characteristic of the BILED circuits can function as latching logic.

More devices and techniques based on state-of-the-art optoelectronic smart pixel technology, which are not described in this short review, are suitable candidates for designing high speed, high density C&E modules. These approaches of using both electronics and optics where optics performs the communication function and electronic circuits compute the C&E operation seem very promising. A more complete review is in [54], [55], [56].

# Chapter 3

# Sorting in 3-D Networks

In this chapter we focus on optical 3-D sorting networks using different free space interconnection structures. Two types of interconnection structures will be discussed in the following sections, namely, one-to-one linear mapping systems and 2-D folded shuffle systems. We first introduce 2-D multistage sorting networks based on these two interconnection patterns, then, we extend 2-D networks into their corresponding 3-D structures. Following that, the idea of expanding smaller networks into a larger one is given. The procedures for the expanding of 2-D and 3-D sorting networks will be presented. This idea can be used to reduce the number of stages required in a sorting network. Finally, we evaluate the hardware requirements and the running time of the sorting networks discussed in this chapter.

## 3.1 Background

### 3.1.1 Notation and Mapping Functions

In the remainder of this chapter, we assume that $N = 2^n$ ($n$ is even) channels needed to be processed. In the case of a 2-D multistage sorting

network, the channels and switching units are arranged in 1-D arrays and the link patterns are arranged in planes. The location of each channel in the 1-D array can be denoted by $P(j)$, where $j$ is the index of channels from 0 to $N-1$. The index $j$ be represented as 2's complement binary numbers, i.e., $j=(x_{n-1}x_{n-2}\ldots x_0)_2$, where $x_i \in \{0,1\}$. On the other hand, a 3-D multistage sorting network occupies a physical volume in space, and has its channels and switching units rearranged in 2-D arrays which are interconnected by link patterns in 3-D space. The $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ 2-D array formed by $N=2^n$ signal channels can be considered as a matrix $A$. The location of each channel in $A$ can be denoted by $a(l,m)$, where $l$ is the row index and $m$ is the column index. We denote $A=[a(l,m)]$. The index $j$ can map to a different pair of indices $(l,m)$ depending on its indexing to $A$. In the following paragraphs, we define the index mapping of a 1-D indexing and its corresponding 2-D row-major indexing. Mapping operations for one-to-one linear mapping and folded shuffles based on row-major indexing are also given.

**Definition 3.1** *Row-Major Indexing* ($I_r$)

This operation maps a 1-D index into its 2-D row-major index as given by

$$\begin{aligned} I_r(P(j)) &= I_r(P(x_{n-1}x_{n-2}\ldots x_0)) \\ &= [a(x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)]. \end{aligned} \qquad (3.1)$$

**Definition 3.2** *1-D and 2-D One-to-One Linear Mapping* ($S_{1l}$ *and* $S_{2l}$)

For the case of the one-to-one imaging, each channel from stage $i$ is linked to the same physical location in stage $i$+1 as given by

$$S_{1l}(P_i(j)) = S_{1l}(P_i(x_{n-1}x_{n-2}\ldots x_0)) = P_{i+1}(x_{n-1}x_{n-2}\ldots x_0), \quad (3.2)$$

$$S_{2l}([a_i(l,m)]) = S_{2l}([a_i(x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)]) = [a_{i+1}(x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)]. \quad (3.3)$$

**Definition 3.3** *1-D Perfect Shuffle and 2-D Folded Shuffle* ($\sigma_{1f}$ *and* $\sigma_{2f}$)

The 1-D shuffling operation is a circular left shift on the bits of an index. The 2-D folded shuffle is actually a 1-D shuffle on the unfolded 2-D data plane. The row and column indices are considered as one binary number, and the 2-D shuffling operation performs a circular left shift on that number.

$$\sigma_{1f}(P_i(j)) = \sigma_{1f}(P_i(x_{n-1}x_{n-2}\ldots x_0)) = P_{i+1}(x_{n-2}x_{n-3}\ldots x_0x_{n-1}), \quad (3.4)$$

$$\sigma_{2f}([a_i(l,m)]) = \sigma_{2f}([a_i(x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)]) = [a_{i+1}(x_{n-2}\ldots x_{\frac{n}{2}}x_{\frac{n}{2}-1}, x_{\frac{n}{2}-2}\ldots x_1x_0x_{n-1})]. \quad (3.5)$$

### 3.1.2 Comparison Criteria

A number of metrics are available to evaluate the performance of a multistage sorting network, namely, *running time*, *node complexity*, and *optical complexity*. These are defined in the next few paragraphs.

**Definition 3.4** *Running Time*

Running time is probably the most important factor in evaluating a

parallel sorting network. It is defined as the time required for a unit of data to pass through the sorting network. Considering a multistage sorting network with $N$ channels, let the propagation time for a single stage optical interconnect be $T_O$, and the C&E module processing time in a single stage be $T_M$. The running time, denoted by $T(N)$, is defined as

$$T(N) = (T_O + T_M) \times (number\ of\ stages). \tag{3.6}$$

**Definition 3.5** *Node Complexity*

Another criterion for evaluating an $N$-channel multistage sorting network is the complexity of the node stage required to sort $N$ data. We denote the number of different functions of the C&E modules required in a sorting network as $F$. These modules are arranged in various combinations at each node stage as shown in Fig. 1.4. We denote by $S$ the number of distinct node plane structures needed to realize a given network. Clearly, the larger the number of C&E modules, $F$, and $S$, the more expensive the solution becomes to obtain. Thus, we define the node complexity $C(N)$ of a sorting network to be the product of the total number of C&E modules, $F$, and $S$, as given by

$$C(N) = F \times S \times (number\ of\ modules). \tag{3.7}$$

**Definition 1.6** *Optical Complexity*

We first define a link complexity $L(N)$ as a measure of the number of distinct point spread functions in the optical interconnection element with

| Optical Interconnection | $L(N)$ |
|---|---|
| Space-invariant | 1 |
| Space-semivariant | $1 < L(N) < N$ |
| Space-variant | $N$ |

Table 3.1: The optical complexity of optical interconnection systems is classified by the degree of space-variance.

respect to the number of channels, assuming no fanout. It can be classified by its degree of space-variance as shown in Table 3.1.

The optical complexity of a $N$-channel sorting network, denoted by $O(N)$, is defined as the product of $L(N)$ and the number of different interconnection patterns in the network as given by

$$O(N) = L(N) \times (number\ of\ stages) \times (number\ of\ patterns). \quad (3.8)$$

It is important to point out that the link complexity of a specific interconnection structure may vary due to different optical system designs. For example, a 2-D folded shuffle can be realized using both space-semivariant and space-variant designs. In this research, we focus on the space-variant design, that is, $L(N) = N$. Another interconnection structure we are interesting in here is one-to-one linear mapping. It can be provided using space-invariant lens systems. Thus, the optical complexity of this type of interconnection is equal to one ($L(N) = 1$).

## 3.2 Cylindrical Spanke Sorting Networks

### 3.2.1 2-D Structures

One of the simplest structures of multistage sorting networks is called an odd-even transposition sorting network [2] shown in Fig. 3.1. It needs $N$ dynamic node stages to sort $N$ inputs, $N \geq 3$. It has also been called a Spanke network [57]. In this type of network, the shortest path through the network passes through $N/2$ C&E modules, while the longest path through the network passes through $N$ C&E modules. The drawback of this arrangement is that there may be path-dependent synchronization, loss, and crosstalk problems.

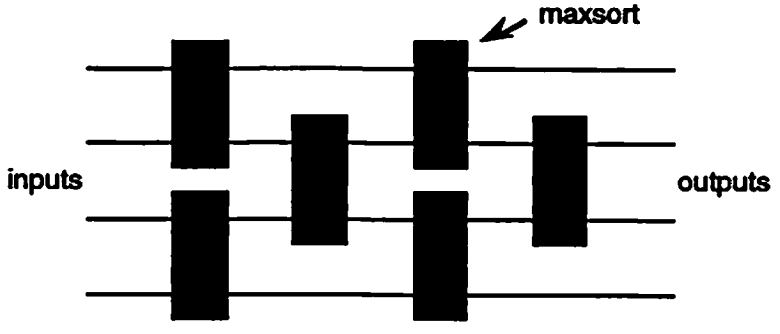



Figure 3.1: Odd-even transposition sorting network with inputs $N$=4.

We can add one C&E module in each even stage of the odd-even transposition sorting network and build a new kind of network called the cylindrical Spanke network [58]. Figure 3.2a and 3.2b show two different architectures for $8 \times 8$ cylindrical Spanke networks. The type I cylindrical Spanke network needs $N$ stages to sort $N$ numbers while the type II

network needs $N$-1 stages to sort $N$ numbers. The half open boxes represent C&E modules that wrap around and connect channel 1 and 8. We can redraw the architecture shown in Fig. 3.2c to present its cylindrical nature.

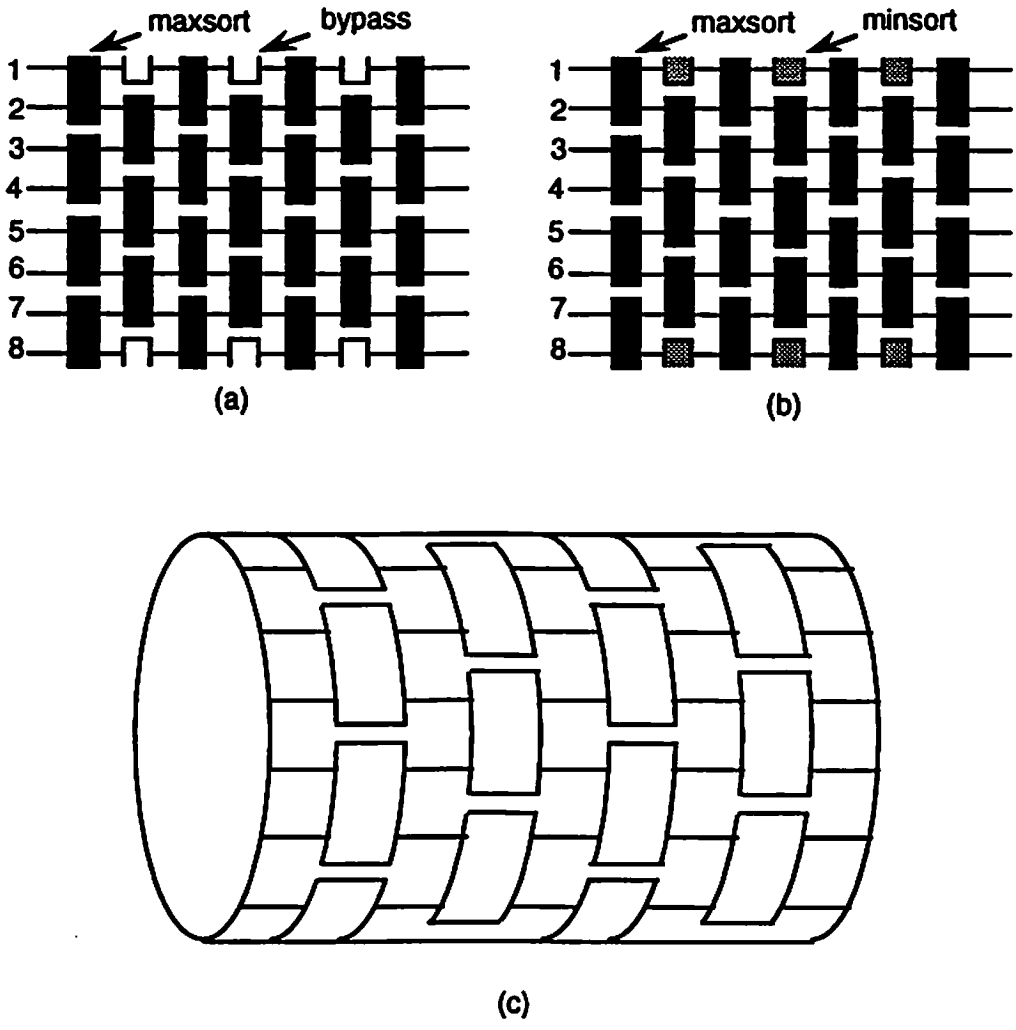



Figure 3.2: Architecture of the cylindrical Spanke network (a) type I drawn in a 2-D structure; (b) type II drawn in a 2-D structure; (c) network wrap around in 3-D space. The boxes in (a) and (b) represent individual 2 input/ 2 output C&E modules, the half open boxes represent modules that wrap around and connect input 1 and 8. The black, gray and white colors represent maxsort, minsort and bypass function respectively.

### 3.2.2 3-D Structures

In order to map a 2-D cylindrical Spanke sorting network onto its 3-D structure, a special type of 2-D indexing called serpentine raster indexing is proposed. The characteristic of this indexing is the adjacent arrangement of the first channel and the last channel and the snake-like arrangement of the rest of the channels. Here, two possible index arrangements are shown in Fig. 3.3.

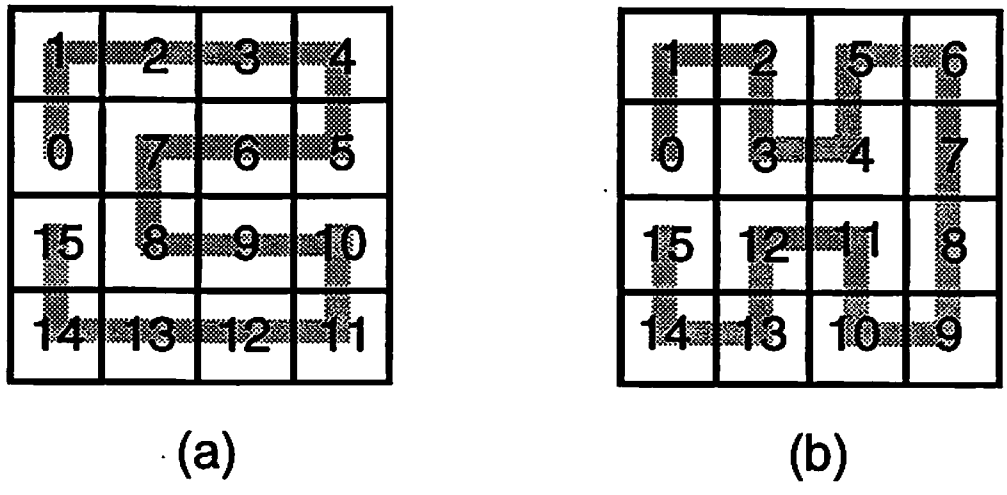



Figure 3.3: Sixteen channels are arranged in two types of serpentine raster indexings.

The index mapping operation $I_{serp}(\bullet)$ of the serpentine arrangement shown in Fig. 3.3b can be generalized as follows:

$$\begin{aligned} I_{serp}(P(j)) &= I_{serp}(P(x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}x_{\frac{n}{2}-1}\ldots x_1x_0)) \\ &= [a(l,m)], \end{aligned} \tag{3.9}$$

where

$$l = \begin{cases} (1 + x_{n-1}) \times 2^{\frac{n}{2}-1} - 1 - (j)_{modulo\ 2^{\frac{n}{2}-1}} & if\ x_{\frac{n}{2}} = 0 \\ x_{n-1} \times 2^{\frac{n}{2}-1} + (j)_{modulo\ 2^{\frac{n}{2}-1}} & if\ x_{\frac{n}{2}} = 1 \end{cases}, \qquad (3.10)$$

and

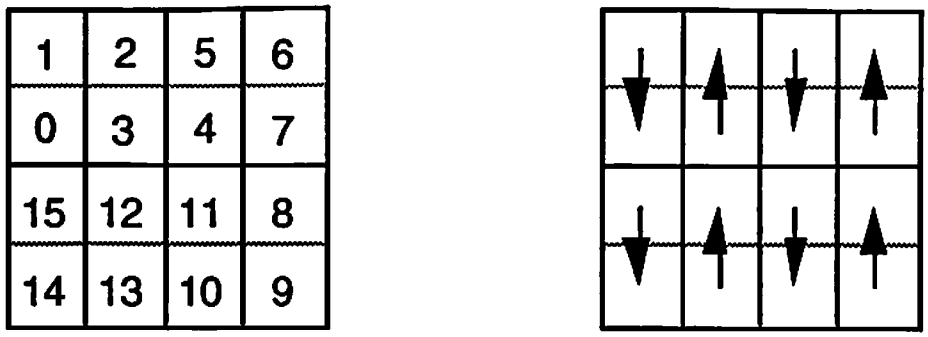
$$m = \begin{cases} decimal\ value\ of\ (x_{n-2} \ldots x_{\frac{n}{2}} x_{\frac{n}{2}-1}) & if\ x_{n-1} = 0 \\ (2^{\frac{n}{2}} - 1 - decimal\ value\ of\ (x_{n-2} \ldots x_{\frac{n}{2}} x_{\frac{n}{2}-1}))_{modulo\ 2^{\frac{n}{2}-1}} & if\ x_{n-1} = 1 \end{cases} \qquad (3.11)$$

Based on the new indexing in Fig. 3.3b, two different node plane arrangements using 2 input/ 2 output C&E modules are needed to implement the even stages and the odd stages of the network shown in Fig. 3.4. The arrows in the 2 input / 2 output C&E modules point to the position of max(x,y).

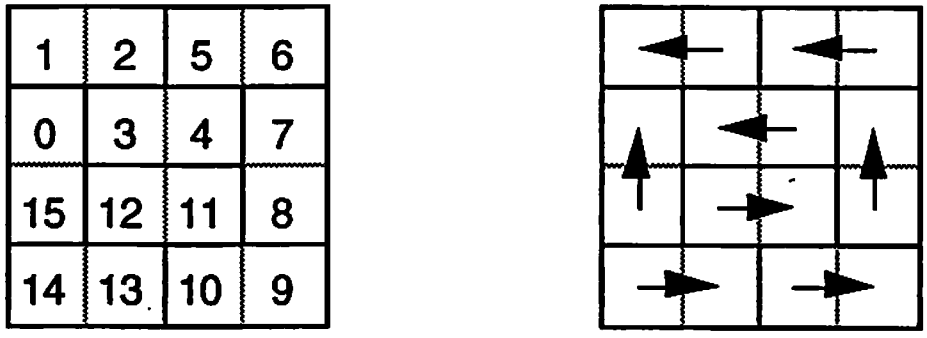
Following the definition, we have $F = 2$ and $S = 2$. The running time $T(N)$ and the node complexity $C(N)$ of this type of network structure can be calculated by

$$T(N) = (N-1) \times (T_O + T_M), \qquad (3.12)$$

$$C(N) = 4 \times N(N-1)/2. \qquad (3.13)$$




(a) odd stage




(b) even stage

Figure 3.4: The 2-D arrangement of 2 input/ 2 output C&E modules for odd and even node stages.

Since the interconnection structure of this type of network is simply a one-to-one linear mapping between two stages, the optical system design is especially easy. Assuming that the node stages in the cylindrical Spanke network are made from OEIC devices that each stage receive light from one side and transmit light from another side, 4*f* imaging systems can be used as link stages shown in Fig. 3.5. Therefore, node planes located at even

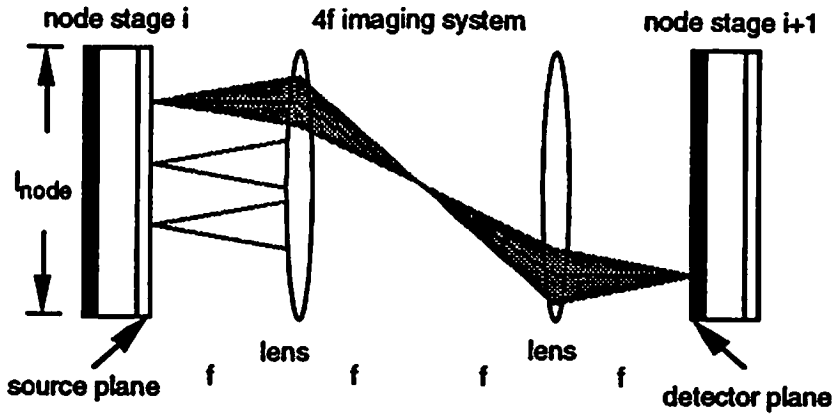



Figure 3.5: One-to-one imaging interconnects using 4f imaging systems.

stages of the network need to be rotated 180 degrees. This type of space-invariant design has an optical complexity of one.

A more simple and dense system setup shown in Fig. 3.6 is to pack all node stages together without any lens system between them [77]. However, the distance between two stages cannot be too small because there is a minimum space required to remove the heat generated by the power dissipation of the node stages. Heat removal from the node stage is very important for the network in order to maintain the temperature of C&E modules within parameters for efficient and reliable operation. Figure 3.7 illustrates the potential of air-cooling technology as a function of power density [59]. Also, the distance between two stages can not be too large. If the transmitting light has a divergence angle, the optical coupling efficiency and the channel crosstalk will be determined by the size of the detectors and the spacing between the adjacent detectors, respectively. These

systems are expected to have very low optical efficiency [77]. In section 4.3 we discuss a related system that uses lens array to image the light from sources to detectors.

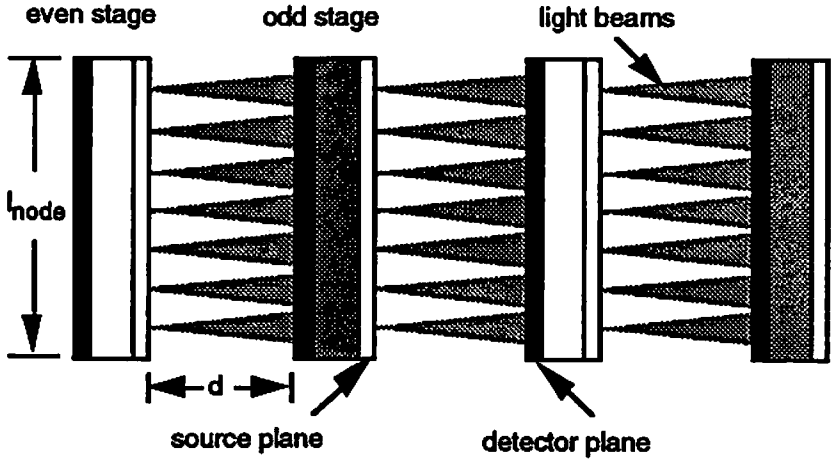



Figure 3.6: Directly coupled structure of a one-to-one linear mapping interconnection system.

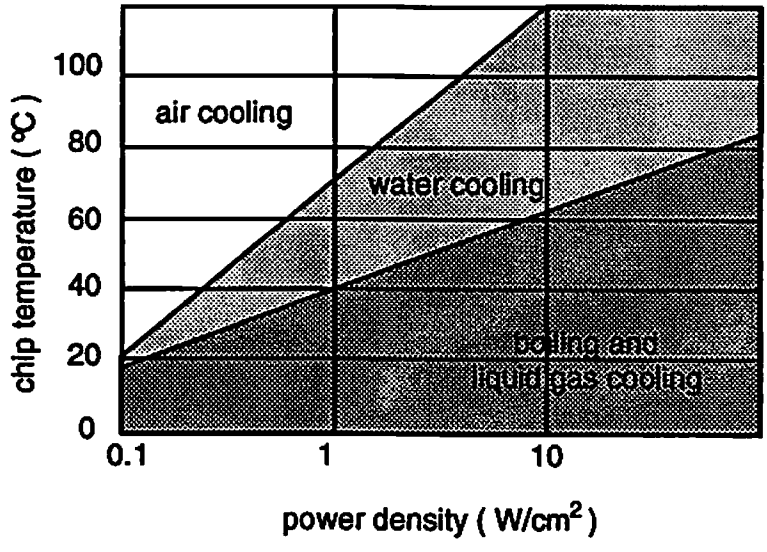



Figure 3.7: Cooling regimes of various cooling technologies.

## 3.3 Bitonic Sorting Networks

Another type of fixed interconnection sorting network of interest in this work is the bitonic sorting network using perfect shuffle interconnections [10]. To sort $N$ inputs, this kind of network is decomposed into $\log N$ levels, with each level having $\log N$ node stages. Thus, the network has a total of $(\log N)^2$ node stages. Figure 3.8 shows a 2-D $8 \times 8$ network structure.

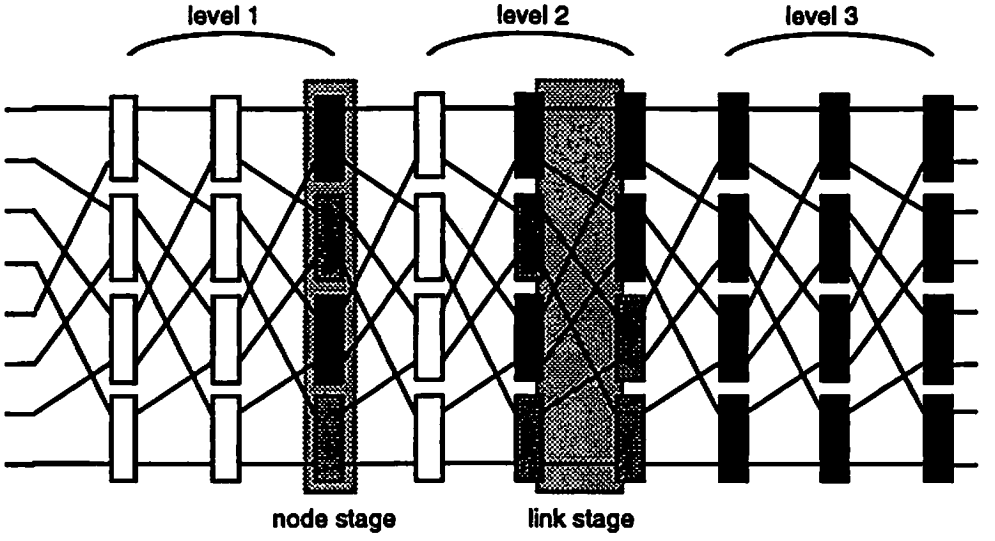



Figure 3.8: Bitonic sorting network based on 1-D perfect shuffle interconnections. Boxes drawn with different shadings represent 2 input/ 2 output C&E modules performing different functions. Black represents the maxsort function, gray represents the minsort function, and white represents the bypass function.

It is clear that there are $1 + \log N$ different node plane structures and that the network requires all three types of 2 input/ 2 output C&E modules arranged in a special sequence. Thus, $F = 3$ and $S = 1 + \log N$. In order to determine the sorting function of a arbitrary C&E module located in a

bitonic sorting network with $N = 2^n$ channels, each C&E module can be considered as an entry in the 2-D switch plane and can be denoted by $m(i,j)$, where $i = 0,1,\ldots,\frac{N}{2}-1$ and $j = 0,1,\ldots,(\log N)^2 - 1$ The index $i$ can also be represented by $N$-1 binary bits, i.e., $i = (x_{n-2}x_{n-3},\ldots,x_0)$. Using the following procedures, we can determine the sorting function of a C&E module by its indices $i$ and $j$.

**Step 1:** First we calculate the value of $r$

$$r = \log N - \left\lceil \frac{j}{\log N} \right\rceil - 1,$$

where $\lceil y \rceil$ represents the biggest integer less or equal to $y$. If $r = 0$, the assigned function of a module in the location $m(i,j)$ is *maxsort*. Otherwise, we go to Step 2.

**Step 2:** In this step, we calculate the value of $u$

$$u = (j)_{modulo\ \log N}.$$

If $u < r$, the assigned function of a module in the location $m(i,j)$ is *bypass*. Otherwise, we go to Step 3.

**Step 3:** Here, we calculate the value of $p$, $p = u - r$. Let us look at the $p+1$ bit of the binary number of $i$, where $i = (x_{n-2}x_{n-3},\ldots,x_p,\ldots,x_0)$. If $x_p = 0$, the assigned function is *maxsort*. If $x_p = 1$, the assigned function is *minsort*.

Again, If we want to extend this 2-D network into its 3-D structure, we can arrange 2 input/ 2 output C&E modules in 2-D arrays and put 2-D folded shuffle interconnections between every two stages to form a 3-D multistage sorting network. Since the structure of multistage interconnection networks using 2-D folded shuffles has been extensively studied in [60] and its interconnections depend on the optical devices chosen by the system designer, we just show a basic 3-D structure in Fig. 3.9.

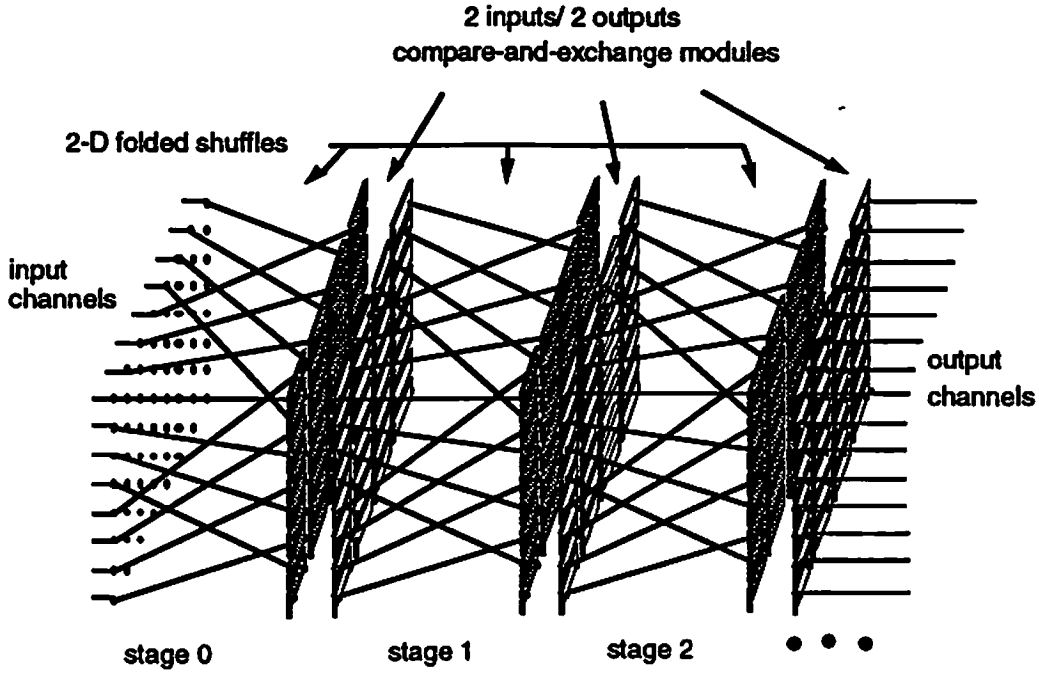



Figure 3.9: 3-D bitonic sorting network.

We can calculate the running time $T(N)$ and the node complexity $C(N)$ of this type of network structure using Eq. (3.6) and Eq. (3.7) as given by

$$T(N) = (\log N)^2 \times (T_O + T_M), \tag{3.14}$$

$$C(N) = 3 \times (\log N) \times \frac{1}{2} N (\log N)^2. \tag{3.15}$$

The optical complexity of the network depends on the optical system designed for the 2-D folded shuffles. Here, we focus on the space-variant design, that is, $O(N) = N$.

## 3.4 Merged Bitonic Sorting Networks

From the previous section, we know that the bitonic sorting network needs $(\log N)^2$ node stages to sort $N$ inputs. Here we discuss techniques for combining four $N$ input / $N$ output bitonic sorting networks to form a larger $4N$ input/ $4N$ output sorting network. Our intent is to find ways to expand the dimensionality of sorting networks. In reality, networks are designed to have a standard number of inputs and outputs. If we can combine small networks to form a larger one without redesigning the entire structure, the cost required to extend the number of channels in a network can be greatly reduced.

### 3.4.1 2-D Structures

The bitonic sorting network has an inherent split-merge design. The technique used to sort a sequence of $N$ numbers is to first generate two sorted subsequences each with $N / 2$ numbers. These two subsequences are

placed together with one subsequence in reverse order and sent into $\log N$ node stages to produce an output sorted in order. Following this structure, $N/2$ sorted outputs can be generated by two $N/4$ sorted inputs along with $(\log N)-1$ node stages, and so on. An example is shown in Fig. 3.10, which shows the 2-D structure of a 16 input/ 16 output sorting network expanded from four 4 input/ 4 output sorting networks. There are $\frac{1}{2}(\log N+1)\log N$ node stages required to sort $N$ input channels. It is important to point out that the interconnections between every two stages are no longer a fixed pattern. This type of sorting network with $N$ channels has $\log N$ different interconnection patterns and node plane structures. We call this type of network a merged bitonic sorting network.

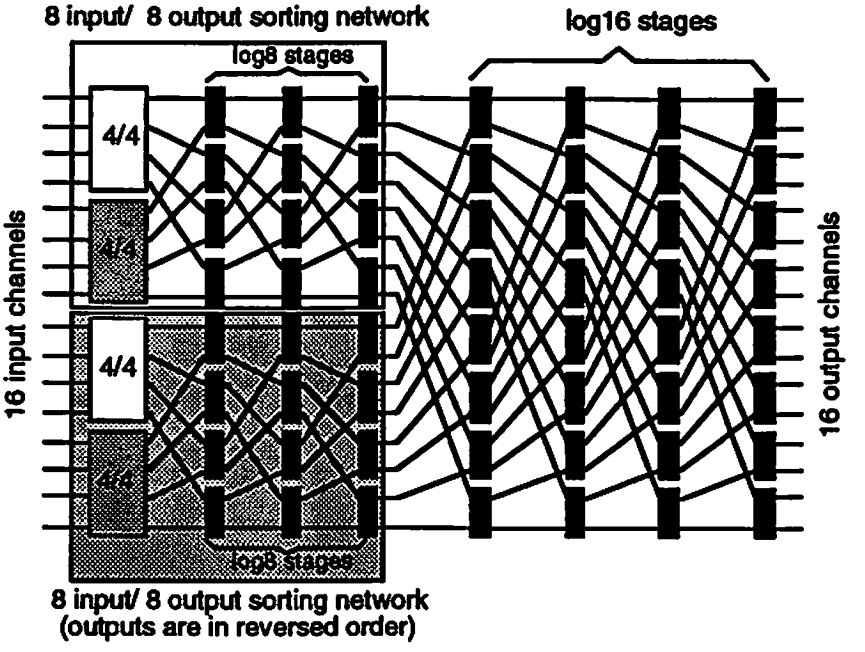



Figure 3.10: A 16 input/ 16 output sorting network merged from four 4 input/ 4 output sorting networks. The sorted outputs of the sorting networks bounded by gray boxes are in reversed order.

### 3.4.2 3-D Structures

A 2-D merged bitonic sorting network can also be extended into a 3-D structure. Again, we assume that there are $N = 2^n$ signal channels arranged in a 2-D plane ($n$ is even) so we have a $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ array. We can use four such arrays to generate a $2^{n+1} \times 2^{n+1}$ array with a total of $4N$ channels as shown in Fig. 3.11.

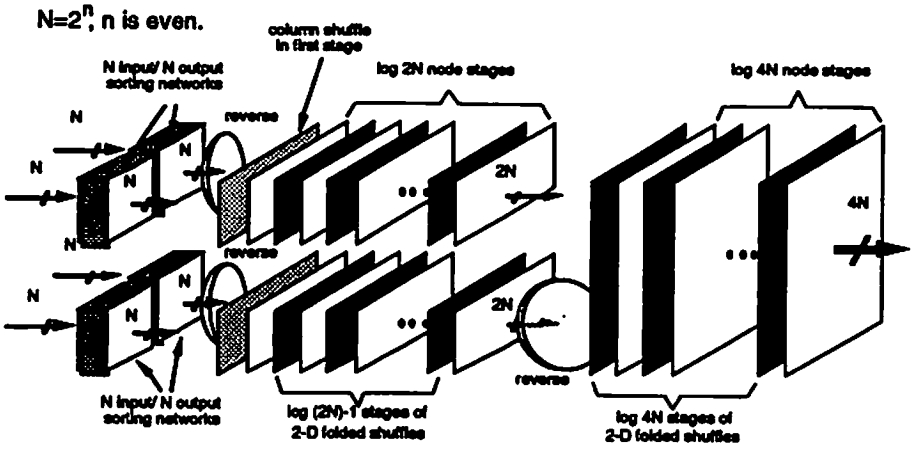



Figure 3.11: A 3-D $4N$ input/ $4N$ output sorting network merged from four 3-D $N$ input/ $N$ output sorting networks.

A new merge algorithm is proposed in this research to perform a 3-D network expansion. In order to analyze the algorithm, we define certain fundamental operations in the following section.

### 3.4.3 Fundamental Operations

Following the description in the previous sections, the $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ 2-D array formed by $N = 2^n$ signal channels can be considered as a matrix $A$ and the

location of each channel can be denoted by $a(l,m)$, where $l$ is the row index and $m$ is the column index. These two indices can be represented as 2's complement binary numbers, i.e., $l = (x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}})_2$, $m = (x_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)_2$, where $x_i \in \{0,1\}$. We denote $A = [a(l,m)]$.

