USC–SIPI REPORT #232

Analysis and Optimization of Lens-Based Optoelectronic Shuffle Networks

by

Andrew S. Miller

February 1993

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 mother and father,

and to Marcela

Acknowledgments

Above all, I wish to thank my advisor, Dr. Alexander Sawchuk, for being everything a student could hope for in an advisor. He has been both supportive and honest, and has consistently made himself available for assistance and consultation. He possesses the invaluable combination of both great knowledge and a willingness to share that knowledge.

I would also like to thank my dissertation committee members, Dr. Keith Jenkins and Dr. Murray Gershenzon, for their assistance in this endeavor.

Heartfelt thanks go to friends and co-workers in the Signal and Image Processing Institute (SIPI): to Charlie Kuznia and Dr. Lily Cheng for being terrific friends and office-mates, and for a great many wonderful discussions; to Dr. Anand Rangarajan for support and guidance through some troubled waters; to my fellow viajeros to the ”sun deal” Sabino Piazzolla (long may you sing) and Dr. Jules Osinski; and to Dr. Jimmy Wang, Dr. Xiao-hong Yan, Dr. Ching-Yih Tseng, Luis Castedo, Dr. Javier Fonollosa, Elisa Sayrol, Michel Gaeta, Greg Petrisor, Clare Waterson, Wei-Feng Hsu, Jeng-Feng Lin, Yibin Lu, Greg Mitchell, Gary Mandyam, Brian Beard, and Ted Dean.

Special thanks to both Dr. Allan Weber and Dr. John Mosher whose assistance to the compuphobic helped avert many a heart attack. I also wish to acknowledge the help and support of the staff at SIPI, Ms. Gloria Halfacre (for the free daily roller coaster), Ms. Mayitta Penoliar, Ms. Linda Varilla and Ms. Mercedes Morente.

Finally, I wish to acknowledge the two people who sparked my interest in science as a youngster, both my father John Miller and my uncle Richard Hermsdorf.

This research was supported by the Defense Advanced Research Projects Agency (DARPA) through the National Center for Integrated Photonic Technology under Contract No. MDA972-90-C-0037, and by the Joint Services Electronics Program (JSEP) through the Air Force Office of Scientific Research under Contract No. F49620-91-C-0028.

Contents

List of Figures

List of Tables

Abstract

We examine in detail the capabilities of the lens-based free-space optical shuffle. Specifically, we characterize the system by developing both the input-to-output and output-to-input relations for the folded and separable shuffles. We examine the physical optics limitations on the resolution capabilities of the lens-based shuffle for the case of simple, single-element lenses. Then we show how the geometry of the system may be optimized to shuffle the greatest number of channels. We show how this method may be extended to the general case of aberration corrected lenses. We then show how deviations from the ideal design affect the performance of the system in terms of crosstalk between the channels. We develop the statistical rules governing optical crosstalk in an array of channels.

Finally, we suggest a pair of ways to improve the design and performance of a shuffle-exchange network. The first is to multiplex many shuffles through a single set of shuffle optics. We show how this can be done, and what the consequences are both in terms of benefits and drawbacks. We then suggest a hybridization of two different shuffle architectures, the lens-based and holographic shuffles, which combines the desirable aspects of both systems while eliminating the undesirable traits.

Chapter 1

Introduction

1.1 Background

The history of computing has been dominated by advances in the fabrication of switching devices. Specifically, the invention of the transistor, its miniaturization and integration onto silicon substrates revolutionized the ability of machines to process, store and communicate information. Until recently, computer designers relied on the tremendous and ever-increasing switching speed of the electronic transistor to build machines which solve more computationally intensive problems. Because the scientific community began to believe in the early 1980's that electronic devices had reached a plateau in performance and size, interest turned to using light as a medium for computation. There was hope that because of the extremely high frequency (THz) of light, devices using light as a medium might be fabricated which would switch at frequencies simply unattainable by electronic devices.

Now, a decade later, the once-feared limits of electronic devices have yet to materialize. Speeds have increased by roughly an order of magnitude, and feature sizes on integrated circuits have decreased by at least a factor of five. Current commercially available custom and semi-custom integrated circuits have reached extremely high data rates (into the 100 MHz range). Dynamic random access memory (DRAM) capable of storing sixteen million bits on a single chip is now available on the market. Despite these impressive figures, the potential for advancement in size and speed is far from exhausted. The smallest achievable lateral feature sizes will soon be measured in nanometers, when it was thought only a decade ago that near one micron there existed a fundamental limit to the gate length of a MOS transistor. Research on heterostructure semiconductor devices based on materials other than silicon have proven an ability to switch at least two, and possibly three *orders of magnitude* faster than the current commercial lot.

Knowing this, one must reevaluate the presumed need for optical switches. If one were to devise an optical transistor, meaning a device which controls light with light, it probably would not have the capabilities of the electronic transistor for fundamental physical reasons. First, in order to get one light beam to control another would require their interaction in some medium over a distance of at least their wavelength [35]. For the high speed optical sources we now have, this implies an interaction distance, and thus a device size, of greater than a micron. Commercially available electronic devices already have smaller interaction regions (e.g. the length under a MOS transistor gate). It is a fundamental fact that it is difficult to get two beams of light to interact with each other. The proven methods of producing a nonlinear interaction between beams of light involve either tremendous optical power or unreasonably large interaction distances, and require a nonlinear optical material. Photons do not readily interact. Electrons, however, do. It is wise, then to use this fundamental difference between light and electricity to some advantage.

Furthermore, a different *approach* to computing is necessary in order to solve extremely complex problems, such as the simulation of airflow, chemical reactions or the weather. The traditional means of executing tasks serially using one central brain (processor) simply is not sufficient for these problems. It is far better to break up a large problem into many smaller pieces and allow a large number of processors to work on these parts in parallel. The number of processors expected to be used in such an arrangement might be as high as 2^{16}, or on the order of ten thousand [34]. The problem at hand, then is communication. In order for these processors to coordinate their efforts and share their results, a large, efficient, and rapidly reconfigurable interconnection network must exist between them.

1.2 Interconnection Networks

A variety of interconnection networks have been proposed [32]. The most powerful interconnection network for one-to-one connections is the crossbar. Its VLSI implementation is simple in concept, and is shown in Fig. 1.1. Wires for the inputs are laid horizontally, while those for the output are laid vertically, and an insulating layer of oxide is placed between the inputs and outputs. A MOS transistor is placed across each intersection. When a voltage is applied to the gate of the transistor, it connects the input to output at that intersection. The power of the crossbar is that any permutation of one-to-one connections from input to output may be realized, and that rearranging the connection pattern to achieve a new permutation is simple and straightforward.

There are, however, a pair of serious drawbacks for the crossbar. First is the number of switch elements, or transistors, needed. If the network must connect N inputs to N outputs, then N^2 switches are needed. As N grows into the thousands, this N^2 complexity becomes prohibitive. Second is the amount of crosstalk between channels generated because of wires running side-by-side for great lengths, and because of the large number of places where wires cross directly over one another. At the frequencies currently used in commercial circuits, the wires on a chip radiate (and receive) like antennas, and this worsens as the frequency increases [10] [40] [13].

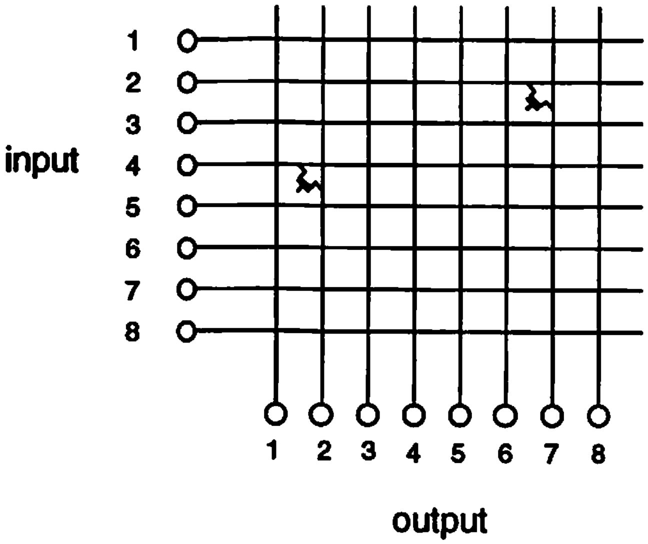


Figure 1.1: The strictly non-blocking crossbar in its VLSI implementation. Here, input 2 is connected to output 7, and input 4 is connected to output 2, each by means of a MOS transistor whose gate has been turned on by an applied voltage.

1.3 Shuffle-Exchange Networks

To alleviate the scaling problem of the crossbar, many architectures have been proposed which require order $N \log_2 N$ switches to achieve arbitrary input-to-output permutations [34]. One such architecture is the shuffle-exchange network, shown in Fig. 1.2. The system consists of duplicate layers, or stages, each composed of a static perfect shuffle connection and an array of $N/2$ bypass-exchange switches. The perfect shuffle connection divides the inputs in half, and interlaces them perfectly in a way analogous to shuffling a deck of cards. The switches take a pair of inputs and either pass them directly through, or exchange them.

It has been shown that if one builds a system with $m = 2\log_2 N - 1$ stages, that the system is capable of achieving any arbitrary permutation of input to output [3]. A network with the ability to achieve any permutation is called *non-blocking.* The system is *fully non-blocking* if an input may be re-routed from its current output to a different (unused) output without disrupting other existing connections. The system is rearrangeably *rearrangeably non-blocking* if other parts

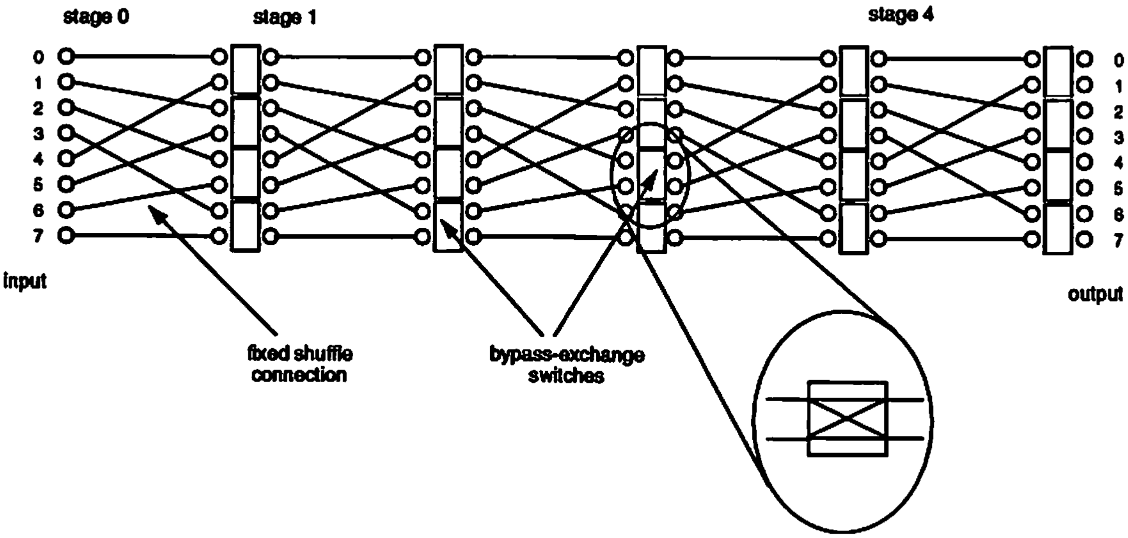


Figure 1.2: A five-stage shuffle-exchange network with N equal eight inputs and eight outputs. Here, $(2\log_2 N - 1)$ stages is sufficient to enable all possible input-to-output permutations.

of the switch fabric must be altered in order to change a single input-to-output connection. It has also been demonstrated that if the number of stages is $m = \log_2 N$, then any input may reach any output, provided no other connections are already established. This system is called *full access*, and it may be sufficient for certain applications which do not require arbitrary input-to-output permutations.

These systems have two basic components. The first is the global, static shuffle connection, and the second is the local, dynamic two-input-to-two-output connection. The distinction between the two functions is relevant to the choice of optics or electronics to execute them. Because the shuffle is both static and global with many points where the communication channels cross each other, it is well suited to the strengths of optics. Beams of light may cross one another with almost no observable interaction. This alleviates the crosstalk problem discussed earlier for the crossbar. Furthermore, optics is capable of executing the shuffle in three dimensions on a 2-D array, as illustrated in Fig. 1.3. The local bypass-exchange operation makes the network dynamic. Electronics accomplishes the

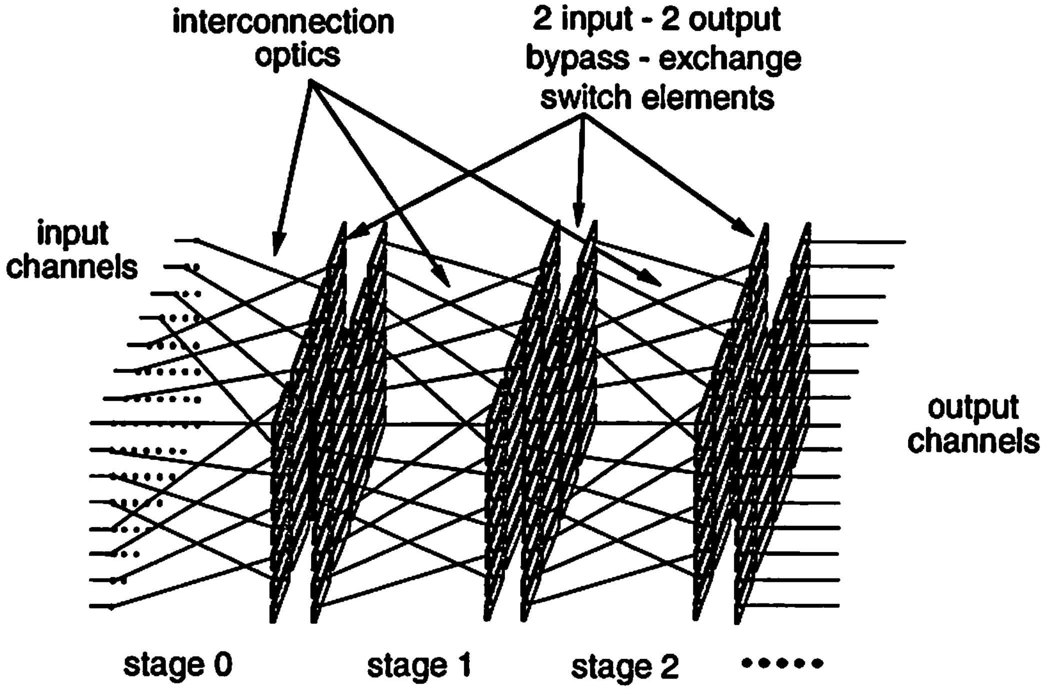


Figure 1.3: A 3-D shuffle-exchange network in which optics executes the static shuffle connections at each stage, and electronics executes the bypass-exchange switching.

task of switching far more easily than optics, and it has been argued that at a fundamental physical level, local (short distance) communications are better carried out electronically [26].

There are relatively simple means of implementing the perfect shuffle using standard bulk optics like lenses, prisms and mirrors [20], [7], [36], [33], [31]. Other free space optical techniques use diffractive or holographic optics [19], [30], [42]. All of these methods may be classified in terms of their need for either collimated or divergent sources, coherent or incoherent sources, high or low resolution capability, whether the system is space-variant, space-invariant or some hybrid, and the simplicity of constructing the system. This list is certainly not exhaustive, but it does encompass the variety of important attributes of any optical interconnection scheme.

1.4 Contributions

Throughout the rest of this work, we shall examine a simple architecture employing conventional bulk optics, specifically four lenses [36] [31]. We shall review the basic architecture and then:

- define its input-to-output and output-to-input characteristics
- show how its geometry may be optimized for greatest channel packing density
- analyze the effects of imperfections in the system construction on its performance, and establish acceptable design tolerances
- describe the statistics of optical crosstalk in a 2-D array of channels, and compare the effect of crosstalk on bit error rate (BER) to that of Gaussian noise
- describe a means of reducing the hardware required to build the multi-stage system by space multiplexing many optical shuffles through a single set of shuffle optics
- compare the channel packing density capability of this lens-based architecture with a holographic beam steering shuffle architecture [43], and suggest means of achieving a truly superior shuffle architecture by combining the strengths of both systems

Chapter 2

Review of Architecture

2.1 Bypass-Exchange Switches

A possible means of executing the bypass-exchange and interfacing with the optical input and output of each stage is shown in Fig. 2.1. The transistors at the input are optically activated heterojunction bipolar transistors, where the base current is provided by the absorption of photons in the base. The transistor amplifies the photocurrent, which is sent to an output laser diode. The control beam provides photocurrent which determines which input goes to which output laser.

This routing mechanism is by active optoelectronics. There are also passive routing methods available, such as that demonstrated by Govindarajan *et al* [21], where the two inputs to a bypass-exchange are routed according to their polarizations. It should be noted that a key element of this switch is that the channels are powered optically. The authors demonstrated that when the laser diode outputs are powered by separate photovoltaic cells, crosstalk between the two channels is reduced by more than two orders of magnitude.

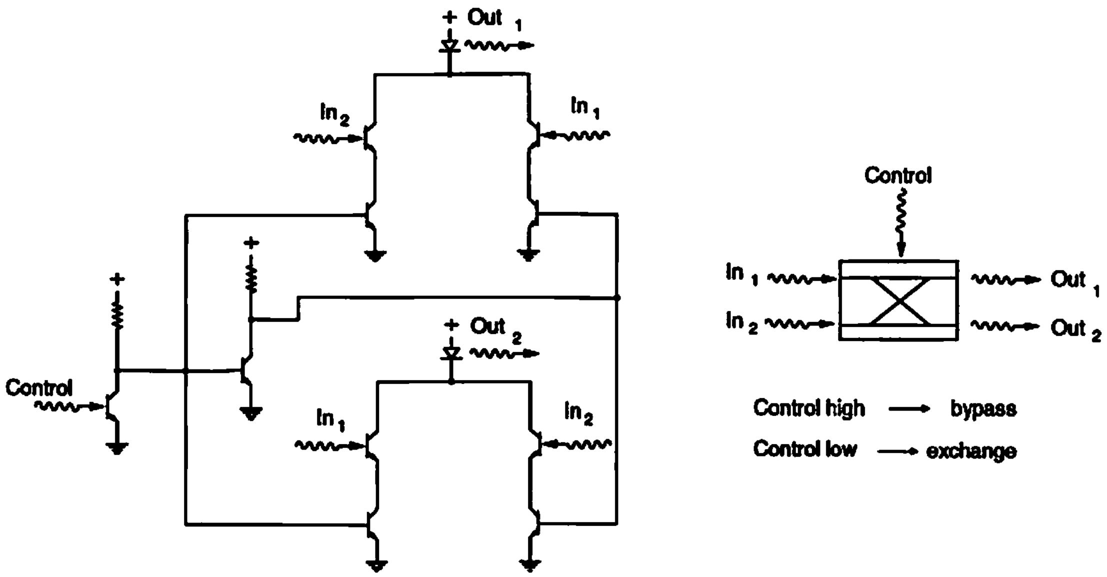


Figure 2.1: Optoelectronic implementation of the bypass-exchange switch, using bipolar phototransistors and laser diodes.

2.2 The Perfect Shuffle

In this work, we examine the capabilities of a shuffle architecture which uses four simple lenses [36] [31]. We stated in the introduction that a perfect shuffle on a column of inputs is analogous to cutting a deck of cards in half, and interlacing the two halves, which is shown in Fig. 2.2. The idea behind the lens-based shuffle is that one may use a pair of lenses to create duplicate images of the input array, and then position the images appropriately to achieve the interlacing necessary for the shuffle.

Knowing that the real image of a simple lens is inverted, we turn the output of the perfect shuffle (at the end of a given stage) upside-down, maintaining the connections from input to output. The result is shown on the right hand side of Fig. 2.2. To find out where the lenses must be located to achieve the shuffle, recall that when a thin lens creates an image, the ray passing through the center of the lens goes undeviated from the source (or object) to the image. Figure 2.2 shows

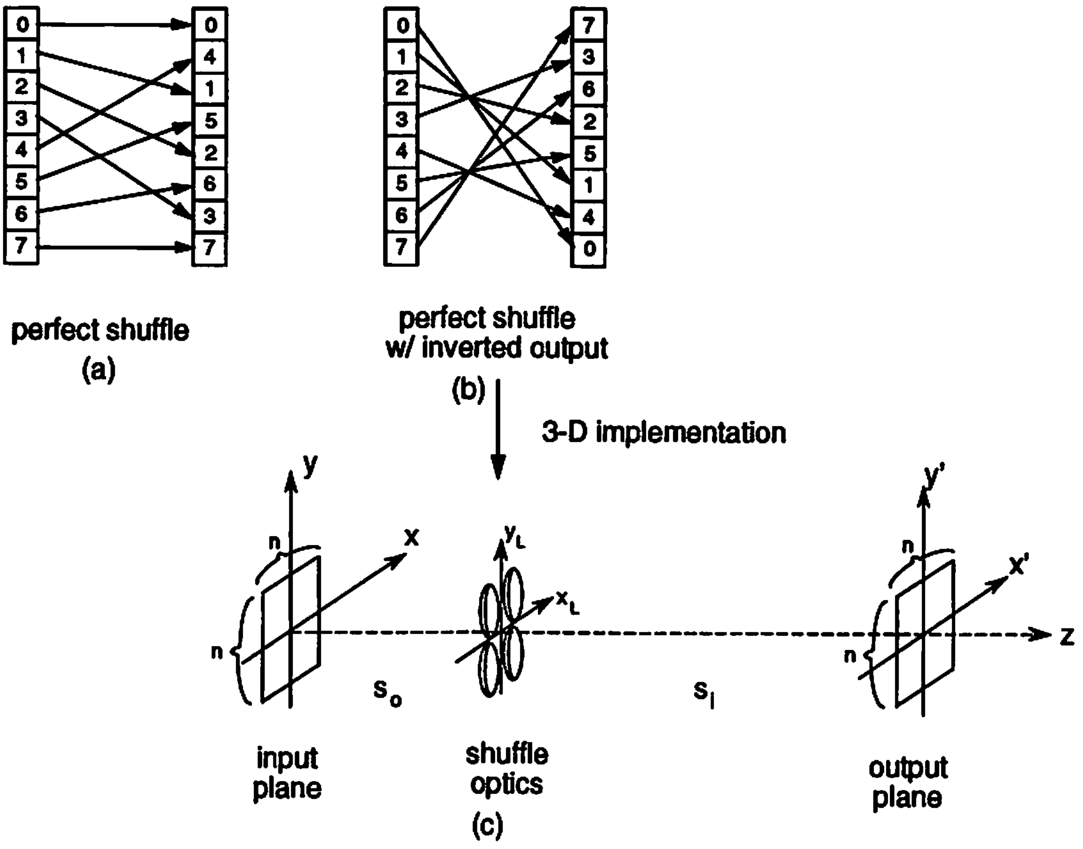


Figure 2.2: A perfect shuffle of eight inputs to eight outputs, and on the right, the same shuffle with its output is inverted as occurs when the shuffle is implemented by imaging with lenses. The two cross points show where the lenses are centered for the planar shuffle, and below is the 3-D version of the shuffle with four lenses.

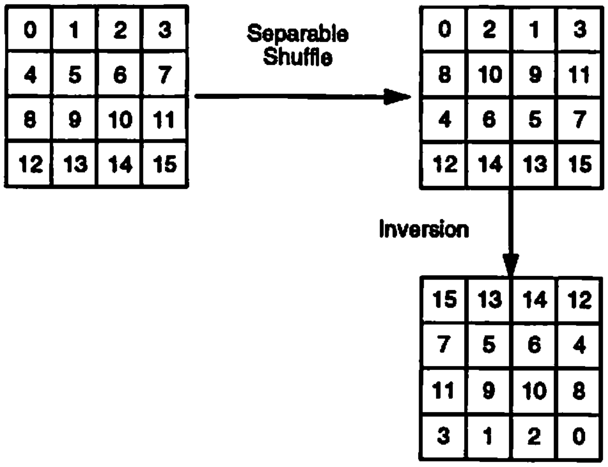


Figure 2.3: Separable shuffle performed independently on the rows and columns of the input. Also shown is the inverted result obtained when the shuffle is executed by our lens-based architecture.

that there are two lenses, each centered at one of the cross points for the central rays.

2.3 Shuffles in 3-D

With optics, we may take advantage of the third dimension to achieve higher parallelism. Simply by using another pair of lenses, the shuffle may be executed simultaneously in the plane orthogonal to the diagram. This yields the 3-D separable perfect shuffle, which shuffles the rows and columns of a 2-D input array independently. The results of a 2-D (planar) shuffle on a linear array of inputs is shown in Fig. 2.2, whereas the results for the 3-D separable shuffle are shown in Fig. 2.3.

A variation on the separable shuffle is the folded shuffle. One may think of this as rastering an $n \times n$ input array into a long $n^2 \times 1$ vector, performing a simple (planar) perfect shuffle on the vector, and then rastering the resulting vector back into an $n \times n$ output array. The results for the folded shuffle are shown in Fig. 2.4.

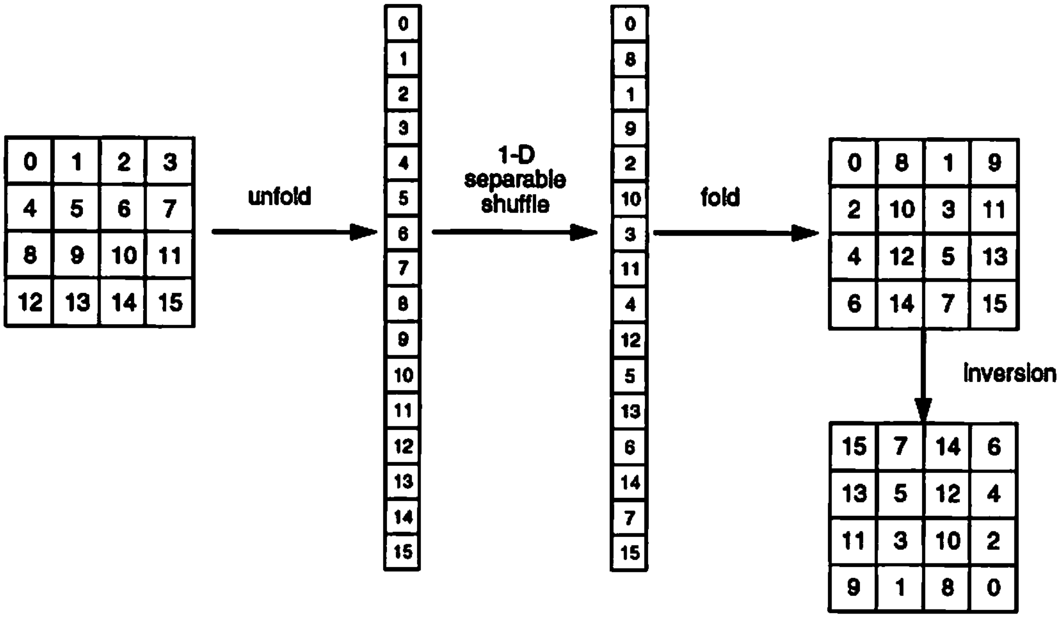


Figure 2.4: The folded shuffle is defined by unfolding the 2-D array into a vector, performing a perfect shuffle, then folding the vector back into the 2-D array.

2.4 Lens Positioning

Each of the four lenses in the optical implementation of the shuffle makes its own image of the input array. The four images overlap in the output plane according to the position of the lenses in x and y. The geometry of this implementation of the shuffle requires that the lenses are displaced from the central axis of the system by one sixth the size L of the input array in 1-D

$$x_l, y_l = \pm\frac{1}{6}L. \tag{2.1}$$

where the subscript l indicates the *lens* position. We also note that the *object distance* s_o (from the source to the lens) is one half the *image distance* s_i (from the lens to the output plane). The magnification M of an image is the ratio of image distance

$$M = -\frac{s_i}{s_o}, \tag{2.2}$$

so the image in the output plane (or copy) of the input array made by each lens is double the size of the input array. So, the actual images in the output plane for the separable and folded shuffles are shown in Fig. 2.5. The input is a 4×4 array, so each copy in the output plane is 8×8. The useful output in each case, outlined in bold, is the central 4×4 array, while the peripheral $3n^2 = 48$ image points are ignored.

Notice that two of the 8×8 images in the folded shuffle are slightly skewed compared to their corresponding positions in the separable shuffle. This is accomplished by slightly skewing the lens placement. If the distance between each input is δ, then

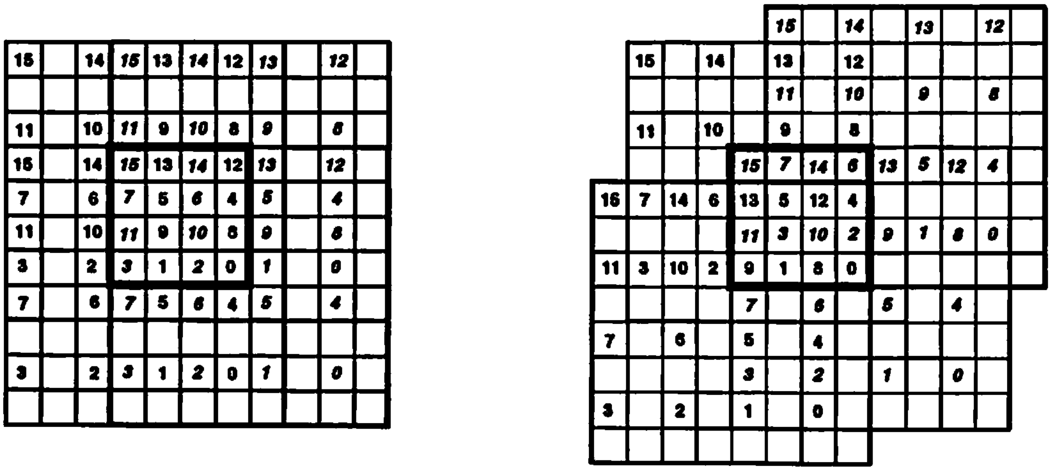
$$L = (N - 1)\delta. \tag{2.3}$$

It will be shown in a later chapter that any lateral shift (x, y) of a lens results in a threefold lateral shift of an image in the output plane. To accomplish the folded shuffle, we must move two of the output copies by δ in both x and y, so we need to shift the lenses by $\delta/3$ in x and y. The resulting skewed lens positions are

$$x, y = \pm\frac{1}{6}(N + 1)\delta. \tag{2.4}$$

2.5 Qualitative Aspects of the Architecture

We choose to examine this architecture because of its simplicity, both in terms of ease of construction and ease of analysis. Its simplicity allows straightforward application of the constraints of physical optics to optimize the system design, and to determine the number of independent channels which the system is capable of resolving. While we believe that its simplicity is an excellent attribute, we make no attempt to prove general superiority of this system over any other. We merely develop in detail its channel packing density capability.



separable shuffle

folded shuffle

Figure 2.5: The lens-based shuffle uses four separate images of the input array, magnified by two, and shifted properly to interlace them to produce the shuffled output in the central square. The folded shuffle uses the same four copies, but with two of them slightly skewed.

Two primary factors influence the packing density limits of any given system, the resolution of the optics and the heat dissipation of the associated optoelectronic elements. We consider both, and find that our architecture performs well in one category and poorly (by comparison) in the other. By extending the current analysis of this system to other architectures, we may make critical comparisons. We have already begun a similar analysis of a holographic arrangement [42], which will be the subject of a future discussion.

Chapter 3

Input to Output Relations in an Optical Shuffle

3.1 Background

A general geometrical relationship between the input and output coordinates of the optical shuffle is useful for simulating the operation of the system, or for the development of an external controller to set the switch fabric. The lens-based shuffle accomplishes the shuffle by interlacing multiple images of the input array. These images are real, and thus inverted. Thus, the shuffled output is also inverted (see Fig. 2.2). The relationships we develop in this chapter incorporate the inversion of the shuffled output, which is inherent in the lens-based shuffle.

3.2 Known Shuffle Relations

For a non-inverted (erect) output, the relationships between input and output for both the folded and separable shuffles are already known. To map an input (i, j) to an output (i', j') is simple if we represent i and j in binary, as a string of binary digits b. The input coordinate is given by

$$(i,j) = (b_{m-1}b_{m-2}\cdots b_{m/2}, b_{(m/2)-1}b_{(m/2)-2}\cdots b_1b_0) \tag{3.1}$$

where b is a binary bit, and $m/2$ is the number of bits needed to specify either i or j in binary. One finds the erect shuffled output by a cyclic rotation of the bits b [5]. For the folded shuffle, the output coordinates are

$$(i',j')_f = (b_{m-2}\cdots b_{m/2}b_{(m/2)-1}, b_{(m/2)-2}\cdots b_1b_0b_{m-1}) \tag{3.2}$$

which is a left cyclic shift of all m bits. The output of the separable shuffle is

$$(i',j')_s = (b_{m-2}\cdots b_{m/2}b_{m-1}, b_{(m/2)-2}\cdots b_1b_0b_{(m/2)-1}) \tag{3.3}$$

which is a left cyclic shift of the $m/2$ bits for i and j separately.

3.3 Motivation for a Different Method

The strength of the binary cyclical shift (BCS) method of Section 3.2 is its simplicity and elegance. It suffers from a few inconvenient, though admittedly minor, restrictions. In order to apply the BCS method, the coordinates i and j must be expressed as binary numbers, instead of decimal. Matrix coordinates (i,j) for an $n \times n$ array usually range from 1 to n; in the BCS method, they must range from 0 to $n-1$. In the BCS method, n must be some power of 2, so that $n = 2^m$. By contrast, any $n \times n$ array where n is *even* may be shuffled, which is much less restrictive than requiring n to be a power of two. Finally, the binary cyclical shift method does not account for the inversion of the output of the lens-based shuffle. We may account for the inversion either by using an inverted output coordinate system, or by taking the complement of the binary representation of the output coordinate (i.e. inverting each bit).

We seek a programmed approach to map an input coordinate (i,j) to its shuffled output coordinate (i',j'), without the inconveniences of the BCS method. We assume an $n \times n$ array, where n is even, where the coordinates (i,j) are expressed as decimal numbers, and where $i,j \in \{1,2,\ldots n\}$. The coordinate systems of input and output are *not* inverted with respect to each other; the inversion inherent in the lens-based shuffle is accounted for in the transformations.

Another weakness of the BCS method is that it reveals little of the physical process of the shuffle. We shall find that defining input-to-output shuffle relations (as well as output-to-input) according to the goals above reveals more physical information about the lens-based shuffle than the BCS method. This is because channel positions at both input and output are specified relative to several different *landmarks* in the system, such as the system's central axis and the optic axis of the imaging lens. The latter, for example, is useful in computing the image quality of any given output point.

3.4 Overview of the New Approach

We describe the source and detector arrays (or input and output) as matrices with an (i,j) coordinate system, where the coordinates satisfy $(1,...,i,...n,1,...,j,...n)$, and n is even. We construct a physical coordinate system (p,q) which specifies the spatial location of channels relative to the system's central axis. We show the coordinate transformation between matrix coordinates (i,j) and intermediate coordinates (p,q). We then develop the relationship between an input (p,q) and its shuffled output (p',q'). In general, the procedure looks like

1. Coordinate transform: $(i,j) \Longrightarrow (p,q)$
2. Optical shuffle: $(p,q) \Longrightarrow (p',q')$
3. Coordinate transform: $(p',q') \Longrightarrow (i',j')$

The reader will see that this method is certainly no simpler than the BCS method of relating input to output coordinates. It does, however, provide the user with results in standard matrix coordinates, without any inversion between input and output coordinate systems. Thus, the choice of methods is purely a matter of taste.

3.4.1 Intermediate Coordinate System

The intermediate coordinate system (p, q) is depicted in Fig. 3.1, where p and q are defined as integer values on a normal (x, y) Cartesian coordinate system. The origin of the (p, q) coordinate system is where the central axis of the shuffle system intersects the input plane. The coordinates p and q of a channel in the array represent its position as the number of channels away from the system axis (i.e. the physical center of the array).

The actual physical location (x, y) of a given channel, normalized to the spacing δ between channels, differs from the channel's coordinate (p, q) by $\pm\frac{1}{2}$. This is because the number of channels n in the array in 1-D must be even in order to perform a perfect shuffle. Therefore, no channel may lie at the origin of the coordinate system, nor at any integer multiple of δ from the origin. The channel nearest the origin in the first quadrant has a physical location of $(x, y) = (\frac{\delta}{2}, \frac{\delta}{2})$. Being one channel away from the origin in each direction, its location in the intermediate coordinate system is $(p, q) = (1, 1)$. In general, one adds or subtracts $\frac{1}{2}$ to convert from (x, y) to (p, q) depending upon the quadrant of the channel. The relationship between (x, y) and (p, q) is

$$p = \frac{1}{\delta}x + \frac{1}{2}\text{sgn}(x) \tag{3.4}$$

$$q = \frac{1}{\delta}y + \frac{1}{2}\text{sgn}(y) \tag{3.5}$$

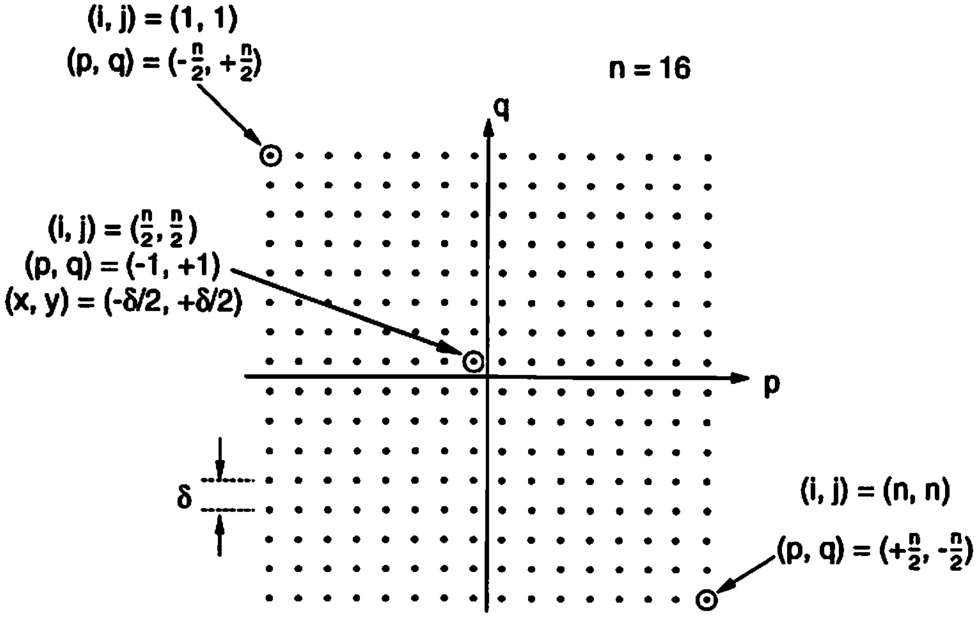


Figure 3.1: Intermediate coordinate system (p, q) and several sample points with their (i, j), (p, q) and (x, y) coordinates.

where the function sgn(·) is the sign of the argument, defined by

$$\operatorname{sgn}(x) = \begin{cases} +1 & ; x > 0 \\ 0 & ; x = 0 \\ -1 & ; x < 0 \end{cases} \tag{3.6}$$

To better understand the relationship between matrix coordinates (i, j), Cartesian coordinates (x, y), and the intermediate coordinates (p, q), consider Fig. 3.1, which shows a few sample points in an array, and their corresponding coordinates in each notation. The array has $n = 16$ channels in each row and column. Matrix coordinates (i, j) ordinarily start at 1 in the upper left corner of the array, and increase to n down and to the right. The point in the upper left of the array is $(i, j) = (1, 1)$. In Cartesian coordinates, with the origin at the center of the array, the upper left point is

$$(x,y) = (-[\frac{n}{2} - \frac{1}{2}]\delta, +[\frac{n}{2} - \frac{1}{2}]\delta) = (-7\frac{1}{2}\delta, +7\frac{1}{2}\delta). \tag{3.7}$$

In the intermediate coordinate system, the same point is $\frac{n}{2}$ channels to the left and above the origin, so

$$(p,q) = (-\frac{n}{2}, +\frac{n}{2}) = (-8, +8). \tag{3.8}$$

A channel close to the origin is

$$(i,j) = (\frac{n}{2}, \frac{n}{2}) = (8, 8), \tag{3.9}$$

whose physical location is given by

$$(x,y) = (-\frac{\delta}{2}, +\frac{\delta}{2}). \tag{3.10}$$

This point is one channel to the left, and one channel above the origin, so its intermediate coordinates are

$$(p,q) = (-1, +1). \tag{3.11}$$

3.4.2 Transformations between Matrix and Intermediate Coordinates

If the coordinates of the input are given as matrix coordinates (i,j), then we seek a transformation both to and from the intermediate coordinates (p,q). The reader can verify that the transformations are given by

$$p = (j - \frac{n}{2}) - \mathrm{u}[\frac{1}{2} - (j - \frac{n}{2})] \tag{3.12}$$

$$q = (\frac{n}{2} - i) + \mathrm{u}[(\frac{n}{2} - i) + \frac{1}{2}] \tag{3.13}$$

$$i = [\mathrm{u}(q) - q] + \frac{n}{2} \tag{3.14}$$

$$j = [\mathrm{u}(-p) + p] + \frac{n}{2}, \tag{3.15}$$

where $\mathrm{u}(x)$ is the unit step function defined by

$$\mathrm{u}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \tag{3.16}$$

3.5 Calculating the Shuffled Output of a Given Input

3.5.1 Separable Perfect Shuffle

We develop the relation between input and shuffled coordinates in the intermediate coordinate system (p, q). The lenses in the optical shuffle are simple positive lenses which form real images with a magnification of two. They take the input (or object) coordinates and reflect them through the optic axis of the lens, and double the distance from the optic axis. So, if we start with a source at (p, q), then to find its shuffled coordinate (p', q'), we need to know the distance of the channel from the optic axis of the lens nearest the channel. We do this because its shuffled output is its image through the lens nearest the input, and measured relative to the optic axis of the lens, the image position is found easily.

Input

Sawchuk and Glaser [31] have shown that, for the separable shuffle, a point at the input position (p, q) is imaged to its shuffled output by a lens centered at

$$(x_{l,s}, y_{l,s}) = \frac{n-1}{6}\delta(\mathrm{sgn}(p), \mathrm{sgn}(q)), \tag{3.17}$$

where n is the number of channels in 1-D (i.e. a row or column) of the array, the subscript l denotes the position of a lens, and the subscript s denotes the separable shuffle. If we know the channel's input position in the intermediate coordinates (p, q), we find its Cartesian coordinates using

$$x = [p - \frac{1}{2}\mathrm{sgn}(p)]\delta \tag{3.18}$$

$$y = [q - \frac{1}{2}\mathrm{sgn}(q)]\delta. \tag{3.19}$$

Positive x indicates that the channel is to the right of the origin, and positive y means that it is above the origin. To find the distance of a channel input to the optic axis, we add the displacement $(x_{l,s}, y_{l,s})$ of the lens from the system axis to the channel position (x, y). This is called the ***object height*** h, and is given by

$$h_{x,s} = x - x_{l,s} \tag{3.20}$$

$$h_{y,s} = y - y_{l,s}. \tag{3.21}$$

where the subscripts x and y indicate the x or y component of the object height, and the subscript s indicates a separable shuffle. When the channel and lens positions are expressed as functions of the channel coordinates p and q, we find the x and y components of the object height also as functions of the input position

$$h_{x,s} = [p - \mathrm{sgn}(p)\frac{n+2}{6}]\delta \tag{3.22}$$

$$h_{y,s} = [q - \mathrm{sgn}(q)\frac{n+2}{6}]\delta, \tag{3.23}$$

Positive h_x means the channel is to the right of the optic axis, and negative to the left. Positive h_y indicates the source is above the optic axis, and negative indicates below.

Output

The optical shuffle inverts the image (output) about the optic axis, and magnifies its distance to the optic axis by two. Thus, the *image height* h' is

$$(h'_{x,s}, h'_{y,s}) = -2(h_{x,s}, h_{y,s}). \tag{3.24}$$

Using Eqs. (3.22) and (3.23), the components of the image height are

$$h'_{x,s} = [-2p + \text{sgn}(p)\frac{n+2}{3}]\delta \tag{3.25}$$

$$h'_{y,s} = [-2q + \text{sgn}(q)\frac{n+2}{3}]\delta. \tag{3.26}$$

We find the output in Cartesian coordinates by adding the lens shift from the system axis to the image height h'

$$x'_s = h'_{x,s} + x_{l,s} \tag{3.27}$$

$$y'_s = h'_{y,s} + y_{l,s}, \tag{3.28}$$

which, using Eq. (3.17), simplify to

$$x'_s = [-2p + \text{sgn}(p)\frac{n+1}{2}]\delta \tag{3.29}$$

$$y'_s = [-2q + \text{sgn}(q)\frac{n+1}{2}]\delta. \tag{3.30}$$

Note that the function sgn($\cdot$) is taken on the input coordinates p and q, because the input quadrant determines which optic axis (or which lens) is relevant. The shuffled output coordinates in terms of p and q are obtained using Eqs. (3.4) and (3.5),

$$p'_s = \frac{x'_s}{\delta} + \frac{1}{2}\text{sgn}(x'_s) \tag{3.31}$$

$$q'_s = \frac{y'_s}{\delta} + \frac{1}{2}\text{sgn}(y'_s). \tag{3.32}$$

General Procedure

We now have the equations to map any input channel to its shuffled output for the optical separable shuffle. The general approach is:

1. Given an input in (i, j) matrix coordinates, convert (i, j) to the intermediate coordinates (p, q) by means of Eqs. (3.12) and (3.13).
2. Use Eqs. (3.29) and (3.30) to calculate the position of the output in x and y.
3. Calculate the separable shuffle output in the intermediate coordinates (p', q') using Eqs. (3.31) and (3.32).
4. Convert the output from intermediate coordinates (p', q') to matrix coordinates (i', j') using Eqs. (3.14) and (3.15).

3.5.2 Folded Perfect Shuffle

Lens Positions

To obtain the output coordinates of the optical folded shuffle, we make minor modifications to the equations presented in Section 3.5.1. Sawchuk and Glaser [31] have shown that a folded shuffle may be also implemented with lenses much like a separable shuffle. The only difference is a slight offset of two of the four lenses. The lens positions in each of the four quadrants is given by

$$(x_{l,f}, y_{l,f}) = \begin{cases} (+\frac{n+1}{6}\delta, +\frac{n+1}{6}\delta) \\ (-\frac{n-1}{6}\delta, +\frac{n-1}{6}\delta) \\ (-\frac{n+1}{6}\delta, -\frac{n+1}{6}\delta) \\ (+\frac{n-1}{6}\delta, -\frac{n-1}{6}\delta) \end{cases} \tag{3.33}$$

where the subscript l indicates the lens position, and the subscript f indicates the folded shuffle. We see that the lens position is $(\frac{n+1}{6})\delta$ in quadrants where x and y

(or p and q) are either both positive or both negative. The lens position is $(\frac{n-1}{6})\delta$ (the same as for the separable shuffle) in quadrants where p and q have opposite signs. A compact notation for this is

$$(x_{l,f}, y_{l,f}) = (\mathrm{sgn}(p)\frac{n + \mathrm{sgn}(pq)}{6}\delta, \mathrm{sgn}(q)\frac{[n + \mathrm{sgn}(pq)]}{6}\delta) \tag{3.34}$$

Input

An input at (p, q) has the same Cartesian coordinates (x, y) as in the separable case, given by Eqs. (3.18) and (3.19). The distance of a source to the closest (and therefore relevant) optic axis is not the same in each quadrant, however, as it is for the separable perfect shuffle. Adding or subtracting the position $(x_{l,f}, y_{l,f})$ of the optic axes for the folded shuffle, according to Eq. (3.34), gives the object height h for the folded shuffle

$$h_{x,f} = x - x_{l,f} = [p - \mathrm{sgn}(p)(\frac{n + 3 + \mathrm{sgn}(pq)}{6})]\delta \tag{3.35}$$

$$h_{y,f} = y - y_{l,f} = [q - \mathrm{sgn}(q)(\frac{n + 3 + \mathrm{sgn}(pq)}{6}]\delta, \tag{3.36}$$

where the subscript f indicates the folded shuffle.