**Definition 3.7** *2-D column shuffle* ($\sigma_{2c}$)

This operation performs 1-D perfect shuffle on the columns of $A$ as given by

$$\begin{aligned}\sigma_{2c}(A) &= \sigma_{2c}([a(x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)]) \\ &= [a(x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-2}\ldots x_1x_0x_{\frac{n}{2}-1})]. \end{aligned} \quad (3.16)$$

**Definition 3.8** *2-D row shuffle* ($\sigma_{2r}$)

This operation performs 1-D perfect shuffle on the rows of $A$ as given by

$$\begin{aligned}\sigma_{2r}(A) &= \sigma_{2r}([a(x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)]) \\ &= [a(x_{n-2}x_{n-3}\ldots x_{\frac{n}{2}}x_{n-1}, x_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)]. \end{aligned} \quad (3.17)$$

**Definition 3.9** *Left-Right Combination* ($C_{lr}$)

The left-right combination of two $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ matrices $A$ and $B$ generates a $2^{\frac{n}{2}} \times 2^{\frac{n}{2}+1}$ matrix $D$. The $l^{th}$ row, $m^{th}$ column entry in $A$ is placed into the $l^{th}$ row, $m^{th}$ column entry in $D$ and the $l^{th}$ row, $m^{th}$ column entry in $B$ is placed into the $l^{th}$ row, $(2^{\frac{n}{2}} + m)^{th}$ column entry in $D$ as given by

$$\begin{aligned} D &= C_{lr}(A,B) \\ &= C_{lr}([a(x_{n-1}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0)\ b(x_{n-1}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0)]) \\ &= [d(x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}, px_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)], \end{aligned} \quad (3.18)$$

where $p \in \{0,1\}$. Thus,

$$d(x_{n-1}\ldots x_{\frac{n}{2}}, px_{\frac{n}{2}-1}\ldots x_1x_0) = \begin{cases} a(x_{n-1}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0) & \text{if } p=0 \\ b(x_{n-1}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0) & \text{if } p=1 \end{cases}. \tag{3.19}$$

**Definition 3.10** *Up-Down Combination* ($C_{ud}$)

The up-down combination of two $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ matrices $A$ and $B$ is to generate a $2^{\frac{n}{2}+1} \times 2^{\frac{n}{2}}$ matrix $E$. The $l^{th}$ row, $m^{th}$ column entry in $A$ is placed into the $l^{th}$ row, $m^{th}$ column entry in $E$ and the $l^{th}$ row, $m^{th}$ column entry in $B$ is placed into the $(2^{\frac{n}{2}}+l)^{th}$ row, $m^{th}$ column entry in $E$ as given by

$$\begin{aligned} E &= C_{ud}(A,B) \\ &= C_{ud}\left(\begin{bmatrix} a(x_{n-1}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0) \\ b(x_{n-1}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0) \end{bmatrix}\right) \\ &= [e(qx_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)], \end{aligned} \tag{3.20}$$

where $q \in \{0,1\}$. Thus,

$$e(qx_{n-1}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0) = \begin{cases} a(x_{n-1}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0) & \text{if } q=0 \\ b(x_{n-1}\ldots x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0) & \text{if } q=1 \end{cases}. \tag{3.21}$$

### 3.4.4 Merging Algorithm

The merging algorithm can be expressed as follows:

**Step 1. reverse:** To reverse the order of $N$ channels in a $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ array, we can use an imaging system built with a lens shown in Fig. 3.12.

$N=2^n$, n is even.

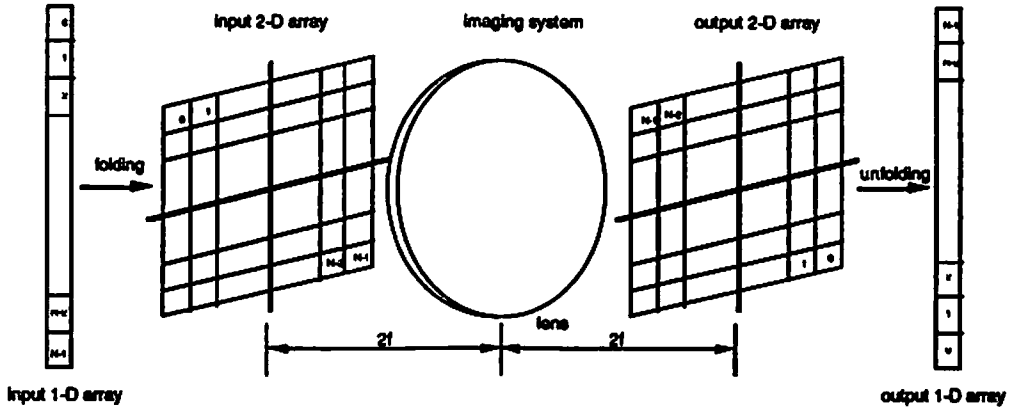



Figure 3.12: 2-D input signal array is reversed using an imaging system.

**Step 2. left-right merge:** One $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ signal array $A$ and one reversed $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ signal array $B$ are first left-right combined into a $2^{\frac{n}{2}} \times 2^{\frac{n}{2}+1}$ signal array $D$. This 2-D signal array $D$ has a shuffled row-major indexing as given by

$$\begin{aligned} D &= C_{lr}(A,B) \\ &= [d(x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}}, px_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)], \end{aligned} \tag{3.22}$$

where $p \in \{0,1\}$.

Following that, $\log_2 2N$ stages of $2^{\frac{n}{2}} \times 2^{\frac{n}{2}+1}$ folded shuffles and node planes are needed to form a $2N$ input/ $2N$ output sorting network.

Problems occur when we try to map a 1-D signal array into a 2-D structure. Let us consider that two 1-D signal arrays which are the 1-D versions of the 2-D signal array $A$ and $B$ are up-down combined and folded into a $2^{\frac{n}{2}} \times 2^{\frac{n}{2}+1}$ signal array $R$ using row-major indexing as shown in Fig. 3.13.

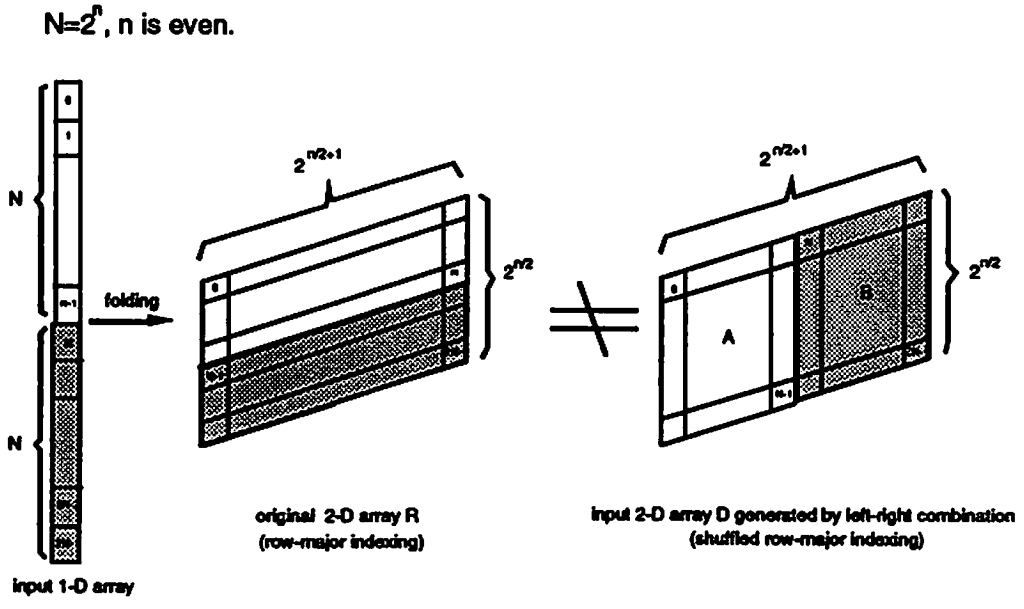



Figure 3.13: The 2-D signal array generated using left-right combination has a shuffled row-major indexing instead of its original row-major indexing.

$$R = \begin{bmatrix} a(x_{n-1}\ldots x_{\frac{n}{2}+1}, x_{\frac{n}{2}}\ldots x_1 x_0) \\ b(x_{n-1}\ldots x_{\frac{n}{2}+1}, x_{\frac{n}{2}}\ldots x_1 x_0) \end{bmatrix}$$
$$= [r(px_{n-1}\ldots x_{\frac{n}{2}+1}, x_{\frac{n}{2}}\ldots x_1 x_0)], \qquad (3.23)$$

where $p \in \{0,1\}$.

If we apply the 2-D folded shuffle to a 2-D signal array $R$, the output 2-D signal array $R_f$ is given by

$$\begin{aligned} R_f &= \sigma_{2f}([r(px_{n-1}\ldots x_{\frac{n}{2}+1}, x_{\frac{n}{2}}\ldots x_1x_0)] \\ &= [r(x_{n-1}\ldots x_{\frac{n}{2}+1}x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0p)]. \end{aligned} \tag{3.24}$$

It is clear that we cannot obtain the same 2-D array by applying the 2-D folded shuffle to the 2-D array $D$. However, if we apply a 2-D column shuffle to the 2-D array $D$, the output 2-D array $D_f$ will be equivalent to $R_f$ as shown in Fig. 3.14.

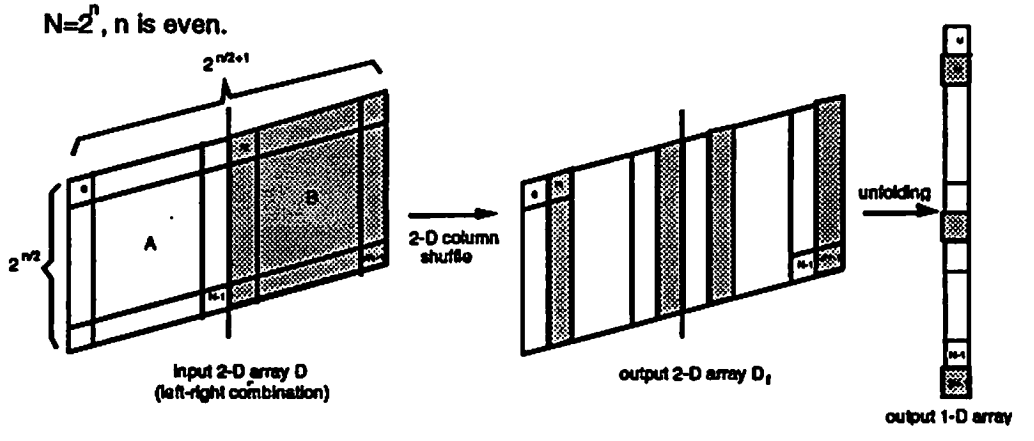



Figure 3.14: Left-right combination of two $N \times N$ networks column shuffled in 3-D space.

$$\begin{aligned} D_f &= \sigma_{2c}([d(x_{n-1}\ldots x_{\frac{n}{2}+1}x_{\frac{n}{2}}, px_{\frac{n}{2}-1}\ldots x_1x_0)]) \\ &= [d(x_{n-1}\ldots x_{\frac{n}{2}+1}x_{\frac{n}{2}}, x_{\frac{n}{2}-1}\ldots x_1x_0p)]. \end{aligned} \tag{3.25}$$

Thus, a 2-D column shuffle is required to replace the 2-D folded shuffle at the first link stage.

**Step 3. reverse:** By repeating step 1 and step 2, two sorting networks with $2^{\frac{n}{2}} \times 2^{\frac{n}{2}+1}$ signal arrays can be generated. Again, to prepare for the final merge step, the output 2-D signal array of one of the sorting networks is reversed by the imaging system.

**Step 4. up-down merge:** In this step, the 2-D signal arrays of two $2N$ input/ $2N$ output sorting networks are combined into a $2^{\frac{n}{2}+1} \times 2^{\frac{n}{2}+1}$ signal array using an up-down arrangement as shown in Fig. 3.15. Since the new array has the same row-major indexing, no special treatment is required at the first link stage. Thus, we add $\log 4N$ stages of $2^{\frac{n}{2}+1} \times 2^{\frac{n}{2}+1}$ 2-D folded shuffles and node planes to finish the 3-D structure of a $4N$ input/ $4N$ output merged bitonic sorting network.

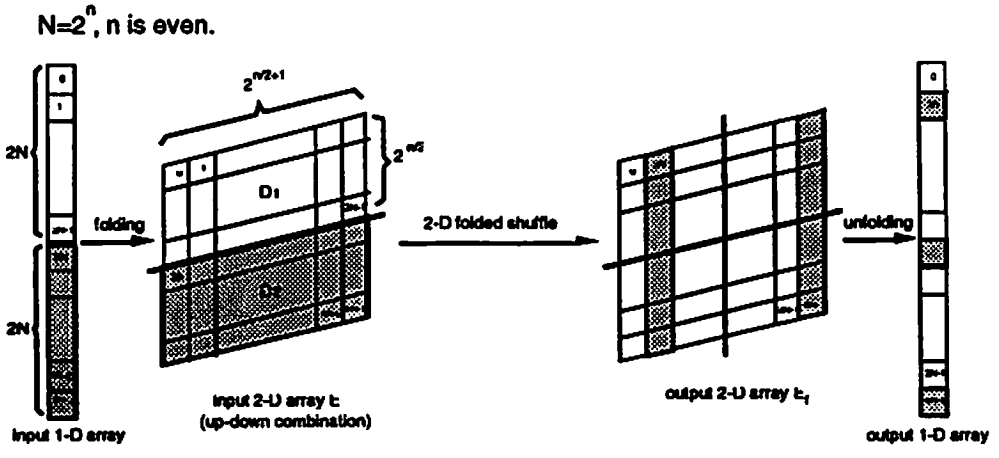



Figure 3.15: System structure of two $N \times 2N$ networks merged into one $2N \times 2N$ network in 3-D space.

## 3.5 Network Evaluation

In this section, we evaluate all three kinds of sorting networks, namely, the cylindrical Spanke sorting network (CSSN), bitonic sorting network (BSN), and merged bitonic sorting network (MBSN), by their running time, node complexity, and optical complexity.

### 3.5.1 Running Time

There are two kinds of time delay within each stage: the propagation time delay caused by interconnection distance, and the switching time delay caused by the C&E module. Let us assume that the distance between two stages of a one-to-one linear mapping system using a directly coupled structure as shown in Fig. 3.6 is $d$. Thus, it has a propagation time delay $T_{O2i}$ equal to the time required for light to travel the distance $d$. State-of-the-art optoelectronic devices have a switching time delay of about 1 ns. Figure 3.16 shows the comparison between the propagation time delay and the switching time delay based on different values of $d$. If the optical system between two node stages requires less than $30\,cm$ space to provide the interconnections, the time delay of each stage is dominated by the switching speed of the node stage.

We assume that the time delay at each stage is identical for all three types of sorting networks, thus, the total running time will be determined by the number of stages required for the networks. The total number of

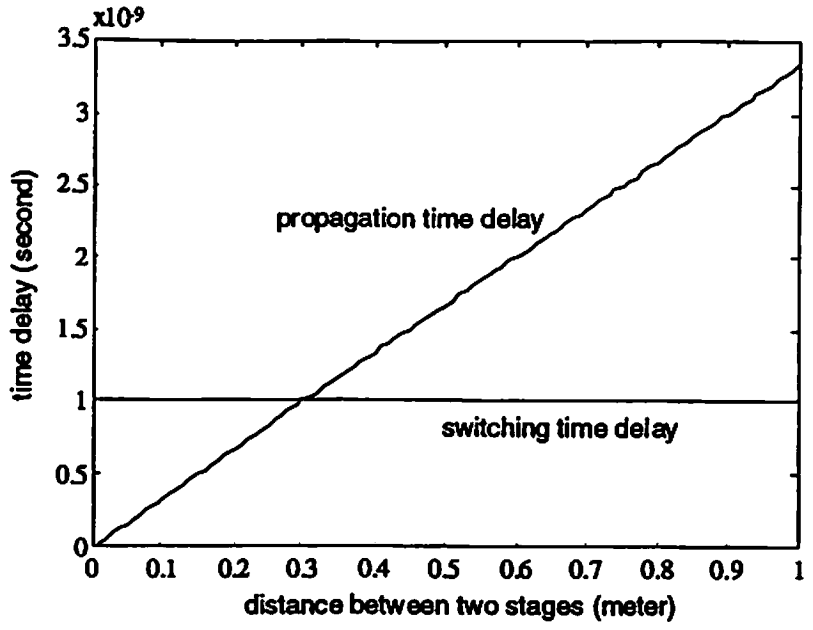



Figure 3.16: The comparison between the propagation time delay and the switching time delay based on different values of $d$.

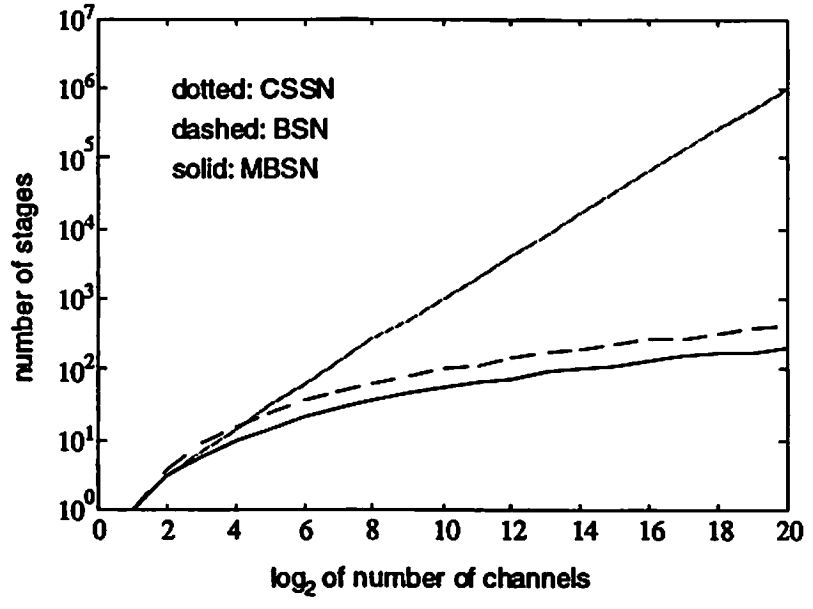



Figure 3.17: Total number of node stages versus $\log_2$ of the number of channels for three kinds of sorting networks. CSSN represents the cylindrical Spanke network; BSN represents the bitonic sorting network; MBSN represents the merged bitonic sorting network.

stages versus number of input channels for all three types of networks are shown in Fig. 3.17.

To sort $N$ input channels, the MBSN requires the lowest number of stages among all three types of sorting networks. When we double the number of input channels, the number of stages required increases by $\log N$ stages. In the same case, the number of stages for the CSSN is doubled. Therefore, the MBSN is less sensitive to an increase in the number of channels. It also has the fastest running time. For $N > 16$, the CSSN has the slowest running time.

### 3.5.2 Node Complexity

In this section, we evaluate all three types of sorting network by their node complexity. Table 3.2 lists their the number of modules, $F$, $S$ and the node complexity based on $N$ input channels.

| Sorting Networks | CSSN | BSN | MBSN |
|---|---|---|---|
| number of types of modules $F$ | 2 | 2 | 1 |
| number of node plane structures $S$ | 2 | $\log N$ | $\log N$ |
| number of modules | $\frac{1}{2}N(N-1)$ | $\frac{N}{2}(\log N)^2$ | $\frac{N}{4}(\log N+1)\log N$ |
| node complexity $C(N)$ | $2N(N-1)$ | $\frac{3N}{2}(\log N)^3$ | $\frac{N}{4}(\log N+1)(\log N)^2$ |

Table 3.2: The node complexity of all three types of sorting networks.

The total number of C&E modules and the node complexity $C(N)$ of all three types of sorting network can be calculated as a function of the number of input channels. The results are shown in Fig. 3.18 and Fig. 3.19.

Again, the MBSN has the lowest node complexity to sort $N$ input channels among all three types of sorting networks. An interesting result is that, when the size of the network is small ($N < 2^{10}$), the CSSN has lower node complexity than the BSN.

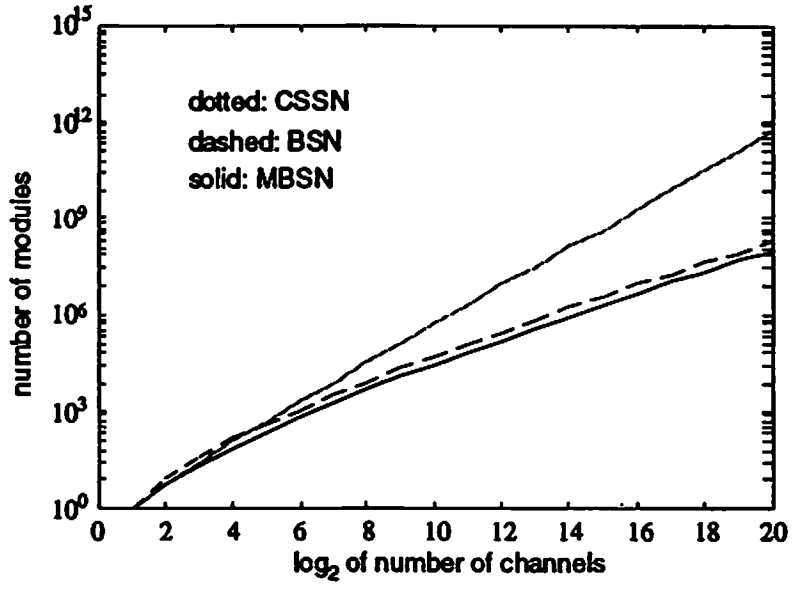



Figure 3.18: Total number of 2 input/ 2 output C&E modules versus $\log_2$ of the number of channels for three kinds of sorting networks. CSSN represents the cylindrical Spanke network; BSN represents the bitonic sorting network; MBSN represents the merged bitonic sorting network.

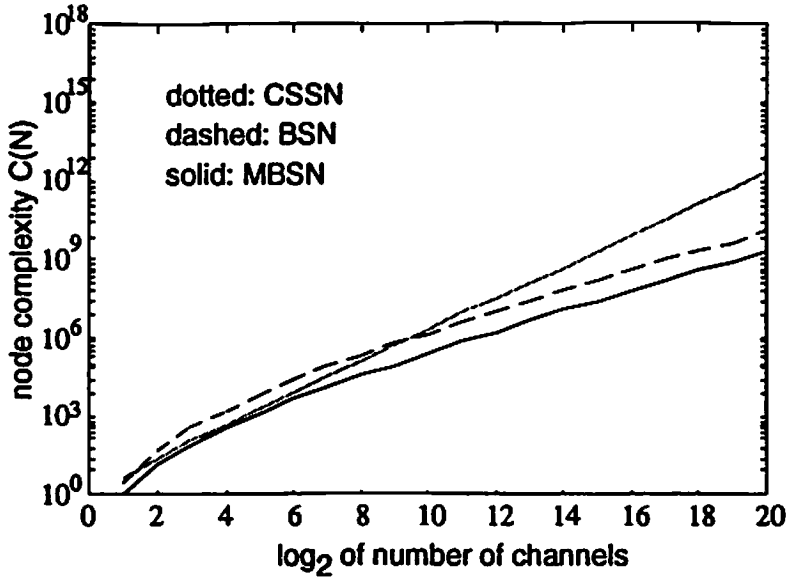



Figure 3.19: The node complexity versus $\log_2$ of the number of channels for three kinds of sorting networks. CSSN represents the cylindrical Spanke network; BSN represents the bitonic sorting network; MBSN represents the merged bitonic sorting network.

### 3.5.3 Optical Complexity

Following the discussion in the previous sections, the number of interconnection patterns, link complexity $L(N)$, and the optical complexity $O(N)$ of all three types of sorting networks based on $N$ input channels are listed in Table 3.3.

Figure 3.20 shows the optical complexity of all types of sorting networks for various values of $N$. It is obvious that the CSSN has the advantage of low optical complexity and the MBSN has worst performance.

| Sorting Networks | CSSN | BSN | MBSN |
|---|---|---|---|
| link complexity $L(N)$ | 1 | $N$ | $N$ |
| number of patterns | 1 | 1 | $\log N$ |
| number of stages | $N-1$ | $(\log N)^2$ | $\frac{1}{2}(\log N+1)\log N$ |
| optical complexity $O(N)$ | $N-1$ | $N(\log N)^2$ | $\frac{N}{2}(\log N+1)(\log N)^2$ |

Table 3.3: The optical complexity of all three types of sorting networks.

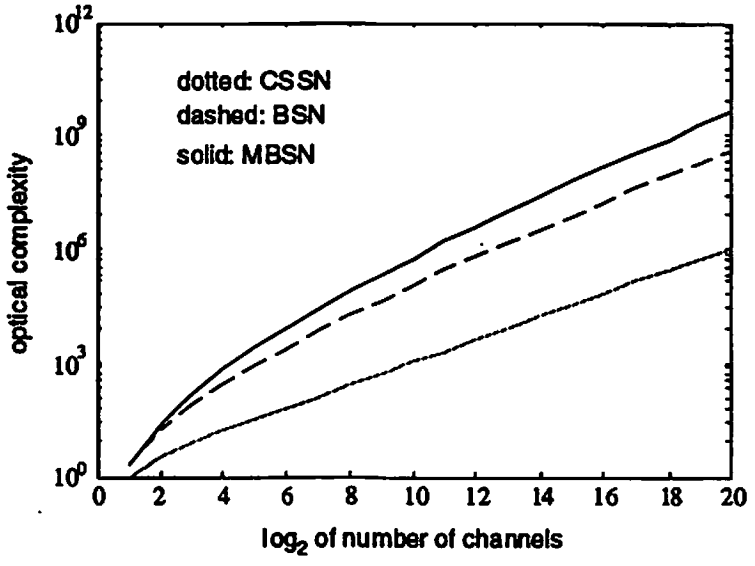



Figure 3.20: The optical complexity versus $\log_2$ of the number of channels for three kinds of sorting networks. CSSN represents the cylindrical Spanke network; BSN represents the bitonic sorting network; MBSN represents the merged bitonic sorting network.

## 3.6 Summary

In this chapter, we systematically study the implementations of multistage sorting networks using 3-D volume structures. An algorithm for expanding

the BSN in 3-D space is proposed and a mathematical proof is given. We find out that during the left-right merging procedures in the 3-D space, a 2-D column shuffle is required to transfer the shuffled row-major indexing of a 2-D signal array to a equivalent indexing performed using a 2-D folded shuffle. The optical interconnection system that performs the 2-D folded shuffle and 2-D column shuffle can be achieved using a hologram or computer-generated diffractive optics devices that will be discussed later. Based on this algorithm, A new type of sorting network, MBSN, which has promising characteristics, is developed and discussed. In the rest of the chapter, we evaluate the CSSN, BSN and MBSN utilizing three kinds of measuring criteria. The result is shown in Table 3.4.

| Sorting Networks | CSSN | BSN | MBSN |
|---|---|---|---|
| Running Time $T(N)$ | high | medium | low |
| Node Complexity $C(N)$ | high | medium | low |
| Optical Complexity $O(N)$ | low | medium | high |

Table 3.4: The evaluating results of all three types of sorting networks.

When the number of input channels is large, a CSSN has drawbacks such as long operating time delay and high switching hardware cost. However, its optical interconnection is very simple and easy to implement. On the other hand, a MBSN has advantages such as a higher sorting rate and low node complexity, but suffers from the problems of complex interconnection architectures in its link stages. System designers should compare the cost of optical elements and switching elements and consider

the complexity and reliability of the facilities needed to support both optical and electronic systems in order to select the most suitable network architecture to use.

# Chapter 4

# Linear Mapping Using Diffractive Microlenses

In this chapter we focus on optical system designs for one-to-one linear mapping using diffractive microlenses. First we study the characteristics of Gaussian microbeams in on-axis diffractive microlens arrays. Following that, we compare the advantages and disadvantages of conventional bulk lens systems and microlens systems. Optical system structures for short and long distance linear mapping interconnections are studied later on based on the previous analysis. Finally, optical packing limitations are calculated based on the Gaussian distribution characteristics of light sources, and the light efficiency of each system structure is given.

## 4.1 Characteristics of Gaussian Beams

We first define the characteristics of Gaussian beams. Let us consider a Gaussian microbeam transformed through a thin lens as shown in Fig. 4.1. The waist $w_0$ of a Gaussian beam source is its radius size at which its amplitude is reduced by $e^{-2}$ from its center peak value. The waist size $w(z)$ of a Gaussian beam in free space propagating over a distance $z$ from its source can be modeled in terms of the waist size $w_0$ of the source [61]

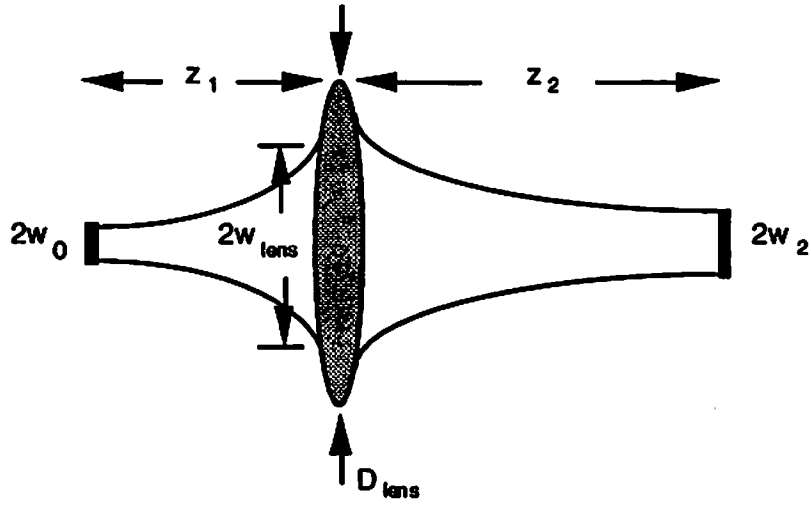



Figure 4.1: The characteristics of Gaussian beam propagating and focusing.

$$w(z)^2 = w_0^2(1+(\frac{z}{z_r})^2), \tag{4.1}$$

where $z_r$ is the scale length for this beam, given by

$$z_r = \frac{\pi w_0^2}{\lambda}, \tag{4.2}$$

and $\lambda$ is the wavelength.

Part of the Gaussian beam is clipped by the microlens as expressed by the clipping ratio, $k$,

$$k = \frac{D_{lens}}{2w_{lens}}, \tag{4.3}$$

where $w_{lens} = w(z_1)$ and $D_{lens}$ is the lens diameter as shown in Fig. 4.1. In Gaussian beam propagation theory, a value of $k = 1.5$ corresponds to a loss

of 1% of the incident power, and a value of $k = 2.12$ corresponds to 0.1% of the incident power [61]. Following the analysis in Ref. [47] we will let $k$ be a variable in the following and realize that $k$ must be $\geq 2$ in practice to avoid significant power loss due to beam truncation. The waist of the beam $w(z')$ at the right hand side of the lens can also be modeled in terms of its propagating distance $z'$ away from the lens as given by

$$\frac{\lambda}{\pi w^2(z')} = \frac{1/z_{r1}}{(1 - z'/f)^2 + (z'/z_{r1})^2}, \tag{4.4}$$

where $z_{r1} = \pi w_{lens}^2/\lambda$ [61]. If we choose $z'$ to be $z' = z_2 = f$, the spot size is given by

$$\begin{aligned} w(z') &= w_2 \\ &= \frac{\lambda f}{\pi w_{lens}}. \end{aligned} \tag{4.5}$$

We use Eq. (4.3) in Eq. (4.5). Hence, Eq. (4.5) becomes

$$\begin{aligned} w_2 &= \frac{\lambda f}{\pi D_{lens}/k} \\ &= \frac{k\lambda}{\pi} \times (F\#). \end{aligned} \tag{4.6}$$

The above equations express the characteristics of light propagation in a Gaussian light source based microlens system. We will use these results in the following sections.

## 4.2 Background

Linear mapping is one of the most simplest interconnection structures. It can be implemented by conventional bulk lenses (macrolenses) in $4f$ imaging systems shown in Fig. 4.2. Note that node stage $i$+1 is rotated 180° (inverted) to achieve a linear mapping operation. Fig. 4.2a is a telecentric imaging system [86], in which the ray is incident perpendicularly to the image plane, thus minimizing the lateral misalignment effects of image defocus. The telecentric imaging system shown in Fig. 4.2a also has lower aberrations than the imaging system shown in Fig. 4.2b because the beams in this system are collimated when they pass through elements such as polarizing beamsplitters, thus aberration accumulation is reduced. Another advantage of the system shown in Fig. 4.2a over the system in Fig. 4.2b is that, if we consider a divergent light source, the system in Fig. 4.2a has a shorter propagation distance between the light sources and lens, thus a smaller lens can be used to capture the beam generated from the light source located at the edge of the node stage $i$.

Macrolens systems have the advantage of high space-bandwidth product (SBWP), low sensitivity to changes in wavelength of the input light, and relatively large numerical apertures such that low aberrations can be expected. However, the relatively long focal lengths generally used (three or four times greater than the size of bulk lenses) result in increasing the physical length and time latency of a system. Here, we denote the vertical dimension of the node plane is $l_{node}$. We assume that each channel on the

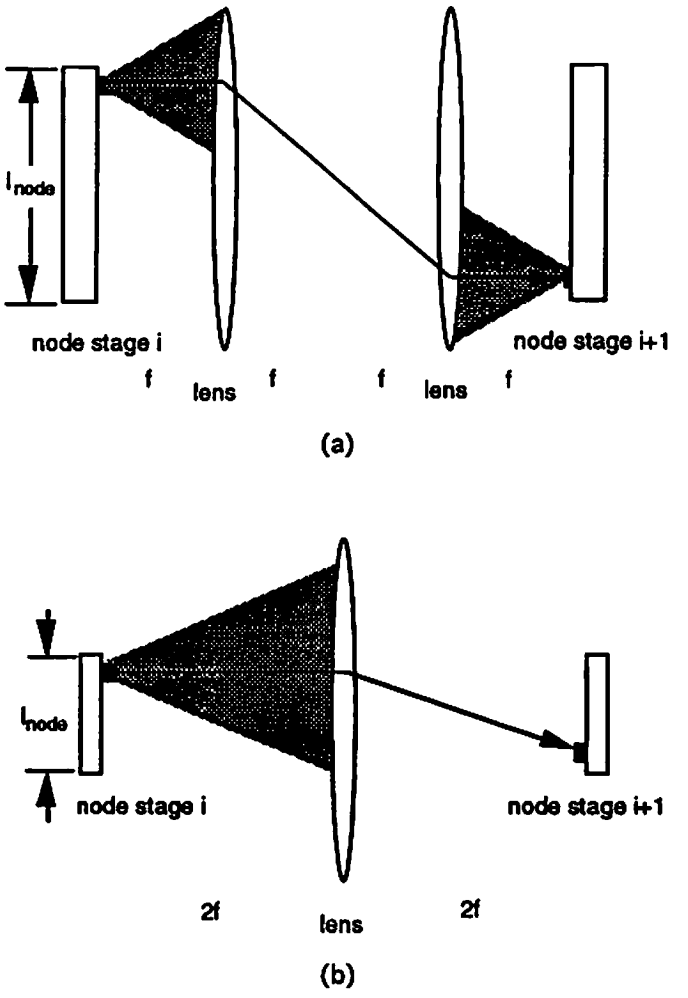



Figure 4.2: Different architectures of imaging systems for linear mapping interconnections having a total length of $4f$.

node stage has a light source with a $5\,\mu m \times 5\,\mu m$ active area and a Gaussian light distribution. One way to reduce the physical length of a macrolens system is to use lenses with low $f$-numbers. However, this approach increases the fabrication complexity and cost of the macrolenses.

Because light from the source in Fig. 4.2 is emitted with a diverging angle, the light wavefront spreads out in a cone before it reaches the lens.

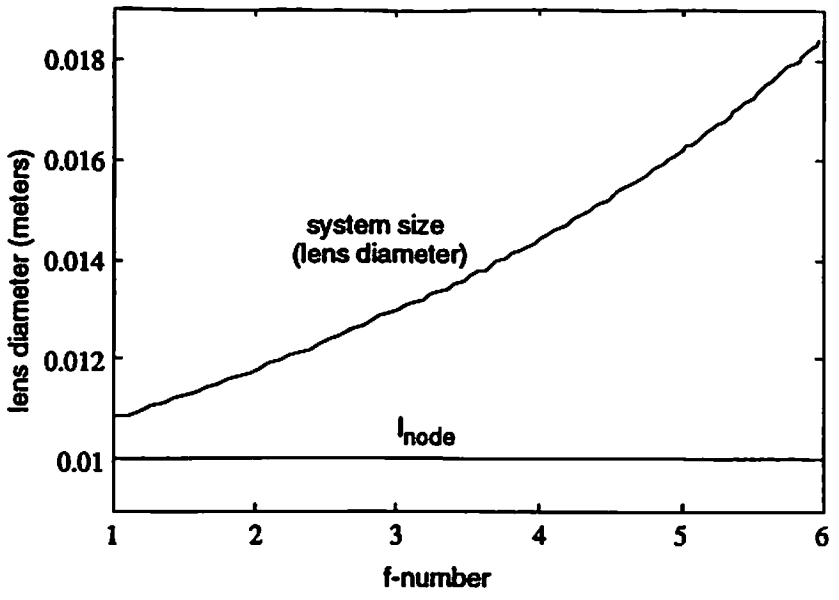



Figure 4.3: The lens diameter required to capture 99% of an incident Gaussian beam emitted by a light source with a $5\,\mu m \times 5\,\mu m$ active area and Gaussian distribution after propagation over a distance $f$. The horizontal axis is plotted as a function of $f$-number. Also shown is a typical vertical size ($10\,mm$) of a node stage ($l_{node}$) which will be used in later discussion.

Therefore, the size of the bulk lenses must be increased in order to maintain the optical efficiency. Figure 4.3 shows the required lens diameter needed to capture 99% of the incident Gaussian beam emitted by a Gaussian source of linear dimension $5\,\mu m$ (a waist size $2w_0 = 5\mu m$) after propagation over a distance $f$ as depicted in the interconnection system of Fig. 4.2a. It is obtained by using Eq. (4.1) for propagation over a distance $z = f$. Figure 4.3 is plotted as a function of the $f$-number of its lens, $F\#$, where $F\#$ is defined as the ratio of the focal length to the diameter of a lens.