Output

The same magnification by two and inversion through the optic axis occurs in the folded shuffle as in the separable shuffle. Thus, the folded shuffle output are slightly different from Eqs. (3.25) and (3.26), and are given by

$$h'_{x,f} = -2h_{x,f} = [-2p + \mathrm{sgn}(p)\frac{n + 3 + \mathrm{sgn}(pq)}{3}]\delta \tag{3.37}$$

$$h'_{y,f} = -2h_{y,f} = [-2q + \mathrm{sgn}(q)\frac{n + 3 + \mathrm{sgn}(pq)}{3}]\delta. \tag{3.38}$$

We find the output in Cartesian coordinates by adding or subtracting the lens shift $(x_{l,f}, y_{l,f})$ for the folded shuffle to the image height h'

$$x'_f = h'_{x,f} + x_{l,f} \tag{3.39}$$

$$y'_f = h'_{y,f} + y_{l,f} \tag{3.40}$$

Using the lens positions in Eq. (3.34), we obtain the folded shuffle output in Cartesian coordinates as a function of input position (p, q)

$$x'_f = [-2p + \mathrm{sgn}(p)(\frac{n+2+\mathrm{sgn}(pq)}{2})]\delta \tag{3.41}$$

$$y'_f = [-2q + \mathrm{sgn}(q)(\frac{n+2+\mathrm{sgn}(pq)}{2})]\delta. \tag{3.42}$$

Just as in the separable case, the function $\mathrm{sgn}(\cdot)$ is taken on the *input* coordinates p and q, because the input quadrant determines which optic axis (or which lens) is relevant. The shuffled output coordinates in terms of p and q are

$$p'_f = \frac{x'_f}{\delta} + \frac{1}{2}\mathrm{sgn}(x'_f) \tag{3.43}$$

$$q'_f = \frac{y'_f}{\delta} + \frac{1}{2}\mathrm{sgn}(y'_f). \tag{3.44}$$

General Procedure

We now have the equations to map any input channel to its shuffled output for the optical folded shuffle. The general approach is:

1. Given an input in (i, j) matrix coordinates, convert (i, j) to the intermediate coordinates (p, q) by means of Eqs. (3.12) and (3.13).
2. Use Eqs. (3.41) and (3.42) to calculate the the folded shuffle output in x and y.

3. Convert the separable shuffle output to the intermediate coordinates (p', q') using Eqs. (3.43) and (3.44).
4. Convert the output from intermediate coordinates (p', q') to matrix coordinates (i', j') using Eqs. (3.14) and (3.15).

3.6 Calculating the Input Coordinate Given an Output

The input-to-output mappings for both the separable and folded shuffles are one-to-one, so in principle, we may write their inverse relations. Such relationships are useful if one wishes to know what input to a shuffle stage produces a given output. For example, if an external controller sets all of the bypass-exchange switches in the system, it must have an algorithm to determine how to connect network inputs with their intended outputs. One approach to the problem is a bidirectional search, in which the controller proceeds forward through the system from the known input, and backward from the desired output, trying to establish the proper connections. Such an algorithm needs both the input-to-output relations, which we have already derived, and the output-to-input relations, which we develop in this section.

3.6.1 Separable Perfect Shuffle

Output

The distance, in number of channels, between an output channel (p', q') and the channel at the edge of the output array determines the corresponding input point (p, q). Consider a vertical slice of the lens-based shuffle, shown in Fig. 3.2. Define the distance k from an output point (p', q') to the nearest edge of the output array as

$$(k_x, k_y) = (p' - \mathrm{sgn}(p')\frac{n}{2}, q' - \mathrm{sgn}(q')\frac{n}{2}) \tag{3.45}$$

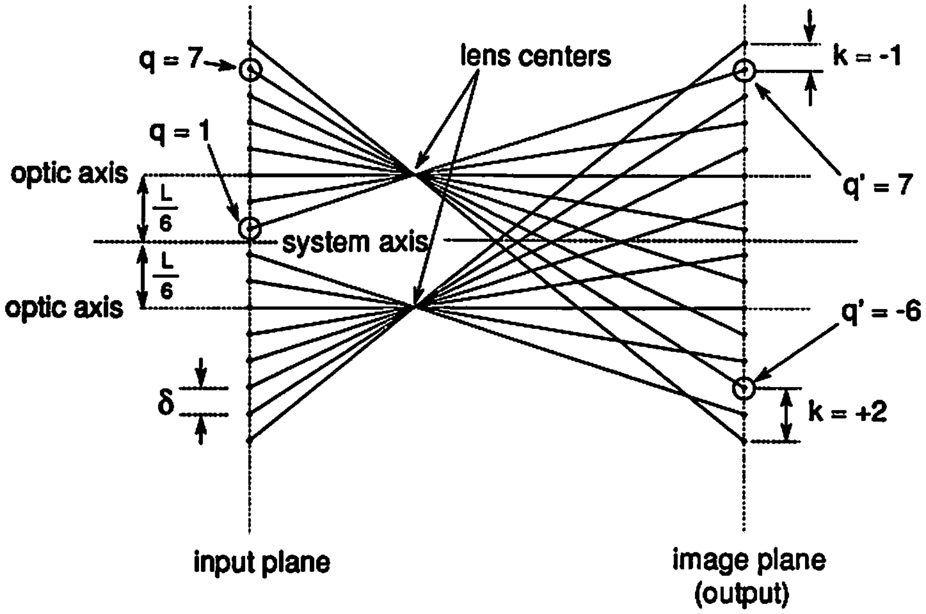


Figure 3.2: Vertical slice of the optical implementation of the perfect shuffle for $n = 16$. Indicated are $q' = +7$ and $q' = -6$, for which $k_y = -1$, meaning one channel below the top of the array, and $k_y = +2$, meaning two channels above the bottom of the array. The corresponding inputs are $q = +1$ and $q = +7$ respectively.

where positive k_x indicates that (p', q') is to the right of (and nearest to) the left edge of the array, and positive k_y indicates that (p', q') is above (and nearest to) the bottom edge of the array.

Input

We see that if k is even for the given output, then the corresponding input is located in the opposite half of the array. So, the sign of the input coordinate is the negative of the output coordinate

$$k_y \text{ even } \Longrightarrow \operatorname{sgn}(q) = -\operatorname{sgn}(q'), \tag{3.46}$$

where we use the vertical coordinate q only to refer to Fig. 3.2; the horizontal coordinate is identically applicable. If k is odd for a given output, then the corresponding input is on the same half of the array, or the input and output coordinates have the same sign

$$k_y \text{ odd } \Longrightarrow \operatorname{sgn}(q) = \operatorname{sgn}(q'). \tag{3.47}$$

We obtain the quantitative relationships from output to input by demagnifying the output by a factor of two and inverting, so that

$$p = \begin{cases} -\operatorname{sgn}(p')\frac{n}{2} - \frac{k_x}{2} & ;\ k_x \text{ even} \\ +\operatorname{sgn}(p')\lceil\frac{|k_x|}{2}\rceil & ;\ k_x \text{ odd} \end{cases} \tag{3.48}$$

$$q = \begin{cases} -\operatorname{sgn}(q')\frac{n}{2} - \frac{k_y}{2} & ;\ k_y \text{ even} \\ +\operatorname{sgn}(q')\lceil\frac{|k_y|}{2}\rceil & ;\ k_y \text{ odd} \end{cases} \tag{3.49}$$

where $\lceil\cdot\rceil$ is the *ceiling* function defined as the lowest integer greater than or equal to the argument.

General Procedure

We now have the equations to determine the input channel which produced a given separable shuffle output. The general approach is:

1. Given the output (i', j') of separable shuffle, convert (i', j') to the intermediate coordinates (p', q') by means of Eqs. (3.12) and (3.13).

2. Calculate the distances k_x and k_y both using Eq. (3.45).

3. Calculate the separable shuffle input in the intermediate coordinates p and q both using Eqs. (3.48) and (3.49).

4. Convert the input from intermediate coordinates (p, q) to matrix coordinates (i, j) using Eqs. (3.14) and (3.15).

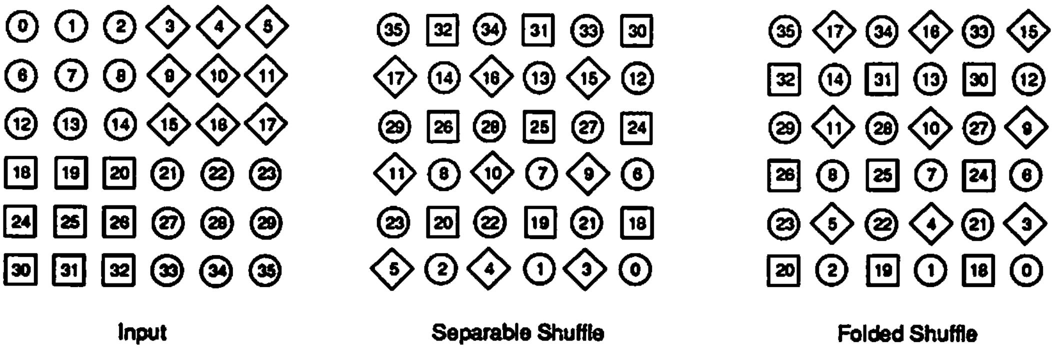


Figure 3.3: Similarities and differences between the optical separable and folded shuffles. Similarities are shown with circles and differences with either squares or diamonds.

3.6.2 Folded Perfect Shuffle

Output Relative to Separable Case

To define the output-to-input relationship for the folded shuffle, we examine Fig. 3.3, which shows both the optically obtained separable shuffle and folded shuffle. There is a simple relationship between the two shuffles. First, half of the points in the folded shuffle output are the same as those in the separable shuffle. These output points are circled in the figure, and their matrix coordinates i'_f and j'_f are either both even, or both odd. In this case, Eqs. (3.48) and (3.49) are immediately applicable, so

$$i', j' \text{ both even or both odd } \Longrightarrow p'_f = p'_s\ , \ p_f = p_s. \tag{3.50}$$

Output points where the indices i'_f and j'_f are different, such that one is even and one is odd, are enclosed in either a square or a diamond. Squares enclose folded output points where the folded shuffle output column number j'_f is odd. In this case, the folded output is the separable output shifted down and to the left one position

$$j'_f \text{ Odd } \Longrightarrow (p'_f, q'_f) = (p'_s, q'_s) - (1, 1). \tag{3.51}$$

Where the folded output column number is even, shown in the figure with diamonds around the output, the output is the separable shuffle shifted up and to the right one position

$$j'_f \text{ Even } \Longrightarrow (p'_f, q'_f) = (p'_s, q'_s) + (1,1). \tag{3.52}$$

General Procedure

We now have the equations to map any input channel to its shuffled output for the optical folded shuffle. The general approach is:

1. If i' and j' are either both even or both odd, then their corresponding inputs i and j are the same as for the separable shuffle. So, the procedure from Section 3.6.1 yields the correct input (i, j).
2. If i' and j' are different (i.e. one even and one odd), then convert the folded output to intermediate coordinates (p', q').
3. For the case where i' and j' are different, shift the intermediate output coordinates (p', q') according to Eqs. (3.51) or (3.52), depending on whether j' is odd or even. Then use the procedure from Section 3.6.1 to find the input coordinates.

3.7 Examples

3.7.1 Separable Shuffle Input to Output

If we have a 16 × 16 source array, and choose to illuminate source $(i, j) = (2, 8)$, as shown in Fig. 3.4 we may find the coordinates of the separable shuffle output (i'_s, j'_s). First, we convert the input (i, j) to the intermediate coordinate system (p, q). Using Eq. (3.12), we see that

$$p = (8 - \frac{16}{2}) - \mathrm{u}[\frac{1}{2} - (8 - \frac{16}{2})] = -1. \tag{3.53}$$

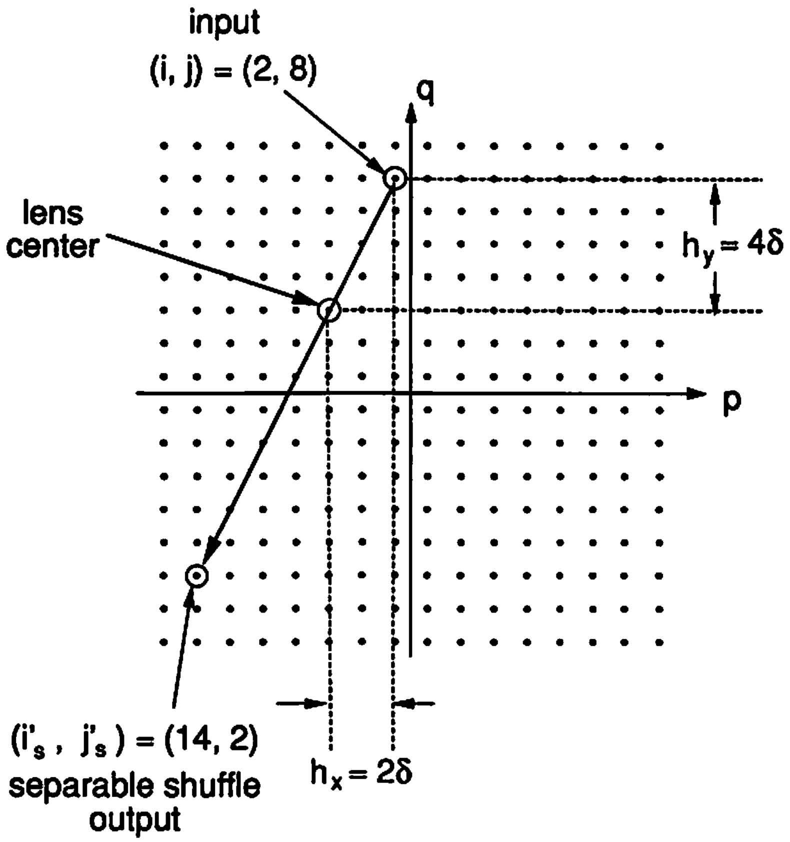


Figure 3.4: Effect of optically shuffling the point $(i, j) = (2, 8)$ to its output at $(i', j') = (14, 2)$. The heights of the object (source) and image (output) are indicated, as well as the (p, q) coordinates of input and output.

Likewise, using Eq. (3.13), q is

$$q = (\frac{16}{2} - 2) + \mathrm{u}[(\frac{16}{2} - 2) + \frac{1}{2}] = 7. \tag{3.54}$$

Note that the conversions from (i, j) to (p, q) and back are simple enough that one may not need Eqs. (3.12) and (3.13) to do the conversion. If the array is not too large to make it impractical, a visual count from the center of the array reveals (p, q).

Next, we calculate the (x, y) coordinates of (p, q), which we get from Eqs. (3.18) and (3.19)

$$x = [-1 - \frac{1}{2}\text{sgn}(-1)]\delta = -\frac{1}{2}\delta \tag{3.55}$$

$$y = [7 - \frac{1}{2}\text{sgn}(7)]\delta = 6\frac{1}{2}\delta. \tag{3.56}$$

We obtain the object height $(h_{x,s}, h_{y,s})$ (of the input) using Eqs. (3.22) and (3.23)

$$h_{x,s} = -\frac{1}{2}\delta - \text{sgn}(-1)\frac{16-1}{6}\delta = +2\delta \tag{3.57}$$

$$h_{y,s} = 6\frac{1}{2}\delta - \text{sgn}(7)\frac{16-1}{6}\delta = +4\delta. \tag{3.58}$$

The horizontal and vertical components of the object height are indicated in Fig. 3.4, and the implied location of the lens. To get the shuffled output coordinates, we find the height of the image using Eqs. (3.25) and (3.26)

$$h'_{x,s} = [-2(-1) + \text{sgn}(-1)\frac{16+2}{3}]\delta = -4\delta \tag{3.59}$$

$$h'_{y,s} = [-2(7) + \text{sgn}(7)\frac{16+2}{3}]\delta = -8\delta. \tag{3.60}$$

The Cartesian coordinates of the output are

$$x'_s = [-2(-1) + \text{sgn}(-1)\frac{16+1}{2}]\delta = -6\frac{1}{2}\delta. \tag{3.61}$$

$$y'_s = [-2(7) + \text{sgn}(7)\frac{16+1}{2}]\delta = -5\frac{1}{2}\delta. \tag{3.62}$$

According to Eqs. (3.31) and (3.32), the output in terms of the intermediate coordinates is

$$p'_s = -6\frac{1}{2} + \frac{1}{2}\text{sgn}(-6\frac{1}{2}) = -7 \tag{3.63}$$

$$q'_s = -5\frac{1}{2} + \frac{1}{2}\text{sgn}(-5\frac{1}{2}) = -6. \tag{3.64}$$

Finally, we convert back to matrix coordinates using Eqs. (3.14) and (3.15)

$$i'_s = \mathrm{u}(-6) - (-6) + \frac{16}{2} = 14 \tag{3.65}$$

$$j'_s = \{\mathrm{u}[-(-7)] + (-7)\} + \frac{16}{2} = 2. \tag{3.66}$$

Figure 3.4 shows the input point $(2,8)$ and the optically shuffled output at point $(14,2)$. If we try another point in the same quadrant, for example $(i,j) = (1,2)$, then by the same procedure we have $(i',j') = (16,14)$. Recall that sources in the same quadrant are imaged through the same lens. In Fig. 3.5, we superimpose the mapping from $(1,2)$ to $(16,14)$ onto the previous mapping of Fig. 3.4. The lines connecting input to output intersect at the point where the lens center is indicated in Fig. 3.4, which makes physical sense because the undeviated rays from input to output go through the center of the lens.

The procedure we have outlined also works for input channels shifted slightly by (ϵ_x, ϵ_y) away from its intended position in the matrix. For example, if we place an input $\frac{1}{3}\delta$ to the right of $(1,2)$ then $(\epsilon_x, \epsilon_y) = (\frac{\delta}{3}, 0)$ and $(i,j) = (1, 2\frac{1}{3})$. Using the fractional value of j in the procedure shows that the input is located at $p = -6\frac{2}{3}$, and the output is located at $p' = 5\frac{1}{3}$, or $j' = 13\frac{1}{3}$.

3.7.2 Folded Shuffle Output to Input

To work backwards from a folded shuffle output of $(i'_f, j'_f) = (12,15)$, which we convert to $(p'_f, q'_f) = (7,-4)$, we note that the output row i' is even, and the output column j' is odd. This means we use Eq. (3.51) to find the intermediate result (p'_s, q'_s). This gives the output had the shuffle been separable. More importantly, it reduces the problem of finding the corresponding input to using Eqs. (3.48) and (3.49), which are applicable to the separable shuffle. If the operation had been a separable shuffle instead of a folded shuffle, the output would have been

$$(p'_s, q'_s) = (p'_f, q'_f) + (1,1) = (8,-3). \tag{3.67}$$

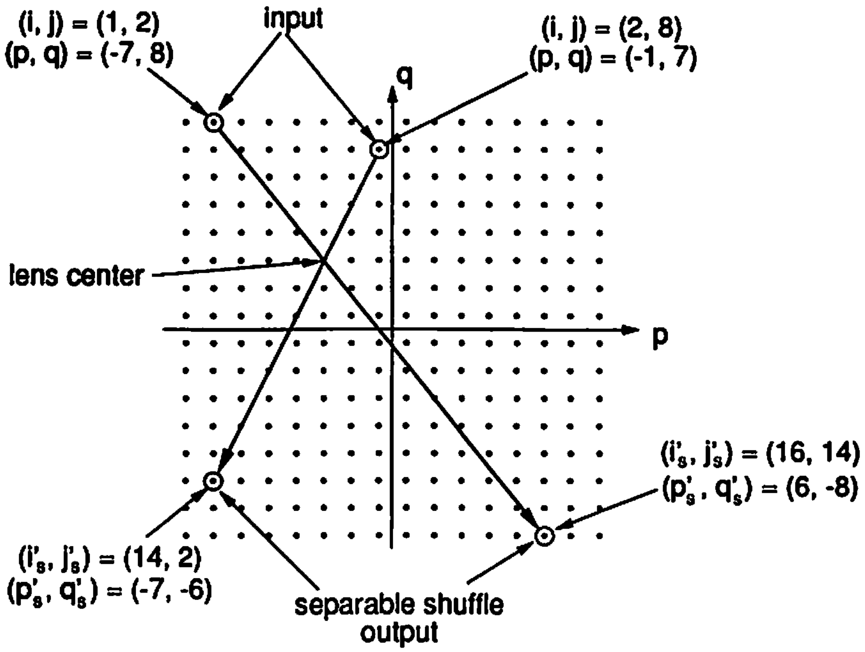


Figure 3.5: Mapping the optical shuffle of two points in the same quadrant shows the position of the lens center where the two lines cross; in this case at $\frac{15}{6}\delta = 2\frac{1}{2}\delta$.

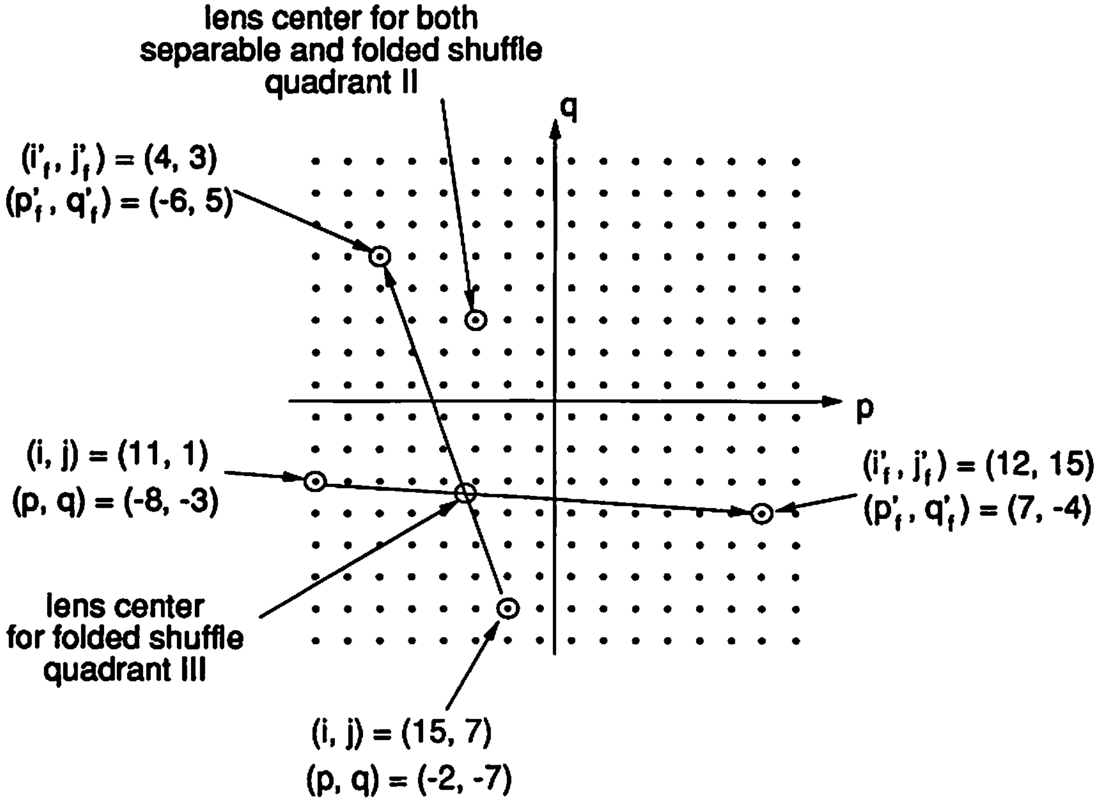


Figure 3.6: Folded shuffle output $(i'_f, j'_f) = (4, 3)$ and $(i'_f, j'_f) = (12, 15)$ and their corresponding inputs $(i, j) = (15, 7)$ and $(i, j) = (11, 1)$. Lines connecting input to output intersect at the lens position in the quadrant $(p < 0, q < 0)$, which shows the slight shift from its position in the separable shuffle, indicated in the quadrant $(p < 0, q > 0)$.

According to Eq. (3.45), the distance of (p'_s, q'_s) to the edge of the array is

$$k_x = 8 - \mathrm{sgn}(8)\frac{16}{2} = 0 \tag{3.68}$$

$$k_y = -3 - \mathrm{sgn}(-3)\frac{16}{2} = 5. \tag{3.69}$$

This means that the input must be located at

$$p = -\mathrm{sgn}(8)\left[\frac{16}{2} - \frac{0}{2}\right] = -8 \tag{3.70}$$

$$q = +\mathrm{sgn}(-3)\left\lceil\frac{5}{2}\right\rceil = -3. \tag{3.71}$$

Converting from the intermediate coordinates (p, q) to the matrix coordinates (i, j), we find that the input which produced the folded shuffle output at $(i'_f, j'_f) = (12, 15)$ is $(i, j) = (11, 1)$.

The folded shuffle output at $(12, 15)$ and its corresponding input are shown in Fig. 3.6, as well as another sample output $(i'_f, j'_f) = (4, 3)$ and its corresponding input. The lines connecting inputs to outputs intersect at the position where the lens must be in quadrant *III* (where p and q are both negative) to implement the folded shuffle. Had the shuffle been separable (on same 16×16 array), the lens would be positioned slightly above and to the right at $(p, q) = (-3, -3)$. In the lens-based implementation of the folded shuffle, however, only two of the four lenses are shifted relative to their positions in the separable shuffle implementation. According to Eq. (3.34), the lenses which are shifted are in quadrants *I* $(p, q > 0)$ and *III*, where p and q are either both positive or both negative. In quadrant *II*, where $p < 0$ and $q > 0$, and quadrant *IV*, where $p > 0$ and $q < 0$, lenses remain in the same position as in the separable shuffle. Figure 3.6 shows the position of the lens in quadrant *II* for both the separable and folded shuffles.

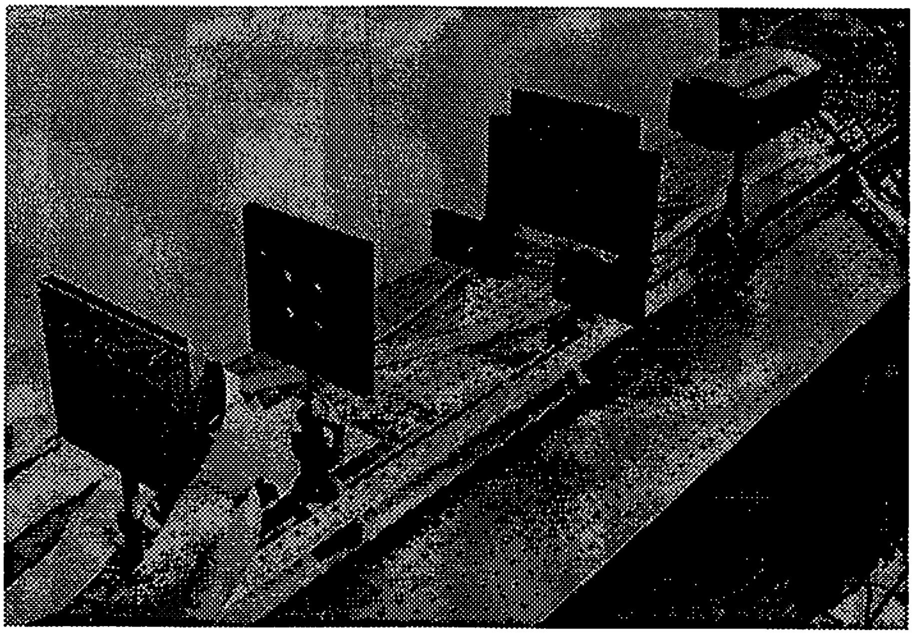


Figure 3.7: Experimental demonstration of the lens-based shuffle.

3.8 Implementation

Both the separable perfect shuffle and the folded shuffle have been implemented in our laboratory by C.B. Kuznia, the laboratory set-up for which is shown in Fig. 3.7. The source array consisted of 256 LEDs in a 16×16 grid, which measured $L = 11.5$ cm on a side, and is pictured in Fig. 3.8. The lens arrays for both the separable and folded shuffles, shown in Figs. 3.9 and 3.10 respectively, all had focal lengths $f = 10$ cm. The input that was calculated to yield "USC SIPI" for both the separable and folded shuffles was injected into the LED array (shown on the left in Figs. 3.11 and 3.12), and shuffled by the optics to obtain the experimental output on the right.

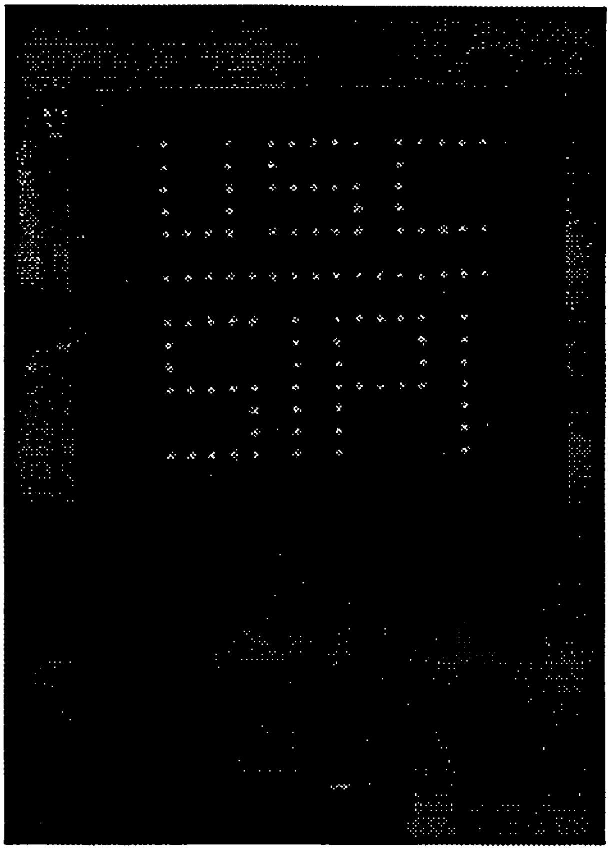


Figure 3.8: 16 × 16 array of LEDs used as input.

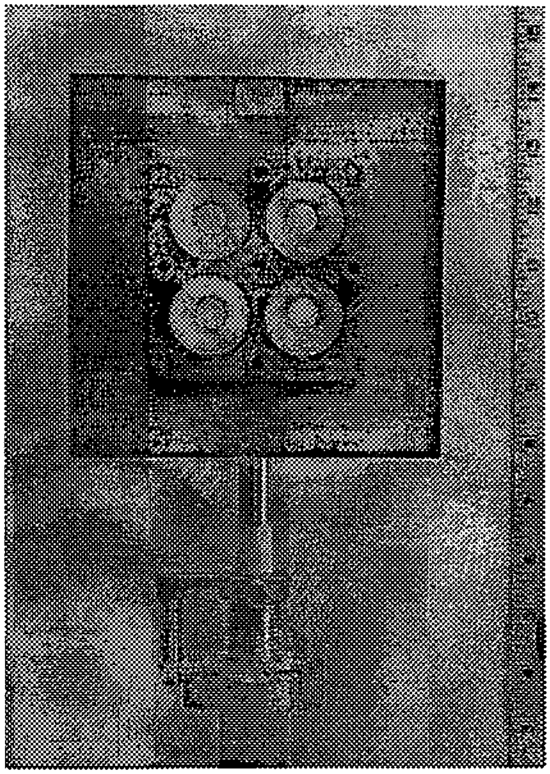


Figure 3.9: 2 × 2 array of lenses for the separable shuffle.

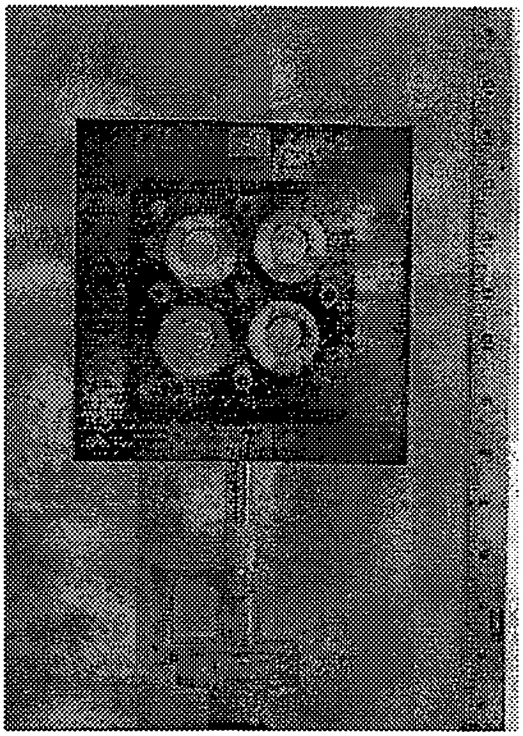


Figure 3.10: 2 × 2 array of lenses for the folded shuffle (with corner lenses slightly skewed compared to the separable case).

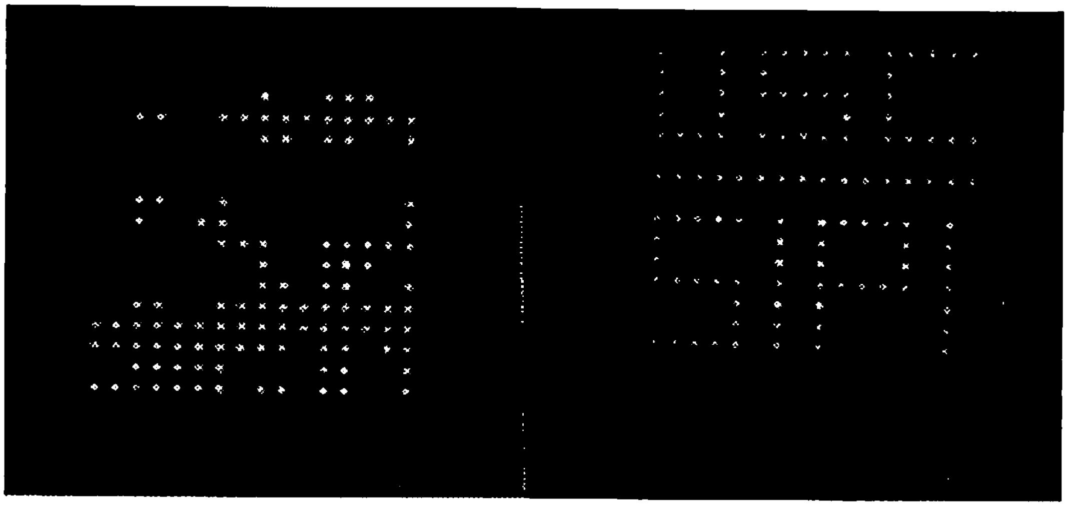


Figure 3.11: On the left is the input calculated to have as its separable shuffle output USC SIPI, and the experimentally obtained result on the right.

3.9 Conclusion

We have reviewed the established binary cyclic shift method of relating input to shuffled output, and discussed a few inconvenient characteristics of this method. We then developed input-to-output and output-to-input rules which incorporate the output inversion inherent in the lens-based optical shuffle. The rules are presented as simple procedures at the end of each section on the separable and folded shuffles. These procedures are easily programmable to simulate either type of optical shuffle. This is useful for developing and testing controller and routing algorithms for multistage shuffle-exchange switching networks. Finally, we have illustrated the use of the procedures with examples.

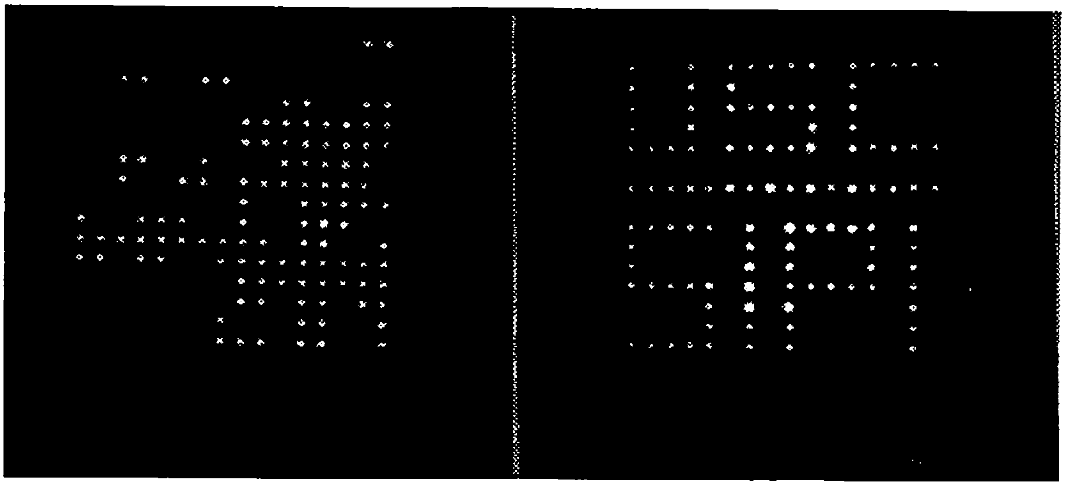


Figure 3.12: On the left is the input calculated to have as its folded shuffle output USC SIPI, and the experimentally obtained result on the right.

Chapter 4

Capabilities of Simple Lenses in a Free Space Perfect Shuffle

4.1 Introduction

Our architecture is the 2-D separable perfect shuffle implemented on an $n \times n$ array of inputs. We assume that the communication channels are input to the fixed optical shuffle stages by an array of optical sources or fibers. The shuffled output signals impinge on an array of detectors that convert the signals to electrical form for bypass-exchange processing. Finally, electronics drives an array of lasers or other optical sources for succeeding optical shuffle stages.

We expect that acousto-optical or opto-mechanical techniques will have lower packing density, as well as lower speed and reliability than the optoelectronic techniques we adopt for our architecture. It is the issue of channel packing density which we shall examine in detail in this chapter. Several interconnection schemes have been proposed for deflecting collimated beams of light to desired output. For example, Nakagami *et al* constructed a system using prisms to steer beams through a Banyan network, and expect to switch between 32 and 64 channels [28]. We expect that the ability of lenses to focus output beams to a point will increase

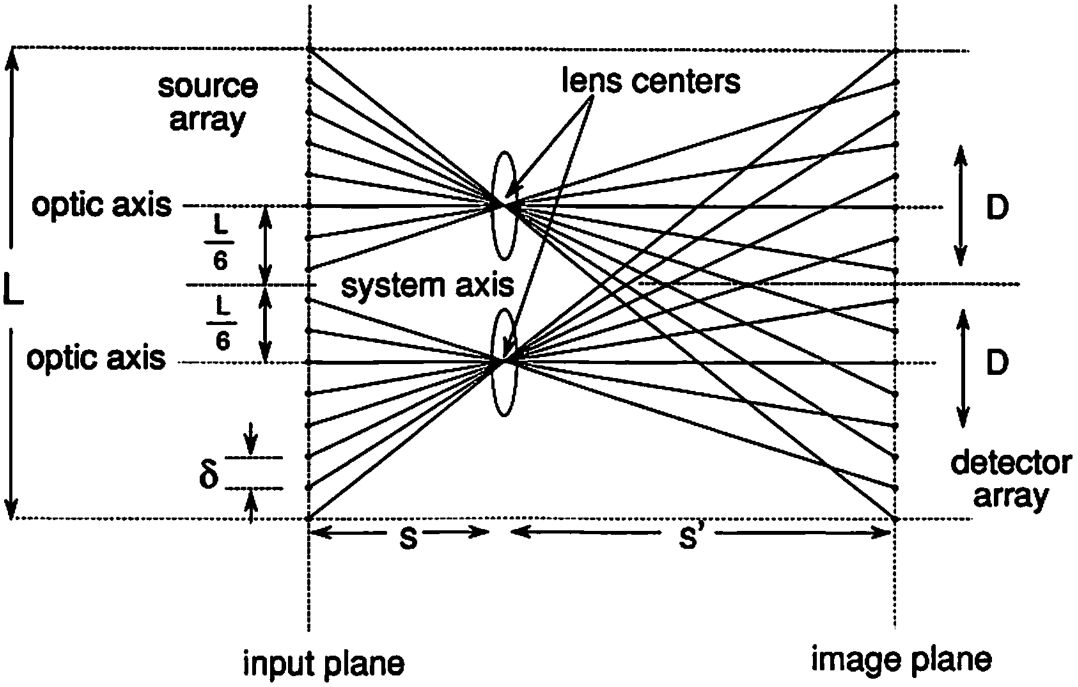


Figure 4.1: Vertical cross-section of perfect shuffle implemented with lenses. δ is the center-to-center spacing of input (and output) channels, L is the physical size of the array in 1-D, D is the diameter of the lenses, s and s' are the object and image distances.

the number of channels resolvable in an interconnection network. Our primary motivation for examining the lens-based architecture, however, is its simplicity.

Figure 4.1 shows a vertical cross-section of the system. As shown by Sawchuk and Glaser [31] (and discussed in Section 3.5.1), the lenses must be centered at horizontal and vertical displacements from the central axis of the system given by

$$x, y = \pm \frac{1}{6}(n-1)\delta, \tag{4.1}$$

where n is the number of channels in 1-D to be connected, and δ is the center-to-center spacing of the sources in the input array. The quantity $(n-1)\delta$ is equal to the physical length, L, of one side of the array, so the lenses are centered $\frac{1}{6}L$ away from the system axis in both x and y. In Fig. 4.1, s is the distance from the source array plane to the lens plane, s' is the distance from the lens plane to the detector array plane, and D is the diameter of each lens. The imaging arrangement is such that the lenses each magnify the source array by two. The relative lens positions given by Eq. (4.1) assure that the four images interlace the quadrants of the array to accomplish the shuffle.

4.2 Optical Limits on Resolvable Spot Size

4.2.1 Architectural Assumptions

We choose to restrict our analysis here to simple, single element lenses with spherical surfaces. Simple lenses are attractive because the analysis which follows is entirely analytic. The rules governing the imaging properties of simple lenses are thoroughly established, and do not require sophisticated computers or software to solve. We simply apply these rules to our architecture.

While the system may be improved by employing multi-element, aberration-corrected lenses, the analysis which follows provides a type of benchmark. For example, if one wishes to construct an optical shuffle which resolves a given number of channels in a given space, our results show whether or not it is possible to do so with simple lenses. Furthermore, if one *can* attain the desired specifications with simple lenses, our results show the optimum way in which to construct the optical shuffle. Thus, the person will know from his demands on resolvable channels and size whether to expend the effort and expense to invest in aberration-corrected lenses. Finally, the very same arguments used to analyze and optimize the simple lens system are equally valid with any sophisticated lens.

4.2.2 Qualitative Issues in Image Resolution

It is well known that increasing the aperture size in a diffraction-limited imaging system will increase the spatial resolution, and hence the number of independent resolvable data lines. Given the system of Fig. 4.1, the maximum aperture diameter of the lenses is

$$D_{max} = \frac{L}{3} \tag{4.2}$$

at which point their edges touch. In the system of Fig. 4.1, collimated beams from sources near the corner of the input array would miss the lenses entirely unless some other means of deflecting them toward the proper lens is employed (e.g. prisms). Otherwise, divergent or pseudo-point sources are necessary for this architecture, which images an input point to an output point.

It has been suggested (e.g. [39]) that an off-axis lens combination such as that shown in Fig. 4.2 might be employed to provide a larger aperture. This might be accomplished by using larger aperture lenses, and grinding away a portion of the aperture of each, so that four lenses could be assembled with their original centers (i.e. before grinding off a portion of the aperture) are still located at $(|x|, |y|) = (\frac{L}{6}, \frac{L}{6})$. This expands the aperture to accommodate collimated pencils of off-axis input light. Even for divergent sources, this arrangement has the advantage of greater light throughput because of the larger aperture. The price, however, is a general increase in the aberrations induced by rays passing through the lens farther from the optic axis. In keeping with our interest in solving the simplest problem in detail, we consider neither the prism option (something like that in [20]) nor the off-axis lens option in Fig. 4.2. Thus, the maximum radius for the lens aperture is one sixth the array size.

So, provided the optics remains diffraction-limited, increasing the aperture size up to the maximum available yields greater resolution. However, as the lens diameter increases while focal length and other system parameters are held constant, the lens aberrations become increasingly significant. In the following analysis, we look at the effective resolution of a single, simple, spherical lens, and analyze the number of channels that can be resolved with arrays of simple lenses. In this work, we use well known relationships on the spot sizes of systems having aberrations.

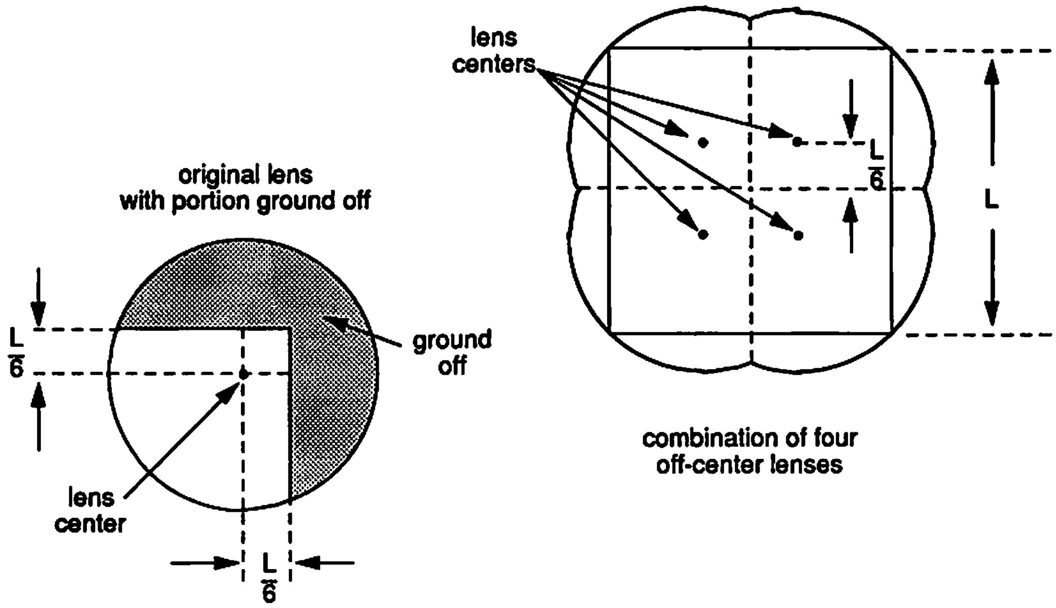


Figure 4.2: Off-center arrangement of lenses (either refractive or diffractive) to gather more light, or to allow use of collimated pencils of light, which for corner sources would otherwise miss the lenses altogether. We do not consider this enhancement in our calculations.