A diffractive microlens array is another approach to implementing linear mapping interconnections in a sorting network. Each microlens has a

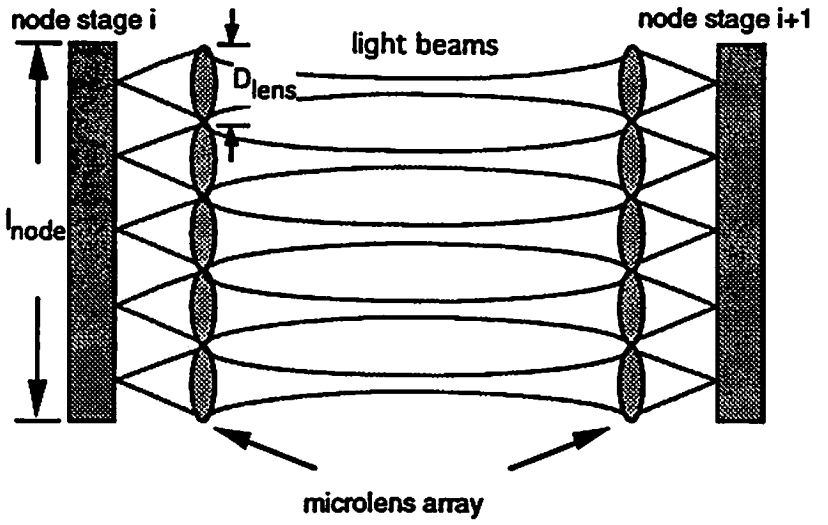



Figure 4.4: A linear mapping interconnection system using microlens arrays.

relatively small SBWP. In order to have self-routing capabilities, the switching nodes in a sorting network must integrated along with additional optoelectronic circuitry, thus the optical I/O density is significantly reduced. A linear mapping system architecture using microlens arrays is shown in Fig. 4.4. The advantage of diffractive microlens systems over conventional ones are its compact size and low cost of fabrication and replication. However, new micro-optomechanical techniques for aligning and mounting microlens arrays are required. In the following sections, we discuss diffractive microlens systems for short and long distance linear mappings based on light sources with a Gaussian intensity distribution.

## 4.3 Short Distance Interconnection

An interconnection system for short distance linear mappings can be realized using a single microlens array as shown in Fig. 4.5. A microlens is

put in front of the cell similar to the structure shown in Fig. 4.5, where $2w_0$ is the size of the light source at node stage $i$, $2w_2$ is the size of the detector at node stage $i+1$, $d$ is the free space propagation distance of the microbeam before it reaches the microlens array, and $z_2$ is the distance between the microlens array and detectors. We now specify the system parameters in order to analyze the system performance. A node plane with a size $l_{node} \times l_{node}$ where $l_{node} = 10\,mm$ is assumed. Each channel on the node plane has a light source which emits a Gaussian intensity microbeam with a waist diameter $5\,\mu m$ $(2w_0 = 5\,\mu m)$ and a wavelength $\lambda = 0.8\,\mu m$. We assume that the lens system has a clipping ratio $k = 2$ and that the detector size on the following node plane is $5\,\mu m \times 5\,\mu m$ $(2w_2 = 5\,\mu m)$. We also arrange the system in a way that $z_2$ is equal to the focal length of the microlens.

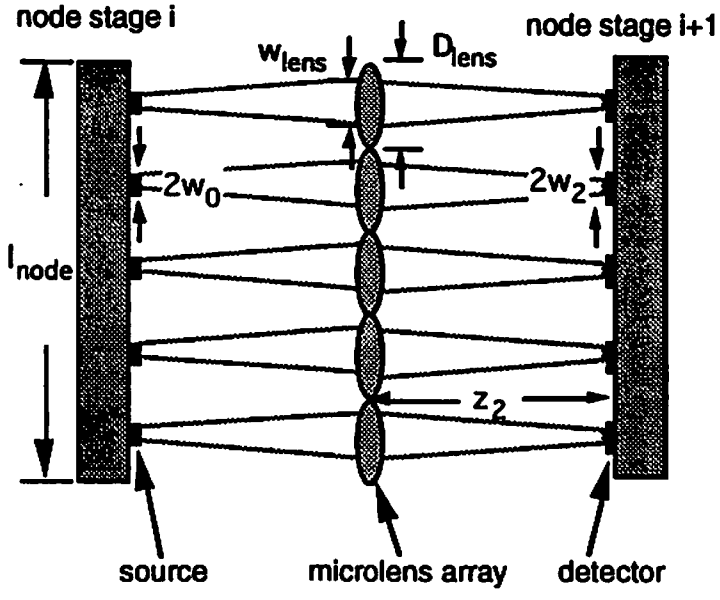



Figure 4.5: Short distance linear mapping system using single microlens array.

To begin, we first use Eq. (4.6) to calculate the upper limits of the *f*-number of the microlens. For $2w_2 = 5\,\mu m$ and $k \geq 2$, the result is that the *f*-number of the microlens needs to be less than 5. Later, in Chapter 5, we discuss multilevel discrete phase approximations to the continuous phase profile of refractive lenses. We now consider a multi-level microlens profile fabricated by $K$ masks. In Chapter 5, it is shown that the efficiency of the lens is proportional to $\mathrm{sinc}^2(1/L)$ where $L = 2^K$ is the number of grating levels [62]. According to this, a binary lens has an efficiency about 41%, and two or three masks can be used to increase the efficiency to 81% ($L = 4$) or 95% ($L = 8$). The local grating period $T_{lens}$ is given by [63]

$$T_{lens} = \lambda\sqrt{1+(f/r)^2}, \qquad (4.7)$$

where $f$ is the focal length and $r$ is the distance from the center optical axis. The minimum size gratings will always appear at the boundary of the lens, $r = D_{lens}/2$. Thus, Eq. (4.7) can be rewritten in terms of the lithographic resolution (or minimum feature size, $\delta$), number of profile levels, $L$, and *f*-number, $F\#$ as given by

$$\delta \times L = \lambda\sqrt{1+(2\times F\#)^2}. \qquad (4.8)$$

If a diffractive lens with $F\# = 5$ is chosen for the system and can be fabricated with a resolution of $\delta = 1\,\mu m$, from Eq. (4.8), an eight-level lens profile can be made. Because the *f*-number of the microlens in the optical system is fixed, the shorter the distance between two stages is, the smaller the size of each channel can be. The total number of channels $N$ is given by

$$N = (\frac{l_{node}}{D_{lens}})^2. \qquad (4.9)$$

Under the conditions that $F\# = 5$ and a node plane with $l_{node} = 10\,mm$, $D_{lens}$ can be calculated from Eq. (4.6). Figure 4.6 shows the results of the packing density and interconnection distance of the system using different focal length microlenses. The number of channels supported by this type of system design is maximized for small interconnection distances. However, the space occupied by the processing circuitry and the optical aberrations caused by the small lens aperture (proportional to the size of each channel) will give a upper limit on the packing density.

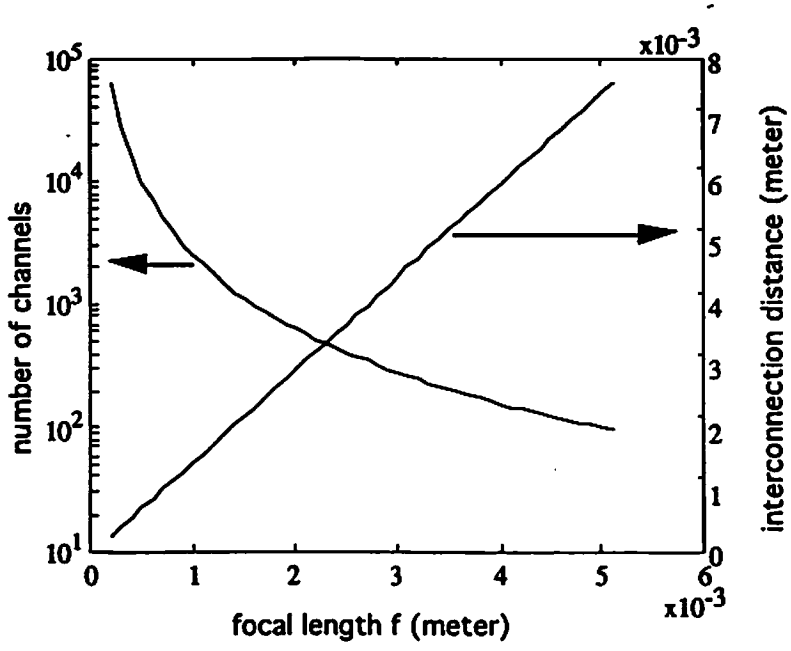



Figure 4.6: The number of channels (left vertical scale), and interconnection distance (right vertical scale) versus focal length $f$. Here we assume the size of node plane is $10\,mm \times 10\,mm$, $f$-number $F\# = 5$, $\lambda = 0.8\,\mu m$, and macrolens clipping ratio $k = 2$.

Let us consider the optical resolution due to the finite size of lenses in a diffraction limited system. We assume that the diameter of each channel is equivalent to the size of each microlens in our design. Light beams transmitted from the sources are focused at the back focal plane of the microlens array at the destination node stage $i$+1. Thus the intensity distribution at the focal plane is determined by the convolution of a Gaussian function (distribution of the light beam) $G(r)$ and an Airy function (Fourier transform of the circular aperture of the lens) as given by

$$I(r) = (G(r) * \frac{2J_1(\alpha r)}{\alpha r})^2, \tag{4.10}$$

where $I$ is the intensity, $J_1$ is the Bessel function of first kind and $r$ is the radius. Here, $\alpha$ is given by

$$\begin{aligned} \alpha &= \frac{\pi D_{lens}}{\lambda f} \\ &= \frac{\pi}{\lambda \times F\#}, \end{aligned} \tag{4.11}$$

where $\lambda = 0.8\,\mu m$ and $F\# = 5$.

The corresponding amplitude distributions of Gaussian and Airy functions at the detector plane are shown in Fig. 4.7. The magnitude squared (intensity) of these two functions after they are convolved is shown in Fig. 4.8. Without considering diffraction due to the lens pupil, a

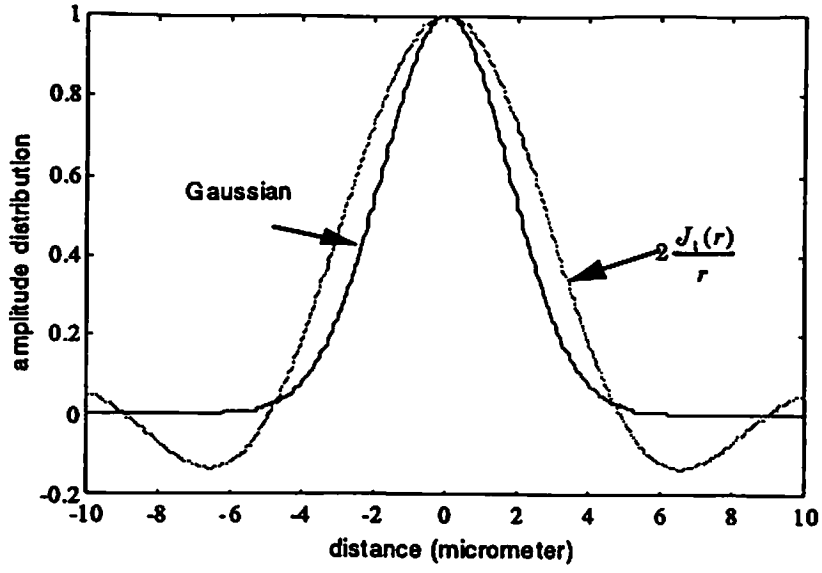



Figure 4.7: Gaussian distribution and lens diffraction curves on the detector plane. Here, $w_0 = 5\,\mu m$, $F\# = 5$, and $\lambda = 0.8\,\mu m$.

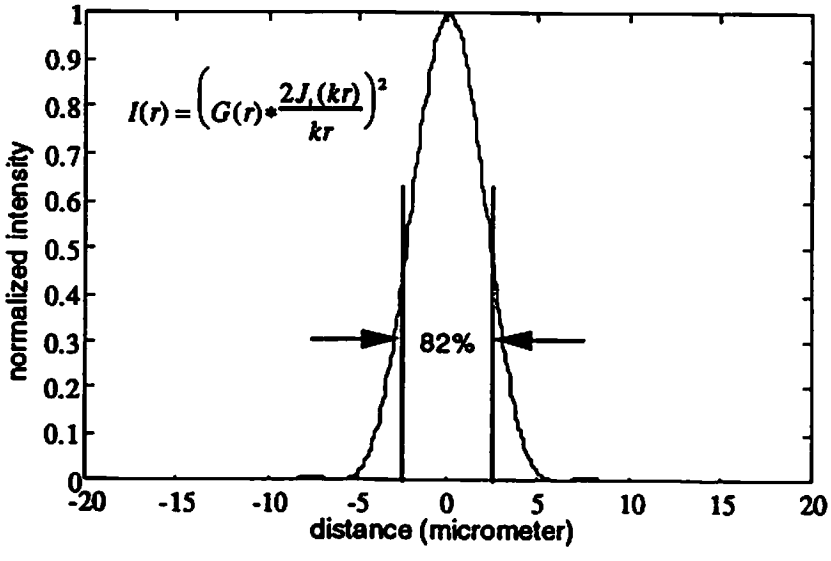



Figure 4.8: The intensity distribution of the convolution result. Here, $w_0 = 5\,\mu m$, $F\# = 5$, and $\lambda = 0.8\,\mu m$. The power focusing on the $5\,\mu m$ center areas is 82%.

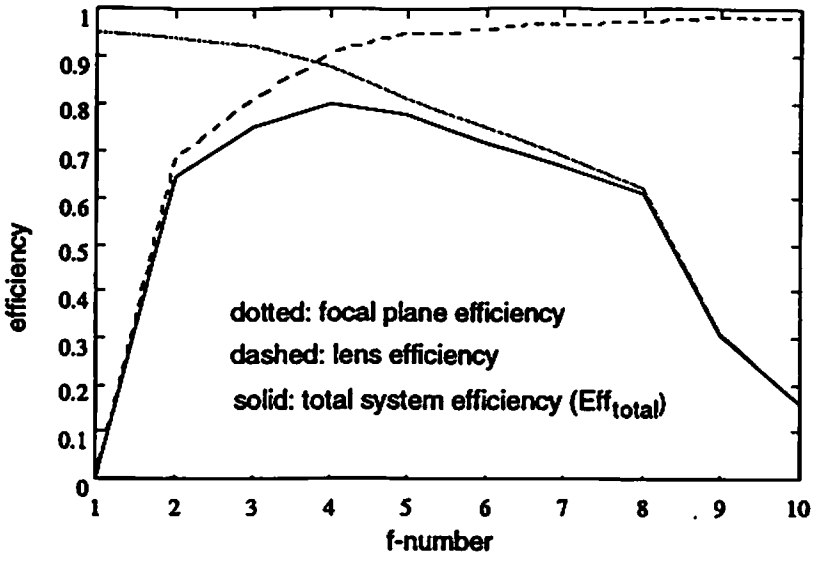



Figure 4.9: The variation of different efficiencies based on *f*-number. Here, $w_0 = 5\,\mu m$, and $\lambda = 0.8\,\mu m$.

Gaussian beam with a $5\,\mu m$ waist diameter has 95% of its power focused on the $5\,\mu m$ diameter detecting area. Diffraction due to the limited size of the lens pupil will further distribute the power of the beam and only 82% of the arriving power will focus on the detecting area. If the *f*-number of the lens is increased, it increases the width of the center lobe of the Airy function and thus the efficiency on the center detecting area is decreased. The total efficiency of the system is given by

$$Eff_{total} = (clipping\ eff.) \times (lens\ eff.) \times (focal - plane\ eff.). \quad (4.11)$$

Figure 4.9 shows the variation of the total system efficiency based on different values of *f*-number. In Fig. 4.9, we have set the clipping efficiency to one for simplicity. We then use Eq. (4.8) to obtain $L$ (the number of microlens profile levels) assuming $\delta = 1\,\mu m$ and $\lambda = 0.8\,\mu m$, and then use

the relation *lens efficiency* = $\mathrm{sinc}^2(1/L)$ to obtain the dashed curve. We Eq. (4.10) to obtain the focal plane efficiency which is plotted as a dotted line in Fig. 4.9. The maximum total system efficiency can be achieved using a lens with $F\# = 4$. From Eq. (4.8), this requires a six-level microlens which needs a special fabrication process.

## 4.4 Long Distance Interconnection

In Section 4.3, a single microlens array imaging system only supports a two-sided OEIC-based node plane structure with a interconnection length less than few hundred micrometers. However, other types of smart pixel devices such as self electro-optic effect devices (SEEDs) are reflection type devices whose sources (or modulators) and detectors are built on the one side of the node plane. Its multistage structure is achieved by arranging each node stage in a planar fashion as shown in Fig. 4.10. A imaging system with a pair of microlens arrays is used to transfer an array of light beams over a distance greater than the size of the node plane $l_{node}$ (10 $mm$).

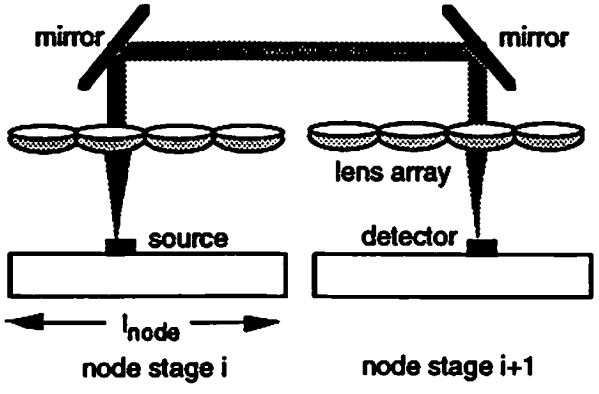
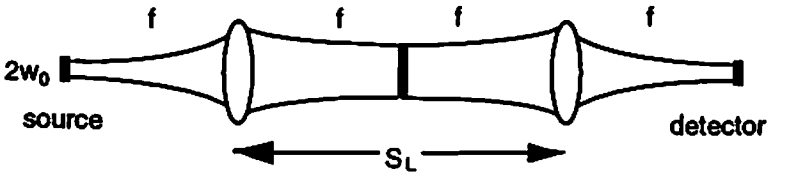
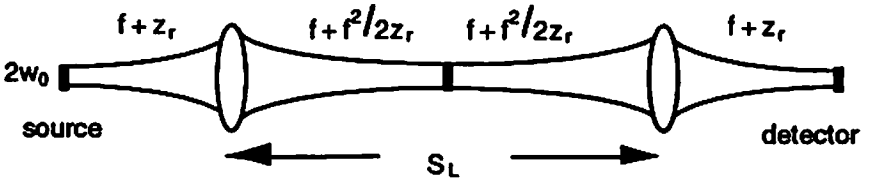



Figure 4.10: Interconnection structure for reflection type devices using an imaging system with a pair of microlens arrays.




(a) system I




(b) system II

Figure 4.11: Two types of long-distance linear mapping systems. Both image the source plane to the detector plane.

Two types of optical systems designed using a pair of microlens arrays with the same focal length have been considered to satisfy this constraint [33]. Their structures are shown in Fig. 4.11. In order to use the optical systems in the applications illustrated in Fig. 4.10, we would like to maximize the length $S_L$ between the microlenses to achieve a minimum required propagation distance. Thus,

$$S_L \geq l_{node}, \tag{4.13}$$

where $l_{node}$ is the size of the node plane.

Considering Gaussian beam illumination, the design space of the system has four parameters: the source waist, $w_0$, the distance between

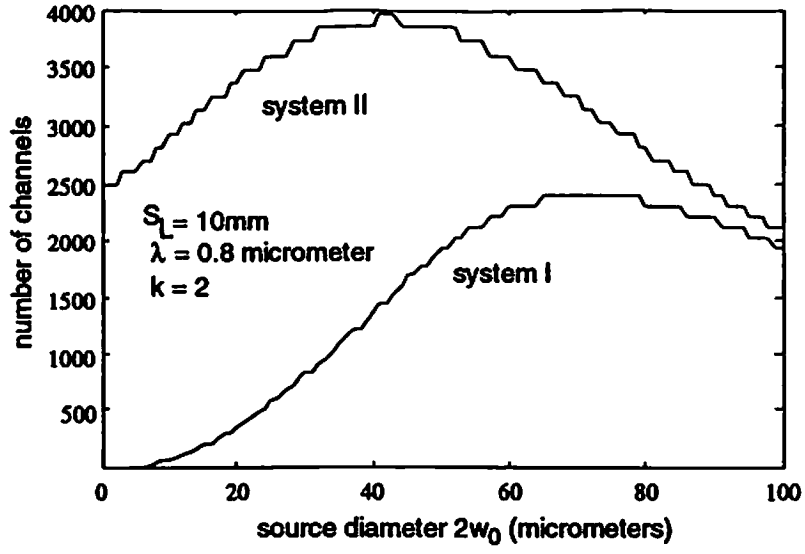



Figure 4.12: The maximum number of channels of two types of systems as a function of source diameter $2w_0$.

microlenses, $S_L$, the clipping ratio, $k$, and the wavelength of the source, $\lambda$. These parameters will affect the diameter and the $f$-number of the microlenses and thus change the total system efficiency and packing density. To simplify the analysis, we divide the constraints into several cases and $k = 2$ is assumed in all cases.

Case 1 : We assume that $S_L = l_{node}$ and $\lambda = 0.8\,\mu m$. The focal length $f$ and $z_r$ can be calculated from the system structure shown in Fig. 4.11. Using Eq. (4.1) and the clipping ratio $k$, the diameter of the microlens $D_{lens}$ can be calculated. Therefore, the number of channels for each system can be calculated from Eq. (4.9). Figure 4.12 shows the maximum number of channels of the two types of imaging system as a function of the source size $2w_0$. From the plot, System I requires a fairly large source diameter ($\approx 60\,\mu m$) to support a large

number of channels, and the maximum number of channels it supports decreases rapidly as the source size gets smaller. System II can support more channels than System I regardless of the size of the source. Using our previous assumption that the diameter of the sources is around 5 $\mu m$, We will not consider System I further.

**Case 2:** We assume $\lambda = 0.8\,\mu m$, source diameter $2w_0 = 5\,\mu m$, and that System II is selected. In Fig. 4.11a, if we know the length of $S_L$, the value of $f$ and $z_r$ can be calculated. Using Eq. (4.1), (4.3), and (4.9), the maximum number of channels System II can support and the total length of the system is plotted as a function of the length $S_L$ in Fig. 4.13. The maximum number of channels drops rapidly as the required length $S_L$ increases. To support more than $10^3$ channels, the interconnection length of System II should be less than 24 $mm$.

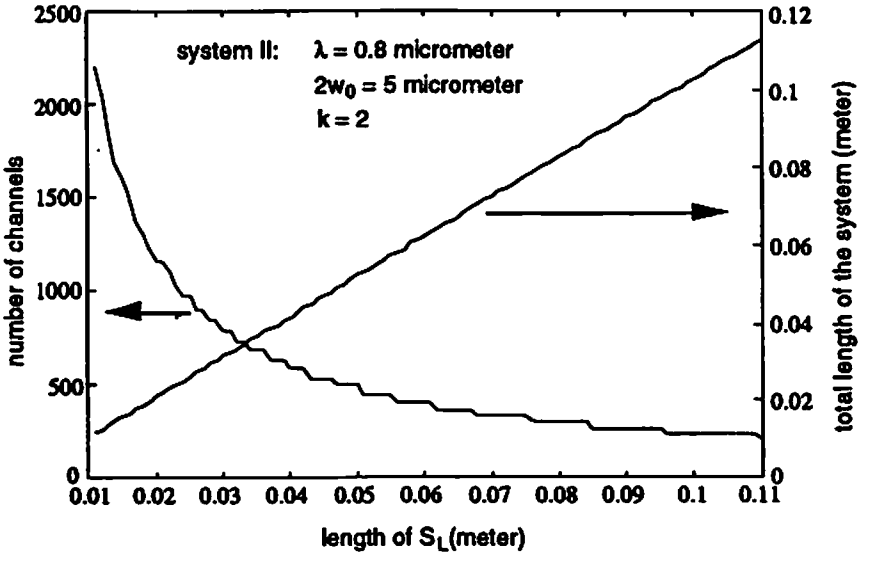



Figure 4.13: The maximum number of channels and the total length of the system is plotted as a function of the length $S_L$.

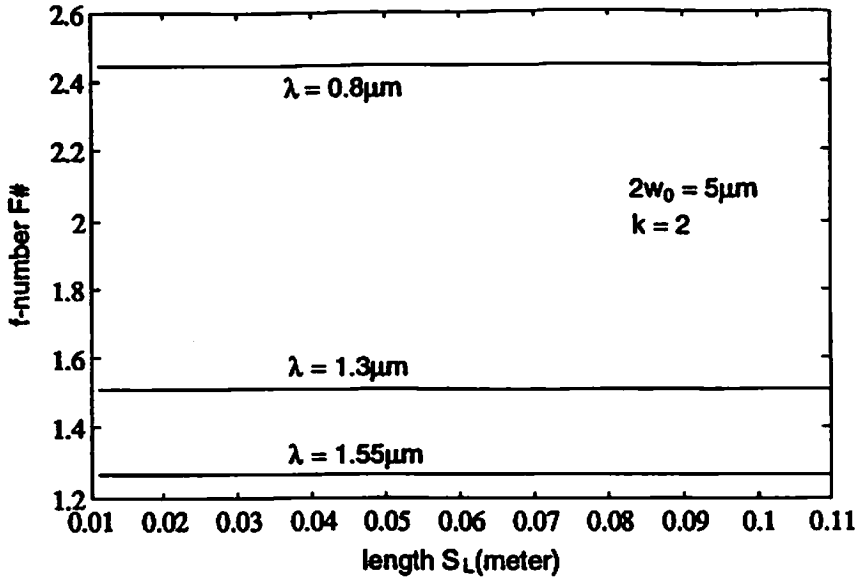



Figure 4.14: The required *f*-number of the microlens as a function of different wavelengths $\lambda$ and length $S_L$.

**Case 3:** We assume that source diameter $2w_0 = 5\,\mu m$ and that System II is selected. The *f*-number of the microlens is calculated based on different values of $S_L$ and wavelength $\lambda$. Three particular wavelengths are considered here: $0.8\,\mu m$, $1.3\,\mu m$, and $1.55\,\mu m$. Lasers with a $0.8\,\mu m$ wavelength are made in large quantities for applications such as compact disk players. The remaining two wavelengths are common in long haul fiber communication links where $1.3\,\mu m$ is the minimum loss wavelength for silica fiber, and $1.55\,\mu m$ is the minimum dispersion wavelength for silica fiber. As previous case, in Fig. 4.11a, if we know the length of $S_L$, the value of $f$ and $z_r$ can be calculated. Using Eq. (4.1) and (4.3), we know the diameter of the microlens ($D_{lens}$), thus, the *f*-number of the microlens can be calculated and plotted as a function of $S_L$ in Fig. 4.14. The *f*-

number of the microlens remains unchanged in the region of $S_L > 10\,mm$, and longer wavelength require a lower $f$-number microlens. However, because we fix the size of the detector $w_2 = 5\,\mu m$, from Eq. (4.6) we have $\lambda \times F\#$ as a constant. Thus, based on Eq. (4.8) and a minimum fabrication feature size $\delta = 1\,\mu m$, a four-level microlens can always be used in a long-distance linear mapping system using these three wavelengths.

## 4.5 Summary

In this chapter, linear mapping interconnections using microlens arrays are described. We analyze a short-distance single microlens array system and two long-distance (System I and System II) designs for different implementations. Under the assumptions of small source diameter, System II is more suitable for our design. We find that the short-distance design is feasible for distances less than $1.5\,mm$ and can support $10^3$ channels. On the other hand, the long-distance design (System II) has an interconnection range

$$10\,mm < interconnction\ range < 24\,mm \tag{4.14}$$

provided that more than $10^3$ channels can be supported. The light efficiency of two different systems are given by

$$Eff_{short} = (lens\ eff.) \times (focal - plane\ eff.), \tag{4.15}$$

$$Eff_{long} = (lens\ eff.)^2 \times (focal - plane\ eff.), \tag{4.16}$$

where we neglect the clipping effects. In Eq. (4.16), the efficiency of a long distance imaging system is proportional to $(lens\ eff.)^2$ because lights emitted from sources need to pass through a pair of microlens array before they reach to the detectors. The characteristics of short-distance and long-distance system designs are summarized in Table 4.1.

| | interconnection range $L_i$ | system efficiency |
|---|---|---|
| short-distance structure | $L_i < 1.5\ mm$ | 80% |
| long-distance structure (System II) | $10\ mm < L_i < 24\ mm$ | 61% |

Table 4.1: Summary of short-distance and long-distance (System II) linear mapping systems using microlens arrays. It is assumed that a minimum number of $10^3$ channels can be supported.

# Chapter 5

# 2-D Perfect Shuffles Using Diffractive Optical Elements

When light illuminates a diffraction grating, the light wavefront is diffracted into a set of waves that propagate in various directions. For a given wavelength and incident angle, a large fraction of the light can be directed to locations which are determined by the duty cycle and period of the grating. In this chapter we explore various architectures for optical interconnections systems for 2-D folded shuffles using diffractive optical elements. The basic concepts of the shifting properties of a 2-D folded shuffle and the design of a phase-only blazed grating are first discussed. Following that, different design methods for 2-D perfect shuffle interconnections are proposed. Finally, the optical packing limitations and system efficiency of each design method are calculated based on the light sources having a Gaussian distribution .

## 5.1 Background

The 2-D folded shuffle implements the 1-D perfect shuffle by folding the 1-D input array into a 2-D array. This arrangement has an inherent advantage in completely using the 2-D space-bandwidth product of the

optical system. In this section, we first describe the general reflected-symmetric property of the 2-D folded shuffle, which will reduce the complexity of space-variant gratings needed to implement it. Finally, the phase profile and multi-level mask design of a blazed grating capable of routing light beams to their destinations will be given.

### 5.1.1 Shifting Properties of 2-D Folded Shuffles

Let there be $N = 2^n$ signal channels arranged in a 2-D plane, where $n$ is even. Using the notation defined in the previous chapter, we represent this 2-D plane as a $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ matrix $A$ denoted by $A_{l,m} = [a(l,m)]$, where $l$ is the row index and $m$ is the column index of this matrix. These two indices can be represented as 2's complement binary numbers, i.e., $l = (x_{n-1}x_{n-2}\ldots x_{\frac{n}{2}})_2$ and $m = (x_{\frac{n}{2}-1}x_{\frac{n}{2}-2}\ldots x_1x_0)_2$, where $x \in \{0,1\}$. Since $n$ is an even number, the matrix $A_{l,m}$ can be further partitioned into four equal sized submatrices (of size $2^{\frac{n}{2}-1} \times 2^{\frac{n}{2}-1}$), denoted as $\beta_{l,m}^{p,q}$.

$$\beta_{l,m}^{p,q} = [a(px_{n-2}x_{n-3}\ldots x_{\frac{n}{2}}, qx_{\frac{n}{2}-2}\ldots x_1x_0)], \tag{5.1}$$

where $p$ and $q$ are 0 or 1. Thus,

$$A_{l,m} = \begin{bmatrix} \beta_{l,m}^{0,0} & \beta_{l,m}^{0,1} \\ \beta_{l,m}^{1,0} & \beta_{l,m}^{1,1} \end{bmatrix}. \tag{5.2}$$

Considering the row and column indices of $a(l,m)$ as one binary number, the 2-D folded shuffle operation $\sigma_{2f}(\bullet)$ is given by

$$\sigma_{2f}(A_{l,m}) = \sigma_{2f}([a(px_{n-2}x_{n-3}\ldots x_{\frac{n}{2}}, qx_{\frac{n}{2}-2}\ldots x_1x_0)])$$
$$= [a(x_{n-2}x_{n-3}\ldots x_{\frac{n}{2}}q, x_{\frac{n}{2}-2}\ldots x_1x_0p)]. \qquad (5.3)$$

The position of the entry $a(l,m)$ shifted by the 2-D folded shuffle operation is proportional to the decimal value of the difference between new and old binary indices. Let $(i, j)$ denote the decimal representation of the matrix indices $(l, m)$. We have

$$\sigma_{2f}(a(i,j)) = [a((2\times(i - p\times 2^{\frac{n}{2}-1}) + q)_{\text{modulo } 2^{\frac{n}{2}}},$$
$$(2\times(j - q\times 2^{\frac{n}{2}-1}) + p)_{\text{modulo } 2^{\frac{n}{2}}})]. \qquad (4.18)$$

An increase or decrease of the index $i$ implies a shift in the positive (down) or negative (up) $i$ index direction, and similar changes in index $j$ implies a shift to the right or left. We summarize the shift properties of four quadrants (submatrices) as follows.

**Quadrant I:** $(p,q) = (0,0)$. After the 2-D folded shuffle operation, each entry $a(i,j)$ in this quadrant shifts down and right to $a(2i,2j)$.

**Quadrant II:** $(p,q) = (0,1)$. After the 2-D folded shuffle operation, each entry $a(i,j)$ in this quadrant shifts down and left to $a(2i+1, 2j-2^{\frac{n}{2}})$.

**Quadrant III:** $(p,q) = (1,0)$. After the 2-D folded shuffle operation, each entry $a(i,j)$ in this quadrant shifts up and right to $a(2i-2^{\frac{n}{2}}, 2j+1)$.

**Quadrant IV:** $(p,q) = (1,1)$. After the 2-D folded shuffle operation, each entry $a(i,j)$ in this quadrant shifts up and left to $a(2i - 2^{\frac{n}{2}} + 1, 2j - 2^{\frac{n}{2}} + 1)$.

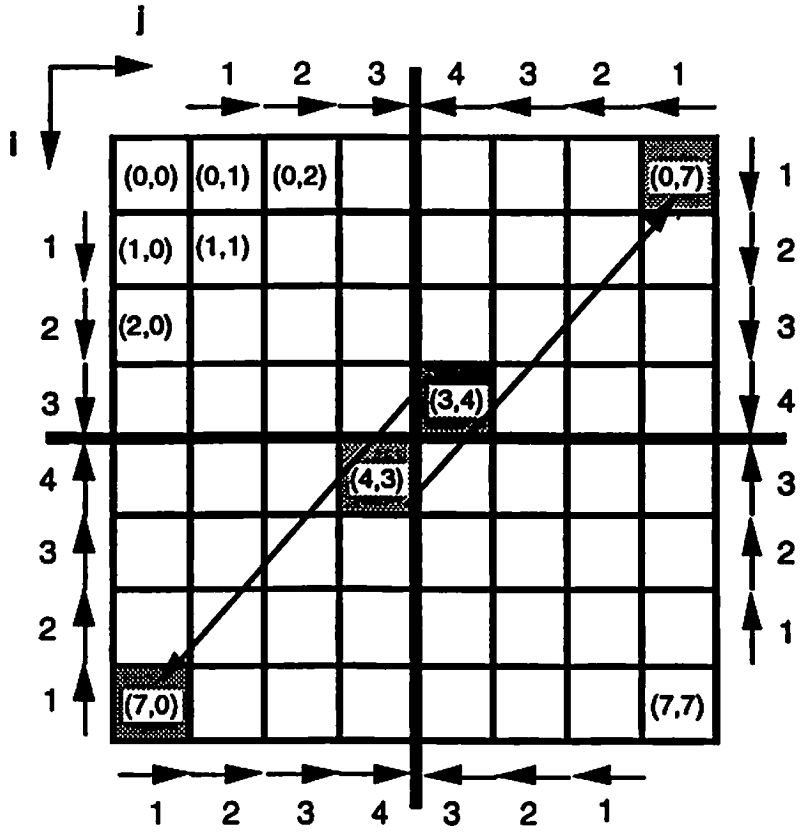



Figure 5.1: The shift property of the 2-D folded shuffle in a plane with 64 channels. The arrows around each quadrant represent the shift directions of the rows or columns in that quadrant and the number on the arrows indicates the distance of the shift.

We illustrate the case of $N = 64 = 2^6$ in Fig. 5.1. The indices shown in the matrix of Fig. 5.1 are decimal values of $i$ and $j$. The small arrows around each quadrant represent the shift directions of rows or columns in

that quadrant and the numbers on the small arrows indicate the distance of the shift. For example, the cell with indices (1,1) in quadrant I will be shifted one unit right and one unit down after the 2-D folded shuffle operation is applied. The new indices will be (2,2). The longest shift occurs at both $(2^{\frac{n}{2}-1}, 2^{\frac{n}{2}-1}-1)$ and $(2^{\frac{n}{2}-1}-1, 2^{\frac{n}{2}-1})$ locations, that is, (3,4) and (4,3). The shift distance is $\sqrt{2}\times 2^{\frac{n}{2}}\times$(cell size) at both locations.