4.2.3 Assumptions on the Analysis

The analysis which follows is not exact. We assume that the lenses are thin, which means the lenses have no physical thickness. This simplifies the analysis of geometrical aberrations. We treat the two primary causes of image degradation, aberrations and diffraction, entirely separately. If one attempts an exact analysis of the problem, the aberrations induced by the lens must be incorporated into a diffraction integral. Numerical calculation of the diffraction integral (over the lens aperture) yields the image intensity at a point in the output plane. Repeating the integration for each point in the output plane for each point in the output plane gives the output intensity profile.

This approach has three fundamental disadvantages. First, it is quite difficult, requiring extensive computation on a reasonably fast computer (if one is to obtain results in a reasonable amount of time). The dramatic increase in computational complexity returns only marginally different results. Finally, the simpler analysis, which examines each of the various components of image degradation (i.e. diffraction and the various aberrations) in isolation actually yields greater physical insight into the design of the system than numerical calculation of a complicated diffraction integral.

4.2.4 Aberrations of a Thin Lens

General Spot Sizes

There are a total of five primary aberrations for any imaging system: spherical, coma, astigmatism, field curvature, and distortion [4], [23]. For a *thin* lens, distortion is zero. The four remaining types have ray distribution profiles shown in Fig. 4.3, and the extent (Δx) of each is given by

$$(\Delta x)_s = 16FA_s \tag{4.3}$$

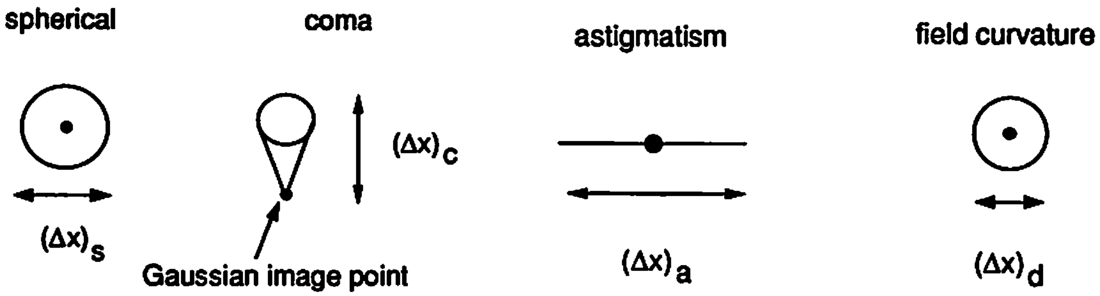


Figure 4.3: Aberration spot profiles. The solid black points represent the Gaussian image point, and the surrounding curve is the region within which geometric optics predicts rays strike the image plane. Δx is the greatest extent of the rays distribution.

for spherical,

$$(\Delta x)_c = 6FA_c \tag{4.4}$$

for coma,

$$(\Delta x)_a = 8FA_a \tag{4.5}$$

for astigmatism, and

$$(\Delta x)_d = 8FA_d \tag{4.6}$$

for field curvature, where

$$F = \frac{f}{D} \tag{4.7}$$

is the f-number of the lens, f and D are the focal length and diameter respectively, and A with an appropriate subscript is the maximum wave aberration of the lens.

Wave Aberrations

The corresponding wave aberrations are given by

$$A_s = a_s r_{max}^4 \tag{4.8}$$

$$A_c = a_c h' r_{max}^3 \tag{4.9}$$

$$A_a = a_a h'^2 r_{max}^2 \tag{4.10}$$

$$A_d = a_d h'^2 r_{max}^2 \tag{4.11}$$

where a is the appropriate aberration coefficient, r_{max} is the lens aperture radius, and h' is the image height, meaning the distance from a point in the image plane to the optic axis of the lens.

Imaging in the Shuffle

The maximum values of image height h' are determined by the channels at the farthest end of the array. For a 2-D array, the maximum occurs at the corner of the array where

$$h'_{max} = \sqrt{2} \cdot (\frac{2}{3} L). \tag{4.12}$$

For the optical shuffle system shown in Fig. 4.1, the *relative* distances, s and s', of the source and detector planes respectively to the lens are fixed because our arrangement requires a magnification of two. In an imaging arrangement, the distance of the source plane to the lens is ordinarily referred to as the object distance s, and the distance from the lens to the detector plane is the image distance s'. In order to obtain a magnification, M, of two, as required by the shuffle implementation, we must have

$$M = 2 = -\frac{s'}{s}, \tag{4.13}$$

where the negative sign indicates an inverted real image.

Optimum Lens Profile for the Shuffle

With the fixed ratio of s' to s needed for the shuffle, we can design the lens to minimize spherical aberration, which also virtually eliminates coma in the process [23]. We can accomplish this by choosing the proper radii of curvature of the lens surfaces, or equivalently, the proper "shape factor", q, which is defined by

$$q = \frac{R_2 + R_1}{R_2 - R_1}. \tag{4.14}$$

Here, R_1 and R_2 are the radii of curvature of the front and back surfaces of the lens, respectively. The q which minimizes spherical aberration is [23]

$$q_{min} = -2(1 - \frac{2f}{s'})(\frac{n^2 - 1}{n + 2}) \tag{4.15}$$

where n is the index of refraction of the lens.

Thin Lens Aberration Coefficients

The aberration coefficients of a lens designed for minimum spherical aberration are given by

$$a_{s-min} = \frac{-1}{32f^3}[(\frac{n}{n-1})^2 - \frac{n}{n+2}(1 - \frac{2f}{s'})^2] \tag{4.16}$$

$$a_c = \frac{1}{4nf^2s'}[(2n+1)(1 - \frac{2f}{s'}) + \frac{n+1}{n-1}q_{min}] \tag{4.17}$$

$$a_a = \frac{-1}{2fs'^2} \tag{4.18}$$

$$a_d = -\frac{n+1}{4nfs'^2}. \tag{4.19}$$

Actual Thin Lens Aberration Spot Sizes

If we substitute Eqs. (4.16)-(4.19) into Eqs. (4.8)-(4.11), and then use those results with Eq. (4.7) in Eqs. (4.3)-(4.6), we can write the various spot sizes as functions of focal length, index of refraction, aperture size, and image distance

$$(\Delta x) = g(f, n, D, s') \tag{4.20}$$

$$(\Delta x)_s = -\frac{1}{32f^2}[(\frac{n}{n-1})^2 - \frac{n}{n+2}(1 - \frac{2f}{s'})^2]D^3 \tag{4.21}$$

$$(\Delta x)_c = \frac{3}{16nfs'}(1 - \frac{2f}{s'})[(2n+1) - 2\frac{(n+1)^2}{n+2}]h'D^2 \tag{4.22}$$

$$(\Delta x)_a = -\frac{1}{s'^2}h'^2 D \tag{4.23}$$

$$(\Delta x)_d = -\frac{n+1}{2ns'^2}h'^2 D \tag{4.24}$$

Recalling the magnification constraint of Eq. (4.13), we use the simple imaging relation

$$\frac{1}{f} = \frac{1}{s} + \frac{1}{s'} \tag{4.25}$$

and substitute

$$s' = 2s \tag{4.26}$$

to obtain

$$s = \frac{3}{2}f \tag{4.27}$$

and

$$s' = 3f. \tag{4.28}$$

Using a typical refractive index of $n = 1.5$ for the lens and the result of Eq. (4.28), the optimum shape factor for our lens is

$$q_{min} = -\frac{5}{21}, \tag{4.29}$$

and the aberration spot sizes for the shuffle arrangement of Fig. 4.1 are

$$(\Delta x)_s = -\frac{47}{168} \cdot (\frac{D^3}{f^2}) \tag{4.30}$$

$$(\Delta x)_c = \frac{1}{168} \cdot (\frac{D^2 h'}{f^2}) \tag{4.31}$$

$$(\Delta x)_a = -\frac{1}{9} \cdot (\frac{D h'^2}{f^2}) \tag{4.32}$$

$$(\Delta x)_d = -\frac{5}{54} \cdot (\frac{Dh'^2}{f^2}) \tag{4.33}$$

We can now calculate the effect of aberrations on the image spot size for any point in the image plane given the lens focal length and the lens radius (or aperture size). We note that it is possible to alter the distance between the lens plane and the output (or detector) plane slightly to minimize either the spherical or the astigmatic spot size. Moving the output plane slightly out of the image plane predicted by Gaussian optics (i.e. away from the image distance s' specified by Eq. (4.25)) allows one to find what is called the circle of least confusion for either spherical or astigmatic aberration. This can be done, however, for only one given point in the output plane. The size of other image spots in the array (i.e. those at different distances h' from the optic axis) may be increased unacceptably in the process. So, for this analysis, we leave the output plane (detector array) in the Gaussian image plane.

The effect of the various aberrations on a ray passing through one point in the aperture of a lens add vectorially. Therefore, the worst case for the net deviation of a ray from the Gaussian image point is the sum of the various contributions

$$(\Delta x)_{net} = (\Delta x)_s + (\Delta x)_c + (\Delta x)_a + (\Delta x)_d, \tag{4.34}$$

which is what we consider to be the net aberration spot size predicted by geometrical aberration theory.

4.2.5 Diffraction Effects

When the aberration spot sizes are small compared to the diffraction-limited spot, we observe an Airy pattern centered on the Gaussian image point. Its intensity pattern is circularly symmetric and its dependence on radial distance from the image point is

$$I(r) = [\frac{2J_1(\pi\frac{r}{\lambda F})}{\pi\frac{r}{\lambda F}}]^2 \tag{4.35}$$

where r is the radial distance from the Gaussian image point, and λ is the wavelength of light. We consider the central disc of the Airy pattern to be the diffraction spot. Its diameter is

$$(\Delta x)_{diff} = 2.44\lambda F. \tag{4.36}$$

4.2.6 Simultaneous Minimization of Diffraction and Aberrations

Qualitative Approach

The maximum packing density of resolvable channels in the optical shuffle arrangement is inversely proportional to the size of the image spots. The spot size for any imaging system is limited by the two primary considerations discussed above: diffraction and aberrations. As summarized by Eq. (4.36), diffraction theory predicts that the image spot size of a point source is proportional to λF, so, from this point of view, it is desirable to use short wavelengths, and large diameter lenses with short focal lengths. In practice, we are limited in our choice of wavelengths. We will likely be using semiconductor lasers, which are presently limited to the red and infrared end of the spectrum. However, as shown by Eqs. (4.21)-(4.24), the aberration spot sizes increase with increasing lens diameter, D.

Next, we recognize that the aberrations themselves may broken down into on- and off-axis types. The on-axis aberration is spherical, and by "on-axis", we mean that it is the only primary aberration which exists on-axis. Furthermore, as shown in Eq. (4.8), spherical aberration is not a function of image height, h', which we defined previously as the distance of an image point from the optic axis of the lens. The remaining primary aberrations are zero (or do not exist) on-axis, and increase with increasing image height.

In summary, the results are: (a) the largest image spot sizes occur at the corners of the output (detector) array; (b) simply reducing the size L of the array will reduce the off-axis aberrations. We cannot, however, reduce L without decreasing the maximum possible diameter of our lenses, due to the constraint of Eq. (4.1). So, while reducing L reduces the aberrations, it also increases the effects of diffraction by limiting available aperture size. Furthermore, reducing the aperture size lowers the collection efficiency of the optics, placing demands on the sources for higher optical power output.

Optimization Procedure

From the qualitative description in Section 4.2.5 of the trade-offs involved, we present an iterative optimization procedure for maximizing the packing density for a given lens.

1. Start with the focal length f of the lens as a given quantity.
2. Arbitrarily pick the array size, L.
3. Calculate the off-axis distance, h', of the detector (or image point) which is farthest from its optic axis (i.e. at the edge of the array, where aberrations are worst).
4. Use the set of Eqs. (4.21)-(4.24), and Eq. (4.34) to determine the aberration spot size versus aperture radius and Eq. (4.36) to determine the diffraction spot size versus aperture radius. Equating these two functions provides a common solution which is the minimum spot size $(\Delta x)_{min}(L)$, for the particular value of L chosen in step 2. The particular aperture radius where the solution occurs achieves the minimum spot size for this L, and is called the optimum radius, $r_{opt}(L)$.
5. Use the constraint on available lens aperture size imposed by the array size, L as expressed in Eq. (4.2) for this particular multi-lens arrangement.
   (a) If the optimum aperture radius $r_{opt}(L)$ determined at step 4 is less than the available aperture space $r_{max}(L) = \frac{L}{6}$, then we have not taken full advantage of available aperture space. In this case, L is reduced to reduce off-axis aberrations. As a rough guideline, choose the new

array size, L', so that $\frac{L'}{6}$ is about half-way between the current $r_{opt}(L)$ and $r_{max}(L)$.

(b) If the optimum aperture size $r_{opt}(L)$ determined at step 4 exceeds the maximum available aperture space $r_{max}(L)$, then we cannot achieve optimum spot size for the current L because the lenses overlap. Here, the array size, L is increased. Again, we choose the new array size, L' so that $\frac{L'}{6}$ falls roughly half-way between the current $r_{max}(L)$ and $r_{opt}(L)$.

6. Repeat steps 3 through 5 until the optimum aperture in step 4 corresponds to the maximum available aperture in step 5. At this point, when $r_{opt}(L) = r_{max}(L)$, we achieve the absolute minimum spot size, $(\Delta x)_{min}(L_{opt})$, possible for the lens focal length chosen in step 1.

Example

To illustrate the procedure, we use the example of an optical shuffle implemented previously by C.B. Kuznia [31] (the photos of which are shown in Section 3.8). He assembled a square source array of 256 LEDs on a 16 × 16 grid, which measured $L = 11.5$ cm on a side, and lenses with focal length $f = 10$ cm. We see from Fig. 4.1 that the location of the detector farthest off-axis in our 2-D implementation of the shuffle is given by Eq. (4.12), which in this case is $h'_{max}(L)|_{L=11.5\text{ cm}} = 10.8$ cm. The plot indicated in step 4 appears in Fig. 4.4, where we have used $\lambda = 0.81\ \mu$m.

Originally, we used an aperture radius of 1.5 cm (indicated by the vertical dotted line) which produced an aberration-limited spot diameter of about $8,000\ \mu$m (indicated by the horizontal dotted line). This calculated value for the size of the spot agrees well with experimental measurements. In fact, being confronted with a spot size of roughly 8 mm, we cut the aperture size roughly in half (which was then used to produce the photographs in Figs. 3.11 and 3.12), and developed the optimization procedure outlined above.

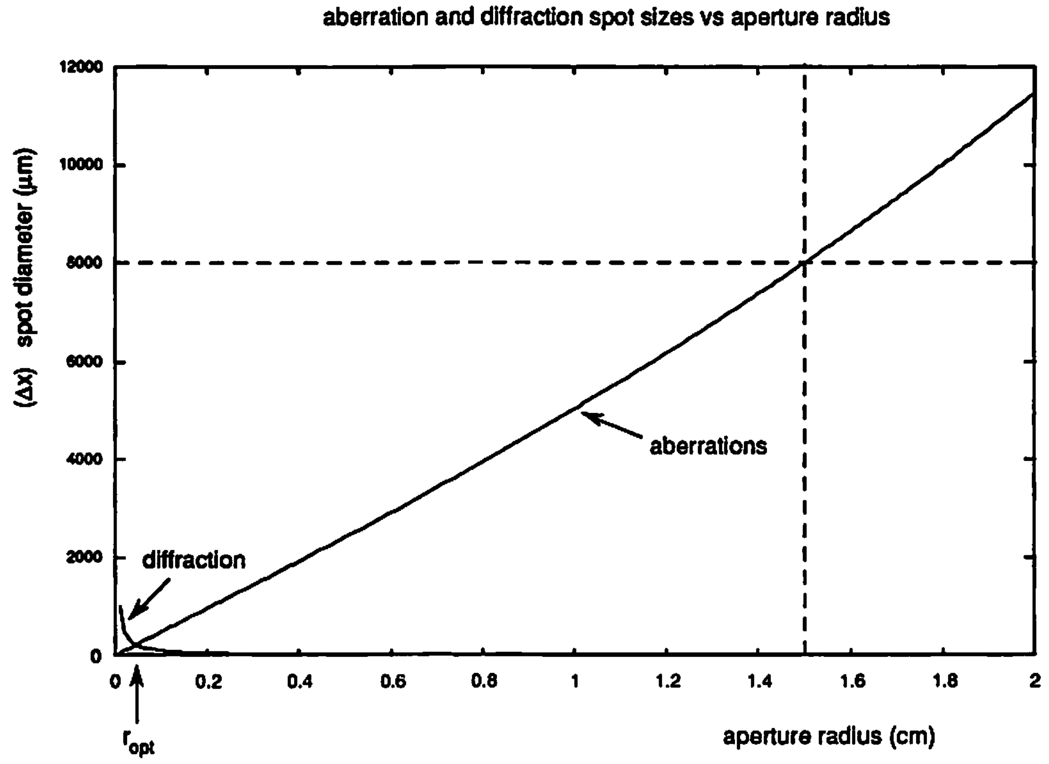


Figure 4.4: Plot of aberration and diffraction spot diameters versus aperture radius for our original experimental demonstration of the shuffle, in which f = 10 cm, L = 11.5 cm, r_{max} = 1.92 cm, $r_{opt} \approx$ 0.045 cm; for the calculations we use λ = 0.81 μm, n = 1.5. The vertical dotted line marks the aperture radius originally used r = 1.5 cm, and the horizontal line the calculated aberration spot size. Note the contribution of diffraction is negligible compared to the aberrations.

f focal length	L_{opt} array size	$r_{opt}(L_{opt})$ lens radius	$(\Delta x)_{min}$ spot diameter	F f-number
15 cm	2.30 cm	0.38 cm	38.6 μm	19.6
10 cm	1.70 cm	0.28 cm	34.9 μm	17.7
5 cm	1.01 cm	0.17 cm	29.4 μm	14.9
1 cm	0.30 cm	0.050 cm	19.6 μm	9.93

Table 4.1: Minimum spot diameter $(\Delta x)_{min}$ achievable for a simple lens with focal length f; corresponding optimum aperture size r_{opt}, array size L_{opt}, and lens f-number F; $\lambda = 0.81$ μm and $n = 1.5$.

We notice that the minimum spot diameter for $L = 11.5$ cm is achieved for a much smaller aperture size, which is actually $r_{opt}(L)|_{L=11.5}$ cm ≈ 0.045 cm. The corresponding optimum spot diameter is $(\Delta x)_{opt}(L)|_{L=11.5}$ cm ≈ 220 μm. The maximum available aperture size is $r_{max}(L)|_{L=11.5}$ cm $= 1.92$ cm. Since the optimum radius is much less than the available aperture space, we follow the iterative procedure reducing L until $r_{opt}(L) = r_{max}(L)$. This occurs for $L = 1.7$ cm, and is illustrated in Fig. 4.5. The $(\Delta x)_{opt}(L)|_{L=1.96}$ cm $= 34.9$ μm and corresponds to the absolute minimum spot diameter $(\Delta x)_{min}$ in Fig. 4.6. The final result is the array size, L_{opt}, and aperture size, $r_{opt}(L_{opt})$, that minimize the spot diameter. For the case of our experimental shuffle, this process enables us to reduce the spot diameter from roughly 8,000 μm to 35 μm. Table 4.1 shows the minimum spot diameter $(\Delta x)_{min}$ and the corresponding system parameters for various focal lengths obtained using this procedure, for $\lambda = 0.81$ μm and $n = 1.5$.

4.2.7 Optoelectronic Limits on Resolvable Spot Density

Achieving minimum spot size does not specify how closely we may space our detectors. Diffraction spreads light out indefinitely from the Gaussian image point. The spreading due to diffraction means that there is overlap, or crosstalk, between neighboring channels in the output plane. Our objective is to space the channels

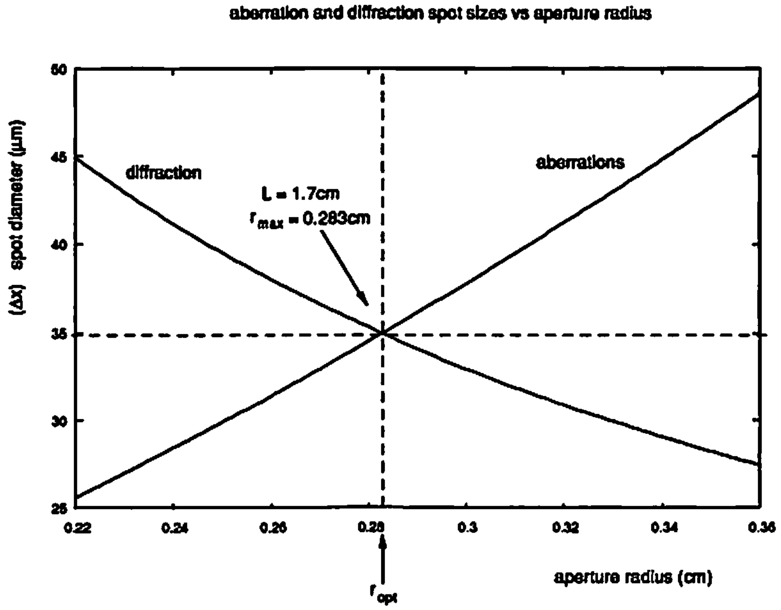


Figure 4.5: Choice of a different size $L = 1.70$ cm for the array which makes the intersection between the diffraction and aberration curves occur at the maximum available aperture radius for the given L, $r_{opt}(L) = r_{max}(L)$. Overall spot diameter is minimized to about 35 μm as indicated by the horizontal dotted line ($n = 1.5$, $\lambda = 0.81$ μm).

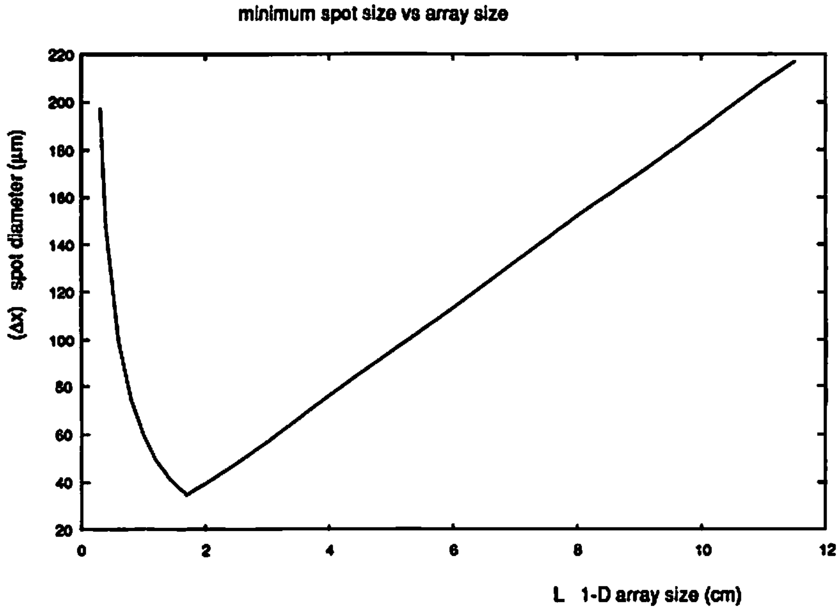


Figure 4.6: For the focal length lens used in our experimental demonstration $f = 10$ cm, we plot the minimum achievable spot diameter versus array size L, confirming the fact that minimum spot size is achieved when $L = 1.7$ cm, such that $r_{opt}(L) = r_{max}(L)$.

to ensure a total optical crosstalk of less than χ_{max}, which we choose somewhat arbitrarily to be $\chi_{max} = -20$ dB. In order to calculate the amount of optical energy spreading into adjacent channels, we must know the intensity distribution around the image point.

In the preceeding analysis, we have applied independently both geometrical aberration theory, which assumes that light travels as a ray from the source to its image, and diffraction theory, which assumes that light behaves as a wave. In fact, these seemingly incompatible approaches may be combined by incorporating the wave aberrations predicted by geometrical optics into a diffraction integral. The result is a rigorous prediction of the intensity distribution in the output plane around the image point. We undertake such an analysis later in Section 5.2.2 for different purposes. The problem is that such an analysis is computationally intensive, and there is a simpler reconciliation of geometrical and diffraction theories suitable for the problem at hand.

The point at which the spot sizes for diffraction and ray aberrations are simultaneously minimized in a plot such as Fig. 4.5 is where the distribution of rays predicted by geometrical aberration theory lies entirely within a circle the size of the diffraction-limited central Airy disc. When the ray aberrations are so restricted, the actual intensity distribution very closely resembles the Airy pattern. Because the Airy pattern has a known intensity distribution, given by Eq. 4.36, we may calculate the light intensity falling on neighboring detectors.

Detector Size and Encircled Power

In order to find the spacing between channels which ensures that optical crosstalk is less than the maximum acceptable χ_{max}, we first choose the size of the detectors in the output plane. Though no absolute best detector radius exists, we choose a reasonably good radius based on two concerns. First is that we want efficient collection of optical power by the detector. This calls for a larger detector radius. The competing concern is to be able to place as many communication channels into as small an area as possible. This calls for smaller detectors.

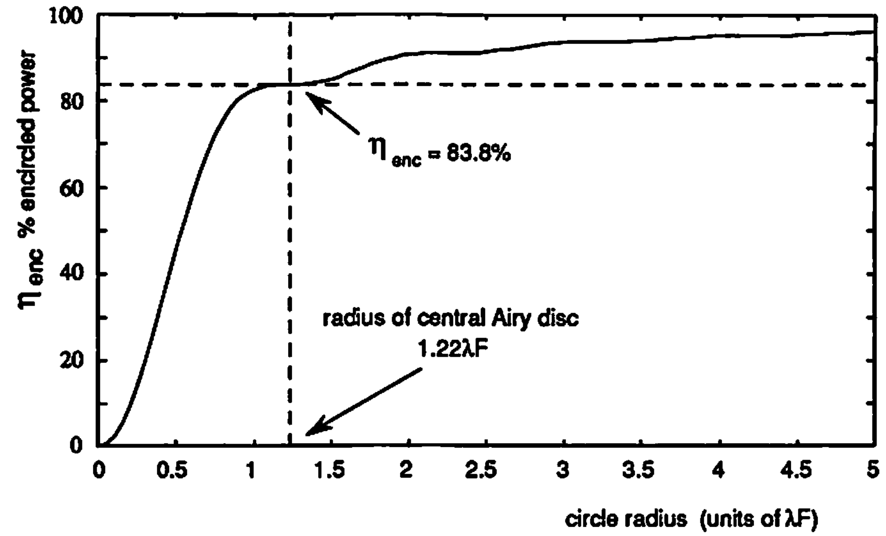


Figure 4.7: Percentage of power in an Airy pattern enclosed (or encircled) within a circle of variable radius. The dotted lines show that the power in the central Airy disc is about 83.8 % of the total power in the image.

In order to calculate how much optical power falls on a detector of given radius, we integrate the intensity distribution of the output image (the Airy pattern) given by Eq. (4.36) over a circle of radius r, centered at the Gaussian image point and measured in units of λF. Encircled power $\eta_{enc}(r)$ is the fraction of the total power in the image contained within that circle.

$$\eta_{enc}(r) = \int_0^r \left[\frac{2J_1(\pi\frac{r'}{\lambda F})}{\pi\frac{r'}{\lambda F}}\right]^2 \frac{1}{\lambda F} dr'. \tag{4.37}$$

The result has a simple, closed form solution [23],

$$\eta_{enc}(r) = 1 - J_0^2(\pi\frac{r}{\lambda F}) - J_1^2(\pi\frac{r}{\lambda F}) \tag{4.38}$$

which is graphed in Fig. 4.7.

By evaluating Eq. (4.38), we see that once the detector radius is the size of the first Airy disc, roughly 83.8 % of the output image power falls on the detector.

$$\eta_{enc}(r)|_{r=1.22\lambda F} \approx 83.8\%. \tag{4.39}$$

Making the detector marginally larger does not increase the encircled power appreciably. It is therefore reasonable to view the first null in the Airy pattern as a point of diminishing returns, and use it as our detector size. Having established the size of the detectors in the output array, we use the center-to-center spacing of the detectors as the remaining variable to determine the amount of crosstalk between neighboring channels in an array.

We note that the approach described above for choosing the size of the detectors is not universal. Others have begun by fixing the center-to-center spacing of the detectors at some reasonable value, and then altering the size of the detector to minimize crosstalk. Wang [37] chose to fix the center-to-center spacing at $2.44\lambda F$, which means that the central Airy discs are placed edge to edge. Then, by reducing the detector size to some value, an acceptable level of crosstalk may be achieved. The difference is that this method achieves a higher packing density at the expense of the collection efficiency of the detector. In contrast, our method assures a reasonably high collection efficiency, but at the expense of requiring more area in the output plane.

Crosstalk Due to Diffraction

Single Neighbor To evaluate optical crosstalk due to a single neighbor, we consider Fig. 4.8, which shows two adjacent detectors D_1 and D_2, with center-to-center spacing δ, and each with radius $1.22\lambda F$. Assume that light is imaged to detector D_1, and that detector D_2 receives crosstalk from optical power falling in the shaded ring. The power received by D_2 is written P_{D_2}, and is the product of the total power within the shaded ring P_{ring} and the fraction κ of the ring

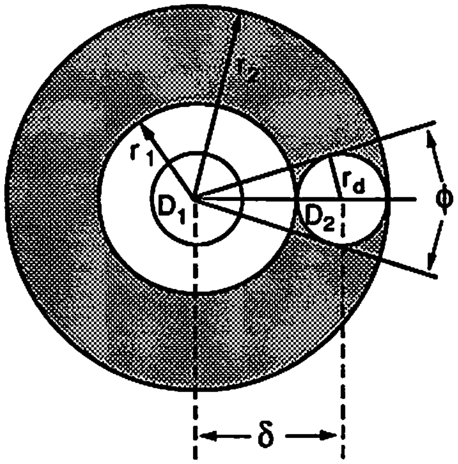


Figure 4.8: Construction for computing crosstalk in adjacent detectors D_1 and D_2. D_2 receives the crosstalk from an image (i.e. an Airy pattern) centered at D_1. δ is the center-to-center spacing between the two detectors, r_d is the radius of the detectors, r_1 and r_2 are the radii of concentric circles around D_1 used to calculate encircled power in the Airy pattern, and ϕ is the angle subtended by D_2 from the center of D_1.

occupied by D_2

$$P_{D_2} = \kappa \cdot P_{ring}. \tag{4.40}$$

The power in the ring is the difference of the powers contained within the two concentric circles from

$$P_{ring} = \eta_{enc}(r_2) - \eta_{enc}(r_1) \tag{4.41}$$

A good approximation for the fraction of the ring κ occupied by D_2 by finding the arc ϕ subtended by D_2, and dividing by 2π,

$$\kappa \approx \frac{\phi}{2\pi}. \tag{4.42}$$

The half-angle,$\frac{\phi}{2}$, formed by the line segment, δ which connects the detector centers, and the tangent line at D_2 from the center of D_1, is shown in Fig. 4.8 and follows the relation

$$\sin(\frac{\phi}{2}) = \frac{r_{det}}{\delta}. \tag{4.43}$$

So, the fraction of the ring occupied by D_2 is

$$\kappa \approx \frac{1}{\pi} \arcsin(\frac{r_{det}}{\delta}), \tag{4.44}$$

and the power at an adjacent detector, P_{adj}, is approximately

$$P_{adj} \approx \frac{1}{\pi} \arcsin(\frac{r_d}{\delta})\{[J_0^2(\pi\frac{r_1}{\lambda F}) + J_1^2(\pi\frac{r_1}{\lambda F})] - [J_0^2(\pi\frac{r_2}{\lambda F}) + J_1^2(\pi\frac{r_2}{\lambda F})]\}, \tag{4.45}$$

where r_d is the detector radius

$$r_d = 1.22\lambda F \tag{4.46}$$

and r_1 and r_2 are the radii of the concentric circles

$$r_{1,2} = \delta \mp r_d. \tag{4.47}$$

We shall later compare the approximation of Eq. (4.45) with the rigorously derived result, showing that the approximation is reliable to within about 1 dB.

Multiple Neighbors In a large array, crosstalk may arise from any or all of the other channels in the array. We consider here the contributions from the first two shells of nearest neighbors, as shown in Fig. 4.9. Neighbors labeled with like numbers are equidistant from the central channel, and U is the unit distance (center-to-center) from the central channel to its closest neighbors, which are labeled 1. The second nearest neighbors are labeled 2, third nearest 3, and so on. The assumed intensity distribution for any image point is the Airy pattern, which is radially symmetric. This implies that neighbors which are the same distance from the central channel contribute equal amounts of crosstalk to the central channel. So, we need to calculate only five separate contributions to be able to obtain the crosstalk from the first two shells of neighbors.

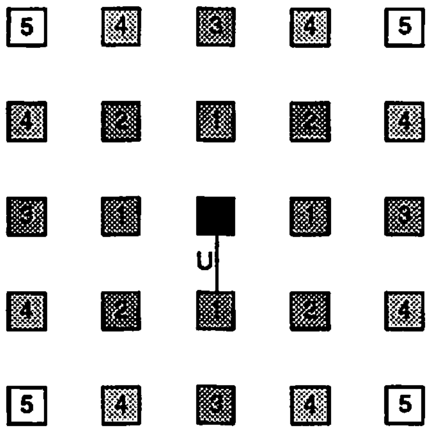


Figure 4.9: Nearest neighbors considered in crosstalk calculation. U is the unit center-to-center separation of nearest neighbors. Groups of neighbors which are equidistant from the central channel are shown with like numbers, where smaller numbers represent closer neighbors. In the first two shells of neighbors shown here, there are five groups of equidistant neighbors (or neighborhoods).

The worst scenario is that all of these neighbors contribute to the crosstalk simultaneously. We apply Eq. (4.45) to the five different contributions of neighbors in the first two shells, using appropriate values of r_1 and r_2. Then, we assume that the contributions of different neighbors add incoherently. Figure 4.10 is a plot of worst case optical crosstalk, χ versus center-to-center detector spacing. Here, χ has been defined as

$$\chi = \frac{1}{P_0} \sum_i P_i \tag{4.48}$$

where P_i is the crosstalk power received from neighbor i (which is within the first two shells of nearest neighbors), and P_0 is the uncorrupted signal power. The parameter α is the center-to-center spacing, δ, normalized to the detector diameter, given by

$$\alpha = \frac{\delta}{2.44\lambda F}. \tag{4.49}$$

To provide a realistic value for our discussion, we set the maximum acceptable crosstalk at $\chi_{max} = -20$ dB. This value may be changed depending on the network application. Figure 4.10 shows that to maintain crosstalk $\chi < -20$ dB, we need

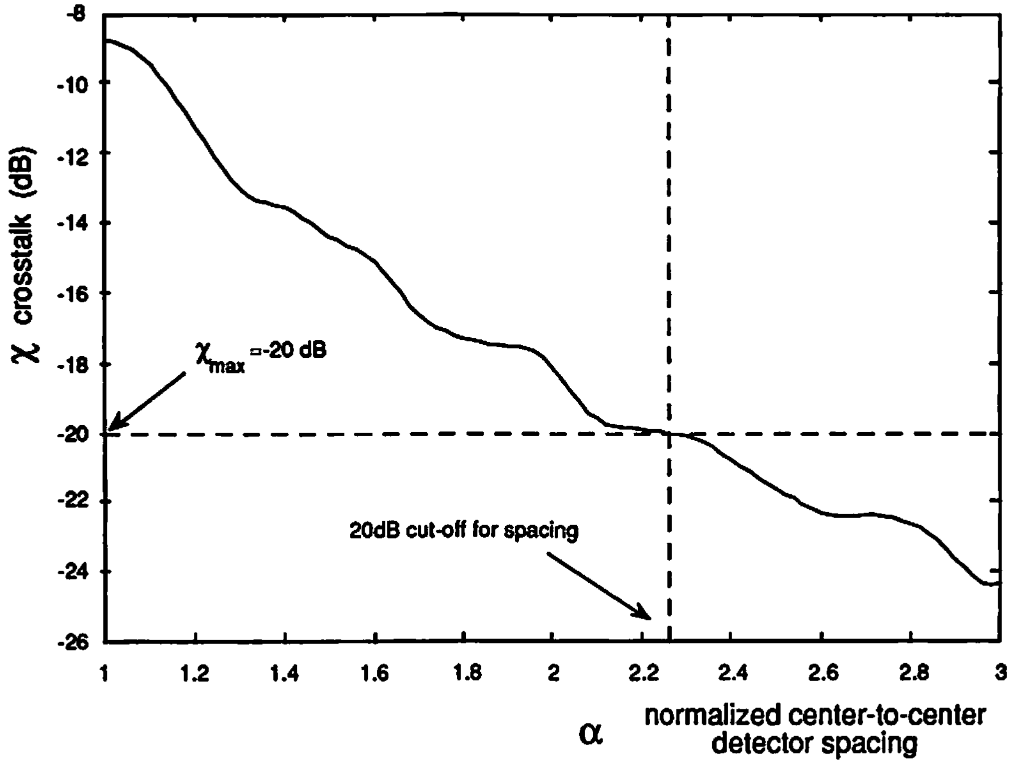


Figure 4.10: Optical crosstalk in dB versus normalized center-to-center detector spacing. $\alpha = 1$ implies the edges of the two detectors are touching. The horizontal dotted line is our chosen tolerance of $\chi_{max} = -20$ dB for crosstalk, and the vertical line is the required spacing to maintain that tolerance. Roughly two and one quarter detector diameters are required between detector centers.

roughly 2.26 detector diameters between detector centers,

$$\alpha \geq 2.26, \tag{4.50}$$

which is equivalent to a physical spacing of

$$\delta \geq 5.514 \lambda F. \tag{4.51}$$

Results

Using Eq. (4.51), we can completely specify the limits of our shuffle design. First, we use the optimization procedure of Section 4.2.5 for a range of focal lengths (as we did to obtain Table 4.1). Next, we use the corresponding f-numbers from Table 4.1 in Eq. (4.51) to calculate the minimum center-to-center detector spacing, δ_{min}, which maintains $\chi_{max} = 20$ dB crosstalk. The maximum packing density, ρ_{max}, achievable in a 2-D is square of the reciprocal of the minimum spacing as given by

$$\rho_{max} = (\frac{1}{\delta_{min}})^2. \tag{4.52}$$

The results of using Eq. (4.52) for a range of focal length lenses are shown in Fig. 4.11. It is not surprising that lenses with shorter focal lengths achieve higher density, since these are the more powerful focusing elements.

If, however, we plot the maximum achievable density versus the (optimized) f-numbers of these various lenses, we see a slightly unexpected result, which is shown in Fig. 4.12. Ordinarily, one associates higher f-numbers with greater resolution (i.e. packing density), but in this case, we see that the opposite occurs. This is because the f-number is not a free variable once the system geometry has been optimized. By this, we mean that for a certain focal length there is a unique combination of aperture size and array size which optimizes the packing density.

The potential confusion here is that we are accustomed to thinking of a given lens (i.e. fixed focal length) with variable aperture. Increasing the f-number

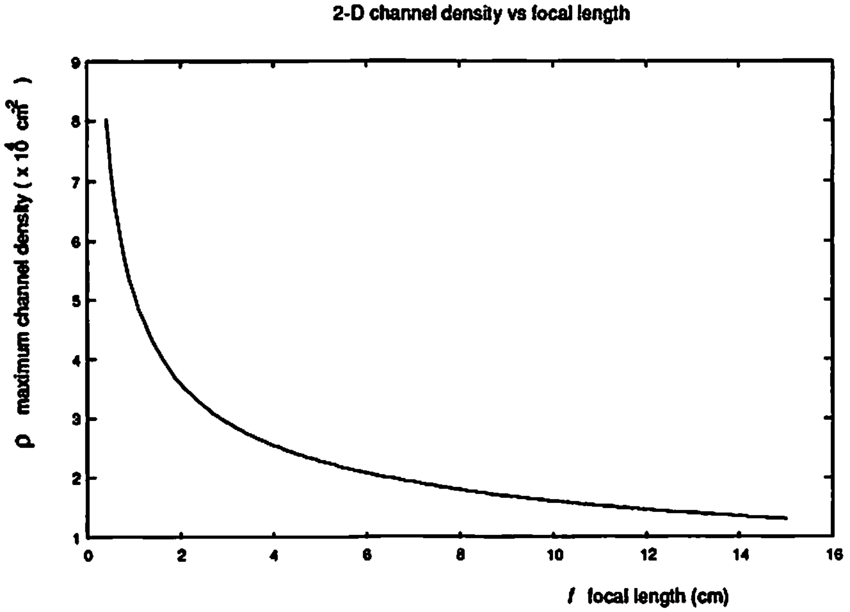


Figure 4.11: Maximum packing density (in units of cm^{-2}) achievable versus lens focal length. Higher densities require shorter focal length (i.e. more powerful) lenses; $\lambda = 0.81$ μm, $n = 1.5$.

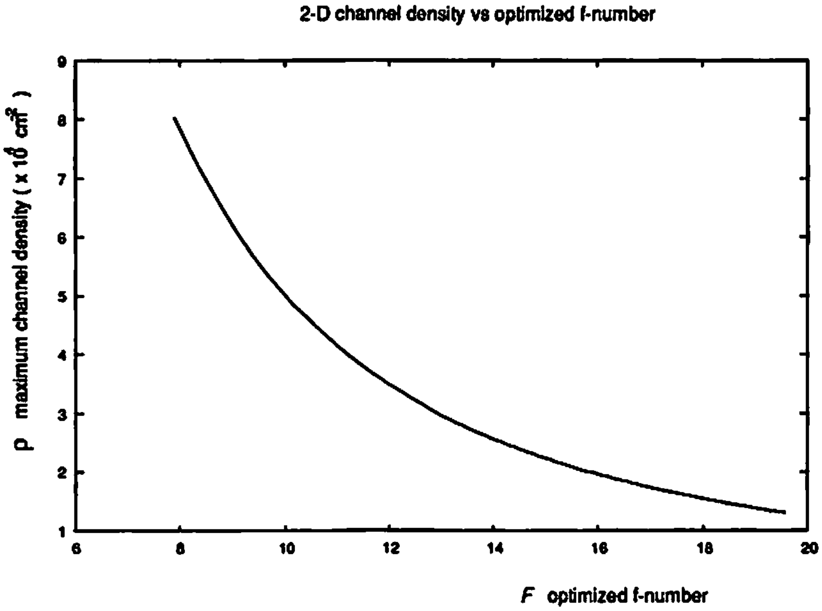


Figure 4.12: Maximum packing density in 2-D (in units of channels per cm^2) versus optimized f-numbers, which are the f-numbers obtained by the optimization procedure for various focal lengths.

means decreasing the aperture size, which under the assumption of aberration-limited imaging implies greater resolution. Aberration-limited imaging, however, occurs in a *region* to the right of the intersection between the aberration and diffraction spot size curves (in Figs. 4.4 and 4.5). Our design procedure fixes the system geometry such that aberrations and diffraction are simultaneously minimized at border between aberration- and diffraction-limited imaging.

The fact that in order to achieve minimum spot size, the array size is constrained to L_{opt} has the added consequence that the maximum number of channels n^2_{max} resolvable over the whole array is

$$n^2_{max} = \rho_{max} \cdot L^2_{opt}. \tag{4.53}$$

Using the ρ_{max} of Fig. 4.11 with the corresponding L_{opt}, Eq. (4.53) yields the graph in Fig. 4.13. The results given in Fig. 4.13 can be used as a design guide to specify the focal length lenses necessary to shuffle n^2 channels. For example, if we wish to shuffle $20,000$ channels, then the shortest focal length capable of accomplishing this is roughly 4.5 cm.

Here, we recognize an important trade-off. Figure 4.11 shows that shorter focal lengths can achieve higher packing densities. Figure 4.13 shows that in order to shuffle more total channels, we must use longer focal lengths. The corollary to the two previous results is that as one shuffles more channels, the achievable packing density goes down. This is shown in Fig. 4.14. We also plot the optimum array size L_{opt} as a function of the number of channels n^2 one wishes to shuffle in Fig. 4.15. This shows that as we shuffle more channels, the array must grow in physical size L (a result which is not unexpected). Thus, the width of the system increases. From Fig. 4.13, we also know that the system must grow in length in order to accommodate more inputs. We also see that for lenses with ordinary focal lengths, we expect the system to be able to shuffle from 10^3 to 10^5 channels with simple lenses alone.

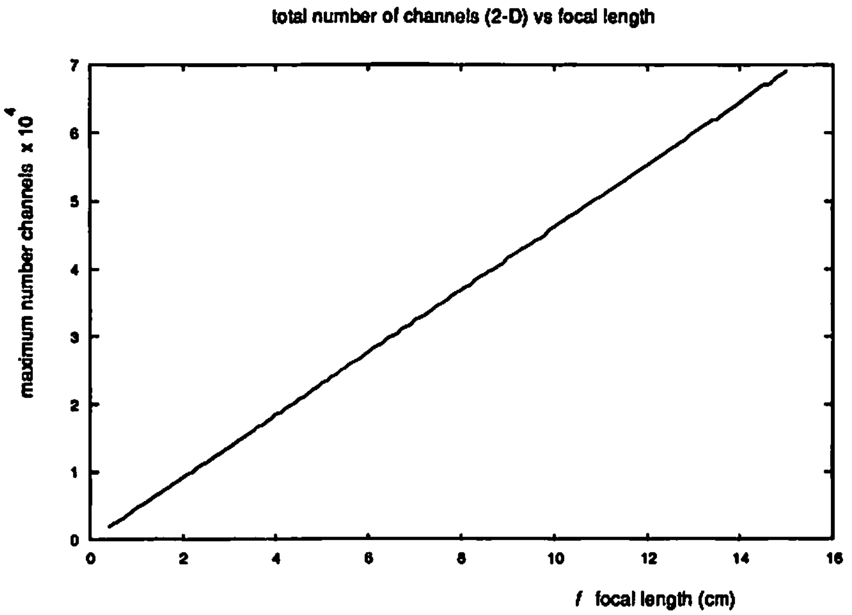


Figure 4.13: Maximum number of channels resolvable in a 2-D array versus the focal length of the lenses. More channels require longer focal length, and correspondingly longer overall system length; $\lambda = 0.81\ \mu\text{m}$, $n = 1.5$ (index of refraction).