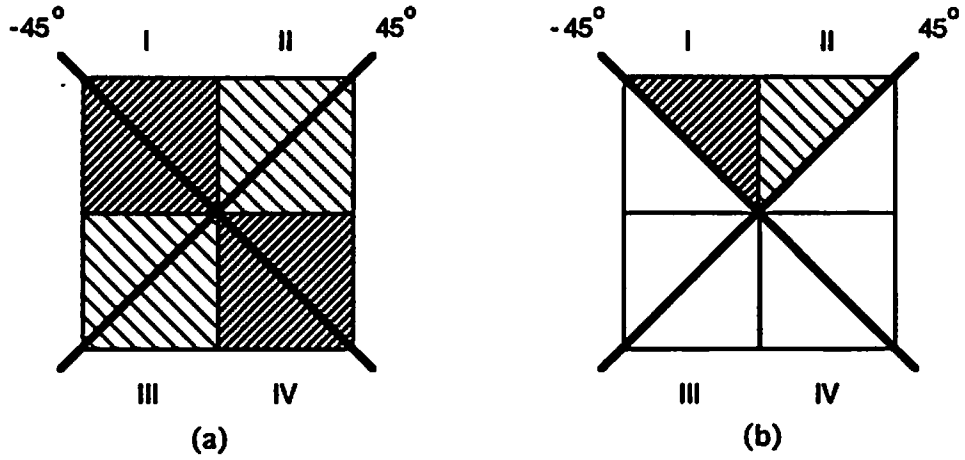



Figure 5.2: (a) Reflected-symmetric property of 2-D folded shuffle. (b) Basic pattern to be duplicated.

Although the 2-D folded shuffle is a space-variant operation, that is, each element in the matrix has a different shift pattern, it still has some symmetry properties. If we study the index changes in four quadrants, the shift pattern is reflected-symmetric along the ±45° lines shown in Fig. 5.2a. The shift pattern in the area shown in Fig. 5.2b can be calculated first, and this pattern is folded along the adjacent lines to generate the pattern of the entire area. The reflected-symmetric partial-invariant property of the 2-D folded shuffle operation studied here is important to the

design of computer-generated gratings because it reduces the complexity of the problem by a factor of 4.

### 5.1.2 Phase-Only Blazed Grating

The primary requirement of interconnection elements is to deflect light to different positions. A blazed grating discussed in [62] can be used to route incident light according to its grating period and the wavelength of the incident light. In order to describe the principle of a blazed grating, we first consider the refracting property of a bulk prism. A normally incident plane wave can be refracted an angle $\theta_p$ by a bulk prism as shown in Fig. 5.3.

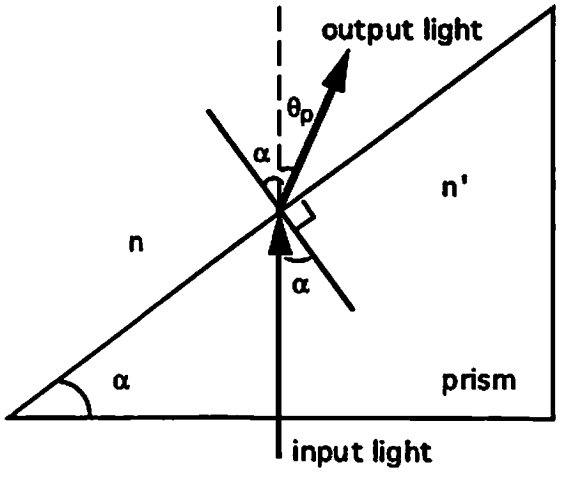



Figure 5.3: Normally incident plane wave refracted by a bulk prism.

Assuming that a prism has an angle $\alpha$ and its index of refraction is $n$, by Snell's law, we have

$$n\sin(\alpha) = \sin(\alpha + \theta_p), \tag{5.5}$$

that is

$$\theta_p = \sin^{-1}(n\sin(\alpha)) - \alpha. \tag{5.6}$$

Thus, a normally incident plane wave is deflected at an angle $\theta_p$, and the exiting wavefront is a plane wave given by

$$U_0(x) = e^{i\frac{2\pi}{\lambda}\sin(\theta_p)x} \tag{5.7}$$

For small prism angle $\alpha$, $\sin(\alpha) \approx \alpha$. We have

$$\theta_p \approx \sin^{-1}(n\alpha) - \alpha \approx (n-1)\alpha, \tag{5.8}$$

$$\tan(\theta_p) \approx \theta_p = \frac{d}{T}, \tag{5.9}$$

and

$$U_0(x) \approx e^{i\frac{2\pi}{\lambda}(n-1)\alpha x}, \tag{5.10}$$

where $T$ is the grating period, $d$ is the thickness of the prism, and $\lambda$ is the wavelength of the incident light.

Consider a 1-D phase-only blazed grating pattern formed by periodic microprisms located along the axis $x$ as shown in Fig. 5.4. When a normally incident plane wave with unit amplitude illuminates the grating, a set of waves with different diffraction order will be generated. The grating equation of the diffraction angle $\theta_d$ is given by

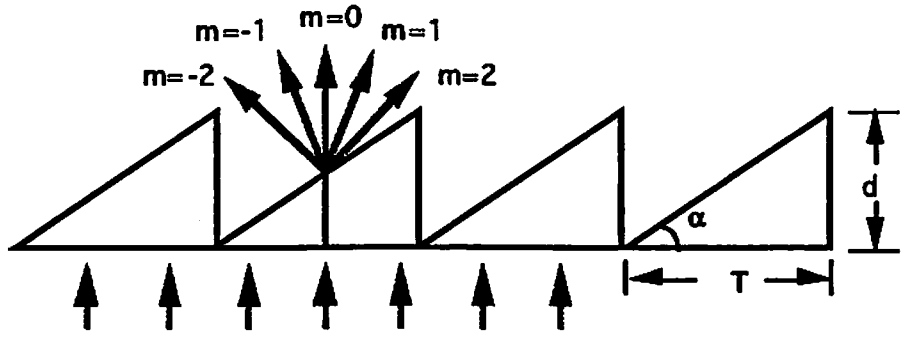



**Figure 5.4: Phase-only blazed grating made by microprisms. Normally incident plane wave is diffracted to a set of beams with different angle.**

$$\sin(\theta_d) = \frac{m\lambda}{T}, \tag{5.11}$$

where $m$ is the diffraction order. For small $\alpha$, the transmittance function $t(x)$ of this blazed grating can be represented by

$$t(x) = \sum_m \delta(x - mT) * \mathrm{rect}(\frac{x}{T}) e^{i2\pi f_0 x}, \tag{5.12}$$

where

$$f_0 = \frac{(n-1)}{\lambda}\alpha = \frac{(n-1)d}{\lambda}. \tag{5.13}$$

If the blazed grating is put in front of a lens, at the back focal plane, we have the Fourier transform of $t(x)$

$$\Im(t(x)) = \sum_m \delta(f - \frac{m}{T}) \frac{\sin(\pi T(f_0 - f))}{\pi T(f_0 - f)} \Bigg|_{f=\frac{x}{\lambda}}. \tag{5.14}$$

The amplitude of the $m^{th}$ diffraction order $a_m$ is given by

$$a_m = \frac{\sin(\pi T(f_0 - \frac{m}{T}))}{\pi T(f_0 - \frac{m}{T})}. \tag{5.15}$$

The diffraction efficiency $\eta_m$ of the $m^{th}$ diffraction order is the magnitude squared of the amplitude $a_m$, that is,

$$\eta_m = |a_m|^2 = \left[ \frac{\sin(\pi T(f_0 - \frac{m}{T}))}{\pi T(f_0 - \frac{m}{T})} \right]^2. \tag{5.16}$$

If we want the incident light to be routed only to the direction of its first diffraction order, that is,

$$a_m = \begin{cases} 1 & m = 1 \\ 0 & m \neq 1 \end{cases}. \tag{5.17}$$

we have first order diffraction efficiency $\eta_1$ as follows

$$\eta_1 = \left[ \frac{\sin(\pi(\frac{n-1}{\lambda} d - 1))}{\pi(\frac{n-1}{\lambda} d - 1)} \right]^2. \tag{5.18}$$

If we want $\eta_1 = 1$, the depth $d$ of the blazed grating must satisfy the condition

$$d = \frac{\lambda}{n-1}. \tag{5.19}$$

From Eq. 5.11, if $\theta_d$ is small, we have

$$\sin(\theta_d) \approx \theta_d = \frac{\lambda}{T}. \tag{5.20}$$

From the equations above, we notice that the light efficiency of the first diffraction order is wavelength and depth dependent. The deflection angle of the first order also depends on the wavelength of the incident light.

### 5.1.3 Binary Mask Design

Microlithographic technology is well developed, and it is possible to fabricate the surface relief profile of a multi-level grating array with high accuracy and quality [38]. The multi-level structure of the blazed grating is shown in Fig. 5.5. The fabrication process typically involves $K$ masking and etching steps to generate $L = 2^K$ phase levels. For example, to fabricate a four-level blazed grating requires two binary masks. The fabrication procedure is shown in Fig. 5.6. The binary pattern of point $R$ on $K^{th}$ mask is given by

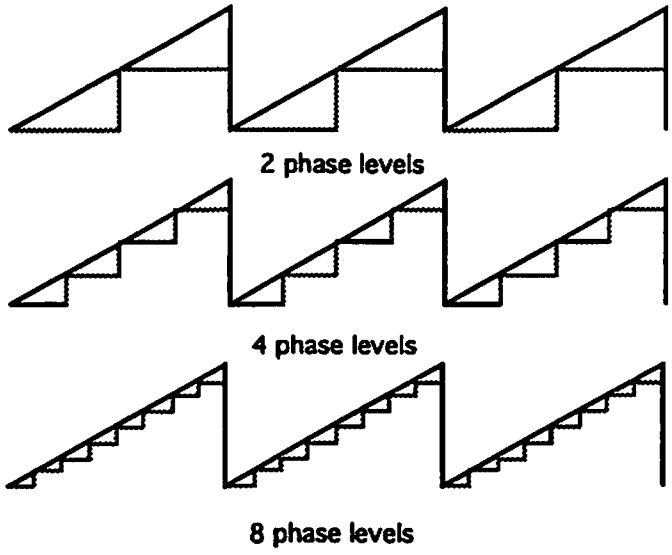



Figure 5.5: Multi-level structures of blazed gratings.

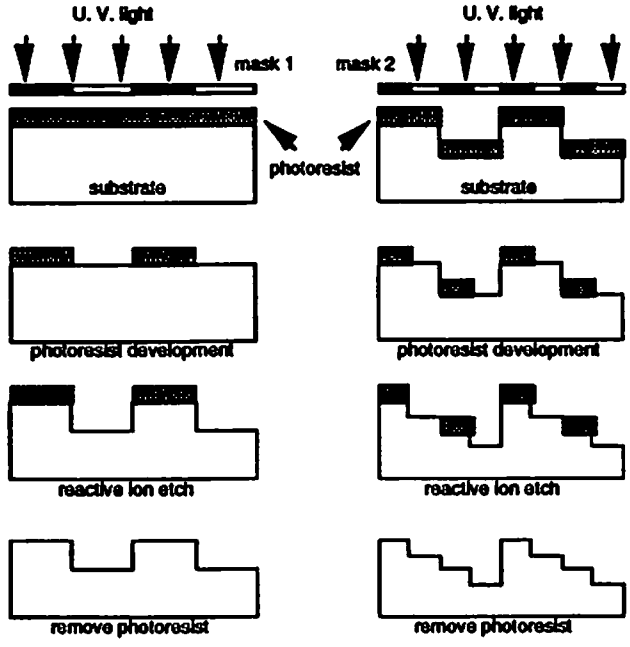



Figure 5.6: The fabrication procedure of a four-level grating pattern utilizing two binary masks.

$$M_K(R) = Quotient(\phi'(R), \frac{\pi}{2^{(K-1)}})_{modulo\,2}, \tag{5.21}$$

where $K$ is the mask number, $Quotient(x,y)$ represents the quotient of the division $x \div y$. If $M_K(R)$ equals zero, the mask is transparent (or white) at point $R$, otherwise, it is opaque (or black). The etch depth $d_K$ of the $K^{th}$ mask is given by

$$d_K = \frac{\lambda}{2^K(n-1)}. \tag{5.22}$$

The first order diffraction efficiency $\eta_1$ of multi-level phase grating $\phi_g$ corresponding to the number of levels $2^K$ is shown in Table 5.1.

| number of phase levels $L = 2^K$ | first order efficiency $\eta_1$ |
|---|---|
| 2 | 0.41 |
| 4 | 0.81 |
| 8 | 0.95 |
| 16 | 0.99 |

Table 5.1: The first order diffraction efficiency $\eta_1$ of multi-level phase grating $\phi_g$ is calculated for a different number of phase levels.

## 5.2 Space-Variant DOE Designs

In this section, 2-D folded shuffle interconnections are implemented by two types of space-variant DOE design methods based on off-axis microlenses

and blazed gratings. The design method based on blazed gratings can be further classified into two different approaches. In the following section, we focus on a system structure with cascaded stages. Each stage has $m \times m = N$ cells of channels and each channel has size $D \times D$. The total size of the chip is fixed at $l_{node} \times l_{node}$, where $l_{node} = mD = 10\,mm$. The shuffled pattern is obtained at the back focal plane of the optical setup. Here, Gaussian beam illumination with waist radius $w_0 = 5\,\mu m$ and wavelength $\lambda = 0.8\,\mu m$ is assumed through all design procedures.

### 5.2.1 Blazed Gratings

2-D folded shuffles can be performed utilizing space-variant DOEs. Each facet of the DOE can be considered a combination of a phase-only blazed grating for light shifting and a lens for Fourier transform operation. In this section, we implement the 2-D folded shuffle operations using two different approaches. First we use space-variant blazed gratings along with a microlens array in the first method to perform the operation. Following that, a new concept is proposed in the second method to modify the performance.

#### 5.2.1.1 First Design Method

**System design:** The 1-D system structure of the first method is shown in Fig. 5.7a. Each light source in node stage $i$ generates a Gaussian beam with a waist radius $w_0$. The beam propagates through a distance $z_1$ and impinges on the blazed grating with a waist radius $w_1$. The size of each cell is

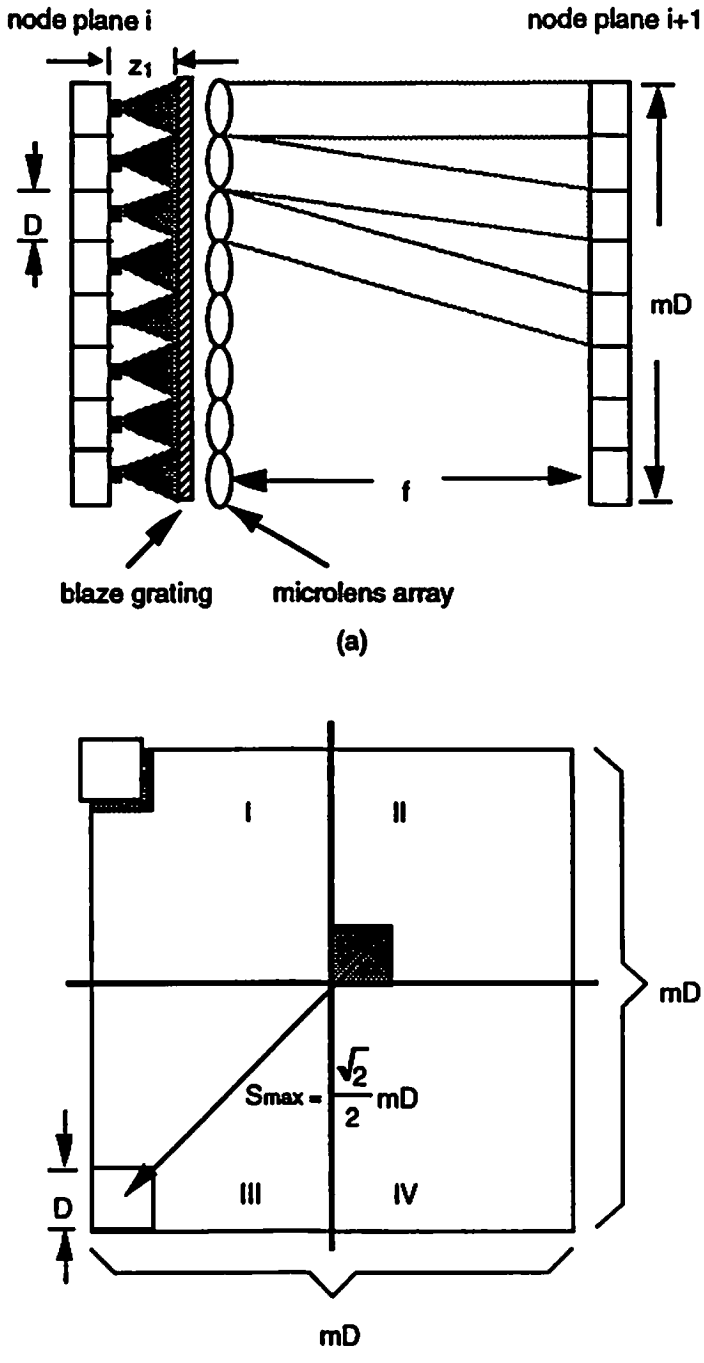



Figure 5.7: Interconnection system utilizing blazed gratings and microlens arrays. (a) 1-D system structure; (b) gray and white boxes represent the location of cells before and after a 2-D folded shuffle operation, respectively.

$D$, where $D = 2kw_0$, and $k$ is the clipping ratio. The number of channels $N$ is given by $N = m \times m$. Again, to avoid truncation and crosstalk, we assume that $k = 2$. At the back focal plane of the microlens, the resulting point spread function (PSF) is the convolution of a Gaussian beam with a waist radius $w_2$, which is determined by the focal length $f$, and the Fourier transform of the grating function.

From Eq. (5.14), if the grating satisfies the constraint in Eq. (5.19), the Fourier transform of the grating function is a delta function which provides a shift operation from the cell center position $\delta(x)$ to a new location represented by $\delta(\frac{x}{\lambda f} - \frac{1}{T})$. The shift distance $S$ of each cell on the Fourier plane is given by

$$\frac{S}{\lambda f} = \frac{1}{T}, \tag{5.23}$$

where $T$ denotes the length of grating period of the cell, $\lambda$ is the wavelength of the light, and $f$ is the focal length of the microlens. The maximum shift length $S_{max}$ produced by a 2-D shuffle operation is equal to $\frac{\sqrt{2}}{2}mD$ as shown in Fig. 5.7b. Thus, the grating period of those cells has minimum length $T_{min}$. If we choose $T_{min} = 10d$ to satisfy the approximation $\sin(\alpha) \approx \alpha$, we have

$$T_{min} = 10 \times \frac{\lambda}{(n-1)}, \tag{5.24}$$

$$\theta_{max} = \frac{\lambda}{T_{min}} = \frac{S_{max}}{f}. \tag{5.25}$$

Consider one particular cell (size $D$) in stage $i$. A Gaussian beam with waist radius $w_0$ propagates through a distance $z_1$ and is incident on the grating plane with a waist radius $w_1 = D/2k$. If the waist radius of the Gaussian beam and the facet size at the back focal plane (stage $i+1$) are equal to $w_2$ and $D$, respectively, using the Gaussian beam propagation theory described in Eq. (4.5), we have the following constraint on the focal length $f$

$$f = \frac{\pi w_1 w_2}{\lambda} = \frac{\pi D w_2}{2k\lambda} = \frac{\pi D w_2}{4\lambda}. \tag{5.26}$$

Combining Eq. (5.25) and Eq. (5.26), we have

$$T_{min} = 10\frac{\lambda}{(n-1)} = \frac{\pi w_2}{2\sqrt{2m}} \tag{5.27}$$

as the required minimum grating period. Equation (5.27) determines the maximum number of channels $N = m^2$ based on the value of $w_2$, the wavelength $\lambda$, and index of refraction $n$. Here, we assume that $\lambda = 0.8\,\mu m$.

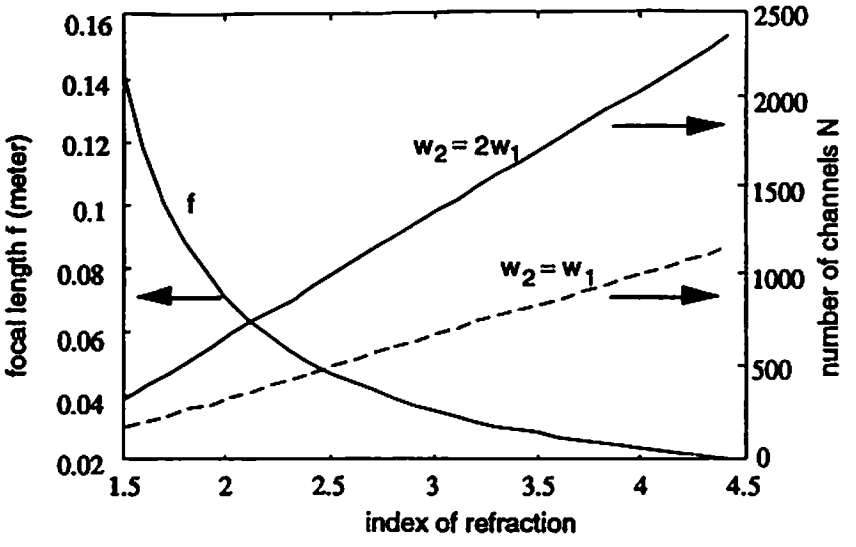



Figure 5.8: Total number of channels $N$ and focal length are plotted as a function of the index of refraction. Two criteria $w_2 = w_1$ and $w_2 = 2w_1$ are used.

Figure 5.8 describes the numerical results of the focal length and number of channels as a function of the index of refraction. Two criteria $w_2 = w_1$ and $w_2 = 2w_1$ are used. Systems using the second constraint can achieve higher packing density but introduce more crosstalk into adjacent channels.

The $f$-numbers of the microlens and the length of the grating period plotted as a function of the index of refraction $n$ are shown in Fig. 5.9 and Fig. 5.10, respectively. From Eq. (4.8), the minimum feature size $\delta_{min}$ in the $L$-level microlens profile is given by

$$\delta_{min} \times L \approx 2\lambda F\#. \tag{5.28}$$

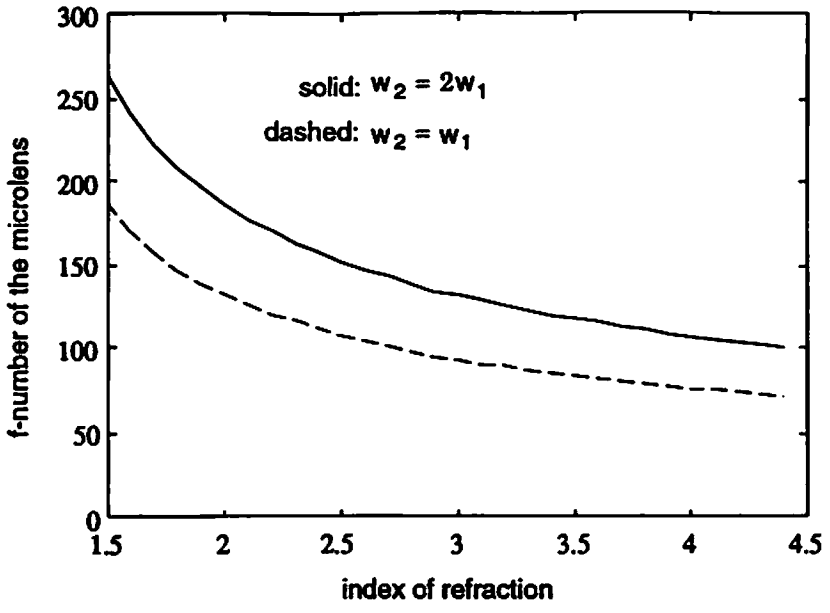



Figure 5.9: *f*-number of the microlens is plotted as a function of the index of refraction. Two criteria $w_2 = w_1$ and $w_2 = 2w_1$ are used.

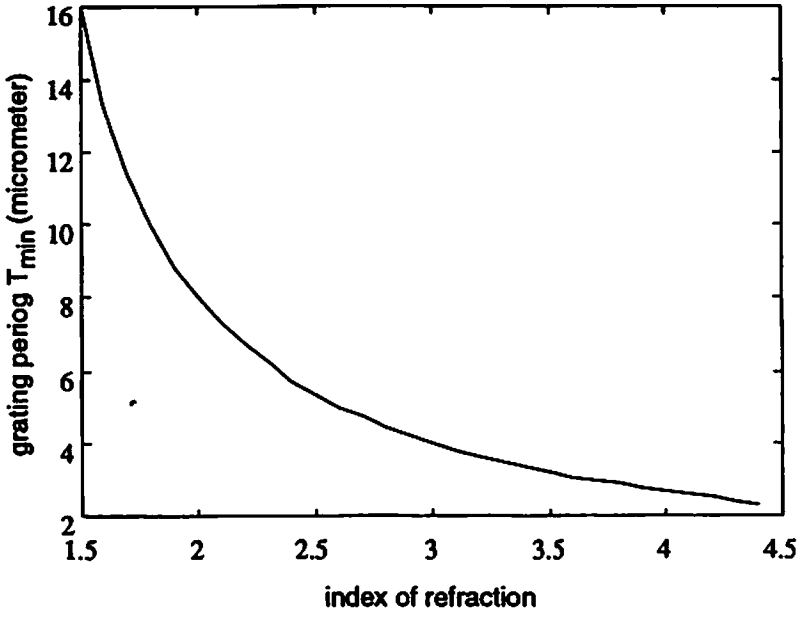



Figure 5.10: Minimum grating period $T_{min}$ is plotted as a function of the index of refraction.

Because the $f$-number of the microlens is relatively large, the fabrication limitations are determined by the length of the minimum grating period $T_{min}$. We assume that $\delta_{min} = 1\,\mu m$. Thus, an $L$-level blazed grating can be made with a minimum grating period $T_{min} = L\,\mu m$. If the number of the levels is increased, the efficiency of the interconnection system is increased.

One important factor which reduces the efficiency of the system is diffraction caused by the finite aperture of the facet. For a square aperture, the point spread function of a facet at the back focal plane is a sinc function with distance $\rho = \lambda f / D$ to its first zero point. Assuming that the criterion $w_2 = w_1$ is used, we can calculate the ratio $r = w_1/\rho$ from Eq. (5.26) and the result is a constant ($r = 2k/\pi$). The PSFs of the Gaussian beam with a waist $w_1$ and sinc function sinc($r$) are shown in Fig. 5.11. With a cell size $D = 2kw_1$, the crosstalk caused by facet aperture is small and more than 99% of the light energy will be distributed in the cell.

The total light efficiency of the system is the product of the grating efficiency and the detection efficiency. The former is determined by the minimum grating period $T_{min}$ from Eq. (5.2) and the latter is strongly dependent on the size of the detectors at the node plane $i+1$. We plot the curve showing the relation between the detection efficiency and detector size in Fig. 5.12. Note that the grating period $T_{min}$ is a function of the index of refraction ($n$) of the material we choose for the grating design and fabrication. For example, for at least 80% detection efficiency, a $70\,\mu m \times 70\,\mu m$ detector size for a four-level grating design and a $150\,\mu m \times 150\,\mu m$ detector size for a sixteen-level grating design are required.

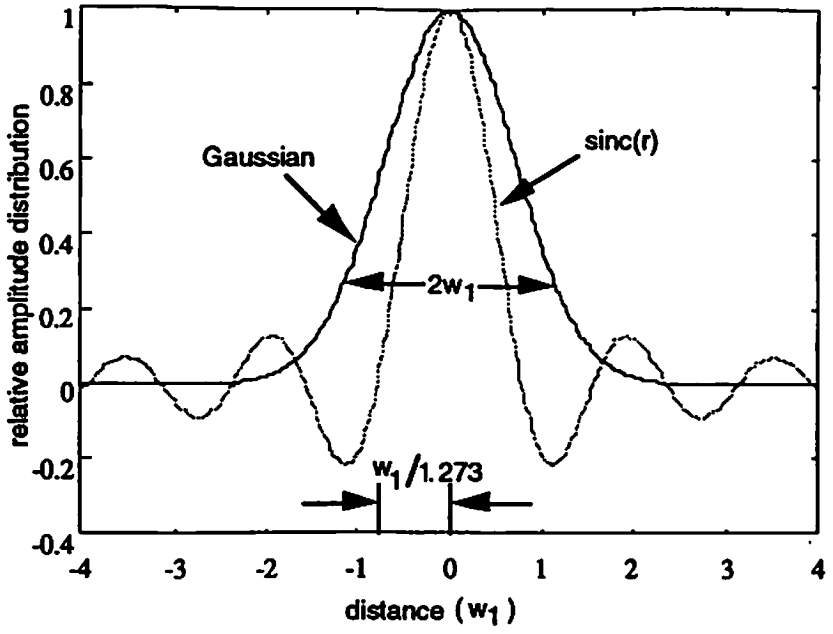



Figure 5.11: The PSFs of the Gaussian beam and sinc function at the back focal plane. Here, $w_1 = 1.273\rho$ is equal to the size of each channel.

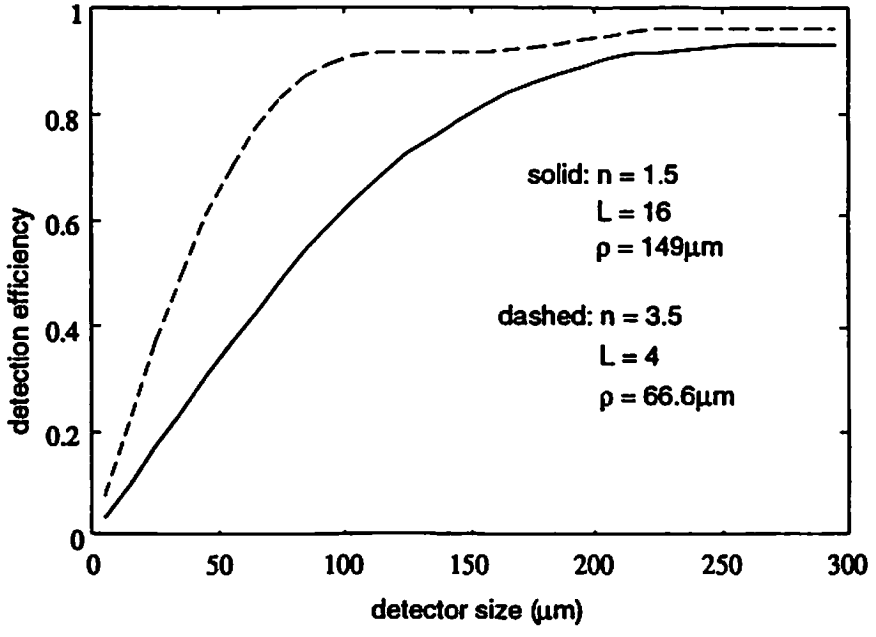



Figure 5.12: The detection efficiency of four-level and sixteen-level gratings versus detector size.

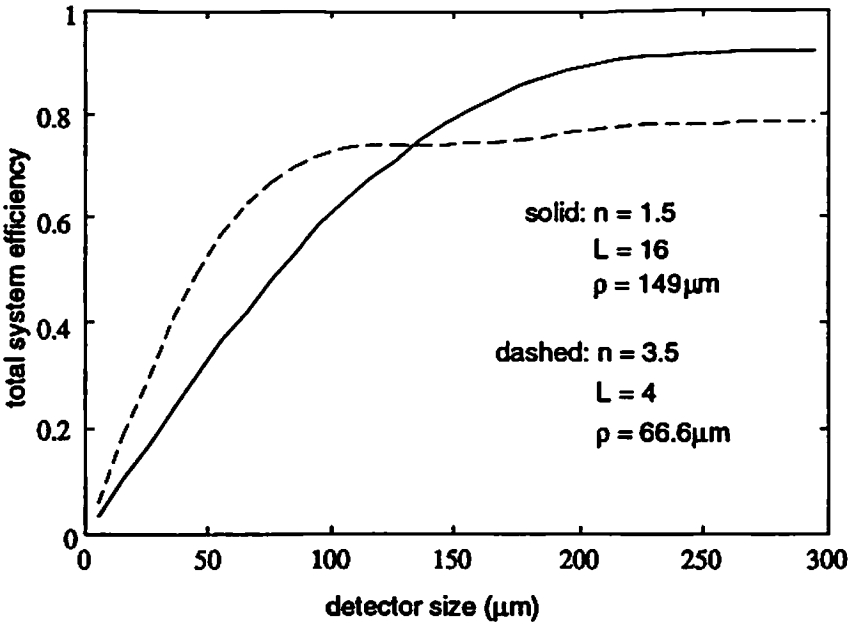



Figure 5.13: The total system efficiency of four-level ($n = 3.5$) and sixteen-level ($n = 1.5$) gratings versus detector size.

Total system efficiency based on four-level and sixteen-level gratings versus detector size is shown in Fig. 5.13. For a system with detector size less than $140\,\mu m$, a 4-level grating is the best choice. Numerical results showing the maximum possible number of channels for four-level and sixteen-level blazed grating designs are shown in Table 5.2. Here we use the results of grating efficiency in Table 5.1 and assume $w_2 = w_1$. The light efficiency loss caused by the clipping factor is neglected.

| Levels | Packing Density | Grating Efficiency | $f$ | $D$ |
|---|---|---|---|---|
| 4 | 694 channels/cm$^2$ | 81% | $35\,mm$ | $380\,\mu m$ |
| 16 | 173 channels/cm$^2$ | 98.7% | $141\,mm$ | $760\,\mu m$ |

Table 5.2: The numerical results of the first method using blazed gratings and microlenses.

**Phase profiles:** The space-variant grating profile of a particular cell is determined solely by its shift distance $S$ and direction vector provided by the 2-D folded shuffle operation. If one cell requires a shift distance $S$ at the back focal plane, its grating period $T$ is given by

$$T = \frac{\lambda f}{S} = \frac{S_{max}}{S} T_{min} \geq T_{min}. \tag{5.29}$$

Associated with each cell $a(i,j)$ are direction vectors which can be calculated from values of the cell indices $i$ and $j$. We define these two vectors as $\vec{u}$ and $\vec{v}$ for cell $a(i,j)$. The positive (down) or negative (up) shift of index $i$ implies a vector $\vec{u}$ whose direction is equivalent to the direction of the shift and whose magnitude is equivalent to the distance of the shift. Similarly, the positive (right) or negative (left) shift of index $j$ implies a vector $\vec{v}$ with same direction and magnitude as the shift. During the phase profile design, a reference point is also associated with each cell $a(i,j)$ and is determined by the directions of the two vectors $\vec{u}$ and $\vec{v}$ as shown in Fig. 5.14.

For example let us consider a cell of size $D \times D$ whose shift vectors $\vec{u}$ and $\vec{v}$ are both positive. The shift direction of this cell is represented by the direction line $P$ as shown in Fig. 5.15. The angle $\theta_{uv}$ is given by

$$\tan(\theta_{uv}) = \frac{|\vec{u}|}{|\vec{v}|}, \tag{5.30}$$

where $|\vec{u}|$ and $|\vec{v}|$ are the magnitude of vector $\vec{u}$ and $\vec{v}$, respectively.

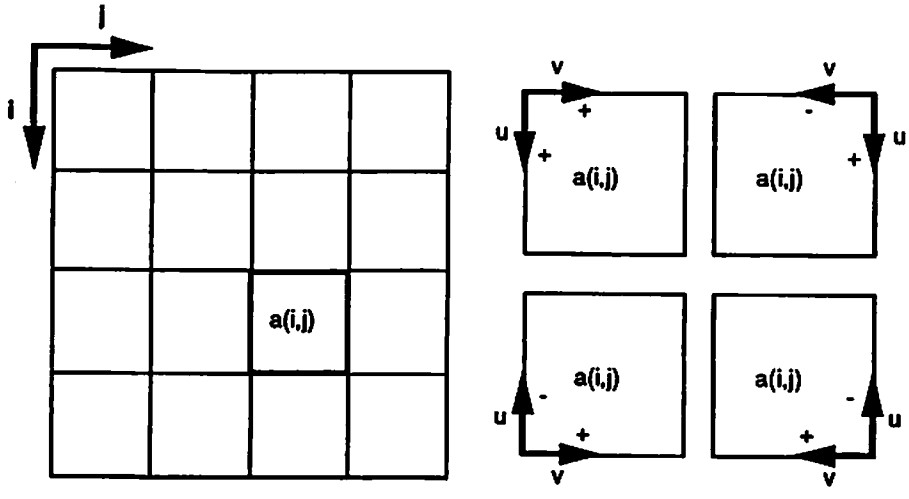



Figure 5.14: The location of reference point in each cell $a(i,j)$ during the phase profile design is determined by the positive or negative directions of two vectors $\vec{u}$ and $\vec{v}$.

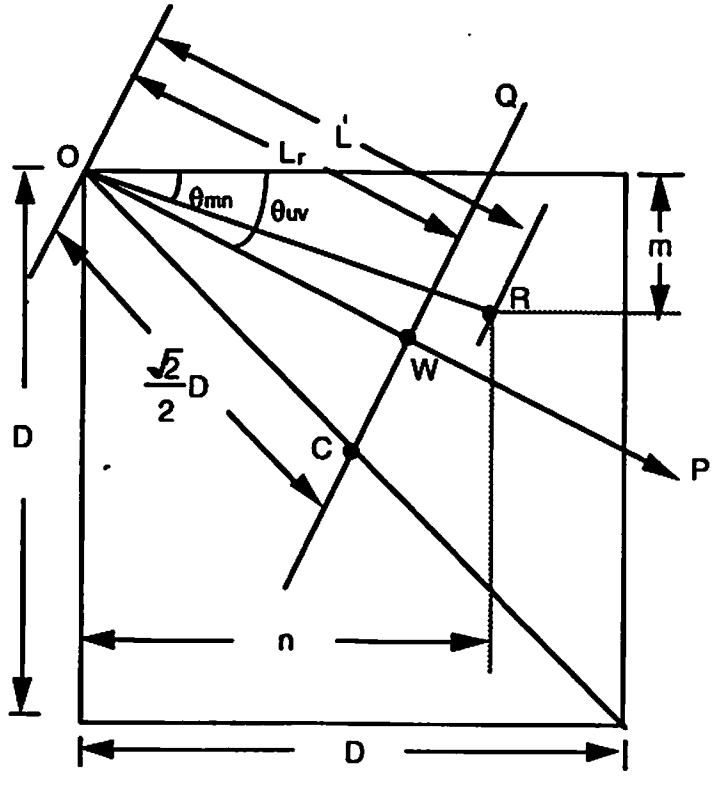



Figure 5.15: The geometric relation of a point $R$ in a particular cell.

The period $T$ and the direction of the blazed grating in this cell can also be determined by the magnitude of $\vec{u}$ and $\vec{v}$. We can draw a straight line $Q$ which passes through cell center point $C$ and is perpendicular to the direction line $P$. The distance from the intersection point $W$ of these two lines to the reference point $O$ is called the reference length, denoted by $L_r$, where

$$L_r = \frac{\sqrt{2}}{2} D\cos(45° - \theta_{uv}). \tag{5.31}$$

Thus, the depth $d'$ of the grating at a certain point $R$ with vertical length $m$ and horizontal length $n$ from the reference point $O$ is given by

$$d' = (L' - L_r)_{modulo\ T} \times \frac{1}{T} \times \frac{\lambda}{(n-1)}, \tag{5.32}$$

where

$$L' = \sqrt{m^2 + n^2}\cos(\theta_{uv} - \theta_{mn}), \tag{5.33}$$

$$\theta_{mn} = \tan^{-1}(\tfrac{m}{n}). \tag{5.34}$$

The phase profile $\phi_g$ at point $R$ is given by

$$\phi_g(R) = d' \times \frac{2\pi(n-1)}{\lambda} = 2\pi\frac{T}{T_{min}}. \tag{5.35}$$

Because a microlens with focal length $f$ is required to perform a Fourier transform, we can combine the phase profiles of microlens $\phi_l$ with $\phi_g$ to create a arbitrary phase profile $\phi'$. At point $R$, we have

$$\phi_l(R) = \frac{\pi}{\lambda f} \times ((m - \frac{D}{2})^2 + (n - \frac{D}{2})^2), \tag{5.36}$$

and

$$\phi'(R) = (\phi_g(R) + \phi_l(R))_{modulo\ 2\pi}. \tag{5.37}$$

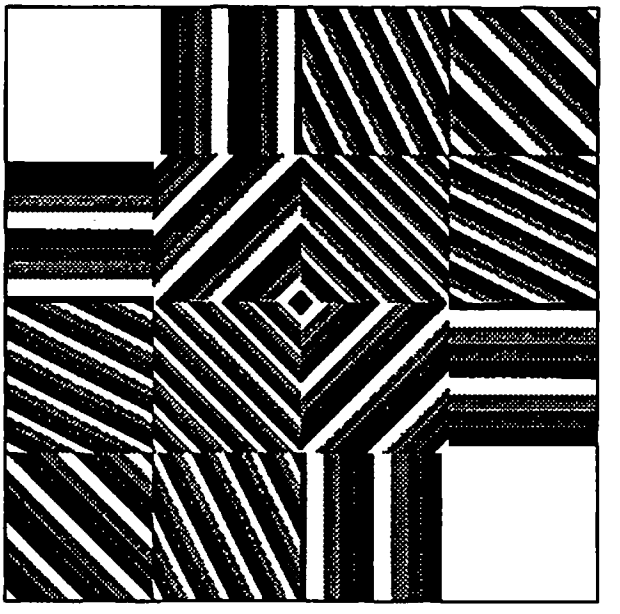


Figure 5.16: An array of computer-generated $4 \times 4$ four-level blazed grating profiles based on the first design method.

In Fig. 5.16, an array of $4 \times 4$ four-level blazed gratings is generated by computer program to perform a 2-D folded shuffle. Four gray levels from

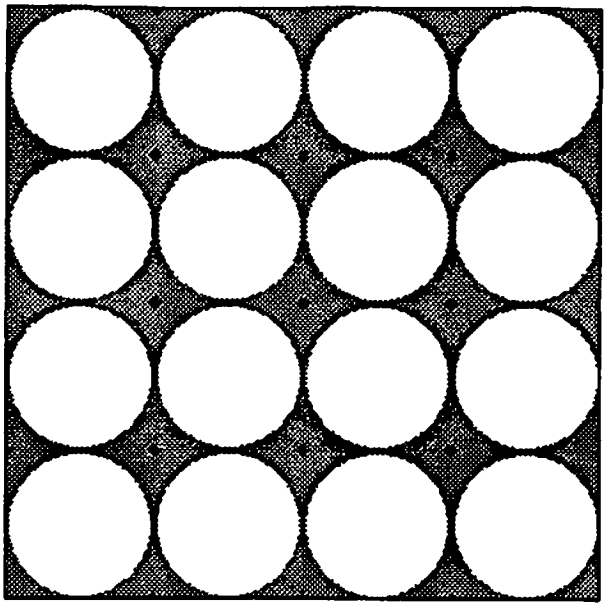


Figure 5.17: An array of computer-generated $4 \times 4$ microlenses profiles. Due to high $f$-number of the microlens, lens profiles can only be quantized into three levels in each cell.

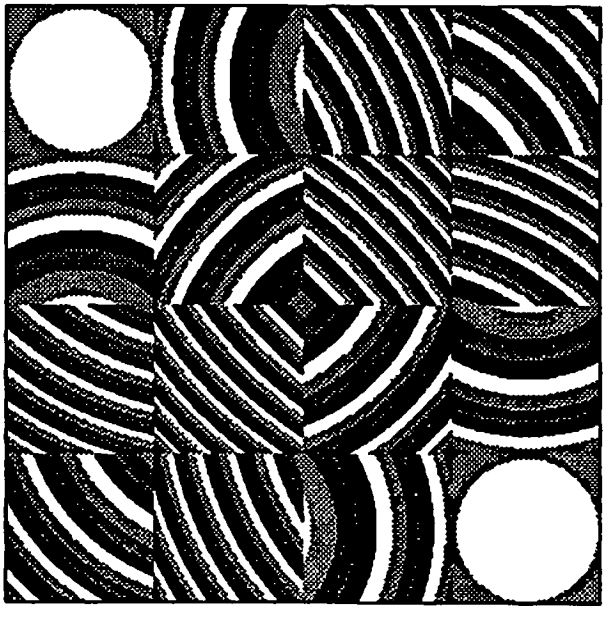


Figure 5.18: Four-level DOE patterns based on first blazed grating design method.

white to black represent four phases: 0, $\pi/2$, $\pi$, and $3\pi/2$, respectively. The four-level profile of the microlens array is shown in Fig. 5.17. The DOE pattern combining grating and lens together is shown in Fig. 5.18.

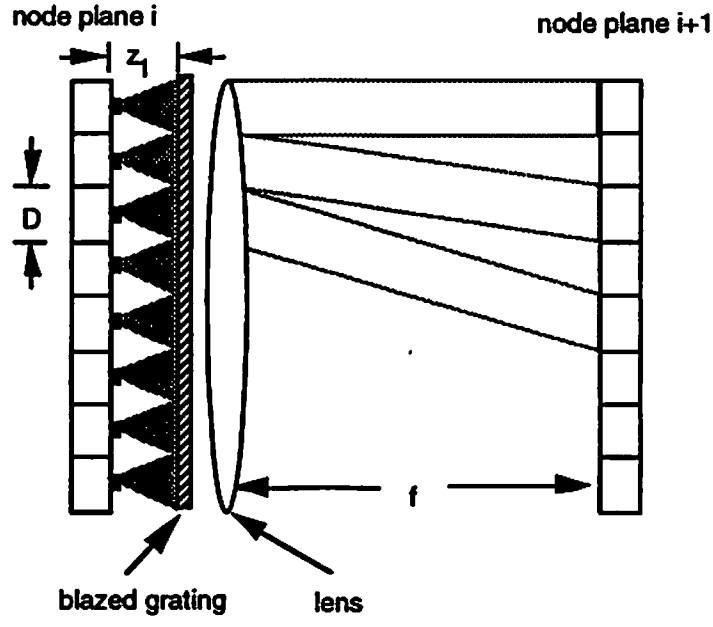
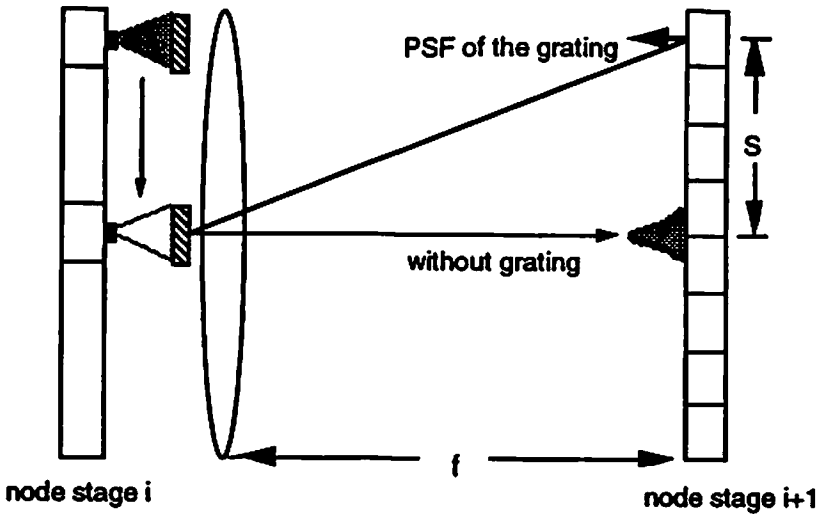


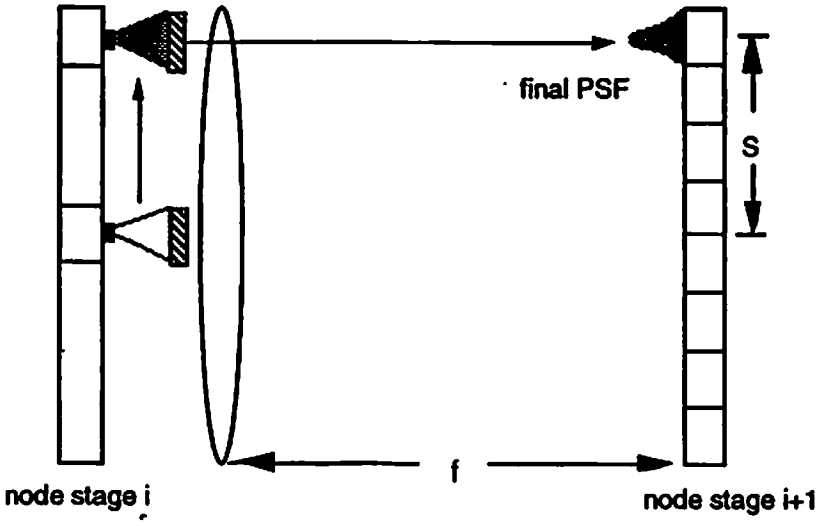

Figure 5.19: 1-D interconnection system structure utilizing blazed gratings and a large lens.

#### 5.2.1.2 Second Design Method

**System design:** A second system is shown as a 1-D cross section in Fig. 5.19. The difference in this system from the previous system is that we replace the microlens array with a macrolens. In order to demonstrate the grating design procedure of each cell, we assume that a particular cell is located at the top of the note stage $i$ and we want to design a grating routing its light to the top cell of the note stage $i+1$ as shown in Fig. 5.20a. Considering that this cell is located at the center of the node stage $i$,




(a)




(b)

Figure 5.20: The grating design method for the new system structure: (a) The grating pattern of the cell is generated by calculating the distance between the center point of the node stage $i+1$ and the center point of the destination cell; (b) cell is moved to its original location and the intensity distribution of PSFs at the Fourier plane remain unchanged.

Without the grating, a Gaussian beam is formed at the center of the back focal plane. The blazed grating is designed such that it performs a shift operation over the distance from the center point of the node stage $i+1$ to the center of its top cell. If we shift the cell from the center of the node stage $i$ to its original location as shown in Fig. 5.20b, the intensity of the PSFs at the back focal plane will not be changed because a shift at the object plane only causes a phase shift at the Fourier plane. Therefore, the grating pattern for this cell is generated by calculating the distance between the center point of the node stage $i+1$ and the center point of the destined cell. We can use this new system structure to calculate the grating patterns of a 2-D folded shuffle.

The shifting of each cell after the 2-D folded shuffle operation is now calculated from the center of the cell array as shown in Fig. 5.21. The maximum shift distance is from the center to four corners, thus $S_{max}$ is given by

$$S_{max} = \frac{\sqrt{2}}{2} D(m-1). \tag{5.38}$$

The maximum shift distance of the second method is less than that of the first method, thus, a higher packing density and shorter longitudinal size can be expected. The two design methods are compared in Fig. 5.22. Numerical results showing the maximum possible number of channels for four-level and sixteen-level blazed grating designs based on the second method are shown in Table 5.3. Here we assume $w_2 = w_1$ and that the light efficiency loss caused by the clipping factor is negligible.

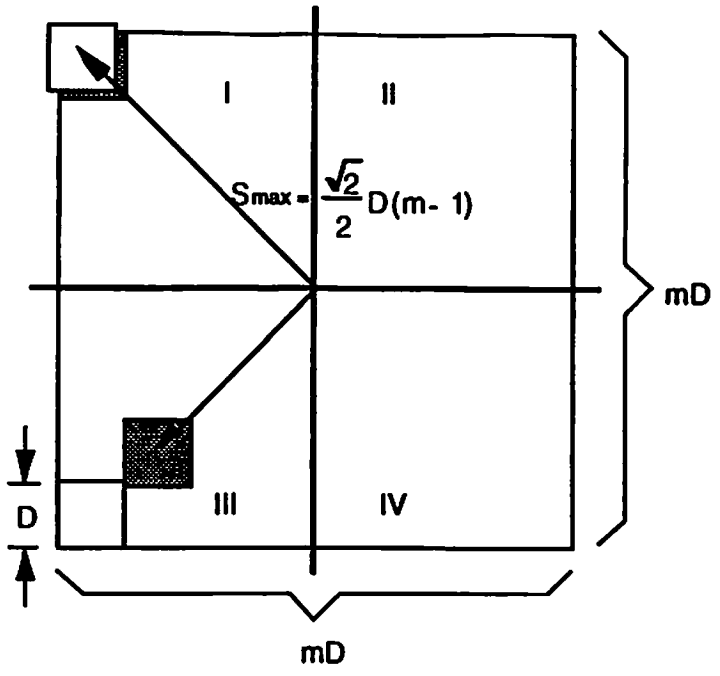



Figure 5.21: The shifting of each cell based on the second design method. Gray and white boxes represent the location of cells before and after a 2-D folded shuffle operation, respectively.

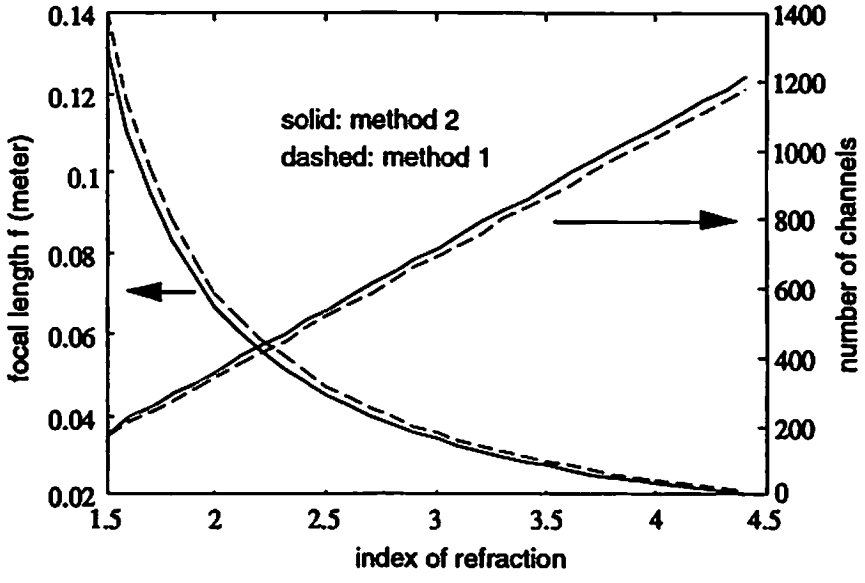



Figure 5.22: The total number of channels and the longitudinal size of the systems based on two design methods are compared.

| Levels | Packing Density | Grating Efficiency | $f$ | $D$ |
|---|---|---|---|---|
| 4 | 897 channels/cm$^2$ | 81% | 27 $mm$ | 334 $\mu m$ |
| 16 | 187 channels/cm$^2$ | 98.7% | 131 $mm$ | 731 $\mu m$ |

Table 5.3: The numerical results of the second method using blazed gratings and a large lens.

**Phase profile:** The phase profile of the grating in each cell can be determined by the same analysis discussed in the first design method. An array of $4 \times 4$ four-level blazed gratings generated by computer program based on second design method is shown in Fig. 5.23,. Four gray levels from white to black represent four phases: 0, $\pi/2$, $\pi$, and $3\pi/2$, respectively. The four-level profile of the lens shown in Fig. 5.24. The DOE pattern combining grating and lens together is shown in Fig. 5.25.

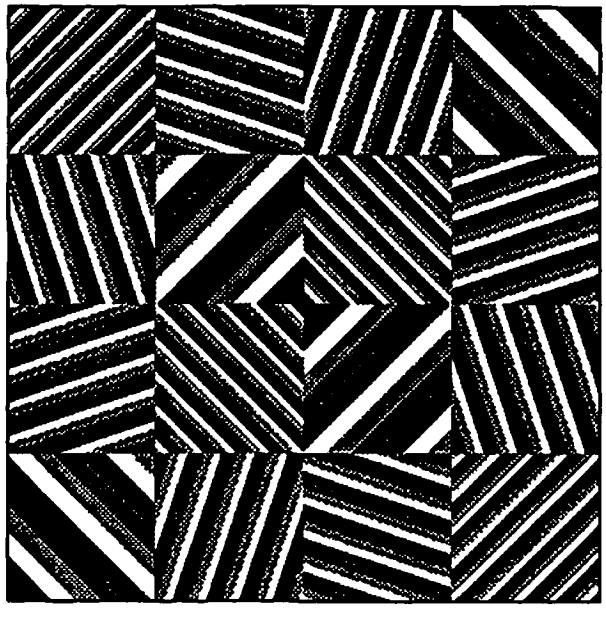


Figure 5.23: An array of computer-generated $4 \times 4$ four-level blazed grating profiles based on the second design method.

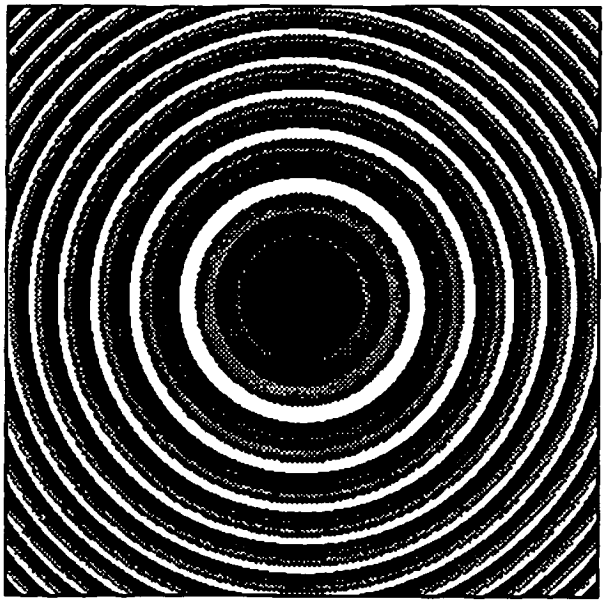


Figure 5.24: A computer-generated $4 \times 4$ four-level macrolens profile.

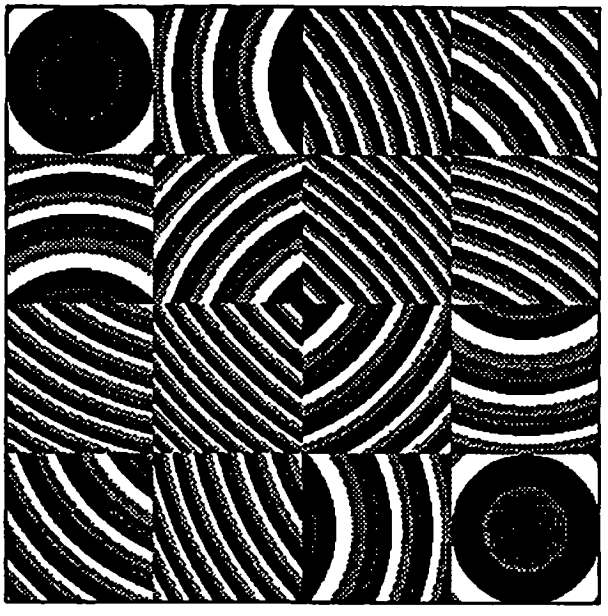


Figure 5.25: Four-level DOE pattern based on the second blazed grating design method.

### 5.2.2 Off-Axis Microlenses

A one-to-one space-variant DOE is essentially a multifaceted optical element consisting of an array of off-axis diffractive microlenses in which each microlens deflects and focuses a single incident beam to its destination. A 1-D perfect shuffle interconnection system structure using off-axis microlenses is shown in Fig. 5.26.

Following the discussions In Section 5.2.1 and using Eq. (5.26) in the blazed grating cases, we have the following constraint on the focal length $f$

$$f = \frac{\pi D w_2}{2k\lambda} \tag{5.39}$$

where $D$ is the size of the channel, $w_2$ is the waist radius of the Gaussian beam at the back focal plane, $k = 2$ is the clipping ratio, and wavelength $\lambda = 0.8\mu m$.

If a particular cell needs to be deflected at angle $\phi$ and shifted a distance $S$ at the back focal plane, a grating is designed as a piece of a macrolens which has a focal length $f$ and a diameter $2(S + D/2)$. The minimum channel size $D$ and maximum shift distance $S$ will determine the minimum feature size in the DOEs. Thus, the fabrication accuracy will limit the packing density of the system. For a fixed chip size $D_L = ND$, the maximum shift distance for a 2-D folded shuffle interconnection is equal to $S_{max} = \frac{\sqrt{2}}{2} mD$. The $f$-number of the lens is given by

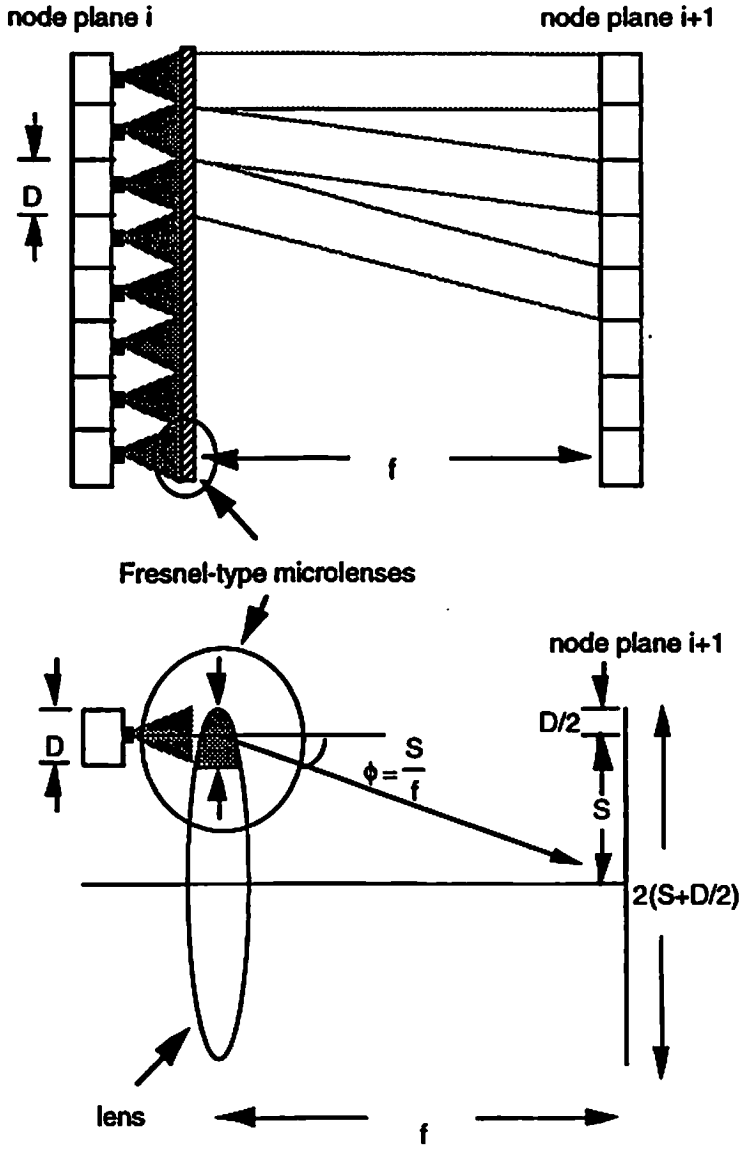



**Figure 5.26: 1-D perfect shuffle interconnection system structure using off-axis microlenses.**

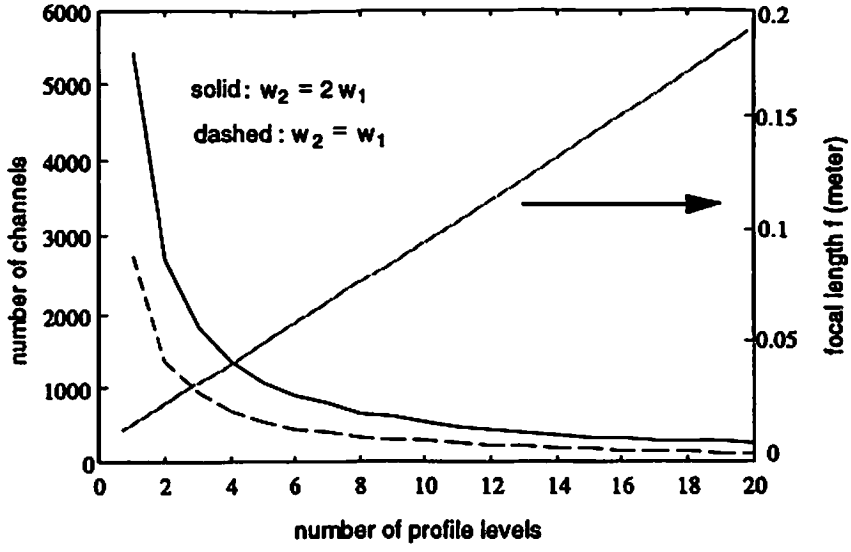



Figure 5.27: Total number of channels and focal length versus number of profile levels. Two criteria $w_2 = w_1$ and $w_2 = 2w_1$ are used.

$$F\# = \frac{f}{2(S_{max} + D/2)} = \frac{\pi w_2}{2k\lambda(1+\sqrt{2N})}. \tag{5.40}$$

From Eq. (5.28), the relation between minimum feature size and the $L$-level lens profile is determined by

$$\delta_{min} \times L \approx 2\lambda F\# = \frac{\pi w_2}{k(1+\sqrt{2N})} \tag{5.41}$$

assuming $\delta_{min} = 1\,\mu m$. Equation (5.41) gives the packing density limitation based on the size of $w_2$, fabrication accuracy and lens efficiency (number of levels). Figure 5.27 shows the total number of channels and the focal length

$f$ versus the number of profile levels for two different value of $w_2$. Again, two criteria $w_2 = w_1$ and $w_2 = 2w_1$ are considered. From Fig. 5.27, higher number of profile levels gives a higher light-efficient system, but requires longer interconnection distance (propagation latency) and lower parallelism (packing density).

Numerical results on the number of channels for four-level and sixteen-level lens designs are shown in Table 5.4. Here we assume $w_2 = w_1$ and neglect the aberrations caused by the clipping factor and the limited aperture. In Fig. 5.28, an array of $4 \times 4$ four-level off-axis microlenses is designed by computer to perform a 2-D folded shuffle. Four gray levels from white to black represent four phases: 0, $\pi/2$, $\pi$, and $3\pi/2$, respectively.

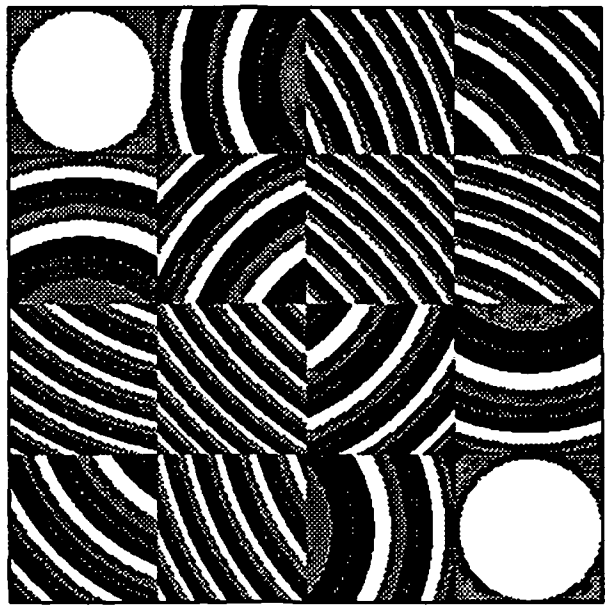


Figure 5.28: An array of computer-generated $4 \times 4$ four-level off-axis microlenses profiles.

| Levels | Packing Density | Grating Efficiency | $f$ | $D$ |
|---|---|---|---|---|
| 4 | 675 channels/cm$^2$ | 81% | 36 $mm$ | 385 $\mu m$ |
| 16 | 164 channels/cm$^2$ | 98.7% | 149 $mm$ | 780 $\mu m$ |

Table 5.4: Numerical results giving the maximum number of channels for the design method using off-axis microlenses.

## 5.3 Space-Semivariant DOE Designs

### 5.3.1 Four-Copy Algorithm

Extensive work has been done on implementing optical 2-D folded shuffles. The space-semivariant design using four simple lenses was first demonstrated by Stirk et. al.[17], and followed by others [18]. These techniques rely on making four copies of the magnified and masked input plane, and are called four-copy algorithms. In this algorithm, the input plane is first magnified and masked. Then, an optical system generates four copies of the preprocessed plane. By shifting and superimposing four copies of this plane properly, the shuffled result appears in the center of the output plane. An illustration of the method is shown in Fig. 5.29. The DOEs implementation of the four-copy algorithm was demonstrated in ref. [48], which used a interferogram-type binary DOE to provide a fanout of four and focusing. There are two drawbacks to this algorithm. First, only 25% light efficiency can be achieved because only one quadrant of the shifted copies is used in the output plane. Second, a spatial encoding scheme (masking) is needed, which further reduces the light efficiency in practice.

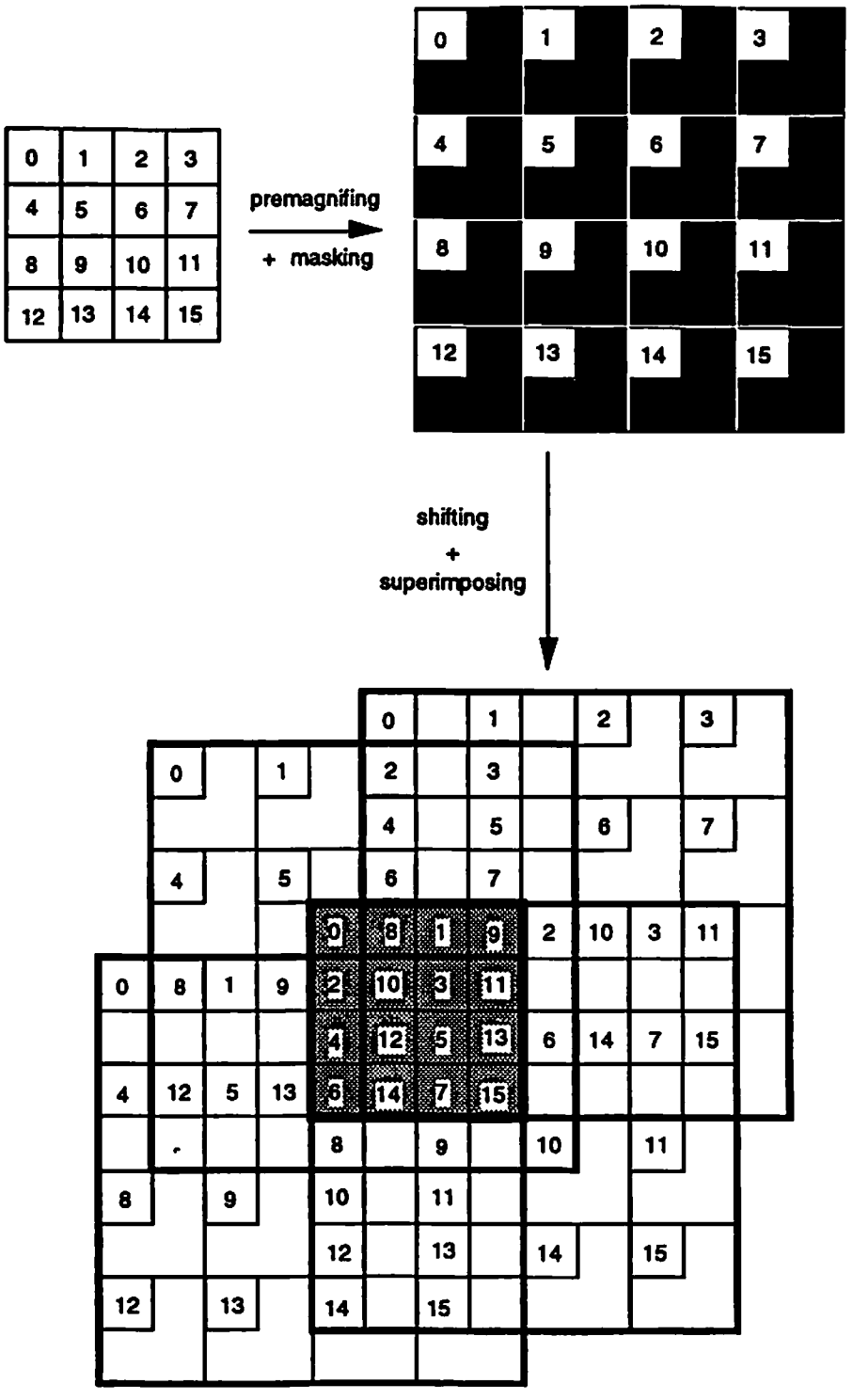



Figure 5.29: Illustration of the four-copy algorithm for 2-D folded shuffle. Only the shaded section in the last part of the figure is retained.

### 5.3.2 One-Copy Algorithm

Based on similar principles as the four-copy algorithm, a simplified algorithm called one-copy algorithm was demonstrated in [64] using holographic elements. The theoretical analysis of different 2-D shuffled implementation schemes based on the one-copy algorithm is studied in [60]. In this algorithm, the input plane is again premagnified and masked. However, as shown in Fig. 5.30, instead of generating and shifting four copies of the input plane, each quadrant of the input plane is properly shifted and superposed by a grating pattern to produce the shuffled result. In principle, 100% light efficiency can be achieved from the preprocessed plane to output plane. The DOE implementation of one-copy algorithm was described in [49]. The one-copy algorithm has the advantage of higher light efficiency than the four-copy algorithm, but it requires a special masking process.

### 5.3.3 Demagnifying Method

A new one-copy approach based on magnifying and demagnifying without a masking process is discussed in this section. The 1-D concept is illustrated in Fig. 5.31. Let us consider that a input plane (channel size $D_2 = D$) is magnified by a telecentric optical system. At the object plane, a DOE (facet size $D_1 = 2D_2$) receives an array of normally incident Gaussian beams, and an array of microlenses demagnifies and focuses the Gaussian beams to the output plane (channel size $D$).