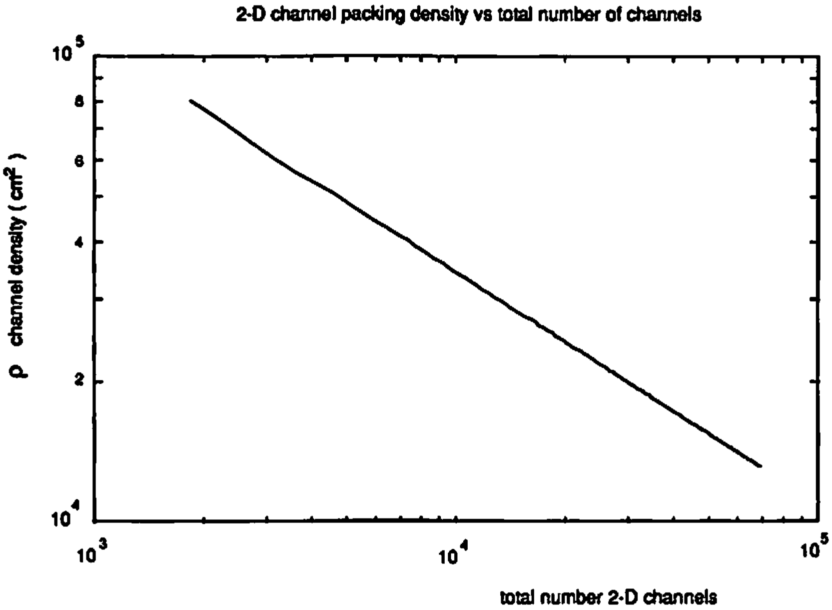


Figure 4.14: Maximum 2-D packing density (in channels per cm^2) versus desired number of channels in the array.

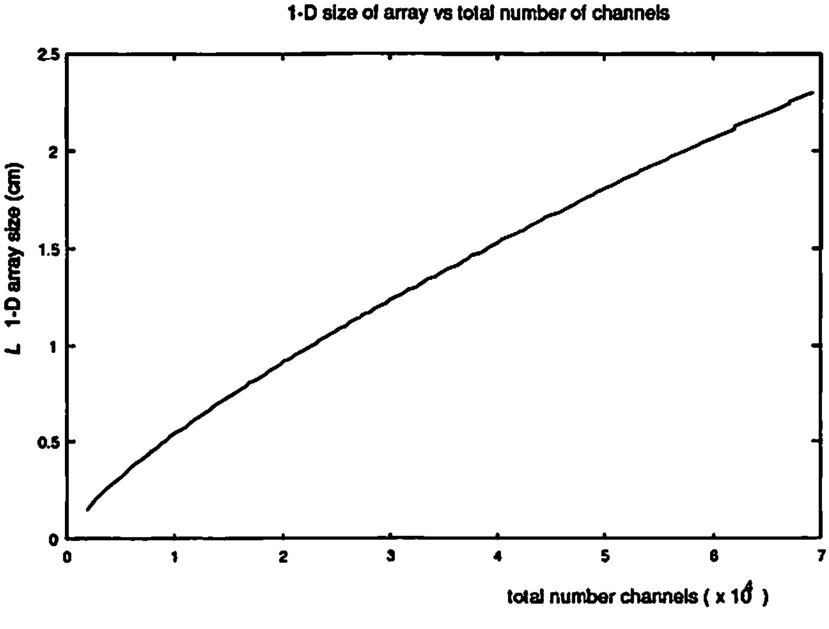


Figure 4.15: Optimum array size L_{opt} versus desired number of channels in 2-D array.

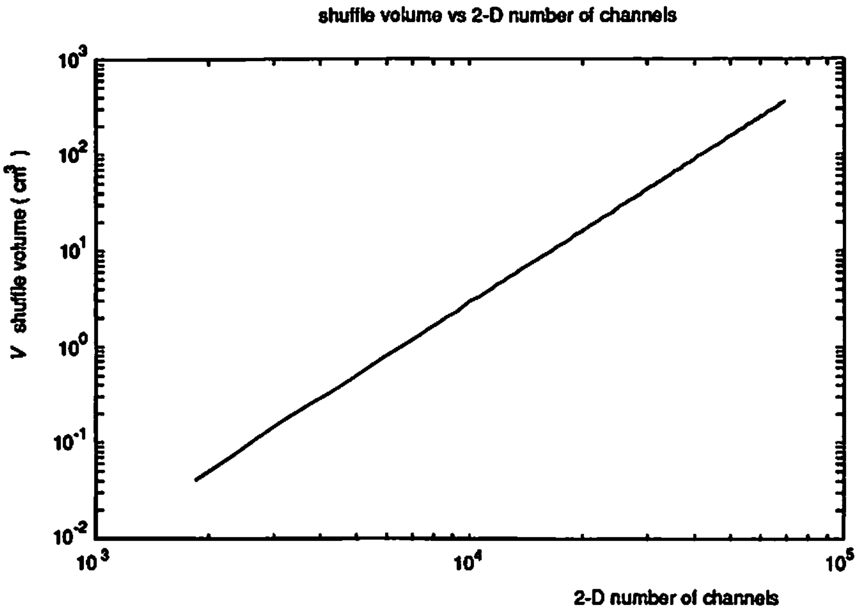


Figure 4.16: Shuffle volume V versus the number of channels N in a 2-D array.

Also characteristic of the shuffle implemented with ordinary lenses is the volume which the system occupies, and the volume which each channel requires. We find that these quantities are not fixed, but rather heavily dependent either on the focal length of the lenses or the number of channels in the network. Figures 4.16 - 4.18 show the shuffle volume V as a function of the number of channels N (in 2-D), the volume density ρ_V versus N, and the volume per channel V/N versus N, respectively.

4.2.8 Generalizations

An important observation is that the procedure for optimizing the system geometry is equally applicable to a system where the lenses are corrected, compound lenses. All one needs is a relationship between the geometrical spot diameter, the image height and the aperture size

$$(\Delta x)_{ab} = f(h', r). \tag{4.54}$$

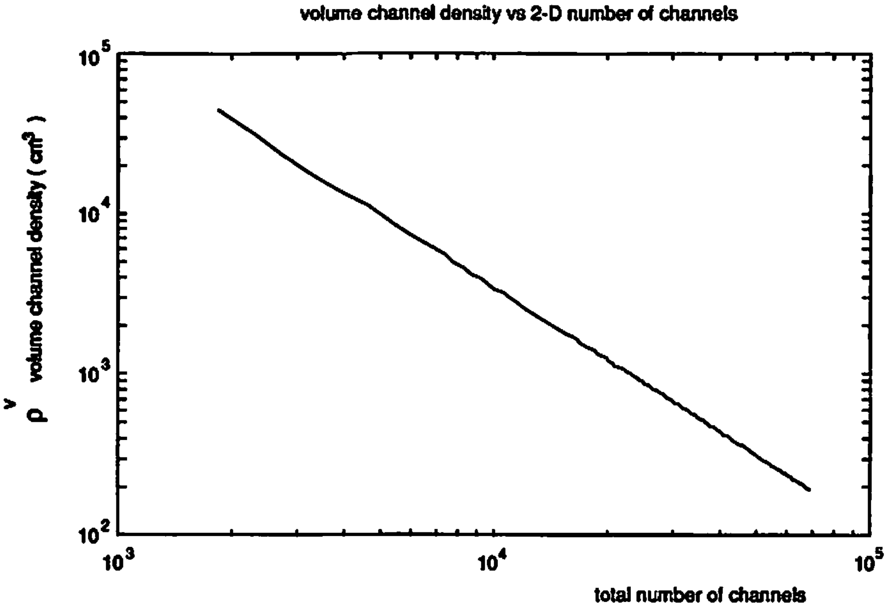


Figure 4.17: Volume density ρ_V versus the number of channels N in a 2-D array.

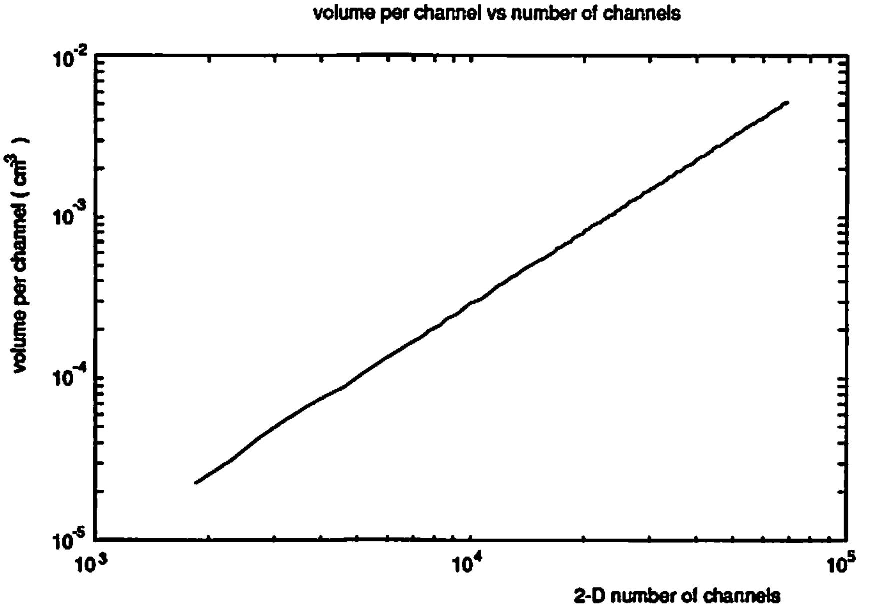


Figure 4.18: Volume occupied by an individual channel V/N versus the number of channels N in a 2-D array.

It matters little whether this relationship is derived analytically or obtained empirically; the optimization procedure is valid regardless. For the case of the simple lens with surfaces chosen for minimum spherical aberration in the shuffle architecture, the aberration spot diameter may be written

$$(\Delta x)_{ab} = \frac{47}{21f^2}r^3 + \frac{1}{42f^2}h'r^2 + \frac{2}{9f^2}h'^2r + \frac{5}{27f^2}h'2r. \tag{4.55}$$

In general, primary aberrations produce a spot diameter given by

$$(\Delta x)_{ab} = C_3r^3 + C_2h'r^2 + C_1h'^2r \tag{4.56}$$

where the C are coefficients which may be determined either emperically (e.g. curve fit of experimental data) or analytically.

The results produced by the optimization procedure may be duplicated analytically without iteration by setting $(\Delta x)_{ab}$ equal to the size of the diffraction spot size $(\Delta x)_{diff}$, given by Eq. 4.36

$$C_3r^3 + C_2h'r^2 + C_1h'^2r = 2.44\lambda\frac{f}{2r}. \tag{4.57}$$

In the optimum system design, both r and h' are known functions of the array size

$$r = \frac{1}{6}L \tag{4.58}$$

$$h' = \frac{2\sqrt{2}}{3}L \tag{4.59}$$

so, we may solve explicitly for the L which simultaneously minimizes diffraction and aberrations, and maximizes the aperture radius. In general, this is given by

$$L_{opt} = \{1.22\lambda f\frac{6^4 3^2}{9C_3 + 36\sqrt{2}C_2 + 288C_1}\}^{1/4}. \tag{4.60}$$

For the case of simple lenses in the shuffle, we use the coefficients

$$C_3 = \frac{47}{21f^2} \tag{4.61}$$

$$C_2 = \frac{1}{42f^2} \tag{4.62}$$

$$C_1 = \frac{11}{27f^2}. \tag{4.63}$$

Equation (4.60) accurately reproduces the results for optimum shuffle geometry by the optimization procedure.

4.2.9 Aberration Corrected Lenses

The result of undertaking a sophisticated lens design is twofold. Though it is a different architecture, the switching system built by McCormick *et al* [24] illustrates the difference. They imaged an array of points onto a field 1.26 mm square using an $f/1.5$ lens with 15.61 mm focal length. The spot sizes achieved were on the order of 5 μm. Using only a simple, uncorrected lens with roughly equal focal length, the geometry for optimum packing density has an f-number of 11.0, and achieves a spot size of 22 μm, obtained using the optimization procedure outlined previously. So, sophisticated lens design achieved (for roughly the same optical system length) more than three times the packing density in 1-D, and greater light throughput due to the faster lens.

4.2.10 Extensions to Practical Implementation

In the preceeding calculations of the resolution capabilities of simple lenses, we have assumed certain ideal circumstances which do not exist in practical systems. Practical systems only approach the capabilities we have described, so we discuss here the effects of deviation from the ideal. In the previous analysis, we have assumed that the sources have zero spatial extent, or that they are effectively point sources, thus the spot diameters as calculated are roughly the extent of the spatial impulse response of the system. Given that the system is linear and shift-invariant, then to find the extent of the image of an extended source, we convolve the source extent with the system's impulse response. As a first order approximation, we simply add the diameter of the impulse response to the diameter of the source profile to obtain the diameter of the actual image spot.

Another assumption we have made is that all four of the lenses in the shuffle system have identical focal lengths. In practice, this is hard to achieve, so there will be some defocus which we have not included in the calculations in at least one (if not all) of the image points.

The preceeding two practical conditions affect the diameters of the image spots, and thus the achievable packing density. Should the lenses be misaligned so that their centers do not lie exactly $L/6$ away from the system axis in both x and y as they should, then the image points are displaced from their intended matrix-like arrangement. This increases the optical crosstalk to certain detectors, and thus requires us to increase the minimum detector spacing δ_{min} required to maintain a maximum -20 dB crosstalk over the entire array. We shall examine these issues in Chapter 5.

4.3 Power Considerations

It is useful to compare system packing density limits due to diffraction and aberrations to those due to heat dissipation. As an illustration, we assume that the sources in the input array are lasers, like the vertical cavity lasers developed by Jewell *el al* [15]. The power inefficiency of the lasers generates excess heat which must be dissipated. For an array of channels with packing density ρ (in units of cm^{-2}), and a heat removal rate, σ_p (in W/cm^2), we calculate the maximum allowed heat dissipation, or critical power P_{crit}, per laser as

$$P_{crit} = \frac{\sigma_p}{\rho}. \tag{4.64}$$

If we use the maximum packing density achievable with our lenses, and assume a common limit of heat dissipation of $\sigma_p \approx 1$ W/cm^2, then the corresponding critical powers, above which the packing density is limited by heat dissipation of the lasers, are given in Table 4.2.

f focal length	ρ_{max} (per cm^2) maximum density	P_{crit} critical power
15 cm	1.31×10^4	76.3 μW
10 cm	1.61×10^4	62.3 μW
5 cm	2.27×10^4	44.0 μW
1 cm	5.08×10^4	19.7 μW

Table 4.2: Critical power P_{crit} per laser. If each laser dissipates more than P_{crit} the packing density is limited by threshold power dissipation; if each laser dissipates less than P_{crit}, then packing density is limited by the resolution of the lenses. $\lambda = 0.81$ μm, $n = 1.5$, $\sigma_p = 1$ W/ cm^2.

4.3.1 Threshold Dissipation

If we use laser diodes as the sources, then we must consider how much power P_{th} is required to bring each laser past threshold, thus enabling it to transmit optical information. This threshold power is just the product of the voltage dropped across the diode junction V_A and the threshold current i_{th}

$$P_{th} = V_A i_{th}. \tag{4.65}$$

For example, V_A, is roughly 1 V, thus we can loosely convert from a threshold current in mA to a power *turn-on penalty* expressed in mW. Even *low* threshold lasers today require about 1 mA before they turn on. So, before any data is sent, the power consumption at threshold of 1 mW well exceeds the breakeven power, P_{crit}, for any packing density over 1,000 channels/cm^2. Because we expect simple lenses to resolve many more channels than this, we conclude that until threshold

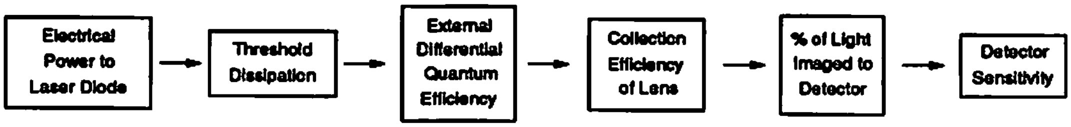


Figure 4.19: Block diagram of power flow, inefficiencies and requirements traced from system input (source) to output (detector).

currents are lowered by more than an order of magnitude, our ability to pack data channels into a free space shuffle is generally limited by heat dissipation.

4.3.2 Flow of Power through our Shuffle Architecture

Budget for Signal Power

Under the assumption that the threshold dissipation problem is either solved or circumvented (e.g. by using a holographically generated spot array in conjunction with an array of modulators), we can trace the flow of signal power through the complete system. Figure 4.19 shows the path of optical power from input (source) to output (detector), and the various points where optical or electrical power is lost as heat due to inefficiencies in the system.

At the outset, the laser converts some fraction η_{ext} of electrical power past threshold into light. The lens collects only some fraction of the light diverging from the source, which we call the collection efficiency η_{lens} of the lens. At the output, light is imaged onto the detector. Assuming the system has been optimized according to the procedure outlined earlier in this chapter, we can assume that the intensity distribution of the output is roughly an Airy pattern. The detector size was chosen to be the size of the central Airy disc, so that $\eta_{enc} = 83.8$ % of the optical power focused by the lens actually falls on the detector. Finally, the detector converts some fraction η_{det} of the light falling onto it into a usable electrical signal.

In the following sections, we discuss the efficiencies which are dependent on the lasers and detectors used in the system. Then, we calculate the optical throughput efficiency which is dependent on the geometry of the shuffle architecture. From this, we obtain an overall power budget for signal power throughout the system. As a result, we expect to find the power lost as heat at the source (laser) array, which will allow us to establish the packing density limit set by the *signal* power (as opposed to the limits set by threshold dissipation and by the optical resolution).

Power Efficiency of the Laser

At the beginning of this chapter, we discussed the power dissipation associated with bringing the laser to threshold. In this section, we examine only the power dissipation in the signal itself (i.e. beyond threshold). The fraction of power injected into the laser past threshold which is converted to light is called the external differential quantum efficiency, η_{ext}. The remaining fraction of injected power $(1-\eta_{ext})$ is lost as heat. The quantity η_{ext} is the number of photons emitted by the laser for every electron of current injected into the laser past threshold, and is given by

$$\eta_{ext} = \frac{\left(\frac{P_{opt}}{h\nu}\right)}{\left(\frac{i-i_{th}}{q}\right)}. \tag{4.66}$$

Here, P_{opt} is the optical signal power, h is Planck's constant, $\nu = c/\lambda$ is the optical frequency, $(i - i_{th})$ is the signal current past threshold i_{sig}, and q is the charge of an electron.

Power injected into the laser past threshold which is not emitted as light is lost as heat. This signal heat dissipation is expressed by

$$P_{sig} = (1 - \eta_{ext}) \cdot (V_A i_{sig}), \tag{4.67}$$

where V_A is the voltage dropped across the laser diode junction. The total heat dissipated in the laser is the sum of the signal and threshold dissipation, or

$$P_{heat} = P_{sig} + P_{th}. \tag{4.68}$$

Because the source is likely to consume the greatest amount of power in the entire communication link, we will later compare the heat dissipated by the source P_{heat} to the breakeven power P_{crit} to determine if the system is limited by lens resolution or by heat dissipation.

Power Throughput of the Lenses

Once the laser emits an amount of optical power P_{opt}, the lens collects the fraction η_{lens} of that light, and images it into the output plane of the shuffle. We denote this by

$$P_{lens} = P_{opt}\eta_{lens} \tag{4.69}$$

where P_{lens} is the optical power emerging from the lens and imaged onto the output plane.

On-Axis Collection Efficiency Next, we evaluate the collection efficiency of the imaging lens and its effective aperture for a source on the optic axis as illustrated in Fig. 4.20. For simplicity, we ignore reflections at the lens and assume that all light impinging on the aperture passes through. We also assume the laser output intensity $I(x, y, z)$ is a beam with a circularly symmetric Gaussian intensity profile in the form [27]

$$I(x, y, z) \approx I_0(\frac{z_0}{z})^2 e^{-2\frac{(x^2+y^2)}{\omega^2}} \tag{4.70}$$

where I_0 is a constant, and the quantity $(\frac{z_0}{z})^2$ expresses the r^{-2} drop-off in peak intensity which results from the divergence of the beam as it propagates in z.

The parameter z_0 is given by

$$z_0 = \frac{\pi\omega_0^2}{\lambda} \tag{4.71}$$

where ω_0 is the beam waist (or the distance from the axis where the beam intensity has fallen to e^{-2} of its peak value on axis) at the laser. The parameter ω is the

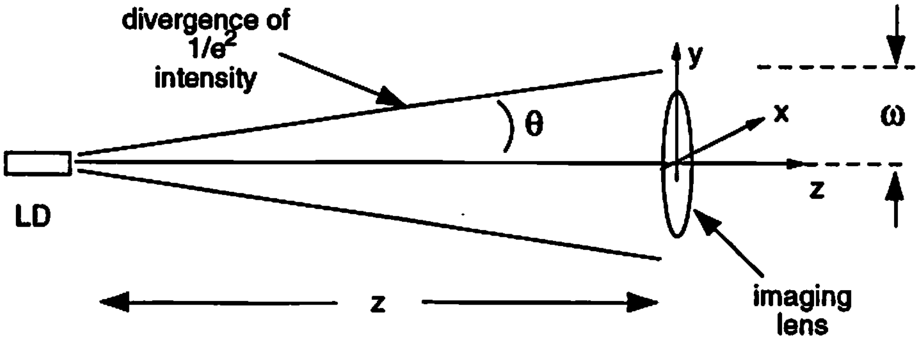


Figure 4.20: Illustration of the on-axis collection efficiency of a lens. Light with a Gaussian intensity profile diverges from a laser diode source (LD) along its propagation in z. The point at which the intensity is $1/e^2$ of the on-axis intensity is shown diverging at half-angle θ.

beam waist after propagation over a distance z (which in our case is the object distance, s, used earlier in the imaging relationships), and is given by

$$\omega \approx \frac{\lambda z}{\pi \omega_0} = \theta \cdot z \tag{4.72}$$

where θ is the half-angle of divergence as shown in Fig. 4.20. For laser diodes with cleaved facets, the divergence is typically not equal in x and y. The new vertical cavity lasers (VCSEL's) [15] have a circularly symmetric profile, and emit a beam which obeys the equation above for $I(x, y, z)$. For the shuffle, the distance z from the source to the lens (on-axis) is the same as the object distance s of the imaging set-up

$$z = s = \frac{3}{2} f. \tag{4.73}$$

So, the intensity profile of the beam given by Eq. (4.70) becomes

$$I(x, y, z) = I_0 \left(\frac{2\pi\omega_0^2}{3\lambda f}\right)^2 e^{-2\frac{x^2+y^2}{\omega^2}} \tag{4.74}$$

when calculated at the lens plane.

We find the collection efficiency η_{lens} of the lens by integrating the intensity profile over the area of the lens according to

$$\eta_{lens} = N \int\int_{lens} I(x,y,z)dxdy, \tag{4.75}$$

where N is a normalizing factor selected so that $\eta_{lens} = 1$ if the lens collects all of the light. In general, we have

$$0 \leq \eta_{lens} \leq 1. \tag{4.76}$$

From probability theory [29], the integration of a zero-mean Gaussian distribution yields

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx = 1, \tag{4.77}$$

thus, the normalization factor for 1-D, N_{1-D}, is

$$N_{1-D} = \frac{1}{\sigma\sqrt{2\pi}}, \tag{4.78}$$

and for a 2-D integration, we have

$$N_{2-D} = N_{1-D}^2 = \frac{1}{2\pi\sigma^2}. \tag{4.79}$$

To relate the parameters of the normal distribution to those of the Gaussian beam, we examine the beam profile at $z = z_0$, and take I_0 to be one. This way, the beam profile in Eq. (4.70) is just the exponential with no additional multiplicative factors, making it perfectly analogous to the integration of the normal distribution in Eq. (4.79). We equate the exponential for the beam profile with that of the normal distribution to determine σ for the beam. Knowing σ, we also know the normalizing factor, N, according to Eq. (4.79). Equating the exponents of both the normal distribution and the Gaussian beam, we have

$$\frac{1}{2\sigma^2} = \frac{2}{\omega_0^2}, \tag{4.80}$$

which implies that the σ of the Gaussian beam is

$$\sigma^2 = \frac{\omega_0^2}{4}. \tag{4.81}$$

Then the 2-D normalizing factor from Eq. (4.79) is just

$$N_{2-D} = \frac{1}{2\pi\frac{\omega_0^2}{4}} = \frac{2}{\pi\omega_0^2}. \tag{4.82}$$

After inserting the limits of integration and simplifying, the collection efficiency becomes

$$\eta_{lens}^{on} = \frac{8\pi}{9}\left(\frac{\omega_0}{\lambda f}\right)^2 \int_{-R}^{+R} dx \int_{-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} dy \cdot e^{-2\frac{(x^2+y^2)}{\omega^2}} \tag{4.83}$$

where the superscript *on* denotes the collection efficiency for a source on the optic axis of the lens.

Off-Axis Collection Efficiency We have so far considered the collection efficiency for a source located on the optic axis of the lens. The collection efficiency for an off-axis source is generally less than that of an on-axis source. In order to calculate η_{lens} for the off-axis case, we imagine the source of Fig. 4.20 moved off-axis to the position of the corner source. The extent of a divergent beam with a Gaussian intensity profile is infinite in principle, so the edge of the beam shown in Fig. 4.20 is not a hard edge. It actually represents the radial distance from the center of the beam at which the intensity falls to e^{-2} of its on-axis value. Thus, the beam from an off-axis source still fills the aperture of the lens just as does the on-axis beam, only with less power.

Figure 4.21 shows a projection of the source array onto the lens array plane, and indicates that the source in the corner is displaced exactly $\frac{L}{3}$, or one lens diameter D away from the optic axis in both the x and y directions. Thus, the center of the Gaussian profile hits the lens plane away from the optic axis, but the center of a Gaussian distribution is just its mean value. So, to calculate the off-axis collection efficiency, we use a non-zero mean in the Gaussian distribution,

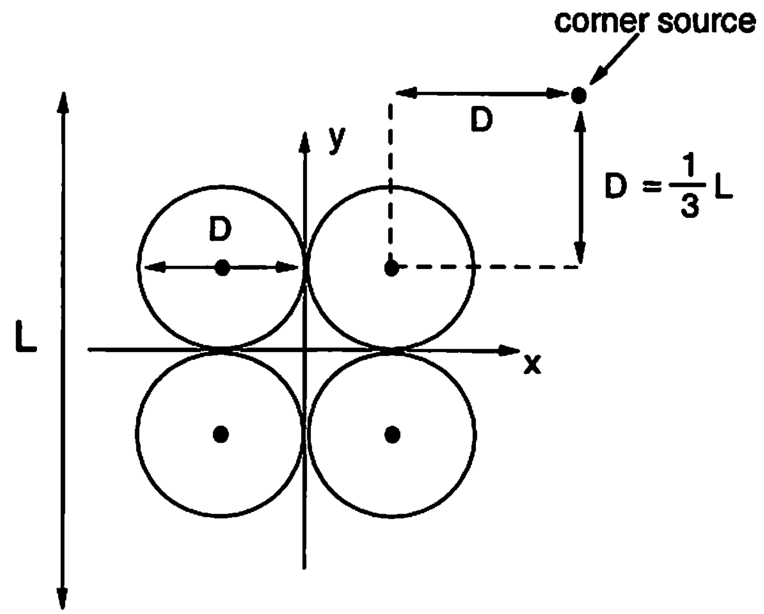


Figure 4.21: The position of a source at the extreme corner of the input array shown projected onto the lens plane of a 2-D shuffle system. The corner source is displaced from the optic axis by $\frac{1}{3}L$ in both x and y.

and repeat the integration given in Eq. (4.83). For the optical shuffle, the non-zero mean is the displacement, D, discussed above for the corner source, which has the worst η_{lens} for any source in the array. The collection efficiency integral is now

$$\eta_{lens}^{off} = \frac{8\pi}{9}(\frac{\omega_0}{\lambda f})^2 \int_{-R}^{+R} dx \int_{-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} dy \cdot e^{-2\frac{[(x-D)^2+(y-D)^2]}{\omega^2}}. \tag{4.84}$$

For the on-axis case, we expect the collection efficiency of the lens to increase as the divergence angle of the beam decreases. The case for off-axis illumination is not as obvious. As with the on-axis case, we expect lower η_{lens} for larger divergence angles. As we reduce the divergence angle, however, we expect that off-axis collection eventually falls. Thus, some optimum divergence angle must exist for off-axis collection. This fact is confirmed by the results of Fig. 4.22, where we used a numerical integration package and Eq. (4.84) to calculate η_{lens}^{off} versus half-angle divergence of the source. Figure 4.22 shows the off-axis collection

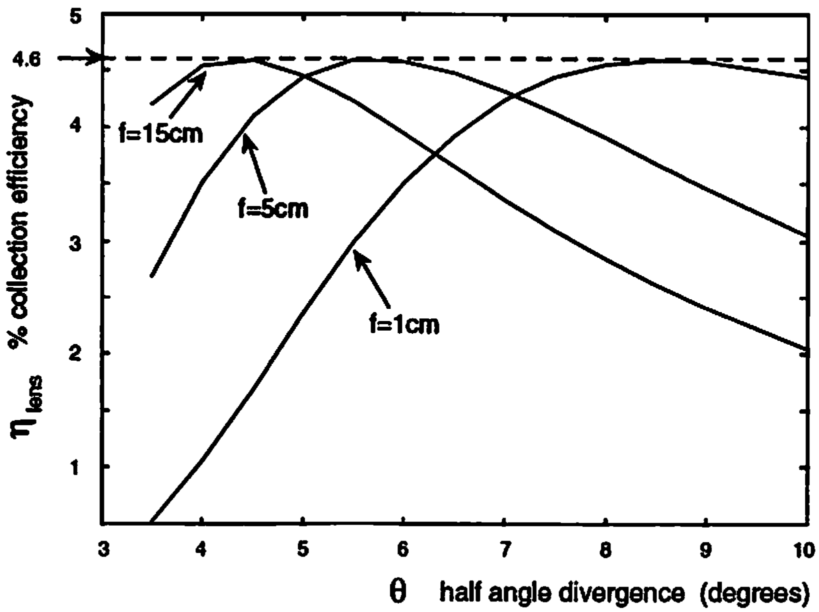


Figure 4.22: Off-axis collection efficiency versus the half-angle divergence θ plotted for optimized shuffle systems with lenses of three different focal lengths, $f = 1, 5, 15$ cm. The collection efficiency peaks at about 4.6 % regardless of focal length, provided that the divergence angle is chosen appropriately; $\lambda = 0.81$ μm.

efficiency for three different focal length lenses, 1, 5, and 15 cm. In all three cases, we have used the corner source in the optimum size array (as determined by the previously described optimization procedure) for each focal length.

Next, we compare the on- and off-axis cases in Fig. 4.23. Because it is always the worse case, off-axis collection efficiency always limits the power throughput of the system. The collection efficiency for an off-axis source consistently peaks at 4.6 %, regardless of the focal length lens we use, as shown in Fig. 4.22. With such a low efficiency, one might wish to steer the beam toward the center of the lens, either by tilting the laser, or with a prism or lenslet. For the sake of this discussion, however, we assume no additional beam steering.

Refering to Fig. 4.21, we see that crosstalk also occurs if light from a source near the edge of the input array is not blocked in the lens plane from proceeding directly through to the output plane. This *straight-through crosstalk* is plotted as

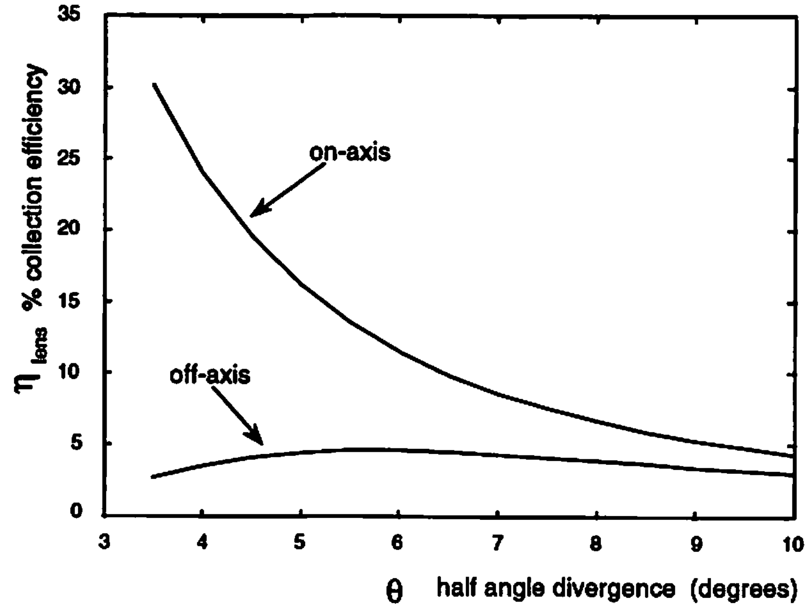


Figure 4.23: Comparison of on- and off-axis collection efficiency for $f = 5$ cm. The on-axis collection efficiency can be increased in principle to 100 % by reducing the divergence angle (i.e. collimating) the beam. Off-axis collection limits the power throughput of the shuffle, because it is always lower than the on-axis collection efficiency; $\lambda = 0.81$ μm.

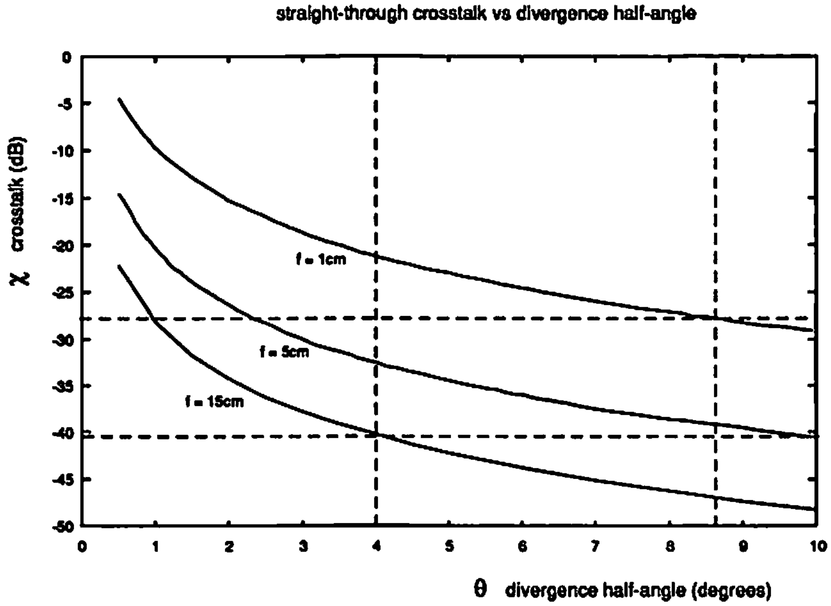


Figure 4.24: Straight-through optical crosstalk versus the half-angle divergence θ of the source for focal lengths $f = 1, 5, 10$ cm. Vertical dotted lines indicate optimum divergence angles for off-axis collection by the lenses of focal lengths 1 and 15 cm.

a function of the divergence half-angle θ of the source for focal lengths of 1, 5 and 15 cm. The vertical dotted lines indicate the optimum divergence angles for the $f = 1, 15$ cm, and show that crosstalk is worse for shorter focal lengths. In the 1 cm focal length case, straight-through crosstalk is roughly -28 dB. Of course, this issue is academic if the shuffle is constructed to block straight-through rays of sources near the edge of the array.

Power Collection at the Detector

We now examine the collection efficiency of the detectors. Recalling Fig. 4.7, we see that a detector collects the fraction η_{enc} of the light imaged to it. We have assumed the detector to be the size of the central Airy disc, the percentage of optical power emerging from the lens which falls on the detector is about 83.8 %.

So, the optical power actually incident on the detector is

$$P_{det} = P_{lens}\eta_{enc}, \tag{4.85}$$

where P_{lens} is the optical power emerging from the lens.

Receiver Sensitivity and Bit Error Rates

The detector converts optical power to a usable electrical signal with an efficiency η_{det}. Thus, the output electrical signal is

$$P_{e-out} = P_{det}\eta_{det}. \tag{4.86}$$

The efficiency with which the detector accomplishes the conversion is strongly dependent on several factors, amongst them the type of detector, the data rate, and the wavelength of light. Our intention is not to discuss these physical dependencies. We shall concentrate, however, on the issue of practical importance to our analysis.

In order for a receiver to detect binary data, and decide correctly which bits are logic 1 and which are logic 0, there is some theoretically required minimum optical power incident on the detector. This minimum power is dependent on the type of detector, the data rate, and the acceptable rate at which errors are made in identifying 1s and 0s. Many available references discuss these requirements (for example, Fig. 14.8 and Table 14.1 of reference [8]). We refer to the minimum optical power at the detector required to achieve a given error rate as $P_{e-out}|_{min}$. Thus, the electrical signal produced by the detector must satisfy

$$P_{det}\eta_{det} > P_{e-out}|_{min} \tag{4.87}$$

in order for the system to function at an acceptable bit error rate.

4.3.3 Example

The critical issue in calculating power dissipation at the source is the optical power available to the detector at the output. The detector must have sufficient optical power available to decide accurately whether the incoming digital signal is a 1 or a 0. So, for the sake of a practical example, we begin with the critical factor, the power at the detector, and work backwards through the steps in Fig. 4.19 toward the source. Starting with the minimum required optical power $P_{det}|_{min}$ for the detector being used will yield exactly the minimum heat dissipated by the source.

We calculate the power budget for the optical shuffle by assuming reasonable values for the system parameters. The applied voltage is roughly 1 V, a reasonable external efficiency for the laser is $\eta_{ext} = 0.20$. A reasonable sensitivity for a receiver operating around 100 Mbits/sec (at an error rate of 10^{-9}) is −40 dBm which implies $P_{min} = 0.1\ \mu$W [8] [2].

We have assumed that our detectors are the size of the central Airy disc, which means they encircle $\eta_{enc} = 83.8\ \%$ of the light in an image point. The best collection efficiency η_{lens}^{off} we can expect for a corner source if we tailor the divergence of the sources to the focal length of the lens being used is 4.6 %. Using lasers with an external differential efficiency η_{ext} of 20 %, we find that the heat dissipation per laser *due to the signal alone* is $P_{sig} \approx 6.8\ \mu$W. This is comparable to the critical powers calculated earlier, and shown in Table 4.2. So, if there were no power wasted as heat in bringing the laser to threshold, power dissipation due to the signal and the resolution of the optics (in this case simple lenses) would place comparable restrictions on the channel packing density of the optical shuffle.

The total heat dissipated at the laser is obtained by adding the threshold power P_{th} dissipation. In current devices, P_{th} is no better than approximately 1 mW, so the heat due to the signal is insignificant by comparison. This is further indication of the importance of lowering the threshold current of the lasers as much as possible.

4.4 Conclusion

We have shown that simple lenses have the capability to shuffle relatively large arrays of data lines. If care is taken to minimize the effects of both diffraction and aberrations simultaneously, such lenses should be able to interconnect tens, or even hundreds of thousands of data lines. The key to the minimization is the proper choice of the size of the array and lens aperture for the focal length lens being used.

These results also demonstrate that the packing density of an array of surface emitting lasers is limited by heat dissipation long before the capabilities of even simple lenses are exploited. In order to utilize fully the resolution available from a lens, either the threshold currents of these lasers must be reduced dramatically, or a different scheme for realizing the source array must be used. Two possibilities might be: (a) having the sources much farther apart, perhaps not even on the same chip or board, and injecting their signals into fibers. The fibers could then be brought together into a small square array so that a compact shuffle could be performed; or (b) use the light from one powerful laser to illuminate a hologram, which could then generate an approximately uniform square array of beams. These constant beams illuminate an array of independent modulators to make an input array with potentially lower power dissipation than an array of lasers.

Chapter 5

Tolerances on the System Construction

5.1 Background

When we optimized the geometry of the lens-based optical shuffle, we assumed certain ideal conditions which do not exist in a real system. Among these assumptions are that the detector plane is not misaligned in any way; that all four lenses have identical focal lengths, which are equal to the design focal length; and that all components are placed exactly as they should be in the design. Since none of these conditions can be met in practice, it is worthwhile to examine the sensitivity of the system to these errors. It is important to know whether any deviations from the ideal are catastrophic. Can we define reasonable limits for deviations from the ideal model, past which system performance becomes unacceptable? These are issues we shall address in this section. When we talk about the degradation of the system's performance, we mean the additional crosstalk incurred as a result of the imperfections just mentioned. In order to calculate the effects of imperfections, we return to the original crosstalk calculations, and determine how perturbations are incorporated.

5.1.1 Original Crosstalk Model

The procedure to optimize the geometry of the lens-based perfect shuffle results in an output spot in the image plane whose intensity distribution closely approximates the Airy pattern, given by Eq. (4.35). We calculated crosstalk due to neighboring channels in Section 4.2.7 by calculating encircled power in the Airy pattern. We integrated the intensity distribution of the Airy pattern over a circle centered at the Gaussian image point. The result, given by Eq. (4.38), was an expression for the fraction of total image power falling within a circle (e.g. a detector) versus the radius of the circle. When used in conjunction with the geometry of Fig. 4.8, we obtained the approximation for single neighbor crosstalk versus center-to-center spacing, given by Eq. (4.45).

5.1.2 Incorporating Perturbations into Crosstalk Calculations

There are two basic perturbations which we include in the crosstalk model. The first is defocus of the output image, which is caused by misplacement of either the input, output or lens planes, or by differences in the focal lengths of the four lenses used to implement the shuffle. In any of these cases, an output image point is blurred, meaning its intensity distribution does not match the ideal Airy pattern for which we had done the crosstalk calculations earlier. In general, more of the optical energy in the output is spread radially outward away from the Gaussian image point, with the likely result of increasing crosstalk in the neighboring channels. We show how the actual output intensity distribution is determined when defocusing occurs, and then use the intensity distribution to calculate crosstalk.

The second perturbation is when there is a slight misalignment between the output channel (i.e. the image) and the detector. Consider the geometry in Fig. 4.8 which we used earlier to calculate power at a detector not centered on the Airy pattern. A slight misalignment of the image and detector is like having

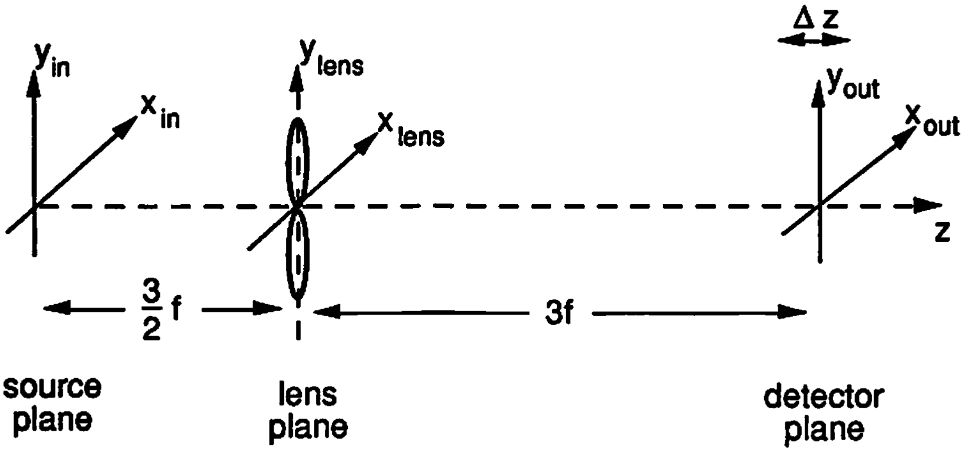


Figure 5.1: Moving either the source plane, lens plane, or detector plane results in a defocused output image.

D_2 partly overlapping D_1, where the normalized center-to-center spacing α is less than one. Equation (4.45), used previously to approximate the crosstalk at an adjacent detector, may not be sufficiently accurate for such small α. So, we discuss an alternative approximation, the area of overlap of two nonconcentric circles. Then, we numerically integrate the intensity distribution in the Airy pattern over an off-center circle to obtain a rigorous result, and compare that result to the approximations.

5.2 Defocus

5.2.1 Cause of Defocus in the Shuffle

The optimization procedure presented in Section 4.2.5 assumed that the output (detector) plane is placed in the image plane specified by Gaussian optics. This means that the output distance from lens plane to output plane is equal to the image distance s' prescribed by the thin lens imaging relation of Eq. (4.25). If, for any reason, the output array is not in the Gaussian image plane, the output channels will be defocused.

If any of the three planes shown in Fig. 5.1 (i.e. the source plane, lens plane, or the detector plane) is misplaced in z. For the sake of simplifying the calculations, we assume that the source (object) plane and the lens plane are fixed in their proper positions, so that the only misplaced item is the detector plane. We show in the following section how to calculate the effect of moving the detector plane out of the image plane. To first order, the effect of misplacing the lens plane in z is directly proportional to that of misplacing the detector plane in z. So, solving one problem effectively solves the other. A second order effect, however, of misplacing the lens plane is that the magnification of the input by 2 is slightly altered. The result is that the image points move laterally from the Gaussian image point. We consider only the first order effects here.

The other cause of defocus which we consider is that the lenses have focal lengths different from the desired focal length f. This is not a problem if all four lenses have the *same* focal length, since in practice s_o and s_i could be adjusted to accommodate the new focal length f'. The real problem arises when the four lenses have different focal lengths f_i, $i \in \{1,2,3,4\}$. Because the calculations of defocused output we shall undertake are computationally intensive and require a great deal of computer time, we shall limit our calculations to a worst case scenario where the focal lengths of all four lenses differ from the intended focal length f by the same amount Δf. Because, in our case the effect of $+\Delta f$ and $-\Delta f$ are nearly identical, it matters little which lenses are $f_{1,2}$ and which are $f_{3,4}$, only that all four are in error by $\pm\Delta f$

$$f_1 = f_2 = f - \Delta f \tag{5.1}$$

$$f_3 = f_4 = f + \Delta f. \tag{5.2}$$

5.2.2 General Effect of Defocus on Image Profile

In order to calculate the effect of these two types of defocus, we must find the intensity profile of a defocused image point. We do this using the diffraction integral, and then incorporate defocus as a phase aberration. We shall follow the analysis by Mahajan [23], [22] of point spread functions for rotationally symmetric aberrations. The intensity distribution in the output plane is given by

$$I_o(\vec{r_o}) = (\frac{P_{ex}}{S_{ex}\lambda^2 z^2})|\int\int e^{i\Phi_d(r_p)} e^{-i\frac{2\pi}{\lambda z}\vec{r_p}\cdot\vec{r_i}} d\vec{r_p}|^2 \tag{5.3}$$

where $\vec{r_o}$ is the radial vector from the Gaussian image point to the point where the intensity is being calculated, $\Phi_d(r_p)$ is the phase aberration of the wavefront (the subscript indicating defocus) at the radial distance r_p from the center of the pupil. P_{ex} is the optical power at the exit pupil (in this case the lens), S_{ex} is the area of the exit pupil, λ is (as always) the wavelength of the light, and z is the distance from the exit pupil (lens) to the detector plane. The integration is performed over the aperture of the lens (i.e. the exit pupil).

Defocus is a radially symmetric wavefront aberration, thus the argument r_p is the scalar distance from the center (or origin) of the pupil plane. The phase aberration may be calculated from the wavefront aberration $W(r_p)$ which is the deviation of the actual wavefront from a perfect converging spherical wave. In the case of defocus, this is

$$W_d(\rho) = \frac{1}{2}(\frac{1}{z} - \frac{1}{R})r_p^2. \tag{5.4}$$

where R is the correct image distance from the exit pupil, and z is actual distance from the exit pupil to the detector plane. The phase aberration Φ_d is obtained simply by multiplying by $2\pi/\lambda$, which yields

$$\Phi_d(\rho) = \frac{\pi a^2}{\lambda}(\frac{1}{z} - \frac{1}{R})\rho^2. \tag{5.5}$$

The argument ρ is the normalized radial distance from the center of the pupil, given by

$$\rho = \frac{r_p}{a} \tag{5.6}$$

where a is the aperture radius.