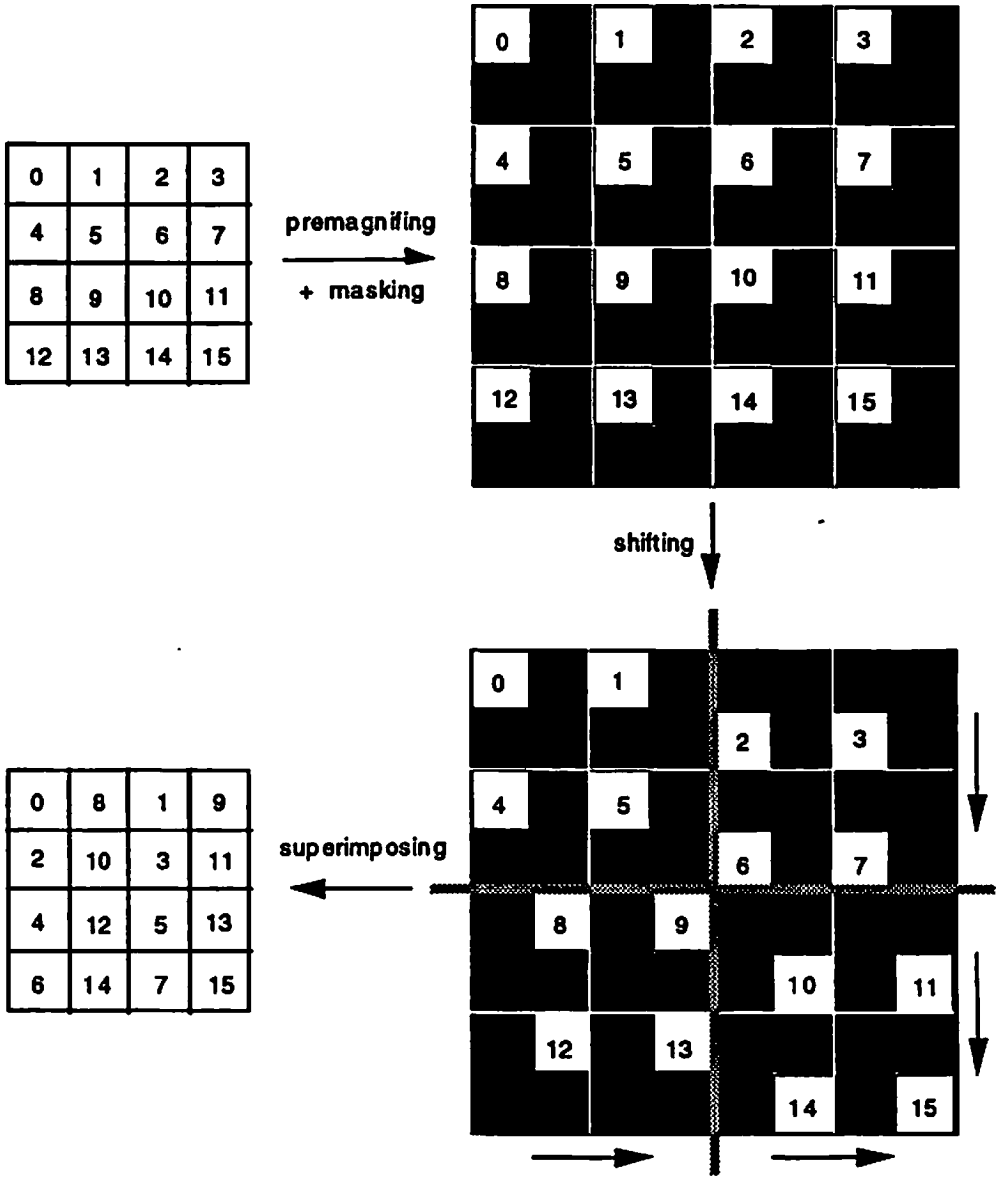



Figure 5.30: Illustration of one-copy algorithm for 2-D folded shuffle.

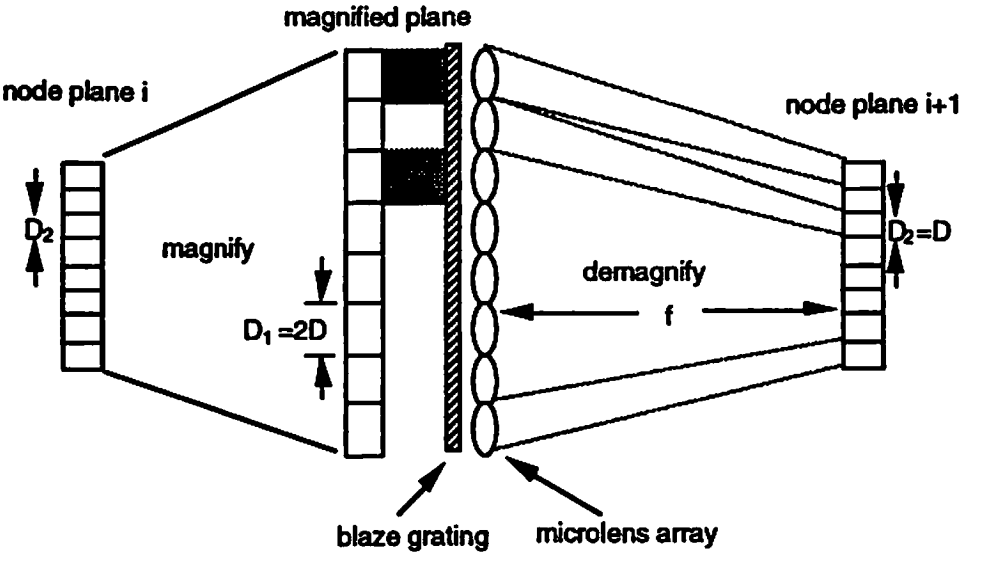



Figure 5.31: Illustration of the demagnifying method for achieving a 1-D perfect shuffle.

Some possible advantages of this method are showing in the 2-D plane. In Fig. 5.32, channels in each quadrant of the object plane are demagnified by a factor of two and shifted the same amount of distance on the output plane. The incident beams in quadrant I and IV are deflected by a distance $\sqrt{2}D(m-1)/2$ over a longitudinal distance $f$, where $f$ is the focal length of the microlens system. The beams in quadrant II and III are deflected by a distance $\sqrt{2}D(m+1)/2$. Thus, only four different designs of grating patterns are needed. One DOE pattern for a channel in each quadrant will be first calculated and the data is stored for the pattern generation of the rest of the facets.

Both design methods discussed previously in section 5.2, the blazed grating design methods and off-axis microlens designs, can be used to

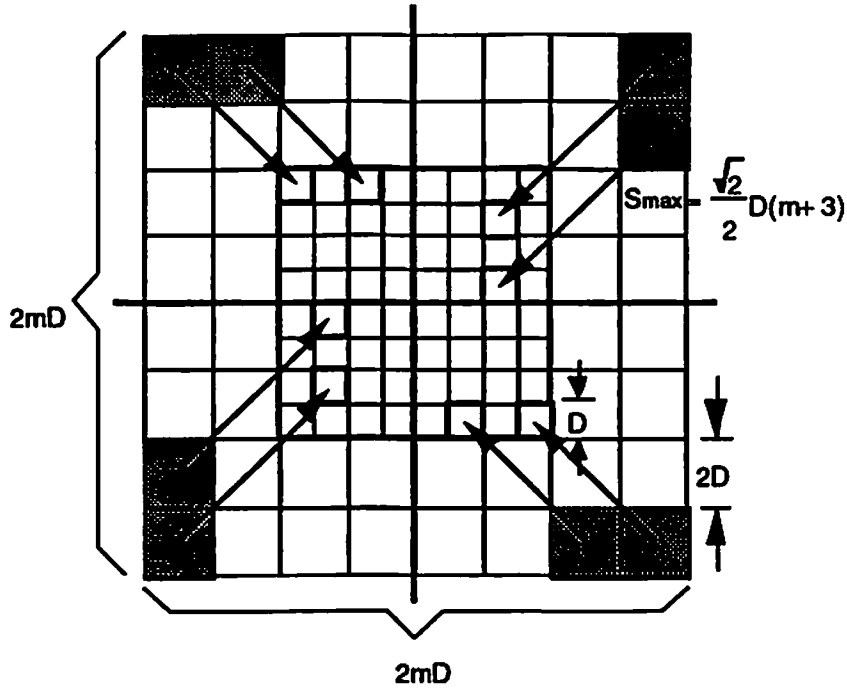



Figure 5.32: Illustration of shifting properties of the demagnifying design method in 2-D plane. Gray and white boxes represent the location of cells before and after the operation, respectively. Note that a demagnification of two is utilized to perform a 2-D folded shuffle.

realize this space-semivariant DOE. Here we choose the first blazed grating design method. A microlens array is used to demagnify and focus the incident Gaussian beams (facet size $2D$) from object plane to output plane (facet size $D$). Thus, using Eq. (5.39), we have the constraint on the focal length of the microlens array

$$f = \frac{2D^2\pi}{4k^2\lambda}, \tag{5.42}$$

and

$$\theta_{max} = \frac{\lambda}{T_{min}} = \frac{S_{max}}{f}, \tag{5.43}$$

where $S_{max} = \sqrt{2}D(m+1)/2$. The packing density of the system can be calculated from Eq. (5.42) and (5.43). Numerical results of four-level and sixteen-level grating designs are shown in Table 5.5. Here we assume $k = 2$, $\lambda = 0.8\,\mu m$, and the demagnification factor is equal to two.

In Fig. 5.33, an array of $4 \times 4$ four-level grating patterns for the demagnifying method are demonstrated. As before, four gray levels from white to black represent four phases: 0, $\pi/2$, $\pi$, and $3\pi/2$, respectively. Note that the four grating patterns within one quadrant are the same. During the mask writing procedure, patterns can be calculated, recorded, and duplicated through each quadrant. There are only four different patterns in this grating. Thus, the degree of space-variance of this semivariant design is equal to four.

| Levels | Packing Density | Grating Efficiency | $f$ | $D$ |
|---|---|---|---|---|
| 4 | 1351 channels/cm$^2$ | 81% | 36.3 $mm$ | 272 $\mu m$ |
| 16 | 329 channels/cm$^2$ | 98.7% | 149 $mm$ | 551 $\mu m$ |

Table 5.5: The numerical results of the demagnification method using first blazed grating design method.

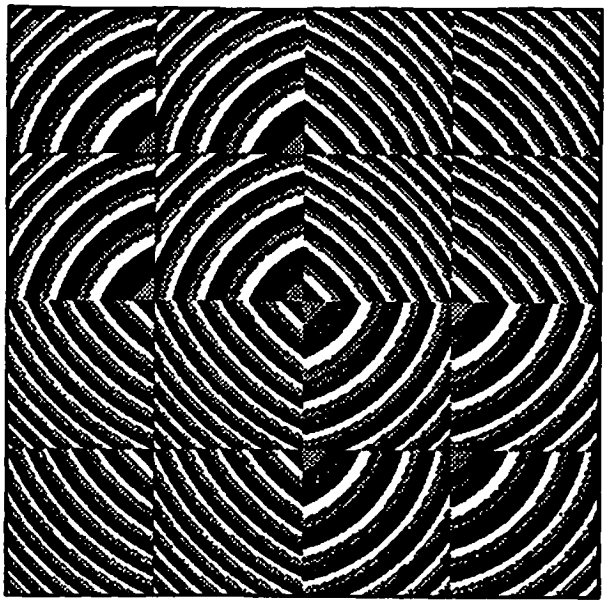


Figure 5.33: An array of computer-generated $4 \times 4$ four-level DOE profiles based on the demagnification method.

## 5.4 Summary

In this chapter, the shifting properties of 2-D folded shuffles were discussed. Based on its reflected-symmetric characteristics, the complexity of a space-variant grating needed to implement the 2-D folded shuffle interconnection can be reduced by a factor of 4. We studied two types of interconnection system designs based on DOEs, i. e., space-variant and space-semivariant designs. The concepts of using blazed gratings and off-axis microlenses was proposed for space-variant system design. A new demagnifying algorithm for space-semivariant system design has been presented. The performance of each design is summarized in Table 5.6 and 5.7. The total light efficiency of the optical system depends on the grating efficiency, system loss, and size of the detector on the next node plane.

| 4-level phase profile | Blazed Grating | | Off-Axis Microlenses | Demagnifying Algorithm |
|---|---|---|---|---|
| | Method 1 | Method 2 | | |
| focal length | 35 $mm$ | 27 $mm$ | 36 $mm$ | 36.3 $mm$ |
| channel size | 380 $\mu m$ | 334 $\mu m$ | 385 $\mu m$ | 272 $\mu m$ |
| grating efficiency | 81% | 81% | 81% | 81% |
| packing density channels/cm$^2$ | 694 | 897 | 675 | 1351 |

Table 5.6: Summary of 2-D perfect shuffle interconnection designs based on a 4-level phase diffractive optical element.

| 16-level phase profile | Blazed Grating | | Off-Axis Microlenses | Demagnifying Algorithm |
|---|---|---|---|---|
| | Method 1 | Method 2 | | |
| focal length | 141 $mm$ | 131 $mm$ | 149 $mm$ | 149 $mm$ |
| channel size | 760 $\mu m$ | 731 $\mu m$ | 780 $\mu m$ | 551 $\mu m$ |
| grating efficiency | 98.7% | 98.7% | 98.7% | 98.7% |
| packing density channels/cm$^2$ | 173 | 187 | 164 | 329 |

Table 5.7: Summary of 2-D perfect shuffle interconnection designs based on a 16-level phase diffractive optical element.

# Chapter 6

# Fabrication Considerations for Multi-level Diffractive Optical Elements

In order to achieve high light efficiency, the multi-level DOE designs discussed in previous chapters should be made using at least four surface-relief levels. In this chapter, we analyze several types of errors occurring in the lithographic fabrication of a multi-level DOE and calculate their effects on the light efficiency of the grating.

## 6.1 Limited Resolution

The fabrication process of a DOE using VLSI microlithographic technology requires a combination of phase-profile generation and material removal. Typically, with standard electron-beam writers, a lateral accuracy of $0.1\,\mu m$ can be achieved during the writing process of the phase-profile generation. That is to say, the transition point of the grating has a maximum location error of $\pm 0.05\,\mu m$. In our case, the effect caused by this position error is negligible because the grating pattern has a relatively long period and a low frequency transition characteristics.

## 6.2 Mask Alignment Errors

When multi-level profiles with $L = 2^K$ levels are desired, it is necessary to create $K$ phase masks and repeat the above fabrication procedure $K$ times. Figure 6.1a shows a 4-level grating profile generated by two phase masks. Although the principle is straightforward, this technique is time-consuming when $L$ increases. Furthermore, the practical difficulties of mask alignment also play a significant role.

During each fabrication cycle, a new mask must be aligned with the existing etching surface. With an optical microscope, an accuracy of $0.5\,\mu m$ can typically be achieved. However, compared to the $0.1\,\mu m$ pattern accuracy, this value is poor. It is obvious that the only component layer with no alignment error is the one which is copied and etched first. If an alignment error occurs in any of the following layers, it will introduce extra features in the surface-relief profile. Figure 6.1b illustrates the 4-level phase profile obtained if the second mask is misaligned by $\Delta_x$.

To analyze alignment errors of a $2^K$-level phase profile, the effect of misalignment occurring in the $m^{th}$ mask, $m = 2,3,\ldots,K$, by an amount $\Delta(m)$ is equivalent to introducing an extra 1-D phase difference profile with a period $T'_m = T/2^{m-1}$ and phases $\pm\phi_a(m)$ where $\phi_a(m) = \pi/2^{m-1}$. We define the duty cycle of this phase difference profile as

$$\alpha(m) = \Delta(m)/T. \tag{6.1}$$

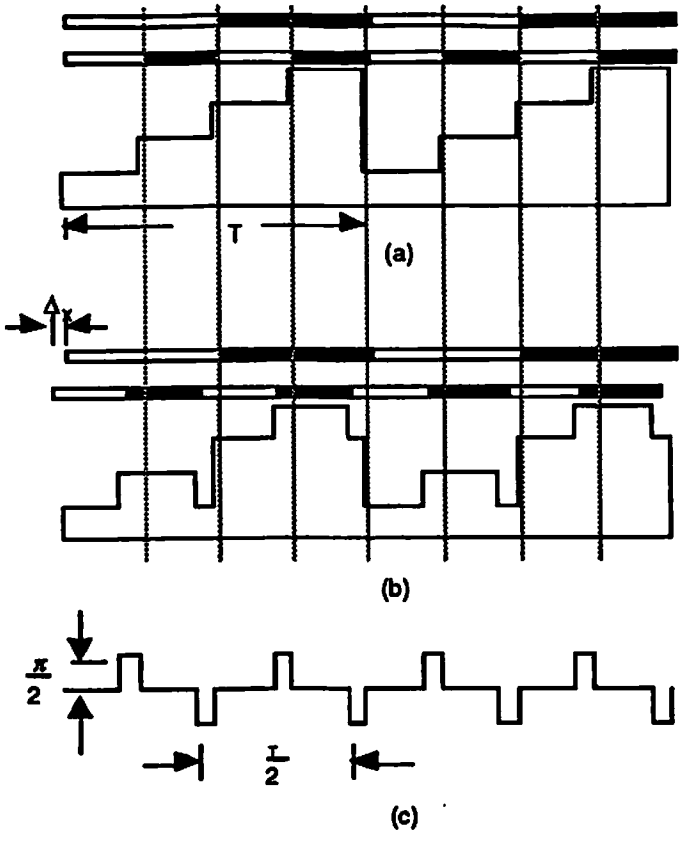



Figure 6.1: Effect of mask alignment in multi-level phase profile generation: (a) the ideal 4-level($K = 2$) phase profile as a function of position; (b) the phase profile obtained if the second mask is misaligned by an amount $\Delta_x$; (c) the effect of the misalignment is equivalent to the phase difference function shown here as a periodic grating with period $T/2$, duty cycle $\alpha = \Delta_x/T$, and phases $\pm\pi/2$.

The transmittance function $U_a^m(x)$ of this grating can be represented by

$$U_a^m(x) = 1 + \{[(e^{j\phi_a(m)} - 1) \times \text{rect}(\frac{x - \alpha(m)T/2}{\alpha(m)T}) + (e^{-j\phi_a(m)} - 1) \times \text{rect}(\frac{x - (\alpha(m)T/2 + T'_m/2)}{\alpha(m)T})] * \sum_p \delta(x - pT'_m)\}, \quad (6.2)$$

where $p$ is the diffraction order. Figure 6.1c shows an example of a phase difference profile with $K=2$ and $\Delta(2)=\Delta_x$. In general, the total transmittance function of the misaligned phase profile is the product of the transmittance function of the ideal phase profile and the transmittance functions of all $m$ phase difference profile for $m \geq 2$ as given by

$$U_{total}(x) = U_{ideal}(x) \times \prod_{m=2}^{K} U_a^m(x). \tag{6.3}$$

In the output plane we have the Fourier transform of the transmittance function

$$\Im\{U_{total}(x)\} = \Im\{U_{ideal}(x)\} * \Im\{U_a^2(x)\} * \cdots * \Im\{U_a^K(x)\}. \tag{6.4}$$

The Fourier transform of the transmittance functions $U_a^m(x)$ can be represented by

$$\begin{aligned}\Im\{U_a^m(x)\} = \delta(f_x) + \{2^{m-1}\alpha(m) \times \\ [(e^{j\phi_a(m)} - 1)\sum_p \delta(f_x - \frac{p}{T'_m})\mathrm{sinc}(K'_m)e^{-j\phi_1} \\ + (e^{-j\phi_a(m)} - 1)\sum_p \delta(f_x - \frac{p}{T'_m})\mathrm{sinc}(K'_m)e^{-j\phi_2}]\}.\end{aligned} \tag{6.5}$$

where $K'_m = p\alpha 2^{m-1}$, $\phi_1 = p\pi 2^{m-1}\alpha(m)$, and $\phi_2 = p\pi(2^{m-1}\alpha(m)+1)$. The $p^{th}$ order grating efficiency of the $\Im\{U_a^m(x)\}$ can be obtained by isolating the $p^{th}$ order terms in the summation of Eq. (6.5) and finding their squared magnitudes to obtain

$$\eta_a^p(m) = \begin{cases} \left|\delta(p) - 2^m \alpha(m)\mathrm{sinc}(K'_m)(1-\cos(\phi_a(m)))e^{-j\phi_1}\right|^2 & \text{if } p \text{ is even} \\ \left|\delta(p) - 2^m \alpha(m)\mathrm{sinc}(K'_m)(\sin(\phi_a(m)))e^{-j(\phi_1-\pi/2)}\right|^2 & \text{if } p \text{ is odd} \end{cases} \tag{6.6}$$

Since $\alpha(m)$ is equal to the value of alignment error divided by period $T$, we have $\alpha(m) << 1$. Thus, from Eq. (6.6), for $p \neq 0$, the magnitude of the $p^{th}$ order grating efficiencies of the $\Im\{U_a^m(x)\}$ are very small and can be neglected. The first order grating efficiency $\eta_{total}^1$ of $\Im\{U_{total}(x)\}$ can be approximated by the product of the first order grating efficiency $\eta_{ideal}^1$ of $\Im\{U_{ideal}(x)\}$ and zero order efficiencies $\eta_a^0(m)$ of all $\Im\{U_a^m(x)\}$, $m = 2,3,\ldots,K$ can be approximated by

$$\eta_{total}^1 \cong \eta_{ideal}^1 \times \prod_{m=2}^{K} \eta_a^0(m). \tag{6.7}$$

From Eq. (6.6), we have

$$\eta_a^0(m) = \left|1 - 2^m \alpha(m)[1-\cos(\phi_a(m))]\right|^2. \tag{6.8}$$

Using an example of 16-level grating profile ($K = 4$), we analyzed the effects of alignment errors of the $m^{th}$ mask for $m = 2,3,4$, assuming perfect etching depth. It was found that the layer with the largest etching depth and features ($m = 2$) is the most critical. In Table 6.1, the efficiency of $\eta_a^0(m)$ are calculated as a function of $\alpha(m)$. The efficiency is reduced significantly in the presence of alignment errors. This is intuitively clear, because small features are introduced that diffract light into high order spatial frequencies of the profile.

| | $\alpha = 0.01$ | $\alpha = 0.02$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
|---|---|---|---|---|
| m=2 | 0.922 | 0.846 | 0.64 | 0.36 |
| m=3 | 0.954 | 0.909 | 0.779 | 0.586 |
| m=4 | 0.976 | 0.952 | 0.882 | 0.771 |

Table 6.1: Zero order efficiencies of extra grating profiles in the last three layers of a 16-level grating profile are calculated as a function of $\alpha$, ratio of the alignment error.

If we assume that misalignment is the only source of error and require that the total efficiency loss should remain $< 5\%$, for a 16-level grating profile with a period length $T = 16\,\mu m$, the tolerances are $\Delta(2) \approx 0.1\,\mu m$ for the second mask alignment, $\Delta(3) \approx 0.2\,\mu m$ for the third mask alignment, and $\Delta(4) \approx 0.3\,\mu m$ for the fourth mask alignment.

## 6.3 Etching Depth Errors

It has been reported that the efficiency loss caused by an etching depth error of ±0.5% is negligible [62]. However, this result assumes that all levels share the same depth error. To be more precise, we again analyze the effect of etching depth errors by decomposing the depth errors of a $2^K$-level grating into its $m$ etching steps, $m = 1,2,\ldots K$. The $m^{th}$ etching error introduces an extra 1-D grating profile with a 50% duty cycle, a period length $T'_m = T/2^{m-1}$, and a phase delay $\phi_\varepsilon(m) = 2\pi \times \varepsilon(m)$, where $\varepsilon(m) = d_\varepsilon(m)/d$ is the error ratio. The transmittance function $U_\varepsilon^m(x)$ of this grating can be represented by

$$U_\varepsilon^m(x) = 1 + \{[(e^{j\phi_\varepsilon(m)} - 1) \times \text{rect}(\frac{x - T'_m/4}{T'_m/2})] * \sum_p \delta(x - pT'_m)\}. \quad (6.9)$$

Figure 6.2 illustrates a 2-level phase profile fabricated with an etching depth error $d_\varepsilon(m) = d_\varepsilon = \varepsilon(m) \times d$, where $d = \lambda/(n-1)$ is the total depth of the grating. Following the previous discussions, at the output plane, we have the Fourier transform of the transmittance function in the form

$$\Im\{U_{total}(x)\} = \Im\{U_{ideal}(x)\} * \Im\{U_\varepsilon^1(x)\} * \cdots * \Im\{U_\varepsilon^K(x)\}. \quad (6.10)$$

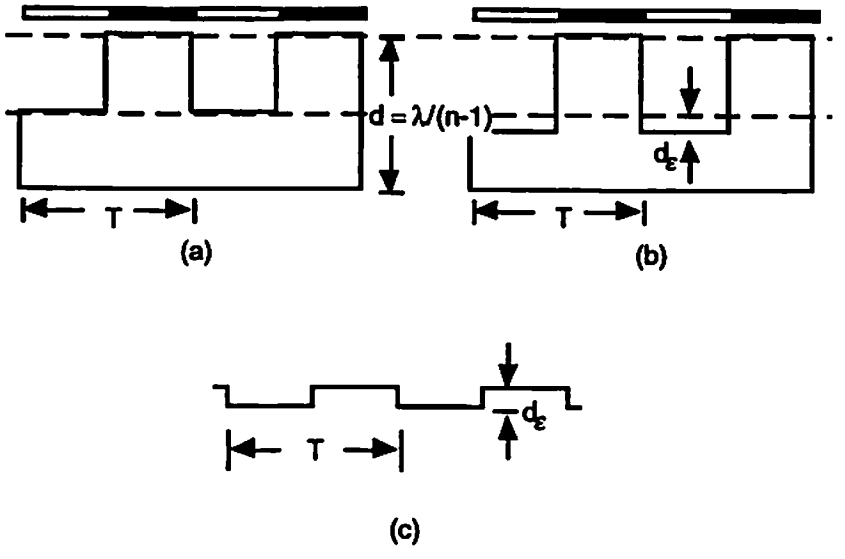



Figure 6.2: Effect of etching depth errors in the multi-level phase profile generation: (a) the ideal 2-level profile, (b) the profile obtained if an etching depth error with an amount of $d_\varepsilon$ is occurred, (c) the effect of the etching depth error is equivalent to adding a periodic grating in the phase domain with period $T$, duty cycle 50% , and a phase delay $2\pi d_\varepsilon/d$, where $d = \lambda/(n-1)$.

The Fourier transform of the transmittance functions $U_e^m(x)$ can be represented by

$$\Im\{U_e^m(x)\} = \delta(f_x) + \frac{1}{2}(e^{j\phi_e(m)} - 1)\sum_p \delta(f_x - \frac{p}{T'_m})\mathrm{sinc}(Q'_m)e^{-j\theta_1} \quad .(6.11)$$

where $Q'_m = p/2$, and $\theta_1 = p\pi/2$. The $p^{th}$ order grating efficiency of the $\Im\{U_e^m(x)\}$ is given by

$$\eta_e^p(m) = \left|\delta(p) + \frac{1}{2}(e^{j\phi_e(m)} - 1)\mathrm{sinc}(p/2)e^{-jp\frac{\pi}{2}}\right|^2. \qquad (6.12)$$

The zero order efficiency of the extra grating profile caused by the etching depth error of the $m^{th}$ mask are given by

$$\eta_e^0(m) = \frac{1}{2} + \frac{1}{2}\cos(\phi_e(m)). \qquad (6.13)$$

We observed that the zero order efficiency of the extra grating in the $m^{th}$ layer is affected only by the depth of the etching error and has no relation to the feature size of the masking. Thus, the depth errors in any layer affect the diffracted components in the same way regardless of the mask level. In Section 6.2 analyzing mask alignment errors, we assumed that the first order efficiency of the total grating is the product of the first order efficiency of the ideal grating and zero order efficiencies of all extra gratings. In our current analysis of the etching depth error, the phase profiles of the extra

gratings are overlaid together and will affect all diffraction orders in general. Thus, the approximation used in Eq. (6.7) is not valid for etching depth errors. Table 6.2 shows the numerical results of $\eta_\varepsilon^0$ of a single layer from Eq. (6.13) as a function of $\varepsilon(m)$, the error ratio.

| | zero order efficiency $\eta_\varepsilon^0$ |
|---|---|
| $\varepsilon(m) = 0.01$ | 0.999 |
| $\varepsilon(m) = 0.02$ | 0.996 |
| $\varepsilon(m) = 0.05$ | 0.976 |
| $\varepsilon(m) = 0.1$ | 0.904 |

Table 6.2: Numerical results of $\eta_\varepsilon^0$ of a single layer as a function of $\varepsilon(m)$, the error ratio.

## 6.4 Summary

In this chapter, we systematically analyze the efficiency loss of DOEs based on different fabrication errors. To summarize the effects caused by misalignments and etching depth errors, we consider the situation of a 16-level grating profile with $n = 1.5$ (glass) operated at a wavelength $\lambda = 0.8\,\mu m$. Thus, the depth $d$ of the profile is $1.6\,\mu m$. By choosing $T = 10d$, we have a period length of $16\,\mu m$. The effects of misalignment and etching depth errors is plotted in Fig. 6.3 as a function of error ratio proportional to $T$ and $d$, respectively. The curves in Fig. 6.3 show that misalignment errors cause more serious effects on output efficiency than etching depth errors. For an error ratio of 0.01, the misalignment error of the second mask

gratings are overlaid together and will affect all diffraction orders in general. Thus, the approximation used in Eq. (6.7) is not valid for etching depth errors. Table 6.2 shows the numerical results of $\eta_\varepsilon^0$ of a single layer from Eq. (6.13) as a function of $\varepsilon(m)$, the error ratio.

| | zero order efficiency $\eta_\varepsilon^0$ |
|---|---|
| $\varepsilon(m) = 0.01$ | 0.999 |
| $\varepsilon(m) = 0.02$ | 0.996 |
| $\varepsilon(m) = 0.05$ | 0.976 |
| $\varepsilon(m) = 0.1$ | 0.904 |

Table 6.2: Numerical results of $\eta_\varepsilon^0$ of a single layer as a function of $\varepsilon(m)$, the error ratio.

## 6.4 Summary

In this chapter, we systematically analyze the efficiency loss of DOEs based on different fabrication errors. To summarize the effects caused by misalignments and etching depth errors, we consider the situation of a 16-level grating profile with $n = 1.5$ (glass) operated at a wavelength $\lambda = 0.8\,\mu m$. Thus, the depth $d$ of the profile is $1.6\,\mu m$. By choosing $T = 10d$, we have a period length of $16\,\mu m$. The effects of misalignment and etching depth errors is plotted in Fig. 6.3 as a function of error ratio proportional to $T$ and $d$, respectively. The curves in Fig. 6.3 show that misalignment errors cause more serious effects on output efficiency than etching depth errors. For an error ratio of 0.01, the misalignment error of the second mask

causes an 8% efficiency loss while etching depth error causes less than 0.1% efficiency loss. Furthermore, an error ratio of 0.01 in misalignment is equivalent to an alignment error of 0.16 $\mu m$, an alignment which is hard to achieve by standard fabrication procedures. On the other hand, an error ratio of 0.01 in etching depth error is equivalent to a depth of 16 $nm$. In etching, an etching depth accuracy of several nanometers can be achieved using standard reactive ion etching techniques. Thus, the constraint of etching depth error is easy to achieve.

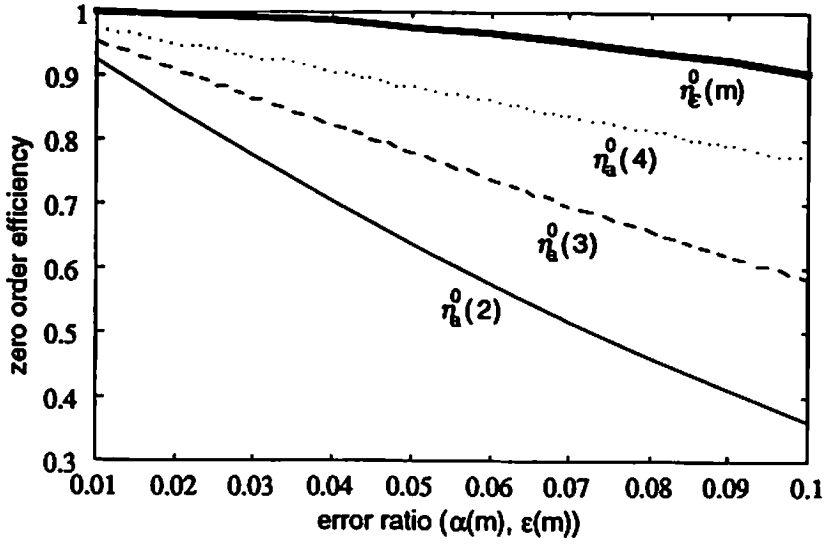



Figure 6.3: The effects of misalignments and etching depth errors is plotted as a function of error ratio proportional to $T$ and $d$, respectively.

Our analysis in the previous paragraphs is based entirely on paraxial Fourier optics. When the feature size of the diffractive element becomes comparable to the wavelength, the accuracy of this theory is reduced. For example, misalignment introduces extra features in the grating profile, whose dimensions are typically of the order of the wavelength of light. The

effects of such small features can be analyzed accurately only by using the rigorous diffraction theory of gratings [65]. However, we believe that the present error analysis provides an adequate estimation of the effects of fabrication errors in a systematic manner.

Overcoming the serious alignment problem in fabrication of multi-level DOEs appears to require that the alignment is performed with the electron-beam pattern generator, thus discarding the use of separate masks. Alternatively, a one-stage procedure based on electron-beam direct write on resist can be used [66, 67, 68]. The main difficulty in direct writing is the control of the relief depth. Moreover, all direct-write techniques are expensive.

# Chapter 7

# Optoelectronic Compare-and-Exchange Modules

A pipelined multistage sorting network needs, in addition to the fixed interconnections between each stage, a self-routing compare-and-exchange (C&E) module whose routing decisions are made by examing the relative magnitude of the local information. Let us consider the packet switching scheme for a sorting network. A stream of binary input data is divided into packets of fixed length, containing destination information (header) multiplexed in the time domain as shown in Fig. 7.1. The input data streams first pass through a preprocessor to synchronize their time slots, then, the synchronized inputs are sent into the C&E module as shown in Fig. 7.2.

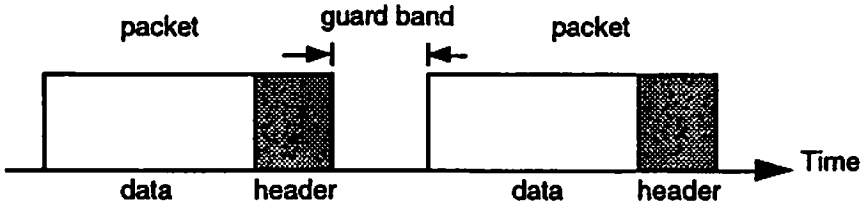



Figure 7.1: Data format for packet switching scheme.

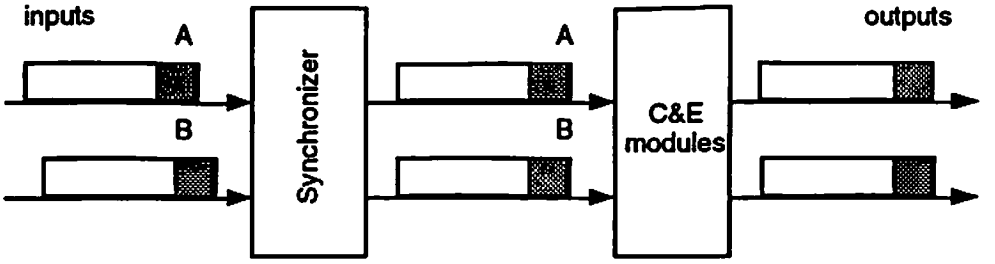



Figure 7.2: Synchronized packet switching system.

Two binary headers $A = A_{n-1}A_{n-2}\ldots A_0$ and $B = B_{n-1}B_{n-2}\ldots B_0$ are sent into the C&E module serially from the most significant bit ($A_{n-1}$, $B_{n-1}$) to the least significant bit ($A_0$, $B_0$), The two channels are compared bit by bit. The output data streams depend on the result of the comparison and the type of module. The most commonly used module type in network design is probably the 2 input/ 2 output compare-and-exchange switching modules shown in Fig. 7.3.

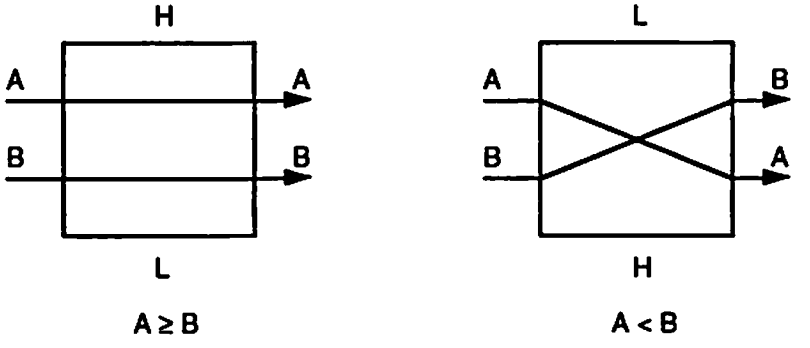



Figure 7.3: The 2 input/ 2 output compare-and-exchange switching module.

Assuming that headers $A$ and $B$ are sent into the "high" input port ($H$) and "low" input port ($L$) of a C&E module respectively, the operation rules of the C&E module [50] can be described as follows:

**Rule 1:** If $A \geq B$, then $A$ is sent to "high" output port $H$ and $B$ is sent to "low" output port $L$. The C&E module is latched in the *bypass* configuration (non-exchange).

**Rule 2:** If $A < B$, then $A$ is sent to "low" output port $L$ and $B$ is sent to "high" output port $H$. The C&E module is latched in the *exchange* configuration.

According to the operation rules, a C&E module should perform three basic operations: comparison, where the calculation of an exchange control signal occurs; exchange, where the data switching operation takes place; and latching, where the exchange switch holds a particular configuration while data passes. This state must be held until a reset signal is received. Generally speaking, a C&E module with single functionality is not feasible for the multistage sorting networks discussed in previous chapters. For example, multistage bitonic sorting networks using perfect shuffle interconnection links need three types of C&E modules to perform bypass, maxsort, and minsort operations in the node stages. In this chapter, we focus on the design of a multifunctional C&E module.

In the following sections, we focus on the multifunctional 2 input/ 2 output C&E module designs based on three types of state-of-the-art devices: L-SEEDs, FET-SEEDs, and OEICs. Design techniques utilizing digital electronic logic theorems will be discussed first. After that, we give

detail circuit designs for C&E modules based on different device characteristics. Finally, we will analyze the performance and packing density of the modules.

# 7.1 Design Techniques

## 7.1.1 Truth Table Minimization

There are many techniques available to design optical comparison circuits [69]. The most straightforward way is to use well-developed digital electronic truth table minimization techniques. Figure 7.4 is a design for a compare-and-exchange unit using conventional NAND gate logic.

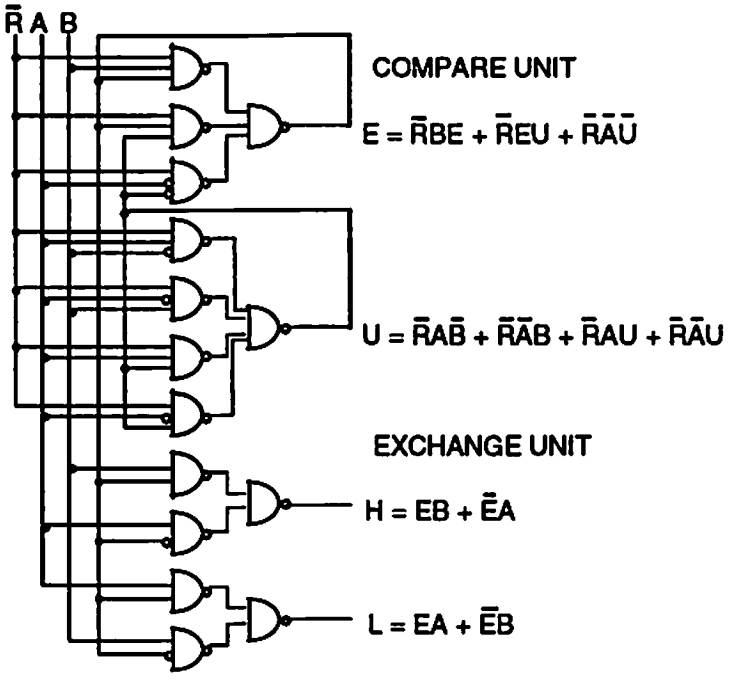
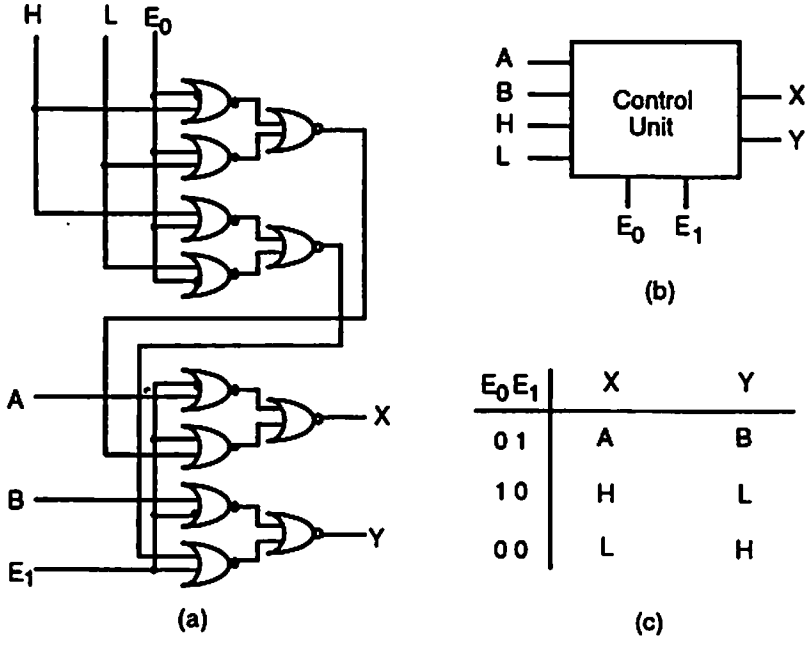



Figure 7.4: Compare-exchange NAND gate logic based on truth table minimization techniques.

Assuming that we know the length of headers $A$ and $B$, in Fig. 7.4, headers $A$ and $B$ are two input bit streams sent into the compare-exchange unit, $\overline{R}$ is a module enable signal, $E$ is a switch enable signal which will be "high" only if $A < B$ occurs, and $U$ is a latching signal which locks the status of signal $E$ if $A < B$ occurred. If $E$ is "high", the unit sends $B$ to output $H$ and $A$ to $L$; if $E$ is "low", it sends $A$ to $H$ and $B$ to $L$.

The control unit is designed to receive the outputs of the compare-exchange unit and two input bit streams and perform various output functions based on control signal inputs as shown in Fig. 7.5. Here, two control signal inputs $E_0$ and $E_1$ are required in our design. If $E_0E_1$ is 01, the module operates as a *bypass* switch; if $E_0E_1$ is 10, the module operates as a *maxsort* switch; if $E_0E_1$ is 00, the module operates as a *minsort* switch.



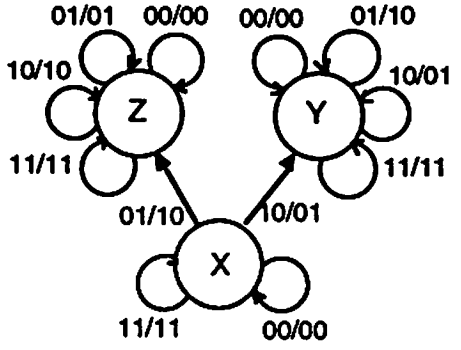

| $E_0E_1$ | X | Y |
|---|---|---|
| 01 | A | B |
| 10 | H | L |
| 00 | L | H |

(c)

Figure 7.5: control unit based on truth table minimization. (a) NOR gate logic, (b) block diagram, and (c) operation table.

### 7.1.2 Finite State Machine

Any computable function can be transformed into a set of Boolean equations and described by a finite state machine [70]. A finite state machine can be represented as a black box with a finite number of internal variables that is connected to the rest of the system through input and output lines. The output at any given time is completely determined by the current input and internal state. A complete representation of a finite state machine can be given by a state transition diagram which shows all possible states and the transitions from each state for every possible input.




(a)

| | next state / min(A,B), max(A,B) | | | |
|---|---|---|---|---|
| current state / $S_0S_1$ | $\bar{A}\bar{B}$ 0 0 | $\bar{A}B$ 0 1 | $A\bar{B}$ 1 0 | $AB$ 1 1 |
| X : 00 | X/00 | Z/01 | Y/01 | X/11 |
| Y : 01 | Y/00 | Y/10 | Y/01 | Y/11 |
| Z : 10 | Z/00 | Z/01 | Z/10 | Z/11 |

(b)

Figure 7.6: State transition diagram (a) and state transition table (b) for a compare-exchange unit. The label pairs in the state transition diagram represent the relation between the logic state of the two inputs $A$ and $B$ and two outputs $\min(A,B)$ and $\max(A,B)$.

Figure 7.6a shows the state transition diagram for a compare-exchange unit. There are two input lines, on which two data streams $A$ and $B$ enter, most significant bit first. There are also two output lines, one is $\min(A,B)$ and another is $\max(A,B)$. The state machine has three internal states, $X$= *waiting*, $Y$ = *bypass*, and $Z$ = *exchange*. We represent these three states

using two binary bits $S_0$, $S_1$. The machine will be in state $X$ if there is no difference between two inputs. If $A > B$ occurs, the machine will change to state $Y$ and stay there. If $A < B$ occurs, it will change to state $Z$ and remain in that state. A mathematically equivalent state transition table shown in Fig. 7.6b is derived from the state transition diagram. The Boolean equations for $S_0$, $S_1$, max($A$,$B$), and min($A$,$B$) can be derived and used to realize the state machine using NOR gate logic as shown in Fig. 7.7.

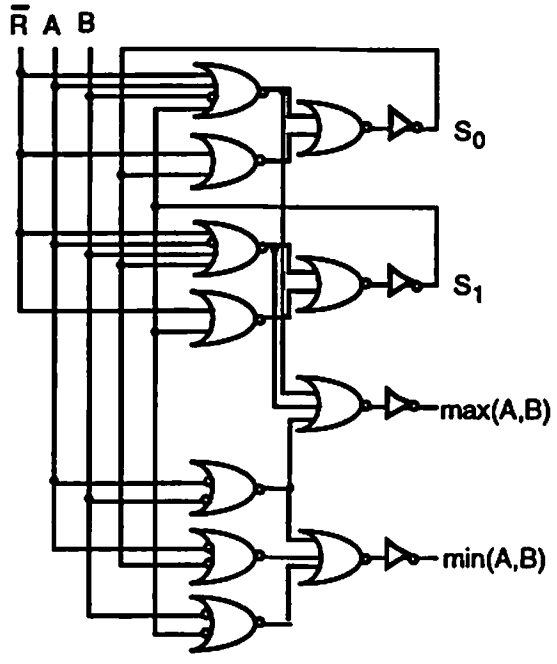



Figure 7.7: NOR logic implementation of the state machine of the compare-exchange unit.

### 7.1.3 Latching Logic

Optical bistable devices that inherently perform the latching operation are ideal candidates to design the comparison circuit in order to reduce the spatial complexity of the implementation.

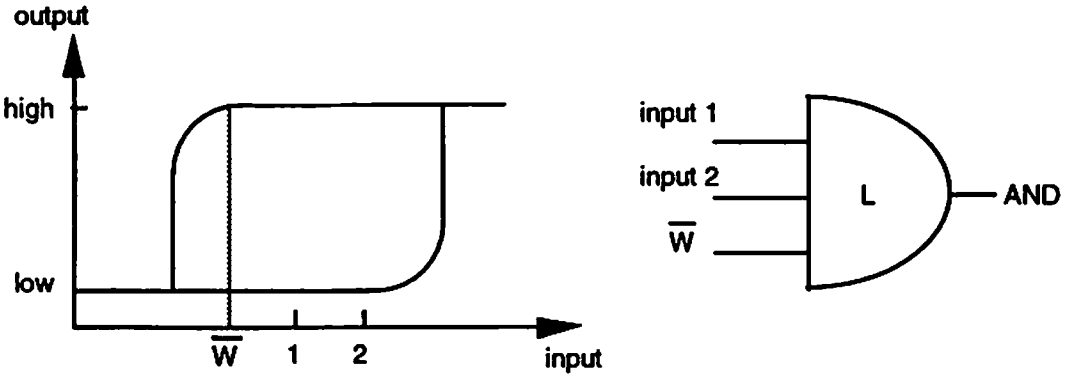



Figure 7.8: Transfer function of a latching AND gate.

A latching AND gate is shown in Fig. 7.8. When both inputs are in a "high" power status, the output of the AND gate will be set to its "high" status which triggers the bias signal $\overline{W}$. The bias signal $\overline{W}$ provides enough power to keep the output of the AND gate in its "high" state. This situation will not change with the removal of the inputs except the bias signal $\overline{W}$, therefore, the gate is effectively latched. Not showing here is a reset line which we discess later.

The comparison logic based on two stages of latching AND gates is shown in Fig. 7.9. When the most significant mismatch of the bit streams is found, one of the first-stage logic gate latches into the "high" state while another remains unlatched. The second-stage gate will receive the first change signal and latch into appropriate switch configuration to prevent later bit comparisons from altering the exchange setting.

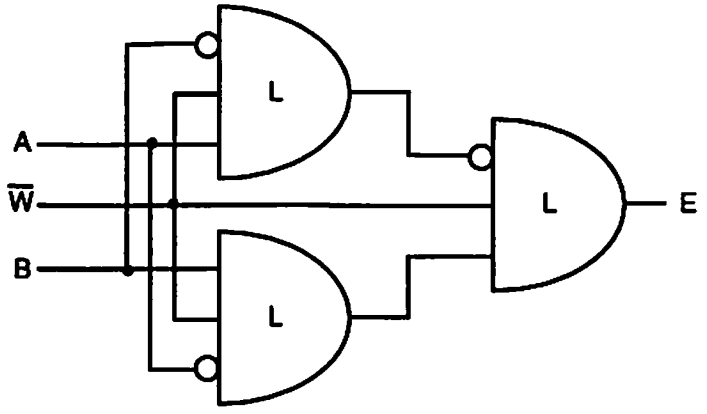



Figure 7.9: Latching AND gate implementation of the comparison unit.

## 7.2 C&E Module Designs

There are many different technologies available for fabricating a smart pixel array which is capable of performing compare-and-exchange operations. The basic functions of such pixel are to receive optical data, process the data, and then send out the final result using an optical format. The receiving part of the pixel can be a detector or modulator based circuitry, the processing part can be either an electronic circuit or a nonlinear optical circuit, and the output part can be a modulator or an optical source. In this section, we investigate three types of state-of-the-art device technologies: L-SEEDs, FET-SEEDs, and OEICs. The performance of the C&E modules based on three technologies will be given.

### 7.2.1 L-SEEDs

Arrays of symmetric self electro-optic effect devices (S-SEEDs) have been made with low operating energy and fast switching speeds [71, 72]. This

type of device, which acts simultaneously as a detector and modulator, has the characteristics of a set-reset latch as shown in Fig. 7.10, and can be made to implement logic functions [73]. SEEDs logic gates (L-SEEDs) that can perform more complicated functions have been realized by using electrically connected multiple quantum wells (MQWs) p-i-n diodes configured like transistors in NMOS or CMOS circuits together with an output S-SEED to provide the output light beams. An integrated array of these logic gates has been successfully demonstrated [74].

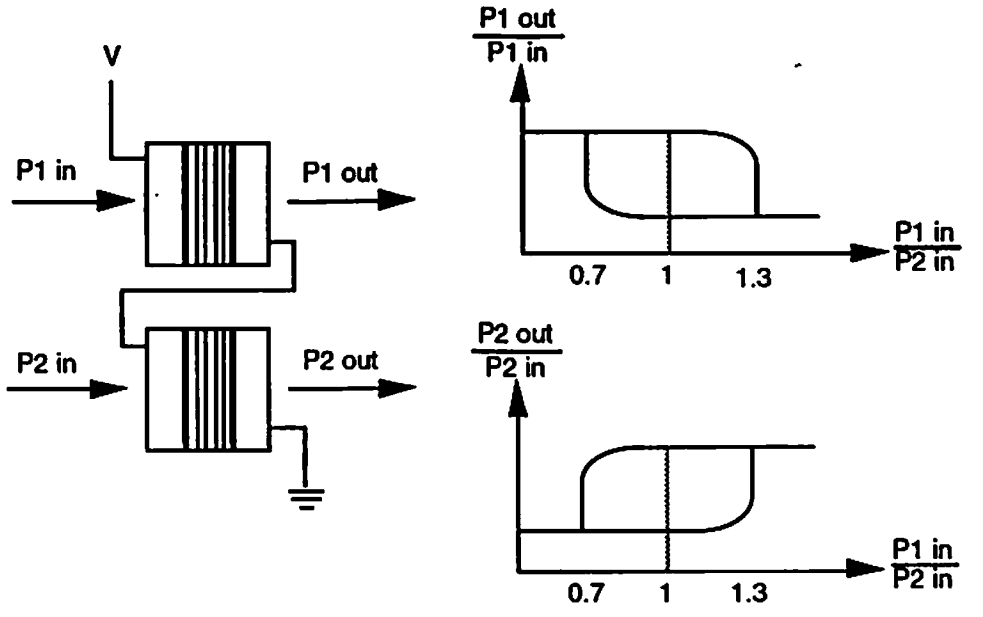



Figure 7.10: Symmetric SEED and its input/output characteristic curve.

Smart pixels utilizing modulator based L-SEEDs require external optical clock beams to read the state of each logic output in the circuitry. All of the data transformation within the pixel is carried out by optical signals.

Electronic power is used only to bias the circuitry. Figure 7.11 shows the block diagram of a multifunctional C&E module based on L-SEEDs. A control unit is built into the C&E module in order to set up different functions upon request. Due to the finite response time of the compare-and-exchange operation in the C&E module, a shift register or buffer may be required to hold the data information until the switch signal is generated. In the following sections, we describe how to design latching AND gates, exchange switches, control circuits and shift registers based on L-SEEDs.

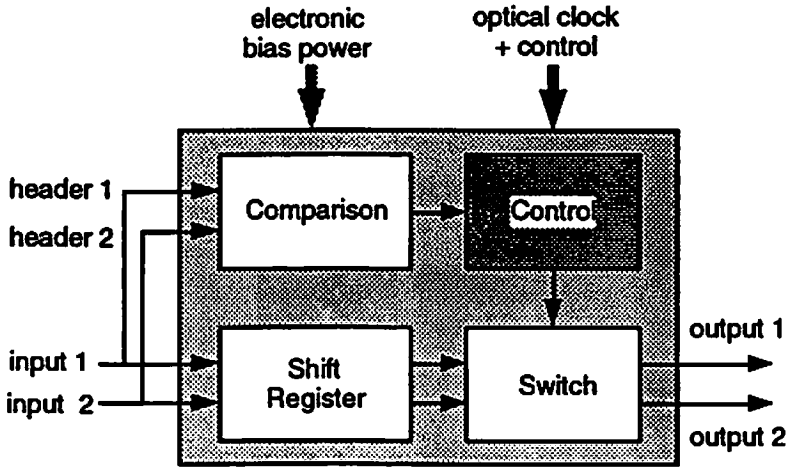



Figure 7.11: Block diagram of a multifunctional C&E module based on L-SEEDs.

#### 7.2.1.1 Comparison Unit

The central element of the comparison unit is the latching AND gate. A typical circuit design based on L-SEEDs is shown in Fig. 7.12a. First a high power optical reset signal $C$ is sent into the circuit. Due to this signal, initially more current will flow through the MQW p-i-n diodes in the upper

half of the circuit than those in the lower half. The node voltage $V_n$ will switch toward $V_0$ if we operate the S-SEEDs at a wavelength where the absorption is less at high voltage. When the clock beams are applied, the output power of $S$ will be "low" and $R$ will be "high". If, after one clock cycle, the reset signal $C$ is set to a low power level, the outputs of the circuit will depend on the power level of the input, which arrives at the same interval as reset signal $C$. If the power level of the input is "low" (or the same as $C$), current which flows through the upper and lower half of the circuit will be the same, the outputs of the circuit remain $S$ = "low" and $R$ = "high". When the input is "high" or greater than $C$, more current will flow through the lower half of the circuit and change the state of the S-SEED. The outputs of the circuit are then latched to $S$ = "high", $R$ = "low" and can only be changed by the reset signal $C$. The timing diagram of the latching circuit is shown in Fig. 7.12b.

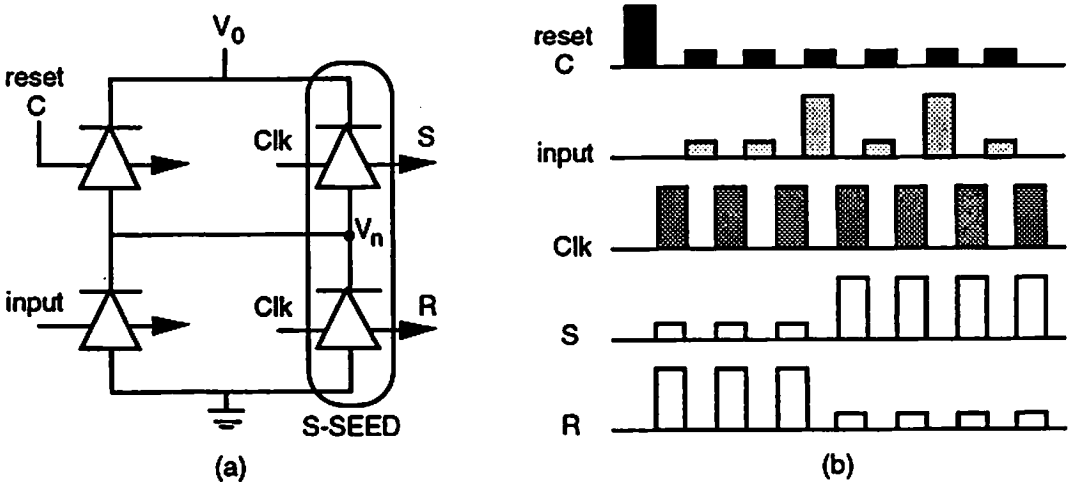



Figure 7.12: (a) Latching circuitry based on L-SEEDs; (b) Timing diagram of the L-SEEDs latching circuitry. The diode symbol in the latching circuit represents MQW diodes as in Fig. 7.10.

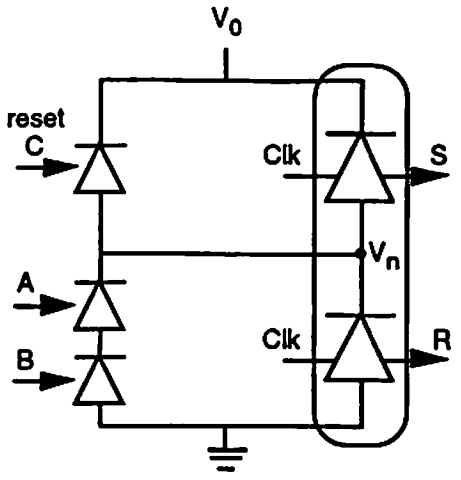



Figure 7.13: Latching AND gate based on L-SEEDs.

Based on the same idea, a design of a latching AND gate utilizing L-SEEDs is shown in Fig. 7.13. If we assume that the contrast ratio of all input beams is identical, when the reset signal *C* is "high", current in the upper half of the circuit will always be larger than that in the lower half of the circuit regardless the states of the input beams *A* and *B*. However, when *C* is set to its low level, only *A*=*B*="high" can change the state of the S-SEED. Therefore, the outputs of this circuit are

$$S = \mathrm{L}[AB], \quad R = \mathrm{L}[\overline{AB}] \tag{7.1}$$

where **L[ ]** represents the latching operation.

Because logic circuits based on L-SEEDs require differential digital signal inputs, we introduce the dual-rail signal generating circuit shown in Fig. 7.14. It translates the input signals into differential signal pairs and

provides inverse operations. In this circuit, the power of the reference signal is set to be 50% of the maximum input signal. Because signals go through the dual-rail signal generators before they enter the sorting networks, they remain in that form as they proceed through the C&E modules.

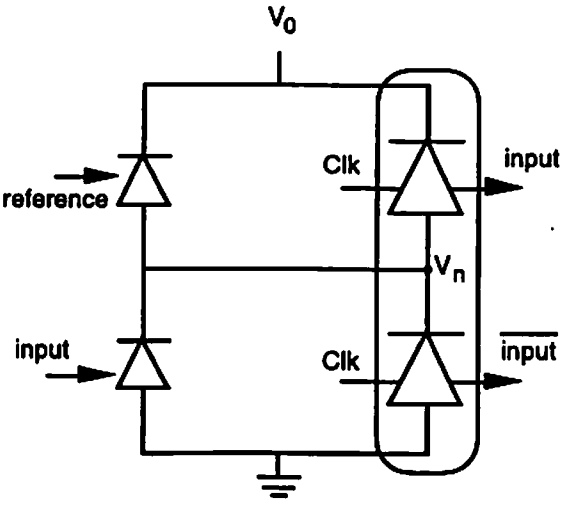



Figure 7.14: Dual-rail signal generator for L-SEEDs circuitry.

A complete design of the comparison unit is shown in Fig. 7.15. This circuit consists of two latching AND gates in the first stage and one latching AND gate combined with an inversion operation in the second stage. We obtain the comparison output $E$ given by

$$E = (\overline{\overline{CBA}})\overline{C}(\overline{CB}\,\overline{A}). \tag{7.2}$$

If $A > B$ occurs before $B > A$, the outputs of the first stage will be $S_1$="high" and $S_2$="low", and the comparison output $E$="low". Conversely, If $B > A$ occurs first, then the output $E$="high".

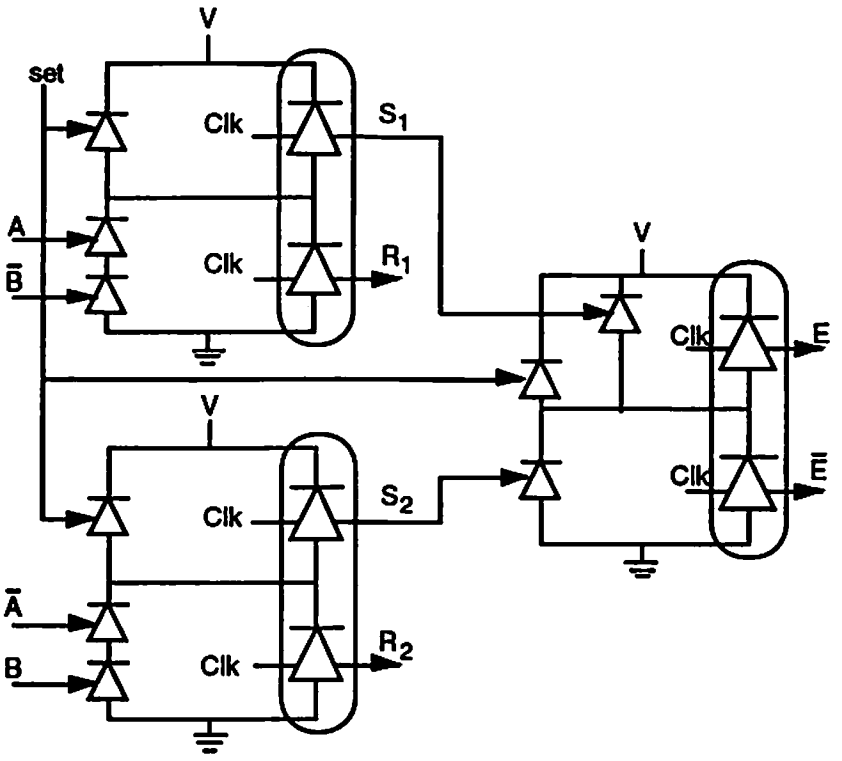



Figure 7.15: The comparison unit based on L-SEEDs circuits.

### 7.2.1.2 Control Unit

Theoretically, circuits that can perform $M$ different logic operations need at least $N$ binary control lines where $N$ is equal to the smallest integer greater than or equal to $\log_2 M$. Here, three different operations need to be built into a C&E module, that is, $M$=3. Therefore, the control unit needs at least two control inputs. The logic circuit for the control unit is shown in Fig. 7.16a. The output $SW$ can be written as

$$SW = (\overline{Set_1} \bullet E \bullet Set_2) + (\overline{Set_1} \bullet E \bullet \overline{Set_2}). \qquad (7.3)$$

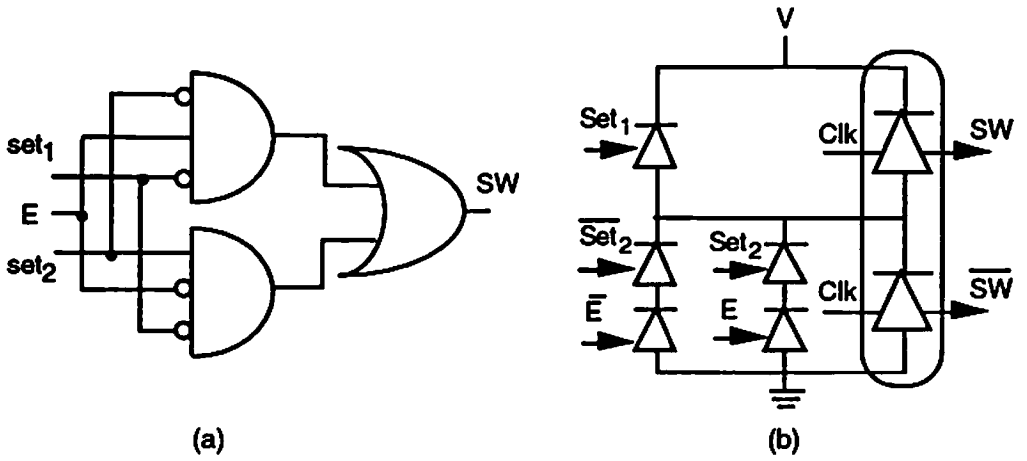



Figure 7.16: (a) Logic circuit for the control unit; (b) The circuit of the control unit based on L-SEEDs.

The circuit design of the control unit is shown in Fig. 7.16b. If we keep the control input $Set_1$="high" all the time, the C&E module will operate as a bypass element. Otherwise, we can switch the power level of $Set_1$ to the "low" state after one clock cycle, and the outputs will be decided by the status of the second control input $Set_2$. The functions of the control circuit are shown in Table 7.1.

| *control* | | *input* | *output* | |
|---|---|---|---|---|
| $Set_1$ | $Set_2$ | | $SW$ | $\overline{SW}$ |
| high | \ | \ | low | high |
| low | high | $E$ and $\overline{E}$ | $E$ | $\overline{E}$ |
| low | low | $E$ and $\overline{E}$ | $\overline{E}$ | $E$ |

Table 7.1: The functions of the control circuit. "\" represents the don't care condition.

### 7.2.1.3 Shift Register

The ability of the S-SEED to be used as a S-R latch was first demonstrated by [75] in their implementation of an all-optical shift register. In their experiment, four S-SEED latches were cascaded in a master-slave configuration to form a two-bit optical shift register. A modified three-stage L-SEEDs shift register design which provides input-output isolation is shown in Fig. 7.17. The stage number of the shift register depends on the time delay of the compare-and-exchange operation of the module. Here, a shift register with three stages has a delay time of 1.5 clock cycles.

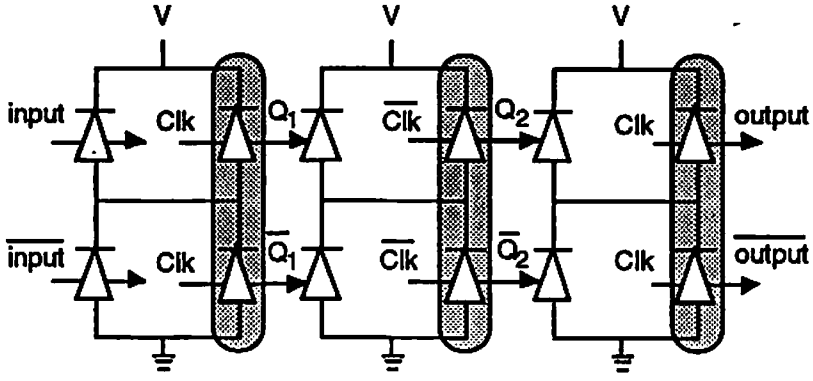



Figure 7.17: A three-stage shift register based on L-SEEDs.

### 7.2.1.4 Exchange Unit

The function $E = AB + CD$ which implements a switching node with two data inputs and one data output (a 2 input/ 1 output node) has been successfully demonstrated using L-SEEDs circuits [76]. Based on the same structure, a 2 input/ 2 output switch was recently proposed [77]. A modified version of this switch structure can be used as a exchange circuit shown in

Fig. 7.18. The input data streams $A$ and $B$ and switch signals $SW$ and $\overline{SW}$ of the exchange unit come from the shift register and the control unit, respectively. Switch signal $SW$ determines the route that data streams $A$ and $B$ pass through. If $SW$="low", then $H$=$A$ and $L$=$B$; if $SW$="high", then $H$=$B$ and $L$=$A$. Mathematically these two outputs can be described as

$$H = A \bullet \overline{SW} + B \bullet SW$$
$$L = A \bullet SW + B \bullet \overline{SW}. \qquad (7.4)$$

The outputs $H$ and $L$ of the exchange unit are the final outputs of the C&E module. In a multistage sorting network, these signals will pass through the interconnection links to the next stage.

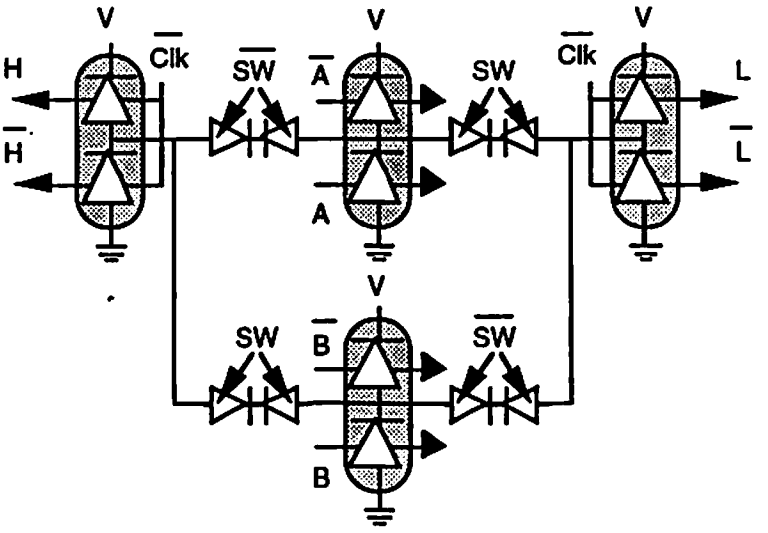



Figure 7.18: The exchange unit based on L-SEEDs.

### 7.2.1.5 Analysis

The combination of four units described in the previous sections forms a complete multifunctional C&E module. In this section, we analyze the following:

**Timing:** We need master clock beams *Clk* and its complement $\overline{Clk}$ to cascade the signals from one stage to another. Two control beams and one reset beam are also required to set up the modules. The timing diagram of the C&E module is shown in Fig. 7.19. The new L-SEEDs-based C&E module has a constant delay consisting of one clock cycle reset time and 2 clock cycles operation time.

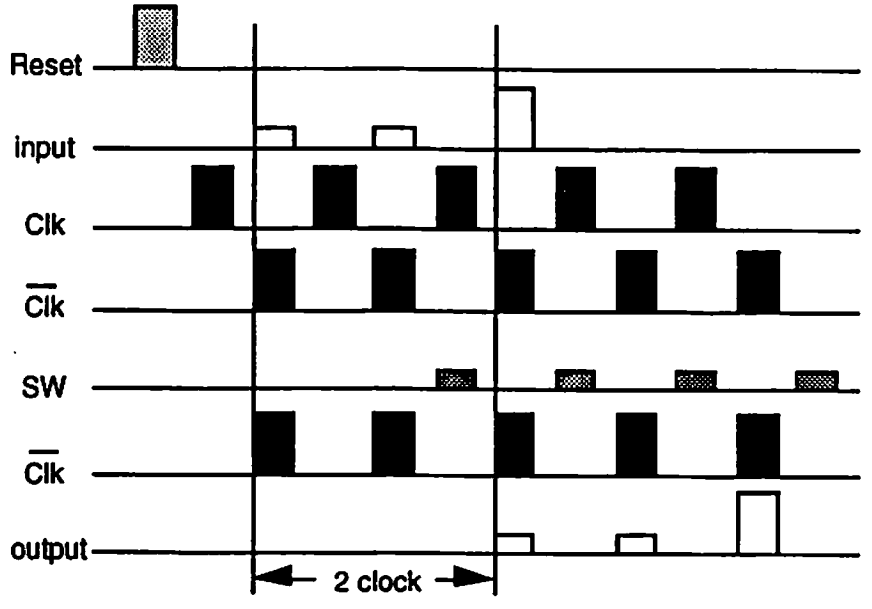



Figure 7.19: Timing diagram of the L-SEEDs C&E module.

**Design constraints:** If we assume that the contrast ratio of all input beams is identical, in Fig. 7.15, when the power level of both inputs at the second stage of the comparison unit is "high" and clock beam is the "low", the power contrast ratio that keeps the output status unchanged is given by

$$P_{high} > 0.7 \times (P_{high} + P_{low}), \tag{7.5}$$

where 0.7 is the lower bound of the characteristic curve of the S-SEED. From Eq. (7.5) we find that the necessary minimum contrast ratio of the signal beams applied to this circuit is 2.4 : 1. This requirement is easy to achieve according to recent technical reports [74]. Another constraint comes from the exchange unit shown in Fig. 7.18, in which the power of the input signals must be greater than the lesser of the switching signals, otherwise the switch beams effectively "short" the input and output S-SEEDs together.

**Operation speed:** Though it has been estimated that the maximum operation speed of SEEDs based circuitry can achieve 100 MHz [78], the maximum switching speed of the module is currently limited by the available laser power and the loss of the optical systems. We can predict the operation rate by the equations

$$\begin{aligned} T_{sw} &= \left(\frac{r_c + 1}{r_c - 1}\right) \times (E_{sw} \times N_d \times N_c) / (P \times (1 - L_{sys})) \\ R_{op} &= 1 / (2 \times T_{sw}) \end{aligned}, \tag{7.6}$$

where $T_{sw}$ is the switching time in second, $R_{op}$ is the operational rate in bits/second, $r_c$ is the contrast ratio, $E_{sw}$ is the device switching energy in joules, $L_{sys}$ is the percentage of the loss of the optical system, $P$ is the laser power in watts, $N_c$ is the number of input channels, and $N_d$ is the maximum number of the devices that need power from the same source. According to current technologies, devices with 1 pJ switching energy and 3:1 power contrast ratio are achievable. In our design, $N_d$ is 10 for one C&E module. If we have a system which has a 90% optical transmission loss, the relations between operational rate, laser power, and number of channels supported by this system is shown in Fig. 7.20.

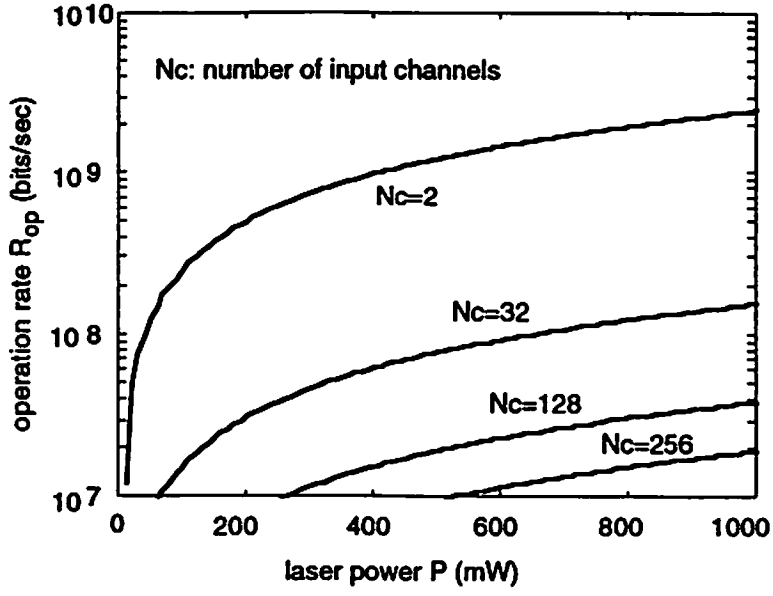



Figure 7.20: Relations between operation rate, laser power, and number of input channels in an L-SEEDs-based sorting system.