It can be shown [22] that the intensity distribution in the detector plane is

$$I(r) = \int_0^1 \int_0^1 \cos[\Phi(\rho) - \Phi(s)] \mathrm{J}_0(\pi r \rho) \mathrm{J}_0(\pi r s) \rho s d\rho ds, \tag{5.7}$$

provided r is measured in units of λF and the intensity is measured in units of $P_{ex}S_{ex}/\lambda^2 R^2$. We obtain the encircled power $P(r_c)$ simply by integrating the intensity distribution over a circle of radius r_c around the presumed image point. The result is

$$P(r_c) = 2\pi^2 \int_0^1 \int_0^1 \cos[\Phi(\rho) - \Phi(s)] Q(\rho, s; r_c) \rho s d\rho ds \tag{5.8}$$

where Q is depends on the integration variables ρ and s. If $\rho = s$, then

$$Q(\rho, s; r_c) = (\frac{r_c^2}{2})[\mathrm{J}_0^2(\pi \rho r_c) + \mathrm{J}_1^2(\pi \rho r_c)]. \tag{5.9}$$

If $\rho \neq s$, then

$$Q(\rho, s; r_c) = [\frac{r_c}{\pi(\rho^2 - s^2)}][\rho \mathrm{J}_1(\pi \rho r_c) \mathrm{J}_0(\pi s r_c) - s \mathrm{J}_1(\pi s r_c) \mathrm{J}_0(\pi \rho r_c)]. \tag{5.10}$$

The critical parameter remaining to be calculated is the phase aberration for defocus $\Phi_d(\rho)$. We calculate Φ_d based on the shuffle geometry.

5.2.3 Application to Lens and Detector Plane Misplacement

Once we compute the integral of Eq. (5.8), then we apply the same method of computing crosstalk employed before when the images were unaberrated. In order to accommodate an integral with the two functional forms of Q given by Eqs. (5.9) and (5.10), we have written a Matlab program to evaluate Eq. (5.8) numerically.

The one remaining parameter is the phase aberration $\Phi(\rho)$ at the lens.

$$\Phi_d(\rho) = \frac{\pi a^2}{\lambda}(\frac{1}{z} - \frac{1}{R})\rho^2. \tag{5.11}$$

If we pick a focal length for the lenses, then the shuffle geometry optimization procedure determines the optimum aperture radius a. Assuming we know the wavelength of light being used, the only remaining variables in Eq. (5.5) are the radius of the converging spherical wave R, and the actual distance z from the lens plane to the detector plane. Misplacing the detector plane simply changes z, so the phase aberration is directly available.

To demonstrate the proportionality between moving the detector plane and moving the lens plane, consider the following. The actual distance from the lens plane to the detector plane is affected only by the relative movement of the lens and detector planes. We adopt the convention that positive shifts are to the right, so if the lens plane is shifted by Δs_o (subscript o for object), and the detector plane is shifted by Δs_i (subscript i for image), then the output distance z is given by

$$z = 3f - \Delta s_o + \Delta s_i \tag{5.12}$$

The perturbation to s_o which results from moving the lens plane, however, affects also the radius R of the converging spherical wave. We can estimate this effect by using the thin lens imaging equation

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f}. \tag{5.13}$$

The radius of curvature R is the same as the correct image distance s_i for the perturbed object distance. So, we solve Eq. (5.13) for s_i in terms of the object distance and the focal length, which gives

$$s_i = \frac{s_o f}{s_o - f} \tag{5.14}$$

To find the effect of misplacing the lens plane, we differentiate s_i with respect to the object distance

$$\frac{\partial s_i}{\partial s_o} = \frac{(s_o - f)f - s_o f \cdot 1}{(s_o - f)^2} = -(\frac{f}{s_o - f})^2 \tag{5.15}$$

Then, for an incremental change Δs_o in the placement of the lens plane, we find the incremental change Δs_i in the distance from the lens plane to the image plane using

$$\Delta s_i \approx \frac{\partial s_i}{\partial s_o} \Delta s_o \tag{5.16}$$

We simplify the equations greatly by noting that the imaging arrangement for the shuffle is fixed so that

$$s_o = \frac{3}{2} f. \tag{5.17}$$

This implies that

$$\Delta s_i \approx -4 \Delta s_o. \tag{5.18}$$

Ordinarily, the image distance, or radius of curvature of the spherical wave converging on the image point, is $3f$. We add the change due to perturbing the lens plane to get

$$R \approx 3f + \frac{\partial s_i}{\partial s_o} \Delta s_o = 3f - 4\Delta s_o. \tag{5.19}$$

This demonstrates the proportionality between moving the lens plane and moving the detector plane.

5.2.4 Application to Errors in Lens Focal Lengths

Similarly, we find the effect of a slight inaccuracy in the focal length of the lenses by differentiating s_i with respect to focal length, which gives

$$\frac{\partial s_i}{\partial f} = (\frac{s_o}{s_o - f})^2. \tag{5.20}$$

Again, using Eq. (5.17) simplifies the result so that

$$\Delta s_i \approx 9\Delta f. \tag{5.21}$$

As before, the ordinary image distance is $3f$. We add to this the perturbation in focal length, which gives

$$R \approx 3f + \frac{\partial s_i}{\partial f}\Delta f = 3f + 9\Delta f. \tag{5.22}$$

The problem of incorrect focal length has two parts. First, the *average* focal length of the four is different from the desired value f. We shall presume that the system may be adjusted through s_o and s_i to bring the output back in focus for $f_a v$, while providing the correct magnification of the input. Deviations between the individual lenses, however, cannot be easily compensated because there is no single image plane. We have stated already that we shall consider only the worst case scenario where all of the lenses are at the maximum deviation $\pm\Delta fmax$ from the average focal length. Mahajan shows [23] that for image geometries with large Fresnel numbers, $N_z \gg 1$, given by

$$N_z = \frac{a^2}{\lambda z} \tag{5.23}$$

that the effect of defocus is nearly symmetric on either side of the correct image plane. So, we shall solve the defocus problem for a single, positive value of $+\Delta f_{max}$, assuming that the intensity distributions for $f + \Delta f_{max}$ and $f - \Delta f_{max}$ are nearly identical.

5.2.5 Results

Knowing the perturbations Δs_o, Δs_i and Δf, we may calculate both the correct and actual distances from the lenses to the detector plane, R (or s_i) and z, respectively. In turn, we may calculate the defocus phase aberration $\Phi_d(\rho)$ over the

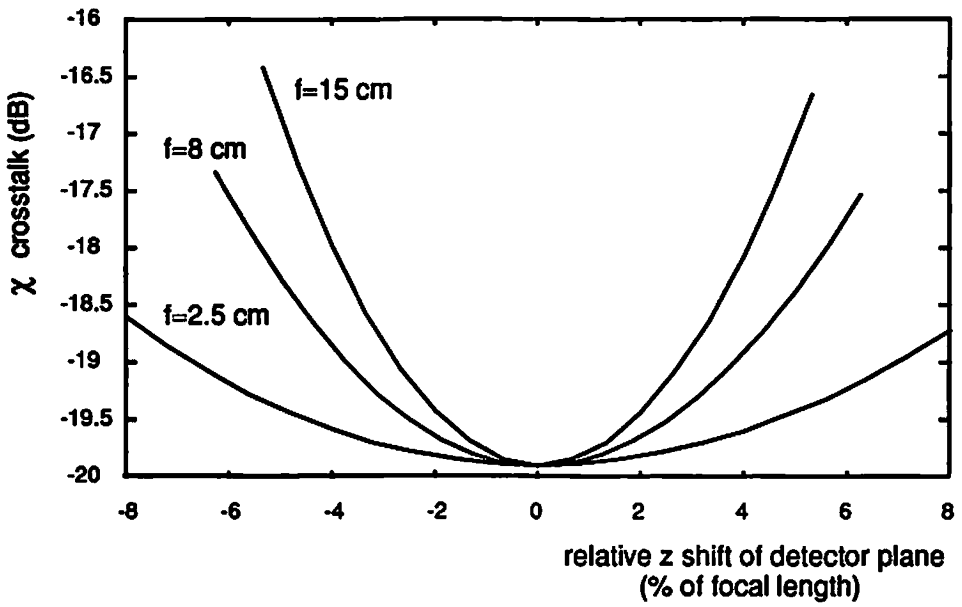


Figure 5.2: Plot of crosstalk versus shift of the detector plane. Crosstalk increases faster for the longer focal lengths, which have the higher optimized f-number, when the shift is measured *relative* to the focal length of the lens.

entire aperture of the lens, $0 \leq \rho \leq 1$. With this, we have Eqs. (5.7) and (5.8) - (5.10) to compute the intensity profile and encircled powers in the detector plane. With the encircled power around a defocused image point, we use the established procedure to compute the crosstalk.

Figure 5.2 shows the effect of misplacing the output plane for three different focal length lenses in an optimized shuffle system. The curves from highest to lowest are for 15, 8 and 2.5 cm focal lengths. The crosstalk is plotted in dB on the vertical axis, and the misplacement of the detector plane relative to the focal length is plotted in percent on the horizontal axis. At first glance, the results seem counterintuitive. If we refer back to the optimized system geometries for the various focal lengths, we see that the longer focal lengths have higher f-numbers. Usually, a higher f-number is associated with a greater depth of field, or tolerance for defocus. The opposite appears to be the case in Fig. 5.2. Crosstalk increases faster for the 15 cm focal length lens, which has an f-number of 16.98 when the

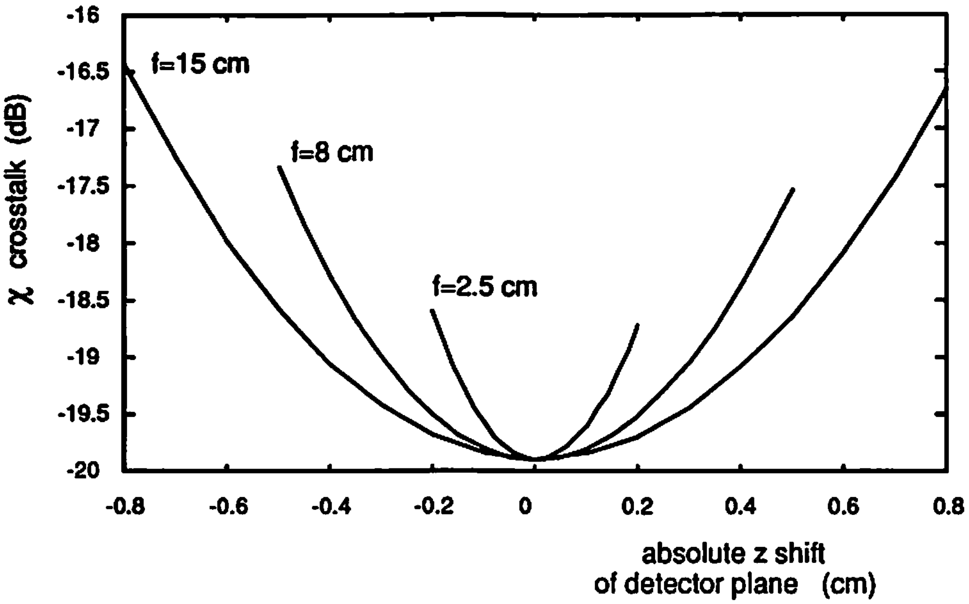


Figure 5.3: Crosstalk versus *absolute* misplacement of detector plane. Crosstalk increases faster for shorter focal lengths (which have lower optimized f-numbers) when the shift of the detector plane is measured as a physical distance (as opposed to a percentage of the lens focal length).

system geometry is optimum, than for the 8 cm focal length lens, which has an optimum f-number of 14.5.

The problem arises in how we plot the misplacement of the detector plane. Here, we have plotted it relative to the focal length of the lens. If we plot the actual physical misplacement of the detector plane, as we do in Fig. 5.3, the situation is reversed and intuition is satisfied. Here, the *bottom* curve is for 15 cm focal length, and the *top* is for 2.5 cm. So, for example, a 2 mm shift of the detector plane yields roughly 1 dB more crosstalk for the 2.5 cm focal length than for the 15 cm focal length. In relative terms, however, the more stringent requirement is on the longer focal lengths. The longest we have considered here is $f = 15$ cm. Figure 5.2 shows that this case exceeds 1 dB degradation in crosstalk after misplacing the output plane by about 2.5 % of the focal length.

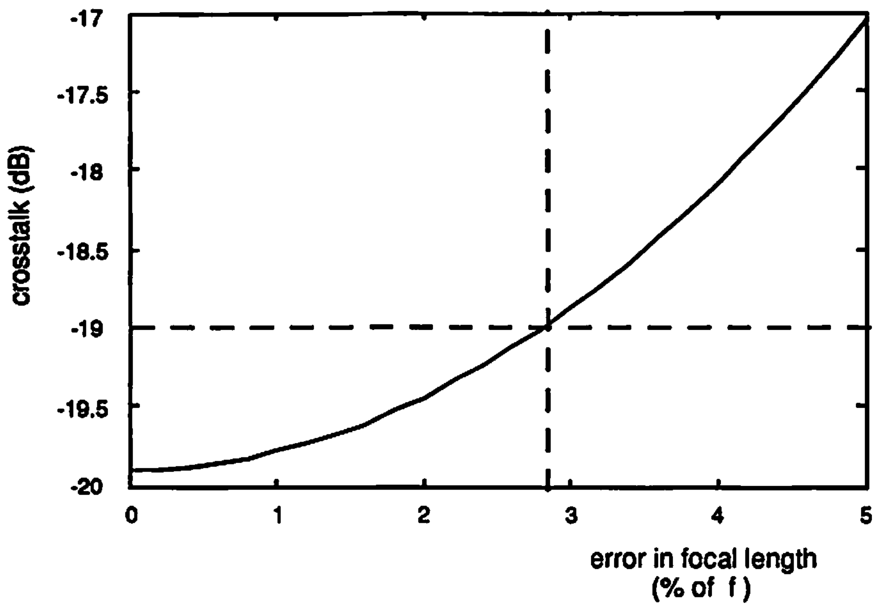


Figure 5.4: Crosstalk vs relative variation in focal length among the four lenses. The variation is the difference between the focal length of an individual lens and the average focal length of the four lenses. The plot represents the worst case where all four lenses differ by $\pm\Delta f$.

If the focal lengths of the lenses are not all exactly f, then there is no *one* image plane. This implies that there must be defocused image points in the detector plane. Our worst case analysis assumes that the variation $\Delta f/f$ around the average focal length is maximum for all four lenses. Figure 5.4 shows that a variation of roughly 3 % of f in the focal lengths will cause a 1 dB degradation in the crosstalk.

5.3 Misalignment of Image Points with Detectors

As with the focusing errors, there are a wide variety of causes for the misalignment of image points and detectors. We shall consider here only two. The first is that the sources and detectors are not perfectly registered on the same grid (i.e.

with the same spacing). This could be the result of small random errors in the placement of sources (and detectors). The other is that the lenses have not been misplaced laterally within the lens plane. According to the geometry developed by Stirk *et al* [36], the lenses must be centered $L/6$ or one-sixth the physical size of the array from the central axis of the system. Any error in lens placement implies a corresponding error in the position of the output images with respect to their intended detectors.

5.3.1 Received Power versus Image Shift

The one remaining issue is that the image intended for the central detector is also slightly misplaced. As such, less power than the original 83.8% will be encircled by the detector. Furthermore, because the displacement of the Airy pattern of the image is slight, the approximation given in Eq. (4.45), which was meant for a detector completely separate from the central Airy disc, may not work well. Therefore, we must either devise an approximation more appropriate for only a slight shift of the image off the detector center, or numerically calculate the encircled power integral, given by

$$P_{enc}(r_c) = \frac{1}{P_{ex}} \int_0^{r_c} \int_0^{2\pi} I_i(r_i) r_i d\theta dr_i \tag{5.24}$$

where P_{ex} is the power at the exit pupil, and the subscript i indicates quantities in the image plane. When the Airy pattern I_i and the detector are concentric (i.e. the image is properly positioned on the detector), this integral turns out to be easy to solve because it has a closed form solution. When the Airy pattern is shifted, however, the circular symmetry of the integral is broken, and a numerical calculation is necessary.

Approximation

Perhaps the simplest approximation one may devise for this case is to calculate the area of overlap of two circles the same size, but with some nonzero center-to-center spacing. Consider Fig. 5.5, where we show the center-to-center spacing as ϵ. The geometry is symmetric about the horizontal and vertical axes, so the problem simplifies to finding the area of the shaded region. This is a circular sector minus the indicated triangle. The area of the circular sector equals the area of the circle πr^2 multiplied by the fraction of the circle subtended by the angle θ, which gives

$$A_{seg} = \frac{\theta}{2\pi} \cdot \pi r^2. \tag{5.25}$$

The area of the triangle is

$$A_{\triangle} = \frac{1}{2}(\frac{\epsilon}{2})h \tag{5.26}$$

where the height of the triangle is

$$h = \sqrt{r^2 - (\frac{1}{2}\epsilon)^2}. \tag{5.27}$$

Subtracting $A_{\triangle}$ from A_{seg} and normalizing by the area of each circle πr^2, we obtain the fractional area of overlap of two shifted circles

$$A_{over} = \frac{2}{\pi}\{\cos^{-1}(\frac{\epsilon}{2r}) - (\frac{\epsilon}{2r})\sqrt{1 - (\frac{\epsilon}{2r})^2}\}, \tag{5.28}$$

which is identical to the optical transfer function (OTF) of a circular pupil [9].

Comparison of Approximations to Rigorous Results

We have calculated both the area of overlap of two nonconcentric circles of the same size, and numerically calculated the actual power received by a detector of radius $1.22\lambda F$ when an Airy pattern is shifted off the detector center. The results are shown in Fig. 5.6. The solid curve shows the approximate result given by the percentage of area overlap, and the dotted curve shows the rigorously derived

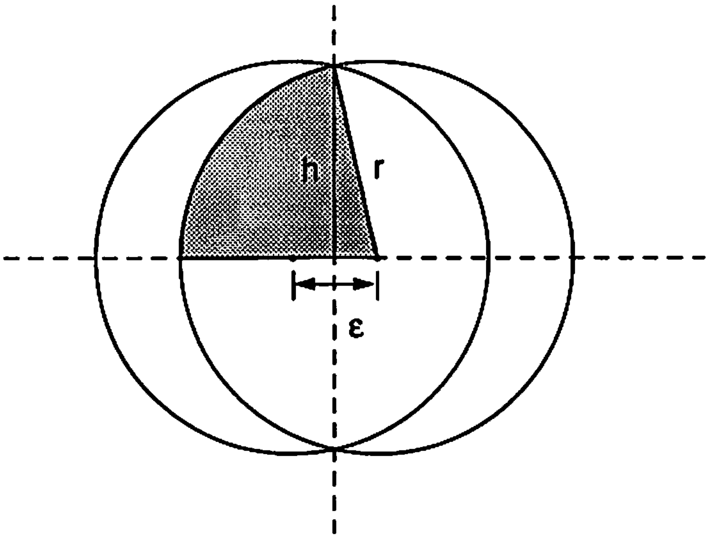


Figure 5.5: Geometry for calculating the overlap area of two equal size circles displaced by ϵ.

actual power striking the detector (as a percentage of when the image falls directly on the detector). The horizontal axis is the shift ϵ as a percentage of the detector diameter $2.44\lambda F$. So, the extent of the graph (100 %) is where the central Airy disc and the detector are edge-to-edge.

Figure 5.7 shows a smaller region, allowing us to see that for purposes of establishing a misalignment tolerance, there is a significant difference between the rigorous result and the area overlap approximation. If, for example, we wish to know how far off center the image point may fall while preserving 95% of the original power on the detector, the area overlap approximation shows that only a 4% misalignment is tolerable. The real situation, however, is more forgiving because power in the Airy pattern is concentrated near the center. The dotted curve in Fig. 5.7 shows that as much as a 23 % misalignment between image and detector may exist before a 5 % loss in power occurs at the detector.

We use the results of the rigorous calculation of power received when the image moves away from the detector center to assess the accuracy of the approximation for crosstalk in Eq. (4.45). Figure 5.8 shows the crosstalk for *one* image point

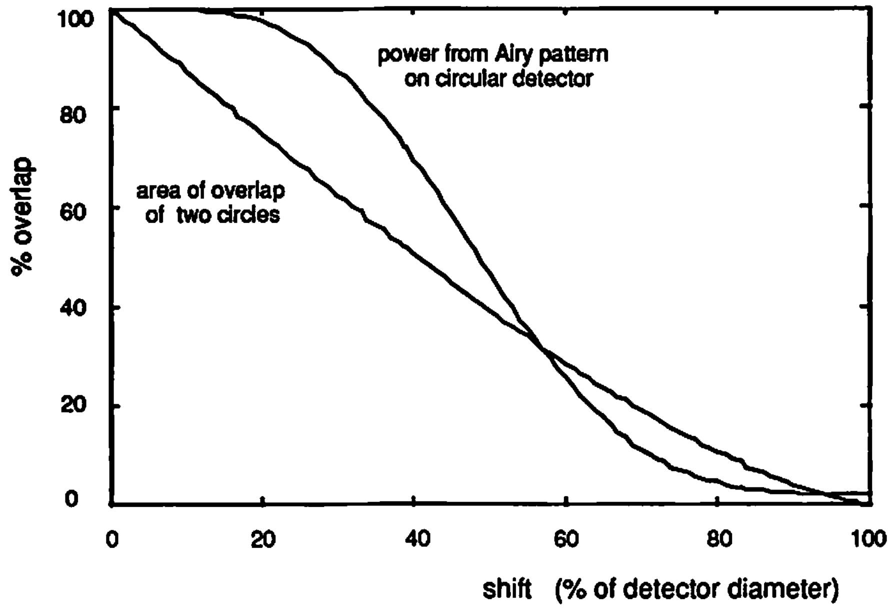


Figure 5.6: Comparison of percentage area of overlap (solid) and actual power in Airy pattern received by detector versus lateral image shift in percent of detector diameter.

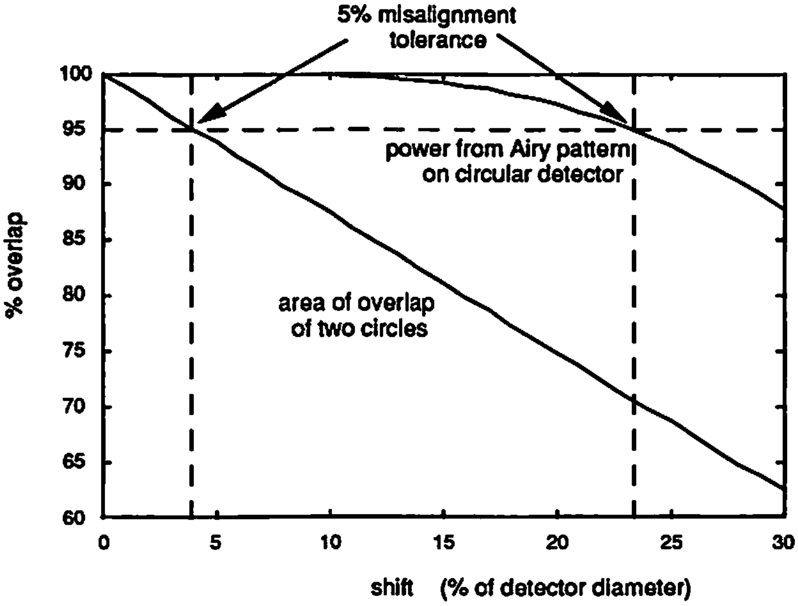


Figure 5.7: Tolerances on image misplacement established by area overlap approximation and rigorous integration.

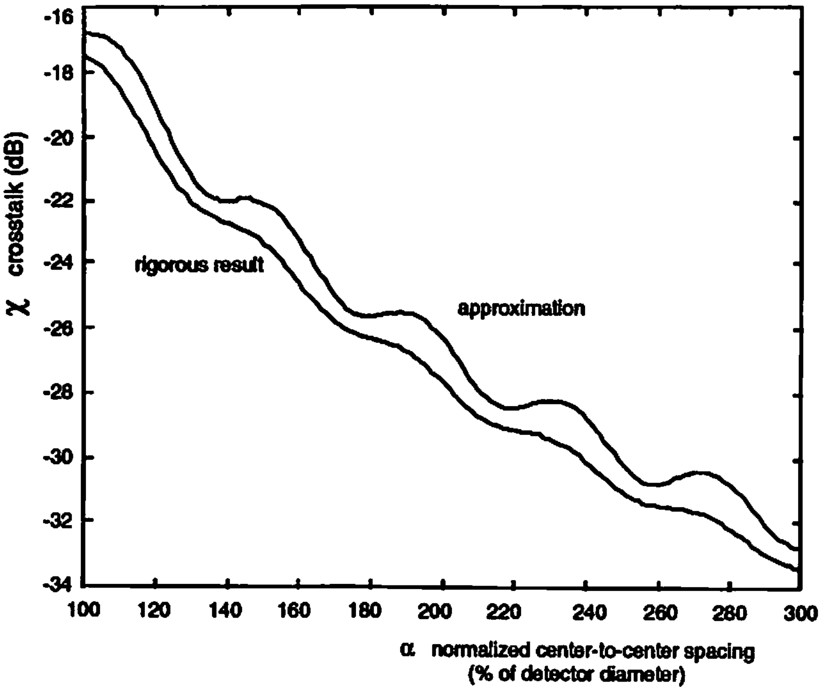


Figure 5.8: Comparison of single-neighbor crosstalk given by approximation and rigorous calculation.

versus center-to-center spacing between the detector and the image point. The center-to-center spacing is plotted as a percentage of the detector diameter, and the crosstalk is given in dB. The approximation is good to within roughly 1 dB over the range of one to three detector diameters. There is a qualitative difference between the approximation (solid curve) and the rigorously derived result (dotted curve). The rigorously derived result is a monotonically decreasing function of spacing, whereas the approximation is not.

We examine the point at which the approximation given by Eq. (4.45) predicts a detector spacing to preserve a net −20 dB crosstalk for all neighbors in the first two shells of neighbors. The requirement is that the detectors be spaced by 2.26 detector diameters. At that separation, a *single* image point contributes −28.29 dB of crosstalk according to the approximation, and −29.28 dB according to the rigorous integration. So, the approximation is inaccurate by roughly

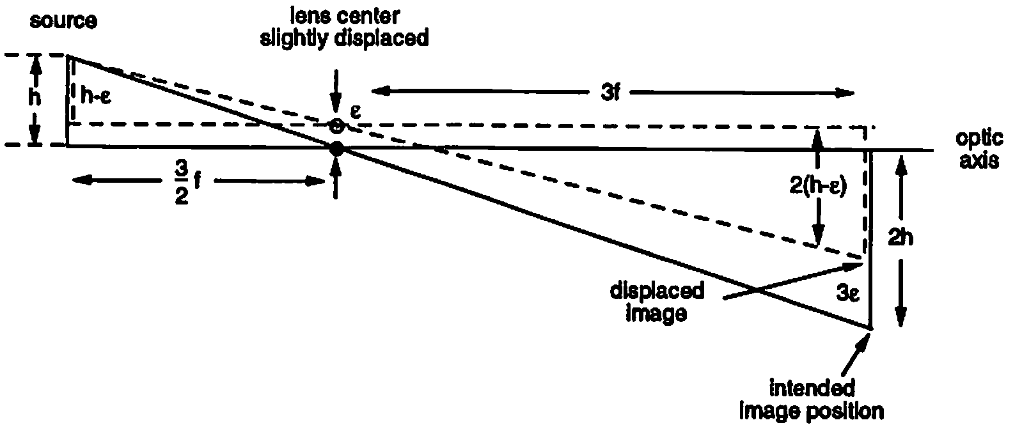


Figure 5.9: Any small lateral error ϵ in placing the lens results in a corresponding shift of the image point by 3ϵ.

1 dB. It overestimates the crosstalk, however, giving a conservative estimate of the requirement for interdetector spacing.

5.3.2 Image Shift due to Lateral Lens Misplacement

As shown by Sawchuk and Glaser [31], the position of the four lenses relative to the central axis of the system in a bulk optics implementation of the shuffle is critical to the interleaving process. Figure 5.9 shows that any lateral error ϵ_{lens} in lens placement results in a $3\epsilon_{lens}$ shift in its image points in the detector plane. Note that if all of the lenses are displaced by the same amount *and in the same direction*, then there is no real problem. The detectors and image points may be realigned simply by shifting the detector array. This is because of the shift-invariance of the lens-based shuffle.

A different problem arises when the lenses are misplaced in different directions. Adjacent image points in the detector plane are formed by different lenses. This means that if two lenses are displaced each by ϵ_{lens} in opposite directions, then their image points will be formed either $6\epsilon_{lens}$ closer to, or farther away from each other.

To analyze the situation, we realize that there is one maximum error ϵ_{max} among the various misplacements of the four lenses. The worst case for crosstalk occurs when all of the lenses are misplaced either outward or inward from their correct positions by the maximum amount in each lateral direction x and y, as shown in Fig. 5.10. Here, the actual lenses are shown shaded, and are displaced outward by $+\epsilon_{lens}$. We illustrate the process of calculating the effect on crosstalk by considering the central channel and its four closest neighbors, which are shown in black and labelled A through D in Fig. 5.10. Each of the four neighboring image points is now at a slightly different distance (we have drawn the figure to scale) than the correct center-to-center distance of 5.514 λF calculated earlier to preserve −20 dB crosstalk.

To calculate the new distances of the neighboring image points from the central detector, we let the central detector lie at the origin of the (x, y) coordinate system. If we call s the distance from the origin to the image center (x_i, y_i), and U the −20 dB spacing of 5.514 λF, then the distances of the various image points are obtained as follows. The distance to the image point A is just the norm of its coordinates x_{iA} and y_{iA} given by

$$S_A = \sqrt{x_{iA}^2 + y_{iA}^2} = \sqrt{(0 + 3\epsilon_{lens})^2 + (U - 3\epsilon_{lens})^2} \tag{5.29}$$

$$s_A = \sqrt{U^2 - 6U\epsilon_{lens} + 18\epsilon_{lens}^2}. \tag{5.30}$$

We normalize everything under the radical to U so that

$$s_A = U\sqrt{1 - 6\epsilon_{rel} + 18\epsilon_{rel}^2} \tag{5.31}$$

where ϵ_{rel} is the shift of the lens relative to the −20 dB image separation of 5.514 λF

$$\epsilon_{rel} \equiv \frac{\epsilon_{lens}}{U}. \tag{5.32}$$

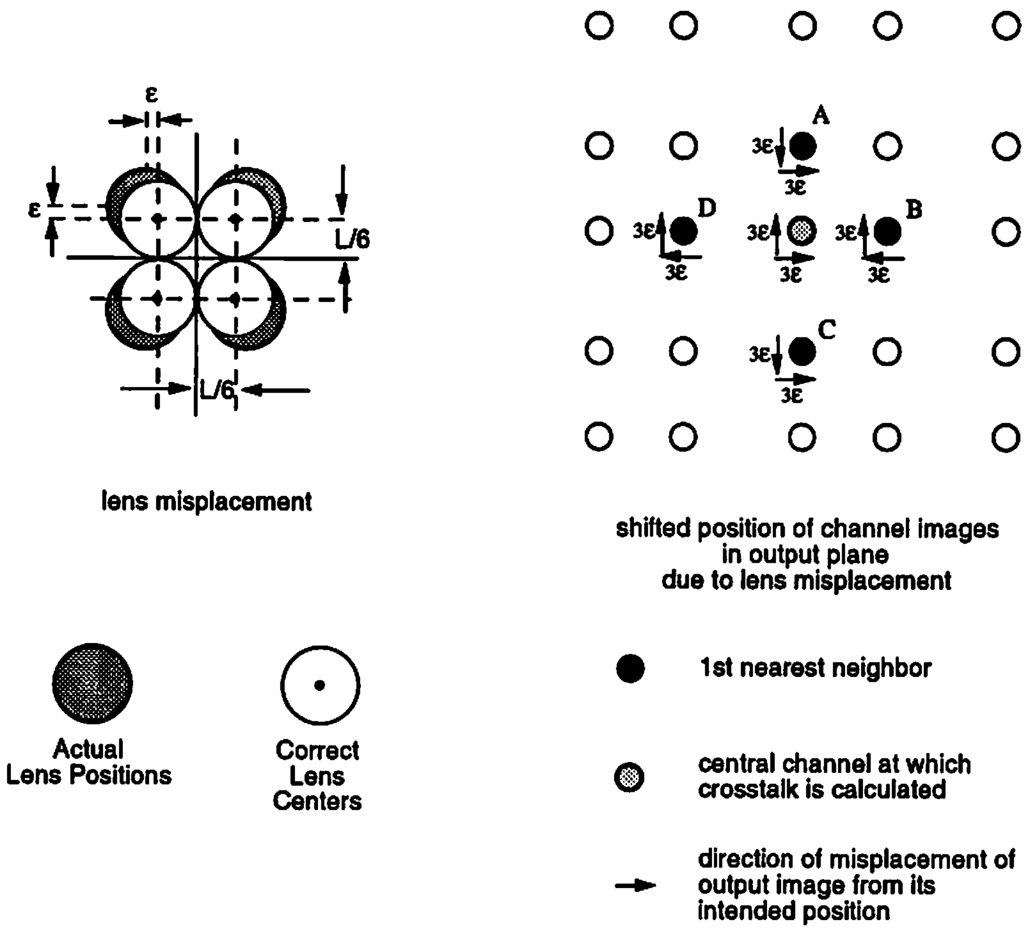


Figure 5.10: The left hand side shows both the correct position of the lenses (in clear circles) and their misplacement (shaded). The right hand side shows *to the correct scale* the resulting skew of the neighboring image points. The arrow next to each channel show the shift in x and y of its image in the output plane relative to its correct position. For example, the central channel (shaded) is 3ϵ above and 3ϵ to the right of where it should be. In general, the equal spacing of the image points is disrupted by the misplacement of the lenses.

Ignoring the terms in ϵ_{rel}^2 and using the binomial approximation, we obtain a first order approximation for the distance of image point A to the central detector

$$s_A \approx 1 - 3\epsilon_{rel} \tag{5.33}$$

where s_A is measured in units of U. We apply the same procedure to the three remaining image points to get

$$s_B \approx 1 - 3\epsilon_{rel} = s_A \tag{5.34}$$

$$s_C \approx 1 + 3\epsilon_{rel} \tag{5.35}$$

$$s_C \approx 1 + 3\epsilon_{rel} = s_C. \tag{5.36}$$

We pursue the same reasoning through the entire set of nearest neighbors in the first two shells. Table 5.1 contains the results. In it, we show the previous correct distances s in units of U of the various neighbors and how many neighbors were located at that distance before being perturbed. Then, we show how the lens misplacement of ϵ_{rel} has changed the distances to s' (also in units of U) and how many image points are now located at each of these new distances. With these perturbations on the center-to-center spacings between neighbor image points and the locally central detector, we may use Eqs. (4.45) and (4.47) to calculate the crosstalk at the central channel.

5.3.3 Crosstalk versus Lens Misalignment

Incorporating the results of the previous sections, we may calculate the effects of lateral misplacement of the four lenses. Figure 5.11 shows the crosstalk in dB versus lens misplacement given in units of λF. If we wish to restrict the degradation in crosstalk to, for example, 1 dB, then the lenses must be placed correctly in both the x and y directions to within 0.24 λF. Considering that the lenses are on the

shell	$\frac{s}{U}$	n	$\frac{s'}{U}$	n'
1	1	4	$1 - 3\epsilon_{rel}$	2
			$1 + 3\epsilon_{rel}$	2
	$\sqrt{2}$	4	$\sqrt{2}$	2
			$\sqrt{2} - 3\sqrt{2}\epsilon_{rel}$	1
			$\sqrt{2} + 3\sqrt{2}\epsilon_{rel}$	1
2	2	4	$2 - 3\epsilon_{rel}$	2
			$2 + 3\epsilon_{rel}$	2
	$\sqrt{5}$	8	$\sqrt{5} - 9\frac{\sqrt{5}}{5}\epsilon_{rel}$	2
			$\sqrt{5} + 9\frac{\sqrt{5}}{5}\epsilon_{rel}$	2
			$\sqrt{5} - 3\frac{\sqrt{5}}{5}\epsilon_{rel}$	2
			$\sqrt{5} + 3\frac{\sqrt{5}}{5}\epsilon_{rel}$	2
	$2\sqrt{2}$	4	$2\sqrt{2}$	2
			$2\sqrt{2} - 3\sqrt{2}\epsilon_{rel}$	1
			$2\sqrt{2} + 3\sqrt{2}\epsilon_{rel}$	1

Table 5.1: New distances s' to the central channel as a result of lateral misplacement of the lenses by ϵ_{rel}, and the number of neighbor images n' at each of these new distances.

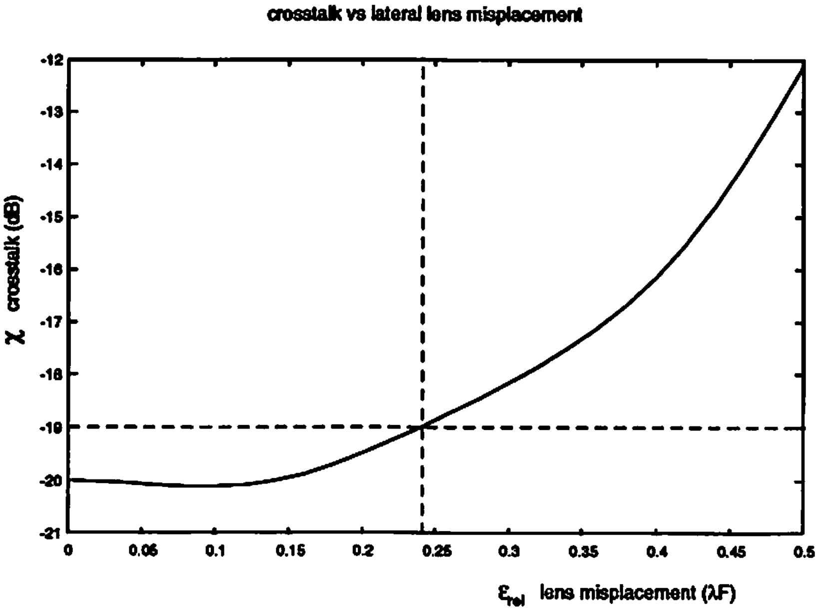


Figure 5.11: Crosstalk versus lateral lens misplacement, and the resultant 1 dB tolerance at 0.24 λF.

order of 1 cm in diameter, and that λF is on the order of a few μm, this constitutes a stringent requirement.

We also note that the graph of crosstalk in Fig. 5.11 is not monotonically increasing in lens misplacement. There is a slight drop between 0 and 0.15 dB. This is likely due to the fact that we used the approximation in Eq. (4.45) to calculate the crosstalk from the neighbors, and only used the exact solution for the misplacement of the image point on the central detector. Recall that the approximation yielded a nonmonotonic graph of crosstalk versus spacing.

5.3.4 Source-Detector Misregistration

Any time a source in the input plane is misplaced laterally from its intended position, then its image in the output plane is also shifted from its intended position. One may reasonably expect small random errors in the placement of the sources. Because of the magnification by two of the optics, any error ϵ_{src} in

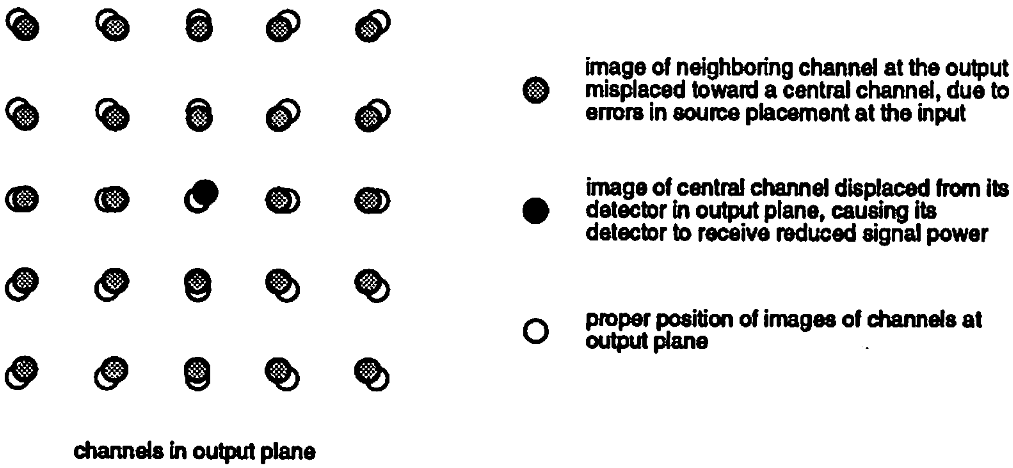


Figure 5.12: Worst case scenario for random misplacement of sources in a local region. Here, sources are misplaced such that their corresponding images in the output plane move inward toward a central point, as shown in the diagram, and the central image is misplaced outward. Displacement of neighboring images toward the central channel increases their contribution to crosstalk. Misplacement of the central image decreases the signal power received.

the placement of a source results in a corresponding shift $2\epsilon_{src}$ of its image in the output plane.

The worst case scenario, illustrated in Fig. 5.12, occurs when sources are misplaced such that their corresponding images in the output plane are all displaced inward by toward a central channel where we measure crosstalk. The inward movement of the neighboring images causes an increase in the crosstalk received by the central detector from neighboring channels. The worst case scenario also contains a shift of the central channel's image, which is shown in black in Fig. 5.12. This decreases the desired power actually received by the central detector. Both factors serve to degrade the crosstalk ratio χ.

We use the approximation in Eq. (4.45) to calculate the crosstalk contributions of the neighboring image points. Then, we use the results of the numerical integration shown in Fig. 5.7 to determine the signal power received at the central channel with the misplacement of the central image point. Figure 5.13 shows the results. The amount of misalignment of the sources ϵ_{src} is plotted in units

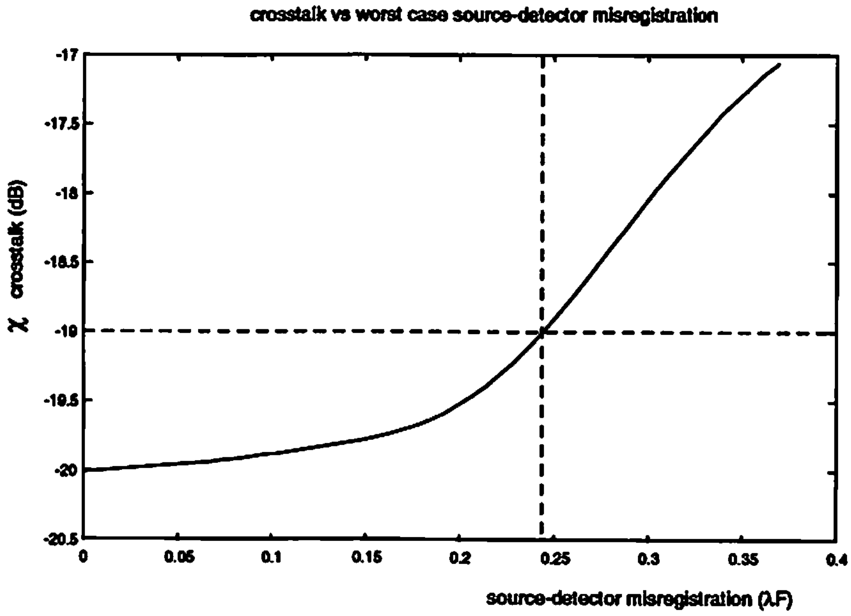


Figure 5.13: Crosstalk χ in dB versus worst case source misplacement ϵ_{src}; all sources are misplaced such that their images in the output plane shift inward; and the 1 dB tolerance of 0.24 λF.

of λF on the horizontal axis, and the resultant crosstalk is plotted in dB on the vertical axis. If we establish a limit of 1 dB degradation in crosstalk, then the maximum source misplacement allowable is roughly 0.24 λF. To better understand the requirement, consider the shuffle implemented with a 1 cm focal length lens. Assuming a wavelength of 0.81 μm, this means that sources must be within 1.9 μm of their intended positions. Given the capabilities of current lithographic processes, this is neither a severe nor a trivial restriction.

5.4 Conclusion

We have extended the analysis of the lens-based shuffle beyond the ideal, optimized geometry developed in Chapter 4. The construction of any real system always involves deviations from the intended design and a corresponding deleterious effect on the system's performance. We have shown how simple perturbations may be applied to the crosstalk model of Section 4.2.7, to show quantitatively the effects of defocus and misalignments on optical crosstalk. This in turn may be used to establish tolerances on the construction of the system.

Chapter 6

Effect of Crosstalk on Bit Error Rate

6.1 Background

We are concerned about the level of crosstalk in our network because crosstalk is a corrupting element which causes errors in digital signal regeneration. Any time a signal is received, routed and re-transmitted, a decision is made whether the signal is a logic 1 or a logic 0. The device responsible for deciding the logic level of an incoming signal is essentially a thresholding device. Any signal below the decision threshold is interpreted as a logic 0, any signal above the decision threshold is a logic 1. In an ideal setting, transmitted signals entering the receiver are quantized at one of the two signal levels shown in Fig. 6.1. In the absence of noise or channel distortion, no received signals could fall between the two logic levels, thus there would be no errors when deciding the logic level of the incoming signal.

Real signals, however, are not completely deterministic. There is always some variation (noise) around the expected signal level due to physical processes which are fundamentally random. Using a simple model for additive thermal noise, each of the two expected signal levels is a random variable having a Gaussian

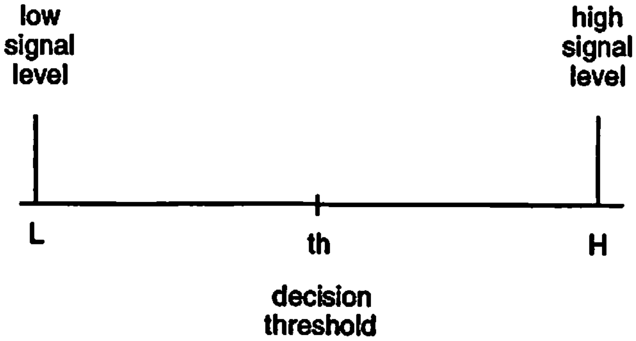


Figure 6.1: The ideal digital signal is one of only two quantized levels, either *high* or *low*, making recognition a simple task.

probability density. Though they may be sharply peaked around the logic high and logic low levels, there is an infinite tail on any Gaussian distribution. Consider Fig. 6.2, where a Gaussian probability density is centered at each of the two signal levels. The infinite extent of the Gaussian curves implies that there is a non-zero probability of making a decision error.

In this chapter, we examine the effect that optical crosstalk in an array of channels has on the error rate in digital signal regeneration. Because there is a straightforward and established method of calculating the effect of Gaussian noise on error rates, it is common to treat crosstalk as an additional Gaussian noise process. We investigate the accuracy of this assumption by rigorously developing the probability distribution of optical crosstalk in isolation (i.e. in the absence of other noise) in an array of free-space optical communication channels. We first review the established method of predicting error rates due to Gaussian noise. Then, we discuss the qualitative aspects of optical crosstalk, some of which distinguish crosstalk from Gaussian noise processes at a fundamental level.

We develop the quantitative rules governing optical crosstalk and obtain its probability distribution. We combine optical crosstalk with Gaussian noise to determine their net effect on error rates. With these results we are able to compare the rigorously-derived effect of optical crosstalk with the error rates predicted by assuming that optical crosstalk is another Gaussian noise process. We find that

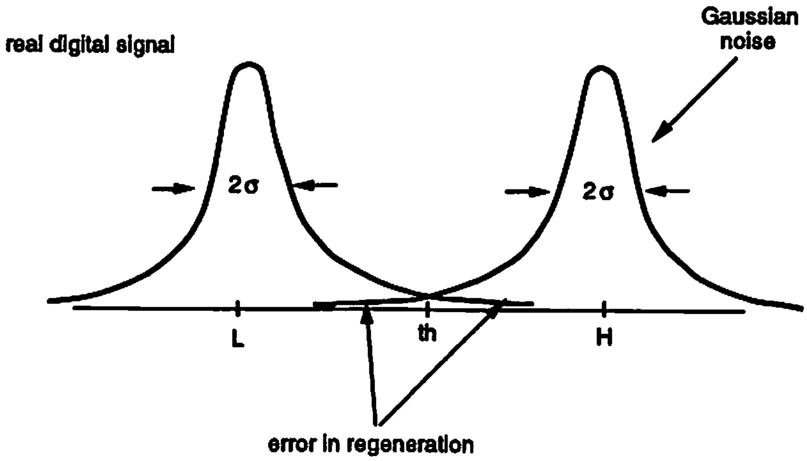


Figure 6.2: Real digital signals do not equal the two ideal levels due to noise imposed upon the signal in transmission or reception. For the example of purely additive Gaussian noise with variance σ as shown here, the tails of the signal distributions cross the decision threshold, giving rise to decision errors.

there is a difference between the two predictions, and as one might expect, the difference is greater when crosstalk is the dominant corrupting influence on the signal. If crosstalk is small compared to Gaussian noise, then treating crosstalk as an additional component of the Gaussian noise is relatively accurate. Finally, we find that while there is a quantitative difference between the rigorous result and the Gaussian approximation, it is insignificant when one wishes to obtain an order-of-magnitude estimate of the bit error rate.

6.2 Review of Error Rates for Gaussian Noise

6.2.1 General Error Probability

For the case of Gaussian noise, there is an established method [11] of predicting the rate at which errors are made in the process of digital signal regeneration. Referring to Fig. 6.2, the net probability of error, or bit error rate (BER), is

$$P(E) = P(1|0) \cdot P(0) + P(0|1) \cdot P(1), \tag{6.1}$$

where $P(1|0)$ is the probability of deciding that a 1 was received when a 0 was actually sent, and $P(0|1)$ is the probability of deciding that a 0 was sent when a 1 was actually sent.

We assume for simplicity that the probabilities of sending a 1 (*high*) and a 0 (*low*) are equal so that

$$P(1) = P(0) = \frac{1}{2}, \tag{6.2}$$

which implies that the decision threshold T is half way between the logic 1 and 0 signal levels

$$T = \frac{1}{2}(H - L) = \frac{1}{2}H, \tag{6.3}$$

and that the variance of the Gaussian noise around each of the two signal levels is identical. The net probability of error is

$$P(E) = P(1|0) \cdot \frac{1}{2} + P(1|0) \cdot \frac{1}{2} = P(1|0) \tag{6.4}$$

6.2.2 Application to Gaussian Noise

To apply the general result of Eq. (6.4) to the case of Gaussian noise, we must obtain both $P(1|0)$ and $P(0|1)$. For identical Gaussian distributions around both the 0 and 1 signal levels, and a decision threshold half way between 0 and 1 as defined in Eq. (6.3), these two probabilities are identical

$$P(1|0) = P(0|1). \tag{6.5}$$

So, the net probability of error in Eq. (6.4) is equal to one (or the other) of the two

$$P(E) = P(1|0). \tag{6.6}$$

This is the sum (or integral) of all probability above the decision threshold on the Gaussian distribution centered at logic 0, given by

$$P(E) = \int_T^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} dx, \tag{6.7}$$

where σ is the root-mean-square (rms) noise level. We define the signal-to-noise ratio as the ratio of the high signal level to the rms noise level

$$SNR \equiv \frac{H}{\sigma}. \tag{6.8}$$

Letting $t = x/\sigma$, the probability of error is the tabulated integral

$$P(E) = \frac{1}{\sqrt{2\pi}} \int_{\frac{1}{2}SNR}^\infty e^{-t^2/2} dt = \mathrm{erfc}(\frac{1}{2}SNR) \tag{6.9}$$

where erfc(x) is the complimentary error function [29].

6.2.3 Bit Error Rate Approximation

To simplify computation of the BER, we relate the integral in Eq. (6.9) to $I(x,p)$ [1]

$$I(x,p) = \int_x^\infty e^{-u} u^{-p} du \tag{6.10}$$

which has a known asymptotic expansion. Integrating $I(x,p)$ by parts repeatedly yields

$$I(x,p) = \frac{e^{-x}}{x^p} - p\frac{e^{-x}}{x^{p+1}} + p(p+1) \int_x^\infty e^{-u} u^{-p-2} du. \tag{6.11}$$

Relating terms in the error function integral, we let

$$u = \frac{1}{2}t^2 \tag{6.12}$$

$$du = tdt, \tag{6.13}$$

which then implies that

$$u^{-p}du = dt. \tag{6.14}$$

Substituting for u (removing the factor of 2) and du, we get

$$\frac{1}{t^{2p}} = \frac{dt}{tdt} = \frac{1}{t} \tag{6.15}$$

$$p = \frac{1}{2}. \tag{6.16}$$

The probability of error is

$$P(E) = \frac{1}{2\sqrt{\pi}} \int_{t=Q}^{\infty} e^{-u} u^{-p} du = \frac{1}{2\sqrt{\pi}} \int_{u=\frac{1}{2}Q^2}^{\infty} e^{-u} u^{-p} du. \tag{6.17}$$

where Q is one half the SNR. The first term in the expansion of $I(x,p)$ given by Eq. (6.11) gives an approximation for the BER

$$P(E) \cong \frac{1}{2\sqrt{\pi}} \frac{e^{-\frac{1}{2}Q^2}}{(\frac{1}{2}Q^2)^{1/2}} \cong \frac{1}{Q\sqrt{2\pi}} e^{-\frac{1}{2}Q^2} \tag{6.18}$$

which is valid for large signal-to-noise ratios ($Q > 5$), this is an excellent approximation for the BER [1].

6.3 Qualitative Description of Crosstalk

Like other noise components, crosstalk contributes to the probability of error in the digital signal regeneration process. It is common to treat crosstalk simply as an additive Gaussian noise process, and use established means for calculating the effect of Gaussian noise on the error rate. This section contains a more rigorous model of the effect of optical crosstalk on the bit error rate, and a discussion of how the results differ from the Gaussian model.

For now, we exclude all other types of noise (e.g. thermal noise, signal distortion) besides crosstalk, which is simply the leakage of signal energy from one channel to a neighboring channel. For the case of optical crosstalk, the portion of energy leaking to a neighboring channel is entirely predictable based on the geometry of the situation. We have already illustrated this fact in great detail in this work. In other words, given the signal level at the input, then based on the geometry of the system, we can model and predict (deterministically) the optical crosstalk. At a fundamental level, its deterministic nature distinguishes crosstalk from Gaussian noise. If one extends the discussion to electrical crosstalk, which is vastly more difficult to calculate than optical crosstalk because it involves far more physical variables, the same principles apply. Electrical crosstalk is predictable based on the geometry of the system and the signal levels on neighboring channels.

Depending on which neighbors are on at a given time, a channel experiences one of many discrete levels of crosstalk. There is some worst-case amount of crosstalk, which occurs when all of the neighboring channels in the system are on. This is the most important qualitative distinction between crosstalk and Gaussian noise. There is no infinite tail in the probability distribution of crosstalk. One may argue that this distinction becomes insignificant in a system where there are many neighbors. The assertion is that the data on neighboring channels is pseudo-random, because any one channel does not know the signal on neighboring channels. Thus, crosstalk is the sum of many pseudo-random variables. The result tends in the limit to a Gaussian distribution, based on the central limit theorem of probability.

The similarity, however, between a Gaussian distribution and the crosstalk probability distribution is greatest near the center of each. Far from the center, the accuracy of the approximation becomes poor. The crosstalk distribution has a sharp cut-off, beyond which the probability of any additional crosstalk is identically zero. The Gaussian tail never becomes identically zero for any finite signal level. Furthermore, for the purposes of calculating error rates, it is precisely the tail of the curve which is most important, as we have seen in Fig. 6.2. That the crosstalk may behave like a Gaussian near its centroid is of little consequence.

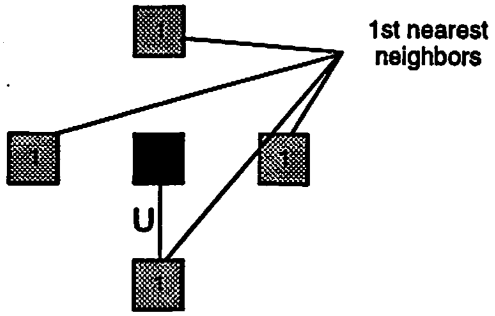


Figure 6.3: The channel receiving crosstalk is indicated in black, and the set of neighbors closest to this channel, its first nearest neighbors, are indicated with partial shading. Each channel in the first neighborhood are a distance U from the central channel.

6.4 Quantitative Description of Crosstalk

6.4.1 Nearest Set of Equidistant Neighbors

In order to determine the statistics of crosstalk, we assume there is *no* Gaussian noise in the signals. The only agent corrupting the signal on a given channel is crosstalk. Then, we break the problem of determining the statistics into smaller parts. Consider Fig. 6.3, which shows a detector (in the middle) which receives crosstalk, and the detectors for the four nearest neighbor channels. Each of the four nearest neighbors is at the same distance U from the central detector, so each contributes an equal amount of crosstalk. We define a ***neighborhood*** as a set of neighbors equidistant from a given channel. So, Fig. 6.3 shows the first nearest neighborhood.

The four nearest neighbors act independently of each other, which implies that the probability $P(k)$ that $k \in 0,1,2,3,4$ out of the four nearest neighbors contribute crosstalk to the central channel is a binomial distribution given by [29]

$$P(k) = \binom{4}{k} p^k q^{4-k}. \tag{6.19}$$

The quantities p and q are complimentary, where p is the probability that a neighbor is *on* (so that it is contributing to the crosstalk), and q is the probability that the neighbor is *off*. We maintain the assumption made in Section 6.2.1, Eq. (6.2) that the signals on neighboring channels are pseudo-random binary sequences, where logic *high* and *low* occur with equal probability. This means $p = q = 1/2$, so that the probability of $k \in\ 0,1,2,3,4$ members of the nearest neighborhood being *on* is

$$P(k) = \begin{pmatrix} 4 \\ k \end{pmatrix} (\frac{1}{2})^4. \tag{6.20}$$

Because all members of any one neighborhood are equidistant from the channel receiving the crosstalk, each member contributes an equal amount to the total crosstalk. For the nearest neighborhood shown in Fig. 6.3, we calculate the contribution χ_1 of a single channel (the subscript 1 indicating the first neighborhood). Assuming the nearest neighbor separation $s_1 = U = 5.514\ \lambda F$ calculated in Section 4.2.7 to ensure a maximum optical crosstalk of χ_{max} of -20 dB, we find that a single nearest neighbor contributes $\chi_1 = -29.1$ dB of optical crosstalk to the central channel.

The net crosstalk received from all first nearest neighbors is the amount χ_1 received from one nearest neighbor multiplied by the number of nearest neighbors that are *on* at a given time. The net crosstalk from the first neighborhood is a set of five discrete values $k_1\chi_1$, where $k_1 \in \{0,1,2,3,4\}$. The probability that net crosstalk is any one of these values is simply the probability of k_1 occurring, so that

$$P(\chi = k_1\chi_1) = P(k_1) = \begin{pmatrix} 4 \\ k_1 \end{pmatrix} (\frac{1}{2})^4. \tag{6.21}$$

The result is plotted in Fig. 6.4. Notice that the probability distribution is discrete, meaning crosstalk does not occur except for the five values specified by Eq. (6.21).

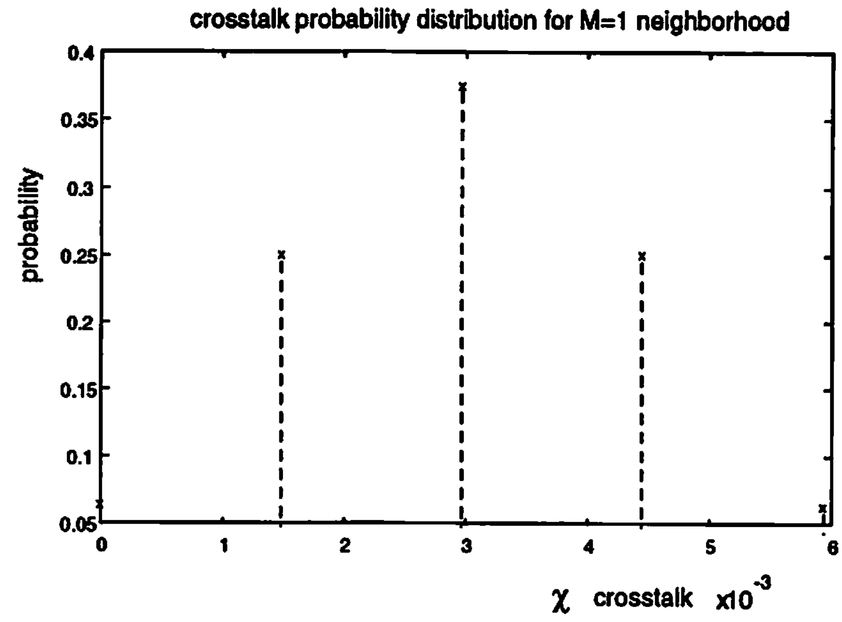


Figure 6.4: Simple binomial distribution for the crosstalk probability for only the first set of nearest neighbors. The five points on the plot correspond to the five possible discrete levels of crosstalk when the number k of neighboring channels contributing is $k \in \{0, 1, 2, 3, 4\}$.

6.4.2 Multiple Neighborhoods

To generalize Eq. (6.21) for an array of channels, we must incorporate the contributions of many neighborhoods of channels. The various neighborhoods differ from each other by their distance from the central channel. If we indicate the neighborhood with the index i, call the total number of channels in the ith neighborhood is n_i and the number contributing to the crosstalk is k_i, then the probability that k_i channels in the ith neighborhood contribute to the crosstalk is the binomial distribution given by

$$P\{k_i\} = \begin{pmatrix} n_i \\ k_i \end{pmatrix} (\frac{1}{2})^{n_i}. \tag{6.22}$$

The actual physical amount of crosstalk χ_i from one channel in the ith neighborhood is the product of k_i and χ_i.

To obtain the net crosstalk from multiple neighborhoods, we sum the contributions $k_i\chi_i$ of each neighborhood

$$\chi_{net} = \sum_{i=1}^{M} k_i\chi_i. \tag{6.23}$$

where M is the total number of neighborhoods. The probability that χ_{net} occurs is the same as the probability that all k_i required to produce χ_{net} (according to Eq. (6.23)) occur simultaneously. The behavior of each neighborhood is independent of the others, so the probability of the k_i occurring simultaneously is the product of their individual probabilities

$$P\{\chi_{net}\} = P\{k_1, k_2, ..., k_M\} = \prod_{i=1}^{M} P\{k_i\}. \tag{6.24}$$

6.4.3 Single-Channel Contributions to Optical Crosstalk

In order to calculate the probability distribution function of optical crosstalk in an array of channels, we must obtain the values χ_i for crosstalk contributions from a single channel in neighborhood i. These are easily calculable using Eq. (4.45)

shell	i	$\frac{s_i}{U}$	n_i	χ_i	ℓ
1	1	1	4	−29.1 dB	5
	2	$\sqrt{2}$	4	−33.3 dB	25
2	3	2	4	−38.5 dB	125
	4	$\sqrt{5}$	8	−40.4 dB	1125
	5	$2\sqrt{2}$	4	−42.6 dB	5625
3	6	3	4	−43.5 dB	28125
	7	$\sqrt{10}$	8	−44.6 dB	253125
	8	$\sqrt{13}$	8	−45.7 dB	2278125
	9	$3\sqrt{2}$	4	−48.6 dB	11390625

Table 6.1: The various neighborhoods i of neighbors, normalized distances s_i/U to the central channel, the number of channels n_i in the neighborhood, the crosstalk contribution χ_i of a single channel from neighborhood i, and the total number ℓ of possible discrete values of crosstalk neighborhoods 1 through i.

and the distance s_i of a channel in the ith neighborhood to the central channel. Figure 4.9 shows the channels (numbered by neighborhood) included in that our previous calculations of worst-case crosstalk. Table 6.1 lists the five neighborhoods i within the first two *shells*, which are the contiguous rings of neighbors around the central channel. The table shows the distances s_i of each neighborhood to the central channel (normalized to the distance $U = 5.514\lambda F$ between nearest neighbors), and the number of channels n_i in each neighborhood.

6.4.4 Crosstalk Probability Distribution

Using Eqs. (6.19), (6.23), and (6.24) with the information from Table 6.1, we extend the results of Fig. 6.4 to include more than the nearest neighborhood (i.e. for $M > 1$). When the first and second neighborhoods are included ($i \in \{1, 2\}$, $M = 2$, and $n_2 = 4$), the number of discrete values of crosstalk increases to $\ell = 25$, and the resultant distribution is shown in Fig. 6.5. As successive neighborhoods are added (up to the fourth), the number of possible discrete values of crosstalk

ℓ increases, while the range (from minimum to maximum net crosstalk) does not change appreciably. This makes crosstalk a pseudo-continuous random variable. The corresponding probability distributions are plotted in Figs. 6.6 and 6.7.

We notice that, despite crosstalk being pseudo-continuous, its probability distribution is highly discontinuous, with an envelope which only loosely resembles a Gaussian curve. The distribution is identically zero (as expected) beyond the worst-case value of crosstalk. Notice that this maximum value is 0.01 when the first five neighborhoods are included, which is a consequence of our design. In Chapter 4, we calculated the worst-case optical crosstalk with five neighborhoods (the first two shells) for various spacings between the detectors versus interchannel spacing. We decided that −20 dB was acceptable, and found that a channel spacing of $U = 5.514\ \lambda F$ guaranteed this amount. We also notice that the maximum crosstalk (0.01H) is less than the decision threshold (0.5H), we conclude that in the absence of Gaussian noise or other corruption, the crosstalk will induce no errors in digital signal regeneration.

6.5 Incorporation of Crosstalk into the BER Model

In this section, we show how crosstalk affects the bit error rate when Gaussian noise is also present, as is the case in most real systems. If the crosstalk exceeds the amount of Gaussian noise, we expect that incorporating crosstalk as additive Gaussian noise in the BER calculations is inaccurate. When crosstalk is small compared to the Gaussian noise, then we expect that modeling crosstalk as a slight increase in the Gaussian noise is a reasonable approximation. We examine these hypotheses after developing a quantitative description of error rates versus crosstalk. At first, we calculate the effect of crosstalk on the bit error rate for the simple case of a single constant value of crosstalk. Then, we apply this result

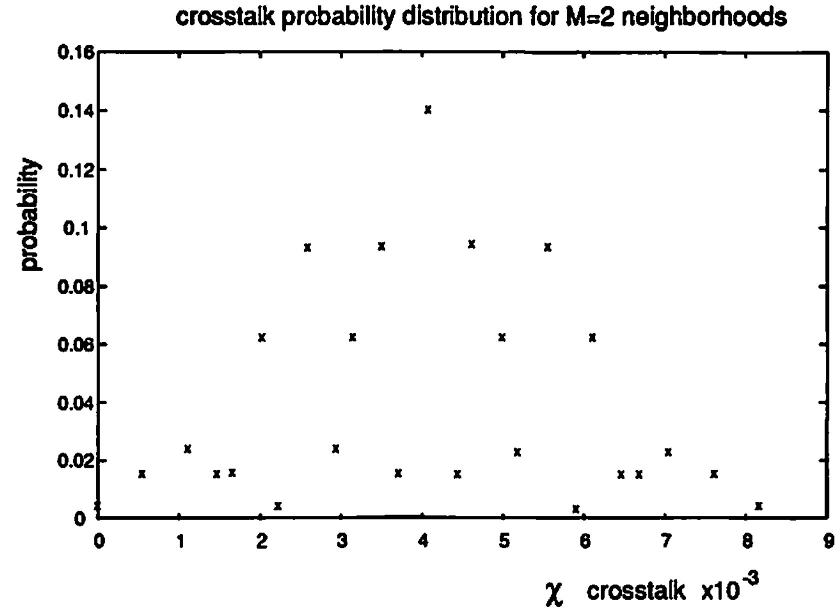


Figure 6.5: Crosstalk probability distribution for the first and second nearest neighborhoods ($M = 2$); the number of possible discrete values crosstalk may assume is $\ell = 25$.

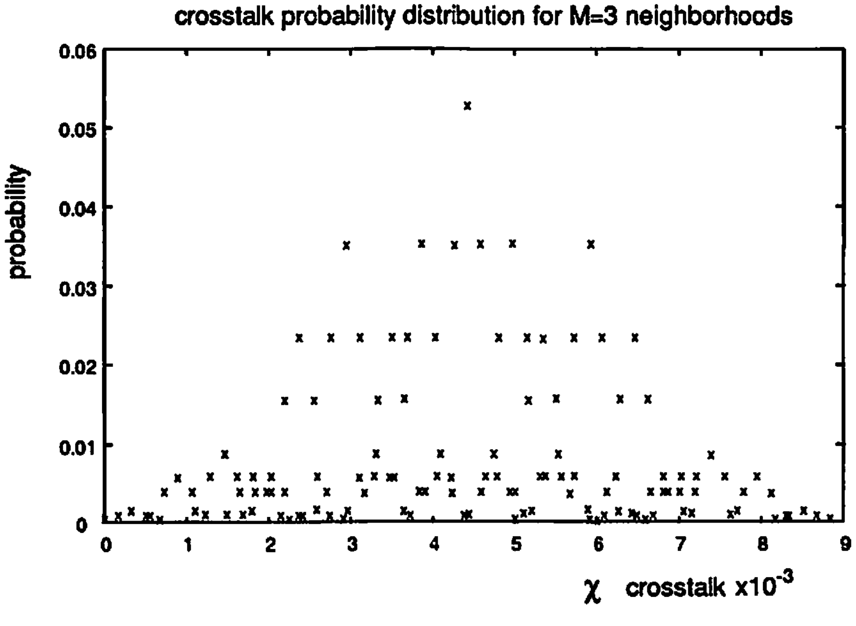


Figure 6.6: Crosstalk probability distribution for three nearest neighborhoods, $M = 3$, $\ell = 125$.

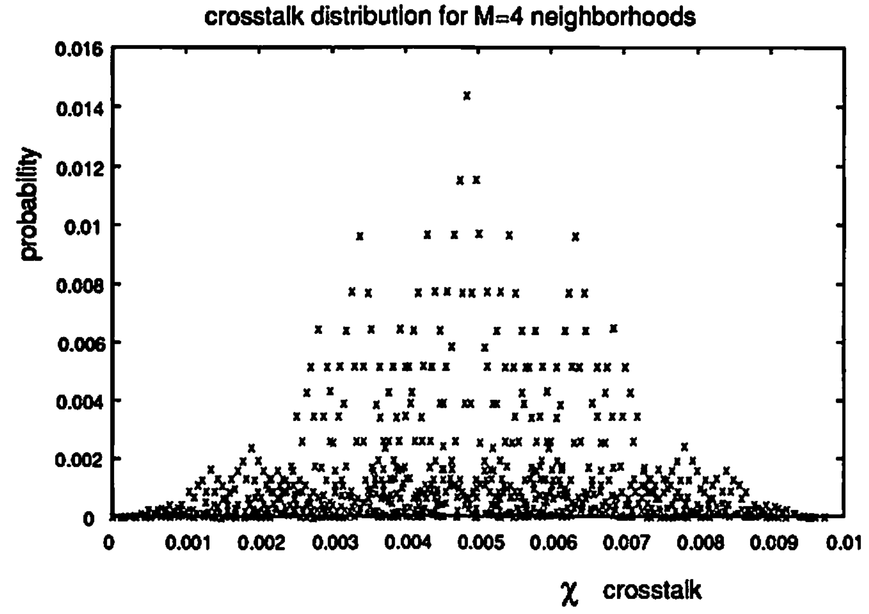


Figure 6.7: Crosstalk probability distribution for four nearest neighborhoods, $M = 4$, $\ell = 1125$.

to the many allowed values of crosstalk to obtain a net probability of error for crosstalk plus Gaussian noise.

6.5.1 Constant Single-value Crosstalk

Before we attempt to develop a detailed description of how crosstalk is incorporated into the model for BER, we consider a simple case, in which the crosstalk is a constant value χ_0. As such, the crosstalk is a (weak) deterministic signal which shifts the Gaussian noise distribution by χ_0, as shown in Fig. 6.8. The error rate in this case is found by a like shift of the lower limit of integration in Eq. (6.7) to obtain

$$P(E) = \int_{\frac{H}{2}-\chi_0}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} dx, \qquad (6.25)$$

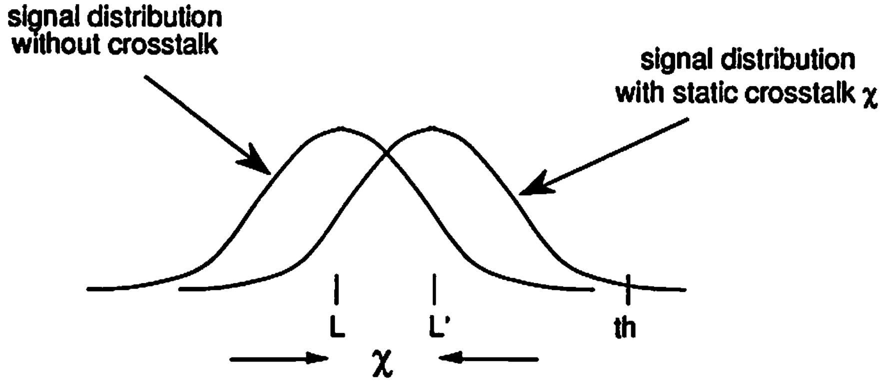


Figure 6.8: Static crosstalk has the effect of shifting the whole Gaussian signal distribution toward the decision threshold, increasing the probability of errors in signal regeneration.

Using the fact that $t = x/\sigma$, we get

$$P(E) = \frac{1}{\sqrt{2\pi}} \int_{t=\frac{H/2-\chi_0}{\sigma}}^{\infty} e^{-t^2/2} dt. \tag{6.26}$$

The lower limit of the integral is the difference of half the SNR and the crosstalk-to-noise ratio

$$P(E) = \frac{1}{\sqrt{2\pi}} \int_{t=Q-\frac{\chi_0}{\sigma}}^{\infty} e^{-t^2/2} dt. \tag{6.27}$$

Now, we use the approximation in Eq. (6.18) to get the probability of error given crosstalk χ_0

$$P(E|\chi_0) = \frac{1}{(Q - \frac{\chi_0}{\sigma})\sqrt{2\pi}} e^{-\frac{1}{2}(Q-\frac{\chi_0}{\sigma})^2}. \tag{6.28}$$

This result shows that the effect of crosstalk on the BER is dependent on the amount of Gaussian noise already in the system. Figure 6.9 shows how the BER varies with signal-to-noise ratio for zero crosstalk (bottom curve) and for crosstalk equal to one tenth the *high* signal level $\chi = 0.1H$ (top curve). There are two frequently used goals for bit error rates, 10^{-9} and 10^{-12}. These require signal-to-noise ratios of 12 and 14, respectively, assuming no crosstalk. In Fig. 6.10, we plot how static crosstalk χ_0 (normalized to the *high* signal level) degrades the

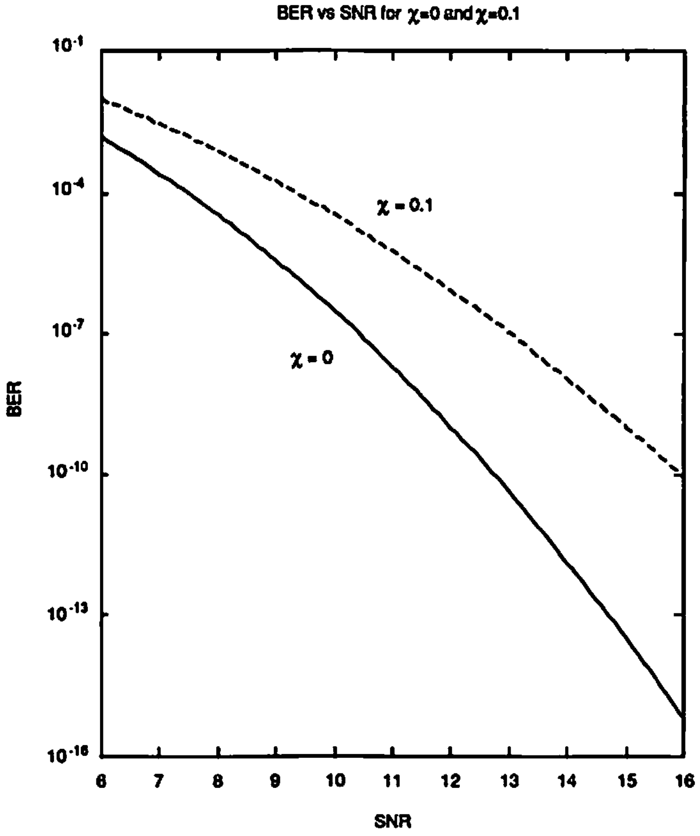


Figure 6.9: BER vs SNR for no crosstalk (lower curve) and for crosstalk one tenth the *high* signal level (upper curve).

BER for both of these SNR. The top curve is for an SNR of 12, and the bottom is for an SNR of 14.

6.5.2 Array of Channels with Random Crosstalk

For a system with multiple channels, each with pseudo-random digital signals, the crosstalk χ_0 varies in time depending on instantaneous signals on other channels. So, the χ_0 in Eq. (6.28) is not the only possible value of crosstalk, and so has a probability $P(\chi_0) < 1$. This means we must find the other possible χ and their

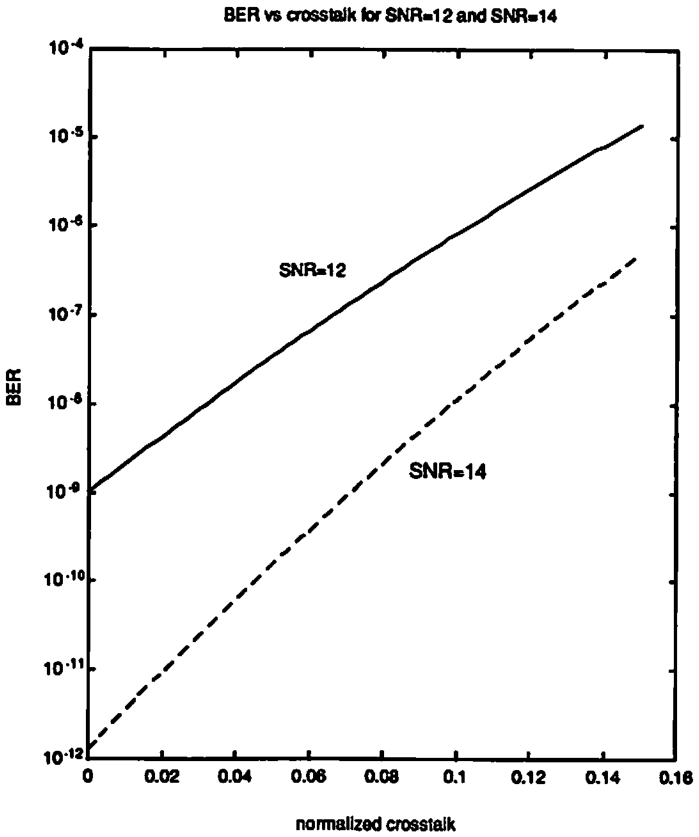


Figure 6.10: BER vs static crosstalk for two common signal-to-noise ratios 12 (upper curve) and 14 (lower curve).

probability of occurrence. Then, the net probability of error is

$$P_{net}(E) = \sum_{j=1}^{\ell} P(E|\chi_j) \cdot P(\chi_j) \tag{6.29}$$

where $P(E|\chi_j)$ is the probability of error given the occurrence of crosstalk value χ_j, and ℓ is the total number of crosstalk levels possible (shown in Table 6.1) given by

$$\ell = \prod_{i=1}^{M}(k_i + 1). \tag{6.30}$$

The probability $P(E|\chi)$ is obtained through Eq. (6.28). The jth possible value of crosstalk (of the ℓ possible values) has a probability of occurring given by Eq. (6.24), which is simply the probability that all the k_i within each of the neighborhoods are correct to produce net crosstalk χ_j. Since the various channels contribute crosstalk independently of the others, Eq. (6.24) says that the net probability of all the k_is occurring simultaneously to produce χ_j is the product of the individual probabilities of each k_i.

6.5.3 Results

Using Eq. (6.29), we calculate the effect of crosstalk on BER assuming the SNR associated with error rates of 10^{-9} and 10^{-12}. The SNR required to achieve these BER are 12 : 1 and 14 : 1, respectively, provided no crosstalk is involved. Table 6.2 shows how the BER rate behaves for an SNR of 12 as the crosstalk contributions of more of the neighborhoods are included in the calculation. The data shows that the most significant contribution to change in the BER occurs with the inclusion of the first two or three nearest neighborhoods. We therefore expect that including more than the first five neighborhoods makes a negligible difference in the BER. This was important in practice because the computation time is unreasonable for $M > 5$ due to the large number ℓ of possible crosstalk (see Table 6.1).

shell	M	$\bar{\chi}$	$\sqrt{(\bar{\chi^2})}$	$P_{Gauss}(E\|\bar{\chi})$	$P_{Gauss}(E\|\sqrt{(\bar{\chi^2})})$	BER
1	1	0.0030	0.0036	1.05×10^{-9}	1.06×10^{-9}	1.28×10^{-9}
	2	0.0041	0.0047	1.07×10^{-9}	1.08×10^{-9}	1.39×10^{-9}
2	3	0.0044	0.0050	1.08×10^{-9}	1.09×10^{-9}	1.42×10^{-9}
	4	0.0049	0.0054	1.09×10^{-9}	1.10×10^{-9}	1.47×10^{-9}
	5	0.0050	0.0055	1.09×10^{-9}	1.11×10^{-9}	1.49×10^{-9}

Table 6.2: $SNR - 12$; Calculation of the mean and rms values of optical crosstalk once the contributions of the first M neighborhoods are included; the corresponding Gaussian approximations $P_{Gauss}(E)$ for the error rate, using the mean and rms values of crosstalk; and the rigorously calculated error rate.

shell	M	$\bar{\chi}$	$\sqrt{(\bar{\chi^2})}$	$P_{Gauss}(E\|\bar{\chi})$	$P_{Gauss}(E\|\sqrt{(\bar{\chi^2})})$	BER
1	1	0.0030	0.0036	1.41×10^{-12}	1.44×10^{-12}	1.83×10^{-12}
	2	0.0041	0.0047	1.46×10^{-12}	1.50×10^{-12}	2.05×10^{-12}
2	3	0.0044	0.0050	1.49×10^{-12}	1.52×10^{-12}	2.12×10^{-12}
	4	0.0049	0.0054	1.52×10^{-12}	1.55×10^{-12}	2.22×10^{-12}
	5	0.0050	0.0055	1.52×10^{-12}	1.56×10^{-12}	2.24×10^{-12}

Table 6.3: $SNR - 14$; Calculation of the mean and rms values of optical crosstalk once the contributions of the first M neighborhoods are included; the corresponding Gaussian approximations $P_{Gauss}(E)$ for the error rate, using the mean and rms values of crosstalk; and the rigorously calculated error rate.

The two columns containing $P_{Gauss}(E)$ in Table 6.2 represent approximations using Eq. (6.18) of the probability of error. These approximations assume crosstalk is an additive Gaussian noise process. In order to use Eq. (6.18), however, one must decide what value of crosstalk (amongst the ℓ possible values) is most suitable for use as σ in Eq. (6.18). The two likely candidates are the expectation (mean) value $\bar{\chi}$ and the rms value $\sqrt{(\bar{\chi^2})}$. The third and fourth columns show the approximate BER using $\bar{\chi}$ and $\sqrt{(\bar{\chi^2})}$, respectively in Eq. (6.18).

In Table 6.3, we plot the same results for $SNR = 14$ (which yields a BER of 10^{-12} in the absence of any crosstalk) for the inclusion of $M = 1$ to $M = 5$ neighborhoods. As before, little change occurs after the first few neighborhoods are included. Notice that for the lower SNR of 12 (i.e. more Gaussian noise), the relative difference between the approximation of the BER and the rigorous result is about 26 %. For the higher SNR of 14, the difference is slightly more at about 31 %.

6.6 Conclusion

In general, the approximations match the rigorous result better for larger noise and smaller crosstalk, as indicated in Table 6.4. The first row of Table 6.4 shows the result of lowering the SNR to 10 (i.e. increasing noise) and lowering crosstalk by one half, ($\Delta\chi$ is the multiplicative change in crosstalk). The reduction of crosstalk by one half is obtained by increasing the nearest neighbor separation so that $U > 5.514\lambda F$. For this case, the difference between the approximate and rigorous results drops to 11%. The second row of Table 6.4 shows that the approximations are worse for higher crosstalk and lower Gaussian noise. Crosstalk is multiplied by a factor $\Delta\chi$ of 5, and the noise is reduced so that $SNR = 16$. In this case, the relative errors in the approximations are 138% and 413% (using $\bar{\chi}$ and $\sqrt{\bar{\chi^2}}$, respectively). Even these larger differences, however, are insignificant when one is interested in a n order of magnitude estimate of the error rate, as Fig. 6.11 shows.

SNR	$\Delta\chi$	$\bar{\chi}$	$\sqrt{(\bar{\chi^2})}$	$P_{Gauss}(E\vert\bar{\chi})$	$P_{Gauss}(E\vert\sqrt{(\bar{\chi^2})})$	BER
10	0.5	0.0025	0.0027	2.94×10^{-7}	2.95×10^{-7}	3.33×10^{-7}
16	5	0.025	0.027	5.70×10^{-14}	1.23×10^{-13}	2.40×10^{-14}

Table 6.4: Calculation of the mean and rms values of optical crosstalk when all $M = 5$ neighborhoods are included for the extreme cases of high Gaussian noise (low SNR) with low crosstalk, and the opposite case when noise is low (high SNR) with high crosstalk.

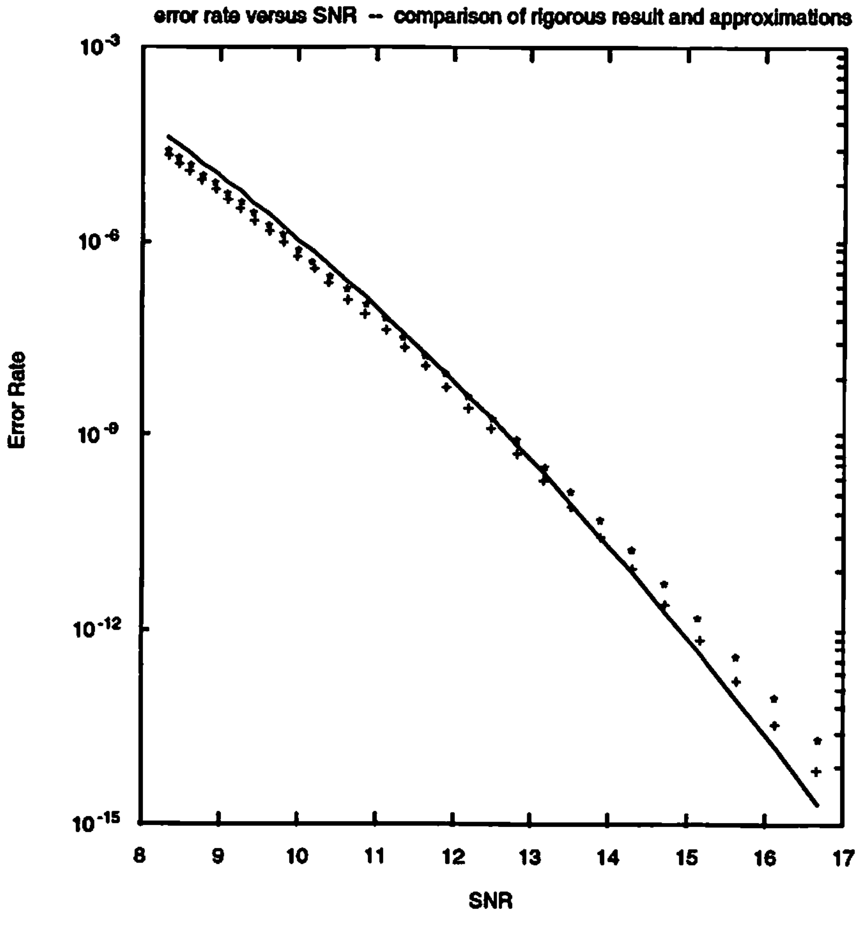


Figure 6.11: Logarithmic plot of the Gaussian approximations (* and +) for the BER and the rigorously calculated error rate (solid curve) versus SNR.

Chapter 7

Multiplexing Many Shuffles through One Set of Optics

7.1 Motivation

The shuffle-exchange network is composed of several identical stages, each with fixed optical perfect shuffles and an array of opto-electronic bypass-exchange switches. The hardware cost and complexity are directly proportional to the number of the stages m. As discussed earlier, to achieve a full access network for N input and output channels, the minimum number of stages required is

$$m = \log_2 N. \tag{7.1}$$

So, if one wishes to connect a thousand inputs to the same number of outputs, at least ten stages are necessary.

Haney has suggested that replicating the same set of optics many times in order to form a multistage network may be unnecessary [12]. At the outset, we hope that if multiplexing is possible, that it improves the interconnection network in some way. The improvement may be either enhanced performance, reduced hardware complexity, or reduced size. In this chapter, we examine the possibility of multiplexing many stages through one set of shuffle optics, the means (both general and specific) of doing it, and the consequences in terms of system speed, hardware complexity, and physical size.

7.2 General Types of Multiplexing

7.2.1 Time Multiplexing

In the original concept of a shuffle-exchange network, data injected into the first stage is shuffled, passes through bypass-exchange switches, and is output to the second stage. In the linear network (one with no multiplexing of stages), data from the first output plane passes directly to another array of sources for injection into the next optical shuffle. The hardware of the first stage is replicated to form the second stage. This is an important feature of the shuffle-exchange network; the fixed connection pattern (the shuffle) is the same in every stage.

Because the shuffle connection pattern does not change from stage to stage, one may re-use the shuffle by means of feedback. This is illustrated schematically in Figure 7.1. An array of sources in the input plane emits data, which passes through the optical shuffle to the bypass-exchange switches at the output plane. With *time multiplexing* the output array from the first stage is fed back and imaged onto the back side of the original input plane. If the array of bypass-exchange switches is reset for second stage operation, the data may be injected into the same optical shuffle to pass through the second stage of the network.

Resetting the array of switches for each subsequent stage requires an amount of time during which data on the array of channels must be held. So, there must be a buffer for each channel capable of storing the data temporarily at either the input or output plane. After resetting the switches, the data is released from the buffers into the shuffle. The process is repeated until the data has traversed all m stages, at which point it emerges from the network.

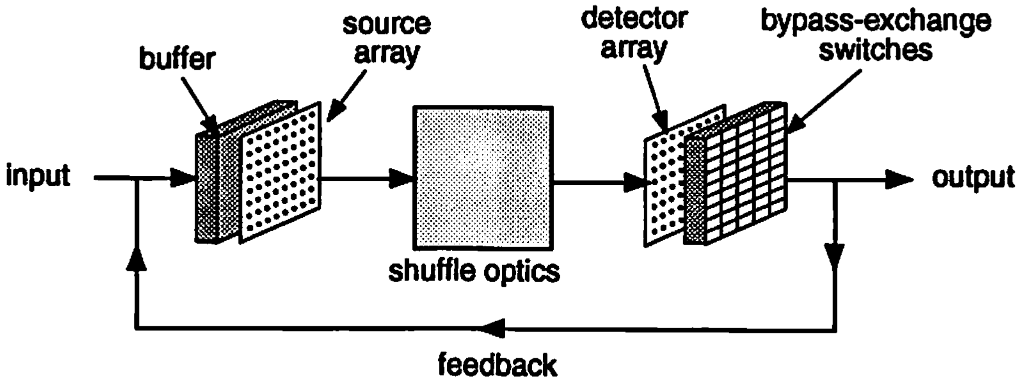


Figure 7.1: Simple feedback to the existing hardware is time multiplexing.

A new packet of data may enter the network only after the previous packet has emerged from the final stage. As such, the network is inherently slower than a linear multistage network in which the optics and the arrays of sources, detectors and switches are all replicated m times. In building a time-multiplexed network, we trade time complexity for hardware complexity.

7.2.2 Space Multiplexing

An alternative to time multiplexing is *space multiplexing* which is shown schematically in Fig. 7.2. There are arrays of sources, detectors and bypass-exchange switches dedicated to each individual stage of the network, so the hardware is essentially the same as that of the linear (non-multiplexed) network, while it is more complicated than the time multiplexed network. The additional hardware eliminates the time delay associated with time multiplexing. A new array of data with a different output permutation may be injected into the first stage as soon as the original data has passed to the second stage. Since there is only a single optical shuffle, however, the optical hardware required to build the network is reduced roughly by a factor of m from the nonmultiplexed network.

Multiplexing many shuffles through a single set of optics requires a feedback system. Because it is a global, fixed connection, we assume that the feedback

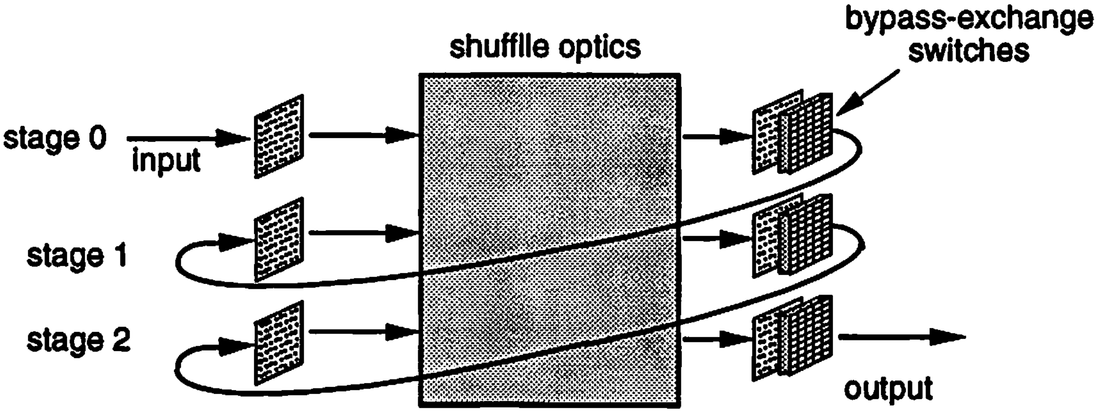


Figure 7.2: Eliminating the time delay in time multiplexing can be solved by having arrays of sources, detectors and switches dedicated to each stage, while still re-using the same set of optics.

is accomplished optically. This means that multiplexing does not reduce the optical hardware by a full factor of m. Furthermore, multiplexing places greater demands on the resolution capability of the optics; a higher space bandwidth product (SBWP) is needed than when a separate set of optics executes the shuffle at each stage.