### 7.2.2 FET-SEEDs

The FET-SEED is a new technology which use several lithographic processing steps to integrate field effect transistors, diodes, multiple quantum well modulators and detectors on a GaAs based substrate [79]. Because the FETs have been fabricated with the MQW modulators and detectors for gain and logic functions, the FET-SEED-based smart pixel offers two main improvements over L-SEEDs circuits. First, the input/output energy requirement is decreased, and circuitry can operate at higher speed under the same laser source power. Second, as the circuitry becomes more complex, L-SEEDs need many optical beams (one per MQW diode) to operate logic functions, so they have a beam complexity proportional to the circuit complexity. Using FET logic greatly reduces the optical beam complexity of the FET-SEED circuitry. Based on current state-of-the-art device technology, a switching energy of 40 fJ and a 650 MHz operation rate has been achieved [80]. Theoretically operation at GHz rates is possible. For batch fabrication, a $4 \times 4$ array of FET-SEED based switching node has been demonstrated [81].

Figure 7.21 shows the block diagram of a multifunctional C&E module based on FET-SEEDs. The optical input signal is detected by an input unit and translated into an electronic signal for subsequent processing. The control unit receives external commands to set up the operation function of the C&E unit. The output unit changes the states of the MQW modulators

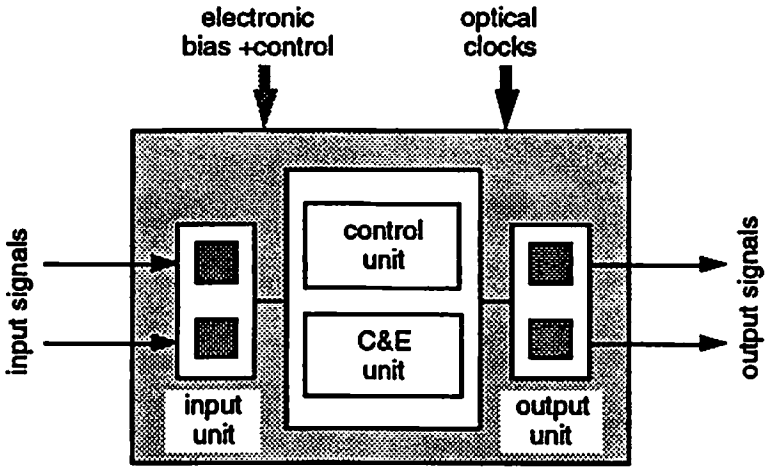



Figure 7.21: Block diagram of a multifunctional C&E module based on FET-SEEDs.

based on processed results, and a set of external optical beams is used to read out the data. The bias of the logic circuits and external control signals of the C&E module are provided by electronic power. The optical power is used only as a read out source.

In a FET-SEED based C&E module, the control unit can be realized using the NOR gate logic circuit shown in Fig. 7.5. Also, the compare-exchange unit can be designed as a finite state machine shown in Fig. 7.7. The contrast ratio of the FET-SEED modulator is low (10:1), thus, the input optical signal and modulated output signal are both represented in a dual-rail format. Figure 7.22 shows one possible circuit design for an input unit, output unit, inverter, and NOR gate based on FET-SEEDs. Diodes in the inverter and NOR gate circuit are used to provide the correct output voltage level.

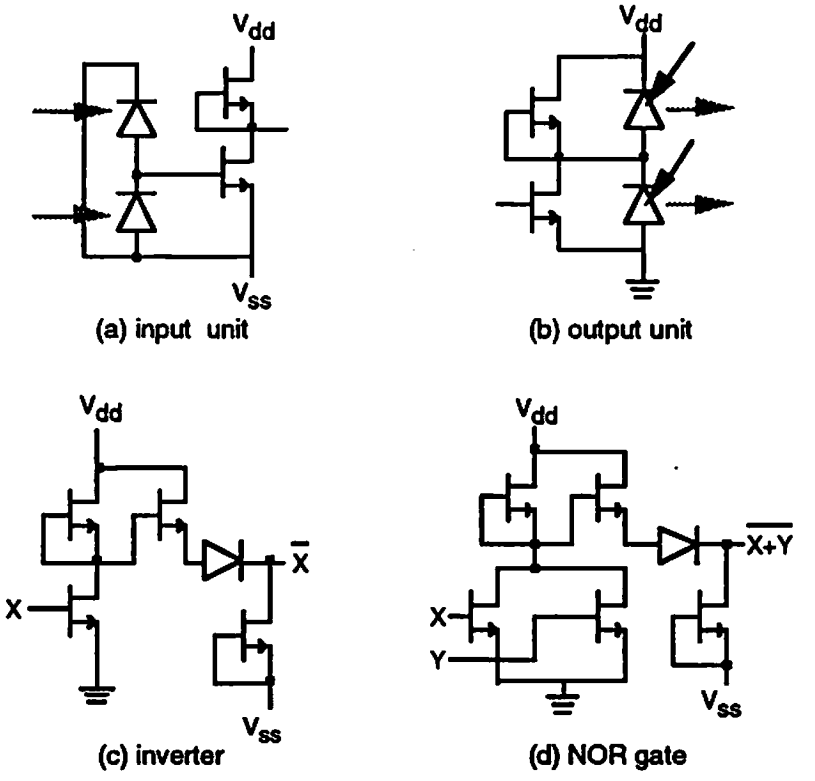



Figure 7.22: Schematic diagrams of FET-SEED circuits; (a) input unit, (b) output unit, (c) inverter, and (d) 2 input/ 1 output NOR gate.

The input unit of each channel includes a pair of MQW detectors and a FET receiver circuit as shown in Fig. 7.22a. The output unit shown in Fig. 7.22b has a pair of MQW modulators biased by FET transmitter circuits. A pair of equal intensity optical beams shine on these modulators and read out the result in a dual-rail format. The inverter circuit in Fig 7.22c is used to provide the inverse logic state of the signal. A 2 input/ 1 output buffered FET NOR gate shown in Fig. 7.22d is used for the control unit and C&E unit. Note that, based on the circuit design, three different levels of electronic bias $V_{dd}$, $V_{ss}$, and *Gnd* are required.

### 7.2.3 OEICs

In this section, we define optoelectronic integrated circuits (OEICs) as a type of optoelectronic device in which separate photodetectors, electronic transistor circuits, and optical sources (laser or LED) are built together. A wide range of fabrication techniques and device technologies are under study for laser-photodetector-based OEICs, and several smart pixel circuits for particular applications have been demonstrated [82, 83, 84].

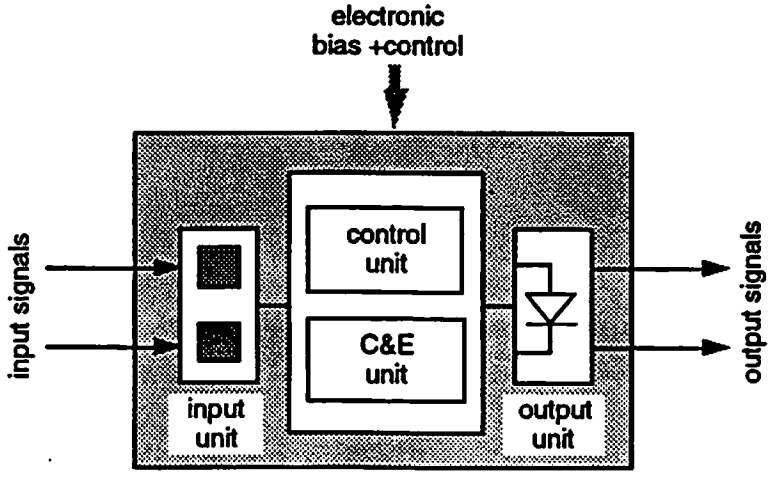



Figure 7.23: Block diagram of a multifunctional C&E module based on OEICs.

Figure 7.23 shows the block diagram of a multifunctional C&E module based on OEICs. The optical input signal is detected and then converted to an electronic signal. The information embodied in the electronic signal is then processed by conventional electronic C&E circuits and subsequently reconverted to an optical signal by using an output light source. Through

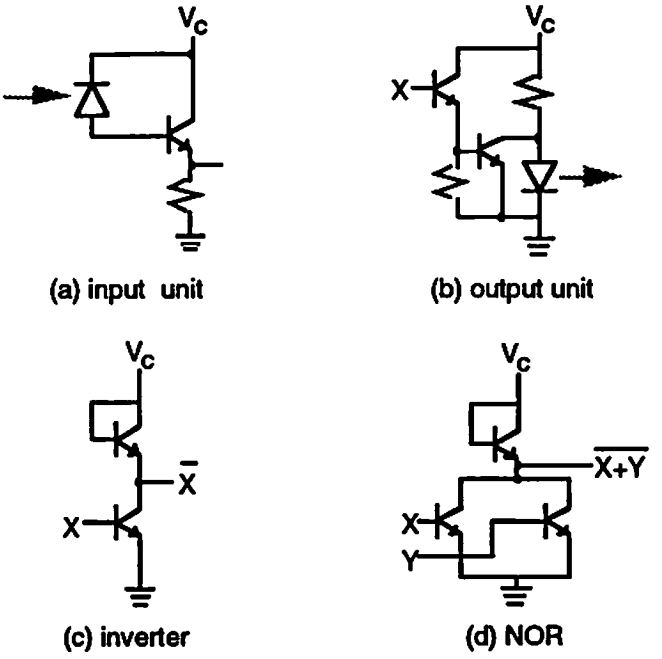



Figure 7.24: Schematic diagrams of OEICs; (a) input unit, (b) output unit, (c) inverter, and (d) 2 input/ 1 output NOR gate.

this optical-electronic-optical cycle, it is the current of the light source that is modulated. In summary, the bias of the logic circuits and external control signals of the C&E module is provided by electronic power, and no optical power is required.

In an OEICs based C&E module, the control unit and control unit can be realized using the same NOR logic designs as those for FET-SEED based C&E module. Figure 7.24 shows one possible circuit design of an input unit, output NOR gate, inverter, and NOR gate based on OEICs. The input unit of each channel includes a photodetector and a heterojunction bipolar transistor (HBT) based circuit as shown in Fig. 7.24a. The output unit shown in Fig. 7.24b has a pair of HBTs to modulate the optical source. The

inverter circuit in Fig 7.24c is used to provide the inverse logic state of the signal. A HBT based 2 input/ 1 output NOR gate shown in Fig. 7 24d is used for the control unit and C&E unit. Based on the circuit design, two different levels of electronic bias $V_c$, and *Gnd* are required.

### 7.2.4 Performance

Table 7.2 summarizes the main characteristics of the C&E module design based on three types of device technologies.

| smart pixel technology | L-SEEDs | FET-SEEDs | OEICs |
|---|---|---|---|
| circuitry design | latching logic | finite state logic | finite state logic |
| switch delay | 2 clock cycles | gate delay | gate delay |
| on chip buffer | yes | no | no |
| control scheme | optical | electronic | electronic |
| data processing | complex (optical) | simple (electronic) | simple (electronic) |
| fabrication | simple | moderate | complex |
| channel bandwidth | 100 MHz | > 1 GHz | > 1 GHz |
| wavelength sensitivity | high | high | low |
| multiple wavelength application | no | no | yes |

Table 7.2: Summary of the performance of the C&E module design based on three device technologies.

OEIC based technology has two main advantages over modulator based devices like L-SEEDs or FET-SEEDs. First, systems based on OEICs do not

require external optical sources and their associated hardware. Thus, the complexity of this part of the system can be greatly reduced. Second, unlike wavelength-sensitive MQW modulators, source-based OEICs can transmit light using tunable laser or several sources with different wavelengths, which makes multi-wavelength applications possible. However, modulator based devices have the advantages of low power consumption and relatively easy fabrication.

## 7.3 Packing Density

In this section, we analyze the packing density of these C&E modules based on their size and power dissipation. A chip size of $10\,mm \times 10\,mm$ is assumed through all calculations.

### 7.3.1 Module Size

**L-SEEDs:** The module contains 28 MQW p-i-n diodes and 11 S-SEEDs. By using current integration technologies, each input and output element can be fabricated into a $20\,\mu m \times 20\,\mu m$ centered cell size with a $5\,\mu m \times 5\,\mu m$ optical window. Therefore, the pixel density for L-SEED based multifunctional C&E module is approximately 5000 pixels/cm$^2$ including all the electrical wires. Since each pixel processes two input channels, the channel density is $10^4$ channels/cm$^2$.

**FET-SEEDs:** The module contains 81 FETs, 15 diodes, and 4 SEEDs for compare-exchange unit and 64 FETs, 12 diodes, and 4 SEEDs for the

control unit. Thus, it has a total of 145 FETs, 27 diodes, and 8 SEEDs. Assuming that the size of each FET, diode, and SEED is $20\,\mu m \times 20\,\mu m$, the pixel density and channel density for FET-SEED based multifunctional C&E module are approximately 1250 pixels/cm$^2$ and 2500 channels/cm$^2$, respectively.

**OEICs:** The module contains 49 HBTs and 2 detectors for the compare-exchange unit and 44 HBTs and 2 lasers for the control unit. Thus, it has a total of 93 HBTs, 2 detectors, and 2 lasers. Assuming that the size of each HBT and detector is $20\,\mu m \times 20\,\mu m$, and that the size of the laser is $20\,\mu m \times 5\,\mu m$ for edge emitting type, the pixel density and channel density for OEIC based multifunctional C&E module are approximately 2500 pixels/cm$^2$ and 5000 channels/cm$^2$, respectively.

### 7.3.2 Power Dissipation

Assuming that the maximum power dissipation of a single chip is no more than 1 W/cm$^2$, we summarize the packing density of L-SEED, FET-SEED, and OEIC based C&E modules in Table 7.3. "I/O" represents the input unit and output unit. In OEIC based module, the power efficiency of the laser is an important factor. Current technology can fabricate a laser operating at low bias voltage (3 V) and drive current (1 mA). We estimate the power dissipation of the FET-SEED and OEIC based circuits at 1 GHz operation rate, and S-SEED based circuits at 100 MHz.

| | L-SEEDs | FET-SEEDs | OEICs |
|---|---|---|---|
| device | Gate: 1 mW at 100 MHz | I/O: 1 mW Gate: 1 mW at 1 GHz | I/O: 4 mW Gate: 1 mW at 1 GHz |
| contrast ratio | <5 | 10 | 10-100 |
| switching energy | 100 fJ | 40 fJ | 20 fJ |
| device count | 10 gates | 27 gates | 27 gates |
| pixel | 10 mW/pixel | 28 mW/pixel | 31 mW/pixel |
| packing density | 200 channels/cm$^2$ | 70 channels/cm$^2$ | 64 channels/cm$^2$ |

Table 7.3: Summary of module packing density based on power dissipation.

## 7.4 Summary

This chapter describes the design of a multifunctional 2 input/ 2 output C&E module for optoelectronic sorting networks. Three state-of-the-art device technologies, L-SEEDs, FET-SEEDs, and OEICs are considered in the hardware design. The packing density based on different constraints is summarized in Table 7.4. Current packing density of the module is limited by the power dissipation of the device. The information flux is defined by the product of operation speed and packing density.

$$\text{information flux} = \text{operation speed} \times \text{packing density} \quad (7.7)$$

Note that FET-SEED and OEIC based modules can obtain higher information flux than that of the L-SEED based module.

| channels/cm$^2$ | L-SEEDs | FET-SEEDs | OEICs |
|---|---|---|---|
| interconnections | $\approx 10^3$ | $\approx 10^3$ | $\approx 10^3$ |
| pixel size | $\approx 10^4$ | $\approx 2500$ | $\approx 5000$ |
| power dissipation | $\approx 200$ | $\approx 70$ | $\approx 64$ |
| information flux | $\approx 20$ GHz/cm$^2$ | $\approx 70$ GHz/cm$^2$ | $\approx 64$ GHz/cm$^2$ |

Table 7.4: Summary of packing density based on various constraints.

To sum up, L-SEED technology is suitable for low complexity, high density operations like image processing. On the other hand, FET-SEEDs and OEICs can provide more complicated operation at high speed. An OEIC based design has the possibility of performing multi-wavelength operations.

# Chapter 8

# Conclusions and Future Work

## 8.1 Conclusions

This dissertation presents the optoelectronic implementation of multistage sorting networks. Volume (3-D) sorting networks which utilize the high spatial bandwidth and parallelism of optics are discussed. In particular diffractive optical elements (DOEs) are used in 3-D sorting networks for the interconnections; and optoelectronic compare-and-exchange (C&E) modules are used for the dynamic sorting nodes.

The first part of this work describes the network properties of various conventional 2-D sorting networks and the development of their corresponding 3-D structures. A merge procedure is proposed for network expansion in 3-D space. The mathematical framework is established in the discussion. A detailed analysis of the network performance is given. The decision of choosing a specific network structure depends on two factors: the cost of optical elements and switching elements; and the complexity and reliability of the network.

Designs of optical interconnection systems for one-to-one linear mappings and 2-D folded shuffles are presented. Our study of the

characteristics of Gaussian microbeams propagating in microscaled diffractive optical elements provides an useful method for analyzing the packing limitation and light efficiency of each designed structures. In our work, short and long distance linear mapping interconnections are accomplished by using microlens arrays. Various approaches have been proposed to implement 2-D folded shuffles. Phase-only blazed gratings and off-axis microlenses are used for space-variant designs, and a demagnifying algorithm is proposed for space-semivariant designs. Considering the practical fabrication problems of the multi-level DOEs, mathematically modeled error functions are developed for performance analysis. The result shows that the misalignment error is the most critical factor of all.

Systematic design methods that take both structure and signal modulation characteristic into account were developed for the optoelectronic 2 input/ 2 output multifunctional C&E modules. The device technologies we choose to implement the logic circuits are L-SEEDs, FET-SEEDs, and OEICs with built-in detectors and laser sources. Concepts and details are given through each design. Packing density limitations due to three main factors are analyzed: optical diffraction limitations, pixel size, and power dissipation. Based on the state-of-the-art technology, the analysis shows that the power dissipation of the device currently limits the number of modules that can be built in a single chip.

In conclusion, this dissertation demonstrates a complete architecture in implementing volume multistage sorting networks optoelectronically.

Through our work, we establish the relationship between the device technology and system consideration, in which provides a direction in evaluating the architecture for any specific application.

## 8.2 Future Work

There are many opportunities for future research suggested by this work. Following are some important areas:

- In this work, we concentrate on special-purpose parallel optoelectronic sorting schemes. The study of multipurpose parallel schemes is also an interesting research area. A single-stage structure using a "vary smart pixel" array for processing and a reflective DOEs for global communication between pixels has been currently studied. Another approach involving cascaded stages of reduced-instruction smart pixel arrays and multiple wavelength technologies is a new field of study.

- Parallel sorting networks can provide various operations like parallel data manipulation and network switching. A self-routing packet network which can be built using sorting networks with additional processing stages is an interesting field of study needed further investigation.

- Diffractive optical elements have played important roles in recent interconnection systems. The proposed computer generated space-variant DOE design of an interconnection system has the advantage of

flexible functionality. However, the limited size of each facet and relatively long distance for light routing may produce severe signal aberration. It is important to investigate new optical system architectures for more efficient, higher resolution optical interconnects.

- The grating design and error analysis of proposed interconnection systems are based entirely on paraxial Fourier optics. This theory is inaccurate for gratings that have a feature size on the order of the wavelength of light. The development of mathematical models based on rigorous diffraction theory is required for accurate analysis of grating characteristics.

- The detailed design of smart pixels with special functions depends on new device characteristics. In practical applications, increasing the complexity of the smart pixel circuitry in a system generally eases the complexity of its control scheme. An important area of research lies in optimizing the schemes for various applications.

- Further investigation on the capabilities and limitations of integrating diffractive elements and smart pixel arrays into a compact packaged system is important.

# Reference


[1] R. W. Keyes, "Fundamental limits in digital information processing," *Proc. IEEE*, vol. 69, pp. 267-278, 1981.

[2] D. E. Knuth, *The Art of Computer Programming: Sorting and Searching*, vol. 3, Addison-Welsley, 1973.

[3] J. T. Schwartz, "Ultracomputers," *ACM Trans. Program Lang. Syst.*, vol. 2, pp. 484-521, 1980.

[4] S. G. Akl, *Parallel Sorting Algorithms*, Academic Press, 1985.

[5] H. P. Moravec, "Fully interconnecting multiple computers with pipelined sorting nets," *IEEE Trans. on Comput.*, vol. C-28, p. 795, 1979.

[6] A. Huang and S. Knauer, "Starlite: A wideband digital switch," Proceedings, *IEEE Global Telecommunications Conference*, (Atlanta, GA), vol. 1, pp. 121-125, 1984.

[7] J. Lenfant, "Parallel permutations of data: A Benes network control algorithm for frequently used permutations," *IEEE Trans. on Comput.*, vol. C-27, pp. 637-647, 1978.

[8] C. B. Kuznia, *Cellular Hypercube Optical Interconnections for Optoelectronic Smart Pixel Cellular Arrays*, Ph. D. thesis, University of Southern California, 1994.

[9] C. D. Thompson and H. T. Kung, "Sorting on a mesh-connected parallel computer," Comm. of ACM, vol. 20, pp. 263-271, 1977.

[10] H. S. Stone, "Parallel processing with the perfect shuffle," *IEEE Trans. on Computers*, pp. 153-161, 1971.

[11] M. Schultz, *Numerical Algorithms for Modern Parallel Computer Architectures*, Springer-Verlag, 1988.

[12] A. A. Sawchuk and B. K. Jenkins, "Dynamic optical interconnections for parallel processors," Proc. SPIE, vol. 625, pp. 143-153, 1986.

[13] K. E. Batcher, "Sorting networks and their applications," in *Proceedings, 1968 Spring Joint Computer Conference*, vol. 32, (AFIPS Press, Reston, VA), pp. 307-314, 1968.

[14] J. W. Goodman, F. I. Leonberger, S. Y. Kung and R. A. Athale, "Optical interconnections for VLSI systems," *Proc. IEEE*, vol. 72, pp. 850-866, 1984.

[15] J. E. Midwinter, "'Light' electronics, myth or reality?," *Proc. IEE*, vol. 132, pp. 371-383, 1985.

[16] A. W. Lohmann, W. Stork and G. Stueke, "Optical perfect shuffle," *Appl. Opt.*, vol. 25, pp. 1530-1531, 1986.

[17] C. W. Stirk, R. A. Athale and M. W. Haney, "Folded perfect shuffle optical processor," *Appl. Opt.*, vol. 27, pp. 202-203, 1988.

[18] A. A. Sawchuk and I. Glaser, "Geometries for optical implementations of the perfect shuffle," *Proc. SPIE*, vol. 963, pp. 270-282, 1988.

[19] H. Dammann and K. Gortler, "High-efficiency in-line multiple imaging by multiple phase hologram," *Opt. Commun.*, vol. 3, pp. 312-315, 1971.

[20] J Jahns and A. Huang, "Planar integration of free-space optical components," *Appl. Opt.*, vol. 28, pp. 1602-1605, 1989.

[21] W.-H. Lee, "Computer-generated holograms: techniques and applications," *Progress in Optics XVI*, E. Wolf, ed., pp. 121-231, 1978.

[22] L. P. Boivin, "Multiple imaging using various types of simple phase gratings," *Appl. Opt.*, vol. 11, pp. 1782-1792, 1972.

[23] U. Killat, G. Rabe and W. Rave, "Binary phase gratings for star couplers with high splitting ratio," *Fiber and Integrated Optics*, vol. 4, pp.

159-167, 1982.

[24] U. Krackhardt, F. Sauer, W. Stork and N. Streibl, "Concept for an optical bus-type interconnection network," *Appl. Opt.*, vol. 31, pp. 1730-1734, 1992.

[25] W. B. Veldkamp and C. J. Kastner, "Beam profile shaping for laser radars that use detector arrays," *Appl. Opt.*, vol. 21, pp. 345-356, 1982.

[26] J. Cordingley, "Application of a binary diffractive optic for beam shaping in semiconductor processing by laser," *Appl. Opt.*, vol. 32, pp. 2538-2542, 1993.

[27] W. B. Veldkamp, J. R. Leger and G. J. Swanson, "Coherent summation of laser beams using binary phase gratings," *Opt. Lett.*, vol. 11, pp.303-305, 1986.

[28] L. D'Auria, J. P. Huignard, A. M. Roy and E. Spitz, "Photolithographic fabrication of thin film lenses," Opt. Commun., vol. 5, pp. 232-235, 1972.

[29] G. J. Swanson and W. B. Veldkamp, "Binary lenses for use at 10.6 micrometers," *Opt. Eng.*, vol. 24, pp. 791-795, 1985.

[30] T. Shiono, K. Setsune, O. Yamazaki and K. Wasa, "Rectangular-apertured micro-fresnel lens array fabricated by electron-beam lithography," *Appl. Opt.*, vol. 26, pp. 587-591, 1987.

[31] J. Turunen, A. Vasara, J. Westerholm, G. Jin and A. Salin, "Optimization and fabrication of grating beamsplitters," *J. Phys. D: Appl. Phys.*, vol. 21, pp. S102-S105, 1988.

[32] U. Krackhardt and N. Streibl, "Design of Dammann-gratings for array generation," *Opt. Commun.*, vol. 74, pp. 31-36, 1989.

[33] F. B. McCormick, F. A. P. Tooley, T. J. Cloonan, J. M. Sasian and H. S. Hinton, "Microbeam interconnections using microlens arrays for free space photonic systems," in *OSA Proceedings on Photonic Switching*, H. S. Hinton and J. W. Goodman, ed., vol. 8, pp. 90-96, 1991.

[34] F. B. McCormick, "Generation of large spot arrays from a single laser beam by multiple imaging with binary phase gratings," *Opt. Eng.*, vol. 28, pp. 299-304, 1989.

[35] S. J. Walker and J. Jahns, "Array generation with multilevel phase gratings," *J. Opt. Soc. Am. A*, vol. 7, pp. 1509-1513, 1990.

[36] J. N. Mait, "Design of binary-phase and multiphase Fourier gratings for array generation," *J. Opt. Soc. Am. A*, vol. 7, pp. 1514-1528, 1990.

[37] M. E. Prise, N. C. Craft, R. E. LaMarche, M. M. Downs, S. J. Walker, L. A. D'Asaro and L. M. F. Chirovsky, "Module for optical logic circuits using symmetric self-electrooptic effect devices," *Appl. Opt.*, vol. 29, pp. 2164-2170, 1990.

[38] J. M. Miller, M. R. Taghizadeh, J. Turunen and N. Ross, "Multilevel-grating array generators: Fabrication error analysis and experiments," *Appl. Opt.*, vol. 32, pp. 2519-2525, 1990.

[39] M. Ekberg, M. Larsson, S. Hard, J. Turunen, M. R. Taghizadeh, J. Westerholm and A. Vasara, "Multilevel grating array illuminator manufactured by electron-beam lithography," *Opt. Commun.*, vol. 88, pp. 37-41, 1992.

[40] T. J. Cloonan, M. J. Herron, F. A. P. Tooley, G. W. Richards, F. B. McCormick, E. Kerbis, J. L. Brubaker and A. L. Lentine, "An all-optical implementation of a 3-D crossover network," *IEEE Photon. Technol. Lett.*, vol. 2, pp. 438-440, 1990.

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

[42] S. Kirkpatrick, C. D. Gelatt Jr., M. P. and Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671-680, 1983.

[43] J. Mait, "Design of Dammann gratings for two-dimensional, nonseperable, noncentrosysmetric responses," *Opt. Lett.*, vol. 14, pp. 196-198, 1989.

[44] H. Heibmeier, U. Krackhardt and N. Streibl, "A Dammann grating with diffraction order of arbitrary intensity etched into $Al_2O_3$," *Opt. Commun.*, vol. 76, pp. 103-106, 1990.

[45] R. L. Morrison, S. L. Walker and T. J. Cloonan, "Beam array generation and holographic interconnections in a free-space optical switching network," *Appl. Opt.*, vol. 32, pp. 2512-2518, 1993.

[46] M. R. Taghizadeh and J. I. B. Wilson, "Optimization and fabrication of grating beamsplitters in silicon nitride," *Appl. Phys. Lett.*, vol. 54, pp. 1492-1494, 1989.

[47] F. B. McCormick, F. A. P. Tooley, T. J. Cloonan, J. M. Sasian, H. S. Hinton, K. O. Mersereau and A. Y. Feldblum, "Optical interconnections using microlens arrays," *Opt. Quantum Electron.*, vol. 24, pp. S465-S477, 1992.

[48] S. H. Song, C. D. Carey, D. R. Selviah, J. E. Midwinter and E. H. Lee, "Optical perfect-shuffle interconnection using a computer-generated hologram," *Appl. Opt.*, vol. 32, pp. 5022-5025, 1993.

[49] Y. Zhan, H. Kang and J. Zhang, "Optical implementation of the folded perfect shuffle interconnection network using quadrant-encoded grating," *Opt. Eng.*, vol. 32, pp. 1657-1661, 1993.

[50] C. W. Stirk and R. A. Athale, "Sorting with optical compare-and-exchange modules," *Appl. Opt.*, vol. 27, pp. 1721-1726, 1988. [Stirk88b]

[51] L. Zhang, R. Jin, C. W. Stirk, G. Khitrova, R. A. Athale, H. M. Gibbs, H. M. Chou, R. W. Sprague and H. A. Macleod, "All-optical compare-and-exchange switches," *IEEE J. Selected Areas Commun., Special Issue on Photonic Switching*, vol. 6, pp. 1273-1279, 1988.

[52] M. J. Murdocca and T. J. Cloonan, "Optical design of a digital switch," *Appl. Opt.*, vol. 28, pp. 2505-2517, 1989.

[53] X. Mao, S. Liu and R. Wang, "Optoelectronic compare-and-exchange switches based on the BILED circuits," *SPIE, Optical Enhancements to Computing Technology*, vol. 1563, pp. 58-63, 1991.

[54] H. S. Hinton, "Photonic switching fabrics," *IEEE Communications Magazine*, vol. 28, pp. 71-89, 1990.

[55] in *LEOS Summer Topical Meeting Digest on Smart Pixels*, (Santa Barbara, CA), 1992.

[56] in *IEEE Journal of Quantum Electronics, Special Issue on Smart Pixel*, vol. 29, 1993.

[57] Ron A. Spanke and V. E. Benes, "N-stage planar optical permutation network," *Appl. Opt.*, vol. 26, pp. 1226-1229, 1987.

[58] L. R. McAdams, *Photonic Switching with Liquid Crystal*, Ph. D. thesis, Stanford University, 1990.

[59] R. R. Tummala, "Multichip packaging-A tutorial," *PIEEE*, vol. 80, pp. 1924-1941, 1992.

[60] Lily Cheng, *Optoelectronic implementation of multistage interconnection networks*, Ph. D. thesis, University of Southern California, 1992.

[61] Joseph T. Verdeyen, *Laser Electronics*, Prentice-Hall, 1981.

[62] G. J. Swanson, "Binary optics technology: The theory and design of multi-level diffractive optical elements," *MIT Lincoln Lab. Tech. Rep.* 854, 1989.

[63] H. Nishihara and T. Suhara, "Micro fresnel lenses," *Progress in Optics XXIV*, E. Wolf, ed., pp. 1-37, 1987.

[64] J. M. Wang, L. Cheng and A. A. Sawchuk, "Holographic implementation of 2-d perfect shuffles based on a one-copy algorithm," *Photonic Switching Technical Digest Series*, pp. 180-183, 1991.

[65] R. Petit, ed., *Electromagnetic Theory of Gratings*, Springer-Verlag, 1980.

[66] T. Fujita, H. Nishihara and J. Koyama, "Blazed gratings and Fresnel lenses fabricated by electron-beam lithography," *Opt. Lett.*, vol. 7, pp. 578-580, 1982.

[67] T. Shiono and H. Ogawa, "Diffraction-limited blazed reflection diffractive microlenses for oblique incidence fabricated by electron-beam lithograthy," *Appl. Opt.*, vol. 30, pp. 3643-3649, 1991.

[68] K. S. Urquhart, R. Stein and S. H. Lee, "Computer-generated holograms fabricated by direct write of positive electron-beam resist," *Opt. Lett.*, vol. 18, pp. 308-310, 1993.

[69] C. W. Stirk, R. A. Athale and C. B. Friedlander, "Optical implementation of the compare-and-exchange operation for applications in symbolic computing," *Proc. Soc. Photo-Opt. Instrum. Eng.*, vol. 754, p. 27, 1987.

[70] M. Minsky, *Computation: Finite and Infinite Machines*, Prentice Hall, 1967.

[71] L. M. F. Chirovsky, L. A. D'Asaro, C. W. Tu, A. L. Lentine, G. D. Boyd and D. A. B. Miller, "Batch fabricated symmetric SEEDs," in *OSA Proceedings on Photonic Switching*, (Optical Society of America, Washington DC), pp. 2-6, 1989.

[72] A. L. Lentine, F. B. McCromick, R. A. Novotny, L. M. F. Chirovsky, L. A. D'Asaro, R. F. Kopf, J. M. Kuo and G. D. Boyd, "A 2 Kbit array of symmetric self electro-optic effect devices," *IEEE Photon Technol. Lett.*, vol. 2, pp. 51-53, 1990.

[73] A. L. Lentine, H. S. Hinton, D. A. B. Miller, J. E. Henry, J. E. Cunningham and L. M. F. Chirovsky, "Symmetric self-electrooptic effect device: optical set-reset latch, differential logic gate, and differential modulator/detector," *IEEE J. Quantum Electron.*, vol. 25, pp. 1928-1936, 1989.

[74] A. L. Lentine, L. M. F. Chirovsky, M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, G. J. Przybylek, L. E. Smith, L. A. D'Asaro and D. A. B. Miller, "Integrated array of self electro-optic effect device logic gate," in *Technical Digest of Topical Meeting on Optical Computing*, (Optical Society of America, Washington DC), paper MA2, 1991.

[75] F. B. McCormick, A. L. Lentine, L. M. F. Chirovsky and L. A. D'Asaro, "An all -optical shift register using symmetric self electro-optic effect devices," in *Technical Digest of Topical Meeting on Optical Computing*, (Optical Society of America, Washington DC), paper ThC5, 1989.

[76] A. L. Lentine, L. M. F. Chirovsky, M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, G. J. Przybylek, L. E. Smith, L. A. D'Asaro and D. A. B. Miller, "Integrated self electro-optic effect device photonic switching nodes," in *Technical Digest of Topical Meeting on Photonic Switching*, (Optical Society of America, Washington DC), paper ThC4, 1991.

[77] A. L. Lentine, T. J. Cloonan and F. B. McCormick, "Photonic switching nodes based on self electro-optic effect devices," *Opt. Quantum Electron.*, vol. 24, pp. S443-S464, 1992.

[78] S. Yu and S. R. Forrest, "Implementations of smart pixels for optoelectronic processors and interconnection system II: SEED-based technology and comparison with optoelectronic gates," *Journal of Lightwave Technology*, vol. 11, pp. 1670-1680, 1993.

[79] A. L. Lentine and D. A. B. Miller, "Evolution of the SEED technology: Bistable logic gates to optoelectronic smart pixels," I*EEE Journal of Quantum Electron*, vol. 29, pp. 665-669, 1993.

[80] T. K. Woodward, A. L. Lentine, L. M. F. Chirovsky, M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, L. E. Smith, L. A. D'Asaro, E. J. Laskowski, and S. S. Pei, "GaAs/AlGaAs FET-SEED receiver/transmitters," in T*echnical Digest of Topical Meeting on Photonic Switching*, (Optical Society of America, Washington DC), paper PMC3, 1993.

[81] A. L. Lentine, T. J. Cloonan, H. S. Hinton, L. M. F. Chirovsky, L. A. D'Asaro, E. J. Laskowski, S. S. Pei, M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, L. E. Smith, G. D. Boyd and T. K. Woodward, "$4 \times 4$ arrays of embedded control FET-SEED $2 \times 1$ switching arrays," in *IEEE Topical Meeting on Smart Pixels*, post-deadline papers, 1992.

[82] G. R. Olbright, R. P. Bryan, K. Lear, T. M. Brennan, G. Poirier, Y. H. Lee and J. L. Jewell, "Cascadable laser logic devices: Discrete integra-

tion of phototransistors with surface-emitting laser diodes," *Electron. Lett.*, vol. 27, pp. 216-217, 1991.

[83] P. Zhou, J. Cheng, C. F. Schaus, S. Z. Sun, C. Hains, K. Zheng, E. Armour, W. Hsin, D. R. Myers and G. A. Vawter, "Cascadable, latching photonic switch with high optical gain by the monolithic integration of vertical-cavity surface-emitting laser and a pn-pn photothyristor," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 1009-1012, 1991.

[84] J. J. Brown, J. T. Gardner and S. R. Forrest, "Optically powered, integrated 'smart' pixels for optical interconnection networks," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 1136-1138, 1991.

[85] see W. F. Hsu device table.

[86] F. B. McCormick, "Free-space interconnection techniques," *Photonic in Switching*, J. E. Midwinter, ed., vol. 2, pp 169-250, 1993.