Thus, we confront the important question of how to define complexity in an optical interconnection network. Is it, for example, more complex to construct and align two separate imaging systems, each containing a lens with a SBWP of X, or a single system containing one lens with SBWP of $2X$? Certainly building the two lower-SBWP systems requires more pieces of hardware. For this reason, it may be considered the more complex option. The system with a single lens of higher SBWP, however, may require greater effort to fabricate and align, and greater mechanical stability to maintain its alignment. As such, it may be the more complex option.

The issue of complexity is simply not as well defined for an optical system as it is for electronics. The complexity of an electronic system is largely determined by the number of active components (i.e. transistors), and the number of process steps to fabricate a given type of transistor. Adding transistors generally makes the system physically larger, more difficult to fabricate and test, and more expensive. The definition of complexity in an optical system is open to question. We make no attempt to define a cost function for optical complexity because it requires subjective judgment.

7.3 Means of Space Multiplexing

7.3.1 Shift of the Whole Input Array

Before proceeding any farther, we must determine whether our lens-based optical shuffle architecture is capable of operating like the system in Fig. 7.2. In other words, if we place more than one input array in the input plane of our shuffle architecture, does the output contain shuffled versions of each of the input arrays? Haney has suggested that since the lens-based architecture of Fig. 4.1 is shift-invariant, then space multiplexing is possible. By shift-invariant, we mean that the shuffle operation is preserved under a lateral shift of the whole input array. Figure 7.3 shows two source arrays, each shifted away from the central axis of the system, along with ray paths to the output plane. We can see that the output has two self-contained shuffled outputs, so the shuffle is preserved under a shift of the input. This means that separate arrays of sources for each stage may be shuffled by the same optics.

7.3.2 Shift-Variance in Spatial Multiplexing

The lens-based shuffle is a shift-invariant system for which space multiplexing is easily accomplished. For architectures which are not shift-invariant, we wish to determine whether the shuffle is preserved under a shift of the whole input array. Intuitively, a shift-invariant system implements a single connection pattern, or in other words, has only one impulse response. A shift-variant system, however, implements more than one connection pattern. We define two new terms [14] .

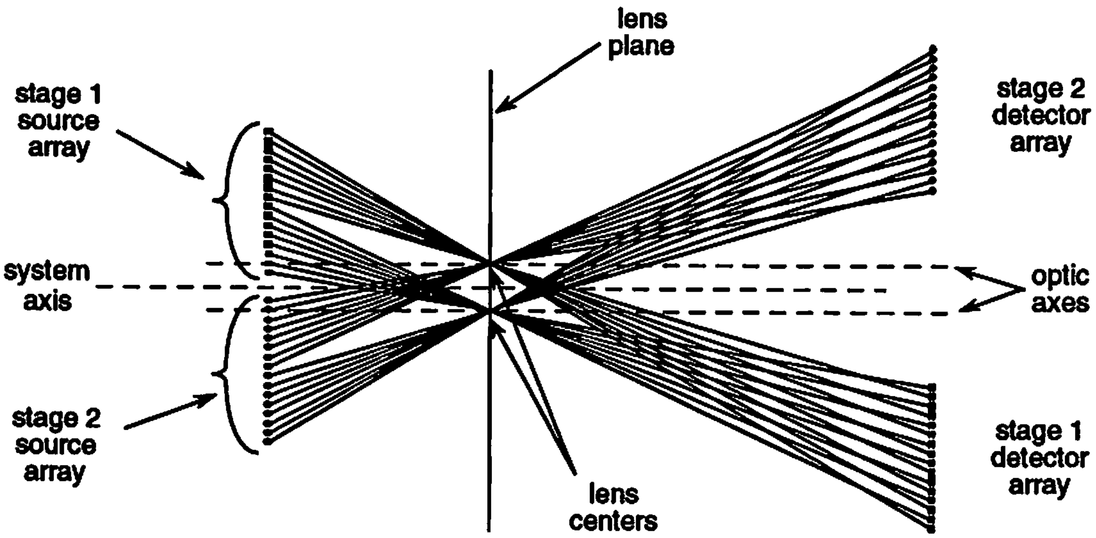


Figure 7.3: Moving an input array up or down from the system axis does not disrupt the shuffle operation of the system as shown by the set of rays through the appropriate lens centers.

A ***strictly shift-variant*** system implements a completely different connection for each of the system input points. For a system with N inputs, the strictly shift-variant system has N different impulse responses. A ***hybrid*** is the case between shift-invariant and shift-variant. It implements more than one and fewer than N connection patterns.

The optics for a strictly space-variant system are inherently more complex than the bulk optics approaches above. Possible implementations include a hologram with exactly one facet per connection [30], an inhomogeneously spaced array of N lenslets, or a prism array. The problem with any of these implementations is that the complexity of the system increases exactly with the number of inputs.

An example of a hybrid system is the optical shuffle implemented with a four-facet hologram [42]. Each quadrant of the input array is imaged onto a different facet (quadrant) of the hologram. The facets redirect the different input quadrants suitably to accomplish the shuffle in the output plane, as shown in Fig. 7.4. The system is hybrid because it employs four distinct impulse responses, which does

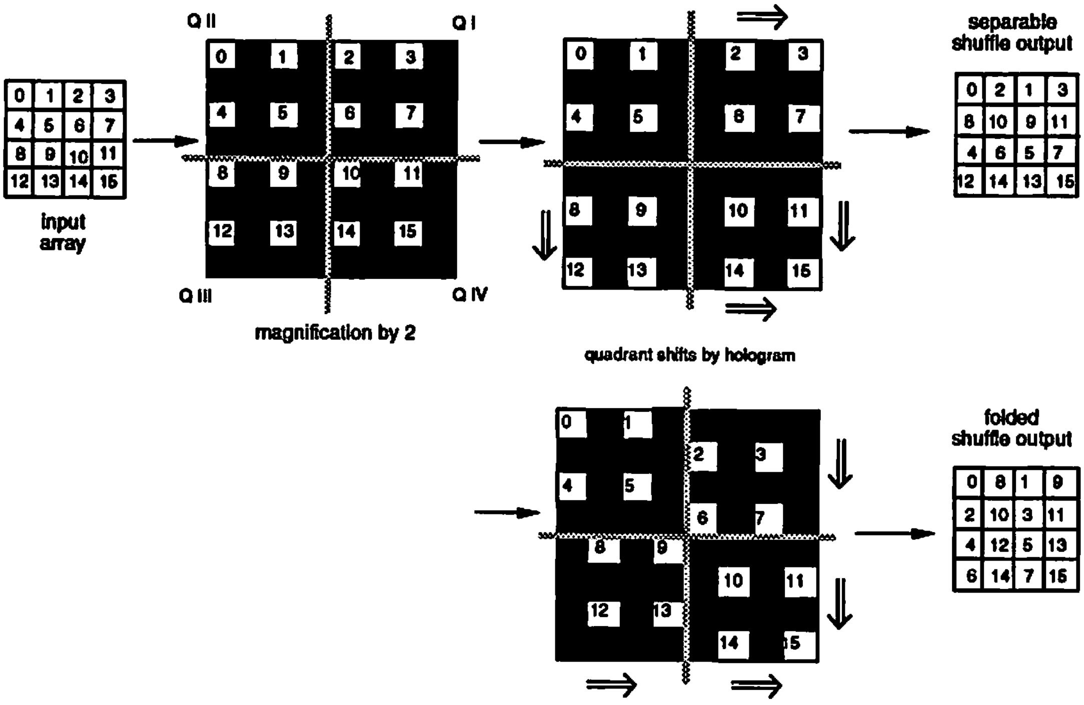


Figure 7.4: One-copy algorithm and 2-D shuffles (from L. Cheng).

not change with the number of channels in the input array. So, for more than four inputs, the four-facet holographic shuffle satisfies the definition of a hybrid system.

One may argue that the lens-based shuffle is similarly a hybrid system because it uses four lenses, each with its own impulse response. The difference arises in the type of source which each system uses. The lens-based shuffle is an imaging system which requires divergent sources (ideally, a point source). If the source array is shifted laterally in the input plane, the shuffled output is preserved (and shifted), as shown in Fig. 7.3. On the other hand, the holographic shuffle uses collimated inputs. So, if the source array is shifted enough for any of the inputs at the edge of a quadrant to miss its intended facet on the hologram, it will not be deflected properly to its shuffled position in the output. Either it misses the hologram altogether, or it hits a different facet, which then directs it to an incorrect position in the output plane. The holographic architecture, therefore, does not tolerate an arbitrary shift of the source array.

7.4 Space Multiplexing by Overlapping Input Arrays

7.4.1 General Description

While the arrangement in Fig. 7.3 demonstrates the principle that space multiplexing is possible with the lens-based shuffle, there are a pair of problems with the method in the figure. First is the potentially large degree of off-axis operation for a system with many stages. Operating far off-axis decreases the light collection efficiency of the optics, as we showed in Section 4.3.2, and degrades the image quality at the output plane. The second problem, however, is fundamentally more serious. Notice that each of the stages requires a separate feedback system. It is thus questionable whether the space multiplexed system is any less complex than the linear (non-multiplexed) network. To ensure that the space-multiplexed

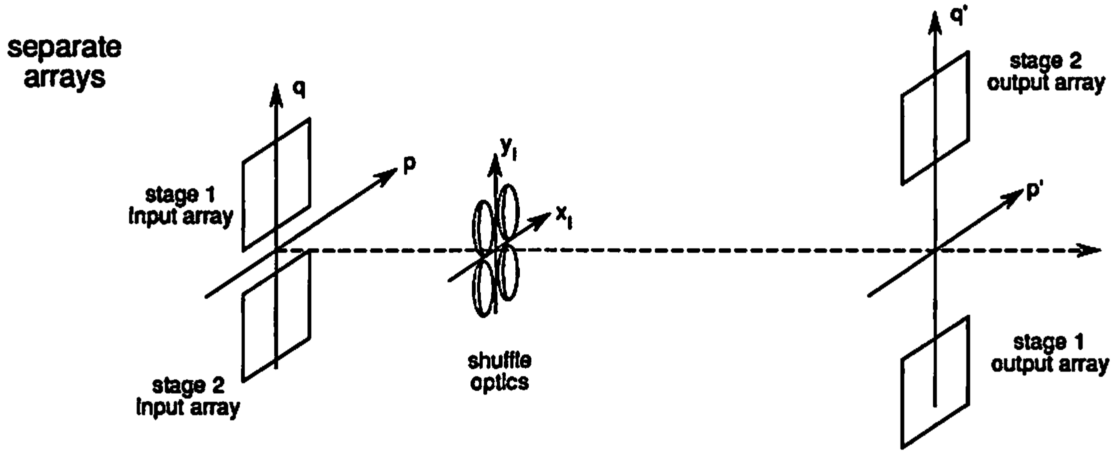


Figure 7.5: A two-stage network space-multiplexed through a single optical shuffle, where the input arrays are entirely separate in the input plane. As a result, the shuffled output for each stage is entirely separate, suggesting the need for a separate feedback system for each stage.

system has lower mechanical complexity than the linear network, we must be able to provide feedback for all stages with a single bulk optics system.

A means of solving both of these problems is to *overlap* input arrays, so that each is slightly shifted relative to the others. Consider the two-stage network shown in Fig. 7.5. The input arrays are entirely separate, and so too are the output arrays. We solve the problems of feedback complexity and extreme off-axis operation by overlapping the source arrays in the input plane, shifted (ϵ_x, ϵ_y) relative to each other, as shown in Fig. 7.6.

To analyze the space multiplexed system, we divide the input plane into non-overlapping squares, as shown in Fig. 7.7, which we refer to as input pixels. For a linear (non-multiplexed) network, each pixel contains a single channel at the pixel center. If the source array is moved laterally by ϵ, then each source is displaced from its pixel center by the same ϵ. If we define the dimensions of the pixel in 2-D as $\Delta \times \Delta$, then we may shift the source array by $\Delta/2$ in either x or y before the sources leave their original input pixels.

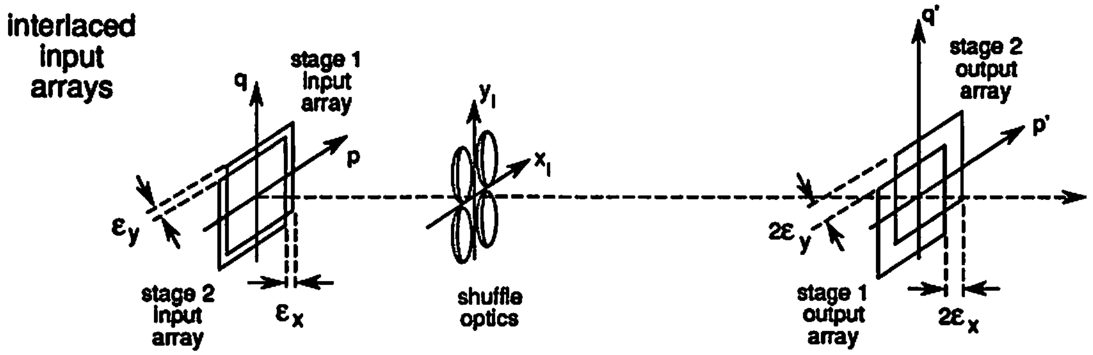


Figure 7.6: Overlapping the input arrays in the input plane, separated by ϵ. The resulting shuffled output for both stages is overlapped, and shifted by 2ϵ horizontally and vertically in the output plane due to the magnification of two in the lens-based shuffle.

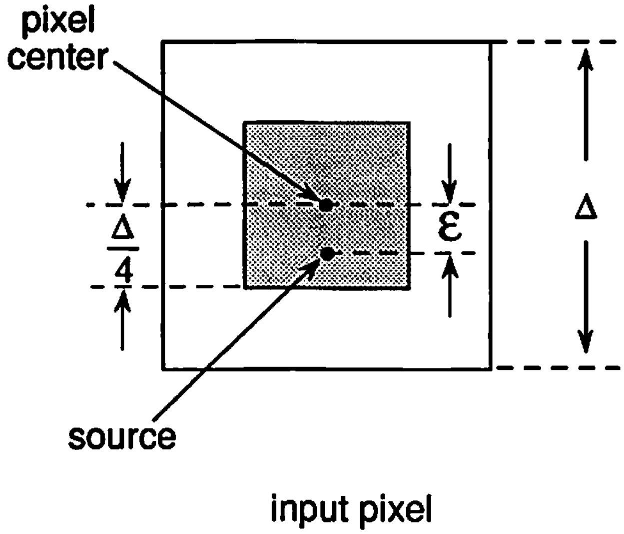


Figure 7.7: The square region surrounding a single channel is a pixel. Here is a pixel in the input plane; the source ordinarily occupies the pixel center in a non-multiplexed network. The diagram shows the source displaced from its pixel center, and the shaded region indicates where the displacement ϵ is small enough to keep its image in the output plane within the magnified boundaries of the input image.

Recall that the lens-based shuffle magnifies the input array by a factor of two. This implies that if a source is displaced by ϵ from its original position (i.e. the center of the pixel), then its image in the output plane is displaced by 2ϵ from the center of its shuffled output pixel. So, if sub-arrays of channels are to remain within the magnified boundaries of their respective pixels in the output plane, then the sources must be displaced by no more than

$$\epsilon_s < \frac{\Delta}{4}, \tag{7.2}$$

where the subscript s indicates a shift of the *source*. This ensures that an channel with shuffled output coordinate (i', j') does not stray into any of the adjacent output pixels at $(i' \pm 1, j' \pm 1)$. It also preserves the shuffle operation as a mapping from input *pixel* to output *pixel*, and leaves the subarray within a given pixel intact (although magnified by two and inverted).

Within a given pixel, the result is a *sub-array* of sources each having the same channel index, as shown in Fig. 7.8. The pixels shown are on opposite sides of the q axis in such a way that they will be horizontal neighbors in the shuffled output. The sub-arrays (within each pixel) in the output plane, which are shown in Fig. 7.9, are different from those of the input plane because of the magnification by two and inversion inherent in the lens-based shuffle. The pixels in the input plane in Fig. 7.8 have coordinates (p_0, q_0), and the corresponding output pixels are indicated in Fig. 7.9 as primed (p_0', q_0'). We note that the source configuration within the pixels (both input and output) are drawn to scale, but that the location of the pixels is not.

7.4.2 Complexity of Electronic Wiring in Space Multiplexing

We assume that, once optics executes the shuffle in a given stage, then electronics executes the bypass-exchange operation between adjacent pairs of channels at (i', j') and $(i', j' + 1)$. In the simple case of a linear network, where the optical

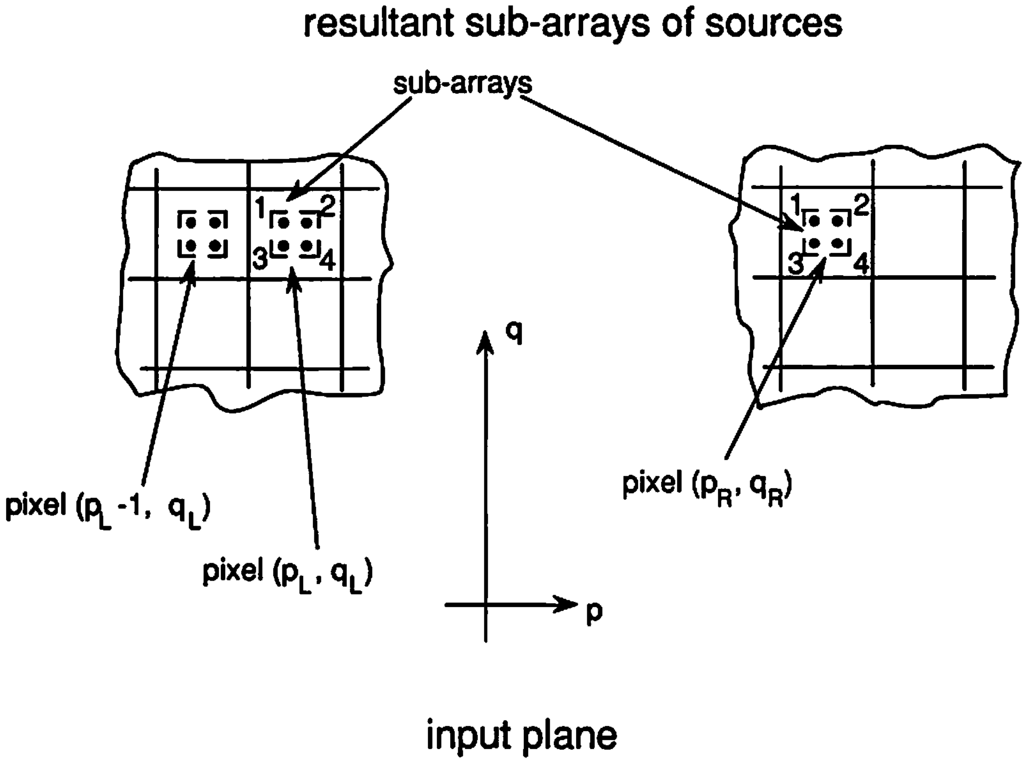


Figure 7.8: Spatially multiplexed source arrays produce sub-arrays of sources within each pixel. Sample pixels on either side of the input plane (p, q) are shown for a network with four space multiplexed stages.

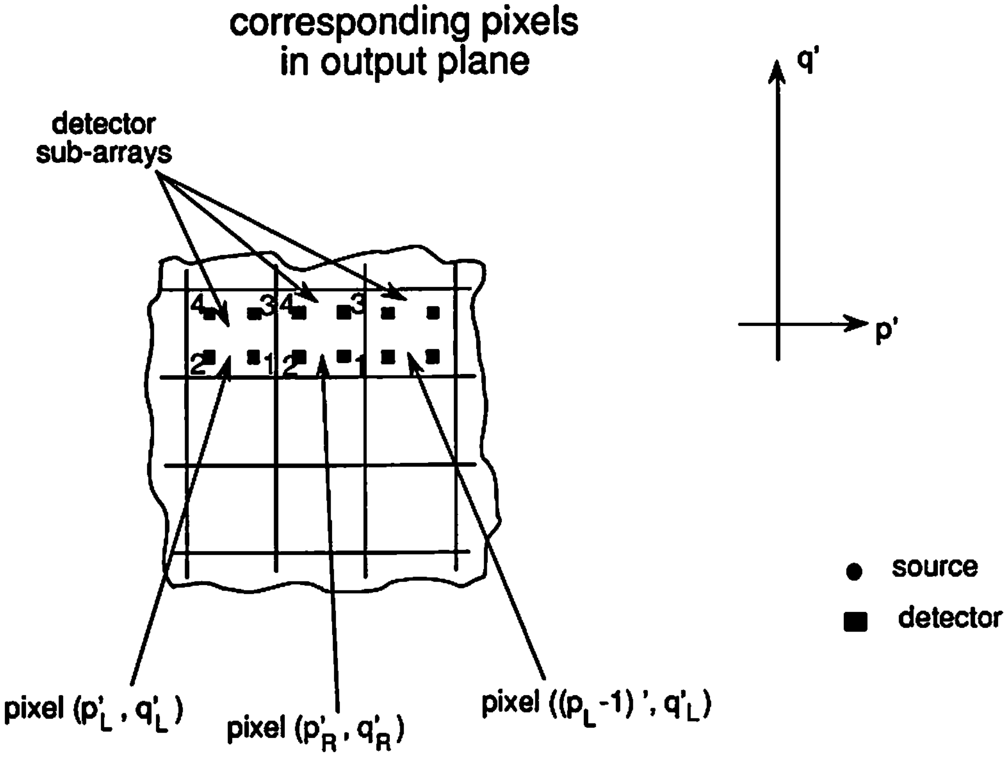


Figure 7.9: Sub-arrays of channels in pixels in the output plane corresponding to those of the input plane in the previous figure.

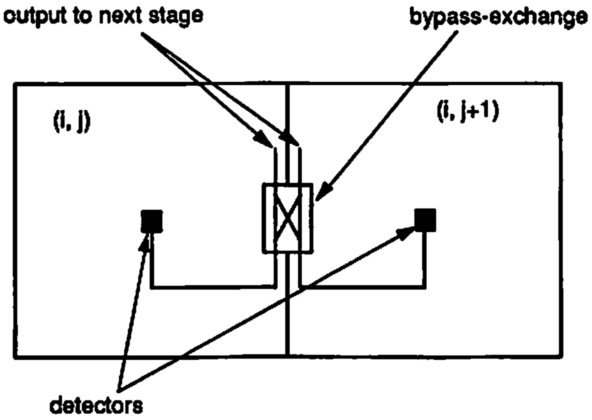


Figure 7.10: Adjacent channels in the detector plane of a linear pipelined system, and the wiring to the bypass-exchange switch.

shuffle is replicated in each stage, the electronic wiring in the output plane is minimal and quite simple, as shown in Fig. 7.10. When we multiplex several stages through a single optical shuffle, the wiring is more complex.

Haney [12] illustrates the situation where *four* stages are spatially multiplexed through the same optics. One of his diagrams is adapted here in Fig. 7.11, and it has been altered to account for the magnification by two and inversion of the output of the lens-based shuffle. The output from an adjacent pair of channels with like stage numbers must be switched. As a result, there are many places where wires cross or run side by side, situations which give rise to electrical crosstalk. It is precisely this type of electrical wiring problem which motivated attempts to build free-space optical interconnection systems in the first place. As a rule, it is best to keep electrical wiring to a minimum.

Govindarajan and Forrest [10] have described how electrical crosstalk limits the packing density of an optoelectronic interconnection network. Others [26], [17] have shown that the relative efficiencies of optics and electronics in communications depends heavily on the distance involved. Each of these analyses leads to the conclusion that simplifying electrical wiring is important. For the issue of

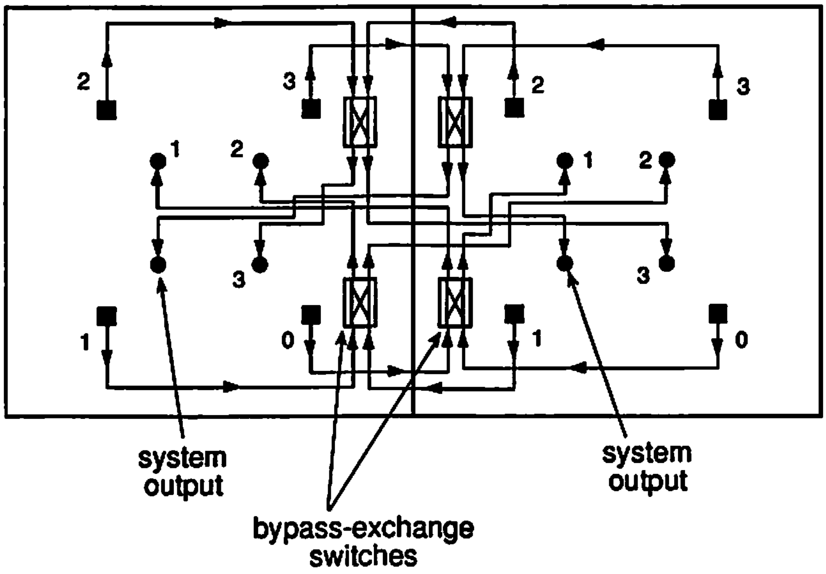


Figure 7.11: Adjacent pixels for a spatially multiplexed system with four stages, and the required wiring to bring the information from adjacent channels of like stages into bypass-exchange switches (after Haney [12]).

crosstalk, wires must cross or run side-by-side as little as possible. To be a more efficient means of communication than optics, electrical wiring must be kept short.

The wiring problem we see in Fig. 7.11 is directly proportional to the number of stages m in the network. Regardless of whether we choose to make the system full access or non-blocking, the number of stages in either case is proportional to $\log_2 N$. Thus, the wiring problem of Fig. 7.11 grows only as $\log_2 N$. While this represents an improvement over implementing the interconnection network as a crossbar, whose complexity grows as N^2, it still represents a problem.

Consider that the situation for adjacent output pixels depicted in Fig. 7.11 is relevant to a network with $m = 4$ stages. If this is a full access network, then the number of stages is $4 = \log_2 N$. Therefore, the number of channels which the network can interconnect is $N = 2^4 = 16$, which is relatively few. To be of much use, an interconnection network must be capable of routing on the order of $N = 1000$ channels. For the network to be full access, the number of stages must then be at least $m = \log_2 N = 10$. The wiring required for the bypass-exchange

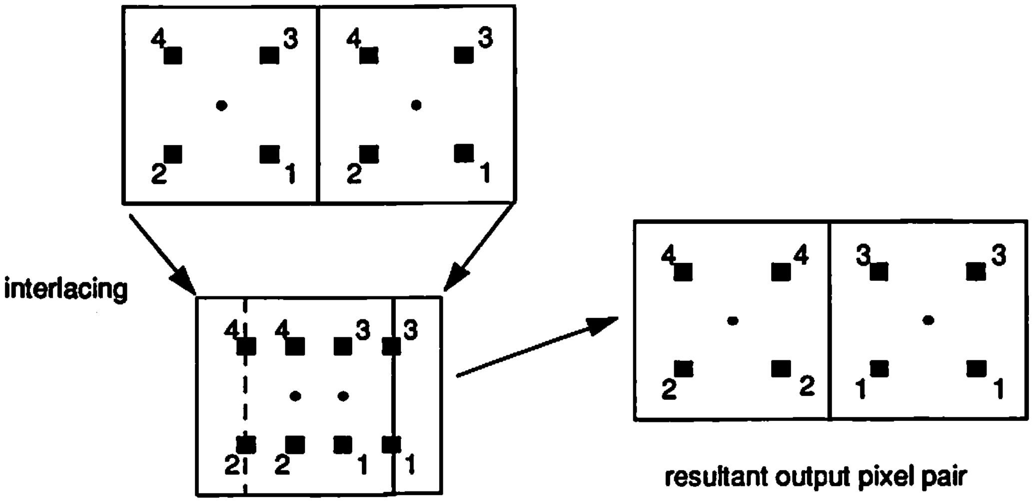


Figure 7.12: Simplifying the OEIC wiring in the shuffle output plane.

operation, depicted for a four-stage network in Fig. 7.11, is proportionately worse for the network with ten stages. So, the increase in wiring complexity is the price for limiting the optical hardware to one set of lenses.

7.4.3 Qualitative Solution

We now investigate the capability of the shuffle optics to circumvent the wiring problem described above. Consider the simple case in Fig. 7.10. The wires bring data from adjacent channels (i', j') and $(i', j' + 1)$ physically close together so that the data may be routed through the bypass-exchange switch. This case is simple because there is only one stage of the network associated with each output pixel. In a space multiplexed system, like in Fig. 7.11, the wiring brings adjacent channels from like stages together for injection into a bypass-exchange switch.

This is analogous to overlapping adjacent pixels so that channels from like stages are nearest neighbors, as shown in Fig. 7.12. Notice that the wiring in this situation is reduced to the simplicity of the non-multiplexed (linear network) case in Fig. 7.10. Even though there are more wires per pixel than in the linear arrangement, there are no overlapping electrical wires (outside the bypass-exchange switches). This is because channels from like stages in adjacent pixels are already nearest neighbors.

We recall that our optical system performs the shuffle by making four images (or copies) of the input plane and interleaving them appropriately. We also note that any two adjacent channels in the output plane of a linear network correspond to images from different lenses, from sources in opposite halves of the input plane. This means that the shuffle optics already interleaves output pixels as part of the shuffle operation.

We use this fact to alleviate the wiring problem in space multiplexing. Figure 7.13 illustrates the situation. In opposite halves of the input plane, we shift sources from their respective pixel centers toward the center of the input array. This causes their images to fall into neighboring pixels in the output plane. In this manner, adjacent pairs of channels from like stages are made nearest neighbors. This means that a channel from stage 2 does not come between two channels from stage 1. Thus, no wires from stage 1 channels must cross wires from stage 2.

7.4.4 General Sub-array Arrangement for m Stages

There are two problems with the method shown in Fig. 7.13. First, the output image points are not evenly spaced to make full use of the space bandwidth product of the optics. Second, it only shows how to accomplish space multiplexing for two stages. To illustrate the idea, consider a system where the number of channels is $N = 5000$. The number of stages required for a full access network is $m = \log_2 N$, which in this case is 13. A possible arrangement for the sub-array of sources in each pixel in the input plane is shown in Fig. 7.14. Also shown is the resultant image of the same pixel in the output plane. The sources and detectors for each of the thirteen stages are numbered in the diagram. Notice the fact that all *image* points are separated by the same distance δ in order to make the best use of available space bandwidth product (SBWP). Furthermore, the wiring from

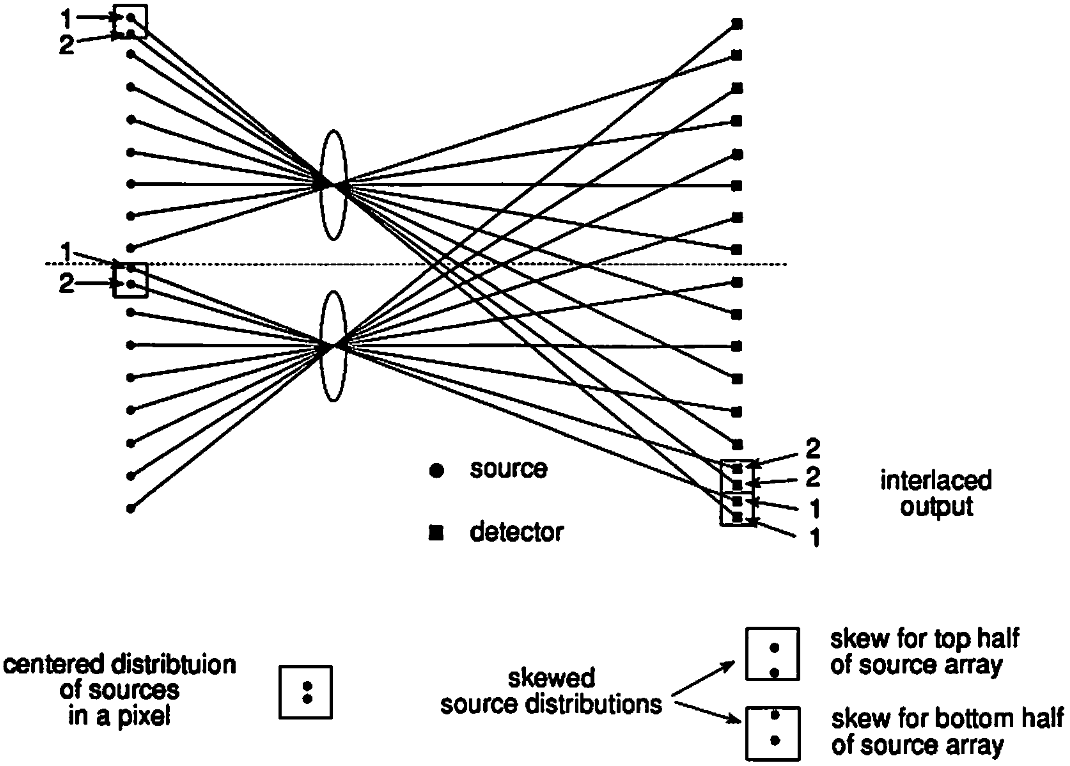


Figure 7.13: To multiplex two channels, we place two sources in each pixel. Instead of the centered distribution shown, however, we may skew the source arrangement to make like stages nearest neighbors in the output plane.

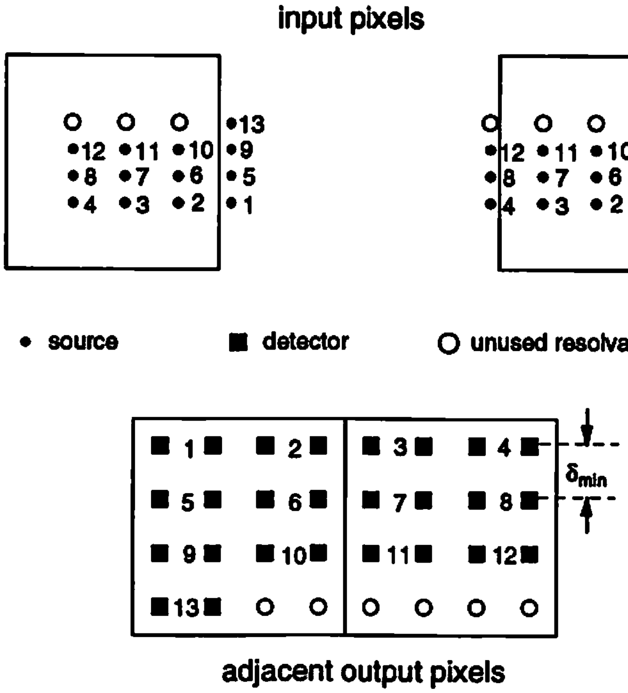


Figure 7.14: Source and detector subarray arrangements for 13 stages, which accomplishes proper interlacing for minimizing wiring.

the detectors to the bypass-exchange switches is minimal because detectors for like stages are nearest neighbors.

Pixel Size

In order to arrange a general number m of sources so that the output points are evenly spaced to make the most efficient use of the SBWP of the optics, consider the example of thirteen stages. In each output pixel, there must be a square subarray with at least as many resolvable spots as the number of stages m. Any additional resolvable spots are unused. So, the subarray must contain τ points such that τ is the first perfect square larger than m

$$\tau = \lceil \sqrt{m} \rceil^2, \tag{7.3}$$

where $\lceil \cdot \rceil$ indicates the nearest integer greater than or equal to the argument. The total space bandwidth product (SBWP) is the product of the total number of channels $N = n \times n$ and τ

$$SBWP = N\tau. \tag{7.4}$$

To determine the size Δ of the pixels, recall that the detectors must be spaced some minimum distance δ apart. The size of the pixel in 1-D is

$$\Delta = \delta\sqrt{\tau}. \tag{7.5}$$

From Chapter 4), we found that the minimum δ needed to keep the optical crosstalk to less than -20 dB is $\delta = 5.514\ \lambda F$.

Approach

We shall refer to the example of a thirteen-stage network to devise a general formula for the positions of sources in the input sub-arrays. Figure 7.15 shows sample pixels on opposite sides of the vertical (or y) axis, which are neighbors in the shuffled output. We see that sub-arrays on opposite sides of the input plane are mirror images of each other. So, we determine the source arrangement which produces correctly interlaced output for pixels on the right hand side of the input plane, and the left-hand side is solved by reflection symmetry.

Interlaced Output Sub-array

In order to specify the proper shift of a source to interlace the output, we first develop a general formula for the position of the interlaced output. We specify the shift of the output image relative to the center of its output pixel. For example, the column labelled $k = 1$ on the right hand side (RHS) of the input array represents the greatest shift in x of any source from the center of its pixel. Figure 7.16 shows

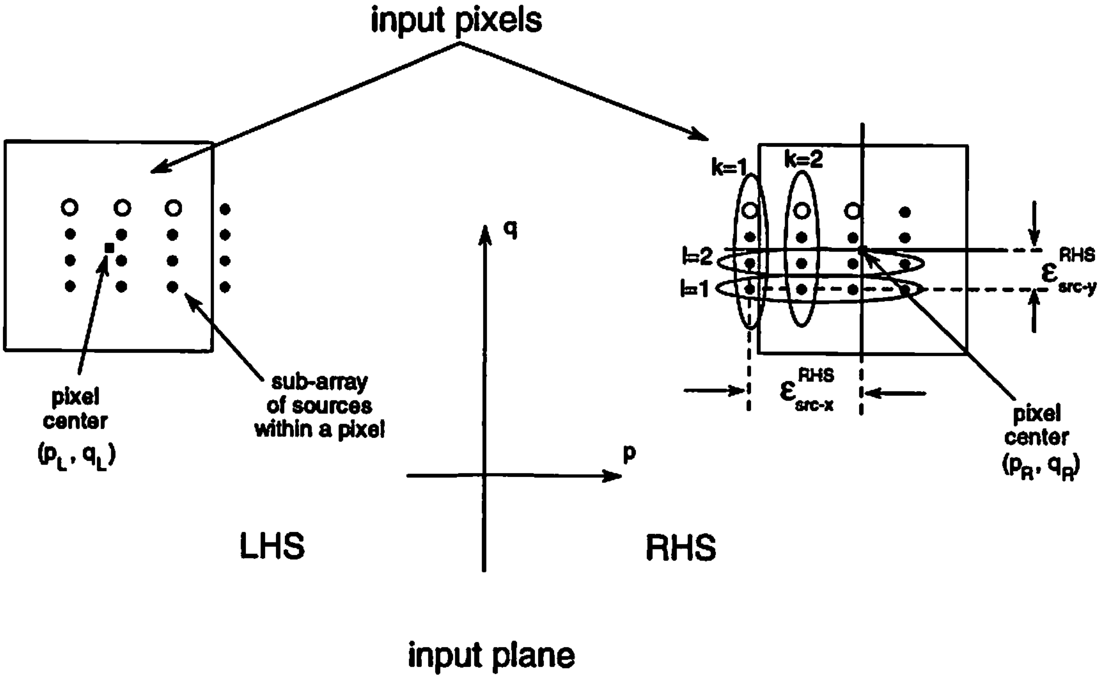


Figure 7.15: Pixels on opposite sides (RHS and LHS) of the input plane contain sub-arrays of sources that are mirror images of each other. The pixel input coordinates (p_R, q_R) and (p_L, q_L) are shown, where the subscripts are for *right* and *left* sides of the array. Column index k and row index l, and their displacement ϵ in x and y, are indicated for the example of 13 stages.

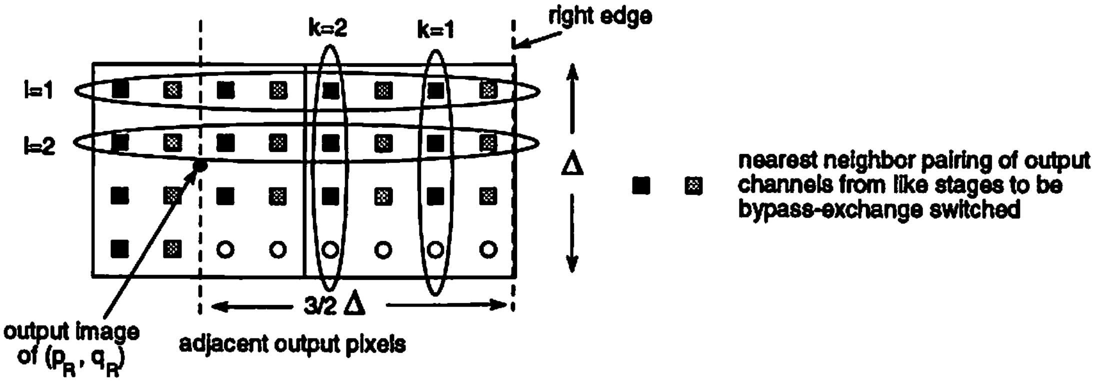


Figure 7.16: Adjacent output pixels for the case of $m = 13$ stages, showing channel pairs of like stages as nearest neighbors. Row and column indices k and l are indicated.

output channels spaced equally by the minimum separation δ needed to maintain -20 dB optical crosstalk.

The corresponding image in the output plane is shown in Fig. 7.16. The output image of $k = 1$ is located in the second column from the right edge of the output pixel. The right edge of the output pixel is three halves the pixel size Δ from the center of the left hand pixel (where the image would appear without had the source not been shifted from the center of its input pixel). The first column is $\delta/2$ from the right edge of the pixel. Each subsequent column to the left is an additional δ from the right edge of the adjacent pixel. So, the column $k = 1$ of detectors in the output is displaced in the x direction by

$$\epsilon_{det-x}^{RHS}(k = 1) = \frac{3}{2}\Delta - \frac{3}{2}\delta \tag{7.6}$$

from the center of the left hand output pixel. The superscript RHS indicates that the shift applies to pixels on the right hand side of the array. The corresponding shift for the left hand side of the array is the same amount, but in the opposite direction

$$\epsilon_{det-x}^{LHS} = -\epsilon_{det-x}^{RHS}. \tag{7.7}$$

In general, we incorporate this into the notation by multiplying the shift by sgn(p)

$$\epsilon_{det-x}(k=1) = \mathrm{sgn}(p)[\frac{3}{2}\Delta - \frac{3}{2}\delta] \tag{7.8}$$

where p is the pixel coordinate in the intermediate coordinate system (p, q) defined in Section 3.4.

In order to generalize Eq. (7.6) for all $k \in \{1, 2, \dots \sqrt{\tau}\}$ output columns, we recognize that subsequent columns (i.e. for $k > 1$), the shift is two columns to the left (for pixels on the RHS of the array). Therefore, the next output column $k = 2$ is displaced from its pixel center in x by

$$\epsilon_{det-x}(k=2) = \mathrm{sgn}(p)[\frac{3}{2}\Delta - \frac{3}{2}\delta - 2\delta]. \tag{7.9}$$

In general, a column k in a sub-array in the output plane (corresponding to a source in the right hand side of the input plane) is shifted from its pixel center by

$$\epsilon_{det-x}(k) = \mathrm{sgn}(p)[\frac{3}{2}\Delta - (2k - \frac{1}{2})\delta] \tag{7.10}$$

where k is an integer such that

$$1 \leq k \leq \sqrt{\tau}. \tag{7.11}$$

Source Sub-array for Interlaced Output

To calculate the required shift of a source to produce an interlaced output, we use the characteristics of the lens-based shuffle. The optics magnifies the input by two and inverts the image. So, a source is displaced in x from its pixel center by half the distance, and in the opposite direction, of its image in the output. Thus, the x displacement of sources (from their pixel centers) in column k of an input sub-array is

$$\epsilon_{src-x}(k) = -\frac{1}{2}\epsilon_{det-x}(k) = \mathrm{sgn}(p)[(k - \frac{1}{4})\delta - \frac{3}{4}\Delta]. \tag{7.12}$$

Source and Detector Position in the Non-interlaced Direction

To calculate the shift of sources and detectors in the y direction, we note that there is no overlap between pixels in y. There is also no difference between the the right and left hand sides of the array

$$\epsilon^{RHS}_{det-y} = \epsilon^{LHS}_{det-y} \tag{7.13}$$

$$\epsilon^{RHS}_{src-y} = \epsilon^{LHS}_{src-y} \tag{7.14}$$

Consider the various rows in the output plane. The first row $l = 0$ of detectors is one half of the detector spacing δ from the top of the pixel. So, its shift in y from the center of the pixel is

$$\epsilon_{det-y}(l = 1) = \frac{1}{2}\Delta - \frac{1}{2}\delta. \tag{7.15}$$

The y displacement of each succeeding row is less by δ than the preceding displacement. In general, the y displacement of row l of a subarray in the output plane is

$$\epsilon_{det-y}(l) = \frac{1}{2}\Delta - (l - \frac{1}{2})\delta \tag{7.16}$$

where l is an integer such that

$$1 \leq l \leq \sqrt{\tau}. \tag{7.17}$$

Accounting for the magnification by 2 and the inversion in the image gives the displacement in y of the corresponding source

$$\epsilon_{src-y}(l) = (\frac{l}{2} - \frac{1}{4})\delta - \frac{1}{4}\Delta \tag{7.18}$$

General Procedure

The following is a design procedure for a system that shuffles an $n \times n$ array and multiplexes m stages through a single set of optics by means of interlacing.

1. Arrange the bypass-exchange switches horizontally.
2. Determine the total space bandwidth product (SBWP) needed using Eqs. (7.4) and (7.3).
3. Use the design procedure described in the Chapter 4 to choose the optimum focal length lens, aperture size, and array size for the above SBWP. This procedure then specifies the minimum detector spacing δ subject to a maximum -20 dB optical crosstalk, or other crosstalk limitation, and is the same δ used in the current calculations.
4. Calculate the pixel size Δ using Eq. (7.5).
5. The displacement of sources from their pixel centers (i.e. the position they occupy in a non-multiplexed network) in x and y can be obtained using Eqs. (7.12) and (7.18). The possible values of the indices k and l are given by Eqs. (7.11) and (7.17).
6. Although their arrangement is more straightforward, one may determine the displacement of corresponding detectors from their pixel centers analytically using Eqs. (7.10) and (7.16). Like indices k and l indicate source and detector locations for like channels.

7.4.5 Analysis of Space Multiplexing using the Lens-based Shuffle

The implicit advantages of multiplexing several stages of shuffles through one set of optics are that the overall system hardware is reduced and the duplicated alignment efforts for the many stages is eliminated. In exchange for these advantages, we place greater demands on the ability of the optics to resolve more data lines. Here we investigate the consequences of greater resolution requirements.

We have shown previously [25] that if simple lenses are used to execute the shuffle, there is a direct relation between the number of data lines the system must shuffle and the shortest focal length lens which can resolve that number of data lines. Figure 4.13 shows that as the number of lines increases, so does the minimum focal length. This implies that when we pipeline many stages of shuffles through one set of lenses, the shortest focal length we may use must be greater than that of the non-pipelined case. So, it is no longer clear if there is a reduction in the overall system length due to space multiplexing.

To examine this question, we look at a specific example. Let the number of channels in the system N be 5000 (or $n = 71$ in 1-D). According to the analysis in Chapter 4, a 1.1 cm focal length lens is sufficient to resolve 71 channels in 1-D. The number of stages in the system is $\log_2 N = 13$. Referring to Fig. 4.1, one stage of the optics requires a minimum length ℓ_{stage} given by

$$\ell_{stage} = s + s' \tag{7.19}$$

where s and s' are the object and image distances, respectively. For the shuffle, we need a magnification of 2, which means that the image distance must be twice the object distance. Using the imaging equation

$$\frac{1}{f} = \frac{1}{s} + \frac{1}{s'}, \tag{7.20}$$

where f is the focal length of the lenses, we find that $s = \frac{3}{2}f$ and $s' = 3f$, so that the length of a stage is

$$\ell_{stage} = \frac{3}{2}f + 3f. \tag{7.21}$$

With $f = 1.1$ cm, the stage length is $\ell \approx 4.95$ cm. Then the whole system, which is composed of thirteen stages, is 13×4.95 cm $= 64.4$ cm.

Using the previous example of Section 7.4.4, in which we multiplex thirteen stages through one set of lenses, then the number of resolvable points must increase by a factor of thirteen. Figure 7.14 showed the source and detector arrangements for adjacent pixels. Notice that there are three unused resolvable spots per pixel, indicated by open circles. So, the actual increase in resolution in 2-D required of the optics is sixteen times the original 5000, or 80,000 resolvable spots. Figure 4.13 is from our previous work [25], and shows that the minimum focal length required to resolve 80,000 channels is 17.4 cm. Since there is only one set of optics, the whole system length is just the length of one stage, given by Eq. (7.21).

$$\ell_{system} = \frac{3}{2}(17.4 \text{ cm}) + 3(17.4 \text{ cm}) = 78.3 \text{ cm}. \tag{7.22}$$

Notice that this is longer than the non-pipelined system by approximately 22 %. The wasted resolvable spots in each pixel roughly account for this difference. In general the length of the multiplexed system is $4.5f$, where f is specified by the analysis of Section 4.2.7 for a SBWP of N_T.

7.5 Space Multiplexing by Blocks

7.5.1 General Description

The one-copy algorithm for the perfect shuffle, which is implemented by a four-facet hologram [42], suggests another means of space multiplexing [6]. In the one-copy algorithm, the input plane is divided into four quadrants (familiar from algebra), which we refer to as Q_i, where $i \in \{I, II, III, IV\}$. The input plane is first magnified by two in each dimension. The four magnified quadrants are shifted appropriately, and stacked to produced the 2-D shuffled results. Both 2-D separable and folded shuffles for a 4 × 4 input array are shown in Fig. 7.4. The optical system which performs the one-copy algorithm is a hologram with four facets, each of which acts like a prism to redirect the collimated input beams. The outputs of the four facets are superimposed in the output plane to yield the shuffled result.

The key to understanding space multiplexing by blocks is that, for multistage networks, like quadrants in the input planes of the various stages must be processed by the same facet; the shift required on like quadrants, regardless of stage number, is the same. So, in order to multiplex many stages through a single

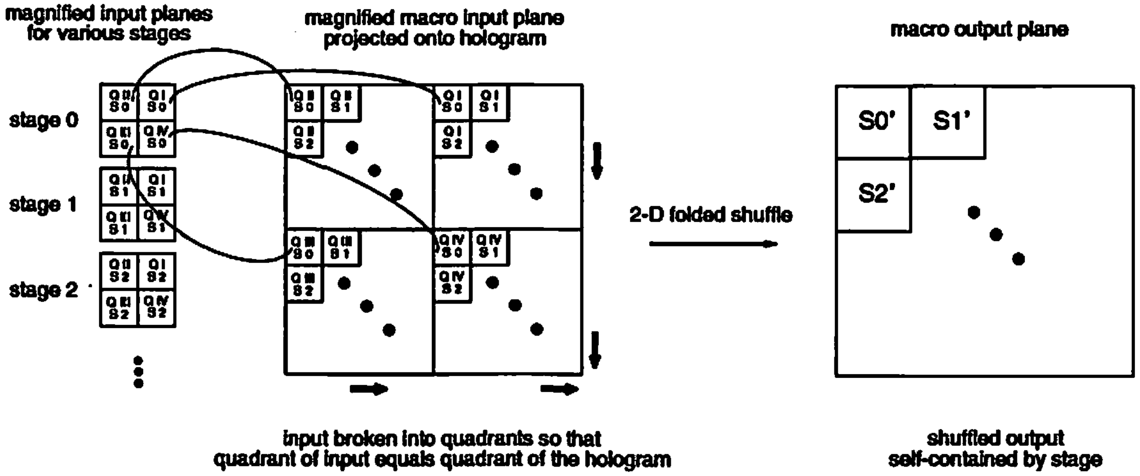


Figure 7.17: Space multiplexing many stages of an Omega network in blocks.

optical shuffle, one simply breaks the input planes for all the stages into quadrants $\{Q_I, Q_{II}, Q_{III}, Q_{IV}\}$, and place all of the Q_I for all stages into the upper left quadrant of a larger input plane, all of the Q_{II} into the upper right quadrant, and so on. The idea is shown schematically in Fig. 7.17.

7.5.2 Consequences

There are two major consequences of using the block method in the above system. First is that the system must be capable of resolving more channels. In other words, the optics must have a higher SBWP. This is likely to mean that a larger hologram and larger lens are necessary. The SBWP requirement for a similar system used for a single stage operation has been analyzed in [16]. Basically, for a single stage network having N channels, the total SBWP of the system using the one-copy algorithm is $4N$. In the space multiplexed m stage network, the total SBWP requirement is $4mN$.

A key feature of the block method of space multiplexing is that the shuffled output of each stage appears in the output plane as a self-contained unit as shown in Fig. 7.17. This means that the neighboring pairs of output channels in like stages are always nearest neighbor pairs when the block method is used. Thus, we circumvent the wiring problem encountered in the interlace method.

The price of this convenience, though, is that a more sophisticated feedback system is required for the block method than for the interlace method. The system must have the ability to break an input array into quadrants and place them appropriately within the larger input plane. This break-and-put function is required both for the initial system input, and after the bypass-exchange operation is performed at each stage. So, the system needs a *pre-processing* step for the initial input, and a *backward processing* step embedded in the feedback loop. The step of performing 2-D shuffles can be refer to as *forward processing.* Thus, the complete system is composed of *pre-processing*, *forward processing*, and *backward processing.*

7.6 General Applicability of Both Multiplexing Techniques

We recall that we used the lens-based shuffle to illustrate the interlace method of space multiplexing, and the facet hologram shuffle to illustrate the block method. In fact, the two different space multiplexing techniques work equally well in either shuffle architecture. Only because we originally suspected that the lens-based shuffle must be tolerant to lateral misalignment of the input plane did we then define the interlace method depicted in Fig. 7.6. The one-copy shuffle algorithm (implemented with a four-facet hologram), on the other hand, is suggestive of the block method.

The intuitive association between each shuffle architecture and the different space multiplexing techniques precludes neither the application of the interlacing method to the holographic shuffle, nor the block method to the lens-based shuffle. The one caveat is that the skewed arrangement of sources designed to minimize wiring problems is impermissible with the holographic shuffle, because of the cross-over of sources from one pixel to another at the central vertical axis of the input plane. The sources which cross the vertical axis would strike the wrong facet on the hologram, and therefore not reach their correctly shuffled position at the output. The simple interlacing of Fig. 7.6, however, is easily achievable with the holographic shuffle.

7.7 Conclusion

We have shown how the amount of optical hardware in a multistage interconnection network may be reduced by space multiplexing many stages of shuffles through a single set of optics. We describe two distinct methods of achieving the space multiplexing, which we call the overlap and block methods. Each of these methods is equally applicable to a shift-invariant optical shuffle implemented with lenses (i.e. bulk optics) and to a partially shift-variant system implemented with a facet hologram. We have also shown the physical justification and limitations of both methods.

Chapter 8

Comparison of Lens-based Shuffle to the Holographic Shuffle

8.1 Background

We have analyzed the capabilities of the lens-based shuffle, and in Chapter 7, we described a holographic design which implements the one-copy algorithm for the shuffle. There are trade-offs between the characteristics of the two architectures. For example, the lens-based optical shuffle is extremely easy to fabricate, has good resolution capability, and the number of lenses required does not change with the number of inputs. It accommodates any divergent pseudo-point source, including LEDs or lasers. The lens-based architecture does not, however, accommodate a modulator array illuminated by a collimated beam.

The shuffle which uses a four-facet hologram is slightly more difficult to manufacture, although its complexity (number of facets) remains the same regardless of the number of inputs, just like the lens system. It has very good light efficiency, in principle approaching 100 %, but only accommodates collimated sources. We would, however, like the interconnection network to function in the infrared, but holographic materials are not available in abundance Furthermore, because the holographic system operates like a prism array, it has no focusing power, and its resolution capability is therefore limited.

In this chapter, we compare the relative abilities of the two methods. In order to analyze the resolution capability of the holographic shuffle (its space bandwidth product), we use simple first order calculations [41]. The comparison leads us to suggest a modified version of the hologram which accommodates point sources and has focusing capability to achieve higher resolution.

8.2 Resolution Capability of the Holographic Shuffle

8.2.1 Diffraction in the Holographic Geometry

As with the lens-based shuffle, the resolution of the holographic shuffle is limited by diffraction and aberrations. Because the holographic shuffle uses collimated inputs, however, aberrations are minimal. The dominant limitation is diffraction. Figure 8.1 shows that diffraction arises in two places, the first being the transit from the source to its image at the hologram. The second is the transit from the spot on the hologram to the output plane.

Within the first region, we identify two sources of diffraction. One is the limited aperture size of the lens, but this is a small effect. The main limit on resolution at the image plane is the limited size a of the source. So to first order, we ignore the effect of diffraction by the lens aperture. Effectively, this means that the lens is large enough, and sufficiently free of aberrations to recreate the wavefront from the source at the image plane. Because this is a 4-f imaging system, we expect the size of the image on the hologram to be nearly equal that of the source.

Light transmitted through the source array mask is roughly collimated, so there is a large depth of focus in a neighborhood on either side of the image plane, allowing us to tilt the hologram while maintaining a good image over the

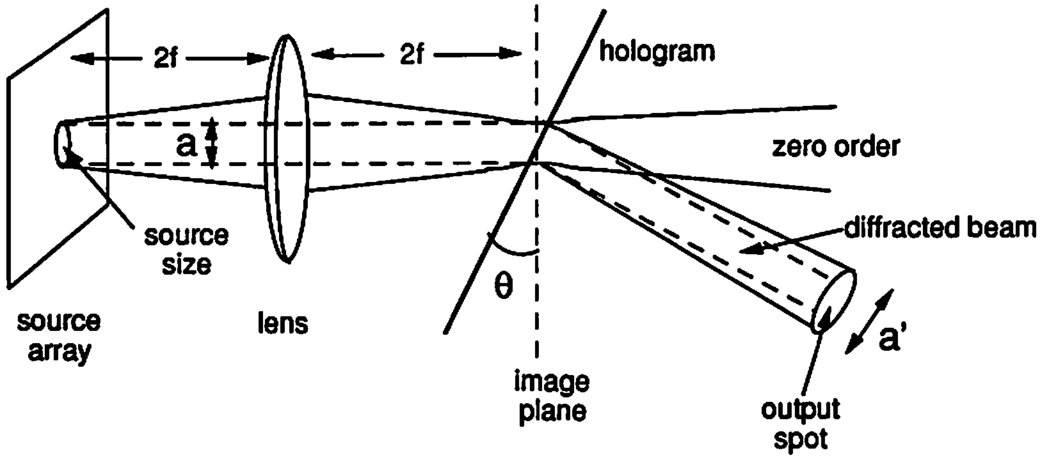


Figure 8.1: Diffraction of the collimated illumination through the "source", imaging by the lens, and diffraction from the spot on the hologram.

whole array. Thus, diffraction from the limited size of the source does not greatly affect the *size* of the image spot on the hologram. Despite the minimal significance of the effect, we retain a simple linear approximation for it in our analysis. The *angular* deviation of the beam due to diffraction (as opposed to the increase in the spot *size* on the hologram), however, affects the reconstruction of the hologram, which we shall examine later.

The second region for diffraction in Fig. 8.1 is from the spot on the hologram to the spot on the output array. We approximate this by first order Fresnel diffraction. If the diameter of the spot on the hologram is a, then the size of the spot on the output plane is approximately

$$a' \cong a + 2\lambda\frac{z}{a} \tag{8.1}$$

where z is the distance between the hologram and the output plane.

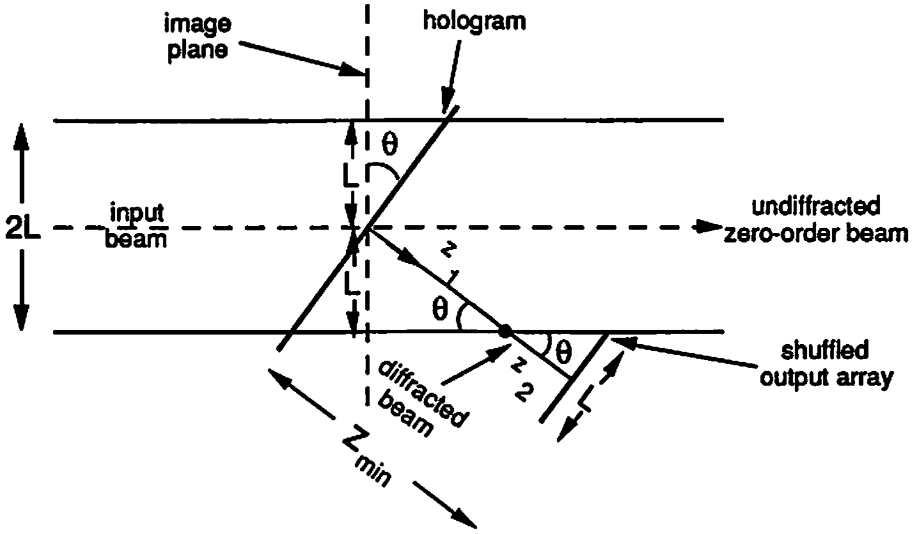


Figure 8.2: Geometry of the separation of the transmitted (zero order) and first order diffracted beams for the holographic shuffle.

8.2.2 Minimizing Diffraction

Distance from Hologram to Output Plane

To minimize the effects of diffraction, the distance z from the hologram to the output plane must be as small as possible. Because we must separate the first order diffracted beams from the zero order (undiffracted) beams, there is a minimum z required between the hologram and the output plane. The situation is shown in Fig. 8.2. The array of input beams has been pre-magnified by a factor of two for projection onto the hologram, so the width of the array of zero order beams emerging from the hologram is $2L$.

The total distance to the output plane is the sum of z_1 and z_2, which are

$$z_1 = \frac{L}{\sin\theta} \tag{8.2}$$

$$z_2 = \frac{L/2}{\tan\theta}. \tag{8.3}$$

The net distance is thus

$$z_{min} = L(\frac{1}{\sin\theta} + \frac{1}{2\tan\theta}). \tag{8.4}$$

The shortest distance possible is L, but this requires a deflection angle of nearly 90 degrees by the hologram. A more reasonable deflection angle, of perhaps 45 degrees, implies a minimum distance of roughly

$$z_{min} \approx 2L. \tag{8.5}$$

So, if we pack the output points directly next to each other, then in 1-D, the array may contain L/a' channels, or

$$n = \frac{L}{a'} = \frac{L}{a + 2\lambda(\frac{2L}{a})} \tag{8.6}$$

$$n = \frac{aL}{a^2 + 4\lambda L}. \tag{8.7}$$

Optimum Source Size

By intuition, we expect that if we increase the size of the source, and therefore its image on the hologram, then diffraction from the hologram to the output plane is reduced. The result is sharper output spots, but fewer of them. Toward the opposite extreme, if we decrease the size of the source so that diffraction becomes significant, the size of the output point increases, thus reducing the number of channels. So, we expect that there is an optimum source size a_{opt} which maximizes the number of channels. In Fig. 8.3, we plot the number of channels n versus source size a using Eq. (8.7), which confirms that this is exactly what happens. We find the optimum source size by differentiating Eq. (8.7) by a, which gives

$$\frac{\partial n}{\partial a} = \frac{4\lambda L^2 - a^2 L}{(a^2 + 4\lambda L)^2}. \tag{8.8}$$

Setting the derivative equal to zero and solving for a gives the source size which maximizes the number of channels,

$$a_{opt} \approx 2\sqrt{\lambda L} \tag{8.9}$$

Using Eq. (8.7), we see that for a given array size, L the maximum number of channels is

$$n_{max}(L) = \frac{(2\sqrt{\lambda L})L}{(2\sqrt{\lambda L})^2 + 4\lambda L} \tag{8.10}$$

$$n_{max} = \frac{1}{4}\sqrt{\frac{L}{\lambda}}. \tag{8.11}$$

For example, if we use red light at $\lambda = 600$ nm, and an array $L = 1$ cm on a side, we calculate that $n_{max} = 32$ for a source size $a = 0.0155$ cm. The curve for $L = 1$ cm in Fig. 8.3 depicts this case.

8.2.3 Effect of Diffraction on the Replay of the Hologram

Recall that we discarded the effect of diffraction from the source to the hologram when calculating the *size* of the image on the hologram. The *angle* at which rays strike the hologram, however, is critical. In order to reconstruct the hologram faithfully, the hologram must see the same wavefront as when it was recorded. Angular deviation of the k-vector of the read beam, $\vec{k}_{read}$, from the k-vector of the write beam, $\vec{k}_{write}$ gives rise to a shift in the output beam, $\vec{k}_{out}$. The effect is shown in Fig. 8.4.

We model the effect of a difference between the read and write beam angles using the grating equation

$$d[\sin\alpha_{out} - \sin\alpha_{in}] = \lambda, \tag{8.12}$$

where d is the spacing of the diffracted orders, and α_{in} and α_{out} are the angle from normal of the input and output beams. With no diffraction, the input (or

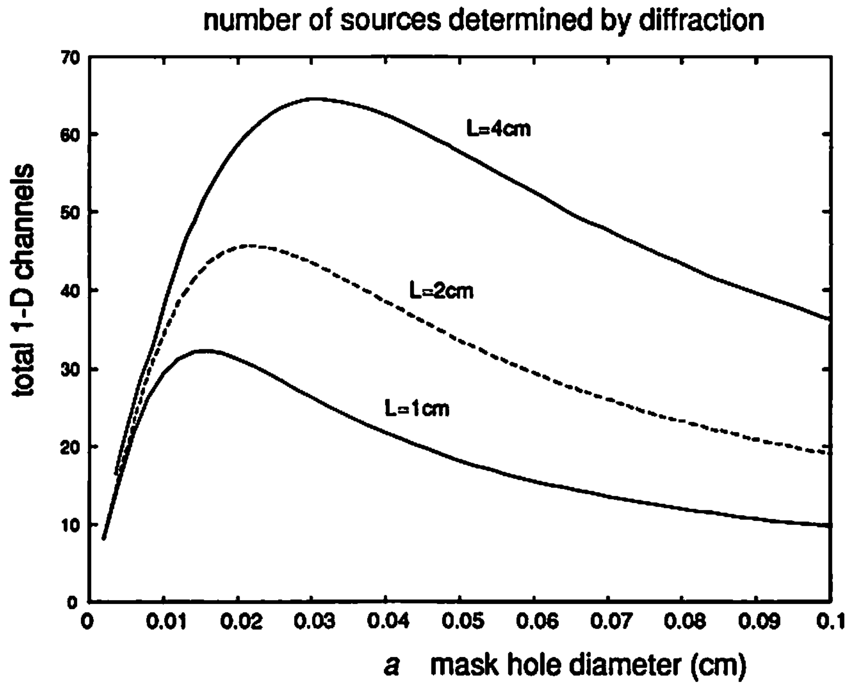


Figure 8.3: Maximum number of channels in 1-D versus source aperture size as a result of diffraction in the holographic shuffle, plotted for 3 different array sizes; from the lowest curve $L = 1, 2, 4$ cm

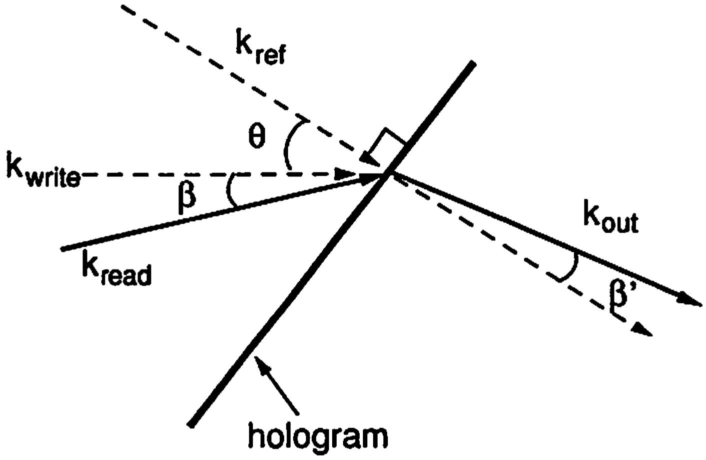


Figure 8.4: Geometry for incident and output angles of the hologram.

read) beam is parallel to $\vec{k}_{write}$, and α_{in} is equal to θ. The output beam is parallel to $\vec{k}_{ref}$, which is perpendicular to the hologram, thus making α_{out} identically zero. To approximate the effect of diffraction from the image upon readout of the hologram, we set the input and output angles to

$$\alpha_{in} = \theta + \beta, \tag{8.13}$$

or the sum of the angle of the hologram and the angle at which the beam diffracts from the horizontal, and

$$\alpha_{out} = \beta' \tag{8.14}$$

the angular deviation of the output beam from the reference beam. The exact solution for the output angle β' is given by

$$\sin\beta' = \frac{\lambda}{d} + \sin(\theta + \beta). \tag{8.15}$$

We wish to find the effect of a small deviation between the write beam and the read beam on the output angle. Equation (8.15) gives the us the output angle β' as a function of wavelength λ, grating spacing d, as well as the relevant input angles θ and β

$$\beta' = f(\lambda, d, \theta, \beta). \tag{8.16}$$

In order to avoid having to specify the parameters d and λ, we differentiate the output angle with respect to the input angle

$$\frac{\partial\beta'}{\partial\beta} = \frac{\cos(\theta + \beta)}{\cos\beta'} \approx \frac{\cos\theta}{1}. \tag{8.17}$$

This gives the change in the output angle $\Delta\beta'$ as a function of an incremental change in the input angle $\Delta\beta$

$$\Delta\beta' \approx \frac{\partial\beta'}{\partial\beta}\Delta\beta. \tag{8.18}$$

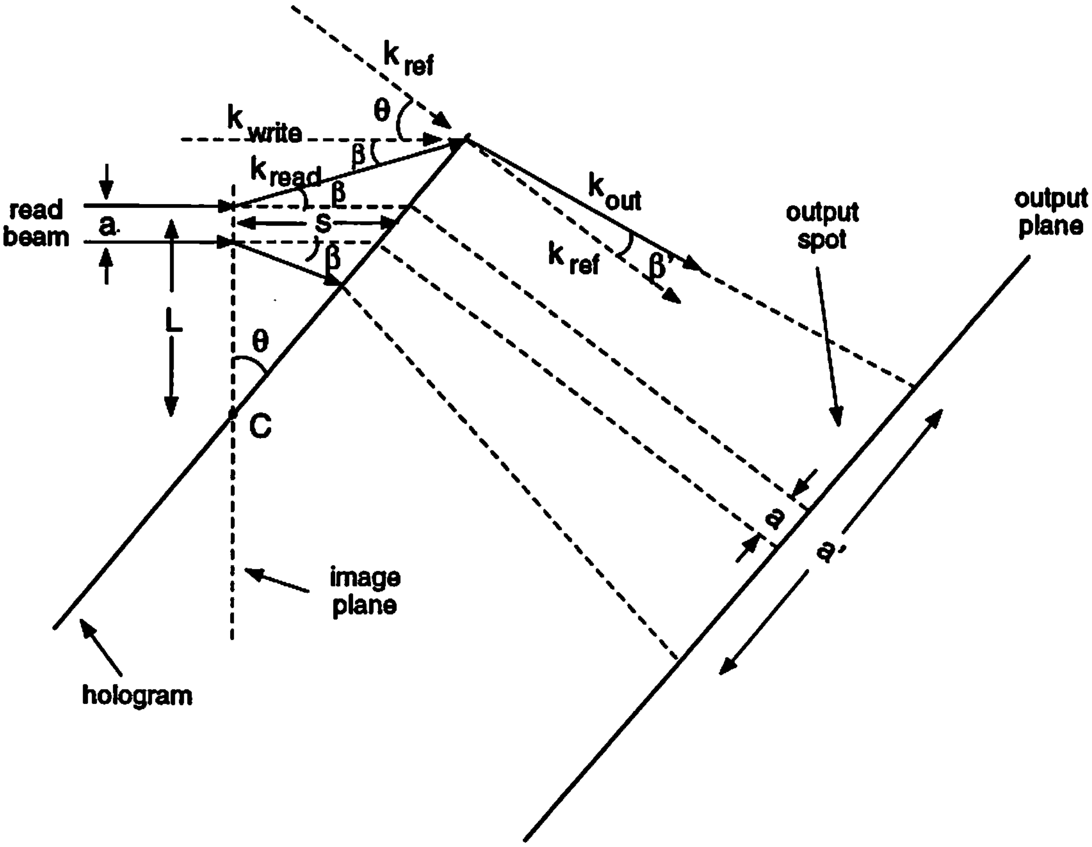


Figure 8.5: k-vector addition of the write vector and hologram vector to produce the output, or reference vector.

Since β' is zero for zero diffraction angle β, then

$$\beta' \cong \beta \cos\theta \tag{8.19}$$

8.2.4 Output Spot Size

We now calculate the effect of the angular deviation β' of the beam on the output spot size. The situation is shown in Fig. 8.5. The ideal input (read) beam is horizontal and perfectly collimated to a diameter of diameter a. With a good lens, we may closely approximate the ideal situation in the image plane; the wavefront at the source is recreated in the image plane by the lens.

Because of its angular tilt θ, the hologram does not lie in the image plane except at the center (labeled C in Fig. 8.5). Diffraction occurs over the small distance s from the image plane to the hologram. There are two effects of the tilt θ. First, there is a marginal increase in the size of the beam to roughly $a + 2r$. This is a minor effect, though we retain it in the calculations. We find the increase r in spot size on the hologram by means of the law of sines, which gives

$$r = s\frac{\sin\beta}{\cos(\theta+\beta)}. \tag{8.20}$$

For a beam at the array edge, the distance between the image plane and the hologram is

$$s = L\tan\theta. \tag{8.21}$$

We approximate the diffraction angle β from the image by

$$\beta = \frac{\lambda}{a}. \tag{8.22}$$

Because the wavelength of light is likely to be much smaller than the size of the source (and therefore the image), we may assume that β is small, so

$$r \cong s\frac{\beta}{\cos\theta} = \frac{\lambda L}{a}\sec\theta\tan\theta. \tag{8.23}$$

The spot size on the hologram is

$$a'' = \frac{a}{\cos\theta} + 2r. \tag{8.24}$$

The more significant effect is the *angular* deviation β of the read beam, and its effect on the replay of the hologram. Without diffraction from the image plane, the beam exits the hologram at angle appropriate to execute the shuffle, which is very close to normal to the hologram. The angular difference β between the read and write beams results in an output beam deflected by β' from the normal. The size a' of the spot at the output is the sum of the spot size at the hologram a'', and the projection of the altered output angle β' over the distance z to the output plane

$$a' = a'' + 2z\beta'. \tag{8.25}$$

Using Eq. (8.4), we obtain the output spot size

$$a' = \sqrt{\lambda L}\{2\sec\theta + \sec\theta\tan\theta + \cot\theta + \frac{1}{2}\cos\theta\cot\theta\} \tag{8.26}$$

We plot the dependence of a' (Eq. (8.26) normalized to $\sqrt{\lambda F}$) on θ in Fig. 8.6, which shows that if the hologram is tilted by roughly $\theta = 36$ degrees, output spot size is minimized. This is close to our original assumption of 45 degrees for the tilt. In order to obtain a more exact answer for the optimum θ and a', we could reiterate the above process using $\theta = 36$ degrees, but we do not do that here. The best obtainable output spot size a', however, is

$$a'_{opt} \approx 5.3\sqrt{\lambda L}, \tag{8.27}$$

as we can see in Fig. 8.6 or by evaluating Eq. (8.26) for $\theta = 36$.

8.3 Comparison with the Lens-based Shuffle

If we use Eq. (8.27), choose a likely wavelength of about 600 nm, and a reasonable array size of roughly 1 cm, we see that the spot size in the output plane is about 0.4 mm. If we tight-pack the output points so that they touch edge-to-edge, we could fit $n = 24$ in 1-D, or $N = 576$ in the 2-D array. In comparison to the lens system, this amounts to a poor capacity. Though we had defined a somewhat stricter requirement on output point spacing [25] (so as to maintain a maximum -20 dB optical crosstalk), the lens system may resolve $n = 181$ points in 1-D, $N = 32,761$ points over the 2-D array. Furthermore, this resolution capability

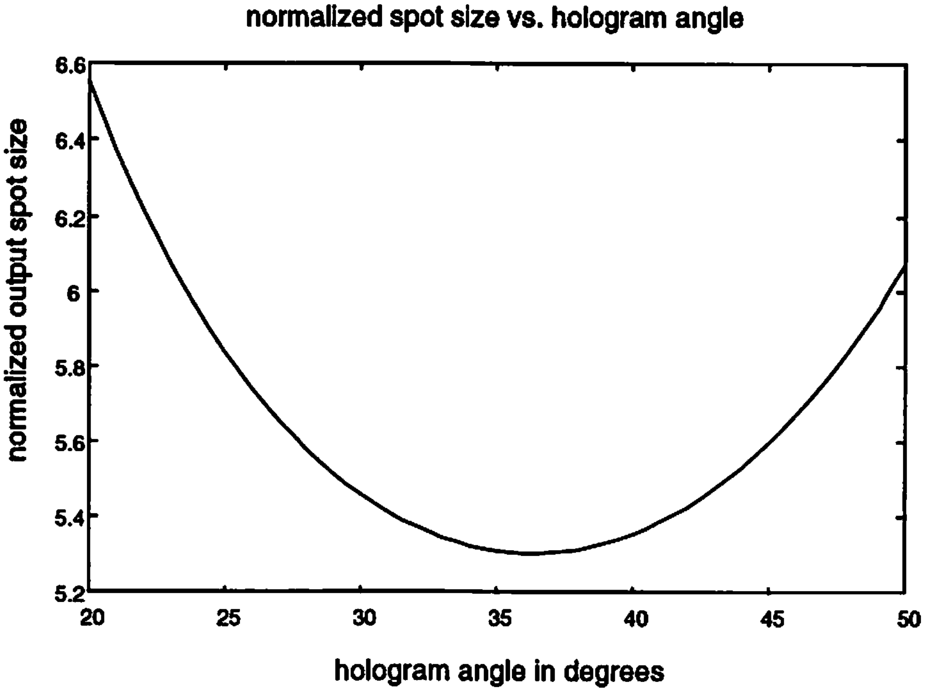


Figure 8.6: Variation of output spot size a', normalized to $(\lambda F)^{1/2}$, with the angular tilt θ of the hologram, which shows that a' is minimum when θ is about 36 degrees.

f focal length	ρ_{max} (per cm^2) maximum density	P_{crit} critical power
15 cm	1.31×10^4	76.3 μW
10 cm	1.61×10^4	62.3 μW
holographic	576	1.74 mW

Table 8.1: Comparison of the channel resolution capability between the lens-based and holographic one-copy implementations of the shuffle, and the corresponding critical power per source above which the packing density of the shuffle is limited by heat dissipation rather than optical resolution. The top two rows are sample focal lengths for the lens-based implementation, and the bottom row is for a 1 cm^2 source array in the holographic shuffle.

is achieved with simple, uncorrected spherical lenses. The performance would increase with aberration correction.

There is a fundamental distinction between the holographic and lens systems which accounts for the difference in capacity. The lens shuffle is an imaging system, whereas the holographic shuffle is merely a beam steering system. So, the hologram lacks focusing power. If we imagine a holographic shuffle system with focusing power, we might expect to raise the resolution capability of the hologram. Calculations show, however, that holographic aberrations for a *simple* holographic imaging system are severe for off-axis sources.

8.4 Future Direction

At first, it may seem an insignificant characteristic that the holographic system employs collimated light, both to write the hologram and to read it. One might imagine that a straightforward conversion is possible to make the system accommodate pseudo-point sources, such as vertical cavity surface emitting lasers. We have found that while the conversion should be possible, it is not a trivial process. The important idea is to create a hologram with the following characteristics:

- the complexity in terms of the number of facets does *not* scale up with the number of channels in the system
- accommodates pseudo-point sources
- has the power to focus the output light to a point
- to maintain high light efficiency, the system would implement the one-copy algorithm for the shuffle

In order to avoid making a completely space-variant hologram, there are two basic options for a new shuffle implementation. The first is similar to the original holographic one-copy approach. We use a pair of pseudo-point sources as the write beam and reference beam to record four different facets in a holographic medium. The medium could be either dichromated gelatin or a photopolymer, but must result in an index grating. The difference from the original holographic shuffle is that the new facets create *images* [38], instead of merely deflecting collimated beams. The center of each quadrant in the input plane is imaged by a corresponding facet on the hologram to a point in the output plane. The location of the point in the output plane is different for each facet of the hologram, and the difference specifies the shift applied to each quadrant of the input to accomplish the shuffle.

The remaining option is to use computer design techniques to create an index grating [18] which accomplishes the shuffle by the one-copy algorithm, and satisfies the other goals listed above. The grating would be discrete, either binary (with zero or π phase shift at any point) or multi-level. There is one fundamental issue relevant to both options. The original holographic shuffle largely avoided aberrations by using collimated input and collimated output. The new hologram (or grating) will map input points to output points. It constitutes a holographic imaging system. It will, therefore, be susceptible to imaging aberrations for any inputs which are off-axis (e.g. shifted relative to the original write beam) [38], just like the lens-based system. So, a significant part of the design investigation, regardless of which option is chosen, is to understand, quantify and minimize the effect of these aberrations.

Chapter 9

Conclusion

We have examined in detail the capabilities of simple bulk optics to implement a three-dimensional shuffle interconnection. The optical shuffle system is a 2×2 array of single element spherical lenses. We chose this architecture primarily for its ease of implementation. There are only four components, and this number is not dependent on the number of inputs in the network. This means that the complexity does not scale up with the size of the input array. Another attractive feature of the optics is its ability to focus light at the output, thus allowing greater spatial resolution than a beam steering optical system, such as prisms or plane mirrors.

We chose to analyze the simplest optics, meaning single element lenses with spherical surfaces, because the rules governing their imaging capabilities are well established. We discussed the two physical limitations on the resolution capability of any imaging system, which are diffraction and aberrations. Using compound lenses, corrected for various aberrations, endows the optics with a combination of greater resolution, higher light throughput and shorter length, as demonstrated by the crossover network implemented at AT&T Bell Labs [24]. To design these lenses, however, requires sophisticated computer design tools, and their fabrication is correspondingly more difficult and expensive. We believe it to be instructive to examine in detail the simplest problem available. This approach allows us to understand the capabilities of simple bulk optics. It provides a foundation for understanding the basic characteristics of the system, and the factors limiting its performance. It also provides us with the limits of simple bulk optics, thus defining when sophisticated lens design is necessary.

Our study shows that simple lenses are capable of resolving on the order of 10^4 to 10^5 channels in a shuffle network. In order to attain these densities, we have specified a simple procedure by which one may iteratively adjust the geometry of the optical interconnection for simultaneous minimization of diffraction and aberrations. The densities achievable with simple bulk optics far exceed the heat dissipation capabilities for comparably dense laser diode arrays. By far the largest amount of heat dissipation of a laser diode occurs at or below threshold. So, until threshold currents of laser diodes are reduced by roughly two orders of magnitude, they will be the factor limiting packing density in an optoelectronic interconnection network. We also found that the lens-based shuffle has relatively low optical throughput for channels at the corner of the array (approximately 4.6 %). Despite this seemingly poor performance, its effect on heat dissipation in the network is insignificant when compared to the threshold dissipation of the sources.

We studied several other quantitative aspects of an optical shuffle interconnection. We assembled the relevant statistical rules to specify the probability distribution of optical crosstalk in an array of optical channels. Then, we combined the effect of optical crosstalk with the amount of Gaussian noise likely to be present in the system in order to predict the effect of crosstalk on BER. We found that modeling crosstalk as additive Gaussian noise is inaccurate, but that the error is insignificant on a logarithmic scale. We also simulated the effect of certain perturbations (or errors) in the system construction. This allows a designer to choose appropriate tolerances on the optics and placement of the hardware.

One of the issues central to any interconnection network is the means by which the switches are set. Because this is such a large problem, we have not addressed the issue. We have, however, developed a simple set of rules which specify the input-to-output and the output-to-input relations for both the separable and folded perfect shuffles. These rules are relevant for anyone developing a routing algorithm for any interconnection network employing a lens-based shuffle.

In order to simplify a multi-stage network, we have proposed a means of multiplexing several stages of shuffles through a single set of shuffle optics. We discovered, in the process, that a key issue in space multiplexing is the complexity of the feedback system. If each stage requires its own separate set of optics, then no savings in system complexity is achieved by multiplexing. This has implications for how one accomplishes the multiplexing, but is not at all an insurmountable problem. We did find that, for a shuffle network employing simple lenses, space multiplexing does not lead to any reduction in length of the overall system. This is because there is a minimum focal length required to achieve a given SBWP which increases as the SBWP increases. The primary advantage of space multiplexing is the reduced complexity of the optics, allowing the fabrication and alignment of only a single set of shuffle optics, plus the feedback optics. We also discovered that the existing shuffle optics is capable of simplifying the electrical wiring between optical components (detectors and sources) and electrical components (the bypass-exchange switches). This is accomplished by a simple perturbation of the source arrangement in the input plane, a process which we have specified.

While the lens-based shuffle has several positive attributes, including simplicity, ease of construction, and good resolution capability, it is also relatively inefficient at transferring light from input to output. The holographic shuffle [42] is based on a one-copy algorithm designed to give it greater light throughput. We have examined the how the lens-based shuffle compares to the one-copy holographic shuffle. The difference is that in the lens-based (four-copy algorithm), the shuffle is accomplished by interlacing four copies (i.e. images) of the *entire* input array. Thus, three quarters of the image points, which are those on the periphery, are thrown away. There is also a large sum of light which simply does not pass through any of the four lenses, and is wasted. The one-copy algorithm produces only *useful* output points by deflecting the four quadrants of the input array appropriately to accomplish the shuffle. So in principle, the one-copy algorithm may attain 100 % light efficiency.

The reason for the enhanced light efficiency of the one-copy algorithm is that it uses collimated input. The lens-based shuffle requires divergent sources (unless the asymmetric lenses of Fig. 4.2 are used) because collimated light from many of the off-axis sources entirely misses the lenses. The distinction between collimated and divergent input has implications for the resolution capability of each system. The holographic shuffle simply steers collimated input, and so to first order, it is limited entirely by diffraction. We have shown that this lack of imaging (or focusing) ability makes the system incapable of the resolution attainable using the lens-based shuffle. So, there is a trade-off between the two architectures in resolution and light throughput. A worthwhile continuation of this work would be to search for a holographic arrangement which would accommodate pseudo-point sources, such as surface emitting laser diodes, have high light efficiency, and high resolution. We have proposed a pair of possible approaches to this problem.

References

[1] G. Arfken, *Mathematical Methods for Physicists*, Orlando: Academic Press, 3rd ed., 1985.

[2] P. Barnsley and P. Chidgey, "All-optical wavelength switching from 1.3 μm to a 1.55 μm WDM wavelength routed network: system results," *Photonic Technology Letters*, Vol. 4, pp. 91–94, January 1992.

[3] R. Boppana and C. Raghavendra, "On self-routing in Benes and (2n-1)-stage shuffle/exchange networks," in *Proc. Int. Conf. Parallel Processing*, pp. 196–200, August 1988.

[4] M. Born and E. Wolf, *Principles of Optics, 6th edition*, Oxford: Pergamon, 1980.

[5] L. Cheng and A. Sawchuk, "Three-dimensional Omega networks for optical implementation," *Applied Optics*, Vol. 31, pp. 5468–5479, September 10, 1992.

[6] L. Cheng, J. Wang, and A. Sawchuk, "Architectural design of space multiplexed 3-D omega networks and their optical implementations." submitted to Applied Optics August 1992.

[7] G. Eichmann and Y. Li, "Compact optical generalized perfect shuffle," *Applied Optics*, Vol. 26, pp. 1167–1169, April 1, 1987.

[8] S. Forrest, "Optical Detectors for Lightwave Communication," in *Optical Fiber Telecommunications II* (S. Miller and I. Kaminow, eds.), ch. 14, Boston: Academic Press, 1988.

[9] J. Goodman, *Introduction to Fourier Optics*, pp. 119–120, San Francisco: McGraw-Hill, 1968.

[10] M. Govindarajan and S. Forrest, "Optically powered arrays for optoelectronic interconnection networks," *Applied Optics*, Vol. 10, pp. 1335–1346, April 1991.

[11] J. Gowar, *Optical Communication Systems*, London: Prentice-Hall, 1984.

[12] M. Haney, "Pipelined optoelectronic free-space permutation network," *Optics Letters*, Vol. 17, No. 4, pp. 282–284, 1992.

[13] G. Heise, "Crosstalk investigation of laser diode pairs," *IEEE Photonics Technology Letters*, Vol. 2, pp. 97–99, February 1990.

[14] B. Jenkins *et al.*, "Architectural implications of a digital optical processor," *Applied Optics*, Vol. 23, No. 19, pp. 3465–3474, 1984.

[15] J. Jewell *et al.*, "Surface-emitting microlasers for photonic switching and interchip connections," *Optical Engineering*, Vol. 29, p. 210, March 1990.

[16] L. C. J.M. Wang and A. Sawchuk, "Optical 2-D perfect shuffles based on a one-copy algorithm." to appear in Applied Optics.

[17] C. Kuznia and A. Sawchuk, "Routing and Control for Free-Space Optical Cellular Hypercube Arrays." submitted to Applied Optics, August 1991.

[18] C. Kuznia and A. Sawchuk, "Cellular Hypercube Interconnections for Optical Processor Arrays," in *in Optical Computing 1991, Technical Digest Series*, vol. 6, (Washington, D.C.), pp. 41–44, Optical Society of America, March 1991.

[19] G. Lohman and A. Lohmann, "Optical interconnection network utilizing diffraction gratings," *Optical Engineering*, Vol. 27, pp. 893–900, October 1988.

[20] A. Lohmann, W. Stork, and G. Stucke, "Optical perfect shuffle," *Applied Optics*, Vol. 25, pp. 1530-1531, May 15, 1986.

[21] L. C. M. Govindarajan, S.R. Forrest and A. Sawchuk, "Optically powered 2×2 optoelectronic bypass/exchange switch with polarization routing," *IEEE Photonic Technical Letters*, Vol. 3, pp. 669–672, July 1991.

[22] V. Mahajan, "Aberrated point spread functions for rotationally symmetric aberrations," *Applied Optics*, Vol. 22, pp. 3035–3041, 1983.

[23] V. Mahajan, *Aberration Theory Made Simple*, Bellingham, WA: SPIE, 1991.

[24] F. McCormick *et al.*, "Optomechanics of a free-space photonic switching fabric: the system," in *Proc. SPIE Conf. on Optomechanics and Dimensional Stability*, pp. 1533–12, 1991.

[25] A. Miller and A. Sawchuk, "Optimization of the Lens-Based Perfect Shuffle." resubmitted to *Applied Optics*, March 1993.

[26] D. Miller, "Optics for low-energy communications inside digital processors," *Optics Letters*, Vol. 14, No. 2, pp. 146–148, 1989.

[27] P. Milonni and J. Eberly, *Lasers*, ch. 14, New York: Wiley, 1988.

[28] T. Nakagami, T. Yamamoto, and H. Itoh, "A multistage reconfigurable optical interconnection network using polarization switch arrays," in *in Photonic Switching 1991, Technical Digest Series*, vol. 6, (Washington, D.C.), pp. 146–149, Optical Society of America, March 1991.

[29] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, New York: McGraw-Hill, 2nd ed., 1984.

[30] B. Robertson *et al.*, "Space-variant holographic optical elements in dichromated gelatin," *Applied Optics*, Vol. 30, pp. 2368–2375, June 10, 1991.

[31] A. Sawchuk and I. Glaser, "Geometries for Optical Implementations of the Perfect Shuffle," in *Proc. ICO Optical Computing*, vol. 963, (Toulon, France), pp. 270–282, SPIE, August-September 1988.

[32] A. Sawchuk and B. Jenkins, "Dynamic optical interconnection for parallel processors," *Proc. SPIE*, Vol. 625, pp. 143–153, 1986.

[33] Y. Sheng, "Light effective 2-D optical perfect shuffle using Fresnel mirrors," *Applied Optics*, Vol. 28, pp. 3290–3292, August 15, 1989.

[34] H. Siegel, *Interconnection Networks for Large-Scale Parallel Processing*, New York: McGraw-Hill, 1990.

[35] P. Smith, "On the physical limits of digital optical switching and logic elements," *Bell System Technical Journal*, Vol. 61, pp. 1975–1993, October 1982.

[36] C. Stirk, R. Athale, and M. Haney, "Folded perfect shuffle optical processor," *Applied Optics*, Vol. 27, pp. 202–203, January 15, 1988.

[37] J. Taboury *et al.*, "Optical cellular processor architecture 2: Illustration and system considerations," *Applied Optics*, Vol. 28, No. 15, pp. 3138–3147, 1989.

[38] J. Taboury, J. Wang, P. Chavel, and F. Devos, "Optical cellular processor architecture 2: illustration and system considerations," *Applied Optics*, Vol. 28, pp. 3138–3147, August 1, 1989.

[39] K. Urquhart *et al.*, "Design of free-space optical interconnection systems utilizing diffractive optics," in *OSA Annual Meeting Technical Digest*, (Washington, D.C.), p. 70, Optical Society of America, November 1991.

[40] A. Varma, J. Ghosh, and C. Georgiou, "Reliable design of large crosspoint switching networks," in *Proc. 18th International Symposium on Fault-Tolerant Computing*, June 1987.

[41] J. Wang, "notes on holographic reconstruction." private communication.

[42] J. Wang, L. Cheng, and A. Sawchuk, "Holographic implementation of 2-D perfect shuffles based on a one-copy algorithm," in *in Photonic Switching 1991, Technical Digest Series*, vol. 6, (Washington, D.C.), Optical Society of America, March 1991.

[43] J. Wang, L. Cheng, and A. Sawchuk, "Optical two-dimensional perfect shuffles based on a one-copy algorithm," *Applied Optics*, Vol. 31, pp. 5464–5467, September 10, 1992.